ELSEVIER

Finite Elements in Analysis and Design 23 (1996) 381–403

FINITE ELEMENTS IN ANALYSIS AND DESIGN

Integrated decision support for aviation safety inspectors[1]

James T. Luxhøj,[a,*] Trefor P. Williams[b]

[a]*Department of Industrial Engineering, Rutgers University, P.O. Box 909, Piscataway, NJ 08855-0909, USA*
[b]*Department of Civil Engineering, Rutgers University, P.O. Box 909, Piscataway, NJ 08855-0909, USA*

Abstract

This paper presents a systems viewpoint for developing an advanced decision support system for aircraft safety inspectors. Research results from a Federal Aviation Administration (FAA) sponsored project to use neural network and expert systems technology to analyze aircraft maintenance databases are summarized. One of the main objectives of this research is to define more refined "alert" indicators for national comparison purposes that can signal potential problem areas by aircraft type for safety inspector consideration.

Integration aspects are addressed on two levels: (1) integration of the various technical components of the decision support system, and (2) integration of the decision support system with individual behavior, management systems and organizational structure, as well as corporate culture across both formal and informal dimensions. The paper summarizes the creation of strategic "inspection profiles" for aging aircraft and reliability curve fitting for structural components both based upon using neural network technology. Also, the potential use of a model-based expert system to facilitate field inspection diagnostics is presented. Finally, a framework for developing an intelligent decision system to support aircraft safety inspections is proposed that links expert systems, neural networks, as well as a paradigm of the decision making process typically used in unstructured situations.

Keywords: Neural networks; Intelligent decision systems; Aircraft safety

1. Introduction to the problem of aircraft safety inspection

Effective and efficient maintenance management is essential not only for production systems but for large-scale service systems, such as air and surface transport systems. 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.

* Corresponding author.

[1] This article is based on research performed at Rutgers University. The contents of this paper reflect the views 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.

PII S0168-874X(96)00009-1

As an example of a large-scale service system, the Federal Aviation Administration (FAA) in the United States is responsible for regulating aircraft traffic and safety. An expected increase in usage of domestic flights in the next few years coupled with an aging population of aircraft has led the FAA to initiate new aircraft safety research efforts. Domestic passenger enplanements increased from 250 to 450 million annually between 1977 and 1987 [1]. The FAA anticipates that domestic enplanements will reach 800 million in the year 2000, and exceed a billion by 2010 for increases of 128% and 272%, respectively [1]. It is also expected that by the year 2010 there will be a 55% increase in aircraft operations including takeoffs and landings at towered airports; a 62% increase in instrument operations in terminal areas; a 73% increase in air carrier hours; and increases of 62% and 75% in the air carrier and commuter fleets, respectively [1]. This steady growth of aircraft transport and aircraft traffic density places increasing pressure on safety inspection activities.

The inspection of aircraft involves a number of complex technical, social, political, economic, and human issues. Inspection frequencies, procedures, and criteria may vary for alternative types of aircraft. Alternative safety equipment and measurement accuracies are required for different components. There may be delays in inspections due to coordination and scheduling conflicts. Expertise is required in diagnosing potential safety problems and in making probability assessments. An aging population of Aviation Safety Inspectors (ASIs) has created concern within the FAA that the expertise associated with aircraft inspection will not be preserved. A sense of urgency exists concerning the capturing and codification of existing aircraft inspection knowledge.

Due to the growth in the number of aircraft, there is an increasing number of structural components 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 engines, assemblies, and components.

Replacement inspections focus on a specific component or components that have been scheduled for replacement at specific intervals. The component that was in service may undergo further testing in the supply area and repaired if necessary and returned as a usable spare. If it is determined that it is not cost effective to repair the worn component, it will be discarded. Also, a replacement inspection may result in the safety inspector making a decision to defer replacement of the inspected component.

New safety indicators need to be defined that will enable inspectors to identify airlines that present a greater safety risk and warrant heightened surveillance. These alert indicators can be used to define upper and lower control limits and to monitor adverse trends. 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 research reported in this paper focuses on the composite *inspection* activities of a regulatory agency that is responsible for ensuring that national safety standards are met.

For modern aircraft systems, there is a high degree of reliability built in which means that there are infrequent failures. When failures are infrequent, it becomes difficult to detect and isolate the problem quickly. The development of a knowledge base for fault detection and isolation for aircraft will enable the codification of existing inspection expertise before this expertise leaves the FAA organization. Once captured, this knowledge can be efficiently applied on a continuous basis via an expert system to enhance the decision-making productivity and consistency of both novice and experienced aircraft safety inspectors.

2. Vision of intelligent decision support for aircraft safety

Currently under development by the FAA, the Safety Performance Analysis System (SPAS) will be an analytical tool that is intended to support FAA inspection activities [2, 3] and will contain indicators of safety performance that can signal potential problem areas for inspector consideration. SPAS will function as a decision support system by enabling inspectors to access existing FAA maintenance databases and by identifying airlines that pose a greater safety risk and merit heightened surveillance. SPAS is novel for FAA inspection activities because it attempts to integrate data on air operator, air agencies, aircraft types, and air personnel components into a unified decision support system and differs from the current use of decentralized databases.

The FAA has established a Center for Computational Modeling of Aircraft Structures (CMAS) at Rutgers University. One CMAS research project concentrates on the Service Difficulty Reporting (SDR) database that contains data related to the identification of abnormal, potentially unsafe conditions in aircraft or aircraft components/equipment. The major objectives of this research are to develop meaningful SDR indicators that establish national air operator "*profiles*" for comparison purposes and to investigate the use of artificial neural networks and expert systems to analyze maintenance databases. The creation of "*inspection profiles*" will assist in the characterization of aircraft with respect to what needs to be inspected, when it needs to be inspected, and how often should inspection occur based upon monitoring operations data from different "*classes*" of aircraft types.

The emergence of Decision Support Systems (DSSs) or computerized information systems that contain domain-specific knowledge and analytical decision models to support decision making for complex and ill-structured tasks began in the 1960s at the Sloan School of Management at the Massachusetts Institute of Technology, the Harvard Business School, and the Business School HEC in France [4]. Although research results are equivocal, a DSS is intended to enhance individual learning by providing easier access to problem recognition, problem structure, information, statistical tools, and methodological knowledge [5]. A DSS is designed to enable easier and faster generation of alternatives and to increase awareness of deficiencies in the current decision making process.

Holtzman [6] uses the new terminology of an "Intelligent Decision System" to describe a hybrid computer-based technology for aiding decision makers in complex decision situations. The tools build upon the methodology of decision analysis and the technology of expert systems. The idea is to use expert systems technology to automate the skills and factual knowledge of the expertise of a few individuals and to use the normative characteristics of decision analysis to improve the quality of the decisions made. The goal is to reduce the time, cost, and training needed to make decisions in complex problem domains. Intelligent decision systems may also be used to make probability assessments in specific situations. Fig. 1 depicts the elements of Holtzman's paradigm for an Intelligent Decision System.

