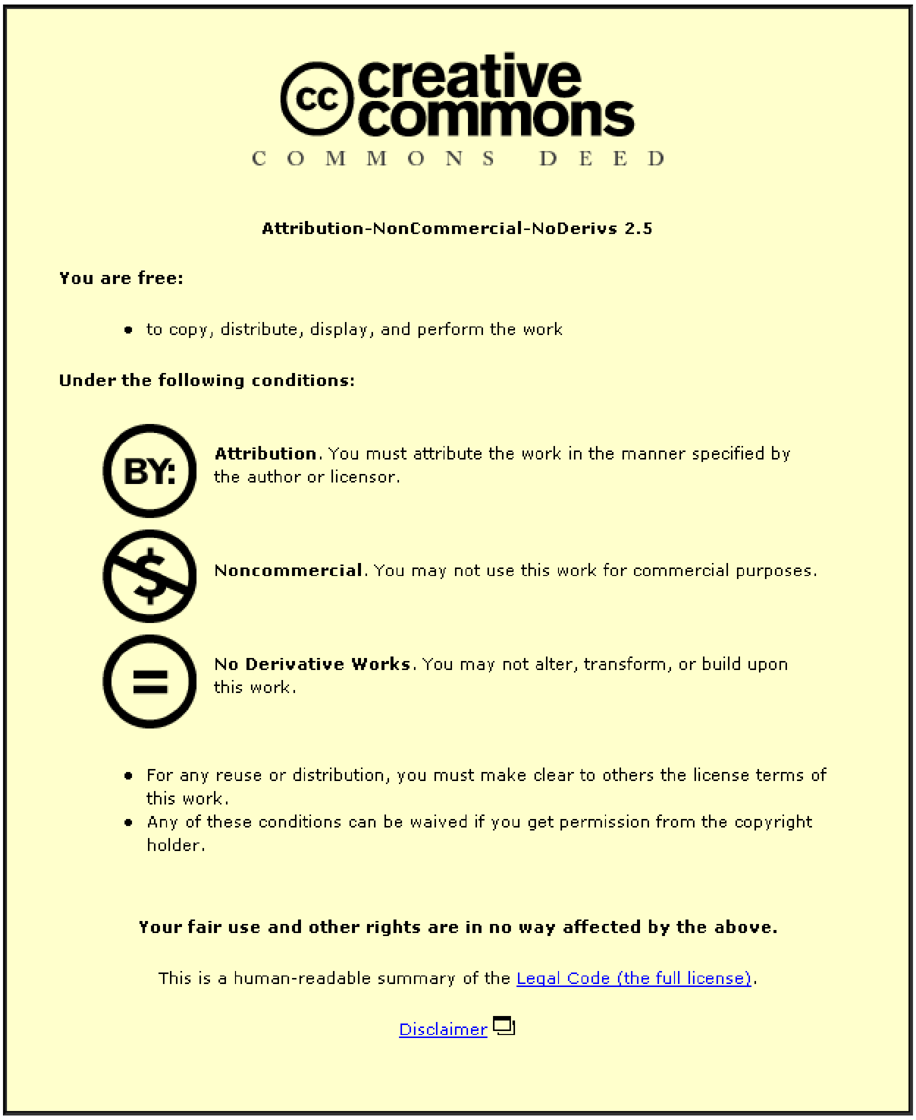


AN INVESTIGATION INTO HAZARD-CENTRIC ANALYSIS OF COMPLEX AUTONOMOUS SYSTEMS

____

by

Clive George Downes

A Doctoral Thesis

Submitted in partial fulfilment of the requirements

for the award of

Doctor of Engineering of Loughborough University

12th of June 2013

AN INVESTIGATION INTO HAZARD-CENTRIC ANALYSIS OF COMPLEX AUTONOMOUS SYSTEMS

by

Clive George Downes

Submitted in partial fulfilment of the requirements for the award of Doctor of Engineering of Loughborough University

Abstract

This thesis proposes a hypothesis that a conventional, and essentially manual, HAZOP process can be improved with information obtained with model-based dynamic simulation, using a Monte Carlo approach, to update a Bayesian Belief model representing the expected relations between cause and effects – and thereby produce an enhanced HAZOP. The work considers how the expertise of a hazard and operability study team might be augmented with access to behavioural models, simulations and belief inference models. This incorporates models of dynamically complex system behaviour, considering where these might contribute to the expertise of a hazard and operability study team, and how these might bolster trust in the portrayal of system behaviour. With a questionnaire containing behavioural outputs from a representative systems model, responses were collected from a group with relevant domain expertise. From this it is argued that the quality of analysis is dependent upon the experience and expertise of the participants but this might be artificially augmented using probabilistic data derived from a system dynamics model. Consequently, Monte Carlo simulations of an improved exemplar system dynamics model are used to condition a behavioural inference model and also to generate measures of emergence associated with the deviation parameter used in the study. A Bayesian approach towards probability is adopted where particular events and combinations of circumstances are effectively unique or hypothetical, and perhaps irreproducible in practice. Therefore, it is shown that a Bayesian model, representing beliefs expressed in a hazard and operability study, conditioned by the likely occurrence of flaw events causing specific deviant behaviour from evidence observed in the system dynamical behaviour, may combine intuitive estimates based upon experience and expertise, with quantitative statistical information representing plausible evidence of safety constraint violation. A further behavioural measure identifies potential emergent behaviour by way of a Lyapunov Exponent. Together these improvements enhance the awareness of potential hazard cases.

Keywords: Hazards, Certification, Autonomous Systems, Systems Engineering, Behavioural Modelling, Bayesian Belief.

Acknowledgements

Four and half years is a noticeable amount of time to take a career detour such as this, especially when viewing the last time you did something like this through the wrong end of the telescope. Then there was just my wife Eileen and myself, and both of my parents to celebrate that occasion with us. Time moved on, as did career, employer and opportunity. When, nearly five years ago I was asked by my boss if I would be interested in "... doing an Engineering Doctorate" as a part of my day job I hesitated for a while – well about fifteen seconds actually. And those past five years now seem to be simultaneously both so short and yet so long!

So I am most grateful to my employer, BAE Systems, in providing me with this opportunity, especially in having the support of Maureen Mccue and John Wright from the start. Other colleagues also deserve specific mention in providing guidance, moral support, or simply showing an interest over the intervening years – thanks go to Dr Fiona Cayzer for suggesting the adoption of a "non-parametric" statistical approach for dealing with the vagaries of humans when responding to questionnaires; Dr Ben Gorry in helping to keep me on my toes with regard to "system safety", as well as helping with some proof reading and in providing general scepticism; and Dr John Pearson for playing the host whenever I visited the System Engineering Innovation Centre at Loughborough, and generally in providing a friendly ear - plus tea, coffee, (occasionally beer), etc. My two industrial supervisors have provided continuing support to the project at different times, appreciation going to both Anne Morris and more recently Stewart Willocks in this regard. There are also many others both at Warton (R&T) and at Loughborough (SEIC) who have been peripherally involved with the work and are also thanked for their contribution and support.

Additional thanks also needs to be extended a little more widely, firstly to the anonymous respondents to the evaluation questionnaire, including colleagues in BAE SYSTEMS, both at Warton and Loughborough, but also at the universities of Aberystwyth, Leicester and Loughborough. Secondly this request for respondents extended further to some of my fellow pilots who operate out of Woodvale; thanks guys even if you thought I was trying to give you an exam! (my apologies also for missing out on some of my committee duties and various hanger sweeps over the past few years)

Clearly my academic supervisor, Professor Paul Chung has played a significant part in supporting and guiding this work – having had plenty of things to read over the years and responding with constructive criticism and plenty of encouragement. In this Paul has provided me with a coherent direction, some of his time, support, and also opportunities for me to present this work overseas!

Finally, Eileen has tolerated my many nights away in Loughborough and at conferences over the past few years, and I very much appreciate her love and support in this. Now the torch so far as education and exams are concerned will be passed to our son, Nicholas, as he embarks on his great adventure.

Table of Contents

Abstract **2**

Acknowledgements **3**

Table of Contents **4**

List of Figures **8**

List of Tables **12**

Glossary **13**

PART 1. INTRODUCTION & LITERATURE REVIEW **15**

Chapter 1. Introduction **16**

1.1. Background, Aim and Objectives 16
1.2. Hypothesis 18
1.3. Motivation in Context 19
1.4. Project Overview 22
1.5. Contribution of Thesis 24
1.6. Thesis Structure 27

Chapter 2. Regulatory Context Review **30**

2.1. Accident Inevitability 30
2.1.1. Accidents Past 30
2.1.2. Human Factors 35
2.1.3. Situational Awareness 36
2.2. Defining Aspects & Constraints 37
2.2.1. Degree of Autonomy 37
2.2.2. Human Equivalence 39
2.2.3. Safety & Reliability 40
2.2.4. Analytic Limitations 41
2.2.5. Regulatory Guidance 42
2.2.6. Hazard Assessment 44
2.2.7. Function or Behaviour Abstraction? 46
2.3. Summary Considerations 48

Chapter 3. Hazard Identification and Analysis Review **49**

3.1. Extending HAZOP 49
3.1.1. Hazard emulation, Identification & Assessment 49

3.1.2. Bayesian Expressions of Risk 50
3.1.3. Scenarios & Deviation Analysis 51
3.2. Systems Theory beyond Normal Accidents 53
3.2.1. Systems Theoretic Accident Model Process 54
3.2.2. Systems Theoretic Process Analysis 56
3.3. Systems Modelling 58
3.3.1. Systems Representation 58
3.3.2. Representing Risks & Hazards 59
3.3.3. Hazard Detection & Complex Behaviours 60
3.3.4. Bayesian & Monte Carlo 61
3.4. Summary Considerations 63

Chapter 4. Systems Engineering Review 64
4.1. Development Life-Cycles 64
4.1.1. Beyond ARP 4754A / 4761 64
4.1.2. Spiral Development 66
4.2. Architecture & Validation 67
4.2.1. Design Intent 67
4.2.2. Validation through Questionnaire 69
4.3. Emergent and Complex Behaviour 69
4.4. Summary Considerations, Influences and Scope 70

PART 2. SYSTEM FRAMEWORK 74

Chapter 5. System Modelling Requirements 75
5.1. Textual Analysis & QFD 75
5.2. Stakeholders 78
5.2.1. Certification Evidence Clients 78
5.2.2. Safety Objective Clients 78
5.2.3. HAZOP Team 78
5.2.4. Engineering 78
5.3. Scenario Requirements 79
5.3.1. Sense And Avoid (SAA) Hazards 79
5.3.2. Scenario Generation 79

Chapter 6. System Model Design 81
6.1. Modelling Framework 81
6.2. Scenario Modelling 84
6.3. Behavioural Intent Modelling 85
6.4. Vignette Modelling 88
6.4.1. Scenario Goal Behaviour – Design Decisions 88

6.4.2. Default Avoidance Behaviour (Intended Behaviour Models) 90
6.4.3. Emergency Avoidance Behaviour (Intended Behaviour Models) 91
6.4.4. Default Heading Maintenance Behaviour (Intended Behaviour Models) 93

Chapter 7. System Constraints Model 96
7.1. System State Modelling 96
7.1.1. Initial State – Defined Environment (Initial Conditions Model) 96
7.1.2. Relative Motion – Defined Environment (Motion Dynamics Model) 100
7.1.3. Integrated Manoeuvring Behaviour (Intended Behaviour Models) 101
7.1.4. Intrusion, Threat & Perception Model (Threat & Flaw Insertion Models) 104
7.1.5. Delay & Offset Direction Model (Flaw Insertion Models) 107
7.1. Safety Constraint Modelling 108
7.1.1. Air Proximity Model (Constraint Violation Detection Models) 108
7.1.2. Emergent System Behaviour Sensitivity Modelling 112
7.2. Summary of Purpose and Recapitulation 115

Chapter 8. Belief and Model Validation 116
8.1. Inference Modelling 116
8.1.1. HAZOP Inference Modelling 118
8.1.2. Defect, Flaw & Deviation Behaviour Capture – Monte Carlo Generated Cases 121
8.1.3. System State and Violation Capture – Defining Monte Carlo Generated Cases 130
8.2. Model Test Case Discussion 134
8.2.1. Validation Test Cases and Hazardous Outcomes 134
8.2.2. Hazard Context 136
8.2.3. System Flaws and Deviations 136
8.2.4. Diagnosed Defects and Data Post Processing 137
8.2.5. Inferred Airprox Events 138
8.2.6. Inferred Flaws and Deviations Post Processing 140
8.2.7. Avoidance and Emergent Behaviour 141
8.2.8. Black Box v White Box Internal Consistency 146
8.2.9. Worst and Improbable Case Example 147
8.3. Final Outcomes – Informing HAZOP 150
8.3.1. Diagnostic Example – Threat-Offset Direction 150
8.3.2. Enhanced HAZOP 151
8.4. Recapitulation of Requirements 153

PART 3. QUESTIONNAIRE, RESULTS & CONCLUSIONS 154

Chapter 9. Questionnaire 155
9.1. Questionnaire design 155
9.2. Questionnaire results 157

9.2.1. General Situation Awareness 157

9.2.2. Anticipation of Interactions 160

9.2.3. Inferences 162

9.2.4. Prioritising Risks 166

9.3. Intermediate Outcomes – Questionnaire 167

Chapter 10. Summary of Results, Outcomes and Future Work 168

10.1. Summary of work 168

10.2. Principal results 169

10.2.1. Questionnaire Responses, Hypothesis and Preferences 169

10.2.2. Avoidance and Emergent Properties 172

10.3. Lessons learnt, alternative approaches and improvements 176

10.3.1. Questionnaire 176

10.3.2. Monte Carlo Modelling 178

10.4. Industry use, benefits, and exploitation 179

10.5. Future work – Frameworks for Behavioural Hazard Assessment 180

Chapter 11. Recommendations and Conclusions 182

11.1. Recommendations 182

11.2. Conclusions 183

ANNEX 185

References 186

Appendix A. Enhanced HAZOP Employing Inference Test Cases 195

Appendix B. Hazards Advising Autonomy: Questionnaire 238

Appendix C. Respondent Results and Analysis 250

List of Figures

Figure 1.1 Certification Considerations for Complex Aircraft Systems [95, 97, 100] ... 19
Figure 1.2 System Development Process with HAZOP, adapted from ARP4754A [101] ... 20
Figure 1.3 Outline HAZOP inference process role [32] ... 21
Figure 1.4 Autonomous Systems HAZOP Context ... 22
Figure 1.5 An Enhanced HAZOP Framework ... 25
Figure 1.6 A Scheme for Incorporating Behavioural Hazard Assessment into ARP 4754A ... 26
Figure 2.1 Living with UAS, adapted from Perrow "Living with High-Risk Systems" [87] ... 34
Figure 2.2 Hazard Model adapted from Reason's "Swiss Cheese" Model [91, 92] ... 35
Figure 2.3 Accident Model & Defensive Barriers after Tribble [108] ... 41
Figure 2.4 Propagation of Effect Through Sub-system Layers after Wilkinson [114] ... 46
Figure 3.1 A Hybrid PRA Framework for Socio-technical Risks [79] ... 51
Figure 3.2 Generic STPA Low-Level Process Control Loop after Stringfellow [104, 105] ... 57
Figure 4.1 Systems Engineering "Vee" Diagram, adapted from ARP4754A [101] ... 65
Figure 4.2 A Simplified Interpretation of Boehm's Model of Spiral Development [8] ... 66
Figure 4.3 Behavioural Subsumption Model after Brooks [12] ... 67
Figure 4.4 SE Integration Equivalent to ARP 4754A for Autonomous & Robotic Systems ... 72
Figure 4.5 A framework compatible with ARP 4754A and DO-178C [97, 101] ... 72
Figure 5.1 Use Case Simplifying the Certification Purpose ... 75
Figure 5.2 Deviation Analysis Requirements as QFD Analysis from Textual Analysis ... 76
Figure 5.3 Requirement Sets Identifying selected Means or "How's" ... 77
Figure 5.4 Use Case Scoping the Hazard Modelling Research Programme ... 78
Figure 5.5 Mid-air Collision Avoidance Scenario ... 80
Figure 6.1 Integrating System Dynamics and Inferred Constraint Violation Perspectives ... 81
Figure 6.2 Generic System Goal Model within a STHAZOP Framework ... 82
Figure 6.3 A framework compatible with ARP 4754A and DO-178C [97, 101] ... 83
Figure 6.4 STHAZOP Framework Sequence Diagram ... 83
Figure 6.5 Scenario Collision Avoidance Process Control Loop, adapted from Stringfellow [104, 105] ... 86
Figure 6.6 Rule Modelling – The Rules of the Air Regulations 2007 [15] ... 88
Figure 6.7 Sense and Avoid Scenario System Goal ... 89
Figure 6.8 Scenario Rules of the Air Avoidance Decision Table ... 90
Figure 6.9 Rules of the Air Avoidance Interpretation as Rate of Turn surface ... 90
Figure 6.10 Scenario Perceived Bearing Fuzzy Sets ... 91
Figure 6.11 Alternative Emergency Manoeuvre and Collision Threat Zone Behaviour ... 92
Figure 6.12 Scenario Perceived Direction of Threat ... 92
Figure 6.13 Scenario Emergency Manoeuvre Internal Block Diagram (dysfunctions in red) ... 92
Figure 6.14 UAS Heading Maintenance Fuzzy Logic as "Causal Flow" Diagram ... 93
Figure 6.15 Scenario Heading Maintenance Decision Table (Column Rules as in Figure 6.8) ... 93
Figure 6.16 Scenario Bearing to Desired Heading Fuzzy Sets ... 94
Figure 6.17 Scenario System Goal Internal Block Diagram, as implemented ... 95

Figure 7.1 UAS and NCT Initial States Specific (Constraints) Parametric Model 96
Figure 7.2 Partial Parametric Definitions for Initial State Constraints 97
Figure 7.3 Generic Physical Constraints Parametric Model 98
Figure 7.4 Generic Dysfunction Constraints Parametric Model 99
Figure 7.5 Generic Proximity Constraints Parametric Model 99
Figure 7.6 UAS and NCT Relative Motions as “Causal Loop – Stock and Flow” Diagram 100
Figure 7.7 Typical Vignette Behaviour 101
Figure 7.8 Typical vignette separation sensitivity distribution 103
Figure 7.9 Typical vignette inference of deviation 103
Figure 7.10 SAA Scenario Threat & Detection Defect Constraints 104
Figure 7.11 Cause – effect chain from Perceived Range Estimation to separation with NCT 106
Figure 7.12 Misestimating Separation – an example 106
Figure 7.13 SAA Scenario Emergency Behaviour Defect Constraints 107
Figure 7.14 SAA Scenario UAS Airprox Constraints 109
Figure 7.15 SAA Scenario UAS Threatened Constraints 111
Figure 7.16 Generic Lyapunov State Monitoring Parametric Model 112
Figure 7.17 SAA Scenario Lyapunov State Monitoring Constraints 114
Figure 7.18 Example Lyapunov “Lambda Tee” plots with similar emergent behaviour (TC1 and TC2) ... 115
Figure 8.1 Simple Example Flaw – Deviation Bayesian Belief Network 116
Figure 8.2 Original HAZOP for Control Flaw Diagnosis and Cause Inference [32] 119
Figure 8.3 Later HAZOP Causal Relationships as a Bayesian Network 119
Figure 8.4 Generic Bayesian Network Parametric Model 120
Figure 8.5 Generic Defect Bayesian Capture Parametric Model 121
Figure 8.6 Generic Flaw Bayesian Capture Parametric Model 122
Figure 8.7 Generic Deviation Bayesian Capture Parametric Model 123
Figure 8.8 Generic Behaviour State Bayesian Capture Parametric Model 123
Figure 8.9 Generic Proximity Bayesian Capture Parametric Model 124
Figure 8.10 SAA Scenario Defect Capture Bayesian Data Constraints 124
Figure 8.11 SAA Scenario Flaw Capture Bayesian Data Constraints 126
Figure 8.12 SAA Scenario Deviation Capture Bayesian Data Constraints 128
Figure 8.13 Airprox constraint violation & system state association in the Bayesian Network 130
Figure 8.14 SAA Scenario Behaviour State Capture Bayesian Data Constraints 131
Figure 8.15 SAA Scenario Proximity Capture Bayesian Data Constraints 132
Figure 8.16 Hazard Modelling – separation as constraint violation [31] 134
Figure 8.17 Hazard Test Case Assignment by Direction of Approach 135
Figure 8.18 Vignette defective design precursor belief association in the Bayesian Network 137
Figure 8.19 Lyapunov Exponent Distributions from 0 to 25 seconds for Test Case 6 (TC6) 142
Figure 8.20 The Corresponding Lyapunov Exponent Distribution Histogram at 25 seconds 142
Figure 8.21 Hazard Test Case 6 (TC6) 143
Figure 8.22 Hazard Test Case 1 (TC1) 143

Figure 8.23 Hazard Test Case 2 (TC2) 144
Figure 8.24 Lyapunov Scatter Plots with respect to Initial Relative Heading 145
Figure 8.25 Airprox Scatter Plots with respect to Initial Relative Heading 145
Figure 8.26 Validate Goal System Behaviour Belief Sequence Diagram 146
Figure 8.27 Black-box versus White-box inference 146
Figure 8.28 Hazard Test Cases 9 and 10 (TC9, TC10) 148
Figure 8.29 Inferred States at the Unique Airprox Event 149
Figure 9.1 Behaviours – interpretation of interactions & intuition [32, 33] 157
Figure 9.2 General situation awareness - general hazardous region at 0.5 NM (Q1) 158
Figure 9.3 General situation awareness - narrower hazards with and without defects (Q2 & Q3) 159
Figure 9.4 Anticipation of interactions - hazardous interactions with respect to defects (Q4 & Q5) 160
Figure 9.5 Anticipation of interactions - hazardous interactions due to speed (Q6 & Q7) 162
Figure 9.6 Inferences – sensitivity of separation to initial variations (Q8) 163
Figure 9.7 Inferences – separation histograms at 50 seconds (Q9) 164
Figure 9.8 A Close Encounter – to derive inferences of deviation from (Q10) 165
Figure 9.9 Inferences – Bayesian inferences in HAZOP & dynamics (Q10) 165
Figure 9.10 Prioritising risks – therefore which list is inferred as true? (Q11) 166
Figure 10.1 Sense and Avoid Scenario System Goal Model 168
Figure 10.2 Questionnaire Respondent Agreement Distributions (for 18 respondents) 169
Figure 10.3 Model versus Consensus – “expert” respondent agreement maps 170
Figure 10.4 Responses to Questions Concerning Anticipation of Interactions 171
Figure 10.5 Comparison of Test Case Rank Order for Diagnosis and Emergence 173
Figure 10.6 Lyapunov Distributions for Test Cases TC1 (left) and TC2 (right) to 25 Seconds 173
Figure 10.7 Lyapunov v Airprox Correlation as Scatter Plot (9000 data points) 175
Figure 10.8 Lyapunov v Airprox Correlation as Linear Regression Plots 175
Figure 10.9 Agreement & Authority by Experience (< 5 years – left, > 5 years – right) 177
Figure 10.10 Observed Control Flaw Relative Likelihoods for Test Case 6 178
Figure A.1 Initial STPA Based HAZOP: Single Deviation Type (Manoeuvre) 195
Figure A.2 Final Enhanced HAZOP: Cases for Approaches from Ahead, Behind and Left 197
Figure A.3 Final Enhanced HAZOP (cont.): Cases for Approaches from the Right 198
Figure A.4 Final Enhanced HAZOP: Consequences for Approaches from Ahead 199
Figure A.5 Final Enhanced HAZOP: Consequences for Approaches from the Left 199
Figure A.6 Final Enhanced HAZOP: Consequences for Approaches from Behind (Left and Right) 199
Figure A.7 Final Enhanced HAZOP: Consequences for Approaches from the Right 200
Figure A.8 Hazard Test Case Assignment by Direction of Approach 202
Figure A.9 Collision Avoidance Process Control Loop context: Hazard Vignettes and Flaws 204
Figure A.10 Control Flaw Relative Likelihoods for Test Cases 6, 1 and 5 from Bayesian Learning 211
Figure A.11 Control Flaw Relative Likelihoods for Test Cases 3 and 4 from Bayesian Learning 212
Figure A.12 Lyapunov Exponent Distributions at 25 seconds for TC 6 213
Figure A.13 Lyapunov Exponent Distributions at 25 seconds for TC 1 214

Figure A.14 Lyapunov Exponent Distributions at 25 seconds for TC 5 215
Figure A.15 Lyapunov Exponent Distributions at 25 seconds for TC 3 216
Figure A.16 Lyapunov Exponent Distributions at 25 seconds for TC 4 217
Figure A.17 Probability Density Functions for Sample Averages of Estimation 219
Figure A.18 Probability Density Functions for Sample Averages of Latency 220
Figure A.19 Vignette Bayesian Monte Carlo parameter sensitivity surfaces 222
Figure A.20 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TC 9 223
Figure A.21 Scatter Plots with respect to Initial Relative Heading including TC 9 224
Figure A.22 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TC 11 225
Figure A.23 Scatter Plots with respect to Initial Relative Heading including TC 9 226
Figure A.24 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TCs 7 & 8 227
Figure A.25 Scatter Plots with respect to Initial Relative Heading including TCs 7 & 8 228
Figure A.26 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TC 12 229
Figure A.27 Scatter Plots with respect to Initial Relative Heading including TC 12 230
Figure A.28 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TC 10 231
Figure A.29 Scatter Plots with respect to Initial Relative Heading including TC 10 232
Figure A.30 Hazardous Outcome Test Cases: Flaws, Deviations and Consequences for Approaches from Ahead and the Right 234
Figure A.31 Hazardous Outcome Test Cases: Flaw, Deviation, Perceived Estimation, and Latency Likelihoods, and Behaviours for Approaches from Ahead and the Right 235
Figure A.32 Hazardous Outcome Test Cases: Flaws, Deviations and Consequences for Approaches from Behind and the Left 236
Figure A.33 Hazardous Outcome Test Cases: Flaw, Deviation, Perceived Estimation, and Latency Likelihoods, and Behaviours for Approaches from Behind and the Left 237
Figure C.1 Questionnaire Responses for 6 : 12 split (any 6 agreements from 18 respondents). 251
Figure C.2 Questionnaire Respondent Agreement Distributions (for 18 respondents). 251
Figure C.3 Questionnaire Responses for 4 : 4 split (any 4 agreements from best 8 respondents). 252
Figure C.4 Questionnaire Respondent Agreement Distributions (for best 8 respondents). 252
Figure C.5 Model v Consensus – respondent agreement matrix (including R9 in best 8). 253
Figure C.6 Normalised average agreement & authority weighted by claimed experience 253
Figure C.7 Model v Consensus – respondent agreement matrix (including R5 in best 8). 254
Figure C.8 Normalised average agreement & authority weighted by claimed experience 254

List of Tables

Table 2.1 Categorising the causes of the Kegworth and Nogales accidents. 32

Table 2.2 Complex v Linear Systems after Perrow [87] 33

Table 2.3 Original Autonomous Control Levels (ACL) [19] 38

Table 2.4 Pilot Authorisation and Control of Task (PACT) Levels [106] 39

Table 2.5 Probability and Severity of Failure Conditions, UAS 1309 [16] 43

Table 3.1 STPA System Intent Specification Levels (simplified) [69, 86] 57

Table 5.1 Hazards identified in the report of NIAG SG-134 [81] 79

Table 6.1 Initial HAZOP Table 85

Table 7.1 Dominant Air Proximity Causal Factors [111] 108

Table 7.2 Air Proximity (Airprox) Classifications [111] 112

Table 8.1 Observed Inadequate Control Actions for a Unique Airprox Event 147

Table 8.2 Enhanced HAZOP Table 151

Table A.1 Hazard Test Case Assignment by Direction of Approach, Speed, etc. 203

Table A.2 Airspace Interaction Airprox Category Rates of Occurrence by Dataset (Test Cases) 203

Glossary

AAIB	Air Accident Investigation Board
ACL	Autonomous Control Level [19, 25]
AI	Artificial Intelligence
AMC	Acceptable Means of Compliance
ANO	Air Navigation Order
ARP	Aerospace Recommended Practice
ASTRAEA	Autonomous Systems Technology Related Airborne Evaluation & Assessment
ATM	Air Traffic Management
CAA	Civil Aviation Authority
CF	Control Flaw
CPA	Closest Point of Approach
CRI	Certification Review Item
CS	Certification Specifications
DD	Dependence Diagram
DoD	Department of Defense
FAA	Federal Aviation Administration
FDAL	Function Development Assurance Level [101]
FHA	Functional Hazard Assessment [101]
FMEA	Failure Modes and Effects Analysis
FMECA	Failure Modes, Effects and Criticality Analysis
FMES	Failure Modes and Effects Summary
FTA	Fault Tree Analysis
FTLE	Finite Time Lyapunov Exponent
GCS	Ground Control Station
HAZID	Hazard Identification
HAZOP	Hazard and Operability study [13]
ICA	Inadequate Control Action [104]
IDAL	Item Development Assurance Level [101]
IFR	Instrument Flight Rules
MBSE	Model Based Systems Engineering
MOD	Ministry of Defence
NCT	Non-Cooperative Traffic
NTSB	National Transportation Safety Board
PACT	Pilot Authority and Control of Task [106]
PASA	Preliminary Aircraft Safety Assessment [101]

PoC	Point of Conflict
PSSA	Preliminary System Safety Assessment [101]
QFD	Quality Function Deployment
RPAS	Remotely Piloted Aircraft System [78]
SCTA	Service du Contrôle du Traffic Aerienne
SD	System Dynamics
SoS	System of Systems
STAMP	System-theoretic Accident Model and Processes [71]
STPA	STAMP-based Analysis / Systems Theoretic Process Analysis [71]
SysML	Systems Modelling Language [83]
TC	Test Case
UAS	Unmanned Aircraft System [78]
UAV	Unmanned Aerial Vehicle [78]
UKAB	United Kingdom Air-proximity Board
UML	Unified Modelling Language
VFR	Visual Flight Rules
V&V	Validation and Verification

Part 1. Introduction & Literature Review

Chapter 1. Introduction

"Ah, this is obviously some strange usage of the word 'safe' that I wasn't previously aware of."

Arthur Dent, "The Hitchhikers Guide to the Galaxy" by Douglas Adams

1.1. Background, Aim and Objectives

An Unmanned Aerial Vehicle (UAV) is expected to be granted a certain amount of autonomy in the execution of its routine behaviour, at the least as a safeguard against any reduction or loss of situational awareness on the part of its human operator, whether through degraded or lost communications or simply due to the physical and temporal remoteness of the operation. In the case of an air vehicle, especially a fixed wing air vehicle, it is not as though this safeguard might be as simple as bringing the vehicle to a complete halt and then to await further instructions. It has its own momentum and intrinsic dynamics driving the continued evolution of the system state, with or without any intelligent external intervention – dictated by the physics of flight at least, with perhaps also on-board sensory perception and intelligence providing a means to continue to obey the rules of the air.

Given that such behaviour is dynamic and interactive, possibly involving another independent entity bound by the same physical laws and rules of behaviour, then a safety assessment of a system capable of such autonomous safeguarding behaviour necessarily entails a degree of complexity. Assuming that the likely complexity of behaviour in an autonomous system has to be considered throughout the development process, from the system concept stage onwards, then it is necessary to take a systematic approach to develop the necessary models of behaviour. This might include considering the formulation of any hazardous behaviour model so as to be used also in a diagnostic role with the potential to diagnose system design flaws. Undoubtedly, particular system behaviours might be surprising when thus modelled and the trust in any such model also needs to be established before it can be used for safety assessment or diagnosis.

Within the United Kingdom "Unmanned Aircraft System Operations in UK Airspace – Guidance" (CAP 722 [17]) is intended to assist those who are involved in the development of Unmanned Aircraft Systems (UAS) in describing the identification of the route to certification for future UK operations. This highlights the safety requirements that have to be met, in terms of airworthiness and operational standards, with the intent to develop a regulatory framework enabling the full integration of UAS activities with manned aircraft operations. In this, autonomy is described as a capability to operate without human control or direct oversight. Within this definition there is then provision for a range of lower level capabilities, from simple advisory systems where humans make the decisions and execute the actions, through intermediate levels of "semi-autonomous" systems. Provision is not yet made for fully autonomous airborne systems. CAP 722 does promote the concept of a decomposed autonomous capability to address the perceived inherent complexity of autonomous operations. It is suggested that an autonomous capability might be comprised of various decision-making sub-systems, each responsible for its own domain of concern. In this view an autonomous capability might be made up from autonomous flight management, sense and avoid, route management, power management and prognostic health management systems, for example [17]. Some, or all, of these systems would collaborate with the human UAS pilot / operator. No specific guidance is given as to how the partitioning and decomposition might be identified, affected or managed.

In this thesis the broad hypothesis that risk events within complex systems are caused by deviations from design or operating intentions and unanticipated non-linear causal interactions among system elements that violate safety constraints is accepted as true. Also it is believed that these interactions cannot be adequately intuitively assessed without a representation of likely system dynamics. It is the author’s view that by augmenting the mental processes and expertise with live models of system dynamics, and drawing from these statistical inferences, helps to validate the assumptions formed within a Hazard and Operability study (HAZOP) and point towards the refinement of the mental models used – in effect providing a necessary degree of self-testing of the hypothetical basis of the HAZOP, with the effect of improving the specification of software that is to exhibit intended behaviour in unexpected situations.

The aim of this work is to investigate the interrelated nature of system modelling and inferred belief with respect to hazards arising within complex behaviour, such that plausible and informative models describing system dynamics and evidence of cause and effect might be systematically employed in the updating of more abstract representations of belief concerning hazard relationships, as captured within a Hazard and Operability study.

So as to formulate a relationship between the HAZOP process, a system dynamics model and an inference mechanism validating[1] the deviation – system flaw / defect – consequence relationships formed within the initial Hazard and Operability study, the following objectives ought to be satisfied:

- Model behaviour, competence and situational awareness with a system dynamics (Monte Carlo) simulation incorporating three dynamically and “autonomously” interacting rule-bases (constituting in effect a three-body problem) – exploring sensitivity to initial conditions and capturing measures of potentially behaviourally complex and emergent behaviour;
- Formulate this model, and the later inference model, within a provisional systems engineering modelling framework – formulating appropriate SysML [83] constructs and attempting to position the two (dynamics and inference) models as an extension to, or alongside, ARP 4754A [101];

[1] If not validating then at least tilting revised belief towards the evidence generated by the dynamics model.

- Consider the views of “experts” and “practitioners” with a questionnaire based validation of observed dynamical behaviour arising from the above model(s);
- Develop an inference based validation of behavioural flaws and defects with respect to anticipated deviations formed within the initial HAZOP;
- Extend the initial HAZOP so as to incorporate the above derived estimated measures of emergent behaviour and inferences of uncertainty or otherwise; incorporating evidence derived from observed system dynamics encapsulating likely and plausible behaviour.

Therefore the following sections outline a scheme to incorporate HAZOP to explore some particular complex and emergent behaviour, and associated representations, of a putative UAV sense and avoid subsystem. The intention is to facilitate an approach to hazard identification as a UAS extension to a UAV Preliminary Aircraft Safety Assessment (PASA) [101].

A question might be addressed at this juncture as to what extent this modelling approach might be justified as representing an “autonomous system”. Current doctrine [78] on the issue of Automated versus Autonomous Systems defines an Automated System as one *“... Knowing the set of rules under which it is operating means that its output is predictable.”* Against this, an Autonomous System is considered to incorporate *“... understanding and its perception of its environment, such a system is able to take appropriate action, from a number of alternatives, without depending on human oversight and control, although these may still be present. Although the overall activity of an autonomous unmanned aircraft will be predictable, individual actions may not be.”* [78]

1.2. Hypothesis

In formulating suitable objectives for this research, and in the context of the previously stated broad hypothesis that risk events within complex systems are caused by deviations from design or operating intentions and unanticipated non-linear interactions, a hypothesis relating specifically to Hazard and Operability (HAZOP) studies has been formulated against which the project outcomes are later assessed. In this it is argued that a conventional, and essentially manual, HAZOP process can be improved with information obtained with model-based dynamic simulation, using a Monte Carlo approach, to update a Bayesian Belief model representing the expected relations between cause and effects – and through this approach thereby produce an enhanced HAZOP.

Additionally, the inclusion of suitable metrics indicative of system behavioural complexity, or rather proxies by way of measures of system stability and complex behaviour, along with domain specific safety-constraint violation detection metrics (e.g. Air Proximity or other near-miss categories), might together further enhance comprehension of the specific system safety concerns arising within the various system behaviours – even where it might later be thought that the observed behaviours were obvious after the event.

1.3. Motivation in Context

Aerospace Recommended Practice (ARP) 4754A [101] describes the information to be produced by the recommended Functional Hazard Assessment (FHA) process, and sits at the core of aircraft system development processes with system safety assessment guidance to one side and item assurance guidance to the other – as depicted in Figure 1.1 below. Guidelines of recommended processes state that information is to be produced both for individual aircraft-level functions and combinations of aircraft-level function. This implies an iterative approach with an incremental incorporation of additional mitigation functions. Specifically, the information to be produced includes:

a) Identification of related failure condition(s) and their effects;

b) Classification of each failure condition based on the identified effects – as defined in UAS 1309 [16] in the case of an unmanned system;

c) Identification of the required system development assurance level;

d) A statement outlining what was considered and what assumptions were made when evaluating each failure condition.

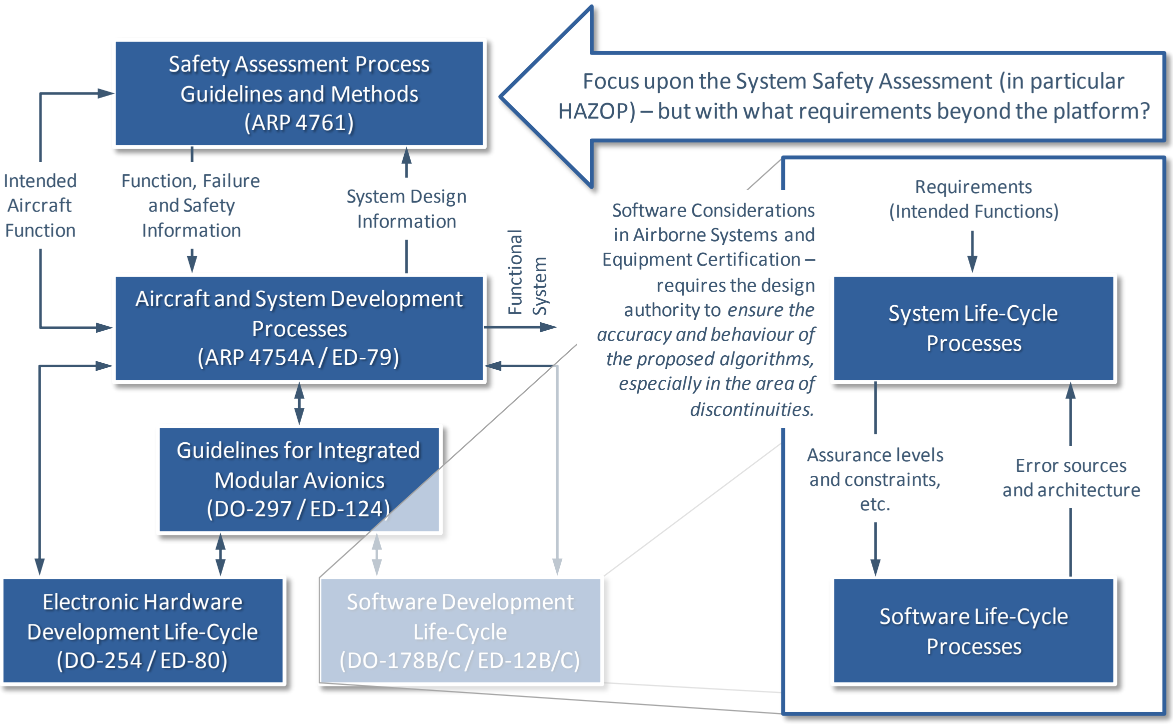

Safety Assessment Process Guidelines and Methods (ARP 4761)
Focus upon the System Safety Assessment (in particular HAZOP) – but with what requirements beyond the platform?
Intended Aircraft Function
Function, Failure and Safety Information
System Design Information
Functional System
Software Considerations in Airborne Systems and Equipment Certification – requires the design authority to ensure the accuracy and behaviour of the proposed algorithms, especially in the area of discontinuities.
Requirements (Intended Functions)
Aircraft and System Development Processes (ARP 4754A / ED-79)
System Life-Cycle Processes
Guidelines for Integrated Modular Avionics (DO-297 / ED-124)
Assurance levels and constraints, etc.
Error sources and architecture
Electronic Hardware Development Life-Cycle (DO-254 / ED-80)
Software Development Life-Cycle (DO-178B/C / ED-12B/C)
Software Life-Cycle Processes

Figure 1.1 Certification Considerations for Complex Aircraft Systems [95, 97, 100]

The goal of these guidelines is to clearly identify the circumstances and severity of each failure condition along with rationale for its classification. However, the system reliability assessment methods that are recommended in ARP 4761 [100] address models of accident causation in what has been described as “chain of events” [70], where Leveson argues that accidents are not necessarily caused by an item failure but often rather through inadequate constraints against bounded complex behaviour. Consequently this research study has as an aim the development of a broader Systems approach towards hazard assessment for evaluating hypothetical deviations from declared intent within a behavioural modelling framework, broadly in the first instance based on the STAMP [67, 70] hazard assessment methodology drawing upon both STPA [86, 105] and HAZOP [13, 66]. It is also proposed that this might facilitate exploratory dynamic hazard assessment alongside a Preliminary Aircraft Safety Assessment (PASA) [101] – which could be seen as fitting into the system development process where depicted below in Figure 1.2.

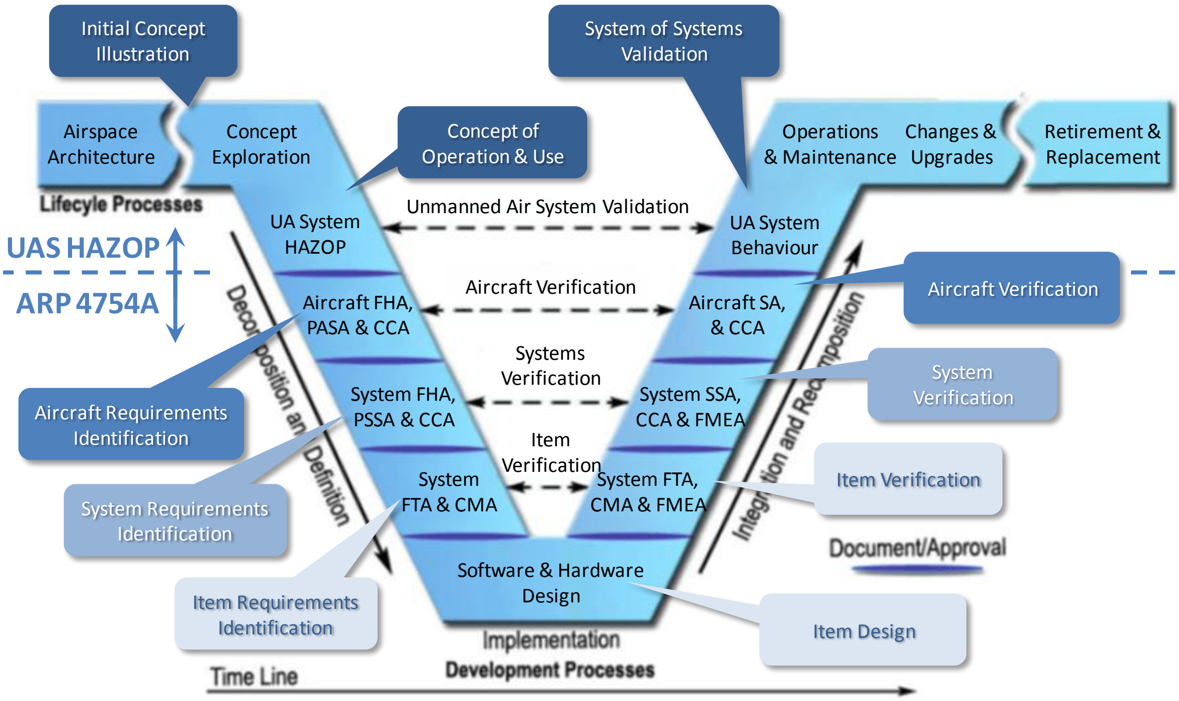


Figure 1.2 System Development Process with HAZOP, adapted from ARP4754A [101]

Therefore the consideration of intent, rather than item failure, ensures that extending FHA with HAZOP is more appropriate for earlier use, perhaps to be applied even before the Preliminary Aircraft Safety Assessment; whereas FMEA requires the existence of a specific and detailed system design. Again, this suggests that a first-cut hazard or accident model ought to facilitate the modelling of a system of interest's intended behaviour whilst also capturing the effect of deviations from declared intent and emergent behaviour. Consequently, when taking into account the argument against "chain of events" models, it is proposed that a behavioural modelling capability encompassing system intent is a requirement of this approach. Aside from a particular system goal, or sub-goal, the system intent is also to enforce a particular safety constraint.

Taking into account these and the previous considerations regarding the application of HAZOP into the domain of FHA, along with the intention to facilitate an inference mechanism for the diagnosis of suspected design faults, together suggests a data-flow model as depicted in Figure 1.3. This model represents the three processes required to: identify a hazard through a variant of STPA to construct an initial HAZOP table, produce the resultant modelled behaviours, and incorporate an inference mechanism that associates the deviations and flaws with the relevant violations. This enhancement of a plain hazard modelling capability is investigated to determine a means to quantify and rank the sensitivity to different causal events, and to explore hazard scenarios and vignettes, through a "what-if" approach. For example this could include false alarm rate effect analysis, or analysis of an allowable tolerance on perceived separation estimates – and within this employ Monte Carlo behavioural simulation to produce distributions of constraint violations associated with degrees of belief or Bayesian probability estimates representing the flaw or defect causal relationships with respect to outcomes.

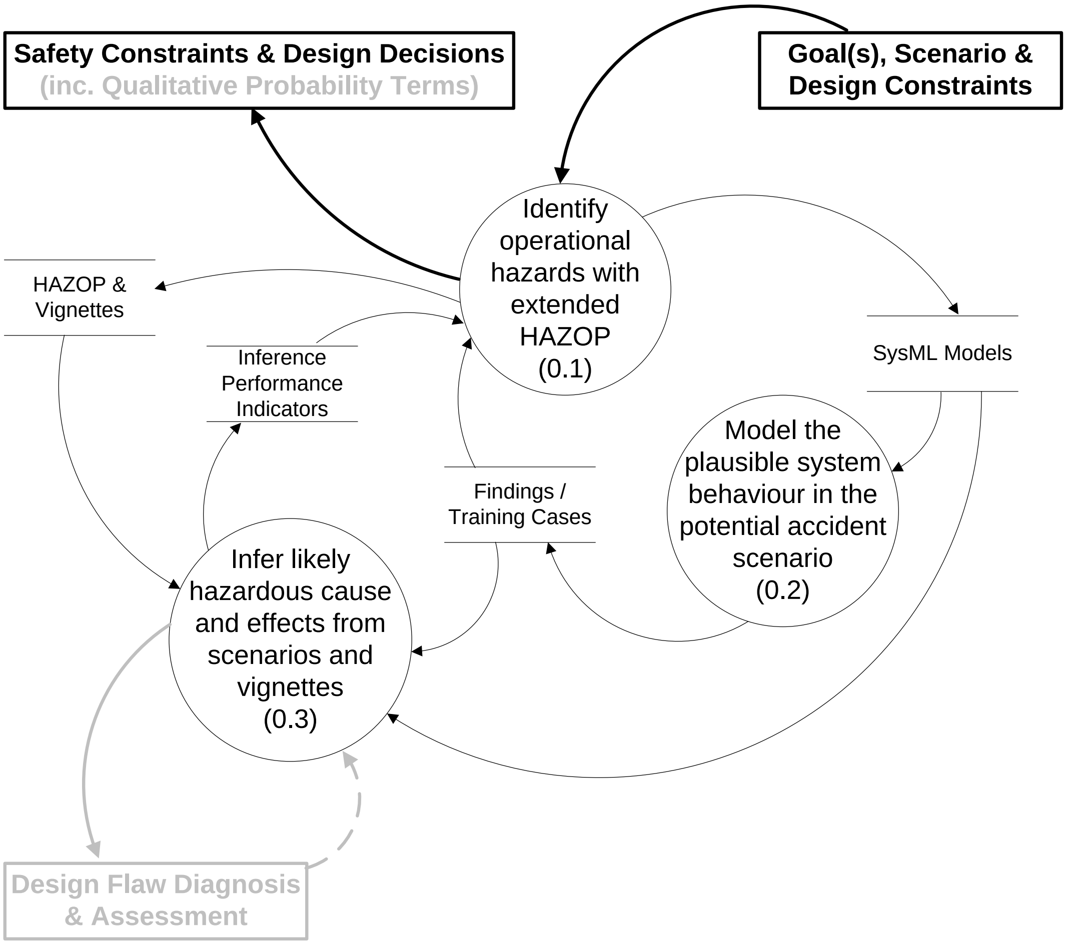

Safety Constraints & Design Decisions
(inc. Qualitative Probability Terms)
Goal(s), Scenario & Design Constraints
HAZOP & Vignettes
Identify operational hazards with extended HAZOP (0.1)
Inference Performance Indicators
SysML Models
Model the plausible system behaviour in the potential accident scenario (0.2)
Findings / Training Cases
Infer likely hazardous cause and effects from scenarios and vignettes (0.3)
Design Flaw Diagnosis & Assessment

Figure 1.3 Outline HAZOP inference process role [32]

Therefore, in the nature and occurrence of an expected defect across a number of slightly differing but otherwise comparable models of systems – differing say in tolerance or performance – sensitivity in the belief relationship ought to be amenable to discovery through Monte Carlo simulation of either simplified system dynamics models and / or discrete event simulations. To obtain a belief value for a qualitative probability statement such as "Extremely Remote" would likely require a larger model including estimates of the likely number of scenario events or vignettes per vehicle or system operational life, and representation of the total number vehicles or systems of the same type.

1.4. Project Overview

A model has been developed for a simplified exemplar describing particular vignettes within a scenario of an air-proximity hazard arising between two air-vehicles. This model is considered to be of sufficient fidelity so as to characterise the inherent complexity and emergent behaviour that might arise with this type of interaction.

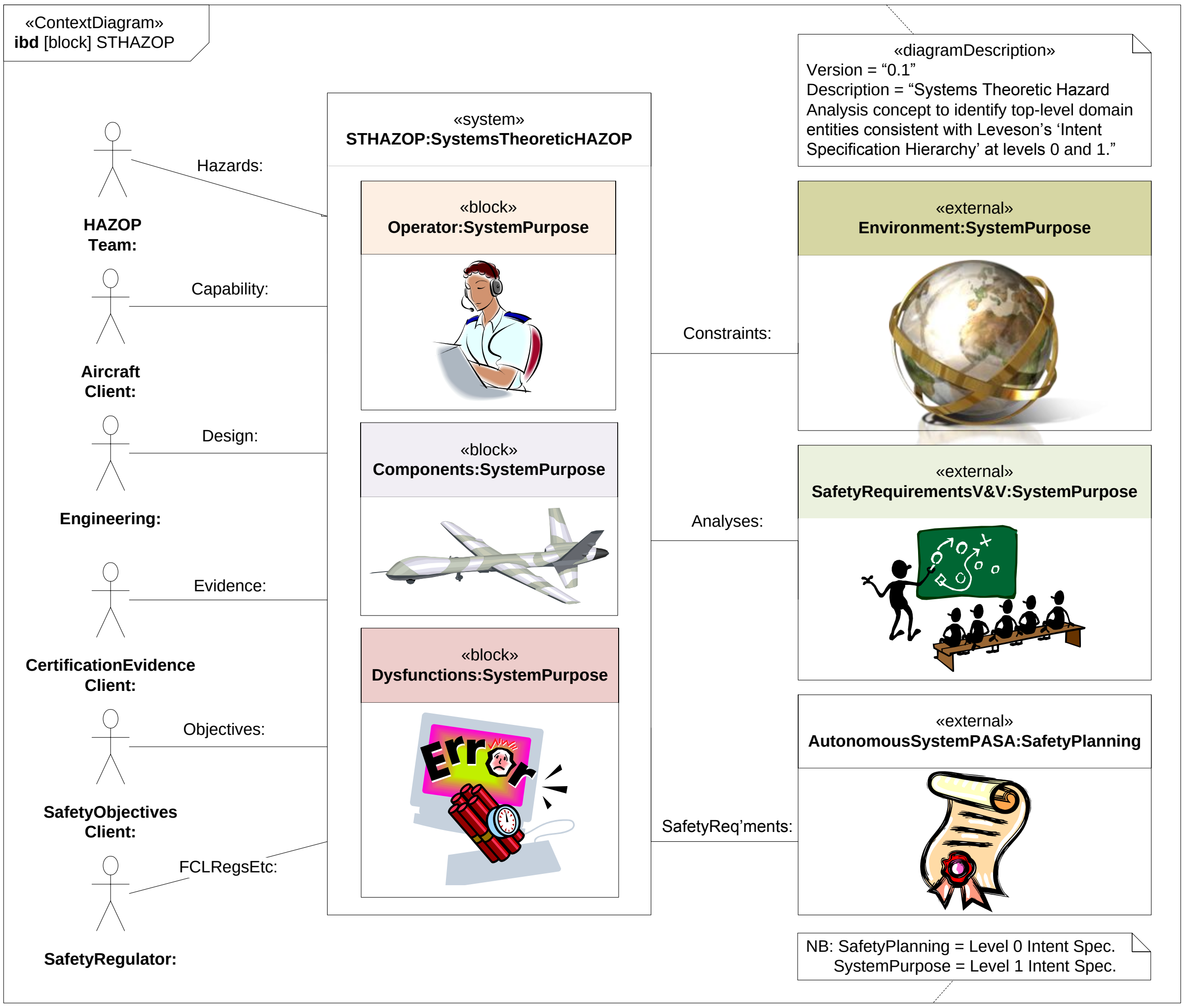

«ContextDiagram»
ibd [block] STHAZOP
«diagramDescription»
Version = "0.1"
Description = "Systems Theoretic Hazard Analysis concept to identify top-level domain entities consistent with Leveson's 'Intent Specification Hierarchy' at levels 0 and 1."
«system»
STHAZOP:SystemsTheoreticHAZOP
HAZOP Team:
Hazards:
Aircraft Client:
Capability:
Engineering:
Design:
CertificationEvidence Client:
Evidence:
SafetyObjectives Client:
Objectives:
SafetyRegulator:
FCLRegsEtc:
«block»
Operator:SystemPurpose
«block»
Components:SystemPurpose
«block»
Dysfunctions:SystemPurpose
Error
Constraints:
Analyses:
SafetyReq'ments:
«external»
Environment:SystemPurpose
«external»
SafetyRequirementsV&V:SystemPurpose
«external»
AutonomousSystemPASA:SafetyPlanning
NB: SafetyPlanning = Level 0 Intent Spec.
SystemPurpose = Level 1 Intent Spec.

Figure 1.4 Autonomous Systems HAZOP Context

This study incorporates a broad Systems Engineering approach towards hazard assessment incorporating system dynamical behaviour to evaluate behaviour as hypothetical deviations from declared intent within a behavioural modelling framework.

The initial hazard assessment methodology is styled upon that derived from the Systems Theoretic Accident Model and Processes (STAMP [71]) as developed by Leveson, and the associated Systems

Theoretic Process Analysis (STPA [71]). Within this Systems Engineering context, the work as presented here together might be seen as a portfolio of four smaller but integrated projects – a specific and novel Requirements Analysis phase, a construction and generalised Systems Dynamics modelling phase, a Hazard Belief Inference modelling phase, and a Validation by Questionnaire and general verification phase. The proposed process model, facilitating the application of HAZOP from the system concept stage onwards, also incorporates dynamical modelling and an associated inference model; together closing the analysis process with the designer in the loop throughout.

It is assumed that in the case of complex system behaviour a difficulty lies in assessing hazard and operability issues with human reasoning, intuition and mental methods alone – and hence the requirement is to effectively provide an additional tool for the "diagnosis" of system defects arising even at the system concept stage. This includes the development of a provisional systems modelling framework, describing the context of the hazard modelling, as in Figure 1.4 along with the context of the scenario and system goal (Figure 5.5, etc.). This modelling framework also encompasses the sequence of actions (Figure 6.4, etc.) required to build comprehensive interrelated HAZOP, dynamics and inference models; plus the various generic physical, dysfunction (defect), safety violation, model state monitoring and behaviour inference capture parametric models (Figure 7.3, etc.), along with the respective specific constraint descriptions (Figure 7.2, etc.).

The actual system dynamics represents constrained simulated avoidance behaviour within a two-dimensional model of the environment (strictly 2.xD, as "bank angle" also has some meaning within the model). There is one fundamental safety requirement, or objective, which requires that each modelled entity avoids the other with a minimum separation of 500 feet, otherwise the fundamental safety constraint is violated. Certain "surprises" within the relative perceptions of behaviour can also be seen to trigger additional safety constraint violations with respect to Air Proximity reporting. Analyses are in effect exported from the dynamics model as Monte Carlo data, to be processed within the inference model – in turn tied to the initial STPA / Hazard and Operability study. Whilst the model intentionally represents various components comprising the system, along with three specific models of dysfunction for investigation, there is no meaningful representation of an operator within the current model.

Cases exercising the model are created from a comprehensive set of combinations of low speed (71.5 – 117 knot), medium speed (117 – 195 knot), and high (or excessive) speed (195 – 305 knot) interactions; combined with three dysfunctions, especially wherever sensitivity to these is either anticipated or later observed. The position of the entity representing the UAV / UAS always forms the initial reference point, with the "non-cooperative traffic" positioned on a converging heading with an initial separation typically of 0.5 nautical miles. Interactions are then investigated where the direction of approach arises within five broad segments encompassing the 360 degrees circumscribing the 0.5 nm radius around the UAV, and with a smaller number of cases within narrower windows of approach where a particular sensitivity has been observed (q.v. Table A.1, Appendix A.

Enhanced HAZOP Employing Inference Test Cases).

The results arsing with the initial system dynamics model and an initial attempt at capturing the inference in this behaviour were presented in the form of an objective multiple-choice questionnaire to a group of “experts” and “practitioners”, representing a HAZOP team. The responses to this questionnaire where analysed using a non-parametric statistical approach (Cohen’s Kappa [20]). In response to the questionnaire some improvements were subsequently made to the dynamics and perceptions within the final model, whilst the initial attempt at inference modelling was completely overhauled as a result (q.v. Chapter 9, along with Appendices B and C; Hazards Advising Autonomy: Questionnaire, and Respondent Results and Analysis).

1.5. Contribution of Thesis

The purpose of this thesis is to investigate the interrelated nature of system modelling and inferred belief with respect to hazards arising within potentially complex behaviour, such that plausible and informative models describing system dynamics might be systematically employed to more fully inform abstract representations of belief concerning hazard relationships. By informing and improving the HAZOP process, the contribution of this thesis is to make unanticipated hazardous conditions more predictable. This is achieved through combining models representing the simulation of the system dynamics and a corresponding representation of the relevant cause and failure effects, as described in the HAZOP process, presented in a Bayesian Belief Network.

In attempting to test the validity and acceptance of this, by employing a properly formed questionnaire wherein a number of “experts” represented a group consensus regarding belief in the observed system behaviour, and whilst they also reasonably disputed certain aspects of the modelled behaviour, where they agreed with each other they also showed a slightly stronger agreement with the modelled behaviour. This implies that whilst certain aspects of the modelled behaviour where accepted a priori by this group, other aspects could surprise even the “more expert” group. Therefore, system dynamics modelling may usefully provide additional insight beyond individual experience and intuition – principally the characteristic sought of a HAZOP team. However this has been tested only with 18 respondents here, a more in-depth investigation would be needed to properly test this assumption. Alternatively, perhaps all that is required is for there to be a properly formulated “reality check” so as to baseline a representative system dynamics model, and then accepting this is to also to at least conditionally accept the other more esoteric and surprising behaviours as possibilities and counter examples for hazardous outcomes in the particular system of interest.

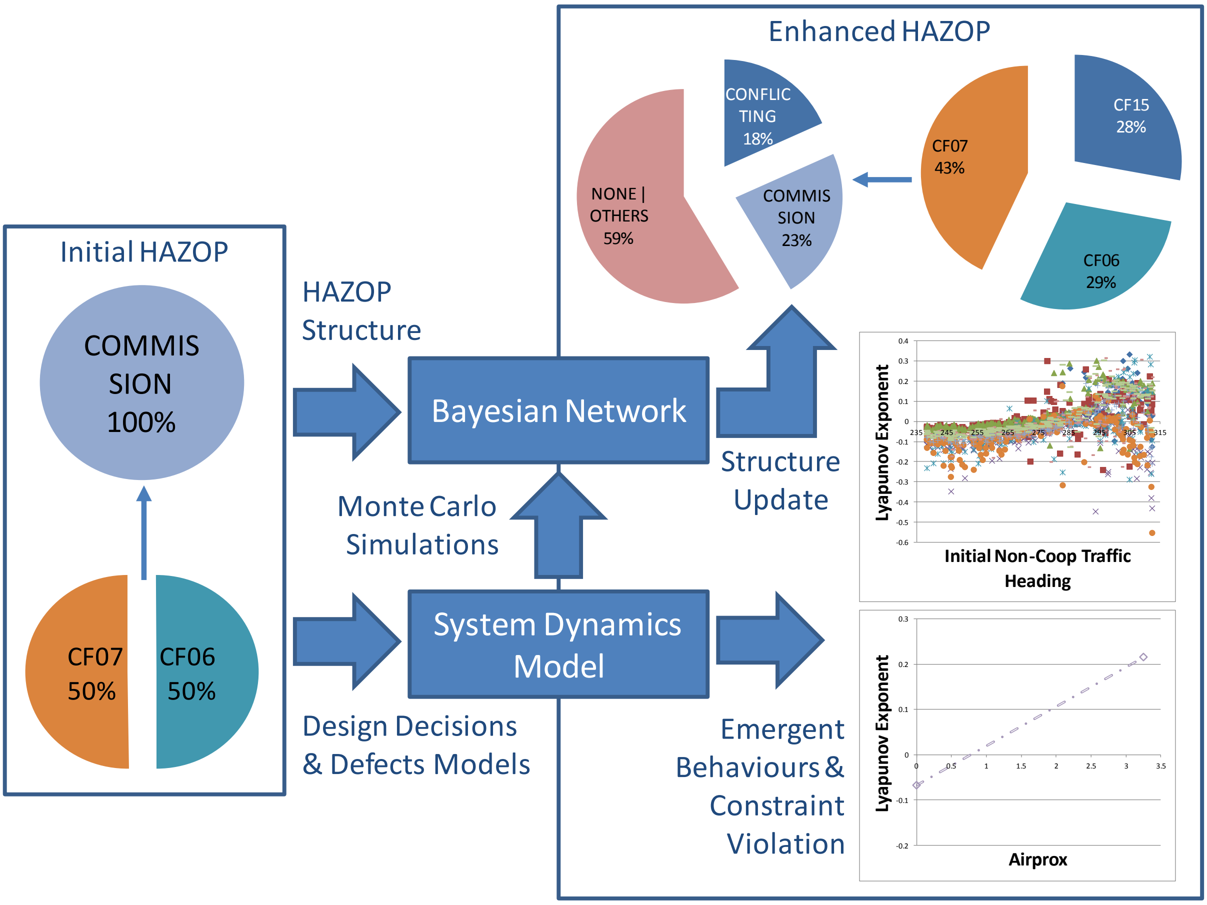


Figure 1.5 An Enhanced HAZOP Framework

Overall, an enhanced HAZOP framework is suggested as depicted in Figure 1.5. This posits that the HAZOP / STPA hazard assessment and mitigation processes might be jointly enhanced with the production of a simplified representative system dynamics model and associated inferred belief model. The system dynamics model captures the defect, control flaw and design decisions provisionally identified in the STPA / HAZOP process, where an inferred belief model captures the provisional belief structure defined with the HAZOP table. Combining the two using Monte Carlo data to update belief in the HAZOP table not only conditions the original belief but also may find new believable associations.

As a Systems Engineering project, addressing the problem with an interdisciplinary and top-down approach viewing the system as a whole, this work consists of the following tangible aspects:

- *Inputs to the project*: A conventional ***HAZOP*** study and a slightly adapted variant of ***STPA*** are incorporated as springboards for the project.
- *Models created within the project*: A ***Systems Dynamics model*** is created, structured in part from information derived from the STPA process. In conjunction with this **a *Bayesian belief model*** is also created, structured in part from structural information derived from the HAZOP process.

- *Outputs arising from the project*: Cases embodying emergent behaviours within the system dynamics are identified with ***Lyapunov Exponent*** values derived directly from Monte Carlo simulations of the system dynamics model. System flaw, deviations and constraint violation case are similarly drawn from Monte Carlo simulations of the system dynamics model, and applied as experience to produce ranked and weighted combinations of these simulation and belief products. These are applied to the deviation cases from the original HAZOP, modifying some beliefs and raising awareness of the more significant risks, and thereby enhancing the hazard assessment by ***incorporating uncertainty and belief*** derived from modelling. Given the problem of validating behaviour of a synthetic model in the absence of real data a trial method employing a ***validation questionnaire*** attempts to draw knowledge from "practitioners" and domain "experts".
- Combining the respective measures of belief within specific test cases with measures of constraint violation and a measure of emergent behaviour helps identify the most hazardous cases within an enhanced HAZOP. A significant aspect of this lies in the choice of the appropriate parameter deviation and guidewords in the first formulation of the HAZOP, and how the choice of a particular parameter subsequently maps onto the system behaviour, emergent or otherwise – which in this case lies in treating the abstract noun "manoeuvre" as the system parameter of interest.

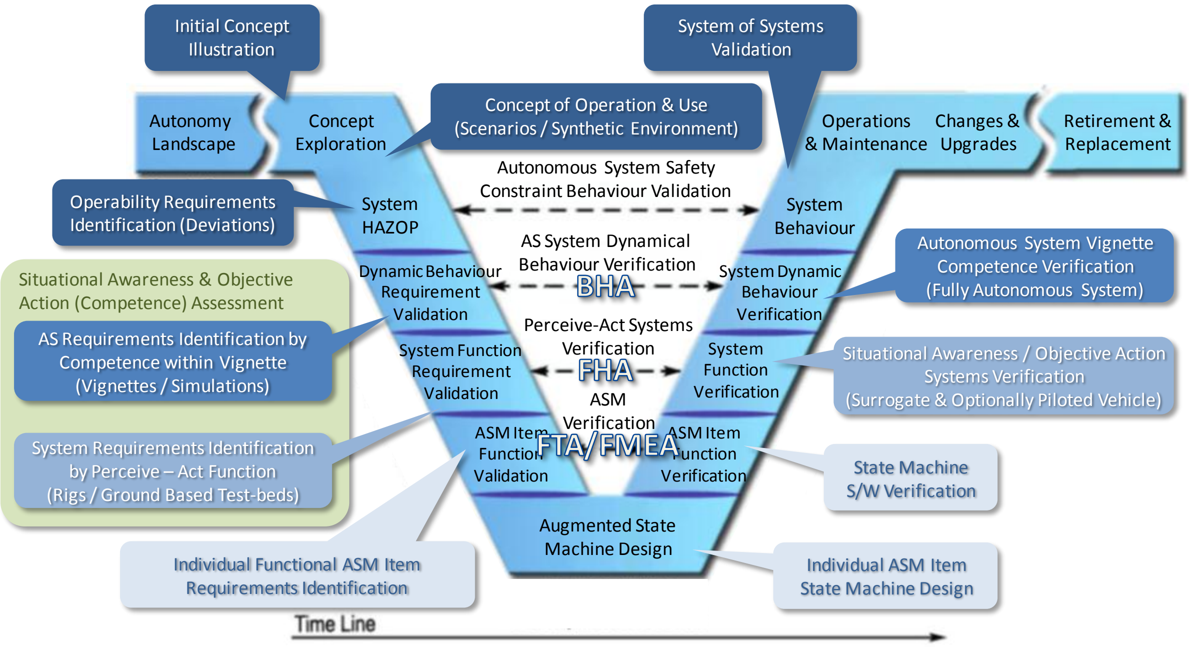

Initial Concept Illustration
System of Systems Validation
Autonomy Landscape
Concept Exploration
Concept of Operation & Use (Scenarios / Synthetic Environment)
Operations & Maintenance
Changes & Upgrades
Retirement & Replacement
Autonomous System Safety Constraint Behaviour Validation
Operability Requirements Identification (Deviations)
System HAZOP
System Behaviour
Situational Awareness & Objective Action (Competence) Assessment
Dynamic Behaviour Requirement Validation
AS System Dynamical Behaviour Verification
BHA
System Dynamic Behaviour Verification
Autonomous System Vignette Competence Verification (Fully Autonomous System)
AS Requirements Identification by Competence within Vignette (Vignettes / Simulations)
Perceive-Act Systems Verification
FHA
System Function Requirement Validation
System Function Verification
Situational Awareness / Objective Action Systems Verification (Surrogate & Optionally Piloted Vehicle)
System Requirements Identification by Perceive – Act Function (Rigs / Ground Based Test-beds)
ASM Verification
FTA/FMEA
ASM Item Function Validation
ASM Item Function Verification
State Machine S/W Verification
Augmented State Machine Design
Individual Functional ASM Item Requirements Identification
Individual ASM Item State Machine Design
Time Line

Figure 1.6 A Scheme for Incorporating Behavioural Hazard Assessment into ARP 4754A

Finally, an tentative view also points in the direction of incorporating behavioural hazard assessment (BHA) as a possible adjunct to the established functional hazard assessment (FHA) methodology as embodied within current guidance regarding the development of certifiable civil aircraft systems – Figure 1.6. This attempts to place the enhanced HAZOP framework in context with the Preliminary Aircraft Safety Assessment (PASA) within ARP 4754A [101].

1.6. Thesis Structure

This thesis is comprised of three larger parts. The first part encompasses a wide ranging review of the background and literature influencing this work. The second part presents and discusses the framework within which this hazard-centric study has been developed. The final part brings together the various outputs and conclusions arising from this study.

Part one, the introduction and literature review, is comprised of four chapters, as follows:

Chapter 1: This chapter provides an introduction and initial overview to the work. Here the scene is set in identifying a need to provide better understanding of hazards as might arise where a designer's account of likely system behaviour and dynamics arising with the autonomous system dynamics is insufficiently understood. As the reader will have read thus far, this broadly describes the motivation, hypothesis, scope and contributions of the work.

Chapter 2: This is the first of the three chapters considering the wider literature enveloping the subject, framed as a reflective regulatory context review; within which past accidents, certain aspects relating to human factors and situational awareness are considered – as all might inform the design of representative system models. In addition further defining aspects and constraints are considered, including models representing varying degrees of Autonomy, human equivalence, the relationship between safety and reliability, analytic limitations, the current regulatory guidance and approaches to hazard assessment; raising questions as to whether the hazard assessment approach ought to be focused on function or behaviour abstractions.

Chapter 3: The progressive issue of the state of the art and understanding in hazard identification is reviewed next. In terms of the aims of this work, the extension of the HAZOP method, work addressing hazard emulation, hazard identification and assessment, Bayesian expression of risk, and scenario and deviation analysis are considered together. Furthermore, how HAZOP and other scenario based methods might be extended with the application of the Systems Theoretic Accident Model Process (STAMP) and Systems Theoretic Process Analysis (STPA) is reviewed, giving way to a discussion of systems representation, the representation of risks and hazards, and the application of Bayesian and Monte Carlo methods to build models of belief.

Chapter 4: Completing the review of state of the art and domain literature the broader context of Systems Engineering is considered. The contrasting development life-cycles as embodied by the recommended systems engineering processes characterised within ARP 4754A / 4761 is compared with reference to the spiral development model applied to software systems development. The relevance of architecture and validation is considered in terms of design intent, and as used within this work, validation through questionnaire. This chapter rounds off with a short overview of what constitutes emergent and complex behaviour as understood in the robotic and autonomous systems community, concluding with a brief summary of the relevant influences and scope to be taken from the whole literature review.

Part two, encompassing the systems framework, comprises three further chapters formed around the conventional system engineering process:

Chapter 5: The first stage of any systems engineering process considers the system requirements. This describes where the broader project requirements were drawn from the (earlier) literature review using textual analysis and QFD. Within this a number of representative stakeholders are identified; including certification evidence clients, safety objective clients, the HAZOP Team, and the engineering function. The specific requirements relating to Sense And Avoid (SAA) Hazards and an approach to hazard scenario generation is then outlined.

Chapter 6: Having obtained or formulated a set of requirements, the system model design proceeds with a description of the modelling framework, with the collision avoidance scenario and behavioural intent modelling, and the system goal and emergency behaviour vignette modelling.

Chapter 7: This chapter continues with the system state modelling, including the parametric representations of the physical constraints, dysfunctions and defect properties, safety constraint and violation models, and the emergent system behaviour sensitivity modelling.

Chapter 8: Having incorporated the system model requirements, composed primarily of the safety and other constraints within the particular scenario, along with the resultant design decisions and expected system dysfunction models, the interface to the Bayesian Network representation of the HAZOP is then described; along with representations of the parametric models of the interfaces used in the capture of the various parameters collected in the Monte Carlo generated cases. The chapter concludes with a discussion of selected model test cases.

Part three, the final three chapters summarise the work and final outcomes of the project

Chapter 9: This chapter describes the design of the questionnaire and presents an interpretation of the results derived from the respondents to this questionnaire. These results are used to partially validate the system model requirements arising from the HAZOP in terms of their manifestation in the resultant system behaviour.

Chapter 10: The penultimate chapter provides a sketched summary of the principal results as arising with the aforementioned questionnaire responses, and offers a possible interpretation of the emergent properties in the results correlating safety constraint violation outcomes and corresponding measurements of the Lyapunov exponent. This is followed by a summary the lessons learnt, alternative approaches and improvements, and a suggestion as to industry use and benefits of exploitation, with a final section on suggested future work.

Chapter 11: The final chapter details the recommendations for implementation, reflecting upon the original stated objectives in Chapter 1, and finally draws together some conclusions relating to the need to consider complexity in system behaviour, a system dynamical model's relationship to hazard assessment, and how this sits within the context of HAZOP.

Annex: Completing the work, following the list of references, three appendices detail the composition of the enhanced HAZOP employing inference test cases, the Questionnaire, and finally the detailed questionnaire respondent results and analysis.

Chapter 2. Regulatory Context Review

"Human beings, who are almost unique in having the ability to learn from the experience of others, are also remarkable for their apparent disinclination to do so."

Douglas Adams, "Last Chance to See"

2.1. Accident Inevitability

This work in systems engineering is concerned with how hazards arise with complex interactions within systems and how these might be identified, especially where high levels of automation or autonomous behaviour are employed, plus how they might be represented and assessed. Often the very existence of some hazard may be obscured from the designer by the inherent complexity of the system, only to be discovered too late in the seemingly bizarre nature of the system behaviour from the point of view of the operator(s) working under pressure in adverse conditions. So as to set the scene, two seemingly different, although in some ways comparable, incidents illustrate where inherent complexity and tight coupling of actions and events have occasioned possibly avoidable air accidents had the operators better understood and correctly interpreted the system behaviour at the time. And yet, both of these accidents were perhaps inevitable given the expectations of the system designers and their understanding of the potential cause and event chains comprising a combination of physical failure, system design limitations, false or misinterpreted cues, situational awareness and human error.

2.1.1. Accidents Past

Seriously unpleasant and surprising things certainly do occasionally happen – especially so when the complexity of whatever unpleasantness then occurring overwhelms the operators.

The evening of the 8th January 1989 just outside Kegworth, Leicestershire, on the approach to runway 27 at East Midlands airport, was one such occasion and place. At 20:25 that evening British Midland Airways Boeing 737-400, G-OBME, came to rest 900 metres short of the runway threshold on the north-bound carriageway of the M1 after a damaged No. 1 engine suffered total mechanical failure on final approach, where the fully serviceable No. 2 engine had earlier been mistakenly shutdown in flight. As might be expected the series of events leading up to this accident started somewhat earlier, at 20:05, as reported by Trimble and Cooper [109] of the Air Accident Investigation Branch of the UK Department of Transport, wherein during the intervening twenty minutes the flight crew experienced severe engine vibrations and shuddering, incorrectly diagnosed the problem, shutdown the serviceable No. 2 engine, and finally attempted to land at East Midlands airport only with the faulty No. 1 engine.

The AAIB report for this accident (Conclusions 3(b), p148 [109]) identifies five factors that were deemed to have contributed to the incorrect response of the flight crew *" ... 1. The combination of heavy engine vibration, noise, shuddering and an associated smell of fire were outside their training and experience. 2. They reacted to the initial engine problem prematurely and in a way that was contrary to their training. 3. They did not assimilate the indications on the engine instrument display before the throttled back the No 2 engine. 4. As the No 2 engine was throttled back, the noise and shuddering associated with the surging of the No 1 engine ceased, persuading them that they had correctly identified the defective engine. 5. They were not informed of the flames which had emanated from the No 1 engine and which had been observed by many on board, including 3 cabin attendants in the aft cabin."* [109]

Consequently the case was made that the serviceable No. 2 engine had been wrongly shutdown by the flight crew as a result of an unapproved diagnostic procedure and a misreading of the relevant engine instruments. However, the No. 2 engine was not fully shutdown until two minutes and seven seconds into the incident, having only been reduced to idle subsequent to the first twenty-two seconds spent diagnosing the problem. At the point where the No. 2 engine throttle was reduced to idle the autothrottle also had to be disengaged, with the coincidental effect that the damaged No. 1 engine recovered from the compressor stalls and settled down to run at a slightly lower fan speed, albeit still with unacceptable vibrations indicated on the instruments, but with an absence of the more severe shuddering felt on the flight deck. It was this coincidence, if indeed it was mere coincidence, which wrongly convinced the flight crew that they had made the correct diagnosis during the hundred or so seconds before finally shutting down the No. 2 engine. Since the publication of the AAIB report it has been suggested[2] that the fuel surges associated with the damaged No. 1 engine, attempting to maintain speed whilst commanded by the autothrottle, abated as a direct consequence of the disengagement of the autothrottle. This action would therefore have produced the same effect of reducing the shuddering irrespective of which engine had been reduced to idle – and therefore providing false positive support to the flight crew's initial, albeit unapproved, diagnosis.

With advances in automation, for example automatic reconfiguration, it might be thought that automation alone might alleviate the prospect of such human error. The events of the early hours of 25th April 2006, Nogales, Arizona, give lie to this view where a General Atomics Predator B unmanned air vehicle (UAV) operating on behalf of the US Border and Customs Protection was lost due to the pilot failing to appreciate configuration differences of an alternate operator console.

The US National Transportation Safety Board describes the probable cause of this accident as *"... The pilot's failure to use checklist procedures when switching operational control from PPO-1 (pilot*

[2] "Motorway Plane Crash", Seconds From Disaster, Season 2 (2005), Episode 3, National Geographic Channel.

payload operator console) to PPO-2, which resulted in the fuel valve inadvertently being shut off and the subsequent total loss of engine power, and lack of a flight instructor in the GCS (ground control station), as required by the CBP's approval to allow the pilot to fly the Predator B. Factors associated with the accident were repeated and unresolved console lockups, inadequate maintenance procedures performed by the manufacturer, and the operator's inadequate surveillance of the UAS program." [82].

Aside from the obvious human error, and the repeated equipment malfunctions causing the pilot to switch consoles initially, the design of the Ground Control Station and the "Lost-Link" procedure are also significant factors. The aircraft control levers at each Pilot Payload Operator console – flaps, condition lever, throttle, and speed lever – appear identical, but may have very different functions. In piloting mode the condition lever controls the fuel valve, engine shutdown and propeller feathering. However, in consoles configured to be used as the observer station the same lever operates the camera iris. As soon as it was noticed that the aircraft was not maintaining altitude, although not knowing why and unsure as to what was happening, the pilot elected to shutdown the ground data terminal which has the effect of initiating the on-board pre-programmed lost-link procedure, which then causes the UAV to autonomously fly a pre-determined profile whilst waiting for the link to be re-established. Unfortunately, with the engine already shutdown the UAV continued to descend.

For each of these accidents it is possible to categorise the Human Factor, Environmental, Hardware and Software aspects, as summarised in Table 2.1. What is less easily characterised is the nature of the interactive complexity of each system as confronted by the operators on the day.

	737 G-OBME Kegworth, January 1989	Predator B UAV Nogales, April 2006
Human Factors & Training	Inappropriate procedure – Premature and incorrect diagnosis leading to incorrect shutdown of RH engine. Lack of communication with cabin crew.	Failure to use checklist when switching operational control between pilot operator stations leading to inadvertent fuel valve shutoff.
Environment	Smoke, fumes, vibration, extreme shudder, night-time, ATC distraction.	Crew located on the ground, flight instructor not co-located.
Hardware Failure	LH Fan Blade Failure.	Repeated and unresolved console lockups.
Software / System Design	Behaviour of auto-throttle affecting fuel flow. Inconspicuous vibration indication and alerting.	Differing control mode configurations depending on whether GCS is being used to pilot or observe.

Table 2.1 Categorising the causes of the Kegworth and Nogales accidents.

Both Stringfellow [104] and Johnson [59] see this incident as a socio-technical issue. Johnson in particular considers this as a problem of oversight in having one agency acting as both regulator and operator (the US Customs & Border Protection agency were operating under FAA delegated authority), along with a belief that there needs to be greater coordination between UAS operation staff and Air Traffic Management, and thereby improve mutual situation awareness. Johnson also suggests that minimum equipment lists and deviations guides be used as means to establish the conditions under which operations should be suspended or describe alternate safe operating procedures – which assumes that all plausible deviations might be evaluated.

Certainly, both of these sample accidents reveal underlying hardware reliability, flight crew training and performance issues. However, whilst allowing for the associated random hardware failure, each unfamiliar situation led to an inability of the respective flight crew to form a correct understanding of what was actually occurring – in short a failure to correctly diagnose the problem such as to inform a correct decision and course of action. In both cases expectations based upon an understanding, or lack of, aspects of the automation within the respective system misled the operator as to what the safe decision should be; from unrecognised behavioural interaction between engine autothrottle and fuel system, to an understanding of the capabilities of an autonomous lost-link procedure inhibited by operator action. Each was tightly coupled producing a prompt impact and interactively complex. Where "complex" reads as "unsafe" Perrow [87] provides a list of characteristics that can be used to distinguish complex systems from more straightforward linear type systems, as detailed in Table 2.2.

Complex Systems	Linear Systems
Proximity	Spatial Segregation
Common-mode Connections	Dedicated Connections
Interconnected Subsystems	Segregated Subsystems
Limited Substitutions	Easy Substitutions
Feedback Loops	Few Feedback Loops
Multiple & Interacting Controls	Single Purpose Segregated Controls
Indirect Information	Direct Information
Limited Understanding	Extensive Understanding

Table 2.2 Complex v Linear Systems after Perrow [87]

Perrow describes a system as being "interactively complex" where multiple discrete failures interact in unexpected ways, typically remaining incomprehensible for some critical period of time. Often this complexity arises as a consequence of the system designer rightly incorporating some redundant sub-system as a backup to preserve continuing safe operation. However, when a failure occurs, or is believed to have occurred, in some other aspect of the system, but which also as a side-effect interacts with this backup system, then often the whole system behaviour appears "mysterious" or "incomprehensible". When the effects propagating through the various system sub-components have a prompt and major impact on each other, and not slowly enough for human thinking, then the system is deemed to be "tightly coupled", and this increases the likelihood that operator intervention might actually make things worse. In particular when an operator makes a tentative choice (No. 2 engine in the case of G-OMBE) this also creates a mental model of the situation and if the results of this choice appear consistent with the mental model then later results that contradict this often do not cause rejection of the model, instead they only give rise to confusion. Systems with these properties are considered to be vulnerable to normal accidents – "normal" as in inevitable at some point.

Overall, Perrow considers that accidents are inevitable for systems of ever greater complexity, as defined by the degree of interaction and coupling. Therefore where the catastrophic potential outweighs the benefit, as depicted to the right in Figure 2.1, the use of these systems ought to be at least restricted, and, in Perrow's view, abandon nuclear power and weapons altogether. As to where an Unmanned Autonomous System might sit within this scheme, along with its associated Unmanned Air Vehicle, it would be reasonable to estimate that the degree of complexity lies somewhere between air traffic control, aircraft and space based systems, with the same catastrophic potential as an aircraft.

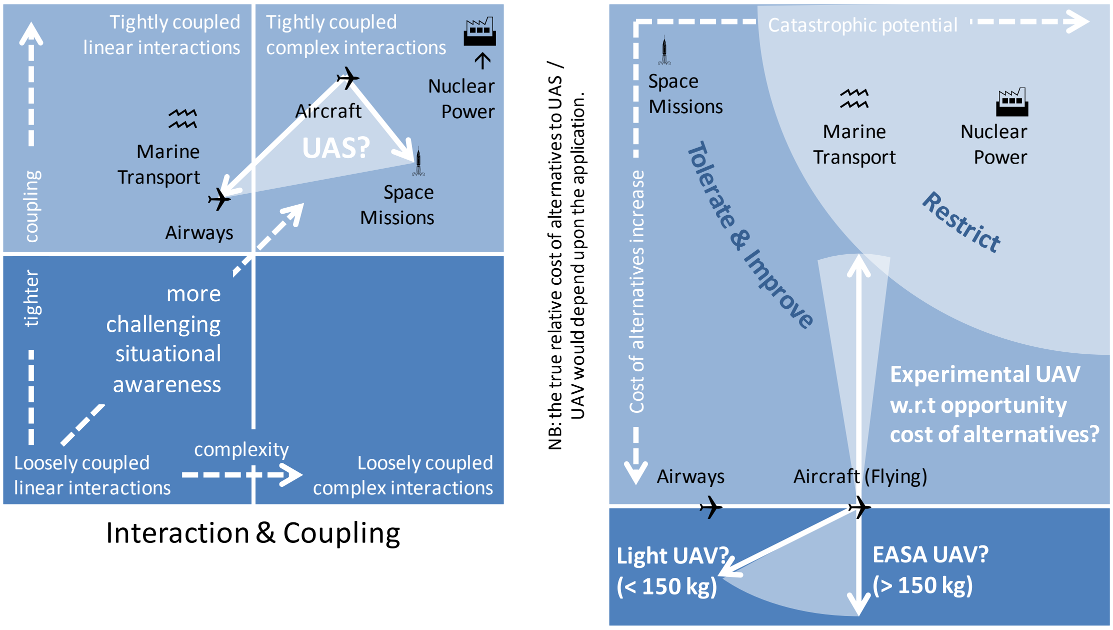


Figure 2.1 Living with UAS, adapted from Perrow "Living with High-Risk Systems" [87]

"To defend against normal accidents, we must understand the complex interactions of our programs, analyze close-calls and mishaps to determine root causes, and USE this knowledge to improve programs and operations." Dr. Michael A. Greenfield, Deputy Associate Administrator (1998), Office of Safety and Mission Assurance [50].

2.1.2. Human Factors

Defensive barriers are considered to be the best protection from the worst effects of the unsafe act. Humans, as operators, designers, managers and regulators are involved with all aspects of the safe operation of a system. This is no less true when the system is designed to run “autonomously”. When we look beyond the immediate causes of any accident involving a complex system, it is necessary to consider what barriers, if any, have been erected to prevent a proximate cause from creating the accident. In the main these barriers are centred upon human behaviour and socio-technical systems, and especially so the further the barrier is situated from the unsafe act.

In a study of Human Error, Reason [91] puts forward a general model for the dynamics of accident causation – more colloquially known as Reason’s “Swiss Cheese” model. In this he postulates that a “trajectory of accident opportunity” penetrates an alignment of gaps in the various defensive systems. This concept has since been applied to various socio-technical accidents, and in later [92] variants these general layers of defence hardened into four categories, as illustrated in Figure 2.2. For example in an airborne encounter between a UAV and another aircraft, the assumed hazardous scenario might be that a procedural violation occurs through the failure of a number of defences. Consequently, both aircraft then might violate the declared minimum separation, possibly due to a failure in air traffic management, where the dynamics of an unsafe act involving a final defensive barrier, say an emergency backup manoeuvre, might then be investigated in terms of decisions and delay.

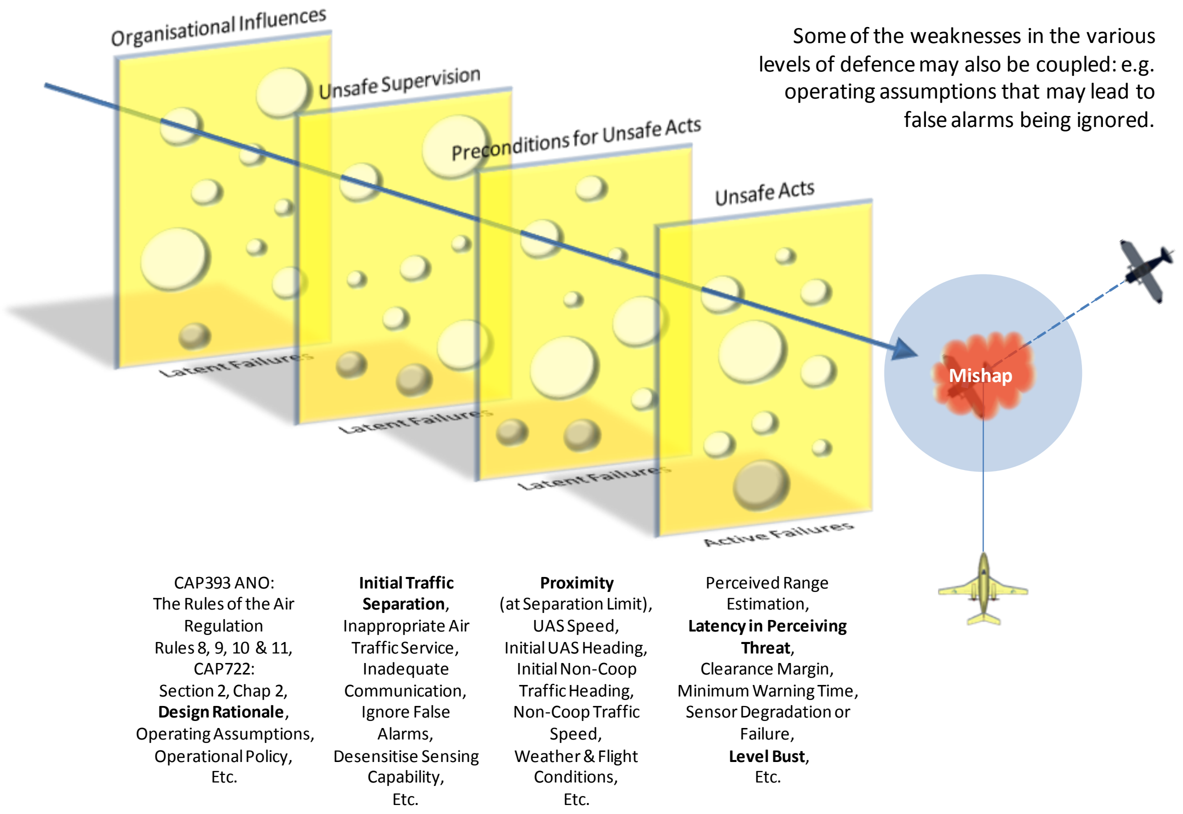

Organisational Influences
Unsafe Supervision
Preconditions for Unsafe Acts
Unsafe Acts
Some of the weaknesses in the various levels of defence may also be coupled: e.g. operating assumptions that may lead to false alarms being ignored.
Latent Failures
Latent Failures
Latent Failures
Active Failures
Mishap
CAP393 ANO: The Rules of the Air Regulation Rules 8, 9, 10 & 11, CAP722: Section 2, Chap 2, Design Rationale, Operating Assumptions, Operational Policy, Etc.
Initial Traffic Separation, Inappropriate Air Traffic Service, Inadequate Communication, Ignore False Alarms, Desensitise Sensing Capability, Etc.
Proximity (at Separation Limit), UAS Speed, Initial UAS Heading, Initial Non-Coop Traffic Heading, Non-Coop Traffic Speed, Weather & Flight Conditions, Etc.
Perceived Range Estimation, Latency in Perceiving Threat, Clearance Margin, Minimum Warning Time, Sensor Degradation or Failure, Level Bust, Etc.

Figure 2.2 Hazard Model adapted from Reason’s “Swiss Cheese” Model [91, 92]

2.1.3. Situational Awareness

Autonomy in systems, which is with autonomous machinery functioning as part of a system, is not treated as all or nothing where safety is concerned. As long as humans (and perhaps corporations) are the only legitimate and legal entities imbued with responsibilities for any decision affecting others, then machine autonomy ought to remain subordinate to human authority. Consequently, bearing in mind both Reason’s defensive barriers and issues of situational awareness, various degrees or levels of autonomy are now often used to define the relative scope and mediation of authority between the autonomous machine and the immediately responsible human agent – the operator. In the case of an autonomous UAV, a major concern is to what degree the operator, or remote pilot, then retains their own situational awareness of the immediate and anticipated behaviour of this machine for which they retain responsibility for the actions that it might take, and therefore the correct assignment of the appropriate level(s) of autonomy to safeguard this.

Endsley argues that situational awareness in the aviation domain can often be degraded by automation distancing the pilot / operator from the control loop and through *“lack of synergy in human and machine decision making”* [40], particularly in the effect of automation and application of artificial intelligence in cockpit systems. The argument for improving the operator situational awareness is therefore directed towards the design of, and assistance given by, situational information displays – as opposed to the full automation of an avoidance task. The particularly concern was where the automation has been designed deliberately so as to put the pilot or operator “out-of-the-loop”, with the possible effect that the operator's mental model of the system does not then enable them to recognise the actual state of the automated system when it fails. Passivity and loss of vigilance is of particular concern. It is suggested that automation introduces new forms of workload, and that this increases with the degree of automation. Therefore cockpit automation ought to be more focused on improving pilot or operator situational awareness rather than automating tasks – where this study concerned the visualisation of “pop-up” threats, demonstrating improved situational awareness of imminent threats although poor for projecting ahead of the threats.

Ruff, et al [98], envision a requirement for systems involving multiple semi-autonomous UAVs controlled by a single supervisor. However previous studies have shown that increased automation can cause sudden significant fluctuations in operator workload along with a loss of situational awareness, and Endsley and others suggest that this reduces the benefit in multiple levels of autonomy. In this workload study, involving automated planning and image selection, it was found that operators exhibit either limited confidence in an autonomous system, and / or prefer to do the work themselves – even if increasing their workload. Within the levels of autonomy available, either managed by operator consent or by exception, it was found that the operators frequently pre-empted autonomous decision making and automated actions were rare; albeit that these were non-safety related decisions.

From these studies it would appear that there remains some doubt as to whether the introduction of different levels of automation and autonomy does improve an operator's situational awareness. However, in the case of a UAS there will inevitably be situations where the pilot or operator will be “out-of-the-loop” to a greater or lesser extent during certain critical events. The challenge for hazard assessment of such a system is in constructing the best mental model encapsulating the possible degraded effect upon situational awareness.

Consequently, additional hazard assessment modelling goals potentially arise with these issues:

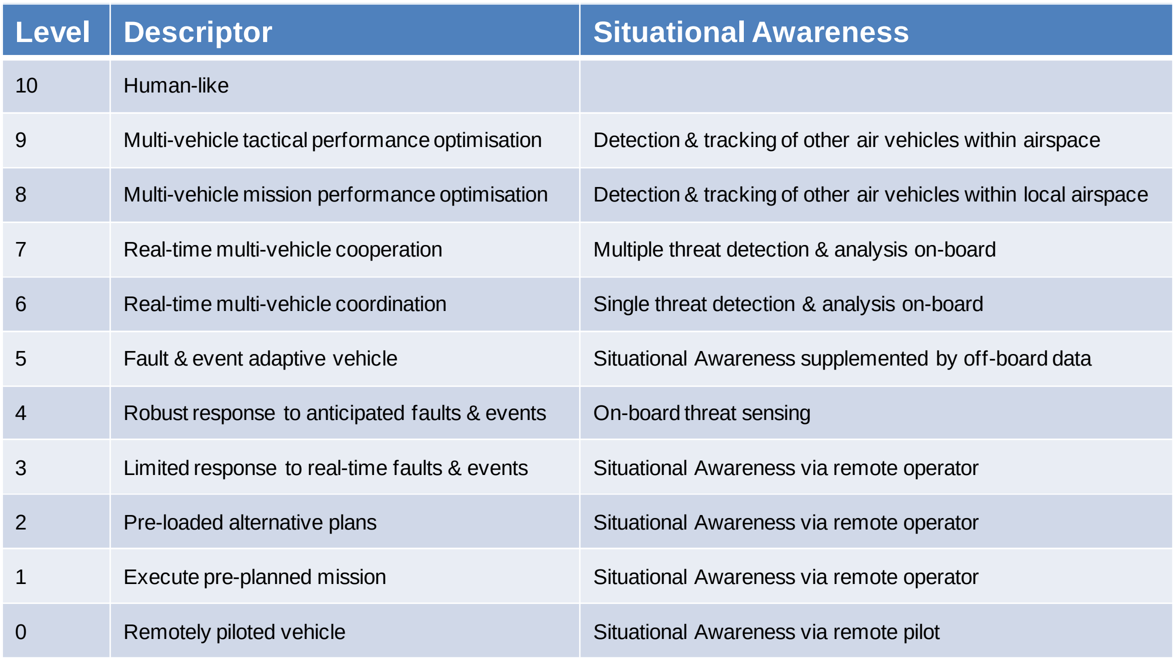
- There is a need to incorporate a formal representation of unsafe supervision or precondition leading to the unsafe act within this work, representing the final barrier to the unsafe act. This might be denoted in the role of the backup emergency avoidance behaviour within the scenario.
- Situational awareness might be modelled as a delay and / or confusion arising with the operator, pilot or controller response. This might possibly be portrayed simply as sensitivity to a variable latency of response and a probabilistic error of omission or commission, and a confusion matrix.
- A “consent or exception”, or other variable authority, model should be incorporated into the representation of the backup emergency avoidance behaviour. Currently “latency” is the only causal parameter in the model that might be treated as a crude proxy for delayed authority.

