

# **AIRLINE SAFETY DATA: HOW PREDICTABLE ARE ACCIDENTS AND FATALITIES?**

**Pamela Rogers**

*Stephen F. Austin State University*  
Nacogdoches, TX 75962-3070  
USA

**Robert Pavur**

*University of North Texas*  
Denton, TX 76203-5017  
USA

## **ABSTRACT**

Aviation safety data is essential to providing guidance for corrective actions. Alternative analyses of safety data is paramount in monitoring the reliability and accuracy of forecasts of aviation accidents to support effective safety strategies. The number of passengers flying and number of flights are currently taxing many airports, hastening the need for effective capacity planning. Understanding potential changes in accidents and fatalities can be useful in helping improve capacity decisions, such as the addition of runways. A reliable predictive model is needed for this purpose. Number of passenger fatalities and accidents are investigated using the Box-Jenkins methodology and artificial neural network models. With the attention given to machine learning techniques, the value of traditional statistical prediction models is called into question. Results reveal that traditional ARIMA models and neural network models have some predictive power in forecasting normalized safety data for a ten-year period using historical data starting in 1938.

**Keywords:** Predictive Analytics, Airline Safety, ARIMA, artificial neural network

## **INTRODUCTION**

Research in predicting airline safety has received increased attention as the civil aviation transportation industry receives immense attention for its safety record. Furthermore, prediction

methods using machine learning techniques, such as artificial neural networks (ANNs), are credited with higher accuracy and providing more confidence in promoting timely preventive measures (Huang, Zeng, Pei, Wong, & Xu, 2016; Altay, Ozkan, & Kayakutlu, 2014). Despite the rise in research using machine learning techniques, traditional statistical approaches often dominate the literature in modeling risk and accidents (Ocampo & Klaus, 2018). For example, autoregressive integrated moving average (ARIMA) models are used as competitor models to ANNs. Zhang (2003) recognized that a hybrid approach could be used to take advantage of the nonlinear capability of ANNs and the statistical properties of ARIMA. The ARIMA models have the advantage of relatively simple understanding and implementation along with strong goodness of fit tests. A major objective of this paper is to examine the performance of ANNs and ARIMA models using airline safety data to explore their capability of contributing to enhanced accuracy of airline safety predictions.

Airline safety comes at a cost and the dynamic prediction of safety in the airline industry provides decision-making information in addressing added resources to the implementation of aviation safety regulations and guidelines (Altay et al., 2014). For example, LaGuardia has been known to make requests to the Federal Aviation Administration (FAA) to increase the number of flights allowed for as much as a 60% increase. Across the country, city officials must seek permission from the FAA to build runways and expand airports. Approvals for these projects requires the examination of a variety of data, the most important being safety data (Murphy, 2001). While commercial aviation has lost its sense of graciousness over the years, deregulation of the industry has led to huge increases in the annual number of airline passengers. This trend is expected to continue, as witnessed by substantial forecasted growth for the nation's airports (Garcia, 2019).

Research on the airline industry has mostly focused on decision-making strategies involving passenger demand and allocation of aircraft and other resources, as well as selling the right seat to the right customer at the right time for the right price to maximize profit (Andersson, 2001). Until recent events, not as much attention has been focused on analyzing safety data collected by the FAA. A frequent statement is that it is safer to travel by air than by car. Perhaps, there is a general sense that the airline industry has made such great strides in safety, and that little is left to do. Researchers such as Kanafani and Keeler (1990) reported that through the 1980s, there was no evidence of any reduction in airline safety. In fact, if one considers that there are over 3,000 take-offs per day at some major airports, the safety record of the airline industry would be considered excellent.

Rose (1989) analyzed accidents per 100,000 departures and fatalities during the period from 1955 to 1986. This study, which included the peak accident years of 1959 and 1960 in which there were 67 accidents per year, found a significant declining trend in accidents (FAA Statistical Handbook of Aviation, 2000). However, the Deregulation Act was passed in 1978 and therefore Rose's study includes only 8 years of data after deregulation. Opponents to airline deregulation predicted there would be an increase in accidents because of the financial losses due to lower airfares created by deregulation (Salpukas, 1982). Additionally it was suggested that deregulation would allow numerous small airlines to operate that lacked the knowledge and skills found in larger airlines to fully maintain aircraft.

