NASA/TM–2015-218791

Bayesian Safety Risk Modeling of Human-Flightdeck Automation Interaction

Ersin Ancel and Ann T. Shih
Langley Research Center, Hampton, Virginia

August 2015

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TM–2015-218791

Bayesian Safety Risk Modeling of Human-Flightdeck Automation Interaction

Ersin Ancel and Ann T. Shih
Langley Research Center, Hampton, Virginia

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

August 2015

Acknowledgments

The authors would like to thank the support of AvSP and Systems Analysis and Methods (SAM) team members; Mary Reveley, Sharon M. Jones, Lawrence Green, Joni Evans (AMA, Inc.), and James Luxhøj (LCR, LLC) for their contributions to the development of the FLAP model. We are also very grateful to the subject matter experts for their participation in the knowledge elicitation sessions and insightful discussions on the model. This work was performed by National Institute of Aerospace (NIA) Task Order No. NNL14AA59T under NASA Contract No: NNL13AA08B.

Available from:

NASA STI Program / Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199
Fax: 757-864-6500

Table of Contents

List of Figures

List of Tables

1 Introduction

Usage of automatic systems in airliners has increased fuel efficiency, added extra capabilities, enhanced safety and reliability, as well as provide improved passenger comfort since its introduction in the late 80's. However, original automation benefits, including reduced flight crew workload, human errors or training requirements, were not achieved as originally expected. Instead, automation introduced new failure modes, redistributed, and sometimes increased workload, brought in new cognitive and attention demands, and increased training requirements [1, 2, 3]. Modern airliners have numerous flight modes, providing more flexibility (and inherently more complexity) to the flight crew. However, the price to pay for the increased flexibility is the need for increased mode awareness, as well as the need to supervise, understand, and predict automated system behavior [4]. Also, over-reliance on automation is linked to manual flight skill degradation and complacency in commercial pilots. As a result, recent accidents involving human errors are often caused by the interactions between humans and the automated systems (e.g., the breakdown in man-machine coordination), deteriorated manual flying skills, and/or loss of situational awareness due to heavy dependence on automated systems [5, 6].

This paper describes the development of the increased complexity and reliance on automation baseline model, named FLAP for FLightdeck Automation Problems. The model development process starts with a comprehensive literature review followed by the construction of a framework comprised of high-level causal factors leading to an automation-related flight anomaly. The framework was then converted into a Bayesian Belief Network (BBN) using the Hugin Software v7.8 [7]. The effects of automation on flight crew are incorporated into the model, including flight skill degradation, increased cognitive demand and training requirements along with their interactions. Besides flight crew deficiencies, automation system failures and anomalies of avionic systems are also incorporated. The resultant model helps simulate the emergence of automation-related issues in today's modern airliners from a top-down, generalized approach, which serves as a platform to evaluate NASA developed technologies.

2 Background Information

2.1 Objectives

The modeling effort discussed in this paper is part of a series of models that serve the NASA Aviation Safety Program's (AvSP) portfolio assessment by providing simulation capability for complex aviation accidents at the system level. These models[1] provide quantitative analysis capability, enabling the AvSP to assess the portfolio impact on the reduction of aviation system risk in current day operations.[2] Besides models, the AvSP synthesizes results of systems analyses, assessments, and studies for programmatic decision making and research portfolio prioritization and communication [9].

The focus of the FLAP model is on the effects of increased complexity and reliance on automation systems in transport category aircraft accidents and incidents. Consequently, the model aims to simulate contributors associated with man-machine interface breakdown, flight crew manual flight skill degradation, automation interface, overconfidence/complacency and simulator training, as well as automated aircraft systems failure and design. Given that the modeling requirements of the current effort are identical to those of the Loss of Control Accident Framework (LOCAF) model, captured in detail in Shih et al. [9], employing Bayesian Networks was deemed the most appropriate approach.

[1] The first model consists of aviation accidents caused by in-flight loss of control, captured in a model named LOCAF [8].

[2] Future versions of these models will provide operations in NextGen environment.

3 Literature Review

A literature review on issues associated with increased automation and its effects on flight crew was conducted. The available literature mostly consisted of anecdotal work; describing main problem areas associated with increased automation usage. One of the most cited works, the Federal Aviation Administration (FAA) Human Factors Team Report [10] addresses flight crew/flightdeck automation interfaces in commercial aircraft, and provides comprehensive information on the issues and recommendations.[3] Studies conducted by Billings [3], Sarter et al. [4], and Orlady et al. [11] also shed light on earlier issues encountered in automated systems as well as evolution of aircraft automation, which assisted in identifying primary issues simulated in this model.

The Commercial Aviation Safety Team (CAST) [12] analyzed 50 Part 121 incidents over the past 5 years involving energy state management and mode awareness. The study identified two root causes – "(1) inadequate training and system knowledge, and (2) unexpected incompatibility of the automation system with the flight regime confronting pilots in their normal duties" [12, p. 3]. In another observational study, the Flightdeck Automation Issues project conducted by Research Integrations, Inc. compiled 94 causes obtained from previous work, accidents and incidents, surveys, and experiments [13]. Another type of automation research involves pilot surveys. Survey-based studies collect information on pilots' attitudes about flightdeck automation and automation usage acceptance in commercial operations [14, 15]. Surveys provide valuable end-user application data and feedback to manufacturers in enhancing guidelines for the design and use of future automated flightdecks [14].

On a parallel effort, studies related to automation modeling were reviewed in order to help determine the approach followed within this study. The reviewed studies were mostly aimed at demonstrating specific automation systems and/or their components, man-machine interfaces, and human performance using both human-in-the-loop and human-out-of-the-loop modeling. Studies reviewed for modeling and simulation included – flight crew performance (involving attention, situational awareness, human cognition, multitask behavior, probability of failure to-complete, time-to master list items, proficiency, repetitions, etc.), automation system behavior, performed task, man-machine communication/interaction, and physical cockpit environment [16, 17, 18, 19, 20, 21, 22]. However, these studies mostly model a specific/unique system or subsystem actively present in the cockpit or the aircraft (similar to case studies).

4 Data Review

In accidents involving late model airliners, the essence of pilot error accidents is no longer related to "stick and rudder" or manual flying skills, rather, it is the efficiency of (system) monitoring of highly automated aircraft [3, 4, 23, p. 1]. Consequently, pilot error usually results in misalignment of automation system/modes, pilots' perceptions and actions, and aircraft state. Currently, common taxonomies and definitions for accident and incident[4] reporting systems such as the CAST/ICAO Common Taxonomy Team's (CICTT) Aviation Occurrence Categories or the accident types in National Transportation Safety Board (NTSB) categories don't have a dedicated group concerned with

---

[3] The Flightdeck Automation Working Group (FDAWG) [5] recently released an updated version of the 1996 report and findings that were compared against the FLAP model.

[4] NTSB defines an accident as an occurrence associated with the operation of an aircraft, which takes place between the time any person boards the aircraft with the intention of flight and all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage and, an incident as an occurrence other than accident, associated with the operation of an aircraft, which affects or could affect the safety of operations [24]. The words "event" and "mishap" are used to describe the collection of accidents and incidents throughout the text.

autonomy/automation-related issues.[5] Lack of a dedicated category prevents a comprehensive search of the dataset of such accidents, which, in turn, limits the scope of statistical analysis capability on automation-related accidents with respect to all accidents within a certain database/timeframe. The available data on automation-related accidents and incidents are not comprehensive or uniformly detailed. The Aviation Safety Reporting System's (ASRS) incident reporting is voluntary, and automation studies "cherry-pick" certain well-known automation-related accidents, which prohibit performing a statistical/correlation analysis over the findings.

The accident/incident data used to help the framework development include 50 mode awareness and energy management related incidents investigated within the CAST study [12], 46 automation-related worldwide accidents and major incidents investigated by the FAA [5], and 63 other accidents and incidents reviewed in other references and internet queries [3, 8, 14]. Given that these accidents and incidents are comprised of events in a large timeframe (1983–2009) and include multiple aircraft make and models with varying levels of automation along with the aforementioned shortcomings, this dataset was strictly used to identify key automation issues and causal factors and to help the SME elicitation process by providing case studies to illustrate node relationships.

5 Overview of the Modeling Steps

The FLAP modeling effort was comprised of two distinct phases. First, a generalized automation-related accident framework was developed, followed by the conversion of the framework to a Bayesian network model. The framework development is provided in Section 6, whereas the details of the FLAP model are given in Section 7.

The framework development is initiated with the literature and data reviews given in the previous sections. Following the reviews, a comprehensive list of causal factors contributing to automation problems was acquired and categorized based on responsible parties (e.g., flight crew, regulatory body, etc.). These causal factor categories were then organized within a hierarchical manner; similar to Reason's Swiss cheese model [25] used in the Human Factors Analysis and Classification System (HFACS) [26] and the previous modeling effort, LOCAF [8]. However, unlike HFACS, which only focuses on human failure/breakdown, this approach also takes automation system failures, as well as the flight crew-automation coordination breakdown into consideration. The causal factors (or nodes) and the connecting links within the framework are supported and documented by both past studies and accidents/incidents alike. The resultant framework is a generalized representation of automation-related accidents/incidents, capable of showing multi-dependencies among various automation stakeholders.

The next step involved the conversion of the framework into a quantitative model using a Bayesian approach via Hugin Software. The draft model was reviewed by Subject Matter Experts (SMEs) in order to obtain feedback, validation, and probabilistic data. Following subsequent calibration, the resultant model, called *baseline model*, is capable of providing preliminary causal factors and their probability values leading to an automation accident and/or incident. Finally, the model will be reviewed by internal and external panels before the AvSP products are inserted for portfolio assessment purposes. The model will then provide the effect of portfolio elements (also called products) in reducing automation-related events in today's aircraft operations. The next section discusses the framework, followed by the model development and node descriptions, AvSP product insertion, and data collection and results.

[5] ASRS incident database includes "human-machine interface" under human factors tab as a direct automation-related factor since June 2009.

6 FLAP Framework Development

The FLAP framework is structured to contain both latent and active factors, similar to HFACS architecture used in the LOCAF model or Reason's Swiss cheese model of accident causation. Latent levels are considered as the supporting components for the active levels. Active failure levels are the "pointy/sharp end" of the spear where the event takes place on the front line operator level, and are often directly linked to the accident or incident. Deficiencies or failures on each level are viewed as "holes." Undesirable events are caused by overlapping of these failures/breakdowns (holes) at latent and active layers [26]. The FLAP framework containing three latent and two active levels and their interactions are given in Fig. 1.