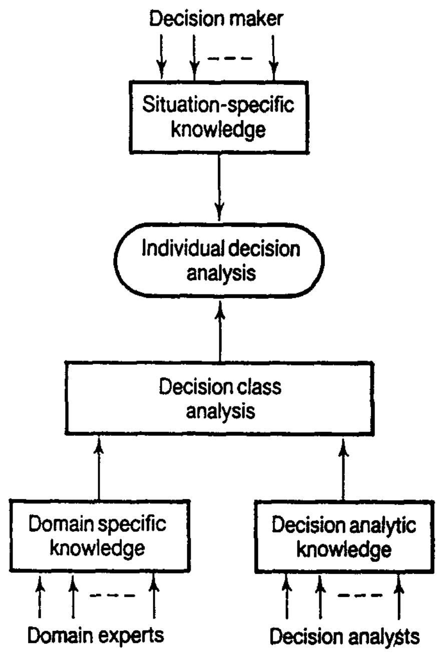


Fig. 1. Holtzman's [6] paradigm of an intelligent decision system.

In complex decision-making situations, research suggests that a decision maker attempts to deal with the unstructuredness by decomposing these situations into familiar, structurable decision elements [7, 8]. Mintzberg et al. [9] develop a model of strategic decision making that attempts to portray the decision-making process as consisting of three phases – identification, development, and selection. Seven distinct, familiar decision "routines" comprise these phases and there are three "supporting" decision routines and six dynamic factors that may influence the decision-making process. Fig. 2 illustrates the Mintzberg et al. model. Although real decision making is not as static or sequential as the figure presents, nevertheless, this model provides insight as to the unique phases of decision making and positions the technology for decision support.

For example, neural networks are especially useful at monitoring data and detecting trends or implicit patterns that can signal potential problems. Thus, neural networks are especially appropriate for supporting the problem recognition and diagnosis decision routines of trying to comprehend external stimuli and assessing cause–effect relationships. Expert systems are better suited for encoding explicit domain knowledge, exploring alternative decisions, and for providing explanations of reasoning processes. Thus, in the Mintzberg et al. decision modeling framework, expert systems are especially appropriate for supporting the search and design routines of the decision-making process that lead to the generation of one or more solutions. Also, expert systems are appropriate for supporting the screening and evaluation-choice routines as these systems are able to search a decision space, develop alternatives, make inferences, and provide explanations and traces of how conclusions were reached.

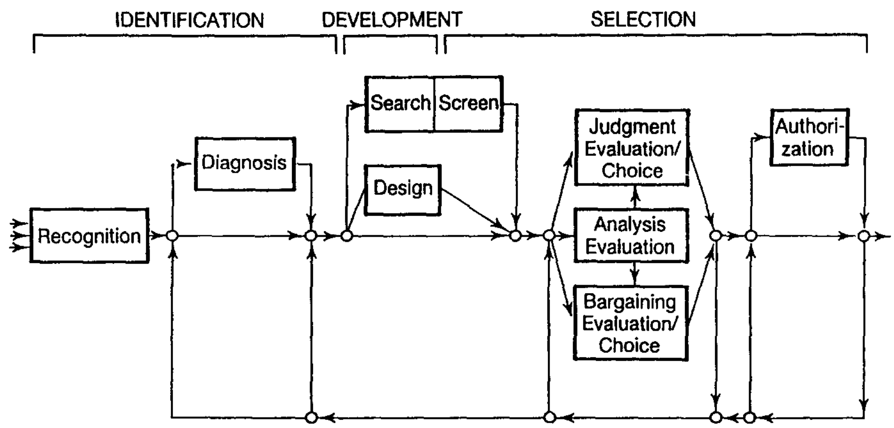


Fig. 2. Mintzberg et al. [9] model of the decision-making process in unstructured situations.

This paper presents a framework or systems viewpoint for developing the components of an intelligent decision system for aircraft safety inspection. Integration aspects are addressed on two levels. First, integration of the various technical components of the intelligent decision system, such as neural networks, expert systems, influence diagrams, and models of decision making are examined. Second, the connections that a decision support system has to individual behavior, management systems and organizational structure, and organizational culture are also discussed.

The eventual goal of the CMAS research is to develop an intelligent decision system that will be a hybrid of expert system and neural network technology supported by aviation databases to facilitate maintenance requirements planning, organizational coordination, and efficient workload scheduling for safety inspectors.

3. Decision support for aircraft safety inspections

One of the first steps in the development of any preventive or predictive maintenance program is to determine what needs to be inspected and the inspection intervals. The inspection frequency is usually a function of the type of equipment, its age and condition, the utilization, the operating environment, and the consequences of equipment unavailability due to failure.

The CMAS research effort attempted to explore the ability of artificial neural networks to capture and retain complex underlying relationships and nonlinearities by investigating the patterns that may exist between an aircraft's operations and maintenance data and SDR reporting profiles. Knowledge of the SDR reporting profiles with respect to an "alert" indicator facilitates a determination of inspection workload requirements for Aviation Safety Inspectors. The current planned SDR performance indicator is S which is simply the number of SDR records for the airline operator for the defined period. The count of records is not normalized. If $S > 0$, the indicator status is set as "*expected*"; if $S = 0$, the indicator status is set as "*advisory*" [3]. This "alert" indicator is too general to be of practical value.

3.1. Neural network-based approaches to inspection

Neural networks are computing systems that imitate intelligent behavior and are composed of a number of simple, highly connected processing elements that process information by a dynamic state response to external inputs [10]. Neural networks are "taught" to give acceptable results. The ability of artificial neural networks to capture and retain complex patterns has been researched and documented in a number of papers since the "rebirth" of neural networks in 1982 when researchers "rediscovered" their important characteristics [11–13]. Fault diagnosis usually requires the collection and processing of large amounts of data which are frequently incomplete. Fault diagnosis is typically composed of fault detection, based upon either off-line or on-line inspection procedures, and then fault isolation. Neural network inspection systems have been developed for detecting and isolating equipment malfunctions in complex aircraft such as the Grumman X-20, the National Aerospace Plane (NASP or X-30), the F-16 Falcon, and for NASA's space shuttle [14]. The use of neural networks for equipment monitoring and fault detection has led to the development of new, on-line "predictive" maintenance paradigms [15, 16]).

The most common type of neural network architecture is backpropagation which is especially useful for pattern recognition. The initial program employs an analog, three layer, backpropagation network. Fig. 3 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]. Backpropagation neural network "learning" parameters include the

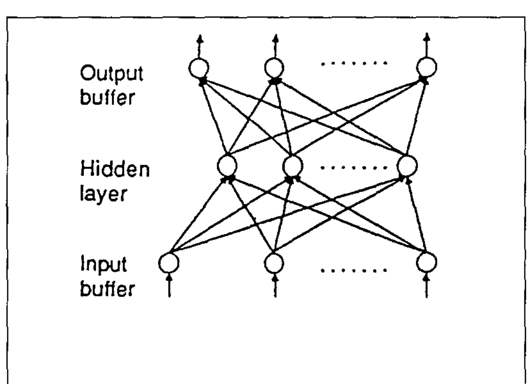


Fig. 3. Three-layer backpropagation neural network.

"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 in regression modeling, nevertheless, the R^2 value from neural network model may be used as an approximation when evaluating model adequacy.