A more difficult task than simply determining trends in safety is the modeling of accidents and fatalities in the airline industry. One approach to modeling the number of accidents occurring over time is to assume that accidents follow a Poisson process (Ross, 1993). This approach would thus assume that the number of accidents are occurring randomly and could be modeled using Markov stochastic processes. However, Foreman (1993) achieved good model fits with

autoregressive models using safety variables such as fatalities per million miles flown and accidents per million miles and used data through 1990. He did not compare his approach to forecasting methods using machine learning.

This paper examines the use of the Box-Jenkins methodology to model airline safety data and compares its performance to the predictive ability of artificial neural network (ANN) models. The recommended number of observations for the Box-Jenkins methodology is a minimum of 50 up to 100 observations (Bowerman, O'Connell, & Koehler, 2005). To examine ARIMA and ANN models with data that is not impacted by a terrorist crisis, data was limited to the time frame before 2000. The data used in this study was limited to data prior to September 11, 2001 as the aviation industry has responded purposefully to increased aviation standards after that date.

This paper identifies ARIMA models that provide an adequate fit to airline safety data and assesses the predictive ability of these models and models developed using neural network techniques with a five year and ten year holdout data set. The importance of these results is to illustrate that airline safety data do not follow a random process and can be effectively modeled with ARIMA and ANN models. Accidents and fatalities appear to be random events, but this research illustrates that Box-Jenkins and neural network models can have very good predictive power with this data. These models may promote the use of these techniques as decision making tools in validating justification for aviation policy as well as for promoting plans to increase airport capacity with additional runways.

### **AIRLINE SAFETY DATA**

While flying remains remarkably safe, it is not good news that each year there are numerous incidences in which extreme action is taken to avoid close collisions (Murphy, 2001). Each year the United States Department of Transportation releases airline safety data in the *FAA Statistical Handbook*. A look at the number of accidents per year since 1938 gives some insight into the history of aviation safety. Figure 1 presents these data. Visually, the data appear to have a downward trend from 1960 to the mid-1980s. But that decline does not appear to be continuing in the 1990s.

