SELECTED PROCEEDINGS

Temporal Analysis of Operational Errors at Air Traffic Control Facilitie

MICHAEL SEELHORST 107 MCLAUGHLIN HALL UNIVERSITY OF CALIFORNIA, BERKELEY BERKELEY, CA 94720 TELEPHONE: (865) 585-0935 EMAIL: MSEELHORST@BERKELEY.EDU MARK HANSEN 114 MCLAUGHLIN HALL UNIVERSITY OF CALIFORNIA, BERKELEY BERKELEY, CA 94720 TELEPHONE: (510) 642-2880 EMAIL: MHANSEN@CE.BERKELEY.EDU

This is an abridged version of the paper presented at the conference. The full version is being submitted elsewhere. Details on the full paper can be obtained from the author.

ISBN: 978-85-285-0232-9

13th World Conference on Transport Research

www.wctr2013rio.com

15-18 JULY 2013
Rio de Janeiro, Brazil

Temporal Analysis of Operational Errors at Air Traffic Control Facilities

Michael Seelhorst
107 McLaughlin Hall
University of California, Berkeley
Berkeley, CA 94720

Telephone: (865) 585-0935
Email: mseelhorst@berkeley.edu

Mark Hansen
114 McLaughlin Hall
University of California, Berkeley
Berkeley, CA 94720

Telephone: (510) 642-2880
Email: mhansen@ce.berkeley.edu

Abstract

This paper focuses on the occurrence of aircraft separation minimum violations as documented in the form of operational errors (OEs) at two types of air traffic controller facilities, terminal radar approach control (TRACON) facilities and air route traffic control centers (ARTCCs). Poisson regression was used to analyze the daily count of OEs at various facilities of both types. The occurrence of OEs was found to increase approximately with the square of daily traffic at TRACON facilities and slightly higher than the square of traffic at ARTCC facilities. At TRACON facilities, where separation violations are not automatically reported, an increase in reporting was seen after a new severity metric was introduced in 2007. It was also found that large, consolidated TRACON facilities tend to behave like a sum of several smaller facilities rather than a single larger facility with respect to the occurrence of OEs vs daily traffic. Weather effects such as visibility and wind were found to influence the occurrence of OEs as well. The model prediction for TRACON facilities is very good for the most severe OE types and very poor for the least severe OE types, indicating many unobserved factors contributing to the reporting of the least severe OE types in the terminal environment. Model prediction for the ARTCC facilities is very good for about half the facilities

Introduction

An operational error (OE) occurs when there is a violation of aircraft separation minimums due to air traffic control or from allowing an aircraft to enter another controller's airspace without notification. Fortunately, these events are infrequent, but each one represents a serious safety risk.

The factors that lie behind the occurrence of these extremely infrequent yet very serious safety risks is of major importance, and significant work has been done to investigate them. Due to the infrequent nature of these events, specific models that handle large sets of sparse data are typically used. One of the most common among these is Poisson regression. Of particular interest to us are the effects of traffic, weather, and policy changes on the rate of operational error occurrences at terminal radar approach control (TRACON) and air route traffic control center (ARTCC) facilities. In this paper we first review some of the studies involving operational errors and then analyze operational errors separately at TRACON and ARTCC facilities using Poisson regression. Conclusions and recommendations for future work will follow.

Background

Although the occurrence of operational errors has been studied extensively, most studies have used aggregation across time or space as a metric for OE occurrence. Hansen and Zhang [1] modeled the daily count of operational errors at all TRACONs with negative binomial regression, Gosling [2] studied the occurrence of OEs at separate Area Route Traffic Control Center (ARTCC) facilities aggregated over an entire year, and Panagiotakopoulos et al [3] modeled the rate of OEs per month at a single facility using extreme value theory. We propose to use the daily count of operational errors at specific facilities between 2004 and 2009. While a daily count at a specific facility is still aggregated across the entire day, it should present insights into the occurrence of OEs that is not seen with more aggregate data.

One of the primary questions of interest that motivated this research is the relationship between OE occurrence and traffic. It has been suggested that the relationship should be roughly quadratic (# OEs ~ traffic^2) because the number of possible path intersections roughly increases with the square of the number of aircraft in a sector. Murphy [4] has shown that the exponent for traffic should be at least 2 for en route facilities, with significant differences across centers. These results were found by using the number of aircraft in the sector at the time each OE occurred as the measure for traffic. Hansen and Zhang [1] have shown that the value of the exponent is around 1.7 for daily operations at all TRACON facilities.

Secondly, the impact of weather on the occurrence of OEs is important. Many authors have included weather effects through an overall subjective metric called traffic complexity, which has been shown to be a significant factor contributing to OEs [3]. Rodgers and Nye [5] found that once the number of operations was accounted for, air traffic complexity was a significant predictor of the total number of operational errors. Air traffic complexity is partially a subjective measure, but is a function of the variety of operations, airspace limitations, and weather. Weather variables such as wind, visibility, and temperature can represent a portion of traffic complexity that could give rise to operational errors.