The neural network models were developed on an IBM-PC compatible computer using the NeuroShell 2 [18] computer program. This program requires Microsoft Windows, a minimum of 4mb of RAM, and at least a 386 microprocessor. The program implements several different types of neural network models. Initial model development focused on the use of backpropagation networks.

3.1.1. SDR prediction

Inputs for the SDR neural networks were flight hours, landings, and the age of the plane. The output was the expected number of SDRs. The data consisted of the 1308 cases of merged SDR and Aircraft Utilization data developed by Battelle for the DC-9. However, when cases with missing data were eliminated, there were a total of 1229 usable data cases. The data were not grouped in any way. All of the training patterns were for individual aircraft.

Backpropagation models can be configured in several different ways. Models can be developed consisting of three, four, or five layers of processing elements. In the standard backpropagation configuration, the processing elements of one node are connected only to the following layer. It is also possible to provide additional connections, called jump connections so that all layers of the network are fully connected.

One problem in developing neural networks is to determine the point in training where the neural network provides the best results. Often, training a neural network to provide a minimum error when presented with the training set produces a neural network that cannot generalize. The NeuroShell 2 software overcomes this problem with a feature called "NET-PERFECT". This feature requires breaking the input data into two different groups. One group is the training set that the neural network trains on. The other group is the test set, that is tested periodically to determine the error produced. The network that produces the minimum error with the test set is saved.

3.1.1.1. 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 for the same set of planes. The SDR database essentially contains *qualitative* descriptions of potentially unsafe components or systems by aircraft serial number. The ARS database contains *quantitative* data related to the operations of the aircraft, such as flight hours and number of landings. The merged database was supplied by Battelle [19] and consisted of 1308 observations for the DC-9 aircraft for

Table 1
Sample "merged" SDR and ARS data [19]

Aircraft model	Serial number[a]	SDR date	Part name	Part location	Part condition	Estimated age	Estimated flight hours	Estimated landing
DC9	333	84-03-22	Skin	E + E COMPT	Cracked	17.74	32619.03	53999.20
DC9	333	84-03-22	Skin	AFT BAG BIN	Cracked	17.74	32619.03	53999.20
DC9	333	86-07-07	Skin	FUSELAGE	Cracked	20.03	36836.23	60980.56
DC9	444	80-06-20	Skin	GALLEY DOOR	Cracked	13.24	34396.44	33888.77
DC9	444	81-12-01	Skin	FS625	Corroded	14.69	38160.55	37597.32
DC9	444	87-05-11	Skin	RT WHEEL WELL	Cracked	20.14	52299.10	51527.19
DC9	444	87-05-11	Skin	STA 580-590	Cracked	20.14	52299.10	51527.19

[a]Fictious serial #'s are used due to confidentiality of data.

the period April 1974 to March 1990. Table 1 displays sample data. The database only contained quantitative data on the following:

- Age
- Estimated flight hours
- Estimated number of landings.

Since actual data on flight hours and landings were not being reported directly in the SDR database, the estimated flight hours and estimated landings are derived values based upon the original delivery date of the plane to the first operator, the date of the ARS data reference, and the SDR date. The equations developed by Battelle for these derived values are reported in [19] and are given below:

$$\text{Estimated flight hours} = [(\text{SDR date-service date})/(\text{ARS date-service date})] * \text{FHSCUM}$$

$$\text{Estimated number of landings} = [(\text{SDR date-service date})/(\text{ARS date-service date})] * \text{LDGSCUM},$$

where SDR date is the date of the SDR (SDR database), Service date is the original delivery date of the plane to the first operator (ARS database), ARS date is the date of the ARS data (ARS database), FHSCUM is the cumulative fuselage flight hours (ARS database) and LDGSCUM is the cumulative fuselage landings (ARS database).

Since the ARS date time lagged the SDR date, Rice extrapolated the quantitative ARS data on flight hours and landings to the SDR date. He developed a multiplier by calculating the ratio of (SDR date-service date/ARS date-service date) and then extrapolated the flight hours and landings at the ARS date to the date of the SDR.

In this initial stage of neural network development it was concluded that neural networks created with ungrouped data did not provide acceptable results across a variety of backpropagation architectures and different learning parameters. It became necessary to transform the input data.

3.1.1.2. Data grouping strategies. In an attempt to create "robust" SDR prediction models that will provide SDR profiles for a "representative" DC-9, different data "grouping" strategies were used. Such an approach was used in [20–26] to create large-scale logistics models for the US Navy. These "population" models were developed to determine both maintenance and system repair/replacement strategies for large groupings of similar equipment based on operating hours, operating environment, failure mode, etc.

Using data grouping strategies of "age", "estimated flight hours", and "estimated landings", neural network models were developed based upon a smaller data set of "averaged" merged data to predict the total expected number of SDRs/year, the number of SDRs/year for cracked cases, and the number of SDRs/year for corrosion cases for the DC-9 aircraft. In this example, fictious aircraft serial numbers are used, since the actual aircraft serial numbers are confidential information.

To provide a means for checking the SDR predictions against existing data, the data were sub-divided into two different sets based on aircraft serial numbers. The first set was used to build the prediction model and the second set was used to evaluate the prediction model's performance on new data. Such an approach is useful for testing prediction model generality [27]. This approach is typically used in neural network modeling to create a "*training*" set of data to build the model and a "*production*" set of data to evaluate model performance on new, unfit data. These terms are used in the paper to distinguish between the two data sets.

After the data has been subdivided into "*training*" and "*production*" sets, then a grouping strategy is similarly applied to each data set. The grouping procedure based on "age" is outlined below:

1. Group the data to create age "*cohorts*".
   $0 \leqslant \text{AGE} < 1$ 1st group NEWAGE = 0
   $1 \leqslant \text{AGE} < 2$ 2nd group NEWAGE = 1
   ... and so on ...

2. Calculate the "*average*" flight hours and number of landings.

Data set

NEWAGE	SERIAL #	EST. FLIGHT HOURS	EST. LANDINGS
9	111	26677.78	27410.47
9	111	25718.65	25947.47
9	222	21383.77	21731.40
9	222	22253.71	17263.51
9	222	24139.99	26998.77

New data set

# OF SDRs	NEWAGE	AVG. FHR	AVG. LDG
5	9	24034.78	23870.324

3. Calculate the average number of SDRs based upon the number of aircraft serial numbers that comprise each age "*cohort*".

Previous data set

# OF SDRs	NEWAGE	AVG. FHR	AVG. LDG
5	9	24034.78	23870.324

Modified data set (e.g. suppose 2 aircraft serial numbers accounted for the 5 SDRs from step 2)

# OF SDRs	NEWAGE	AVG. FHR	AVG. LDG
2.5	9	24034.78	23870.324

As a result of the grouping strategy, all interpretations are now with respect to the *average* number of SDRs per year. The output variable becomes the average number of SDRs for a "*representative*" DC-9 with a "*profile*" of estimated flight hours and estimated landings as defined by its associated age cohort. An important point to remember when using this data grouping procedure is that one must have a sufficiently large data sample for the DC-9 in order to compute "averages" of estimated landings and flight hours for a specified aircraft age. The more data that one has, the better one can model a "*representative*" aircraft using the data grouping strategy as previously discussed.

