Pergamon

Computers ind. Engng Vol. 30, No. 2, pp. 257–267, 1996
0360-8352(95)00170-0

0360-8352/96 $15.00 + 0.00

USING NEURAL NETWORKS TO PREDICT COMPONENT INSPECTION REQUIREMENTS FOR AGING AIRCRAFT*

HUAN-JYH SHYUR,[1] JAMES T. LUXHOJ[1] and TREFOR P. WILLIAMS[2]

[1]Department of Industrial Engineering and [2]Department of Civil Engineering, Rutgers University, P.O. Box 909, Piscataway, NJ 08855-0909, U.S.A.

Abstract—Currently under development by the Federal Aviation Administration (FAA), the Safety Performance Analysis System (SPAS) will contain indicators of aircraft safety performance that can identify potential problem areas for inspectors. The Service Difficulty Reporting (SDR) system is one data source for SPAS and contains data related to the identification of abnormal, potentially unsafe conditions in aircraft or aircraft components/equipment. A higher expected number of SDRs suggests a greater possibility of a maintenance problem and may be used to alert Aviation Safety Inspectors (ASIs) of the need for preemptive safety or repair actions.

The preliminary SDR performance indicator in SPAS is not well defined and is too general to be of practical value. In this study, an artificial neural network model is created to predict the number of SDRs that could be expected by part location using sample data from the SDR database that have been merged with aircraft utilization data. The predictions from the neural network models are then compared with results from multiple regression models. The methodological comparison suggests that artificial neural networks offer a promising technology in predicting component inspection requirements for aging aircraft.

1. INTRODUCTION

Large scale service systems, such as air and surface transport systems, require well designed maintenance management programs to effectively compete in a global economy. These repairable systems are subject to aging mechanisms, such as wear, fatigue, creep and stress corrosion. Inspection and diagnostic activities are integral components of an effective maintenance strategy in an attempt to ensure system safety, reliability and availability.

Due to the significant growth in the number of aircraft in the U.S., there is an increasing number of structural components for the Federal Aviation Administration (FAA) to monitor. There is a need to develop new techniques for maintaining airworthiness of aging aircraft and for improved methods for accurate prediction of residual life of repaired structures. The use of new prediction methods, such as artificial neural networks, may prove useful for forecasting of removal and inspection dates for aircraft engines, assemblies and components.

The Safety Performance Analysis System (SPAS), currently under development by the FAA, will be an analytical tool designed to support Aviation Safety Inspectors (ASIs) [1, 2] and will contain indicators of safety performance that can signal potential problem areas for inspector consideration [3–5]. It will also enable inspectors to access existing FAA data sources in a timely manner, and will function as a decision support tool to assist inspectors in identifying airlines and/or aircraft that present a greater safety risk and warrant further surveillance.

The Flight Standards Information System (FSIS) is a comprehensive program that contains all Flight Standards automation efforts. SPAS and Flight Standards Automation are components of FSIS. The SPAS Project Management Plan was proposed in March 1991. SPAS is a novel research program for FAA inspection activities, since it will integrate data relating to air operator, air agencies, aircraft types and personnel and shifts away from the current use of decentralized databases.

*This article is based on research performed at Rutgers University. The contents of this paper reflect the view of the authors who are solely responsible for the accuracy of the facts, analyses, conclusions, and recommendations presented herein, and do not necessarily reflect the official view or policy of the Federal Aviation Administration.

The FAA has established a Center for Computational Modeling of Aircraft Structures (CMAS) at Rutgers University. One CMAS research project has focused on analyzing the contribution of the Service Difficulty Reporting (SDR) database to SPAS. The SDR subsystem contains data related to the identification of abnormal, potentially unsafe conditions in aircraft or aircraft components/equipment. The *major objectives* of this research project are:

—To develop meaningful indicators that establish national air operator profiles for comparison purposes.
—To identify national SDR trends and inputs to improve the FAA's surveillance system.
—To provide guidelines for the efficient scheduling of FAA safety inspectors.

A significant effort of SPAS is to develop safety performance indicators that will identify and define critical safety characteristics for airline operators. The 13 performance indicators that are currently defined in SPAS will assist in diagnosing an airline's safety profile compared with others in the same class. The preliminary SDR performance indicator in SPAS is too general to be of practical value. The result of this CMAS research effort is that more refined, specific SDR performance indicators have been generated. The tracking of performance indicators facilitates the identification of unfavorable trends, thus enabling a safety inspector to focus attention on airlines most in need of closer examination. Such heightened tracking enhances efficient scheduling of inspections under budgetary and staffing constraints.

Models to predict the *overall* number of SDRs, and the number or SDRs for cracking and corrosion cases by aircraft type have been developed using both multiple regression analysis and neural networks. The "best" models from these two modeling approaches have been compared and are reported in Luxhoj *et al.* [6]. Neural networks have proven to be a very effective model-free regression methodology to predict the expected number of SDRs. For creating more useful safety performance indicators, this paper attempts to predict the number of SDRs that could be expected by *part location* based on the SDR database. This is the first step to construct an indicator to signal potential problem areas by *component type* for inspector consideration. These alert indicators can be used to define upper and lower control limits and to monitor adverse trends in component performance. Efficient inspection activities will facilitate timely aircraft maintenance and minimize the cost of aircraft unavailability.