**Figure 1. Number of airline accidents from 1938 to 1999.**

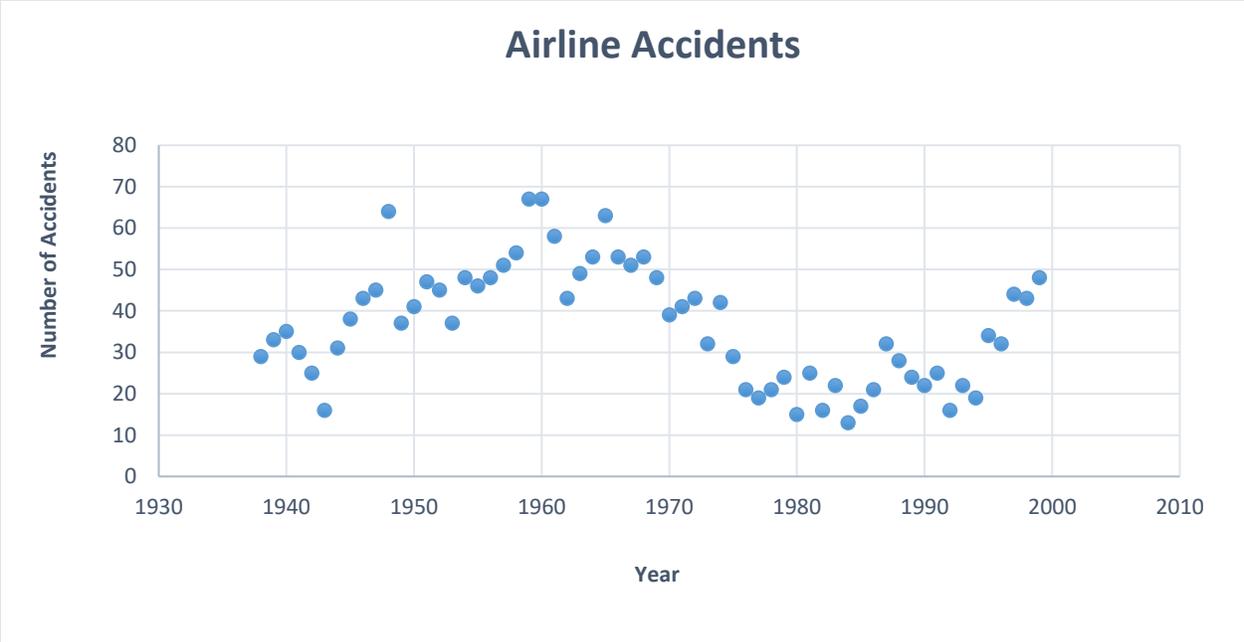


Figure 1 may not accurately portray the airline industry's safety record. Since the number of passengers and the number of miles flown by passengers has more than doubled from 1980 to 1999, the safety record is indeed impressive (Murphy, 2001). To normalize the variables used in this study, we divide the variables by either millions of departures or millions of miles flown. The variables selected for this study are listed in Table 1. If these variables are plotted over time, a declining trend is obvious. Figure 2 provides an example of how one of these normalized variables performs over time. Most of the declining trend appears to have occurred from 1938 to the early 1980s. There were only three years when there were 0 passenger fatalities. Those years were 1980, 1993, and 1998.

**Table 1. Variables used in forecasting safety data.**

| Variable Name | Description of Variable  |
|---------------|--|
| PASSFMDEP     | Number of passenger fatalities yearly divided by millions of departures.   |
| PASSFMMF      | Number of passengers fatalities yearly divided by millions of miles flown. |
| FACCMDEP      | Number of fatal accidents yearly divided by millions of departures.        |
| FACCMMF       | Number of fatal accidents yearly divided by millions of miles flown.       |
| ACCMMF        | Number of accidents yearly divided by millions of miles flown.             |

**Figure 2. The number of passenger fatalities per million miles of passenger travel.**



### **TIME SERIES MODELS FOR AIRLINE SAFETY DATA**

To assess the predictive power of both the Box-Jenkins and neural network models, a holdout of 10 observations from 1990 to 1999, for each variable, is used in determining the predictive error. For each variable considered, the Box-Jenkins model that is the best fit for the data from 1938 to 1989 is used. These same years are used to determine an acceptable ANN model. The root mean squared error is computed for the years 1990 through 1994 and for the years 1990 through 1999 for both the Box-Jenkins model and the ANN model. That is, a five year and a ten year forecast is assessed for each model.

The Box-Jenkins methodology has been criticized for not having much better forecasting accuracy than simpler exponential smoothing techniques. Makridakis and Hibon (1997) examine data to determine why the post-sample forecasting for the Box-Jenkins models was sometimes worse than simpler models. Their paper illustrates that the simpler ARIMA models with only one or two autoregressive or moving average terms typically produce better results than more complicated models produced by the Box-Jenkins methodology. Therefore, ARIMA models in this study are developed to be as parsimonious as possible, with no more than three terms, and are based on the Box-Jenkins methodology.

The ANN models are composed of interconnected neurons linked together through weighted directed arcs and organized into layers. The ANN model used in this study is the commonly used backpropagation model using three layers. The lowest (or bottom) layer is called the input layer, and the last layer (or top) is the output layer. The middle layer is often termed the hidden layer. A sigmoid function is used in the neural network to create a non-linear model to fit the data. As training set data are presented to the input layer, the neural network adjusts the connection weights until the error between the expected output and the network output is minimized. Through the connection weights the network tries to memorize the input. The backpropagation neural network uses the backpropagation of error during learning. These networks most often use a generalized-delta learning rule. This learning rule, generally attributed to Rumelhart, Hinton, and Williams (1986), is a nontrivial extension of the learning rule for a simple Adaline network (Widrow & Lehr, 1990).