The grouping procedure resulted in the following:

	# of data records		"Grouped" # of data records	
Model	Training	Production	Training	Production
Overall # of SDRs	805	424	16	14
# SDRs (cracking)	572	306	16	16
# SDRs (corrosion)	242	127	10	9

3.1.1.3. SDR neural network models for the whole of the DC-9 aircraft. The results from the SDR backpropagtion neural networks models are summarized in Tables 2–4. Training times for the backpropagation models were insignificant. Since prediction accuracy was deemed to be most important, the Mean square error (MSE) was used to select the "*best*" neural network configuration. These neural network models may be used to predict the average number of SDRs using a data grouping strategy of one year time increments for the overall number of SDRs and for the number of corrosion cases. To predict the average number of SDRs for cracking cases, it was determined that the best data grouping strategy was based on increments of 4000 flight hours. Note that although the neural network for the corrosion case performs well on the training data set ($R^2 = 0.9411$, MSE $= 0.086$), the MSE on the production set increased significantly (MSE $=$ 3.125). It should also be observed that the model for corrosion cases had the least number of training and production patterns derived from data groupings with the least number of observations of the three models constructed. Thus, this model should be used with caution on new, unfit data as it does not appear to generalize well.

As in regression modeling, 90% or 95% "confidence intervals" could be developed for the overall number of SDRs and the number of SDRs for cracking and corrosion cases. These confidence intervals could be displayed in a fashion analogous to quality control charts serving as more refined "alert" indicators that specify upper and lower safety control limits by aircraft type.

Luxhøj et al. [28] report on the promising development of a two stage "hybrid" neural network model for SDR prediction. Tables 2–4 also summarize the results of these hybrid neural networks. The first stage uses a Probabilistic Neural Network (PNN) to classify the average age of a DC-9 aircraft into its corresponding class for the expected number of SDRs. A PNN is a supervised neural network that is used to train quickly on sparse data sets [29–31]. Training a PNN is very fast because it requires that each pattern be presented to the network only once during training.

Table 2
Neural network models for overall number of SDRs (data grouped by age, time increment = 1 year)

Input variables		AGE, FHR, LDG	
Output variables		Number of SDR/airplan	
Number of training patterns		16	
Number of production patterns		14	
Backpropagation neural network model			
Number of hidden node	6	Learning time	00:06:32
Learning rate	0.1	R^2 (training data)	0.9452
Momentum	0.1	MSE (training data)	0.152
Initial weight	0.3	MSE (production data)	0.541
Hybrid neural network model			
	Step 1 - PNN		
Input	AGE, FHR, LDG	Patterns classified correctly (training data)	16
Output	Class 1 → $0 \leqslant s < 2$ Class 2 → $2 \leqslant s < 4$ Class 3 → $4 \leqslant s < 6$ Class 4 → $6 \leqslant s < 8$	Patterns classified incorrectly (training data)	0
Learning time	00:00:03	Patterns classified correctly (production data)	12
Smoothing factor	0.02	Patterns classified incorrectly (production data)	2
	Step 2 - Backpropagation		
Input	AGE, FHR, LDG, Class 1, Class 2, Class 3, Class 4		
Output	Number of SDR	Learning time	00:22:55
Number of hidden node	7	R^2 (training data)	0.9603
Learning rate	0.1	MSE (training data)	0.110
Momentum	0.1	MSE (production data)	2.626
Initial weight	0.3		

The neural network separates input patterns into some defined output categories. In the process of training, the PNN clusters patterns by producing activations in the output layer. The value of the activations correspond to the probability density function estimate for that category. All output values of a PNN should be either 0 or 1 and only the output value in the most probable category is 1. Increasing the "smoothing factor" of a PNN causes more relaxed surface fits through the data.

The PNN in this study is used to classify SDRs into one of 4 classes, *class 1* for $0 \leqslant S \leqslant 2$, *class 2* for $2 < S \leqslant 4$, *class 3* for $4 < S \leqslant 6$, and *class 4* for $6 < S \leqslant 8$ where S represents the number of SDRs. The PNN is used in the first stage to classify the average age of a DC-9 aircraft into its corresponding class for expected number of SDRs. This vector of average age and class then is fed into a backpropagation neural network to predict the number of SDRs. The second stage then feeds the classified output along with the above quantitative data to a backpropagation neural

Table 3
Neural network models for SDR cracking cases (data grouped by flight hours, increment = 4000 h)

Input variables		AGE, FHR, LDG	
Output variables		Number of SDR/airplan	
Number of training patterns		16	
Number of production patterns		14	
Backpropagation neural network model			
Number of hidden node	5	Learning time	00:03:45
Learning rate	0.1	R^2 (training data)	0.6899
Momentum	0.1	MSE (training data)	0.009
Initial weight	0.3	MSE (production data)	0.409
Hybrid neural network model			
	Step 1 - PNN		
Input	AGE, FHR, LDG	Patterns classified correctly (training data)	16
Output	Class 1 $\rightarrow 0 \leqslant s < 1.5$ Class 2 $\rightarrow 2 \leqslant s \geqslant 1.5$	Patterns classified incorrectly (training data)	0
Learning time	00:00:02	Patterns classified correctly (production data)	10
Smoothing factor	0.02	Patterns classified incorrectly (production data)	5
	Step 2 - Backpropagation		
Input	AGE, FHR, LDG, Class 1, Class 2		
Output	Number of SDR	Learning time	00:17:35
Number of hidden node	7	R^2 (training data)	0.8404
Learning rate	0.1	MSE (training data)	0.005
Momentum	0.1	MSE (production data)	0.019
Initial weight	0.3		

network to predict the number of SDRs for combined cracking and corrosion cases. SDR prediction results using multiple regression techniques, backpropagation, and "hybrid" neural networks are compared in [28]. In all cases, the prediction results were better on the training data sets than from solely using a backpropagation architecture. However, the MSEs for the production data only improved in the cracking case. Further investigations are required with larger data sets to determine the extent of the benefits of a two-stage approach, as the training time significantly increases with a hybrid model.

3.1.1.4. SDR neural network models for components of the DC-9 aircraft. In an attempt to explore further the use of neural networks to create "safety alerts", Shyur et al. [32] report on the development of SDR prediction models for the DC-9 aircraft that use neural networks for 19 major structural groupings, such as the cargo door, elevator, radome, spoiler, tail cone, etc. The neural network models use the three-layer backpropagation learning architecture to predict the expected

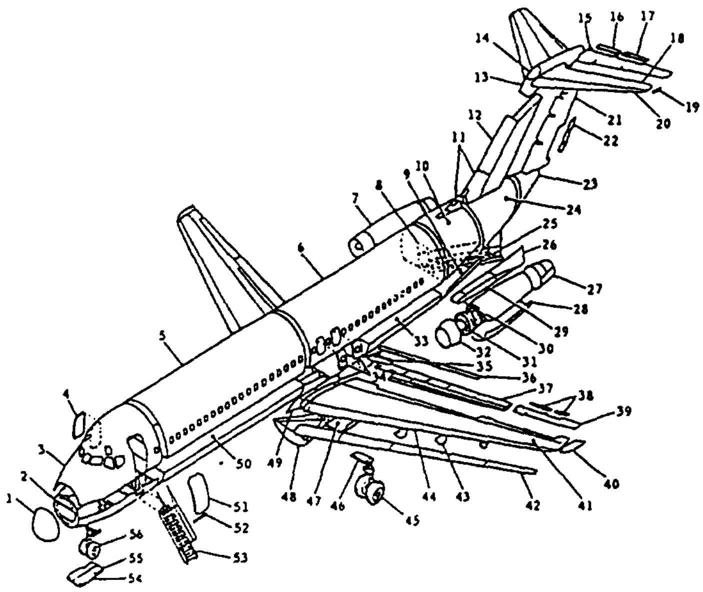
Table 4
Neural network models for SDR corrosion cases (data grouped by age, time increment = 1 year)