Finally, another factor of interest is the effect of policy changes on OE reporting. At TRACON facilities, no automated tool is currently universally used to detect losses of separation like the Operational Error Detection Patch is used in the en route environment [6]. As a result, no matter how perceptive the controllers are, some errors will go unreported. The specific policy change that we investigated is the adoption of the Separation Conformance as a metric for OE severity. In June 25, 2007, an FAA order was sent out that specified a new measure of OE severity that would replace the OE Severity Index as the official measure. A key component of the new safety measure was the introduction of proximity events which would no longer be classified as operational errors, although they are still violations of separation minimums. Because they are no longer considered operational errors, controllers should not be penalized for reporting them as they would for a normal OE. Thus one would expect the number of reported errors in this category to increase.

The OE severity index was a range from 0-100 of many weighted factors, such as horizontal separation, closure rates, and converging / diverging paths. The separation conformance metric is much simpler, and relies only on horizontal and vertical separation retained at the closest point of proximity. Depending on the relative percentage of vertical and horizontal separation retained, the OE is classified into four categories, A, B, C and PE (proximity event), with A being the most severe type and a PE being a very minor breach of separation minimums.

Figure 1, shown below, illustrates the categories of separation conformance as a function of the horizontal and vertical separation retained at the point of minimum separation. Only the percentages of the minimum separation requirements matter, so the metric is the same regardless of the magnitude of the separation required in each direction.

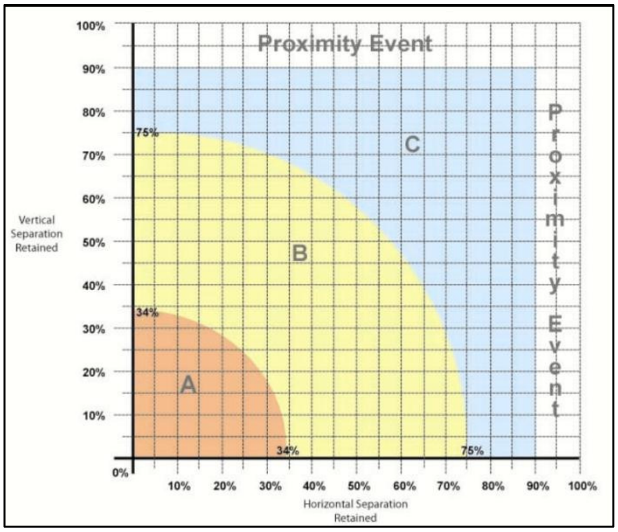

Figure 1. Separation Conformance Severity Metric (Source: FAA)

TRACON Analysis

The OE data used for this study is the daily count of errors at each of the largest 16 stand-alone TRACON facilities from October 1, 2004 to June 28, 2009 (see Table 1, below), resulting in a total of 27,710 TRACON-days. A total of 1,798 operational errors were observed over this time period, which is roughly a rate of 0.06 OEs/TRACON-day, or 1 OE/day for all TRACONs studied here. For purposes of comparison between the time period before and after the adoption of the separation conformance metric, all OEs will be classified as A, B, C, and PEs using the separation conformance metric even if the OEs occurred before the measure went into effect.

TABLE 1 TRACON Facilities

TRACON Facility	Primary City
N90	New York
D10	Dallas
A80	Atlanta
C90	Chicago
PCT	Washington D.C.
SCT	Los Angeles
D21	Detroit
I90	Houston
NCT	San Francisco
D01	Denver
L30	Las Vegas
P50	Phoenix
S46	Seattle
M98	Minneapolis
S56	Salt Lake City
A90	Boston

TRACON Model

Poisson regression is a common model that is used to study count data, and can be used for sparse data like we observe with operational errors. One observation in our model will be defined as the number of OEs at a particular TRACON on a particular day. Thus, each facility will have many different observations for each day in our time period of study, which we will assume are independent of each other. This type of model will allow us to capture longitudinal changes in OE occurrence as well as cross-sectional variation across facilities.

For our model, we assume that the occurrence of daily operational errors at any TRACON follows the Poisson distribution. The probability of observing a specific number of OEs for a given facility and date is shown by the following equation:

$$P(y_i = Y_i) = \frac{e^{-\lambda_i}{\lambda_i}^{Y_i}}{Y_i!} \quad (1)$$

where: y_i = OE number at date and facility I
Y_i = Observed OE number at date and facility i
λ_i = Average number of OEs to be expected at date and facility i

We will model the expected number of OEs, λ_i, with a logarithm link function of our explanatory variables:

$$ln(\lambda_i) = \alpha + \sum_j \beta_j \, x_{ji} \quad (2)$$

where: x_{ji} = explanatory variable j for date / facility i
β_j = coefficient for explanatory variable j
α = The model intercept

A list of our explanatory variables is given below in Table 2.

TABLE 2 TRACON Model Variable Descriptions

Explanatory Variable	TRACON Model – Variable Description
Log_Traffic	Natural logarithm of daily TRACON operations
Percent_Itinerant	Percent of TRACON operations that are itinerant
IMC	Percent of the time daily operations are in IMC conditions
TempF	Average daily temperature (deg F)
WindSpd	Average daily wind speed (knots)
Vis	Average daily visibility (miles)
N90_Audit	Dummy for N90 TRACON during 2005 audit
Old_Rule	Dummy for dates prior to June 25, 2007
Spring	Dummy for March, April, May
Summer	Dummy for June, July, August
Fall	Dummy for September, October, November

Previous research has investigated the quantity and type of operations and how they influence the occurrence of operational errors. We use two traffic measures to model these metrics: the natural logarithm of the daily traffic and the percentage of daily traffic that is listed as itinerant. The natural log of traffic is used so the coefficient obtained will represent the elasticity between OE rate and traffic. Itinerant operations are flights that are departing or arriving to an airport within the TRACON facility being observed, rather than being a through flight, which originates and exits the observed TRACON without landing. The percent of itinerant operations is used because we assume that controlling non-itinerant traffic is fundamentally different than controlling aircraft arriving or departing from airports within the area. The percentage of flights that are designated as itinerant is a proxy for the complexity of traffic in the sector, due to the complicated trajectories of departing and arriving flights. The source of the traffic data is the OPSNET database within ASPM.