According to Jain and Nag (1997), a backpropagation with one hidden layer of neurons can “closely” approximate any decision surface. The degree of closeness depends on the relationship between the architecture of the network and the decision surface being approximated. One important architectural decision involves the determination of the number of hidden-layer neurons. While increasing the number of hidden-layer neurons can lead to improved performance on the observations used to train a backpropagation, it may also lead to a lack of generalizability of the network (Jain & Nag, 1997). Because there are no theoretical guidelines for determining the appropriate number of hidden-layer neurons, a trial-and-error process has generally been used by researchers and practitioners to determine the best architecture.

**Model 1. Passenger Fatalities per Million Departures (PASSFMDEP).**

Because of the declining trend in PASSFMDEP, a first difference is used with the Box-Jenkins methodology. After elimination of the trend, the data are stationary and the first order autoregressive model with first and second moving average terms appears to be a reasonable fit. The residuals autocorrelation check reveals only small autocorrelations. The Box-Jung statistics all have large p-values with the exception of the third Chi-Square which has a p-value of .0491. This model is considered the best fit after examining several alternate models. The statistical analysis of this model is as follows.

**Table 2: Statistical Analysis of Model 1**

| Parameter | Estimate | Standard Error | t-Value | Pr >  t | Approximate Lag |
|-----------|----------|----------------|---------|---------|-----------------|
| MA1,1     | -0.26931 | 0.13207        | -2.04   | 0.0470  | 1               |
| MA1,2     | 0.67374  | 0.10687        | 6.30    | <.0001  | 2               |
| AR1,1     | -0.83142 | 0.12916        | -6.44   | <.0001  | 1               |

**Autocorrelation Check of Residuals**

| To Lag | Chi Square | DF | Pr > ChiSq | Autocorrelations |        |        |        |        |        |
|--------|------------|----|------------|------------------|--------|--------|--------|--------|--------|
| 6      | 4.31       | 3  | 0.2294     | -0.027           | -0.035 | -0.029 | 0.155  | 0.143  | -0.161 |
| 12     | 10.57      | 9  | 0.3065     | -0.174           | 0.102  | 0.219  | -0.036 | -0.091 | 0.015  |
| 18     | 25.06      | 15 | 0.0491     | 0.290            | 0.220  | 0.177  | -0.035 | 0.070  | 0.162  |
| 24     | 28.68      | 21 | 0.1220     | -0.083           | -0.054 | -0.093 | 0.094  | 0.107  | 0.027  |

The root mean squared error for the five year and ten year holdout data sets are 21.44 and 21.12, respectively. The backpropagation neural network model uses 9 inputs, 18 neurons in the hidden layer, and one output. This model yields a value of 17.07 and 18.65 for the five year and ten year holdout data sets, respectively. Thus, the neural network model performs considerably better than the Box-Jenkins model in predicting passenger fatalities per million departures (PASSFMDEP).

**Model 2. Passenger Fatalities per Million Miles Flown (PASSFMMF).**

Because of the declining trend in PASSFMMF, a first difference is used with the Box-Jenkins methodology. After elimination of the trend, the data are stationary and the first order autoregressive model with a second order moving average term appears to be a reasonable fit. The residuals autocorrelation check reveals very small autocorrelations. The Box-Jung statistics have

large p-values. This model is considered the best fit after examining several alternate models. The statistical analysis of this model is as follows.

