HAZARDS IN ADVISING AUTONOMY: INCORPORATING HAZARD MODELLING WITH SYSTEM DYNAMICS INTO THE AEROSPACE SAFETY ASSESSMENT PROCESS FOR UAS

C. G. Downes*, P. W. H. Chung†

** BAE SYSTEMS Military Air & Information, R&T, Future Capability, Warton Aerodrome, PR4 1AX, UK, clive.downes@baesystems.com, 01772 856379, † Loughborough University, Dept. of Computer Science, LE11 3TU, UK, P.W.H.Chung@lboro.ac.uk*

Keywords: HAZOP, Hazard Modelling, Safety Assessment, UAV, Autonomous Systems.

Abstract

This paper describes the further continuation of an investigation to identify and develop tools for the identification and management of hazards likely to arise with the quality and behavioural aspects in and resulting from automatic advice – such as might arise with an automated system advisory function facilitating critical decision-making with an autonomous vehicle. An example of a representative critical advisory function is identified in that supporting a necessary "Sense & Avoid" capability, as embodied within an airborne autonomous system. In consideration then of how might a model driven approach, combining physical and dynamical models, statistical data and belief be combined to aid system evaluation, work has so far been undertaken to investigate the nature of suitable models to provide representations of the control structure and system dynamics. Whilst the system engineering methods are to be generic, the context of "Sense & Avoid" provides a relevant framework within which to pose a "toy-problem" with complex behaviour, against which to judge the methods and models.

1 Introduction

In the first instance representations of dynamical models have been developed in the context of an outline exemplar describing an air-proximity hazard arising between two air-vehicles; a model of sufficient fidelity so as to capture a representation of the complexity and emergent behaviour that can arise with this interaction. Furthermore, this study has been pursued to develop a broader Systems approach towards hazard assessment so as to evaluate hypothetical deviations from declared intent – wherein a behavioural modelling framework is to be styled upon that of a STAMP [9] (Systems Theoretic Accident Model and Processes) based hazard assessment methodology and drawing upon STPA [15, 18]. It is proposed that one might combine the associated system models, undertake exploratory dynamic hazard assessment, and conduct this within the context of a Preliminary Aircraft Safety Assessment (PASA); as a possible extension to the process guidelines as described in Aerospace Recommended Practice (ARP) 4754A [3]; as outlined in Figure 1. It is also suggested that this improved process model ought to facilitate HAZOP (Hazard and Operability) studies from the system concept stage onwards, along with an inference approach indentifying likely design faults arising with dynamic hazards; effectively supporting the "diagnosis" of design faults during concept development, and then through iteration around this process loop, validate the plausibility of the particular suspected design faults and consistency with the Hazard model.

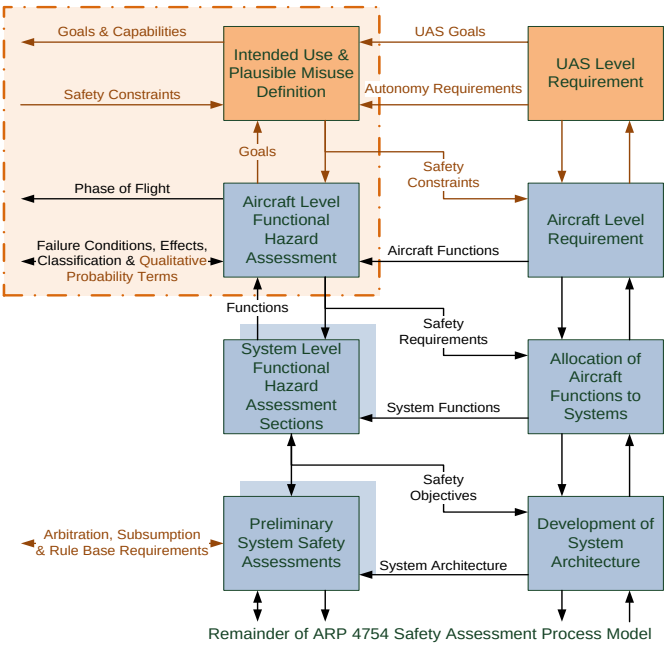

Figure 1: Considering a possible extension to the ARP 4754.

Currently this model incorporates only the constraint of a minimum permitted horizontal separation of 500 ft at closest approach, and in the case of an Unmanned Air Vehicle (UAV) therefore embody a principal safety constraint that the UAV must be capable of manoeuvring (autonomously), as required, to satisfy this minimum safe separation requirement.

