Expert Systems with Applications 36 (2009) 6512–6519

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A fuzzy expert system for aviation risk assessment

Michael Hadjimichael *

Naval Research Laboratory, Marine Meteorlogy Division, 7 Grace Hopper Avenue, Monterey, CA 93943-5502, United States

ARTICLE INFO

Keywords:
Fuzzy rule-based systems
Risk assessment
Aviation safety

ABSTRACT

The Flight Operations Risk Assessment System (FORAS) is a risk modeling methodology which represents risk factors and their interrelationships as a fuzzy expert system. A FORAS risk model provides a quantitative relative risk index representing an estimate of the cumulative effects of potential hazards on a single flight operation. FORAS systematizes the process of eliciting human expertise, provides for a natural representation of the knowledge in an expert system, and automates the process of risk assessment. The FORAS tool is valuable to airline safety departments for examining risk trends, to pilots and dispatchers for assessing risks associated with each flight, and to airline management for quantifying the effects of making safety-related changes. The quantitative relative risk index generated by FORAS allows comparisons between flights, and facilitates the communication of safety issues throughout the organization.

Published by Elsevier Ltd.

1. Introduction

Quantitative assessment of risk is particularly challenging in domains where undesired events are extremely rare, and the causal factors are difficult to quantify and non-linearly related. One such domain is aviation safety. Of major concern to the commercial aviation community is Approach and Landing accident risk. A significant portion of hull-loss accidents occur during this phase of flight. Accidents which may occur at this phase include collisions with terrain, and runway undershoot, overrun, or excursion (Flight Safety Foundation, 2001). In current practice, the primary method of proactively assessing aviation risk, in an operational setting and on a per-flight basis, is by the use of checklists, which itemize a small set of weighted risk factors, and are completed by the flight crew before departure (Flight Safety Foundation, 2003).

In this article, the Flight Operations Risk Assessment System (FORAS), a fuzzy expert system for proactive risk assessment, will be presented. FORAS is a risk management tool that will assess various mishap risks associated with flight operations. It is a methodology for producing aviation risk models for air carriers, and includes software for the representation and application of those models. FORAS can be used by an airline to develop models for most types of risks, if there is a sufficient knowledge base and understanding of the risk within the organization. Such risks may include Approach and Landing accidents, turbulence-induced injuries, runway incursion, etc. We refer to these as risk categories or risk types.

Hundreds of combined years of safety experience and knowledge are embodied in the personnel of an aviation organization. The FORAS methodology systematizes the elicitation and encapsulation of human expertise in a risk model, provides for a natural representation of that knowledge in an expert system, and automates the proactive risk assessment process.

A FORAS risk model for any category is uniquely developed for each airline, as each airline is unique in its operations and understanding of risk. Such a model is an encoding of the human knowledge within the airline about a specific type of risk. It is designed to give safety managers and other users a quantitative, relative, assessment of a specific risk for an operation, which can be examined either by individual flight, or as a variety of subgroups, e.g., by fleet, region, or route. The assessment is performed using a mathematical model, expressed as a fuzzy expert system (Kandel, 1992; Zadeh, 1965) which synthesizes a variety of inputs associated with a particular flight, including information on crew, weather, management policy and procedures, airports, traffic flow, aircraft, and dispatch operations. Output per-flight is a single relative risk index inferred by the application of a complex set of rules to the input data. The system can also provide an indication of factors which contribute to greater than expected index values.

A FORAS risk model output is thus a measurement system that can determine the relative risk of a mishap. However, FORAS is not a *go/no-go* tool. It is a decision aid and a means of identifying flights with greater than typical risk level, by quantifying the complex interaction of factors which influence risk.

Whereas an absolute risk index might be interpreted as a probability of mishap, and can be considered in isolation from other flights, a *relative* risk index is an indicator useful only in comparison, and hence relative, to other indicators generated by the same risk model, for the same flight operator. The assessment is relative, because the system output is not an absolute measure of accident risk, but rather it is a number which increases and decreases as the

* Tel.: +1 831 656 5151; fax: +1 831 656 5025.
E-mail address: mike.hadjimichael@nrlmry.navy.mil.

0957-4174/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.eswa.2008.07.081

risk of a mishap increases and decreases. While all flights have a very low risk of serious mishap, flights can be compared to identify those flights where mishap risk is greater. With this analysis available, flight dispatchers can increase crew awareness and, given an indication of contributing factors, initiate mitigative actions. Operator management can implement strategies to manage risk to an acceptable level.

A FORAS risk assessment model is a fuzzy expert system, created by knowledge elicitation from the subject matter experts within the airline organization. The model represents risk as a hierarchical decomposition of contributing factors, whose interrelationships are represented by a fuzzy rule set. The decomposition of risk can help to identify those elements that contribute most significantly to the calculated risk.