2.2. Defining Aspects & Constraints

2.2.1. Degree of Autonomy

Levels of autonomy are open to interpretation, and therefore where these to be accommodated in the system hazard assessment, for example representing a dynamic change of relative authority occurring within a scenario, it is expected that it would be necessary to declare which scheme is to be employed, and then also assign the appropriate range of levels to the various defensive barriers. Currently there are two reasonably well established schemes denoting variable levels of autonomy as used in the domain of piloted aircraft and unmanned autonomous airborne systems. These are Taylor’s, et al, Pilot Authority and Control of Task (PACT [106]) levels and Clough’s alternative Autonomous Control Levels [19].

Clough attempted to address the subjectivity inherent in the likely response to requests such as *“tell us how autonomous a UAV is ...”* [19]. In doing so ten Autonomous Control Levels (ACL) where defined, and adopted in their more tactical form from 2001 onwards in US Department of Defence UAV road-maps [25]. In this form these levels chart the progress of UAV development from simple remotely guided vehicles to fully autonomous swarms. For example, the DoD road-map charts a route from systems real time health diagnosis (L2), through UAVs that can adapt to failures and flight conditions (L3), and on to the development of an onboard route re-planning capability (L4).



Level	Descriptor	Situational Awareness
10	Human-like	
9	Multi-vehicle tactical performance optimisation	Detection & tracking of other air vehicles within airspace
8	Multi-vehicle mission performance optimisation	Detection & tracking of other air vehicles within local airspace
7	Real-time multi-vehicle cooperation	Multiple threat detection & analysis on-board
6	Real-time multi-vehicle coordination	Single threat detection & analysis on-board
5	Fault & event adaptive vehicle	Situational Awareness supplemented by off-board data
4	Robust response to anticipated faults & events	On-board threat sensing
3	Limited response to real-time faults & events	Situational Awareness via remote operator
2	Pre-loaded alternative plans	Situational Awareness via remote operator
1	Execute pre-planned mission	Situational Awareness via remote operator
0	Remotely piloted vehicle	Situational Awareness via remote pilot

Table 2.3 Original Autonomous Control Levels (ACL) [19]

Were these levels to be applied to a civil UAS and correspondingly to a civil certification route it may be more appropriate to adopt the initially proposed scheme, as abridged in Table 2.3. For example, the degree of devolved autonomy associated with a UAV Sense & Avoid sub-system might be located between Autonomous Control Level 4 – robust response to anticipated faults and events, and Level 6 - real-time multi-vehicle coordination. Level 8 and above might be deemed suitable to provide for operation in controlled airspace without continuous supervision, and arguably by extension, Level 9 and above would be required of an autonomous vehicle so to enable it to operate outside of controlled airspace (class G airspace) without supervision.

The UK Civil Aviation Authority's guidance in CAP 722 [17] also promotes the concept of a decomposed autonomous capability, so as to address the perceived inherent complexity of autonomous operations. True autonomy is described as a capability to operate without human control or direct oversight. It is suggested that an autonomous capability might be comprised of various decision-making sub-systems, each responsible for its own domain of authority. Within this definition there is then provision for a range of lower level capabilities, from simple advisory systems where humans make the decisions and execute the actions, through the intermediate levels of "semi-autonomous" systems. In this view an autonomous capability might be made up from autonomous flight management, sense & avoid, route management, power management and prognostic health management systems, for example. Some, or all, of these would collaborate with the human UAS pilot / operator. For the definition of these levels one might then elect to align these with the ACLs or the PACT levels, Table 2.4, as used within ASTRAEA [29], however, no specific guidance is provided as to how the partitioning and decomposition might be identified or managed.

Level	Operational Relationship	Computer Autonomy	Pilot Authority
5	Automatic	Full	Interrupt
4	Direct Support	Advise and action unless revoked	Revoke action
3	In Support	Advise and action if authorised	Accept advice and authorise (or not)
2	Advisory	Advise	Accept advice and act accordingly
1	At Call	Advise only if requested	Full
0	Under Command	None	Full

Table 2.4 Pilot Authorisation and Control of Task (PACT) Levels [106]

The vignettes to be described later in this work, in the context of a hazard perception questionnaire, would likely be assigned to PACT levels. Where a vignette follows the rules of the air the governing sub-system model will be operating at PACT 4 – that is committed to advised action unless revoked – but when responding with an emergency evasion manoeuvre this becomes PACT 5 – fully automatic.

2.2.2. Human Equivalence

Equivalence, or at the very least a measure thereof, appears to be the grail toward which the national regulator would appear to be working. However, what this implies in terms of system behaviour requires some further consideration. For example there is the “one in a hundred” argument, where Johnson, et al [60], express a view that the presence of a pilot “makes up the difference”. In this view the presence of the pilot typically allows for a 100 fold difference in erroneous detection rates when compared with a safety critical fully automatic high redundancy electronic control system. So where Fault Diagnosis and Isolation requirements are identified through reliability analysis methods, such as FMECA and Functional Hazard Analysis, and certification specifications CS-23 [37] and CS-25 [36] declare that false alarms should not be a “nuisance” to the crew, liable to distract them during critical phases of flight, and that they should not introduce new hazards to the system, this of itself does not help in the determination of an acceptable metric for the verification of a specific advisory function supporting real-time autonomous decision making.

Equating autonomy and ‘human equivalence’ is a cause for concern according to Alexander, et al [1], in their view of the appropriateness of UK military safety standards for the certification of autonomous systems, questioning the emphasis given to achieving human equivalence rather than optimum safety. Rather, the authors believe that scales used to categorise differing degrees of ‘autonomy’ might be more useful in determining an exact understanding of the degree of challenge and difficulty in certification. Consequently, the authors suggest that the ascribed level of autonomy might be used as an indicator of the difficulty that should be expected in certification. How this might be specifically measured and used in practice is still not described. There is a hint that specific hazard vignettes might be aligned to each level of autonomy, for a particular class of vehicle. However, whether the lower category levels truly describe autonomous behaviour is also questioned (for example ACL 0 to 3 – simple remotely piloted through to systems with a robust response to real time faults). It might be argued though that the higher levels of autonomous control do contribute to a better understanding of the aims of an enhanced hazard assessment as at this point they merge with goal-directed behaviour – a necessary consideration if one is to properly grasp the intent of the system design.

Alexander, et al. also consider the focus of the software engineering discipline as deficient in limiting itself to analysis of the software system without sufficient regard to the operating environment. Understanding the interrelationship of software systems to the extant environment is deemed as being essential to safety. It is proposed therefore that models should be built of the environment that an autonomous system is likely to encounter – including sensors, actuators, and phenomena – and that the correspondence of these models to the real environment should be evaluated. It is acknowledged that such an endeavour is a difficult task.

2.2.3. Safety & Reliability

Dissimilarity is a recurring theme in the guidance to the specification and design of fault tolerant systems embodying either functional or analytic redundancy. Beware though of interactive complexity – for example an aircraft might employ a secondary effect of control to mitigate a control loop failure, as such effects couple orthogonal axis of control it would very likely be dangerous then to make decisions on an independent diagnosis employing a different mental model.

From the perspective of the committee charged with developing ARP 4761 [100], Portwood [90], explains how the shift towards highly integrated systems, away from the more traditional federated architectures, has driven analyses now to be conducted at the aircraft or whole system level – which has even broader scope in the case of a UAS. Such a broadening of scope fits with Alexander's, et al. [1], view that hazard assessment ought to incorporate a holistic model of the UAS operating environment. While in the analysis of the possible safety consequences arising in the loss or reduction of system functionality, Saglimbene [102] suggests that the application of service experience, engineering and operational judgement is used to classify by severity and qualitatively describe the consequent hazard effect. These classified hazard effects are then tied to the relevant system safety objectives, thereby determining the maximum tolerable probability of hazard occurrence and tolerable level of risk. Glavaski, et al [48], suggests that as in the earlier example where some types of failure might be accommodated by reconfiguring and altering control allocation so as to employ dissimilar system or control surfaces in mitigation, these would incorporate means to maintain situational awareness of the revised limits of control authority.

The “black-box” nature of “off-the-shelf” software components renders their trustworthiness and modifiability as limited in the view of Ye and Kelly [115], and consequently the choice of mitigation strategies is therefore also restricted. The authors consider how this problem might be addressed through the use of a range of fault tolerant software techniques, and how failures in these might be represented and analysed. In the identification of likely failure modes it is suggested that deviation analysis might prompt consideration of all plausible deviations that may occur in the functions provided by any particular component. This would likely form a pessimistic view as some failures might not possibly occur due to the unknown internal implementation of the component.

2.2.4. Analytic Limitations

Explosions in the system state-space and complexity in the scientific sense pose fundamental challenges to the adoption of formal methods in the analysis of whole system dynamical behaviour – as seen in demonstrations of deterministic chaos in Poincaré's restricted three-body problem. Therefore, where Tribble and Miller [108] simulate the mode logic of a Flight Guidance System and a Vertical Navigation function using formal executable models, they recognise that the single variable for “altitude” alone results in the equivalent state space of more than 40 independent Boolean variables (10^{12}); with the altitude parameter truncated to the required 12-bit numerical accuracy. In passing, it is worthwhile to note Tribble and Miller’s interpretation of a lower level defensive barrier model as used in the safety assessment of this avionic software, as in Figure 2.3.

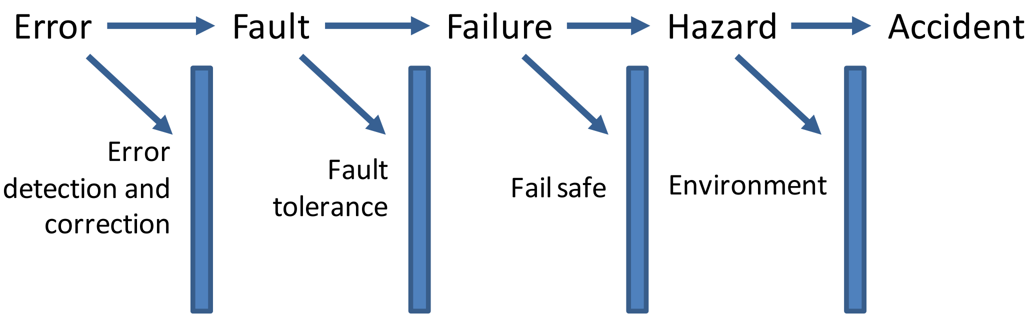


Figure 2.3 Accident Model & Defensive Barriers after Tribble [108]

In this software safety analysis process they employed Functional Hazard Assessment (FHA), Fault Tree Analysis (FTA), Failure Mode Effects Analysis (FMEA), model checking and theorem proving – but had not completed the analysis of the larger state-space of the vertical navigation function.

Similarly, Meenakshi, et al [75], describe a formal software verification methodology for an Autopilot Mode Transition Logic that is compatible with the objectives of DO-178B [95]. In this they evaluate an on-the-fly explicit state model checker, a symbolic model checker, and a combination of symbolic model checker, formal model simulation and a bounded model checker. From this study it is concluded that state space explosion is a major problem for explicit state model checkers. Symbolic model checkers need to create the whole state-space prior to verification, whilst bounded model checkers are used for falsification not verification. From this guidance given is that explicit model checkers are suitable for verification of state spaces no greater than 10^7 states, symbolic model checkers are effective up to 10^{90} states, while anything larger could be verified through bounded model checking (falsification).

In the case of Systems of Systems, Porter, et al [89], 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 – all of which would likely exceed the limits of a formal proof of safety. Therefore in capturing system dynamics for the hazard assessment the goal ought to be the falsification, rather than proof, of safe behaviour.

2.2.5. Regulatory Guidance

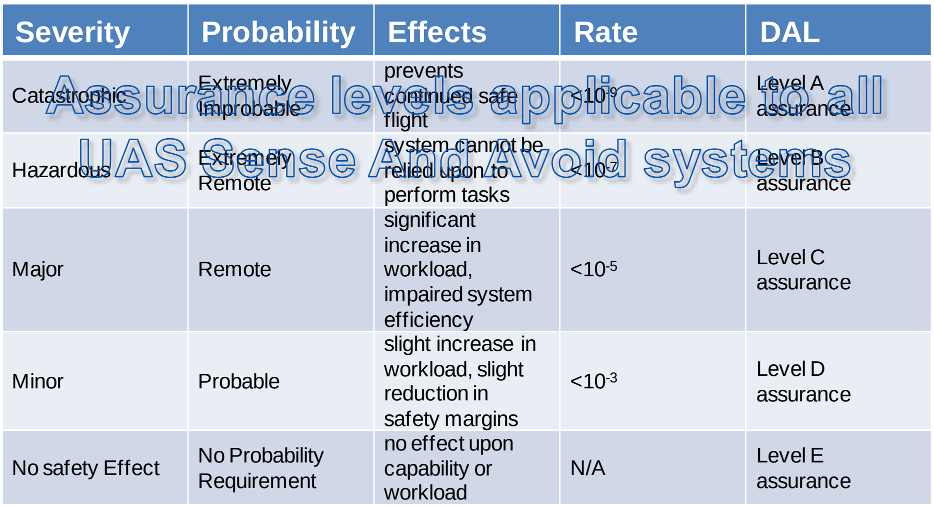
Failure categorisation and developmental assurance are at the heart of the current civil aircraft certification standards and regulatory guidance for airborne systems, although standards for military systems may allow the designer to adopt a more risk based approach to safety assurance. DO-178C [97] describes the recommended airborne software development life-cycle, whilst DO-254 [96] performs the same purpose for airborne electronic hardware.

In the certification of airborne software, DO-178B [95] directs that each potential failure must be categorised according to a Failure Condition. Further guidance as to applicability is also to be found in the source definitions arising within AC25-1309-1A [44] and AMC/AMJ 25-1309 [36]. Alongside failure rates, qualitative probability classifications may also be used where appropriate, defined thus:

1) Probable – likely to occur one or more times during the entire operational life of each aircraft;
2) Remote – unlikely to occur to each aircraft during its total life, but several times in the life of a number of aircraft;
3) Extremely Remote – not anticipated to occur to each aircraft during its total life, but a few times when considering the life of all aircraft of this type;
4) Extremely Improbable – so unlikely that they are not anticipated to occur during the entire operational life of all aircraft of this type.

With the advent of UAVs expected to be developed within a civil certification regime this guidance had been further adapted with draft AMC UAS.1309 [16], as detailed in Table 2.5, wherein a UAV sense and avoid sub-system would be treated as a specific “Certification Review Item” and subject to these definitions. Consequently, as the vignettes adopted in this work relate to the modelling of the behaviour of a putative UAV sense and avoid subsystem, then any aberrant behaviour of interest would be classed at best as Hazardous (a large reduction in safety margins or functional capabilities) to Catastrophic (failure conditions that could result in multiple fatalities).

This sub-system would be assigned a Development Assurance Level of A, or B if deemed to be only Hazardous. Adopting Alexander's [1] principle of associating this with the relevant level of autonomy, we might judge that a high level of difficulty is expected in certifying this system operating at ACL L6 and PACT 4/5, unsurprisingly. For this approach to be useful a matrix of autonomous level to development assurance level ought to be mapped for a range of representative UAS sub-systems.



Assurance levels applicable to all UAS Sense And Avoid systems

Severity	Probability	Effects	Rate	DAL
Catastrophic	Extremely Improbable	prevents continued safe flight	$<10^{-9}$	Level A assurance
Hazardous	Extremely Remote	system cannot be relied upon to perform tasks	$<10^{-7}$	Level B assurance
Major	Remote	significant increase in workload, impaired system efficiency	$<10^{-5}$	Level C assurance
Minor	Probable	slight increase in workload, slight reduction in safety margins	$<10^{-3}$	Level D assurance
No safety Effect	No Probability Requirement	no effect upon capability or workload	N/A	Level E assurance

Table 2.5 Probability and Severity of Failure Conditions, UAS 1309 [16]

In this scheme raising the required assurance level also raises the demands on the objectives for the various development processes. These processes include: software planning, software development, output verification of the requirements, output verification of the design, output verification of the software coding and integration, output testing of integration, verification of verification, configuration management, quality assurance, and certification liaison. However, in the period since DO-178B [95] was published, there have been dramatic changes in the breadth of application, software languages and tools, and a revolution in software methods themselves; such that this has recently been supplanted by DO-178C [97]. This now aims to provide more consistency, whilst accommodating separate parameter data items that change the behaviour of executable code without modifying it, addressing model-based development and verification techniques, supporting object-based technologies and supporting formal methods, plus a few others. However, Eastaughffe, et al [38], note that in DO-178B [95] design safety and human factors only get limited treatment, and that formal methods are not mandated.

It can be argued that an assurance based approach predicated upon simple determinism will not address the exploration of the safety issues inherent in the inevitable emergence of behaviour present at a higher level of system abstraction, inevitably encompassing complex behaviour. Rather emergence is currently rejected in the assurance of the higher integrity levels. Unfortunately, behavioural systems exhibit properties determined by the physics of an interaction and complexity within the wider system space, where surprising behaviours do emerge – authentic machine-based autonomous behaviour may ultimately need to encompass the characteristics of apparent cognizance (seemingly aware) and apperceptive behaviour (with seeming awareness relating to past experience).

2.2.6. Hazard Assessment

Whole system level developmental guidance, expectations and requirements are provided within ARP 4754A [101], where the aircraft system developer is given basic tools for use throughout the system integration process. However, the entire system scope of an autonomous air vehicle might be wider than that currently envisaged within ARP 4754A.

ARP 4754A [101], and its recently superseded predecessor ARP 4754 [99], describe the information to be produced by a largely prescribed functional hazard assessment process. However it is stressed that these are recommended practices and not regulatory requirements. Information is to be produced both for individual aircraft-level functions and combinations of aircraft-level function. This prescribes an iterative top-down approach applying the systems "V" integration life-cycle, from aircraft requirements through to aircraft verification, with system and item allocation, and the incremental incorporation of any additional mitigation functions as required. For example, in the case of a UAV these might include additional discrete autonomous system state detection and system reconfiguration functions. In general, the following information is to be produced:

a) Identification of related failure condition(s);
b) Identification of the effects of the failure condition(s);
c) Classification of each failure condition based on the identified effects and assignment of the necessary safety objectives;
d) Identification of the required system development assurance level;
e) A statement outlining the scope and assumptions made when evaluating each failure condition – operational or environmental conditions and phase of flight.

Also item members of a functional failure set comprising a given system function need not be assured to a single highest level of assurance, where parallel, dissimilar, multiple channel architectural mitigation provides protection from both random physical failures and design error anomalies. This does not provide a free choice of level of assurance, and guidance is given (ARP 4754A [101], p44, Table 3). At best the two most critical items leading to the top-level failure condition in any set might be assured at a level lower than the assurance requirement for the whole function. Unlike the earlier edition of these guidelines the term “functional independence” is now used to describe functions that might have previously described as “dissimilar”, and also a clearer distinction is now made between functional and item assurance.

The approach taken with regard to a preliminary system safety assessment, as defined in ARP 4754A, refers to a list of acceptable methods to be found in the companion document, ARP 4761 [100], which in turn describes Fault Tree Analysis (FTA), Dependence Diagrams (DD), and Markov Analysis (MA) as appropriate analytic methods, amongst others. The focus of the work here is to address novelty in the field of system hazard modelling and analysis, sufficient at least to aid the derivation of UAS advisory function performance requirements, therefore these guidelines and further developments in the field of hazard identification warrant close consideration.

In determining the requirements for an overall safe aircraft system design, the guidelines require that a Preliminary System Safety Assessment (PSSA) is first applied at the aircraft level. Arguably, in the case of an Unmanned Air System (UAS), this ought to be assessed at the whole system level incorporating the human operator, ground control station, autonomous airborne elements and interactions with others; along with the necessary means of conveying information and maintaining a path for authority and control. In all cases this preliminary assessment is required to produce a failure conditions list along with an identification of the corresponding safety requirements. A further use of the assessment is to demonstrate how the system will meet both qualitative and quantitative requirements for the identified hazards. Out of this preliminary assessment the protective strategies should then be identified, with account given to the fail-safe concepts and architectural attributes needed to meet the specified safety objectives. This includes the identification and capture of all derived system safety requirements: e.g. protective strategies such as partitioning, built-in-test, dissimilarity, monitoring, and safety related tasks and intervals, etc.

Consequently, the acceptable means of compliance for aircraft systems is predicated upon the relationship between probabilistic failure rates and severity classification of failure conditions, but in application to UAS this does not provide for the human aspects or issues of equivalence with human operation. Conversely, the perspective on Air Traffic Management (ATM) and Air Traffic Control (ATC) does recognise the human and organisational aspects of safety. For example, for EUROCONTROL [42] there is guidance on the harmonisation of risk assessment in this field, wherein also Reason’s [91] approach to accident causation is recognised by the French Service du Contrôle du Traffic Aerienne (SCTA).

In the UK, guidance on the conduct of hazard assessment for aerodrome operators and air traffic service, CAP 760 [14], contains guidance on the application and conduct of Hazard and Operability Studies (HAZOP), recognising the need to incorporate the human roles and procedures within the safety analysis. Given the obviously different relationship between the human and machine that exists for an unmanned autonomous vehicle, as compared with a conventional piloted vehicle, it would be worthwhile to consider the adoption of these safety assessment, hazard analysis forms and techniques as being of at least as much relevance to a situation where the role of the pilot is likely to migrate towards that of a remotely situated controller.

Evans and Kelly [43], view that the guidelines and methods in ARP 4761 [100] not only address the safety assessment and hazard analysis processes for avionic systems but also support the identification of all aircraft functional hazards as required by 00-56 [76, 77]. However, 00-56 [76] indicates an even broader scope than is covered by ARP 4754A [101] / 4761 [100], whereby a contractor is also required to ensure that the techniques selected are also suitable for identifying hazards and accidents arising from a wide range of potential causes. These wider causes also include any credible failures arising from normal and abnormal use, interactions between systems, procedures, managerial and human factors – amongst others. Consequently, it would appear to be quite appropriate, if not in fact necessary, to consider also the role of HAZOP, or something very much like it, and dynamical behavioural system analysis, alongside the application of FHA as described by ARP 4754A / 4761.

2.2.7. Function or Behaviour Abstraction?

Working solely within the current Functional Hazard Assessment framework is therefore unlikely to be sufficient in assessing hazards relating to the environmental interaction and high level goals of an unmanned autonomous airborne system. Wilkinson and Kelly [114], propose that a more pro-active approach should be adopted in the identification of potential hazards such that Hazard and Operability analysis (HAZOP) might be usefully combined with Functional Hazard Assessment when assessing airborne system safety.

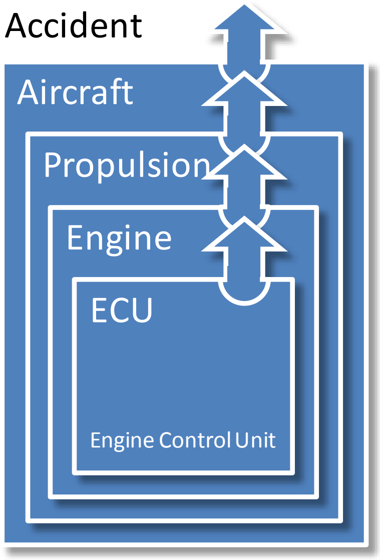


Figure 2.4 Propagation of Effect Through Sub-system Layers after Wilkinson [114]

It is proposed that where complexity and issues of integration with other systems exist, rather than analysis through a “chain-of-event” model, an analysis considering the effects of “deviations” from intended function or behaviour should apply.

A concern is how the effect of functional failure might be determined where an item is several layers removed from the environment, as depicted in Figure 2.4. Whilst FHA focuses upon the potential failure of system functions, HAZOP is more flexible in considering the propagation of down-stream effects within a mental model of a system incorporating system deviations, as well as functional failure. A necessary step towards the incorporation of HAZOP alongside FHA has been identified in the grouping of functions, wherein the different types of failure condition are associated with each failure type. Hence control functions might be associated with failures arising in deviations. Such an approach might be focused only upon critical combinations of functions and phases of flight so as to develop models of the “consequence chain” rather than the production of hazard checklists or an exhaustive FHA of computationally complex combinations of functions and phase.

Nevertheless, higher level abstraction of design and design process is slowly being accommodated. Jaffe, et al [58], reported on the progress towards proposed upgrades now embodied in DO-178C [97]. In this source code is just another, very detailed representation of a design, whereas DO-178B [95] cumbersomely and repetitively employed the phrase “design representation”. Therefore an overall design representation may be many tiered, in which each tier comprises either, or both, source code not requiring further design, or a representation of further captured requirements defining a further, lower, tier. This more abstract perspective ought to remove the repetitiveness in the use of the term “design representation” by making the representation iterative, where the components are elements of some machine-translatable language. This approach lends itself better to a concept of the modelling software and the modern software development process.

In the application of Object Oriented Technology to aviation system software, Rierson [93] considered the issues to be addressed for accommodation within DO-178B. One of the issues is to understand how Object behaviour might be modelled. In the case of C++ functions it is believed that extensive verification and testing will be required in order to understand their behaviour.

From a more conservative perspective, Haddon [52] outlines the safety assessment equivalent for a UAV in terms of consequence and impact kinetic energy, as now defined in AMC UAS 1309 [16]. In this it is recognised that any related remote ground station equipment must also be subject to the certification review and treated as a part of the aircraft. Haddon suggests that applying a safety code provides the aircraft designer with a target for the minimum acceptable standards applicable to all aspects of the design. Paragraph “1309” of the aircraft airworthiness codes [16, 36, 37, 44, 45] is an example of a basic safety target. In military systems other safety targets might be captured in a safety case. However, Haddon states that *“... current codes of civil airworthiness requirements do not* *permit certification on the basis of such safety assessments alone"* as *"... with the safety case approach a complete reassessment of the aircraft and its operating environment may be required for every change of role."* In so far as autonomous behaviour is discussed it is apparent that the emphasis is on system reliability rather than intended system behaviour, although it is noted that *"... the remoteness of the pilot / controller of a UAV raises major issues for aircraft operations in terms of air traffic management, compliance with the Rules of the Air etc, ..."* Therefore, it is likely that in embedding the Rules of the Air within the system, for which there is a strict interpretation for collision avoidance and rights of way, then on the principle of "no hazard, no benefit", alternative and unconventional avoidance strategies may be permitted, but would not provide a justification for accepting non-compliance with the applicable airworthiness requirement. Also, such alternative strategies would be required to demonstrate that no additional hazards were introduced.

Under EASA regulations aircraft less than 150 kg are subject to national aviation authorities [53]. Civil light UAV systems may be exempt from the principles of CAP 722 [17] if it can be demonstrated that the system poses no worse safety risk than that equivalent to existing model aircraft. As such this places a restriction upon the operation of these vehicles so that they may only be operated on a direct line of sight no further than 500 metres from the operator. It is unlikely that a smaller UAV operated in this manner would embody any significant degree of autonomy – consistent with Clough's ACL L0 [19].

2.3. Summary Considerations

Taking the foregoing discussion in this section as the basis for further "exam questions" to be considered in this work:

- As safety validation and assurance tools are also to be qualified to assurance criteria commensurate with the system model that they are to assure, arguably a hazard model used for this purpose might also be subjected to an equivalent degree of verification – and in some way fit with a model representation compatible with DO-178C.
- This work should consider to what extent the required expertise in HAZOP might be extended with behavioural models and simulation where complex system behaviour results. Part of this is to consider where exactly these models might be placed with respect to the experts, and what role might they play in either bolstering trust, in the portrayal of likely system behaviour, or alternatively prompting further investigation where scepticism exists.

Chapter 3. Hazard Identification and Analysis Review

"I'll be more enthusiastic about encouraging thinking outside the box when there's evidence of any thinking going on inside it."

Terry Pratchett

3.1. Extending HAZOP

3.1.1. Hazard emulation, Identification & Assessment

Hazard and operability study (HAZOP) [13, 66] and hazard identification (HAZID) are established techniques arising originally with the chemical and process industries, and more recently also optionally applied to the development of safety cases in Air Traffic Management.

For process industries computer based qualitative reasoning models supporting the HAZOP process have been developed, by in effect providing "HAZOP emulation" to the HAZOP team; for example the STOPHAZ and HAZID codes [72]. 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 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, with the software then capturing all of the permitted propagations of cause and effect between components, both down and up-stream. In the nature of the problem space of these models, a typical HAZID model is constructed from a model library within a fluid modelling system, representing feasible flow bounded within qualitative constraints within which the propagation of deviations is modelled. Innovative features might then extend this to include asymmetric influences, detection of configuration problems, fluid compatibility checks, classification of consequence types, etc. [73]

Therefore, propagation methods of qualitative flow deviation and effect have been addressed to a greater extent, at least for static and quasi-static flow models, so further consideration ought to be given to the encapsulation and propagation of behaviour of essentially dynamic systems. The process plant models as described above are treated as closed systems, operating at particular set-points. However, autonomous and robotic systems with goal seeking behaviours aiming to maintain a state of dynamic equilibrium with respect to their environment cannot be safely treated as closed systems.

3.1.2. Bayesian Expressions of Risk

Bayesian forms for the expression of risks and hazards are described by Hu, et al [56], wherein a Bayesian network models risk relationships learnt from statistics of ship navigation accident data. Once trained, this network is used for the identification of hazards in the navigation of ships. A network structure is derived from ship accident data, producing a specific network model from a naïve Bayes model with learning facilitated by the accident data. This incorporates sets of defined conjugate evidence to form the network nodes.

Derived from Bayes' original work on the probability distribution of an unobserved variable [6], a Bayesian inference gives the posterior probability of hypothesis *(θ)* being true where evidence *(a)* is observed as the product of the posterior probability that evidence *(a)* is observed given hypothesis *(θ)* true, and the prior probability of the hypothesis *(θ)*, divided by the model evidence for *(a)* independent of the hypothesis, which is therefore constant for all hypotheses – equation 3.1 below. The probability term *P(a|θ)* is also known as the "likelihood" value of *(a)* given hypothesis *(θ)*.

$$\boldsymbol{P(\theta|a) = \frac{P(a|\theta)\cdot P(\theta)}{P(a)}} \qquad 3.1$$

This likelihood value is calculated from the joint probability distributions of parameter *(a_j)* where *(θ)* true. In Hu's, et al., model this represents the joint probability of a series of *(m)* hazard attributes *(a_j)* where hazard *(θ)* is true, each hypothesis being represented by a node in the network – equation 3.2.

$$\boldsymbol{L(\theta|a) = \prod_{j=1}^{m} f\left(a_j\middle|\theta\right)} \qquad 3.2$$

Therefore this likelihood function is used within the Bayesian interpretation to update the posterior probabilities of the hazard hypothesis – equation 3.3. This is the basis for an update function to train the naïve Bayes model and thereby derive the resultant Bayesian network. This likelihood function is also used with the resultant network to facilitate a Bayesian inference model for quantitative risk assessment, where is calculated as Risk = Likelihood × Impact – equation 3.4.

$$\boldsymbol{P(\theta|a) = \alpha\cdot\prod_{j=1}^{m} f\left(a_j\middle|\theta\right)\cdot P(\theta)} \qquad 3.3$$

$$\boldsymbol{R_k(x,y,t) = \widehat{\theta}_{MD}\cdot P_k(\theta|a)\cdot\sum_{i=1}^{n}\omega_i C_i(x,y,t)} \qquad 3.4$$

In calculating the severity of an accident [56], equation 3.4 gives the risk for a particular hazard (*k*) as a product of the mean accident rate, the probability that the hazardous event (*k*) occurs due to the associated hazard attributes (a_j), within the joint probability of the (*m*) hazards, and the sum of causal relationships for this event for the (*n*) types of accident cause class (C_i) weighted by an impact factor (ω_i) for each accident cause class.

Trucco [110] reports a similar study linking a Bayesian Belief Network representing Human and Organisational Factors in accidents at sea to importance measures and basic events in a fault tree representation of maritime accident causation. Bayesian methods applied to system safety assessment are also presented in a proposed methodology for the safety assessment of socio-technical systems in the form of a "hybrid" approach by Mohaghegh, et al [79]. This illustrates how a variety of deterministic, qualitative, quantitative and statistical techniques might be integrated within a hybrid Probability Risk Assessment – Figure 3.1. The techniques include Systems Dynamics modelling, Qualitative-Quantitative inference models in the form of Bayesian Belief Networks, and Fault Tree models represented as Binary Decision Diagrams. An example is developed consisting of the impact of airline maintenance operations on aviation safety risk consisting of system dynamics modules representing management commitment, financial pressure, training, hiring, human reliability, and technicians' commitment. In the example the systems dynamics model provides a technician error probability to a Bayesian belief network representing technical activities within aircraft maintenance. The Bayesian belief network then in turn generates beliefs about likely failures due to technician error to a fault tree or binary decision diagram representation of the system. Top event probabilities are then passed to an event sequence diagram, returning an evaluation of risk. The resultant risk values then are used to drive the relevant system dynamics models – and hence close the modelling loop.

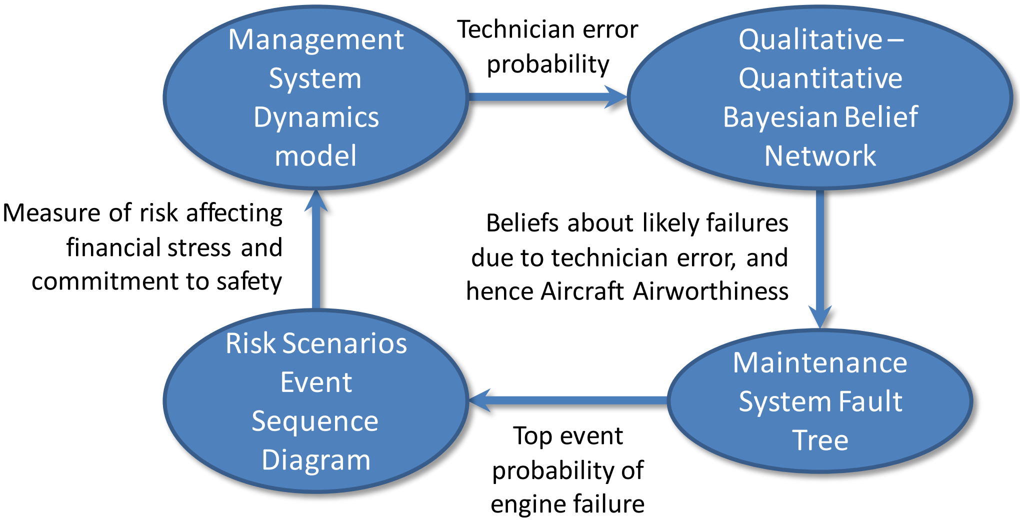

Management System Dynamics model
Technician error probability
Qualitative – Quantitative Bayesian Belief Network
Measure of risk affecting financial stress and commitment to safety
Beliefs about likely failures due to technician error, and hence Aircraft Airworthiness
Risk Scenarios Event Sequence Diagram
Maintenance System Fault Tree
Top event probability of engine failure

Figure 3.1 A Hybrid PRA Framework for Socio-technical Risks [79]

For this hybrid approach further work remained with "... future research to extend the model to include other aspects of operations (e.g. flight crew model), and to analyze the common effects of organizational factors on both the operation and maintenance of an airline." [79]

3.1.3. Scenarios & Deviation Analysis

Scenarios typically form the basis of deviation analyses such as HAZOP, wherein guidewords are applied to selected parameters, and through which the propagation of effects are considered by a team of system experts, preferably including operators, and possibly supported by scenario emulation.

Allenby and Kelly [2], note that the flow based technique of HAZOP and the function based approach of FHA are complementary in that both consider deviations from “design intent”, and these might be combined in a hazard identification approach *“... using a suitably expressive requirements representation.”* [2] For Allenby and Kelly, the articulation of these safety requirements take the form of a Use Case diagram, a scenario, or perhaps a vignette within a larger scenario, a definition of any preconditions, the intended system response, and the expected post-conditions. Such an expressive requirements representation, capturing the intended system behaviour, might be adequately described as a SysML model [83]. This modelling language facilitates the capturing of functional decomposition, information and event flow behaviour, wherein models might be used to describe both the extant system and the hazards considered to arise within it. Allenby and Kelly also note that safety analysis techniques such as FMEA assess the effects of known behaviour. Conversely, deviation based techniques can be employed with more abstract models of system behaviour, making them appropriate for use early in the system development lifecycle, whereas FMEA requires the existence of a more specific and detailed system design.

This suggests that a first-cut hazard or accident model ought to facilitate the modelling of a system of interest’s desired behaviour whilst also capturing the effect of deviations from declared intent and emergent behaviour. Consequently, it is argued that a behavioural modelling capability is required, and the highest level of abstraction for this suggests the development of a meta-model facilitating the exploration of potential hazards at the earliest possible point within the system lifecycle.

Despotou and Kelly [22], consider the extension of such analyses addressing system dependability – security, trustworthiness, etc., with a view to the systematic definition of a specific set of guidewords augmenting the list of hazards. Typically these hazards would be identified within a FHA, which together with additional guidewords might identify the totality of the system failure conditions. For example the failure categories “loss of function”, “function provided when not required”, “incorrect operation of function”, arising from a FHA, should to be taken as guidewords incorporated into a HAZOP. This approach needs to be specific with regard to the guidewords used and with which models these are to be associated. Examples appear to show that the application of a guideword on a different element of a model can reveal different concerns. Consequently, an additional proposal is that the projected effects of a deviation should not be restricted to a single dimension focused on a single attribute; rather the wider consequences ought to be assessed to gain an understanding as to how deviations affect the operation of the system as a whole. The same principle should be applied during identification of causes. Hence in the context of operation, a security failure condition such as “denial of service” (a security attribute) may reveal causes in failures regarding other attributes such as “limited processing rate of requests” (a performance attribute). In order to identify hazards that are difficult to anticipate, and thereby identify the appropriate system interactions to be modelled, it is likely that a combination of formal closed modelling and somewhat more exploratory open modelling be considered together.

The wider realm of autonomous systems is perhaps the most complex for safety assessment, where the decomposed and devolved nature of any autonomous system should perhaps be considered to be either within, or in its entirety, a system of systems. In order to obtain an understanding as to how possible deviations might cause accidents in System-of-Systems (SoS) Despotou, et al [23] suggest an alternative to the manual and mental reductive processing techniques in the determination of hazards – as is the established HAZOP method. In a SoS the logical steps, between a hazard and its causes may be complex and range over a significant distance in time and space. The proposed alternative combines these with, or derives these from, simulation-based approaches. In the case of a SoS several different vignettes may be defined – varying the mission, environment, opposition (conflicts and incursions perhaps in the civil sense) and other parameters. In the proposed approach a number of agents are employed and together these address a number of possible defined deviations, with a particular vignette being run repeatedly with different combinations of these. In this scheme whenever simulated accidents, or near-misses, occur these are to be logged with respect to the specific deviation and hence then direct further simulation within this region so as to study the particular case.

Various aspects of the foregoing suggested approaches for the extension of FHA into HAZOP, along with modelling and simulation, have been considered for incorporation into the scheme developed in this work. Specifically in the hazard vignette, guidewords have been applied to an abstracted behavioural attribute – namely that described by the abstract noun ‘manoeuvre’. In addition, the concern in the vignette is the violation of a separation constraint, within which the sensitivity to nearer misses can then be explored with Monte Carlo simulations.

3.2. Systems Theory beyond Normal Accidents

System dynamics modelling using computers has been around almost since the dawn of digital computing itself, where from the 1950s to the early 1960s Jay Forrester and a team of graduate students at MIT took system dynamics modelling from hand-simulation to formal computer modelling. Forrester [46] developed system dynamics modelling with the intention of studying industrial system behaviour, and how the stability of these systems are influenced by policies, structures and delays. Subsequently Phyllis Fox and Alexander Pugh wrote the first version of DYNAMO (DYNAmic MOdels), a system dynamics computer modelling language that became a standard until the 1980’s.

In their *Introduction to Simulation* in 1983 Roberts, Andersen, et al. [94] positioned DYNAMO as an educational tool and in a foreword by Forrester as *“... a door leading to understanding of change.”* In this book examples are used that demonstrate how things are interrelated and how the states of these things change as a consequence of the dynamics within these relationships. These are used to illustrate where political processes effect change, and how rising and falling economic activity occur with physical and social behaviour. Whilst DYNAMO exists now as a relic of an era alongside FORTRAN, more contemporary tools such as Vensim™, and others, have since displaced it.

Taking inspiration from System Dynamics, and attempting to articulate the limitations of "chain of event" models of accident causation where tight coupling and interactive complexity appear as safety factors, Leveson [67] argues for the application of a systems approach to system safety analysis, especially applied to safety in socio-technical systems wherein such models are rooted. While Normal Accident [87] theory and High Reliability Organisation theory have focused attention on the important issues of complexity and coupling, Leveson, et al. [68] view that they limit progress by defining the problem too narrowly, whether through excessive pessimism or optimism, and in confusing reliability with safety. Engineers and sociologists should work more closely together with shared definitions and assumptions, applying a systems approach at the organizational and cultural level demonstrating how tradeoffs among safety, performance, schedule, and budget can be evaluated.

3.2.1. Systems Theoretic Accident Model Process

Theoretic models of systems accident processes are currently attracting a degree of interest from a number of quarters. The claim is that system safety constraint violation and system dynamics should be the foundation of system safety analysis, potentially encompassing an entire socio-technical system from the role of the state regulator to the final defensive barrier and the unsafe act. Systems-Theoretic Accident Model and Processes (STAMP) [67, 70] hypothesizes that risk events arise from unanticipated interactions among components, whether failed or working as designed, violating safety constraints within systems characterised by interactive complexity, and that safety culture can be modelled, formally analysed, and engineered.

Leveson [71] argues that only systems exhibiting Organised Simplicity are amenable analytic reduction, where it is feasible to decompose a system into components that operate independently, and system behaviour is governed only by discrete events over time. This holds where behaviour of individual components is the same whether viewed singularly or as a part of the whole – ruling out systems with feedback loops and non-linear interactions. Systems that have no known underlying structure, treated as aggregates, and therefore not amenable to reduction, are said to exhibit Unorganised Complexity. Between these two there is Organised Complexity, less than Unorganised Complexity in terms of randomness of behaviour, and more than Organised Simplicity on any measure of complexity. This type of system is found to be too organised for statistical analysis, without consideration of the underlying structure, and too complex for complete analysis or reduction. Leveson sees systems exhibiting Organised Complexity as the most challenging to system safety.

Leveson [67, 70, 71] considered how the system theoretic approach aids understanding in an accident scenario where the third US Milstar satellite was placed in an unusable low orbit due to a complex interaction of mistakes and decisions. It was judged that this accident occurred within a flawed development and operational process not amenable to the usual explanation as being caused by a simple chain of events. Leveson argues for a generic model of socio-technical control within which hierarchical development and operations control structures, and specific system control loops might be modelled. The argument is that the simple chain-of-event accident model is inadequate to deal with such complexity as it cannot account for indirect and nonlinear relationships, rather the entire safety control structure and accident process determine the role each part plays, and that at a given level of complexity certain emergent properties are irreducible. Also it is observed that this approach does rather more to prevent future accidents than merely attributing blame at the point of failure.

Leveson [67, 71] notes that the systems of interest are typically Open Systems, maintained in a state of dynamic equilibrium through closed loop feedback control. Therefore, we might judge that alongside the other recommended "chain-of-event" methods provided by ARP 4761 [100] Markov Analysis is insufficient to predict complex forms of accident causation, as Markov models represent Closed Systems wherein equilibrium is effectively synonymous with Invariant or Stationary; when the probability of a subsequent system state is then identical to that of the preceding state.

Considering the NASA Space Shuttle Columbia accident, Duclac, et al [35], illustrate how static control structures and dynamic behavioural models might be used together within the STAMP process. A 200-run Monte Carlo sensitivity analysis was performed with a systems dynamics model representing an indicator of effectiveness and credibility of the Independent Technical Authority (ITA), and effect on the System Technical Risk; representing sensitivity to quantifiable variables such as risk, resource allocation, knowledge, shuttle aging, launch rate, safety efficacy, learning, and perceived success over a 1000 month period. The resultant dynamic behaviours revealed the possibility of two modes relating to the credibility of the ITA and technical risk; one largely successful due the implementation of the recommendations of the ITA, although declining slowly later due to complacency; and an alternative mode starting with insufficient resources such that credibility rapidly declines towards a tipping point thereafter poor behaviours are reinforced and the system enters a permanent high risk state. However, it is suggested that the creation and use of the resultant system dynamics models requires a high degree of effort and substantial domain expertise.

STAMP has also been applied to the evaluation of security risk. Laracy & Leveson [65] consider the application of STAMP to the safety analysis and protection of critical infrastructure. In this they suggest that Probabilistic Risk Assessment (PRA) is unsatisfactory as the rate of intent and capability on the part of the attacker is unknown and not measurable – the likelihood of rare events cannot be accurately estimated, and especially where there is hostile intent. Furthermore it is irrational to assume that attacks will not be simultaneous, and therefore an attack upon security cannot be construed as being equivalent to a single point of failure. The resultant security model considers incidents as a result of inadequate control, rather than as a system component failure event, requiring the enforcement of suitable constraints to shape the emergent properties providing security.

Any independent entity may potentially threaten deleterious effects without hostile intent within a system of systems, for example in the reduction of safe separation due to the incursive behaviour of a ‘non-cooperative’ vehicle in a ‘sense and avoid’ scenario. For an accident to occur there need not be hostile intent in failing to observe or correctly respond to the behaviour of the controlled entity within the system of interest. In this work it is proposed that an enhanced form of HAZOP might be created in a fusion with STAMP – coupling Bayesian representations with Monte Carlo data arising from the dynamics of system interactions. It is argued that inferences drawn from this statistical and observed behaviour might usefully validate assumptions in the mental model used during the HAZOP process.

The STAMP framework [67, 70], and associated system dynamics, may potentially span interactions between all entities, from legislature to those operating at point of the unsafe act, thereby accommodating defence in depth as described by Reason [91, 92]. In simulation however, Leveson [71] suggests wariness of the role it plays in systems safety, stating that *“... care must be taken that any simulation or other planning tools to assist human problem solving do not rest on the same incorrect assumptions about the system that led to the problems in the first place.”* Nevertheless, SpecTRM-RL [69, 71, 86], a specification and requirements tool used to formulate black-box functional descriptions of system components, does incorporate a system simulation environment – although no examples of sensitivity simulation results revealing safety constraint violation appear in *Engineering a Safer World* [71].

3.2.2. Systems Theoretic Process Analysis

Process analyses based upon the systems theoretic approach are now being developed to fulfil much the same role as HAZOP. Owens, et al [86], and Stringfellow, et al [105], describe a hazard analysis methodology that encompasses not only component failure but also requirements and design errors. Systems Theoretic Process Analysis (or sometimes STAMP-based hazard Analysis – STPA) is the basis for a new hazard identification technique, wherein instances of inadequate control and the necessary safety constraints are identified, and how these constraints may be violated are represented. STPA is intended to be undertaken from the beginning of any system design process, and used to facilitate safety related design trade-off within a design space incorporating information regarding possible inadequate control actions and control flaws across a range of design options. It is posited that unlike methods such as HAZOP, this new hazard analysis method does not require a pre-existing design description – rather it is formed around a system Intent Specification. The concept of an Intent

Specification framework, as in Table 3.1, was introduced by Leveson [69] in developing the socio-technical viewpoint as a central tenet in the integration of safety modelling into systems engineering.

Level	Descriptor	Features
0	Project management plans, safety plans, etc.	
1	System Purpose	Assumptions, constraints, responsibilities, goals, hazard analysis, etc.
2	System Design	External interfaces, tasks, controls, functional decomposition, validation, etc.
3	Black-box models	Environment model, task / HCI models, functional models, analysis, etc.
4	Design Representation	HCI design, Hardware & Software specifications, test plans & results, etc.
5	Physical Representation	GUI & physical control designs, software code & hardware assembly, test plans, etc.
6	Operations	Audit procedures, operator manuals, error reports, change requests, monitoring, etc.

Table 3.1 STPA System Intent Specification Levels (simplified) [69, 86]

It is argued that accidents occur when safety controls fail to enforce safety constraints, resulting in unhandled disturbances, uncontrolled failures, or unsafe interactions, and that the potential existence of these defects might be indentified during the system design and options selection phase; wherein architectural and behavioural requirements are formed and decomposed into lower level purposes. Furthermore, within this safety-driven design approach it is suggested that the regulatory, design, and operational controls might also be modelled in order to identify any migration towards states of higher risk due to external economic pressures, along with any safety issues in potential lack of cooperation. For example a "sense and avoid" scenario, modelling UAS systems dynamics, might also model the degree of cooperation between the two controllers where multiple controllers are present.

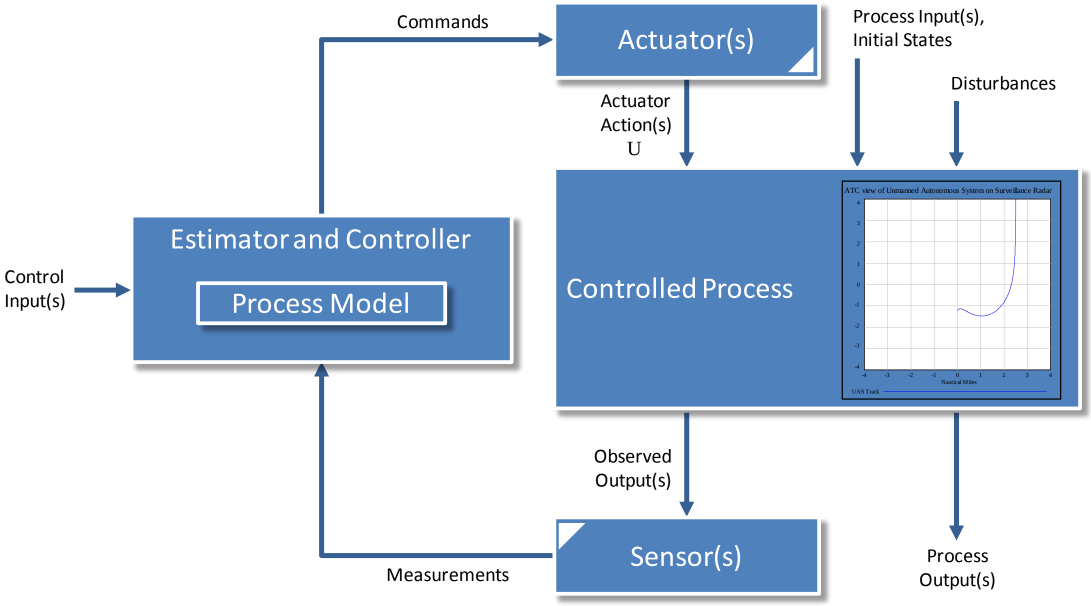


Figure 3.2 Generic STPA Low-Level Process Control Loop after Stringfellow [104, 105]

A safety-driven design process identifies potential accidents and the hazards that initiate them. In the case of STPA this aims to identify the safety constraints, design decisions, inadequate control actions, control flaws, and inadequate control executions. These are identified by adopting a generic Process Control Loop model, as in Figure 3.2, wherein a dynamic feedback process maintaining separation with respect to a safety constraint is represented by a closed-loop system.

In cases where this implies physical separation within a dynamic process, for example collision avoidance, such loops are readily identifiable. Where the control process and nature of the safety constraint separation is more abstract, further work may be required to identify this loop. Having identified the safety parameter subject to deviation within this loop, usually a Command parameter, specific deviations known as Inadequate Control Actions are identified – for example a command to "manoeuvre" if deviated would affect the output in terms of physical separation. Potential Control Flaws are identified for each block within the control loop, working around the loop for each deviation in turn, and these flaws are associated with the relevant deviation as potential causes. These help identify safety constraints, and the process is applied iteratively to any inner loops. These descriptors may also be used in a Model Based Systems Engineering (MBSE) formulation of a SysML representation of an Intent Specification.

3.3. Systems Modelling

3.3.1. Systems Representation

Representations of systems take many forms. Plain text, tables, schematics, block diagrams, assembly drawings, illustrations, models, and structural diagrams in a variety of forms are all used to represent the many facets of a real system. Formal modelling languages provide a particular class of system representations with the intent of ensuring unambiguous interpretation and the possibility of manipulation and transformation for use by machines in automatic composition, analysis and documentation – all which is useful in maintaining accuracy and consistency whilst aiding system safety assessment A formal language that encompasses system requirements, constraint and parameter modelling, amongst others, and continues to be developed in application to Systems Engineering is SysML [55, 83], a development of the Unified Modelling Language (UML) [84, 85], and accommodating descriptions of certain system viewpoints defined within DoDAF / MODAF [26].

In the mid-1990s Friendenthal & Lykins [47] discussed parameter-based representations, concerned with the apparent inadequacy and inconsistency in the then current systems engineering methods' ability to encompass the parameterised behaviour of a system and capture the logical partitioning of multiple system attributes. It was viewed that this required an encompassing representation of both the system of interest with its environment, each captured as models of processes; having inputs, outputs and mechanism, and each characterised by parameters. Subsequently, Cantor, et al [18] argued the case for developing the Unified Modelling Language, extending its use from the software domain into the systems domain. It was argued that a need exists to "bridge the semantic gap" between software and the related systems engineering disciplines, and that this was also to capture system characteristics such as safety and reliability, and parametric relationships – as well as facilitating common standards for the exchange of information between tools. All essential system attributes of hierarchical behaviour, continuous time, inputs and outputs, physical structure, performance and physical characteristics were to be included. The result of this is the systems modelling language (SysML).

Therefore, the adequacy of a representation capturing both intended and dysfunctional system behaviour, described using SysML [83], should be considered as framework within which to cast a safety assessment. The SysML system modelling language facilitates the capture of functional decomposition, information and event flow behaviour. Models in this form might be used to describe both the extant system (model), and the location and causal nature and likely effect of hazards considered to arise within it. The system Intent Specification framework as described by Leveson [69], Owens, et al [86], and mentioned by Stringfellow, et al [105], is presented as a table, as in Table 3.1, spanning the environment, operator, system & components, and validation & verification intent. Each type of entity within this table may be assigned a particular sub-set of SysML diagram types, effectively defining an intent framework in a form similar to DoDAF, an architectural framework.

3.3.2. Representing Risks & Hazards

Model Based Systems Engineering (MBSE) is an approach to systems engineering employing formal modelling techniques that encompass system requirements, design, analysis, validation and verification. Increasingly the aerospace and defence industry recognises that "something is missing" [112] in the current systems engineering processes, something that fails to properly take account of interactive and irreducible complexity, something required also to capture likely system side-effects within system-modelling tools, where it is suggested that this is fundamentally a modelling problem.

Hazard analysis per se rarely appears to be the subject of a formal MBSE approach. Soyler and Sala-Diakanda [103] describe the beginning of a proposed SysML framework for defining and capturing disaster management systems. The challenge in modelling disaster management systems is to quantify the impact, model evacuation, damage estimations and the state of preparedness, with a view to establishing new safety standards. It is anticipated that a more formal generalised approach will address perceived failures to recognize the full impact of the many interdependencies between the agencies, or the length of time between successive disasters. Soyler and Sala-Diakanda propose a MBSE methodology capturing the architectures of complex disaster management system and so encourage a holistic approach to problem solving by enforcing traceability. These models are to consist of state machines representing transitions in the incident state, Use Case diagrams identifying the role of government and non-governmental organisations, activity charts to model planning and preparedness, and block definition and internal block diagrams representing the structure and connections between the various management entities.

Similarly, Guiochet, et al [51], argue that HAZOP applied to robotic systems might also be facilitated with UML Use Cases and Sequence Diagrams formally defining system requirements specifications embodying some behavioural characteristics with which to reason about deviations. Consequently, a model for an interactive form of HAZOP might accommodate similar formalism in a unified modelling approach, extended to include Parametric diagrams and Constraint Blocks encapsulating the causal dynamics models as proposed in STAMP. Each Use Case and Sequence Diagram might also represent the various types of element comprising an “Intent Specification” [69], encompassing the various perspectives as described previously whilst introducing the concept of a representation for any anticipated specific system “dysfunction models” also. Therefore the beginning of a MBSE approach is partially adopted in this work to formally capture the system safety design intent, within a framework of an Intent Specification.

3.3.3. Hazard Detection & Complex Behaviours

It would appear that the great majority of works addressing hazard assessment in engineered system prefer to avoid safety issues arising within the regime of deterministic chaos. Work in this area appears to be relatively sparse, although the notion of Lyapunov exponents, strange attractors and the like as measures of emergent and complex behaviour are well established and accessible through standard textbooks on Control Theory (e.g. Baker and Gollub [4], Kapitaniak [61], and many more).

Zhu, et al [116], describe a method to autonomous detection and avoid surface hazards based on Lyapunov control methods for a future planetary-lander. The first part of this describes an optical and computational method of determining a median landing plane, surface normal and plane intercept point using optically observed statistical data and least-squares method. The second part combines a position-velocity state-space model with a statistical hazard model based upon position (reflecting physical positions with respect to hazards in the unevenness of the reachable terrain) and processing these states produces a Lyapunov function surface, with aim then being to direct the trajectory towards the nearest local minimum derivative (flattest and least hazardous) point in this surface. In this the Lyapunov function represents derivatives with respect to the two-dimensional spatial coordinate system describing the surface and statistical evidence of surface unevenness.

Perhaps of more relevance to this project, which is concerned with behavioural hazards emerging over time, McCue, et al [74], employ a Finite Time Lyapunov Exponent (FTLE) in the detection of instabilities leading to ship capsize in real-time. In this it is recognised that small variations in the initial conditions, described in terms of wave size and position, can result in large variations in behaviour. Expressing a predictor of this behaviour, the Lyapunov exponent is conventionally calculated along a trajectory *x*, as shown in equation 3.5.

$$\boldsymbol{\lambda_T\big(x(t), \delta x(0)\big) = \frac{1}{T}\ln\frac{\|\delta x(t+T)\|}{\|\delta x(t)\|}} \qquad 3.5$$

In this case the trajectory described by *x* represents the extent of roll of the ship, and the finite time interval is fixed window shorter than the entire history of motion, where resultant positive Lyapunov values anticipate ship motions of a more chaotic nature and hence tendency to capsize, whilst negative values anticipate stable behaviour.

3.3.4. Bayesian & Monte Carlo

Bayesian modelling, either as qualitative "belief networks" or as quantifiable statistical inference models, has a role in capturing probabilistic, likelihood and belief relationships between hazardous deviations, potential system failure causes, dynamical behaviour and externally imposed safety constraints. But how might this best be achieved?

Dilks, et al, [24], provides an example of the application of Bayes theorem to improving the estimation of uncertain parameters in water quality models, as used in the water treatment industry. Employing the Bayesian relationship for the posterior probability that an uncertain model parameter *θ* is correctly estimated given the observed data *x* (3.6), and that the effect of the observed evidence is constant (3.7), then substituting the value obtained from the "Likelihood function" *L(θ|x)* provides the general relationship for the probability that an uncertain parameter is given by the observed data (3.8).

$$\boldsymbol{P(\theta|x) = \frac{P(x|\theta)\cdot P(\theta)}{P(x)}} \qquad 3.6$$

$$\boldsymbol{P(\theta|x) = c\cdot P(x|\theta)\cdot P(\theta)} \qquad 3.7$$

$$\boldsymbol{P(\theta|x) = c\cdot L(\theta|x)\cdot P(\theta)} \qquad 3.8$$

$$\boldsymbol{L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta)} \qquad 3.9$$

The joint probability expressed in the above relationship (3.9) yields a single valued "Likelihood" that data value x_i will be observed for a particular value of uncertain model parameter *θ*, where *f(*x_i*|θ)* is the probability density function of *x* given *θ*. Assuming a Gaussian (normal) and independently distributed error model, with an estimate for the standard deviation *σ* representing the expected deviation due to errors, the following expression yields a likelihood value from the joint probabilities arising with each observed data item with respect to common mean *μ* (3.10).

$$\boldsymbol{L(\theta|x) = \prod_{i=1}^{n} f(x_i|\mu,\sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right],\ \theta = (\mu,\sigma^2)} \qquad 3.10$$

Substituting this expression of the "Likelihood", calculated directly from the observed data x_i with respect to the model prediction for this data, u_i, gives the updated posterior probability that the uncertain model parameter correctly estimates any observation, in terms of the measured errors in prior observed data with respect to model predictions, assuming zero mean in the error distribution (3.11). The prior probability distribution *P(θ)* is assumed uniform (rectangular) over some interval.

$$\boldsymbol{P(\theta|x) = c \cdot \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\epsilon_i}{\sigma}\right)^2\right] \cdot P(\theta),\ \epsilon_i = x_i - u_i,\ \mu_\epsilon = 0} \qquad 3.11$$

With this Dilks, et al, propose that data points used to calculate the likelihood[3] be obtained using Monte Carlo simulations of measurement error applied to model input data derived from observed field data within the relevant model, to sample from preliminary estimates of parameter distributions, and thereby improve the estimated uncertainty in a given set of model parameters. It is necessary however to ensure that any covariance is identified.

In another related application of Bayesian methods, Dawsey, et al, [21] propose to combine evidence to better characterise contamination events and reduce false positive detections in water distribution systems, using a Bayesian Belief Network (BBN) to integrate sensor data with validation evidence from contamination scenarios. A particular concern is where the true probability of an occurrence is much less than the false positive rate of a contaminant detection sensor. They propose that a BBN might represent the causal relationships between events and observations that comprise a contamination scenario. Simulations are used to estimate the probability of detection given that a contaminant release has occurred. Again the Bayesian network uses the Likelihood function (3.12) to determine the likelihood that a sensor value x_i belong to *Parents(X_i)*, for example a surrogate water quality measure, represented in preceding nodes of the network.

$$\boldsymbol{P(x_1, \dots x_n) = \prod_{i=1}^{n} P\big(x_i \big| Parents(X_i)\big)} \qquad 3.12$$

In a network model where the causal events represent contaminants in a number of geographical locations, along with operation change events – *Parents(X_i)* – and imply observable evidence in a number of sensor outputs and operation record (x_i), the authors claim some improvement in the discrimination of false positive sensor outputs. However, the authors also note that time differences between detections and timing itself *"... is relevant to the probability that observations are evidence of the same causal event"* [21], where time is not represented in the current model.

[3] Mysteriously Dilks, et al, express the likelihood calculation without raising the term $1/\sqrt{(2\pi\sigma^2)}$ to the power *n*.

3.4. Summary Considerations

The Bayesian expression described so far contribute to the determination of a Likelihood function and point to the inference required to identify the probability that a cause, or combination of causes, lie behind a particular effect. However, the principle power of the Bayesian approach lies in a capacity to update a Bayesian model of inference using new evidence.

Returning to consider also Hu's, et al [56] Bayesian risk assessment modelling for ship navigation, this approach combined with the other foregoing uncertainty modelling approaches, suggest an integration of dynamics modelling with Bayesian inference, coupled through Monte Carlo simulation of key causal parameters. However, questions arise:

- Is there adequate accident data available with which to validate this approach?
- How otherwise might a learnt model be validated?
- Consequently, would a system hazard model embodying a learnt structure provide acceptable evidence in the aerospace and defence domain?

In the case of Bayesian networks, so far as learning from data is concerned, it is known that where data is only partly observed estimates of uncertain model parameters θ may be found using the expectation–maximisation (EM) algorithm. This algorithm iteratively seeks the expected log-likelihood with respect to the conditional distribution of the unobserved variable using a current estimate of the uncertain model parameter(s), and then within an iteration computes the parameter values that maximise the expected log-likelihood. In this, the initial "Expectation" step uses current values for the observed variables (X) and the uncertain model parameter (θ) to calculate $\mathrm{P}(Z \mid X, \theta)$ – where Z is the unobserved variable(s); then subsequently in the "Maximisation" step the current θ is replaced by the expression $\theta \leftarrow \arg\max_{\theta'} Q(\theta'|\theta)$. With iteration, this process is guaranteed to find the local maximum. This is the method that is to be used to train the Bayesian network representing the inferred relationships between flaws and defects to the expected deviations derived from the HAZOP process. This training data is to be generated with Monte Carlo simulations of a System Dynamics model. As the particular network is to be derived from the structure and hierarchy of the HAZOP table representing the initial beliefs of the HAZOP team, in terms of relating believed cause and effect, then with such a "known" graph, the EM algorithm is an appropriate machine learning mechanism to be applied so as to learn and update belief from partly observed data.

Chapter 4. Systems Engineering Review

"The major difference between a thing that might go wrong and a thing that cannot possibly go wrong is that when a thing that cannot possibly go wrong goes wrong it usually turns out to be impossible to get at or repair."

Douglas Adams, "Mostly Harmless"

4.1. Development Life-Cycles

4.1.1. Beyond ARP 4754A / 4761

Close upon the time when the first *Certification Considerations for Highly-Integrated or Complex Aircraft Systems* [99] was published, Hugge and Lang [57] considered how producers and integrators of military avionic systems had then for a while already addressed the necessary discipline and structure required to safely create highly integrated or complex avionic systems – putting forth the view that the Advanced Design for Quality Avionic Systems (ADQAS) process satisfied the same system development assurance processes as ARP 4754. Within the ADQAS process six phases are identified: Avionics Requirements Definition, Avionics Preliminary Design, Avionics Detailed Design, Avionics Engineering and Manufacturing Development, Production, and Operation and Support. Overall the system development process would appear to adopt the "waterfall" model, with the only obvious feedback being in the Reliability Centred Maintenance and analysis of results used to inform the next system development cycle. It is also noted that *"... Quality Function Deployment (QFD) provides an excellent tool to derive product requirements in a disciplined, documented way, and to ensure early conflict resolution so that the required balance of performance, supportability, and cost is defined up front."*

In describing the systems engineering process, Blanchard and Fabrycky [7] detail and illustrate the general system acquisition process, with Conceptual Design, Preliminary Design, Detail Design and Development, Production, and Operational Use and Support phases; where these separate phases support the development at the System Level, Subsystem Level, Component Level, and in Modifications for Improvement, each within their own "waterfall" cascade of requirement, allocation, synthesis, and evaluation. This broadly forms the basis of other processes that adapt this sequence of events with different approaches as to what, where and when validation and verification applies. From this, approaches have been developed so as to facilitate development of either self-contained integrated system architectures (e.g. aircraft), predicated upon decomposition and hierarchy, or incremental approaches allowing system developmental risk to be minimised. The "waterfall" model is perhaps only best suited to the development of single, non-integrated, stand-alone software.

As might be assumed from the foregoing, ARP 4754A [99, 101] adopts the top-down self-contained integrated system architectural approach predicated upon decomposition and hierarchy, commonly understood to be represented in the Systems Engineering "Vee" process model, as depicted in Figure 4.1 (with a small extension also to consider a place for HAZOP within this).

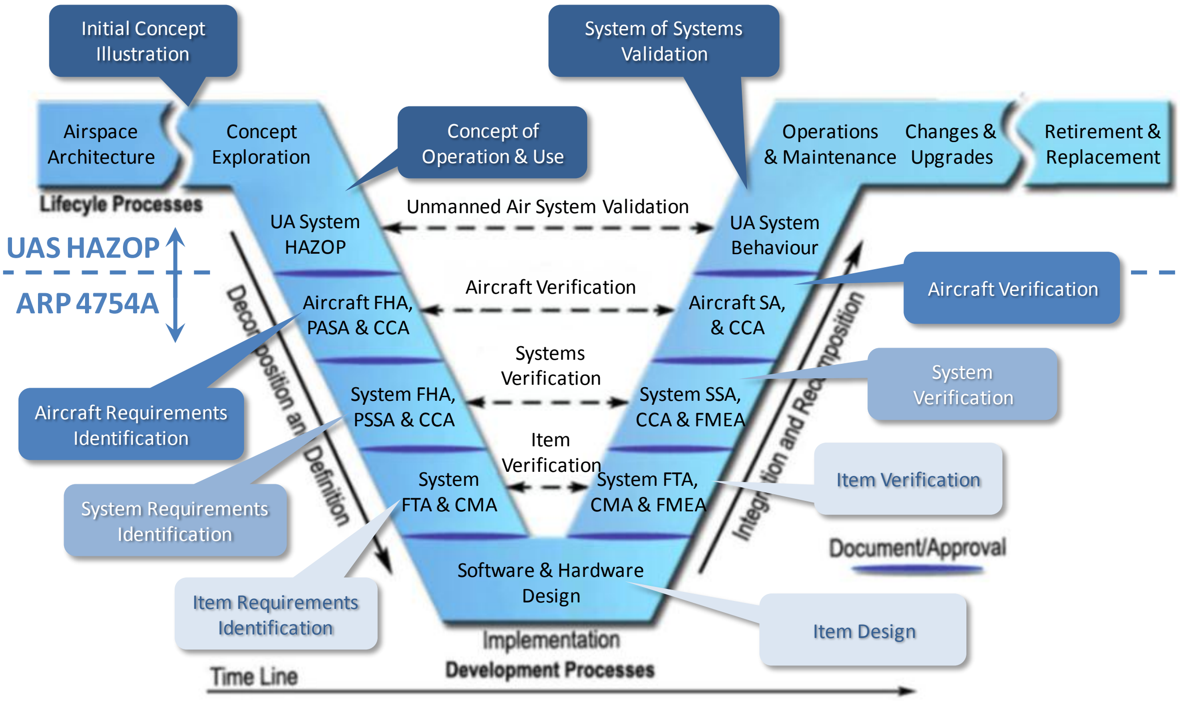


Figure 4.1 Systems Engineering "Vee" Diagram, adapted from ARP4754A [101]

Within the above process, the decomposition leg to the left produces aircraft, system and item requirements, with each lower level assessment being used to validate the decomposed requirements. Correspondingly, the integration leg to the right provides for the verification of the resulting system as it is composed. Within Systems Engineering it is also important that language be used carefully, otherwise requirements may be misunderstood, therefore with reference to the US Project Management Institute guide [88] the following definitions apply:

- Validation. The assurance that a product, service, or system meets the needs of the customer and other identified stakeholders. It often involves acceptance and suitability with external customers;
- Verification. The evaluation of whether or not a product, service, or system complies with a regulation, requirement, specification, or imposed condition. It is often an internal process [88].
- Where this work proposes a closed-loop safety assessment process, incorporating HAZOP styled upon STPA, system dynamics modelling, and Bayesian inference, it is also proposed that the process developed with this work is considered as an approach to extend the Preliminary Aircraft Safety Assessment (PASA) for a UAS / UAV – hence the upward extension of the "Vee" in Figure 4.1.

4.1.2. Spiral Development

With a spiral development process, Boehm [8] posits that a risk-driven rather than a document-driven or code-driven approach to software development ought to be used where the overall requirement is unknown or embodies a great deal of uncertainty – upon reflection, a good example of such a challenge might be revealed with the National Programme for Information Technology (NPfIT) for the NHS. *"A primary source of difficulty with the waterfall model has been its emphasis on fully elaborated documents as completion criteria for early requirements and design phases. ... Document-driven standards have pushed many projects to write elaborate specifications of poorly understood user interfaces and decision support functions ...",* Boehm [8].

Boehm also challenges the belief that the automatic conversion of a formal specification into a program alleviates the problem of addressing risk with evolving and uncertain requirements, *"The transform model also shares some of the difficulties of the evolutionary development model, such as the assumption that users' operational systems will always be flexible enough to support unplanned evolution paths."* [8] With this point in mind consideration ought to be given to what the framework ought to be to support the evolution of requirements and the overall modelling approach.

With the spiral model, all iterations of evolutionary development incorporate a risk analysis phase involving prototypes to evaluate alternatives, and to identify and resolve risks, Figure 4.2. Also, beyond the application of the spiral model to the principles of software development, the same principles apply to more general complex system engineering developments, for much the same reasons that SysML is presented as "systems" extension to the software oriented Unified Modelling Language – principally modern complex systems usually incorporate complex software elements and within this the hardware, software, and even “liveware”, aspects ultimately merge.

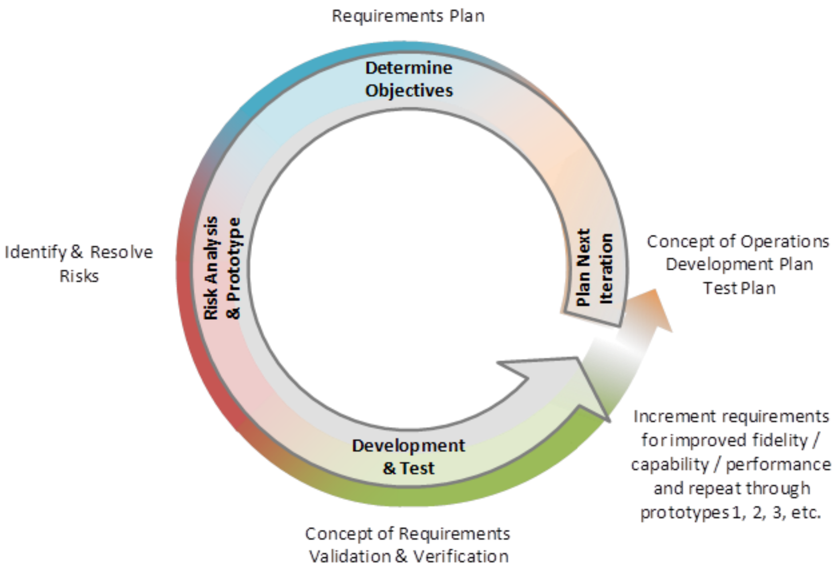

Requirements Plan
Determine Objectives
Identify & Resolve Risks
Risk Analysis & Prototype
Plan Next Iteration
Concept of Operations Development Plan Test Plan
Development & Test
Increment requirements for improved fidelity / capability / performance and repeat through prototypes 1, 2, 3, etc.
Concept of Requirements Validation & Verification

Figure 4.2 A Simplified Interpretation of Boehm's Model of Spiral Development [8]

4.2. Architecture & Validation

4.2.1. Design Intent

A consideration also in developing the model architecture specification is the requirement to capture design intent with respect to behaviour. For this it is proposed that a suitable model might be found in the field of behaviour-based robotics, for example with the Subsumption Architecture, after Brooks [12]; at least for machine based elements of the behavioural intent.

In this work collision avoidance has been modelled as a layered competence and behaviour-based robotics architecture – as illustrated in Figure 4.3. This depicts an intent to suppress a particular interpretation of the General Flight Rules within the Rules of the Air [15] in the event that an emergency back-up behaviour is required to avoid an imminent collision with threatening traffic. Whilst this representation does not perfectly model the actual software implementation, it can be argued that it does capture design intent, and might therefore be used to capture a behavioural requirement rather than a particular design solution.

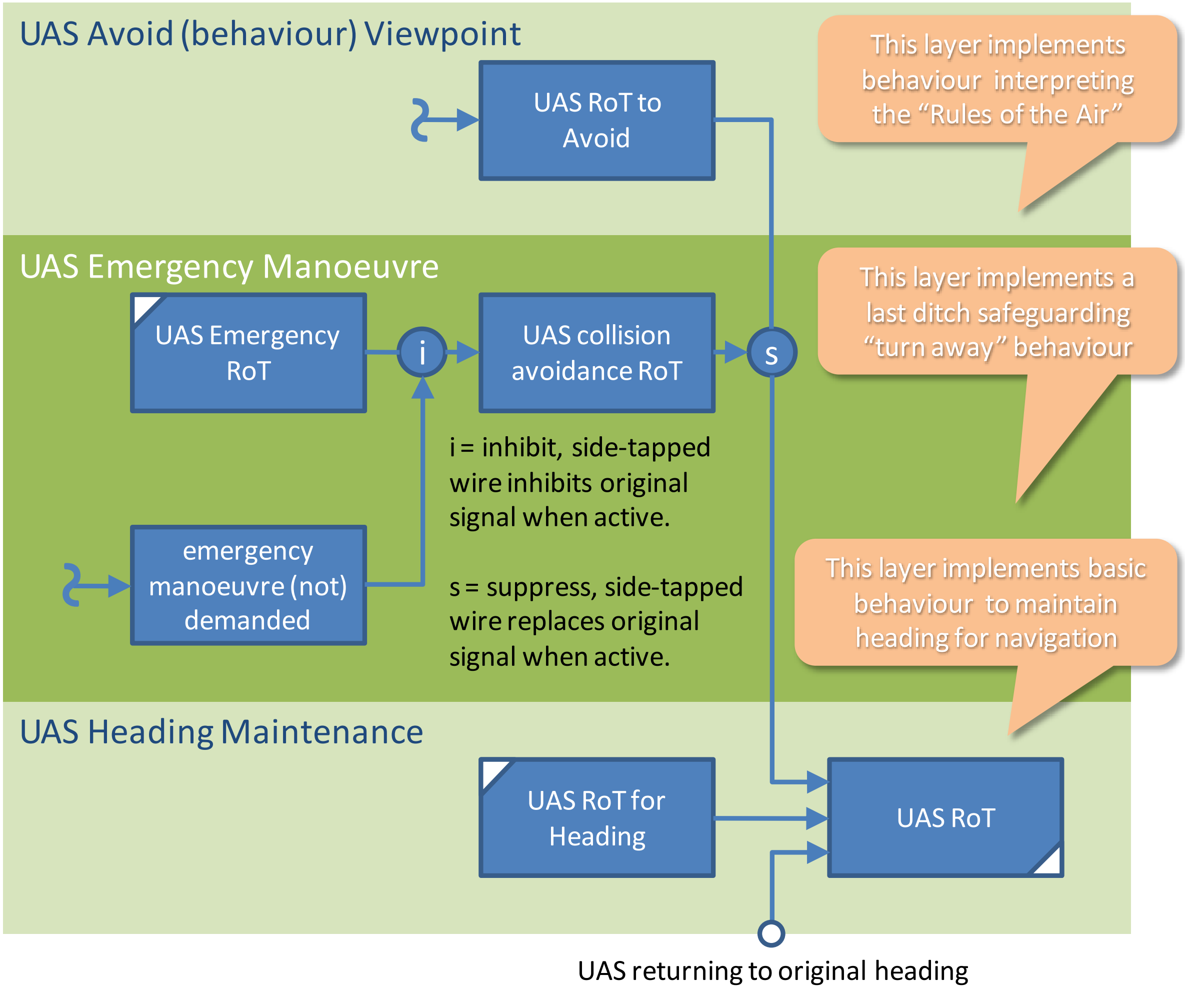

UAS Avoid (behaviour) Viewpoint
UAS RoT to Avoid
This layer implements behaviour interpreting the "Rules of the Air"
UAS Emergency Manoeuvre
UAS Emergency RoT
i
UAS collision avoidance RoT
s
This layer implements a last ditch safeguarding "turn away" behaviour
i = inhibit, side-tapped wire inhibits original signal when active.
emergency manoeuvre (not) demanded
s = suppress, side-tapped wire replaces original signal when active.
This layer implements basic behaviour to maintain heading for navigation
UAS Heading Maintenance
UAS RoT for Heading
UAS RoT
UAS returning to original heading

Figure 4.3 Behavioural Subsumption Model after Brooks [12]

Pursuant to the broad objective of developing a "unified" hazard analysis method, incorporating the techniques of HAZOP, STAMP and STPA, a collision avoidance scenario has been constructed complying with the Rules of the Air including a series of interaction vignettes,. Incorporating these rules is clearly another aspect of design intent, in the adoption of operational rules generally, and also at higher levels of abstraction in the design rules. Within the system dynamics model, both vehicles therefore are to embody the Rules of the Air Regulations 2010, as described in CAP 393, Section 4, General Flight Rules [15]; and as detailed in the following excerpt:

> *"Avoiding aerial collisions – an aircraft which has the right-of-way under this rule shall maintain its course and speed.*
>
> *Converging – when two aircraft are converging in the air at approximately the same altitude, the aircraft which has the other on its right shall give way.*
>
> *Approaching head-on – when two aircraft are approaching head-on, or approximately so, in the air and there is a danger of collision, each shall alter its course to the right.*
>
> *Overtaking – an aircraft which is being overtaken in the air shall have the right-of-way and the overtaking aircraft, whether climbing, descending or in horizontal flight, shall keep the out of the way of the other aircraft by altering its course to the right."*
>
> The Air Navigation Order, The Rules of the Air Regulations 2007 [15].

The entity representing the UAV additionally embodies an emergency manoeuvre behaviour that provides an alternative interpretation of the above rules when immediately threatened by the course taken by the other aircraft, as suggested by Figure 4.3 – in general this back-up behaviour will cause the "UAV" to attempt to turn away from the direction of the threat. Also within these rules, each vehicle will either take the shortest direction of turn towards its original heading, where a rule allows return to heading, or by default takes a turn to the left where the angle of turn required is approximately equal either way. For each interaction vignette, dynamic behaviour arises partly as a consequence of the internal interactions between the intended behaviours, as defined by layers of subsumed competences (three in this case) along with the external interactions between the entities as defined by the above control rules, and also the system dynamics. All are to be contained within an Intent Specification [69], and taken together also in effect articulate a requirement for a legal airborne collision avoidance rule.

Opposed to avoidance, definitions of proximity should also be considered. The UKAB [111] provides some formality of understanding in defining safety constraints through a definition of Air Proximity.

4.2.2. Validation through Questionnaire

In the absence of actual UAS collision avoidance data, or other such accessible models, it is difficult (at least) to objectively validate the models created within this work. Therefore a quantifiable statistical approach has to be used, from subjective data derived through a questionnaire. However, with the likelihood of only having access to quite low numbers of people with any relevant experience of expertise, and having to reconcile potentially widely varying opinions, also rules out conventional statistical parametric methods. Instead a non-parametric statistical method has been chosen. This method enables the reconciliation of opinions, given a fixed set of options to choose from, by seeking to measure the inter-rater agreement between respondents, and also in this case the outputs arising from the simulations. Much as in setting a multiple-choice examination, greater effort and care has to be devoted in the construction of the questions than in the processing and assessment of the responses that can in fact be mechanised. Cohen's Kappa [20], is used as the requisite measure, calculated on the choice of options on a per question basis in the questionnaire – generating a pair of agreement values for each question, with respect to both the "correct" model based option and a "floating" option identified from the consensus as it emerges.

In the Measurement of Observer Agreement for Categorical Data, Landis & Koch [64] suggest a refinement to Cohen's method in that they propose a measure and categorisation for the *Strength of Agreement*, where Poor ≤ 0.0 < Slight ≤ 0.2 < Fair ≤ 0.4 < Moderate ≤ 0.6 < Substantial ≤ 0.8 < Almost Perfect ≤ 1.0. Whilst it is recognised that these divisions are arbitrary, the authors felt that these categories do provide useful "benchmarks"; likewise here for want of any better benchmark.

4.3. Emergent and Complex Behaviour

Whilst we have the popularisation of the science of deterministic chaos [49] misunderstanding may arise where the terms emergent or complex behaviour are conjoined with the expectations of the behaviour of deterministic systems, and especially where faults, failures and hazardous behaviour are being considered. However, in the field of Behaviour-based Robotics it is understood that these two concepts are intimately connected, frequently arise with even quite simple interacting control system, and are indeed a necessary aspect in the consideration of the engineering, tailoring or shaping of useful behaviours from such robotic systems. In particular Arkin [3] summarises the views of a number of other established and renowned researchers in the field of machine intelligence *"Emergence is: 'the appearance of novel properties in whole systems' (Moravec); 'Global functionality emerges from the parallel interaction of local behaviours' (Steels); 'Intelligence emerges from the interaction of the components of the system' (Brooks); 'Emergent functionality arises by virtue of interaction between components not themselves designed with the particular function in mind' (McFarland and Bosser);"* and a common thread is that *"... emergence is a property of a collection of interacting components (here, behaviours)"* [3, 9-11, 80]. These are the definitions and the nature of the interpretation that ought to be placed upon the application of the term *emergent* as it applies in this work – and also as applied in other relevant earlier work in Robotics by this author [5, 27, 28].

In considering the relationship between *emergence* and *complexity* a similar perspective comes from John Holland, a pioneer in the field of Genetic Algorithms, through observations made in the context of the execution of strict and deterministic rules within a game such as 'checkers'. In this Holland observes that *"Emergence occurs in systems that are generated (where) the systems are composed of copies of a relatively small number of components that obey simple laws"*, *"Emergent phenomena in generated systems are, typically, persistent patterns with changing components",* and that *"Generated complexity ... organized complexity can result from simple rules and procedures."* [54] Holland provides a simile for this interpretation of complexity by comparing it with the persistent similarity of patterns formed by a standing wave in front of a rock in a fast-moving stream – ever changing and with no two patterns exactly the same, but all having common distinguishable characteristics – or "seeing the same thing twice" in the formations of pieces in the games of checkers or chess, not exactly the same formations but having the same salient characteristics – or waves of neurons firing in the brain, and so on. This is the interpretation of emergence to be adopted in this work.

Similarly, in the *Origins of Order*, Kaufmann [62, 63] considers the nature of self-organisation and adaptation in complex systems by exploring the behaviour of computer models of randomly connected '*NK*' Boolean networks (logical switching networks having N nodes with K connections per node) from which study it is hypothesised that certain complex systems tend toward self-organisation – *"The transition from the ordered regime to the chaotic regime constitutes a phase transition, which occurs as a variety of parameters are changed. The transition region, on the edge between order and chaos, is the complex regime."* [63] Kauffman also hypothesises that adaption to the edge of chaos may ultimately become a general evolutionary principle in that systems poised between order and chaos are the natural culmination of selective evolution – if true one might imply that this form of complexity is essential in the functioning of an adaptive system, and it therefore becomes inevitably necessary to understand how hazards arising in such systems might be modelled.

4.4. Summary Considerations, Influences and Scope

Where Hu, et al [56] demonstrate an assessment of risk in ship navigation using Bayesian learning, wherein networks are wholly derived by learning from accident data, no equivalent a priori data is available for UAV / UAS near-misses and collisions – or indeed for any type of risk event involving an autonomous system. A challenge for this work is to reconcile any such learnt data or structures with information derived from a HAZOP process. This suggests that at least a part of any such inference model must be derived from the hazard structure derived from the HAZOP process itself. For this investigation into hazard-centric analysis of complex autonomous systems, the focus is the identification of a route towards "deviation" based risk assessment akin to HAZOP applied to a system (or subsystem) operating within its environment, as suggested by Wilkinson and Kelly [114], and Despotou, et al [22, 23], within scenarios as identified by Allenby and Kelly [2].

In taking a more systems theoretic approach towards the system and risk modelling, Leveson [67, 69-71] provides a suitable framework for consideration with the "Intent Specification" framework and STAMP, further developed by Owens, et al [86] and Stringfellow [105] to realise the beginnings of a new HAZOP type process by way of STPA. Leveson [67, 71] also prompts us to explore the value in describing the dynamics of any potential accident system in terms as inspired by Forrester [46]. Causal-loop and system dynamics modelling can be readily accomplished with a variety of simulation tools. Vensim™ has been adopted in this work, and though DYNAMO[4] is now obsolete Roberts, Andersen, et al [94] still provide some additional guidance when creating models of system dynamics.

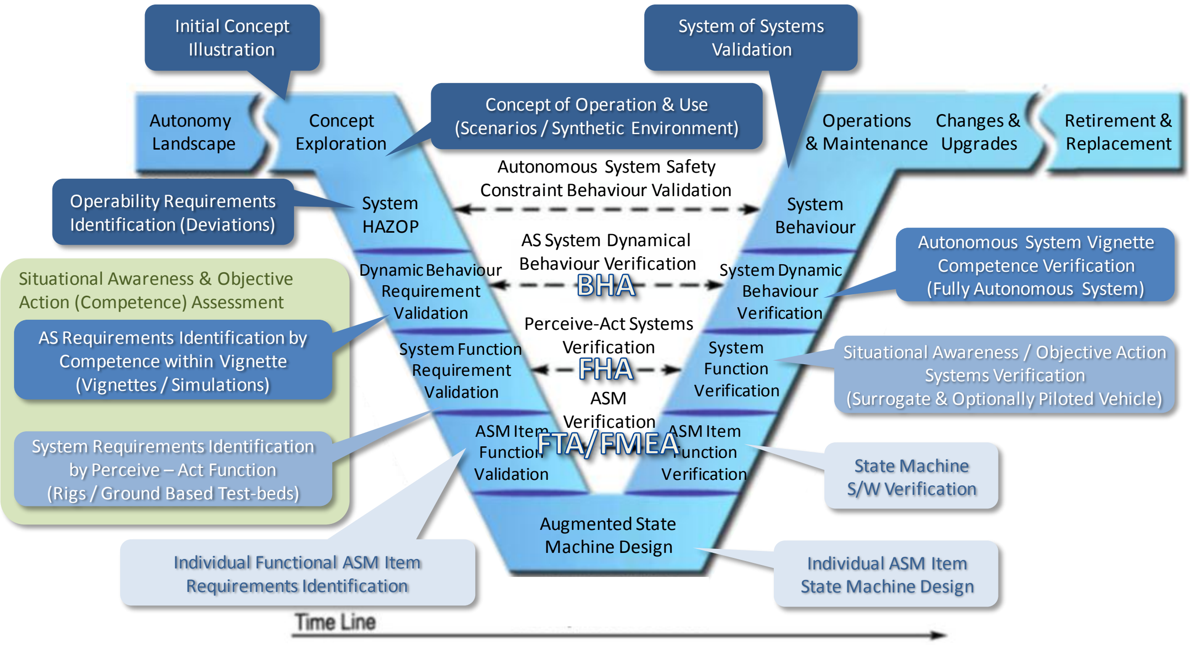
The purpose overall is to identify a plausible hybrid approach to system dynamic hazard assessment. This should integrate a mental model from within a HAZOP process with a Bayesian inference model relating cause, deviation, and effect. In the absence of historic data or observations, this might derive belief for particular hazardous relationships from stochastically generated sensitivity data for parameters associated with hazardous behaviour within a Systems Dynamic model. Hu, et al [56], Mohaghegh, et al [79], Trucco [110], and Dilks, et al [24] all provide different examples of Bayesian centred hybrid approaches to risk and uncertainty involving different sources of data, some associated with established Functional Hazard Assessment methods; for example fault trees and FMEA.

In consideration of a suitable form of representation for the integration of modular behaviours one might look to the field of behaviour-based robotics, for example in Brooks' decomposition of robot competences, known as the Subsumption Architecture [12]. This architectural approach provides a representation in the form of relatively simple controllers, defined as augmented finite state machines, interconnected so as to form layers of competence, each intended to realise a particular behaviour, and within which internal connections and mechanisms acting between layers act to subsume the lower primitive behaviours within higher and usually more purposeful behaviours. Hence this approach defines a reactive robotic scheme that embodies a layering of behavioural control strategies, each potentially having different goals and priorities, and that these overall constitute the feedback control model for the robotic entity's interaction with its environment – thereby providing a generic architecture with which to define behavioural interaction in autonomous systems.

Combining Boyd's Observe, Orientate, Decide, and Act (OODA) loop [34], Brooks robot behavioural competence levels [12], Taylor's, et al, PACT authority levels [106], and Endsley's situational

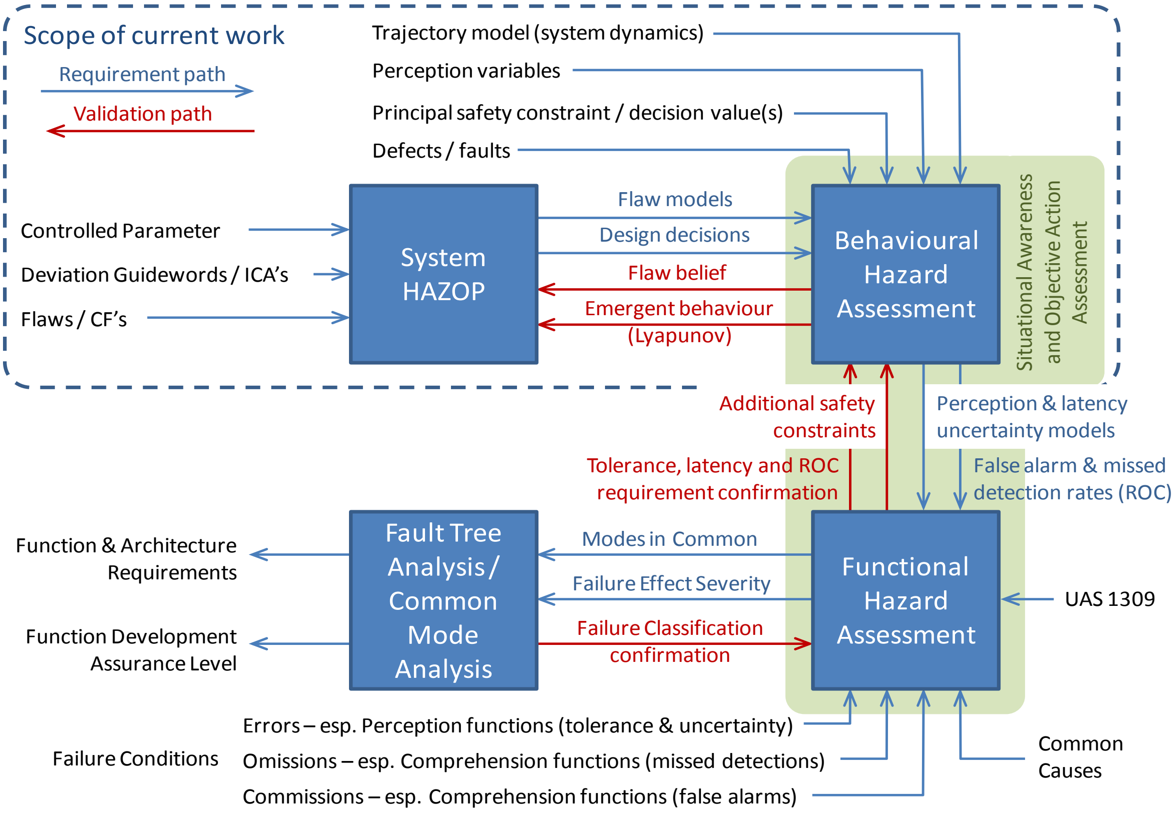
[4] For those with System Dynamic models in the DYNAMO language "Vensim" embodies a conversion utility.

awareness model [39, 41], creates the possibility of defining a general framework within which to consider autonomous system behaviour. Such an approach might also be integrated with the processes for the development of civil aircraft systems as we introduce autonomous behaviour as a requirement into these systems, Figure 4.4. Whilst this remains beyond the scope of the current work, potential connections within this wider development framework could be considered, Figure 4.5.



Initial Concept Illustration
System of Systems Validation
Autonomy Landscape
Concept Exploration
Concept of Operation & Use (Scenarios / Synthetic Environment)
Operations & Maintenance
Changes & Upgrades
Retirement & Replacement
Autonomous System Safety Constraint Behaviour Validation
Operability Requirements Identification (Deviations)
System HAZOP
System Behaviour
AS System Dynamical Behaviour Verification
Autonomous System Vignette Competence Verification (Fully Autonomous System)
Situational Awareness & Objective Action (Competence) Assessment
Dynamic Behaviour Requirement Validation
BHA
System Dynamic Behaviour Verification
AS Requirements Identification by Competence within Vignette (Vignettes / Simulations)
Perceive-Act Systems Verification
System Function Requirement Validation
FHA
System Function Verification
Situational Awareness / Objective Action Systems Verification (Surrogate & Optionally Piloted Vehicle)
ASM Verification
System Requirements Identification by Perceive – Act Function (Rigs / Ground Based Test-beds)
ASM Item Function Validation
FTA/FMEA
ASM Item Function Verification
State Machine S/W Verification
Augmented State Machine Design
Individual Functional ASM Item Requirements Identification
Individual ASM Item State Machine Design
Time Line

Figure 4.4 SE Integration Equivalent to ARP 4754A for Autonomous & Robotic Systems



Scope of current work
Requirement path
Validation path
Trajectory model (system dynamics)
Perception variables
Principal safety constraint / decision value(s)
Defects / faults
Controlled Parameter
Deviation Guidewords / ICA's
Flaws / CF's
System HAZOP
Flaw models
Design decisions
Flaw belief
Emergent behaviour (Lyapunov)
Behavioural Hazard Assessment
Situational Awareness and Objective Action Assessment
Additional safety constraints
Perception & latency uncertainty models
Tolerance, latency and ROC requirement confirmation
False alarm & missed detection rates (ROC)
Function & Architecture Requirements
Function Development Assurance Level
Fault Tree Analysis / Common Mode Analysis
Modes in Common
Failure Effect Severity
Failure Classification confirmation
Functional Hazard Assessment
UAS 1309
Failure Conditions
Errors – esp. Perception functions (tolerance & uncertainty)
Omissions – esp. Comprehension functions (missed detections)
Commissions – esp. Comprehension functions (false alarms)
Common Causes

Figure 4.5 A framework compatible with ARP 4754A and DO-178C [97, 101]

Comprehensive guidance is available to the aerospace systems engineer, providing for the development of safe manned aircraft and their associated systems, incorporating many decades of experience gained in the development of relevant systems and their operation. Central to this guidance, Aerospace Recommended Practice (ARP) 4754A [101] provides the basic framework within which function and item developmental assurance is framed. Also within this framework, ARP 4761 [100] then provides the additional guidelines and recommended hazard and reliability assessment methods considered to be appropriate to the undertaking of both qualitative and quantitative risk and reliability analysis; necessary to determine whether systems can meet their stated safety objectives. These recommended reliability methodologies quantify statistical risk within a frequentist model of probability – that is a measure of the relative frequency of a (hazardous) event arising in a series of trials. However, where particular events and combinations of circumstances are effectively unique and irreproducible in practice it may be more appropriate to adopt a Bayesian (belief based) approach towards probability. With this approach a subjective basis in the degree of belief in the likelihood of an event occurring is modified by incorporating evidence arising with actual occurrences of that event; combining intuitive estimates based upon experience and expertise with quantitative statistical data. Depending upon the reader’s point of view, one might perceive that the belief based approach sits on a spectrum between qualitative and quantitative perceptions, between subjective and objective measures of perceived risk. Any general behavioural description and predictions of behaviour arising with a complex autonomous system is likely to embody qualities best described in a Bayesian form. The question is what form should that take in practice?