While it is true that prediction models for determining aircraft *maintenance* requirements could be based on simply forecasting aggregate failure rates by aircraft type for all planes repaired at the same depot or forecasting failure rates for each plane assigned to a different, regional repair facility, the primary purpose of the CMAS research focuses on the composite *inspection* activities of a regulatory agency.

2. RESEARCH METHODOLOGY

While there exist many prediction methods in the literature, this research focuses on two modeling approaches to develop more refined SDR performance indicators for aircraft component types: multiple regression and neural networks. Multiple regression represents a "classical" approach to multivariate data analysis while the emerging field of neural networks represents a "new" approach to nonlinear data analysis. Multiple regression is a general statistical technique used to analyze the relationship between a single dependent (predicted) variable and several independent (predictor or regressor) variables. Multiple linear regression produces a linear approximation to the data. Variable transformations allow, to some extent, the linear regression methods to handle nonlinear cases as well. However, such transformations may make the interpretation of the results difficult. One could always find a polynomial of higher degree that would give a perfect fit to a specified data set. However, this results in overfitting and an inability of the regression model to generalize. Also, regression models do not learn incrementally, and must be re-estimated periodically.

Neural networks attempt to simulate the functioning of human biological neurons. Neural networks have been particularly useful in pattern recognition problems that involve capturing and learning complex underlying (but consistent) trends in data. Neural networks are highly nonlinear, and in some cases, are capable of producing better approximations than multiple regression which

Table 1. Sample SDR and ARS "merged" data [10]

Aircraft model	SDR date	Part name	Part location	Part condition	Estimated age	Estimated flight hr	Estimated landing
DC9	84-03-22	Skin	E + E compt	Cracked	17.74	32,619.03	53,999.20
DC9	84-03-22	Skin	Aft bag bin	Cracked	17.74	32,619.03	53,999.20
DC9	86-07-07	Skin	Fuselage	Cracked	20.03	36,836.23	60,980.56
DC9	80-06-20	Skin	Galley door	Cracked	13.24	34,396.44	33,888.77
DC9	81-12-01	Skin	FS625	Corroded	14.69	38,160.55	37,597.32
DC9	87-05-11	Skin	RT wheel well	Cracked	20.14	52,299.10	51,527.19
DC9	87-05-11	Skin	STA 580-590	Cracked	20.14	52,299.10	51,527.19

produces a linear approximation [7–9]. However, as noted above, variable transformations do allow, to some extent, the regression methods to handle nonlinearity. Neural network learning supports incremental updating and is easier to embed in an intelligent decision system since batch processing is not required. While neural networks offer an alternative to regression that will learn functional relationships among variables to predict an outcome measure, neural network outcomes lack a simple interpretation of results. For instance, the modeling technique does not provide objective criteria to decide what set of predictors is more important for the prediction. Neural networks may also suffer from overfitting of the data and lack of prediction generality. The limitations of neural networks with respect to outliers, multicollinearity and other problems inherent in real world data have received scant attention.

Knowledge of the expected numbers of SDRs by aircraft component type will have value to Aviation Safety Inspectors (ASIs) when attempting to efficiently schedule field inspection workload requirements. Moreover, the identification of unfavorable inspection trends will enable the FAA to specify that the airlines take preemptive maintenance measures.

2.1. Data description

The CMAS research team was provided with a subset of the SDR database that had been merged with the Aircraft Utilization (ARS) database. This merged database was created by Rice [10] and consisted of 1308 observations for the DC-9 aircraft for the period April 1974 to March 1990. Table 1 displays sample data. Estimated flight hours and estimated landings are derived values based upon the original delivery date of the first operator, the date of the ARS data reference and the SDR date. The equations developed by Rice [10] for these derived values are presented in Fig. 1.

3. NEURAL NETWORK MODELS

Neural networks are computing systems that incorporate a simplified model of the human neuron, organized into networks similar to those found in the human brain [7]. Instead of

. AGE
. ESTFHRS = (SDRDATE–SERVICE)/(ARSDATE–SERVICE)*FHSCUM
. ESTLDGS = (SDRDATE–SERVICE)/(ARSDATE–SERVICE)*LDGSCUM

where

ESTFHRS = Estimated flying hours at SDRDATE
ESTLDGS = Estimated landings at SDRDATE
SDRDATE = Date of SDR (from SDR database)
SERVICE = Original delivery date (first operator, ARS)
ARSDATE = Date or ARS data reference
FHSCUM = Cumulative fuselage flying hours (ARS)
LDGSCUM = Cumulative fuselage landings (ARS)

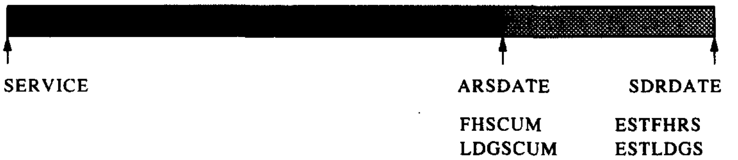


Fig. 1. Derived predictor variables [10].