The FORAS project was initiated by the Icarus Committee, affiliated with the Flight Safety Foundation, Washington, DC. The foundation's mission is to serve as an aviation safety "think tank" and to stimulate improvements in safety by influencing those in industry who can effect change. The committee is composed of individuals who have distinguished themselves in their particular fields of expertise and represents various segments of industry and regions of the world. The Icarus Committee formed a "Safety Index Working Group" whose purpose was to develop a working model of a safety metric with which an airline or other aircraft operator might manage, monitor and measure operational safety performance. The FORAS project, initiated in 1997, is the result of that initiative. Its goal is to develop a quantitative index for proactively assessing aviation risk, addressing a maximal number of features and moving the emphasis of risk reduction away from the measurement of accident rates, and focusing instead on the recognition of the risk factors involved in the aviation process. FORAS has been in development for several years, and has recently been used to implement a risk assessment model for the Approach and Landing phase of flight for EVA Airways (Hadjimichael & McCarthy, 2005, 2006).

The FORAS methodology described here is being applied to aviation risk, but in fact it is applicable to any domain where there is little probabilistic information available, but there is a great deal of human expertise and experience available. Such settings might include, for example, chemical and nuclear facilities.

The paper is organized as follows: in Section 2, relevant background material on aviation risk and aviation risk assessment is presented. Section 3 presents brief foundations of FORAS, followed by an in-depth description of the methodology and model development process. Section 4 concludes the paper and describes the continuing research track.

2. Background

2.1. Aviation risk

Flight operations are extremely complex, involving many components: human, mechanical, technological, and environmental. Consequently, the risks associated with flight operations are equally complex and diverse. Extensive research has been devoted to the analysis and management of these risks (Reason, 1997; Wells & Rodrigues, 2004; Wood, 2003). In the first application of this methodology, research will focus on modeling the risk associated with operations during the Approach and Landing phase of flight which includes three phases: approach, final approach, and landing. This phase is considered here to begin as the aircraft descends below 10,000 feet, and end when the aircraft comes to a stop on the runway. The most common types of incidents which occur during this phase are controlled flight into terrain (CFIT), loss of control, landing overrun, runway excursion, and unstabilized approach (Flight Safety Foundation, 2001). Approach and Landing accidents comprise a significant portion of world-wide commercial jet major accidents, as seen in Table 1 (Burin, 2007).

Studies by the Flight Safety Foundation have identified numerous direct and contributing risk factors (Flight Safety Foundation, 2001). Direct causes of Approach and Landing incidents include excessive speed, vertical position, failure to follow Standard Operating Procedures, and failure to go around. Common causal factors include: difficulties interacting with automation, disorientation/visual illusion, lack of training/experience/qualification, high/fast on approach, ATC incorrect advice/service/instruction, low/slow on approach, procedural violations, "press-on-itis," improper flight handling, lack of positional awareness, inadequate CRM (Crew Resource Management), procedural errors, and inadequate judgment/airmanship. Contributing circumstantial factors include: runway condition, inadequate ground aids, inadequate regulatory oversight, lack of safety equipment, inadequate regulation, inadequate training, management failure, inadequate procedures, inadequate CRM, and poor visibility. Most of these factors can be categorized broadly using terms such as "fatigue," "experience," "airport issues," and "weather." Many risk factors in aviation have been tied to human factors issues, and are studied and classified, although not always well-understood (Wiegmann & Shappell, 2003).

The goal of a risk assessment system is to identify these factors, weigh their relative influence, and provide enough information to raise awareness and prompt mitigative action.

2.2. Commonly used risk assessment in aviation

In aviation, much understanding of risk arises from accident analysis, as well as flight and operations modeling and simulation. Accident analysis may yield a great deal of knowledge about causal factors, but it is reactive, and potentially at great human and/or financial cost. Risk modeling approaches are typically aggregations of the collected knowledge resulting from accident and incident analysis, theoretical and empirical studies (e.g., effects of fatigue on human performance), and human experience.

A major challenge in aviation risk assessment is to be proactive, timely, and comprehensive. Most aviation risk assessment methods generally fall into three main categories:

1. Checklists used in the cockpit.
2. Probabilistic approaches which currently are providing more theoretical analyses, and
3. Digital flight data recorder data analysis which may indicate aircraft operation beyond accepted performance thresholds.

The most well-known risk assessment checklist is the Flight Safety Foundations CFIT Checklist (Flight Safety Foundation, 2003). This consists of weighted *yes/no* questions, the sum of which yields an index. Although relatively simple to complete, checklists are relatively unsophisticated: limited in scope and representative power, restricted to information immediately available to the pilot, and representing a linear model of risk. More in-depth reviews of risk assessment techniques for flight operations and aviation safety can be found in Place (xxxx), GAIN (2003).

Table 1
World-wide, major commercial aviation accidents, 2005–2007 (Burin, 2007)

Year	Major accidents	Occurred during Approach and Landing
2005	16	10
2006	11	6
2007 (through October)	13	9

The most well-known probabilistic approach considers the probability and severity of a set of outcomes. Risk is defined as the product of *likelihood* and *severity*. Deficiencies in these methods include the difficulty of determining both the probabilities of rare events (such as CFIT), and their severity. Further, the probabilities may be dynamic, and vary with a variety of factors which are not known in advance. Another probabilistic approach uses a Bayesian network of causal factors to related input factors to possible outcomes (Pearl, 1988). Such causal networks are dependent on the estimation or derivation of conditional probabilities. This approach and similar ones are well described in Luxhoj (2003).