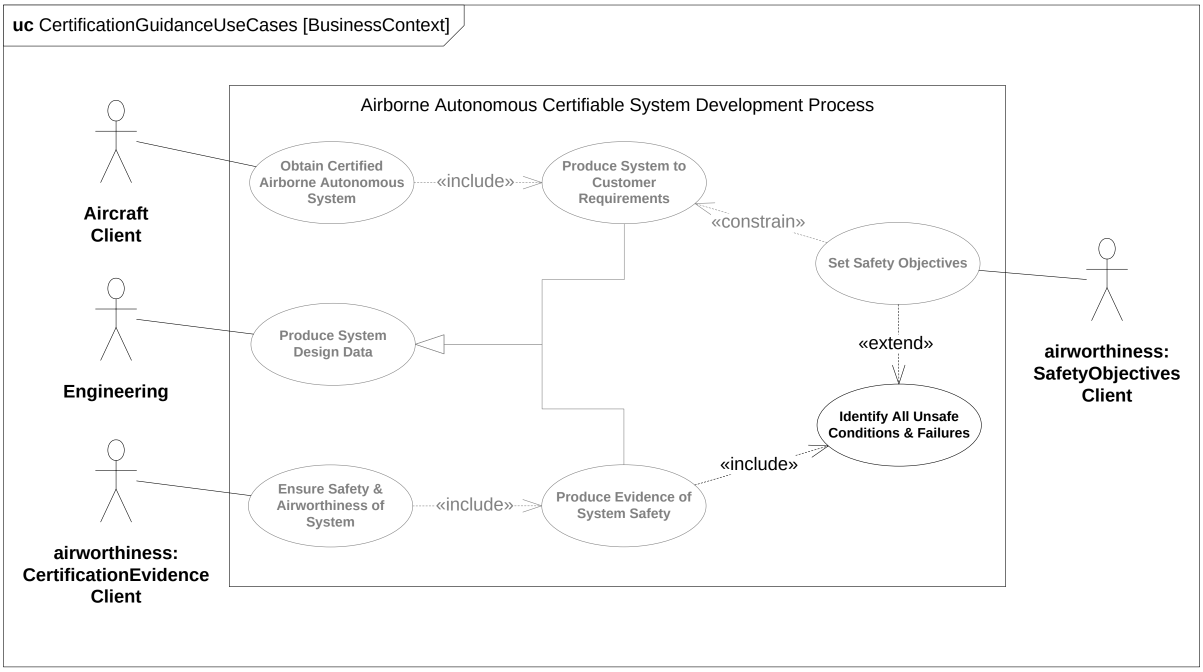
Part 2. System Framework

Chapter 5. System Modelling Requirements

"Everything makes sense a bit at a time. But when you try to think of it all at once, it comes out wrong."

Terry Pratchett, "Only You Can Save Mankind"

5.1. Textual Analysis & QFD



uc CertificationGuidanceUseCases [BusinessContext]
Airborne Autonomous Certifiable System Development Process
Aircraft Client
Engineering
airworthiness: CertificationEvidence Client
airworthiness: SafetyObjectives Client
Obtain Certified Airborne Autonomous System
«include»
Produce System to Customer Requirements
«constrain»
Set Safety Objectives
Produce System Design Data
«extend»
Identify All Unsafe Conditions & Failures
«include»
Ensure Safety & Airworthiness of System
«include»
Produce Evidence of System Safety

Figure 5.1 Use Case Simplifying the Certification Purpose

As this work is conducted within a systems engineering framework, a degree of rigour in the requirements analysis is necessary, and yet is also difficult to conduct in practice for this topic with the scarcity and relative inaccessibility of representative stakeholders – given the newness of the evolving niche being created by the discipline of Autonomous Systems. Therefore, a different approach was proposed, as much is already published on the topic – and in some respects this provides a better, more informed and global viewpoint. Therefore, relevant topic items were drawn from the literature to stand as proxies for stakeholder views. Interpretation and the exercising of personal judgement upon the information so derived produced lists of "means" and "needs" through a process of textual analysis. Together with a "differential" QFD analysis, Figure 5.2, this was used in place of a "live" requirements elicitation process – to be placed into the context shown in Figure 5.1.

The QFD analysis technique is a systematic and structured method for turning customer requirements – primarily non-functional system requirements (and often tied to specific non-functional performance requirements) – into measurable requirements covering all aspects of delivering a product or service from marketing through design, manufacturing, and production. In order to be properly effective the process needs to clearly identify the customer and it is also complex and time consuming. In this project the customer has necessarily been abstracted away, instead adopting a structured reading of the relevant literature – effectively making this published community the proxy customer.

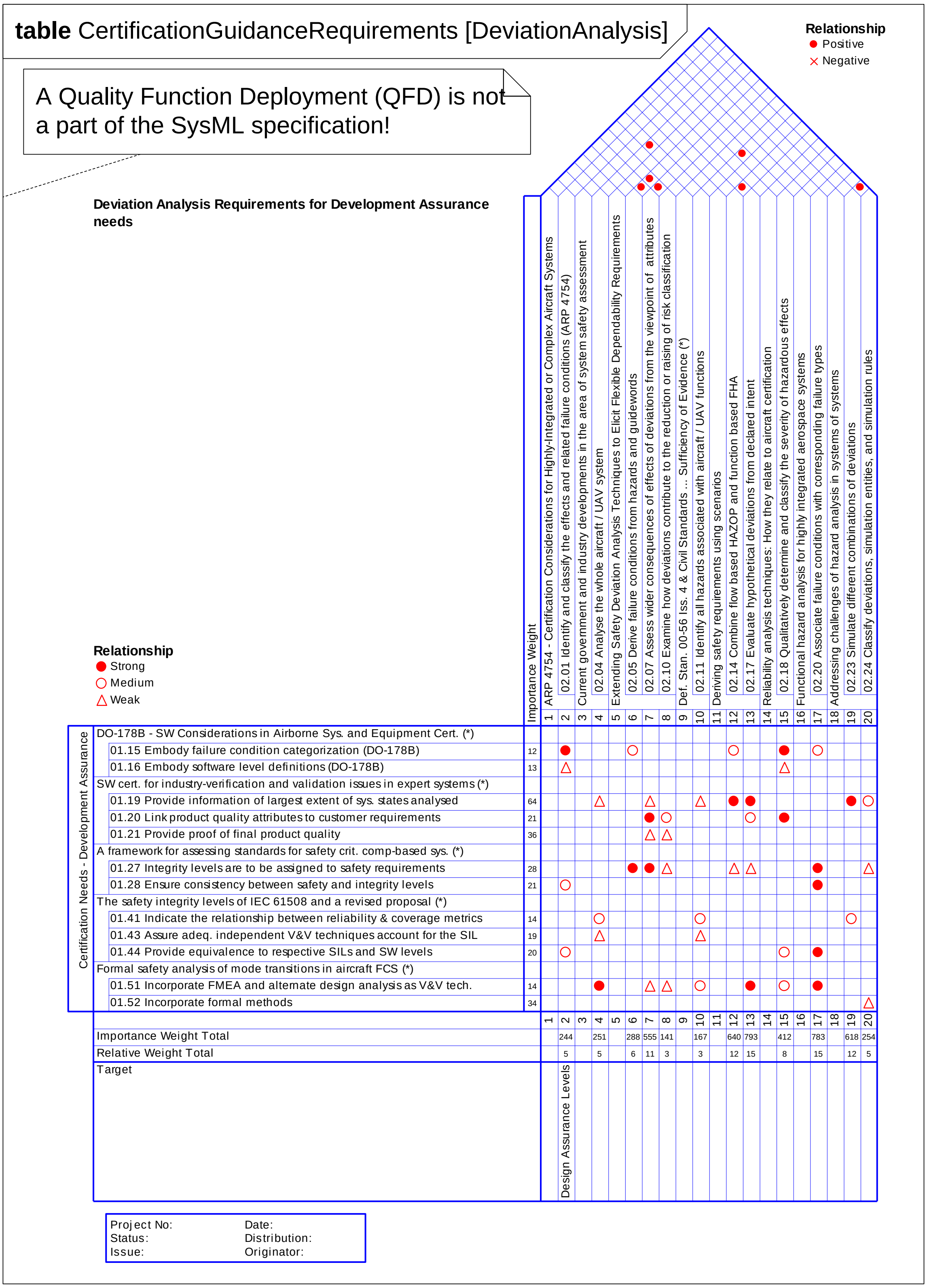


table CertificationGuidanceRequirements [DeviationAnalysis]
A Quality Function Deployment (QFD) is not a part of the SysML specification!
Relationship
● Positive
× Negative
Deviation Analysis Requirements for Development Assurance needs
Relationship
● Strong
○ Medium
△ Weak

		Importance Weight	1 ARP 4754 - Certification Considerations for Highly-Integrated or Complex Aircraft Systems	2 02.01 Identify and classify the effects and related failure conditions (ARP 4754)	3 Current government and industry developments in the area of system safety assessment	4 02.04 Analyse the whole aircraft / UAV system	5 Extending Safety Deviation Analysis Techniques to Elicit Flexible Dependability Requirements	6 02.05 Derive failure conditions from hazards and guidewords	7 02.07 Assess wider consequences of effects of deviations from the viewpoint of attributes	8 02.10 Examine how deviations contribute to the reduction or raising of risk classification	9 Def. Stan. 00-56 Iss. 4 & Civil Standards ... Sufficiency of Evidence (*)	10 02.11 Identify all hazards associated with aircraft / UAV functions	11 Deriving safety requirements using scenarios	12 02.14 Combine flow based HAZOP and function based FHA	13 02.17 Evaluate hypothetical deviations from declared intent	14 Reliability analysis techniques: How they relate to aircraft certification	15 02.18 Qualitatively determine and classify the severity of hazardous effects	16 Functional hazard analysis for highly integrated aerospace systems	17 02.20 Associate failure conditions with corresponding failure types	18 Addressing challenges of hazard analysis in systems of systems	19 02.23 Simulate different combinations of deviations	20 02.24 Classify deviations, simulation entities, and simulation rules
Certification Needs - Development Assurance	DO-178B - SW Considerations in Airborne Sys. and Equipment Cert. (*)																					
	01.15 Embody failure condition categorization (DO-178B)	12		●				○						○			●		○			
	01.16 Embody software level definitions (DO-178B)	13		△													△					
	SW cert. for industry-verification and validation issues in expert systems (*)																					
	01.19 Provide information of largest extent of sys. states analysed	64				△			△			△		●	●						●	○
	01.20 Link product quality attributes to customer requirements	21							●	○					○		●					
	01.21 Provide proof of final product quality	36							△	△												
	A framework for assessing standards for safety crit. comp-based sys. (*)																					
	01.27 Integrity levels are to be assigned to safety requirements	28						●	●	△				△	△				●			△
	01.28 Ensure consistency between safety and integrity levels	21		○															●			
	The safety integrity levels of IEC 61508 and a revised proposal (*)																					
	01.41 Indicate the relationship between reliability & coverage metrics	14				○						○									○	
	01.43 Assure adeq. independent V&V techniques account for the SIL	19				△						△										
	01.44 Provide equivalence to respective SILs and SW levels	20		○													○		●			
	Formal safety analysis of mode transitions in aircraft FCS (*)																					
	01.51 Incorporate FMEA and alternate design analysis as V&V tech.	14				●			△	△		○			●		○		●			
	01.52 Incorporate formal methods	34																				△
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Importance Weight Total			244		251		288	555	141		167		640	793		412		783		618	254
	Relative Weight Total			5		5		6	11	3		3		12	15		8		15		12	5
	Target			Design Assurance Levels																		

Project No:
Status:
Issue:

Date:
Distribution:
Originator:

Figure 5.2 Deviation Analysis Requirements as QFD Analysis from Textual Analysis

With plenty of discussion and opinions regarding certification and safety of autonomous airborne systems in the literature, it appeared that a search using appropriately selected keywords could be used as the basis for a Systemic Textual Analysis. Therefore, a series of keywords were identified and classified within four levels of abstraction of the certification process; UAV Context, Airworthiness & Certification, System Safety Assessment, and Safety Assurance. A QFD analysis was performed with sentences found containing these identified keywords re-phrased to express potential operational or non-functional system requirements (WHAT or needs) and functional or non-functional implementation requirements (HOW or means), with some shown in Figure 5.2.

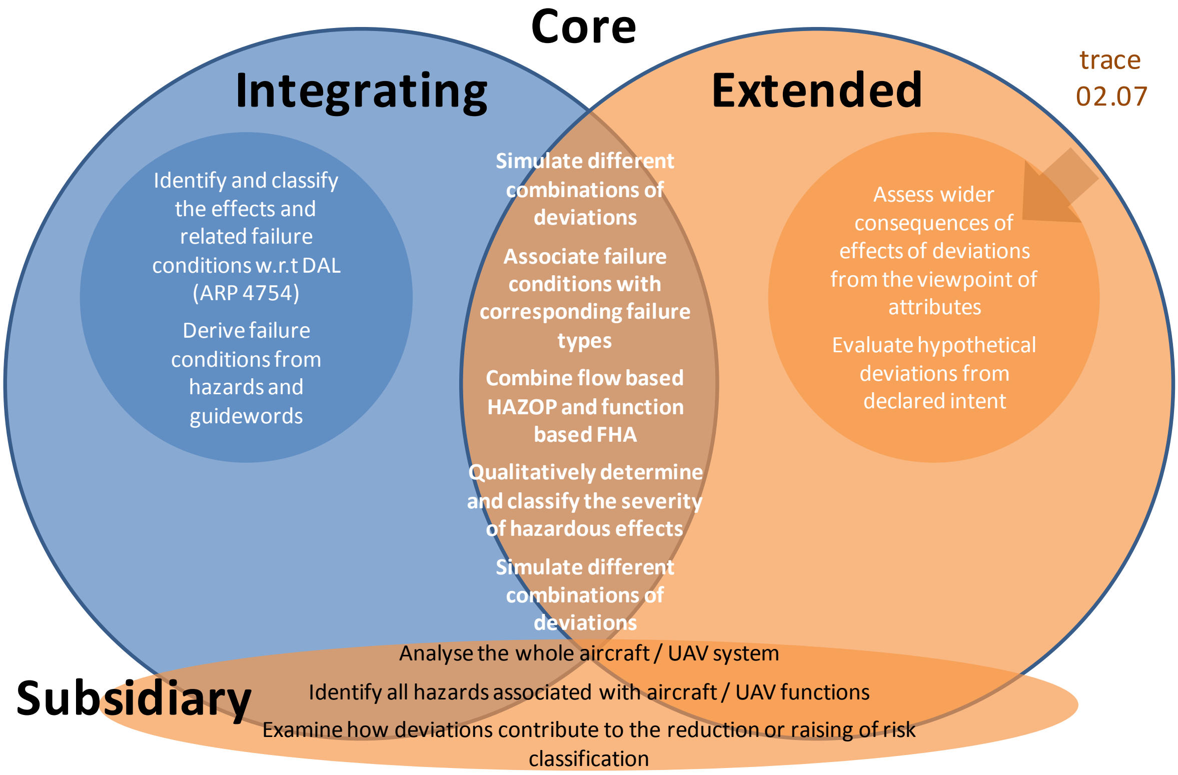

Core
Integrating
Extended
trace
02.07
Identify and classify the effects and related failure conditions w.r.t DAL (ARP 4754)
Derive failure conditions from hazards and guidewords
Simulate different combinations of deviations
Associate failure conditions with corresponding failure types
Combine flow based HAZOP and function based FHA
Qualitatively determine and classify the severity of hazardous effects
Simulate different combinations of deviations
Assess wider consequences of effects of deviations from the viewpoint of attributes
Evaluate hypothetical deviations from declared intent
Subsidiary
Analyse the whole aircraft / UAV system
Identify all hazards associated with aircraft / UAV functions
Examine how deviations contribute to the reduction or raising of risk classification

Figure 5.3 Requirement Sets Identifying selected Means or “How’s”

In order to separate requirements from the literature a “differential” method was developed, where analysis is performed at two levels, in this case Safety Assurance (what) – Airworthiness & Certification (how), and Airworthiness & Certification (what) – System Safety Assessment (how). The process is repeated, first with the priority of classifying hazardous effects satisfying regulatory guidelines, and then prioritising the provision of information as to the extent of states analysed and academic challenges. With these results, taking cuts from the upper halves of each set of functional requirements produces two sets of outputs. Depicted to the left in Figure 5.3 are the requirements for a programme focused upon identifying and classifying hazardous effects, whilst to the right focuses upon providing information of the extent of states analysed. Between these two, the conjunction identifies common requirements satisfying both of these sets. The final result is then formulated as a Use Case, in Figure 5.4. Further detail with regard to this approach is contained in the paper [30].

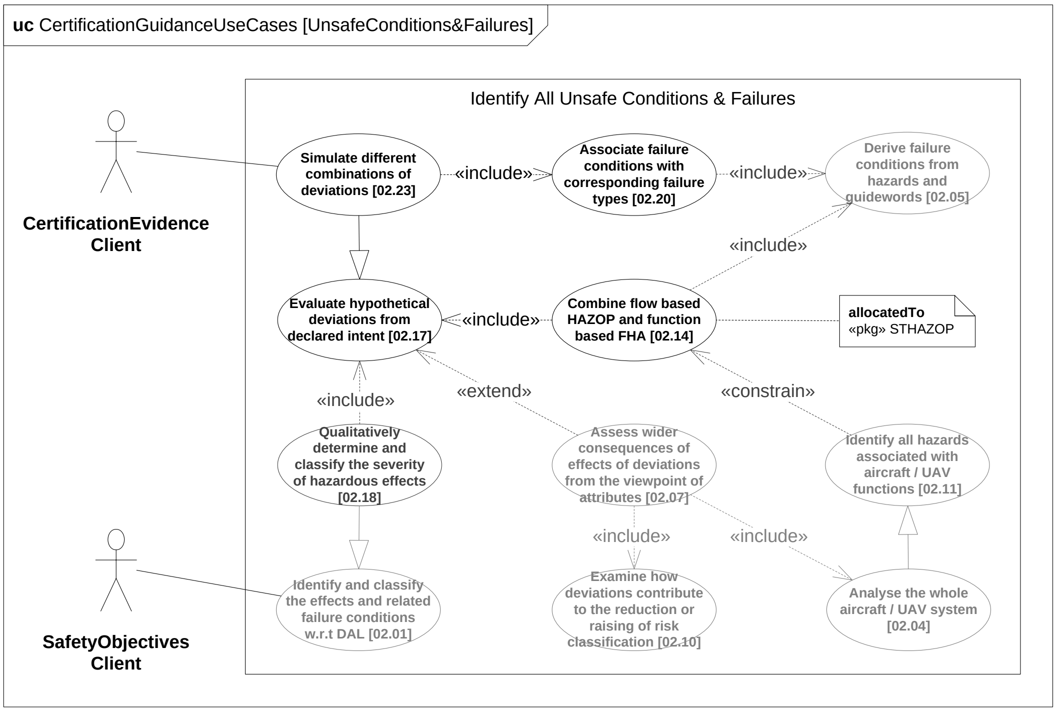

uc CertificationGuidanceUseCases [UnsafeConditions&Failures]
Identify All Unsafe Conditions & Failures
CertificationEvidence Client
SafetyObjectives Client
Simulate different combinations of deviations [02.23]
Associate failure conditions with corresponding failure types [02.20]
Derive failure conditions from hazards and guidewords [02.05]
Evaluate hypothetical deviations from declared intent [02.17]
Combine flow based HAZOP and function based FHA [02.14]
allocatedTo
«pkg» STHAZOP
Qualitatively determine and classify the severity of hazardous effects [02.18]
Assess wider consequences of effects of deviations from the viewpoint of attributes [02.07]
Identify all hazards associated with aircraft / UAV functions [02.11]
Identify and classify the effects and related failure conditions w.r.t DAL [02.01]
Examine how deviations contribute to the reduction or raising of risk classification [02.10]
Analyse the whole aircraft / UAV system [02.04]
«include»
«include»
«include»
«include»
«extend»
«constrain»
«include»
«include»
«include»

Figure 5.4 Use Case Scoping the Hazard Modelling Research Programme

5.2. Stakeholders

5.2.1. Certification Evidence Clients

Evidence is required by design assessors from regulatory bodies such as EASA and the UK Civil Aviation Authority, or for military aircraft and systems, the UK Military Aviation Authority.

5.2.2. Safety Objective Clients

As the safety objective has to be realised by a design authority, this group comprises airworthiness engineers and departments, *system and software safety engineers*, and chief engineers.

5.2.3. HAZOP Team

HAZOP requires the expertise and insight of a relevant group comprising some or all of – *system engineers and designers*, system operators, *experienced pilots* and air traffic controllers, experienced system maintainers, *human factor engineers*, *system safety engineers* and *reliability engineers*, the customer or other authoritative “user”. At least a Study Leader, a Recorder, a System Expert or Designer, and Users of the System are recommended [14].

5.2.4. Engineering

This group represents the fundamental capability required to create a system, including chief engineers, assistant chief engineers, discipline heads, engineering professional leaders, *systems engineers*, *human system engineering*, and other engineering functions relevant to the system of interest – those *highlighted* have had at least some input into this work through the questionnaire.

5.3. Scenario Requirements

5.3.1. Sense And Avoid (SAA) Hazards

SAA Hazard Scenarios [NATO Industrial Advisory Group SG-134]		Scenario 1 (IFR)	Scenario 2 (VFR)
Loss of Separation Provision			
HAZ001	UAV does not comply with separation provision instruction from ATC	Yes	No
HAZ005	Loss of separation provision from ATC	Yes	No
HAZ007	UAV does not maintain safe separation from other aircraft	No	Yes
Separation Provision Error			
HAZ002	UAV incorrectly responds to separation provision instruction from ATC	Yes	No
HAZ006	ATC separation provision error	Yes	No
HAZ008	UAV breaches separation minima	No	Yes
HAZ010	UAV separation minima is breached by other aircraft	No	Yes ☑
Delayed Separation Provision			
HAZ004	Delayed response to separation provision instruction from ATC	Yes	No
HAZ009	DUO (UAV pilot) issues separation provision instruction too late	No	Yes
Intentional Deviation from Separation Provision Instruction			
HAZ003	Excessive number of intentional deviations from separation provision instruction	Yes	No

Table 5.1 Hazards identified in the report of NIAG SG-134 [81]

Sense and Avoid (SAA) is the "hot topic" upon which a significant degree of attention and effort is currently focused, unsurprising given the importance of this as indicated by guidelines [16, 17]. Consequently professional committees, such as NATO Industrial Advisory Group SG-134, have already devoted a fair amount of effort in identifying a number of hazard scenarios, Table 5.1. Of these hazard scenarios the simplest to model is HAZ10, given an assumption of a pre-existing failure of separation provision, and the VFR scenario opens up the interaction to all comers, where only the General Flight Rules of the Rules of the Air might be considered to apply – as such the scenario has the fewest constraints and is therefore likely to represent the most hazardous. Therefore, this hazard forms the basis of the scenario, as it is simply represented yet provides scope for emergent behaviour.

5.3.2. Scenario Generation

Consequently, an outline exemplar describing an air-proximity hazard arising between two air-vehicles has been developed, representing the control structure and system dynamics, and representations of system behaviour across a range of vignettes within the scenario. This poses a simple yet sufficiently challenging problem embodying representative complex behaviour. The purpose of the scenario and model is not to furnish an exact and veritable representation of the putative system and its behaviour, but rather to represent what might be quintessential in the likely behaviour of comparable systems and thereby provide insight into where probable hazards and their actual causes lie, and how these might be measured and informed. With this caveat, the system model represents a two-dimensional interaction space with a simplified implementation of the general flight rules, physics and perception.

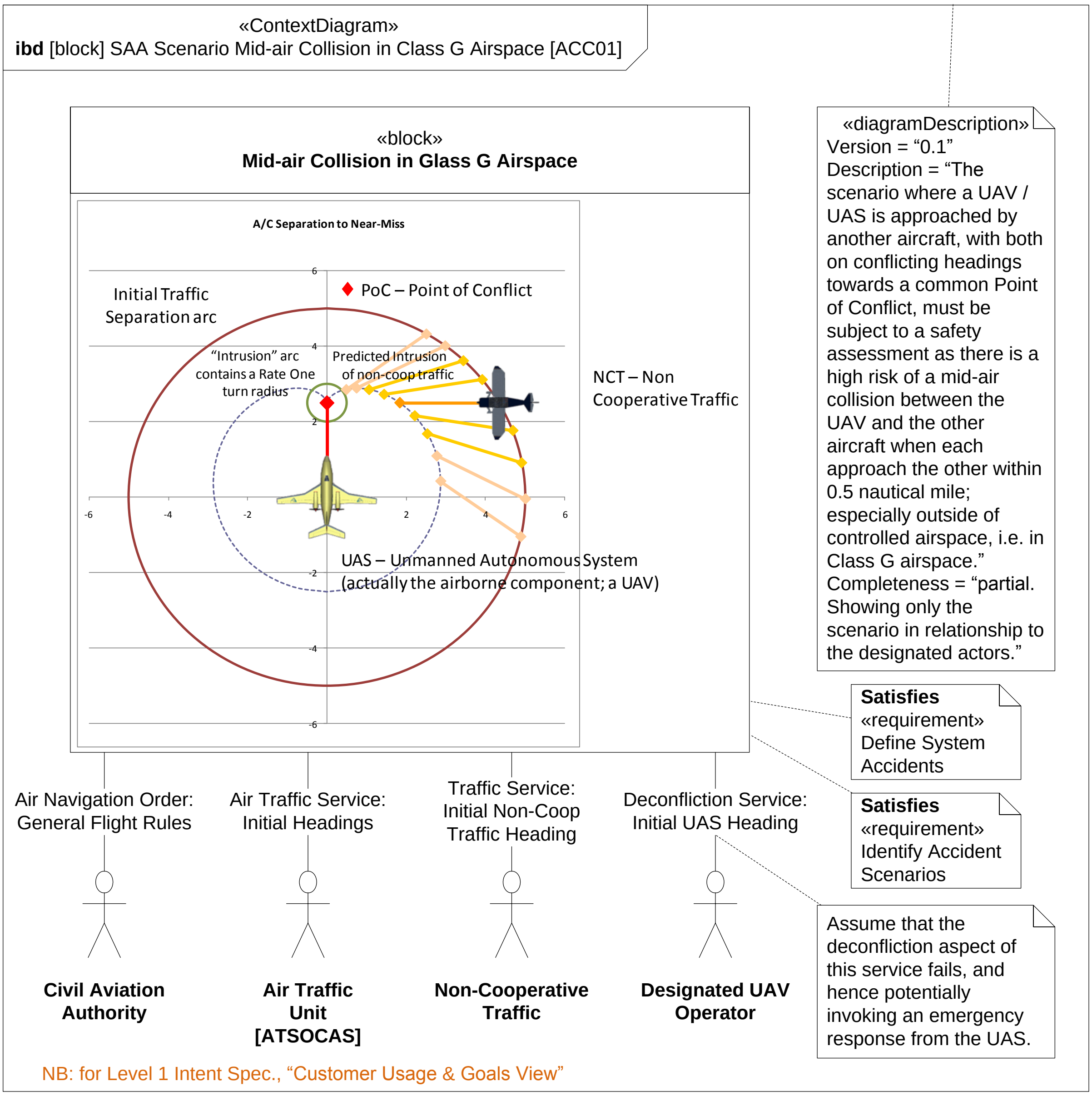

«ContextDiagram»
ibd [block] SAA Scenario Mid-air Collision in Class G Airspace [ACC01]
«block»
Mid-air Collision in Glass G Airspace
A/C Separation to Near-Miss
Initial Traffic Separation arc
PoC – Point of Conflict
"Intrusion" arc contains a Rate One turn radius
Predicted Intrusion of non-coop traffic
NCT – Non Cooperative Traffic
UAS – Unmanned Autonomous System (actually the airborne component; a UAV)
6
4
2
-2
-4
-6
-6
-4
-2
2
4
6
«diagramDescription»
Version = "0.1"
Description = "The scenario where a UAV / UAS is approached by another aircraft, with both on conflicting headings towards a common Point of Conflict, must be subject to a safety assessment as there is a high risk of a mid-air collision between the UAV and the other aircraft when each approach the other within 0.5 nautical mile; especially outside of controlled airspace, i.e. in Class G airspace."
Completeness = "partial. Showing only the scenario in relationship to the designated actors."
Satisfies
«requirement»
Define System Accidents
Satisfies
«requirement»
Identify Accident Scenarios
Air Navigation Order: General Flight Rules
Air Traffic Service: Initial Headings
Traffic Service: Initial Non-Coop Traffic Heading
Deconfliction Service: Initial UAS Heading
Assume that the deconfliction aspect of this service fails, and hence potentially invoking an emergency response from the UAS.
Civil Aviation Authority
Air Traffic Unit [ATSOCAS]
Non-Cooperative Traffic
Designated UAV Operator
NB: for Level 1 Intent Spec., "Customer Usage & Goals View"

Figure 5.5 Mid-air Collision Avoidance Scenario

The basic scenario is as presented in Figure 5.5. As shown in this figure, and for all of the vignettes within the scenario, it is assumed that both vehicles are initially heading toward a common point of conflict (PoC), as indicated by the diamond shaped marker positioned directly ahead of the UAS.

Chapter 6. System Model Design

“Real stupidity beats artificial intelligence every time.”

Terry Pratchett, "Hogfather"

6.1. Modelling Framework

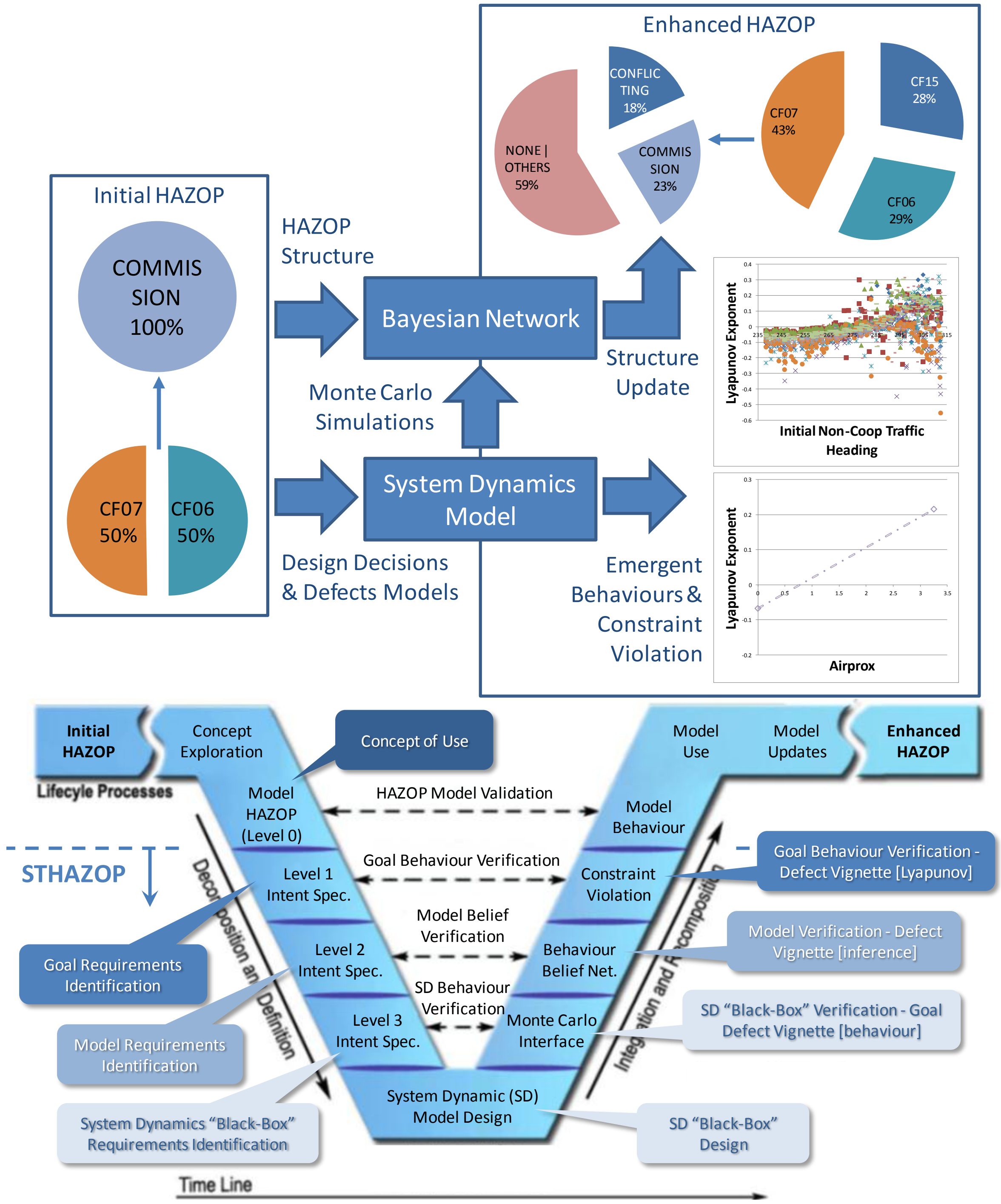


Figure 6.1 Integrating System Dynamics and Inferred Constraint Violation Perspectives

The modelling framework, Figure 6.1, is describable on a number of levels. In terms of context, the framework comprises two mental processes, HAZOP and Systems Theoretic Process Analysis, integrated with two computable models, Goal System Dynamics and Goal Bayesian Belief, the results from which update the HAZOP table with revised belief of causes and effects, Figure 6.2.

In the context of Leveson's Intent Specification [69], the purpose is not to decompose and refine system requirements but rather to formulate the nature of the system safety requirements in terms of a hierarchy of design and stakeholder intent. Thereby, the modelling framework can also be described in this form as a collection of system requirements, process elements and viewpoints, within interactions spanning the system context across four levels of decomposed stakeholder intent. With such a view of design intent and purpose, a hierarchy of lower systems models can be identified and formulated, so as to be described and if possible generalised with the systems modelling language, SysML – ultimate to lead to a meta-model accommodating "Intent".

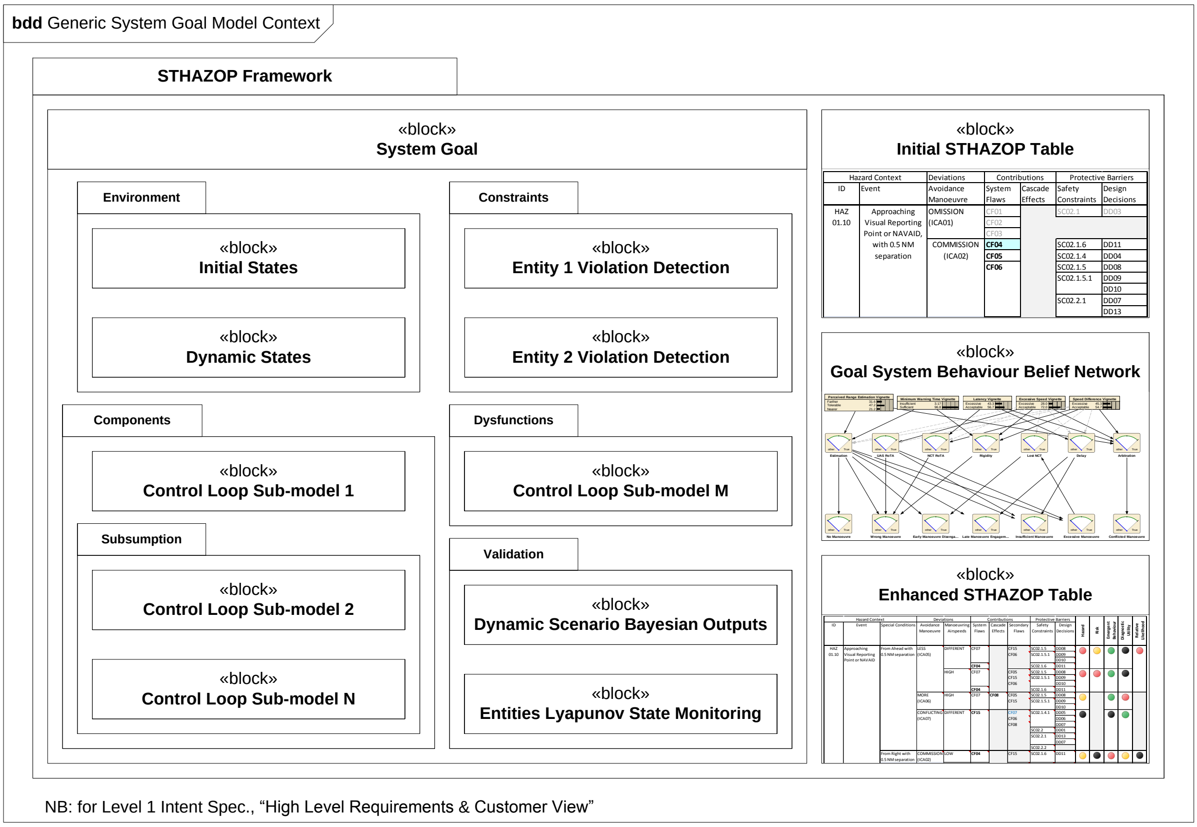


Figure 6.2 Generic System Goal Model within a STHAZOP Framework

So far the generalisation of the model has produced a definition of a generic system goal context, as shown above in Figure 6.2. This generic model indicates where various packages comprising the framework are contained within each other, decomposing the framework into identifiable, discrete and manageable parts. This model identifies the elements taken from the system context that form the system goal, in turn derived from a general form of the Intent Specification framework [86], and positions these alongside the external tool elements to be used to create an enhanced HAZOP output.

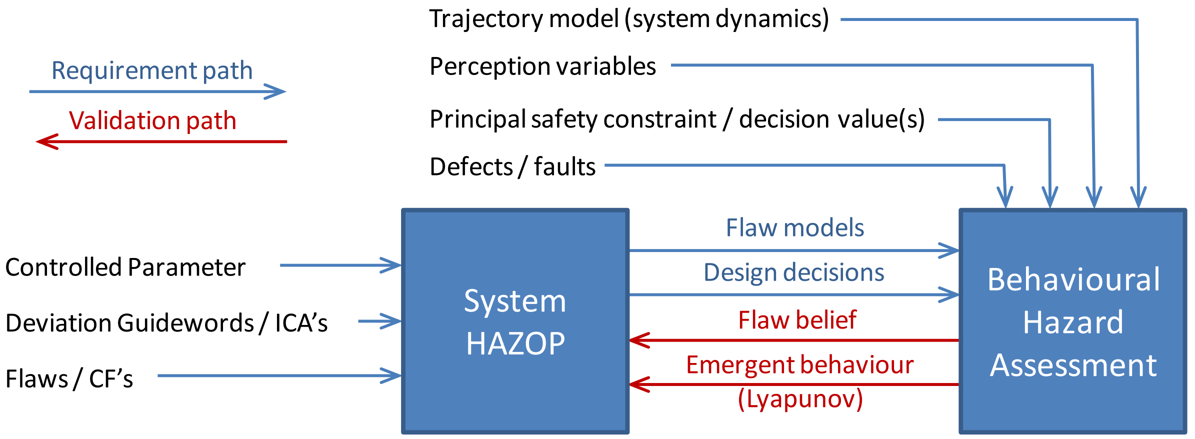

Requirement path
Validation path
Trajectory model (system dynamics)
Perception variables
Principal safety constraint / decision value(s)
Defects / faults
Controlled Parameter
Deviation Guidewords / ICA's
Flaws / CF's
System HAZOP
Flaw models
Design decisions
Flaw belief
Emergent behaviour (Lyapunov)
Behavioural Hazard Assessment

Figure 6.3 A framework compatible with ARP 4754A and DO-178C [97, 101]

The purpose of this form is to succinctly and completely capture all the necessary model components required to construct a system dynamics representation, and also recognise the embodiment of this along with HAZOP and a Bayesian inference methodology – Figure 6.3. Extensions are yet required to incorporate Operators, and with further abstraction, socio-technical elements too.

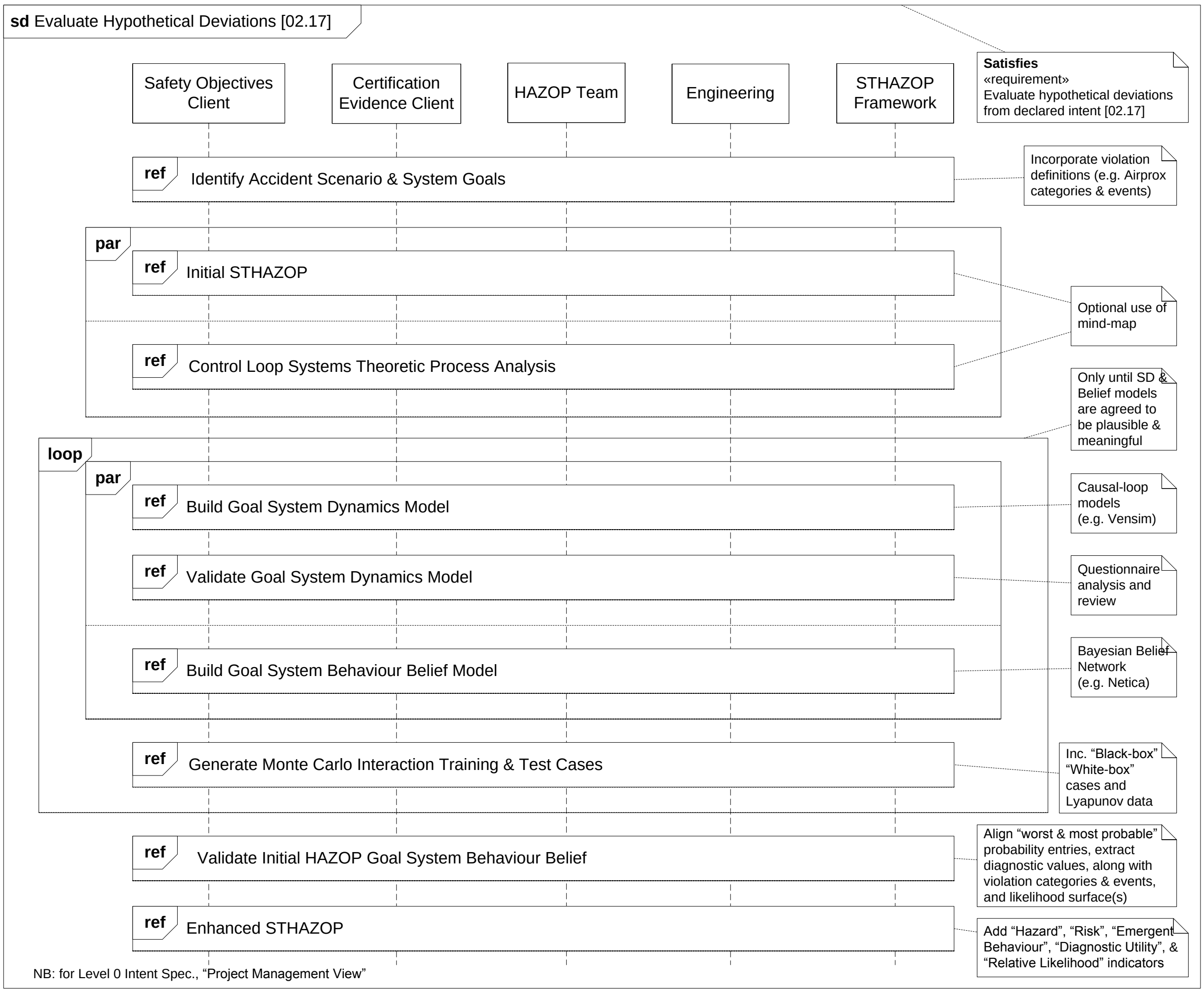

sd Evaluate Hypothetical Deviations [02.17]
Safety Objectives Client
Certification Evidence Client
HAZOP Team
Engineering
STHAZOP Framework
Satisfies «requirement» Evaluate hypothetical deviations from declared intent [02.17]
ref Identify Accident Scenario & System Goals
Incorporate violation definitions (e.g. Airprox categories & events)
par
ref Initial STHAZOP
ref Control Loop Systems Theoretic Process Analysis
Optional use of mind-map
Only until SD & Belief models are agreed to be plausible & meaningful
loop
par
ref Build Goal System Dynamics Model
Causal-loop models (e.g. Vensim)
ref Validate Goal System Dynamics Model
Questionnaire analysis and review
ref Build Goal System Behaviour Belief Model
Bayesian Belief Network (e.g. Netica)
ref Generate Monte Carlo Interaction Training & Test Cases
Inc. "Black-box" "White-box" cases and Lyapunov data
ref Validate Initial HAZOP Goal System Behaviour Belief
Align "worst & most probable" probability entries, extract diagnostic values, along with violation categories & events, and likelihood surface(s)
ref Enhanced STHAZOP
Add "Hazard", "Risk", "Emergent Behaviour", "Diagnostic Utility", & "Relative Likelihood" indicators
NB: for Level 0 Intent Spec., "Project Management View"

Figure 6.4 STHAZOP Framework Sequence Diagram

The final viewpoint upon the modelling framework considers the order in which the various model aspects ought to be created and completed. The sequence of activities and events encompasses the same group of stakeholders as identified in section 5.2, although out of necessity many of the parts representing these stakeholders have been executed by the author in this work – with particular stakeholder participation reserved for the validation of the model behaviour through questionnaire. As can be seen from the sequence diagram, Figure 6.4, certain activities that are to be performed in parallel, such as formulating the initial HAZOP whilst performing a STPA, or building and validating the system dynamics model whilst also building the behaviour belief model, and then to iterate both of these models in the light of new information gained from the analysis and validation of the system dynamics. Having completed the system dynamics modelling, this then is used to formulate the final behaviour belief model and the outputs from this, along with measures of emergent behaviour, are finally used to update the initial HAZOP table. It is conceivable that where more complex systems are to be considered the whole sequence as described above becomes an embedded activity within a larger iterative process, contain numbers of hazard assessments composed within many HAZOP tables and most probably contained within a hierarchy of models. That has not been considered yet.

6.2. Scenario Modelling

Within the sequence of scenario modelling activities, an initial HAZOP table is constructed and the skeleton of the system model is defined in the context of the system goals. This skeleton model is constrained by a minimalist model of the environment.

Additionally obvious dysfunctions, identified during or before the initial HAZOP, and a means by which constraint violation is to be detected, should now be identified and incorporated into the skeleton Goal System Dynamics Model (see Chapter 7). For this scenario, Air Proximity already stands as a framework within which to declare near-misses, with accessible reports providing examples of near-misses; as transcripts, diagrams and analysis, along with other relevant data [111]. The reason that a model of near-misses or safety constraint violation is important, and should be made explicit in the hazard assessment process, is that when it comes to the validation of potentially hazardous behaviour or expressing the perceived internal states of the model in dangerous configurations, near-misses occur more often than collisions and are accidents waiting to happen – and are more often than not documented by those involved in the incident.

Subsequently, having defined the system of interest and accident scenario, a command or control parameter is chosen, which in this case is the term or command “Manoeuvre”, with deviation guidewords then applied to this, with the result producing the initial HAZOP, in Table 6.1.

Hazard Context		Deviations	Contributions		Protective Barriers	
ID	Event	Avoidance Manoeuvre	System Flaws	Cascade Effects	Safety Constraints	Design Decisions
HAZ 01.10	Approaching Visual Reporting Point or NAVAID, with 0.5 NM separation	OMISSION (ICA01)	CF01		SC02.1	DD03
			CF02			
			CF03			
		COMMISSION (ICA02)	**CF04**		SC02.1.6	DD11
			CF05		SC02.1.4	DD04
			CF06		SC02.1.5	DD08
					SC02.1.5.1	DD09
						DD10
					SC02.2.1	DD07
						DD13
			CF07		SC02.1.5	DD08
					SC02.1.5.1	DD09
						DD10
		EARLY (ICA03)	**CF04**		SC02.1.6	DD11
			CF08		SC02.1.3	
					SC02.3	
		LATE (ICA04)	**CF04**		SC02.1.6	DD11
			CF09		SC02.1.2	
			CF10		SC02.1.7	DD12
			CF11			
		LESS (ICA05)	**CF04**		SC02.1.6	DD11
			CF05		SC02.1.4	DD04
			CF10		SC02.1.7	DD12
			CF11			
			CF12			
			CF13			
		MORE (ICA06)	**CF04**	**CF08**	SC02.1.6	DD11
			CF05		SC02.1.4	DD04
			CF14			
		CONFLICTING (ICA07)	**CF04**		SC02.1.6	DD11
			CF15		SC02.1.4.1	DD05
						DD06
						DD07
					SC02.2	DD01
					SC02.2.1	DD07
						DD13
					SC02.2.2	

Table 6.1 Initial HAZOP Table

6.3. Behavioural Intent Modelling

Before the behavioural model is created, the original design intent in the system safety control loop is assessed with an initial HAZOP as described. As previously proposed, an adapted variant of STPA is also performed in conjunction with this HAZOP to identify a number of potential control flaws or defects for each functional element within a generic control loop – with these flaws identified with CF** designations in Figure 6.5. By considering each deviation for each functional element around the control loop in turn, control flaws, safety constraints and mitigating design decisions are identified in a recursive process.

Representations of command failure, delay, and misestimating perceptions are all likely flaws to be incorporated into a system model along with the nature of protective mechanisms identified at this stage. These representations might be termed “Flaw Modelling Requirements”, in that the purpose of the system dynamics and inference models is to explore worst case behaviour, and therefore appropriate representation of these potential defects should be formally captured as modelling requirements. Specific system dysfunction models are also identified at this stage by the HAZOP team, and help to identify the direction of the exploration for hazards with the behavioural model (Figure 6.2). New hazard cases and system dysfunctions may also be revealed later with the output of the behavioural model, and these too can be incorporated later into the HAZOP and modelling requirements.

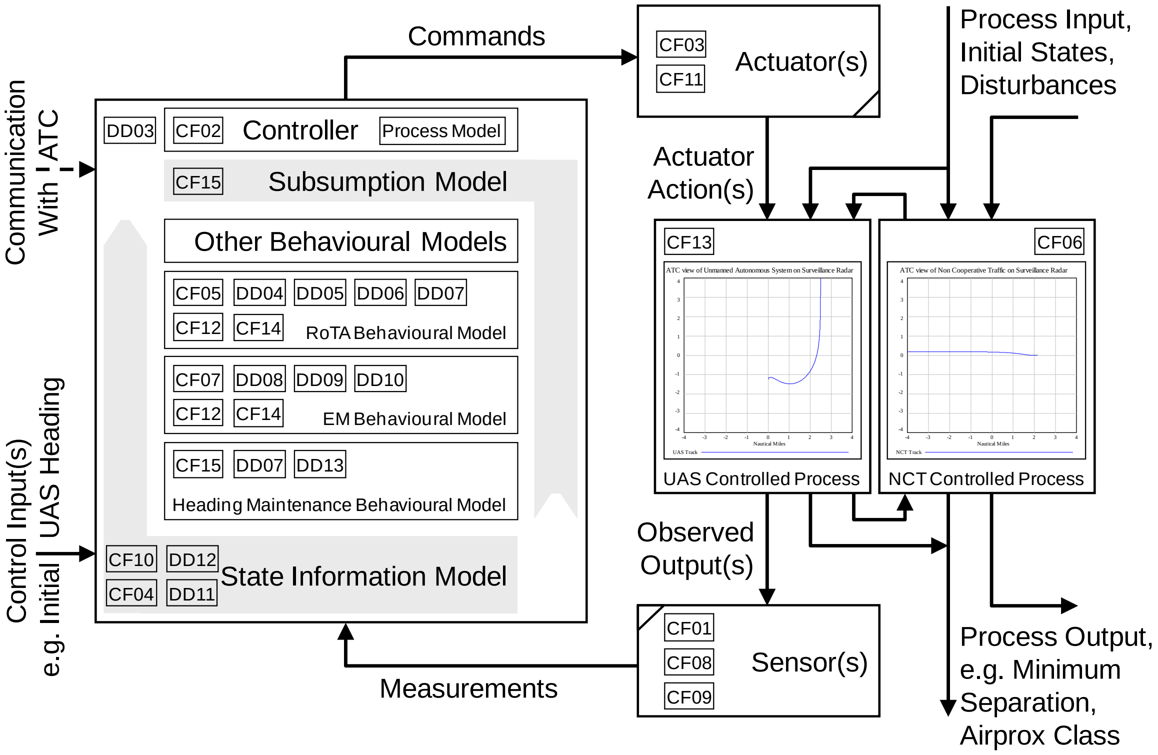


Figure 6.5 **Scenario Collision Avoidance Process Control Loop, adapted from Stringfellow [104, 105]**

Obviously dangerous configurations, such as head-on encounters at speed, might be easily anticipated as a cause of system dysfunction. Other dysfunctions such as arise with misperceived estimation of separation, or a delayed interaction leading to possible entrainment, all of these may not be so obvious or have an effect which is not properly understood at this preliminary stage. Subsequently, equipped with the various models of behaviour, further safety constraints may be devised, internal and external parameters altered, and newly designed protective mechanisms included for a further iteration of exploration. Each safety constraint and design decision is recorded and mapped into the HAZOP table for incorporation in the actual system design – as described for STPA by Owens, et al [86], and Stringfellow, et al [105].

As stated previously, this study employs the abstract behavioural parameter “Manoeuvre” to which guidewords are applied, creating seven deviations or “Inadequate Control Actions in the language of STPA, as described below:

ICA01 – Omission The UAV (Unmanned Aerial Vehicle) does not manoeuvre to avoid the other aircraft.

ICA02 – Commission The UAV manoeuvres in the wrong direction attempting to avoid the other aircraft.

ICA03 – Early The UAV prematurely ceases a manoeuvre to avoid the other aircraft and returns to the given heading.

ICA04 – Late The UAV belatedly initiates a manoeuvre to avoid the other aircraft.

ICA05 – Less The UAV rate of manoeuvre is insufficient to avoid the other aircraft.

ICA06 – More The UAV rate of manoeuvre is excessive in attempting to avoid the other aircraft.

ICA07 – Conflicting The UAV is unable to resolve the conflict between realising the System Goal and avoiding the other aircraft.

In this work the deviation “Omission” has not been explicitly implemented. However, manoeuvres that appear to deviate in not attempting to avoid the other aircraft nevertheless emerge. From consideration of these deviations, fifteen potential control flaws (CF), or system defects have been identified through this STPA-HAZOP process. Later, together with Monte Carlo outputs generated from the system dynamical behaviour, the relationship between these deviations and control flaws are used to configure a flaw belief model – aiding system hazard identification, inference and validation. However, only the control flaws shown in bold in Table 6.1 are legitimate in the context of the current model – incorporating the specific defect representations as described in the scenario previously:

CF01 The UAV fails to detect the other aircraft altogether.

CF02 The UAV does not command a turning manoeuvre from the flight control system (FCS).

CF03 The UAV flight control surfaces do not respond to demands from the FCS.

CF04 The UAV incorrectly estimates the location and direction to the predicted point of conflict.

CF05 The UAV fails to follow established Rules of the Air.

CF06 The other aircraft fails to follow established Rules of the Air.

CF07 The UAV interprets the Rules of the Air too rigidly when closing within 500ft of the other aircraft.

CF08 The UAV loses detection of the other aircraft.

CF09 The UAV fails to detect the other aircraft at sufficient range.

CF10 The UAV delays the commanding of a turning manoeuvre from the flight control system (FCS).

CF11 The UAV flight control surfaces are slow to respond to demands from the FCS.

CF12 The UAV demanded rate of turn is insufficient for the closing angle, speed and range to the other aircraft.

CF13 The UAV maximum demanded rate of turn exceeds the maximum achievable rate of turn.

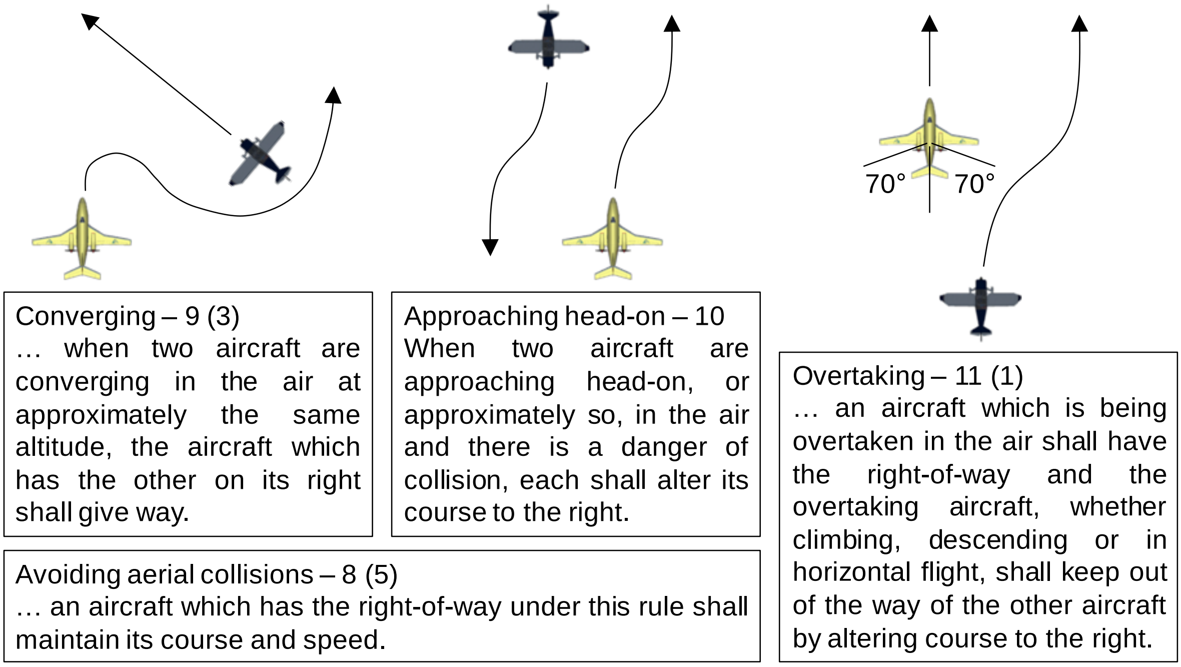
CF14 The UAV demanded rate of turn is excessive for the closing angle, speed and range to the other aircraft.

CF15 The UAV does not arbitrate the competing heading demands appropriately.

6.4. Vignette Modelling

Within the sequence of vignette modelling activities above, the skeleton system dynamics model is configured around the system goals and formal rules, defining the various viewpoints within the model in conjunction with the framework defined previously, in Figure 6.2. Subsequently, definitions of the initial conditions for the simulations (initial separations and heading, speeds, etc.), the motion dynamics, sensing models and the specific flaws as identified previously, are all added. Finally, models to detect safety constraint violations are added to the system dynamics model, along with mechanisms to capture designated Monte Carlo outputs and the calculation of the Lyapunov exponent of the deviated control parameter, in this case "Manoeuvre". The following sections detail the form of the system dynamics control model, describing the implementation of rules and other assumptions.

6.4.1. Scenario Goal Behaviour – Design Decisions



Converging – 9 (3)
… when two aircraft are converging in the air at approximately the same altitude, the aircraft which has the other on its right shall give way.

Approaching head-on – 10
When two aircraft are approaching head-on, or approximately so, in the air and there is a danger of collision, each shall alter its course to the right.

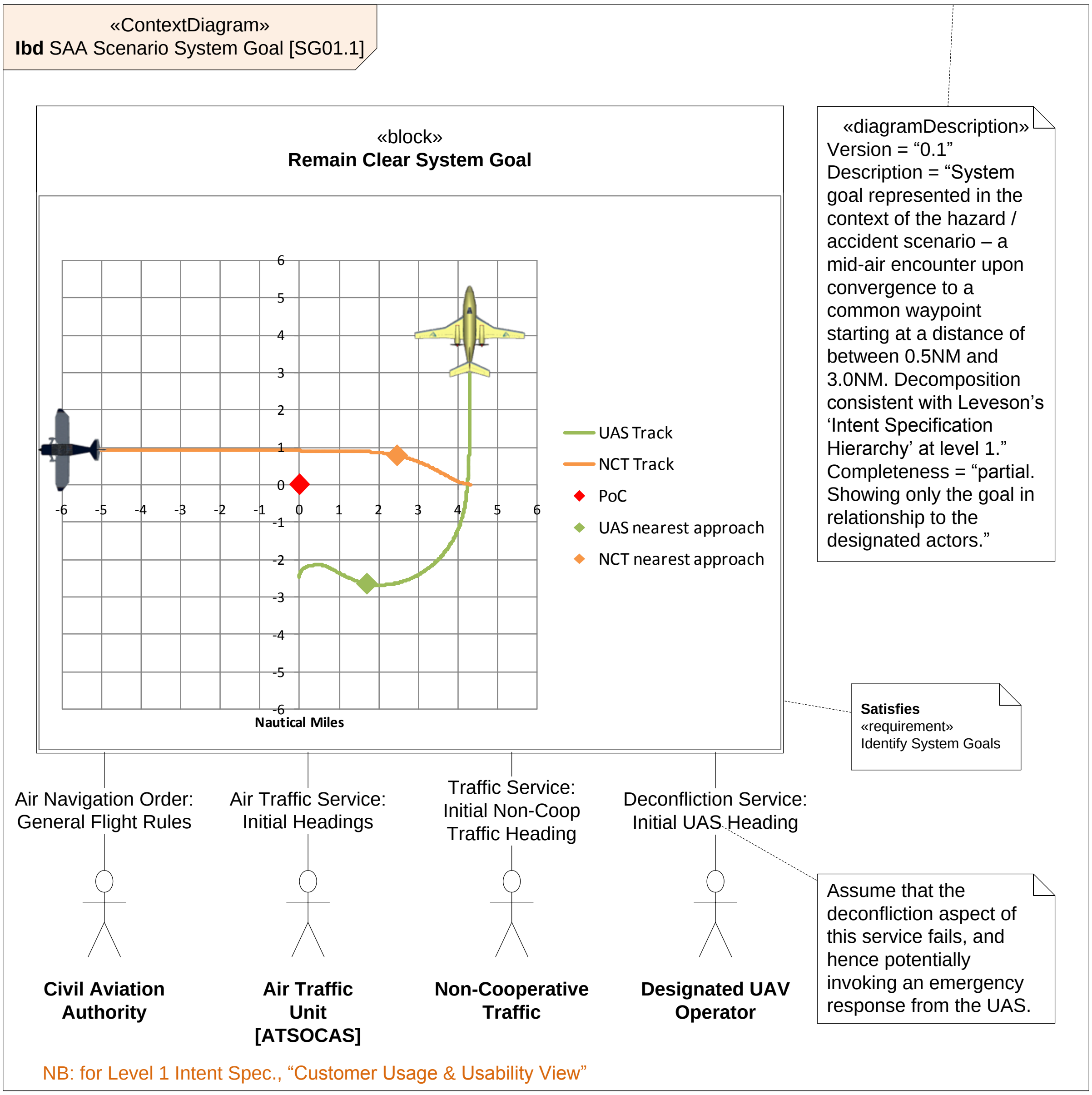
Overtaking – 11 (1)
… an aircraft which is being overtaken in the air shall have the right-of-way and the overtaking aircraft, whether climbing, descending or in horizontal flight, shall keep out of the way of the other aircraft by altering course to the right.

Avoiding aerial collisions – 8 (5)
… an aircraft which has the right-of-way under this rule shall maintain its course and speed.

Figure 6.6 Rule Modelling – The Rules of the Air Regulations 2007 [15]

The representations of both vehicles should embody the Rules of the Air, as described in CAP 393 [15] and illustrated in Figure 6.6. The entity representing the UAV also embodies an additional emergency manoeuvre behaviour that modifies the interpretation of these rules of the air when immediately threatened by the other intruding aircraft with very little room for manoeuvre – in general this will cause the "UAV" to attempt to turn away from the direction of the threat. Also within these rules, when starting to return to course, each vehicle either takes the shortest turn towards its original heading or by default turns left where the angle is approximately equal either way.

Where previously the context diagram representing the scenario, in Figure 5.5, describes a potentially hazardous interaction due to vehicles on a converging course, the goal of the two vehicles avoiding each other, whilst attempting to maintain their respective headings, is summarised in Figure 6.7. From this statement and as shown below, it can be inferred that the dynamics model is required to embody at least two behaviours, or competences, for each entity. The default behaviour is to maintain headings, as that is the goal in the absence of the hazard. Avoidance behaviour is required also.


«ContextDiagram»
Ibd SAA Scenario System Goal [SG01.1]
«block»
Remain Clear System Goal
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
Nautical Miles
UAS Track
NCT Track
PoC
UAS nearest approach
NCT nearest approach
Satisfies
«requirement»
Identify System Goals
Air Navigation Order: General Flight Rules
Air Traffic Service: Initial Headings
Traffic Service: Initial Non-Coop Traffic Heading
Deconfliction Service: Initial UAS Heading
Civil Aviation Authority
Air Traffic Unit [ATSOCAS]
Non-Cooperative Traffic
Designated UAV Operator

«diagramDescription»
Version = "0.1"
Description = "System goal represented in the context of the hazard / accident scenario – a mid-air encounter upon convergence to a common waypoint starting at a distance of between 0.5NM and 3.0NM. Decomposition consistent with Leveson's 'Intent Specification Hierarchy' at level 1."
Completeness = "partial. Showing only the goal in relationship to the designated actors."

Assume that the deconfliction aspect of this service fails, and hence potentially invoking an emergency response from the UAS.

NB: for Level 1 Intent Spec., "Customer Usage & Usability View"

Figure 6.7 Sense and Avoid Scenario System Goal

6.4.2. Default Avoidance Behaviour (Intended Behaviour Models)

Arbitrating between avoidance and maintaining heading an interpretation of the Rules of the Air may be characterised by a table of rate-of-turn values, as shown in Figure 6.8. Assuming also a need to minimise the arbitrary introduction of instability and limit-cycling, this might be represented in a more continuous fashion, as shown in Figure 6.9; facilitating a fuzzy control interpretation of the rule selection logic where this output function surface de-fuzzifies a commanded rate of turn.

IF → AND ↓	destination	to Port	Ahead	to Starboard	Astern
Right of Way, Hold Course (HC)	UAS / NCT on left	No Rate of Turn (0° s^{-1})	No Rate of Turn (0° s^{-1})	Rate One turn to the right (+3° s^{-1})	Rate Two turn to the right (+6° s^{-1})
Give Way (GW)	UAS / NCT head on	Rate Four turn to the right (+24° s^{-1})	Rate Two turn to the right (+6° s^{-1})	Rate Three turn to the right (+12° s^{-1})	Rate Four turn to the right (+24° s^{-1})
	UAS / NCT on right	Rate Four turn to the right (+24° s^{-1})	Rate Four turn to the right (+24° s^{-1})	Rate Two turn to the right (+6° s^{-1})	Rate Three turn to the right (+12° s^{-1})
Right of Way, Hold Course (HC)	UAS / NCT overtaking	Rate One turn to the left (-3° s^{-1})	No Rate of Turn (0° s^{-1})	Rate Three turn to the left (-12° s^{-1})	Rate Two turn to the left (-6° s^{-1})
		Column “A”	Column “B”	Column “C”	Column “D”

Figure 6.8 Scenario Rules of the Air Avoidance Decision Table

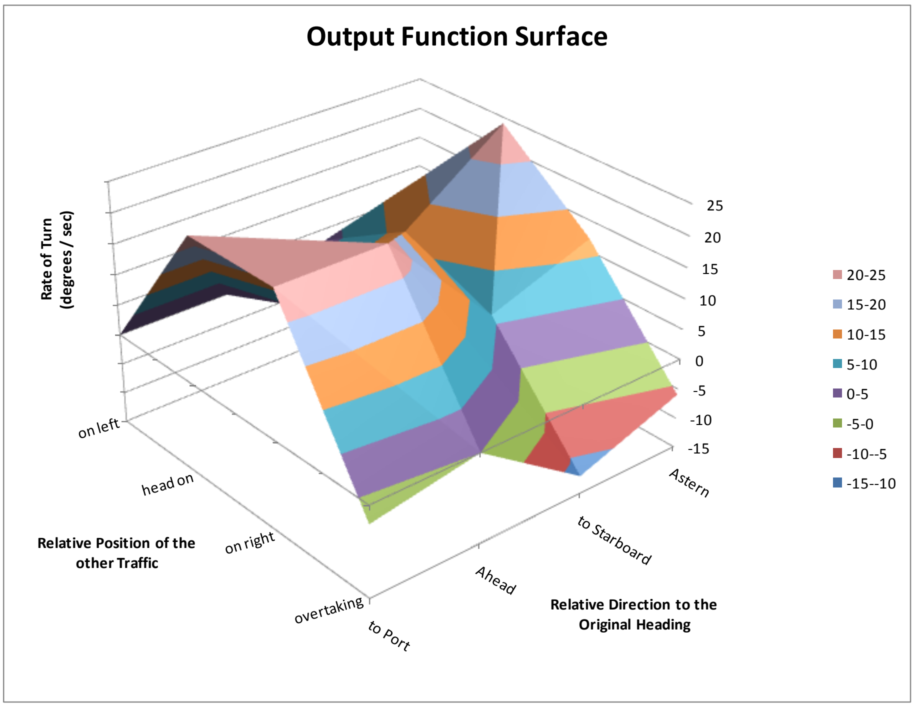
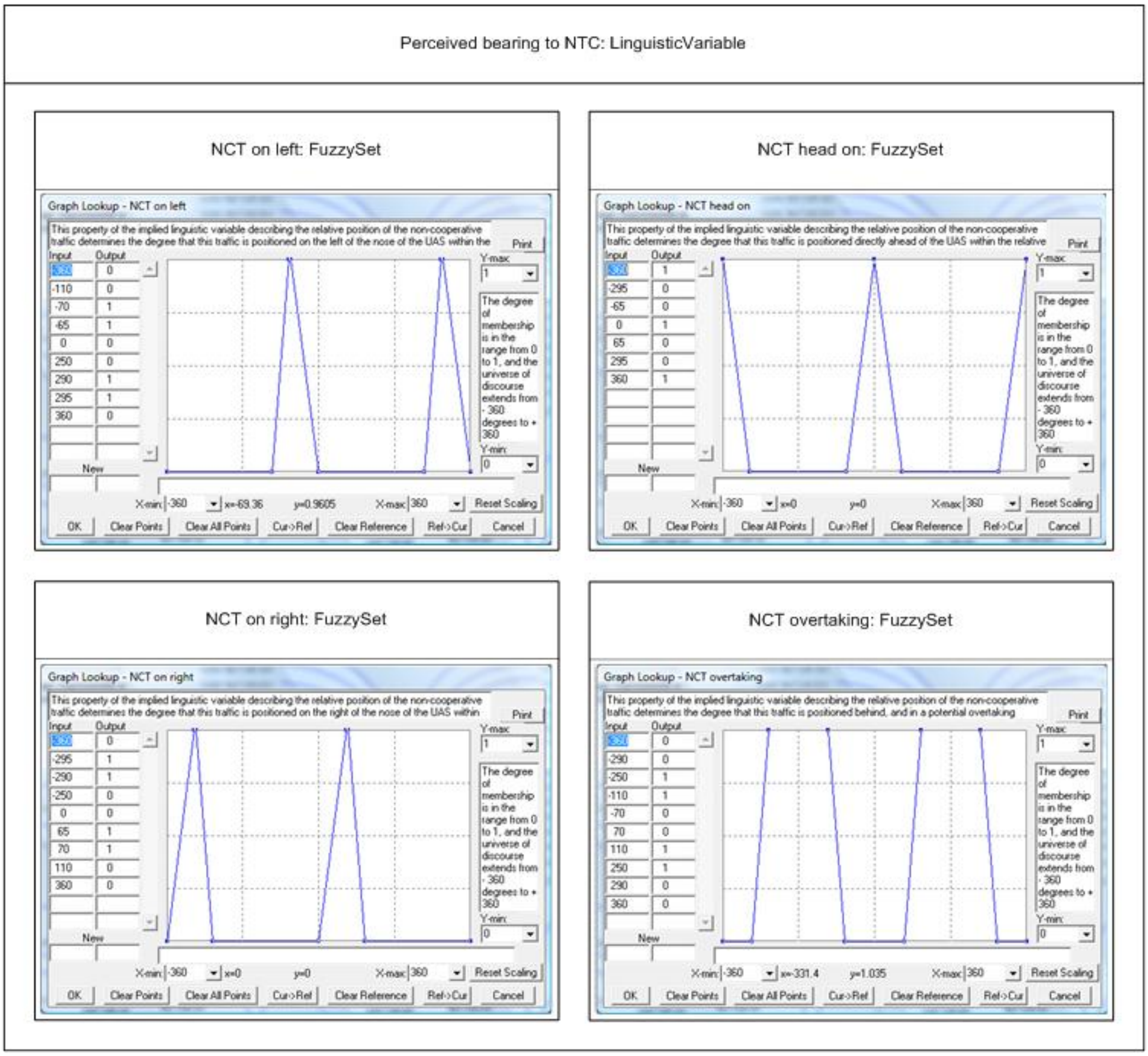


Figure 6.9 Rules of the Air Avoidance Interpretation as Rate of Turn surface



Perceived bearing to NTC: LinguisticVariable

NCT on left: FuzzySet

Graph Lookup - NCT on left

This property of the implied linguistic variable describing the relative position of the non-cooperative traffic determines the degree that this traffic is positioned on the left of the nose of the UAS within the

Input	Output
-360	0
-110	0
-70	1
-65	1
0	0
250	0
290	1
295	1
360	0

Y-max: 1

The degree of membership is in the range from 0 to 1, and the universe of discourse extends from - 360 degrees to + 360

Y-min: 0

X-min: -360 x=-69.36 y=0.9605 X-max: 360

NCT head on: FuzzySet

Graph Lookup - NCT head on

This property of the implied linguistic variable describing the relative position of the non-cooperative traffic determines the degree that this traffic is positioned directly ahead of the UAS within the relative

Input	Output
-360	1
-295	0
-65	0
0	1
65	0
295	0
360	1

Y-max: 1

The degree of membership is in the range from 0 to 1, and the universe of discourse extends from - 360 degrees to + 360

Y-min: 0

X-min: -360 x=0 y=0 X-max: 360

NCT on right: FuzzySet

Graph Lookup - NCT on right

This property of the implied linguistic variable describing the relative position of the non-cooperative traffic determines the degree that this traffic is positioned on the right of the nose of the UAS within

Input	Output
-360	0
-295	1
-290	1
-250	0
0	0
65	1
70	1
110	0
360	0

Y-max: 1

The degree of membership is in the range from 0 to 1, and the universe of discourse extends from - 360 degrees to + 360

Y-min: 0

X-min: -360 x=0 y=0 X-max: 360

NCT overtaking: FuzzySet

Graph Lookup - NCT overtaking

This property of the implied linguistic variable describing the relative position of the non-cooperative traffic determines the degree that this traffic is positioned behind, and in a potential overtaking

Input	Output
-360	0
-290	0
-250	1
-110	1
-70	0
70	0
110	1
250	1
290	0
360	0

Y-max: 1

The degree of membership is in the range from 0 to 1, and the universe of discourse extends from - 360 degrees to + 360

Y-min: 0

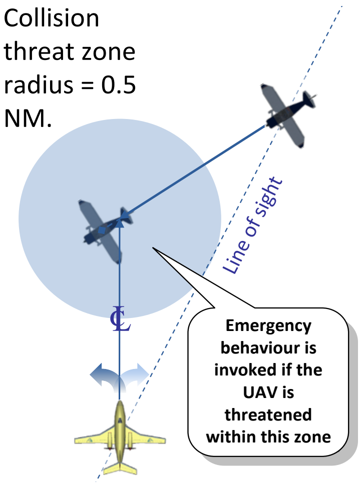
X-min: -360 x=-331.4 y=1.035 X-max: 360

Figure 6.10 Scenario Perceived Bearing Fuzzy Sets

6.4.3. Emergency Avoidance Behaviour (Intended Behaviour Models)

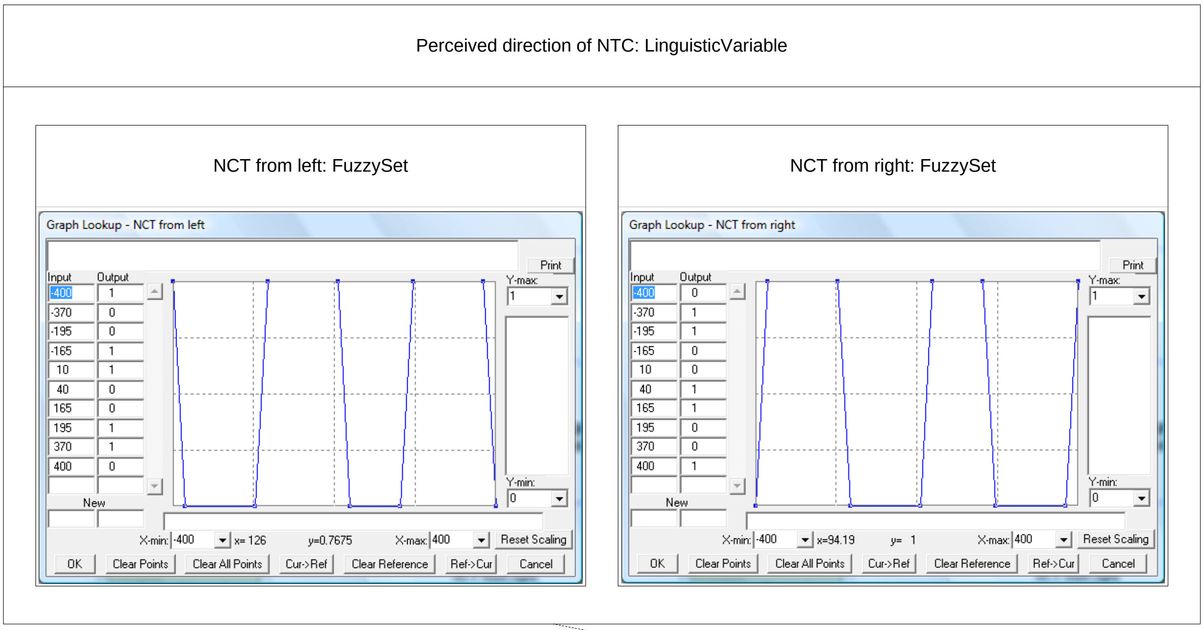
For the emergency behaviour, the behavioural model for the UAV is presented with a threat zone, of radius equivalent to that of a standard rate one turn (3° per second) at the current UAV airspeed and centred directly ahead along the extended Centre Line (℄), as illustrated in Figure 6.11.

If the UAV model perceives that a threat exists, and with sufficient warning time, then an emergency manoeuvre is initiated to turn the UAV away from the centre-line – but this decision point may also be offset left or right by a number of degrees. This decision point or threshold acts as a discontinuity with respect to the default "rules of the air" behaviour, and if this decision behaviour is incorrectly specified then the hazard may not be mitigated but worsened. Otherwise, the UAV complies with the rules as specified in Figure 6.6, and manoeuvres are based solely upon detections along the line of sight to the intruding aircraft. When safe separation is achieved, both model entities turn towards their original headings. The intruder (NCT) always follows the general flight rules as specified.


Collision threat zone radius = 0.5 NM.
Line of sight
℄
Emergency behaviour is invoked if the UAV is threatened within this zone

Where applying the Rules of the Air as represented in Figure 6.8 and Figure 6.9, the required avoidance fuzzy implementation employs linguistic variable terms describing the perceived bearing to the intruder, detailed in Figure 6.10. For the simpler emergency manoeuvre behaviour an alternative two-valued linguistic variable is used to classify the relative threat, left or right, as represented in Figure 6.12, providing values arising with "Track Error Limits" (dys)function providing the manoeuvre direction – Figure 6.13.

Figure 6.11 Alternative Emergency Manoeuvre and Collision Threat Zone Behaviour



Graph Lookup - NCT from left

Input	Output
-400	1
-370	0
-195	0
-165	1
10	1
40	0
165	0
195	1
370	1
400	0

Y-max: 1 Y-min: 0
X-min: -400 x= 126 y=0.7675 X-max: 400

Graph Lookup - NCT from right

Input	Output
-400	0
-370	1
-195	1
-165	0
10	0
40	1
165	1
195	0
370	0
400	1

Y-max: 1 Y-min: 0
X-min: -400 x=94.19 y= 1 X-max: 400

Note that "ahead" for the purposes of an emergency manoeuvre decision, in this version, has been offset to the right by 25 degrees.

Figure 6.12 Scenario Perceived Direction of Threat

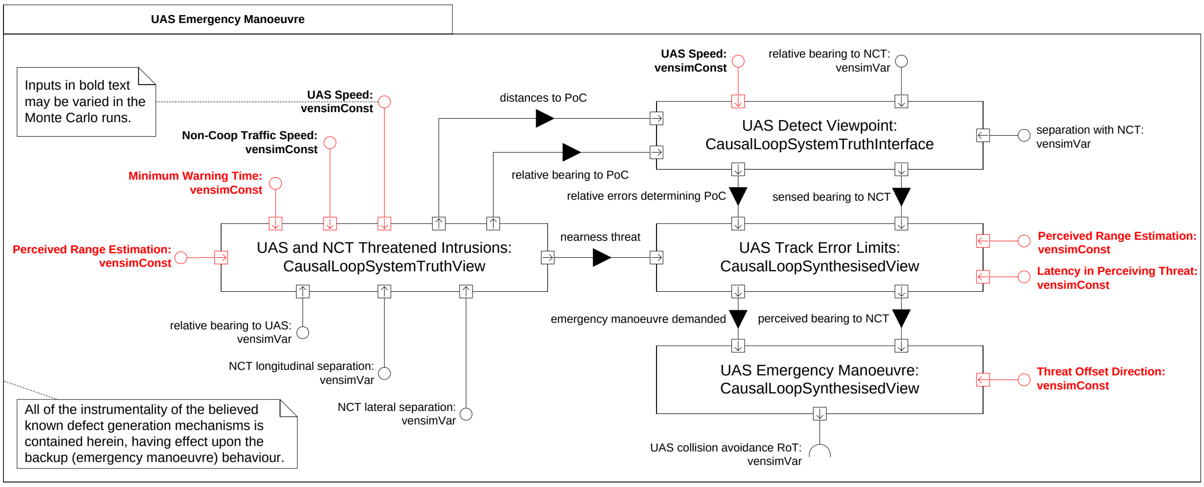

UAS Emergency Manoeuvre
Inputs in bold text may be varied in the Monte Carlo runs.
UAS Speed: vensimConst
Non-Coop Traffic Speed: vensimConst
Minimum Warning Time: vensimConst
Perceived Range Estimation: vensimConst
UAS and NCT Threatened Intrusions: CausalLoopSystemTruthView
relative bearing to UAS: vensimVar
NCT longitudinal separation: vensimVar
NCT lateral separation: vensimVar
All of the instrumentality of the believed known defect generation mechanisms is contained herein, having effect upon the backup (emergency manoeuvre) behaviour.
distances to PoC
relative bearing to PoC
nearness threat
UAS Speed: vensimConst
relative bearing to NCT: vensimVar
UAS Detect Viewpoint: CausalLoopSystemTruthInterface
separation with NCT: vensimVar
relative errors determining PoC
sensed bearing to NCT
UAS Track Error Limits: CausalLoopSynthesisedView
Perceived Range Estimation: vensimConst
Latency in Perceiving Threat: vensimConst
emergency manoeuvre demanded
perceived bearing to NCT
UAS Emergency Manoeuvre: CausalLoopSynthesisedView
Threat Offset Direction: vensimConst
UAS collision avoidance RoT: vensimVar

Figure 6.13 Scenario Emergency Manoeuvre Internal Block Diagram (dysfunctions in red)

6.4.4. Default Heading Maintenance Behaviour (Intended Behaviour Models)

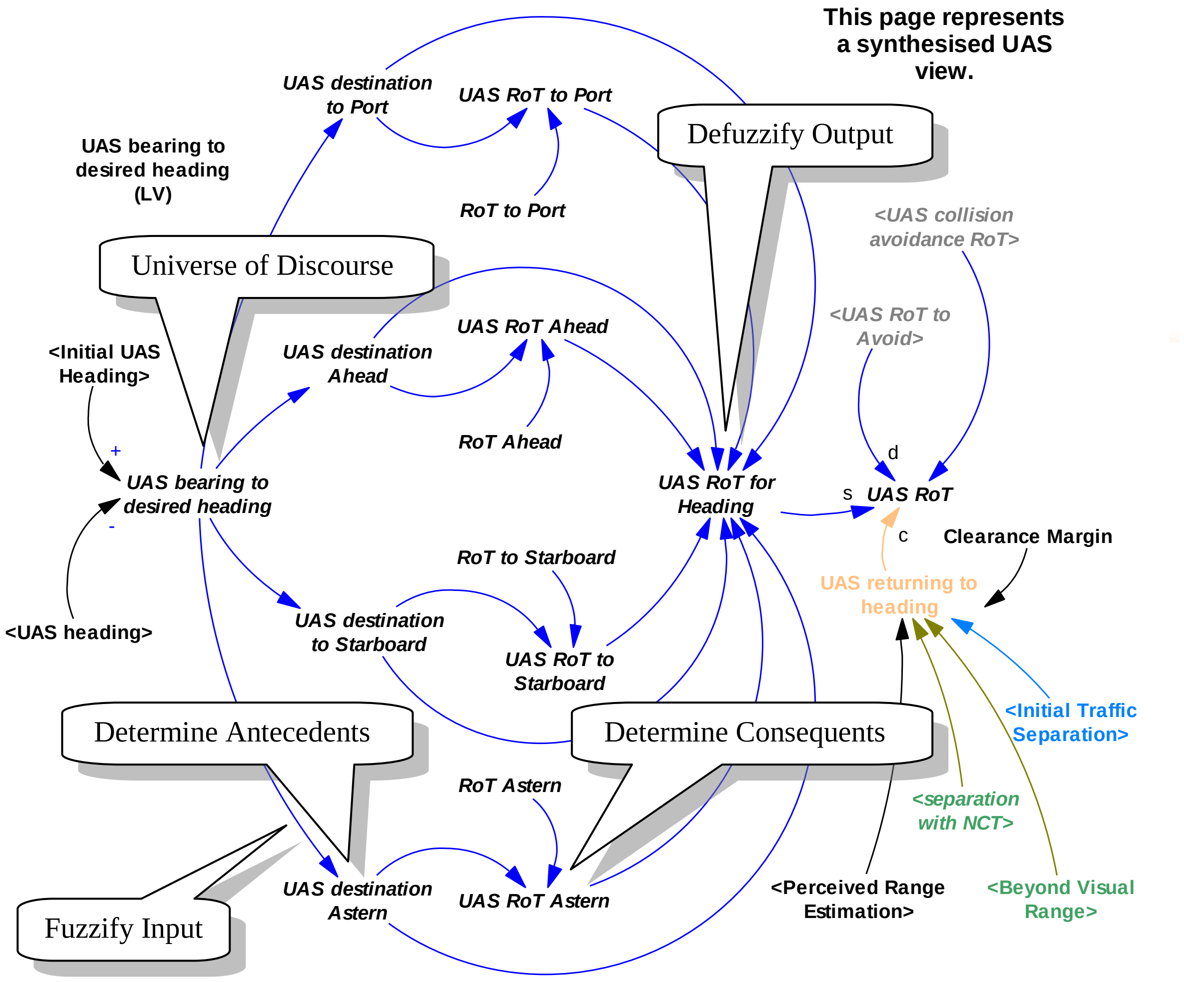
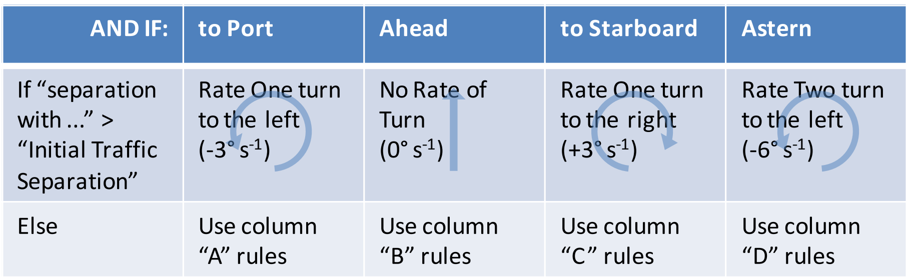


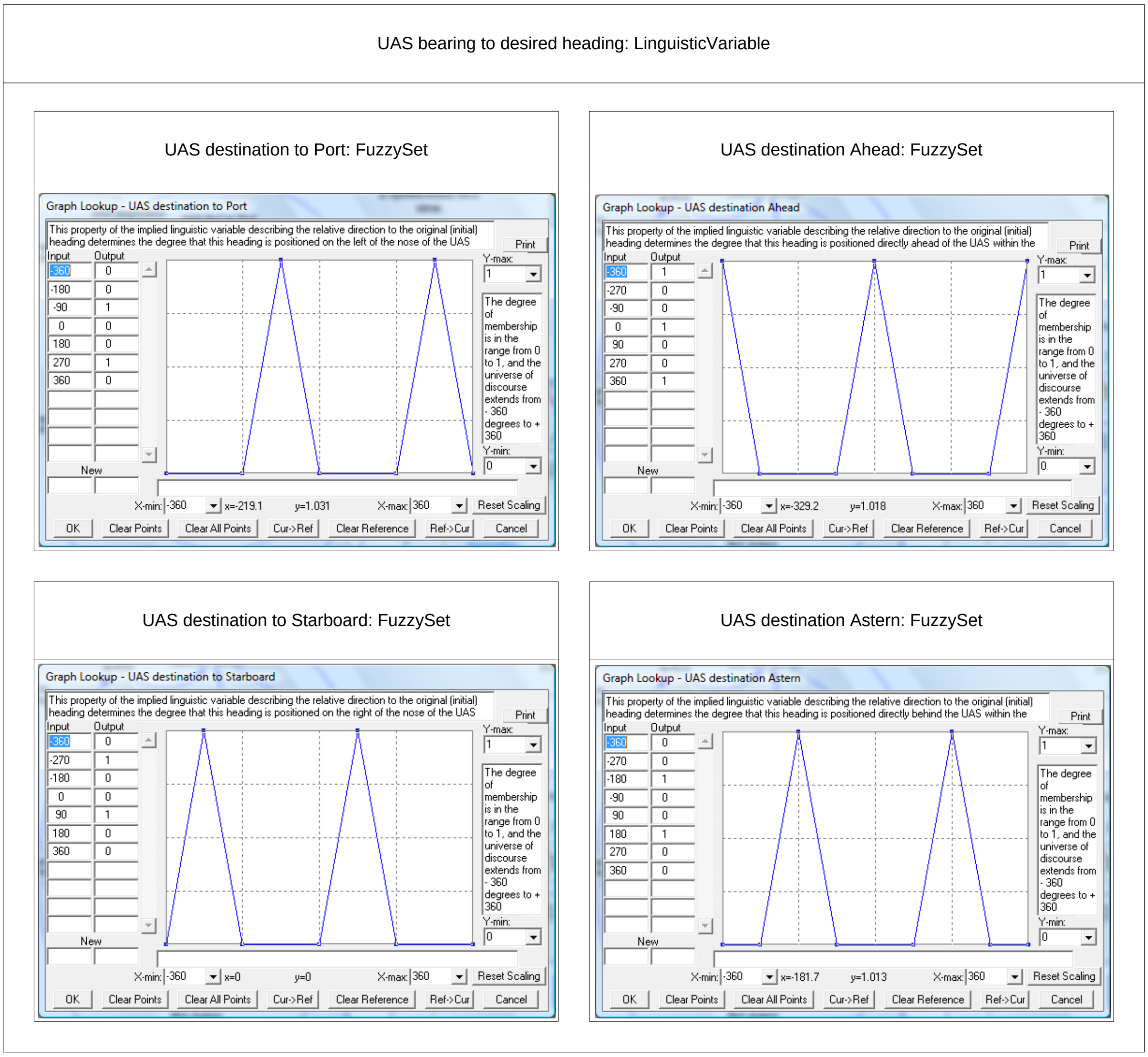
Figure 6.14 UAS Heading Maintenance Fuzzy Logic as "Causal Flow" Diagram

The default heading maintenance logic and controller is implemented as a simple fuzzy logic system, as outlined above in Figure 6.14. A number of the rule systems and controllers within the model are implemented as either single or two input fuzzy logic controllers, as described previously. In this case Figure 6.15 and Figure 6.16 respectively represent the single input heading maintenance control model defuzzification value table, and the input linguistic variable within the "Universe of Discourse" of bearing to desired heading – both used to steer the UAS towards its original heading.



AND IF:	to Port	Ahead	to Starboard	Astern
If "separation with ..." > "Initial Traffic Separation"	Rate One turn to the left (-3° s^{-1})	No Rate of Turn (0° s^{-1})	Rate One turn to the right (+3° s^{-1})	Rate Two turn to the left (-6° s^{-1})
Else	Use column "A" rules	Use column "B" rules	Use column "C" rules	Use column "D" rules

Figure 6.15 Scenario Heading Maintenance Decision Table (Column Rules as in Figure 6.8)



UAS bearing to desired heading: LinguisticVariable

UAS destination to Port: FuzzySet

Graph Lookup - UAS destination to Port

This property of the implied linguistic variable describing the relative direction to the original (initial) heading determines the degree that this heading is positioned on the left of the nose of the UAS

Input	Output
-360	0
-180	0
-90	1
0	0
180	0
270	1
360	0

The degree of membership is in the range from 0 to 1, and the universe of discourse extends from - 360 degrees to + 360

Y-max: 1
Y-min: 0
X-min: -360 x=-219.1 y=1.031 X-max: 360

UAS destination Ahead: FuzzySet

Graph Lookup - UAS destination Ahead

This property of the implied linguistic variable describing the relative direction to the original (initial) heading determines the degree that this heading is positioned directly ahead of the UAS within the

Input	Output
-360	1
-270	0
-90	0
0	1
90	0
270	0
360	1

The degree of membership is in the range from 0 to 1, and the universe of discourse extends from - 360 degrees to + 360

Y-max: 1
Y-min: 0
X-min: -360 x=-329.2 y=1.018 X-max: 360

UAS destination to Starboard: FuzzySet

Graph Lookup - UAS destination to Starboard

This property of the implied linguistic variable describing the relative direction to the original (initial) heading determines the degree that this heading is positioned on the right of the nose of the UAS

Input	Output
-360	0
-270	1
-180	0
0	0
90	1
180	0
360	0

The degree of membership is in the range from 0 to 1, and the universe of discourse extends from - 360 degrees to + 360

Y-max: 1
Y-min: 0
X-min: -360 x=0 y=0 X-max: 360

UAS destination Astern: FuzzySet

Graph Lookup - UAS destination Astern

This property of the implied linguistic variable describing the relative direction to the original (initial) heading determines the degree that this heading is positioned directly behind the UAS within the

Input	Output
-360	0
-270	0
-180	1
-90	0
90	0
180	1
270	0
360	0

The degree of membership is in the range from 0 to 1, and the universe of discourse extends from - 360 degrees to + 360

Y-max: 1
Y-min: 0
X-min: -360 x=-181.7 y=1.013 X-max: 360

Figure 6.16 Scenario Bearing to Desired Heading Fuzzy Sets

A systematic model such as this lends itself to a definition as a standard part within a library; configured with appropriate fuzzy membership functions and constants where required to fit a particular role. Fuzzy logic and control generally fits well as a generic component in a hierarchical, behavioural and competence scheme of control and management; configurable for a wide range of representations and uses. However, no succinct model representation of fuzzy logic has been created for SysML. Arguably the increasing prevalence of fuzzy logic, along with neural networks, genetic algorithms and Bayesian networks, suggests that standard succinct models for all of these are required to be expressed and made clear in the wider systems engineering context – currently in SysML this amounts to a block labelled "fuzzy controller", etc. It will remain difficult to automate model construction without consistent and proper representations of generic items such as these.