Design Decision			The UAV shall incorporate a Collision Avoidance System to prevent separations closer than 500ft horizontally, remaining "well clear" of other aircraft.										
Principal Safety Constraint			The UAV must manoeuvre to maintain a safe horizontal separation from other aircraft where appropriate.										
Assumption			The hazardous incident starts with the UAV in class G airspace operating in marginal Visual Meteorological Conditions and involves a light aircraft routing for the same Visual Rep										
Hazard Context				**Deviations**		**Hazard Risks**					**System Actions / Safety Constraints**	**Defensive Barriers / Design Decisions**	
Phase of Flight	Airspace Class	Event	Inherited Hazard	Guideword	Parameter / Command	Cause	Consequence	Cascade Effect	Classification	Likelihood	UAV Action	Design Action	
ENR	G	Approaching VRP @ Class G / D / C boundary	UAV separation minima is breached		Avoidance Manoeuvre Command								
				NO / OMISSION		The UAV fails to detect the other aircraft altogether.	The UAV does not manoeuvre to avoid the other aircraft.		Safety not assured	Extremely Remote	The UAV Collision Avoidance System shall provide the capability to ensure that other aircraft are avoided by a minimum of 500ft in the horizontal plane (NIAG SG-134 CAS7: Miss Distance / Closest Point of Approach (CPA) - horizontal).	The UAV shall incorporate a Collision Avoidance System to prevent separations closer than 500ft horizontally, remaining "well clear" of other aircraft.	
											The collision avoidance system shall not rely on ATC input or intervention to protect the UAV from collisions with other aircraft (NIAG SG-134 CAS8: Independence from ATC Separation Provision).	TBD	

Figure 2: An Extended HAZOP Table including Phase of Flight and distinguishing between protective System Actions (Safety Constraints) and Defensive Barriers (in the system Design Decisions).

Any higher-level constraint model is to be layered upon the current physical-layer model, and developed into appropriate representational system models in SysML.

1.1 The challenge for Functional Hazard Assessment

It should be noted that within the aerospace field, wherever an operator or other human agency is essential to the safe operation of an air system, then Hazard and Operability (HAZOP) studies are already recommended practice; especially in the case of Air Traffic Control (ATC) operations [2]. However, ARP 4754/4761 [1, 3] make no mention of HAZOP, as these guidelines are concerned with the development of **aircraft** systems; where in fact the configuration and distributed nature of Unmanned Air Systems (UAS) tend to blur the boundaries in the roles and autonomy of the airborne avionics and control systems, the software, data-links and hardware in Ground Control Stations (GCS), and the Designated UAV Operator (DUO), ATC, etc., and indeed potentially also in the expected interaction behaviour of any other airspace users (manned or unmanned).

Consequently, the approach taken with regard to a preliminary system safety assessment, as defined in ARP 4754 [3], refers to the list of acceptable methods to be found in ARP 4761 [1], which in turn describes Fault Tree Analysis (FTA), Dependence Diagrams (DD), and Markov Analysis (MA) as appropriate methods. Both FTA and DD are examples of what Leveson describes as "chain-of-event" models of accident causation [9], and posits that these types of models "... cannot account for indirect and nonlinear relationships". Leveson describes that the systems of interest, in accident causation, are typically Open Systems, and maintained in a state of dynamic equilibrium through closed loop control and feedback. Therefore, in this world-view neither can MA provide sufficient completeness in the prediction of accident causation, as these models are representative of Closed Systems wherein equilibrium is treated as a synonym for Invariant or Stationary; i.e. when the probability of a subsequent system state is identical to that of the preceding state. Furthermore, each of these methods also relies upon knowing or estimating a specific probabilistic value assigned to any individual failure or state. For a while now, it has been suggested by Wilkinson and Kelly [19] that a more pro-active approach should be adopted in the identification of potential hazards than provided for in the process of Functional Hazard Assessment (FHA); as specified by ARP 4754/4761. In particular where complexity and issues of integration with other system apply it is proposed that, rather than the production of hazard checklists and analysis through a "chain-of-event" model, a process akin to HAZOP ought to be applied [19] with consideration of the effects of "deviations" from intended function or behaviour. More recently, in the case of Systems of Systems, Porter, et al [16], suggest that special attention ought to be given to the integration issues manifest in networks of manned and unmanned systems (amongst others); with a need to consider the challenges of semi-automatic behaviour, functional allocation to operator and automation, complex non-linear interactions, and so on.