Another contributor to traffic complexity is the weather in the sector on the day the operational error occurred, which we obtained from ASPM as well. It would be difficult to quantify the weather over the airspace of the entire TRACON, so the largest airport within the facility's airspace was used as a proxy for the entire region. For example, the weather at JFK was used to represent the New York TRACON and the weather at SFO was used to represent the weather at the Northern California TRACON. Although this is a somewhat crude measurement, the airspace around TRACONs are small with low flight levels, so the airport weather measurements will be used as a first approximation for TRACON weather.

The weather variables obtained were hourly measures of temperature, wind speed, visibility, and whether the operations were in IMC or VMC. Because hourly weather observations are obtained, each metric was averaged across a day for each facility, weighted by the hourly operations at the airport of interest. Weighing the weather measurements by the hourly operations will give us average measures that are representative of the weather an average flight departing or arriving at the airport of interest will experience. Although this is not entirely indicative of all traffic in the TRACON, the primary airport within each facility handles a majority of the facility's traffic.

The dummy variable N90_Audit was equal to 1 if the observation was at the New York TRACON during the time period of the 45 day audit in 2005. The reason this is included is that the audit revealed a very large number of unreported OEs during this time period that is not representative of the same reporting behavior at other facilities. Seasonal dummy variables were included to capture the variation across seasons. The Old_Rule dummy is a measure of policy changes in the system. We set this variable equal to 1 for all time periods prior to June 25, 2007, when the separation conformance measure went into effect.

TRACON Results

Three models were run for counts of different groups of OEs. The first model used the counts of all OEs, including proximity events. The second model did not include the proximity events (only A, B, and C errors), and the third model uses only the two most severe error types, A and Bs. The results from all three models are presented in 32 below.

TABLE 3 TRACON Model Regression Estimates

	All OEs Model			A,B&C Model			A&B Model		
Variable	Estimate		Std. Error	Estimate		Std. Error	Estimate		Std. Error
Intercept	-15.9	**	1.13	-16.7	**	1.29	-16.3	**	1.95
Log_Traffic	1.31	**	0.06	1.35	**	0.07	1.48	**	0.10
Percent_Itinerant	2.54	**	0.93	2.66	*	1.05	0.90		1.54
IMC	0.80	**	0.11	0.73	**	0.12	0.20		0.21
TempF	0.0015		0.002	0.011	**	0.002	0.009	*	0.004
WindSpd	0.056	**	0.006	0.058	**	0.007	0.04	**	0.01
Vis	-0.12	**	0.02	-0.11	**	0.02	-0.16	**	0.03
N90_Audit	3.81	**	0.11	3.76	**	0.12	3.13	**	0.22
Old_Rule	-0.57	**	0.05	-0.43	**	0.06	-0.27	**	0.09
Spring	-0.12		0.07	-0.07		0.08	0.22		0.14
Summer	-0.32	**	0.09	-0.16		0.11	0.12		0.18
Fall	-0.06		0.08	0.02	**	0.09	0.32	*	0.15

** Significant at 1% level
* Significant at 5% level

The first thing to notice is that the Log_Traffic coefficient is less than 2 and highly significant for all three models. It ranges from 1.31 for all OEs up to 1.48 for A & B errors. This suggests that the most severe errors are more sensitive to traffic than the less severe errors. The second thing to notice is that the Old_Rule coefficient is negative and significant for all three cases, indicating that all types of OEs have increased after the Separation Conformance metric went into effect. The seasonal dummy variables do not reveal any obvious trends, as most of the variables have large standard errors.

Much of the differences in seasonal effects are likely captured in the weather variables. The weather variables that are significant in each model include temperature, wind speed, and visibility. Increasing wind speed and decreasing visibility are both likely to increase the number of OEs by creating a more complicated airspace to navigate. Temperature, which can be an indicator of overall good weather, has a positive sign, which indicates higher temperature increases the occurrence of all types of OEs. The IMC variable is positive, as expected, but only significant for the models including the least severe types of OEs.

All three traffic coefficients are lower than we expected based on intuition and previous work. One consideration that we left out of the first set of models was the distinction between the consolidated TRACON facilities and non-consolidated ones. Consolidated TRACONs effectively function as a group of smaller TRACONs that are located in the same building. The operational characteristics of these facilities differs enough from the smaller TRACONs that our model is not capturing the true effect of traffic on the occurrence of OEs.