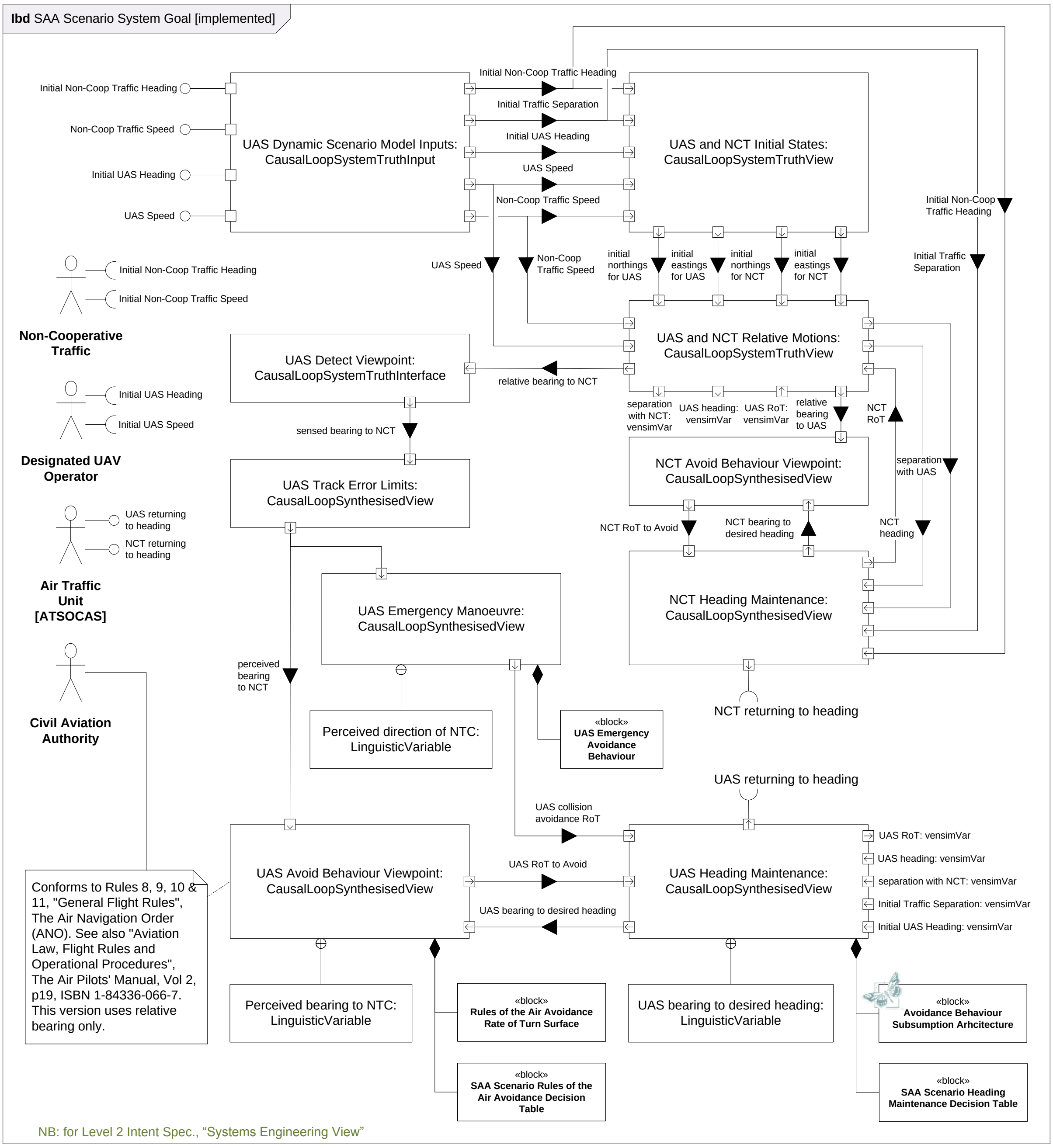

ibd SAA Scenario System Goal [implemented]
Initial Non-Coop Traffic Heading
Non-Coop Traffic Speed
Initial UAS Heading
UAS Speed
UAS Dynamic Scenario Model Inputs: CausalLoopSystemTruthInput
Initial Non-Coop Traffic Heading
Initial Traffic Separation
Initial UAS Heading
UAS Speed
Non-Coop Traffic Speed
UAS and NCT Initial States: CausalLoopSystemTruthView
Initial Non-Coop Traffic Heading
Initial Traffic Separation
UAS Speed
Non-Coop Traffic Speed
initial northings for UAS
initial eastings for UAS
initial northings for NCT
initial eastings for NCT
Initial Non-Coop Traffic Heading
Initial Non-Coop Traffic Speed
Non-Cooperative Traffic
UAS and NCT Relative Motions: CausalLoopSystemTruthView
UAS Detect Viewpoint: CausalLoopSystemTruthInterface
relative bearing to NCT
Initial UAS Heading
Initial UAS Speed
Designated UAV Operator
sensed bearing to NCT
separation with NCT: vensimVar
UAS heading: vensimVar
UAS RoT: vensimVar
relative bearing to UAS
NCT RoT
UAS Track Error Limits: CausalLoopSynthesisedView
NCT Avoid Behaviour Viewpoint: CausalLoopSynthesisedView
separation with UAS
UAS returning to heading
NCT returning to heading
Air Traffic Unit [ATSOCAS]
NCT RoT to Avoid
NCT bearing to desired heading
NCT heading
UAS Emergency Manoeuvre: CausalLoopSynthesisedView
NCT Heading Maintenance: CausalLoopSynthesisedView
Civil Aviation Authority
perceived bearing to NCT
Perceived direction of NTC: LinguisticVariable
«block» UAS Emergency Avoidance Behaviour
NCT returning to heading
UAS returning to heading
UAS collision avoidance RoT
UAS Avoid Behaviour Viewpoint: CausalLoopSynthesisedView
UAS RoT to Avoid
UAS bearing to desired heading
UAS Heading Maintenance: CausalLoopSynthesisedView
UAS RoT: vensimVar
UAS heading: vensimVar
separation with NCT: vensimVar
Initial Traffic Separation: vensimVar
Initial UAS Heading: vensimVar
Conforms to Rules 8, 9, 10 & 11, "General Flight Rules", The Air Navigation Order (ANO). See also "Aviation Law, Flight Rules and Operational Procedures", The Air Pilots' Manual, Vol 2, p19, ISBN 1-84336-066-7. This version uses relative bearing only.
Perceived bearing to NTC: LinguisticVariable
«block» Rules of the Air Avoidance Rate of Turn Surface
UAS bearing to desired heading: LinguisticVariable
«block» Avoidance Behaviour Subsumption Arhcitecture
«block» SAA Scenario Rules of the Air Avoidance Decision Table
«block» SAA Scenario Heading Maintenance Decision Table
NB: for Level 2 Intent Spec., "Systems Engineering View"

Figure 6.17 Scenario System Goal Internal Block Diagram, as implemented

The entire system dynamics simulation model has been created as a Vensim™ model and hence the "CausalLoop" block type referenced in Figure 6.17, above. This illustration shows the internal relationship between all of the major components in the model, albeit with some simplification. The model itself is layered into "SystemTruth" views, representing the underlying physical motions and actual bearings and ranges, and "Synthesised" views, representing the internal system logic to which error and dysfunction may be introduced. Chapter 7 continues to describe the physical model infrastructure that constrains the "truthful" behaviour of the above system, the model elements that introduce or describe the Control Flaws, and against these the descriptions of the safety constraints, including the means to capture any safety constraint violations as they arise during simulations.

Chapter 7. System Constraints Model

"Where is the boundary, that encompasses the smallest number of components, within which dynamic behaviour under study is generated?"

Jay W Forrester, "Principles of Systems"

7.1. System State Modelling

7.1.1. Initial State – Defined Environment (Initial Conditions Model)

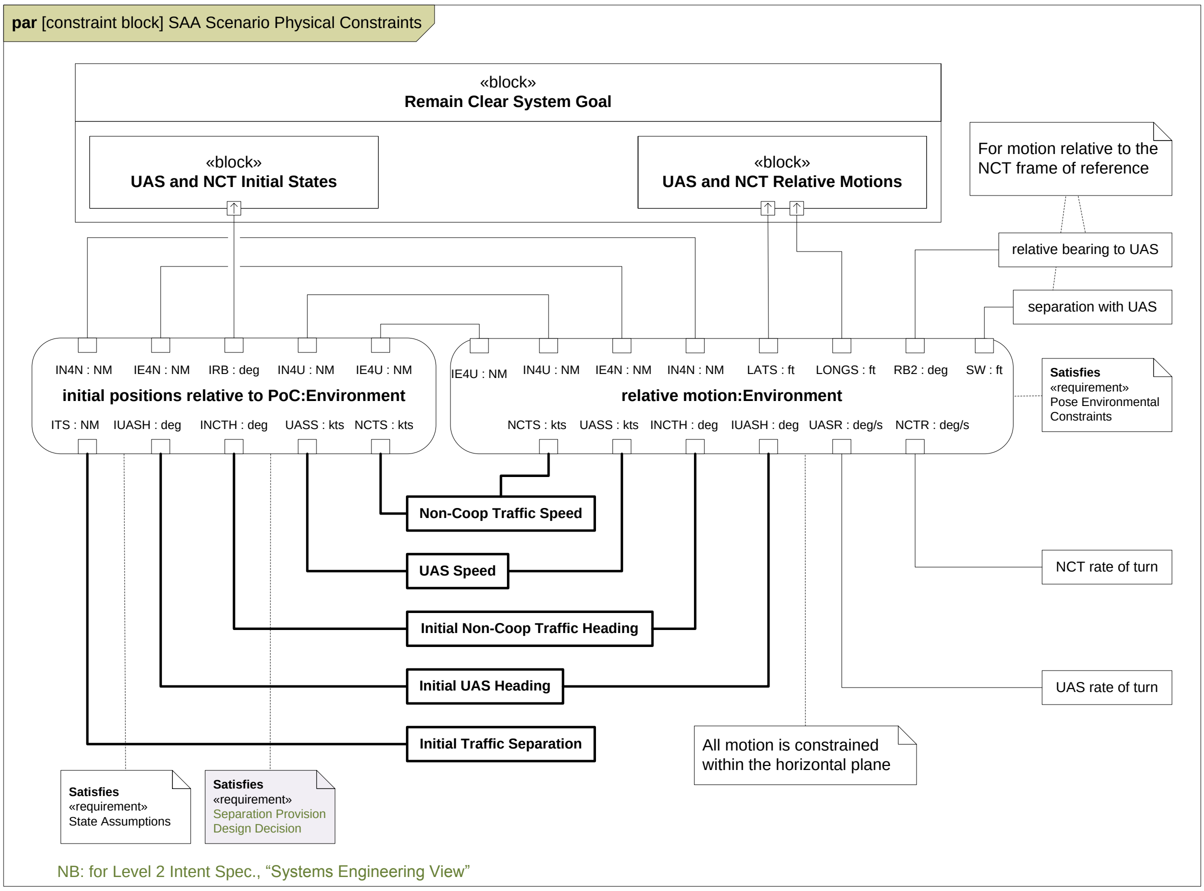


Figure 7.1 UAS and NCT Initial States Specific (Constraints) Parametric Model

Constraint models define the limits of behaviour of the system dynamics simulations. Here these models are defined both in adapted SysML parametric diagrams, and also as individual Constraint blocks. The first of these constraints describes part of the model initial conditions and relative states – defining some of the initial physical states of the model – as shown in Figure 7.1 above.

Expressions, such as those shown overleaf, are used to describe the simulation model equations that propagate state values throughout the model. In general these expressions may take two distinct forms;

«constraint» representing simple expressions and assignments; and

«heuristic» representing conditional assignments, alternatives and rules.

$$\boldsymbol{UASInitialDistanceToTravel} = \frac{\boldsymbol{ITS} \cdot \dfrac{\boldsymbol{UASS}}{\boldsymbol{NCTS}}}{\boldsymbol{1} + \dfrac{\boldsymbol{UASS}}{\boldsymbol{NCTS}}}$$

…

$$\boldsymbol{IRB} = \begin{cases} \boldsymbol{IRA}, & \boldsymbol{BearingRelativeToPoC} > \boldsymbol{180} \\ -\boldsymbol{IRA}, & \boldsymbol{otherwise} \end{cases}$$

For example, the partial initial state expressions as presented above are presented as equivalent constraint definitions in Figure 7.2, the first being an example of a «constraint» and the other an example of a representative «heuristic». As a further illustration, the implementation of these two example expressions within the system dynamics model, with equations (323) and (102), are detailed in the small excerpt taken from the Vensim model simulation code, and shown further below.

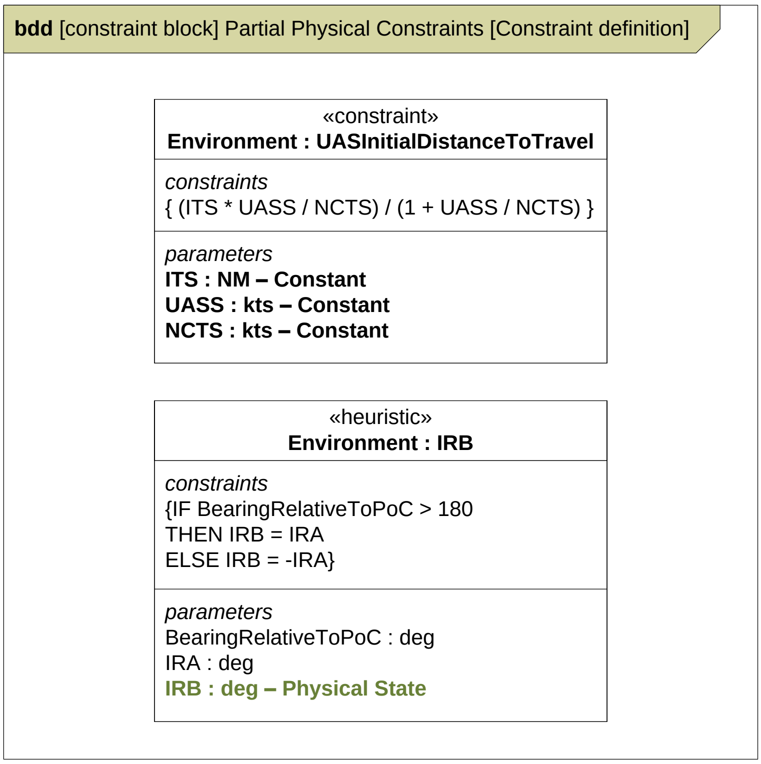
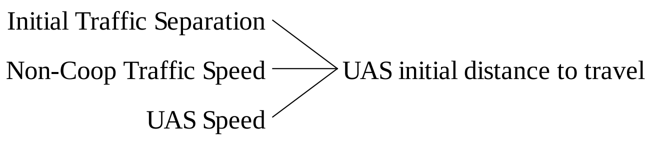
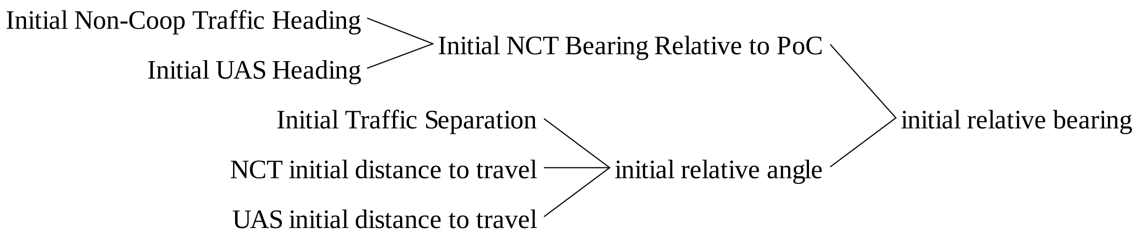

bdd [constraint block] Partial Physical Constraints [Constraint definition]
«constraint»
Environment : UASInitialDistanceToTravel
constraints
{ (ITS * UASS / NCTS) / (1 + UASS / NCTS) }
parameters
ITS : NM – Constant
UASS : kts – Constant
NCTS : kts – Constant
«heuristic»
Environment : IRB
constraints
{IF BearingRelativeToPoC > 180
THEN IRB = IRA
ELSE IRB = -IRA}
parameters
BearingRelativeToPoC : deg
IRA : deg
IRB : deg – Physical State

Figure 7.2 Partial Parametric Definitions for Initial State Constraints

```
(323) UAS initial distance to travel = (Initial Traffic Separation * UAS Speed / "Non-Coop Traffic Speed") / (1 + UAS Speed / "Non-Coop Traffic Speed")
(102) initial relative bearing = IF THEN ELSE (Initial NCT Bearing Relative to PoC > 180, initial relative angle, -initial relative angle)
```

Note also that within the parametric block "initial positions relative to the PoC" (Figure 7.1) there are also additional expressions not shown here for the sake of brevity. However, some of the underlying connectivity is shown over the page, detailing the immediate source terms for the above constraints.




Initial Non-Coop Traffic Heading
Initial UAS Heading
Initial NCT Bearing Relative to PoC
Initial Traffic Separation
NCT initial distance to travel
UAS initial distance to travel
initial relative angle
initial relative bearing

More generally, a number of types and forms of constraints appear to be an essential generic aspect of the system dynamics model definition, and each of these specific types ought to be applied at least once each within the model definition process. Together, five generic types of constraint have been identified in the course of this work:

a) Physical constraints including state initialisation, motions, transforms, etc. – Figure 7.3;
b) Dysfunctions and defect properties – Figure 7.4;
c) Safety constraint violations as proximity, exceedance, near-miss, etc. – Figure 7.5;
d) Emergent property monitoring, with Lyapunov Exponent, entropy or other – Figure 7.16; and
e) Test and training case (Monte Carlo) data capture interface – Figure 8.5 to Figure 8.9; as depicted below and later.

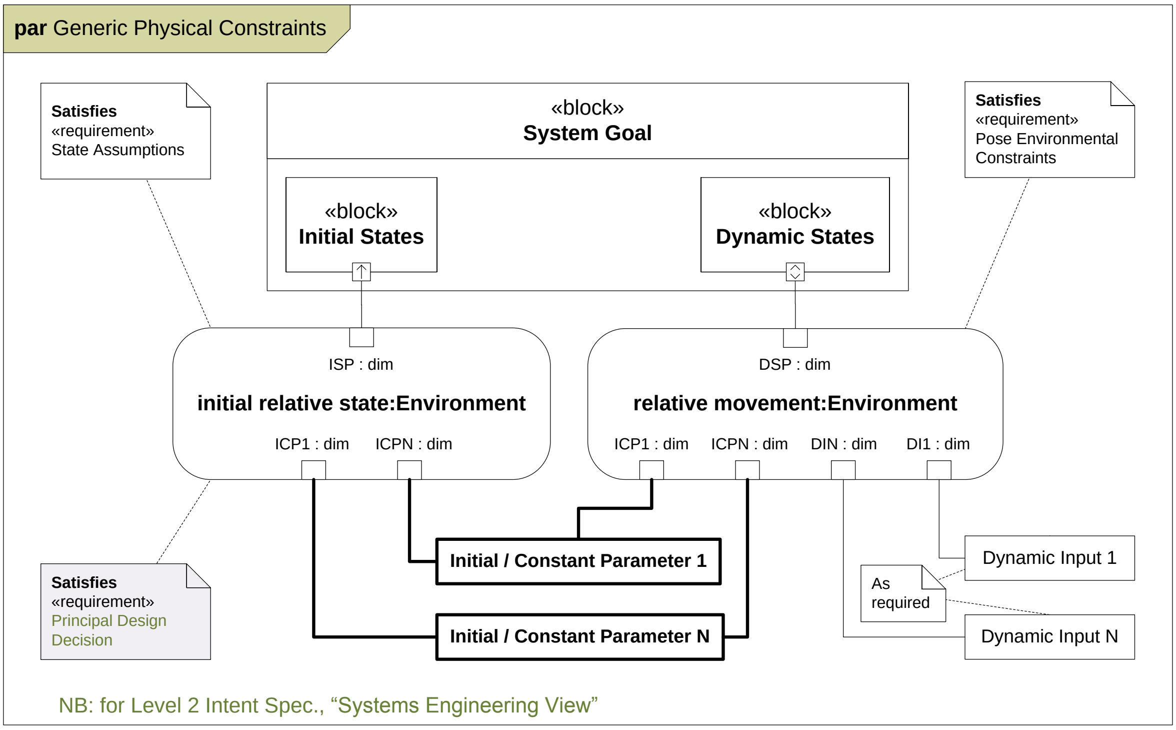

par Generic Physical Constraints
Satisfies
«requirement»
State Assumptions
«block»
System Goal
«block»
Initial States
«block»
Dynamic States
Satisfies
«requirement»
Pose Environmental Constraints
ISP : dim
initial relative state:Environment
ICP1 : dim
ICPN : dim
DSP : dim
relative movement:Environment
ICP1 : dim
ICPN : dim
DIN : dim
DI1 : dim
Initial / Constant Parameter 1
Initial / Constant Parameter N
As required
Dynamic Input 1
Dynamic Input N
Satisfies
«requirement»
Principal Design Decision
NB: for Level 2 Intent Spec., "Systems Engineering View"

Figure 7.3 Generic Physical Constraints Parametric Model

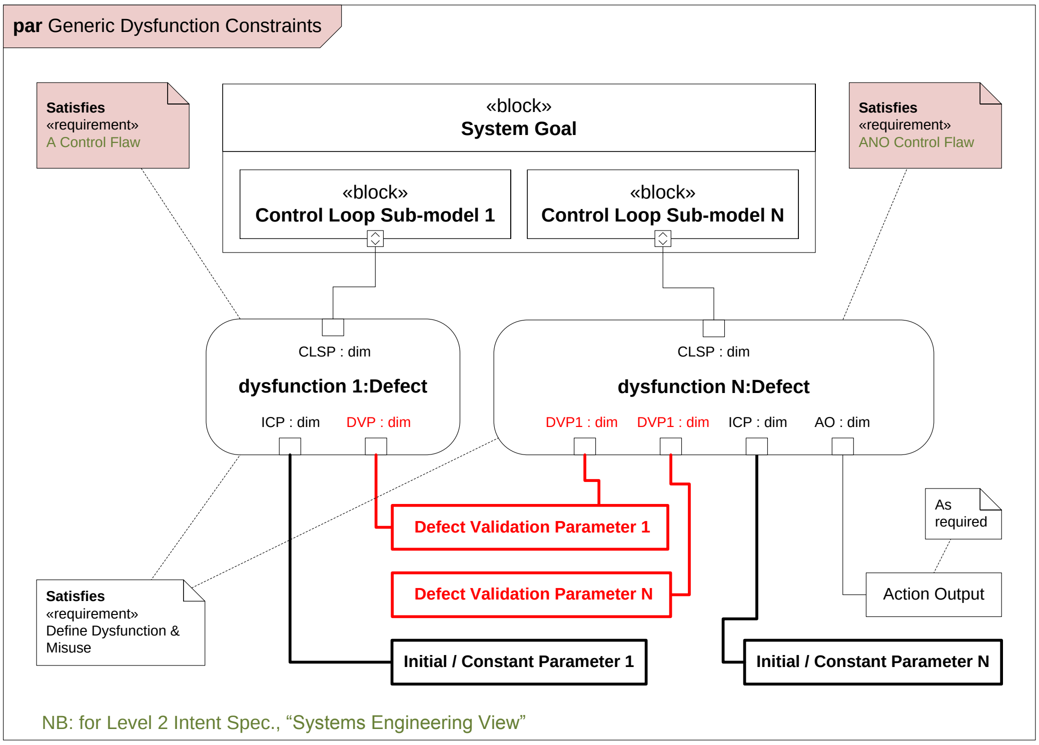


Figure 7.4 Generic Dysfunction Constraints Parametric Model

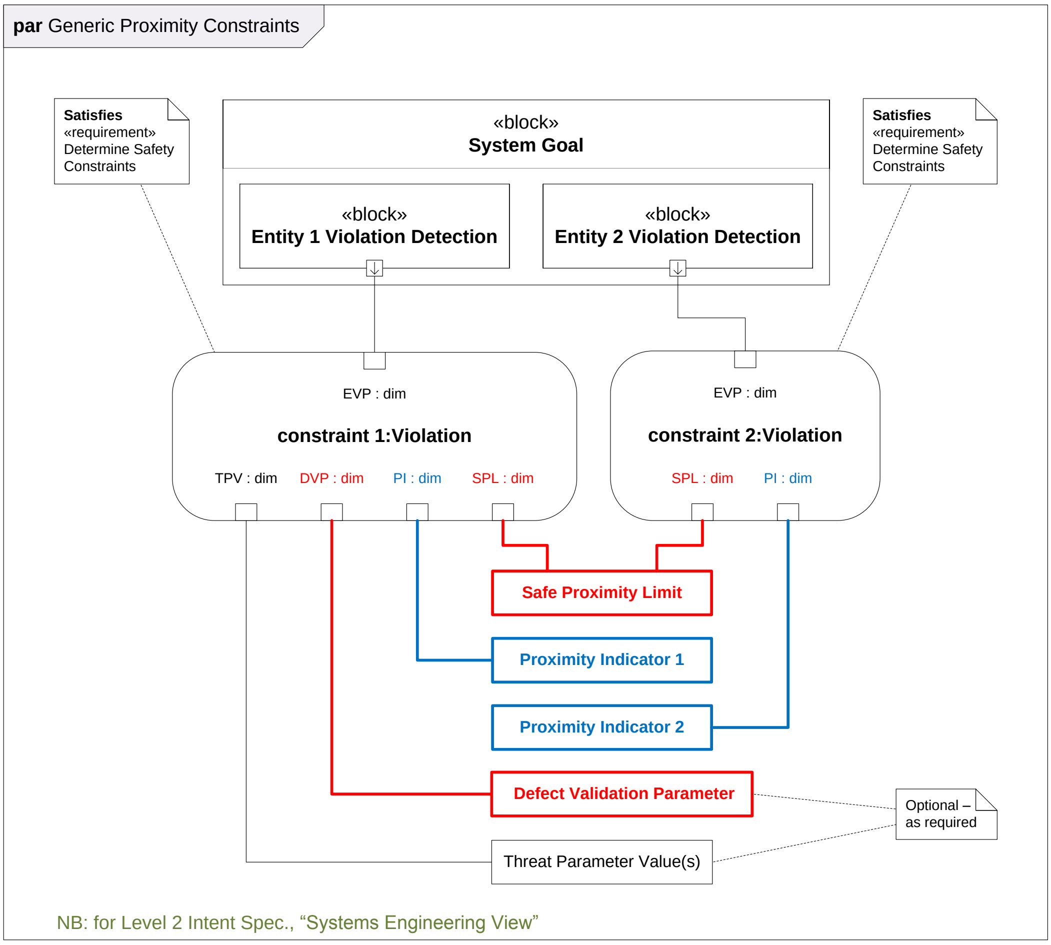


Figure 7.5 Generic Proximity Constraints Parametric Model

7.1.2. Relative Motion – Defined Environment (Motion Dynamics Model)

The complexity of the model occasionally defies succinct readable interpretation as constraint blocks, as in the axis transforms, translations and equations of motions model shown below, Figure 7.6.

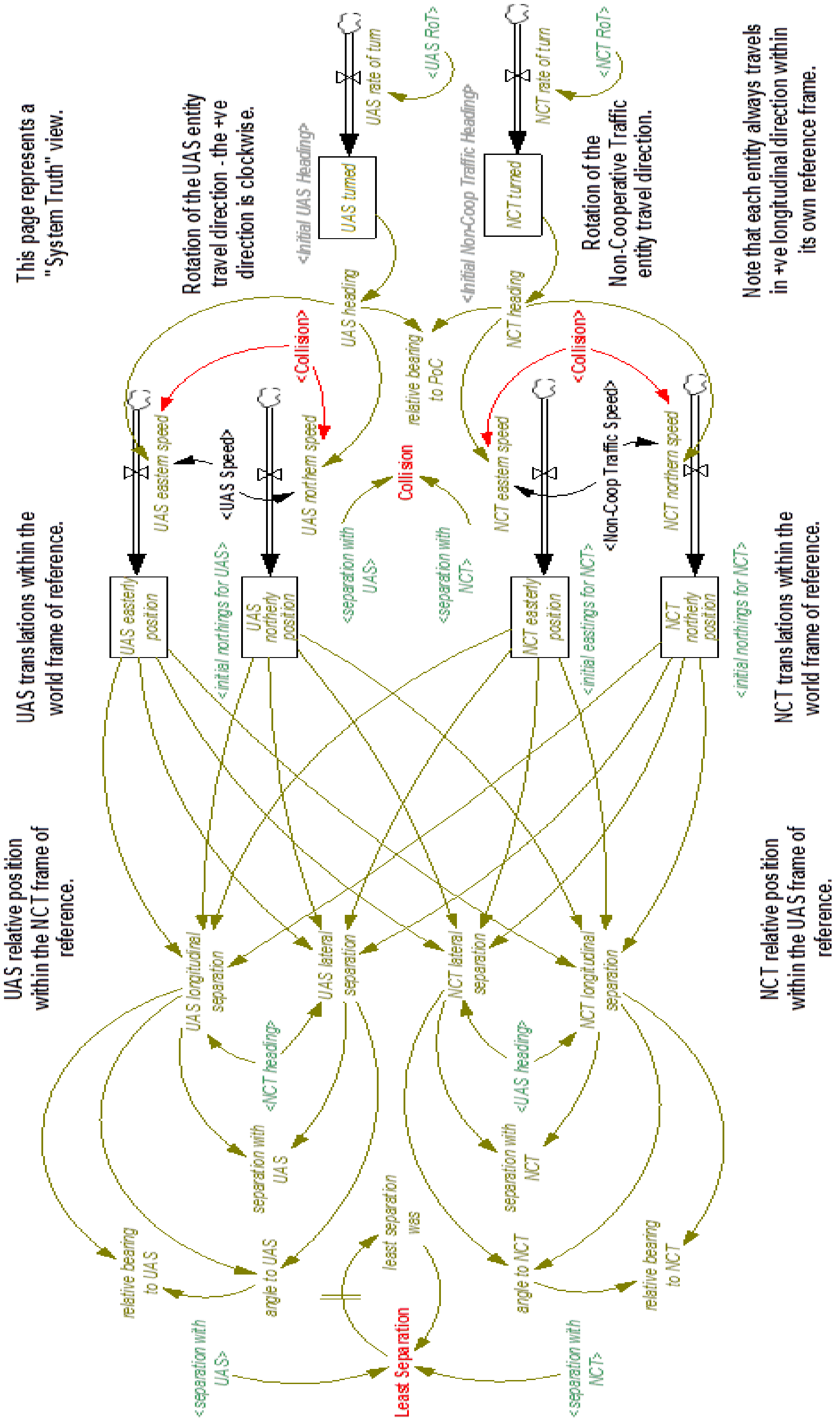

This page represents a "System Truth" view.
Rotation of the UAS entity travel direction - the +ve direction is clockwise.
Rotation of the Non-Cooperative Traffic entity travel direction.
Note that each entity always travels in +ve longitudinal direction within its own reference frame.
UAS translations within the world frame of reference.
NCT translations within the world frame of reference.
UAS relative position within the NCT frame of reference.
NCT relative position within the UAS frame of reference.
UAS rate of turn
<UAS RoT>
UAS turned
<Initial UAS Heading>
UAS heading
NCT rate of turn
<NCT RoT>
NCT turned
<Initial Non-Coop Traffic Heading>
NCT heading
relative bearing to PoC
<Collision>
<Collision>
UAS eastern speed
UAS northern speed
<UAS Speed>
Collision
<separation with UAS>
<separation with NCT>
NCT eastern speed
NCT northern speed
<Non-Coop Traffic Speed>
UAS easterly position
UAS northerly position
<initial northings for UAS>
NCT easterly position
<initial eastings for NCT>
NCT northerly position
<initial northings for NCT>
UAS longitudinal separation
UAS lateral separation
NCT lateral separation
NCT longitudinal separation
<NCT heading>
<UAS heading>
separation with UAS
separation with NCT
relative bearing to UAS
angle to UAS
least separation was
Least Separation
angle to NCT
relative bearing to NCT
<separation with UAS>
<separation with NCT>

Figure 7.6 UAS and NCT Relative Motions as "Causal Loop – Stock and Flow" Diagram

7.1.3. Integrated Manoeuvring Behaviour (Intended Behaviour Models)

The behaviours shown in Figure 7.7 depict two typical sets of trajectories over 50 second periods starting from breach of separation – which in these instances starts at 3 nautical miles. As depicted in the "intrusion" case, the UAV (shown initially heading in a northerly direction) continuously misestimates the range to the aircraft breaching separation – a Perceived Range Estimation (PRE) dysfunction. In this encounter, the UAV attempts to manoeuvre behind the non-cooperative aircraft, as this is the aircraft that has the other on its right and so shall give way. At a certain point the aircraft breaching separation observes that the UAV is now directly ahead, and so gives way to the right, whilst the UAV still attempts to manoeuvre behind it. Conversely, an acceptable behaviour is depicted in the "ok" case, albeit with the UAV perhaps overreacting as it is only emergency behaviours that include a proper sense of separation in this model; otherwise the avoidance is largely determined by any visible detection perceived along a line-of-sight.

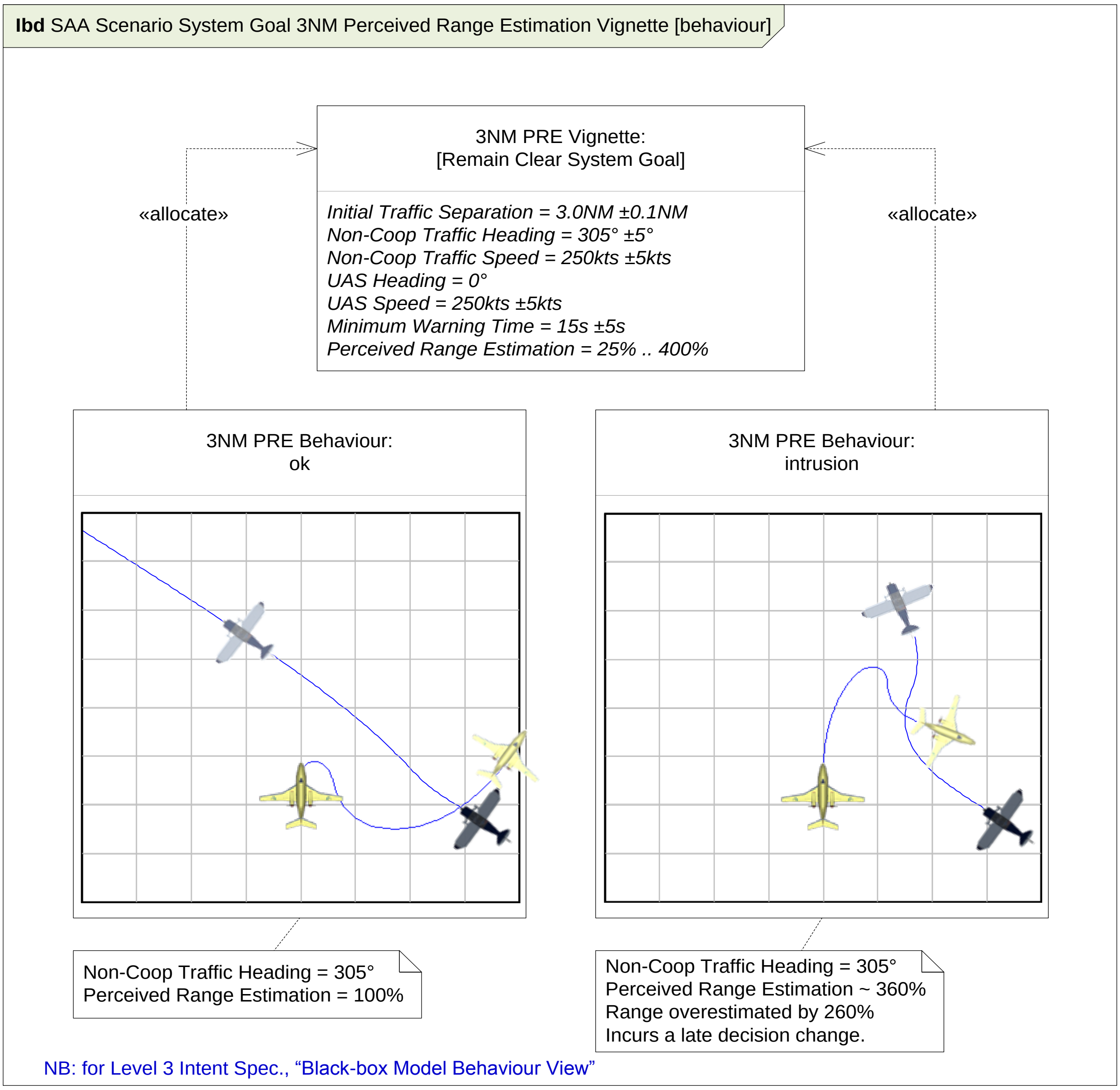
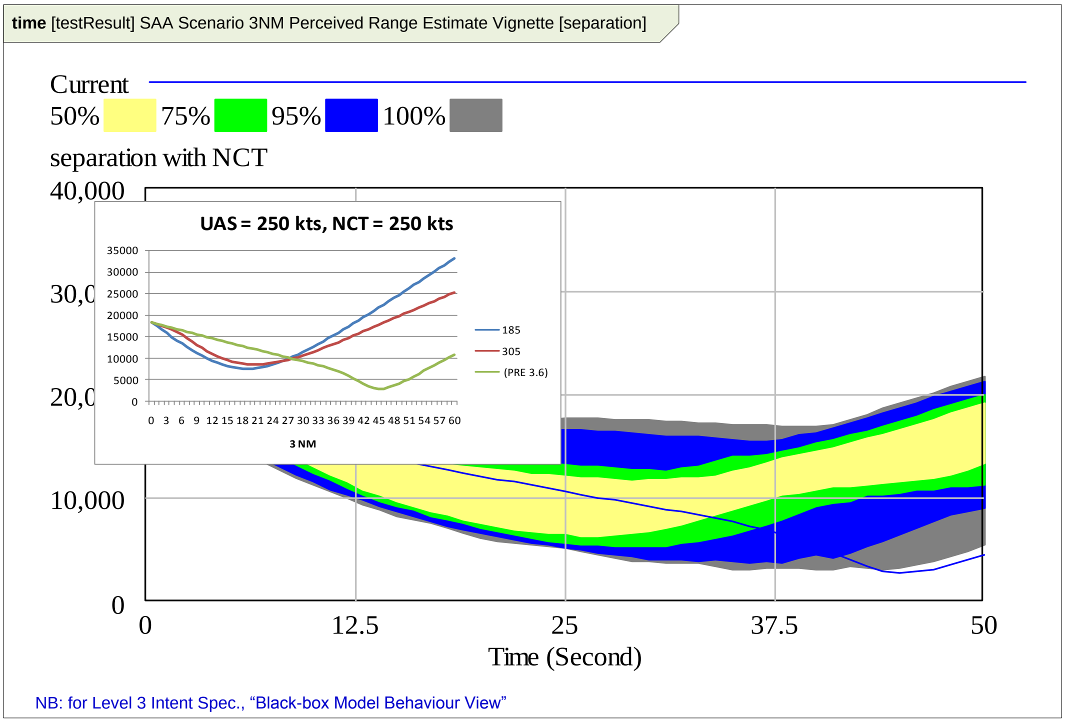

Ibd SAA Scenario System Goal 3NM Perceived Range Estimation Vignette [behaviour]
3NM PRE Vignette:
[Remain Clear System Goal]
Initial Traffic Separation = 3.0NM ±0.1NM
Non-Coop Traffic Heading = 305° ±5°
Non-Coop Traffic Speed = 250kts ±5kts
UAS Heading = 0°
UAS Speed = 250kts ±5kts
Minimum Warning Time = 15s ±5s
Perceived Range Estimation = 25% .. 400%
«allocate»
«allocate»
3NM PRE Behaviour:
ok
3NM PRE Behaviour:
intrusion
Non-Coop Traffic Heading = 305°
Perceived Range Estimation = 100%
Non-Coop Traffic Heading = 305°
Perceived Range Estimation ~ 360%
Range overestimated by 260%
Incurs a late decision change.
NB: for Level 3 Intent Spec., "Black-box Model Behaviour View"

Figure 7.7 Typical Vignette Behaviour

Expressing the Monte Carlo results relating to these behaviours, using the Monte Carlo plot tool in Vensim and a plot (using Excel) representing a stacking of the different beliefs with respect to "Inadequate Control Actions", produces perhaps unfamiliar depictions of the relevant time-based data; as are employed in the inference questions described later in Chapter 9, and illustrated here in Figure 7.8 and Figure 7.9. In the time-series plots corresponding to this vignette, the separation distribution plot shown in Figure 7.8 reveals sensitivity information contained in the encounter as a spread of statistical distributions depicting the many separation distances over elapsed time. This representation employs a native presentation format within Vensim. From this estimates can be made as to how likely a given separation might be at a particular point in elapsed time. Therefore with closer inspection of Figure 7.8 we can see for example that the closest point of approach (<3000 ft at > 37.5 seconds) lies in an outlier of the distribution of separations, sitting near to the 100% limit – note also the slight double-dip in the distribution plot at between approximately 35 and 40 seconds. Then with respect to these separation plots, the lower of the two time-series plots in Figure 7.9 provide a representation of the degrees of "belief" that particular deviations, or Inadequate Control Actions (ICA), are responsible for the behaviour. Although each belief value is expressed in a range from 0 to 1, as these beliefs are not necessarily mutually exclusive, the combination of these beliefs is represented as a stacked plot in Figure 7.9 – the total of which can exceed unity. Notice also the twin peaks having some degree of correlation with the double-dip in the Monte Carlo distribution plot.

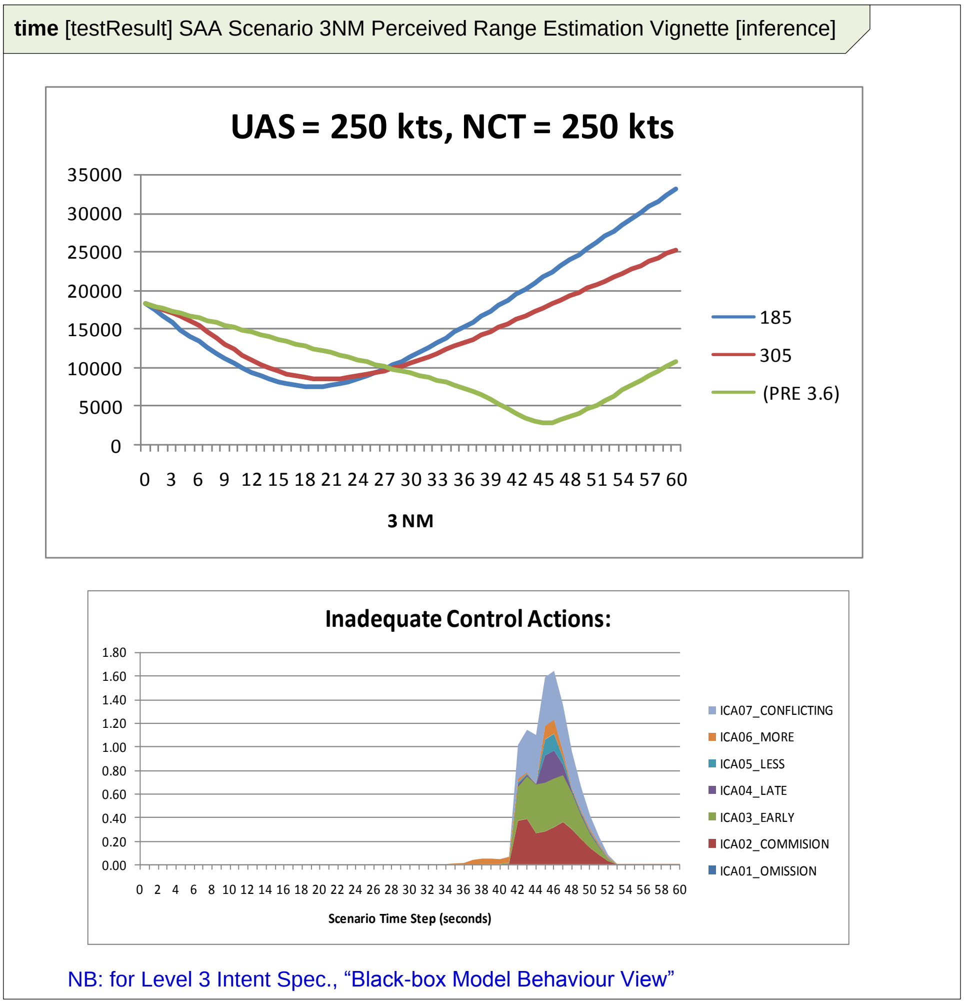
Interpreting this combination of beliefs of possible Inadequate Control Actions, we might infer that at the first closest approach in the "intrusion" case shown in Figure 7.8, and at the corresponding initial peak in the lower plot in Figure 7.9, the UAV is manoeuvring the wrong way (ICA02), coupled with a belief that it is attempting also to prematurely return to heading (ICA03), whilst employing an emergency behaviour that is "conflicted" – or unresolved (ICA07). A moment later, at the second peak, additionally it is inferred that the necessary manoeuvring has now occurred rather late (ICA04), is less than required (ICA05), but is also excessive (ICA06) considering the need to maintain line of sight contact – noting that these two interpretations span a decision point.

Later, this type of interpretation is employed in certain questionnaire cases, for the respondents to correctly identify interpretations. Eventually, and following the responses in the questionnaire, in the final version representation of the relative beliefs using this stacked plot representation was abandoned in favour of a pie-chart representation, which provides single cuts across the respective beliefs at a particular point in time sequence and providing a better relative representation of the respective beliefs.



NB: for Level 3 Intent Spec., “Black-box Model Behaviour View”

Figure 7.8 Typical vignette separation sensitivity distribution



NB: for Level 3 Intent Spec., “Black-box Model Behaviour View”

Figure 7.9 Typical vignette inference of deviation

7.1.4. Intrusion, Threat & Perception Model (Threat & Flaw Insertion Models)

Three fundamental and representative system defects have been incorporated into the model, including: errors in the perceived range estimation, variation of the warning time of threats, and variation of the latency in reacting to threats. Sensor models, fields and depth of view, explicit aerodynamic behaviour, inertial properties and delays in controller and actuator mechanisms are all either ignored, assumed to be "perfect", or otherwise considerably simplified. Only dynamics essential for avoidance are represented. Separations less than a minimum (35') are deemed collisions.

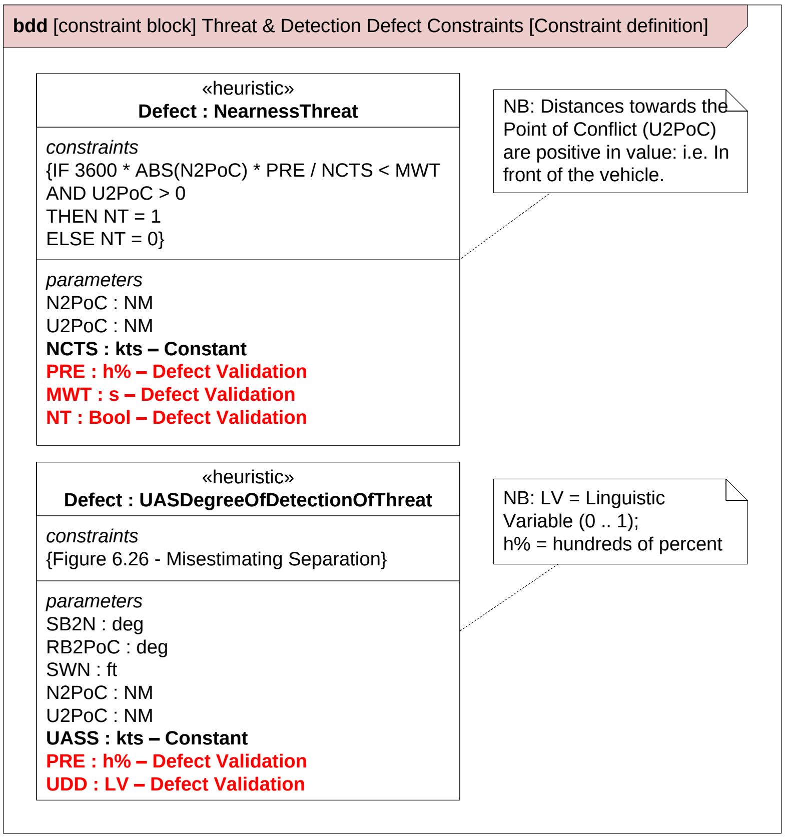
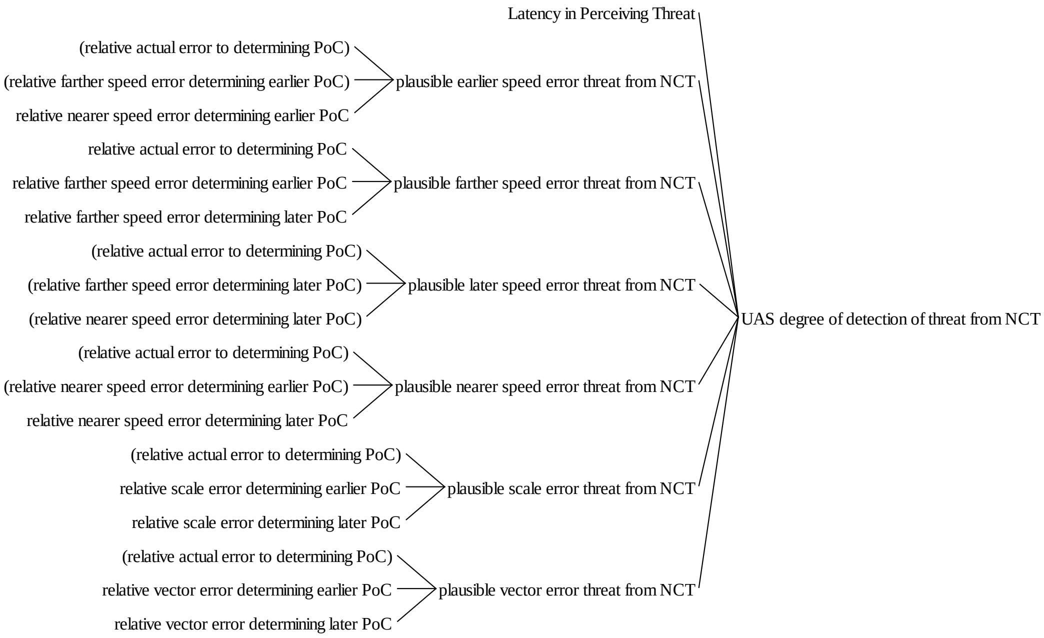

bdd [constraint block] Threat & Detection Defect Constraints [Constraint definition]
«heuristic»
Defect : NearnessThreat
constraints
{IF 3600 * ABS(N2PoC) * PRE / NCTS < MWT
AND U2PoC > 0
THEN NT = 1
ELSE NT = 0}
parameters
N2PoC : NM
U2PoC : NM
NCTS : kts – Constant
PRE : h% – Defect Validation
MWT : s – Defect Validation
NT : Bool – Defect Validation
NB: Distances towards the Point of Conflict (U2PoC) are positive in value: i.e. In front of the vehicle.
«heuristic»
Defect : UASDegreeOfDetectionOfThreat
constraints
{Figure 6.26 - Misestimating Separation}
parameters
SB2N : deg
RB2PoC : deg
SWN : ft
N2PoC : NM
U2PoC : NM
UASS : kts – Constant
PRE : h% – Defect Validation
UDD : LV – Defect Validation
NB: LV = Linguistic Variable (0 .. 1);
h% = hundreds of percent

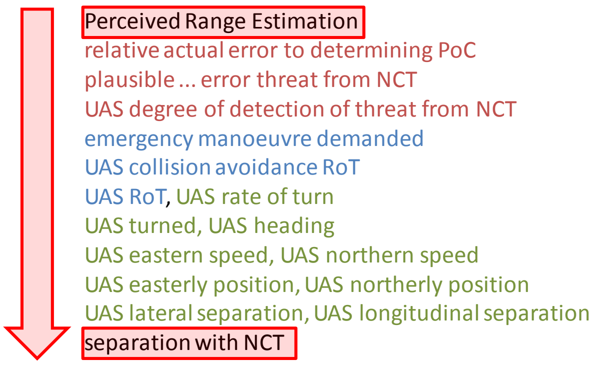
Figure 7.10 SAA Scenario Threat & Detection Defect Constraints

To this end, constraints have been defined representing the dysfunction expressions arising with these defects. In the case of the dysfunctional threat detection mechanism this is divided into two parts, as depicted in Figure 7.10. One part considers the behavioural relationship and source of dysfunction in misperceptions of perceived range (PRE) and insufficient minimum warning time (MWT), giving rise to potential failure to recognise the nearness or imminence of a threat. The other part considers the effect of perceived range estimation also in how the system might interpret possible systematic errors, in the miscalculation of the vector along which it is believed that the intruder is travelling. Note also, that the nearness threat determination is dependent upon the NCT speed – which defines the time to the point of conflict, whilst the vector dysfunction is dependent upon the UAS speed – which defines the radius of the threat zone as shown in Figure 7.12. These constraint definitions are implemented in the system model with equations (198), and equations (293), (204-9), (240), (244-51) and (218-9) – showing examples of only equations (198) and (293) in the following extract from the Vensim model simulation code:

```
(198) nearness threat = IF THEN ELSE ((3600 * ABS (NCT distance to PoC) * Perceived Range
Estimation / "Non-Coop Traffic Speed") < Minimum Warning Time :AND: UAS distance to PoC
> 0, 1, 0)
(293)   UAS degree of detection of threat from NCT =   DELAY FIXED (
          MIN (
              MAX (
                IF THEN ELSE (plausible farther speed error threat from NCT <> :NA:,
                  plausible farther speed error threat from NCT,
                  IF THEN ELSE (plausible earlier speed error threat from NCT <> :NA:,
                      plausible earlier speed error threat from NCT, :NA:)),
              IF THEN ELSE (plausible nearer speed error threat from NCT <> :NA:,
                  plausible nearer speed error threat from NCT,
                  IF THEN ELSE (plausible later speed error threat from NCT <> :NA:,
                      plausible later speed error threat from NCT, :NA:))
              ),
                MIN (plausible scale error threat from NCT, plausible vector error threat from NCT)
          ), Latency in Perceiving Threat, 0)
```

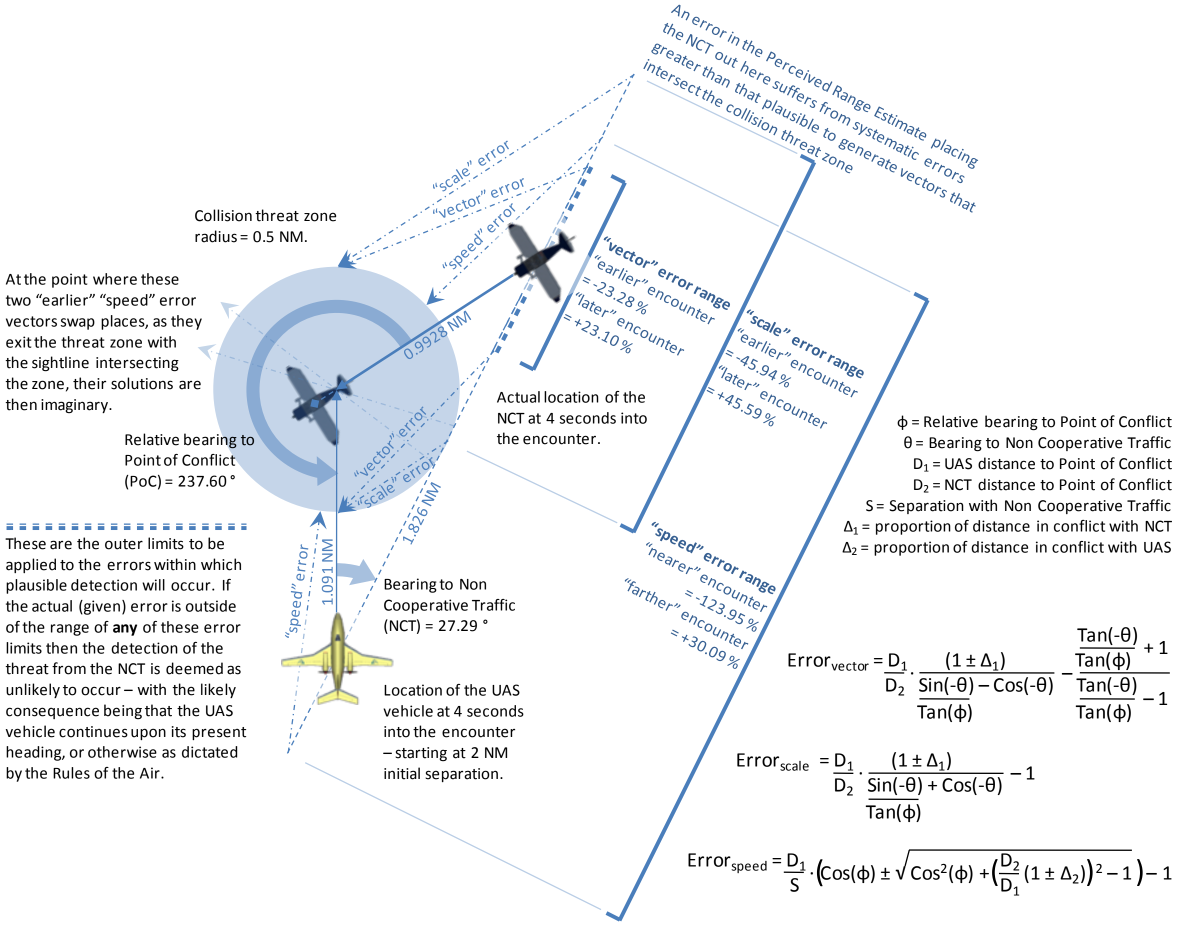
The relationship of the various component calculations that propagate this particular error state through the model are shown below and specifically the detail the logical flow of the Perceived Range Estimation (PRE) error from its origin to its consequence in shown in Figure 7.11.


Latency in Perceiving Threat
(relative actual error to determining PoC)
(relative farther speed error determining earlier PoC)
relative nearer speed error determining earlier PoC
plausible earlier speed error threat from NCT
relative actual error to determining PoC
relative farther speed error determining earlier PoC
relative farther speed error determining later PoC
plausible farther speed error threat from NCT
(relative actual error to determining PoC)
(relative farther speed error determining later PoC)
(relative nearer speed error determining later PoC)
plausible later speed error threat from NCT
UAS degree of detection of threat from NCT
(relative actual error to determining PoC)
(relative nearer speed error determining earlier PoC)
relative nearer speed error determining later PoC
plausible nearer speed error threat from NCT
(relative actual error to determining PoC)
relative scale error determining earlier PoC
relative scale error determining later PoC
plausible scale error threat from NCT
(relative actual error to determining PoC)
relative vector error determining earlier PoC
relative vector error determining later PoC
plausible vector error threat from NCT


Perceived Range Estimation
relative actual error to determining PoC
plausible ... error threat from NCT
UAS degree of detection of threat from NCT
emergency manoeuvre demanded
UAS collision avoidance RoT
UAS RoT, UAS rate of turn
UAS turned, UAS heading
UAS eastern speed, UAS northern speed
UAS easterly position, UAS northerly position
UAS lateral separation, UAS longitudinal separation
separation with NCT

Colour Key: Sense, decide and act behaviour abstraction, controller behaviour, physical behaviour

Figure 7.11 Cause – effect chain from Perceived Range Estimation to separation with NCT


An error in the Perceived Range Estimate placing the NCT out here suffers from systematic errors greater than that plausible to generate vectors that intersect the collision threat zone
"scale" error
"vector" error
"speed" error
Collision threat zone radius = 0.5 NM.
At the point where these two "earlier" "speed" error vectors swap places, as they exit the threat zone with the sightline intersecting the zone, their solutions are then imaginary.
0.9928 NM
"vector" error range
"earlier" encounter = -23.28 %
"later" encounter = +23.10 %
"scale" error range
"earlier" encounter = -45.94 %
"later" encounter = +45.59 %
Actual location of the NCT at 4 seconds into the encounter.
Relative bearing to Point of Conflict (PoC) = 237.60 °
"vector" error
"scale" error
1.826 NM
"speed" error
1.091 NM
These are the outer limits to be applied to the errors within which plausible detection will occur. If the actual (given) error is outside of the range of **any** of these error limits then the detection of the threat from the NCT is deemed as unlikely to occur – with the likely consequence being that the UAS vehicle continues upon its present heading, or otherwise as dictated by the Rules of the Air.
Bearing to Non Cooperative Traffic (NCT) = 27.29 °
Location of the UAS vehicle at 4 seconds into the encounter – starting at 2 NM initial separation.
"speed" error range
"nearer" encounter = -123.95 %
"farther" encounter = +30.09 %
ϕ = Relative bearing to Point of Conflict
θ = Bearing to Non Cooperative Traffic
D1 = UAS distance to Point of Conflict
D2 = NCT distance to Point of Conflict
S = Separation with Non Cooperative Traffic
Δ1 = proportion of distance in conflict with NCT
Δ2 = proportion of distance in conflict with UAS

$$\mathrm{Error}_{\mathrm{vector}} = \frac{D_1}{D_2} \cdot \frac{(1 \pm \Delta_1)}{\frac{\mathrm{Sin}(-\theta)}{\mathrm{Tan}(\phi)} - \mathrm{Cos}(-\theta)} - \frac{\frac{\mathrm{Tan}(-\theta)}{\mathrm{Tan}(\phi)} + 1}{\frac{\mathrm{Tan}(-\theta)}{\mathrm{Tan}(\phi)} - 1}$$

$$\mathrm{Error}_{\mathrm{scale}} = \frac{D_1}{D_2} \cdot \frac{(1 \pm \Delta_1)}{\frac{\mathrm{Sin}(-\theta)}{\mathrm{Tan}(\phi)} + \mathrm{Cos}(-\theta)} - 1$$

$$\mathrm{Error}_{\mathrm{speed}} = \frac{D_1}{S} \cdot \left(\mathrm{Cos}(\phi) \pm \sqrt{\mathrm{Cos}^2(\phi) + \left(\frac{D_2}{D_1}(1 \pm \Delta_2)\right)^2 - 1}\right) - 1$$

Figure 7.12 Misestimating Separation – an example

The determination of the effect of misestimating separation, illustrated as three different types of plausible error behaviour in Figure 7.12, can be abstracted to these three cases: miscalculating the distance to the threat zone ("speed" error), skewing the perceived direction of travel ("vector" error), or applying an incorrectly scaled triangle of velocities ("scale" error). Together this involves an extended series of conditional equations, as alluded to above (in the square roots), but stands as a generic worst case (Min-Max) solution encompassing these plausible error cases.

7.1.5. Delay & Offset Direction Model (Flaw Insertion Models)

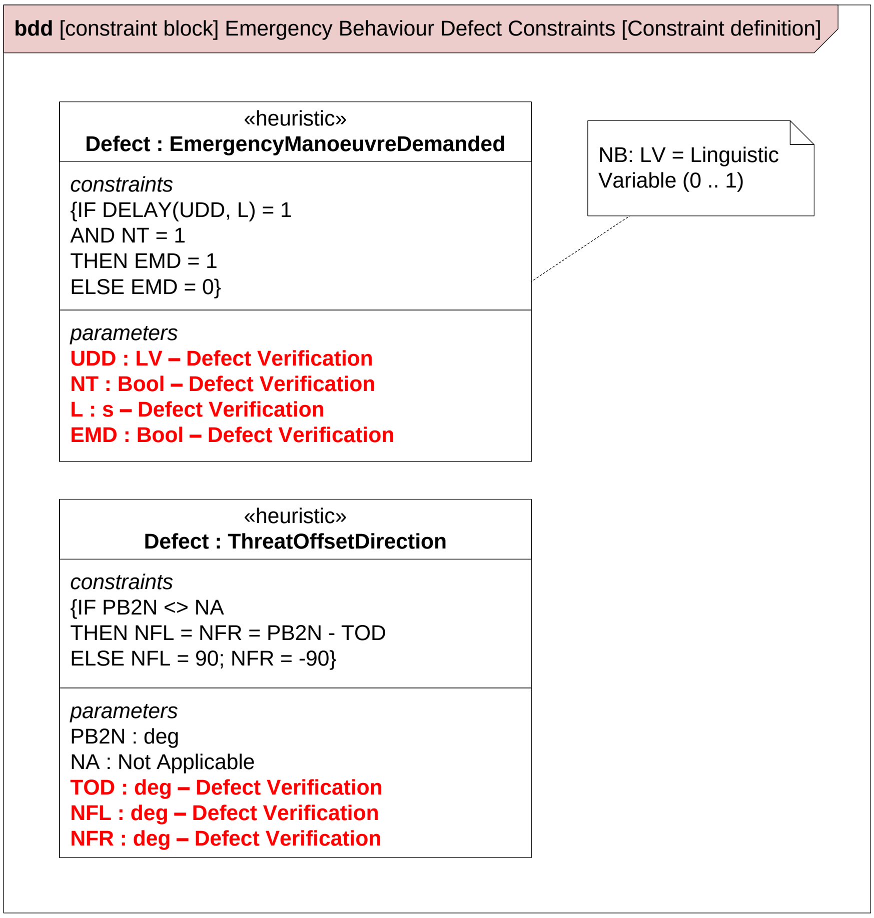

bdd [constraint block] Emergency Behaviour Defect Constraints [Constraint definition]
«heuristic»
Defect : EmergencyManoeuvreDemanded
constraints
{IF DELAY(UDD, L) = 1
AND NT = 1
THEN EMD = 1
ELSE EMD = 0}
parameters
UDD : LV – Defect Verification
NT : Bool – Defect Verification
L : s – Defect Verification
EMD : Bool – Defect Verification
NB: LV = Linguistic Variable (0 .. 1)
«heuristic»
Defect : ThreatOffsetDirection
constraints
{IF PB2N <> NA
THEN NFL = NFR = PB2N - TOD
ELSE NFL = 90; NFR = -90}
parameters
PB2N : deg
NA : Not Applicable
TOD : deg – Defect Verification
NFL : deg – Defect Verification
NFR : deg – Defect Verification

Figure 7.13 SAA Scenario Emergency Behaviour Defect Constraints

Delays and offsets are simpler, with the emergency behaviour defect constraint definitions in Figure 7.13 – implemented with equations (073), and with equations (131-2) and (282):

(073) emergency manoeuvre demanded = IF THEN ELSE ((UAS degree of detection of threat from NCT = 1) :AND: (nearness threat = 1), 1, 0)

(131) NCT from left = WITH LOOKUP (IF THEN ELSE (perceived bearing to NCT <> :NA:, perceived bearing to NCT - Threat Offset Direction, 90), ([(-375,0)-(375,1)], (-375,1), (-345,0), (-195,0), (-165,1), (-15,1), (15,0), (165,0), (195,1), (345,1), (375,0)))

(132) NCT from right = WITH LOOKUP (IF THEN ELSE (perceived bearing to NCT <> :NA:, perceived bearing to NCT - Threat Offset Direction, -90), ([(-375,0)-(375,1)], (-375,0), (-345,1), (-195,1), (-165,0), (-15,0), (15,1), (165,1), (195,0), (345,0), (375,1)))

(282) Threat Offset Direction = 47 Units: **undefined** [-90,90]

For all of the defects there are traceable chains of effect or in some cases multiple paths through the model arising with a defect. Each of these paths might be traced and can help in the validation of the model. Defects arise in dysfunction constraints defined within the sense, decide and act abstraction, as below.

7.1. Safety Constraint Modelling

As stated previously, Air Proximity (Airprox) exists as a framework within which to declare near-misses between two aircraft, with biannually published compilations of reports providing adequate background information, guidance and other data [111]. Codifying the near-miss as an air proximity event will help to provide a recognised formal framework within which to couch safety constraint violation. Another, additional or alternative measure of hazardous behaviour demonstrating a tendency to violate safety constraints may be found in a suitable measurement of emergent behaviour. Later it will be shown that a degree of correlation exists between safety constraint categorisation of near-miss behaviour and the observed emergent properties within a group of nearby or similar cases.

7.1.1. Air Proximity Model (Constraint Violation Detection Models)

The dysfunctions provided in the model can be considered as representations of the “did not see”, “inadequate avoiding action”, and “late sighting” cases in Table 7.1, from a UAS viewpoint.

Cause	Attributed to
Did not separate / poor judgement	Controller
Climbed / descended through assigned level	Pilot
Did not see conflicting traffic	Pilot
Inadequate avoiding action / flew too close	Pilot
Late sighting of conflicting traffic	Pilot
Misinterpretation of ATC message	Pilot
Not obeying orders / following advice from ATC	Pilot
Penetration of CAS/SRZ/ATZ without clearance	Pilot
Conflict in other type of airspace	Other
FIR conflict	Other
Sighting Report	Other

Table 7.1 Dominant Air Proximity Causal Factors [111]

The design intent of the collision avoidance system is as likely to be judged as much on the possibilities of it giving rise to air proximity events, as it is for any speculative probabilistic collision rate. Therefore the modelled safety constraint ought to relate appropriately to the likelihood of near-misses, wherein various categories of causal factors and potential hazard might be defined – wherein *"... An Airprox is a situation in which, in the opinion of a pilot or controller, the distance between aircraft as well as their relative positions and speed was such that the safety of the aircraft involved* *was or may have been compromised."* [111]. Consequently, this is not only a question of measuring the separation distance, but other relative indicators from the perspective of each participant are also required. Note also that this judgement is subjective and expressible as belief.

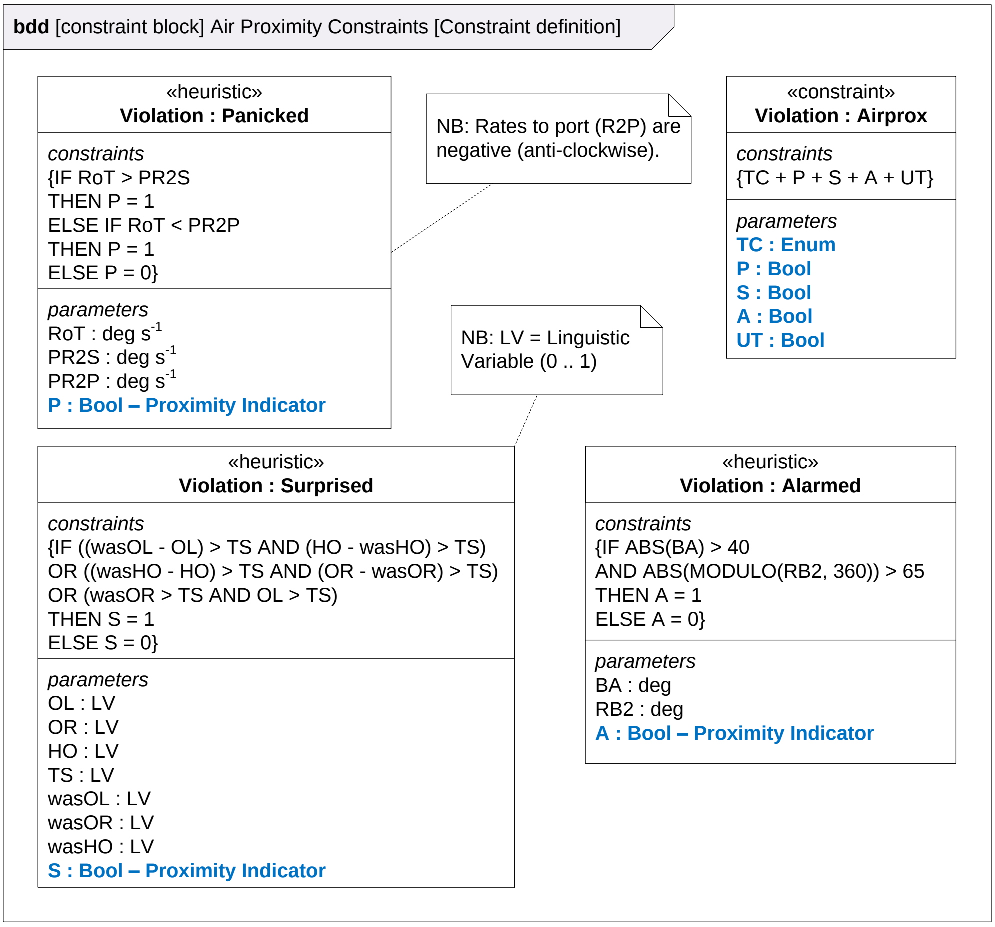


Figure 7.14 SAA Scenario UAS Airprox Constraints

Therefore, in addition to a separation function – itself producing an inverse square relationship with separation distance, incorporating misperceived distances, normalised around the minimum acceptable separation distance (500') – four additional constraints also apply:

- Alarmed – whenever the vehicle banks beyond 40 degrees and outside of a forward looking cone with an internal angle of 130 degrees. In these conditions the vehicle loses sight of the other vehicle, and would be justifiably alarmed;
- Panicked – the vehicle is forced to perform excessive manoeuvring, with a rate of turn greater than Rate Three (12 degrees per second);
- Surprised – the other vehicle appears not to be applying the Rules of the Air, either by crossing ahead from left to right at a rate greater than a certain Transit Sensitivity value (x 65 degrees per second azimuth), or behind from right to left within half a second – how far from right to left is determined by the value of the Transit Sensitivity (from 140 degrees at minimum to 220 degrees at maximum);

- Threatened – the vehicle is forced into behaviour other than the default interpretation of the Rules of the Air, with a turn to port greater than Rate Two (-6 degrees per second).

Together these are combined (added) to generate an "Airprox" value, as described in the air proximity constraint definitions in Figure 7.14 – implemented in the system dynamics model with equations (287-8), (336), (365):

```
(287)  UAS Airprox = NCT Too Close + UAS Panicked + UAS Surprised + UAS Threatened +
       UAS Alarmed
(288)  UAS Alarmed = IF THEN ELSE ((ABS(UAS banked) > 40)
       :AND: ABS (MODULO (relative bearing to NCT, 360)) > 65, 1, 0)
(336)  UAS Panicked = IF THEN ELSE (UAS RoT > Panic Rate to Stbd, 1,
       IF THEN ELSE (UAS RoT < Panic Rate to Port, 1, 0))
(365)  UAS Surprised = IF THEN ELSE ((((NCT was on left - NCT on left)
       > Transit Sensitivity) :AND: ((NCT head on - NCT was head on) > Transit Sensitivity))
       :OR: (((NCT was head on - NCT head on) > Transit Sensitivity) :AND:
       ((NCT on right - NCT was on right) > Transit Sensitivity))
       :OR:    ((NCT was on right  > Transit Sensitivity) :AND:
       (NCT on left > Transit Sensitivity)), 1, 0)
```

And with equations (118-9), (163), and (191) for the NCT:

```
(118)  NCT Airprox = NCT Panicked + NCT Surprised + UAS Too Close + NCT Alarmed
(119)  NCT Alarmed = IF THEN ELSE ((ABS (NCT banked) > 40)
       :AND: ABS (MODULO (relative bearing to UAS, 360)) > 65, 1, 0)
(163)  NCT Panicked = IF THEN ELSE (NCT RoT > Panic Rate to Stbd, 1,
       IF THEN ELSE (NCT RoT < Panic Rate to Port, 1, 0))
(191)  NCT Surprised = IF THEN ELSE ((((UAS was on left - UAS on left)
       > Transit Sensitivity) :AND: ((UAS head on - UAS was head on) > Transit Sensitivity))
       :OR: (((UAS was head on - UAS head on) > Transit Sensitivity) :AND:
       ((UAS on right - UAS was on right) > Transit Sensitivity))
       :OR: ((UAS was on right > Transit Sensitivity) :AND:
       (UAS on left > Transit Sensitivity)), 1, 0)
```

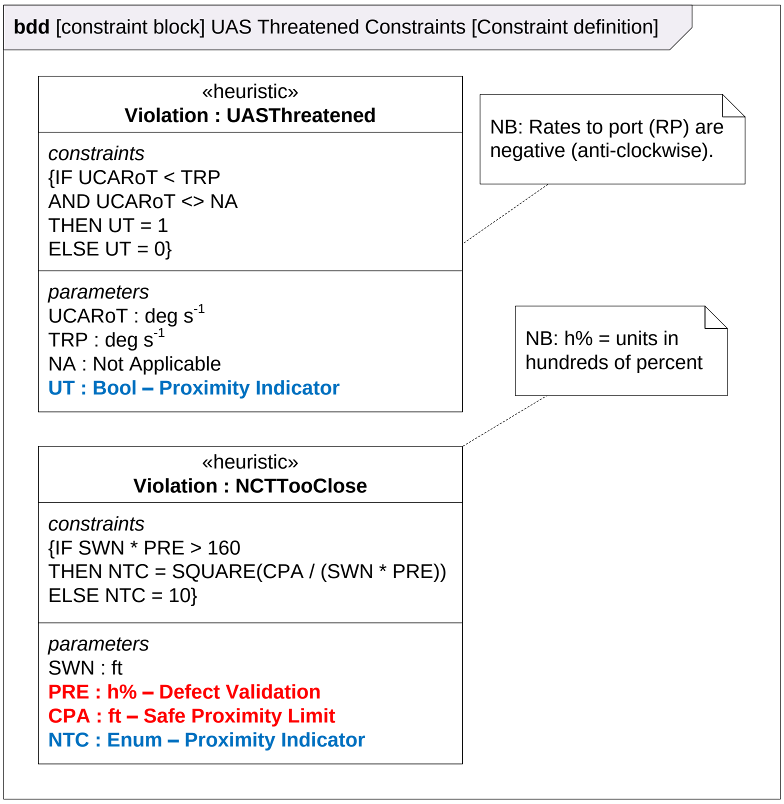

bdd [constraint block] UAS Threatened Constraints [Constraint definition]
«heuristic»
Violation : UASThreatened
constraints
{IF UCARoT < TRP
AND UCARoT <> NA
THEN UT = 1
ELSE UT = 0}
parameters
UCARoT : deg s-1
TRP : deg s-1
NA : Not Applicable
UT : Bool – Proximity Indicator
NB: Rates to port (RP) are negative (anti-clockwise).
NB: h% = units in hundreds of percent
«heuristic»
Violation : NCTTooClose
constraints
{IF SWN * PRE > 160
THEN NTC = SQUARE(CPA / (SWN * PRE))
ELSE NTC = 10}
parameters
SWN : ft
PRE : h% – Defect Validation
CPA : ft – Safe Proximity Limit
NTC : Enum – Proximity Indicator

Figure 7.15 SAA Scenario UAS Threatened Constraints

The UAS only threatened constraint definitions presented in Figure 7.15 are implemented in the system dynamics model with equations (366) and (193):

(366) UAS Threatened = IF THEN ELSE (UAS collision avoidance RoT
< Threatened Rate to Port :AND: UAS collision avoidance RoT <> :NA:, 1, 0)

(193) NCT Too Close = IF THEN ELSE (separation with NCT * Perceived Range Estimation
> 160, (Closest Point of Approach /
(separation with NCT * Perceived Range Estimation)) ^ 2, 10)
Too Close – lack of safe separation, as judged according to the perceived range estimate in the case of the UAS.

Summations of the UAS and NCT proximity threat producing values above 6.5 are placed in Category A – defined in Table 7.2, values between 4.5 and 6.5 are placed in Category B, values between 2.5 and 4.5 are placed in Category C, and values between 0.5 and 2.5 are placed in Category D – values less than this are categorised as None. In the case of category D, whilst this may not be a strictly correct interpretation of the evidence, or lack of, it appears to work sensibly in practice within the model. Together then a selection of these constraint violation indicators are also presented later within the Bayesian belief model.

Airprox	Risk	Interpretation	Occurrence
Category A	Risk of collision	an actual risk of collision existed	7 pa in UK on a 5 year average
Category B	Safety not assured	the safety of the aircraft was compromised	24 pa in UK on a 5 year average
Category C	No risk of collision	no risk of collision existed	56 pa in UK on a 5 year average
Category D	Risk not determined	insufficient information, inconclusive	2 pa in UK on a 5 year average

Table 7.2 Air Proximity (Airprox) Classifications [111]

7.1.2. Emergent System Behaviour Sensitivity Modelling

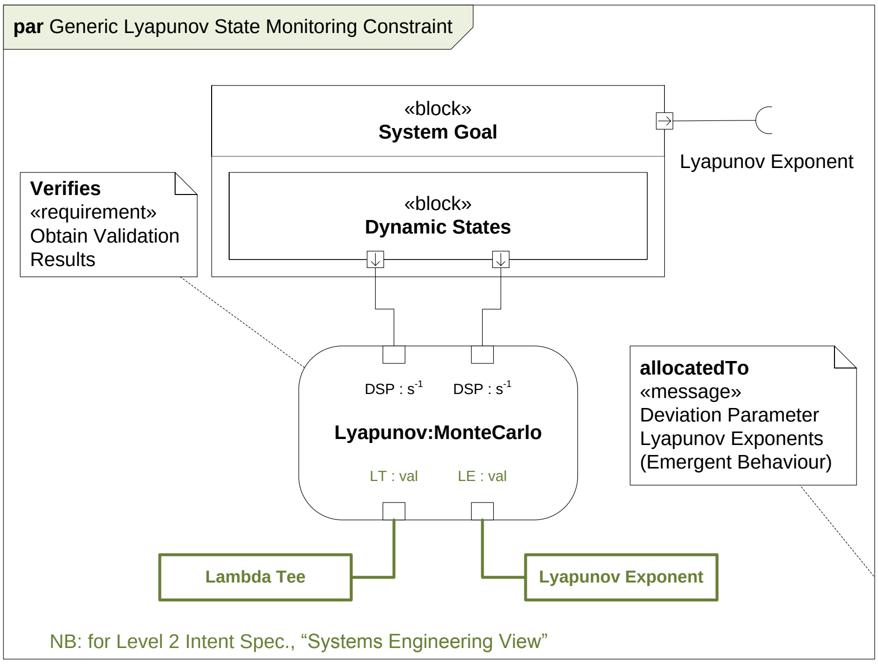


Figure 7.16 Generic Lyapunov State Monitoring Parametric Model

An established measure of the emergent behaviour of a system is its Lyapunov Exponent [4, 61, 107] – equation 7.1. This is a measure of the rate of separation of initially nearby trajectories, which in this instance are the differences in headings due to the initial respective rates of turn.

$$|\boldsymbol{\delta\theta_T}| = |\boldsymbol{\delta\theta_0}| \cdot \boldsymbol{e^{\lambda T}} \qquad 7.1$$

This measure provides a comparison between this initial movement of the system, and that observed at a later time *T*, such that if the modulus (absolute value) of the average the rate of change decreases over time then the system must be dissipative, otherwise divergent rates of change indicate that the system is potentially chaotic. If the average of the rates of change remains constant then the system is conservative and stable. Inspection of a family of Lyapunov exponents, generated by Monte Carlo simulation, provides insight into the likelihood of a particular case being a member of a family with more or less emergent behaviour.

The exponent λ is formulated as the averaged logarithm of the rates of change being measured, taking advantage of the "chain rule" for a chained series of derivative ratios, producing the exponent from the running average of the logarithms of the modulus of the ratios of successive rates of change – equations 7.2 and 7.3. As λ is an exponent, then positive values are indicative of divergent behaviour and negative values are indicative of convergent behaviour, whilst values tending towards zero indicate stable oscillatory behaviour. For the purposes of this study the Lyapunov exponents have been calculated for the change of heading $\delta\theta$ of the UAV (equivalent to the UAV rate of turn with fixed time steps), and similarly $\delta\varphi$ for the NCT.

As the prior histories of these two values are unknown, and a need to scale the two initial exponent values to appropriately capture the respective influences of the UAV and NCT on the exponent, then an assumption is made that the first value will be used as the first ratio. In this manner the often more aggressive initial manoeuvring of the UAV is allowed dominate, unless the NCT is the first to manoeuvre. This appropriately scales the remainder of the Lyapunov calculation – formed by adding both UAV and NCT values together, and from which a maximum value is obtained – equation 7.4.

$$\boldsymbol{\lambda_{\theta_T}} = \begin{cases} \frac{1}{T}\left[\sum_{t=t_i,\Delta t}^{T} \ln\left(\frac{|\dot{\theta}_t|}{|\dot{\theta}_{t-\Delta t}|}\right) + t_i \cdot \lambda_{\theta_{t_i}}\right], & \dot{\theta}_t \neq 0, \quad \dot{\theta}_{t-\Delta t} \neq 0 \\ \frac{1}{t_i}\ln\left(|\dot{\theta}_{t_i}|\right), & \dot{\theta}_{t_i} \neq 0, \quad \dot{\theta}_{t_i-\Delta t} = 0 \\ 0, & \dot{\theta}_t = 0, \quad \dot{\theta}_{t-\Delta t} = 0 \end{cases} \qquad 7.2$$

$$\boldsymbol{\lambda_{\varphi_T}} = \begin{cases} \frac{1}{T}\left[\sum_{t=t_k,\Delta t}^{T} \ln\left(\frac{|\dot{\varphi}_t|}{|\dot{\varphi}_{t-\Delta t}|}\right) + t_k \cdot \lambda_{\varphi_{t_k}}\right], & \dot{\varphi}_t \neq 0, \quad \dot{\varphi}_{t-\Delta t} \neq 0 \\ \frac{1}{t_k}\ln\left(|\dot{\varphi}_{t_k}|\right), & \dot{\varphi}_{t_k} \neq 0, \quad \dot{\varphi}_{t_k-\Delta t} = 0 \\ 0, & \dot{\varphi}_t = 0, \quad \dot{\varphi}_{t-\Delta t} = 0 \end{cases} \qquad 7.3$$

$$\lambda_T = \lambda_{\theta_T} + \lambda_{\varphi_T}, \; \hat{\lambda}_T = \max_{r=1\rightarrow 200}\left(\lambda_{T_r}\right) \qquad 7.4$$

For the constraint definitions represented in Figure 7.16 and Figure 7.17, the Lyapunov exponent is formed in the system dynamics model with equations (115), (109), (328), (157), (360) and (187), implementing these constraints. Note that as a short cut, derivatives with respect to heading are obtained directly from the respective rates of turn. The relationship between heading and rate of turn is also directly associated with the parameter term "Manoeuvre".

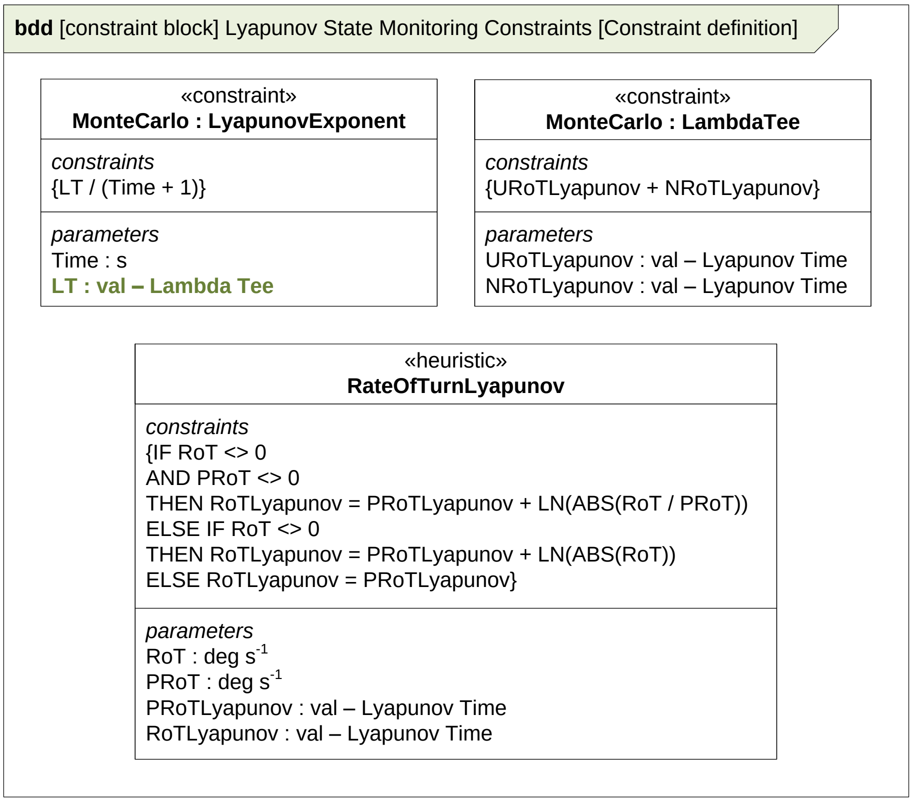

bdd [constraint block] Lyapunov State Monitoring Constraints [Constraint definition]
«constraint»
MonteCarlo : LyapunovExponent
constraints
{LT / (Time + 1)}
parameters
Time : s
LT : val – Lambda Tee
«constraint»
MonteCarlo : LambdaTee
constraints
{URoTLyapunov + NRoTLyapunov}
parameters
URoTLyapunov : val – Lyapunov Time
NRoTLyapunov : val – Lyapunov Time
«heuristic»
RateOfTurnLyapunov
constraints
{IF RoT <> 0
AND PRoT <> 0
THEN RoTLyapunov = PRoTLyapunov + LN(ABS(RoT / PRoT))
ELSE IF RoT <> 0
THEN RoTLyapunov = PRoTLyapunov + LN(ABS(RoT))
ELSE RoTLyapunov = PRoTLyapunov}
parameters
RoT : deg s-1
PRoT : deg s-1
PRoTLyapunov : val – Lyapunov Time
RoTLyapunov : val – Lyapunov Time

Figure 7.17 SAA Scenario Lyapunov State Monitoring Constraints

(115) Lyapunov Exponent = (UAS Lyapunov + NCT Lyapunov) / (Time + 1)

(109) Lambda Tee = UAS Lyapunov + NCT Lyapunov

(328) UAS Lyapunov = UAS RoT Lyapunov

(157) NCT Lyapunov = NCT RoT Lyapunov

(360) UAS RoT Lyapunov = IF THEN ELSE (((UAS rate of turn <> 0) :AND:
(UAS previous RoT <> 0)),
(previous UAS RoT Lyapunov + LN (ABS (UAS rate of turn / UAS previous RoT))),
IF THEN ELSE((UAS rate of turn <> 0),
previous UAS RoT Lyapunov + LN (ABS (UAS rate of turn)),
previous UAS RoT Lyapunov))

(187) NCT RoT Lyapunov = IF THEN ELSE (((NCT rate of turn <> 0) :AND:
(NCT previous RoT <> 0)),
(previous NCT RoT Lyapunov + LN (ABS (NCT rate of turn / NCT previous RoT))),
IF THEN ELSE((NCT rate of turn <> 0),
previous NCT RoT Lyapunov + LN (ABS (NCT rate of turn)),
previous NCT RoT Lyapunov))

For the visualisation of these exponents, directly accessing the exponent value prior to the division of the integral by Time (T) – or “Lambda Tee” (λT) – shows behaviour more clearly. This viewpoint facilitates easier observation of changes in the direction of the integral, which itself provides useful insight into the nature of the manoeuvring and with familiarity also allows interpretation of this behaviour directly from the calculation of the Lyapunov exponent. For example upward inflections from negative values indicate increases in the rate of manoeuvre and often also indicate changes in the direction of the turn, as illustrated with the examples in Figure 7.18, both of which embody "S" shaped turns by the UAV, and clearly recognisable as occurring at around 12.5 seconds with the dip in the right-hand example; both of which capture a particular characteristic emergent behaviour with a manoeuvre not explicitly defined in the model.

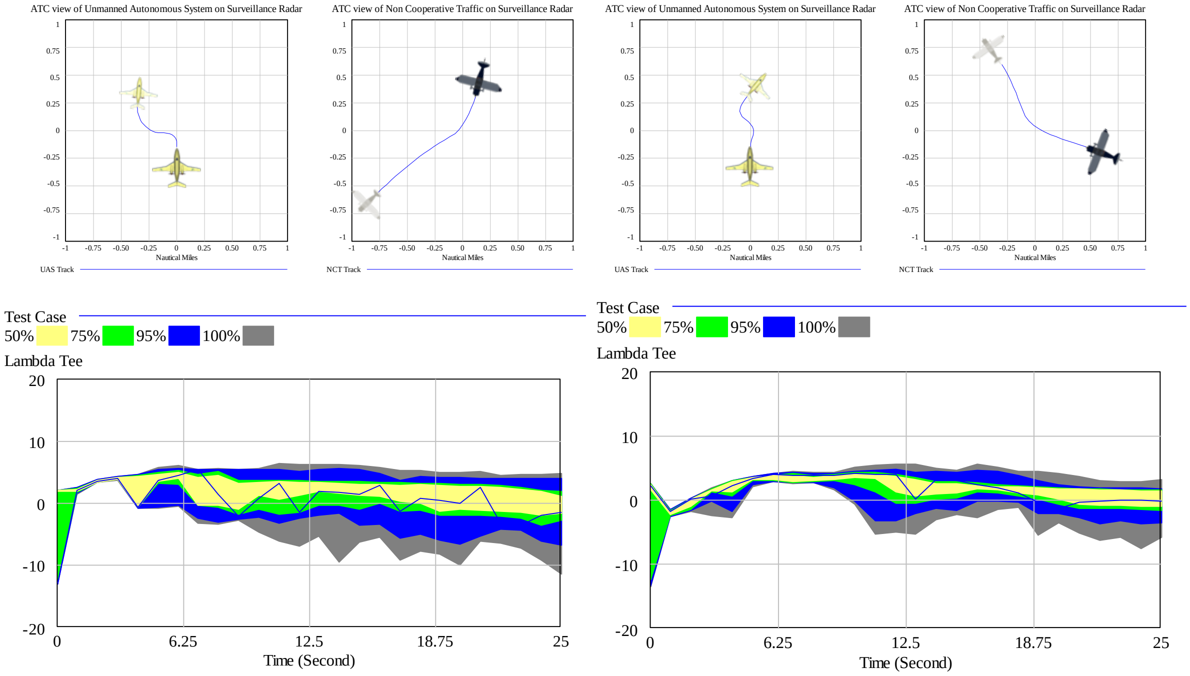


Figure 7.18 Example Lyapunov "Lambda Tee" plots with similar emergent behaviour (TC1 and TC2)

7.2. Summary of Purpose and Recapitulation

To recap, the purpose in the definition and creation of this model is to satisfy the following broad requirements, and with the aim of refuting the null hypotheses and seeking evidence for the alternative hypothesis. The next stage is to see to what extent these objectives have been achieved.

02.23 Simulate different combinations of deviations;

- 02.23.01 Model exemplar hazard HAZ010 for the VFR case, from the NIAG SG-134 Hazard Scenarios;
- 02.23.02 Define a simplified but believable and informative mid-air collision avoidance scenario with which to evaluate the hazard model.

02.17 Evaluate hypothetical deviations from intent;

02.20 Associate failure conditions with corresponding failure types;

02.14 Combine flow based HAZOP and function based FHA.

Chapter 8. Belief and Model Validation

"A plausible impossibility is always preferable to an unconvincing possibility."

Aristotle

8.1. Inference Modelling

The overall modelling approach is constructed within the scope of two reactive models, a System Dynamics model and a Bayesian Inference model. The remainder of the models are static descriptive models, mostly formulated in SysML some of which are meta-models, as for example with the. The aim here is to validate plausibility of the reactive models, and thereby furnish measures of the satisfaction of requirements 02.23 and 02.17, as were drawn from the earlier textual and QFD analysis. As the Bayesian model also serves a continuing purpose of providing internal validation and verification of the assumptions formed within the HAZOP process, how this is constructed should be dealt with first.

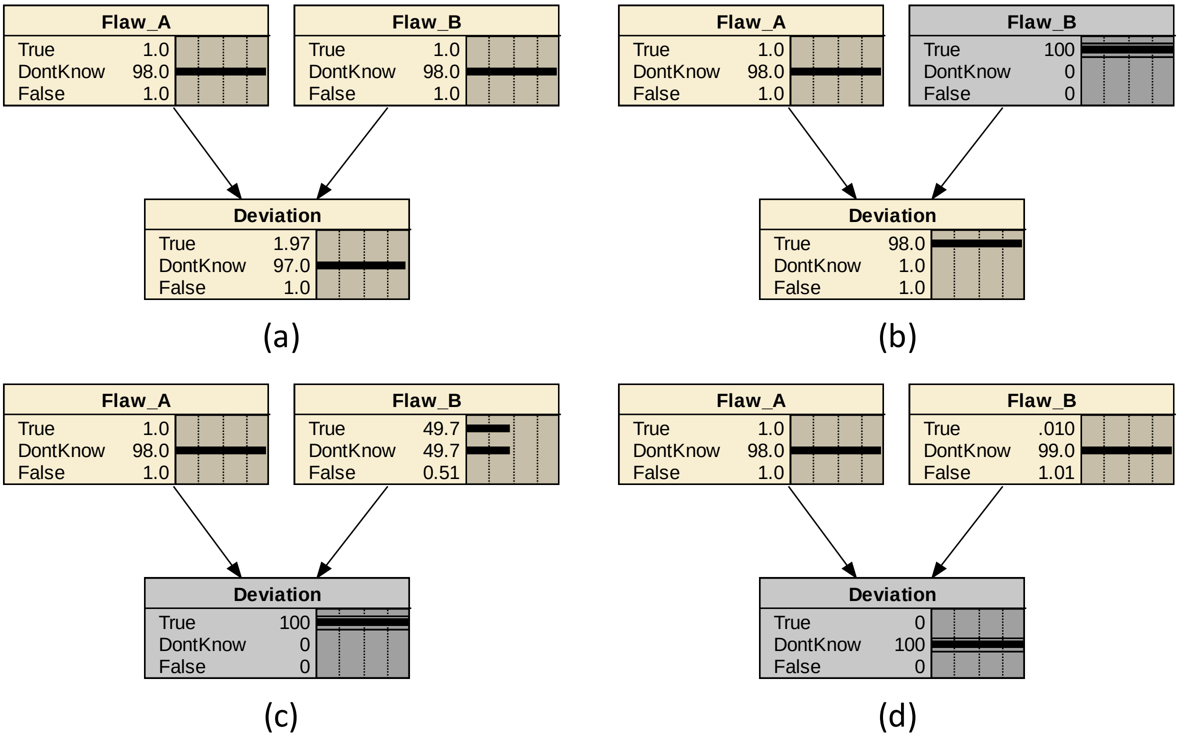


Figure 8.1 Simple Example Flaw – Deviation Bayesian Belief Network

Within the sequence of inference modelling activities, a Bayesian network describing the belief that flaws cause deviations is created; in a form as shown in the simplified example illustrated in Figure 8.1. So as to represent absence of certainty, three states are defined for each node – True, False and "Don't Know". The initial bias is towards not knowing or not having very much certainty of belief – either True or False, but greater belief that the truth of this is largely unknown.