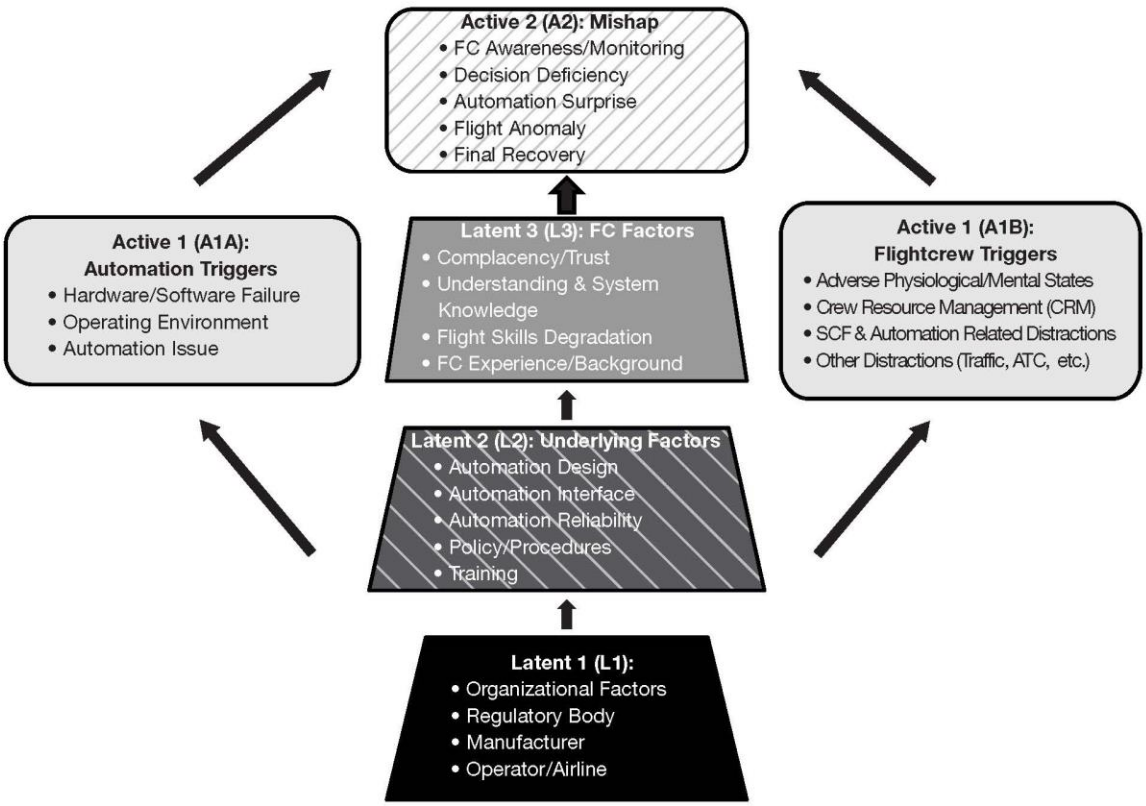

Fig. 1. FLAP framework overview.

The first latent level (L1) includes the major stakeholders within commercial airline operations; i.e., regulatory body (FAA, Directorate General for Civil Aviation, etc.), aircraft manufacturers, and operators. Second level latent (L2) factors include issues related to high-level underlying factors, including automation characteristics such as design, interface, and reliability, as well as airline policy/procedures and training practices. The third level consists of latent flight crew (FC) related factors such as complacency/trust, understanding and system knowledge, flight skills degradation, and experience/background.

There are two active causal factor levels in the framework. The first level active causal factors, (also considered as "triggers" or precursors), are divided into two sub-sections, Active A1A and Active A1B. These active failures take place during flight and stem from either automation system anomalies (A1A) or the underperforming flight crew (A1B) due to several reasons. The active causal factors directly affecting the outcome of the accident/incident are given in the Active 2 (A2) level. The underlying causes in A2 can stem from A1A, A1B, and/or L3 levels. For instance, the *Automation Surprise* node (A2 level) can be caused by automation triggers (A1A), latent flight crew factors (L3) or precursors affecting flight crew performance (A1B). The A2 level includes flight crew deficiency in system awareness, decision deficiency, and automation surprise, which can result in the flight anomaly and recovery. In addition to the links present in Fig.1, factors in active and latent layers may influence each other.

7 FLAP Model Overview

The framework presented in the previous section was used as the basis of the Bayesian Belief model using Hugin Expert software. In order to facilitate its representation and discussion, the model was divided into three sections using Hugin's object-oriented feature, which allows encapsulation of certain parts of the model. The FLAP model consists of the Top-Level in Fig. 2, as well as the encapsulated Automation and Flight Crew Conditions sections (called subnets) given in Fig.3 and Fig 4, respectively. This section provides the modeling method with the brief overview of the BBNs followed by the node definitions of all the encapsulated subnets.

7.1 Modeling Method Overview

7.1.1 BBNs

A Bayesian belief approach was used to model the complex flightdeck man-machine environment. Today's aircraft operate within a large, complex and safe air transportation system, where accidents rarely result from a single linear causal sequence. Instead, accidents and incidents usually result from the deficiencies among the interactions between aircraft system, humans, and external environment variables. The linear causal or time order approach present in the event tree/fault tree methods are not suitable to model such multi-dependent causal factors, whereas a BBN is a more intuitive and appropriate method, easily capturing the multiple non-linear dependencies [27].

A BBN is a directed acyclic graph representation of a network-based framework. BBNs contain a set of nodes that represent random variables and these nodes are connected via links designating the causal dependencies [28]. Within BBNs, random variables are represented via discrete (finite) or continuous (infinite) chance nodes.[6] For the discrete chance nodes, the function describing the dependency of the node on its parent nodes is given with a conditional probability table (CPT). Each node's CPT includes all the possible combinations of its parent nodes. The probability calculation is done by using Bayes' Theorem and the conditional probabilities obtained from the SMEs. In the BBN framework, probabilistic and causal relationships among variables are flexibly represented and executed as graphs, and can thus be visualized and easily modified. This facilitates model building and rapid interactive manipulations of the model to explore the causal features, which are particularly helpful when interacting with SMEs in the probability elicitation process [27].

7.2 Software Selection

A variety of commercially available BBN software packages are available in the market. An extensive review of this list of options revealed that the Hugin Expert v7.8, a commercial off-the-shelf software marketed as a decision support tool was found to be most compatible with the intended goal of evaluating various safety technologies [9]. The Hugin software is ideal for this task because it features a node category named "decision node," which represents a decision to be made by the user/modeler. Decision nodes are used to enable/disable certain technologies in the AvSP portfolio in the FLAP model.

---

[6] Only discrete (finite) chance nodes were used in the FLAP model.

One of the most prominent features of the Hugin software is the ability to create object-oriented Bayesian networks (OOBNs). The OOBNs include an instance of another network, also called a subnet. Instance nodes connect to other nodes using input and output modes, enabling a modular representation of the encapsulated network. This modularity not only allows a simpler representation of complex models but also enables the reuse of various individual subnets from different models within a larger more complex construct [9, 27, 29]. Hugin's object-oriented capability was used to represent the Automation and Flight Crew Conditions subnet, both linked to the Top-Level model.[7]

7.3 Conditional Probability Tables and Data Collection

The Hugin software uses CPTs to calculate downstream node probabilities. In the FLAP model, each causal factor node has a CPT with either two or three states, representing the node's outcome. The conditional probability values are gathered during the SME meetings (Section 13-15). In order to illustrate the data elicitation process, the CPT of the Top-Level *Manufacturer Management* node (with notional values) is shown in Table 1. Since this node has two parent nodes, there are 2^2, or four combinations of outcomes with two states (i.e., adequate or inadequate manufacturer management), yielding to a CPT size of eight.[8] The degree of beliefs of the SMEs is collected for each causal alternative where the sum of the two states adds up to 1.0. The experts were asked questions to provide probabilities considering the state of parent nodes that make up the four unique combinations. An example for such a query is constructed as "Considering that there is evidence to suggest that airline operator has inadequate management and regulatory body oversight is also inappropriate, how likely is that the supervision issues arise?" (i.e., the second column in Table 1). Similarly, the query for the last column will be a negatively constructed question stating that there is no evidence to suggest that there are operator management and regulatory oversight issues, etc. The alternating combination of the parent nodes' states comprises the values between the far left and far right columns [27].

Table 1. Manufacturer Management CPT

Operators/ Airlines	Inadequate Operator Management		Adequate Operator Management	
Regulatory Body	Inappropriate Regulatory Oversight	Appropriate Regulatory Oversight	Inappropriate Regulatory Oversight	Appropriate Regulatory Oversight
Inadequate Manufacturer Management	0.70	0.55	0.50	0.25
Adequate Manufacturer Management	0.30	0.45	0.50	0.75

[7] Note that Hugin Software allows encapsulation only if subnets do not receive any inputs from top level nodes, a requirement of directed acyclic graph structure of OOBNs to prevent directed cycles. In order to overcome this constraint and for the sake of discussion, the common causal factors nodes are duplicated in all three model sections/subnets. Actual model calculations are performed on a "flat" model, without subnets.

[8] For nodes with three states, with two parent nodes, the CPT size will be 12.

8 Top-Level FLAP Model

The Top-Level model is used to integrate the subnets and convey the causal factors to the event of interest; i.e., probability of an automation-related accident and incident. Similar to the structure shown in Fig. 1, the Top-Level model, shown in Fig. 2, includes several layers of latent and active factors and their interactions, described in detail in the sections below.

8.1 L1 Level: Latent Organizational Factors

The bottom three nodes include the L1 level organizational latent factors such as the *Regulatory Body*, *Manufacturer*, and *Operators/Airlines*. These nodes and their respective links are duplicated in Automation and Flight Crew Conditions subnets, but are only described in this section. The *Regulatory Body* node represents deficiencies within the regulatory process in both aircraft certification and flight standards of commercial transport operations. Example entities for this node include the FAA, the French Directorate General (Direction Générale de l'Aviation Civile, DGAC) or the United Kingdom Civil Aviation Authority. When evaluating flightdeck components or aircraft at large, issues like inappropriate representation of operational environment (e.g., workload, experience levels of typical line pilots, training of flight test pilot, etc.) or insufficient/deficient regulatory processes (e.g., slow, burdensome certification or reluctance to modify unnecessary requirements, etc.) are included in this node [5]. Due to heavy influence and oversight, which delineates a cause-effect relationship, the *Regulatory Body* node is linked to five other nodes – *Manufacturer*, *Training*, and *Operators/Airlines* (Fig. 2), *Automation Design*, *Automation Interface* (links shown in Fig. 3). Consequently, deficiencies found in regulatory practices can result in inadequate automation characteristics or manufacturer/airline management deficiencies [10]. The *Regulatory Body* node has two states – "adequate/inadequate oversight."