Input variables		AGE, FHR, LDG	
Output variables		Number of SDR/airplan	
Number of training patterns		10	
Number of production patterns		9	
Backpropagation neural network model			
Number of hidden node	5	Learning time	00:02:40
Learning rate	0.1	R^2 (training data)	0.9411
Momentum	0.1	MSE (training data)	0.086
Initial weight	0.3	MSE (production data)	3.125
Hybrid neural network model			
	Step 1 - PNN		
Input	AGE, FHR, LDG	Patterns classified correctly (training data)	10
Output	Class 1 → $0 \leqslant s < 2$ Class 2 → $2 \leqslant s < 4$ Class 3 → $s \geqslant 4$	Patterns classified incorrectly (training data)	0
Learning time	00:00:01	Patterns classified correctly (production data)	6
Smoothing factor	0.1	Patterns classified incorrectly (production data)	3
	Step 2 - Backpropagation		
Input	AGE, FHR, LDG, Class 1, Class 2, Class 3		
Output	Number of SDR	Learning time	00:05:44
Number of hidden node	7	R^2 (training data)	0.9727
Learning rate	0.1	MSE (training data)	0.04
Momentum	0.1	MSE (production data)	3.502
Initial weight	0.3		

number of SDRs for cracking cases. A structural schematic of the DC-9 Model 30 aircraft is presented in Fig. 4.

For the 1308 sample data observations, there are only 569 data observations for the DC-9 Model 30 aircraft, and only 390 observations identify the part location. As there were insufficient and incomplete input data for each part location, the part locations were categorized into 19 larger "groupings" as presented in Table 5. Note that the part location numbers in Table 5 do not correspond to the part location numbers in Fig. 4 due to the "grouping" strategy. Approximately, 70% of the cracking cases for the sample data occurred in the aircraft main fuselage area and the "Fuselage STA 588 to 996" includes 20.8% of the cracking cases.

A three-layer backpropagation architecture is used to classify the SDR cracking cases for data grouped by age in increments of 0.5 years. Moreover, the number of SDRs for one aircraft in a certain age group is calculated. Due to the age "grouping" strategy, only 18 input patterns can be used to train the neural network model. The model includes 3 input neurons (aircraft age, flight



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. 4. Schematic of the DC-9 model 30 aircraft (*Source: DC-9 structure repair manual*, Douglas Aircraft Co., Inc.).

Table 5
SDR component neural networks for cracking cases

Initial parameters	
Learning rate = 0.05	Input layers = 3 (AVG-AGE AVG-FHR AVG-LDG)
Momentum = 0.05	Hidden layers = 15
Initial weight = 0.3	Ouput layer = 19 (# of SDR for each part location)
Patterns = 18	

Part #	Part description	R^2-value
1	Radome	0.9432
2	Fuselage nose structure	0.8001
3	Fuselage station 229 to 588	0.7095
4	Fuselage station 588 to 996	0.7961
5	Fuselage station 996 to 1087	0.8767
6	Tail cone	0.9137
7	Fuselage tail structure	0.8467
8	Rudder	0.9600
9	Pylon AFT panel	0.9035
10	Wing	0.7988
11	Passenger forward entrance door	0.8901
12	Elevator	0.9814
13	Main gear door	0.7605
14	Cargo door	0.9654
15	Vertical stabilizer	0.7017
16	AFT nose gear door	0.4891
17	Horizontal stabilizer	0.9413
18	AFT press BLKHD	0.8484
19	Spoiler	0.9743

Data grouped by age (range = 0.5 years)

hours, and number of landings) and 19 output neurons that identify the number of SDRs in 19 different part locations.

As displayed in Table 5, 13 of the 19 models have R^2 values above 0.800 which suggests that a backpropagation neural network is very effective in predicting the number of SDRs for major structural groupings of part locations. Eight of the 19 models have R^2 values of 0.9000 or higher. The "best" part location backpropagation models in this study are for the elevator ($R^2 = 0.9814$), spoiler ($R^2 = 0.9743$), cargo door ($R^2 = 0.9654$), rudder ($R^2 = 0.96$), randome ($R^2 = 0.9432$), horizontal stabilizer ($R^2 = 0.9413$), tail cone ($R^2 = 0.9137$), and pylon AFT panel ($R^2 = 0.9035$). However, the model cannot predict well in the "AFT Nose Gear Door" case ($R^2 = 0.4891$). The number of observations for each of the 19 part locations is one major factor that has an influence on the accuracy and efficacy of the model.

3.1.2. *Component reliability curve fitting*

A related CMAS study involved investigating the use of neural networks for reliability curve fitting of aging aircraft structural components. DC-9 component failure data, such as cumulative

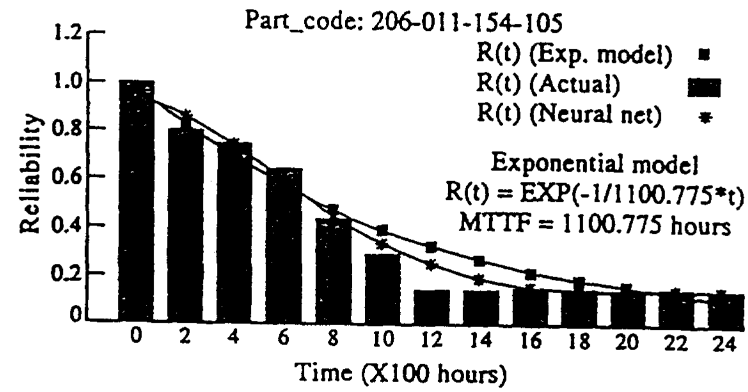


Fig. 5. Comparison of reliability curve-fitting techniques for part code 105-31754 (Helicopter Model Type B).

time before failure, were not available from either the SDR or ARS databases. However, Luxhøj and Shyur [33] report on the use of artificial neural networks for reliability curve fitting for aging helicopter components in the presence of sparse data sets. Both mathematical-function based and neural network models were investigated. The sample data included part code, cumulative operational hours of the helicopter, and cumulative time before the part failed. Although the models are developed for aging helicopter components, the modeling techniques presented in this paper may be extended to the DC-9 aircraft.

Backpropagation neural networks were investigated and the results suggest that the neural networks appear to capture the nonlinearities in the data better than mathematical-function-based approaches. Fig. 5 displays comparative curve fitting techniques for a Helicopter Model Type B component. This ongoing research appears very encouraging for using neural networks for estimating component reliability and for predicting component removal and inspection dates.