Let the Flaws (*Fx*) in this example, have three states “True” (*Fx$_T$*), “Don’t Know” (*Fx$_D$*) and “False” (*Fx$_F$*), as defined below.

$$P(FA_T) = P(FB_T) = 0.01, \qquad P(FA_D) = P(FB_D) = 0.98, \qquad P(FA_F) = P(FB_F) = 0.01$$

Also let us assume that in the case of Flaw B (*FB*), we are quite sure that in 98% of cases this flaw, when true, does directly causes the deviation *D* to be true also – as indicated in example case (b).

$$P(D_T|FB_T) = 0.98, \qquad P(D_T|\neg FB_T) = 0.01$$

$$P(D_D|FB_T) = 0.01, \qquad P(D_D|\neg FB_T) = 0.98$$

Consequently as the deviation state, either as true or unknown, is solely dependent the state of Flaw B, we can directly calculate the probabilities of these states as the sums of the conditional probabilities below – with the result shown also in example case (a), above.

$$P(D_T) = P(D_T|FB_T) \cdot P(FB_T) + P(D_T|FB_D) \cdot P(FB_D) + P(D_T|FB_F) \cdot P(FB_F) = 0.0197$$

$$P(D_D) = P(D_D|FB_T) \cdot P(FB_T) + P(D_D|FB_D) \cdot P(FB_D) + P(D_D|FB_F) \cdot P(FB_F) = 0.9703$$

$$P(D_F) = 0.01$$

From this we might now infer the likelihood of either Flaw A or B being a cause of deviation, whenever a deviation is observed as “True” – as is the illustrated in case (c) above, and as expressed overleaf. As representations of belief, the outcome can be expressed thus: if it is known that 98% of cases of Flaw B cause the Deviation, but also with a possibility that the cause of the deviation might be with Flaw A, then the only evidence being a single observation of the (true) existence of the Deviation suggests that the likelihood that Flaw B is actually the cause may be equally true as "Don't Know" – as in the absence of any other evidence, Flaw B is just as likely to be present as Flaw A where the belief in Flaw A is mostly "Don't Know". With the certainty of Flaw B as a known cause close to 100%, then this split of likelihoods from the point of view of B is close to 50:50, the remainder covering the possibility of the small number of known false cases in this situation.

$$P(FA_T|D_T) = \frac{P(D_T|FA_T) \cdot P(FA_T)}{P(D_T)} = \frac{P(D_T)}{P(D_T)} \cdot P(FA_T) = P(FA_T) = 0.01$$

$$likewise\ \ P(FA_D|D_D) = P(FA_D) = 0.98\ \ and\ \ P(FA_F|D_F) = P(FA_F) = 0.01$$

$$P(FB_T|D_T) = \frac{P(D_T|FB_T,) \cdot P(FB_T)}{P(D_T)} = 0.49746$$

$$likewise\ \ P(FB_D|D_T) = \frac{P(D_T|FB_D) \cdot P(FB_D)}{P(D_T)} = 0.49746$$

$$likewise\ \ P(FB_F|D_T) = \frac{P(D_T|FB_F) \cdot P(FB_F)}{P(D_T)} = 0.0050761$$

Assuming that no evidence of the certainty of a deviation or otherwise is presented, such that the deviation adopts an uncertain "Don't Know" state, then the likelihood that Flaw B exists is reduced by an order of magnitude as the evidence is counterfactual, the possibility that the state of Flaw B is unknown increases slightly, and the likelihood that it is false increases slightly also; as expressed below[5] and shown in case (d), Figure 8.1. Such counterfactual evidence also has an effect on the revised probabilistic belief update propagated through the Bayesian model during training using evidence derived from the Monte Carlo simulations applied to the System Dynamics model.

$$P(FB_T|D_D) = \frac{P(D_D|FB_T) \cdot P(FB_T)}{P(D_D)} = 0.0010306$$

$$likewise\ \ P(FB_D|D_D) = \frac{P(D_D|FB_D) \cdot P(FB_D)}{P(D_D)} = 0.98980$$

$$likewise\ \ P(FB_F|D_D) = \frac{P(D_D|FB_F) \cdot P(FB_F)}{P(D_D)} = 0.010100$$

With this method all of the nodes and relationships in the HAZOP Bayesian belief network are preconfigured with a 98% bias towards "Don't Know", and after training from data extracted through the Monte Carlo simulations, any still unobserved cases remain clearly represented as "Don't Know".

8.1.1. HAZOP Inference Modelling

The purpose of this is to provide a mechanism whereby belief as formulated with the HAZOP table can be updated with evidence obtained from a System Dynamics model. Therefore this belief inference model may be used to validate the beliefs formed within the HAZOP process, assuming that we might also find another and independent means of validating the System Dynamics model.

With this methodology in mind, the initial HAZOP as produced in Table 6.1, and as represented also in Figure A.1, is employed in the construction of a Bayesian belief network as shown in Figure 8.3. So as to tie the control flaw likelihood to measurable physical effects, additional causative assumptions may also be wired into the belief model; for example assuming a cause and effect model where a latency defect may be expected to produce particular flaws – in this case where the non-cooperative traffic's interpretation of the rules of the air does not allow sufficient time to respond (ICA05), or a delayed response causes the UAS to alter its decision failing to arbitrate correctly (IAC07), or simply the UAS just acts after too much delay (ICA04). The flow of cause and effect is

[5] The sum of these probabilities ought to be unity, as can be seen here these truncated values produce a small apparent cumulative rounding error of 0.093%.

in the direction of the arrows; Figure 8.3. However, the earlier version of the composition of HAZOP table and Bayesian Network, shown below in Figure 8.2 [32], reversed this CF – ICA relationship and was used to produce the inferences for the questionnaire, but this proved ultimately unsatisfactory.

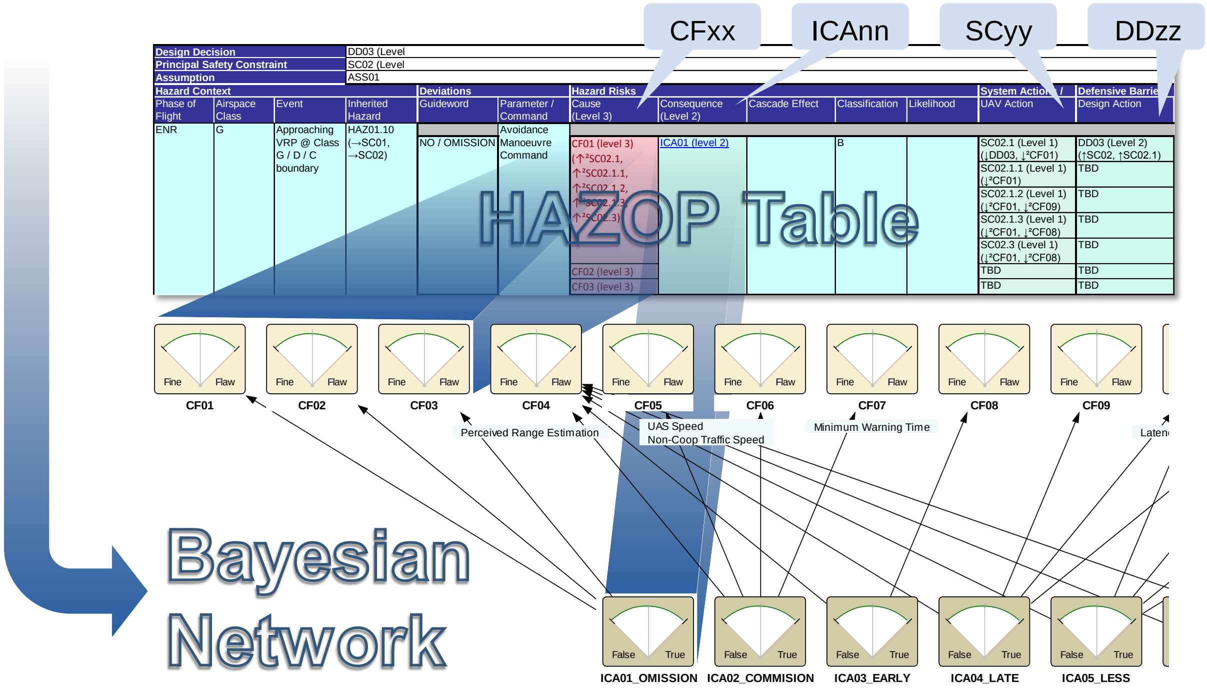
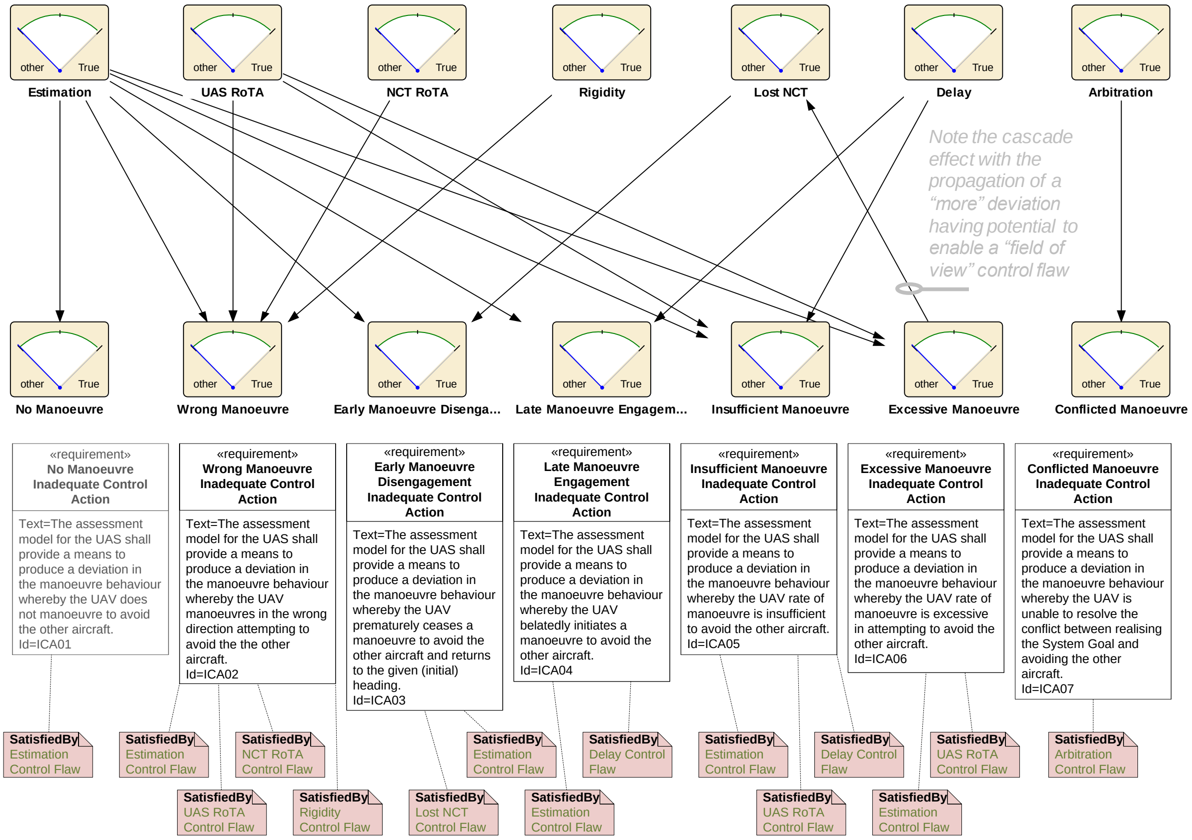


Figure 8.2 Original HAZOP for Control Flaw Diagnosis and Cause Inference [32]


other True
Estimation
UAS RoTA
NCT RoTA
Rigidity
Lost NCT
Delay
Arbitration
Note the cascade effect with the propagation of a "more" deviation having potential to enable a "field of view" control flaw
No Manoeuvre
Wrong Manoeuvre
Early Manoeuvre Disenga...
Late Manoeuvre Engagem...
Insufficient Manoeuvre
Excessive Manoeuvre
Conflicted Manoeuvre
SatisfiedBy Estimation Control Flaw
SatisfiedBy UAS RoTA Control Flaw
SatisfiedBy NCT RoTA Control Flaw
SatisfiedBy Rigidity Control Flaw
SatisfiedBy Lost NCT Control Flaw
SatisfiedBy Estimation Control Flaw
SatisfiedBy Delay Control Flaw
SatisfiedBy Estimation Control Flaw
SatisfiedBy Estimation Control Flaw
SatisfiedBy Delay Control Flaw
SatisfiedBy UAS RoTA Control Flaw
SatisfiedBy UAS RoTA Control Flaw
SatisfiedBy Estimation Control Flaw
SatisfiedBy Arbitration Control Flaw

«requirement»
No Manoeuvre Inadequate Control Action

Text=The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV does not manoeuvre to avoid the other aircraft.
Id=ICA01

«requirement»
Wrong Manoeuvre Inadequate Control Action

Text=The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV manoeuvres in the wrong direction attempting to avoid the the other aircraft.
Id=ICA02

«requirement»
Early Manoeuvre Disengagement Inadequate Control Action

Text=The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV prematurely ceases a manoeuvre to avoid the other aircraft and returns to the given (initial) heading.
Id=ICA03

«requirement»
Late Manoeuvre Engagement Inadequate Control Action

Text=The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV belatedly initiates a manoeuvre to avoid the other aircraft.
Id=ICA04

«requirement»
Insufficient Manoeuvre Inadequate Control Action

Text=The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV rate of manoeuvre is insufficient to avoid the other aircraft.
Id=ICA05

«requirement»
Excessive Manoeuvre Inadequate Control Action

Text=The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV rate of manoeuvre is excessive in attempting to avoid the other aircraft.
Id=ICA06

«requirement»
Conflicted Manoeuvre Inadequate Control Action

Text=The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV is unable to resolve the conflict between realising the System Goal and avoiding the other aircraft.
Id=ICA07

Figure 8.3 Later HAZOP Causal Relationships as a Bayesian Network

There may also be occasion to believe that a deviation associated with a particular guideword might also have a cascade effect, in that a particular deviation variant gives rise to a new flaw, and thereby cascades to realise other deviations – for example where excessive (MORE / ICA06) manoeuvring causes the UAS to lose sight of the other aircraft (LostNCT), potentially for the manoeuvre behaviour to further deviate with premature (EARLY / ICA03) disengagement of the avoidance behaviour; as in Figure 8.3. Note that in this case the direction of the cause and effect arrow is from the specific deviation, and to the newly introduced flaw, and cascading back to further possible deviations.

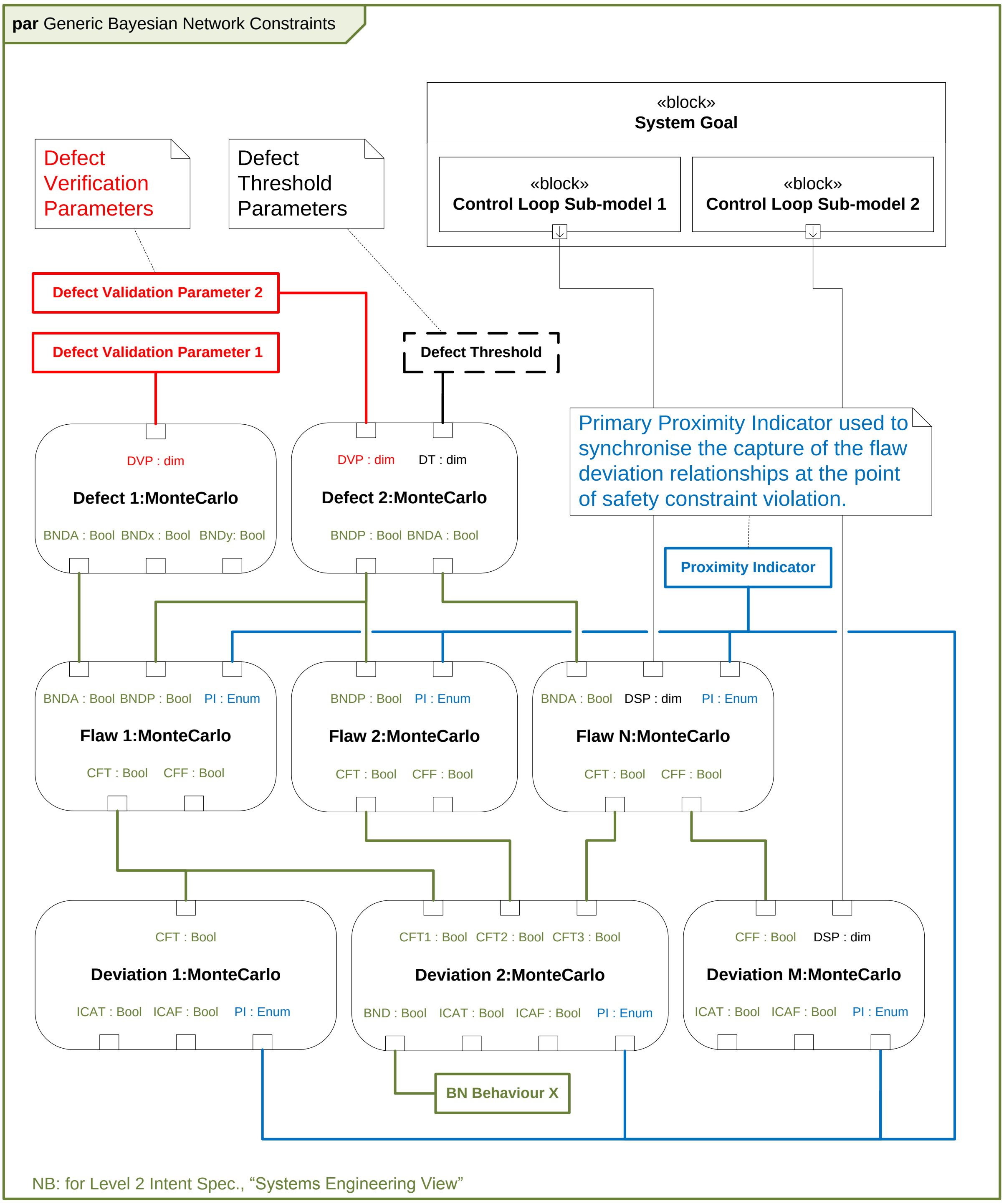

par Generic Bayesian Network Constraints
Defect Verification Parameters
Defect Threshold Parameters
«block» System Goal
«block» Control Loop Sub-model 1
«block» Control Loop Sub-model 2
Defect Validation Parameter 2
Defect Validation Parameter 1
Defect Threshold
Primary Proximity Indicator used to synchronise the capture of the flaw deviation relationships at the point of safety constraint violation.
Proximity Indicator
DVP : dim
Defect 1:MonteCarlo
BNDA : Bool BNDx : Bool BNDy: Bool
DVP : dim DT : dim
Defect 2:MonteCarlo
BNDP : Bool BNDA : Bool
BNDA : Bool BNDP : Bool PI : Enum
Flaw 1:MonteCarlo
CFT : Bool CFF : Bool
BNDP : Bool PI : Enum
Flaw 2:MonteCarlo
CFT : Bool CFF : Bool
BNDA : Bool DSP : dim PI : Enum
Flaw N:MonteCarlo
CFT : Bool CFF : Bool
CFT : Bool
Deviation 1:MonteCarlo
ICAT : Bool ICAF : Bool PI : Enum
CFT1 : Bool CFT2 : Bool CFT3 : Bool
Deviation 2:MonteCarlo
BND : Bool ICAT : Bool ICAF : Bool PI : Enum
CFF : Bool DSP : dim
Deviation M:MonteCarlo
ICAT : Bool ICAF : Bool PI : Enum
BN Behaviour X
NB: for Level 2 Intent Spec., "Systems Engineering View"

Figure 8.4 Generic Bayesian Network Parametric Model

As with previous aspects describing the constraints surrounding the System Dynamics model, the need to define a consistent interface between the data to be captured from this model, through Monte Carlo simulation, and the presentation of this revealed data to the Bayesian network model is also provided in a framework of constraints. In its generic form (Figure 8.4) the Bayesian network interface model describes three sets of interconnected constraint blocks, representing Defect, Flaw and Deviation capture constraints respectively, in a parametric relationship to models within the System Goal model.

Deviation capture constraints may be related to one or more Flaw capture constraints, and may also exhibit direct linkage to specific physical behaviours. Flaw capture constraints may also be related to one or more Defect and Deviation capture constraints. Defect capture constraints are used to validate defect and decision threshold parameters. Flaws and Deviations are captured for training, evaluation and later inference, triggered by a valid proximity or other safety constraint violation – with this "Proximity Indicator" trigger synchronising these events as shown on the previous page in Figure 8.4.

8.1.2. Defect, Flaw & Deviation Behaviour Capture – Monte Carlo Generated Cases

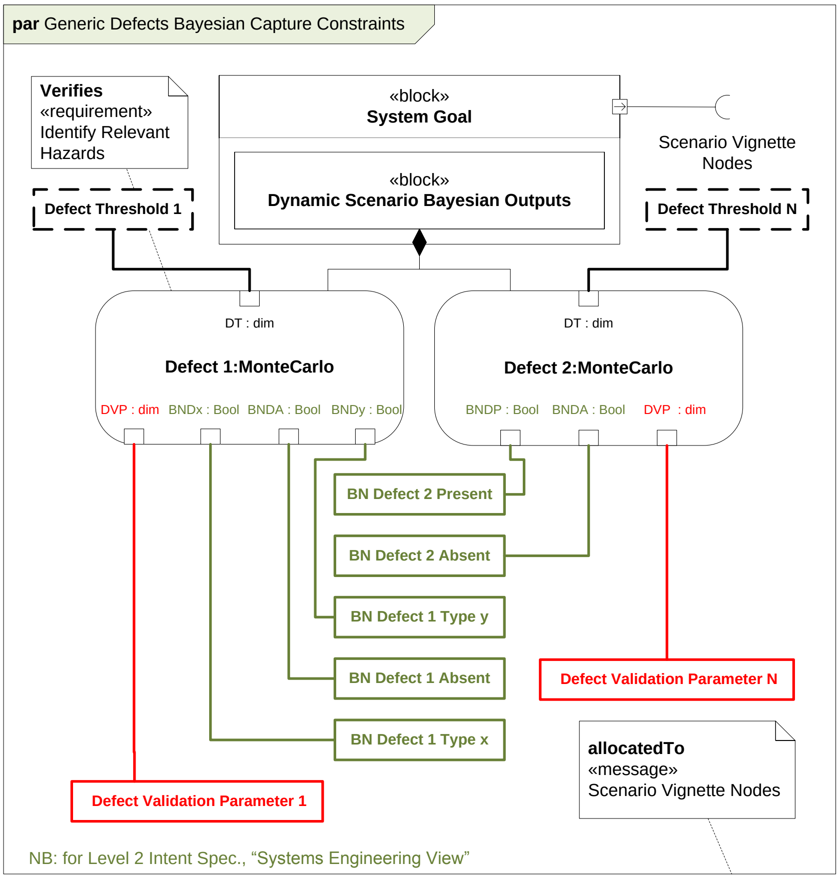


Figure 8.5   Generic Defect Bayesian Capture Parametric Model

As with the earlier defined Physical, Dysfunction, Safety Constraint Violation, and Emergent Property constraint models, the Test and Training case data capture interfaces can also be similarly described –

defining the limits of behaviour of the system dynamics model to be captured through Monte Carlo simulations. Again these models are defined both in adapted SysML Parametric diagrams, and also as individual Constraint blocks.

In the case of the system Defect events, Figure 8.5, input values are derived from constants representing decision thresholds and tolerances driving the model behaviour, although the constraints themselves are associated with the dynamic scenario and the Bayesian interface model they do not take values from the internal dynamics of the model itself. The Flaw event capture constraints respond to Defect events and observable physical behaviour, Figure 8.6. Deviation event capture is constrained by Flaws, and occasionally also by certain physical behaviours, and by the broader model of safety constraint violation and detection, Figure 8.7. Observable physical behaviour and Proximity violation events are also separately captured, Figure 8.8 and Figure 8.9. All of these constraint models employ «heuristic» block models as each represents a decision as to whether to capture a particular event value or not – based either upon the violation of a system safety constraint, an exceedance event, or the occurrence of an out of tolerance value driving the model in that case.

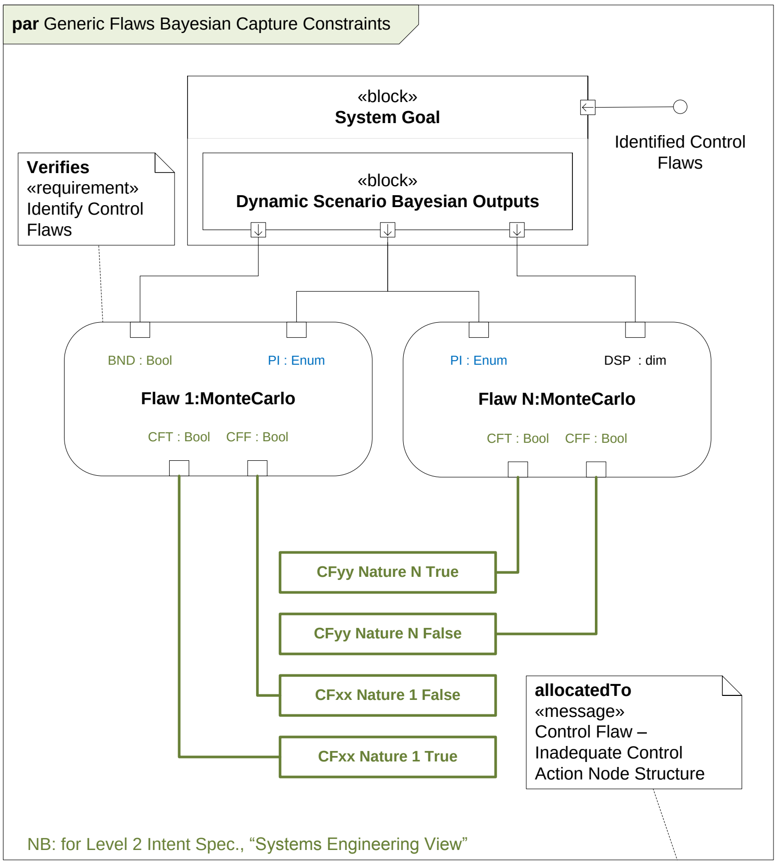


Figure 8.6 Generic Flaw Bayesian Capture Parametric Model

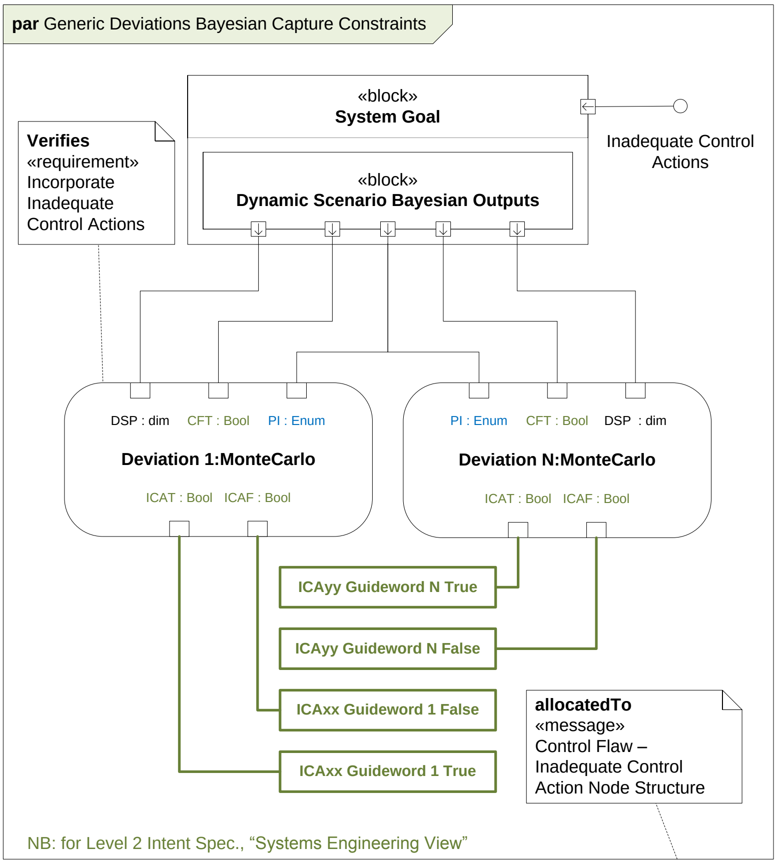


Figure 8.7 Generic Deviation Bayesian Capture Parametric Model

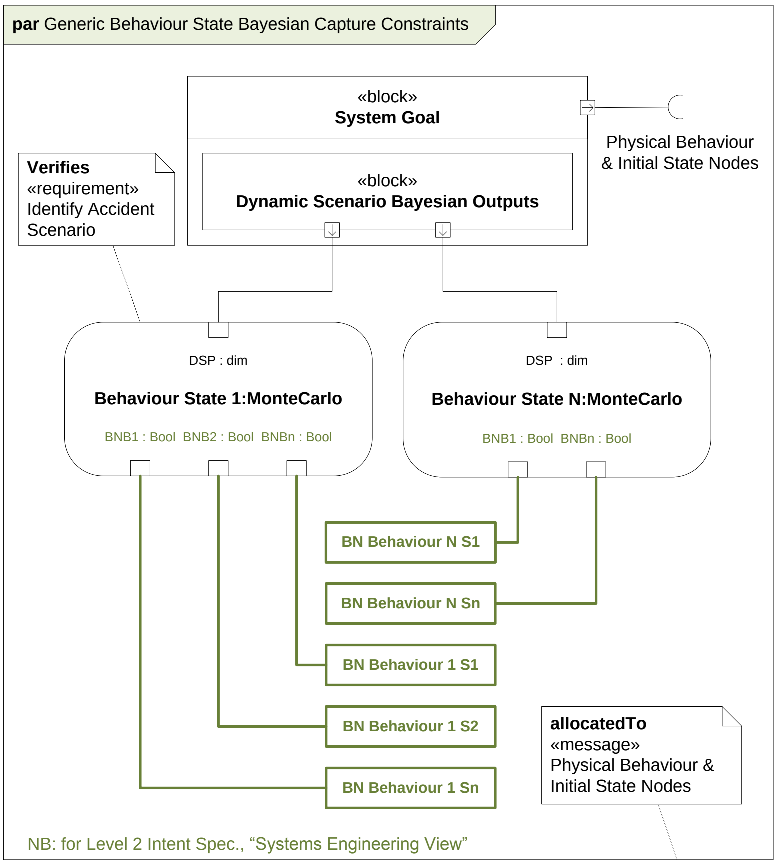

par Generic Behaviour State Bayesian Capture Constraints
«block»
System Goal
«block»
Dynamic Scenario Bayesian Outputs
Physical Behaviour & Initial State Nodes
Verifies
«requirement»
Identify Accident Scenario
DSP : dim
Behaviour State 1:MonteCarlo
BNB1 : Bool
BNB2 : Bool
BNBn : Bool
DSP : dim
Behaviour State N:MonteCarlo
BNB1 : Bool
BNBn : Bool
BN Behaviour N S1
BN Behaviour N Sn
BN Behaviour 1 S1
BN Behaviour 1 S2
BN Behaviour 1 Sn
allocatedTo
«message»
Physical Behaviour & Initial State Nodes
NB: for Level 2 Intent Spec., “Systems Engineering View”

Figure 8.8 Generic Behaviour State Bayesian Capture Parametric Model

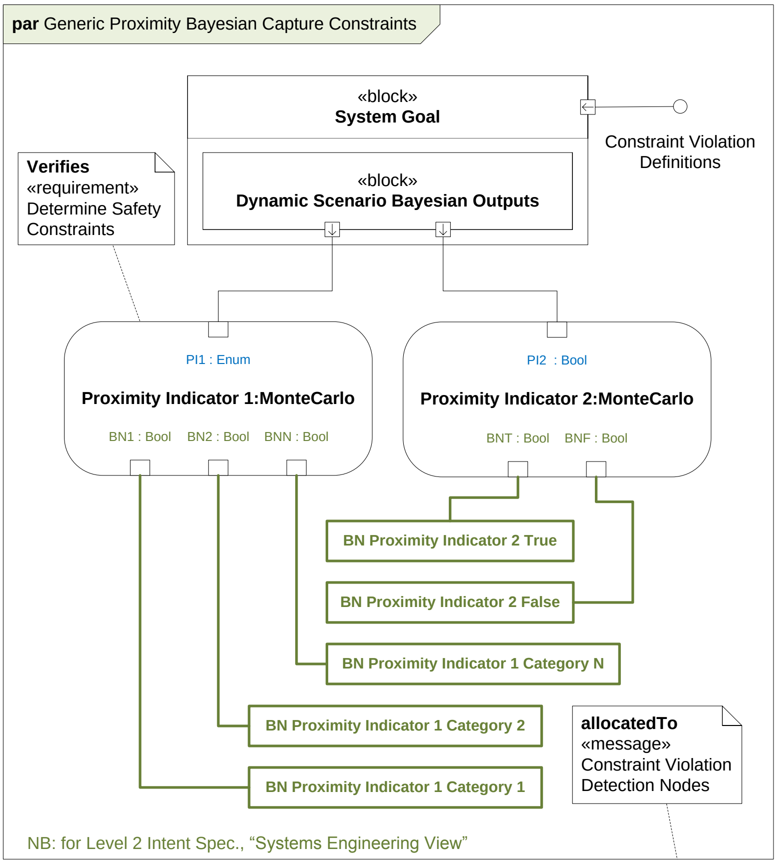

par Generic Proximity Bayesian Capture Constraints
«block»
System Goal
«block»
Dynamic Scenario Bayesian Outputs
Constraint Violation Definitions
Verifies
«requirement»
Determine Safety Constraints
PI1 : Enum
Proximity Indicator 1:MonteCarlo
BN1 : Bool
BN2 : Bool
BNN : Bool
PI2 : Bool
Proximity Indicator 2:MonteCarlo
BNT : Bool
BNF : Bool
BN Proximity Indicator 2 True
BN Proximity Indicator 2 False
BN Proximity Indicator 1 Category N
BN Proximity Indicator 1 Category 2
BN Proximity Indicator 1 Category 1
allocatedTo
«message»
Constraint Violation Detection Nodes
NB: for Level 2 Intent Spec., "Systems Engineering View"

Figure 8.9 Generic Proximity Bayesian Capture Parametric Model

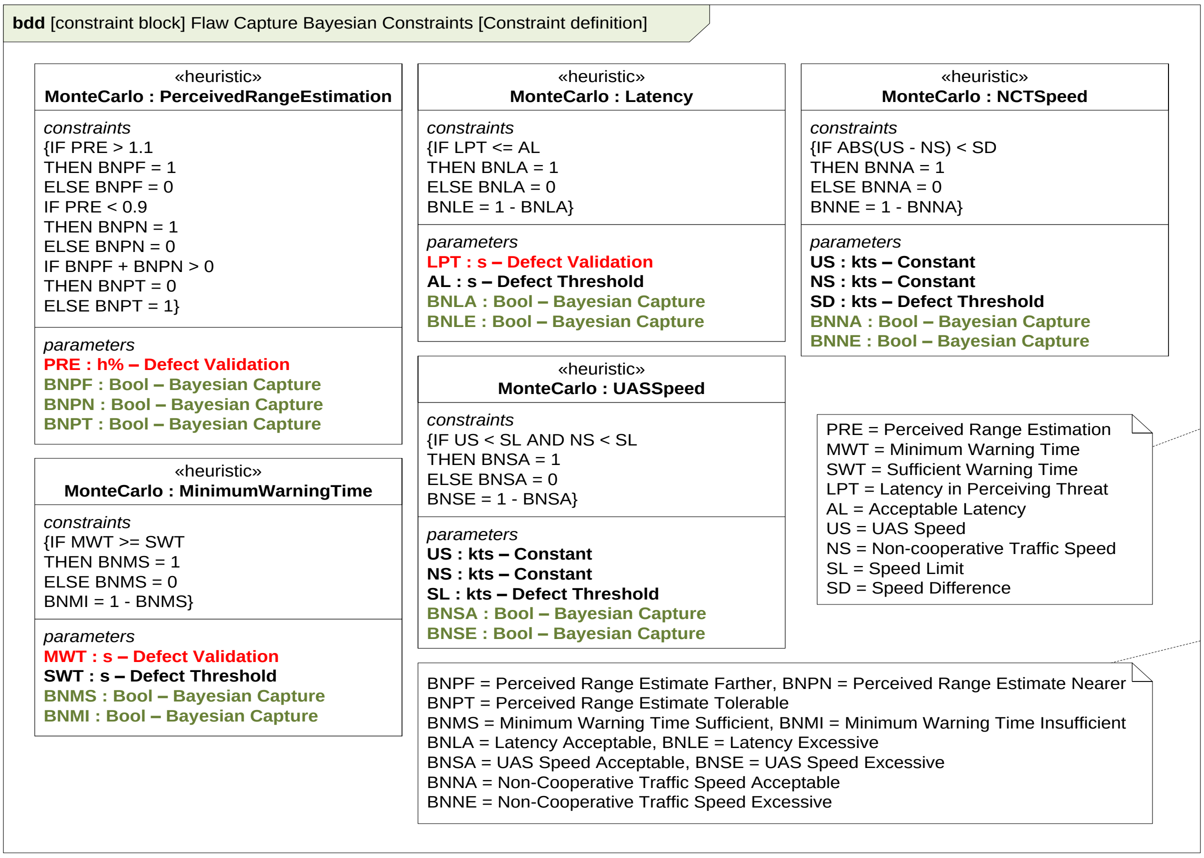

bdd [constraint block] Flaw Capture Bayesian Constraints [Constraint definition]
«heuristic»
MonteCarlo : PerceivedRangeEstimation
constraints
{IF PRE > 1.1
THEN BNPF = 1
ELSE BNPF = 0
IF PRE < 0.9
THEN BNPN = 1
ELSE BNPN = 0
IF BNPF + BNPN > 0
THEN BNPT = 0
ELSE BNPT = 1}
parameters
PRE : h% – Defect Validation
BNPF : Bool – Bayesian Capture
BNPN : Bool – Bayesian Capture
BNPT : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Latency
constraints
{IF LPT <= AL
THEN BNLA = 1
ELSE BNLA = 0
BNLE = 1 - BNLA}
parameters
LPT : s – Defect Validation
AL : s – Defect Threshold
BNLA : Bool – Bayesian Capture
BNLE : Bool – Bayesian Capture
«heuristic»
MonteCarlo : NCTSpeed
constraints
{IF ABS(US - NS) < SD
THEN BNNA = 1
ELSE BNNA = 0
BNNE = 1 - BNNA}
parameters
US : kts – Constant
NS : kts – Constant
SD : kts – Defect Threshold
BNNA : Bool – Bayesian Capture
BNNE : Bool – Bayesian Capture
«heuristic»
MonteCarlo : UASSpeed
constraints
{IF US < SL AND NS < SL
THEN BNSA = 1
ELSE BNSA = 0
BNSE = 1 - BNSA}
parameters
US : kts – Constant
NS : kts – Constant
SL : kts – Defect Threshold
BNSA : Bool – Bayesian Capture
BNSE : Bool – Bayesian Capture
«heuristic»
MonteCarlo : MinimumWarningTime
constraints
{IF MWT >= SWT
THEN BNMS = 1
ELSE BNMS = 0
BNMI = 1 - BNMS}
parameters
MWT : s – Defect Validation
SWT : s – Defect Threshold
BNMS : Bool – Bayesian Capture
BNMI : Bool – Bayesian Capture
PRE = Perceived Range Estimation
MWT = Minimum Warning Time
SWT = Sufficient Warning Time
LPT = Latency in Perceiving Threat
AL = Acceptable Latency
US = UAS Speed
NS = Non-cooperative Traffic Speed
SL = Speed Limit
SD = Speed Difference
BNPF = Perceived Range Estimate Farther, BNPN = Perceived Range Estimate Nearer
BNPT = Perceived Range Estimate Tolerable
BNMS = Minimum Warning Time Sufficient, BNMI = Minimum Warning Time Insufficient
BNLA = Latency Acceptable, BNLE = Latency Excessive
BNSA = UAS Speed Acceptable, BNSE = UAS Speed Excessive
BNNA = Non-Cooperative Traffic Speed Acceptable
BNNE = Non-Cooperative Traffic Speed Excessive

Figure 8.10 SAA Scenario Defect Capture Bayesian Data Constraints

Each heuristic model usually returns two values – "True" and "False", where each respectively takes a value of nought or one, and obviously for each value taken one output should complement the other. Whilst this approach appears to be redundant, this obvious disambiguation does help guarantee a consistent interface to the Bayesian model where nodes in general might embody an arbitrary number of states. In the case of the system defect capture constraint definitions these are represented on the previous page in Figure 8.10, and are also implemented in the system dynamics model with equations (041-3), (024-5), (022-3), (048-9) and (028-9):

```
(041)   BN PRE Farther = IF THEN ELSE (Perceived Range Estimation > 1.1, 1, 0)
(042)   BN PRE Nearer = IF THEN ELSE (Perceived Range Estimation < 0.9, 1, 0)
(043)   BN PRE Tolerable = IF THEN ELSE (BN PRE Farther + BN PRE Nearer > 0, 0, 1)
(024)   BN MWT Insufficient = 1 - BN MWT Sufficient
(025)   BN MWT Sufficient =
        IF THEN ELSE (Minimum Warning Time >= Sufficient Warning Time, 1, 0)
```

```
(022)   BN L Acceptable =
        IF THEN ELSE (Latency in Perceiving Threat <= Acceptable Latency, 1, 0)
(023)   BN L Excessive = 1 - BN L Acceptable
(048)   BN Speed Acceptable =
        IF THEN ELSE ((UAS Speed < Speed Limit)
        :AND: ("Non-Coop Traffic Speed" < Speed Limit), 1, 0)
(049)   BN Speed Excessive = 1 - BN Speed Acceptable
(028)   BN NCTS Acceptable =
        IF THEN ELSE (ABS(UAS Speed - "Non-Coop Traffic Speed") < Speed Difference, 1, 0)
(029)   BN NCTS Excessive = 1 - BN NCTS Acceptable
```

However, of particular note in this case is the fact that these heuristic blocks define these output values slightly differently from the general cases as applied to the Flaw and Deviation capture models. Notably, one of the above blocks contains a tri-state output representing events that may occur above, below, and within a range of tolerance – for example: too far, too near, and tolerable. Also, although the remainder of the above blocks do represent bi-state outputs these are more descriptively described as the defective behaviour being captured is not necessarily best described as logical values, True or False. Taken together these five heuristic blocks describe the requirement to capture the three explicit defective design precursors, as related to the Bayesian nodes shown in Figure 8.10, along with the two implied speed related defect precursors which are related to observable physical behaviour nodes.

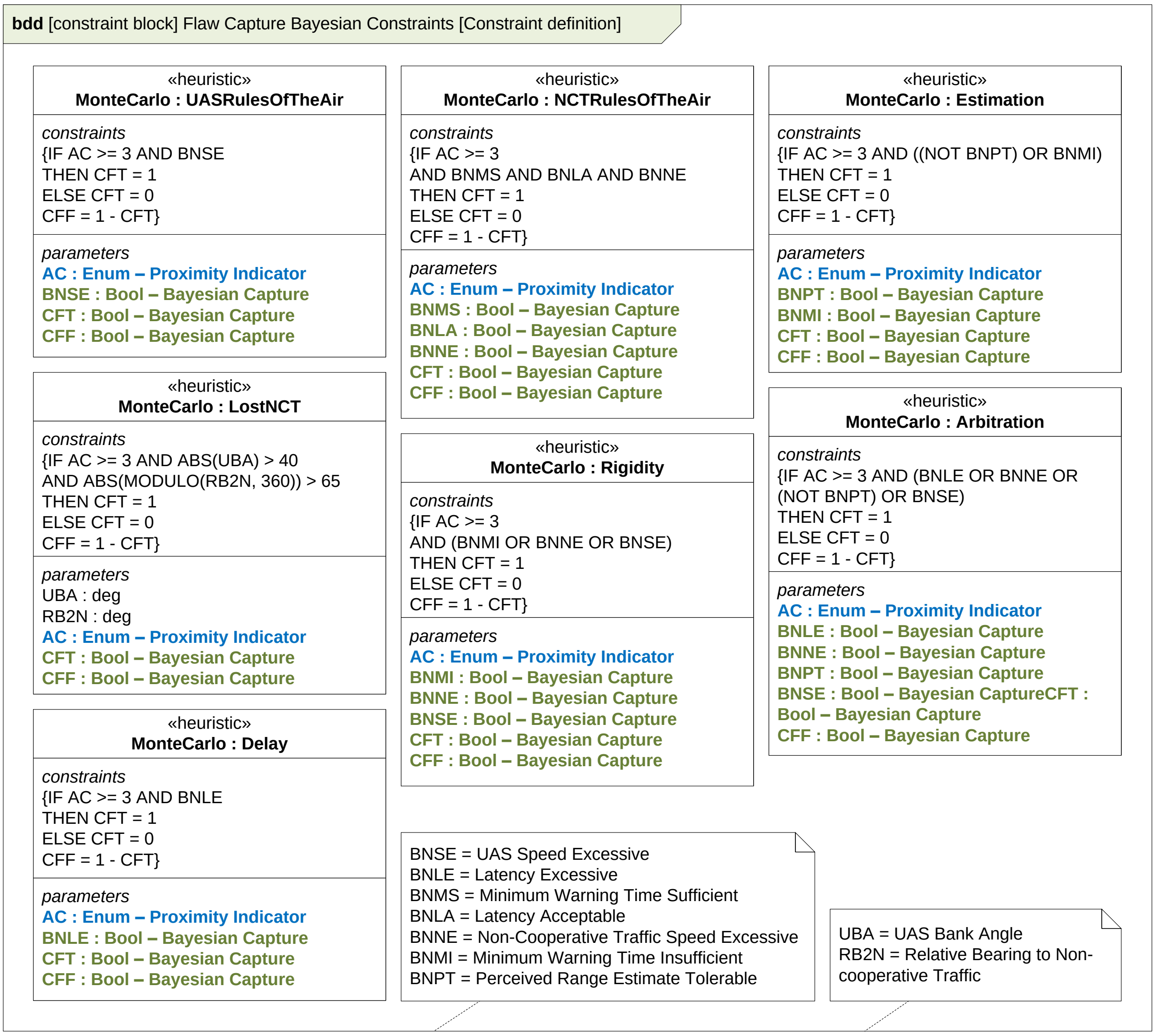

bdd [constraint block] Flaw Capture Bayesian Constraints [Constraint definition]
«heuristic»
MonteCarlo : UASRulesOfTheAir
constraints
{IF AC >= 3 AND BNSE
THEN CFT = 1
ELSE CFT = 0
CFF = 1 - CFT}
parameters
AC : Enum – Proximity Indicator
BNSE : Bool – Bayesian Capture
CFT : Bool – Bayesian Capture
CFF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : NCTRulesOfTheAir
constraints
{IF AC >= 3
AND BNMS AND BNLA AND BNNE
THEN CFT = 1
ELSE CFT = 0
CFF = 1 - CFT}
parameters
AC : Enum – Proximity Indicator
BNMS : Bool – Bayesian Capture
BNLA : Bool – Bayesian Capture
BNNE : Bool – Bayesian Capture
CFT : Bool – Bayesian Capture
CFF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Estimation
constraints
{IF AC >= 3 AND ((NOT BNPT) OR BNMI)
THEN CFT = 1
ELSE CFT = 0
CFF = 1 - CFT}
parameters
AC : Enum – Proximity Indicator
BNPT : Bool – Bayesian Capture
BNMI : Bool – Bayesian Capture
CFT : Bool – Bayesian Capture
CFF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : LostNCT
constraints
{IF AC >= 3 AND ABS(UBA) > 40
AND ABS(MODULO(RB2N, 360)) > 65
THEN CFT = 1
ELSE CFT = 0
CFF = 1 - CFT}
parameters
UBA : deg
RB2N : deg
AC : Enum – Proximity Indicator
CFT : Bool – Bayesian Capture
CFF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Rigidity
constraints
{IF AC >= 3
AND (BNMI OR BNNE OR BNSE)
THEN CFT = 1
ELSE CFT = 0
CFF = 1 - CFT}
parameters
AC : Enum – Proximity Indicator
BNMI : Bool – Bayesian Capture
BNNE : Bool – Bayesian Capture
BNSE : Bool – Bayesian Capture
CFT : Bool – Bayesian Capture
CFF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Arbitration
constraints
{IF AC >= 3 AND (BNLE OR BNNE OR
(NOT BNPT) OR BNSE)
THEN CFT = 1
ELSE CFT = 0
CFF = 1 - CFT}
parameters
AC : Enum – Proximity Indicator
BNLE : Bool – Bayesian Capture
BNNE : Bool – Bayesian Capture
BNPT : Bool – Bayesian Capture
BNSE : Bool – Bayesian CaptureCFT :
Bool – Bayesian Capture
CFF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Delay
constraints
{IF AC >= 3 AND BNLE
THEN CFT = 1
ELSE CFT = 0
CFF = 1 - CFT}
parameters
AC : Enum – Proximity Indicator
BNLE : Bool – Bayesian Capture
CFT : Bool – Bayesian Capture
CFF : Bool – Bayesian Capture
BNSE = UAS Speed Excessive
BNLE = Latency Excessive
BNMS = Minimum Warning Time Sufficient
BNLA = Latency Acceptable
BNNE = Non-Cooperative Traffic Speed Excessive
BNMI = Minimum Warning Time Insufficient
BNPT = Perceived Range Estimate Tolerable
UBA = UAS Bank Angle
RB2N = Relative Bearing to Non-cooperative Traffic

Figure 8.11 SAA Scenario Flaw Capture Bayesian Data Constraints

Selected Control Flaws as defined in Chapter 6 are now added as event capture heuristics encompassing: CF04, "Estimation"; CF05, "UASRulesOfTheAir"; CF06, "NCTRulesOfTheAir"; CF07, "Rigidity"; CF08, "LostNCT"; CF10, "Delay"; and CF15, "Arbitration", as detailed above in Figure 8.11. Note that any behaviour indicating "Don't Know" or unknown cases is not captured from the simulations as by definition it is not directly observable and only exists by implication. It is for this reason that the untrained Bayesian network is preconfigured with a significant bias towards the unknown and uncertain – although not entirely naive; thereby the conditioning of the respective Bayesian nodes is later biased away from the uncertain model by the training data derived only from "known" correlated events and data, as observed within the Monte Carlo outputs taken from the model and as captured above.

The flaw (CFxx) capture constraint definitions as represented in Figure 8.11 are implemented in the system dynamics model with equations (056-7), (062-3), (064-5), (058-9), (060-1), (054-5) and (066-7), respectively:

(056) CF05 UASRoTA False = 1 - CF05 UASRoTA True

(057) CF05 UASRoTA True = IF THEN ELSE ((Airprox Categories >= 3) :AND: BN Speed Excessive, 1, 0)

CF05: The UAV assessment model shall incorporate a mechanism that fails to follow established Rules of the Air.

(062) CF08 LOSTNCT False = 1 - CF08 LOSTNCT True

(063) CF08 LOSTNCT True = IF THEN ELSE ((Airprox Categories >= 3) :AND: (ABS(UAS Bank Angle) > 40) :AND: (ABS (MODULO (relative bearing to NCT, 360)) > 65), 1, 0)

CF08: The UAV assessment model shall incorporate a mechanism that causes the loss of detection of the other aircraft when manoeuvring.

(064) CF10 DELAY False = 1 - CF10 DELAY True

(065) CF10 DELAY True = IF THEN ELSE ((Airprox Categories >= 3) :AND: BN L Excessive, 1, 0)

CF10: The UAV assessment model shall incorporate a mechanism that delays the commanding of a turning manoeuvre from the flight control system.

(058) CF06 NCTRoTA False = 1 - CF06 NCTRoTA True

(059) CF06 NCTRoTA True = IF THEN ELSE ((Airprox Categories >= 3) :AND: BN MWT Sufficient :AND: BN L Acceptable :AND: BN NCTS Excessive, 1, 0)

CF06: The UAV assessment model shall incorporate a mechanism whereby the other aircraft fails to follow established Rules of the Air.

(060) CF07 RIGIDITY False = 1 - CF07 RIGIDITY True

(061) CF07 RIGIDITY True = IF THEN ELSE ((Airprox Categories >= 3) :AND: (BN MWT Insufficient :OR: BN NCTS Excessive :OR: BN Speed Excessive), 1, 0)

CF07: The UAV assessment model shall incorporate a mechanism whereby the UAV interprets the Rules of the Air too rigidly when closing within 500ft of the other aircraft.

(054) CF04 ESTIMATION False = 1 - CF04 ESTIMATION True

(055) CF04 ESTIMATION True = IF THEN ELSE ((Airprox Categories >= 3) :AND: ((:NOT: BN PRE Tolerable) :OR: BN MWT Insufficient), 1, 0)

CF04: The UAV assessment model shall incorporate a mechanism that incorrectly estimates the location and direction to the predicted point of conflict.

(066) CF15 ARBITRATION False = 1 - CF15 ARBITRATION True

(067) CF15 ARBITRATION True = IF THEN ELSE ((Airprox Categories >= 3) :AND: (BN L Excessive :OR: BN NCTS Excessive :OR: (:NOT: BN PRE Tolerable) :OR: BN Speed Excessive), 1, 0)

CF15: The UAV assessment model shall incorporate a mechanism whereby the UAV does not arbitrate the competing heading demands appropriately.

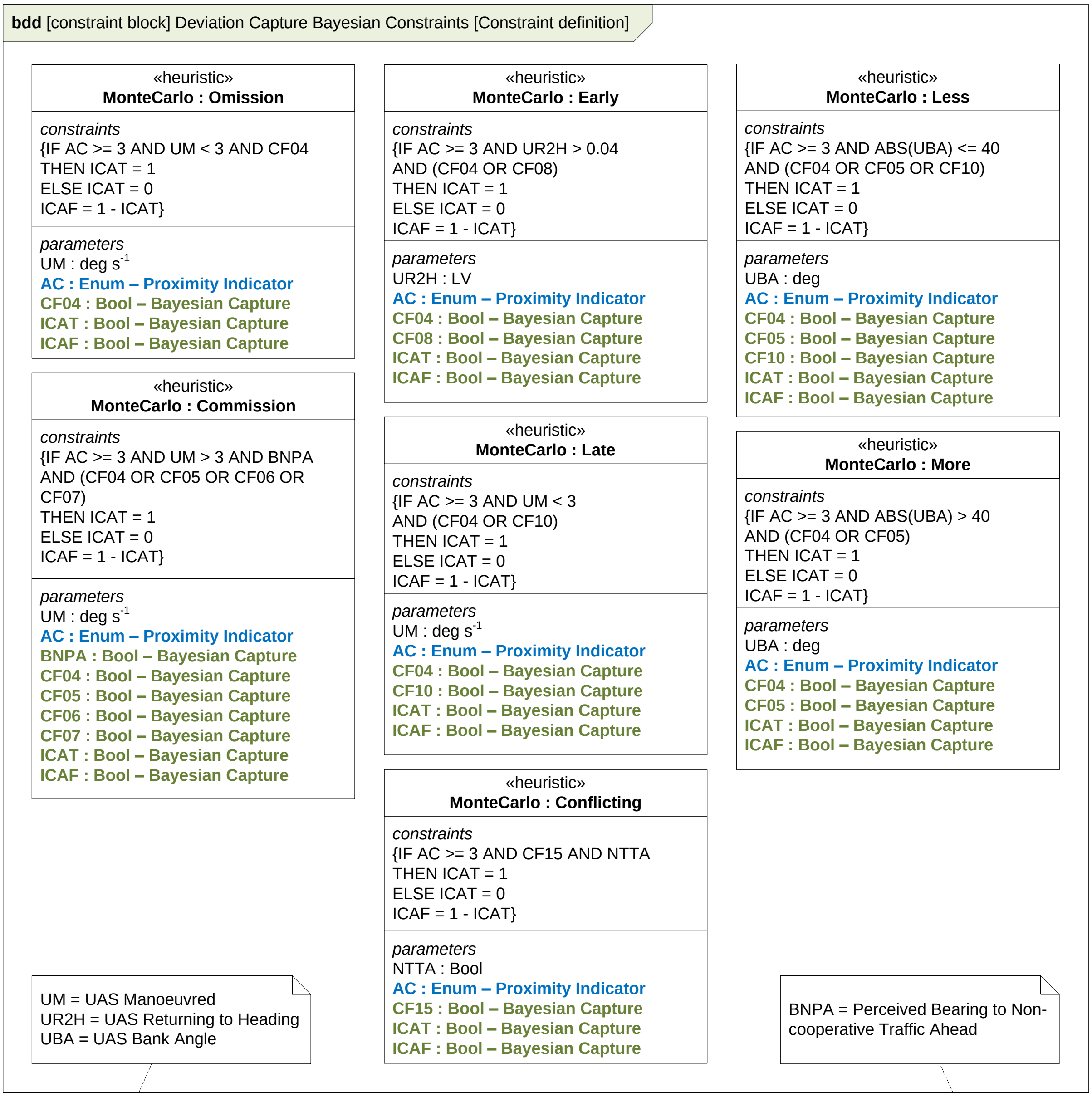

bdd [constraint block] Deviation Capture Bayesian Constraints [Constraint definition]
«heuristic»
MonteCarlo : Omission
constraints
{IF AC >= 3 AND UM < 3 AND CF04
THEN ICAT = 1
ELSE ICAT = 0
ICAF = 1 - ICAT}
parameters
UM : deg s-1
AC : Enum – Proximity Indicator
CF04 : Bool – Bayesian Capture
ICAT : Bool – Bayesian Capture
ICAF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Commission
constraints
{IF AC >= 3 AND UM > 3 AND BNPA
AND (CF04 OR CF05 OR CF06 OR
CF07)
THEN ICAT = 1
ELSE ICAT = 0
ICAF = 1 - ICAT}
parameters
UM : deg s-1
AC : Enum – Proximity Indicator
BNPA : Bool – Bayesian Capture
CF04 : Bool – Bayesian Capture
CF05 : Bool – Bayesian Capture
CF06 : Bool – Bayesian Capture
CF07 : Bool – Bayesian Capture
ICAT : Bool – Bayesian Capture
ICAF : Bool – Bayesian Capture
UM = UAS Manoeuvred
UR2H = UAS Returning to Heading
UBA = UAS Bank Angle
«heuristic»
MonteCarlo : Early
constraints
{IF AC >= 3 AND UR2H > 0.04
AND (CF04 OR CF08)
THEN ICAT = 1
ELSE ICAT = 0
ICAF = 1 - ICAT}
parameters
UR2H : LV
AC : Enum – Proximity Indicator
CF04 : Bool – Bayesian Capture
CF08 : Bool – Bayesian Capture
ICAT : Bool – Bayesian Capture
ICAF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Late
constraints
{IF AC >= 3 AND UM < 3
AND (CF04 OR CF10)
THEN ICAT = 1
ELSE ICAT = 0
ICAF = 1 - ICAT}
parameters
UM : deg s-1
AC : Enum – Proximity Indicator
CF04 : Bool – Bayesian Capture
CF10 : Bool – Bayesian Capture
ICAT : Bool – Bayesian Capture
ICAF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Conflicting
constraints
{IF AC >= 3 AND CF15 AND NTTA
THEN ICAT = 1
ELSE ICAT = 0
ICAF = 1 - ICAT}
parameters
NTTA : Bool
AC : Enum – Proximity Indicator
CF15 : Bool – Bayesian Capture
ICAT : Bool – Bayesian Capture
ICAF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : Less
constraints
{IF AC >= 3 AND ABS(UBA) <= 40
AND (CF04 OR CF05 OR CF10)
THEN ICAT = 1
ELSE ICAT = 0
ICAF = 1 - ICAT}
parameters
UBA : deg
AC : Enum – Proximity Indicator
CF04 : Bool – Bayesian Capture
CF05 : Bool – Bayesian Capture
CF10 : Bool – Bayesian Capture
ICAT : Bool – Bayesian Capture
ICAF : Bool – Bayesian Capture
«heuristic»
MonteCarlo : More
constraints
{IF AC >= 3 AND ABS(UBA) > 40
AND (CF04 OR CF05)
THEN ICAT = 1
ELSE ICAT = 0
ICAF = 1 - ICAT}
parameters
UBA : deg
AC : Enum – Proximity Indicator
CF04 : Bool – Bayesian Capture
CF05 : Bool – Bayesian Capture
ICAT : Bool – Bayesian Capture
ICAF : Bool – Bayesian Capture
BNPA = Perceived Bearing to Non-cooperative Traffic Ahead

Figure 8.12 SAA Scenario Deviation Capture Bayesian Data Constraints

Likewise the Inadequate Control Actions, otherwise known as Deviations, also as defined in Chapter 6, are now added as event capture heuristics encompassing: ICA01 – Omission; ICA02 – Commission; ICA03 – Early; ICA04 – Late; ICA05 – Less; ICA06 – More; and ICA07 – Conflicting. As can be seen above, in Figure 8.12, these blocks also reference the corresponding Control Flaw capture parameters with which these deviation behaviours are to be associated. Again, the "Don't Know" state is implied in the absence of an event for any particular node, and as also stated earlier both these heuristics, and those for the Control Flaws in Figure 8.11, utilise an active proximity safety constraint violation indicator to trigger the capture of an event – otherwise at other times within the context of simulation the event relationships remain unknown.

The deviation (ICAxx) capture constraint definitions as represented in Figure 8.12 are implemented in the system dynamics model with equations (078-91):

(078) ICA01 OMISSION False = 1 - ICA01 OMISSION True

(079) ICA01 OMISSION True = IF THEN ELSE ((Airprox Categories >= 3) :AND: (UAS Manoeuvred < 3) :AND: CF04 ESTIMATION True, 1, 0)

ICA01: The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV does not manoeuvre to avoid the other aircraft.

(080) ICA02 COMMISION False = 1 - ICA02 COMMISION True

(081) ICA02 COMMISION True = IF THEN ELSE ((Airprox Categories >= 3) :AND: (UAS Manoeuvred > 3) :AND: BN perceivedBearingToNCT Ahead :AND: (CF04 ESTIMATION True :OR: CF05 UASRoTA True :OR: CF06 NCTRoTA True :OR: CF07 RIGIDITY True), 1, 0)

ICA02: The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV manoeuvres in the wrong direction attempting to avoid the the other aircraft.

(082) ICA03 EARLY False = 1 - ICA03 EARLY True

(083) ICA03 EARLY True = IF THEN ELSE ((Airprox Categories >= 3) :AND: (UAS returning to heading > 0.04) :AND: (CF04 ESTIMATION True :OR: CF08 LOSTNCT True), 1, 0)

ICA03: The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV prematurely ceases a manoeuvre to avoid the other aircraft and returns to the given (initial) heading.

(084) ICA04 LATE False = 1 - ICA04 LATE True

(085) ICA04 LATE True = IF THEN ELSE ((Airprox Categories >= 3) :AND: (UAS Manoeuvred < 3) :AND: (CF04 ESTIMATION True :OR: CF10 DELAY True), 1, 0)

ICA04: The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV belatedly initiates a manoeuvre to avoid the other aircraft.

(086) ICA05 LESS False = 1 - ICA05 LESS True

(087) ICA05 LESS True = IF THEN ELSE ((Airprox Categories >= 3) :AND: (ABS (UAS Bank Angle) <= 40) :AND: (CF04 ESTIMATION True :OR: CF05 UASRoTA True :OR: CF10 DELAY True), 1, 0)

ICA05: The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV rate of manoeuvre is insufficient to avoid the other aircraft.

(088) ICA06 MORE False = 1 - ICA06 MORE True

(089) ICA06 MORE True = IF THEN ELSE ((Airprox Categories >= 3) :AND:
(ABS (UAS Bank Angle) > 40) :AND: (CF04 ESTIMATION True :OR:
CF05 UASRoTA True), 1, 0)

ICA06: The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV rate of manoeuvre is excessive in attempting to avoid the other aircraft.

(090) ICA07 CONFLICTING False = 1 - ICA07 CONFLICTING True

(091) ICA07 CONFLICTING True = IF THEN ELSE ((Airprox Categories >= 3) :AND:
CF15 ARBITRATION True :AND: NCT threat arises again, 1, 0)

ICA07: The assessment model for the UAS shall provide a means to produce a deviation in the manoeuvre behaviour whereby the UAV is unable to resolve the conflict between realising the System Goal and avoiding the other aircraft.

8.1.3. System State and Violation Capture – Defining Monte Carlo Generated Cases

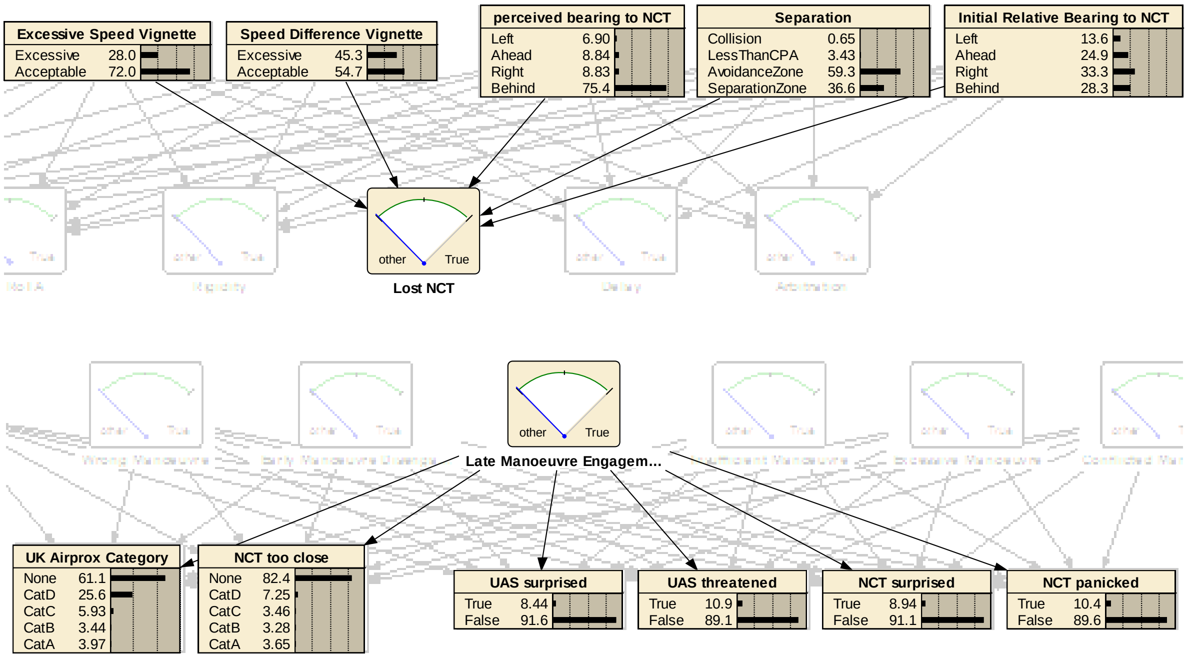
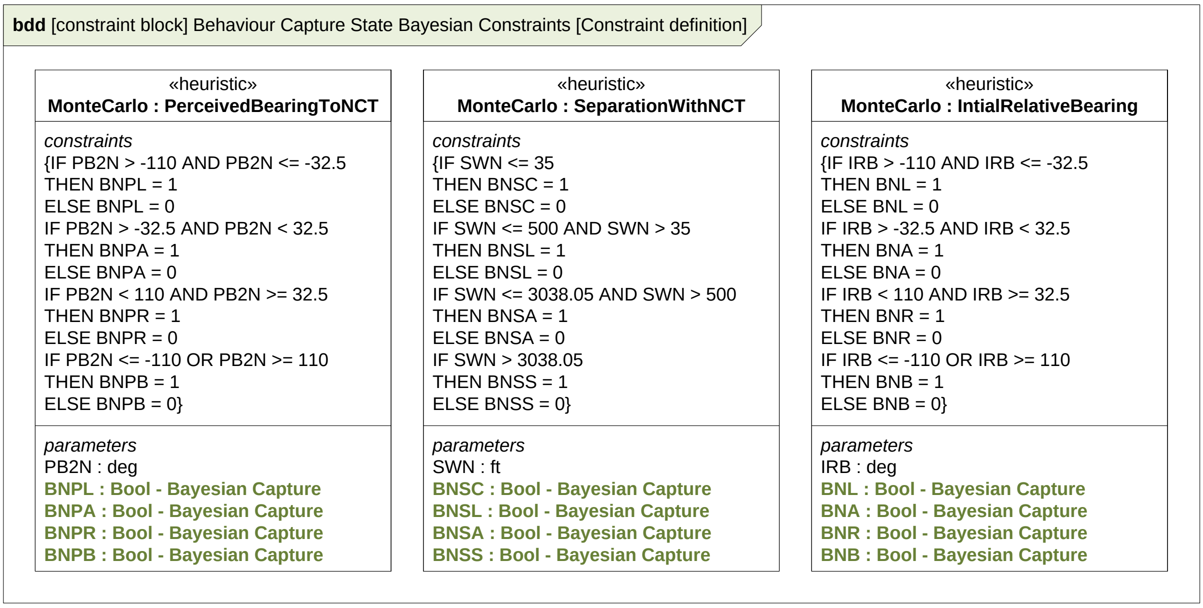


Figure 8.13 Airprox constraint violation & system state association in the Bayesian Network

For the purposes of this study it has been decided not to use a fully connected naïve Bayesian network, but rather structure the skeleton network around beliefs as expressed within a HAZOP study. However, particular physical relationships observed throughout a sequence of events might be used to contextualise the whole range of system flaws with implied, but unknown, casual relationships, as represented in Figure 8.13 above. Therefore, the states of these five observable physical properties or behaviours are coupled directly to all of the identified flaws with initially unknown probabilities, and with an initially equally likely probability across each set of states.



bdd [constraint block] Behaviour Capture State Bayesian Constraints [Constraint definition]

«heuristic»
MonteCarlo : PerceivedBearingToNCT

constraints
{IF PB2N > -110 AND PB2N <= -32.5
THEN BNPL = 1
ELSE BNPL = 0
IF PB2N > -32.5 AND PB2N < 32.5
THEN BNPA = 1
ELSE BNPA = 0
IF PB2N < 110 AND PB2N >= 32.5
THEN BNPR = 1
ELSE BNPR = 0
IF PB2N <= -110 OR PB2N >= 110
THEN BNPB = 1
ELSE BNPB = 0}

parameters
PB2N : deg
BNPL : Bool - Bayesian Capture
BNPA : Bool - Bayesian Capture
BNPR : Bool - Bayesian Capture
BNPB : Bool - Bayesian Capture

«heuristic»
MonteCarlo : SeparationWithNCT

constraints
{IF SWN <= 35
THEN BNSC = 1
ELSE BNSC = 0
IF SWN <= 500 AND SWN > 35
THEN BNSL = 1
ELSE BNSL = 0
IF SWN <= 3038.05 AND SWN > 500
THEN BNSA = 1
ELSE BNSA = 0
IF SWN > 3038.05
THEN BNSS = 1
ELSE BNSS = 0}

parameters
SWN : ft
BNSC : Bool - Bayesian Capture
BNSL : Bool - Bayesian Capture
BNSA : Bool - Bayesian Capture
BNSS : Bool - Bayesian Capture

«heuristic»
MonteCarlo : IntialRelativeBearing

constraints
{IF IRB > -110 AND IRB <= -32.5
THEN BNL = 1
ELSE BNL = 0
IF IRB > -32.5 AND IRB < 32.5
THEN BNA = 1
ELSE BNA = 0
IF IRB < 110 AND IRB >= 32.5
THEN BNR = 1
ELSE BNR = 0
IF IRB <= -110 OR IRB >= 110
THEN BNB = 1
ELSE BNB = 0}

parameters
IRB : deg
BNL : Bool - Bayesian Capture
BNA : Bool - Bayesian Capture
BNR : Bool - Bayesian Capture
BNB : Bool - Bayesian Capture

Figure 8.14 SAA Scenario Behaviour State Capture Bayesian Data Constraints

For the purpose of behaviour capture, the constraining heuristics for three assumed possible physical precursors are set out in Figure 8.14, above – noting that the speed related events have already been captured as potential defects in Figure 8.10. The system relative physical behaviour constraint definitions as represented above are implemented in the system dynamics model with equations (037-40), (044-7) and (018-21), respectively:

(037) BN perceivedBearingToNCT Ahead = IF THEN ELSE (perceived bearing to NCT > -32.5 :AND: perceived bearing to NCT < 32.5, 1, 0)

(038) BN perceivedBearingToNCT Behind = IF THEN ELSE (perceived bearing to NCT <= -110 :OR: perceived bearing to NCT >= 110, 1, 0)

(039) BN perceivedBearingToNCT Left = IF THEN ELSE (perceived bearing to NCT > -110 :AND: perceived bearing to NCT <= -32.5, 1, 0)

(040) BN perceivedBearingToNCT Right = IF THEN ELSE (perceived bearing to NCT < 110 :AND: perceived bearing to NCT >= 32.5, 1, 0)

(044) BN Separation AvoidanceZone = IF THEN ELSE (separation with NCT <= 3038.05 :AND: separation with NCT > 500, 1, 0)

(045) BN Separation Collision = IF THEN ELSE (separation with NCT <= 35, 1, 0)

(046) BN Separation LessThanCPA = IF THEN ELSE (separation with NCT <= 500 :AND: separation with NCT > 35, 1, 0)

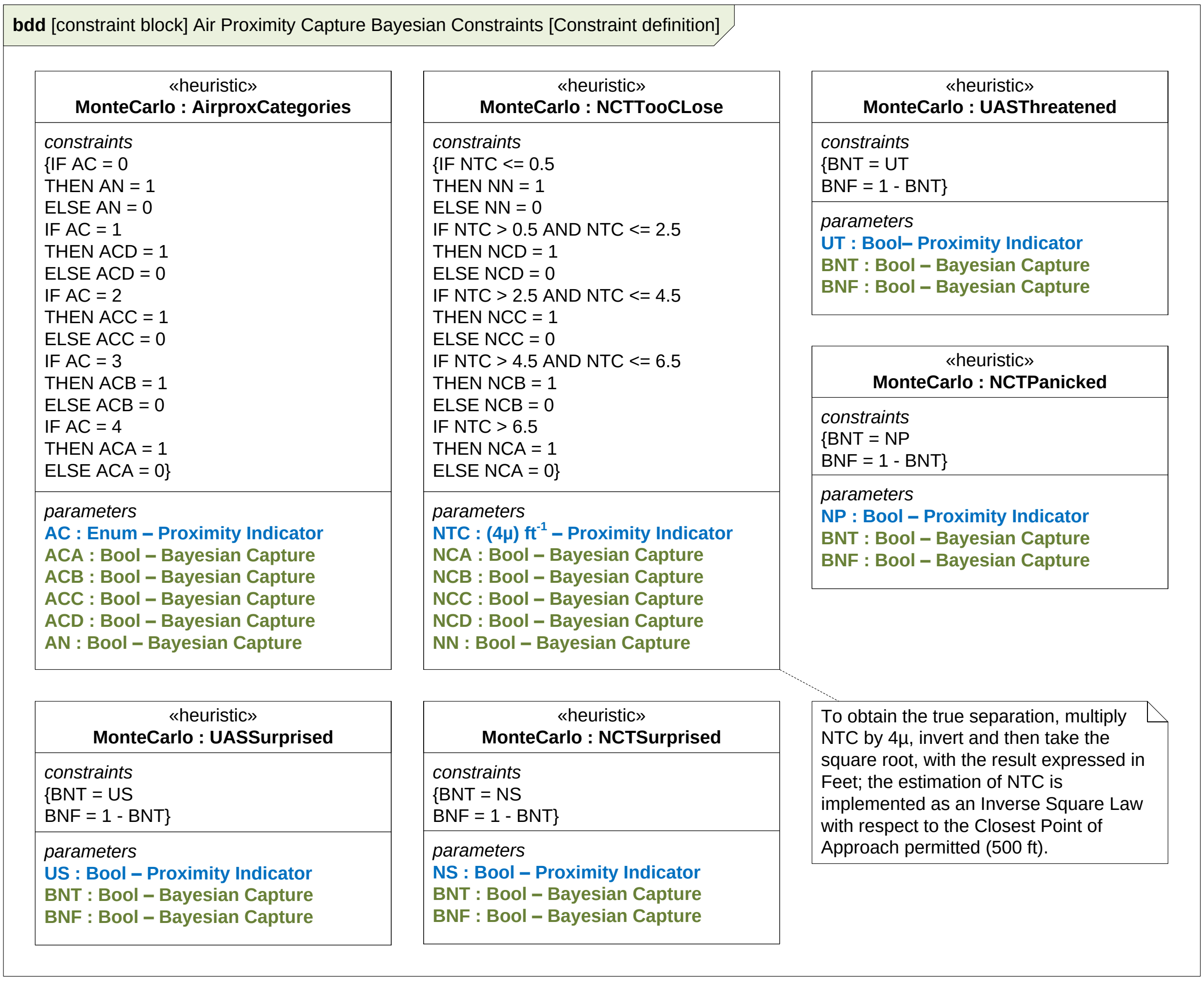
(047) BN Separation SeparationZone = IF THEN ELSE (separation with NCT > 3038.05, 1, 0)
6076.11 feet per nautical mile.

(018) BN InitialBearing Ahead = IF THEN ELSE (initial relative bearing > -32.5 :AND: initial relative bearing < 32.5, 1, 0)

(019) BN InitialBearing Behind = IF THEN ELSE (initial relative bearing <= -110 :OR: initial relative bearing >= 110, 1, 0)

(020) BN InitialBearing Left = IF THEN ELSE (initial relative bearing > -110 :AND: initial relative bearing <= -32.5, 1, 0)

(021) BN InitialBearing Right = IF THEN ELSE (initial relative bearing < 110 :AND: initial relative bearing >= 32.5, 1, 0)



bdd [constraint block] Air Proximity Capture Bayesian Constraints [Constraint definition]

«heuristic»
MonteCarlo : AirproxCategories

constraints
{IF AC = 0
THEN AN = 1
ELSE AN = 0
IF AC = 1
THEN ACD = 1
ELSE ACD = 0
IF AC = 2
THEN ACC = 1
ELSE ACC = 0
IF AC = 3
THEN ACB = 1
ELSE ACB = 0
IF AC = 4
THEN ACA = 1
ELSE ACA = 0}

parameters
AC : Enum – Proximity Indicator
ACA : Bool – Bayesian Capture
ACB : Bool – Bayesian Capture
ACC : Bool – Bayesian Capture
ACD : Bool – Bayesian Capture
AN : Bool – Bayesian Capture

«heuristic»
MonteCarlo : NCTTooCLose

constraints
{IF NTC <= 0.5
THEN NN = 1
ELSE NN = 0
IF NTC > 0.5 AND NTC <= 2.5
THEN NCD = 1
ELSE NCD = 0
IF NTC > 2.5 AND NTC <= 4.5
THEN NCC = 1
ELSE NCC = 0
IF NTC > 4.5 AND NTC <= 6.5
THEN NCB = 1
ELSE NCB = 0
IF NTC > 6.5
THEN NCA = 1
ELSE NCA = 0}

parameters
NTC : $(4\mu)\ \mathrm{ft}^{-1}$ – Proximity Indicator
NCA : Bool – Bayesian Capture
NCB : Bool – Bayesian Capture
NCC : Bool – Bayesian Capture
NCD : Bool – Bayesian Capture
NN : Bool – Bayesian Capture

«heuristic»
MonteCarlo : UASThreatened

constraints
{BNT = UT
BNF = 1 - BNT}

parameters
UT : Bool– Proximity Indicator
BNT : Bool – Bayesian Capture
BNF : Bool – Bayesian Capture

«heuristic»
MonteCarlo : NCTPanicked

constraints
{BNT = NP
BNF = 1 - BNT}

parameters
NP : Bool – Proximity Indicator
BNT : Bool – Bayesian Capture
BNF : Bool – Bayesian Capture

«heuristic»
MonteCarlo : UASSurprised

constraints
{BNT = US
BNF = 1 - BNT}

parameters
US : Bool – Proximity Indicator
BNT : Bool – Bayesian Capture
BNF : Bool – Bayesian Capture

«heuristic»
MonteCarlo : NCTSurprised

constraints
{BNT = NS
BNF = 1 - BNT}

parameters
NS : Bool – Proximity Indicator
BNT : Bool – Bayesian Capture
BNF : Bool – Bayesian Capture

To obtain the true separation, multiply NTC by 4μ, invert and then take the square root, with the result expressed in Feet; the estimation of NTC is implemented as an Inverse Square Law with respect to the Closest Point of Approach permitted (500 ft).

Figure 8.15 SAA Scenario Proximity Capture Bayesian Data Constraints

With the majority of the assumed flow of cause and effect directed from top to bottom in the diagrams shown so far as representing the Bayesian network model, this now leads us to the final assumed effects – the safety constraint violation. As stated previously, the form of the safety constraint is chosen to reflect the nature of near-miss cases arising with system misbehaviour. In this example, in addition to the prime event categorising the severity of the near-miss event, only four out of the seven available “subjective” events are also included separately in the Bayesian network. A more thorough investigation might include all seven. The “Too Close” heuristic is actually a categorisation of separation distance through the violation event sequence. These air proximity categorisation and capture constraint definitions, as represented in Figure 8.15, are implemented in the system dynamics model with equations (013-7), (050-1), (032-6), (030-1), (052-3) and (026-7), respectively:

(013) BN Airprox CatA = IF THEN ELSE (Airprox Categories = 4, 1, 0)

A - Risk of collision: an actual risk of collision existed.

(014) BN Airprox CatB = IF THEN ELSE (Airprox Categories = 3, 1, 0)

B - Safety not assured: the safety of the aircraft was compromised.

(015) BN Airprox CatC = IF THEN ELSE (Airprox Categories = 2, 1, 0)

C - No risk of collision: no risk of collision existed.

(016) BN Airprox CatD = IF THEN ELSE (Airprox Categories = 1, 1, 0)

D - Risk not determined: insufficient information was available to determine the risk involved, or inconclusive or conflicting evidence precluded such determination.

(017) BN Airprox None = IF THEN ELSE (Airprox Categories = 0, 1, 0)

(050) BN UASSurprised False = 1 - BN UASSurprised True

(051) BN UASSurprised True = UAS Surprised

(032) BN NCTTooClose CatA = IF THEN ELSE (NCT Too Close > 6.5, 1, 0)

(033) BN NCTTooClose CatB =
IF THEN ELSE((NCT Too Close > 4.5) :AND: (NCT Too Close <= 6.5), 1, 0)

(034) BN NCTTooClose CatC =
IF THEN ELSE ((NCT Too Close > 2.5) :AND: (NCT Too Close <= 4.5), 1, 0)

(035) BN NCTTooClose CatD =
IF THEN ELSE ((NCT Too Close > 0.5) :AND: (NCT Too Close <= 2.5), 1, 0)

(036) BN NCTTooClose None = IF THEN ELSE (NCT Too Close <= 0.5, 1, 0)

(030) BN NCTSurprised False = 1 - BN NCTSurprised True

(031) BN NCTSurprised True = NCT Surprised

(052) BN UASThreatened False = 1 - BN UASThreatened True

(053) BN UASThreatened True = UAS Threatened

(026) BN NCTPanicked False = 1 - BN NCTPanicked True

(027) BN NCTPanicked True = NCT Panicked

8.2. Model Test Case Discussion

In the construction, and simulation, of a suitably representative system dynamics model, and in the subsequent reduction of the inherent complexity in the output(s) using Bayesian methods and simple measures of complexity such as a Lyapunov exponent, the analysed results may be simplified and incorporated as an enhancement to the initial HAZOP. Within the sequence of activities updating behaviour belief with evidence derived from the system dynamics by carefully selecting a number of representative vignettes, a set of test cases can be identified providing data on specific critical outcomes that ought to be explicitly developed in the HAZOP encompassing complex and emergent behaviour. With this information a hazard scheme is finally introduced denoting measures of hazard, risk, and expectation of emergent behaviour, diagnostic utility, and overall event likelihood, as calculated by equation 8.1:

$$\boldsymbol{H = CatA \times CPA,\ R = H \times (M - E) \times D \times L} \tag{8.1}$$

Where *H* represents the final hazard value, *CatA* the relative rate of category A airprox events, *CPA* the rate of events observed at the closest classified point of approach. *R* represents the risk, with *M* being the total number of test cases, *E* the rank within the test cases of the emergent behaviour, *D* represents the rank of the diagnostic consistency measure, and finally *L* represents the relative likelihood.

8.2.1. Validation Test Cases and Hazardous Outcomes

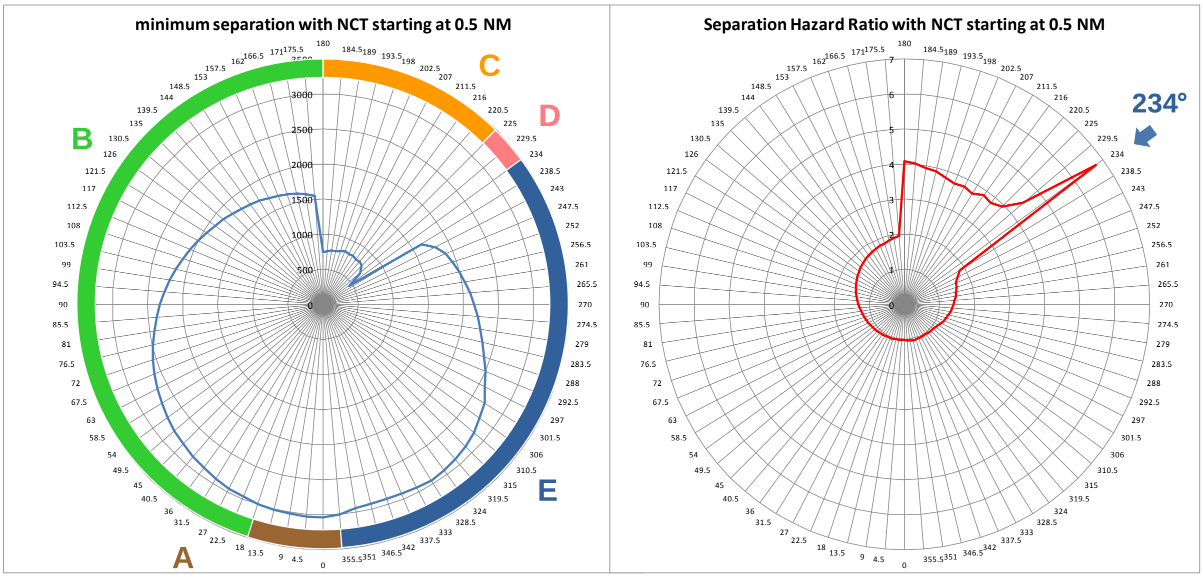


Figure 8.16 Hazard Modelling – separation as constraint violation [31]

With the system dynamics modelling undertaken earlier within the project, at that point employing a more naïve and un-damped model of the system dynamics, it was readily apparent that initial conditions with respect to direction of approach produced clear discontinuities in system behaviour, as indicated in Figure 8.16 [31]. Where the modelling of the overall effect upon the separation hazard produced a violation of the safety constraint, complex behaviour was often exhibited and could be seen as a means of obtaining insight into model and the system dynamics. For example from earlier versions of the 0.5 NM scenario, these exhibit five distinct dynamic behaviours dependent upon the initial relative heading: those approaching from within sector (A) often demonstrate a degree of entrainment (when both vehicles are at or near to the same speed); approach from direction (B) largely demonstrates direct and compliant avoidance; (C) demonstrates compliant avoidance with reduced separation; (D) demonstrates significant loss of safe separation (with insufficient space for an avoidance turn to the right); and finally (E) demonstrates suppression of the default interpretation of the Rules of the Air, with the substitution instead of the back-up emergency avoidance manoeuvre.

Consequently, in taking account of these effects, these influenced both the form of the Bayesian Belief Network in part, and the selection and partitioning of the test cases. Thereby, a comprehensive set of test cases are defined both for training and illustration purposes, Figure 8.17 below.

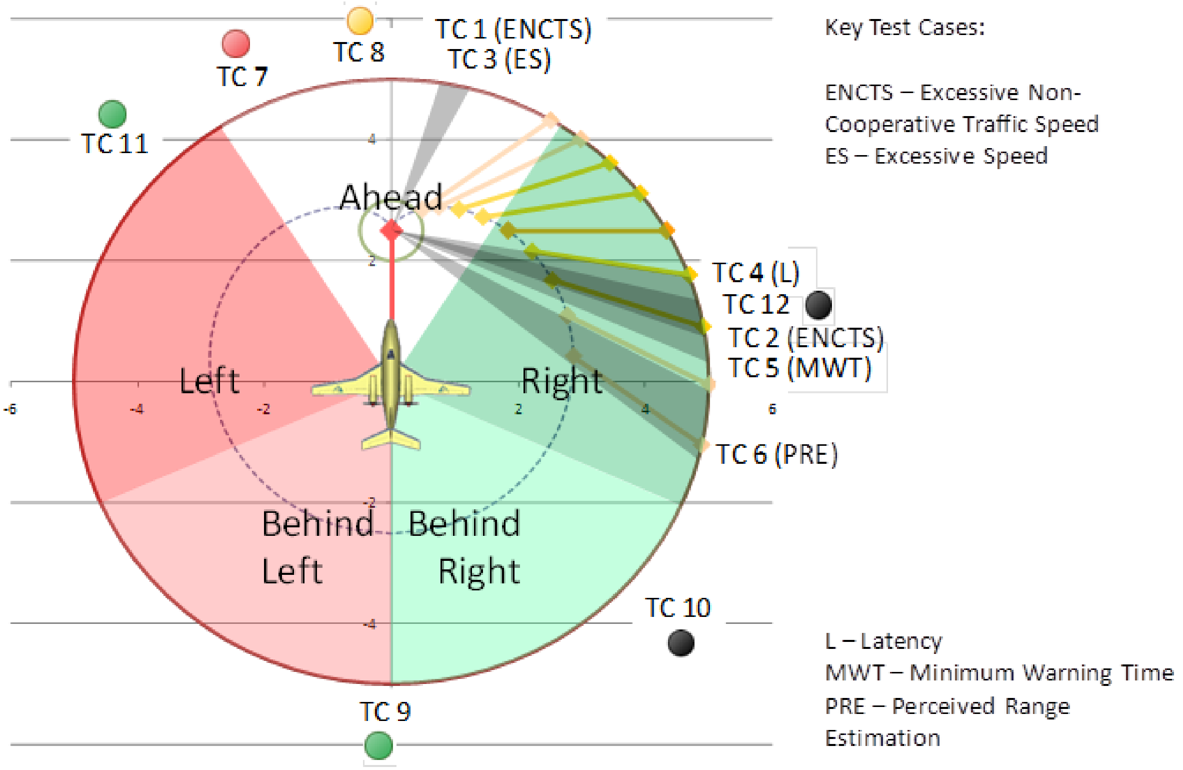


Figure 8.17 Hazard Test Case Assignment by Direction of Approach

Following a preliminary investigation of the behaviour of the model, twelve test cases were selected as itemised above, and as detailed in Table A.1 in Appendix A. Encompassing these 5+ sets of test cases, each vignette was selected as an appropriate example of a particular sequence of events, within which the hazard is realised as a violated constraint. Each was drawn from different sets of 200 Monte Carlo simulation runs; wherein the maximum and minimum values indicated in Table A.1 denote the range over which values were randomly selected within a uniform distribution at the start of each simulation run (and where the darker sectors shown in Figure 8.17 graphically approximately represent the variations for initial heading and the principal vignettes).

Additionally, two of the test cases are also decomposed into interesting subsidiary events. Test Case 12 includes an event immediately prior to collision as well as the event at the point of collision, whilst Test Case 2 also includes a precursor event occurring some seconds earlier, triggering a category "B" ("surprised" and "panicked") air proximity event, as well as the entry and exit events within the subsequent near-miss. Forty-five datasets have been generated using the final behavioural model representing a continuous distribution from all directions of approach and permitted speed ranges, wherein the range estimation error and latency are varied. In addition to these datasets there are also six datasets representing special cases using an earlier variant of the model (from which cases for the questionnaire were selected), plus a further thirteen datasets more thoroughly exploring the "give-way to traffic from the right" behaviour, derived from the final variant of the system dynamics behavioural model. In order to select a few representative cases with which to evaluate the Bayesian inference concept for internal self-consistency, all cases generating category "A" Airprox events are counted, as detailed in Table A.2, Appendix A. Beyond this, all of the datasets have been used to train the Bayesian network, providing 201,600 training cases in 12,600 sixteen second sample windows.

8.2.2. Hazard Context

The notable difference between the initial STPA – based HAZOP and the enhanced HAZOP is the recognition that the context varies considerably according to the direction of approach, or interaction, which might be inferred directly from an interpretation of the rules of the air. However, the splitting of the "behind" cases into "left" and "right" sub-cases is more due to convenience when defining the upper and lower bounds upon initial headings in the Monte Carlo cases, as this approach sector passes through 360° / 0° threshold.

8.2.3. System Flaws and Deviations

As the originally assigned relationships between flaws and deviations (inadequate control actions) formulated in the initial STPA – based HAZOP were a matter of conjecture, the realised constraint violations observed in the model, along with the simultaneously observed flaw and deviation events, allows these conjectured relationships to be confirmed, or not. Combining the inferred flaws with the observed flaws for a particular combination of observed deviations facilitates an arrangement of these observed flaws into a primary flaw and a series of possible secondary flaws for this combination of deviations within the specific special condition.

8.2.4. Diagnosed Defects and Data Post Processing

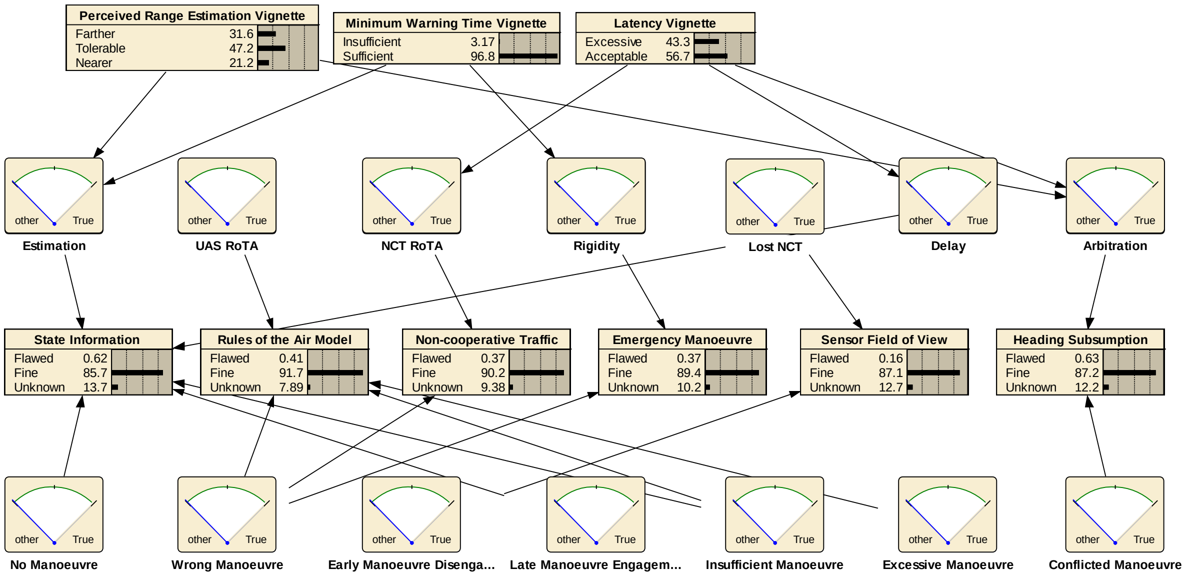


Figure 8.18 Vignette defective design precursor belief association in the Bayesian Network

Diagnostic nodes were also added to the inference model so as to define a representation of inference for forms of misbehaviour associated with the six sub-system models identified in the STPA process; State Information, Rules of the Air Model, Non-Cooperative Traffic, Emergency Manoeuvre, Sensor Field of View, and Heading Subsumption. Although not strictly necessary for the inference of flaw causes leading to deviation effects, these were constructed with hand-coded belief tables and adopted the same approach with regard to unknown states, although these nodes were not subsequently presented with training data. So far this diagnostic approach has shown only limited success. Having created a description of the HAZOP as a network of belief, the requirement then is to describe a constrained interface model linking the System Dynamics model and Bayesian forms, Figure 8.18.

The measure of the diagnostic consistency evident in each test case starts with a calculation of a diagnostic belief threshold heuristic *b* for each sub-system. These sub-systems may be diagnosed as being fine (*ok*) or flawed (*fl*), based upon the Bayesian inference yielding the probability *p* of the sub-system being fine, flawed or unknown (*uk*), as returned in the "black-box" test results for each test case – equations 8.2:

$$\boldsymbol{b_{ok}} = \begin{cases} \boldsymbol{1}, & \boldsymbol{p_{ok} > \frac{2}{3}(p_{fl} + p_{uk})} \\ \boldsymbol{0}, & \boldsymbol{otherwise} \end{cases}, \quad \boldsymbol{b_{fl}} = \begin{cases} \boldsymbol{1}, & \boldsymbol{p_{fl} > \frac{2}{3}(p_{ok} + p_{uk})} \\ \boldsymbol{0}, & \boldsymbol{otherwise} \end{cases} \tag{8.2}$$

From this a measure of consistency c in each test case, across *S* sub-systems, is calculated – equations 8.4, where equations 8.3 yield the numbers of sub-systems fine (OK), flawed and unknown:

$$\boldsymbol{N_{ok} = \sum_{s=1}^{S} b_{ok_s}, \; N_{fl} = \sum_{s=1}^{S} b_{fl_s}, \; N_{uk} = S - N_{ok} - N_{fl}} \tag{8.3}$$

$$c_{ok} = \begin{cases} \frac{1}{N_{ok}} \sum_{s=1}^{S} b_{ok_s} \cdot p_{ok_s}, & N_{ok} \neq 0 \\ 0, & otherwise \end{cases}, \quad c_{fl} = \begin{cases} \frac{1}{N_{fl}} \sum_{s=1}^{S} b_{fl_s} \cdot p_{fl_s}, & N_{fl} \neq 0 \\ 0, & otherwise \end{cases} \qquad 8.4$$

This consistency measure is then used to calculate a diagnostic index *i* – equation 8.5 – and utility *u* for each test case, across *M* test cases – equation 8.6. *K* is an arbitrary scale coefficient:

$$i = \frac{K}{2} \cdot \frac{c_{ok}+c_{fl}}{\sum_{m=1}^{M}(c_{ok_m}+c_{fl_m})} \cdot \left(1 + \frac{N-N_{uk}}{N}\right) \qquad 8.5$$

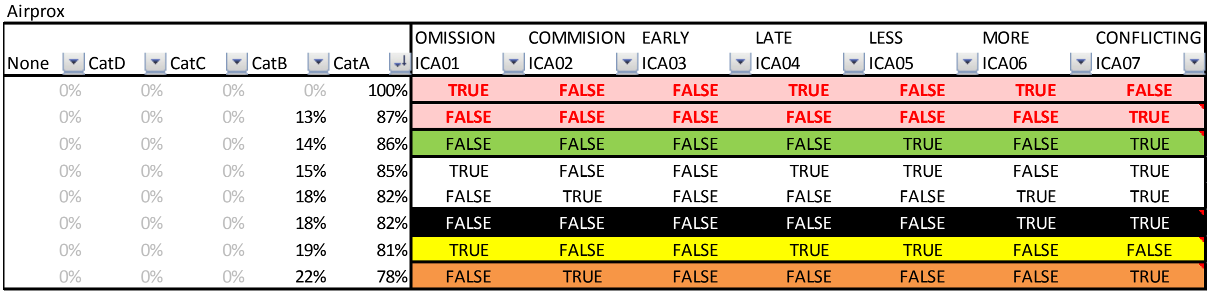
$$u_m = \mathbf{rank}_{i_1 \to i_M}(i_m) \qquad 8.6$$

In applying the adapted STPA process, six sub-system models were identified (State Information, Rules of the Air Model, Non-Cooperative Traffic, Emergency Manoeuvre, Sensor Field of View, and Heading Subsumption), to which a variety of “Control Flaws” and “Inadequate Control Actions” (deviations) were assigned, and explicitly associated within the Bayesian network. The diagnostic utility value arising with each test case indicates (by rank) the best, the worst and middling of these in their ability to consistently (rather than necessarily accurately) diagnose potentially faulty sub-system models. From this evaluation the higher speed directly ahead test cases (7 & 8) appear to yield little useful diagnostic information, ranking 10th and 12th out of 12 for the high speed (TC7) and speed difference (TC8) vignettes respectively. This is most likely a consequence of the speed related states in the Bayesian network being formulated as weak discriminators of faults in being associated with all control flaws, and hence treated (equally) as potential parents of all flaws, rather than the more discriminatory parent-child relationships defined in the other more specific defect causes (estimation, latency and warning time). This poor discrimination might be viewed as an argument against a naïve Bayesian approach in that little appears to have been added through the learning process from cases. Conversely, it might be argued that the learning process has revealed a stronger association with the “rigidity” flaw than first conceived, and that this deserves further consideration to determine whether or not this is merely an incidental artefact. This finding begs the question as to whether the “rigidity” flaw has been assigned to the most appropriate sub-system, or whether the consequence of a “rigidity” flaw might be that the deviation in manoeuvre response is more likely to be “less” rather that “wrong” in these circumstances. In any event this particular diagnostic case appears to be poorly formulated.