Digital flight data recorder data is generally monitored by a Flight Operational Quality Assurance (FOQA) program. This data, generated by the aircraft during routine flight operations, is analyzed in order to reveal situations that may require corrective action, to enable early intervention to correct adverse safety trends before they can lead to accidents, and to provide an objective means of following-up on corrective action to determine whether it has been effective (Flight Safety Digest, 1998).

2.3. The FORAS approach

In contrast, a FORAS risk model is knowledge-driven and non-probabilistic. FORAS risk models do not depend directly on statistical probabilities. They are primarily based on the knowledge derived through extensive discussions (knowledge elicitation sessions) with subject matter experts. Furthermore, the emphasis is on representing the *process* of the flight operation (in terms of risks factors), rather than the *outcome.* A FORAS risk model represents risk as a hierarchical decomposition of risk factors. This representation is intuitive, easy to interpret, and facilitates the identification of primary causal factors.

The FORAS approach is similar to the work of Kangari & Riggs (1989) in its use of fuzzy methods for the expression of risk in the construction industry, although that work computes risk using probability and severity, and is not intended for low probability, high severity events. It is also similar to the work of Carreno & Jani (1993) in its use of hierarchical combinations of risk factors in a fuzzy expert system for insurance risk assessment. In that work, the risk model is much simpler, and limited to a uniform two-level hierarchy for all input variables. Fuzzy logic has also been applied to human factors and human reliability analysis in three-level hierarchies (Marseguerra, Zio, & Librizzi, 2007).

3. Flight operations risk assessment system

3.1. Foundations

A knowledge-based expert system is a collection of facts representing the knowledge of subject matter experts (Dym & Levitt, 1991; Kandel, 1992). Their knowledge is expressed as a set of inference rules in the form

if *antecedent* **then** *consequent*.

The antecedent clause is a test, and may take the form of an expression formed of logical conjunctions and disjunctions of (*variable, value*) pairs, or other logical expressions which evaluate to True or False. Typically, antecedent clauses are written as conjunctive expressions. A rule whose antecedent is in any other form can be re-written as a set of rules with only conjunctive form antecedent clauses (Andrews, 1986). If the antecedent clause evaluates to True, then we say that the clause is satisfied and the rule "fires," or is activated. As a forward-chaining process, facts asserted by the rule consequent may trigger additional rules to fire.

In a fuzzy expert system, the variables in the antecedent and consequent may be *linguistic*. That is, their values are expressed in natural language terms which easily represent the knowledge of the subject matter experts (Kandel, 1992; Zadeh, 1965). Fuzzy set theory specifies a method for the mathematical specification and interpretation of such values, and for performing logical inference using these values.

A fuzzy expert system is an ideal method for the representation and application of knowledge in a domain such as aviation safety, in which knowledge may be highly subjective and empirical, resulting from years of experience, accident investigations, psychological studies, simulations, and modeling. Subject matter experts prefer to describe their knowledge using terms such as "high" to describe the linguistic variable *experience*, or "severe" to describe the variable *work amount*, or "low" to describe visibility. Such terms are easily represented by fuzzy set membership functions, over universes of discourse such as *years flying*, *hours flown*, or *miles visibility*.

3.2. Overview

The FORAS methodology has two components: model development, and risk inference. In the development process, a risk model is created using domain knowledge, and is based on variables available in the organization's databases. The risk inference process then inputs individual flight operational data from those databases into a software representation of the model, and computes a risk index for each flight.

The FORAS methodology is based on five principles:

1. The focus of a FORAS risk model is prevention, taking a proactive approach which identifies mishap precursors. By focusing on process and precursors, a FORAS risk index reflects those variables which lead to an unsafe situation. Identifying these variables in advance leads to a raised level of awareness of risk, and the potential for mitigative actions.
2. The risk models are based on human expertise. They are created from the collective, unified knowledge and expertise of an organization's subject matter experts and their understanding of the underlying processes which may lead to accidents or incidents. This expertise may be in the specific operations and procedures of the organization, as well as general knowledge of arising from theoretical or empirical research (such as fatigue effects on human performance).
3. Risk analyses generated by the model must be rapid, consistent, and independent of individual program user bias. A FORAS risk model computes an index from a great variety of variables representing risk precursors, as identified by subject matter experts. If the data for an individual flight is available in "real-time," then the risk index for that flight will be available as quickly. Furthermore, a FORAS model is consistent. An analysis of the same flight data will always yield the same risk index, unlike the varying opinions of subjective human operators involved in the flight operation (dispatchers and crew) who have varying degrees of skill and experience.
4. Risk assessments must be quantitative, for ease of comparison and communication. A risk model outputs an index number for each analyzed flight representing the relative risk for a specific risk category. The index is expressed on a scale of 1–10, with higher numbers indicating greater relative risk. This number represents an estimate of the specific risk by considering all quantifiable contributing and mitigating factors. It is not a probability, but rather it is considered relative to a baseline value or the assessment of another flight or group of flights. Thus, a greater value indicates a greater likelihood of mishap. Such a

quantitative relative index allows comparisons among flights, and facilitates the communication of safety issues throughout the organization.