3.2. *A potential model-based expert system for safety diagnostics*

There have been numerous expert systems developed in the maintenance and fault diagnosis problem area. Maintenance of complex equipment involves a number of diagnostic procedures that utilize rules and judgements. The large number of rule-based expert systems developed for fault diagnosis prohibit their documentation here, but a survey of applications is provided in [34]. However, classical rule-based expert systems for diagnostics have been recently criticised since the large number of rules for commercial applications result in knowledge bases that frequently are unmaintainable, untestable, and unreliable [35].

With the increased computational power of modern computers, the use of Bayesian probability theory to construct expert systems has been revived. As reported in [36] current expert systems for fault diagnosis suffer from an inability to handle new faults, an inability to recognize when a fault is beyond the consultation system's scope, inadequate explanation of the final diagnosis, excessive requests for new information, and difficulties in construction.

HUGIN [37] is a software for the construction of knowledge-based systems based on causal probabilistic networks or CPNs. The software incorporates new, efficient algorithms to support

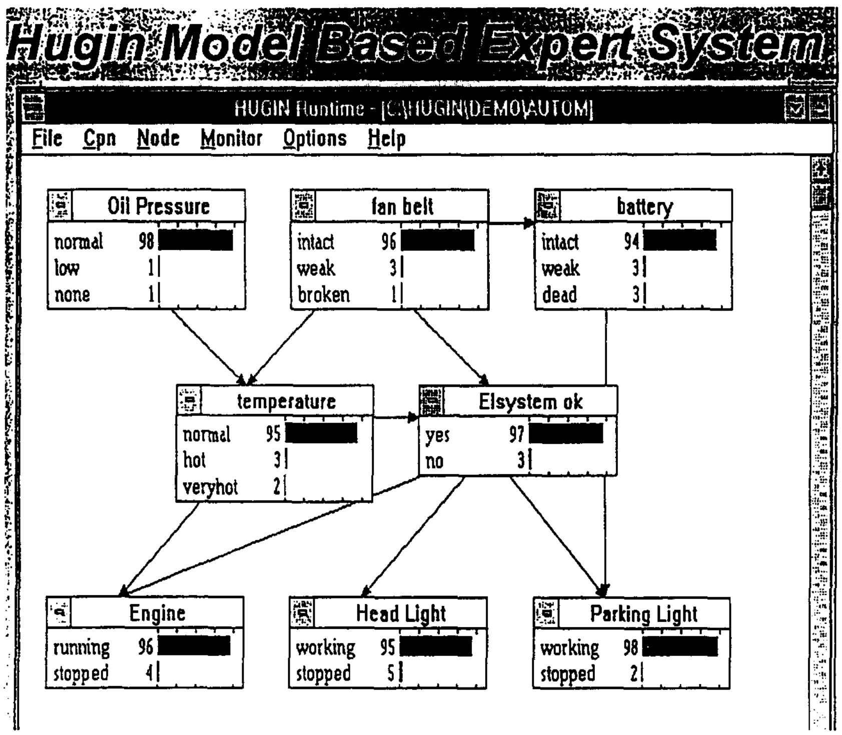


Fig. 6. HUGIN network for maintenance diagnostics.

Bayesian probability calculations and offers an alternative to traditional rule-based programming. The CPNs, also known as belief networks or influence diagrams, represent a possible means to model efficiently the uncertain relationships among components of a system. Moreover, model-based expert systems incorporate causal knowledge by including a representation of a system's structure, function, and behavior.

The model uses a number of statements about the problem domain (e.g. "The patient has lung cancer") and a number of causal relationships between such statements. Each statement is assigned a number of states (e.g. "yes" and "no"), and each state is assigned a probability. Causal dependencies are given as conditional probabilities for a state given the states of the parent node.

In a safety diagnostics model, for example, the knowledge embedded in the cause–effect links between nodes in the CPN will be answers to questions such as "If the direct cause represented by node X is known to have a given value, what is the probability that the effects, given in node Y, will have a certain outcome?" In the CPN illustrated in Fig. 6, one could ask "If the engine in the car gets hot, what is the probability that the carburetor will stop working?" In normal rule-based systems, the question would probably be "If the carburetor stops working, will the engine then get hot (yes/no)?" With HUGIN, the inference engine allows evidence to be entered into nodes and the effect of such evidence to be propagated to other nodes which provides for a very efficient reasoning process.

Horvitz et al. [38] describe an application of HUGIN to develop a probabilistic diagnostic model for NASA's space shuttle propulsion-system engines. A causal probabilistic network for the shuttle's Orbital Maneuvering System (OMS) was developed with the support of experienced flight controllers at Space Shuttle Mission Control in Houston, Texas. The "belief network" shows how the values of helium pressure affect the pressure readings as reported by the two independent pressure sensors on an OMS helium tank. However, these pressure readings can also be affected, with uncertainty, by the errors in the sensor mechanisms themselves. An experienced user in sensor failures can code his or her belief about the relative rate of failure of alternative critical sensors in the system.

The use of such a model-based expert system is being investigated as a possible computerized technique to support aircraft safety inspectors. Such a system would provide the ability to consider alternative hypotheses under uncertainty when diagnosing aircraft systems. The use of a Bayesian model could provide two types of assistance to the safety inspector. First, information related to the status of the aircraft could be presented and safety alert information could be displayed. Second, the conditional reasoning properties of the Bayesian network will enable the safety inspector to formulate "What if?" questions on the current condition of the aircraft and experiment with possible causes for the observed symptoms.

4. A framework for developing intelligent decision support

There are many proposed components to the SPAS research effort. A general framework or systems viewpoint for integrating the use of artificial neural networks for SDR prediction and the use of a model-based diagnostics tool for aircraft safety inspections is needed. Holtzman's [6] notion of an "Intelligent Decision System (IDS)" is useful for designing a decision aid for aircraft safety inspectors that integrates both domain specific knowledge and decision theoretic knowledge (such as influence diagrams).

The *knowledge base* of an IDS for Aviation Safety Inspectors could be divided into the following components:

(1) Domain knowledge related to specific aircraft or to a class of aircraft.

(2) Preference knowledge which is used to elicit a certain type of preference model regarding decision criteria from the safety inspector.

(3) Probabilistic knowledge which addresses the problem of probability assessments in specific situations.

(4) User data that contains facts to define the circumstances of the safety inspector

(5) Process knowledge that guides the safety inspector through the process of decision analysis for aircraft inspections.