The *Manufacturer* node represents large aircraft manufacturers as well as automation system/avionic equipment manufacturing companies. The node includes deficiencies in manufacturers' organizational climate, resource management, regulation implementation, and supervisory/oversight, which could result in deficiencies in their final product (aircraft or automation systems) [8]. The node also covers deficiencies in automation design philosophy and approach [11, p. 233], level of automation – including over-automation [14, p. 3], economic benefit [30, p. 232], and standardization and cultural diversity [10, pp. 46–50, 13] that could eventually lead to human factors issues. The *Manufacturer* node inherently influences characteristics of *Automation Design*, *Automation Interface* (links shown in Fig. 3), and also affects *Training* and *Policy/Procedures* nodes [10, pp. 48–56,81,95,105] (Fig. 2). This node is a binary node with "adequate/inadequate manufacturer management" states.

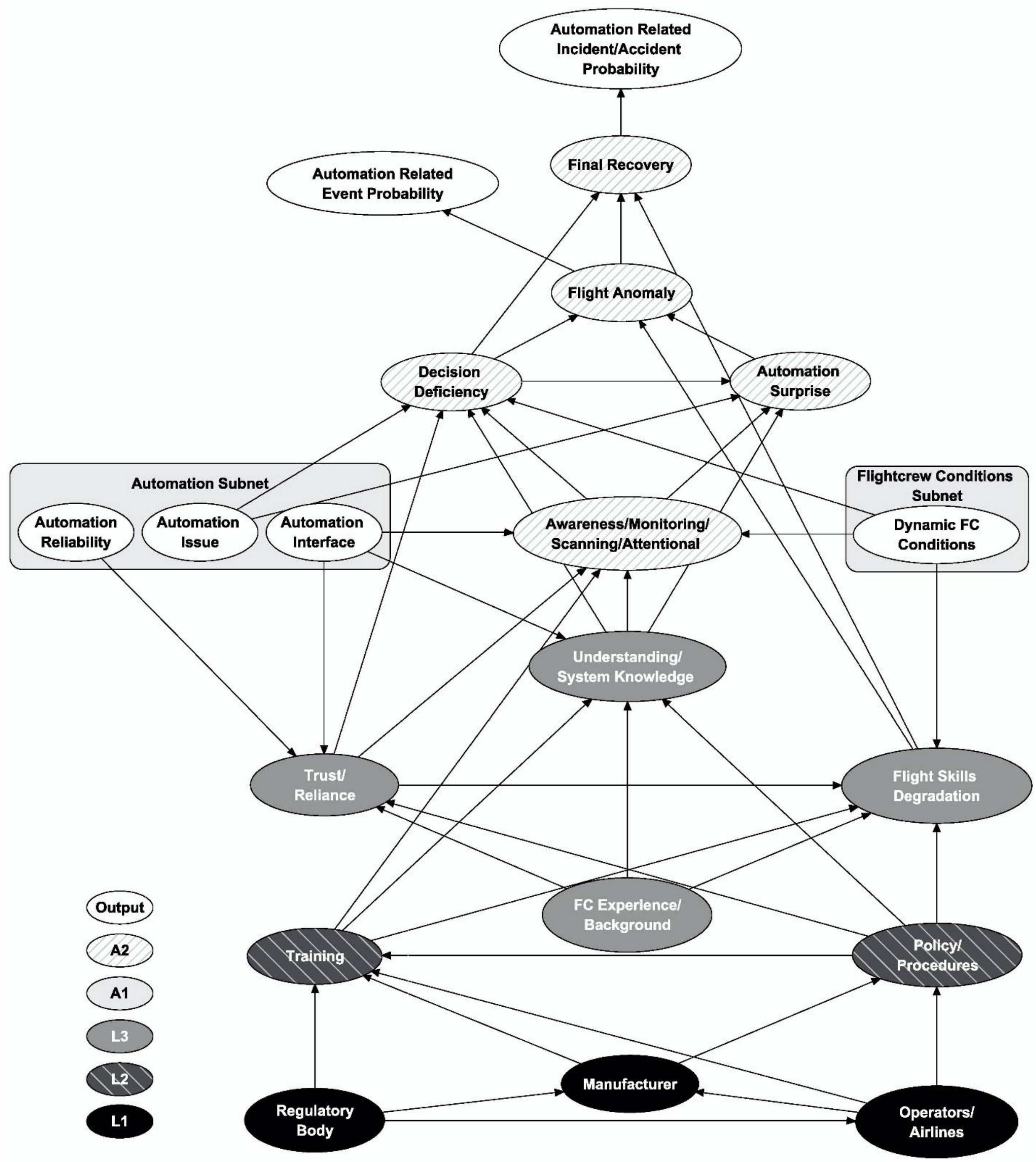

Fig. 2. Top-level FLAP model.

The *Operators/Airlines* node delineates the organizational aspects of corporate airlines as causal or contributing factors in automation accidents/incidents. This node includes deficiencies in organizational climate, resource management, supervision, training philosophy, and operational practices (crew scheduling, operational tempo), etc. These organizational deficiencies can trickle down and materialize as *Training* [10, p. 105; 15, pp. 313, 324; 30, p. 240] or *Policy/Procedure* [10, p. 65] issues. Also, lack of adequate supervision and management guidance can result in *Adverse Physiological* or *Mental States* in flight crews, links shown in Fig. 4. Finally, during the SME meeting, an additional link was added between the *Operator/Airlines* and *Manufacturer* nodes, representing unrealistic airline expectations/requirements imposed on the manufacturers, driven by economic motivation and operational efficiency. The *Operators/Airlines* node states include “adequate/inadequate operator management.”

8.2 L2 Level: Latent Underlying Factors

The second latent layer includes underlying factors affecting both automation systems and airline operations. These factors are *Policy/Procedures* and *Training* in the Top-Level Flap model in Fig. 2, and *Automation Design*, *Automation Interface*, and *Automation Reliability* (given in the Automation subnet, Fig. 3).

The *Training* node includes deficiencies associated with inadequate training due to cost cutting, scheduling conflicts, instructional errors or inappropriate planning resulting in the pilot not receiving the appropriate knowledge and skill set necessary to safely fly the aircraft [8]. Within the automation context, training issues are generally associated with limited resources and the common practice of “on-the-job” training of the remainder of automation functions that were left out during initial training [4, 10, 11]. This node is connected to several causal factors in the model. Different components of training are represented via the node it is linked to. For instance, *Training* to *Flight Skills Degradation* designates the basic flight (stick and rudder) component of training. Similarly, *Training* is connected to *Awareness/Monitoring/Scanning/Attentional* (*Awareness/Monitoring* for short) and *Understanding/System Knowledge* nodes as shown in Fig. 2 and Crew Resource Management *(CRM)* node in Fig. 4, representing respective training constituents. The binary states of this node are “adequate/inadequate training.”

The *Policy/Procedures* node covers deficiencies associated with inappropriate flight crew guidance caused by issues stemming from *Manufacturer* and *Operators/Airlines.* For instance, some examples of inadequately determined procedures are – operator procedures inconsistent with manufacturer recommendations and design philosophy (e.g., use of autobrakes or flight directors as designed), incorrect modification of procedures for economic or fuel saving reasons, or inappropriate level of proceduralization (too prescriptive or general) [5, pp. 53, 55, 10]. The *Policy/Procedures* node is connected to the *Flight Skills Degradation* and *Trust/Reliance* nodes to represent cases where operator guidance on automation usage level solely promotes automatic flight aircraft handling. Similarly, inappropriate guidance on aircraft control could interfere with pilot understanding and system knowledge and affect training procedures. The states of the *Policy/Procedures* node are “adequate/inadequate procedures.”

8.3 L3 Level: Latent Flight Crew Factors

The nodes *Flight Crew Experience/Background, Trust/Reliance*, *Flight Skill Degradation,* and *Understanding/ System Knowledge,* constitute the third level of latent causal factors.

The *FC Experience/Background* node is a leaf node (i.e., not affected by the presence of other causal factors) and is used to distinguish flight crew personal factors involving varying levels of experience. Using her survey data, Rudisill [14] highlights the differences between young, inexperienced and experienced pilots. Although younger pilots grasped automation fundamentals faster than their counterparts, their lack of experience caused them to be fixated on automation failures, causing loss of situational awareness and failure to monitor flight parameters like airspeed, altitude, navigation, etc. However, with extensive automation usage, experienced pilots may demonstrate degraded flight skills as well as over-reliance on automation since opportunities to practice manual skills are scarce. This node also includes differences of commercial pilots’ training backgrounds (military vs. civil aviation). Pilots with a military background transition to civil aviation with higher average flight time and exposure to a wider flight envelope than civil counterparts. In contrast, civil aviation collegiate level programs offer comprehensive highly structured programs with advanced flightdeck technology [5]. The two states of this node are “high/low exposure,” aimed at capturing pilots with various experience levels as well as backgrounds. The high exposure pilot is assumed to have received structured training, experienced with broader spectrum of flight attitudes and/or accrued high flight hours. The pilots with low exposure might possess low flight hours or high but routine operations mostly accrued in commercial operations with limited stick and rudder control and extensive automation emphasis. The FC *Experience/Background* node is linked to the *Flight Skills Degradation*, *Understanding/System Knowledge*, and *Trust/Reliance* nodes [4, 30].

The *Trust/Reliance* node includes FC complacency and inappropriate confidence level assigned to autoflight systems. Complacency is defined as the false sense of security as operators of a system come to rely on automation, generally carrying an unrealistic confidence in their personal efficacy [4, 15]. Accident and incident investigations point to pilot over-reliance on automation as contributor roughly in one out of four events where pilots wrongly believed they understand all aspects of the system [5]. The *Trust/Reliance* node is a parent for three causal factor nodes; namely, *Awareness/Monitoring* [4], *Flight Skills Degradation* [31], and *Decision Deficiency* [30]. The causal relationships between the *Trust/Reliance* node and the *Awareness/Monitoring* as well as the *Decision Deficiency* nodes are illustrated in FAA study on flight path management systems as “automation reliance reduces system awareness of the present and projected state of the aircraft and its environment, resulting in incorrect decisions and actions [5, pp. 36–37].” Similarly, as the FC relies more on the automation system, it is less likely that they will check on system and aircraft status or that they will intervene should any unexpected/undesirable behavior arise. Increased automation reliability, which possibly leads to complacency, was also linked to deterioration of basic position awareness skills [10, p. 43]. Additionally, pilots’ lack of confidence in their own skills on manual handling was also cited as the underlying reason for automation overreliance (linking *FC Experience/Background* to *Trust/Reliance)*. The states of this node are threefold – “overconfidence, under-confidence, and adequate trust,” which are mutually exclusive.