8.2.5. Inferred Airprox Events

The constraints violated are taken, and adapted, from the definitions of air proximity categories and the “closest point of approach” required to “stay well clear” at 500 feet. However, even at or beyond 500 feet, if the observed behaviour of either vehicle is such as to appear to violate the rules of the air, from the other vehicle’s perspective, or incur an excessive response, or (visual) contact with the other is lost, then each of these events may escalate the violation category – taken here as the air proximity category. As each test case has already been selected on the basis that it contains either a category A or B violation, then what remains to be determined is how likely would the particular combination of deviations present in the test case at the point of violation be categorised as being either one or the other – clearly the category A events embody the higher risk of actual collision.

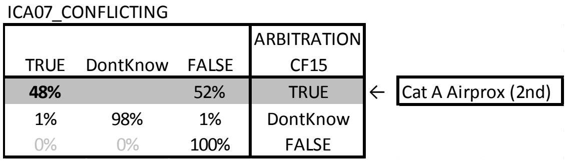
The process to obtain the relative air proximity category rates takes the row from the Bayesian probability table for "Airprox", contained within the trained Bayesian model, having input states matching the particular combination of deviations observed as active in the particular test case:



Airprox

					OMISSION	COMMISION	EARLY	LATE	LESS	MORE	CONFLICTING
None	CatD	CatC	CatB	CatA	ICA01	ICA02	ICA03	ICA04	ICA05	ICA06	ICA07
0%	0%	0%	0%	100%	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE
0%	0%	0%	13%	87%	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE
0%	0%	0%	14%	86%	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE
0%	0%	0%	15%	85%	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE
0%	0%	0%	18%	82%	FALSE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE
0%	0%	0%	18%	82%	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE
0%	0%	0%	19%	81%	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE
0%	0%	0%	22%	78%	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE

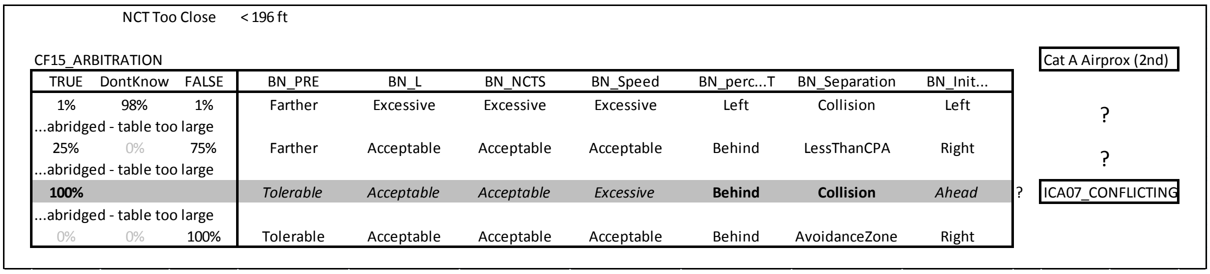
The rows of the above table contain the distribution of probability across the five Airprox categories (including "None"), which in this case ought to be split mostly between A and B categories, with only small residual values for the remaining categories. The combination of ICA=True detailed in the second row down of the table above is seen in both Test Case TC1, at the closest point of approach, and in Test Case TC2 at entry, both representing "Excessive Non-Cooperative Traffic Speed" hazard cases – redefined as a DIFFERENT manoeuvring airspeed deviation in the Enhanced HAZOP. TC1 is an "Ahead" vignette and TC2 a "Right" vignettes – Figure 8.17 Additionally, the same combination of events is also seen in sample runs (not specific test cases) S161 at collision, and S85 at collision, each drawn from the same respective vignette data sets – see also Figure A.30.



ICA07_CONFLICTING

			ARBITRATION
TRUE	DontKnow	FALSE	CF15
48%		52%	TRUE
1%	98%	1%	DontKnow
0%	0%	100%	FALSE

← Cat A Airprox (2nd)



NCT Too Close < 196 ft

CF15_ARBITRATION

TRUE	DontKnow	FALSE	BN_PRE	BN_L	BN_NCTS	BN_Speed	BN_perc...T	BN_Separation	BN_Init...
1%	98%	1%	Farther	Excessive	Excessive	Excessive	Left	Collision	Left
...abridged - table too large									
25%	0%	75%	Farther	Acceptable	Acceptable	Acceptable	Behind	LessThanCPA	Right
...abridged - table too large									
100%			*Tolerable*	*Acceptable*	*Acceptable*	*Excessive*	**Behind**	**Collision**	*Ahead*
...abridged - table too large									
0%	0%	100%	Tolerable	Acceptable	Acceptable	Acceptable	Behind	AvoidanceZone	Right

Cat A Airprox (2nd)

?

?

? ICA07_CONFLICTING

From the above table, an inference made from the Bayesian probability tables after training, we can see that a CONFLICTING manoeuvre deviation, simultaneous with a DIFFERENT airspeed deviation, likely due to an "arbitration" flaw in both the ahead and right approach conditions, appears to be particular likely to produce a category A violation – and also likely to be subject to a collision from behind. This last observation might also suggest that the NCT "pilot" behaviour could be just too dumb as expressed in the model, and therefore something to be investigated further.

In the next lower category A case, a LESS and CONFLICTING manoeuvre deviation combined with a LOW airspeed deviation, likely due to "arbitration" and "estimation" flaws in the right condition, also produces a similar outcome – see also Figure A.30, Test Case TC6, and also the third row down in the earlier Airprox Bayesian probability table. It is appropriate therefore to classify both sets of these conditions and precursors as likely to produce the most hazardous behaviour.

Similarly, the probability distributions for "NTC Too Close", "UAS Surprised", "UAS Threatened", "NCT Surprised" and "NCT Panicked" are obtained in a similar manner from their respective probability tables within the trained Bayesian model.

8.2.6. Inferred Flaws and Deviations Post Processing

As with the determination of the diagnostic behaviour for each test case, the calculation updating the belief that the flaws identified in the initial HAZOP remain true, or whether indeed these ought to be replaced, is undertaken using another belief heuristic *b*, for both flaws *f* and deviations *d* – equations 8.7 and 8.8. Again these revised belief values are derived from the Bayesian inference returning the probabilities *p* that a flaw *f* or deviation *d* is present, set against the probabilities that the response from the Bayesian model is "don't know" (p_{dkf} and p_{dkd}). This revised belief is then factored by the enabling event *e* for the given flaw or deviation; where the term e_f takes the value "1" where the system dynamics model reveals the actual existence of the flaw, zero otherwise – likewise with e_d in the case of deviations.

$$\boldsymbol{b_f = p_f\left(1 - p_{dkf}\right) \cdot e_f} \qquad 8.7$$

$$\boldsymbol{b_d = p_d(1 - p_{dkd}) \cdot e_d,\ \ others = 1 - \sum b_d} \qquad 8.8$$

These revised beliefs are used subsequently to reorder the appearance of the control flaw information within the final enhanced HAZOP table. The flaw having the now the highest belief value is classified as the "primary" flaw, with the remainder of flaws having non-zero belief values being classified as "secondary" flaws, and these can be assembled in descending order of belief. In some cases (TC 8, 7 & 3 for example) the belief heuristic may be close in value and some judgement may be applied to this reordering, or perhaps the control flaw definitions themselves ought to be revisited.

This process therefore matches the evidence that flaws and deviations as defined in the system dynamics model match those as originally identified in the initial HAZOP and thence embodied within the structure of the Bayesian Network. From the detailed results (Appendix A) we can see that the "estimation" flaw (CF04) remains reasonably well matched with the deviations as originally associated, with the exception of the deviation EARLY; where this appears now more likely to be a consequence of the “arbitration” flaw. However, the “estimation” flaw – deviation relationships now also varies as a consequence of the additional airspeed deviations. Conversely, the “rigidity” flaw (CF07) now appears associated with many more deviations other than COMMISSION, as originally determined, and is now the dominant flaw where the airspeeds are either high or significantly different. The “latency” flaw (CF10) now rarely appears (in these test cases) apart from deviations of LESS from the left, and MORE from the right; the first appearing not to be turning sufficiently when threatened whilst having the right of way, the second in turning hard and late when required to give way, both embodying threat response delays in the range of 4 to 5 seconds – recognised as flaw CF10.

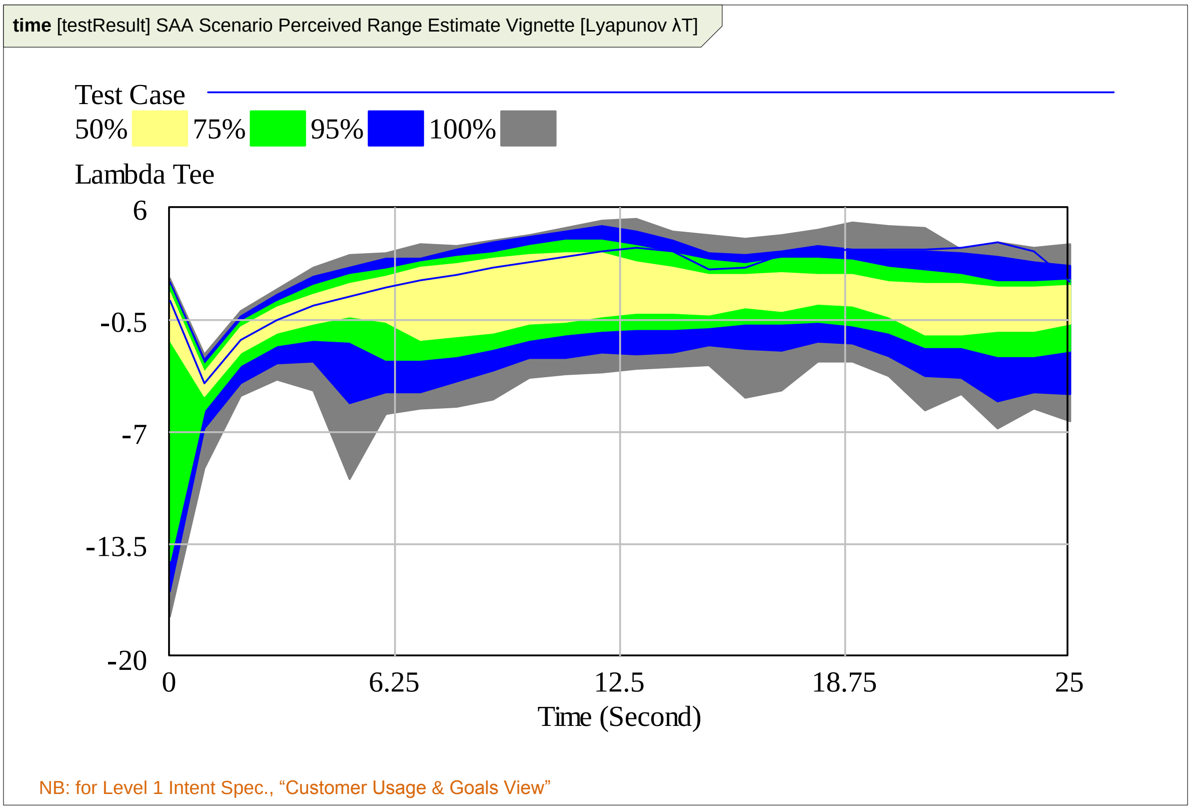
8.2.7. Avoidance and Emergent Behaviour

Whilst the individual trajectories of the respective entities clearly reveal apparent hazards in their behaviours these nevertheless represent only single cases drawn from the larger sets of 200 training cases arising with each Monte Carlo vignette group. As the intention is to capture the tendency towards emergent behaviour in a quantifiable manner, and then present this in a succinct form within the updated HAZOP process, then a more general measure is required than observations made on the occasional interesting trajectory.

For this it was decided that an established measure of emergent behaviour would be employed, wherein the Lyapunov Exponent should suffice, as defined previously in Figure 7.17.

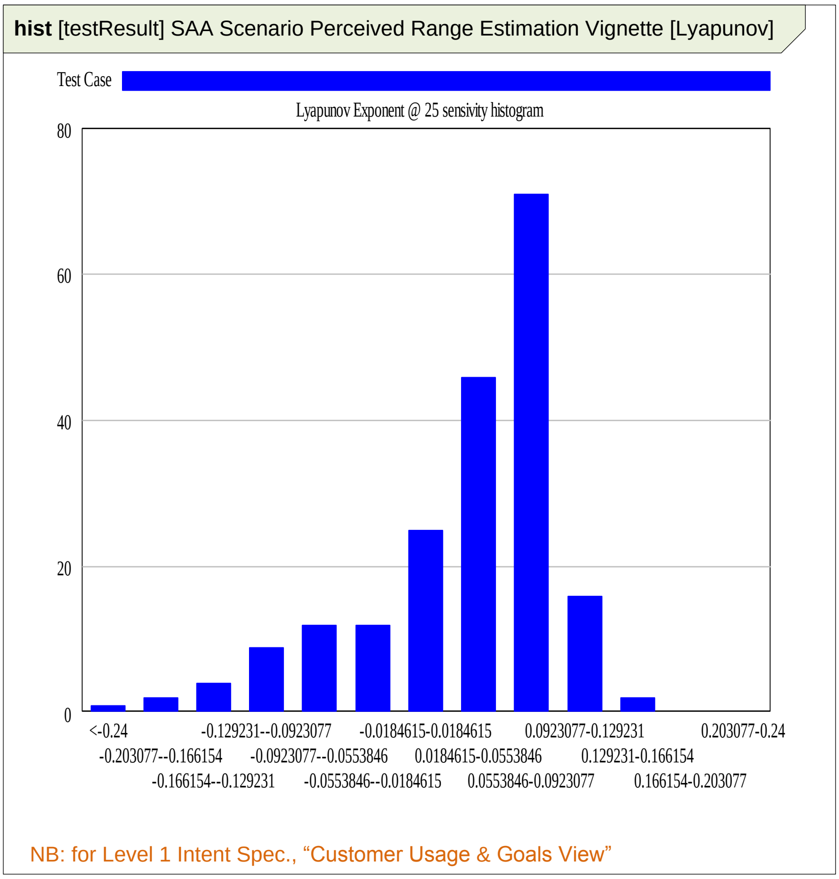
However certain peculiarities ought to be noted. Inspecting the “Lambda Tee – λT” distribution plots in Figure 8.19, indicates that where the NCT is the first to manoeuvre then the value is initially negative, whereas where the UAV manoeuvres first or more aggressively, then the initial value is usually positive, but in either case both stabilise within a “bottle-neck” within a couple of seconds (within approximately the first 8 - 16 samples) – which occurs well before the main sequence of the interaction. Notice also that for Test Case TC6 the majority of the trajectories return positive exponent values at 25 seconds from the start of the encounter, as revealed in the corresponding histogram cross section shown in Figure 8.20, indicating that the encounter is not yet over – as the NCT has taken extreme avoiding action and has yet to resume its original heading, Figure 8.21.

In general, for this model, the manoeuvres of the UAV are often initially performed at rates greater than that of the NCT. Negative dips, or spikes, in the λT distribution plots reveal where significant reversals of direction occur, whilst smoothly rising trajectories indicate a progressively increasing rate of manoeuvre, typically peaking where the two successfully avoid one-another – then to continue to settle back towards their original headings as the Lyapunov trajectories continue to fall. This behaviour is most clearly seen in three test cases (TC8, 7 & 3 – see Figure A.31, uppermost three cases) representing conflicts from directly ahead with each vehicle quickly passing the other with minimum deflection – and not colliding.



NB: for Level 1 Intent Spec., "Customer Usage & Goals View"

Figure 8.19 Lyapunov Exponent Distributions from 0 to 25 seconds for Test Case 6 (TC6)



NB: for Level 1 Intent Spec., "Customer Usage & Goals View"

Figure 8.20 The Corresponding Lyapunov Exponent Distribution Histogram at 25 seconds

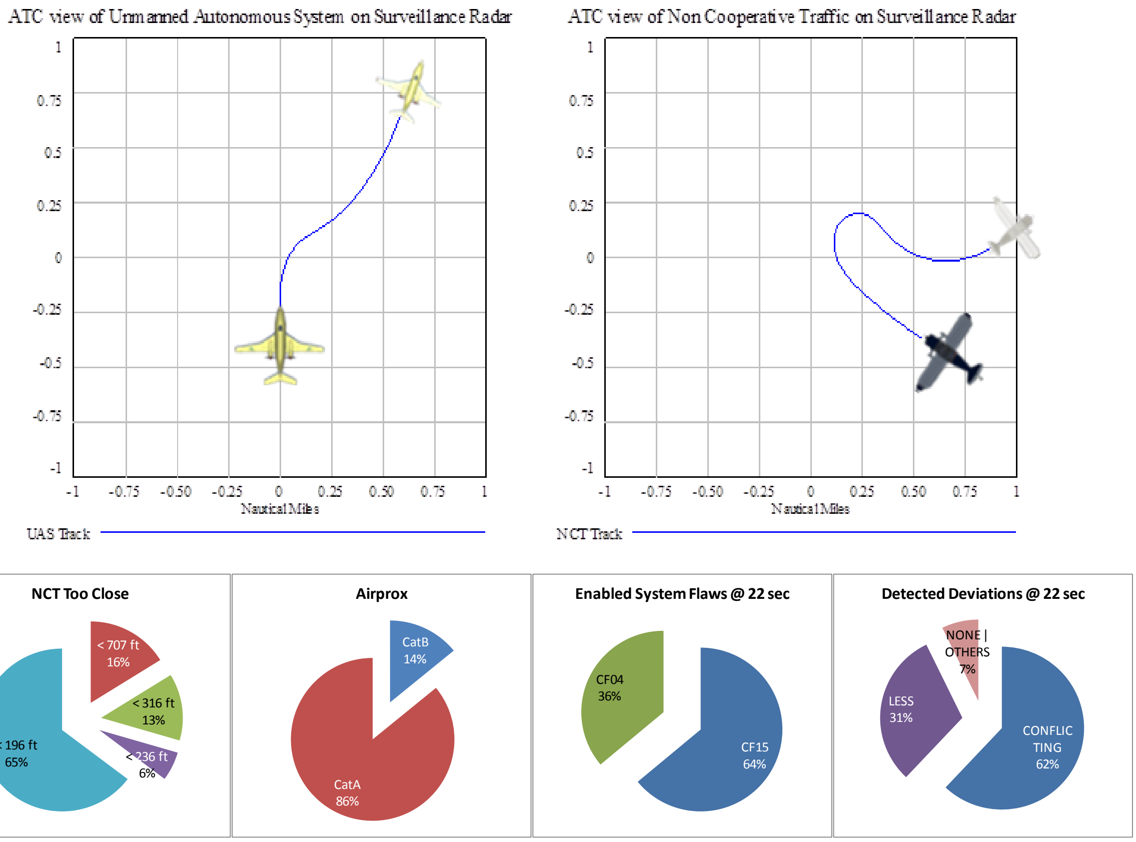


Figure 8.21 Hazard Test Case 6 (TC6)

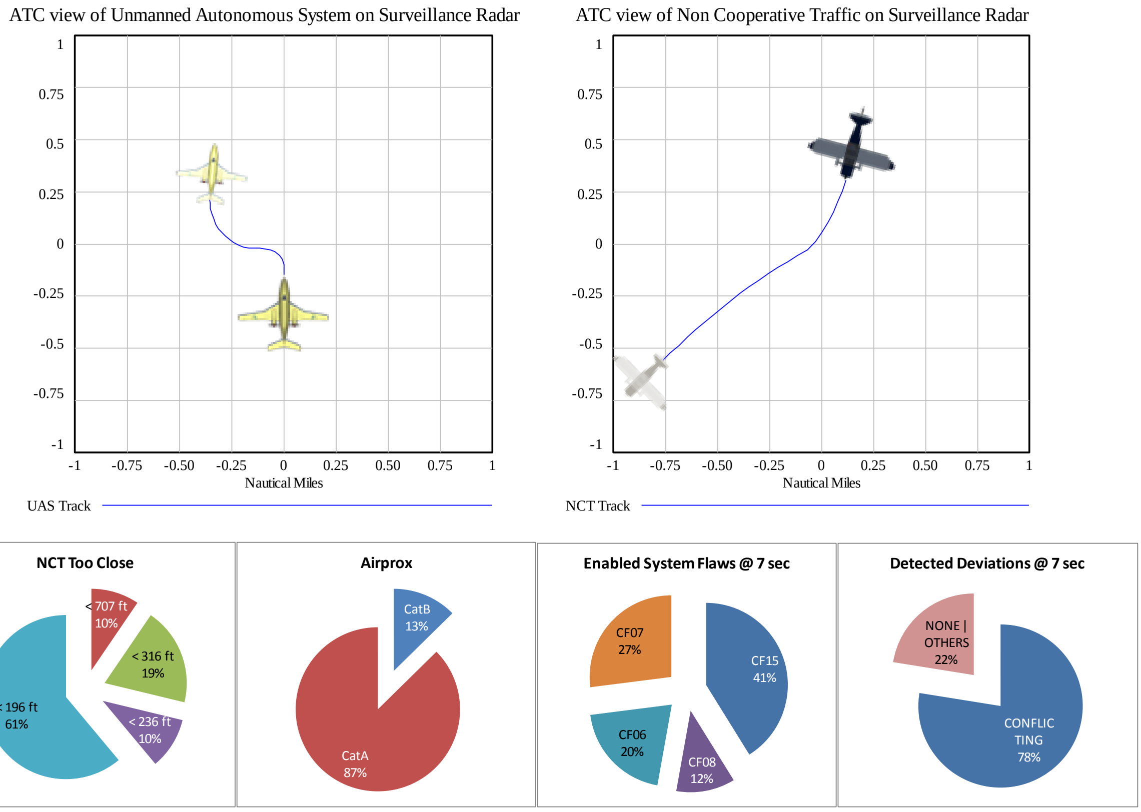


Figure 8.22 Hazard Test Case 1 (TC1)

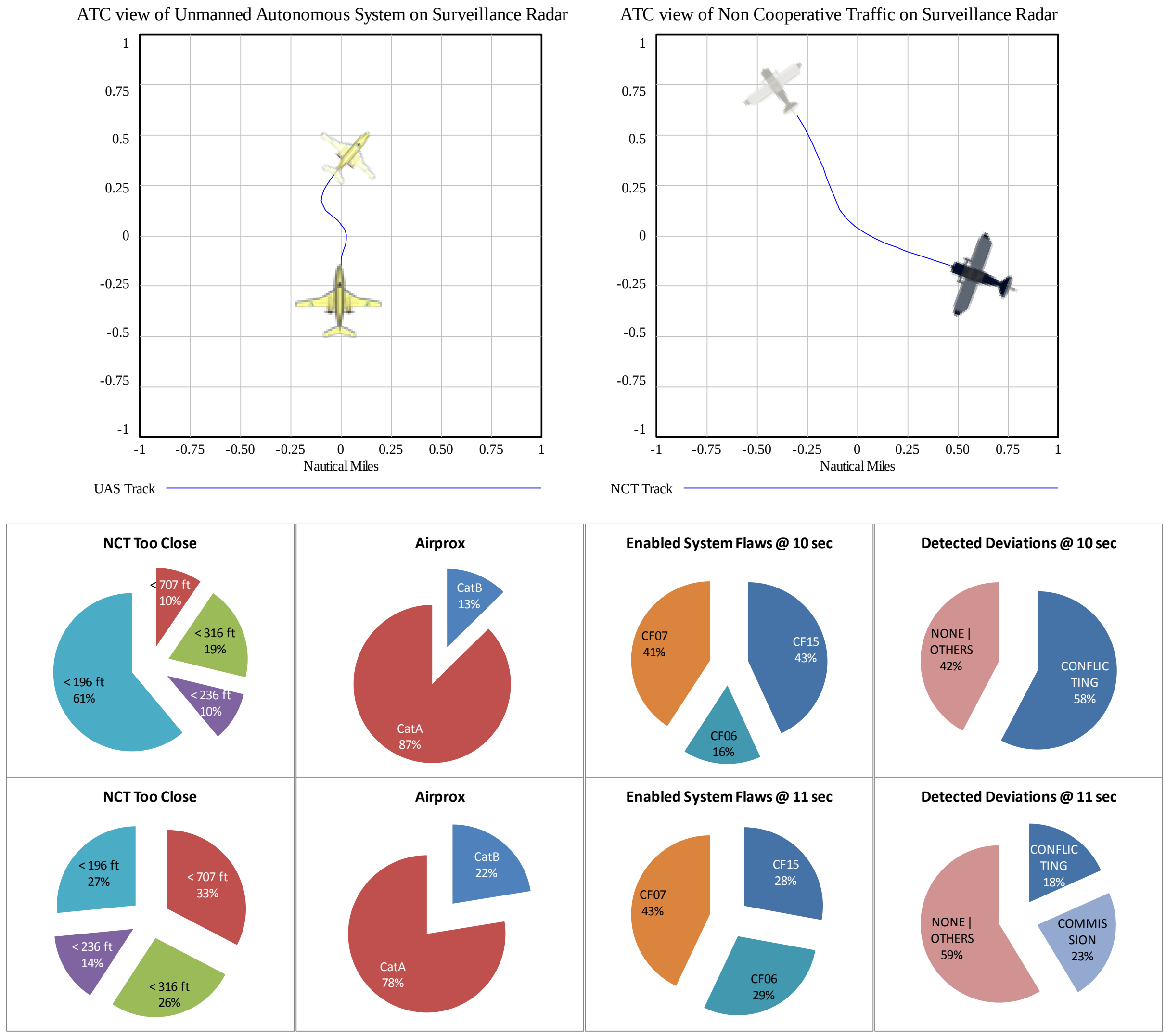


Figure 8.23 Hazard Test Case 2 (TC2)

Of more interest, if it is accepted that "emergence" can be defined in the saliency of a persistent similarity of patterns arising with the interaction of simple underlying rules [54], test cases TC1 and TC2 exhibit a degree of emergence caught in the Monte Carlo system dynamics modelling process. Comparing the pie-charts in Figure 8.22 and Figure 8.23, it can be seen that at 7 and 10 seconds respectively both reveal a similar distribution of separation hazard and air proximity violation, and similar distribution of flaws – with the exception that in the case of TC1 the UAS is also likely to lose track (along lines of sight) of the Non-Cooperative Traffic (NCT). It can also be seen that in the case of TC2 at 11 seconds that the UAS exhibits a COMMISSION deviation, in that it momentarily turns towards the NCT as it passes behind it. Both cases demonstrate a distinct similarity of behaviour but with differences based upon the context (principally the relative bearings starting the encounters).

Together test cases TC1, TC2 and TC6 stand as counter examples that may be used to falsify any assumptions made about the safety of this dynamical system, where that assumption is based solely upon a straight forward application of the Rules of the Air. More generally correlations between the initial relative heading, the distribution of Lyapunov Exponents, and Airprox are revealed in Figure 8.24 and Figure 8.25, and these are investigated in detail with regression analysis in Appendix A.

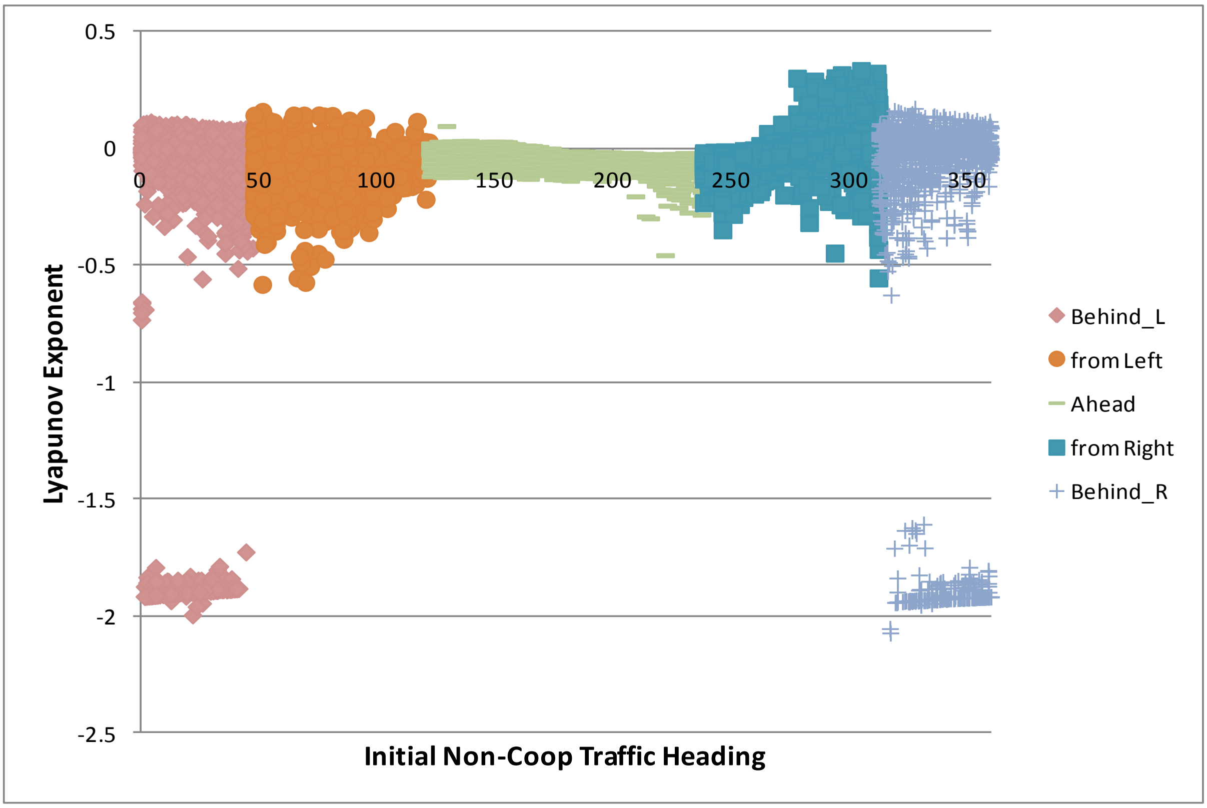


Figure 8.24 Lyapunov Scatter Plots with respect to Initial Relative Heading

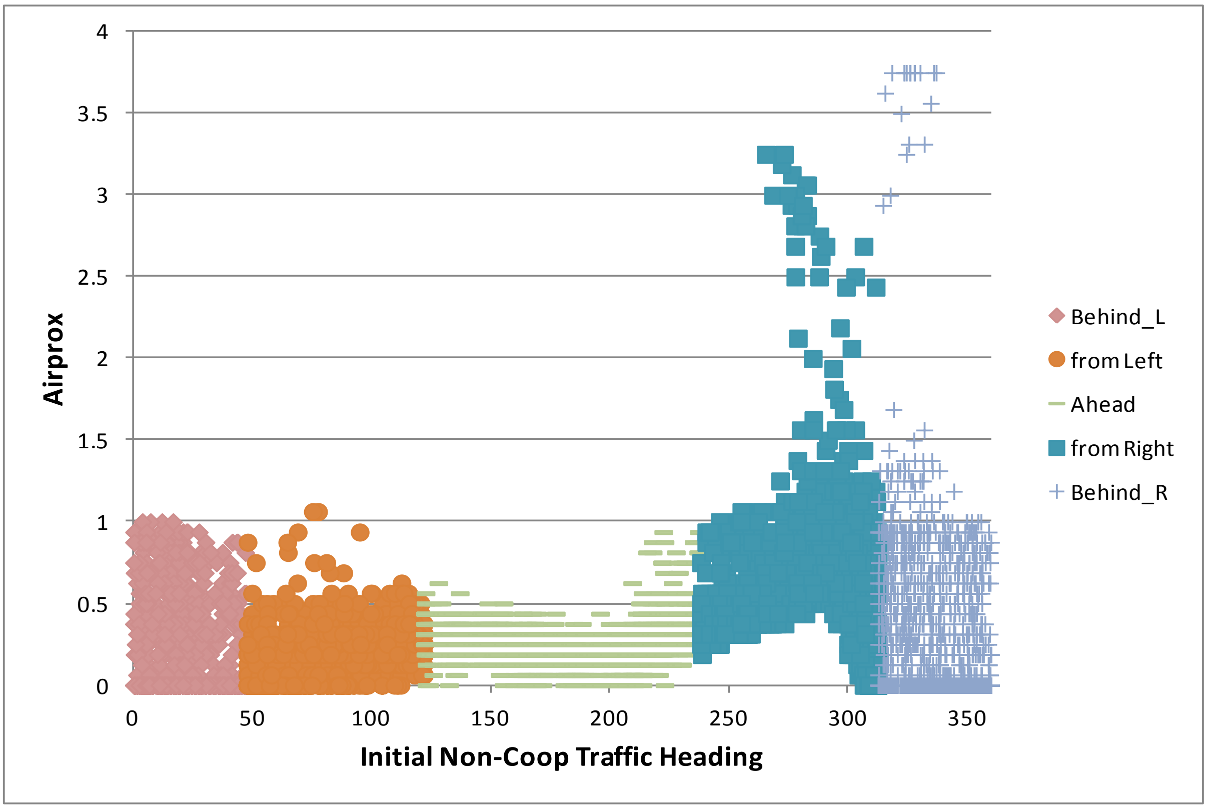


Figure 8.25 Airprox Scatter Plots with respect to Initial Relative Heading

8.2.8. Black Box v White Box Internal Consistency

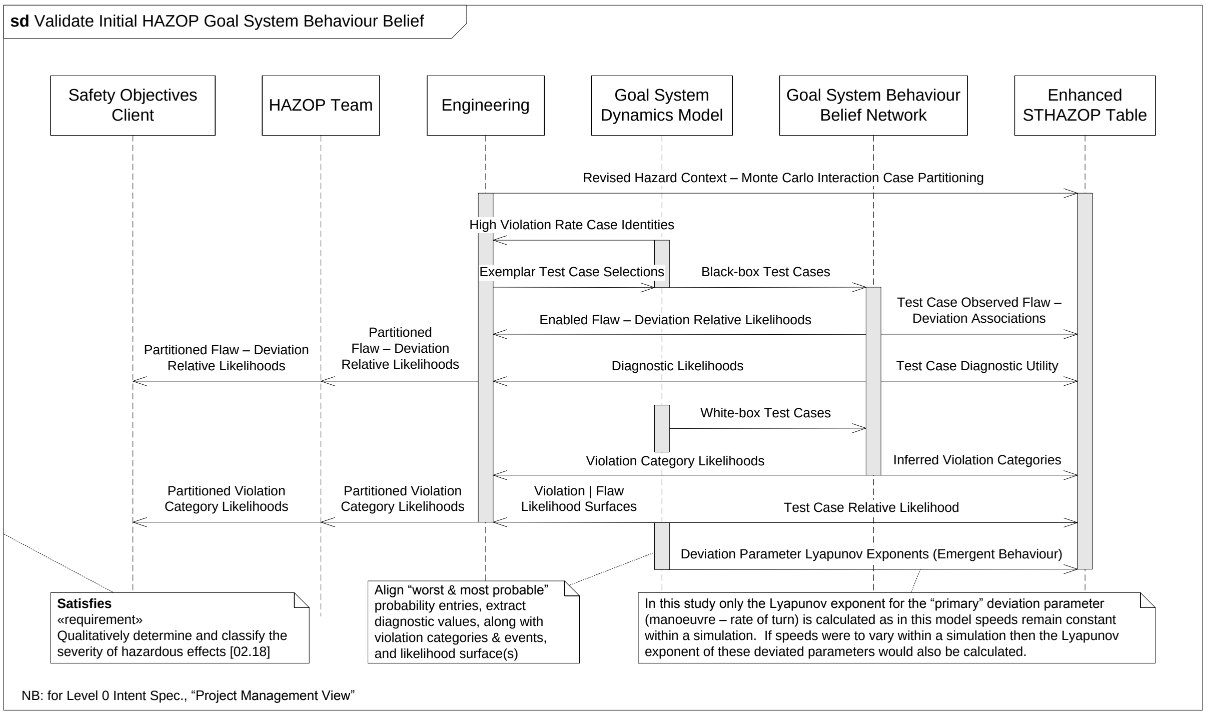

sd Validate Initial HAZOP Goal System Behaviour Belief
Safety Objectives Client
HAZOP Team
Engineering
Goal System Dynamics Model
Goal System Behaviour Belief Network
Enhanced STHAZOP Table
Revised Hazard Context – Monte Carlo Interaction Case Partitioning
High Violation Rate Case Identities
Exemplar Test Case Selections
Black-box Test Cases
Test Case Observed Flaw – Deviation Associations
Enabled Flaw – Deviation Relative Likelihoods
Partitioned Flaw – Deviation Relative Likelihoods
Partitioned Flaw – Deviation Relative Likelihoods
Diagnostic Likelihoods
Test Case Diagnostic Utility
White-box Test Cases
Violation Category Likelihoods
Inferred Violation Categories
Partitioned Violation Category Likelihoods
Partitioned Violation Category Likelihoods
Violation | Flaw Likelihood Surfaces
Test Case Relative Likelihood
Deviation Parameter Lyapunov Exponents (Emergent Behaviour)
Satisfies
«requirement»
Qualitatively determine and classify the severity of hazardous effects [02.18]
Align "worst & most probable" probability entries, extract diagnostic values, along with violation categories & events, and likelihood surface(s)
In this study only the Lyapunov exponent for the "primary" deviation parameter (manoeuvre – rate of turn) is calculated as in this model speeds remain constant within a simulation. If speeds were to vary within a simulation then the Lyapunov exponent of these deviated parameters would also be calculated.
NB: for Level 0 Intent Spec., "Project Management View"

Figure 8.26 Validate Goal System Behaviour Belief Sequence Diagram

The method by which the diagnostic consistency measure (section 8.2.4) and the final inferred relationship between flaws and deviations (section 8.2.6) are extracted from the Bayesian model is through a series of tests to determine the belief model internal self-consistency – sequentially described in Figure 8.26. Initially we assume that an observed combination of physical and constraint violation states might accurately, or even uniquely, infer corresponding flaw–deviation combinations.

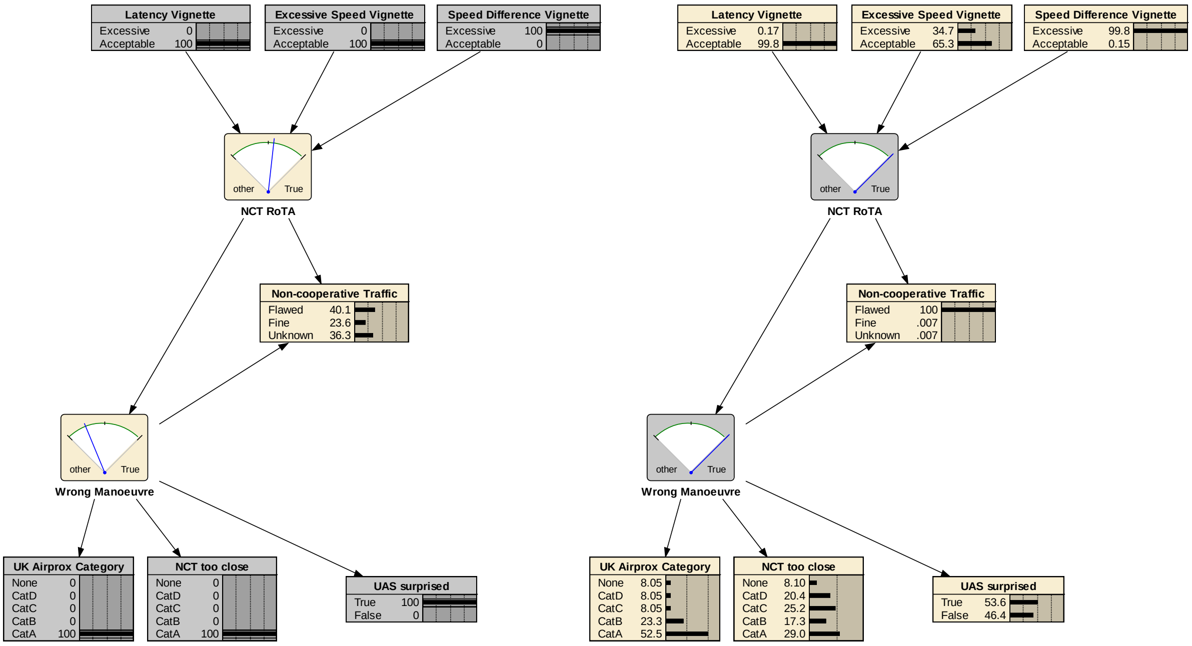
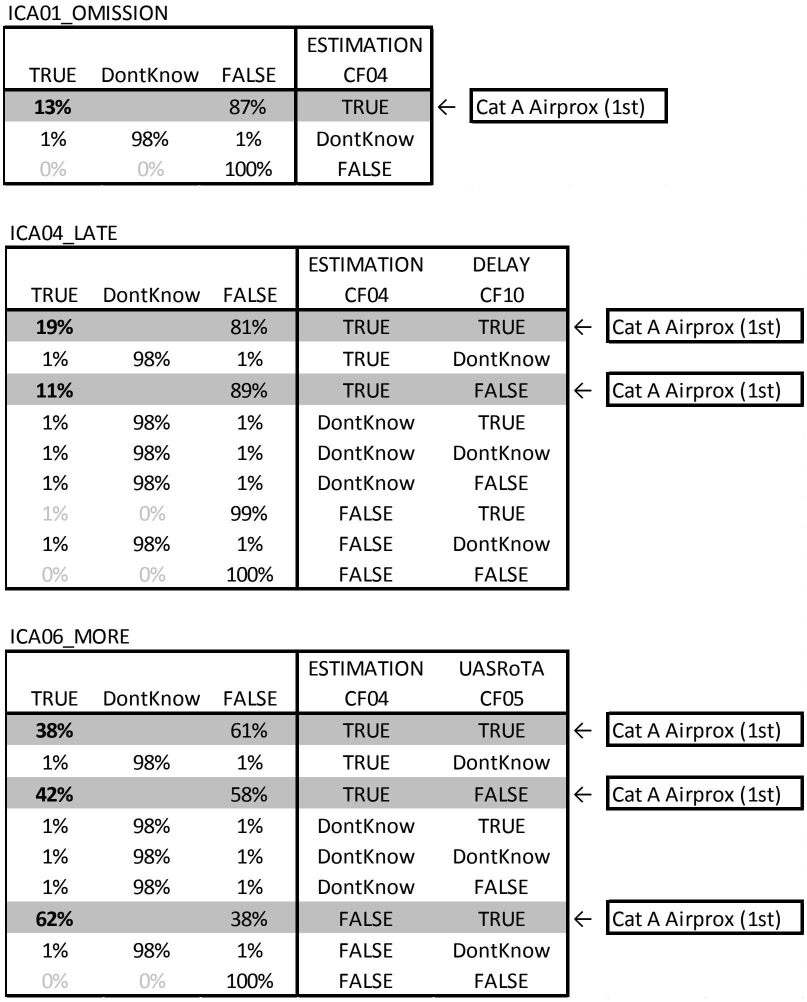


Figure 8.27 Black-box versus White-box inference

This is the black box model that assumes that we know nothing about the mechanisms internal to the system linking flaws to deviations. The second part is to apply only the observed and captured flaw – deviation state combinations derived from the test cases. This is the white box model that assumes that we have complete faith in the flaw – deviation relationship, whatever the mechanism within the system that connects the two, to be revealed in external behaviour. The comparison of the two reveals varying degrees of consistency and residual uncertainty within the inference model from which to judge the quality (i.e. uniqueness, but not necessarily accuracy of interpretation) of diagnosis and belief in the flaw – deviation relationships; subsequently used to update the HAZOP study model.

8.2.9. Worst and Improbable Case Example



ICA01_OMISSION

TRUE	DontKnow	FALSE	ESTIMATION CF04	
13%		87%	TRUE	← Cat A Airprox (1st)
1%	98%	1%	DontKnow	
0%	0%	100%	FALSE	

ICA04_LATE

TRUE	DontKnow	FALSE	ESTIMATION CF04	DELAY CF10	
19%		81%	TRUE	TRUE	← Cat A Airprox (1st)
1%	98%	1%	TRUE	DontKnow	
11%		89%	TRUE	FALSE	← Cat A Airprox (1st)
1%	98%	1%	DontKnow	TRUE	
1%	98%	1%	DontKnow	DontKnow	
1%	98%	1%	DontKnow	FALSE	
1%	0%	99%	FALSE	TRUE	
1%	98%	1%	FALSE	DontKnow	
0%	0%	100%	FALSE	FALSE	

ICA06_MORE

TRUE	DontKnow	FALSE	ESTIMATION CF04	UASRoTA CF05	
38%		61%	TRUE	TRUE	← Cat A Airprox (1st)
1%	98%	1%	TRUE	DontKnow	
42%		58%	TRUE	FALSE	← Cat A Airprox (1st)
1%	98%	1%	DontKnow	TRUE	
1%	98%	1%	DontKnow	DontKnow	
1%	98%	1%	DontKnow	FALSE	
62%		38%	FALSE	TRUE	← Cat A Airprox (1st)
1%	98%	1%	FALSE	DontKnow	
0%	0%	100%	FALSE	FALSE	

Table 8.1 Observed Inadequate Control Actions for a Unique Airprox Event

Upon a detailed inspection of the Bayesian probability tables a unique event is revealed whereby only one combination of deviations, or inadequate control actions, gives rise to a single category A air proximity event and appears in no other category event – identified in Table 8.1 above, and in the first row of the earlier Airprox Bayesian probability table. On the occasion of this event, deviation states representing OMISSION, LATE and MORE occur simultaneously. By applying these three "True" states, along with the remainder of the equivalent "False" states, and from the closest inferred approximation to the observed control flaws it can be discovered that this event most probably occurred with an initial encounter arriving from behind, and passing the UAV most closely by on the left, and clearly within the limit of the CPA. This combination can be identified in the inference results within the Bayesian Network shown in Figure 8.29 – in this the shaded (grey) nodes represent the above assumed observations entered as findings with the inferences remaining on the flaws, confirming a likely strong failure in ESTIMATION (and also in RIGIDTY), along with degraded UASRoTA behaviour and some likely DELAY.

Within the limited selection of test cases, the nearest equivalent available cases are found in TC9 and TC10, Figure 8.28, although neither of these test cases represents the speed combination inferred in the unique example, where in the unique case both vehicles would appear to have been travelling with excessive speed (circa 250 kts). In these two potential alternative counter-examples each has the UAS being overtaken, and then "taken by surprise" so then to attempt to suddenly "give way" to the traffic on the right – likely due to an over simplified interpretation of the perspective on the NCT as seen from the UAS (not representing the relative velocity as inferring a particular intent on the part of the NCT). Of itself this is of little importance in the current model. However, what it does reveal is that rare but plausible behaviours may be discovered in a methodology combining system dynamics modelling, Monte Carlo sensitivity analysis and Bayesian reasoning, and so provides a tool to further explore rare near-miss cases which may serve as counter-examples to the safety assumptions – and hence falsify a particular safety hypotheses in support of a HAZOP study.

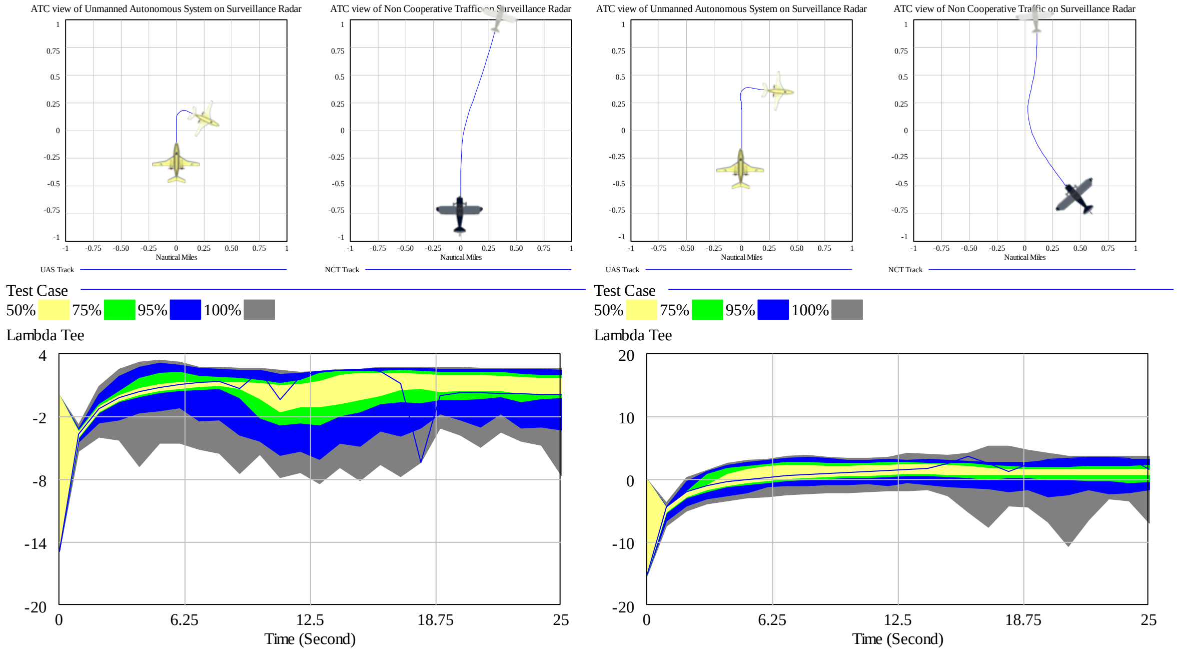


Figure 8.28 Hazard Test Cases 9 and 10 (TC9, TC10)

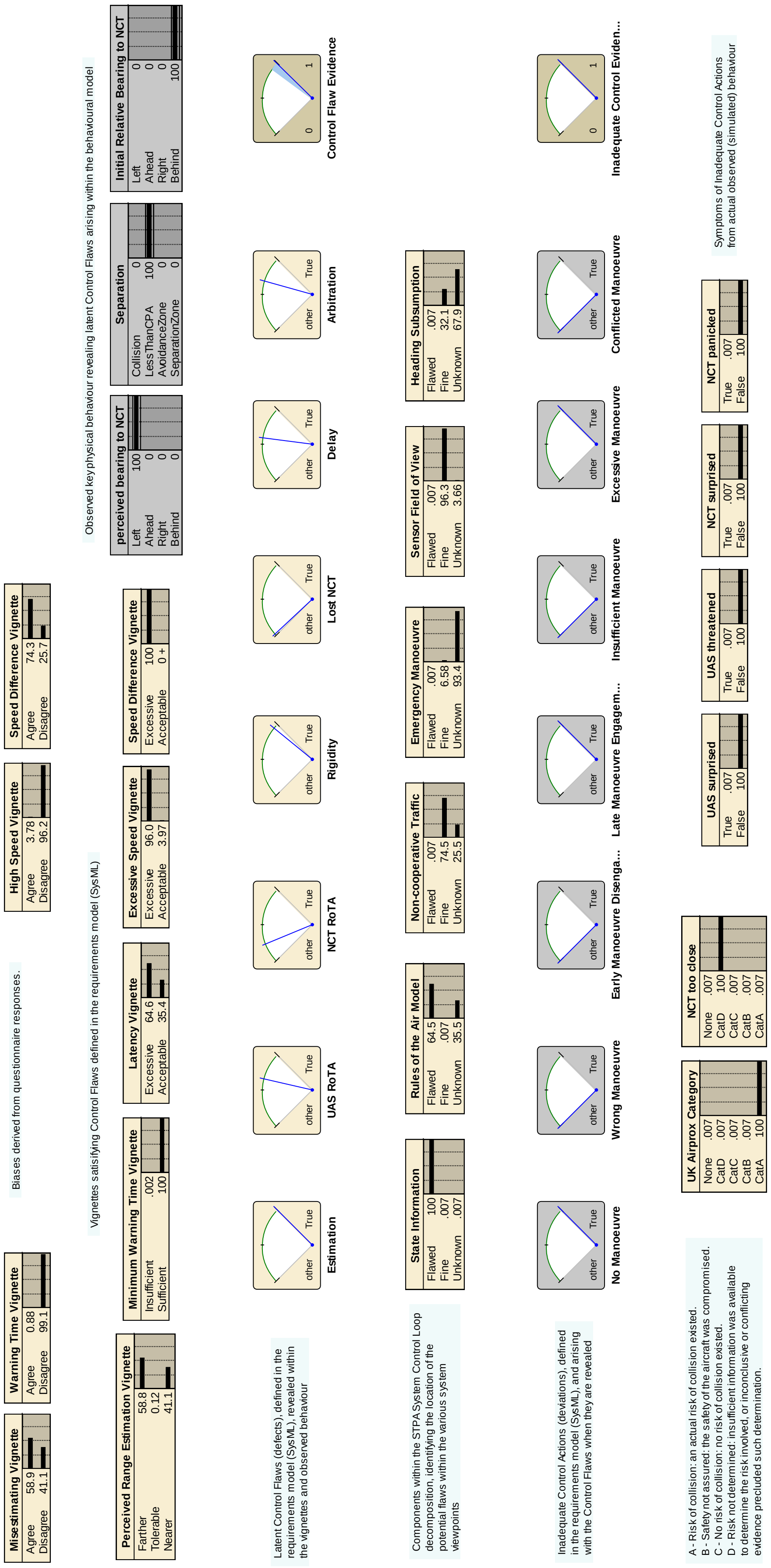
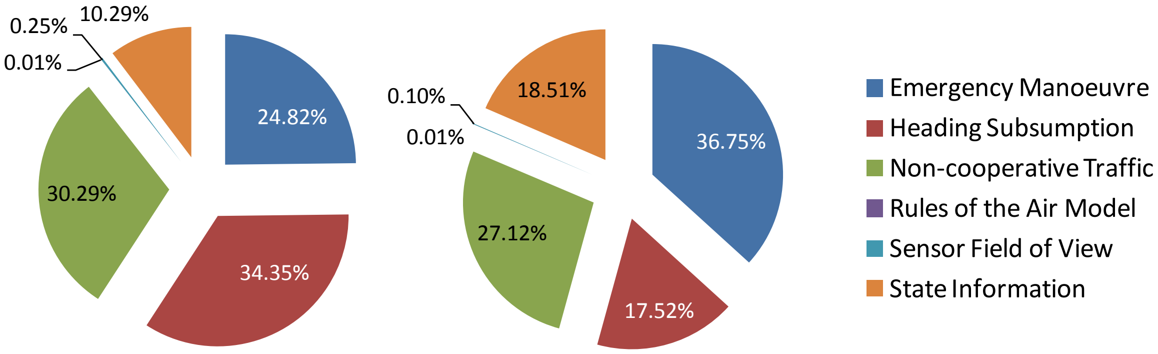

Misestimating Vignette
Agree 58.9
Disagree 41.1
Warning Time Vignette
Agree 0.88
Disagree 99.1
Biases derived from questionnaire responses.
High Speed Vignette
Agree 3.78
Disagree 96.2
Speed Difference Vignette
Agree 74.3
Disagree 25.7
Perceived Range Estimation Vignette
Farther 58.8
Tolerable 0.12
Nearer 41.1
Minimum Warning Time Vignette
Insufficient .002
Sufficient 100
Vignettes satisfying Control Flaws defined in the requirements model (SysML)
Latency Vignette
Excessive 64.6
Acceptable 35.4
Excessive Speed Vignette
Excessive 96.0
Acceptable 3.97
Speed Difference Vignette
Excessive 100
Acceptable 0 +
Observed key physical behaviour revealing latent Control Flaws arising within the behavioural model
perceived bearing to NCT
Left 100
Ahead 0
Right 0
Behind 0
Separation
Collision 0
LessThanCPA 100
AvoidanceZone 0
SeparationZone 0
Initial Relative Bearing to NCT
Left 0
Ahead 0
Right 0
Behind 100
Latent Control Flaws (defects), defined in the requirements model (SysML), revealed within the vignettes and observed behaviour
other True
Estimation
UAS RoTA
NCT RoTA
Rigidity
Lost NCT
Delay
Arbitration
0 1
Control Flaw Evidence
Components within the STPA System Control Loop decomposition, identifying the location of the potential flaws within the various system viewpoints
State Information
Flawed 100
Fine .007
Unknown .007
Rules of the Air Model
Flawed 64.5
Fine .007
Unknown 35.5
Non-cooperative Traffic
Flawed .007
Fine 74.5
Unknown 25.5
Emergency Manoeuvre
Flawed .007
Fine 6.58
Unknown 93.4
Sensor Field of View
Flawed .007
Fine 96.3
Unknown 3.66
Heading Subsumption
Flawed .007
Fine 32.1
Unknown 67.9
Inadequate Control Actions (deviations), defined in the requirements model (SysML), and arising with the Control Flaws when they are revealed
No Manoeuvre
Wrong Manoeuvre
Early Manoeuvre Disenga...
Late Manoeuvre Engagem...
Insufficient Manoeuvre
Excessive Manoeuvre
Conflicted Manoeuvre
Inadequate Control Eviden...
A - Risk of collision: an actual risk of collision existed.
B - Safety not assured: the safety of the aircraft was compromised.
C - No risk of collision: no risk of collision existed.
D - Risk not determined: insufficient information was available to determine the risk involved, or inconclusive or conflicting evidence precluded such determination.
UK Airprox Category
None .007
CatD .007
CatC .007
CatB .007
CatA 100
NCT too close
None .007
CatD 100
CatC .007
CatB .007
CatA .007
UAS surprised
True .007
False 100
UAS threatened
True .007
False 100
NCT surprised
True .007
False 100
NCT panicked
True .007
False 100
Symptoms of Inadequate Control Actions from actual observed (simulated) behaviour

Figure 8.29 Inferred States at the Unique Airprox Event

8.3. Final Outcomes – Informing HAZOP

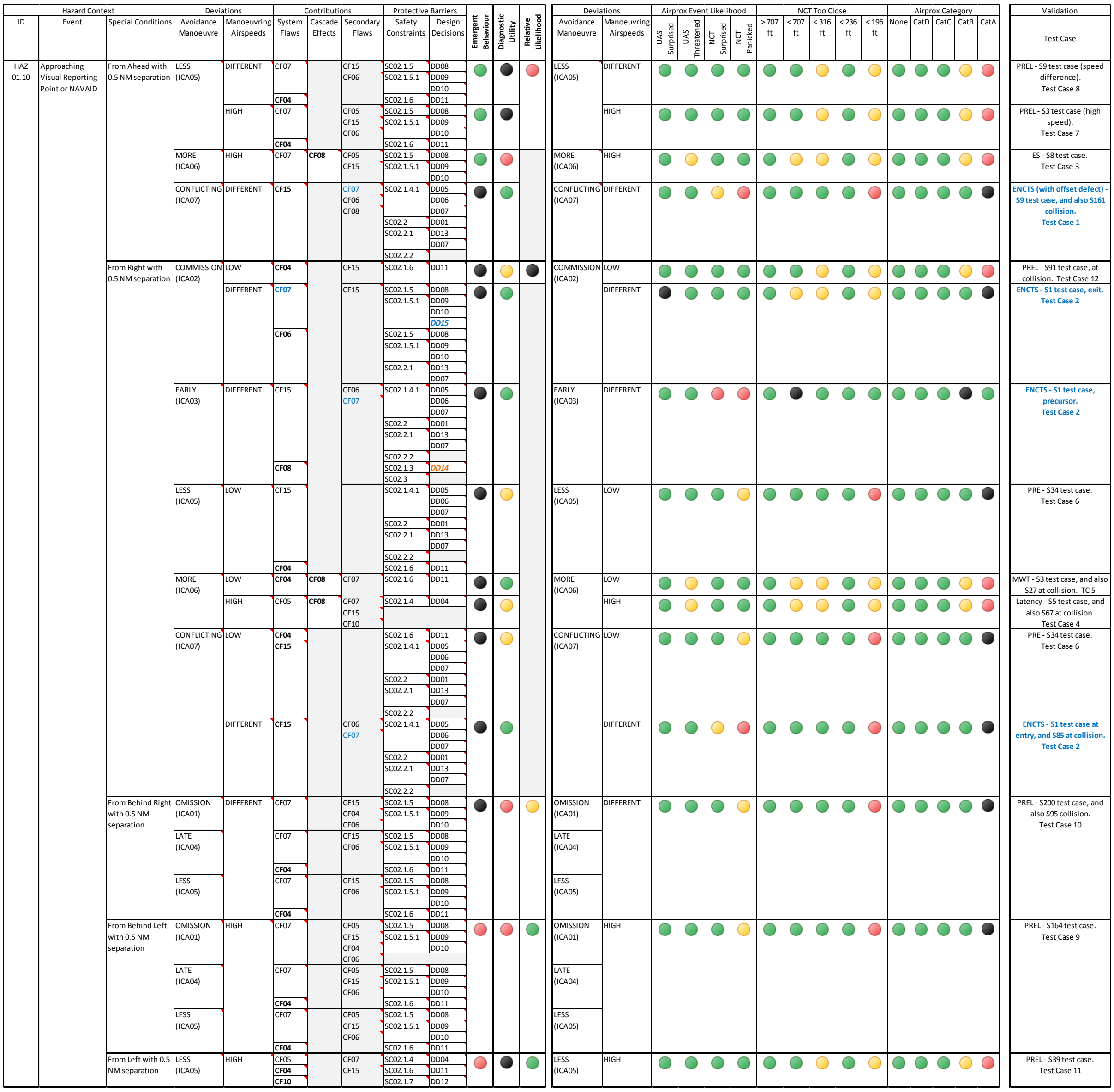
8.3.1. Diagnostic Example – Threat-Offset Direction

As a consequence of observations and investigations using the system dynamics model, additional and modified safety constraints and design decisions may also arise with improved understanding of system behaviour and diagnosable causes. For example, in test cases TC1 and TC2 where the UAV reveals an "S" shaped manoeuvre, and in briefly turning towards the other aircraft, passing immediately in front or behind at 8 and 11 seconds respectively, as the UAV changes direction both CONFLICTED and COMISSION deviations are flagged (Figure 8.22, Figure 8.23). Simultaneously to this the Bayesian model reveals that, "NCT RoTA", "Rigidity" and "Arbitration" control flaws are likely to be flagged, with the "Rigidity" control flaw (CF07) dominating as the most or very likely; and also immediately prior to the near-miss the "Lost NCT" control flaw (CF08) significantly appears too. Interrogating the diagnostic Bayesian nodes, presented earlier in Figure 8.18, the "Emergency Manoeuvre", "Heading Subsumption", and "Non-cooperative Traffic" sub-system models emerge as the most probable causes of this defective behaviour. The "Rigidity" system control flaw occurs where the UAS appears to interpret the Rules of the Air too rigidly when closing within 500ft and is associated with test-cases involving approaches at significantly different airspeeds from ahead and to the right of the centre-line. An alleviation of this form of rigidity on the part of the UAS suggests an additional design decision (DD15) altering the direction to the perceived threat so as be offset by +47° (to the right) to allow for the other aircraft if passing from right to left ahead of the UAV. This decision varies the decision-threshold angular offset value from straight-ahead, and the point where the back-up emergency manoeuvre invokes a specific direction for the emergency turn. Comparing the diagnoses for test-case TC1 (below left) and TC2 (below right), it can be seen that the revised diagnoses does appear to reduce the likelihood that the "Heading Subsumption" model is at fault, but at the same time the "Emergency Manoeuvre" would appear now to be more likely to be at fault.



This is unsurprising as the action performed by the emergency manoeuvre direction decision threshold is on the boundary between emergency manoeuvring **and** subsuming other behaviours, and hence the arbitration of other manoeuvres. In both cases the real culprit is the Non-cooperative Traffic which is travelling too fast for the encounter. It can be argued that this example reveals a potential to diagnose design defects with the application of an appropriate model and an integrated diagnostic methodology.

8.3.2. Enhanced HAZOP



Hazard Context			Deviations		Contributions			Protective Barriers					Deviations		Airprox Event Likelihood				NCT Too Close					Airprox Category					Validation
ID	Event	Special Conditions	Avoidance Manoeuvre	Manoeuvring Airspeeds	System Flaws	Cascade Effects	Secondary Flaws	Safety Constraints	Design Decisions	Emergent Behaviour	Diagnostic Utility	Relative Likelihood	Avoidance Manoeuvre	Manoeuvring Airspeeds	UAS Surprised	UAS Threatened	NCT Surprised	NCT Panicked	> 707 ft	< 707 ft	< 316 ft	< 236 ft	< 196 ft	None	CatD	CatC	CatB	CatA	Test Case
HAZ 01.10	Approaching Visual Reporting Point or NAVAID	From Ahead with 0.5 NM separation	LESS (ICA05)	DIFFERENT	CF07 **CF04**		CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11	🟢	⚫	🔴	LESS (ICA05)	DIFFERENT	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	PREL - S9 test case (speed difference). Test Case 8
				HIGH	CF07 **CF04**		CF05 CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11	🟢	⚫			HIGH	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	PREL - S3 test case (high speed). Test Case 7
			MORE (ICA06)	HIGH	CF07	**CF08**	CF05 CF15	SC02.1.5 SC02.1.5.1	DD08 DD09 DD10	🟢	🔴		MORE (ICA06)	HIGH	🟢	🟡	🟢	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	ES - S8 test case. Test Case 3
			CONFLICTING (ICA07)	DIFFERENT	**CF15**		CF07 CF06 CF08	SC02.1.4.1 SC02.2 SC02.2.1 SC02.2.2	DD05 DD06 DD07 DD01 DD13 DD07	⚫	🟢		CONFLICTING (ICA07)	DIFFERENT	🟢	🟢	🟡	🔴	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	**ENCTS (with offset defect) - S9 test case, and also S161 collision. Test Case 1**
		From Right with 0.5 NM separation	COMMISSION (ICA02)	LOW	**CF04**		CF15	SC02.1.6	DD11	⚫	🟡	⚫	COMMISSION (ICA02)	LOW	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	PREL - S91 test case, at collision. Test Case 12
				DIFFERENT	**CF07** **CF06**		CF15	SC02.1.5 SC02.1.5.1 SC02.1.5 SC02.1.5.1 SC02.2.1	DD08 DD09 DD10 ***DD15*** DD08 DD09 DD10 DD13 DD07	⚫	🟢			DIFFERENT	⚫	🟢	🟢	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🟢	🟢	⚫	**ENCTS - S1 test case, exit. Test Case 2**
			EARLY (ICA03)	DIFFERENT	CF15 **CF08**		CF06 CF07	SC02.1.4.1 SC02.2 SC02.2.1 SC02.2.2 SC02.1.3 SC02.3	DD05 DD06 DD07 DD01 DD13 DD07 ***DD14***	⚫	🟢		EARLY (ICA03)	DIFFERENT	🟢	🟢	🔴	🔴	🟢	⚫	🟢	🟢	🟢	🟢	🟢	🟢	⚫	🟢	**ENCTS - S1 test case, precursor. Test Case 2**
			LESS (ICA05)	LOW	CF15 **CF04**			SC02.1.4.1 SC02.2 SC02.2.1 SC02.2.2 SC02.1.6	DD05 DD06 DD07 DD01 DD13 DD07 DD11	⚫	🟡		LESS (ICA05)	LOW	🟢	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	PRE - S34 test case. Test Case 6
			MORE (ICA06)	LOW	**CF04**	**CF08**	CF07	SC02.1.6	DD11	⚫	🟢		MORE (ICA06)	LOW	🟢	🟡	🟢	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	MWT - S3 test case, and also S27 at collision. TC 5
				HIGH	CF05	**CF08**	CF07 CF15 CF10	SC02.1.4	DD04	⚫	🟡			HIGH	🟢	🟡	🟢	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	Latency - S5 test case, and also S67 at collision. Test Case 4
			CONFLICTING (ICA07)	LOW	**CF04** **CF15**			SC02.1.6 SC02.1.4.1 SC02.2 SC02.2.1 SC02.2.2	DD11 DD05 DD06 DD07 DD01 DD13 DD07	⚫	🟡		CONFLICTING (ICA07)	LOW	🟢	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	PRE - S34 test case. Test Case 6
				DIFFERENT	**CF15**		CF06 CF07	SC02.1.4.1 SC02.2 SC02.2.1 SC02.2.2	DD05 DD06 DD07 DD01 DD13 DD07	⚫	🟢			DIFFERENT	🟢	🟢	🟡	🔴	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	**ENCTS - S1 test case at entry, and S85 at collision. Test Case 2**
		From Behind Right with 0.5 NM separation	OMISSION (ICA01)	DIFFERENT	CF07		CF15 CF04 CF06	SC02.1.5 SC02.1.5.1	DD08 DD09 DD10	⚫	🔴	🟡	OMISSION (ICA01)	DIFFERENT	🟢	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	PREL - S200 test case, and also S95 collision. Test Case 10
			LATE (ICA04)		CF07 **CF04**		CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11				LATE (ICA04)																
			LESS (ICA05)		CF07 **CF04**		CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11				LESS (ICA05)																
		From Behind Left with 0.5 NM separation	OMISSION (ICA01)	HIGH	CF07		CF05 CF15 CF04 CF06	SC02.1.5 SC02.1.5.1	DD08 DD09 DD10	🔴	🔴	🟢	OMISSION (ICA01)	HIGH	🟢	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	PREL - S164 test case. Test Case 9
			LATE (ICA04)		CF07 **CF04**		CF05 CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11				LATE (ICA04)																
			LESS (ICA05)		CF07 **CF04**		CF05 CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11				LESS (ICA05)																
		From Left with 0.5 NM separation	LESS (ICA05)	HIGH	CF05 **CF04** **CF10**		CF07 CF15	SC02.1.4 SC02.1.6 SC02.1.7	DD04 DD11 DD12	🔴	⚫	🟢	LESS (ICA05)	HIGH	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	PREL - S39 test case. Test Case 11

Table 8.2 Enhanced HAZOP Table

In the enhanced HAZOP, as shown above in Table 8.2, four additional features have been incorporated, arising with the analysis of observed behaviour captured from the system dynamics and as incorporated into the Bayesian model:

1) The key hazard event is now decomposed also by a Special Condition – in this case describing each of the five approach sectors;
2) An additional deviation parameter has been identified – Airspeed, with guidewords LOW, HIGH and DIFFERENT;

3) Deviations and Flaws have been rationalised to cover only the more plausible and dangerous cases;
4) A traffic-light scheme has been introduced to denote measures of hazard, risk, expectation of emergent behaviour, diagnostic utility, and overall event likelihood – but in this example only for cases arising with misestimating range and delays in enacting the emergency backup behaviour.

A more detailed view of each of the separate sections of the Enhanced HAZOP can be found in Appendix A – Figure A.2 to Figure A.7, inclusive.

With this traffic-light scheme, green stands for a best or better measure, yellow for medium, red for significantly degraded, and any black "traffic-light" indicates the worst behaviour or outcome for the particular measure, as expressed by equation 8.9 below:

$$\boldsymbol{H = CatA \times CPA,\ R = H \times (M - E) \times D \times L} \tag{8.9}$$

In this expression, *H* represents the final hazard value, *CatA* the relative rate of category A Airprox events, and *CPA* the rate of events observed at the closest classified point of approach. *R* represents the risk, with *M* being the total number of test cases, and *E* expresses the rank within the test cases of the emergent behaviour – with the most emergent test cases awarded the lower ranks, as depicted later in Figure 10.5. Finally, parameter *D* represents the rank of the diagnostic consistency measure, and *L* represents the relative likelihood. Whilst this is the final form in which the HAZOP is to be presented the utility and efficacy of this has not yet been assessed with potential users.

8.4. Recapitulation of Requirements

Although the requirements for this project are relatively broadly defined, derived as they are from the earlier literature review, the purpose of the validation process is to confirm system requirements. Therefore in the context of the project system modelling requirements we might consider to what extent these have been satisfied in the evidence presented so far.

02.23 Simulate different combinations of deviations – with the revision of the initial HAZOP, as previously described, the addition of "Airspeed" as a deviation parameter with new guidewords LOW, HIGH and DIFFERENT demonstrate that at least pair-wise combinations are feasible. Furthermore, in the sequence of events observed through any particular vignette it is clear that more than one guideword might also apply at any time. The Bayesian representation accommodates this. An increase in the dimensionality of the hazard assessment would suggest a further requirement to adopt a methodology such as an experimental Parameter Design method to guide the selection of parameters in multiple parameter investigations;

02.23.01 Model exemplar hazard HAZ010 for the VFR case, from the NIAG SG-134 Hazard Scenarios – a VFR scenario implies that both vehicles are responsible for their own separation, as is the case in the system dynamics model;

02.23.02 Define a simplified but believable and informative mid-air collision avoidance scenario with which to evaluate the hazard model – consequently whilst the VFR case presents the more challenging Sense and Avoid hazard scenario in practice, the design of the simulation is actually greatly simplified in having no need to necessarily represent involvement with Air Traffic Management. The response from the questionnaire suggests that the consensus opinion considers the two situations where the UAV changes direction at the last moment are believable, caused by either misestimating the separation distance or misjudging the intercept point due differences in speed.

02.17 Evaluate hypothetical deviations from intent – avoidance design intent with respect to arbitration is principally embodied within the Subsumption model, with this also featuring in the diagnostic example. This demonstrates that conventional deviations such as OMISSION, COMISSION, LESS, etc. can be applied to manoeuvring, but also abstract hypothetical deviations with guidewords such as ARBITRATION might be evaluated too.

02.20 Associate failure conditions with corresponding failure types – relationships between system control flaws and deviations are updated in the Bayesian representation of the HAZOP with data derived from Monte Carlo simulations of the system dynamics model;

02.14 Combine flow based HAZOP and function based FHA – addressed in part.

Part 3. Questionnaire, Results & Conclusions

Chapter 9. Questionnaire

"I love deadlines. I like the whooshing sound they make as they fly by."

Douglas Adams, "The Salmon of Doubt"

9.1. Questionnaire design

It is generally recognised that the validation of a hazard analysis is difficult to accomplish objectively in the absence of meaningful hazard statistics. Consequently, the hazard modelling method is to be validated, at least in part, by drawing upon the knowledge and experience of a group of subject matter specialists. Within the sequence of activities to validate the system dynamics model, by executing the dynamical and inference models within a series of hazardous vignettes certain subtleties were identified and embodied as multiple-choice questions for the specialists.

General hypothesis – risk events within complex adaptive systems are often caused by deviations from design or operating intentions and unanticipated non-linear causal interactions among system elements that violate safety constraints, and that these interactions cannot be adequately intuitively assessed without an appropriate representation of likely system dynamics. Also to specifically refute that measures and representations of system dynamics do not enhance nor advance the understanding of complex safety related interactions in the beliefs and processes employed in a Hazard and Operability study as applied to a complex system – i.e. a system likely to exhibit some degree of complex and emergent behaviour.

In order to determine where exactly an exemplar model might reveal either deviations from intentions or unanticipated interactions, and whether this is perceived to be true might be evaluated with respect to the expertise of domain experts capable of judging the relevant hazard. An attempt to confirm, validate, or falsify (reject), at least a part of the above hypothesis has been carried out. The responses to a questionnaire describing a series of modelled interactions are recorded and analysed. The respondents have acted in effect as a virtual HAZOP team and judge the advice of the model. Through this we effectively obtain a judgement on at least the first step towards the combination of STAMP and HAZOP. This questionnaire has been developed to assess to what extent dynamic behaviour for a given system might be intuitively grasped by specialists drawn from the field. With the outline system description given it is not the intention that the respondent necessarily attempts to mentally emulate the model so much as to employ whatever experience, engineering judgement, and even "gut-feeling", as they might if considering the manoeuvre "deviations" described therein as though a member of a HAZOP team – the description serves as a guide only to intended behaviour. It was suggested that respondents should allow from one to a maximum of two hours for the reading and completion of the questionnaire.

The questionnaire was anonymous. However, respondents were asked to indicate the extent of their experience (no experience, less than a year of effective experience, one or more but less than five years of effective experience, five or more but less than twenty years of effective experience, more than twenty years of experience) in fields related to this study:

- Operator, piloting, air crew and / or air traffic;
- System safety engineering;
- Robotics or autonomous systems engineering;
- Applying HAZOP to hazard identification;
- System simulation and / or modelling;
- System diagnostics.

This was to help cross-reference, correlate and benchmark the responses to the multiple-choice questions.

Each respondent was also asked to select exactly four terms or expressions that might best qualify their approach and perspective regarding the assessment of system safety, with this information to be used to aid the development of further requirements and help adjust perspectives:

Component Failure Mode; Functional Failure Mode; Mission Phase Analysis; Classified Severity of Effect; Probabilistic Occurrence Rates; Release and Flow of Energy; System Operating State; Causal Propagation; Parameter Deviations; Human Reliability Error; Captured Expertise; Captured Socio-technical Behaviour; Captured Complex System Dynamics; Adaptation and Migration of Behaviour; Active Exploration; Safety Constraint Violation.

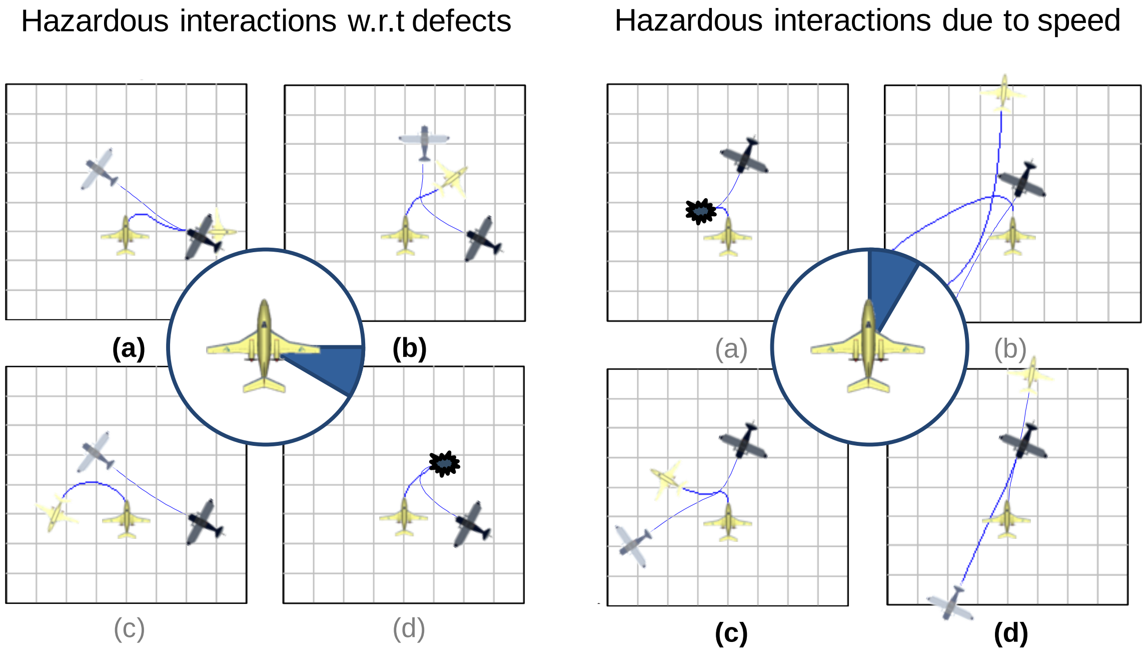
The main body of the questionnaire comprises eleven multiple choice questions contained within four sub-sections. These questions were designed to explore what is, or is not, intuitive in an understanding of the one-on-one collision avoidance scenario, based upon the simplified system model as previously described. Each question is based on a situation with four plausible outcomes, of which only one is completely correct according to the model and its observed behaviour. The aim was to determine how close the respondents' specialist assessment of the situation is to matching the model. So as to provide wider scope in this assessment hedging with the selection of multiple options was allowed wherever it was believed that the true answer lay in more than one choice. Clearly, selecting all four options for a question reveals nothing, this being identical to having selected none at all. The measure of inter-rater agreement, Cohen's Kappa [20], is calculated on the choice of options on a per question basis – generating a pair of agreement values for each question with respect to both the "correct" model based option and an option identified from the consensus.

In the discussion below, options which are correct with respect to the design intent or the model behaviour are outlined with a box, and additional commentary is provided beneath each figure. Figures may relate to more than one question – as described in the commentary. The aggregate analysis and overall results are presented in section 9.2, below. The wording of the actual questions is contained in Appendix B.

9.2. Questionnaire results

9.2.1. General Situation Awareness

The first group of multiple-choice questions endeavour to determine whether each respondent correctly understands the fundamental nature of the interaction as implied by the stated rules of the air within CAP 393 [15]. These were mostly contained within two categories of interactions; a set of interactions caused by the explicit introduction of defects, and another set of interactions with hazards apparently implicit in the speed, or the difference in speeds, of the two vehicles – Figure 9.1.



Hazardous interactions w.r.t defects
Hazardous interactions due to speed
(a)
(b)
(c)
(d)
(a)
(b)
(c)
(d)

Figure 9.1 Behaviours – interpretation of interactions & intuition [32, 33]

The "defects" set contained interactions arising with the NCT initially appearing from the right, and perhaps not too surprisingly the "speed" hazard set all start with interactions with NCT appearing from straight ahead. At the stage were the questionnaire was created the system dynamics modelling employed a somewhat naïve and un-damped model of the system dynamics in relation to the rate at which a turn would be initiated, essentially demanding an instantaneous change in bank angle. This had the effect of producing abrupt manoeuvres that effectively pre-empted the initial direction of turn.

With the basic situations depicted in Figure 9.2, fourteen of the eighteen respondents correctly identify option (a) as the situation where the greatest hazard is most likely to occur. This is commented upon and properly understood by Respondent 6 (R6) as being the case where Rule 10 is invoked with the vehicles approaching head-on, and hence where both vehicles are then obliged to give way by altering course to the right in order to pass with the other on the left. R8 also correctly observes that this further implies that dependent upon how far to the right each is to the other at the start of the encounter; a turn to the right may turn each into the direct path of the other. The opportunity to avoid each other in this case is limited by the achievable rate of turn, closing velocity and initial separation distance – and thence quality of judgement.

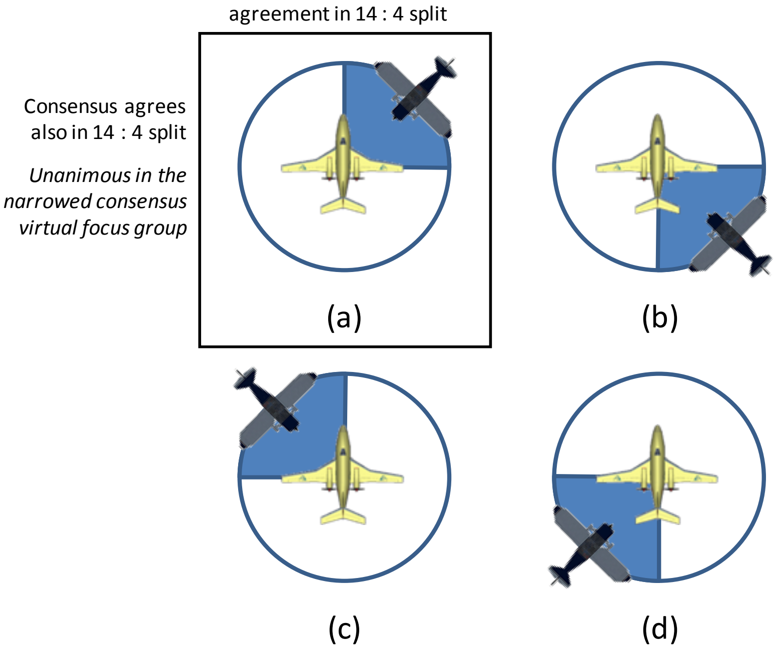

agreement in 14 : 4 split
Consensus agrees also in 14 : 4 split
Unanimous in the narrowed consensus virtual focus group
(a)
(b)
(c)
(d)

Figure 9.2 General situation awareness - general hazardous region at 0.5 NM (Q1)

Therefore for the first question and without applying anything other than a mental model of the situation, the majority of respondents correctly perceive the relevant hazard scenario. With a subsequent narrowing of the respondents to a smaller consensual virtual focus group of 8 people, this smaller group is unanimous in their agreement for this question and there is general consensus around all the questions. Helpfully this smaller group might be taken to be a better indicator of the measure of understanding and agreement in relation to the subsequent questions – in effect reinforcing the identity of the more expert respondents against which any benchmark might be set. Therefore the views of both the larger cohort of 18 respondents and the narrower consensus of 8 respondents are compared across this review of responses.

The next two questions and responses, as depicted in Figure 9.3, relate to narrower versions of the general case presented in the first question. Option (a) in Figure 9.3 is representative of where there is no specific internal defect; rather it is more likely to result in a systemic failure with greater approach speeds – assuming initial separation and manoeuvrability are constants. The correct answer encompasses both rules 9 and 10 of the rules of the air (Figure 6.6), where the UAV is obliged to give way or adjust course, and hence encompassing potential failures due both to systemic and specific internal defects. For option (a) in Q2 only respondent R6 dissented from the (more expert) consensus focus group, although the consensus within the wider group of 18 respondents was 11 in favour of (a), versus 7 against. Again, generally those within the core of the consensus, and arguably therefore with greater convergence of expertise, readily identify the off-set head-on hazard – in the absence of specific defects, with speed not treated as a defect.

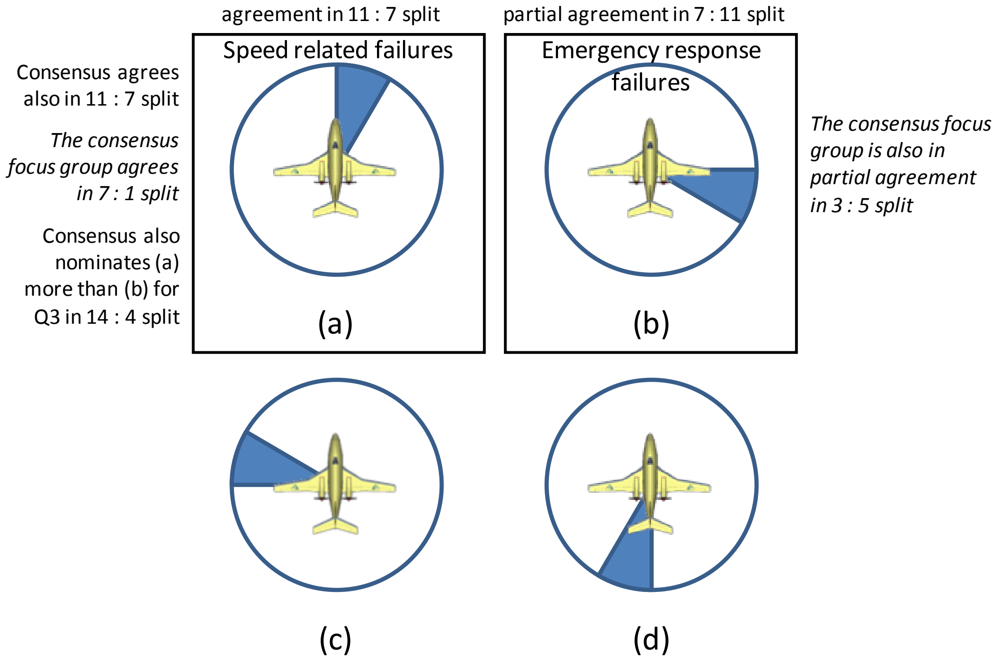

agreement in 11 : 7 split
partial agreement in 7 : 11 split
Speed related failures
Emergency response failures
Consensus agrees also in 11 : 7 split
The consensus focus group agrees in 7 : 1 split
Consensus also nominates (a) more than (b) for Q3 in 14 : 4 split
The consensus focus group is also in partial agreement in 3 : 5 split
(a)
(b)
(c)
(d)

Figure 9.3 General situation awareness - narrower hazards with and without defects (Q2 & Q3)

Option (b) in Q3 represents the case where a variety of internal defects might occur in such that the UAV attempts to interpret Rule 9 but then fails part-way in not correctly invoking the emergency behaviour, but having now turned inappropriately toward the incursion. This situation is not so readily perceived with only 7 out of 18 in partial agreement. In fact 14 out of the 18 prefer option (a) seeing the situation in Q3 the same as in Q2. However, R12 correctly observes that in all options other than (b) the other aircraft should manoeuvre, such that only in case (b) does a failure in the UAV make a collision more likely – but such deduction arises only with this lone observation. The more typical view is expressed by R14 that in a failure case the behaviour is not predictable and therefore conflict in any forward segment could be hazardous. What appears not to be generally appreciated in the distinction between the speed related and emergency response failure cases is the nature of the interaction of the two rule-based behaviours of the UAV. The rules of the air and the emergency manoeuvre behaviours may be in conflict, a threat may be deferred or misperceived and that this is likely to be a greater hazard where an avoidance manoeuvre under a single rigid interpretation of the rules of the air has already been (inappropriately) invoked. The emergency manoeuvre behaviour is in effect an additional safety barrier that might fail.

9.2.2. Anticipation of Interactions

This next section deals with four specific failure or near-miss vignettes within the scenario. Unbeknownst to the respondents, each vertical pair of cases in Figure 9.4 and Figure 9.5 form one of the four specific vignettes, each then with slight variations of direction or warning time to produce the pairs of outcomes as shown. The purpose of these questions was to challenge the respondent to identify the more plausible behavioural outcome for the type of defect or flaw modelled.

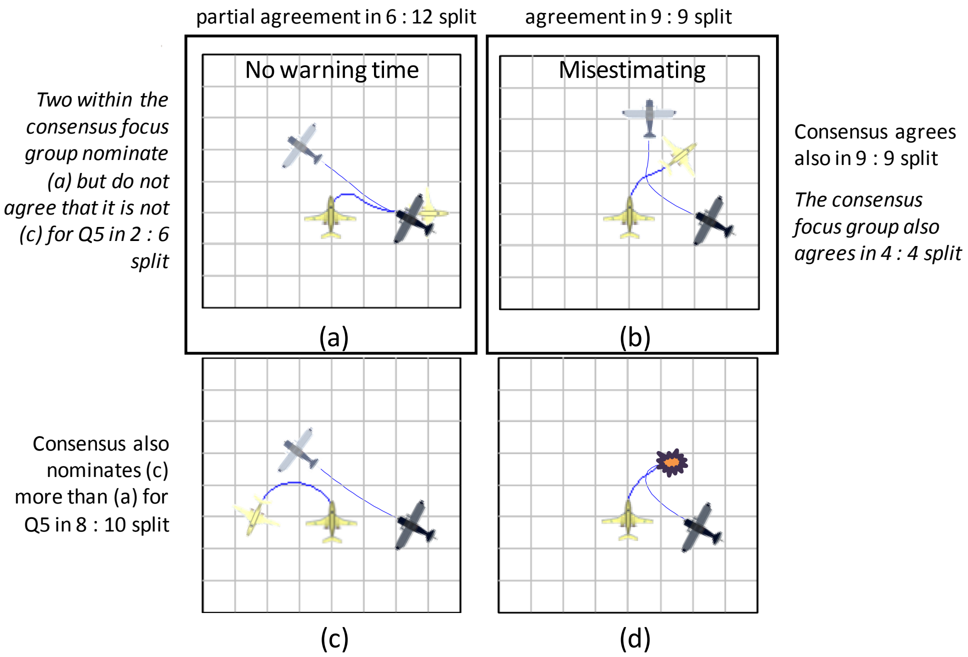

partial agreement in 6 : 12 split
agreement in 9 : 9 split
No warning time
Misestimating
Two within the consensus focus group nominate (a) but do not agree that it is not (c) for Q5 in 2 : 6 split
Consensus agrees also in 9 : 9 split
The consensus focus group also agrees in 4 : 4 split
(a)
(b)
Consensus also nominates (c) more than (a) for Q5 in 8 : 10 split
(c)
(d)

Figure 9.4 Anticipation of interactions - hazardous interactions with respect to defects (Q4 & Q5)

Both questions Q4 and Q5, as presented in Figure 9.4, represent variations of behaviour with regard to the case to give way to traffic from the right. Q4, option (b) represents the outcome where due to a failure to correctly estimate the distance to the threat the UAV implements an interpretation of Rule 9 that causes it to delay invoking the alternative emergency avoidance safeguard. Agreement amongst the respondents that this is the case is evenly split. Of the four within the consensus virtual focus group who accepted this option, R1 and R5 did not agree that option (d) is not true also. However, R1 positively recognised that the situation includes a discontinuity of behaviour – with a delayed switching of behaviour between a rigid (simple) interpretation of Rule 9 and a turn to the left. The wider group of respondents responded similarly in half identifying the correct option is (b), with R3 and R4 identifying the relationship between (b) and (d) – although R4 did not manage to correctly also identify the general situation in any of the three preceding questions! As options (b) and (d) are close variations of the same causal situation, with (b) being the near-miss for case (d), it might be judged that this likely vignette has a 50:50 chance of being recognised through inference and deduction without the aid of a dynamical model – and also not be very dependent upon a participant’s degree of expertise.

Question Q5, option (a) represents the outcome where no warning of an increased threat is given having implemented a rigid interpretation of Rule 9, and therefore inhibiting the overriding safeguard of the emergency avoidance behaviour. However, agreement amongst the respondents that believed that options (a) represented the outcome is partial at best. Of the two who selected option (a), R1 and R6, within the more expert consensus focus group, R1 was unable to also rule out cases (b) and (d). In fact four respondents within the more expert group selected option (c) as the most appropriate outcome – where this is actually the behaviour paired with (a) with the system responding safely as intended. However, R1 expected that there would be no discontinuity of behaviour in this vignette – as the behaviour in changing from a rigid (simple) interpretation of Rule 9 to an otherwise sharp left turn does not occur. Conversely, R8, also a member of the more expert group, considered that none of the options represented the outcome, stating that the UAV would not change course if unaware of the impending conflict. This comment suggests that the wording of the question is perhaps defective. It suggests that the system description ought to have been clearer in describing the differences in the model of detection for a simple line-of-sight detection, sufficient to implement a simple interpretation of the rules of the air, and the perceived range based estimate employed against the more imminent threat. R16 did recognise, however, that the UAV has to give way to the aircraft approaching from the right.

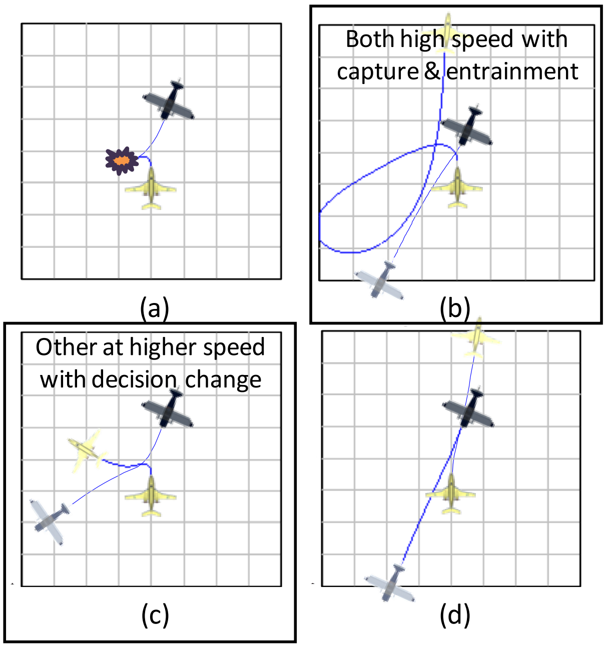
The next two questions, Q6 and Q7 in Figure 9.5 (over the page) represent variations of the off-set head-on case where both vehicles would be expected to alter course to the right under Rule 10 of the rules of the air. Furthermore, in each of these vignettes at least one of the vehicles is travelling faster than 140 kts; the maximum permitted speed at the VFR operating minima of 1500 metres (~0.8 NM) visibility at altitudes less than 3000 feet. Q6 describes the situation where both vehicles are travelling towards each other at the maximum IFR speed limit of 250 kts beneath 10,000 feet. For this question, whilst incorrectly chosen (a), R4 notes that the intruder would turn left according to the emergency manoeuvre rule and that the intruder would turn right according to the RoTA (Rules of the Air). Conversely, R14 queries why the UAV might evade to the left in any case.