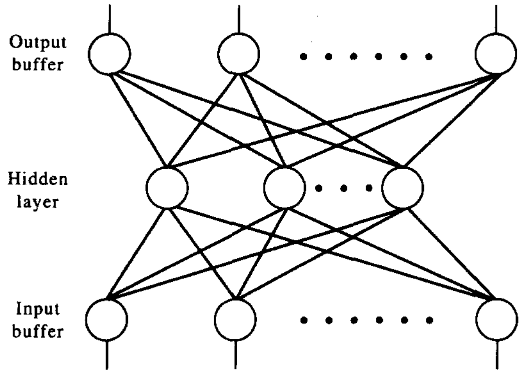


Fig. 2. An example three-layer backpropagation neural network.

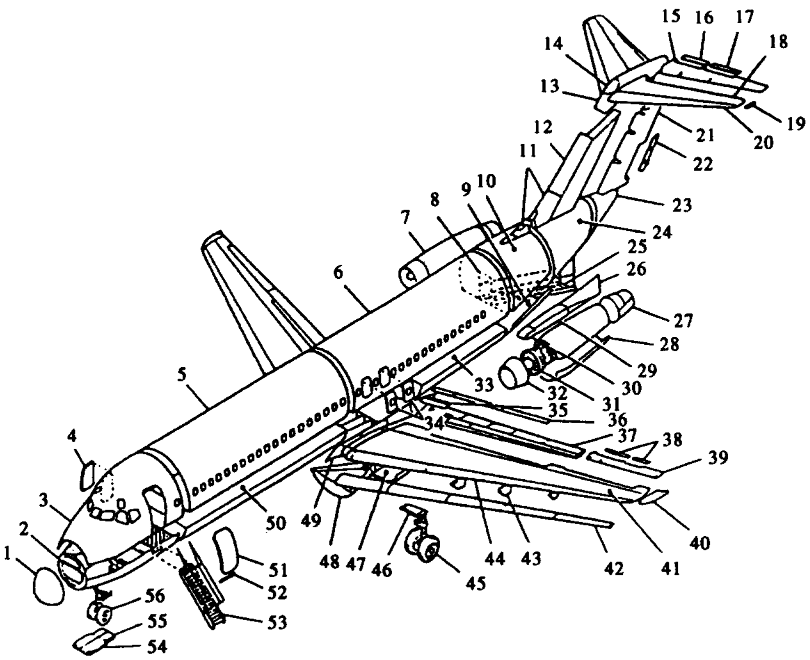
programming the neural network, it is "taught" to give acceptable results. Artificial neural networks are computer simulations of biological neurons. Neural networks can be layered into many levels, with or without hidden layers exhibited between an input and an output layer. Figure 2 displays a network of neurons that are organized into a three-layer hierarchy. The ability of artificial neural networks to capture underlying, complex trends in data has been researched and documented in a significant number of research papers since the "rebirth" of neural networks in 1982 when researchers "rediscovered" their important characteristics [11–15]. The large number of research papers available on these characteristics prohibits their documentation here, but as an indication of their diverse cognitive power, there have been applications of neural networks in varied areas from stock market price prediction and credit rating approval to engineering applications such as pattern/image recognition, digital signal processing and automated vehicle guidance [13].

This research attempts to take advantage of the ability of artificial neural networks to capture and retain complex underlying relationships that exist between an aircraft's operations data and SDR inspection reporting profiles.

There are six broad categories and approx. 50 different types of neural network architectures in use today [16]. Backpropagation neural networks are the most commonly used neural network architectures. These neural networks are especially good for pattern recognition. Figure 2 shows the basic configuration of the three-layer backpropagation network. To develop a backpropagation model, a training set of data patterns which consist of both inputs and the actual outputs observed must be developed. During training, the neural network processes patterns in a two-step procedure. In the first or forward phase of backpropagation learning, an input pattern is applied to the network, and the resulting activity is allowed to spread through the network to the output layer. The program compares the actual output pattern generated for the given input to the corresponding training set output. This comparison results in an error for each neurode in the output layer, In the second, or backward phase, the error from the output layer is propagated back through the network to adjust the interconnection weights between layers. This learning process is repeated until the error between the actual and desired output converges to a predefined threshold [17]. A trained neural network is expected to predict the output when a new input pattern is provided to it.

3.1. SDR part location neural networks

A structural schematic of the DC-9 aircraft is presented in Fig. 3. For the 1308 sample data observations from the period 1974–90, there are only 569 data observations for cracking cases for the DC-9, and only 390 observations identify the part location. There is insufficient and incomplete input data for each part location, so the part locations were categorized into 11 larger "groupings" as presented in Table 2. The corresponding sample sizes for each major part grouping is superscripted in parentheses. Note that the part location numbers in Table 2 do not correspond to the part location numbers in Fig. 3 due to the "grouping" strategy. The frequency of reporting by part location is graphically portrayed in the frequency histogram of Fig. 4. For the given data, approx. 70% of cracking cases occurred in the aircraft main fuselage areas and the "Fuselage STA 588 to 996 (recoded as "Part 3")" includes 22.2% of cracking cases.