5. A FORAS analysis will be useful for the communication of risk to all levels of the organization. A quantitative risk assessment which is consistent and independent of user bias is easier to communicate to all levels of the organization. Depending on actual implementation, dispatchers may receive per-flight risk assessments, while safety managers and higher level management can receive monthly summaries and trends, consider the effects of policy changes and equipment investments, etc.

In summary, the FORAS process systematizes the process of eliciting human expertise, provides for a natural representation of the knowledge in an expert system, and automates the process of risk assessment. This model provides a quantitative relative risk index representing an estimate of the exposure of a flight system to a set of hazards. It is valuable tool to airline safety departments for examining risk trends, to pilots and dispatchers for assessing risks associated with each flight, and to airline management for quantifying the effects of making safety-related changes.

3.3. Representing risk

A risk assessment model is an organized set of causal factors which are all related to the risk being modeled. Some organizational schemes include checklists, causal influence diagrams and fault trees (Luxhoj, 2003), and hierarchical decompositions (e.g., FORAS) (Hadjimichael & McCarthy, 2005, 2006; Kangari & Riggs, 1989; Marseguerra et al., 2007). Causal factors are usually quantifiable variables and the organizational structure represents a computational model for the estimation and quantification of risk.

3.3.1. Hierarchical decomposition

The FORAS system represents the risk structure as a conceptual hierarchy of risk factors. In this representation, the decomposed risk category, e.g. Approach and Landing Accident Risk, is at the top (the root node) of the tree. High level concept terms, such as *crew functionality*, appear near the top of the hierarchy, while low-level data terms such as *length of duty period* are at the bottom, most specific, part of the hierarchy. Fig. 1 shows the top levels of a risk structure for Approach and Landing accident risk. Fig. 2 shows a portion of the low-level risk structure.

The hierarchical risk structure shows how a particular risk decomposes into its component parts. The hierarchy is represented by the mathematical construct known as a *tree*, and we adopt tree terminology when discussing the properties of the hierarchical decomposition (Ore, 1963). In such a hierarchy, each factor is represented by a *node*, and the factors below each node and connected to it denote that factor's decomposition, or sub-factors. The decomposition nodes are referred to as *child nodes*. Nodes with no children are referred to as *leaves*. The node above any node is its *parent node*. The node at the top of the decomposition tree, which has no parent, is called the *root* node. Thus, the node representing the risk category is root node of the tree. Given any node, X, in the tree, the length of the path between the root and X is referred to as the *depth* of node X.

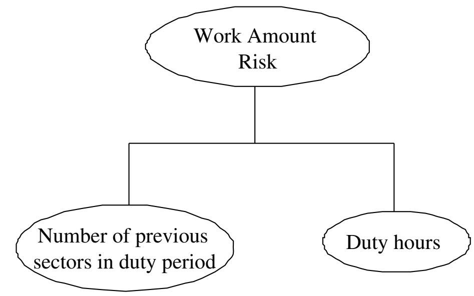

Fig. 2. Low-level portion of the risk structure decomposing Work Amount Risk.

A set of rules is associated with each node and its children. These rules represent the relationships and relationship strengths between risk factors. They indicate how the child factors influence the parent factor, and thus how to compute the value of the parent risk factor using the values of its decomposition sub-factors. Child nodes represent the rule antecedent variables while the parent node represents the consequent variable. We denote a set of rules as a sequence of tables. For a rule set with two input variables, A_1 (with m possible values $v_{1,1},\ldots,v_{1,m}$), and A_2 (with n possible values $v_{2,1},\ldots,v_{2,n}$), the set of possible output values for output variable R is represented by the $m \times n$ table, T. For rule $A_1 = v_{1,i}$ and $A_2 = v_{2,j}$, output $R = T_{i,j}$. A two-variable rule set is represented by one $m \times n$ table. A three-variable rule set (where A_3 has l possible values) is represented by l tables of size $m \times n$. This table notation is easily understood by subject matter experts as they assign output values for each input combination.

Fig. 3 demonstrates a rule set with nine rules, where antecedent variables *Previous sectors in duty period* and *Duty hours* combine to form consequent variable *Work amount Risk*. The antecedent variables have possible values *low, medium,* and *high*. These values have been selected by the experts as the most convenient for describing their knowledge. For example, the value *low* of variable *Duty hours* describes what a low number of hours flown is, in the context of work amount risk. The consequent variable has a value in the range 1–10 (10 represents the greatest amount of workload-related risk). In the typical rule set shown in Fig. 3, the upper left corner of this table (*Previous Sectors = Low, Duty Hours = Low*) indicates the rule:

If *Previous sectors* is low and *Duty Hours* is low then *Work amount Risk* is 1

Each decomposition is limited to at most three factors, in order to limit the number of rules for that node, and hence the complexity of the knowledge elicitation process. Assuming a maximum of three values per factor variable, three factors will require three

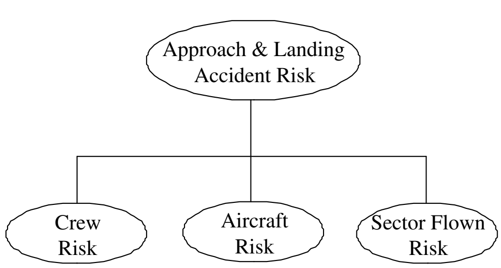

Fig. 1. High-level portion of a risk structure.