In fact in all of the options shown in Figure 9.5 an outcome similar to (a) is quite likely, with variations only in the relative lengths of the two trajectories tracking towards the collision; all of the other options shown represent different near-miss cases. However, there are differences in each of these that the observant respondent ought to have detected. Vignettes (a) and (c) reveal a relative difference in the length of the two vehicle trajectories to the point of crossover, as should be expected from a reading of the wording of Q7 (q.v. Appendix B). For Q7 R3 correctly nominates option (c) but hedges this selection with (a), with the observation that this would depend on deficient behaviour on the part of the intruder. Clearly R3 has seen the relationship between the two outcomes for Q7. Likewise respondent R4 makes the same nomination and hedge, observing that the UAV appears to decide to turn left according to the (emergency) evasion manoeuvre whilst the intruder appears to decide to turn right according to the RoTA, but the difference in speed may make the decision of UAV to turn left a poor one.

identified only once, by R6

Consensus also nominates (a) in addition to (c) for Q7 in 11 : 7 split

Within the consensus focus group (a) and (c) receive five nominations for Q6



(a)

Both high speed with capture & entrainment

(b)

Other at higher speed with decision change

(c)

(d)

Also identified only once, by R6, in the consensus virtual focus group

Consensus also nominates (c) more than (b) for Q6 in 10 : 8 split

The consensus focus group agrees on (c) for Q7 in a 6 : 2 split

partial agreement in 11 : 7 split

Figure 9.5 Anticipation of interactions - hazardous interactions due to speed (Q6 & Q7)

Interestingly, whilst nearly all of the respondents focused on options (a) and (c), some for correct reasons and some not, only one respondent saw the likelihood of outcome (b) in certain interactions with near matching speeds. For Q6 the model illustrates that this type of interaction may result in a period of entrainment before an escape on the part of the autonomous vehicle. This is clearly an emergent behaviour outside the design intent, being typical of that found at or near the "edge of chaos"; as such this is the kind of behaviour that might not otherwise be revealed without a dynamic model – and even then possibly disbelieved when observed in the output of the model.

9.2.3. Inferences

Each of the four vignettes shown in Figure 9.6 and Figure 9.7 represents outcomes with respect to the "separation" variable in the results of four Monte Carlo simulations of 200 runs apiece, with the vertical axis in Figure 9.6 representing distributions of separation trajectory distance in feet. Figure 9.7 presents the same distributions with time slices at 50 seconds as histograms. In both questions, each vignette embodies a variance of ±5° on the headings of both vehicles, ±5 kts on each airspeed, and an initial separation in the range of 0.4 – 0.6 NM. Some vignettes also embody variances of 0 to 2 seconds emergency warning and -75% to +300% tolerances on the perceived range estimation.

The intention with Q8, Q9 and Q10 is to explore the related issues of the respondents’ perception of raw separation and hazard data presentation, in conjunction with model interpretation (q.v. Appendix B, along with Figure 9.6, Figure 9.7, Figure 9.8 and Figure 9.9). Again, unbeknownst to the respondents the outputs presented in these questions are matched together in a specific manner, both with each other and also to the four questions from the preceding section. In these three questions options (a), (b), (c) and (d) each consistently correspond with the same four vignettes – specifically the Minimum Warning Time, Perceived Range Estimation, NCT Excessive Speed (203 kts), and the Both Excessive Speed (250 kts) vignettes, respectively. In addition, options (a) and (b) correspond also to options (a) and (b) in questions Q4 and Q5, Figure 9.4. Likewise, options (c) and (d) correspond also to options (c) and (d) in questions Q6 and Q7, Figure 9.5. All of which are therefore drawn from just these four vignettes, each developed with two cases.

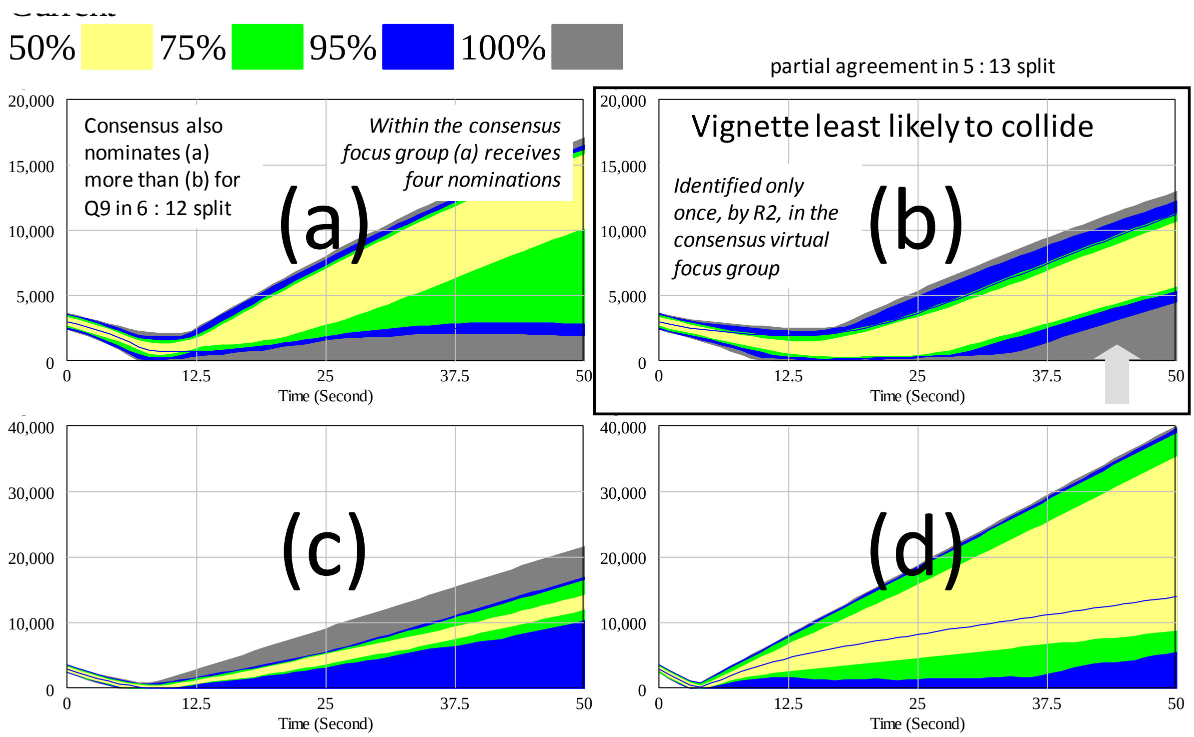
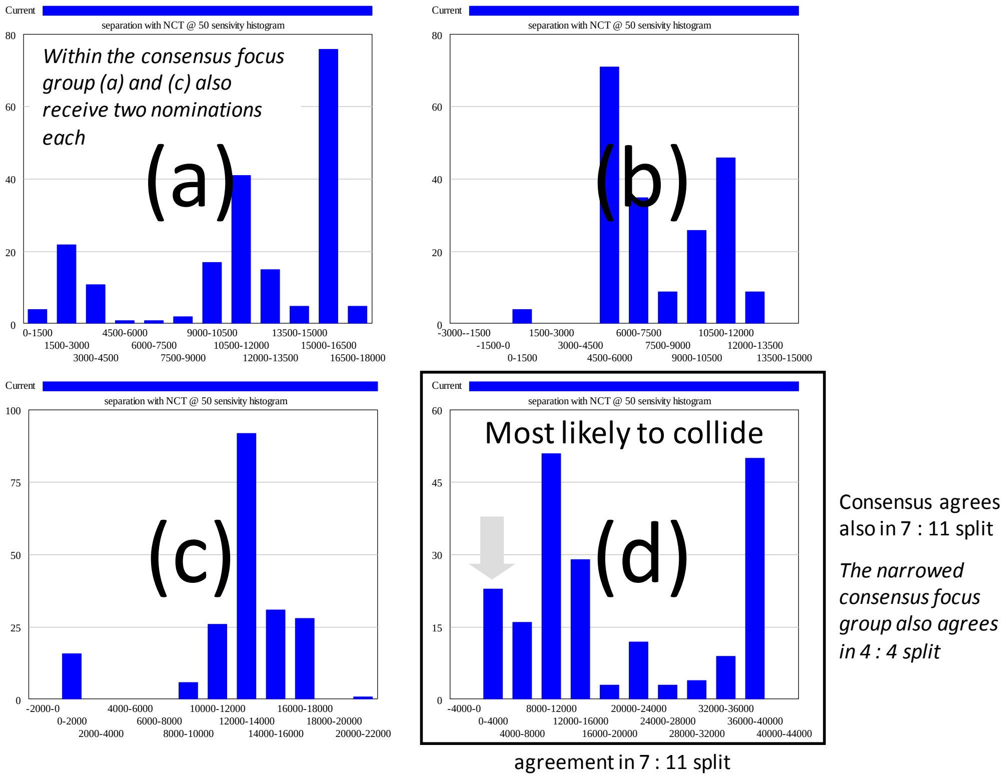


Figure 9.6 Inferences – sensitivity of separation to initial variations (Q8)

However, the respondents were not made explicitly aware of these direct correlations, although a clue does arise later with Q10. Instead in Q8 and Q9 the respondents were asked to make judgements on this more-or-less raw statistical data alone. For the separation trajectory distributions, Figure 9.6, R11 nominated option (a) for Q8, as “there appeared to be a residual risk of collision as the vehicles don't separate as quickly as in option (b)”. By comparing these two options with the corresponding histograms in Figure 9.7, in can be seen that both options (a) and (b) exhibit very similar probabilities where around 2.5% of the interactions result in collisions by 50 seconds, these having occurring at between 8 and 16 seconds respectively (Figure 9.6) – although option (b) is marginally less frequent.

However, as suggested by R11, other factors such as the overall separation profile ought to inform the choice as well. It might be inferred that outcome (b) would seem to indicate clean departures from a set of near-misses, as flagged by the arrow indicated in (b), Figure 9.6. From this one might infer that as profile (a) appears to be indicative of a tendency for the two entities to become entwined – even appearing to threaten to converge again beyond 50 seconds – that (a) is the more risky. And this is the view as expressed by R11, a member of the more expert group, but it would appear that respondent R11 has also perhaps misread the question, as Q8 asked for the identification of the data set exhibiting the least likely risk (q.v. Appendix B) – which is option (b). Other expert group respondents, such as R16, also suggested that the minimum separation distance arising with the simulation ought to be examined. In all cases it would have been reasonable to hedge one's choices in response to Q8 by selecting both options (a) and (b), given the limited information available in these separation trajectory views.



Within the consensus focus group (a) and (c) also receive two nominations each

Most likely to collide

Consensus agrees also in 7 : 11 split

The narrowed consensus focus group also agrees in 4 : 4 split

agreement in 7 : 11 split

Figure 9.7 Inferences – separation histograms at 50 seconds (Q9)

As stated, the histograms in Figure 9.7 represent slices of the separation distributions at 50 seconds – with the one special (although non-obvious) proviso that collisions occurring earlier than this have their separation frozen at collision. Unfortunately, R6 and R8, both members of the expert group felt that these questions were "not very clear" and that they "don't understand the graphs". Nevertheless, the expert group revealed a 50:50 likelihood of correct response, along with a degree of consensus.

But overall this was a poor response, and further consideration needs to be given as to how this type of data might be best presented – e.g. Bayesian data is now presented as pie-chart representations

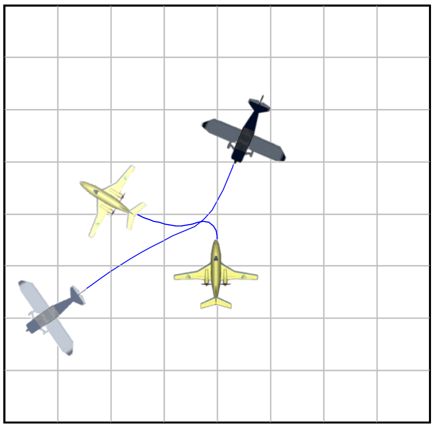
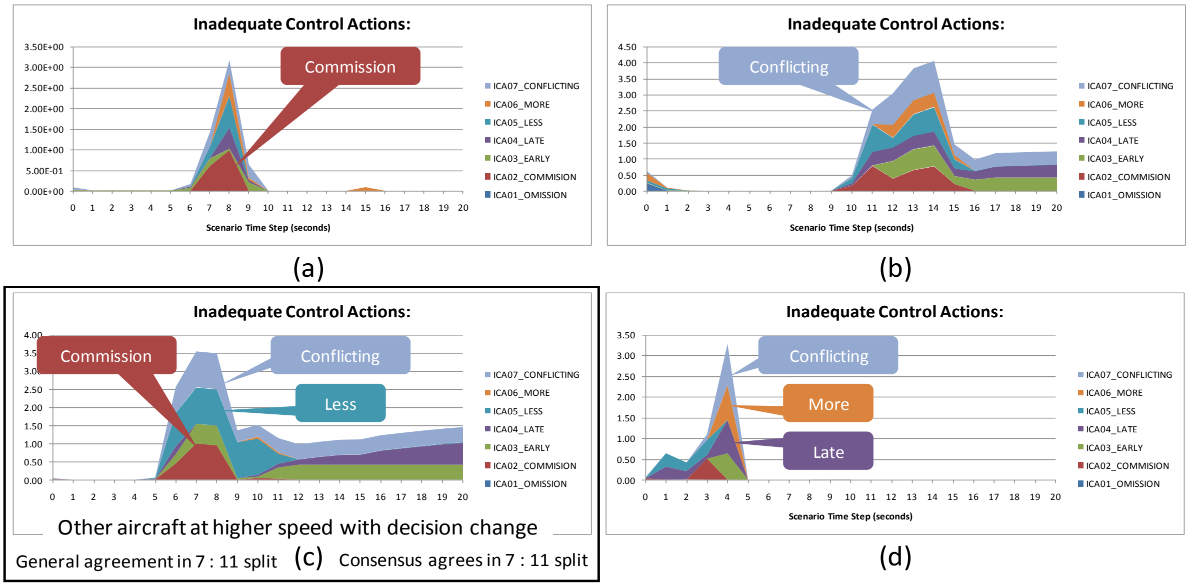


Figure 9.8 A Close Encounter – to derive inferences of deviation from (Q10)

Referring also to the graphical depictions of data described earlier by example in section 7.1.3, the time-series plots presented in Figure 9.9 represent four inferences taken from the previous vignettes that were to be compared with the UAV trajectory in Figure 9.8, above. Each of these time-series based inferences represent the varying degree of belief in different likely system defects as the hazard escalates and then abates, with each band of 'belief' based upon a HAZOP deviation, or Inadequate Control Action, as previously described as 'Hazard Identification' – with these deviations applied to an abstract parameter "Manoeuvre". The question posed to the respondents was: which of the inference plots then represents the close encounter as depicted in Figure 9.8, where there is a failure of the UAV to anticipate the appropriate change in heading on the part of the other aircraft, and thereby then incorrectly to initiate a poorly resolved dangerous emergency manoeuvre taking it toward the other aircraft and also too close then to ensure that it might safely turn away?



Other aircraft at higher speed with decision change

General agreement in 7 : 11 split Consensus agrees in 7 : 11 split

The narrowed consensus focus group also agrees in a 5 : 3 split

Figure 9.9 Inferences – Bayesian inferences in HAZOP & dynamics (Q10)

Perhaps this question is phrased in a slightly leading manner (“poorly resolved” = Conflicting, “toward” = Commission, “too close” = Less), but with this assistance a limited consensus does emerge with the correct selection of (c), although R16 expressed “no confidence”.

9.2.4. Prioritising Risks

a) A minimum warning time < 1 sec
Wider tolerance on perceived range estimation to 25% - 400%
Higher speeds of up to 250 kts in the case of both vehicles
A higher differential speed for the non-cooperative traffic

Consensus nominates (a) more than (c) for Q11 in 7 : 11 split

b) Wider tolerance on perceived range estimation to 25% - 400%
A minimum warning time < 1 sec
A higher differential speed for the non-cooperative traffic
Higher speeds of up to 250 kts in the case of both vehicles

Within the consensus focus group (a) and (b) receive three nominations each

c) Higher speeds of up to 250 kts in the case of both vehicles
A higher differential speed for the non-cooperative traffic
A minimum warning time < 1 sec
Wider tolerance on perceived range estimation to 25% - 400%

identified only twice, by R1 and R15

Also identified only once, by R1, in the consensus virtual focus group

d) Wider tolerance on perceived range estimation to 25% - 400%
A higher differential speed for the non-cooperative traffic
Higher speeds of up to 250 kts in the case of both vehicles
A minimum warning time < 1 sec

Figure 9.10 Prioritising risks – therefore which list is inferred as true? (Q11)

This final question of the questionnaire was intended to reveal whether any lessons might have been taken from the previous questions and applied to the final choice here – where a supposed riskiest consideration is at the top of each of the four lists in Figure 9.10. If the respondent had by now determined that the greater hazard beliefs appear to have arisen with the high-speed off-set head-on vignettes, then the conclusion must be that (c) is the correct answer. In itself, this was also intended to reveal a lesson, in that internal defects may have less effect upon the risk outcome than the relationship between speed (viz. energy) and the first safety barrier – namely the initial safe separation constraint. Within the more expert group only R1 chose option (c), having previously clearly identified the head-on approach options at the start of the questionnaire, also correctly identifying three out four of the interactions – albeit with an equal amount of hedging, and identified two out of three of the inferences, but did hedge this last question with option (d) as well. So unfortunately, only R1 responded as expected, whereas R15 answered Q11 perfectly, but was wrong on nearly all of the other questions – although the general risk direction, minimum warning time and differential speed interactions were correctly identified, with some hedging. Overall, the respondents’ interpretation of the effect of a reduced warning time appeared to be inconsistent. For example R4 appears to believe that a greater warning time might be more dangerous, whilst R12 correctly observes that a warning of less than one second is likely to be as bad as no warning at all.

9.3. Intermediate Outcomes – Questionnaire

The more general lessons to be taken from the responses to the questionnaire at this point are:

- The narrowed consensus (expert) group exhibited greater correct general situation awareness;
- The more extreme emergent behaviour was likely to be disbelieved, even by most experts;
- Inferences drawn from raw data and Bayesian interpretations are unlikely to be correctly interpreted, although for an appropriate representation the expert group may correctly out-vote dissenters;
- Internal defects are seen as a greater hazard than increases in system energy v safety constraints.

Chapter 10. Summary of Results, Outcomes and Future Work

“I may not have gone where I intended to go, but I think I have ended up where I needed to be.”

Dirk Gently, "The Long Dark Tea-Time of the Soul" by Douglas Adams

10.1. Summary of work

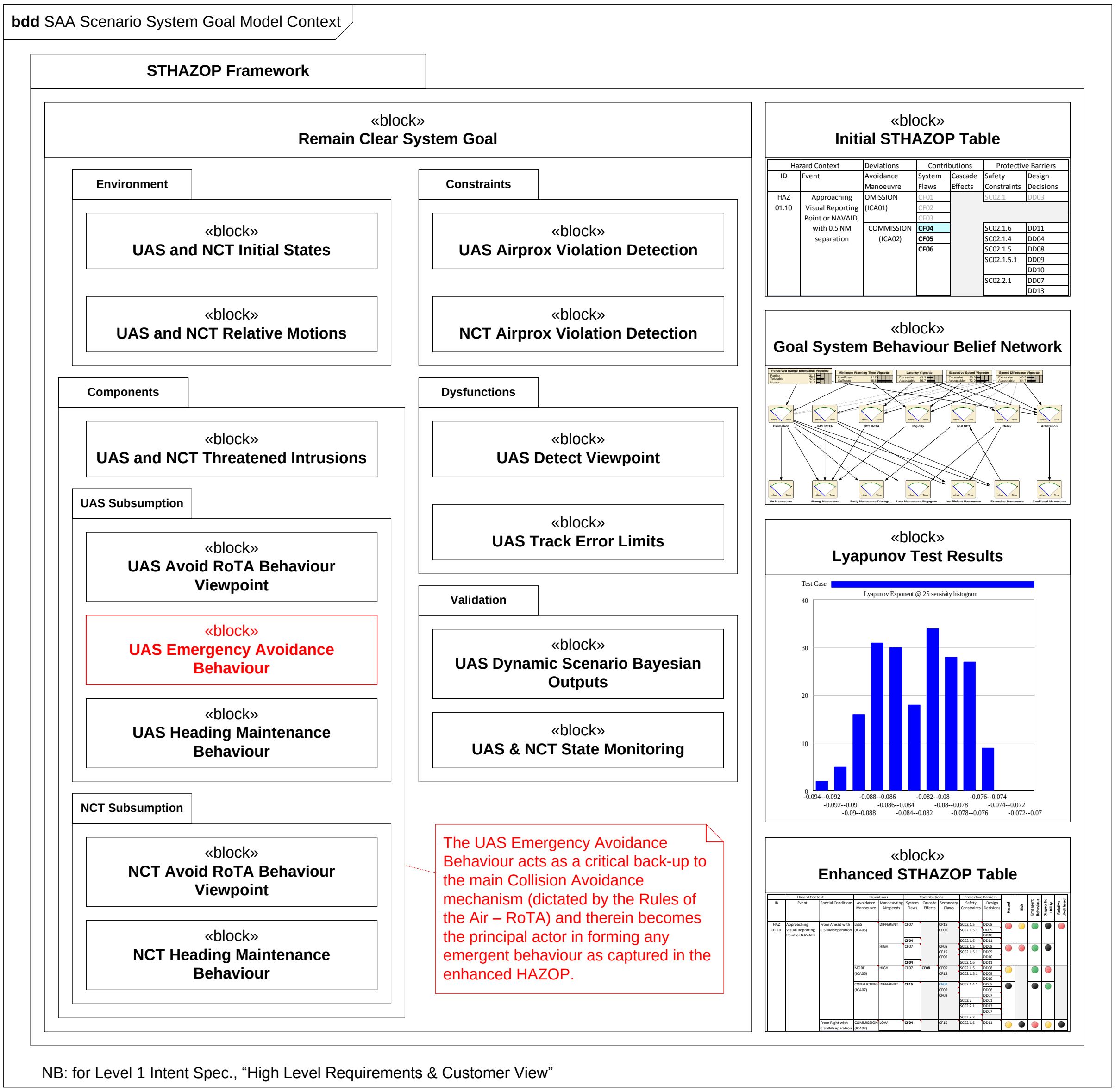

bdd SAA Scenario System Goal Model Context
STHAZOP Framework
«block» Remain Clear System Goal
Environment
«block» UAS and NCT Initial States
«block» UAS and NCT Relative Motions
Constraints
«block» UAS Airprox Violation Detection
«block» NCT Airprox Violation Detection
Components
«block» UAS and NCT Threatened Intrusions
UAS Subsumption
«block» UAS Avoid RoTA Behaviour Viewpoint
«block» UAS Emergency Avoidance Behaviour
«block» UAS Heading Maintenance Behaviour
NCT Subsumption
«block» NCT Avoid RoTA Behaviour Viewpoint
«block» NCT Heading Maintenance Behaviour
Dysfunctions
«block» UAS Detect Viewpoint
«block» UAS Track Error Limits
Validation
«block» UAS Dynamic Scenario Bayesian Outputs
«block» UAS & NCT State Monitoring
The UAS Emergency Avoidance Behaviour acts as a critical back-up to the main Collision Avoidance mechanism (dictated by the Rules of the Air – RoTA) and therein becomes the principal actor in forming any emergent behaviour as captured in the enhanced HAZOP.
«block» Initial STHAZOP Table
«block» Goal System Behaviour Belief Network
«block» Lyapunov Test Results
«block» Enhanced STHAZOP Table
NB: for Level 1 Intent Spec., “High Level Requirements & Customer View”

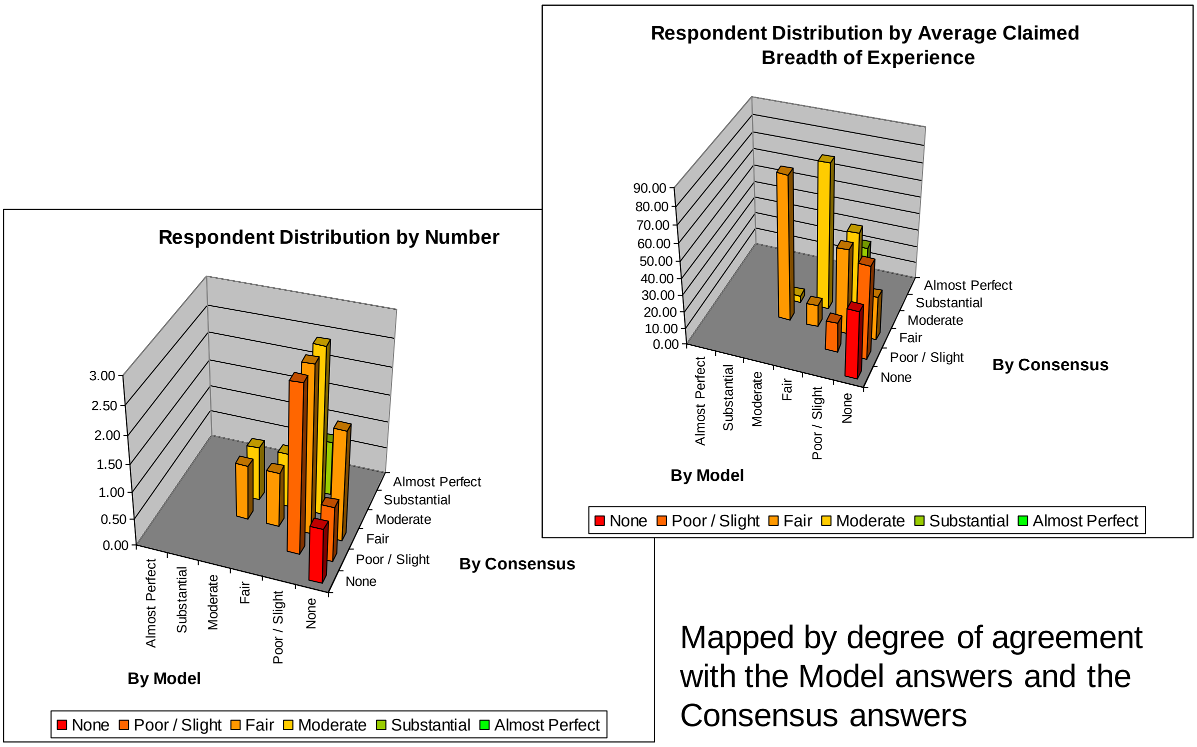
Figure 10.1 Sense and Avoid Scenario System Goal Model

This work proposes a building-block approach to the construction of an integrated system dynamics and hazard inference model, incorporating the generic elements as identified first in Figure 6.4, and now in detail in Figure 10.1 for the specific “Sense and Avoid” exemplar. A broad purpose, as declared with the objective to formulate this within a provisional systems engineering modelling framework, is to explore the entirety of the likely modelling domain so that such a framework might be broadly identified; in so far as that is possible within the scope of the project.

10.2. Principal results

10.2.1. Questionnaire Responses, Hypothesis and Preferences

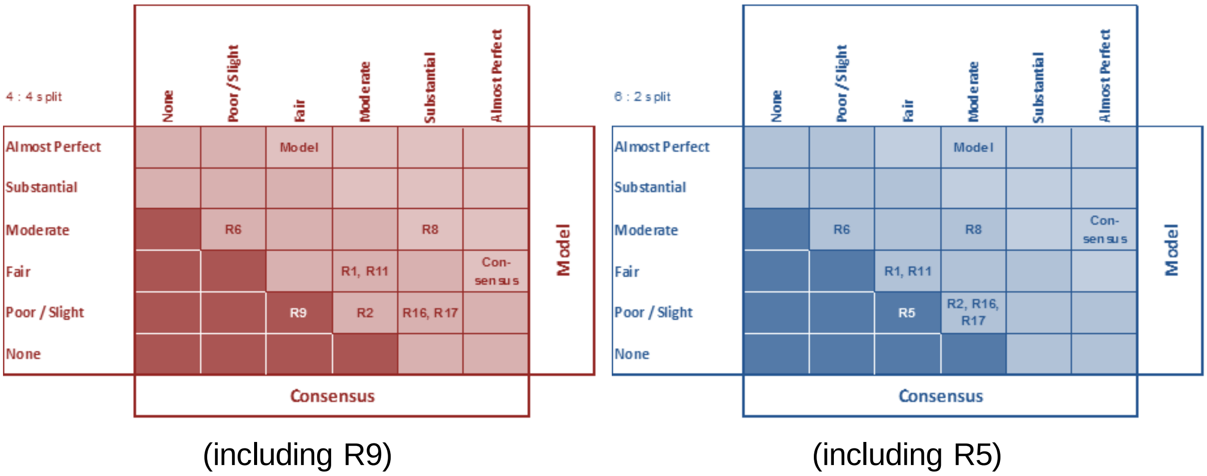
The assumption that risk events within complex adaptive systems are caused by deviations from design or operating intentions and unanticipated non-linear causal interactions among system elements that violate safety constraints, as suggested by Leveson [67, 70], and that these interactions cannot be adequately intuitively assessed without an adequate representation of likely system dynamics – appears at least not to have been rejected. That this is untrue is refuted in the responses to the questionnaire – even where the respondents disagree, and especially where individual respondents also disagree with the consensus. The belief in the modelled interactions is supported to some degree with the responses to all of the questions; in that half of the modelled interactions are often correctly identified, with the less believed interactions nearly so when considered only within the consensus group of the marginally more expert, isolated from the less expert respondents. However, this view is slightly undermined by a degree of interpretative bias arising with the phrasing and representation in a few of the questions.



Mapped by degree of agreement with the Model answers and the Consensus answers

Figure 10.2 Questionnaire Respondent Agreement Distributions (for 18 respondents)

The questionnaire was distributed in an electronic form to more than thirty recipients, each enabled with a semi-automated e-mail return. As a result of this, eighteen recipients responded, usually after a degree of encouragement by way of follow-up calls, and then with these responses being rendered anonymous for processing. These responses were processed with a question by question comparison of multiple-choice responses – compared both against the model-based benchmark and against the emergent consensus responses derived for each question. With the overall respondent agreement measure then classified and counted in each category, along with a count weighted by the average respondent relevant experience in each category, the results as depicted in Figure 10.2 were produced. In this figure, note also that where respondents are revealed to be in poor agreement with the consensus, in general they also exhibit more-or-less equally poor agreement with the model. The distribution is of a generally diagonal nature – with "poor" responses being grouped in the lower corner.



	None	Poor / Slight	Fair	Moderate	Substantial	Almost Perfect
Almost Perfect			Model			
Substantial						
Moderate		R6			R8	
Fair				R1, R11		Con-sensus
Poor / Slight			R9	R2	R16, R17	
None						

	None	Poor / Slight	Fair	Moderate	Substantial	Almost Perfect
Almost Perfect				Model		
Substantial						
Moderate		R6		R8		Con-sensus
Fair			R1, R11			
Poor / Slight			R5	R2, R16, R17		
None						

Figure 10.3 Model versus Consensus – "expert" respondent agreement maps

In two sub-groups, of eight of the more "expert" respondents, extracting their respective measures of agreement separately produced the two consensus-agreement maps in Figure 10.3. Note that for these nearer-to-consensus sub-groups the replacement of one marginal respondent (R9), with an almost equally marginal (but different) respondent (R5) improved the consensus – indicating that respondent R5 is not only more in agreement with the group, but also moved the consensus towards the model. The detailed respondent results, along with additional comments, form the basis of the commentary to each question response previously presented in Chapter 9 – the Questionnaire itself is given in full in Appendix B. Overall, the eighteen respondents' opinions of the hazardous interactions, in Figure 10.4, identified the "correct" behaviours as shown; noting that ***all*** of the interactions were produced from model behaviours, but each with initial conditions other than those specified in the questions

A subsidiary aspect of the questionnaire also asked which of four terms or expressions might best qualify the respondent's approach and perspective toward system safety assessment. Combining this with the general hypothesis appears to reveal some limited support for the first part of the basic hypothesis (belief in the effect of deviations and unanticipated non-linear interactions) but only a weak requirement for a non-chain-of-events, behaviour-constraint-based approach.

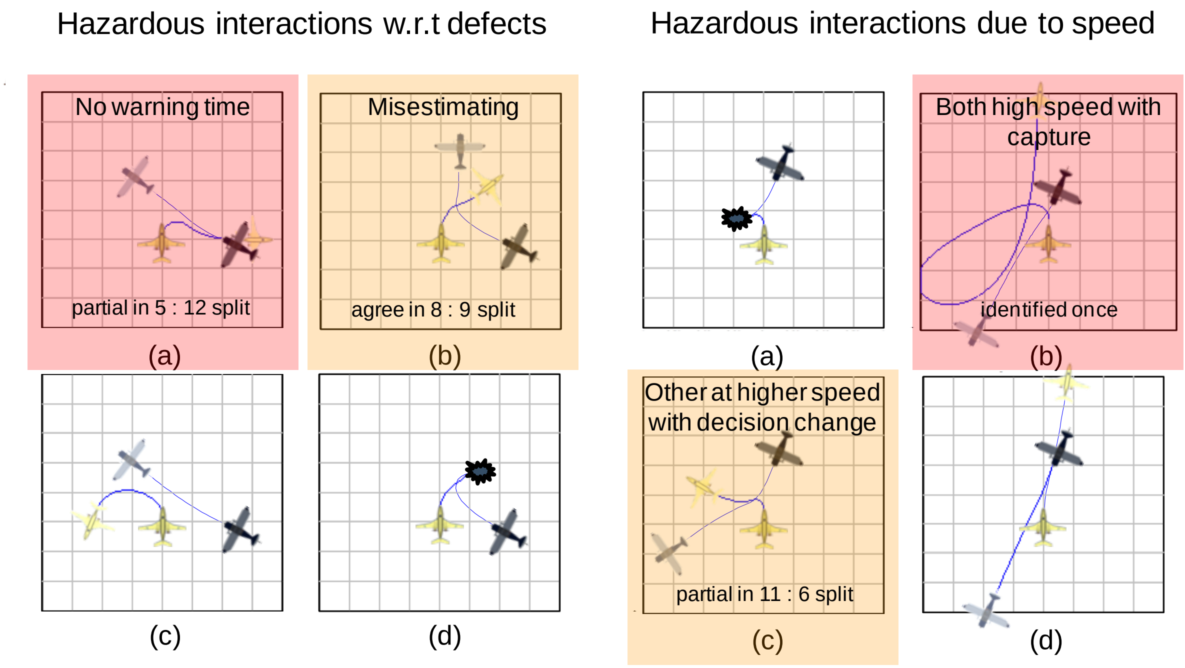

Hazardous interactions w.r.t defects
Hazardous interactions due to speed
No warning time
partial in 5 : 12 split
(a)
Misestimating
agree in 8 : 9 split
(b)
(a)
Both high speed with capture
identified once
(b)
(c)
(d)
Other at higher speed with decision change
partial in 11 : 6 split
(c)
(d)

Figure 10.4 Responses to Questions Concerning Anticipation of Interactions

In these responses there appears to be only minority support for *Causal Propagation* (5 out of 18 respondents, and 2 out of 8 in the more expert consensus), and *Behavioural Modelling and Adaptation* (3:18 in general and 2:8 expert) as priority system safety assessment considerations. Conversely, the general bias of the respondents is very much more inclined towards conventional chain-of-event safety assessment priorities of *Component Failure Modes* (9:18 and 5:8), *Functional Failure Modes* (12:18 and 5:8), and *Probabilistic Occurrence Rates* (8:18 and 5:8). Therefore in applying this modelling approach to an enhanced HAZOP process, consideration might be given to the following observations:

- More experienced system safety engineers and least experienced ‘operators’ (private pilots) exhibit the greater measures of agreement here, both with the model and consensus; although with this respondent sample, the less experienced HAZOP group and the more experienced System Simulation cohorts would likely have more influence upon the associated stakeholder requirements and safety assessment approach due to their majority.
- The apparent requirement for HAZOP and STAMP-like features in complex system safety analysis is judged to be substantially less important to this group of respondents than features usually associated with conventional Functional Hazard Assessment, Fault Tree and Failure Modes and Effects analysis, etc. (re: “chain-of-events”)
- *Functional Failure Modelling* – has by far a greater degree of support. Hence the hazard assessment process ought to at least accommodate deviations of behaviour that include the guidewords Omission, Commission, and Error applied to relevant parameters and signals.

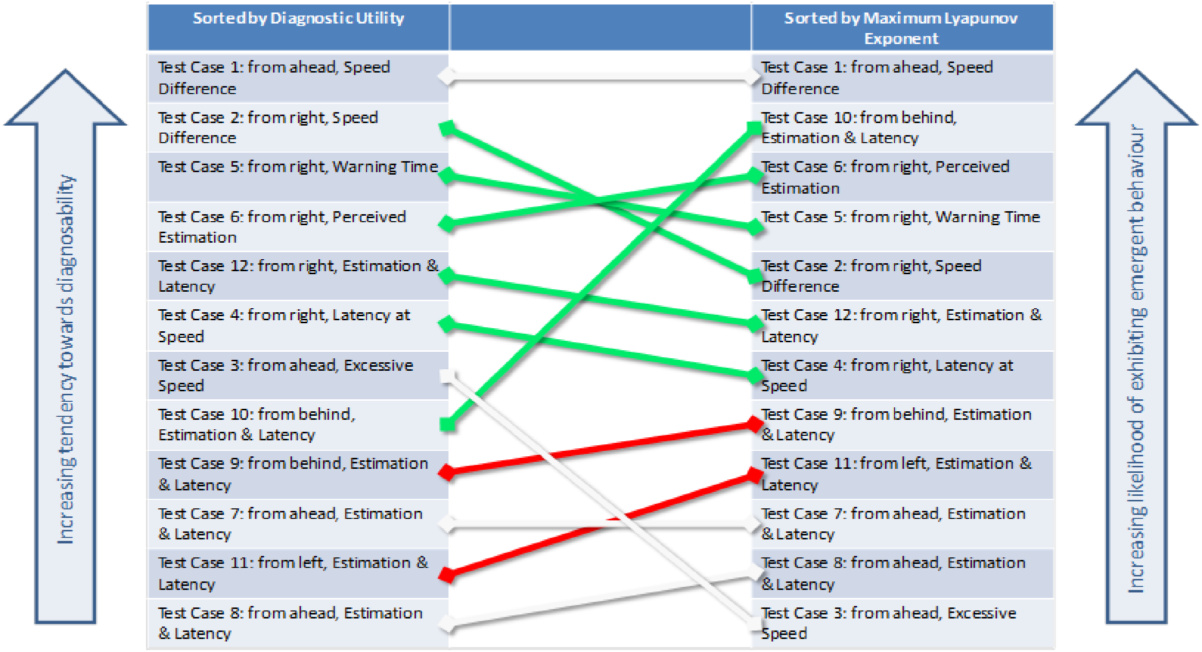
- *Causal Propagation Modelling* – is supported by more than a third of those respondents claiming some experience in system safety engineering, with half of these being in the more experienced group, both of whom demonstrate ‘fair’ agreement with the model; this ought to be built upon.
- *Behaviour Adaptation & Migration Modelling* – is seen as less important by the respondents, but is also an important constituent of STAMP, and so trust needs to be created where complex and emergent behaviour is likely and otherwise difficult to comprehend or intuit from design intent.
- However, *Complex System Dynamics Capture* – appears as a preference of those less encumbered by their degree of apparent expertise in this scenario, although this does include a holder of a private pilot’s licence – perhaps these are more inclined toward new perspectives.
- Finally, *System Expertise Capture* – emerges with just two respondents, of whom one exhibits ‘moderate’ agreement with the model, as high a degree of agreement as any exhibited here.

10.2.2. Avoidance and Emergent Properties

Close inspection of the Lyapunov exponent distributions, showing the spectrum of exponents as histograms, reveal that some of these approximate to Gaussian distributions, or skewed distributions as for example in Figure 8.20, whilst others appear to be bi-modal (q.v. Figure A.12 to Figure A.16, Figure A.31 and Figure A.33, Appendix A). Where these modes are split between negative and positive groups, this likely indicates that certain Monte Carlo sets contain both a more dissipative (negative Lyapunov) group wherein after a simple manoeuvre each then quickly settles towards their respective original headings, and possibly a more emergent (positive Lyapunov) group, involving a degree of entrainment and / or entailing prolonged or excessive manoeuvring.

In ordering the test cases by rank of the most positive (maximum) Lyapunov exponent values, as calculated over 25 seconds from initiation, wherein there ought either to be resolution in the collision avoidance or continuing dangerous interaction[6], and comparing this rank with a similar ranking by the diagnostic utility of the same cases, reveals an intriguing result – Figure 10.5. This suggests an additional hypothesis for the possible existence of a correlation between “diagnosability” (equation 8.6) and emergent behaviour. Within this correlation there also appears to be some grouping of the “approaching from the left” (red), “approaching from the right” (green) and “approaching from ahead” (white) cases. If such correlation actually exists it might be postulated that this occurs because emergent cases are an embodiment of more complex behaviour within which to form a diagnosis. For example the “S” shaped avoidance behaviour observed in test cases TC1 and TC2 – Figure 10.6 – are both indicative of the same control flaw as each arise with the same combination of deviations, although from slightly different directions and also where the earlier system dynamics model used in TC1 embodies different and less realistic (effectively instantaneous) responses.

[6] Test case 12 contains an outlier Lyapunov exponent that produces a positive value until 250 seconds.



Increasing tendency towards diagnosability

Sorted by Diagnostic Utility

- Test Case 1: from ahead, Speed Difference
- Test Case 2: from right, Speed Difference
- Test Case 5: from right, Warning Time
- Test Case 6: from right, Perceived Estimation
- Test Case 12: from right, Estimation & Latency
- Test Case 4: from right, Latency at Speed
- Test Case 3: from ahead, Excessive Speed
- Test Case 10: from behind, Estimation & Latency
- Test Case 9: from behind, Estimation & Latency
- Test Case 7: from ahead, Estimation & Latency
- Test Case 11: from left, Estimation & Latency
- Test Case 8: from ahead, Estimation & Latency

Sorted by Maximum Lyapunov Exponent

- Test Case 1: from ahead, Speed Difference
- Test Case 10: from behind, Estimation & Latency
- Test Case 6: from right, Perceived Estimation
- Test Case 5: from right, Warning Time
- Test Case 2: from right, Speed Difference
- Test Case 12: from right, Estimation & Latency
- Test Case 4: from right, Latency at Speed
- Test Case 9: from behind, Estimation & Latency
- Test Case 11: from left, Estimation & Latency
- Test Case 7: from ahead, Estimation & Latency
- Test Case 8: from ahead, Estimation & Latency
- Test Case 3: from ahead, Excessive Speed

Increasing likelihood of exhibiting emergent behaviour

Figure 10.5 Comparison of Test Case Rank Order for Diagnosis and Emergence

Comparing the respective λT distribution profiles of these two test cases also reveals a similarity between the two, although the much faster response of the model used in TC1 produces a greater range and a higher value (0.23) for the maximum exponent at 25 seconds, Figure 10.6, below.

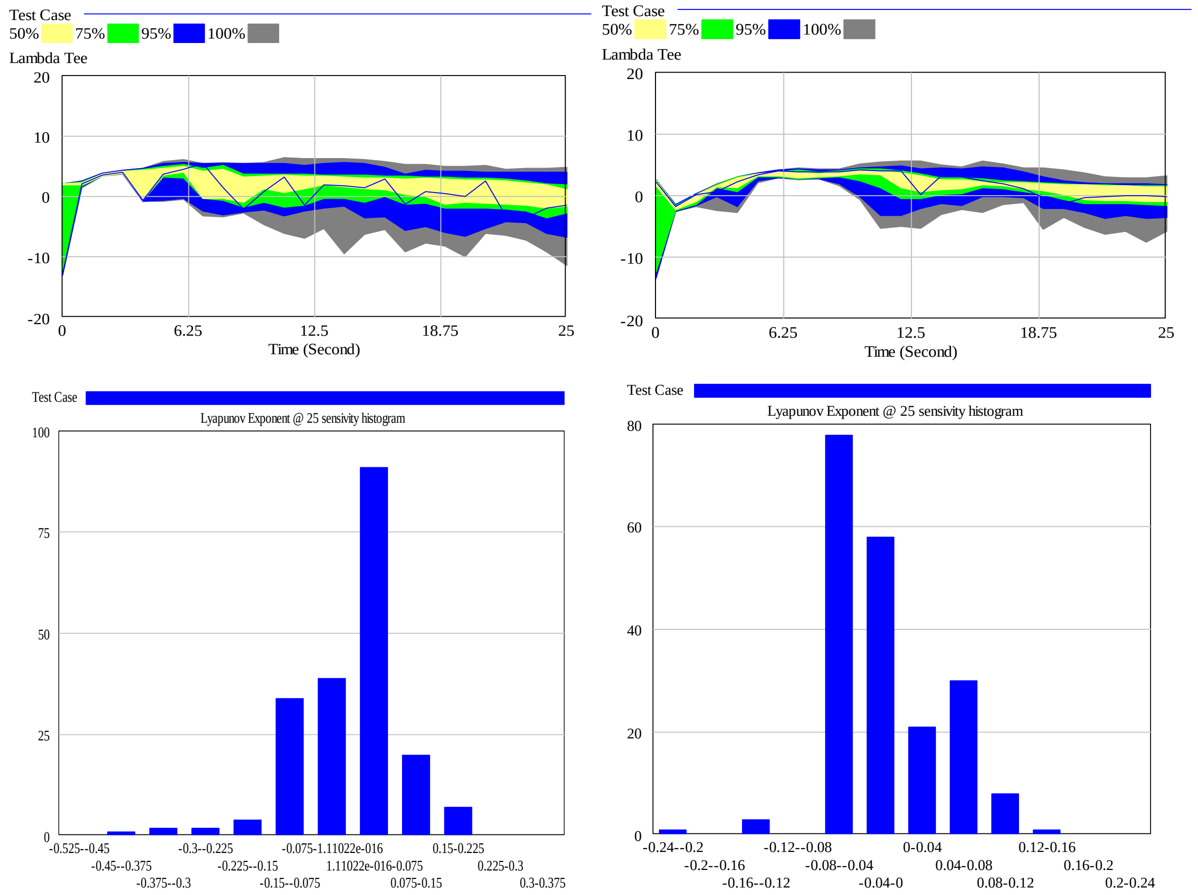


Figure 10.6 Lyapunov Distributions for Test Cases TC1 (left) and TC2 (right) to 25 Seconds

The 9000 sampled data points reveal a degree of a positive correlation between the Lyapunov exponent value and constraint violation – where these data points arise with the 200 Monte Carlo runs, each for nine different speed combinations and from five different initial directions of approach – where for each sample run there is also a corresponding Lyapunov exponent and Airprox (constraint violation) value. Illustrating this previously in Figure 8.24, this Lyapunov exponent was shown plotted against initial relative heading, and then also in Figure 8.25 the Airprox violation value is likewise plotted against initial relative heading. Comparing these two scatter plots suggests a possible correlation between Lyapunov exponent and Airprox. This suggests that even a provisional mapping of the distribution of Lyapunov Exponents – as in these scatter plots and also in Figure 10.7 – reveal a visible relationship in air-proximity constraint violation detections and initial relative headings, especially in the most hazardous situations where the other aircraft approaches from the right (~270°).

In support of the inclusion of a measure of the maximum Lyapunov exponent as a proxy for safety constraint violation, a plotted regression analysis reveals various degrees of correlation dependent upon the initial relative bearing, Figure 10.8. This suggests a hidden relationship between Airprox and Lyapunov within the actual system behaviour; although this may be explainable with the scenario (and perhaps thereby more generally) with proper understanding of the nature of the interactions. Head-on interactions are obviously dangerous but entail virtually no complex manoeuvring behaviour as having insufficient space within which to manoeuvre, revealing in general a negative correlation. However all of the remaining interactions reveal positive correlations, especially with the interactions arising from the behind, both right and left, where violations are rare but usually prolonged where they involve entrainment – in this model. Therefore, it can be argued that both the Lyapunov exponent combined with a domain specific measure of safety constraint violation (e.g. Airprox) might usefully enhance and contextually filter the final HAZOP output. A breakdown of the individual regression plots with respect to all heading and speed combination vignettes are to be found in Figure A.20 to Figure A.29, in Appendix A.

Note in this “Sense and Avoid” exemplar scenario that the "Initial Relative Heading" is the critical parameter (that is the heading of the other aircraft with respect to the vehicle of interest – the putative UAV), where "Manoeuvre" (i.e. variations from this initial setup by the UAV) is used as the parameter to which the deviations and guidewords apply, whilst the combined "rates of turn" (of both entities) provide the basis of measurement for the Lyapunov Exponent – in effect this measure preserves the history of the degree of manoeuvring up to a particular point in time: where negative Lyapunov values indicate a system settling down into a new steady-state, and positive values indicate the opposite, and with this likely emergent and possibly dangerous behaviour. Clearly the measures of "Initial Relative Heading", "Manoeuvre", and "rates of turn" obviously have direct relationships to each other, and the nature of this relationship ought to be considered in the identification of any parameter and guideword to be used within an enhanced HAZOP.

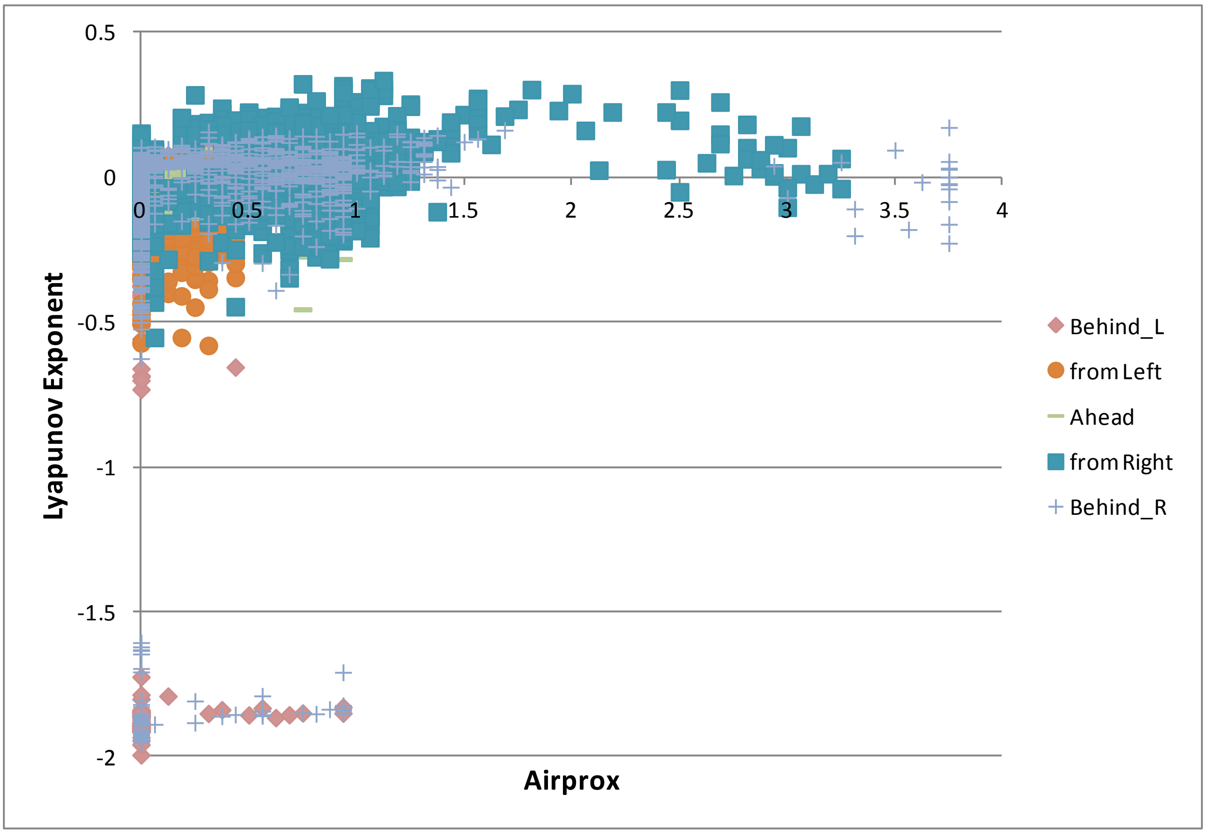


Figure 10.7 Lyapunov v Airprox Correlation as Scatter Plot (9000 data points)

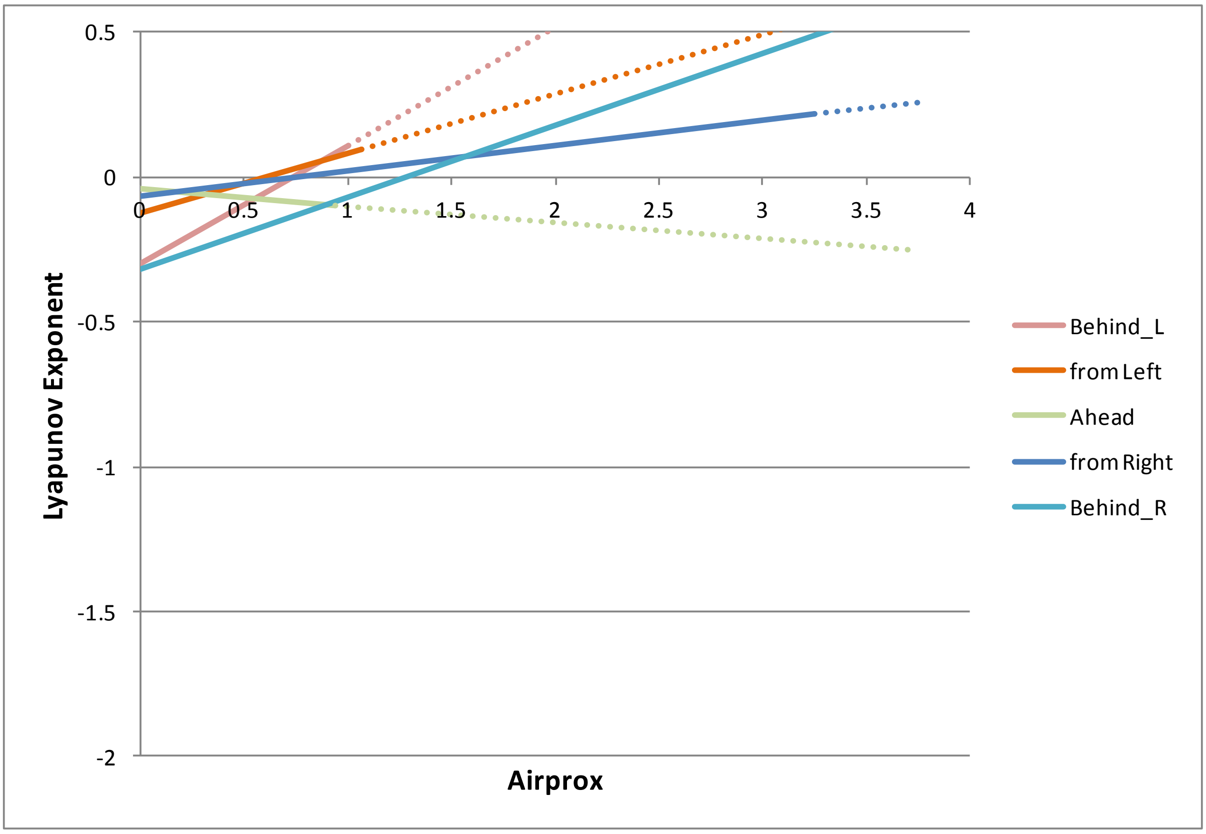


Figure 10.8 Lyapunov v Airprox Correlation as Linear Regression Plots

10.3. Lessons learnt, alternative approaches and improvements

10.3.1. Questionnaire

The aim of this work overall is to investigate the development of Hazard Assessment methods employing the simulation of behavioural models within a range of plausible deviations. In a similar context, Alexander, et al [1], briefly describe how simulation models representing SoS have been used to identify hazards with deviations within the HIRTS DARP project. Alexander's approach also employs machine learning to extract patterns from the simulation outputs. The expectation is that such techniques may aid the construction of goal structured representations, as required in the production of safety-cases satisfying Def-Stan 00-56 [76], by providing a 'vector' for the description of complex environments and interactions, and thereby offering advantage over free-text policies in being a more clearly expressed representation of plausible outcomes [1]. However, the issue of plausibility and acceptance of these representations itself also ought to be addressed. Therefore, the aim of the questionnaire, as developed and analysed in this work, is to assist in the formulation of a requirement deeper than that initially identified for the project [30], in considering the terms of plausibility and extent that the required expertise within a HAZOP team might be enhanced with behavioural models and simulation where complex system behaviour results. Part of this is also to understand where exactly these models might be placed with respect to the position of the experts, and what role they might play in either bolstering trust, in the likely system behaviour, or alternatively in prompting additional appropriate questions of the team.

The results of the questionnaire make it clear that for a range of suitably technically informed respondents, the overall responses to the putative outcomes expressed by the model vary widely. Where a more narrowly focused group of respondents has been considered, containing slightly more of the relevant expertise in piloting, system safety, simulation and HAZOP, the view narrows, becoming more consensual whilst exhibiting 'fair' to 'moderate' degree of agreement with the model. More specifically, where a respondent with a system reliability background and some pilot experience is included in the expert group the model-consensus agreement becomes only 'fair', but rises to 'moderate' when replaced by a respondent having considerably greater simulation experience – but none of piloting an aircraft. From this and the general result for the more expert group, one might conclude that this marginally more expert group exhibits slightly more trust in the model of behaviour, and that having a simulation and modelling viewpoint is also likely to enhance this more than experience in the piloting of an aeroplane – although the sample is too small to deduce this with any certainty. Nevertheless, it is interesting to observe that the responses of a respondent with significant additional experience in simulation and modelling may have had some influence in shifting the 'expert' consensus towards the model, Figure 10.9 (q.v. Figure C.6 and Figure C.8, Appendix C – Respondent Results and Analysis).

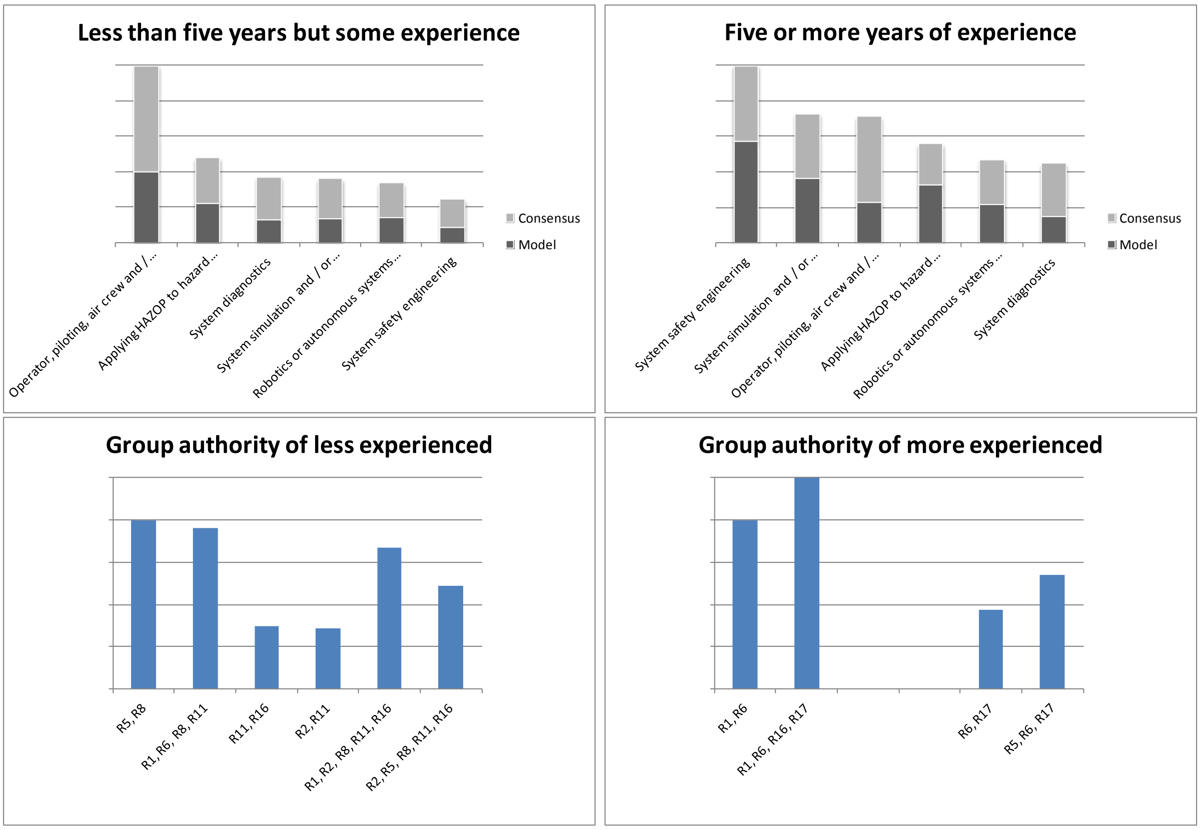


Figure 10.9 Agreement & Authority by Experience (< 5 years – left, > 5 years – right)

In the context of the responses to the questionnaire, a perspective of equivalence is revealed in the interpretation of the modelled behaviours, in that one of group of respondents appeared to have attempted to apply a purely logical process to the behaviour whilst others appear to rationalise their own beliefs or experience in selecting behaviours according to their expectations – which itself might be considered as a perspective of 'human equivalence'. Within the first group the logical process extends, at least in the case of one respondent, in attempting to model the rules described so as to almost mentally simulate the outcomes in their mind's eye as objectively as possible; i.e. not conditioned by other expectations. Of course this processing is what the model itself exists for – however if the respondent's own 'gut-feeling' flies against the objectivity as represented by the modelled behaviour then a process is required to bring about a better understanding of this likely behaviour through a closed-loop iteration of the safety constraint design and continual refinement of the modelled behaviour.

From the responses to the questionnaire, the more expert cohort appears to share a greater consensus of understanding and interpretation of behaviour then an otherwise similarly experienced but more broadly versed technical cohort. This might be revealed immediately in a degree of identification with a statement regarding the general situational awareness related to the hazard in question. However, even experts are likely to find particular forms of emergent behaviour incredible, and so would require adequate access to the model to determine why such behaviour might occur – noting that the models might be necessarily simplified in such a manner so as to be tractable whilst also attempting to identify likely worst-case behaviours. The consideration of worst-case behaviours better suits the philosophy of the scientific method in attempting to invalidate a particular hypothesis, in this case a system safety hypothesis, rather than the more typical developmental approach of attempting to validate a specific intended design requirement. Many engineers appear to downgrade the outcome of a central tenet of system complexity, in that failures of ordered behaviour often occur where systems are driven too hard towards and over the "edge of chaos" [62, 63]; preferring rather to believe that accidents are simply due to single point failures and proximate causes. More effort appears to be required so as to persuade creators and operators of systems the truth of this.

Perhaps at the core of this approach, considering behavioural modelling and the simulation of hazardous deviations, there is a fundamental challenge to the concept of demonstrating (pilot) 'equivalence' with behavioural models, or indeed any other hazard assessment methodology applied to autonomous systems exhibiting complex behaviour. Alexander, et al [1], argues that an emphasis on achieving human equivalence in the case of the external perception of the operation and behaviour of a UAV is perhaps misplaced when considering the actual goal of achieving optimum safety.

10.3.2. Monte Carlo Modelling

Through the belief validation study various graphical representations of system behaviour were introduced, representing Monte Carlo behaviour and probabilistic data, beyond trajectories and one-dimensional time series data. The first of these presents the varied and rich behaviour captured succinctly by the expression of the Lyapunov exponent as probability distributions over time and as a snap shot (Figure 8.20, Figure 10.6). The second representation presents the sensitivity of system violation events to a combination of defects by way of the violation likelihood, either as a stacked plot (Figure 7.7, Figure 9.9), or as a pie-chart (as shown within Figure 6.1, and Figure 10.10, below).

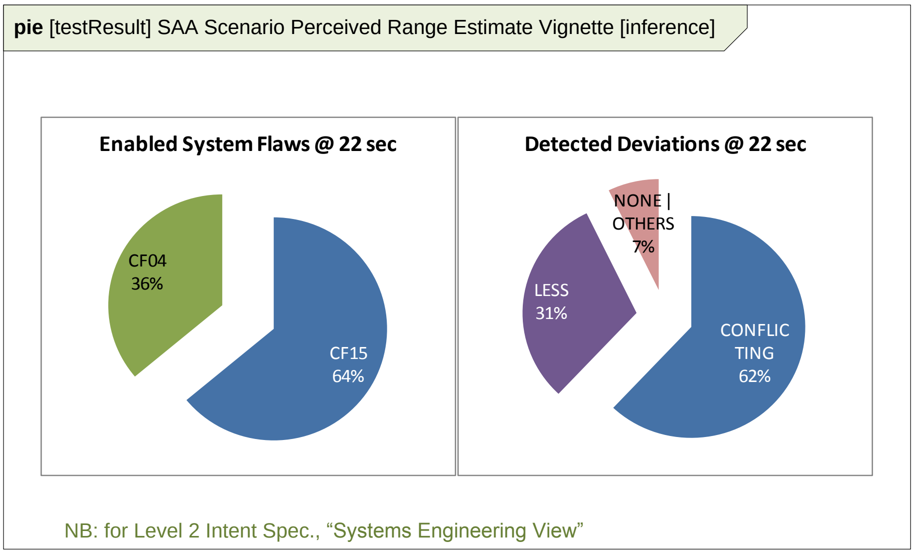


Figure 10.10 Observed Control Flaw Relative Likelihoods for Test Case 6

At present it is uncertain as to how the richness of these representations might be easily presented to and then interpreted by a HAZOP team, especially in the case of any greater dimensionality as may arise with a more complicated scenario. However, maximum peak values and breadth of distributions may be relatively easily obtained from these forms, thereby providing representative single value items that can then be incorporated directly into a HAZOP table – resulting in the more readily assimilated "traffic-light" scheme as employed in the final enhanced HAZOP representation, Table 8.2. These schemes should provide additional guidance and focus to particular areas of concern regarding both emergent behaviour and sensitivity to defective design, all to the benefit of the system engineer and their efforts to produce safer systems. Furthermore, in order to generate these categorisations of events it is necessary to define these in terms of constraint violation rates, and not only in terms of accident rates. However, before being incorporated into a more general HAZOP tool, further investigation with potential users is required, whilst consideration should also be given to how representations containing many more parameters, deviations and flaws might be best represented.

Finally, a more systematic approach should be considered for the selection of parameters to be exercised (randomised) in the Monte Carlo simulations and sensitivity analysis, especially in the case of a more complex, hierarchical and higher dimensionality system; otherwise the modelling approach and analysis cannot be scaled to larger problems. Design of experiments, experimental design, or the related Parameter Design methods developed by Taguchi [7] are all methods that might be considered.

10.4. Industry use, benefits, and exploitation

Throughout this work, every effort has been made to consider and accommodate the Guidelines for Development of Civil Aircraft and Systems [101], the Guidance for Unmanned Aircraft System Operations in UK Airspace [17], the Guidance material for UAS system safety requirements [16], and the Sense and Avoid hazard assessment work by the NATO Industrial Advisory Group [81].

Since 2005, the UK aerospace industry's ASTRAEA programme purpose has been to identify the requirements for technologies so as to enable the integration of UAVs in routine operation within un-segregated airspace. At inception this programme consisted of a number of technology themes, including themes addressing investigations and demonstrations of technologies to provide a transparent approach to "Sense and Avoidance", and a "pilot equivalent" approach and framework for UAV systems health management [29], amongst others. Within the systems health management theme a need for an increased focus on the development of software requirements for safety was identified, believed to be necessary to help prevent modes of failure due to new airborne technologies, such as machine-based decision making systems. Within this programme, West, et al, [113] proposed that the uncertainty in autonomous decision making be handled as a risk where the outcomes are mitigated by actions to be determined through hazard analysis. It is also recognised that new accident models based on systems theory (for example STAMP) suggest that constraints on systems behaviour act as barriers to the occurrence of accidents. West also notes that this is especially true of software intensive systems. It is apparent that industrial expertise in this area is perhaps ready to incorporate new accident theories and assessment methods; especially where the industry is now considering the incorporation of models of behaviour normally attributable to humans to be embodied as machine-based decision making systems. This work continues.

10.5. Future work – Frameworks for Behavioural Hazard Assessment

Further work yet needs to be done so as to determine the most usable documented form for this style of hazard assessment, along with efforts to define the requisite inter-tool and model capture interfaces of the causal modelling and Bayesian tool sets as used to produce these behavioural results. More fundamentally perhaps, a significant issue yet to be addressed is in the scalability and modularity of the proposed modelling approach, in terms of both hierarchical decomposition and propagation models in an approach encompassing more than a single isolated scenario with a small number of associated dysfunction vignettes. The approach taken in the creation of the model employed here has been described in earlier work, Downes, et al [31-33], and it ought to be noted that the modelling does not automatically identify paths between consequences and faults, as is the case for HAZID for example, McCoy, et al [73]. However, unlike systems and industrial plants formed within relatively confined architectures, such as those described with Pipe & Instrumentation Diagrams or with circuit schematics, it is not obvious as to how automatic hazard identification might be achieved through similar propagation methods of qualitative deviation and effect as addressed by HAZID. Such methods adequately address static and quasi-static behavioural models, but further consideration has to be given to the encapsulation and propagation of behaviour of systems that are quintessentially dynamic and embody evolutionary potential. Certainly process plant models are treated as closed systems, operating at particular set-points, with deviations applied to specific parameters around these set-points, but systems embodying goal seeking behaviours cannot be safely treated as closed systems.

Rather the method of hazard identification employed here adapts the STPA process [86, 104, 105], as described by Owens, Stringfellow, et al, and then to apply behavioural sensitivity simulation and Bayesian belief capture so as to explore the likely dynamic constraint violation behaviours, and in essence validate the hazards identified in the employment of this adapted form of STPA.

The purpose at this preliminary stage of the hazard assessment is to identify the adequacy of the top-level safety constraints and protective barriers for the autonomous system, whilst initiating the Functional Hazard Assessment (FHA) processes by providing a view of the likely uncertainty inherent in the system with regard to control flaws, false alarm and misdiagnosis rates, incorporating also the likely allowable tolerance upon machine perception and latency characteristics. The formulation and initial description of the safety constraints and protective barriers are likely to be of an essentially qualitative nature as derived through either a HAZOP [66] or STPA [71] process. The Functional

Hazard Assessment enables a transition from a qualitative to a quantitative safety assessment processes, through the application of FTA, DD, FMES methods [100] at the far end of the process. Otherwise, in the case of FHA, behavioural analysis is described only in the context of Markov Analysis (finite state modelling), and sensitivity analysis appears to be confined to Fault Tree Analysis, and no apparent consideration is given to the possibility that complex system dynamic behaviour is an intrinsic aspect in the design of an autonomous, machine-based decision making, system. Therefore applying sensitivity analysis to the conditions, states, and safety constraints (or constraints more generally), and specifically then measuring the stability and convergence of the key control parameter states, provides the necessary, and otherwise missing insight into the likelihood of emergent behaviour – i.e. behaviour operating "at the edge of chaos" [62, 63] and with the saliency of a persistent similarity of patterns arising with the interaction of simple underlying rules [54].

It is expected by the regulatory authorities that any decision-making by an Autonomous Systems should be repeatable, except where a machine-made decision might cause the trajectory of a system to pass beyond a go / no-go decision threshold – *"The decisions made by an autonomous system are made on a rational basis. In addition, to ensure consistent behaviour that will encourage human trust the system's decision-making should be repeatable. That is, the system should exhibit the same behaviour each time it is exposed to identical circumstances and it should not produce large changes in behaviour for small changes in inputs. An obvious exception to this is where the input to the system results in a 'yes/no' decision, such as a point of no return (e.g. deciding to return to the departure airfield instead of continuing to the destination due to a very small difference in the amount of fuel remaining). Such behaviours can be evaluated using sensitivity analysis, relating system inputs to output."* CAP 722, Section 2, Chp 7, Para 3.4.1. [17] And so the issue of developing a hazard assessment process that systematically takes account of uncertainty in machine-based situational awareness and system sensitivity to initial conditions remains.

Chapter 11. Recommendations and Conclusions

"But it's turtles all the way down!"

an old lady, anecdotal Anonymous – possibly by William James or Bertrand Russell
referenced by Stephen Hawking, "A Brief History of Time"

11.1. Recommendations

So as to facilitate the incorporation of this methodology into useful industrial applications, the following tasks ought to be considered:

- Assemble a library of representative dynamical model fragments (Molecules in the parlance of Vensim, or masked blocks and s-functions in the case of Matlab / Simulink). These fragments might include model fragments representing rule structures, fuzzy logic, heuristics, motion models and transforms, Lyapunov exponent and entropy calculations, and constraint representations, amongst others.
- Construct a family of models of constraints encompassing a) physical behaviours and limits, b) models of dysfunction and defect, c) proximity and safety violation, d) system state monitoring, including FTLE, and e) Bayesian Monte Carlo model capture; all ought to be identified.
- Investigate the development of mechanisms facilitating automated construction of models from a single representation, possibly with a SysML representation as the reference form. In the first instance a means to automate the production of the initial un-trained Bayesian network consistent with the initial HAZOP formulation would be most helpful. Next useful would be the connection of the deviations, design decisions and flaws identified in the STPA process to the corresponding sub-models within the System Dynamics model, as well as identifying and managing the model elements represented by nodes within the Bayesian network. Together this would close a loop around the whole model, at least in a provisional manner to facilitate more rapid prototyping.
- Define data interfaces managing the conversion, reformatting and archiving of large quantities of data between the different software tools, also as required to facilitate rapid prototyping, especially with consideration to Monte Carlo simulation.
- Develop improved visualisation mechanisms.
- Embed questionnaire based validation, or "reality checking" of observed dynamical behaviour, perhaps with guidelines on "cherry picking" comparable general, typical and extreme cases.
- Investigate the creation and incorporation of acceptable human and socio-technical interaction models, perhaps linked with questionnaire based validation and knowledge capture techniques.

11.2. Conclusions

In the introduction to this thesis a hypothesis was put that a conventional, and essentially manual, HAZOP process can be improved with information obtained with model-based dynamic simulation, using a Monte Carlo approach, to update a Bayesian Belief model representing the expected relations between cause and effects – and through this approach thereby produce an enhanced HAZOP. Through this it is claimed that by informing and improving the HAZOP process in this manner, the contribution of this thesis is to make unanticipated hazardous conditions more predictable.

However, it is necessary to make clear that an approach towards hazard assessment based upon simulation cannot exhaustively determine all potentially unanticipated hazards even within a simple system as the method only represents a selective exploration of the infinite state space of an open system. Clearly it is impossible to address combinatorial complexity with simulation alone, where for example even a simple system with 1000 failure modes will have 500,000 combinations of any two of those failure modes. Rather what this study attempts to reveal is that rare but plausible behaviours may be discovered by a method of simulation and inference, potentially providing an additional tool to further explore rare near-miss cases which may serve as counter-examples to a given set of safety assumptions – and hence falsify a particular safety hypotheses in support of a HAZOP study.

Therefore, as the scope of the problem space is too large for induction and proof, the goal in the capture of system dynamics for the hazard assessment ought to be in the falsification, rather than the proof, of safe behaviour. As has been shown, test cases TC1, TC2 and TC6 stand as counter examples that may be used to falsify certain assumptions made about the safety of this particular dynamical system; where that assumption is based solely upon a straight forward application of the Rules of the Air. In particular, test cases TC1 and TC2 both exhibit a salient persistent similarity of patterns arising with the interaction of a few simple underlying rules, and the introduction of a simple perceptual flaw. Whilst similar, these two cases feature notably different underlying dynamics, in so far as their respective models incorporate different significantly damping properties (i.e. one essentially has no damping). It may therefore be reasonably argued that this similarity is due to the dynamic interaction of the underlying rules, rather than the more basic physical properties. Furthermore, all three test cases also exhibit a range of consistently positive Lyapunov exponent values even many seconds after the initial encounter has occurred – which in this model indicates that one or both vehicles are still manoeuvring, and appear not have yet started to move towards a new equilibrium state. In the context of this work unexpected patterns of salient persistent similarity are taken as examples of Emergent Behaviour, whilst positive Lyapunov exponent values are taken as indicators of Behavioural Complexity – with the two being directly correlated.

There is also some evidence of (less direct) correlation between Air-proximity safety constraint violation category values and corresponding Lyapunov exponent values, as revealed in the regression analysis of the scatter plots arising with these two measures taken from the Monte Carlo simulation data sets. This correlation perhaps only exists where the more significant behavioural complexity is exhibited – therefore it is argued that whilst the Lyapunov exponent may be used as a proxy for the detection of safety constraint violation in detecting behavioural complexity and potentially emergent (i.e. non-explicit but similar salient) behaviour, it would perhaps be best used along with an explicit measure of safety constraint violation.

Consequently, it is proposed that the probabilities and likelihoods produced by the Bayesian Belief and Bayesian Monte Carlo models may stand as reasonable estimates with the purpose of guiding the system engineer and designer towards safer design decisions. In the modification to the HAZOP method, the inferred likelihood and maximum Lyapunov rank values are returned to an enhanced HAZOP to be used to code a "Relative Likelihood" traffic-light indicator and the calculation of a deviation-based combined hazard risk value. Where specific dysfunction, defect and flaw models are considered these might need only represent a model of worst case behaviour, as the objective is to produce plausible counter-examples, which of itself facilitates simplification of the whole model.

Currently Leveson and the research group at MIT [65, 67-71, 86, 104, 105], with others elsewhere, are developing a Systems Approach to Safety that incorporates the concepts of dynamical systems, complexity and un-intended consequences. These works reveal mathematically, conceptually and philosophically challenging ideas, perhaps particularly to the conservative nature of the wider engineering profession – and maybe more so in the case of system safety engineering. However, the very nature of machine intelligence, autonomous decision making and inherent uncertainty within the operating environment appears to provide sufficient justification to challenge the status quo. Consequently, it would be reasonable to expect that further research is required in these directions, and a part of this ought to be to explore the challenge to systematise and assure these processes, and where possible also to simplify or otherwise embody these techniques so that they might find adoption by the wider professional engineering community.

In the meantime it might be expected that the lessons learnt within this work, applied as it is here to collision avoidance, may also be generalised in due course to any behavioural system wherein perceptual errors, decision thresholds, and latencies potentially give rise to unanticipated behaviour, and so provide opportunities to identify and mitigate any dangerous behaviour through a formalised iterative HAZOP – System Dynamics modelling approach, incorporating Bayesian inference.

Annex

References

[1] R. D. Alexander, M. Hall-May, and T. P. Kelly, "Certification of Autonomous Systems under UK Military Safety Standards," in *25th International System Safety Conference. ISSC 07, Proceedings of*, , (2007).

[2] K. Allenby and T. Kelly, "Deriving safety requirements using scenarios," in *Requirements Engineering, 2001. Proceedings. Fifth IEEE International Symposium on*, , (2001), pp. 228-235.

[3] R. C. Arkin, *Behavior-Based Robotics*: Bradford Books, The MIT Press, (1998).

[4] G. L. Baker and J. P. Gollub, *Chaotic Dynamics: an introduction*. Cambridge: Cambridge University Press, (1990).

[5] D. P. Barnes, C. G. Downes, and J. O. Gray, "A Parallel Distributed Control Architecture for a Hexapodal Robot," *Parallel and Distributed Computing in Engineering Systems,* pp. 341 - 346, (1992).

[6] M. Bayes and M. Price, "An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S," *Philosophical Transactions,* vol. 53, pp. 370-418, January 1, 1763 (1763).

[7] B. S. Blanchard and W. J. Fabrycky, *Systems Engineering and Analysis*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, Inc., (1998).

[8] B. W. Boehm, "A spiral model of software development and enhancement," *Computer,* vol. 21, pp. 61-72, (1988).

[9] R. A. Brooks, "A robot that walks; emergent behaviors from a carefully evolved network," in *Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on*, (1989), pp. 692-4+2 vol.2.

[10] R. A. Brooks, "Intelligence without Representation," *Artificial Intelligence,* vol. 47, pp. 139-159, January 1991 (1991).

[11] R. A. Brooks, "New Approaches to Robotics," *Science,* vol. 253, pp. 1227-1232, 13 September 1991 (1991).

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

[13] BSI, *Hazard and Operability Studies (HAZOP studies) - Application Guide* vol. BS: IEC61882:2002: British Standards Institute, (2002).

[14] CAA, *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).

[15] CAA, *CAP 393 - Air Navigation: The Order and the Regulations*. London, UK: TSO (The Stationery Office) on behalf of the UK Civil Aviation Authority, Safety Regulation Group, (2010).

[16] CAA, *Draft AMC UAS.1309 - Guidance material for UAS system safety requirements*, Amendment 9 ed. London, UK: TSO (The Stationery Office) on behalf of the UK Civil Aviation Authority, Safety Regulation Group, (2010).

[17] CAA, *CAP 722 - Unmanned Aircraft System Operations in UK Airspace – Guidance*, 5th ed. London, UK: TSO (The Stationery Office) on behalf of the UK Civil Aviation Authority, Flight Operations Policy, Safety Regulation Group, (2012).

[18] M. Cantor, S. Friedenthal, C. Kobryn, and B. Purves, "Extending UML from software to systems engineering," in *Engineering of Computer-Based Systems, 2003. Proceedings. 10th IEEE International Conference and Workshop on the*, , (2003), pp. 271-273.

[19] B. T. Clough, "Metrics, Schmetrics! How The Heck Do You Determine A UAV's Autonomy Anyway?," in *2002 PerMis Workshop, Proceedings of*, NIST, Gaithersburg, MD., (2002), pp. 1-7.

[20] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, pp. 37-46, (1960).

[21] W. J. Dawsey, B. S. Minsker, and V. L. VanBlaricum, "Bayesian Belief Networks to Integrate Monitoring Evidence of Water Distribution System Contamination," *Journal of Water Resources Planning and Management. Vol. 132*, vol. 132, pp. 234-241, (2006).

[22] G. Despotou and T. Kelly, "Extending Safety Deviation Analysis Techniques to Elicit Flexible Dependability Requirements," in *System Safety, 2006. The 1st Institution of Engineering and Technology International Conference on*, , (2006), pp. 29-38.

[23] G. Despotou, R. Alexander, and T. Kelly, "Addressing challenges of hazard analysis in systems of systems," in *Systems Conference, 2009 3rd Annual IEEE*, , (2009), pp. 167-172.

[24] D. W. Dilks, R. P. Canale, and P. G. Meier, "Development of Bayesian Monte Carlo techniques for water quality model uncertainty," *Ecological Modelling*, vol. 62, pp. 149-162, (1992).

[25] DOD, *Unmanned Aircraft Systems Roadmap 2005 - 2030*. Washington D.C., USA: Office of the Secretary of Defence, Department of Defense, (2005).

[26] DOD, *DoDAF Architecture Framework Version 2.0*. Washington D.C., USA: Office of the Secretary of Defence, Department of Defense, (2010).

[27] C. G. Downes, D. P. Barnes, and J. O. Gray, "First Steps for a Real Time Control Demonstrator; showing walking behaviour & distributed control," *Parallel Computing and Transputer Applications*, vol. 2, pp. 1405 - 1414, (1992).

[28] C. G. Downes and J. O. Gray, "Towards Behavioural Control; taking steps using transputers," *IEE Colloquium Digest: Applications of Parallel and Distributed Processing in Automation and Control,* vol. 1992/204, pp. 9/1 - 9/8, (1992).

[29] C. G. Downes, "ASTRAEA T7; An architectural outline for system health management on civil UAVs," in *Autonomous Systems, 2007 Institution of Engineering and Technology Conference on*, IEE, London, UK. , (2007), pp. 1-4.

[30] 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*, Loughborough, UK. , (2010), pp. 1-7.

[31] 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*, Saarbrucken, Saarland, Germany. , (2011), pp. 115-120.

[32] C. G. Downes and P. W. H. Chung, "Hazards in advising autonomy: Incorporating hazard modelling with system dynamics into the aerospace safety assessment process for UAS," in *System Safety, 2011, 6th IET International Conference on*, Birmingham, UK. , (2011), pp. 1-6.

[33] C. G. Downes and P. W. H. Chung, "Hazards in advising autonomy: Inferring hazard causes in UAS dynamics," in *Reliability and Maintainability Symposium (RAMS), 2012 Proceedings - Annual*, Reno, NV, USA. , (2012), pp. 1-6.

[34] A. S. Dreier, *Strategy, Planning & Litigating to Win: Orchestrating Trial Outcomes with Systems Theory, Psychology, Military Science and Utility Theory*: Conatus Press, (2012).

[35] N. Dulac, N. Leveson, D. Zipkin, S. Friedenthal, J. Cutcher-Gershenfeld, J. Carroll, and B. Barrett, "Using system dynamics for safety and risk management in complex engineering systems," in *Simulation Conference, 2005 Proceedings of the Winter* M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, Eds., (2005), pp. 1311-1320.

[36] EASA, *CS-25 - Certification Specifications for Large Aeroplanes*. Brussels: European Aviation Safety Agency, (2003).

[37] EASA, *CS-23 - Certification Specifications for Normal, Utility, Aerobatic, and Commuter Category Aeroplanes*. Brussels: European Aviation Safety Agency, (2003).

[38] K. A. Eastaughffe, A. Cant, and M. A. Ozols, "A framework for assessing standards for safety critical computer-based systems," in *Software Engineering Standards, 1999. Proceedings. Fourth IEEE International Symposium and Forum on*, (1999), pp. 33-44.

[39] M. R. Endsley, "Situation awareness global assessment technique (SAGAT)," in *Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National*, (1988), pp. 789-795 vol.3.

[40] M. R. Endsley, "Supporting situation awareness in aviation systems," in *Systems, Man, and Cybernetics, 1997. 'Computational Cybernetics and Simulation'., 1997 IEEE International Conference on*, (1997), pp. 4177-4181 vol.5.

[41] M. R. Endsley and E. S. Connors, "Situation awareness: State of the art," in *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, (2008), pp. 1-4.

[42] EUROCONTROL, "EAM 2 / GUI 5 - Harmonisation of Safety Occurrence Severity and Risk Assessment," European Organisation for the Safety of Air Navigation, EUROCONTROL EAM 2 / GUI 5, (2005).

[43] J. R. Evans and T. P. Kelly, "Defence Standard 00-56 Issue 4 and Civil Standards - Appropriateness and Sufficiency of Evidence," in *System Safety, 2008 3rd IET International Conference on*, , (2008), pp. 1-6.

[44] FAA, *AC 25-1309-1A - Advisory Circular: System Design and Analysis*. Washington D.C., USA: US Department of Transport, Federal Aviation Administration, (1988).

[45] FAA, *AC 23-1309-1D - Advisory Circular: System Safety Analysis and Assessment for Part 23 Airplanes*. Washington D.C., USA: US Department of Transport, Federal Aviation Administration, (2009).

[46] J. W. Forrester, *Industrial Dynamics*. Cambridge, MA: The MIT Press, (1961).

[47] S. A. Friedenthal and H. Lykins, "Parameter-based representation for modeling complex systems. 2," in *Engineering of Computer-Based Systems,1996. Proceedings., IEEE Symposium and Workshop on*, , (1996), pp. 65-71.

[48] S. Glavaski, M. Elgersma, M. Dorneich, and P. Lommel, "Failure accommodating aircraft control," in *American Control Conference, 2002. Proceedings of the 2002*, (2002), pp. 3624-3630 vol.5.

[49] J. Gleick, *Chaos: Making a New Science*. London, UK: Cardinal by Sphere Books, (1988).

[50] M. A. Greenfield, "Normal Accident Theory: The Changing Face of NASA and Aerospace," Hagerstown, Maryland: NASA, (1998).

[51] 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.

[52] D. R. Haddon, "Aircraft airworthiness certification standards for civil UAVs," *AERONAUTICAL JOURNAL*, vol. 107, pp. 79-86, (2003).

[53] D. R. Haddon and C. J. Whittaker, "UK-CAA policy for light UAV Systems," *Navigation France*, vol. 54, pp. 73-84, (2006).

[54] J. H. Holland, *Emergence: from chaos to order*: Oxford University Press, (1998).

[55] J. Holt and S. Perry, *SysML for Systems Engineers*. London: The Institution of Engineering and Technology, (2008).

[56] 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.

[57] P. B. Hugge and J. D. Lang, "Results of implementing a disciplined avionic development process: advanced design for quality avionic systems (ADQAS)," in *Aerospace and Electronics Conference, 1995. NAECON 1995., Proceedings of the IEEE 1995 National*, (1995), pp. 220-226 vol.1.

[58] M. S. Jaffe, R. Busser, D. Daniels, H. Delseny, and G. Romanski, "Progress Report on Some Proposed Upgrades to the Conceptual Underpinnings of DO-178B/ED-12B," in *System Safety, 2008 3rd IET International Conference on*, (2008), pp. 1-6.

[59] C. W. Johnson, "Insights from the Nogales Predator Crash for the Integration of UAVs into the National Airspace System under FAA Interim Operational Guidance 08-01," in *Proceedings of the 27th International Conference on Systems Safety*, J. M. Livingston, R. Barnes, D. Swallom, and W. Pottraz, Eds. Huntsville, Alabama, USA: International Systems Safety Society, (2009), pp. 3066-3076.

[60] T. L. Johnson, R. Koneck, and S. F. Bush, "Improving UAV mission success rate through software enabled control design," in *Aerospace Conference Proceedings, 2000 IEEE*, (2000), pp. 373-378.

[61] T. Kapitaniak, *Chaos for Engineers*, Second, Revised ed. Berlin: Springer-Verlag, (2000).

[62] S. A. Kauffman, "Antichaos and Adaptation," *Scientific American*, vol. 265, pp. 78 - 84, August 1991 (1991).

[63] S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*. New York, USA: Oxford University Press, (1993).

[64] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, vol. 33, pp. 159-174, (1977).

[65] J. R. Laracy and N. G. Leveson, "Apply STAMP to Critical Infrastructure Protection," in *Technologies for Homeland Security, 2007 IEEE Conference on*, , (2007), pp. 215-220.

[66] H. G. Lawley, "Operability studies and hazard analysis," *Chemical Engineering Progress, AIChE*, vol. 70, p. 45, (1974).

[67] N. Leveson, "A New Accident Model for Engineering Safer Systems," *Safety Science*, vol. 42, pp. 237-270, April 2004 (2004).

[68] N. Leveson, N. Dulac, K. Marais, and J. Carroll, "Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems," *Organization Studies*, vol. 30, pp. 227-249, February/March 2009 (2009).

[69] N. G. Leveson, "Intent specifications: an approach to building human-centered specifications," *Software Engineering, IEEE Transactions on*, vol. 26, pp. 15-35, (2000).

[70] 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).

[71] N. G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. Cambridge, MA: The MIT Press, (2011).

[72] 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).

[73] 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).

[74] L. McCue, C. Bassler, and W. Belknap, "Real-time Identification of Behavior Leading to Capsize," in *STAB2006, 9th International Conference on Stability of Ships and Ocean Vehicles,* (2006).

[75] B. Meenakshi, K. Das Barman, K. G. Babu, and K. Sehgal, "Formal safety analysis of mode transitions in aircraft flight control system," in *Digital Avionics Systems Conference, 2007. DASC '07. IEEE/AIAA 26th,* (2007), pp. 2.C.1-1-2.C.1-11.

[76] MOD, *Def Stan 00-56: Safety Management Requirements for Defence Systems, Part 2: Guidance on Establishing a Means of Complying with Part 1*. Glasgow, UK: Ministry of Defence, UK Defence Standardization, (2007).

[77] MOD, *Def Stan 00-56: Safety Management Requirements for Defence Systems, Part 1: Requirements*. Glasgow, UK: Ministry of Defence, UK Defence Standardization, (2007).

[78] MOD, *Unmanned Aircraft Systems: Terminology, Definitions and Classification* vol. JDN 3/10. Bicester, UK: Ministry of Defence, Development, Concepts and Doctrine Centre, (2010).

[79] Z. Mohaghegh, R. Kazemi, and A. Mosleh, "Incorporating organizational factors into Probabilistic Risk Assessment (PRA) of complex socio-technical systems: A hybrid technique formalization," *RELIABILITY ENGINEERING & SYSTEM SAFETY,* vol. 94, pp. 1000-1018, (2009).

[80] H. P. Moravec, *Mind Children: The Future of Robot and Human Intelligence*: Harvard University Press, (1990).

[81] NIAG, "SG-134 SENSE AND AVOID - Main Report," NATO Industrial Advisory Group, March 2010 (2010).

[82] NTSB, "NTSB - Accident - CHI06MA121, Predator B, Nogales AZ, 2006 (LOC HF)," 31 October 2007 (2007).

[83] OMG, "OMG Systems Modeling Language (OMG SysML™)," in *formal/2010-06-02,* Version 1.2 ed Needham, MA 02494, USA: Object Management Group, Inc., (2010).

[84] OMG, "OMG Unified Modeling LanguageTM (OMG UML), Superstructure," in *formal2011-08-06*, Version 2.4.1 ed Needham, MA 02494, USA: Object Management Group, Inc., (2011).

[85] OMG, "OMG Unified Modeling LanguageTM (OMG UML), Infrastructure," in *formal2011-08-05*, Version 2.4.1 ed Needham, MA 02494, USA: Object Management Group, Inc., (2011).

[86] 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.

[87] C. Perrow, *Normal Accidents: Living with High-Risk Technologies*. Princeton, New Jersey, USA: Princeton University Press, (1999).

[88] PMI, *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)*, 4 ed. Newtown Square, PA, USA: Project Management Institute, Inc., (2008).

[89] 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).

[90] B. E. Portwood, "Current government and industry developments in the area of system safety assessment," in *Digital Avionics Systems Conference, 1998. Proceedings., 17th DASC. The AIAA/IEEE/SAE*, (1998), pp. B31-1-6 vol.1.

[91] J. Reason, *Human Error*. Cambridge, UK: Cambridge University Press, (1990).

[92] J. Reason, "Beyond the organisational accident: the need for "error wisdom" on the frontline," *Quality and Safety in Health Care*, vol. 13, pp. ii28-ii33, December 2004 (2004).

[93] L. K. Rierson, "Object-oriented technology (OOT) in civil aviation projects: certification concerns," in *Digital Avionics Systems Conference, 1999. Proceedings. 18th*, (1999), pp. 2.C.4-1-2.C.4-8 vol.1.

[94] N. Roberts, D. Andersen, R. M. Deal, M. S. Garet, and W. A. Shaffer, *Introduction to Computer Simulation: A System Dynamics Modelling Approach*. Reading, MA: Addison-Wesley Publishing Company, Inc., (1983).

[95] RTCA, *DO-178B, Software Considerations in Airborne Systems and Equipment Certification*. Washington D.C., USA: RTCA Inc., (1992).

[96] RTCA, *DO-254, Design Assurance Guidance for Airborne Electronic Hardware*. Washington D.C., USA: RTCA Inc., (2000).

[97] RTCA, *DO-178C, Software Considerations in Airborne Systems and Equipment Certification*. Washington D.C., USA: RTCA Inc., Prepared by SC-205, (2011).

[98] H. A. Ruff, "Exploring automation issues in supervisory control of multiple UAVs," *HUMAN PERFORMANCE, SITUATION AWARENESS AND AUTOMATION: CURRENT RESEARCH AND TRENDS, VOL 2,* pp. 218-222, (2004).

[99] SAE, *ARP 4754 - Certification Considerations for Highly-Integrated or Complex Aircraft Systems*. Warrendale, Pennsylvania, USA: Society of Automotive Engineers, SAE International, (1996).

[100] SAE, *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).

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

[102] M. S. Saglimbene, "Reliability analysis techniques: How they relate to aircraft certification," in *Reliability and Maintainability Symposium, 2009. RAMS 2009. Annual*, (2009), pp. 218-222.

[103] A. Soyler and S. Sala-Diakanda, "A model-based systems engineering approach to capturing disaster management systems," in *Systems Conference, 2010 4th Annual IEEE*, , (2010), pp. 283-287.

[104] M. V. Stringfellow, "ACCIDENT ANALYSIS AND HAZARD ANALYSIS FOR HUMAN AND ORGANIZATIONAL FACTORS," in *Department of Aeronautics and Astronautics*. vol. Doctor of Philosophy Boston: Massachusetts Institute of Technology, (2010), p. 283.

[105] 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).

[106] R. M. Taylor, S. Abdi, R. Dru-Drury, and M. C. Bonner, "Cognitive Cockpit Systems: Information Requirements Analysis for Pilot Control of Automation," *Engineering Psychology and Cognitive Ergonomics*, vol. 5, Aerospace and Transportation Systems, pp. 81-88, (2001).

[107] R. Temam, *Infinite Dimensional Dynamical Systems in Mechanics and Physics* vol. 68: Springer, (1997).

[108] A. C. Tribble and S. P. Miller, "Software intensive systems safety analysis," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 19, pp. 21-26, (2004).

[109] E. J. Trimble and D. A. Cooper, "Report on the accident to Boeing 737-400, G-OBME, near Kegworth, Leicestershire on 8 January 1989," AAIB, Aldershot, Hampshire, UK. 4/1990, 25 August (1990).

[110] P. Trucco, "A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation," *RELIABILITY ENGINEERING & SYSTEM SAFETY*, vol. 93, pp. 845-856, (2008).

[111] UKAB, "Twenty Second Report by the UK Airprox Board (January 2009 to June 2009)," *Analysis of Airprox in UK Airspace*, vol. 22, (2009).

[112] G. Warwick and G. Norris, "Designs for Success - systems engineering must be rethought if program performance is to improve," in *Aviation Week & Space Technology*. vol. 172: McGraw-Hill Co., (2010), pp. 72-75.

[113] J. West, S. Conlan, and P. Smith, "Autonomous Health Management Equivalence in the Civil Regulatory Environment - Final Report," BAE SYSTEMS (Operations) Ltd., Prestwick BAE/P/SOP1439/RP/00003, November 2008 (2008).