NO.	Description	NO.	Description	NO.	Description
1.	Radome	20.	Horizontal stabilizer leading edge	39.	Aileron
2.	Fuselage nose lower structure	21.	Rudder	40.	Wing tip
3.	Fuselage nose upper structure	22.	Rudder tab	41.	Wing main structure
4.	Forward service door	23.	Tail cone	42.	Wing slat
5.	Fuselage STA 229 to 588 upper structure	24.	Fuselage tail structure	43.	Flap hinge fairing
6.	Fuselage STA 588 to 996 upper structure	25.	Passenger AFT entrance door stairway	44.	Wing leading edge
7.	Upper cowl door	26.	Pylon AFT panel	45.	Main gear
8.	Passenger AFT entrance stairwell door	27.	Thrust reverser cowling	46.	Main gear outboard door
9.	Fuselage STA 996 to 1087 lower structure	28.	Lower cowl door	47.	Main gear inboard door
10.	Fuselage STA 996 to 1087 upper structure	29.	Pylon center panel	48.	Keel
11.	Dorsal fin	30.	Pylon leading edge	49.	Wing-to-fuselage fillet
12.	Vertical stabilizer	31.	Engine	50.	Fuselage STA 229 to 588 lower structure
13.	Vertical stabilizer tip	32.	Nose cowl	51.	Passenger forward entrance door
14.	Removable tip fairing	33.	Fuselage STA 756 to 996 lower structure	52.	Forward stairwell door
15.	Elevator	34.	Overwing emergency exits	53.	Passenger forward entrance stairwar
16.	Elevator control tab	35.	Flap vane	54.	Forward nose gear doors
17.	Elevator geared tab	36.	Spoiler	55.	AFT nose gear doors
18.	Horizontal stabilizer AFT section	37.	Wing flap	56.	Nose gear
19.	Horizontal stabilizer tip assembly	38.	Aileron tabs		

Fig. 3. Structural schematic of the DC-9 aircraft (Douglas Aircraft Co., Inc, *DC-9 Structure Repair Manual*).

Table 2. Backpropagation neural networks for DC-9 part locations (cracking cases)

Initial parameters	
Learning rate = 0.01 Momentum = 0.05 Intitial weight = 0.3 Patterns = 20	Input layers = 3 (Avg-age, Avg-flight hours, Avg-landings) Hidden layers = 21 Output layer = 11 (No. of SDR for each part location)

Part No.	Part description (·): No. of observations	R^2 value
1	Fusealge nose structure(51)	0.8563
2	Fuselage station 229 to 588(37)	0.8329
3	Fuselage station 588 to 996(81)	0.8504
4	Fuselage station 996 to 1087(68)	0.7382
5	Fuselage tail structure(39)	0.6837
6	Rudder(13)	0.7832
7	Pylon aft panel(7)	0.4926
8	Wing(15)	0.5771
9	Passenger fwd entrance door(27)	0.7948
10	Cargo door(7)	0.8371
11	Aft press blkhd(20)	0.8744

Based on the current SDR data base which presents one SDR for one aircraft record, a prediction model is created that can be used to signal potential problem areas by homogeneous aircraft type. The expected number of SDRs in a certain part location is used as an index that indicates the possibility of failure in the area. A higher expected number of SDRs suggests a greater possibility that a maintenance problem exists.

To generate the model, current data have been grouped by aircraft age or flight hours to calculate the average number of SDRs in each part location. For example, in the age "cohort" or grouping of 10 yr $\leqslant$ aircraft age $\leqslant$ 10.5 yr, if it is observed that there are two aircraft with three part No. 2 failures and three aircraft with six part No. 3 failures, then we can argue that one aircraft in that age range has 1.5 SDRs (3 SDRs/2 aircraft) for part No. 2 or 2 SDRs (i.e. 6 SDRs/3 aircraft) for part No. 3. Moreover, the average age, flight hours and number of landings in this age "cohort" is also calculated to create a complete training data record.

In an attempt to create "robust" SDR prediction models, different data "grouping" strategies were surveyed. Currently, the data with aircraft age less than 16 yr are grouped by age in increments of 0.5 yr and the remaining data are grouped by flight hours in increments of 4000 hr. This grouping strategy is suggested by a multiple regression approach as reported by Luxhoj *et al.* [6]. The SAS regression procedure "STEPWISE" with significance level 0.05 is used in Ref. 6 to find explanatory variables (such as age, flight hours and number of landings) for an *overall* SDR prediction model that included both cracking and corrosion cases. This overall model did not identify part locations. If the observed data with aircraft age less than 16 yr are grouped by flight hours, then no independent variables are selected for a regression model using the backwards stepwise procedure.