2 Wherefore HAZOP?

Figure 2 presents a form of HAZOP table as developed in earlier work within the ASTRAEA project, with the purpose of providing a list of hazards as top-events within the development of another phased-mission reliability analysis methodology [17] – therefore supporting an innovative but otherwise traditional "chain-of-events" failure model subsequently. However, the HAZOP process itself was primary to this, with the development of a more open representation of the system and its intended behaviour as a UML model, and as usual conducted as a mental-process with

a group of experts reasoning about the likely effects of deviations in the described intended behaviours. This hazard assessment process was explored through the application of a set of guidewords, and then using the UML model as the framework within which to describe the likely propagation of effects – with a view also to later development of a machine based reasoning support mechanism utilising these models. Hazard and Operability studies are used to consider every deviation of a process system from its design intent and are generally undertaken as a group mental process, involving manual construction and recording, and performed by a team of experts. Overall, it is often a time consuming approach to hazard analysis, due to this manual construction. The key to a successful HAZOP study lies in a greater part in the appropriate selection of suitable Guidewords, and suitable Parameters to which these guidewords might be applied, and so form a specific and narrowly defined set of Deviations. For example in this study the following Guidewords and Parameters have been applied (in part):

Guidewords (used)	Guidewords (unused)
NONE / NO / OMISSION	PART OF
WRONG / COMMISION	OTHER THAN
EARLY	
LATE	
LESS	
MORE	
CONFLICTING	
Parameters	**Commands as parameter**
Data	Manoeuvre
Direction	
Distance	
Speed	
Thrust	

The point of HAZOP studies is that it is not an attempt to calculate or quantify the likelihood of occurrence, but consider qualitatively what will happen in the case of a parameter, measure or quantity being outside of its normal operating range (a deviation). Therefore the study considers every stage of a process and applies relevant guidewords to each stage; evaluating what the effect would be to the system and process should that guideword occur. Hence the method considers a continuous process in steady state, and considers all potential deviations from the intended design.

2.1 HAZOP as used in the aerospace sector

In the case of aerodrome operators and air traffic service providers, HAZOP can be used as a systematic approach for identifying hazards in the operation of these services. In this case the system representation would show the system as a set of Components, Interconnections, Entities (within the components or interconnections), and Attributes to which the Guidewords are to be applied. It is advised that the number of participants should be limited to between five and seven people. Such a team of experts would include a Study Leader, a Recorder, an Operation of System Expert, Users of the System, and other experts [2].

2.2 HAZOP as (semi)automated propagation method

Computer based qualitative reasoning models supporting the HAZOP process have been developed, by in effect providing "HAZOP emulation" to the HAZOP team; for example as with STOPHAZ and HAZID codes [10-14]. These have been developed in the main for the chemical, process and power generation plant industries, and utilise libraries of components embodying qualitative propagation models with associated effects attached, supporting the construction of complex plant models derived from their Pipe & Instrumentation Diagrams (P&ID). Using these software based tools and libraries the HAZOP team can discuss, reason and apply their expertise to the believed behaviour of individual components, some of which might be quite complex of themselves, with the software then capturing all of the permitted propagations of cause and effect between components; both down-stream and up-stream. With the propagation methods of qualitative deviation and effect already addressed to a greater extent, at least for static behavioural models, further consideration might be given to the encapsulation and propagation behaviour of essentially dynamic systems. Certainly the process plant models are treated as closed system, operating at particular set-points. However, autonomous and mobile robotic systems cannot be safely treated as closed systems, due to their homeostatic goal seeking behaviour within a state of dynamic equilibrium.