The *Flight Skills Degradation* node contains the erosion of manual flight skills (e.g., basic stick-and-rudder capabilities, flight control errors, instrument scan, etc.) due to continuous operation of autoflight systems and lack of practice, especially in FAR Parts 121 and 135. As a result of the transition from classic flight instruments and ground-based navigation to modern flightdecks, manual, visual or non-precision approaches are no longer employed, except when more advanced approaches (e.g., Global Positioning System/Required Navigation Performance or GPS/RNP with vertical guidance) are not available. Mainly due to lack of practice (but also due to lack of motivation and scheduling), an increase in manual handling errors was identified in a recent accident/incident analysis and was associated with continuous operation of autoflight systems [6, 11]. Examples of manual flight operation vulnerabilities include:

- Failure to “prevent, recognize and recover from upset conditions, stalls or unusual attitudes
- Inappropriate manual handling after transition from automated control
- Inadequate energy management
- Inappropriate control inputs for the situation [5, p. 31].”

Besides manual skills, the definition of this node expands to include decision errors associated with appropriate level of automation use and mixed-mode flying as well as usage of work-arounds implemented by the FC to address workload or inadequate procedures. Since the *Flight Skills Degradation* node receives input from the *Dynamic FC Conditions* node, it represents the pilots’ current status including physiological, mental factors and distractions besides latent factors like *Trust/Reliance*, *Training*, and *Policy/Procedures*. The *Flight Skills Degradation* node is connected to both the *Flight Anomaly* and *Final Recovery* nodes to represent the cases where degraded skills can cause flight anomalies or unsuccessful flight recovery, where automation support is not employed for various reasons (e.g., late runway change, Flight Management System (FMS) reprogramming restriction below 10,000 ft. or when the automation system is not reliable or available). The states for this node are “degraded/not degraded flight skills."

The *Understanding/System Knowledge* node refers to issues related to flight crew knowledge of aircraft systems or presence of gaps and misconceptions in their mental model of the system. Over the years, the proliferation of new flightdeck technologies substantially increased flight crew knowledge requirements. Consequently, over 40 percent of accidents and 30 percent of major incidents identified by a FAA study showed knowledge deficit [5]. Deficiencies considered in this node include understanding of autoflight modes, autopilot, autothrottle/autothrust, and flight management computer and their complex interactions, system couplings, operating procedures as well as diagnosing and debugging automation problems. Besides the flightdeck components, pilots' insufficient information on operating limits and aerodynamic capabilities of the aircraft also is included in this node. The *Understanding/System Knowledge* node is connected to the *Decision Deficiency* node since issues with pilot understanding are one of the primary sources of errors captured in the this node. Pilot understanding is also linked to the *Automation Surprise* and *Awareness/Monitoring* nodes. Lack of overall system understanding and knowledge can possibly prevent effective pilot mitigations to such hazards in cases where automation behavior is not well understood (e.g., automation surprises). Similarly, system awareness is lacking due to inadequate knowledge of that system (i.e., knowing where to look) [1, 10]. The states for this node are "adequate/inadequate system knowledge."

8.4 A2 Level: Active Factors/Mishap

The A2 or active failure level includes nodes associated with FC *Awareness/Monitoring*, *Decision Deficiency*, *Automation Surprise,* which lead to *Flight Anomaly* and *Final Recovery*.

The first node in the A2 level is the *Awareness/Monitoring* causal factor. Situational awareness was determined as a factor in over 50 percent of the accidents reviewed by FAA [5]. Situational (or system) awareness is defined as the "perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future [32, p. 5]." In the aviation environment, the definition translates as the ability of a FC to track and anticipate the behavior of a) the automation system variables and controls, b) the aircraft state and flight parameters (position, speed, flight path, energy state), and c) the operating environment (terrain, air traffic clearances, and traffic) [4, 10, p. 43]. Besides deficiencies in situational awareness, this node also encapsulates issues related to attentional deficiencies such as inadequate instrument scanning and inadvertent flight crew activation of aircraft controls, which potentially result in automation surprises. Several factors contribute to insufficient system awareness, including the automation interface (e.g., failure to notify indirect mode changes), training (e.g., insufficient methods/policies for monitoring mode changes), understanding/system knowledge (proliferation in the number of modes), flight crew complacency, workload status and flight crew conditions (presence of distractions and/or physiological mental states). Illustrating the issue, Parasuraman states that "gaps and misconceptions in an operator's mental model of a system as well as inadequate system feedback design can result in breakdowns in attention allocation which, in turn, can contribute to a loss of situation, or more specifically, system and mode awareness" [30]. The *Awareness/Monitoring* node is linked to two nodes – *Decision Deficiency* [10, p. 35, 30] and *Automation Surprises* [4, p. 6]. Node states are "inadequate/adequate system awareness."

The *Decision Deficiency* node includes all cognitive errors made by the flight crew. Examples of such errors include mode selection error (appropriate use of vertical speed mode, programming for a vertical navigation (VNAV) descent, etc.), mode confusion error (caused by indirect mode changes due to decreased awareness), flight programming error (lateral route, vertical restrictions), FMS error (using FMS for runway change, programming for FMS departure/arrival), checklist use/procedures errors (checklist workarounds/omissions, continue landing with unstabilized approach, failure to go-around, etc.), misdiagnosis of faults, following inappropriate automation generated directive, etc. These cognitive errors can stem from both flightdeck systems and pilot related issues alike. The *Decision Deficiency* node influences three nodes; *Automation Surprise* (via improper programming or mode errors), *Flight*

Anomaly, and *Final Recovery* (deficiencies in recognizing the anomaly and selecting proper mitigation strategy). The two states for this node are "decision deficiency/no decision deficiency."

The *Automation Surprise* term is coined by Sarter et al. [2], where the operator is surprised by the automation systems, unable to comprehend its current behavior or estimate future occurrences. Pilot survey studies identified the phenomenon via pilots asking questions like "what is it doing now, why did it do that, or what is it going to do next? [14]" *Automation Surprises* surface as the result of inadvertent pilot activation, unannounced/indirect changes in modes and subsequent aircraft behavior, or as decompensation incidents where the automation disconnects and transfers the control back to the flight crew due to exceedance of control limits. In cases where the flight crew's situational awareness is less than adequate, decompensation can lead to an automation surprise. The presence of automation surprises is one of the prominent causal factors for the *Flight Anomaly* node where the flightcrew recognizes that the aircraft is outside its flight envelope or restrictions via cues from aircraft systems (stick-shaker/pusher, bank limiter, configuration alerts, etc.) or air traffic control (ATC) interventions. The node states are "surprise/no surprise."

The *Flight Anomaly* node designates any departure from the intended flight plan or safe flight envelope that qualifies as an incident, caused by flight crew decision deficiency, flight skill degradation, and/or presence of automation surprises. The anomalies include aerospace deviation in altitude, speed, position, aircraft performance parameters as well as energy management deficiency, and aircraft entering a flight state without being properly configured. Depending on the anomaly, the aircraft could potentially experience stall, LOC, over-speed, loss of separation (and consequent near mid-air collision, mid-air collision), controlled flight into terrain, hard landing, or other accidents/incidents (crew/passenger injuries and/or damage to the aircraft). The node contains two states – automation-related "anomaly/no anomaly."

The *Final Recovery* node refers to the ability of the flight crew to recover from an abnormal flight condition defined in the *Flight Anomaly* node. Given that the model simulates an accident/incident environment, the *Final Recovery* node plays a decisive role in whether the incident turns into an accident. Factors affecting flight crew's ability to recovery are a) cognitive ability to recognize and determine the correct mitigation action (represented via *Decision Deficiency* node) and b) *Flight Skills Degradation* to capture flight crew's current physical and mental attitude as well as experience and communication (CRM) abilities, which are necessary to implement the necessary mitigation action. The *Final Recovery* node contains two states – "successful/not successful recovery." Both *Flight Anomaly* and the *Flight Recovery* nodes are connected to FLAP model output nodes, covered in Section 7-11.

9 Automation Subnet

As previously discussed, issues stemming from automation systems are compiled under the Automation subnet given in Fig. 3. The subnet includes two levels of latent factors and one active factor level, and provides three outputs nodes to the Top-Level FLAP model.

9.1 L1 Level: Latent Organizational Factors

Similar to the Top-Level model, the Automation subnet also follows the active/latent causation structure where the major stakeholders (*Regulatory Body*, *Operators/Airlines*, and *Manufacturers*) are represented at the bottom, as latent causal factors. The L1 level nodes are identical to those covered in the Top-Level FLAP model and are provided in Section 7-8-8.1.

9.2 L2 Level: Latent Underlying Factors

The underlying causal factor of the Automation subnet is the *Automation Design* node. Due to its complexity, automation designers and programmers are unable to reveal all the possible flight regimes, failure modes, and scenarios during tests. Issues like man-machine coordination breakdown often surface only after extensive operational experience mainly because system designers and developers are unaware of potential human factor problems [3, 11, 22]. For instance, Airbus A320/A330/A340's feedback issues due to uncoupled sidestick design was highlighted by the FAA's Human Factors Team in 1996 [10], yet 13 years after, this issue was a factor in Air France Flight 447 accident.[9] This node encompasses issues within the automation system design process including system, hardware, and software designs from preliminary phase to flight hardware including assumptions, requirements, testing/debugging, implementation, verification and validation, configuration management, quality assurance, etc. [33]. More specifically, the *Automation Design* node includes deficiencies associated with:

- system complexity where the details of the automation are difficult to understand and analyze in the event of a failure, due to tightly-coupled automation systems;
- proliferation in the number of automation modes with many of them achieving similar goals, which increases the training burden and interface complexity leading to increased risk of flight crew error;
- system coupling, referred to the internal relationships or interdependencies between or among automation functions, which are rarely documented, often leading to automation surprises;
- designation of authority/autonomy; and
- automation feedback design process including human-centered automation [3, 5, 10].

The *Automation Design* node is connected to all the downstream nodes in the Automation subnet since improper planning and execution of automation requirements can result in failures and unexpected behavior throughout the system. The states for this node are "adequate/inadequate automation design."