To illustrate this concept, imagine that the number of OEs is proportional to the square of traffic at all facilities. If you double the traffic at any facility, the number of OEs at that facility should increase by a factor of 4. Assume we have two identical facilities, with the same traffic and number of OEs. If we combine these two facilities into one, we will now have double the traffic, but only double the number of OEs, which is not consistent with our assumption that the number of OEs rises with the square of traffic at each facility. Thus, if the consolidated TRACONs are actually behaving like the sum of two or more smaller TRACONs, the quadratic behavior that we believe exists is being masked by linearly combining the traffic and OEs at each sub-TRACON facility.

To categorize the TRACONs into consolidated and stand-alone facilities, we used the definition of a consolidated TRACON from ASPM. These facilities provide approach control for two or more large hub airports where no single airport accounts for more than 60 percent of the total TRACON traffic count. This metric fits for four different facilities: Southern California (SCT), Northern California (NCT), New York (N90), and Potomac (PCT) TRACONs. To correct for this difference across facility types, we included a dummy variable for the consolidated TRACONs, and recalculated the coefficient estimates. The results are shown below in Table 4.

Interestingly, the traffic coefficient is very close to 2 for each of the models, and is highly significant. This suggests that the occurrence of operational errors of all severity levels roughly increases with the square of traffic, all else equal. The Consolidated dummy variable is negative and highly significant for all three models, suggesting that these facilities have fewer OEs than the other facilities, all else equal. This is consistent with our thought experiment about linearly combining facilities where OEs increase with the square of traffic.

Another interesting change in this model is the sign of the Percent_Itinerant variable, which is now negative. The change is likely due to lower average percentage of operations that are itinerant for traffic at consolidated TRACONs compared with the stand-alone TRACONs. The lower percentage of itinerant operations at consolidated TRACON facilities is another reason to treat them separately from the stand-alone facilities.

TABLE 4 TRACON Model Regression Estimates

	All OEs Model			**A,B,&C Model**			**A&B Model**		
Variable	**Estimate**		**Std. Error**	**Estimate**		**Std. Error**	**Estimate**		**Std. Error**
Intercept	-17.0	**	1.05	-17.8	**	1.23	-17.5	**	1.90
Log_Traffic	2.04	**	0.08	2.00	**	0.10	1.95	**	0.16
Percent_Itinerant	-1.75	*	0.88	-1.00		1.04	-1.40		1.55
IMC	0.90	**	0.11	0.82	**	0.13	0.27		0.21
TempF	0.0011		0.002	0.008	**	0.002	0.006		0.004
WindSpd	0.055	**	0.006	0.057	**	0.007	0.04	**	0.01
Vis	-0.13	**	0.02	-0.12	**	0.02	-0.17	**	0.03
N90_Audit	3.83	**	0.11	3.77	**	0.12	3.13	**	0.22
Old_Rule	-0.64	**	0.05	-0.49	**	0.06	-0.31	**	0.09
Spring	-0.11		0.07	-0.06		0.08	0.23		0.14
Summer	-0.29	**	0.09	-0.14		0.11	0.13		0.18
Fall	-0.05		0.08	0.02		0.09	0.32	*	0.15
Consolidated	-0.98	**	0.08	-0.85	**	0.09	-0.60	**	0.15

** Significant at 1% level
* Significant at 5% level

TRACON Model Fit

Two common measures of goodness-of-fit for Poisson regression models are the deviance and the Pearson Chi-Square statistics. These statistics provide a rough estimate if the assumption in the Poisson model of equal mean and variance is valid. If the deviance and the Pearson statistics divided by the degrees of freedom in the model are both close to 1, then the Poisson model assumption is generally accepted. If these statistics are greater than 1, this is an indication that the model is over-dispersed (e.g. the variance is actually greater than the mean) and the Poisson model is not valid. Typically in these situations a more general model, such as Negative Binomial regression is used.

The other case, where the statistics divided by the degrees of freedom are less than 1, indicating under-dispersion, is not as commonly seen and as a result fewer methods have been developed to deal with these situations. The deviance statistics divided by the number of degrees of freedom for our models range from 0.330 for the All OE model to 0.137 for the A&B model. The results for the Pearson statistics are 1.33 for the All OE model and 1.07 for the A&B model. The low deviance numbers suggest a poor fit due to under-dispersion but the Pearson numbers suggests a good fit. Using these two numbers alone is not enough evidence to suggest a fit or lack of fit due to under-dispersion.

Boyle and Flowerdew [7] have shown that using Poisson regression on very sparse data sets can lead to low deviance values due to the lack of asymptotic converge of the deviance statistic to the Chi-Square distribution. A simulation method has been developed to determine if the low deviance is a proper indicator of lack of fit due to under-dispersion or is simply a result of very sparse data [8]. The simulation begins with using the fitted values from the original model as the means of a set of Poisson random variables that represent the true distribution of OE occurrences. These Poisson random variables are then used to create a new set of observed values by taking a random draw from each Poisson random variable for each observation. For each new set of observations, we run the same model and calculate the new deviance. If the model is a proper fit for the data, then the mean of these simulated deviances will be close to the original deviance.

For the first model using all OEs, the simulated deviance mean is 8309 with a standard deviation of about 200. The actual deviance of 9147 is much larger than the simulated mean, thus indicating true under-dispersion. The A,B & C model has a very similar distribution to the All OEs model, indicating under-dispersion and a poor fit as well. The model for A&B OEs has a true deviance of 3797 with a simulated mean and standard deviation of 3720 and 185, respectively. The similarity between the simulated mean and the actual deviance suggests that the low deviance value arose simply due to highly sparse data and is not an indication of under-dispersion. A lack of fit in the first two models does not necessarily rule out the legitimacy of the parameter estimates, however.