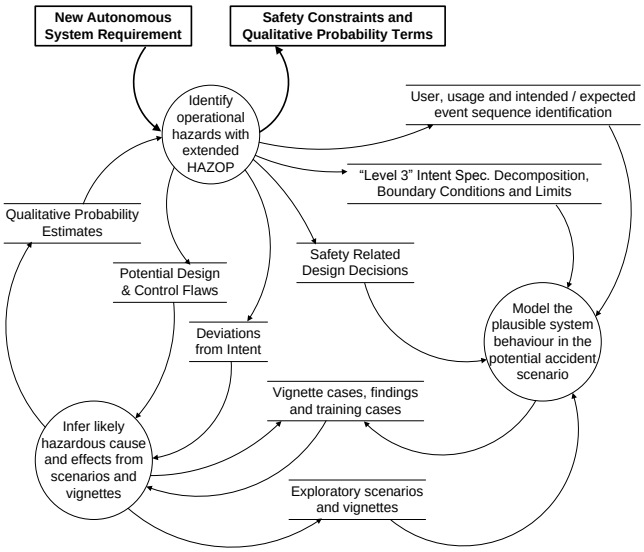

Figure 3: Hazard Modelling with System Dynamics Model.

To this end it is proposed that a hazard modelling and inference process be developed, as described in Figure 3. In this process the actual HAZOP is to be styled upon the approach as adopted for STPA [15, 18], whilst an additional method is to be developed to cater for the iterative nature of hazard assessment with dynamical modelling supported by an inference mechanism connecting and validating the apparent dynamical behaviour with respect to the assessed causes.

	"coverage"	Component Failure Mode	Functional Failure Mode	Mission Phase	Classify Severity of Effect	Propagation	Deviations	Capture Expertise	Probabilistic Occurrence Rates	Safety Constraint Violation	Capture Complex Dynamics	Energy Release or Flow	Human Reliability or Error	System Operating State	Capture Socio-Technical Behaviour	Adaptive Migratory Behaviour	Interactive Exploration
HAZOP & STAMP	64 – 18 (46)	3	3	-	3	3	9	9	-	9	9	3	1	3	3	3	3
STAMP & FMEA	58 – 30 (28)	9	-	1	3	3	-	-	9	9	9	3	3	-	3	3	3
HAZOP & FMEA	55 – 27 (28)	9	3	1	3	3	9	9	9	-	-	3	3	3	-	-	-
STAMP & FFA	52 – 36 (16)	1	9	9	3	3	-	-	-	9	9	-	-	-	3	3	3
HAZOP & FFA	52 – 36 (16)	3	9	9	3	3	9	9	-	-	-	3	1	3	-	-	-
FMEA & FFA	48 – 48 (0)	9	9	9	3	3	-	-	9	-	-	3	3	-	-	-	-

The Employment of FMEA, FTA and FFA together constitutes FHA (Functional Hazard Analysis) as described in ARP 4761

Figure 4: Hazard Analysis – options for combining methods, a pseudo QFD ranking.

3 Incorporating STAMP – STPA

Considering the various features or "coverage" of the different hazard assessment methods, from "classification of severity effect" to "socio-technical behaviour" (Figure 4), a case might be made that taken together a HAZOP methodology and a STAMP framework [9] forms a combined methodology that complements the coverage of the approaches adopted by FHA; albeit somewhat subjectively.

In adopting the STAMP framework, and in particular the "Systems Theoretic Process Analysis" (STPA) [15, 18], an approach can be developed that considers the likely nature of an applicable range of "Inadequate Control Actions" (ICA), effectively deviations, in the context of a simplified model of the Low-level Process Control Loop [18].

3.1 STPA and the low-level process control loop

Applying this approach then in the case of the work described here, certain simple extensions can prove useful. For example, two further fundamental modifications to the generic process control loop have been added, as shown in Figure 5; both related to behavioural control and interactions relevant to mobile robotics and autonomous systems. First, these types of systems generally do not produce outputs typical of a production or set-point control system. Often the useful output arises in the context of interactions with other independently controlled entities. Second, reactive robotic systems usually embody a layering of behavioural control strategies, each with different goals and priorities, and that these constitute the overall feedback control model for the robotic entity's interaction with its environment [4]. To this end, Figure 5 includes additional places representing such layered behavioural models, the further inclusion of a "state information" model to represent the 'perceptions' that the behavioural models will have of the state of the environment, and a representation of the coupling of the process output space with any other related entities to the right.