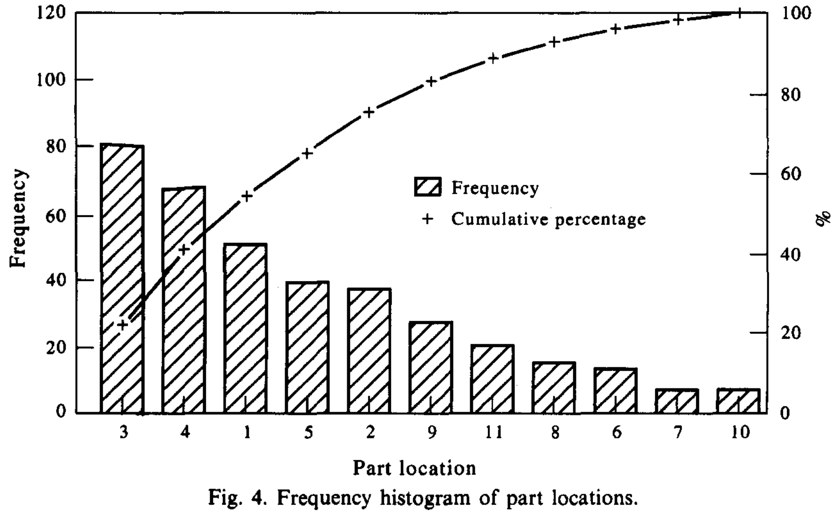


Fig. 4. Frequency histogram of part locations.

Table 3. Sample training patterns for neural network models

Age (input)	Flight hours (input)	Landings (input)	Part-1 (output)	Part-2 (output)	Part-3 (output)	...	Part-11 (output)
7.85	18,690.47	20,802.41	0	0	0		0
9.89	23,196.85	22,131.14	0	0	1		0
10.22	26,628.15	29,868.82	0	0	1		0
14.54	38,342.68	39,529.86	0	0	2		0
15.48	20,389.20	21,082.51	0	1	0		0
16.10	7648.51	8747.93	0	0	0		0
16.59	9847.38	11,650.02	0	0	0		0
19.71	14,890.32	17,608.92	1	1	1		0
20.49	18,417.62	22,897.46	0	1	1	...	1
21.27	21,193.82	28,295.16	0	0	1		0
21.22	25,934.51	30,647.01	0	1	1		1
21.18	30,354.34	33,997.60	1	0	1		1
20.93	34,201.58	41,745.07	1	1	1		1
20.47	37,618.85	41,610.96	1	1	1		1
19.95	41,831.14	46,925.22	1.333	1	1		1
20.85	46,138.41	50,187.29	1	1.333	1.083		1
21.28	49,973.60	56,181.93	1.25	1.5	1		1
21.41	53,329.71	60,088.01	1	1	1.583		1
21.84	57,753.52	65,016.48	1.222	1	1.273		0
21.87	61,351.04	56,383.35	1	1	1		0

However, if the data are grouped by age, then it appears that age is a significant explanatory variable for this model. Furthermore, we can also identify that data with aircraft age greater than 16 yr should be grouped by flight hours as this can be a significant explanatory variable in the prediction model. When using the data grouping strategy of flight hours in the SDR data base, the suggested interval grouping size is 4000 flight hours [6].

A three-layer backpropagation architecture is used to classify SDR cracking cases for part location data grouped by the above data partitioning strategy. Previous research by Luxhoj *et al.* [6] on the sensitivity of using different neural network architectures for the SDR problem domain suggests that the three-layer backpropagation architecture yields the best results for the given SDR data. This observation is consistent with Maren *et al.* [13] who report on successful applications of three-layer backpropagation architectures for fault diagnosis. For the component data, the number of SDRs for one aircraft in a certain age group is calculated. As Luxhoj *et al.* [6] demonstrate, the use of "ungrouped" data results in poor prediction models for both the multiple regression and neural network techniques. However, due to the age "grouping" strategy, only 18 input patterns can be used to train the neural network model. The model includes three input neurons (i.e. operations data, such as aircraft age, flight hours and number of landings) and 11 output neurons that identify the number of SDRs in 11 different part locations. These patterns are created in an attempt to capture underlying relationships between aircraft operations data and the "mix" of expected number of SDRs by component type. The sample training patterns are presented in Table 3, and the designed backpropagation neural network model is shown in Fig. 5. The numbers in the 11 "output" columns represent the "average" number of SDRs per airplane for

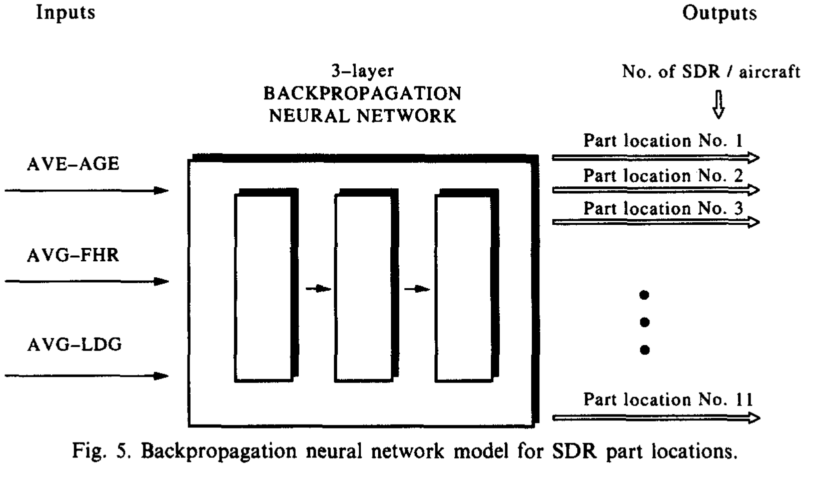


Fig. 5. Backpropagation neural network model for SDR part locations.