TRACON Prediction

The final models were used to predict the number of OEs at each facility over the time period studied. For each facility we wanted to test the null hypothesis that the observed data were produced by a distribution defined by the results from our model. If we reject the null, we can conclude that model is not a good fit for the true process. A common method of evaluating this goodness-of-fit is to use Pearson's Chi-Square statistic, shown by the following equation:

$$TS = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{3}$$

Where k = number of categories

O_i = number of observations in category i

E_i = expected number in category i

The test statistic is distributed Chi-Squared for large sample sizes. Also, the expected number of counts in each category should be larger than 5. However, the Chi-Square distribution assumption can be invalid when any category has a much larger observed count than expected count, which we have in many of our facility predictions. Rather than use Pearson's statistic, we will use the G-Test for goodness of fit. The G-Test statistic is based on likelihood-ratio and is approximately Chi-Square distributed with k-1 degrees of freedom. The equation for calculation of the G-Statistic is shown below.

$$G = 2\sum_{i=1}^{k} O_i \cdot ln\left({O_i}/{E_i} \right) \tag{4}$$

The number of categories for some of our facilities is very small (2 or 3), so we will use simulation to determine the exact p-value of the G-Test. This method is common when the asymptotic behavior of the test statistic is in question. The method followed is to assume that the null hypothesis is true and draw a new set of observed data using the predicted results as the true distribution. A new G-statistic is calculated and the compared to the original G-statistic. This process is repeated 10,000 times for each facility, and the p-value for the G-Test is the percentage of times the simulated G-statistic is greater than the original G-statistic. The simulation results for each facility are shown in Table 5.

The p-values in the table above correspond to the null hypothesis that the model results represent the real distribution the observed data came from. Thus, a good model fit will have a large p-value in this table, because we will not be able to reject the null hypothesis at a high level of significance. Lack of rejection of the null is not the same as accepting the null, and thus we must be careful when interpreting these values as acceptance of a good model fit. For the very large p-values shown above, however, these at least suggest that the model fit is adequate. Notice that the model fit for the severe errors (A & B model) is much better than for all the errors together (All OE Model). Perhaps many of the less severe errors are caused by variables we neglected to include, or are somehow impossible to accurately model. The weather and operational measures included in our model appear to predict well the most severe errors at most of the facilities. The prediction is very poor when looking at the aggregate, however.

TABLE 5 TRACON Facility Prediction Results

	G-Test P-Values			
TRACON Facility	**All OEs Model**	**A,B, & C Model**	**A & B Model**	**Associated Airport(s)**
N90	0.001	0.014	0.003	LGA / JFK
D10	0.000	0.000	0.961*	DFW
A80	0.000	0.000	0.019	ATL
C90	0.000	0.000	0.419*	ORD
PCT	0.000	0.000	0.000	IAD / BWI
SCT	0.000	0.000	0.092*	LAX / SAN
D21	0.000	0.000	0.195*	DTW
I90	0.000	0.000	1.000*	IAH
NCT	0.005	0.029	0.808*	SFO / OAK
D01	0.009	0.750*	0.408*	DEN
L30	0.002	0.041	0.776*	LAS
P50	0.008	0.008	0.037	PHX
S46	0.037	0.324*	0.602*	SEA
M98	0.001	0.011	0.381*	MSP
S56	0.150*	0.056*	0.440*	SLC
A90	0.000	0.000	0.002	BOS
All Facilities	0.000	0.000	0.000	

* Not significant at 5% level

ARTCC Analysis

ARTCC Model

We also investigated the occurrence of operational errors at air route traffic control centers (ARTCCs). Total daily operations in each center were used as the measure for traffic, just like the TRACON model. However, the weather variables had to be introduced in a different way from the TRACONs due to the large size of the ARTCC airspaces. Weather data was taken from NOAA surface stations [9] and aggregated for each day across each facility's airspace. The number of stations in each ARTCC ranged from 50 to over 700 depending on the facility and the day. We will assume that the sample size of these is large enough to get a rough estimate of the weather conditions across the entire airspace. The weather variables used in this model were average visibility, average temperature, and the percentage of stations within each facility that experienced fog and rain for each day. Yearly dummy variables are used instead of an indicator of the change in reporting policy because the ARTCC facilities do not rely on self-reporting the way TRACONs do. Seasonal dummy variables are included as well.

TABLE 6 ARTCC Model Explanatory Variables

Explanatory Variable	**ARTCC Model – Variable Description**
Log_Traffic	Natural logarithm of daily ARTCC operations
Temperature	Average daily temperature (deg F)
Visibility	Average daily visibility (miles)
Fog	Percentage of stations experiencing fog
Rain	Percentage of stations experiencing rain
Y2004	Dummy for the year 2004
Y2005	Dummy for the year 2005
Y2006	Dummy for the year 2006
Y2007	Dummy for the year 2007
Y2008	Dummy for the year 2008
Spring	Dummy for March, April, May
Summer	Dummy for June, July, August
Fall	Dummy for September, October, November

ARTCC Results

The model was run for the three categories of OEs, just like for the TRACONs (All OEs, A, B & Cs, and A & Bs). The results for all three models are shown below in Table 7.