**Table 3: Statistical Analysis of Model 2**

| <b>Conditional Least Squares Estimation</b> |                 |                       |                |                    |                        |
|---|-----------------|-----------------------|----------------|--------------------|------------------------|
| <b>Parameter</b>                            | <b>Estimate</b> | <b>Standard Error</b> | <b>t-Value</b> | <b>Pr &gt;  t </b> | <b>Approximate Lag</b> |
| MA1,1                                       | 0.93283         | 0.09378               | 9.95           | <.0001             | 2                      |
| AR1,1                                       | -0.45345        | 0.14309               | -3.38          | 0.0017             | 1                      |

| <b>Autocorrelation Check of Residuals</b> |                   |           |                      |                         |        |        |        |        |        |
|---|-------------------|-----------|----------------------|-------------------------|--------|--------|--------|--------|--------|
| <b>To Lag</b>                             | <b>Chi Square</b> | <b>DF</b> | <b>Pr &gt; ChiSq</b> | <b>Autocorrelations</b> |        |        |        |        |        |
| 6   | 3.88              | 4         | 0.4226               | -0.081                  | -0.084 | 0.118  | 0.003  | 0.231  | -0.041 |
| 12  | 5.79              | 10        | 0.8322               | -0.017                  | 0.099  | 0.146  | -0.058 | 0.027  | -0.007 |
| 18  | 11.11             | 16        | 0.8025               | -0.003                  | 0.281  | -0.021 | -0.017 | -0.016 | -0.042 |
| 24  | 15.99             | 22        | 0.8163               | 0.092                   | -0.086 | -0.129 | -0.022 | 0.140  | -0.037 |

The root mean squared error for the five year and ten year holdout data sets are .0312 and .0318, respectively. The backpropagation neural network model uses 9 inputs, 11 neurons in the hidden layer, and one output. This model yields a value of .0311 and .0307 for the five year and ten year holdout data sets, respectively. Thus, the neural network model performs slightly better than the Box-Jenkins model in predicting passenger fatalities per million miles flown (PASSFMMF) for the ten year data set and about same for the five year holdout data set.

**Model 3. Number of Fatal Accidents Yearly Divided by Millions of Departures (FACCMDEP).**

Because of the declining trend in FACCMDEP, a first difference is used with the Box-Jenkins methodology. After elimination of the trend, the data are stationary and the first order autoregressive model with the second order moving average term appears to be a reasonable fit. The residuals autocorrelation check reveals small autocorrelations. Two of the p-values for the Box-Jung statistics are small (near .05). However, this model is considered the best fit after examining several alternate models. The statistical analysis of this model is as follows.

**Table 4: Statistical Analysis of Model 3**

| <b>Conditional Least Squares Estimation</b> |                 |                       |                |                    |                        |
|---|-----------------|-----------------------|----------------|--------------------|------------------------|
| <b>Parameter</b>                            | <b>Estimate</b> | <b>Standard Error</b> | <b>t-Value</b> | <b>Pr &gt;  t </b> | <b>Approximate Lag</b> |
| MU  | -0.05922        | 0.01052               | -5.63          | <.0001             | 0                      |
| MA1,1                                       | 1.00000         | 0.15818               | 6.32           | <.0001             | 2                      |
| AR1,1                                       | -0.94470        | 0.15292               | -6.18          | <.0001             | 1                      |

| <b>Autocorrelation Check of Residuals</b> |                   |           |                      |                         |  |  |  |  |  |
|---|-------------------|-----------|----------------------|-------------------------|--|--|--|--|--|
| <b>To Lag</b>                             | <b>Chi Square</b> | <b>DF</b> | <b>Pr &gt; ChiSq</b> | <b>Autocorrelations</b> |  |  |  |  |  |

|    |       |    |        |        |        |        |        |        |        |
|----|-------|----|--------|--------|--------|--------|--------|--------|--------|
| 6  | 9.62  | 4  | 0.0474 | 0.060  | -0.230 | -0.169 | 0.295  | 0.058  | -0.180 |
| 12 | 17.34 | 10 | 0.0672 | -0.069 | 0.173  | 0.203  | -0.061 | -0.219 | -0.107 |
| 18 | 20.02 | 16 | 0.2191 | 0.129  | 0.114  | -0.089 | -0.008 | 0.057  | 0.023  |
| 24 | 23.01 | 22 | 0.4004 | -0.002 | -0.059 | 0.083  | 0.117  | 0.088  | -0.022 |

The root mean squared error for the five year and ten year holdout data sets are .7664 and .4634, respectively. Interestingly, the ten year prediction is better than the five year prediction for the Box-Jenkins model. The backpropagation neural network model uses 9 inputs, 18 neurons in the hidden layer, and one output. This model yields a value of .2325 and .2548 for the five year and ten year holdout data sets, respectively. Thus, the neural network model performs much better than the Box-Jenkins model in predicting the number of fatal accidents yearly divided by millions of departures (FACCMDEP).