Fig. 7 illustrates a conceptual drawing of an IDS for the strategic level in aircraft safety inspections. Based upon input which is obtained from existing FAA data sources, such as the Service Difficulty Reporting (SDR) and Aircraft Utilization (ARS) databases, data are entered into a computer-based model for SDR prediction. The decision model could be either a neural network or multiple regression model. Based on the given inputs, the model then creates "inspection profiles" for the overall expected number of SDRs, and the number of SDRs for cracking and

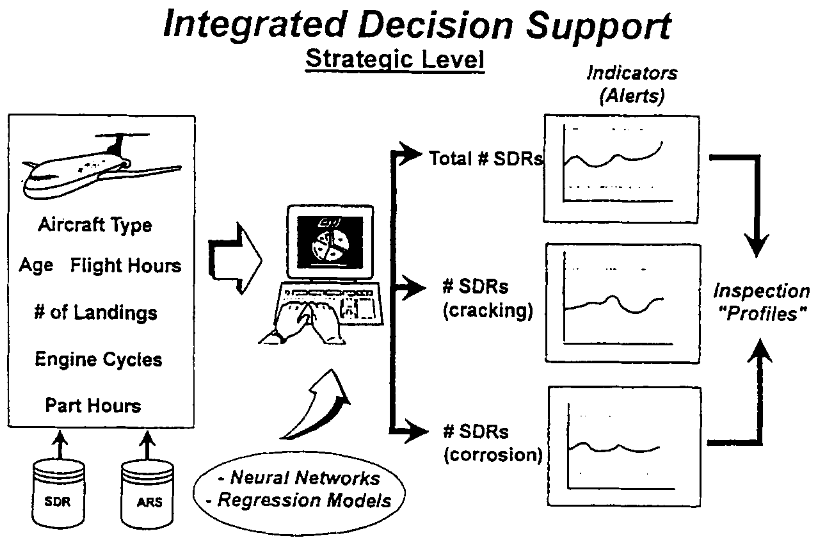


Fig. 7. Strategic level for intelligent decision system (IDS).

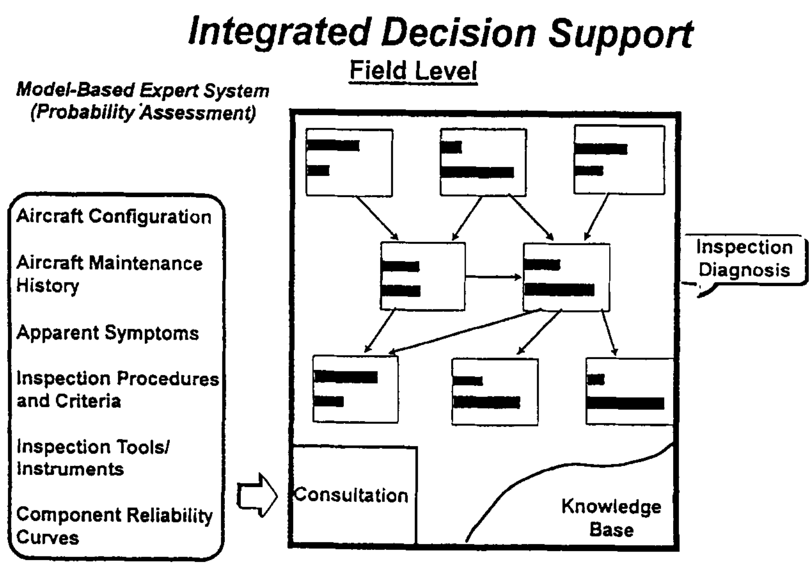


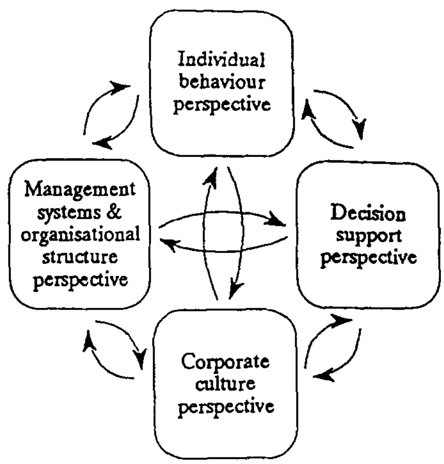
Fig. 8. Field level for intelligent decision system (IDS).

corrosion cases by aircraft type. This strategic analysis facilitates workload estimation for safety inspectors by aircraft type and is essential for the efficient scheduling of inspection routes at terminals.

The next phase of the IDS involves a field-level analysis and uses a model-based expert system, such as HUGIN, to assist the safety inspector with making probability assessments in specific aircraft diagnostic situations. As illustrated in Fig. 8, the expert system could also provide a consultation service on aircraft configuration, maintenance history, apparent symptoms, inspection procedures and criteria, inspection tools/instruments, and component reliability curves. The integration of both the strategic and field levels of decision support comprises an IDS for aircraft safety insepctions.

5. Decision support connected to organizational learning

One of the critical issues in designing decision support systems for modern organizations is to ensure that these systems are integrated and foster cross-functional learning. Riis and Neergard [39] present a new paradigm that considers the perspectives of individual behavior, decision support systems, management systems and organizational structure, as well as corporate culture. These multi-perspective learning model, presented in Fig. 9, also includes the formal and informal dimensions of an organization.



	Individual Dimension	Collective Dimension
Formal Dimension	The Decision Support Perspective	The Management Systems and Organisational Structure Perspective
Informal Dimension	The Individual Behaviour Perspective	The Corporate Culture Perspective

Fig. 9. Riis–Neergaard [39] model of multi-perspective learning.

If the intelligent decision system for Aviation Safety Inspectors is to have a successful implementation, these additional perspectives need to considered during the system design process. A Danish–Norwegian survey [40] reports that most enterprises tended to introduce technological means first and afterwards adjust the imbalance by applying organizational means.

To ensure that technology better, organizational development, and corporate strategy are coupled, the Riis–Neergaard model can be used to ask key questions during the system development process such as:

- What will be the consequences for the four types of learning of introducing the new decision support technology for Aviation Safety Inspectors?
- Which changes are required in the FAA's management systems, information technology, organizational structure and culture, as well as individual behavior in order to fully utilize the new decision support technology?

Such an approach will prevent the situation where the new decision support technology is developed in isolation and views the decision system development as an interactive, collaborative process with due regards to the mutual interrelationships to the other three learning perspectives. This integrative approach encourages the development of individual qualifications for Aviation Safety Inspectors rather than exclusively relying on the knowledge embedded in an expert system. The approach also places awareness on collective learning processes as implemented by formal organizational structure and management systems for planning and control of aircraft safety inspections and recognizes the role that informal systems play in operationalizing decision support concepts. Such an understanding is essential for a successful implementation of new decision support technology in large, complex organizations, such as the FAA.

6. Conclusions

The value or contribution of this research exists in the methods or techniques used to develop an integrated decision support system for aircraft inspection activities. Integration was examined on two levels: (1) integration of the various technical components of the decision support system and (2) integration of the decision support system with individual behavior, management systems and organizational structure, and organizational culture across both formal and informal dimensions.

The issue of multiple fault diagnostics where one symptom leads to several faults, many symptoms lead to one fault, or many symptoms lead to many faults creates a challenging problem for Aviation Safety Inspectors. Available symptom data may be misinterpreted or unused which may lead to the incorrect removal of an aircraft's component. This paper presents research into several aspects of safety inspections – predicting workload requirements, defining safety indicators, estimating component reliability, and making probability assessments during diagnostic procedures.

The Mintzberg et al. model was presented as one possible framework for understanding the inherent elements of the decision-making process in complex problem domains. This model is viewed as providing insight as to what decision making phases can best be supported by advances in computerized technology. The proposed framework presented in this paper for the development of an Intelligent Decision System for aircraft inspection activities has general applicability and attempts to combine the reasoning features of artificial intelligence techniques with decision-theoretic models such as influence diagrams. Research is underway to combine the components of neural networks and model-based expert systems to create a prototype IDS for Aviation Safety Inspectors.