TABLE 7 ARTCC Models Regression Estimates

	All OEs Model			A, B, & C Model			A & B Model		
Variable	**Estimate**		**Std. Error**	**Estimate**		**Std. Error**	**Estimate**		**Std. Error**
Intercept	-25.4	**	0.81	-26.7	**	1.01	-28.7	**	2.08
Log_Traffic	2.53	**	0.08	2.59	**	0.11	2.62	**	0.22
Temperature	0.004	*	0.002	0.006	**	0.002	0.015	**	0.004
Visibility	0.054	**	0.017	0.067	**	0.021	0.088	*	0.042
Fog	0.97	**	0.23	1.20	**	0.28	1.23	*	0.56
Rain	0.41	**	0.08	0.46	**	0.10	0.31		0.20
Y2004	-0.07		0.11	0.12		0.13	-0.30		0.26
Y2005	-0.30	**	0.08	-0.10		0.10	-0.15		0.18
Y2006	-0.19	*	0.08	-0.20	*	0.10	-0.53	**	0.19
Y2007	-0.14	**	0.08	-0.10		0.10	-0.34		0.19
Y2008	-0.06		0.08	-0.05		0.10	-0.32		0.19
Spring	-0.03		0.06	0.01		0.07	-0.14		0.14
Summer	-0.01		0.07	-0.002		0.09	-0.44	*	0.18
Fall	-0.01		0.06	0.01		0.08	-0.21		0.15

** Significant at 1% level
* Significant at 5% level

The coefficient for the Log_Traffic variable is between 2.5 and 2.6, depending on the error severity used, and is highly significant in each. This is higher than was seen in the case of the TRACONs, even when accounting for the operational differences between facility types. Our model results suggest that in the en route airspace more traffic, all else equal, increases the rate of OE occurrence faster than in the TRACON airspace. This could be an indicator of differences in airspace complexity caused by an increase in traffic.

Also, note the significant effects of weather on OEs. The coefficient estimates for temperature and visibility are consistent with the results found in the TRACON models. Fog and rain both appear to increase the number of OEs, which is to be expected.

ARTCC Model Fit

The problem of low deviance is prevalent in the ARTCC model as well. The values for deviance / degrees of freedom range from 0.386 for the All OE model to 0.107 for the A&B model. The same simulation was performed as with the TRACON model to determine if the low deviance represents a potential problem of under-dispersion or if it arises simply due to sparse data. The simulated mean for each of the three models is very close to the actual mean. This indicates that the low model deviance is likely due to highly sparse data, rather than any under-dispersion.

ARTCC Prediction

We also predicted the number of OEs at each ARTCC facility and performed the same goodness-of-fit tests. The prediction results are shown below in Table 8.

The model predictions are better on a per-facility basis for the en route centers. Over half of the facilities have a very large p-value, indicating a lack of evidence to reject the hypothesis that the model is a good fit for the data. Note that the spread across facilities is very large. For some facilities, such as ZDV, the model does a very good job predicting the occurrence of all OE types, while for others, such as ZID, the model does a poor job. Unlike the TRACON prediction results, the p-values for the ARTCC facilities appear to be consistent across all OE types. That is, the ARTCC model does just as good, or poor, of a job predicting serious OEs for any given facility as is does predicting the less severe errors for the same facility. This is likely because of the high variability in reporting minor errors at TRACON facilities due to self-reporting.

TABLE 8 ARTCC Facility Prediction Results

	G-Test P-Values		
ARTCC Facility	**All OEs Model**	**A,B,&C Model**	**A&B Model**
ZAB	0.834*	0.839*	0.838*
ZAN	0.154*	0.154*	0.152*
ZAU	0.002	0.001	0.001
ZBW	0.000	0.000	0.000
ZDC	0.000	0.000	0.000
ZDV	0.939*	0.940*	0.938*
ZFW	0.000	0.001	0.000
ZHU	0.000	0.000	0.000
ZID	0.000	0.000	0.001
ZIX	0.019	0.020	0.019
ZKC	0.089*	0.087*	0.088*
ZLA	0.117*	0.123*	0.119*
ZLC	0.262*	0.258*	0.251*
ZMA	0.099*	0.092*	0.093*
ZME	0.685*	0.676*	0.675*
ZMP	0.000	0.000	0.000
ZNY	0.120*	0.120*	0.126*
ZOA	0.520*	0.510*	0.514*
ZOB	0.002	0.002	0.002
ZSE	0.251*	0.243*	0.247*
ZTL	0.035	0.033	0.034
All Facilities	0.000	0.000	0.000

* Not significant at 5% level

OE Severity Logit Model

Finally, we wanted to investigate how operating conditions and weather influence the severity of a given operational error, rather than simply exploring the occurrence of the errors themselves. In the highway accident literature [10], a logit model has been used to determine the relationship between the severity of an accident and various road and operating conditions at the time of the accident. A similar analysis will be performed here. In our case, we have grouped the OE severity levels into three categories: Proximity Event, C, and A&B. One observation was defined as a single OE-TRACON-day, and the explanatory variables used in the Poission Regression models earlier in the paper were recorded. Note that the vast majority of the days did not have any operational errors, which we are not analyzing in this section. We are developing a model to explain the severity of the operational error, conditioning on the operational error occurring. For the case where a single day at a TRACON has multiple OEs, the operating conditions were repeated into multiple observations. We realize these errors are likely correlated, but for now we will assume that each separate error severity allocation is independent. A multinomial logit model was estimated for the data at TRACONs and at ARTCCs. The alternative used as the base is the Proximity Event alternative. The estimation results are shown below in Table 9 and Table 10.