Work amount		***Previous sectors in duty period***		
		Low	*Medium*	*High*
Duty hours	*Low*	1	3	8
	Medium	3	6	9
	High	6	8	10

Fig. 3. Sample rule set defining Work amount Risk.

3×3 rule tables (27 rules). A fourth factor could require nine such tables (81 rules) which would be overly complex for knowledge elicitation purposes. To minimize complexity, decomposition into more than three nodes is restated by using an intermediate node, as in Fig. 4. The intermediate node represents a semantically meaningful variable which can be considered a "higher-level" conceptual parent of the two child nodes.

3.3.2. Representation by fuzzy expert system

Subject matter experts select the possible values for each of the risk factor linguistic variables, and provide corresponding definitions for the fuzzy values, such as *low, medium,* and *high* defined on corresponding universe of discourse. The universe of discourse of leaf node variables is determined by the semantic meaning of the variable, and a reasonable range of possible inputs. For example, a reasonable range of on-duty hours might be 0–24. Fig. 5 shows reasonable membership functions for the values of *Duty hours*.

The output variable range of integers 1–10 is similarly selected by the experts as being the most convenient language to discuss the consequent variable of each rule set. Each of these numbers is represented by a fuzzy set, a triangle membership function centered on the integer, as in Fig. 6 (Dubois & Prade, 1988). Thus, for all non-leaf nodes, the universe of discourse for the represented variables is the set of real numbers between 0.5 and 10.5. Of course, subject matter experts are free to select any range of output values for a rule consequent variable. Whatever the range selected, two or three fuzzy sets are defined on that range for use in the forward-chaining process, when the output variable value is used as an input higher in the hierarchy.

For example, from Fig. 3, *Work amount* is the consequent variable, and membership functions are defined to translate the calculated result of *Work amount* into various degrees of *low, medium*, and *high*, for the further calculation of the parent node of *Work amount* (Fig. 7).

Note that all non-leaf factors will be decomposed into at most three sub-factors each, with at most three possible values each, so at most $3 \times 3 \times 3 = 27$ rules are included in any single rule set.

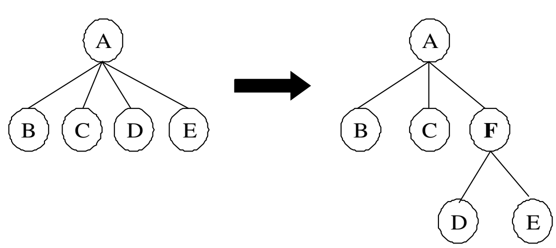

Fig. 4. One possible transformation of decompositions, to maintain a maximum of three child nodes per decomposition, by adding intermediate node **F**.

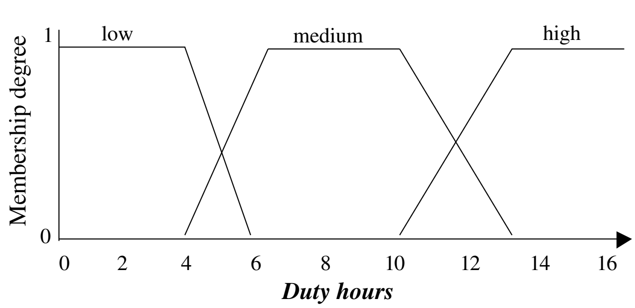

Fig. 5. Membership functions for fuzzy set values *low*, *medium*, and *high* of the linguistic variable for factor *Duty hours.*

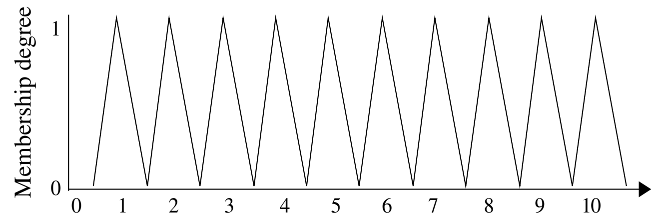

Fig. 6. Ten membership functions for fuzzy numbers 1–10. Each fuzzy number i represented by a membership function which has value 0 everywhere except from $i - 0.5$ to $i + 0.5$.

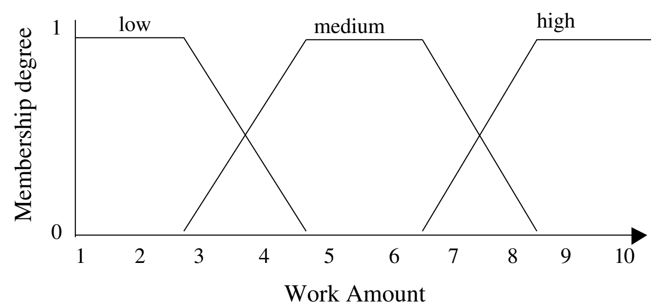

Fig. 7. Membership functions for fuzzy set values *low*, *medium*, and *high* of the linguistic variable *Work amount*.