[9] http://www.bea.aero/en/enquetes/flight.af.447/rapport.final.en.php

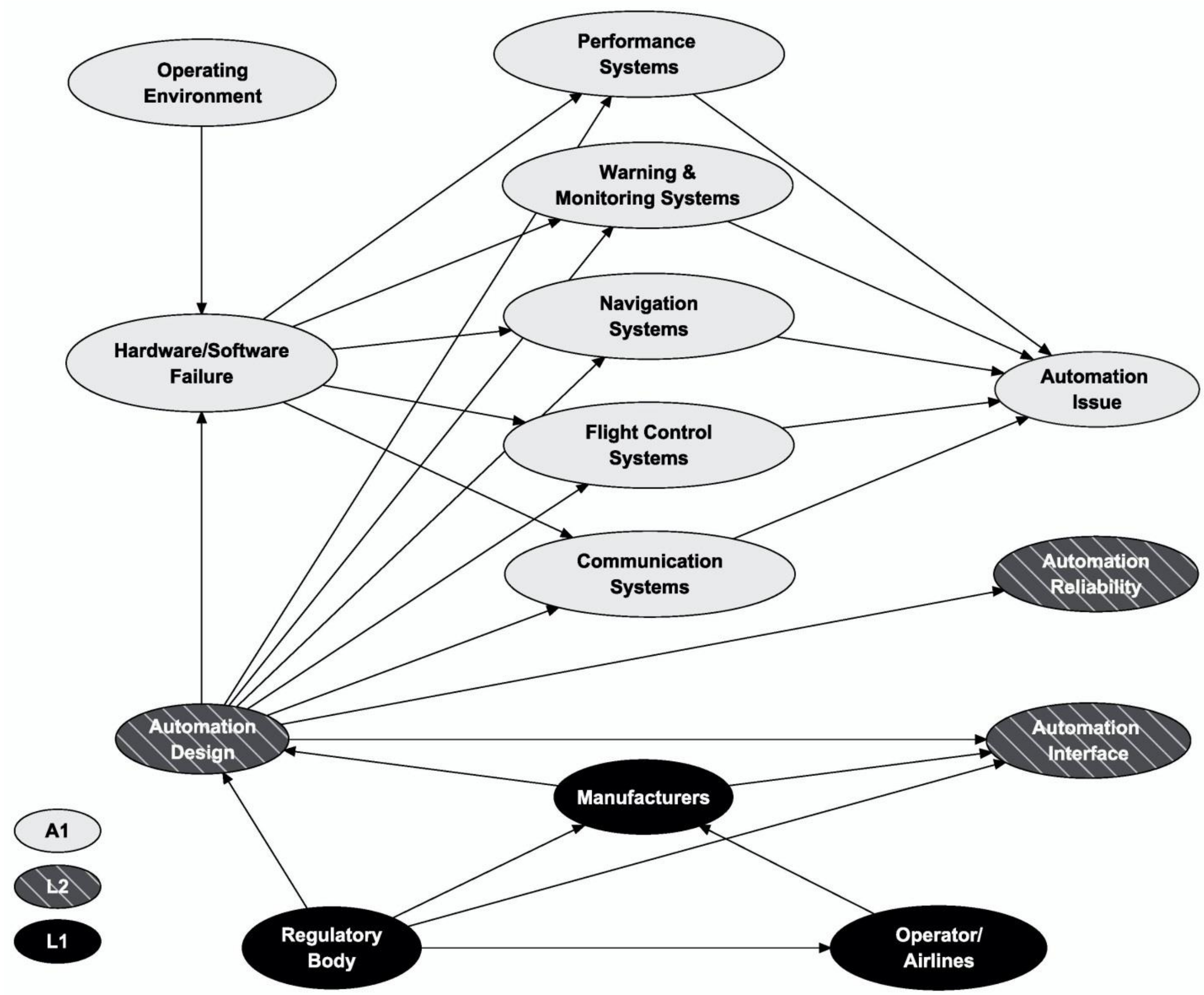

Fig. 3. Automation subnet.

9.3 A1 Level: Active Factors/Automation Triggers

The first active layer within the Automation subnet includes the *Operating Environment* node, the *Hardware/Software Failure* node, and five automation function nodes. The *Operating Environment* node provides external causes that potentially affect the operation of Hardware/Software (HW/SW) either by disrupting sensor outputs (e.g., cold weather, ice, volcanic ash) or by damaging aircraft systems directly (e.g., lightning, impact damage). Examples of mishaps involving cold weather include the Air France Flight 447 accident where ice crystal build-up on the pitot tubes resulted in the flight control system degrading to alternate law,[9] or the ValuJet Flight 558 accident, where cold weather prevented proper extension of nosegear struts that actuate the mechanisms to shift aircraft systems to the flight mode.[10] The states of the *Operating Environment* node are "environment issue/no environment issue."

---

[10] http://www.ntsb.gov/doclib/reports/1996/AAR9607.pdf

The *HW/SW Failure* node includes all glitches and malfunctions of the systems that were not anticipated by the designers, including malfunctions of antennas, sensors, or other measuring equipment that provide information to the automation systems downstream nodes. Presences of failures, malfunctions, or glitches affect the respective automation functions downstream by definition. The states of this node are "HW/SW failure/No HW/SW failure."

The *Performance Systems* function node includes issues associated with on-board performance computer (e.g., electronic flight bag system) function of the FMS, specifically, weight and balance, fuel weight, engine thrust limits, take off reference data (engine failure recognition speed – V_1, rotation speed – V_R, takeoff safety speed – V_2, and runway length requirements), maximum/optimum altitude calculations, or carrying out the projected altitude or speed targets (climb power, maintaining speed, etc.). The node also includes the improper sensor inputs (due to design or hardware/software failure) used in calculation and measurement of the performance goals [34]. The two states for this (and the following automation function nodes) are "failed/not failed."

The *Warning and Monitoring Systems* include three types of automated warning systems; a) aircraft configuration for current or upcoming flight phase (landing gear warning if the gear is not down while throttles are closed or flap warnings if not set up properly before takeoff), b) monitoring of aircraft systems (including hydraulic, fuel, pressurization or other systems), and c) presence of environmental threats (ground proximity warning system, traffic collision avoidance system, windshear avoidance systems, ice protection systems, etc.)[11]. The *Warning and Monitoring Systems* function node covers the failure of these systems due to both faulty design and/or HW/SW failure. Note that this node solely includes the design and implementation of the alarm system and assumptions whereas the issues related to ergonomic aspects of the warning systems (i.e., human perception and workload, de-cluttering, selection and characteristics of visual, auditory or tactile alerting systems) are included in the *Automation Interface* output node.

The *Navigation Systems* function node includes components and systems used in navigation such as GPS, Very High Frequency omnidirectional radio range, distance measurement equipment, area navigation, vertical navigation, instrument landing system, and other precision approach system components. This node also includes all the receivers used to capture frequency, range, bearing, localizer deviation, GPS position/ground speed, and time information [34].

The *Flight Control Systems* function node encompasses all the systems involved in automatic flight within the FMS and implementation of inputs via flight control surfaces. This node includes failures associated with flight control surfaces (ailerons, rudder, elevator, flaps, slats, spoilers, trims, etc.), flight control commands (roll axis, pitch axis, thrust axis), and modes (automatic pilot, autothrottle/thrust). This node also includes authority and autonomy related issues and assumptions such as envelope protection and stall/bank limiters. Additionally, inertial reference data (providing position, velocities, vertical speed, pitch, roll, heading, acceleration) as well as air data (supplying altitude, speeds, and temperature) are encapsulated within the *Flight Control Systems* node [34].

Finally, the *Communication Systems* function node includes data link (flight plans, clearance, weather, etc.) and surveillance systems (flight ID, aircraft state, trajectory conflicts, etc.). Systems like the Aircraft Communications Addressing and Reporting System, telemetry, communication radios, satellite links, telemetry, and Automatic Dependent Surveillance – Broadcast/Contract (ADS-B/C) are included in this node [34].

9.4 Automation Subnet Output Nodes

The Automation subnet provides three output nodes that are transferred to the Top-Level FLAP model – *Automation Issue, Automation Reliability,* and *Automation Interface*. The first output node, *Automation Issue,* provides the probability of an automation system exhibiting malfunction or failure, stemming from any of the five functional systems described above. Besides failures, inconsistent or unexpected automation system behavior itself can also be the root of flight crew confusion or decision deficiency and it is represented within the Automation Issue node. Unexpected automation behavior, which is captured in the *Automation Issue* node is also linked to *Automation Surprise* node [4]. The states of this node are "automation issue/no automation issue."

The *Automation Reliability* node is another direct output of the subnet, stemming from the *Automation Design* node. Automation reliability primarily affects the flightcrew's perception where highly reliable automation systems inherently increase reliance on automation. However, deficiencies found during automation design processes could potentially result in higher number of failures, which, in turn, affect the perceived system reliability by pilots [10], therefore, connecting the *Automation Reliability* node to the *Trust/Reliance* node at the top level. The two states for this node are "reliable/unreliable automation."

The final output node of this subnet is *Automation Interface.* This node is identified as one of the most prominent causes of man-machine breakdown due to its effects on flight crew situational awareness, pilot saturation and/or confusion [3, 4, 5, 10, 35]. As previously stated, this node contains the human-factor related aspects of the cockpit design. In order to prevent pilot distraction and nuisance during critical operating regimes (takeoff and landing), manufacturers prioritize alerts and inhibit the transmission of lower importance/unrelated alerts at certain flight phases. However, when performed improperly, the prioritization process has caused issues in the past [3, 35].[11] Inappropriate determination of the characteristics of visual (location, size, brightness, color), auditory (sound level, loudness, frequency deafness, location), and tactile alerting systems (intensity, vibration, area of body) as well as alert categorization issues (human perception and workload, de-cluttering) are captured in this node [35]. This node also includes non-intuitive flight crew interface including inadequate feedback (visual, aural, tactile regarding aircraft status, mode selections,[12] methods for annunciating of direct/indirect mode changes, etc.), and standardization (display symbology, nomenclature, and content on system synoptic and warning displays as well as differences in mode nomenclature and display among different aircraft types).[13] The *Automation Interface* node is linked to three causal factors; *Trust/Reliance*, *Awareness/Monitoring* [4, 10, p. 43], and *Understanding/System Knowledge* [5, p. 78] as shown in Fig. 2. The states for this node are "adequate/inadequate interface."

10 Flight Crew Conditions Subnet

The Flight Crew (FC) Conditions subnet provides dynamic FC conditions to the Top-Level FLAP model, in an approach similar to Reason's model of accident causation [25]. The subnet considers external and internal distraction sources as well as psychological and physiological aspects of FC performance. As in the Automation subnet, the FC Conditions subnet includes active and latent layers, providing one output node to the Top-Level FLAP model (Fig. 4).

[11] Decluttering/simplification of Primary Flight Display (PFD) on Airbus A330 was one of the causal factors in the accident during the test flight in Toulouse Blagnac Airport, France, since it decreased the system observability by hiding the active automation mode from the flightcrew [4].

[12] Different methods to identify the active mode such as pushed button illumination or mode annunciator indicator in PFD.