Acknowledgements

The authors would like to acknowledge the support of the Federal Aviation Administration's SPAS program and Mr. John Lapointe and Mr. Michael Vu. Knowledge of the Riis–Neergaard [40] model of organizational learning was obtained during a sabbatical leave by the lead author at Aalborg University in Denmark.

References

[1] The Federal Aviation Administration Plan for Research, Engineering, and Development, Vol. I: Program Plan, Report # 100-591, US Department of Transportation, January 1989.
[2] Safety Performance Analysis System, "Functional Description Document," Report # VS FSIS-D-1106, US Department of Transportation, Volpe National Transportation Systems Center, Cambridge, MA, March 1992.
[3] Safety Performance Analysis System, "SPAS Performance Measure Definitions (Version 1.0)", Report # VR FSIS-D-1134, Department of Transportation, Volpe National Transportation Systems Center, Cambridge, MA, October 1995.
[4] M. Klein and L.B. Methlie, *Expert Systems: A Decision Support Approach*, Addison Wesley, New York, 1990.
[5] M. Santana, "Managerial learning: A neglected dimension in decision support systems", *Proc. of the 28th Ann. Hawaii Int. Conf. on System Sciences*, Vol. IV, pp. 82–91, 1995.
[6] S. Holtzman, *Intelligent Decision Systems*, Addison-Wesley, Reading, MA, 1989.
[7] H.A. Simon, *The New Science of Management Decisions*, Harper & Row, New York, 1960.
[8] A. Newell, and H.A. Simon, *Human Problem Solving*, Prentice-Hall, Englewood-Cliffs, NJ, 1972.
[9] H. Mintzberg, D. Rasinghani and A. Theoret, "The structure of the unstructured decision processes", *Administrative Sci. Quart.*, **21**, pp. 246–275, 1976.
[10] R.C. Eberhart and R.W. Dobbins, *Neural Network PC Tools: A Practical Guide*, Academic Press, New York, 1990.
[11] W.S. McCulloch and W. Pitts, "A logical calculus of ideas immanent in nervous activity", *Bull. Math. Biophys.* **5**, pp. 115–133, 1943.
[12] J.J. Hopfield, "Neural networks and physical systems with emergent collective abilities", *Proc. of the National Academy of Science*, Vol. 79, p. 2554–2558, 1982.
[13] J.J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons", *Proc. of the National Academy of Science*, Vol. 81, pp. 3088–3092, 1984.
[14] A. Maren, C. Harston and R. Pap, *Handbook of Neural Computing Applications*, Academic Press, New York, 1990.
[15] Wireman, T., *World Class Maintenance Management*, Industrial Press, New York, 1990.
[16] E. Page, "Spurring wider use of on-line condition monitoring for predictive maintenance", *Ind. Eng.*, **26**(11), pp. 32–34, 1994.
[17] Caudil, M., "Neural network training tips and techniques", *AI Expert*, **6**(1), pp. 56–61, 1991.
[18] *NeuroShell 2 User's Manual*, Ward Systems Group, Inc., Frederick, MD, 1993.
[19] R.C. Rice, "Repair database assessment", Report # VA-0013 (Battelle Contract No. DTRS-57-89-C-00006), Volpe Transportation Systems Center, March 1991.
[20] J.T. Luxhøj and T.P. Rizzo, "Probabilistic spares provisioning for population models, *J. Business Logisitcs*, **9**(1), pp. 95–117, 1988.

[21] W.J. Fabrycky, "Logistics systems design using finite queueing analysis", *Proc. of the Int. Logistics Congress*, Vol. II, pp. 65–7, 1981.
[22] W.J. Fabrycky, C.J. Malmborg, T.P. Moore and K.W. Brammer, "Repairable equipment population systems (REPS) demonstrator user's guide (IBM-PC)", Virginia Polytechnic Institute and State University, October 1984.
[23] F. Frisch, "Mortality and spareparts: A conceptual analysis", *Proc. of the 1983 Federal Acquisition Research Symp.*, pp. 467–480, 1983.
[24] J.T. Luxhøj and M.S. Jones, "A computerized population model for system repair/replacement", *Comput. Ind. Eng.*, **14**, pp. 345–359, 1988.
[25] J.T. Luxhøj, "Sensitivity analysis of maintained systems using a population model: A case study", *Int. J. Quality Reliability Management*, **8**(1), pp. 56–70, 1991.
[26] J.T. Luxhøj, "Replacement analysis for components of large scale production systems", *Int. J. Prod. Econ.*, **27**, pp. 97–110, 1992.
[27] R.S. Pindyck and D.L. Rubinfield, *Econometric Models and Economic Forecasts*, McGraw-Hill, New York, 1991.
[28] J.T. Luxhøj, T.P. Williams and H. Shyur, "Comparison of regression and neural network models for prediction of inspection profiles for aging aircraft", *IIE Trans.* (submitted).
[29] D. Specht, "Probabilistic neural networks for classification, mapping, or associative memory", *Proc. of the IEEE Int. Conf. on Neural Networks*, Vol. 1, pp. 525–532, 1988.
[30] D. Specht, "Probabilistic neural networks", *Neural Networks*, **3**, pp. 109–118, 1990.
[31] D. Specht, "A general regression neural network", *IEEE Trans. Neural Networks*, **2**, pp. 568–576, 1991.
[32] H. Shyur, J.T. Luxhøj and T.P. Williams, "Using neural networks to predict component maintenance requirements for aging aircraft", *Comput. Ind. Eng.*, in press.
[33] J.T. Luxhøj and H. Shyur, "Reliability curve fitting for aging helicopter components", *Reliab. Eng. and System Safety*, **46**, pp. 229–234, 1995.
[34] A. Badiru, *Expert Systems Application in Engineering and Manufacturing*, Prentice-Hall, Englewood Cliffs, NJ, 1992.
[35] X. Li, "What's so bad about rule-based programming", *IEEE Software*, pp. 103–105, 1994.
[36] S.R. Kumara, R.L. Kashyap and A.L. Soyster, "Artificial intelligence and manufacturing: An introduction", in: *Artificial Intelligence: Manufacturing Theory and Practice*, Institute of Industrial Engineers, Norcross, GA, 1989.
[37] F.V. Jensen, *Introduction to Bayesian Networks: HUGIN*, Aalborg University Press, Aalborg, Denmark, 1993.
[38] E. Horvitz, C. Ruokangas, S. Srinivas and M. Barry, "Project vista: Display of information for time-critical decisions", *Proc. of the 4th Rockwell Int. Conf. on Control and Signal Processing*, Anaheim, CA, 1992.
[39] J.O. Riis and C. Neergaard, "Managing technology projects: An organizational learning approach", *Proc. of the 28th Ann. Hawaii Int. Conf. on System Sciences*, pp. 514–522, 1995.
[40] J. Frick, F. Gersten, P.H.K. Hansen, J.O. Riis and H. Sun, "Evolutionary CIM implementation – An empirical study of technological-organizational development and market dynamics, *Proc. 8th CIM Europe-Annual Conf.*, Birmingham, UK, 1992.