3.3.3. Inference and risk analysis

Overall risk is calculated by applying all inference rules in a forward-chaining process, where the order of evaluation is bottom–up, following the order denoted by the risk hierarchy. Nodes are evaluated in an order according to their depth in the tree. Thus, all nodes at depth i are evaluated before evaluating any nodes at depth $i - 1$. This assures that all the necessary inputs are available for each node's calculation. Nodes of equal depth may be evaluated in arbitrary order. See the example of Fig. 8.

Values for the leaf node risk factors of the hierarchy are assigned directly from the input database variables specified by those risk factors. For the example of Fig. 2, crew roster data are retrieved from the database, and used to calculate the initial values of nodes *Number of previous sectors in duty period* and *Duty hours.* From these is calculated the membership function for the fuzzy value of parent node *Work amount*.

A fuzzy risk value is calculated for every non-leaf node by first calculating the values of its child nodes and then evaluating the rule set for that node. Child node values are generally crisp values when the child node is a leaf node, or fuzzy membership functions when the child node is non-leaf.

We use the typical definitions to determine fuzzy set membership, logical operators, and rule inference (Cayrol, Farency, & Prade, 1982; Cox, 1994; Zadeh, 1965). For a child node which has crisp value v, then for each fuzzy value F_i (e.g., *low, medium,* or *high)* with

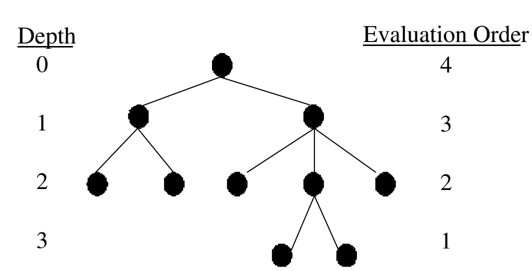

Fig. 8. Evaluation order is determined by node depth. Nodes of equal depth are evaluated in arbitrary order.

membership function $\mu_{F_i}(x)$ in the rule antecedent, the membership degree is simply $\mu_{F_i}(v)$. If the child node has a fuzzy value F_v, then the resulting membership function is defined as $\min(\mu_{F_v}(x), \mu_{F_i}(x))$ The fuzzy logical *and* operator is defined by min(*x*,*y*) (Zadeh, 1965). Define the activation degree, *a*, of the antecedent of the rule by $a = \mu_{F_i}(v)$ in the case of crisp child node values, and $a = \max_x(\min(\mu_{F_v}(x), \mu_{F_i}(x)))$ otherwise. Then for rule consequent *C*, the rule output value *C′* is defined by $\mu_{C'}(x) = \min(a, \mu_{F_c}x)$. The output membership functions for every rule in a rule set are combined using the *max* function to yield a single fuzzy value for the parent node.

In the example of Fig. 1, membership functions are first calculated for the fuzzy values of factors *Crew Risk, Aircraft Risk,* and *Sector Risk*. A rule set for *Approach and Landing Risk* specifies the calculation of the final membership function indicating overall risk.

Defuzzification is used to convert that final fuzzy set value *R* of the root node into a single risk index value. The *center of mass* method is chosen for defuzzification, as it gives the best representation to all the contributing components of risk, rather than only the greatest (as done by *mean of maxima*) (Zadeh, 1965). The center of mass method locates the *x*-coordinate of the center of mass of the area under $\mu_R(x)$:

$$\bar{x} = \frac{\int_{x\in U}(x \cdot \mu_R(x))\,\mathrm{d}x}{\int_{x\in U}\mu_R(x)\,\mathrm{d}x},$$

where $\bar{x}$ is the defuzzified output which best represents contributions from all fuzzy rules. The final model output risk index is equal to $\bar{x}$.

3.3.4. Critical parameters identification and mitigative actions

A valuable component of a risk assessment system is the ability to identify "critical parameters" – those model parameters which either singly or in combination, strongly and negatively influence the risk assessment. This is the subject of continuing research, but several observations may be made at this point.

Each model hierarchy component (subtree) evaluates to a fuzzy membership function on the universe of discourse defined by the interval [0.5, 10.5] (see Fig. 6). Defuzzification of this membership function will give a risk index for that particular component (a subindex). Using these intermediate indices, it is a simple matter to descend the hierarchy from the root, following the branch with the greatest subindex. Unfortunately, as each subindex is a defuzzified intermediate value, and because defuzzification is an imperfect representation of a membership function, comparisons between defuzzification values from different parts of the hierarchy, while useful, may lead to misleading results.

In general, in a particular risk assessment, critical parameters are risk factors which are likely "highly causal" to the assessment. Because any risk assessment is in fact a complex combination of influencing factors, the concept of criticality is not well-defined. An important ongoing effort is to develop a meaningful and useful definition of critical parameters, such that their identification in a risk assessment can lead to actions with the greatest possible mitigative effect. Possible definitions include "greatest contributors to risk assessment," "most sensitive input parameters," and "most deviating from baseline values." Implementations under consideration include:

1. Determination of parameters yielding a local risk minimum within a single rule set, using a risk index gradient. The risk index gradient is a gradient matrix indicating how risk changes as each rule set input parameter changes.
2. Determination of parameters yielding a global risk minimum within a single rule set, using risk index gradient (identifying the best possible combination of rule set input parameters).
3. Determination of parameters yielding a global risk minimum within the entire model, using sensitivity testing on global parameter set (re-evaluating a flight while changing individual model input parameters).