#### **Model 4. Number of Fatal Accidents Yearly Divided by Millions of Miles Flown (FACCMMF).**

Because of the declining trend in FACCMMF, a first difference is used with the Box-Jenkins methodology. After elimination of the trend, the data are stationary and an autoregressive model using only the second lag appears to be a reasonable fit, however, the p-value for the autoregressive term is .1171. Strictly speaking, this term would not be considered significant at a reasonable alpha level. Since this is the best model using Box-Jenkins methodology, we will assume that it is at least a reasonable fit. The residual autocorrelation check reveals very small autocorrelations. All of the p-values for the Box-Jung statistics are large. This model is considered the best fit after examining several alternate models. The statistical analysis of this model is as follows.

**Table 5: Statistical Analysis of Model 4**

| <b>Conditional Least Squares Estimation</b> |                 |                       |                |                    |                        |
|---|-----------------|-----------------------|----------------|--------------------|------------------------|
| <b>Parameter</b>                            | <b>Estimate</b> | <b>Standard Error</b> | <b>t-Value</b> | <b>Pr &gt;  t </b> | <b>Approximate Lag</b> |
| AR1,1                                       | -0.22001        | 0.13798               | -1.59          | 0.1171             | 2                      |

| <b>Autocorrelation Check of Residuals</b> |                   |           |                      |                         |        |        |        |        |        |
|---|-------------------|-----------|----------------------|-------------------------|--------|--------|--------|--------|--------|
| <b>To Lag</b>                             | <b>Chi Square</b> | <b>DF</b> | <b>Pr &gt; ChiSq</b> | <b>Autocorrelations</b> |        |        |        |        |        |
| 6   | 3.18              | 5         | 0.6727               | 0.102                   | 0.046  | -0.022 | 0.184  | -0.012 | -0.092 |
| 12  | 6.83              | 11        | 0.8130               | 0.134                   | 0.113  | 0.135  | 0.053  | 0.002  | -0.075 |
| 18  | 9.44              | 17        | 0.9255               | 0.096                   | 0.057  | 0.071  | -0.074 | 0.106  | -0.015 |
| 24  | 10.79             | 23        | 0.9853               | 0.008                   | -0.054 | -0.033 | 0.101  | 0.016  | 0.009  |

The root mean squared error for the five year and ten year holdout data sets are .0013 and .0012, respectively. Interestingly, the ten year prediction is better than the five year prediction for the Box-Jenkins model. The backpropagation neural network model uses 5 inputs, 5 neurons in the hidden layer, and one output. This model yields a value of .0012 and .0015 for the five year and ten year holdout data sets, respectively. Thus, for predicting the number of fatal accidents yearly divided by millions of miles flown (FACCMMF), the neural network model performs only slightly better than the Box-Jenkins model for the ten year holdout data set, but does not outperform the Box-Jenkins model for the ten year holdout data set.

**Model 5. Number of Accidents Yearly Divided by Millions of Miles Flown (ACCMMF).**

Because of the declining trend in FACCMMF, a first difference is used with the Box-Jenkins methodology. After elimination of the trend, the data are stationary and the first order autoregressive model with the first order term of the moving average model appears to be a reasonable fit. The residuals autocorrelation check reveals small autocorrelations. The p-values for the Box-Jung statistics are large except for the first Chi-square. This model is considered the best fit after examining several alternate models. The statistical analysis of this model is as follows.