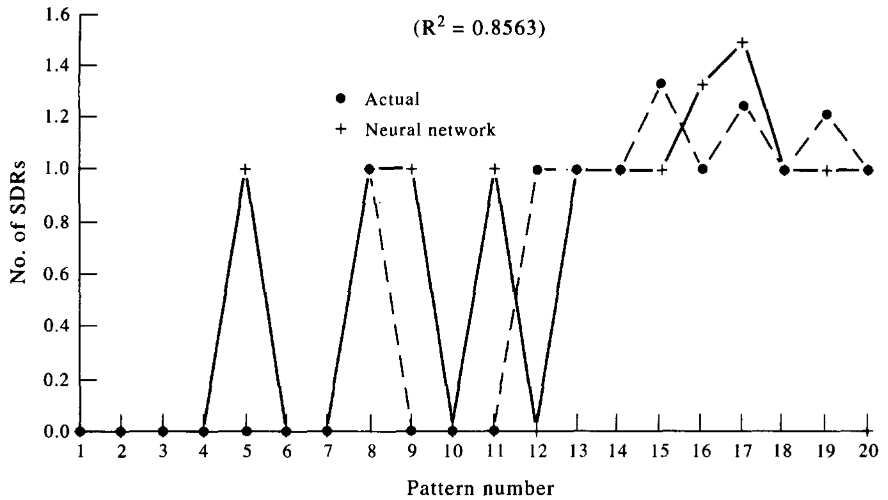


Fig. 6. Model performance of the fuselage nose structure (part location No. 1).

that part location and were obtained using the "cohort" data grouping strategy as previously discussed.

The prediction model is trained using the *NeuroShell* 2 [18] computer program, which requires Microsoft Windows, a minimum of 4 MB of RAM, and at least a 386SX microprocessor. In addition, this commercial software program implements several different types of neural network architectures, and supports easy import/export of files, model building and testing, and the creation of run-time versions of trained networks all via a friendly graphical user interface. Backpropagation neural network "learning" parameters include the "learning rate" which is used to specify the magnitude of the weight changes, the "momentum" factor which specifies the proportion of the last weight change that is added to the new weight change, and an "initial weight" that is used to initialize the weights between the network's connections prior to "training".

In neural network modeling, the R^2 value compares the accuracy of the model to the accuracy of a trivial benchmark model where the prediction is simply the mean of all the sample patterns. A perfect fit would result in an R^2 value of 1, a very good fit near 1, and a poor fit near 0. If the neural network model predictions are worse than one could predict by just using the mean of the sample case outputs, the R^2 will be 0. Although not precisely interpreted in the same manner as the R^2 value or coefficient of multiple determination in regression modeling, nevertheless, the R^2 value from a neural network model may be used as an approximation when evaluating model adequacy.

Plots of the actual value vs the predicted value from the network for two of the part locations are exhibited in Figs 6 and 7. These parts are the Fuselage Nose Structure (Part No. 1) and the Fuselage Station 588–996 (Part No. 3). Corresponding R^2 values for each of the 11 part location models are provided in Table 2. Eight of the 11 models have R^2 values about 0.7 which suggests that a backpropagation neural network is very effective in predicting the number of SDRs for major structural groupings of part locations. Five of the 11 models have R^2 values of 0.8 or higher. The "best" part location backpropagation models in this study are for the AFT Press BULKHEAD ($R^2 = 0.8744$), fuselage nose structure ($R^2 = 0.8563$), fuselage stations 588 to 996 ($R^2 = 0.8504$), cargo door ($R^2 = 0.8371$) and fuselage stations 229 to 588 ($R^2 = 0.8329$).

The accuracy of neural network models will be affected by the input training patterns. The number of the observations for each of the 11 part locations is one major factor that has influence on the accuracy and efficiency of the model. For a small number of observations, the neural network model cannot provide an effective prediction for some part locations. For example, the Pylon AFT panel model ($R^2 = 0.4926$) is based on only seven observations. Even for some cases that have a relatively high R^2 value but with only few observations, there is still insufficient evidence to conclude that the model provides a reasonable prediction in this part location. The cargo door neural network model is one of these cases.

3.2. Comparison with multiple regression models

An alternative approach to estimate the number of SDRs by part location is the use of multiple regression. However, one multiple regression model can only predict one dependent variable. To predict the number of SDRs for 11 different part locations, 11 different multiple regression models should be created. The following list of possible explanatory variables was used in creating the multiple regression models: age, flight hours, number of landings, age^2, flight hours^2, number of landings^2, age × flight hours, age × number of landings, flight hours × number of landings, flight hours/age, and number of landings/age. Quadratic terms were considered in an inherently linear model to evaluate any nonlinear relationships. The regression models were examined for multicollinearity, since a high degree of multicollinearity makes the results not generalizable as the parameter estimates in the model may not be stable due to the high variance of the estimated coefficients. Since flight hours, number of landings and age of an aircraft are interrelated, multicollinearity is inherent in the independent variables.