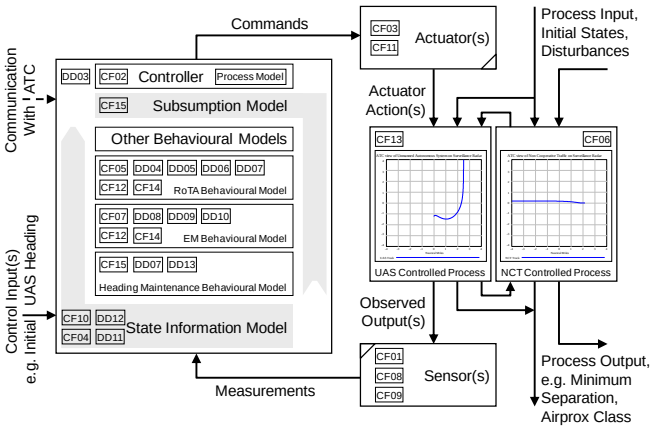

Figure 5: UAS Collision Avoidance Low-level Process Control Loop, styled on the generic STPA process [18].

Otherwise, the manner by which this representation is used to systematically identify hazards remains the same as for the described STPA process [18]. Start by identifying the controlled parameter ("manoeuvre" – in the case of the Collision Avoidance System) along with the applicable set of guidewords; for each deviation identify all likely design or "control flaws" (CFxx) associated with each of the relevant parts of the model by applying each deviation (ICAnn) whilst working around the loop (including any connecting components that might be assigned to other related entities); for each flaw then identify any additional safety constraint

(SCyy) as mitigation; and finally detail the associated design decisions (DDzz) required to implement or enforce these additional safety constraints. Repeat the process around the control-loop until safety assurance is satisfied. Additionally, annotate the HAZOP and models with these as shown below.

4 Closing the Loop with Inference

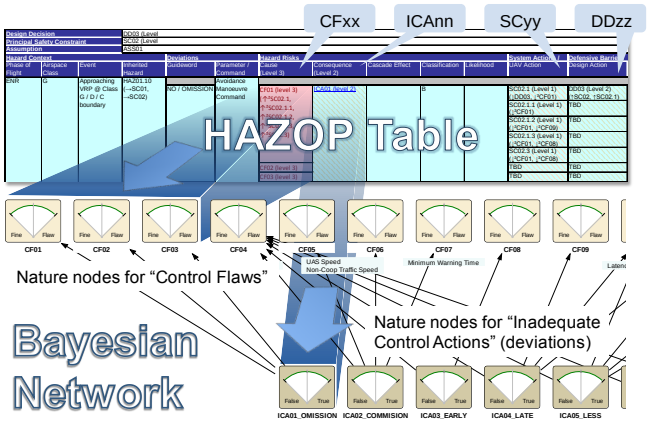

Figure 6: HAZOP for potential Design and Control Flaw Diagnosis, and Cause Inference using Bayesian Network.

The HAZOP process facilitates reasoning about the possibility and likelihood of hazardous consequences occurring due to deviations from design intent. However, as it is generally conducted as a mental-process, or at best supported with software models for quasi-static cases, more is needed to support reasoning and validation with regard to complex dynamical outcomes. For example an automated assessment of risk in the navigation of shipping using Bayesian Learning [8] has been demonstrated with networks wholly derived by learning from accident data. However, with no a priori data available, and the problem then to reconcile any such learnt data structures with information derived from a HAZOP process, suggests that at least a part of such an inference model be derived from the structure of the HAZOP table itself; as in Figure 6. Subsequently time-series data from the system dynamics model may then be processed, indicating degrees of belief in the causes, facilitating closed-loop model-based assessment combining risk analysis and risk evaluation, as illustrated in Figure 7.

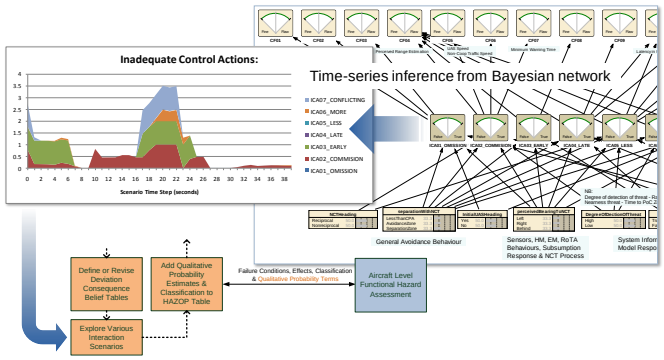

Figure 7: Closing the Loop with Time-series Inference – Validation and Verification.