**Table 6: Statistical Analysis of Model 5**

| Conditional Least Squares Estimation |          |                |         |         |                 |
|--------------------------------------|----------|----------------|---------|---------|-----------------|
| Parameter                            | Estimate | Standard Error | t-Value | Pr >  t | Approximate Lag |
| MA1,1                                | -0.86505 | 0.25061        | -3.45   | 0.0012  | 1               |
| AR1,1                                | -0.92237 | 0.19229        | -4.80   | <.0001  | 1               |

| Autocorrelation Check of Residuals |            |    |            |                  |        |       |       |        |        |
|------------------------------------|------------|----|------------|------------------|--------|-------|-------|--------|--------|
| To Lag                             | Chi Square | DF | Pr > ChiSq | Autocorrelations |        |       |       |        |        |
| 6                                  | 9.94       | 4  | 0.0415     | 0.144            | 0.127  | 0.193 | 0.134 | -0.028 | 0.281  |
| 12                                 | 12.44      | 10 | 0.2566     | 0.036            | 0.037  | 0.157 | 0.058 | 0.034  | 0.081  |
| 18                                 | 13.67      | 16 | 0.6231     | 0.065            | -0.018 | 0.080 | 0.022 | 0.032  | -0.061 |
| 24                                 | 14.93      | 22 | 0.8651     | 0.007            | 0.035  | 0.102 | 0.019 | 0.030  | 0.029  |

The root mean squared error for the five year and ten year holdout data sets are .0013 and .0012, respectively. Interestingly, the ten year prediction is better than the five year prediction for the Box-Jenkins model. The backpropagation neural network model uses 9 inputs, 3 neurons in the hidden layer, and one output. This model yields a value of .0024 and .0018 for the five year and ten year holdout data sets, respectively. Thus, for predicting the number of accidents yearly divided by millions of miles flown (ACCMMF), the neural network model does not outperform the Box-Jenkins model for either the five year or ten year holdout data sets.

**CONCLUSIONS**

Airline accidents and fatalities may be thought to be purely random and rather difficult to forecast. However, the ARIMA models and the ANN models developed in this study reveal that models can be developed with reasonable predictive power to forecast safety data. How well should the models be able to forecast the data? This question may have a subjective answer. Perhaps one could use Theil's U to determine if the forecasts are substantially better than the naive model. To obtain a feel for the accuracy of the forecasts, the root mean squared error can be compared to the range of values for the data over the period of the forecast. For the ten years of holdout data, the range of the data for PASSFMDEP, PASSFMMF, FACCMDEP, FACCMMF, and ACCMMF are 41.56, .062, .746, .0012, and .0044, respectively. The root mean squared error of the ARIMA (neural net) models are 21.12 (18.65), .032 (.031), .46 (.25), .0012 (.0015), .0013 (.0012), respectively. Since these results indicate that the root mean squared error is generally much less than the range of the data, the models can be considered to be viable in assessing future values. However, further research is still needed to obtain more accurate forecasts.

While the safety numbers appear to be small, one must keep in mind that these figures are

based on a normalization of the data and hence the millions of passengers flying make these numbers look especially low. If the number of near collisions and incursions involving the pilots taking extra action to avoid collisions are included in the data, the numbers may not look as rosy. A reliable model for safety data could assist in understanding changes in airline accidents and fatalities. This information may be useful to the airline industry and government administrators in determining needed changes in airport capacity and airline safety procedures. In addition, the comparison of a machine learning technique and a traditional statistical prediction method illustrate that valuable insights can be gained by using both approaches.