Two statistical measures of multicollinearity are the tolerance (TOL) value and the variance inflation factor (VIF) (Hair *et al.* [19]). The tolerance value is equal to one minus the proportion of a variable's variance that is explained by the other predictors. A low tolerance value indicates a high degree of collinearity. The variance inflation factor is the reciprocal of the tolerance value, so a high variance inflation factor suggests a high degree of collinearity present in the model. The VIF and TOL measures assume normality and are typically relative measures. A high tolerance value (above 0.10) and a low VIF value (below 10) usually suggest a relatively small degree of multicollinearity (Hair *et al.* [19]).

Due to the lack of observed data in two cases (i.e. sample sizes of seven), only nine multiple regression models are created to estimate the number of SDRs for major structural groupings of part locations. The VIF and TOL values were examined for each model to reduce the multicollinearity. A sample model to predict the expected number of SDRs for the major structural grouping of "Fuselage Station 588–996 (Part No. 3)" is provided below:

$$\text{No. of SDRs (Part No. 3)} = -6.192868 + (0.341556 \times \text{age}) - (0.000004097 \times \text{age} \times \text{flight hours}) + (0.001897 \times \text{age/flight hours}).$$

This model has an R^2 value of 0.7844 and the relative measures of multicollinearity are reported as:

Independent variable	**TOL**	**VIF**
age	0.05040	19.839
age × flight hours	0.03395	29.455
age/flight hours	0.06052	16.522

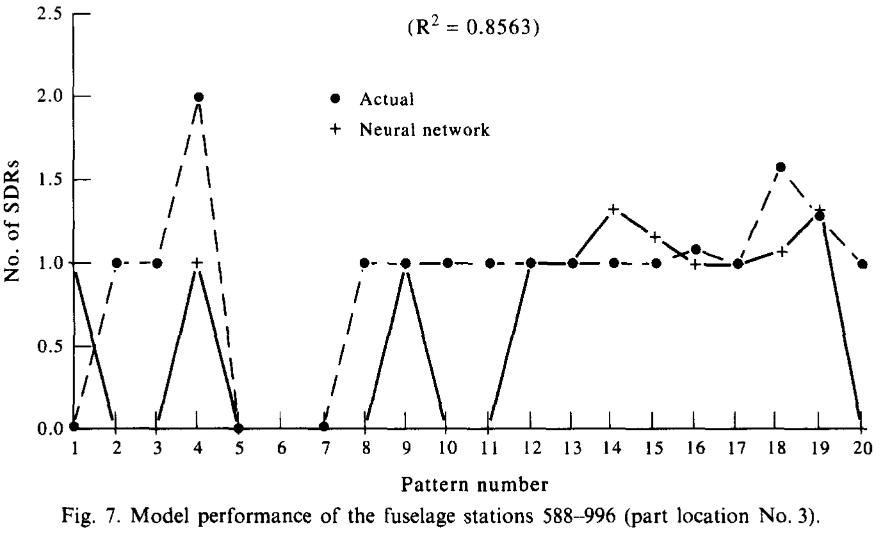


Fig. 7. Model performance of the fuselage stations 588–996 (part location No. 3).

Table 4. Comparison of neural network and multiple regression models

Part No.	Part description (·): No. of observations	NN R^2	NN MSE	MR R^2	MR MSE
1	Fuselage nose structure (51)	0.8563	0.043	0.6723	0.123
2	Fuselage station 229 to 588 (37)	0.8329	0.048	0.5314	0.170
3	Fuselage station 588 to 996 (81)	0.8504	0.039	0.7844	0.070
4	Fuselage station 996 to 1087 (68)	0.7382	0.078	0.5435	0.171
5	Fuselage tail structure (39)	0.6837	0.099	0.6303	0.145
6	Rudder (13)	0.7832	0.054	0.4699	0.164
7	Pylon aft panel (7)	0.4926	0.107	—	—
8	Wing (15)	0.5771	0.101	0.4453	0.166
9	Passenger fwd entrance door (27)	0.7948	0.051	0.6801	0.099
10	Cargo door (7)	0.8371	0.034	—	—
11	Aft press blkhd (20)	0.8744	0.047	0.6593	0.158

The multicollinearity measures suggest that some degree of multicollinearity is present in this model, but these values are close to the generally recommended threshold values of VIF $\leqslant 10$ and TOL $\geqslant 0.10$.

Table 4 presents the R^2 values for the nine regression models as compared with the R^2 values for the corresponding neural network models. Although not interpreted precisely in the same manner, nevertheless, analyzing the R^2 values from neural network and multiple regression facilitates approximate comparisons. Only one multiple regression model has an R^2 value above 0.7.

Since the goal is to maximize the precision of the SDR predictions, the mean square error (MSE) is also used for comparative purposes. The MSE is defined as:

$$\text{MSE}\,(\hat{\beta}) = E(\hat{\beta} - \beta)^2,$$

where β is some arbitrary parameter. The above expression is equivalent to:

$$\text{MSE} = [\text{bias}\,(\hat{\beta})]^2 + \text{var}\,(\hat{\beta}).$$

The criterion of minimizing the MSE thus considers the variance and the square of the bias of the estimator. Comparing the results of the two approaches reveals that neural network modeling provides better fits to the SDR part location data in all cases than does regression modeling. Also, it appears that neural networks may be more useful for predictive purposes, in some cases, when there are sparse data sets, an observation that has been reported in Luxhoj and Shyur [20] in their analysis of helicopter part reliability data. However, the general predictive capability and statistical confidence of neural networks given small sample sizes still requires investigation across varied characterizations of data sets.