Once a suitable method has been established, a second expert system may be created to suggest mitigative actions whose purpose is to reduce the risk index for a flight, based on its risk assessment. For example, if high crew fatigue and crew inexperience are identified as a critical issue for a flight, then suggested action might be a change of crew.

3.4. Model development

3.4.1. Knowledge elicitation process

Hundreds of combined years of safety experience and knowledge are embodied in the personnel of an aviation organization. That experience is distributed throughout various departments, such as Safety, Flight Operations, Crew Rostering, Maintenance, Medical, Management, *etc*. The premise of the knowledge-based FORAS system is that such experience may be elicited from these aviation safety experts, combined into one unified risk model, represented mathematically in a software program, and used to evaluate the risk factors affecting current and upcoming flights. Thus, a basic and central component of the FORAS methodology is the knowledge elicitation process.

The FORAS knowledge elicitation process pulls together the distributed expertise into a single risk model by extensive interviews of subject matter experts. FORAS knowledge elicitation is a semi-structured interview process, generally proceeding in a top–down manner, decomposing each risk factor hierarchically into its contributing components, and in a bottom–up manner to classify known risk factors into higher level categories. Each discussion group explores general risk concepts and specific cases which highlight those concepts in order to establish or refine a set of primary risk factor categories such as Crew Factors, or Environmental Factors. Each primary risk factor discussed is decomposed into its contributing components, recursively decomposing subcomponents down to the quantifiable data variable level.

The decomposition is arranged hierarchically, as demonstrated in the example of Figs. 1 and 2. This method of structuring is comfortable for the experts, and is easily modeled in the expert system. The selection of risk factors represented in the model is constrained by the availability of data. Only those factors which will be quantifiable using available data from the organization's databases are retained in the model. Thus, once the risk structure is well-defined, it is pruned and revised to remove variables which are not accessible.

Rules are specified as demonstrated in the example of Fig. 3. As the risk hierarchy and rules are developed, they are reviewed multiple times by the contributing experts. In the aviation domain, significant contributions may be made by the instructor pilots, who have the advantage of a great deal of experience, as well as a familiarity with the issues of greatest concern to junior, less experienced pilots.

Several groups of experts within the organization are interviewed at least two times: an initial session, and a follow-up session to review and confirm that the knowledge elicited has been properly captured. The groups are drawn from various airline departments (pilots, dispatchers, safety managers), depending on the type of risk category being studied. Follow-up sessions continue as the model is continually enhanced and refined. Each group's membership is selected in order to interview experts in similar positions at the same time. This allows all members of the interview group to feel more comfortable by speaking in a group of their peers.

During interviews, differences of opinion may arise occasionally. Typically, such differences may be resolved by continued discussion and a careful definition of terms as differences may reflect different understandings of the issue in question. Differences may also reflect a poorly chosen decomposition, and may indicate that the higher level risk factor needs re-evaluation.

Once risk factors and the risk structure have been established, the rules are specified (as in Fig. 3). This is done using an interview process as before. Typically, a range of output values is selected (e.g., 1–10), and the extreme cases in the rule set table are assigned first. The remainder of the table is then completed. Each rule set is systematically considered, and later reconsidered for validation.

3.4.2. Model overview

The risk model resulting from the elicitation process at EVA Airways is extensive. The complexity and proprietary nature of some of the information in the model prohibit its complete inclusion in the paper. However, we may discuss an overview of the model variables whose presence in the model is common to most organizations' risk models. As illustrated in Fig. 1, the model variables affecting Approach and Landing Accident Risk are grouped according to three major categories: crew factors, aircraft factors, and departure-arrival city-pair factors. "Crew factors" refers to all factors which influence crew performance. These include the high level categories of intercrew communication, experience, and stress level. A sample intercrew communication variable is the quality of the crew pairing, while experience is measured in terms of both flying experience and airport familiarity. Stress level is determined primarily by considering the amount of short-term and long-term fatigue levels.

Aircraft factors depend primarily on the maintenance database, considering primarily any known malfunctioning equipment.

City-pair factors consider primarily the complexity of flying to, and landing at, the arrival airport. These complexities are dependent on the route flown, airport factors, runway factors, and environmental (weather) factors (which include visibility issues).

In total, about 44 non-leaf nodes and 80 leaf nodes represent the risk model structure. Variable relationships are expressed in approximately 200 tables representing approximately 1500 expert system rules.

Software has been written to compile an expert-specified risk model into a software implementation. Also, software and a graphical user interface have been written to tie the implemented risk model to databases supplying input data, accept user queries, perform the risk inference automatically, and display risk indices back to the user.

3.4.3. Testing and validation proposal

A functional Approach and Landing model is operational at EVA Airways. Validation of a knowledge-based system requires that we show that the knowledge represented by the risk model accurately represents the knowledge of the subject matter experts whose expertise was used to create the model. The first phase of validation is to be performed on a small set of flights. The FORAS risk model is used to generate an assessment for a set of flights which are independently ranked by the experts. The risk index is valid if the FORAS ranking of flights is similar to the expert ranking.