## REFERENCES

- Altay, A., Ozkan, O., & Kayakutlu, G. (2014). Prediction of aircraft failure times using artificial neural networks and genetic algorithms. *Journal of Aircraft*, 51(1), 47-53. DOI: 10.2514/1.C031793
- Andersson, S. E. (2001). Forecasting and optimization for airline operational planning. *International Journal of Services Technology and Management*, 2(1), 160-172. DOI: 10.1504/IJSTM.2001.001597
- Bowerman B.L., O'Connell, R. T., & Koehler, A. (2004). *Forecasting, Time Series, and Regression* (4th ed.). Boston, MA: Cengage Learning.
- Murphy, C. (2001). How to fix the air-traffic mess. *Fortune*, 43(14), 116-123.
- Foreman, S. E. (1993). An application of Box-Jenkins ARIMA techniques to airline safety data. *Logistics and Transportation Review*, 29(3), 221-240.
- Garcia M. (2019, May 6). "FAA 20-year forecast predicts further air traffic concentration at hubs." [Web log post] Retrieved from <https://runwaygirlnetwork.com/2019/05/06/faa-20-year-forecast-predicts-further-air-traffic-concentration-at-hubs/>
- Huang, H., Zeng, Q., Pei X., Wong, S. C., & Xu, P. (2016). Predicting crash frequency using an optimised radial basis function neural network model. *Transportmetrica A: Transport Science*, 12(4), 330-345. DOI: 10.1080/23249935.2015.1136008
- Jain, B. A. & Nag B. N. (1997). A performance evaluation of neural network decision models. *Journal of Management Information Systems*, 14, 201-216. DOI: 10.1080/07421222.1997.11518171
- Makridakis, S. & Hibon, M. (1997). ARMA models and the Box-Jenkins Methodology. *Journal of Forecasting*, 16, 147-163. DOI: 10.1002/(SICI)1099-131X(199705)16:3<147::AID-FOR652>3.0.CO;2-X
- Kanafani, A. & Keeler, T. E. (1990). Air deregulation and safety: Some econometric evidence from time series. *The Logistics and Transportation Review*, 26, 203-210.
- Ocampo, R. P. & Klaus, D. M. (2018). Applying regression analysis to model the risk of space flight and terrestrial activities. *The Journal of Space Safety Engineering*, 5, 135-139. DOI: 10.1016/j.jsse.2018.08.001
- Rose, N. L. (1989). Financial influences on airline safety. In L.N. Moses & I. Savage (Eds.) *Transportation Safety in an Age of Deregulation* (93-114). New York: Oxford University Press.
- Ross, M. S. (1993). *Probability Models* (5th ed.). New York: Academic Press.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J., (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536. DOI: 10.1038/323533a0
- Salpukas, A. (1982, May 8). "Future of airline deregulation; issue and debate." *New York Times*, Section 2, page 29. Retrieved from <https://www.nytimes.com/1982/05/08/business/future->

[of-airline-deregulation-issue-and-debate.html](#).

United States Department of Transportation, (2000). *FAA Statistical Handbook of Aviation*, United States Government Printing Office (published annually).

Widrow, B. & Lehr, M. A., (1990). 30 years of adaptive neural networks: Perceptron, Madaline, and backpropagation. *Proceedings of the IEEE*, 78, 1415-1442.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175. DOI: 10.1016/S0925-2312(01)00702-0

## ABOUT THE AUTHORS

**Pamela Rogers**, assistant professor at Stephen F. Austin State University, earned her PhD in operations and supply chain management from University of North Texas. She has more than 20 years of teaching experience in higher education and for the U.S. Air Force. Professional experience includes inventory control, corporate training, advertising, project management and website design. Her research interests include manufacturing, supply chain management, quality and business education, which is published in journals such as *the International Journal of Production Research and Decision Sciences Journal of Innovative Education*. She is a member of the DSI, Southwest DSI, ISM, ASCM/APICS and Alpha Iota Delta.

**Robert J. Pavur** is a Professor in the Information Technology & Decision Sciences Department at the University of North Texas. He holds a masters and doctorate degree in Statistics from Texas Tech University. He has published in such journals as the *Annals of Operations Research*, *IEEE Transactions on Reliability*, *European Journal of Operational Research*, *Journal of the Operational Research Society*, and *International Journal of Operations and Quantitative Management*. He is a co-author of the textbook *Introduction to Business Statistics: A Computer Integrated, Data Analysis Approach* published by Cengage. He is a long-time member of the Decision Sciences Institute and Southwest Decision Sciences and has held several officer positions.