[13] Modes intended to accomplish a similar object might have different names and use different nomenclature for the flight crew interface. For instance, in some aircraft, vertical navigation used within FMS is referred to as "VNAV," while in others, these modes are called "PROF" for profile or managed navigation.

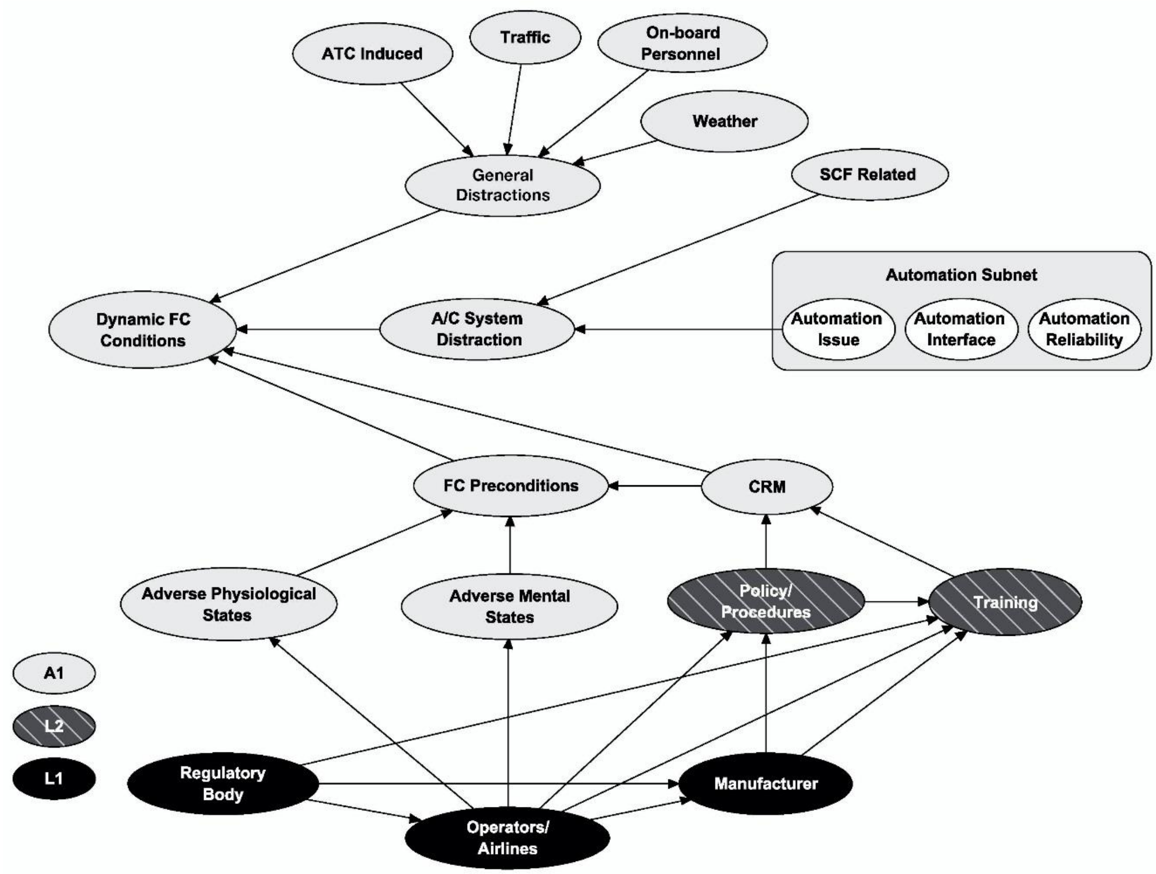

Fig. 4. Flight crew conditions subnet.

10.1 L1 Level: Latent Organizational Factors

Similar in structure to the Automation subnet, the FC Conditions subnet share the same three latent organizational factors; *Regulatory Body*, *Operators/Airlines*, and *Manufacturers*, which are placed at the bottom of the subnet. These nodes are identical to those covered in at the Top-Level FLAP model and Automation subnet, and are provided in Section 7-8-8.1.

10.2 L2 Level: Latent Underlying Factors

The FC Conditions subnet contains two latent underlying factors; *Policy/Procedures* and *Training*. Both of these nodes are parent causal factors for the CRM node, which is considered as A1 level, both pre-flight and in-flight sections. The *Policy/Procedures* and *Training* node descriptions are given in Section 7-8-8.2.

10.3 A1 Level: Flight Crew Conditions Active Factors – Pre-Flight

The group of causal factors in this layer helps determine the probability of FC readiness before the flight takes place. The *FC Preconditions* node is an aggregation node that considers flight crew adverse physiological and mental states as well as pre-flight CRM deficiencies and determines how fit the crew is for the upcoming flight. This node is linked to the *Dynamic FC Conditions* node, providing a baseline, which is then updated by considering the presence of distractions throughout the flight. The states of this node are "ready/not ready crew."

The *Adverse Physiological States* node includes physical fatigue (lack of sleep or demanding flying schedule), medical illness, excessive physical training, impaired physiological state, physiological incapacitation, self-medication, violation of crew rest requirement, and violation of bottle-to-throttle requirement [8]. The sole parent of this node is the *Operators/Airlines* node since inadequate supervision and operations within the company can result in insufficient crew rest period. The node is connected to the *FC Preconditions* node, which then connects to *Awareness/ Monitoring* node, which is greatly affected by the presence of physiological issues. The states of this node are "physiological issue/no physiological issue."

The *Adverse Mental States* node include complacency, distraction, *get-home-itis*, misplaced motivation, mental tiredness, distraction, confusion, depression, and/or alcoholism [8]. The presence of adverse mental states can be the cause of decision deficiencies (via the *Dynamic FC Conditions* node), which include omission errors. The states of this node are "adverse mental states/no adverse mental states."

The *Crew Resource Management (CRM)* node includes deficiencies like communication skills and coordination that take place among the flightcrew as well as between other entities (i.e., other aircraft, ATC, maintenance facility and other support personnel) before, during, and after the flight [8]. The CRM node covers formalized confirmation and cross-verification of selected modes via verbalizing, verifying, monitoring as well as confirming, analyzing, monitoring and intervening. Improper CRM (as erroneous communication between pilots) was one of the major contributors in around 40 percent of accidents identified by FAA study [5]. CRM deficiencies surface as cross-verification errors and crew coordination problems, including workload management. The node is linked to both *FC Preconditions* and *Dynamic FC Conditions*, indicating issues associated with the pre-flight briefing as well as in-flight communication and coordination, respectively. The states for this node are "adequate/inadequate CRM."

10.4 A1 Level: Flight Crew Conditions Active Factors-In-Flight

This part of the subnet provides the model with the updated FC conditions that are present during the flight. Along with the input from the *FC Preconditions* node and several sources of external distraction, the ability of the crew to perform flying duties is aggregated in the *Dynamic FC Conditions* node and outputted to the Top-Level model (node details are covered in Section 7-10-10.5). Possible distractions the flight crew may encounter during a flight are divided into two categories – general distractions and aircraft system related distractions. The in-flight component of the *CRM* node is also a causal factor in *Dynamic FC Conditions*, especially considering the crew's ability to manage emergencies and distractions.

The *Aircraft (A/C) Systems Related Distractions* node is an aggregation node that provides the probability of flight crew to get distracted by the presence of a) *Automation Issues* and b) *SCF*. Distractions stemmed from troubleshooting autoflight system anomalies/behavior as well as reprogramming the FMS are captured in this node. The *SCF* related node provides all other failures within the aircraft systems that are not associated with autoflight systems, such as the minimum equipment list. Presence of aircraft system distractions inherently increases probability of flight crew experiencing fixation or saturation since they require identification and mitigation. The *SCF* node within this model, however, is not aimed to cover accidents/incidents where system component failures are determined as cause or contributors. The node states are "presence/absence of A/C systems related distractions."

Another aggregation node, *General Distractions,* takes into consideration all other major sources of mental disturbance that potentially result in fixation or absorption. The causal factors considered are fourfold; *Traffic, ATC, On-board Personnel,* and *Weather.* The presence of traffic causing the crew to actively search for the surrounding aircraft was identified as a source of distraction in many mode awareness and energy management incidents [12]. Similarly, issues within the information flow to and from ATCs were identified in several references and accidents/incidents [12]. During the high workload phase of flight, frequent changes in the flight trajectory, where the pilots have to interpret, plan, and execute new clearances or issues with ATC communications, were deemed disruptive to crew situational awareness [2]. Besides ATC and traffic, numerous cases where cabin crew interference with cockpit (passenger emergencies/requests) or presence of the FAA Enroute to check the pilot during an initial operating experience flight were also identified as causes for the flight crew overlooking the autoflight systems [12]. One last causal factor for distraction was determined to be the presence of weather. "Adverse weather is a threat that is present in almost 60 percent of all flights" and it is managed by pilot mitigation [5, p. 30]. The presence of icing, fog/visibility issues, thunderstorms, rain, wind, and low ceiling are examples of adverse weather that cause additional mental work to the flight crew [12]. All four causal factors linked to the *General Distractions* node are comprised of two-state nodes (i.e., presence/absence of traffic, ATC, On-board personnel, and weather).

10.5 Flight Crew Conditions Output Node

As previously mentioned, the *Dynamic Flight Crew Conditions* node is the sole output of this subnet and it is linked to key causal factors including *Decision Deficiency, Flight Skills Degradation,* and *Awareness/Monitoring.* This node takes into consideration the presence of distractions (both aircraft related and other distraction sources) and crew related issues (CRM, as well as presence of on-board personnel, such as check airman or cabin crew) in order to calculate the probability of a crew member suffering fixation or absorption. Fixation is described by "being locked into one task or one view of a situation even as evidence accumulates that […] the particular view is incorrect" [10, p. 59] or "failure to revise situation in presence of new conflicting information" [4, p. 5]. A fixated pilot may still believe that an unstabilized approach can be salvageable even when the rest of the crew, ATC, and aircraft instrumentation suggest otherwise. On the other hand, absorption is a state of mind where the pilot is focused on one single task such as FMS programming or flight management computer troubleshooting while discarding others (also referred to as task shedding). In the cases of inadequate CRM with poor workload management, presence of distractions could potentially lead to higher probability of fixation and absorption. The *Dynamic Flight Crew Conditions* node is comprised of three mutually exclusive states; "fixation, absorption, and no effect."

11 FLAP Model Output Nodes and Preliminary Results

As shown in Fig.2, the Top-Level model output nodes are *Automation-Related Event Probability* and *Automation-Related Incident/Accident Probability.* Given the risk is defined as the product of likelihood of event and its severity, the model provides risks associated with heavy automation in today's aircraft operations. This is achieved by providing likelihood (probabilities) of such events occurring (via *Automation-Related Event Probability* node) and their severity (via *Automation-Related Incident/Accident Probability* node, which distinguishes between accidents and incidents).