5 Models

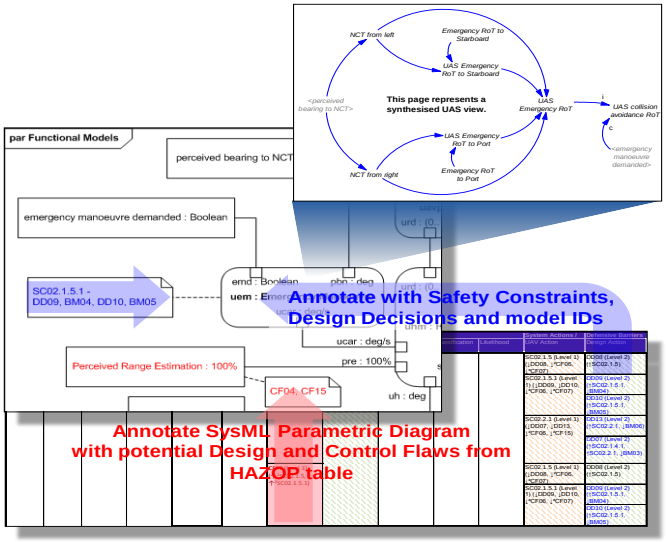

Figure 8: Functional Models – System and Components.

Finally, Guiochet, et al, [7] argue that HAZOP might also be facilitated with UML Use Cases and Sequence Diagrams as a more formal system requirements specification embodying some behavioural characteristics with which to reason about deviations. Then to provide an interface model with similar formalism this unified modelling approach might be extended to include Parametric diagrams and Constraint Blocks from SysML (Figure 8) for the causal loop models as proposed in STAMP. Each of these constructs might also represent the various types of element comprising an "Intent Specification" [15]; including the Environment, Supervisory / Operator Task, Functional, Validation & Verification (analysis) models, whilst introducing the concept of a representation for any anticipated specific system Dysfunction models also.

6 Conclusions & Future Work

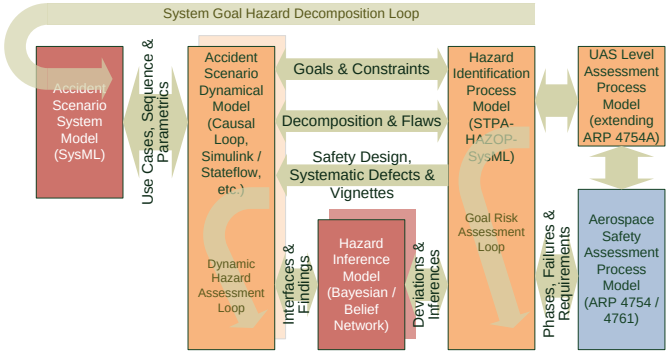

Figure 9: Incorporating HAZOP alongside ARP 4754A..

A proposal for a closed-loop safety assessment process, incorporating HAZOP styled upon STPA, system dynamics modelling, and Bayesian inference, is proposed as an approach to extend the PASA for a UAS/UAV; as shown in Figure 9. This arises from an earlier provisional requirements analysis and a discussion of bottom-up model building [5, 6]. To date, a dynamical model incorporating representative

behaviours has been created, with a selection of hazardous vignettes also identified and explored. The results from these behavioural interactions have been collated with the corresponding Bayesian inference results in the form of a questionnaire to be presented to appropriate system specialists in the next stage of this work. This future validation by questionnaire is intended to reveal whether the same outcomes, as produced by the models, might be intuited directly by those with sufficient expertise, both in terms of the behavioural dynamics and the inferences made from these.

Acknowledgements

The authors wish to thank BAE SYSTEMS and EPSRC for supporting this project, operating through the Systems Engineering Doctorate Centre at Loughborough University.

References

[1] *ARP 4761 - Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment*. Warrendale, Pennsylvania, USA: Society of Automotive Engineers, SAE International, (1996).

[2] *CAP 760 - Guidance on the Conduct of Hazard Identification, Risk Assessment and the Production of Safety Cases*. London, UK: TSO (The Stationery Office) on behalf of the UK Civil Aviation Authority, Air Traffic Standards Department, Safety Regulation Group, (2006).

[3] *ARP 4754A - Guidelines for Development of Civil Aircraft and Systems*. Warrendale, Pennsylvania, USA: Society of Automotive Engineers, SAE Aerospace, SAE International Group, (2010).

[4] R. A. Brooks, *Cambrian Intelligence: The Early History of the New AI*: Bradford Books, The MIT Press, (1999).