TABLE 9 TRACON Logit Model Results

	Variable	Estimate		Std. Error
Category C OEs	ASC	-3.73		2.87
	Traffic (Daily Operations)	-0.11		0.22
	Percent Itinerant	4.74		2.45
	Percent IMC	0.048	*	0.27
	Average Temp (F)	-0.013		0.005
	Average Wind Speed (knots)	0.024	*	0.016
	Average Visibility (miles)	0.074		0.042
	N90 Audit	0.33		0.29
	Old Rule	0.47		0.14
	Spring	0.15	**	0.18
	Summer	0.56		0.24
	Fall	0.22	*	0.20
	Consolidated TRACON	0.31		0.21
Category A&B OEs	ASC	-3.61		3.09
	Traffic (Daily Operations)	-0.17		0.24
	Percent Itinerant	5.48	*	2.59
	Percent IMC	-0.86	**	0.31
	Average Temp (F)	-0.015	*	0.006
	Average Wind Speed (knots)	-0.013		0.018
	Average Visibility (miles)	0.003		0.048
	N90 Audit	-0.54		0.33
	Old Rule	0.68	**	0.15
	Spring	0.58	**	0.21
	Summer	0.89	**	0.28
	Fall	0.67	**	0.22
	Consolidated TRACON	0.73	**	0.23

** Significant at 1% level
* Significant at 5% level

Some things to note from these results are that the traffic coefficient for both alternatives are not significant, which means that the number of operations does not appear to influence the severity of an OE, holding all else constant. This is consistent with the results from the previous models, where the traffic coefficient was roughly the same for each of the three models, indicating that all errors increase with traffic at the roughly the same rate.

The percent itinerant variable is positive and significant for the A&B alternative. This implies that a higher percentage of itinerant (vs through) operations more likely results in a category A or B error rather than a proximity event or C error. This follows our intuition, because itinerant operations are a fundamentally different type of traffic compared to through operations and are likely more difficult to control. Another interesting result is that the IMC variable is negative and significant for the A&B alternative, but in our Poission Regression model (Table 3), we found that the IMC variable was not significant in the A&B model. This suggests that better weather (more time operating in IMC) does not influence the occurrence of category A&B operational errors. However, given an OE occurs, if it occurs in poor weather (higher percentage IMC conditions), then it is less likely to be a severe error. This might be somewhat counter-intuitive, but one possible explanation for this is that controllers may be more alert during times of poor weather, and thus any errors that occur are likely to be minor in severity.

The Old Rule variables are positive and significant, which is consistent with the increase in proximity events the separation conformance metric went into effect. The relative magnitude of these variables represent the relative composition of errors before the rule went into effect, which tended to have more severe errors than after the rule went into effect. The Consolidated TRACON variable has a coefficient that is positive and significant only for the A&B category, indicating that errors at these facilities tend to be more severe than the errors at smaller facilities, even after controlling for traffic, weather, and percentage of itinerant operations. This is likely an indication of the different traffic composition at these facilities, due to their large nature, but we did not investigate this explicitly in this study.

TABLE 10 ARTCC Logit Model Results

	Variable	Estimate		Std. Error
Category C OEs	ASC	-0.77		0.52
	Traffic (Daily Operations)	0.027		0.030
	Average Temp (F)	0.002		0.004
	Average Visibility (Miles)	0.04		0.04
	Percentage Stations with Fog	0.94	*	0.54
	Percentage Stations with Rain	0.20		0.18
	Year 2004	0.69	**	0.24
	Year 2005	0.65	**	0.18
	Year 2006	0.08		0.17
	Year 2007	0.17		0.17
	Year 2008	0.12		0.17
	Spring	0.10		0.13
	Summer	0.14		0.17
	Fall	0.11		0.14
Category A&B OEs	ASC	-2.54	**	0.72
	Traffic (Daily Operations)	0.04		0.04
	Average Temp (F)	0.02	**	0.005
	Average Visibility (Miles)	0.08		0.05
	Percentage Stations with Fog	1.14		0.72
	Percentage Stations with Rain	-0.04		0.25
	Year 2004	0.14		0.33
	Year 2005	0.54	**	0.23
	Year 2006	-0.40	*	0.23
	Year 2007	-0.17		0.23
	Year 2008	-0.25		0.22
	Spring	-0.08		0.18
	Summer	-0.47	*	0.23
	Fall	-0.19		0.19

** Significant at 1% level
* Significant at 5% level

The results from the ARTCC model indicate that number of daily operations does not affect the severity of a given OE, similar to what was shown in the TRACON results. Since the airspace regions for ARTCCs are much larger than those for TRACONs, we use average temperature is our best proxy for overall good weather, where higher temperatures indicate better weather, all else equal. The percentage stations reporting fog and rain are somewhat indicative of poor weather, but are not very significant in our model. The primary significant result is that the average temperature for the category A&B errors is positive with a low standard error. This indicates that higher temperature (good weather) will result in an error more likely being a severe category A or B error, rather than a category C error or proximity event. This possibly counter-intuitive result is consistent with the TRACON model. Controllers could be more alert during times of poor weather, and thus the errors that do occur are less likely to be severe.