The model is designed to identify the prominence of automation-related issues among all foreseeable accidents and incidents. For that reason, by definition, the *Flight Anomaly* node provides the probability of an in-flight upset resulting from the combination of the upstream automation-related nodes. Consequently, the *Automation-Related Event Probability* node reflects the ratio of automation-related events among all accidents and incidents. This node's states are "automation anomaly/no automation anomaly." The uncalibrated preliminary results of the FLAP model indicate that around 78 percent of all U.S. based accidents and incidents in today's aircraft are related to pilot's automation usage (probability of the "automation anomaly" state of the *Automation-Related Event Probability* node). The remaining 22 percent of these events are not tied to automation.

The other output node, *Automation-Related Incident/Accident Probability,* is used to provide the accident/incident ratio using the *Final Recovery* node. The *Flight Anomaly* node probability, combined with that of the *Decision Deficiency* and *Flight Skills Degradation* nodes, are the inputs to the *Final Recovery* node, which determines the probability of an incident evolving into an accident. The assumption is such that recovered incidents remain as incidents (e.g., correction of over-speed or stall situation) where unsuccessful recovery efforts may result in injury to crew and passengers and/or damage to the aircraft, hence an accident. The states of the *Automation-Related Incident/Accident Probability* node are "accident/incident." The preliminary results indicate that around 2.7 percent of all incidents would result in an accident, and the remaining 97.3 percent of all automation events will remain as incidents.

Besides the output nodes, all other causal factors nodes can be accessed individually. The preliminary results show that around 80 percent of all events involve flight crew *Decision Deficiencies,* whereas 72 percent of the cases were tied to failure in situational awareness (*Awareness/Monitoring* node). The *Flight Skills Degradation* node played a role in around 70 percent of all the cases and the *Automation Issue* node, which includes malfunctions and unexpected automation behavior, was present in 50 percent of the cases. It is important to note that these values do not represent the final values of the modeling effort, nor are they intended to be used in flightdeck automation, policy or decision-making processes. The next section provides an overview on how the model data is employed.

12 AvSP Product Insertion Process

NASA's AvSP is responsible for developing methodologies and technologies (referred to as products) to improve air transportation safety within the NextGen environment. The AvSP is comprised of three projects, namely, Vehicle Systems Safety Technologies project, System-Wide Safety and Assurance Technologies project, and Atmospheric Environment Safety Technologies project [36]. As previously discussed, the goal of the modeling effort is to gauge the impact of the products developed within these projects on current and future aviation risks. In order to do so, the products will be inserted into the model by the SMEs and then verified by involved stakeholders to ensure proper placement. The next step involves SME re-evaluation of affected nodes' probability by considering the effect of the product. Owing to the Bayesian Belief structure, the benefits of the products are propagated downstream from the affected node. For instance, a new methodology that potentially improves pilots' situational awareness is applied at the *Awareness/Monitoring* node and can prevent a *Flight Anomaly,* hence lowering the automation-related accident/incident probability. The AvSP portfolio analysis employs the relative impacts of various products to determine their individual and collective benefits. The rectangular nodes, shown in Fig. 5, represent notional products and the arrows designate the affected nodes.

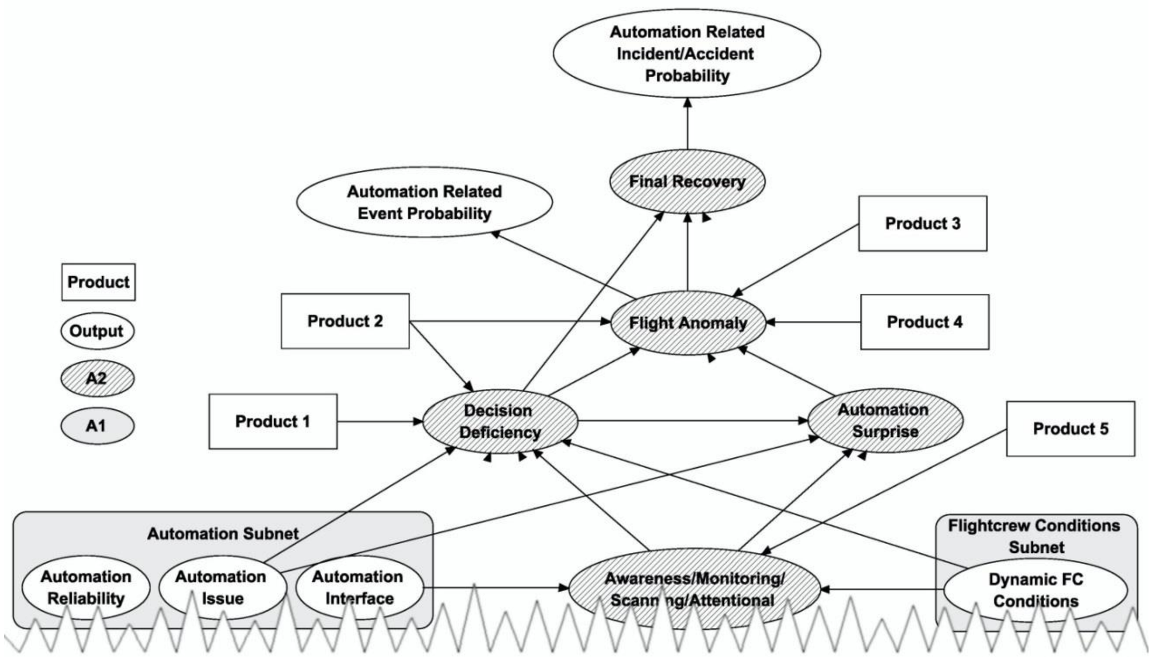

Fig. 5. Nominal product insertion.

13 Data Collection Process

The data collection process for this model was similar to the previous modeling approach, LOCAF, as described by Luxhøj et al. [37]. In summary, the process consists of a series of SME elicitation panels for initial model development and probability data gathering, followed by internal and external reviews and conferences/publications to seek validation. This section highlights the boundary conditions for the FLAP model, assumptions, elicitation process, and finally preliminary results.

13.1 Boundary Conditions

13.1.1 Model Perspective

As previously stated, the model is intended to solely capture the probability of automation-related flight anomalies that could result in incidents and accidents. The two Top-Level output nodes provide accident and incident probabilities of automation-related anomalies with respect to all conceivable accidents and incidents. For that reason, the elicitation process strictly considers probabilities within the accident/incident perspective instead of all aviation operations when querying causal factor probabilities in the model. This assumption renders the probabilities more tangible; e.g., probability of situational awareness deficiency with respect to all yearly U.S.-based flights (over 8 million departures in 2013[14] alone) versus all accidents and incidents within the last 10 years (over 1200 accidents[15] in Part 121 and 135 and 35,000 incidents in ASRS database[16]).

[14] http://www.transtats.bts.gov/
[15] http://www.ntsb.gov/aviationquery/
[16] http://akama.arc.nasa.gov/ASRSDBOnline/QueryWizard_Filter.aspx

13.1.2 Aircraft and Timeframe

The model considers today's aircraft operating within the United States, under FAR Part 121 & FAR 135 with considerable but varying degrees of automation usage, such as Boeing B-737, B-747, B-757, B-767, B-777, B-787 families and Airbus A300, A310, A320, A330, A340, A350 as well as regional jets like Embraer and Bombardier CRJs. The timeframe for this study was a 10-year period including aircraft commissioned in year 2003 through 2013. In order to determine and communicate the intended level of automation among experts, a taxonomy developed by Endsley & Kaber [38] was employed. The taxonomy categorizes automation *level one* as fully manual, *level ten* as fully automatic control and the combinations falling between the two. Although the aircraft cited above have varying automation levels for different tasks, an automation *level-six,* labeled as "Blended Decision Making," was considered as the norm within the model and was communicated throughout the elicitation sessions. In a level-six automation aircraft, "[t]he computer generates a list of decision options that it selects from and carries out if the human consents. The human may approve of the computer's selected option or select one from among those generated by the computer or the operator. The computer will then carry out selected option. This level represents a higher level decision support system that is capable of selecting among alternatives as well as implementing second option [38, p. 465]."

14 Assumptions and Limitations

There were several assumptions made throughout the modeling effort due to limited resources. As in the previous modeling effort and some other BBN approaches, the FLAP model embraces SME opinion for the sole source for data generation primarily due to lack of statistically meaningful data. Although several accidents, incidents and studies were used to develop the model structure, the required probability values in the model were acquired from the SMEs. Additionally, in order to keep the model size manageable and still achieve a generalized automation problem model, nodes like *Awareness/Monitoring* and *Decision Deficiency*[17] contain several assumptions and error types, lowering the model resolution. However, since the model is primarily used to evaluate future NASA technologies' impact as part of a system-level comprehensive portfolio analysis study, the model resolution and fidelity satisfy the analysis purpose. Consequently, the model output is not intended to assist manufacturer flightdeck design or regulatory body (e.g., FAA) policy decision-making processes. Also, due to the variance in application within the industry, the automation system was developed based on major functions instead of actual system components.

15 Elicitation Process and Expert Profiles

The probability data elicitation process consists of a series of SME sessions. The model structure and probability values as well as the impacts of AvSP portfolio elements are acquired from the same set of SMEs, called operational SMEs. For the FLAP model, the operational SMEs consist of two commercial pilots – experienced in Part 121 and 135 operations and airline management/training as well as one human factors expert specialized in flightdeck automation. The additional SME panels held in-between operational SME meetings are used to ensure the model assumptions and structure are sound for the given study purpose. A typical elicitation process for the modeling effort is highlighted in Luxhøj [37].

At the time of writing, the baseline model was established following the first operational SME meeting. The goal of this four-day meeting was to review the preliminary model developed by the team, make revisions, and populate probabilities for the causal factors. The baseline model is typically

[17] The high probability value of these nodes can be attributed to the large number of deficiencies/errors considered in node assumptions.

reviewed by an external panel before proceeding with AvSP product insertions. Depending on the application, the model with inserted products can also undergo an internal review by NASA technologists and other interested parties to ensure the products were placed appropriately. During the second operational SME meeting, experts are expected to review the comments obtained by external and internal panels and also revisit the baseline model results to calibrate and ensure findings are matching experts' mental models as well as limited historical accident/incident data.

16 Conclusions and Next Steps

This paper highlights the development process of a high-level automation-related accident/incident model aimed at serving as a platform for AvSP portfolio assessment. In order to do so, past automation studies and accidents/incidents were reviewed and key issues were identified. Similar to the LOCAF modeling effort, these key issues are then represented in a hierarchical manner and their interdependencies were mapped within the FLAP framework. The network was then modeled using the Hugin Software. In order to populate the model, SME opinions were employed due to lack of a comprehensive historical dataset.