4. CONCLUSIONS AND RECOMMENDATIONS

In this study, promising results are achieved when using three-layer backpropagation neural networks to predict SDR reporting profiles by part location for major structural components for the DC-9 aircraft. Prediction of the number of SDRs for each part location is helpful to Aviation Safety Inspectors (ASIs) and may be used to signal potential problem areas based upon aircraft operating conditions and age. Using neural networks, one can obtain the expected number of SDRs for a major structural grouping of components whenever the operating conditions are known. These current results are encouraging, but additional data are required to validate the generality of the modeling approach.

The part location data "grouping" strategy is useful to provide efficient input patterns to create a reasonable neural network model. However, some information hidden in the data set is lost in the process of "grouping". Due to this strategy, the difficulty to maintain the current neural network will also increase. A new approach to replace the current "grouping" strategy should be developed.

Another recommendation is to develop the entire probability mass function from the neural network for major part locations. This would then provide aircraft safety inspectors with the most likely category of failure, the second most likely etc. based upon the operating conditions and age of the aircraft. Attempts to refine the management reporting from the neural network models are underway.

Neural network modeling, a model-free regression technique, is easy to develop, maintain and use. However, it should possess one set of reasonable and useful input buffers. In the SDR reporting profiles, only three different input buffers (age, flight hours and number of landings) can be used in this model. Other factors that affect the components, such as engine hours and flight cycles, are still not exhibited in the SDR reporting profiles. If these data can be collected and merged to the current data base, then more promising and meaningful models can be created.

Acknowledgements—The authors of this report would like to acknowledge the support of the SPAS program and Mr John Lapointe and Mr Michael Var.

REFERENCES

1. *Safety Performance Analysis Subsystem.* Functional description document. U.S. Department of Transportation, Volpe National Transportation Systems Center, Cambridge, MA, March (1992).
2. *Safety Performance Analysis Subsystem.* Prototype concept document. U.S. Department of Transportation, Volpe National Transportation Systems Center, Cambridge, MA, April (1992).
3. *Safety Performance Analysis Subsystem.* Indicators suggested for SPAS prototype (DRAFT). U.S. Department of Transportation, Volpe National Transportation Systems Center, Cambridge, MA, August (1992).
4. *Safety Performance Analysis Subsystem.* Continuing analysis: indicator graphs and tables. U.S. Department of Transportation, Volpe National Transportation Systems Center, Cambridge, MA, October (1992).
5. *Safety Performance Analysis Subsystem.* Continuing analysis: additional indicator definitions (DRAFT). U.S. Department of Transportation, Volpe National Transportation Systems Center, Cambridge, MA, October (1992).
6. T. Luxhoj, T. P. Williams and H.-J. Shyur. Prediction of inspection profiles for aging aircraft. Rutgers University, Center for Computational Modeling of Aircraft Structures, Technical Report (1994).
7. P. Simpson. *Artificial Neural Systems.* Pergamon Press, New York (1990).
8. P. D. Wasserman and T. Schwartz. Neural network, part 1: what are they and why is everybody so interested in them now? *IEEE Expert* **2,** 10–13 (1987).
9. A. Weigend, B. A. Huberman and D. E. Rumelhart. Predicting the future: a connectionist approach. *Int. J. Neural Syst.* **1,** 193–210 (1990).
10. R. C. Rice. Repair Database Assessment. Battelle Summary Report, Contract No. DTRS-57-89-C-00006, March (1991).
11. J. J. Hopfield. Neural networks and physical systems with emergent collective abilities. *Proc. Natl Acad. Sci.* **79,** 2554–2558 (1982).
12. J. J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl Acad. Sci.* **81,** 3088–3092 (1984).
13. A. J. Maren, C. T. Harston and R. M. Pap. *Handbook of Neural Computing Applications.* Academic Press, New York (1990).
14. W. S. McCulloch and W. Pitts. A logical calculus of ideas immanent in nervous activity. *Bull. Mathl Biophys.* **5,** 115–133 (1943).
15. D. L. Reilly, L. N. Cooper and C. Elbaum. A neural model for category learning. *Biol. Cybernet.* **45,** 35–41 (1982).
16. R. Hecht-Nielsen. Neurocomputing: picking the human brain. *IEEE Spectrum* **25,** 36–41 (1988).
17. M. Caudil. Neural network training tips and techniques. *AI Expert* **6,** 56–61 (1991).
18. *NeuroShell 2.* Ward Systems Group, Inc., Frederick, MD (1993).
19. J. F. Hair, R. E. Anderson, R. L. Tatham and W. C. Black. *Multivariate Data Analysis,* 3rd edn. Macmillan, New York (1992).
20. J. T. Luxhoj and H.-J. Shyur. Reliability curve fitting for aging helicopter components. *Reliab. Eng. Syst. Safety* **48,** 229–234 (1995).