Conclusions

Operational errors in terminal radar approach control (TRACON) an air route traffic control center (ARTCC) facilities were modeled using Poisson regression. The daily OE count at each facility was used as the dependent variable while operational and weather measures were used as the independent variables. The rate of daily OE occurrences at TRACON facilities was found to increase with the square of daily traffic, which is consistent with previous research and general intuition. It was also found that consolidated TRACON facilities behave effectively as a sum of several other, smaller TRACON facilities in terms of how the number of OEs is influenced by traffic. The rate of occurrence of all types of OEs at TRACONs has been found to increase after the introduction of the separation conformance severity metric.

Possible under-dispersion exists for the models of the two least severe OE types for the TRACON facilities. The A & B model has a very low deviance likely because of very sparse data. Our Poisson regression model provides a good fit for only one facility for all error types, but fits most of the facilities well for the most severe error types. That is, the most severe error types are the easiest to predict.

The ARTCC models indicate a traffic exponent between 2.5 and 2.6, which is higher than the value of 2 obtained for the TRACON models. This could be an indication of a different relationship between traffic and controller workload in the en route airspace. The way OEs are identified in the TRACON airspace is far different than the en route airspace, however, so there could be many unidentified factors that contribute to this difference. Weather variables such as fog and rain both appear to increase the occurrence of all types of OEs, although fog at a higher rate than rain.

The results from the two logit models suggest that the number of daily operations does not significantly affect the severity level of an OE. A higher percentage of itinerant operations at TRACON facilities imply a higher chance of an OE being severe (category A or B). A high percentage of itinerant operations is indicative of high traffic complexity. We might expect a high traffic complexity to increase the occurrence of OEs, but that is not the case from the Poission Regression results. Instead, we find that high traffic complexity increases the severity of an OE, given one occurs. For both TRACON and ARTCC facilities, we find that poor weather decreases the likelihood of an error being severe (category A or B). Poor weather conditions could be capturing higher controller awareness, which might influence the severity of errors more than simply the operating conditions.

Future Work

Future work could include using more information about each operational error than we included. Although our goal was to disaggregate our data as much as possible, we still have daily counts of operational errors, traffic, and average weather variables. Using the exact conditions in the sector at the time the OEs occurred could be more representative of the true causes of these rare events.

Additional models could be used, such as zero-inflated Poisson regression, to more accurately model the large amount of zeros in the data set. Our model could also be modified to somehow consider the exact way the consolidated TRACONs are acting like the sum of a number of smaller TRACONs. Detailed operational characteristics of these large consolidated facilities would be needed to perform this analysis, however. Also, other effects of policy changes could be investigated if different time periods were used. For example, the implementation of the Air Traffic Safety Action Program or the Traffic Analysis Review Program will both affect the way OEs are reported in the TRACON environment.

REFERENCES

[1] Hansen, M. and Y. Zhang. Safety of Efficiency: Link between Operational Performance and Operational Errors in the National Airspace System, *Transportation Research Record*, Vol. 1888, 2004, pp. 15-21.

[2] Gosling, G. D. Analysis of Factors Affecting the Occurrence and Severity of Air Traffic Control Operational Errors. *Transportation Research Record*, Vol.1788, 2003, pp. 49–57.

[3] Panagiotakopoulos, D., A. Majumdar, and W. Ochieng. Characterising the Distribution of Safety Occurrences in Aviation: An Approach Using Extreme Value Theory, *Trans portation Research Record* Vol. 2106, 2009, pp. 129–140.

[4] Murphy, D. and G. Shapiro. A Quantitative Model for En Route Error Rate Analysis, *ATM* 2007.

[5] Rodgers, M.D., & L.G. Nye. Factors Associated With the Severity of Operational Errors at Air Route Traffic Control Centers. In M.D. Rodgers, An Examination of the Operational Error Database for Air Route Traffic Control Centers, *DOT/FAA/AM-93/22,* 1993, pp. 11-25.

[6] Schroeder, D., L. Bailey, J. Pounds, and C. Manning. A Human Factors Review of the Operational Error Literature, *DOT/FAA/AM-06/21*, Civil Aerospace Medical Institute, Oklahoma City, OK, 2006.

[7] Boyle P J, Flowerdew R. Modelling Sparse Interaction Matrices: Interward Migration in Hereford and Worcester, and the Underdispersion Problem. *Environ Planning A* 1993; 25:1201–09.

[8] Boyle, P., R. Flowerdew, and A. Williams. Evaluating the Goodness of Fit in Models of Sparse Medical Data: A Simulation Approach. *International Journal of Epidemiology*. Vol. 26, No. 3., 1997.

[9] NOAA National Climatic Data Center. Global Surface Summary of Day. http://www.ncdc.noaa.gov/cgi-bin/res40.pl?page=gsod.html.2006

[10] Patil, S., S. R. Geedipally, D. Lord. Analysis of Crash Severities Using Nested Logit Model – Accounting for the Underreporting of Crashes. *Accident Analysis and Prevention*. Vol. 45 p.646-653.