The next steps in the current modeling effort consist of an external review to check the baseline model soundness and validity, followed by the second operational SME session allocated for model calibration, review of concerns/comments provided by the external panel, and insertion of AvSP portfolio products into the model. Following a set of internal review meetings, the third operational SME meeting is planned to revisit the model and provide updated probabilities for inserted AvSP products and their impacts. The analysis of the data stemming from the model provides insight on a) increased automation dependence on today's aircraft and its implications and b) impact of NASA products in mitigating such issues. Finally, the FLAP model will be integrated into the past and future modeling efforts owing to the OOBN modeling techniques, and will be further used in portfolio prioritization efforts.

References

[1] Sarter, N. B. "Investigating Mode Errors on Automated Flight Decks: Illustrating the Problem-Driven, Cumulative, and Interdisciplinary Nature of Human Factors Research," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* Vol. 50, No. 3, 2008, pp. 506-510. doi: 10.1518/001872008x312233

[2] Sarter, N. B., Mumaw, R. J., and Wickens, C. D. "Pilots' Monitoring Strategies and Performance on Automated Flight Decks: An Empirical Study Combining Behavioral and Eye-Tracking Data," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* Vol. 49, No. 3, 2007, pp. 347-357. doi: 10.1518/001872007x196685

[3] Billings, C. E., "Human-Centered Aviation Automation: Principles and Guidelines," NASA TM-110381, 1996.

[4] Sarter, N. B., Woods, D., and Billings, C. E. "Automation Surprises," *Handbook of Human Factors & Ergonomics*, 2nd ed., edited by G. Salvendy, Wiley, New York, NY, 1997, pp. 1926-1941, Chap. 57.

[5] FAA, "Operational Use of Flight Path Management Systems: Final Report of the Performance-based Operations Aviation Rulemaking Commitee/Commercial Aviation Safety Team Flight Deck Automation Working Group," 2013.

[6] FAA, "Safety Alert for Operators (13002)," January 2013.

[7] Hugin Expert, Version 7.8, 2013.

[8] Ancel, E., and Shih, A. T., "The Analysis of the Contribution of Human Factors to the In-flight Loss of Control Accidents," *12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Indianapolis, IN, Sept 17-19, 2012.

[9] Shih, A. T., Ancel, E., and Jones, S. M., "Object-oriented Bayesian Networks (OOBN) for Aviation Accident Modeling and Technology Portfolio Impact Assessment," *American Society for Engineering Management (ASEM) 33rd International Annual Conference*, Virginia Beach, VA, Oct 17–20, 2012.

[10] FAA, "The Interfaces Between Flightcrews and Modern Flight Deck Systems," June 1996.

[11] Orlady, H. W., Orlady, L. M., and Lauber, J. K. "Automation," *Human Factors in Multi-Crew Flight Operations*. edited by H. Orlady, and L. Orlady, University Press, Cambridge, UK, 1999, pp. 223-255, Chap. 11.

[12] CAST, "Mode Awareness and Energy State Management Aspects of Flight Deck Automation – Final Report Safety Enhancement 30, Revision 5," 2008.
[13] Funk, K., Lyall, B., and Riley, V., "Perceived Human Factors Problems of Flightdeck Automation (Phase 1 Final Report)," February 1996.
[14] Rudisill, M., "Line Pilots' Attitudes About and Experience with Flight Deck Automation: Results of an International Survey and Proposed Guidelines," *8th International Symposium on Aviation Psychology*, Columbus, OH, Apr. 24–27, 1995.
[15] Sherman, P. J., Helmreich, R. L., and Merritt, A. C. "National Culture and Flight Deck Automation: Results of a Multination Survey," *The International Journal of Aviation Psychology,* Vol. 7, No. 4, 1997, pp. 311-329. doi: 10.1207/s15327108ijap0704_4
[16] Degani, A., "Modeling Human-Machine Systems: On Modes, Error, and Patterns of Interaction," Ph.D. Dissertation, Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 1996.
[17] Feigh, K., and Pritchett, A., "Modeling the Work of Humans and Automation in Complex Operations," *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, Grapevine, TX, Jan 7–10, 2013.
[18] Foyle, D. C., and Hooey, B. L., eds. *Human Performance Modeling in Aviation*, Boca Ratorn, FL: CRC Press, 2008, pp. 67–77.
[19] Gore, B. F. "Human Performance Cognitive-behavioral Modeling: A Benefit for Occupational Safety," *International Journal of Occupational Safety and Ergonomics (JOSE),* Vol. 8, No. 3, 2002, pp. 339–351.
[20] Gore, B. F. "The Use of Behavior Models for Predicting Complex Operations," *Conference on Behavior Representation in Modeling and Simulation*. Charleston, SC, Mar 21–24, 2010.
[21] Parasuraman, R., Sheridan, T. B., and Wickens, C. D. "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* Vol. 30, No. 3, 2000, pp. 286–297. doi: 10.1109/3468.844354
[22] Riley, V. A., "Human Use of Automation," Ph.D. Dissertation, University of Minnesota, Minneapolis, MN, 1994.
[23] Civil Aviation Authority, "Flight Crew Reliance on Automation," CAA Paper 2004/10, 2004.
[24] Evans, J. K., "An Application of CICTT Accident Categories to Aviation Accidents in 1988-2004," NASA CR-2007-214888, 2007.
[25] Reason, J. T. *Managing the Risks of Organizational Accidents*, Hampshire, UK: Ashgate, 1997.
[26] Wiegmann, D. A., and Shappell, S. A. *A Human Error Approach to Aviation Accident Analysis: The Human Factors Analysis and Classification System*, Surrey, UK: Ashgate, 2003.
[27] Ancel, E., Shih, A. T., Jones, S. M., Reveley, M. S., Luxhøj, J. T., and Evans, J. K. "Predictive Safety Analytics: Inferring Aviation Accident Shaping Factors and Causation," *Journal of Risk Research*, 2014. doi: 10.1080/13669877.2014.896402
[28] Korb, K. B., and Nicholson, A. E. *Bayesian Artificial Intelligence*, London: Chapman Hall/CRC Press LLC 2004.
[29] Koller, D., and Pfeffer, A., "Object-oriented Bayesian Networks," *Proceeding of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-97)*, Providence, RI, August 1–3, pp. 302-313, 1997.
[30] Parasuraman, R., and Riley, V. "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* Vol. 39, No. 2, 1997, pp. 230-253. doi: 10.1518/001872097778543886
[31] Parasuraman, R., and Manzey, D. H. "Complacency and Bias in Human Use of Automation: An Attentional Integration," *Human Factors: The Journal of the Human Factors and Ergonomics Society,* Vol. 52, No. 3, 2010, pp. 381–410. doi: 10.1177/0018720810376055
[32] Mouloua, M., Hancock, P., Jones, L., and Vincenzi, D. "Automation in Aviation Systems: Issues and Considerations," *Handbook of Aviation Human Factors*, 2nd ed., edited by J A. Wise, D. V. Hopkin, and D. J. Garland, CRC Press, Boca Raton, FL, 2010, Chap. 8.
[33] Briere, D., Favre, C., and Traverse, P. "Electrical Flight Controls, From Airbus A320/330/340 to Future Military Transport Aircraft: A Family of Fault-Tolerant Systems," *The Aviation Handbook.* edited by C. R. Spitzer, CRC Press, LLC, Boca Raton, FL, 2001, Chap. 12.
[34] Walter, R. "Flight Management Systems," *The Avionics Handbook.* edited by C. R. Spitzer, CRC Press LLC, Boca Raton, FL, 2001, pp. 264–288, Chap. 15.
[35] Veitengruber, J. E. "Design Criteria for Aircraft Warning, Caution, and Advisory Alerting Systems," *Journal of Aircraft,* Vol. 15, No. 9, 1978, pp. 574–581. doi: 10.2514/3.58409

[36] Jones, S. M., Reveley, M. S., Withrow, C., Evans, J. K., Barr, L. C., and Leone, K., "Systems Analysis of NASA Aviation Safety Program - Final Report, NASA/TM-2013-218059," NASA/TM-2013-218059, June 2010.

[37] Luxhøj, J. T., Shih, A. T., Jones, S. M., Ancel, E., and Reveley, M. S., "Safety Risk Knowledge Elicitation in Support of Aeronautical R&D Portfolio Management: A Case Study," *American Society for Engineering Management (ASEM) 33rd International Annual Conference*, Virginia Beach, VA, Oct 17–20, 2012.

[38] Endsley, M. R., and Kaber, D. B. "Level of Automation Effects on Performance, Situation Awareness and Workload in a Dynamic Control Task," *Ergonomics,* Vol. 42, No. 3, 1999, pp. 462–492. doi: 10.1080/001401399185595

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01- 08 - 2015	Technical Memorandum	

4. TITLE AND SUBTITLE	**5a. CONTRACT NUMBER**
Bayesian Safety Risk Modeling of Human-Flightdeck Automation Interaction	**5b. GRANT NUMBER**
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Ancel, Ersin; Shih Ann T.	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER 999182.02.20.07.01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
NASA Langley Research Center Hampton, VA 23681-2199	L-20598
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
National Aeronautics and Space Administration Washington, DC 20546-0001	NASA
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
	NASA-TM-2015-218791

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified - Unlimited
Subject Category 03
Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The introduction of flightdeck automation systems improved safety, economy, and comfort in air transportation. However, these systems are also associated with events related to pilot handling and automation errors due to over-reliance on automatic systems, and particularly, deficiencies found in man-machine interface including collaboration and joint decision making, among others. This report highlights the development of a Bayesian safety risk model that is focused on increased automation usage in today's aircraft. The goal of this modeling effort is to provide NASA's Aviation Safety Program (AvSP) with a platform capable of assessing the projected impacts of AvSP technologies and products towards reducing the relative risk of automation-related accidents and incidents. In order to do so, a generic framework, capable of mapping both latent and active causal factors leading to automation problems, is developed. Next, the framework is converted into a Bayesian Belief Network that is populated with conditional probabilities provided by a panel of Subject Matter Experts. The resultant model captures causal relationships between stakeholders, automation design, pilot background and their interfaces in an intuitive graphical representation. With the insertion of technologies and products, the model provides estimated individual and collective risk reductions from the AvSP technologies and methodologies.

15. SUBJECT TERMS

Aviation safety; Bayesian belief networks; Flightdeck automation; Portfolio assessment

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			STI Help Desk (email: help@sti.nasa.gov)
					19b. TELEPHONE NUMBER ***(Include area code)***
U	U	U	UU	33	(757) 864-9658