[5] C. G. Downes, P. W. H. Chung, and A. Morris, "Hazards in advising autonomy: A structured approach seeking novelty in developing the requirements for an exemplar," in *System of Systems Engineering (SoSE), 2010 5th International Conference on*, (2010), pp. 1-7.

[6] C. G. Downes and P. W. H. Chung, "Hazards in advising autonomy: Developing requirements for a hazard modelling methodology incorporating system dynamics," in *Dependable Control of Discrete Systems (DCDS), 2011 3rd International Workshop on*, (2011), pp. 115-120.

[7] J. Guiochet, D. Martin-Guillerez, and D. Powell, "Experience with Model-Based User-Centered Risk Assessment for Service Robots," in *High-Assurance Systems Engineering (HASE), 2010 IEEE 12th International Symposium on*, (2010), pp. 104-113.

[8] S. Hu, C. Cai, and Q. Fang, "Risk assessment of ship navigation using Bayesian learning," in *Industrial Engineering and Engineering Management, 2007 IEEE International Conference on*, (2007), pp. 1878-1882.

[9] N. G. Leveson, "A systems-theoretic approach to safety in software-intensive systems," *Dependable and Secure Computing, IEEE Transactions on*, vol. 1, pp. 66-86, (2004).

[10] S. A. McCoy, S. J. Wakeman, F. D. Larkin, P. W. H. Chung, A. G. Rushton, and F. P. Lees, "HAZID, A Computer Aid for Hazard Identification: 2. Unit Model System," *Process Safety and Environmental Protection*, vol. 77, pp. 328-334, (1999).

[11] S. A. McCoy, S. J. Wakeman, F. D. Larkin, P. W. H. Chung, A. G. Rushton, F. P. Lees, and P. M. Heino, "HAZID, A Computer Aid for Hazard Identification: 3. The Fluid Model and Consequence Evaluation Systems," *Process Safety and Environmental Protection*, vol. 77, pp. 335-353, (1999).

[12] S. A. McCoy, S. J. Wakeman, F. D. Larkin, M. L. Jefferson, P. W. H. Chung, A. G. Rushton, F. P. Lees, and P. M. Heino, "HAZID, A Computer Aid for Hazard Identification: 1. The Stophaz Package and the Hazid Code: An Overview, the Issues and the Structure," *Process Safety and Environmental Protection*, vol. 77, pp. 317-327, (1999).

[13] S. A. McCoy, S. J. Wakeman, F. D. Larkin, P. W. H. Chung, A. G. Rushton, and F. P. Lees, "Hazid, A Computer Aid for Hazard Identification: 4. Learning Set, Main Study System, Output Quality and Validation Trials," *Process Safety and Environmental Protection*, vol. 78, pp. 91-119, (2000).

[14] S. A. McCoy, S. J. Wakeman, F. D. Larkin, P. W. H. Chung, A. G. Rushton, and F. P. Lees, "Hazid, a Computer Aid for Hazard Identification: 5. Future Development Topics and Conclusions," *Process Safety and Environmental Protection*, vol. 78, pp. 120-142, (2000).

[15] B. D. Owens, M. S. Herring, N. Dulac, N. G. Leveson, M. D. Ingham, and K. A. Weiss, "Application of a Safety-Driven Design Methodology to an Outer Planet Exploration Mission," in *Aerospace Conference, 2008 IEEE*, (2008), pp. 1-24.

[16] J. Porter, M. Squair, and A. Singh, "Risk & Safety Aspects of Systems of Systems," *44th AIAA Aerospace Sciences Meeting and Exhibit; Reno, NV; USA; 9-12 Jan*, pp. 1-15, (2006).

[17] R. Remenyte-Prescott, J. D. Andrews, and P. W. H. Chung, "An efficient phased mission reliability analysis for autonomous vehicles," *RELIABILITY ENGINEERING & SYSTEM SAFETY*, vol. 95, pp. 226-235, (2009).

[18] M. V. Stringfellow, N. G. Leveson, and B. D. Owens, "Safety-Driven Design for Software-Intensive Aerospace and Automotive Systems," *Proceedings of the IEEE*, vol. 98, pp. 515-525, (2010).

[19] P. J. Wilkinson and T. P. Kelly, "Functional hazard analysis for highly integrated aerospace systems," in *Certification of Ground/Air Systems Seminar (Ref. No. 1998/255), IEE*, (1998), pp. 4/1-4/6.