Further validation will be performed over a longer period, at least one year, with a close monitoring of results by the subject matter experts whose expertise is reflected by the model, and by an independent set of experts. Validation may also be possible in conjunction with digital flight data. Routine monitoring (post-flight) of digital flight data parameters identifies exceedance events during which certain parameters exceed allowable thresholds. These exceedances may be used as an indication of higher risk flights, and ideally there will be a positive and significant correlation between FORAS higher risk flights, and flights registering higher than normal exceedance rates.

During an extended validation period, the model will be tuned as necessary, and baseline values will be established for each sector flown. Such tuning may range from adjusting membership functions to re-writing rule sets.

4. Conclusion

This paper has presented a methodology by which the safety knowledge inherent in an organization such as an airline can be elicited, represented, and used for operational risk analysis on a flight-by-flight basis. The knowledge is represented as a risk model, based on human expertise, and thus representative of and specialized the specific experience of a particular organization. Because the model focuses on causal factors, it is useful as a proactive risk reduction tool and decision aid. Output of the model is normalized risk index which can be compared to a baseline value to determine relative risk. Such a rapid, quantitative, and consistent analysis is a valuable aid to communication of risk and safety issues within the organization.

Further research follows three tracks. A study of an operational implementation within an airline (EVA Airways) is now in progress. This study will provide a more in-depth analysis and validation of results, and determine the true usefulness of such a system. In addition, a more robust method of determining the "most causal" risk factors is necessary. This is a complex issue, as finding a meaningful and useful definition of "most causal" is a significant research challenge. Finally, when a useful definition has been found, a second expert system can be developed to suggest mitigative actions to the users.

Acknowledgements

The support of the sponsors NASA Aviation Safety Program, Naval Research Laboratory, EVA Airways, and TAP Air Portugal is gratefully acknowledged. The FORAS project was initiated by the Flight Safety Foundation Icarus committee, and its support, particularly via the steering committee of John McCarthy, Jim Burin, Jack Enders, and Doug Schwartz has been invaluable to the development of the project. The support of Irv Statler is also gratefully acknowledged.

References

Andrews, P. B. (1986). *An introduction to mathematical logic and type*. Orlando, FL: The Academic Press, Inc.
Burin, J. (2007). Aviation safety 2007: The year in review. In *CDROM proceedings, international air safety seminar, Seoul, October 1–4*.
Carreno, L. A., & Jani, Y. (1993). A fuzzy expert system approach to insurance risk assessment using FuzzyCLIP. In *WESCON/'93 conference record, San Francisco, 28–30 September*.
Cayrol, M., Farency, H., & Prade, H. (1982). Fuzzy pattern matching. *Kybernetes*(11).
Cox, E. (1994). *The fuzzy systems handbook*. Boston: AP Professional.
Dubois, D., & Prade, H. (1988). Fuzzy numbers: An overview. *Analysis of Fuzzy Information*, 9–10.
Dym, C., & Levitt, R. (1991). *Knowledge-based systems in engineering*. New York: McGraw-Hill.
Flight Safety Digest (1998). *Flight Safety Foundation, 17*(7–9).
Flight Safety Foundation (2001). Approach and Landing accident reduction toolkit. Suite 300, 601 Madison St., Alexandria, VA 22314.
Flight Safety Foundation (2003). CFIT checklist: Evaluate the risk and take action. *Accident Prevention, 60*.
Gain Working Group B (2003). Analytical methods and tools. *Guide to methods & tools for airline flight safety analysis*.
Hadjimichael, M., & McCarthy, J. (2005). Using a fuzzy expert system to represent risk: FORAS. *2004 SAE Transactions Journal of Aerospace*.
Hadjimichael, M., & McCarthy, J. (2006). Development of a fuzzy expert system for aviation risk modeling. *2005 SAE Transactions Journal of Aerospace*.
Kandel, A. (1992). *Fuzzy expert systems*. Boca Raton, FL: CRC Press.
Kangari, R., & Riggs, L. S. (1989). Construction risk assessment by linguistics. *Transactions on Engineering Management, 36*(2).

Luxhoj, J. (2003). Probabilistic causal analysis for system safety risk assessments. In *Proceedings, investigating and reporting of incidents and accidents, Williamsburg, VA, September 16–19*.
Marseguerra, M., Zio, E., & Librizzi, M. (2007). Human reliability analysis by fuzzy "CREAM". *Risk Analysis, 27*(1).
Ore, O. (1963). *Graphs and their uses*. Washington, DC: Mathematical Association of America.
Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Francisco: Morgan Kaufmann.
Place, S. Review of risk assessment techniques for flight operations. *Airport International*. <http://www.airport-int.com>.
Reason, J. (1997). *Managing the risk of organizational accidents*. Aldershot, UK: Ashgate.
Wells, A. T., & Rodrigues, C. C. (2004). *Commercial aviation safety*. New York: McGraw-Hill.
Wiegmann, D. A., & Shappell, S. A. (2003). *A human error approach to aviation accident analysis*. Burlington, VT: Ashgate.
Wood, R. H. (2003). *Aviation safety programs: A management handbook*. Englewood, CO: Jeppesen Sanderson.
Zadeh, L. (1965). Fuzzy sets. *Information Control, 8*.