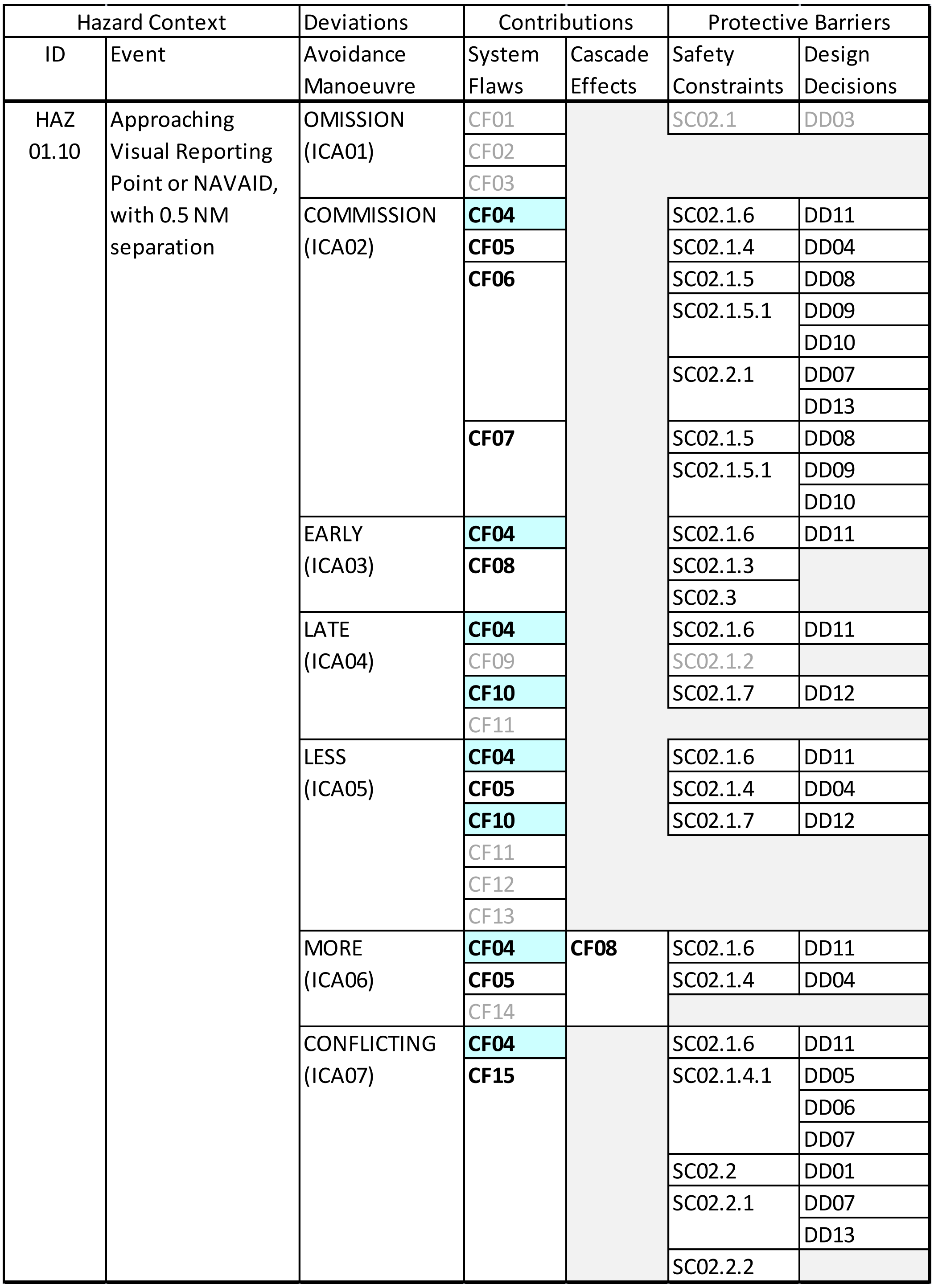
[114] P. J. Wilkinson and T. P. Kelly, "Functional hazard analysis for highly integrated aerospace systems," in *Certification of Ground/Air Systems Seminar (Ref. No. 1998/255), IEE*, , (1998), pp. 4/1-4/6.

[115] F. Ye and T. Kelly, "Component failure mitigation according to failure type," in *Computer Software and Applications Conference, 2004. COMPSAC 2004. Proceedings of the 28th Annual International*, (2004), pp. 258-264 vol.1.

[116] S. Zhu, P. Cui, and H. Hu, "Hazard detection and avoidance for planetary landing based on Lyapunov control method," in *Intelligent Control and Automation (WCICA), 10th World Congress on*, (2012), pp. 2822-2826.

Appendix A. Enhanced HAZOP Employing Inference Test Cases

A.1. Initial STPA – based HAZOP



Hazard Context		Deviations	Contributions		Protective Barriers	
ID	Event	Avoidance Manoeuvre	System Flaws	Cascade Effects	Safety Constraints	Design Decisions
HAZ 01.10	Approaching Visual Reporting Point or NAVAID, with 0.5 NM separation	OMISSION (ICA01)	CF01		SC02.1	DD03
			CF02			
			CF03			
		COMMISSION (ICA02)	**CF04**		SC02.1.6	DD11
			CF05		SC02.1.4	DD04
			CF06		SC02.1.5	DD08
					SC02.1.5.1	DD09
						DD10
					SC02.2.1	DD07
						DD13
			CF07		SC02.1.5	DD08
					SC02.1.5.1	DD09
						DD10
		EARLY (ICA03)	**CF04**		SC02.1.6	DD11
			CF08		SC02.1.3	
					SC02.3	
		LATE (ICA04)	**CF04**		SC02.1.6	DD11
			CF09		SC02.1.2	
			CF10		SC02.1.7	DD12
			CF11			
		LESS (ICA05)	**CF04**		SC02.1.6	DD11
			CF05		SC02.1.4	DD04
			CF10		SC02.1.7	DD12
			CF11			
			CF12			
			CF13			
		MORE (ICA06)	**CF04**	**CF08**	SC02.1.6	DD11
			CF05		SC02.1.4	DD04
			CF14			
		CONFLICTING (ICA07)	**CF04**		SC02.1.6	DD11
			CF15		SC02.1.4.1	DD05
						DD06
						DD07
					SC02.2	DD01
					SC02.2.1	DD07
						DD13
					SC02.2.2	

Figure A.1 Initial STPA Based HAZOP: Single Deviation Type (Manoeuvre)

System (control) flaw entries shown as **bold** text in Figure A.1 are explicitly implemented in the model, whereas the remaining flaws are not explicitly implemented. Entries highlighted with a light blue background indicate the primary flaws, Perceived Range Estimation and Latency; used in the construction of the Category A/B Air Proximity constraint violation conditional likelihood function. Items in grey in the following text are also not explicitly implemented.

A.2. Original HAZOP Definitions

Hazard

HAZ 01.10 – UAV separation minima is breached by the other aircraft (NIAG SG-134 Separation Provision Error HAZ010).

Deviations (Avoidance Manoeuvre)

OMISSION (ICA01) – The UAV does not manoeuvre to avoid the other aircraft.

COMMISSION (ICA02) – The UAV manoeuvres in the wrong direction attempting to avoid the other aircraft.

EARLY (ICA03) – The UAV prematurely ceases a manoeuvre to avoid the other aircraft and returns to the given heading.

LATE (ICA04) – The UAV belatedly initiates a manoeuvre to avoid the other aircraft.

LESS (ICA05) – The UAV rate of manoeuvre is insufficient to avoid the other aircraft.

MORE (ICA06) – The UAV rate of manoeuvre is excessive in attempting to avoid the other aircraft.

CONFLICTING (ICA07) – The UAV is unable to resolve the conflict between realising the System Goal and avoiding the other aircraft.

Possible Contributory System (Control) Flaws – grey unimplemented

CF01 – The UAV fails to detect the other aircraft altogether.

CF02 – The UAV does not command a turning manoeuvre from the flight control system (FCS).

CF03 – The UAV flight control surfaces do not respond to demands from the FCS.

CF04 – The UAV incorrectly estimates the location and direction to the predicted point of conflict.

CF05 – The UAV fails to follow established Rules of the Air.

CF06 – The other aircraft fails to follow established Rules of the Air.

CF07 – The UAV interprets the Rules of the Air too rigidly when closing within 500ft of the other aircraft.

CF08 – The UAV loses detection of the other aircraft.

CF09 – The UAV fails to detect the other aircraft at sufficient range.

CF10 – The UAV delays the commanding of a turning manoeuvre from the flight control system (FCS).

CF11 – The UAV flight control surfaces are slow to respond to demands from the FCS.

CF12 – The UAV demanded rate of turn is insufficient for the closing angle, speed and range to the other aircraft.

CF13 – The UAV maximum demanded rate of turn exceeds the maximum achievable rate of turn.

CF14 – The UAV demanded rate of turn is excessive for the closing angle, speed and range to the other aircraft.

CF15 – The UAV does not arbitrate the competing heading demands appropriately.

A.3. Final Enhanced HAZOP

Hazard Context			Deviations		Contributions			Protective Barriers						
ID	Event	Special Conditions	Avoidance Manoeuvre	Manoeuvring Airspeeds	System Flaws	Cascade Effects	Secondary Flaws	Safety Constraints	Design Decisions	Hazard	Risk	Emergent Behaviour	Diagnostic Utility	Relative Likelihood
HAZ 01.10	Approaching Visual Reporting Point or NAVAID	From Ahead with 0.5 NM separation	LESS (ICA05)	DIFFERENT	CF07 **CF04**		CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11	🔴	🟡	🟢	⚫	🔴
				HIGH	CF07 **CF04**		CF05 CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11	🔴	🔴	🟢	⚫	
			MORE (ICA06)	HIGH	CF07	**CF08**	CF05 CF15	SC02.1.5 SC02.1.5.1	DD08 DD09 DD10	🟡		🟢	🔴	
			CONFLICTING (ICA07)	DIFFERENT	**CF15**		CF07 CF06 CF08	SC02.1.4.1 SC02.2 SC02.2.1 SC02.2.2	DD05 DD06 DD07 DD01 DD13 DD07	⚫		⚫	🟢	
		From Behind Right with 0.5 NM separation	OMISSION (ICA01)	DIFFERENT	CF07		CF15 CF04 CF06	SC02.1.5 SC02.1.5.1	DD08 DD09 DD10	🔴	⚫	⚫	🔴	🟡
			LATE (ICA04)		CF07 **CF04**		CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11					
			LESS (ICA05)		CF07 **CF04**		CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11					
		From Behind Left with 0.5 NM separation	OMISSION (ICA01)	HIGH	CF07		CF05 CF15 CF04 CF06	SC02.1.5 SC02.1.5.1	DD08 DD09 DD10	🔴	🟢	🟡	🔴	🟢
			LATE (ICA04)		CF07 **CF04**		CF05 CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11					
			LESS (ICA05)		CF07 **CF04**		CF05 CF15 CF06	SC02.1.5 SC02.1.5.1 SC02.1.6	DD08 DD09 DD10 DD11					
		From Left with 0.5 NM separation	LESS (ICA05)	HIGH	CF05 **CF04** **CF10**		CF07 CF15	SC02.1.4 SC02.1.6 SC02.1.7	DD04 DD11 DD12	🔴	🟢	🟡	⚫	🟢

Figure A.2 Final Enhanced HAZOP: Cases for Approaches from Ahead, Behind and Left

In the above traffic-light scheme, green stands for a best or better measure, yellow for medium, red for significantly degraded and any black indicates the worst behaviour or outcome for the particular measure; calculated as follows:

$$H = CatA \times CPA, \qquad R = H \times (M - E) \times D \times L \tag{A.1}$$

Where *H* represents the hazard value, *CatA* the relative rate of category A airprox events, *CPA* the rate observed at the closest classified point of approach. *R* represents the risk, with *M* being the total number of test cases, *E* the rank of the emergent behaviour, *D* the rank of the diagnostic consistency measure, and with *L* representing the relative likelihood.

Hazard Context			Deviations		Contributions			Protective Barriers						
ID	Event	Special Conditions	Avoidance Manoeuvre	Manoeuvring Airspeeds	System Flaws	Cascade Effects	Secondary Flaws	Safety Constraints	Design Decisions	Hazard	Risk	Emergent Behaviour	Diagnostic Utility	Relative Likelihood
HAZ 01.10	Approaching Visual Reporting Point or NAVAID	From Right with 0.5 NM separation	COMMISSION (ICA02)	LOW	**CF04**		CF15	SC02.1.6	DD11	🟡	⚫	🔴	🟡	⚫
				DIFFERENT	**CF07**		CF15	SC02.1.5	DD08	🟡		🔴	🟢	
								SC02.1.5.1	DD09					
									DD10					
									DD15					
					CF06			SC02.1.5	DD08					
								SC02.1.5.1	DD09					
									DD10					
								SC02.2.1	DD13					
									DD07					
			EARLY (ICA03)	DIFFERENT	CF15		CF06 CF07	SC02.1.4.1	DD05	🟢		🔴	🟢	
									DD06					
									DD07					
								SC02.2	DD01					
								SC02.2.1	DD13					
									DD07					
								SC02.2.2						
					CF08			SC02.1.3	***DD14***					
								SC02.3						
			LESS (ICA05)	LOW	CF15			SC02.1.4.1	DD05	⚫		⚫	🟡	
									DD06					
									DD07					
								SC02.2	DD01					
								SC02.2.1	DD13					
									DD07					
								SC02.2.2						
					CF04			SC02.1.6	DD11					
			MORE (ICA06)	LOW	**CF04**	**CF08**	CF07	SC02.1.6	DD11	🟡		🔴	🟢	
				HIGH	CF05	**CF08**	CF07 CF15 CF10	SC02.1.4	DD04	🟡		🟡	🟡	
			CONFLICTING (ICA07)	LOW	**CF04**			SC02.1.6	DD11	⚫		⚫	🟡	
					CF15			SC02.1.4.1	DD05					
									DD06					
									DD07					
								SC02.2	DD01					
								SC02.2.1	DD13					
									DD07					
								SC02.2.2						
				DIFFERENT	**CF15**		CF06 CF07	SC02.1.4.1	DD05	⚫		🔴	🟢	
									DD06					
									DD07					
								SC02.2	DD01					
								SC02.2.1	DD13					
									DD07					
								SC02.2.2						

Figure A.3 Final Enhanced HAZOP (cont.): Cases for Approaches from the Right

Deviations		Airprox Event Likelihood				NCT Too Close					Airprox Category					Validation
Avoidance Manoeuvre	Manoeuvring Airspeeds	UAS Surprised	UAS Threatened	NCT Surprised	NCT Panicked	> 707 ft	< 707 ft	< 316 ft	< 236 ft	< 196 ft	None	CatD	CatC	CatB	CatA	Test Case
LESS (ICA05)	DIFFERENT	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	PREL - S9 test case (speed difference). Test Case 8
	HIGH	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	PREL - S3 test case (high speed). Test Case 7
MORE (ICA06)	HIGH	🟢	🟡	🟢	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	ES - S8 test case. Test Case 3
CONFLICTING (ICA07)	DIFFERENT	🟢	🟢	🟡	🔴	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	**ENCTS (with offset defect) - S9 test case, and also S161 collision.** **Test Case 1**

Figure A.4 Final Enhanced HAZOP: Consequences for Approaches from Ahead

Deviations		Airprox Event Likelihood				NCT Too Close					Airprox Category					Validation
Avoidance Manoeuvre	Manoeuvring Airspeeds	UAS Surprised	UAS Threatened	NCT Surprised	NCT Panicked	> 707 ft	< 707 ft	< 316 ft	< 236 ft	< 196 ft	None	CatD	CatC	CatB	CatA	Test Case
LESS (ICA05)	HIGH	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	PREL - S39 test case. Test Case 11

Figure A.5 Final Enhanced HAZOP: Consequences for Approaches from the Left

Deviations		Airprox Event Likelihood				NCT Too Close					Airprox Category					Validation
Avoidance Manoeuvre	Manoeuvring Airspeeds	UAS Surprised	UAS Threatened	NCT Surprised	NCT Panicked	> 707 ft	< 707 ft	< 316 ft	< 236 ft	< 196 ft	None	CatD	CatC	CatB	CatA	Test Case
OMISSION (ICA01) LATE (ICA04) LESS (ICA05)	HIGH	🟢	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	PREL - S164 test case. Test Case 9
OMISSION (ICA01) LATE (ICA04) LESS (ICA05)	DIFFERENT	🟢	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	PREL - S200 test case, and also S95 collision. Test Case 10

Figure A.6 Final Enhanced HAZOP: Consequences for Approaches from Behind (Left and Right)

Deviations		Airprox Event Likelihood				NCT Too Close					Airprox Category					Validation
Avoidance Manoeuvre	Manoeuvring Airspeeds	UAS Surprised	UAS Threatened	NCT Surprised	NCT Panicked	> 707 ft	< 707 ft	< 316 ft	< 236 ft	< 196 ft	None	CatD	CatC	CatB	CatA	Test Case
COMMISSION (ICA02)	LOW	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	PREL - S91 test case, at collision. Test Case 12
	DIFFERENT	⚫	🟢	🟢	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🟢	🟢	⚫	**ENCTS - S1 test case, exit. Test Case 2**
EARLY (ICA03)	DIFFERENT	🟢	🟢	🔴	🔴	🟢	⚫	🟢	🟢	🟢	🟢	🟢	🟢	⚫	🟢	**ENCTS - S1 test case, precursor. Test Case 2**
LESS (ICA05)	LOW	🟢	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	PRE - S34 test case. Test Case 6
MORE (ICA06)	LOW	🟢	🟡	🟢	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	MWT - S3 test case, and also S27 at collision. TC 5
	HIGH	🟢	🟡	🟢	🟢	🟢	🟡	🟡	🟢	🟡	🟢	🟢	🟢	🟡	🔴	Latency - S5 test case, and also S67 at collision. Test Case 4
CONFLICTING (ICA07)	LOW	🟢	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	PRE - S34 test case. Test Case 6
	DIFFERENT	🟢	🟢	🟡	🔴	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🟢	⚫	**ENCTS - S1 test case at entry, and S85 at collision. Test Case 2**

Figure A.7 Final Enhanced HAZOP: Consequences for Approaches from the Right

A.4. Additional HAZOP Definitions

Additional Deviations (Manoeuvring Airspeeds)

LOW – The airspeeds of both aircraft generally do not exceed the speed for which the absolute minimum permitted VFR (Visual Flight Rule) flight visibility might apply.

HIGH – The airspeed of either or both vehicles approaches or exceeds the permitted speed limit for aircraft operating below 10,000 ft.

DIFFERENT – The difference in airspeed between the two vehicles approaches or exceeds one quarter of the permitted speed limit below 10,000 ft.

New Design Decisions

DD14 – The field of regard of the onboard sensor system shall be a minimum of ± 65° with respect to the forward longitudinal axis in the plane of the turn when banked at ± 40°. (SC02.1.3)

DD15 – The direction to the perceived threat shall be offset by + 47° (to the right) to allow for the other aircraft passing from right to left ahead of the UAV. (SC02.1.5.1)

Air Proximity Definitions

Items in grey in the following text are implemented in the system dynamics model but are not represented in the corresponding Bayesian Network model. Items in **bold** text indicate the primary constraint violation indicator(s).

Category A – Risk of collision: an actual risk of collision existed.

Category B – Safety not assured: the safety of the aircraft was compromised.

Category C – No risk of collision: no risk of collision existed.

Category D – Risk not determined: insufficient information was available to determine the risk involved, or inconclusive or conflicting evidence precluded such determination.

UAS / NCT Surprised – The NCT / UAS appears not to be applying the Rules of the Air by either crossing ahead from left to right at a rate greater than the Transit Sensitivity (0.075) x 65° per second azimuth, or behind from right to left within half a second - how far from right to left is determined by the value of the Transit Sensitivity (from 140° at minimum to 220° at maximum).

UAS Threatened – The UAS is forced into behaviour other than a strict interpretation of the Rules of the Air, with a turn to port greater than Rate Two (-6° per second).

UAS / NCT Panicked – The UAS / NCT is forced to perform excessive manoeuvring, with a rate of turn greater than Rate Three (12° per second).

UAS / NCT Alarmed – The UAS / NCT manoeuvres aggressively and whilst banking excessively (> ± 40°) the NCT is out of view.

A.5. Validation Test Cases and Hazardous Outcomes

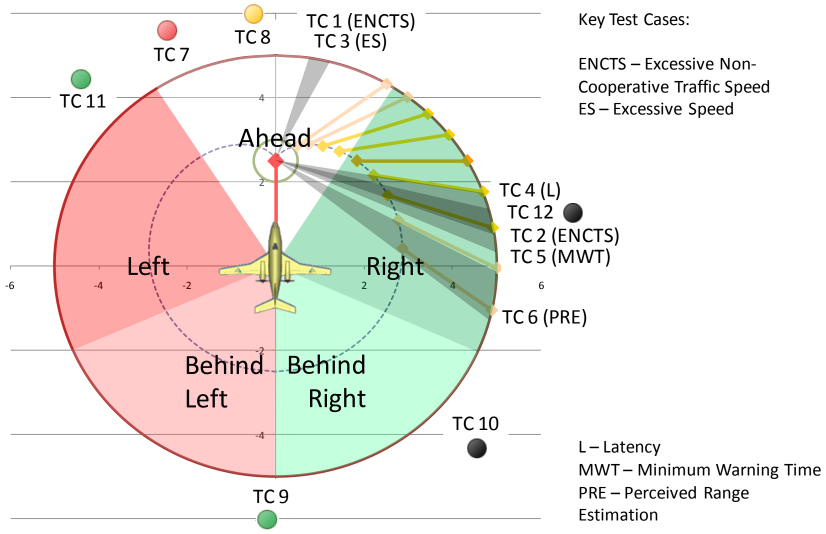


Figure A.8 Hazard Test Case Assignment by Direction of Approach

Test Case	Initial Traffic Separation		Initial NCT Heading		NCT Speed		UAV Speed		Perceived Range Estimation		Minimum Warning Time		Latency in Perceiving Threat		Threat Offset Direction	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
8	0.5 NM		122°	238°	195 kts	305 kts	71.5 kts	117 kts	70%	140%	15 s		0 s	5 s	47°	
			169.50		**210.45 kts**		**103.18 kts**		**72.13%**				**1.51 s**			
7			122°	238°	195 kts	305 kts	195 kts	305 kts	70%	140%			0 s	5 s		
			136.41°		**276.94 kts**		**204.20 kts**		**135.95%**				**0.76 s**			
3	0.4 NM	0.6 NM	197.5°	207.5°	245 kts	255 kts	245 kts	255 kts	100%		10 s	20 s	1 s		47°	
ES	**0.42 NM**		**207.49°**		**253.53 kts**		**252.96 kts**				**15.13 s**					
1	0.4 NM	0.6 NM	197.5°	207.5°	198 kts	208 kts	89.25 kts	99.25 kts			10 s	20 s			0°	
ENCTS	**0.47 NM**		**201.60°**		**198.30 kts**		**92.28 kts**				**16.96 s**					
12	0.5 NM		238°	312.5°	71.5 kts	117 kts	71.5 kts	117 kts	70%	140%	15 s		0 s	5 s	47°	
			285.46°		**114.62 kts**		**72.99 kts**		**88.41%**				**1.47 s**			
2	0.4 NM	0.6 NM	283°	293°	198 kts	208 kts	89.25 kts	99.25 kts	100%				1 s		42°	52°
ENCTS	**0.46 NM**		**287.53°**		**205.89 kts**		**96.62 kts**								**51.98°**	
6	0.4 NM	0.6 NM	296.5°	306.5°	135 kts	145 kts	89.25 kts	99.25 kts	25%	400%	5 s				47°	
PRE	**0.56 NM**		**304.41°**		**137.34 kts**		**91.24 kts**		**75.51%**							
5	0.4 NM	0.6 NM	283°	293°	135 kts	145 kts	135 kts	145 kts	100%		0 s	10 s				
MWT	**0.55 NM**		**290.45°**		**136.24 kts**		**144.42 kts**				**1.52 s**					

4	0.4 NM	0.6 NM	278.5°	288.5°	245 kts	255 kts	245 kts	255 kts			15 s	0 s	5 s	
L	**0.44 NM**		**281.99°**		**253.16 kts**		**245.22 kts**					**4.99 s**		
10	0.5 NM		312.5°	360°	117 kts	195 kts	71.5 kts	117 kts	70%	140%	15 s	0 s	5 s	47°
			323.38°		**183.86 kts**		**88.84 kts**		**75.26%**			**1.86 s**		
9	0.5 NM		0°	47.5°	195 kts	305 kts	71.5 kts	117 kts	70%	140%	15 s	0 s	5 s	47°
			0.25°		**289.69 kts**		**75.34 kts**		**89.99%**			**1.40 s**		
11	0.5 NM		47.5°	122°	195 kts	305 kts	71.5 kts	117 kts	70%	140%	15 s	0 s	5 s	47°
			105.73°		**257.52 kts**		**99.87 kts**		**73.88%**			**3.98 s**		

Table A.1 Hazard Test Case Assignment by Direction of Approach, Speed, etc.

Airprox CatA	Airprox CatB	Airprox CatC	Airprox CatD	Airprox None		
0.44%	0.53%	2.05%	18.07%	78.91%	**Airspace Interaction Ahead**	
0.00%	0.19%	1.56%	19.91%	78.34%	PREL-1800-1400-1400-0nm5	
1.31%	1.28%	3.34%	16.22%	77.84%	PREL-1800-1400-2500-0nm5	
0.00%	0.00%	0.31%	17.09%	82.59%	PREL-1800-1400-9425-0nm5	
0.03%	0.69%	2.91%	18.84%	77.53%	PREL-1800-2500-1400-0nm5	
1.06%	**1.31%**	**4.09%**	**17.97%**	**75.56%**	**PREL-1800-2500-2500-0nm5**	TC7
0.00%	0.00%	0.81%	18.34%	80.84%	PREL-1800-2500-9425-0nm5	
0.00%	0.19%	1.25%	21.44%	77.13%	PREL-1800-9425-1400-0nm5	
1.53%	**1.13%**	**3.69%**	**17.84%**	**75.81%**	**PREL-1800-9425-2500-0nm5**	TC8
0.00%	0.00%	0.44%	15.00%	84.56%	PREL-1800-9425-9425-0nm5	

Airprox CatA	Airprox CatB	Airprox CatC	Airprox CatD	Airprox None		
0.01%	0.02%	0.28%	8.81%	90.89%	**Airspace Interaction Behind_L**	
0.00%	0.00%	0.00%	7.44%	92.56%	PREL-0240-1400-1400-0nm5	
0.00%	0.00%	0.47%	16.25%	83.28%	PREL-0240-1400-2500-0nm5	
0.00%	0.00%	0.00%	0.00%	100.00%	PREL-0240-1400-9425-0nm5	
0.00%	0.00%	0.00%	5.88%	94.13%	PREL-0240-2500-1400-0nm5	
0.00%	0.00%	0.00%	8.03%	91.97%	PREL-0240-2500-2500-0nm5	
0.00%	0.00%	0.00%	0.00%	100.00%	PREL-0240-2500-9425-0nm5	
0.00%	0.00%	0.00%	2.81%	97.19%	PREL-0240-9425-1400-0nm5	
0.06%	**0.19%**	**2.06%**	**38.84%**	**58.84%**	**PREL-0240-9425-2500-0nm5**	TC9
0.00%	0.00%	0.00%	0.00%	100.00%	PREL-0240-9425-9425-0nm5	

Airprox CatA	Airprox CatB	Airprox CatC	Airprox CatD	Airprox None		
0.05%	0.13%	3.13%	14.83%	81.88%	**Airspace Interaction from Left**	
0.00%	0.00%	3.25%	15.81%	80.94%	PREL-0850-1400-1400-0nm5	
0.06%	0.25%	3.25%	13.72%	82.72%	PREL-0850-1400-2500-0nm5	
0.00%	0.00%	3.38%	12.28%	84.34%	PREL-0850-1400-9425-0nm5	
0.00%	0.00%	2.41%	13.31%	84.28%	PREL-0850-2500-1400-0nm5	
0.00%	0.09%	2.16%	12.31%	85.44%	PREL-0850-2500-2500-0nm5	
0.00%	0.00%	1.94%	12.28%	85.78%	PREL-0850-2500-9425-0nm5	
0.00%	0.00%	3.72%	18.34%	77.94%	PREL-0850-9425-1400-0nm5	
0.34%	**0.78%**	**4.22%**	**21.91%**	**72.75%**	**PREL-0850-9425-2500-0nm5**	TC11
0.00%	0.00%	3.81%	13.50%	82.69%	PREL-0850-9425-9425-0nm5	

Airprox CatA	Airprox CatB	Airprox CatC	Airprox CatD	Airprox None		
8.10%	2.15%	10.21%	50.56%	28.98%	**Airspace Interaction with EM Flaw**	
18.50%	**2.94%**	**23.03%**	**48.69%**	**6.84%**	**ENCTS-2025-9425-2030-0nm5**	TC1
14.19%	1.66%	7.03%	31.75%	45.38%	ES-2025-2500-2500-0nm5	
2.13%	1.53%	7.34%	76.66%	12.34%	L-2835-2500-2500-0nm5	
5.94%	2.78%	11.84%	64.88%	14.56%	MWT-2880-1400-1400-0nm5	
0.03%	0.03%	3.66%	50.09%	46.19%	MWT-3015-1400-1400-0nm5	
7.81%	3.97%	8.34%	31.28%	48.59%	PRE-3015-9425-1400-0nm5	

Original Model Datasets as used in the "User Test" Questionnaire

Airprox CatA	Airprox CatB	Airprox CatC	Airprox CatD	Airprox None		
1.50%	0.23%	1.08%	17.58%	79.61%	**Airspace Interaction Behind_R**	
0.00%	0.00%	0.03%	25.53%	74.44%	PREL-3360-1400-1400-0nm5	
2.56%	0.75%	2.47%	40.69%	53.53%	PREL-3360-1400-2500-0nm5	
0.00%	0.00%	0.00%	1.59%	98.41%	PREL-3360-1400-9425-0nm5	
0.00%	0.00%	0.13%	16.53%	83.34%	PREL-3360-2500-1400-0nm5	
0.00%	0.00%	1.50%	34.19%	64.31%	PREL-3360-2500-2500-0nm5	
0.00%	0.00%	0.00%	0.13%	99.88%	PREL-3360-2500-9425-0nm5	
0.88%	**0.28%**	**1.00%**	**11.91%**	**85.94%**	**PREL-3360-9425-1400-0nm5**	TC10
10.03%	1.06%	4.47%	26.34%	58.09%	PREL-3360-9425-2500-0nm5	
0.00%	0.00%	0.16%	1.31%	98.53%	PREL-3360-9425-9425-0nm5	

Airprox CatA	Airprox CatB	Airprox CatC	Airprox CatD	Airprox None		
2.18%	0.95%	3.91%	43.22%	49.75%	**Airspace Interaction from Right**	
1.25%	0.38%	3.25%	48.00%	47.13%	PREL-2750-1400-1400-0nm5	
3.44%	2.06%	6.16%	43.19%	45.16%	PREL-2750-1400-2500-0nm5	
0.41%	0.09%	1.25%	44.94%	53.31%	PREL-2750-1400-9425-0nm5	
1.84%	1.25%	4.31%	42.63%	49.97%	PREL-2750-2500-1400-0nm5	
2.59%	1.31%	5.56%	41.81%	48.72%	PREL-2750-2500-2500-0nm5	
0.72%	0.66%	2.25%	46.34%	50.03%	PREL-2750-2500-9425-0nm5	
1.44%	0.59%	4.56%	43.22%	50.19%	PREL-2750-9425-1400-0nm5	
7.59%	2.00%	5.66%	37.16%	47.59%	PREL-2750-9425-2500-0nm5	
0.34%	**0.19%**	**2.19%**	**41.66%**	**55.63%**	**PREL-2750-9425-9425-0nm5**	TC12

Airprox CatA	Airprox CatB	Airprox CatC	Airprox CatD	Airprox None		
2.15%	0.86%	4.05%	35.29%	57.65%	**Airspace Interaction without EM Flaw**	
0.00%	0.00%	1.88%	23.06%	75.06%	ENCTS-2025-9425-2030-0nm5	
0.00%	0.00%	1.78%	31.84%	66.38%	ENCTS-2205-9425-2030-0nm5	
0.00%	0.00%	1.63%	31.78%	66.59%	ENCTS-2275-9425-2030-0nm5	
2.53%	0.97%	4.38%	41.34%	50.78%	ENCTS-2700-9425-2030-0nm5	
12.75%	**3.59%**	**15.94%**	**53.22%**	**14.50%**	**ENCTS-2880-9425-2030-0nm5**	TC2
1.44%	**0.16%**	**2.34%**	**24.84%**	**71.22%**	**ES-2025-2500-2500-0nm5**	TC3
2.72%	**2.00%**	**5.69%**	**36.84%**	**52.75%**	**L-2835-2500-2500-0nm5**	TC4
6.16%	**2.03%**	**6.69%**	**53.88%**	**31.25%**	**MWT-2880-1400-1400-0nm5**	TC5
0.00%	0.00%	2.53%	59.25%	38.22%	MWT-3015-1400-1400-0nm5	
0.00%	0.00%	0.00%	0.88%	99.13%	NONE-1850-2500-2500-3nm0	
0.00%	0.00%	0.66%	36.88%	62.47%	NONE-3050-2500-2500-3nm0	
2.38%	**2.38%**	**8.91%**	**38.25%**	**48.09%**	**PRE-3015-9425-1400-0nm5**	TC6
0.00%	0.00%	0.28%	26.69%	73.03%	PRE-3050-2500-2500-3nm0	

Table A.2 Airspace Interaction Airprox Category Rates of Occurrence by Dataset (Test Cases)

In selecting the test cases for validation and subsequent incorporation into the enhanced HAZOP, highlighted in Table A.2, the aim was to identify those with the highest rated generating category "A" occurrences whilst also involving the slowest combination of interaction speeds. In the case of the approaches from ahead a second test case was chosen at maximum speeds also. Whilst it is expected that head-on interactions at higher closing velocities are obviously hazardous, which appears to be borne out by the incidence category "A" events revealed by the table, this particular case (TC7) is of interest in considering whether it might be more or less diagnosable using the Bayesian inference model. The remaining special cases were selected on the basis of their representation of specific flaws and in producing a higher relative rate of category "A" occurrences. Individual test cases were then subsequently drawn from each dataset by identifying an appropriate category "A" air proximity event within each dataset (plus a few representative collisions). Having identified, selected and obtained the specific test case results, both directly from the system dynamics model output data and as processed by the Bayesian learning, these event series are then processed to reveal the extra information subsequently used to enhance the HAZOP capturing the likely complex behaviour in a succinct form.

A.6. Validation Test Cases and Results

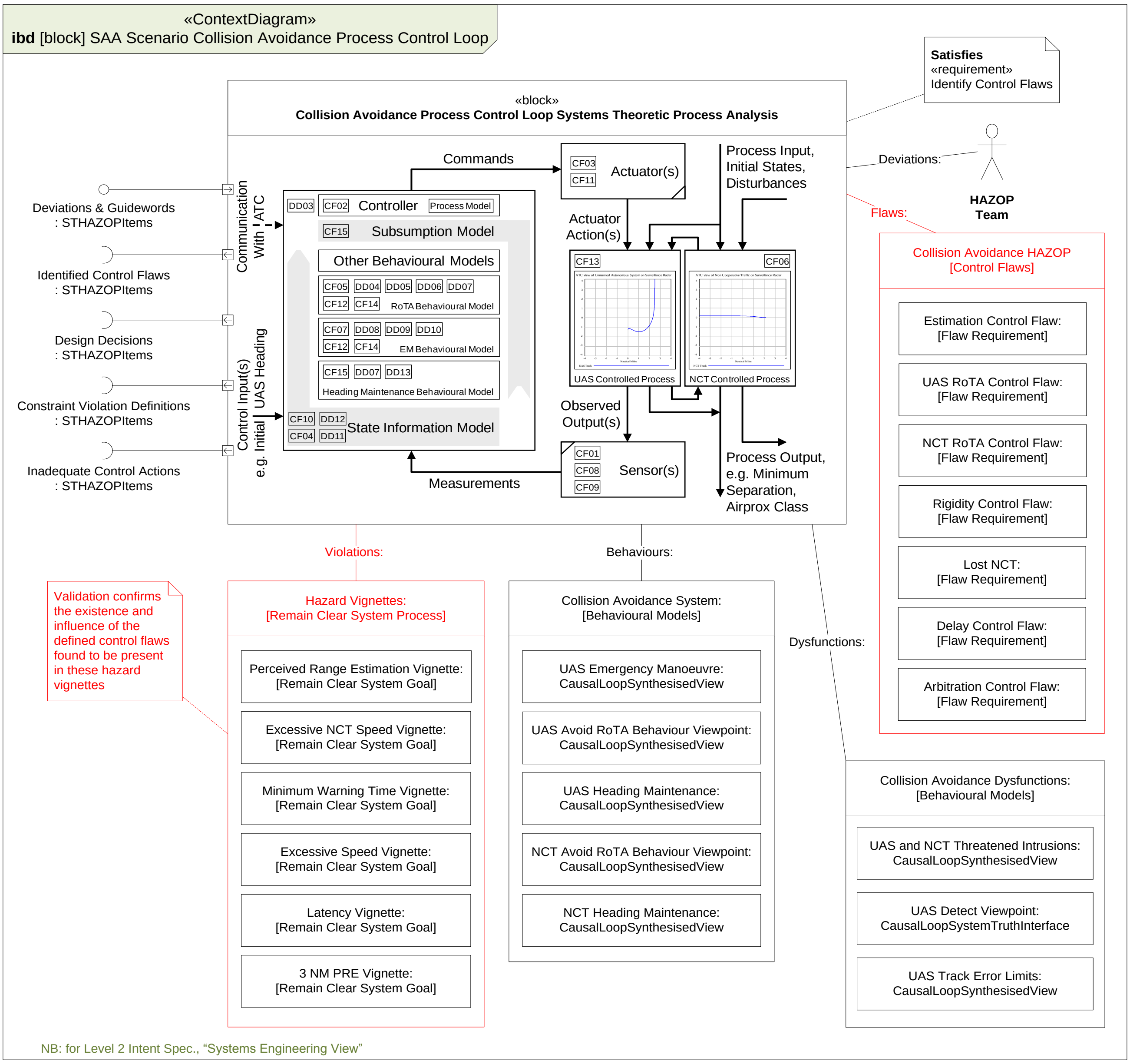
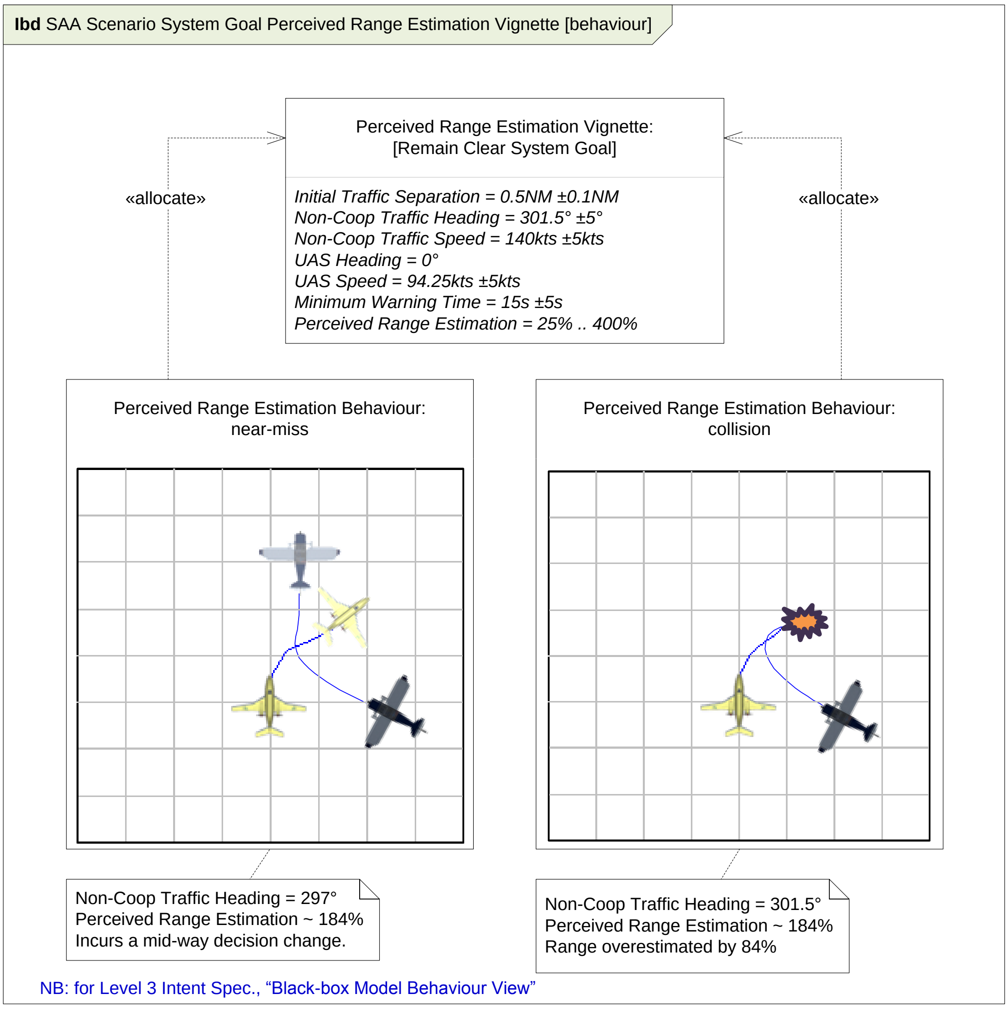
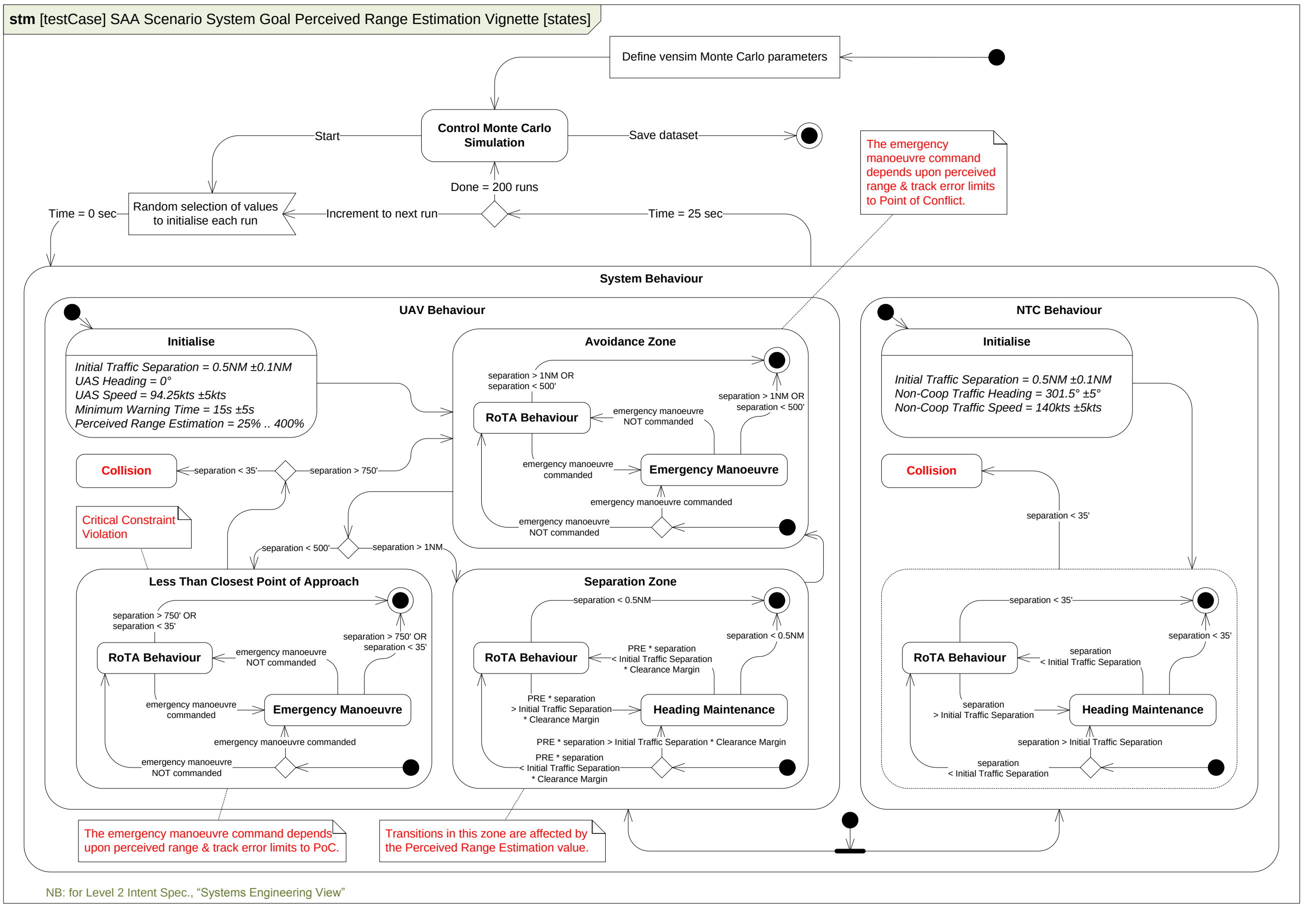


Figure A.9 Collision Avoidance Process Control Loop context: Hazard Vignettes and Flaws


ibd SAA Scenario System Goal Perceived Range Estimation Vignette [behaviour]

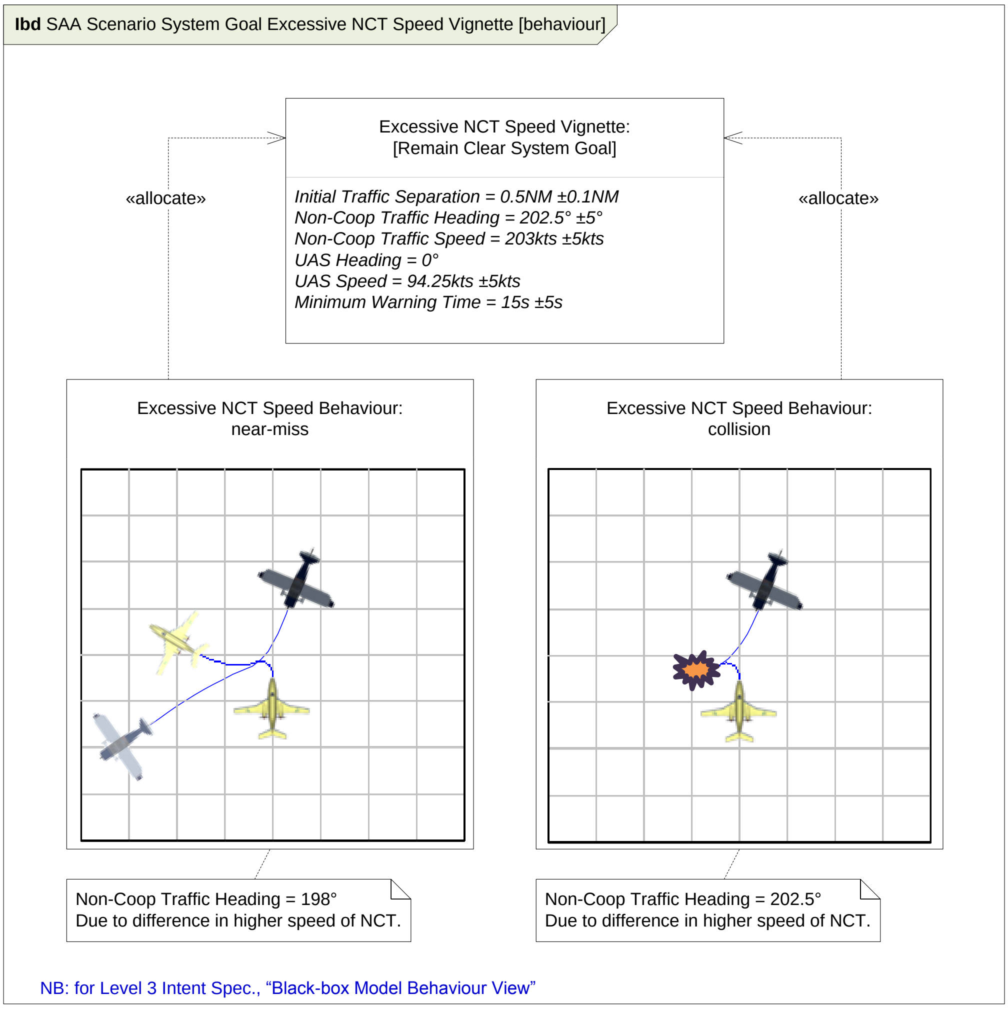
Perceived Range Estimation Vignette:
[Remain Clear System Goal]
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 301.5° ±5°
Non-Coop Traffic Speed = 140kts ±5kts
UAS Heading = 0°
UAS Speed = 94.25kts ±5kts
Minimum Warning Time = 15s ±5s
Perceived Range Estimation = 25% .. 400%
«allocate»
«allocate»
Perceived Range Estimation Behaviour:
near-miss
Perceived Range Estimation Behaviour:
collision
Non-Coop Traffic Heading = 297°
Perceived Range Estimation ~ 184%
Incurs a mid-way decision change.
Non-Coop Traffic Heading = 301.5°
Perceived Range Estimation ~ 184%
Range overestimated by 84%

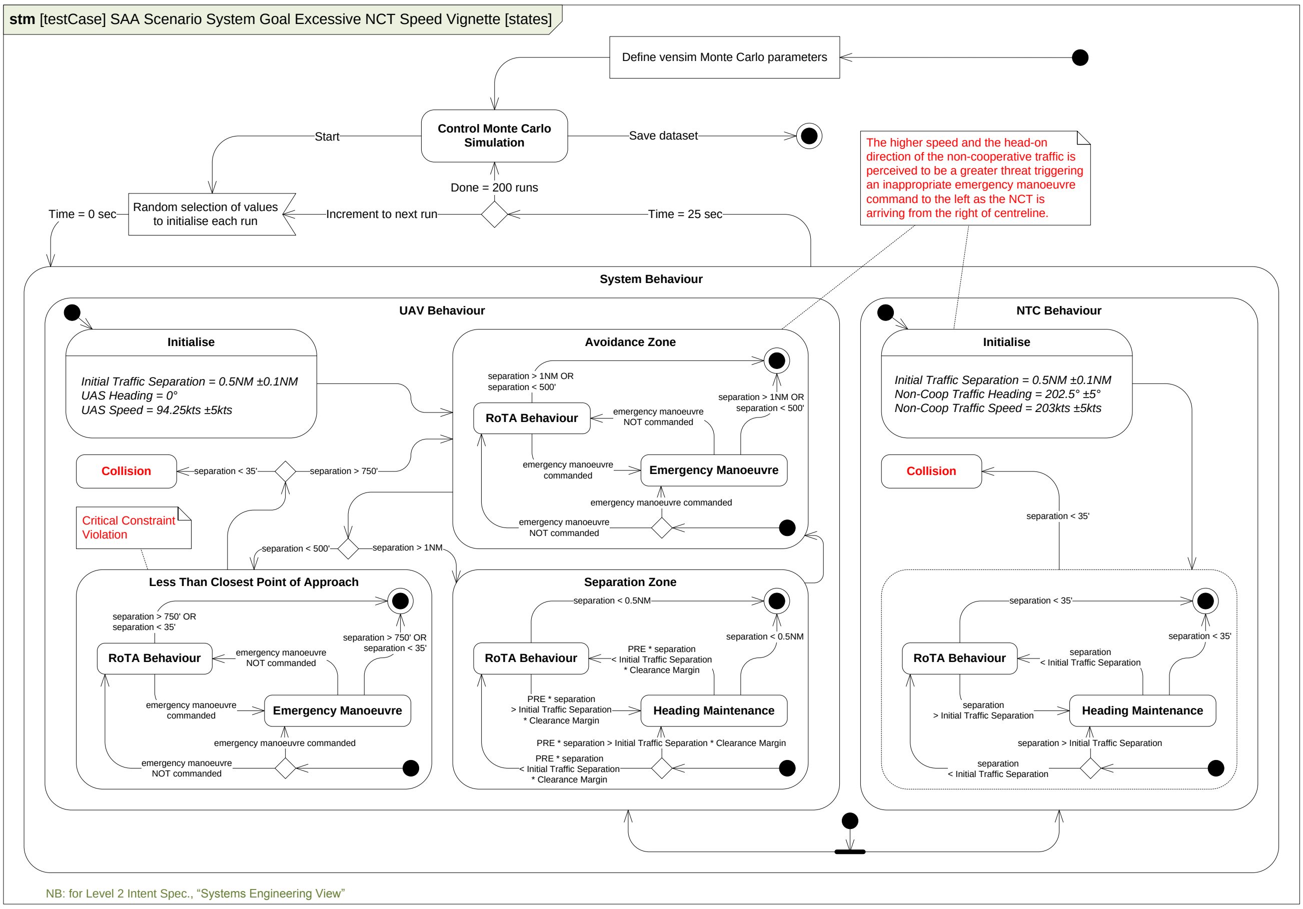
NB: for Level 3 Intent Spec., “Black-box Model Behaviour View”

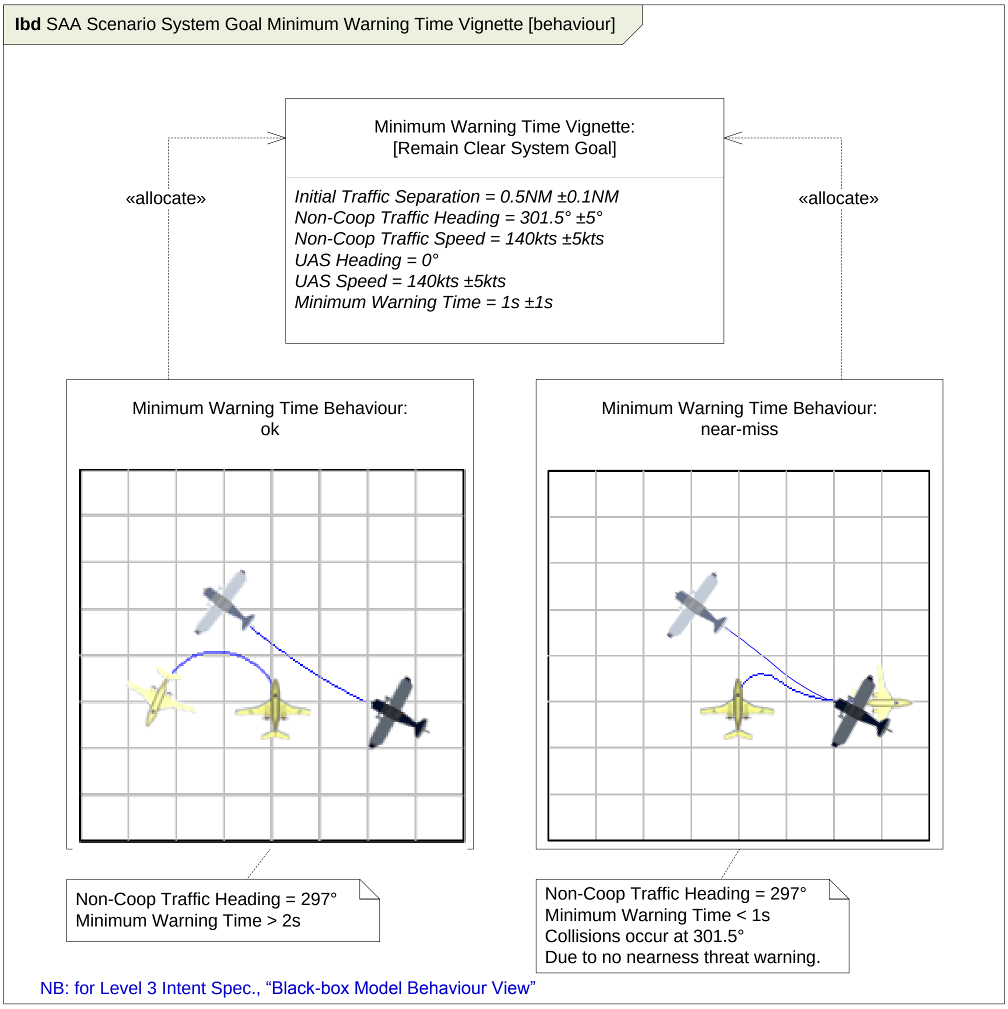

stm [testCase] SAA Scenario System Goal Perceived Range Estimation Vignette [states]

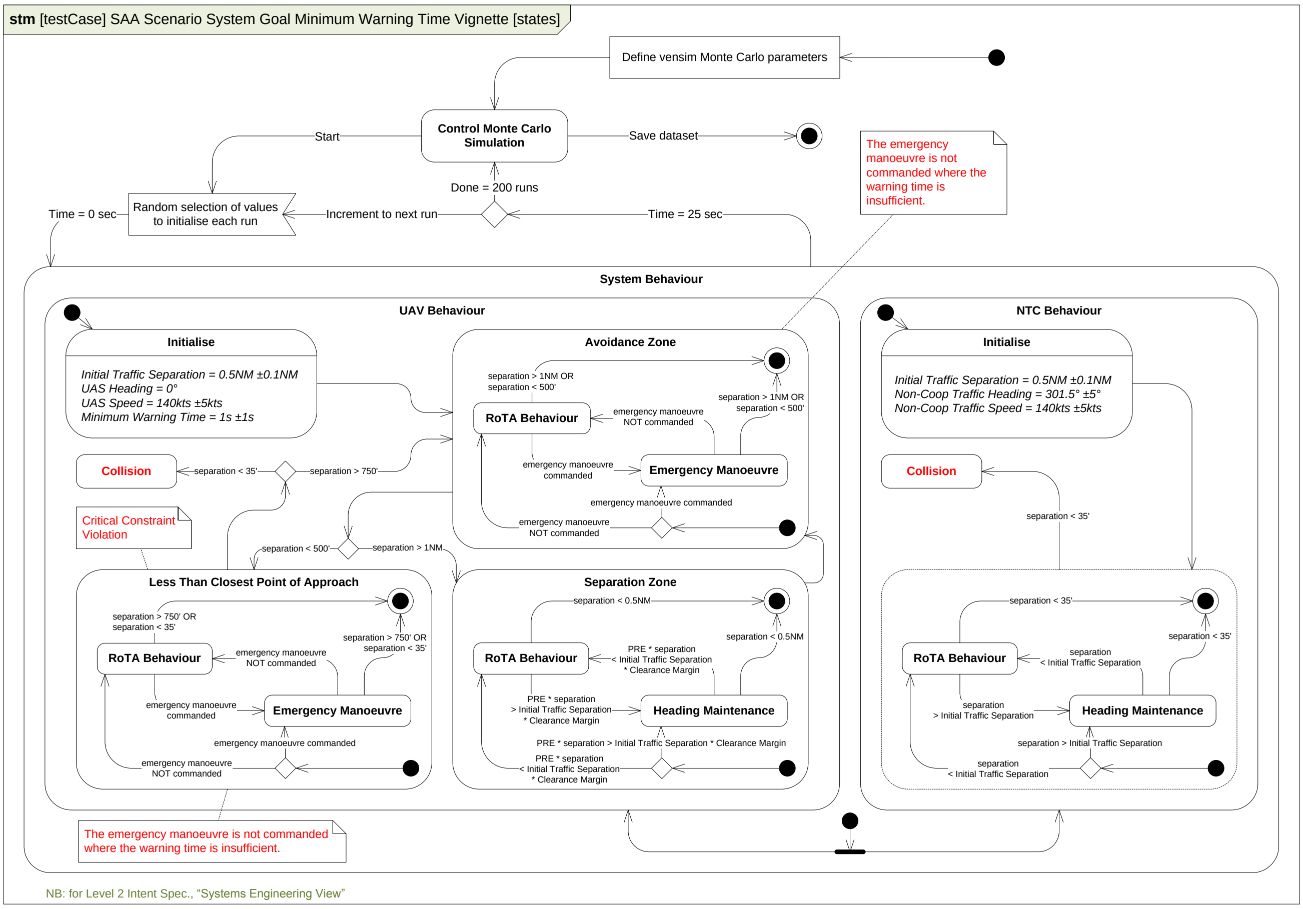
Define vensim Monte Carlo parameters
Control Monte Carlo Simulation
Start
Save dataset
Done = 200 runs
Random selection of values to initialise each run
Time = 0 sec
Increment to next run
Time = 25 sec
The emergency manoeuvre command depends upon perceived range & track error limits to Point of Conflict.
System Behaviour
UAV Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
UAS Heading = 0°
UAS Speed = 94.25kts ±5kts
Minimum Warning Time = 15s ±5s
Perceived Range Estimation = 25% .. 400%
Collision
separation < 35'
separation > 750'
Critical Constraint Violation
separation < 500'
separation > 1NM
Avoidance Zone
separation > 1NM OR separation < 500'
RoTA Behaviour
emergency manoeuvre NOT commanded
separation > 1NM OR separation < 500'
emergency manoeuvre commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre NOT commanded
Less Than Closest Point of Approach
separation > 750' OR separation < 35'
RoTA Behaviour
emergency manoeuvre NOT commanded
separation > 750' OR separation < 35'
emergency manoeuvre commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre NOT commanded
Separation Zone
separation < 0.5NM
separation < 0.5NM
RoTA Behaviour
PRE * separation < Initial Traffic Separation * Clearance Margin
PRE * separation > Initial Traffic Separation * Clearance Margin
Heading Maintenance
PRE * separation > Initial Traffic Separation * Clearance Margin
PRE * separation < Initial Traffic Separation * Clearance Margin
NTC Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 301.5° ±5°
Non-Coop Traffic Speed = 140kts ±5kts
Collision
separation < 35'
separation < 35'
separation < 35'
RoTA Behaviour
separation < Initial Traffic Separation
separation > Initial Traffic Separation
Heading Maintenance
separation > Initial Traffic Separation
separation < Initial Traffic Separation
The emergency manoeuvre command depends upon perceived range & track error limits to PoC.
Transitions in this zone are affected by the Perceived Range Estimation value.

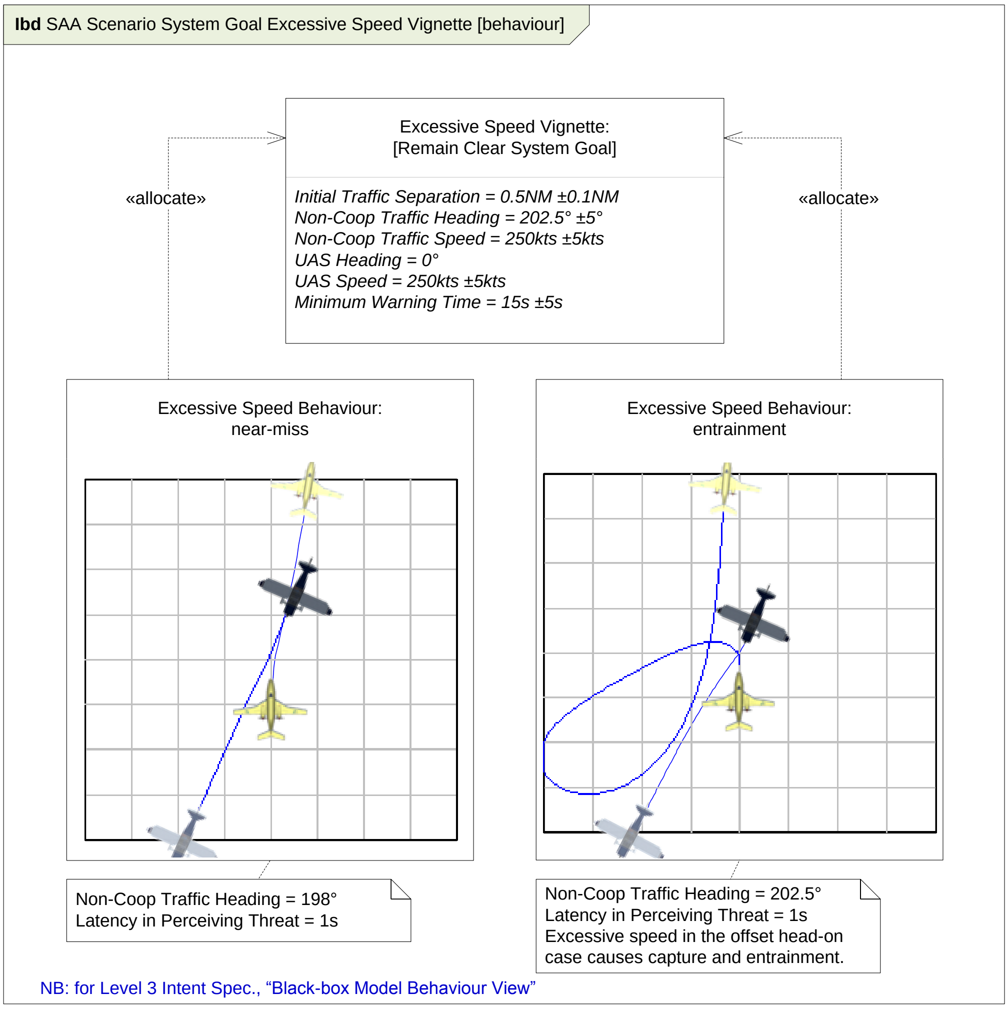
NB: for Level 2 Intent Spec., “Systems Engineering View”


ibd SAA Scenario System Goal Excessive NCT Speed Vignette [behaviour]
Excessive NCT Speed Vignette:
[Remain Clear System Goal]
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 202.5° ±5°
Non-Coop Traffic Speed = 203kts ±5kts
UAS Heading = 0°
UAS Speed = 94.25kts ±5kts
Minimum Warning Time = 15s ±5s
«allocate»
«allocate»
Excessive NCT Speed Behaviour:
near-miss
Excessive NCT Speed Behaviour:
collision
Non-Coop Traffic Heading = 198°
Due to difference in higher speed of NCT.
Non-Coop Traffic Heading = 202.5°
Due to difference in higher speed of NCT.
NB: for Level 3 Intent Spec., “Black-box Model Behaviour View”


stm [testCase] SAA Scenario System Goal Excessive NCT Speed Vignette [states]
Define vensim Monte Carlo parameters
Control Monte Carlo Simulation
Start
Save dataset
Done = 200 runs
Random selection of values to initialise each run
Time = 0 sec
Increment to next run
Time = 25 sec
The higher speed and the head-on direction of the non-cooperative traffic is perceived to be a greater threat triggering an inappropriate emergency manoeuvre command to the left as the NCT is arriving from the right of centreline.
System Behaviour
UAV Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
UAS Heading = 0°
UAS Speed = 94.25kts ±5kts
Collision
separation < 35'
separation > 750'
Critical Constraint Violation
separation < 500'
separation > 1NM
Avoidance Zone
separation > 1NM OR separation < 500'
RoTA Behaviour
emergency manoeuvre NOT commanded
separation > 1NM OR separation < 500'
emergency manoeuvre commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre NOT commanded
Less Than Closest Point of Approach
separation > 750' OR separation < 35'
RoTA Behaviour
emergency manoeuvre NOT commanded
separation > 750' OR separation < 35'
emergency manoeuvre commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre NOT commanded
Separation Zone
separation < 0.5NM
RoTA Behaviour
PRE * separation < Initial Traffic Separation * Clearance Margin
separation < 0.5NM
PRE * separation > Initial Traffic Separation * Clearance Margin
Heading Maintenance
PRE * separation > Initial Traffic Separation * Clearance Margin
PRE * separation < Initial Traffic Separation * Clearance Margin
NTC Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 202.5° ±5°
Non-Coop Traffic Speed = 203kts ±5kts
Collision
separation < 35'
separation < 35'
RoTA Behaviour
separation < Initial Traffic Separation
separation < 35'
separation > Initial Traffic Separation
Heading Maintenance
separation > Initial Traffic Separation
separation < Initial Traffic Separation
NB: for Level 2 Intent Spec., “Systems Engineering View”


ibd SAA Scenario System Goal Minimum Warning Time Vignette [behaviour]
Minimum Warning Time Vignette:
[Remain Clear System Goal]
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 301.5° ±5°
Non-Coop Traffic Speed = 140kts ±5kts
UAS Heading = 0°
UAS Speed = 140kts ±5kts
Minimum Warning Time = 1s ±1s
«allocate»
«allocate»
Minimum Warning Time Behaviour:
ok
Minimum Warning Time Behaviour:
near-miss
Non-Coop Traffic Heading = 297°
Minimum Warning Time > 2s
Non-Coop Traffic Heading = 297°
Minimum Warning Time < 1s
Collisions occur at 301.5°
Due to no nearness threat warning.
NB: for Level 3 Intent Spec., "Black-box Model Behaviour View"

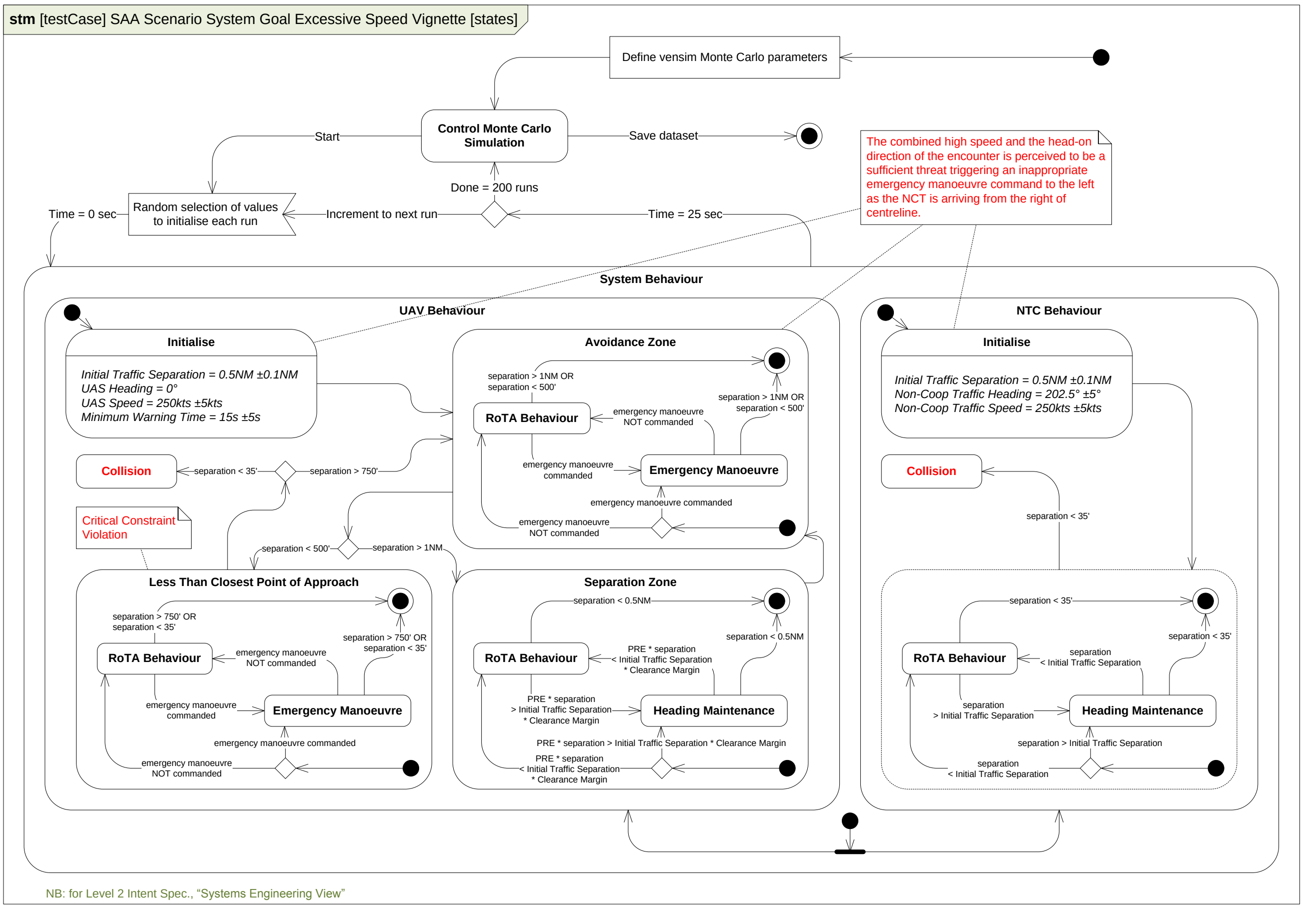

stm [testCase] SAA Scenario System Goal Minimum Warning Time Vignette [states]
Define vensim Monte Carlo parameters
Control Monte Carlo Simulation
Start
Save dataset
The emergency manoeuvre is not commanded where the warning time is insufficient.
Done = 200 runs
Time = 0 sec
Random selection of values to initialise each run
Increment to next run
Time = 25 sec
System Behaviour
UAV Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
UAS Heading = 0°
UAS Speed = 140kts ±5kts
Minimum Warning Time = 1s ±1s
Avoidance Zone
separation > 1NM OR
separation < 500'
RoTA Behaviour
emergency manoeuvre
NOT commanded
separation > 1NM OR
separation < 500'
emergency manoeuvre
commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre
NOT commanded
Collision
separation < 35'
separation > 750'
Critical Constraint Violation
separation < 500'
separation > 1NM
Less Than Closest Point of Approach
separation > 750' OR
separation < 35'
separation > 750' OR
separation < 35'
RoTA Behaviour
emergency manoeuvre
NOT commanded
emergency manoeuvre
commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre
NOT commanded
Separation Zone
separation < 0.5NM
separation < 0.5NM
RoTA Behaviour
PRE * separation
< Initial Traffic Separation
* Clearance Margin
PRE * separation
> Initial Traffic Separation
* Clearance Margin
Heading Maintenance
PRE * separation > Initial Traffic Separation * Clearance Margin
PRE * separation
< Initial Traffic Separation
* Clearance Margin
The emergency manoeuvre is not commanded where the warning time is insufficient.
NTC Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 301.5° ±5°
Non-Coop Traffic Speed = 140kts ±5kts
Collision
separation < 35'
separation < 35'
separation < 35'
RoTA Behaviour
separation
< Initial Traffic Separation
separation
> Initial Traffic Separation
Heading Maintenance
separation > Initial Traffic Separation
separation
< Initial Traffic Separation
NB: for Level 2 Intent Spec., "Systems Engineering View"



ibd SAA Scenario System Goal Excessive Speed Vignette [behaviour]

Excessive Speed Vignette:
[Remain Clear System Goal]
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 202.5° ±5°
Non-Coop Traffic Speed = 250kts ±5kts
UAS Heading = 0°
UAS Speed = 250kts ±5kts
Minimum Warning Time = 15s ±5s
«allocate»
«allocate»
Excessive Speed Behaviour:
near-miss
Excessive Speed Behaviour:
entrainment
Non-Coop Traffic Heading = 198°
Latency in Perceiving Threat = 1s
Non-Coop Traffic Heading = 202.5°
Latency in Perceiving Threat = 1s
Excessive speed in the offset head-on case causes capture and entrainment.

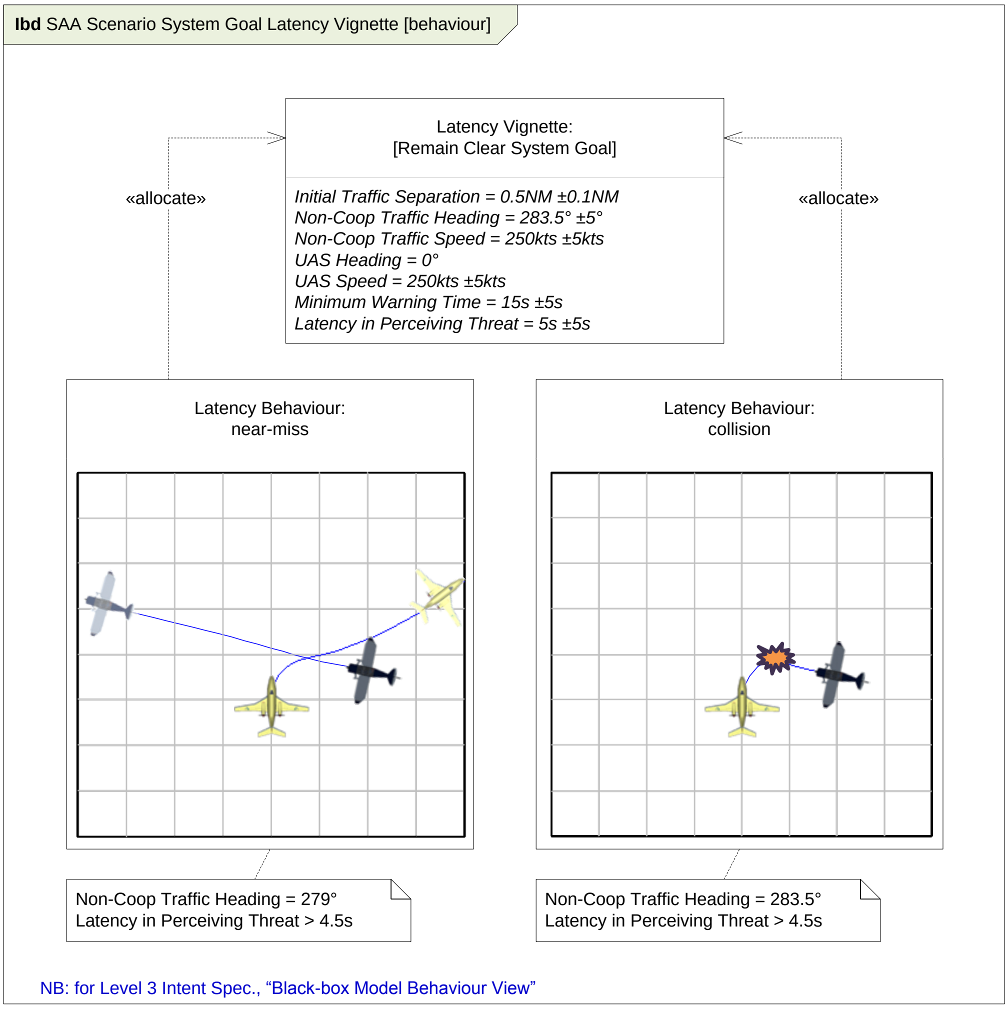
NB: for Level 3 Intent Spec., “Black-box Model Behaviour View”

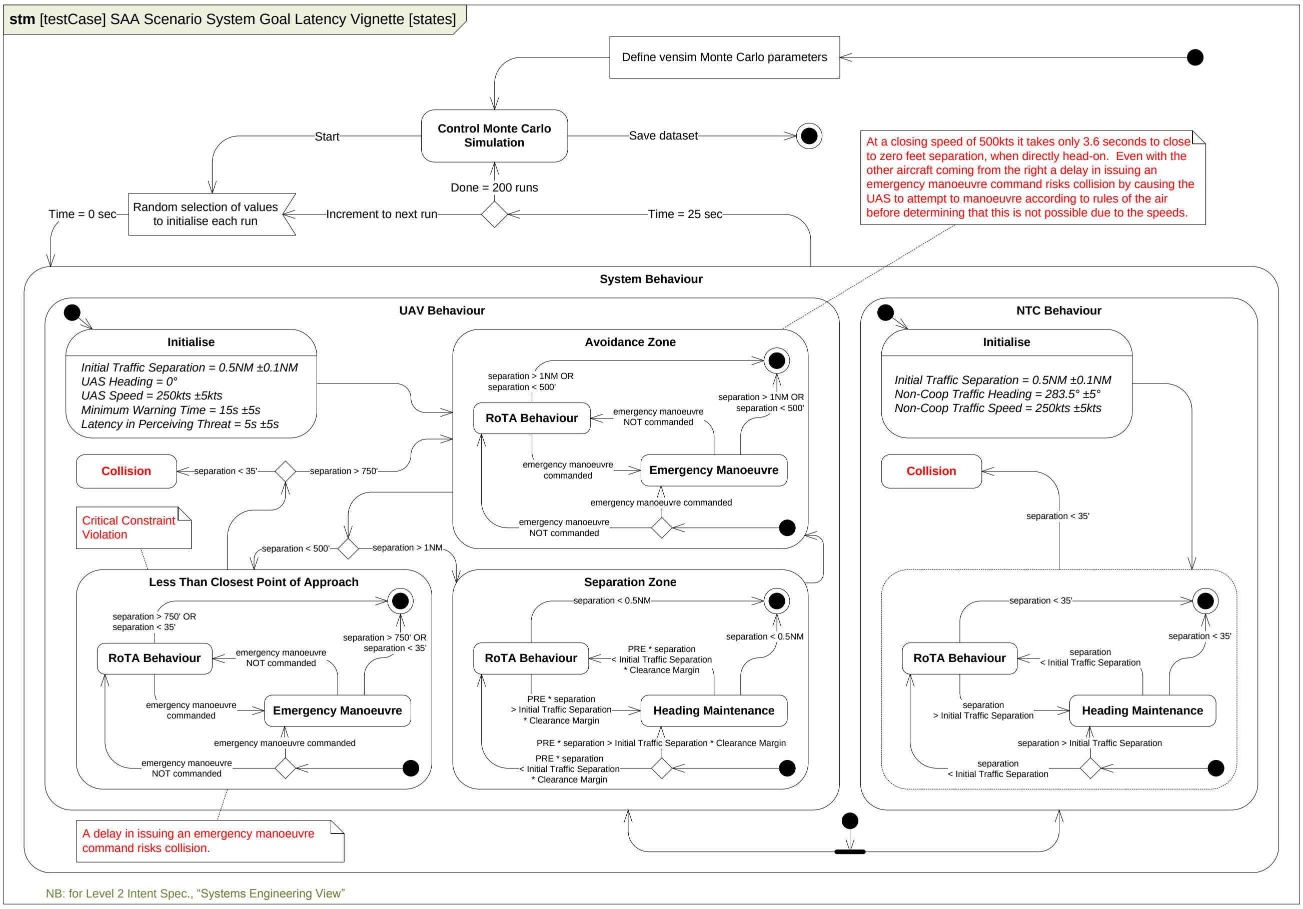


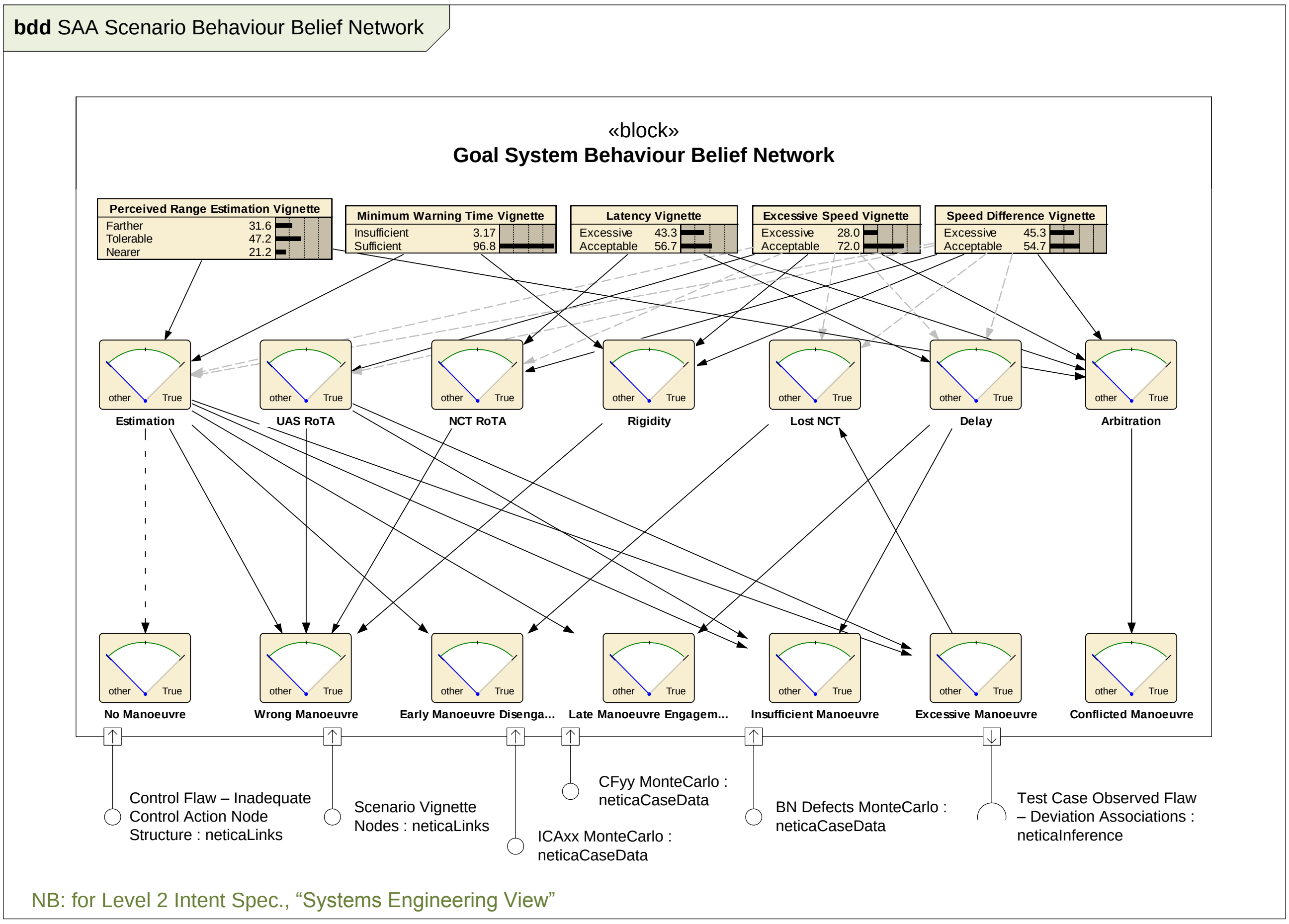
stm [testCase] SAA Scenario System Goal Excessive Speed Vignette [states]

Define vensim Monte Carlo parameters
Control Monte Carlo Simulation
Start
Save dataset
Done = 200 runs
Random selection of values to initialise each run
Time = 0 sec
Increment to next run
Time = 25 sec
The combined high speed and the head-on direction of the encounter is perceived to be a sufficient threat triggering an inappropriate emergency manoeuvre command to the left as the NCT is arriving from the right of centreline.
System Behaviour
UAV Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
UAS Heading = 0°
UAS Speed = 250kts ±5kts
Minimum Warning Time = 15s ±5s
Collision
separation < 35'
separation > 750'
Critical Constraint Violation
separation < 500'
separation > 1NM
Avoidance Zone
separation > 1NM OR separation < 500'
RoTA Behaviour
emergency manoeuvre NOT commanded
separation > 1NM OR separation < 500'
emergency manoeuvre commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre NOT commanded
Less Than Closest Point of Approach
separation > 750' OR separation < 35'
RoTA Behaviour
emergency manoeuvre NOT commanded
separation > 750' OR separation < 35'
emergency manoeuvre commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre NOT commanded
Separation Zone
separation < 0.5NM
RoTA Behaviour
PRE * separation < Initial Traffic Separation * Clearance Margin
separation < 0.5NM
PRE * separation > Initial Traffic Separation * Clearance Margin
Heading Maintenance
PRE * separation > Initial Traffic Separation * Clearance Margin
PRE * separation < Initial Traffic Separation * Clearance Margin
NTC Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 202.5° ±5°
Non-Coop Traffic Speed = 250kts ±5kts
Collision
separation < 35'
separation < 35'
RoTA Behaviour
separation < Initial Traffic Separation
separation < 35'
separation > Initial Traffic Separation
Heading Maintenance
separation > Initial Traffic Separation
separation < Initial Traffic Separation

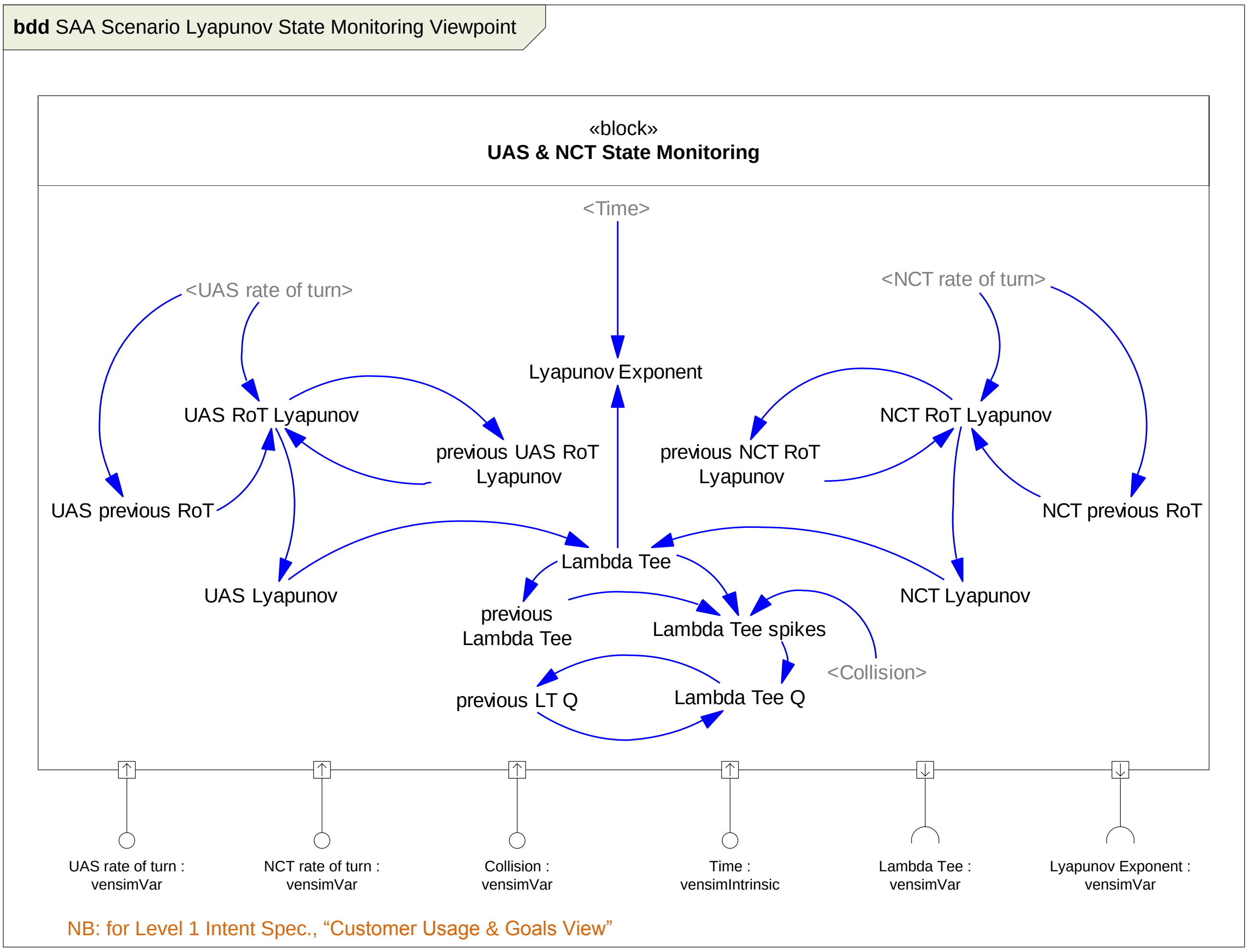
NB: for Level 2 Intent Spec., “Systems Engineering View”


ibd SAA Scenario System Goal Latency Vignette [behaviour]
Latency Vignette:
[Remain Clear System Goal]
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 283.5° ±5°
Non-Coop Traffic Speed = 250kts ±5kts
UAS Heading = 0°
UAS Speed = 250kts ±5kts
Minimum Warning Time = 15s ±5s
Latency in Perceiving Threat = 5s ±5s
«allocate»
«allocate»
Latency Behaviour:
near-miss
Latency Behaviour:
collision
Non-Coop Traffic Heading = 279°
Latency in Perceiving Threat > 4.5s
Non-Coop Traffic Heading = 283.5°
Latency in Perceiving Threat > 4.5s
NB: for Level 3 Intent Spec., “Black-box Model Behaviour View”

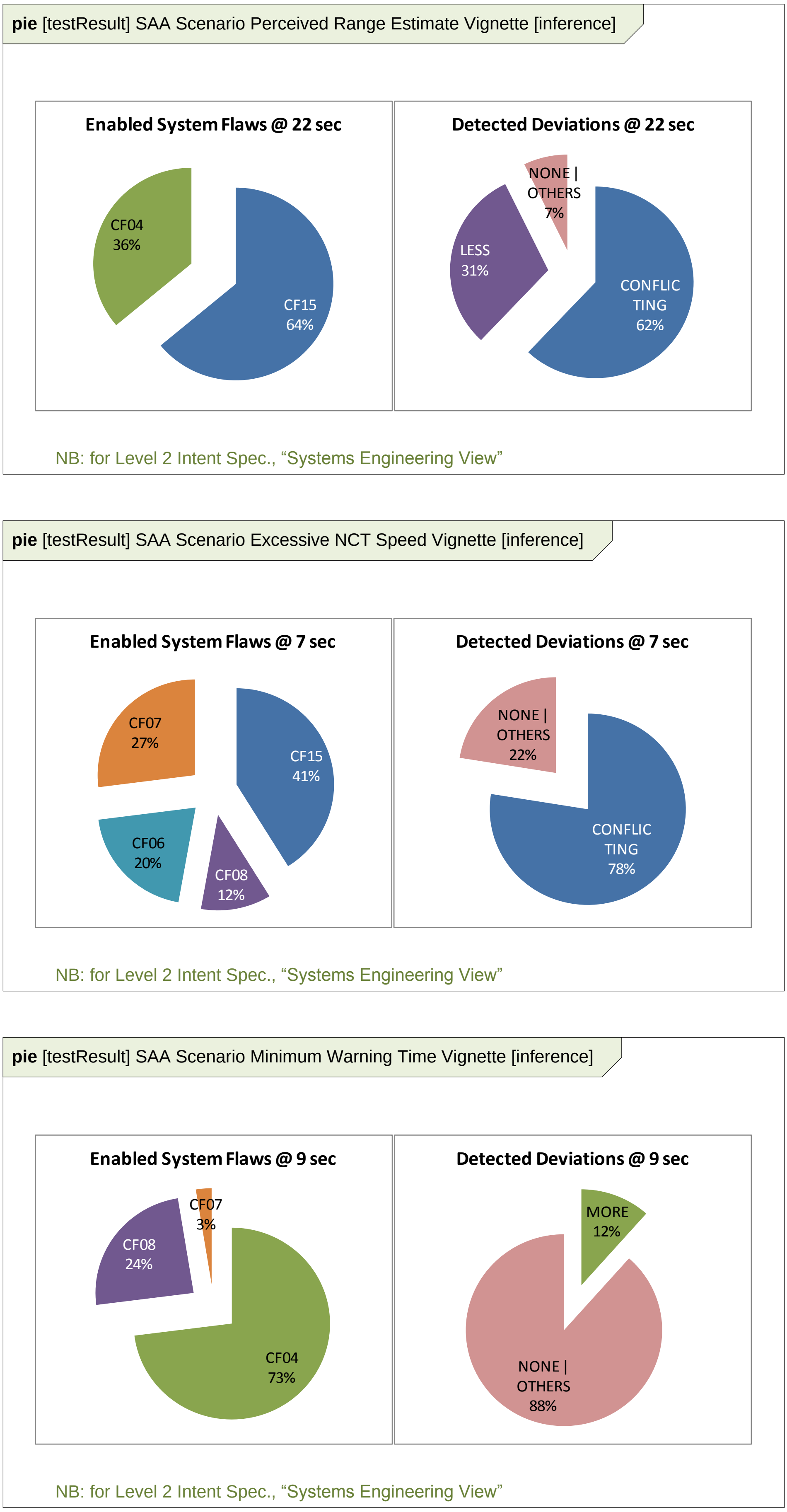

stm [testCase] SAA Scenario System Goal Latency Vignette [states]
Define vensim Monte Carlo parameters
Control Monte Carlo Simulation
Start
Save dataset
Done = 200 runs
Random selection of values to initialise each run
Time = 0 sec
Increment to next run
Time = 25 sec
At a closing speed of 500kts it takes only 3.6 seconds to close to zero feet separation, when directly head-on. Even with the other aircraft coming from the right a delay in issuing an emergency manoeuvre command risks collision by causing the UAS to attempt to manoeuvre according to rules of the air before determining that this is not possible due to the speeds.
System Behaviour
UAV Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
UAS Heading = 0°
UAS Speed = 250kts ±5kts
Minimum Warning Time = 15s ±5s
Latency in Perceiving Threat = 5s ±5s
Avoidance Zone
separation > 1NM OR separation < 500'
RoTA Behaviour
emergency manoeuvre NOT commanded
separation > 1NM OR separation < 500'
emergency manoeuvre commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre NOT commanded
Collision
separation < 35'
separation > 750'
Critical Constraint Violation
separation < 500'
separation > 1NM
Less Than Closest Point of Approach
separation > 750' OR separation < 35'
RoTA Behaviour
emergency manoeuvre NOT commanded
separation > 750' OR separation < 35'
emergency manoeuvre commanded
Emergency Manoeuvre
emergency manoeuvre commanded
emergency manoeuvre NOT commanded
Separation Zone
separation < 0.5NM
separation < 0.5NM
RoTA Behaviour
PRE * separation < Initial Traffic Separation * Clearance Margin
PRE * separation > Initial Traffic Separation * Clearance Margin
Heading Maintenance
PRE * separation > Initial Traffic Separation * Clearance Margin
PRE * separation < Initial Traffic Separation * Clearance Margin
A delay in issuing an emergency manoeuvre command risks collision.
NTC Behaviour
Initialise
Initial Traffic Separation = 0.5NM ±0.1NM
Non-Coop Traffic Heading = 283.5° ±5°
Non-Coop Traffic Speed = 250kts ±5kts
Collision
separation < 35'
separation < 35'
separation < 35'
RoTA Behaviour
separation < Initial Traffic Separation
separation > Initial Traffic Separation
Heading Maintenance
separation > Initial Traffic Separation
separation < Initial Traffic Separation
NB: for Level 2 Intent Spec., “Systems Engineering View”


bdd SAA Scenario Behaviour Belief Network
«block»
Goal System Behaviour Belief Network
Perceived Range Estimation Vignette
Farther 31.6
Tolerable 47.2
Nearer 21.2
Minimum Warning Time Vignette
Insufficient 3.17
Sufficient 96.8
Latency Vignette
Excessive 43.3
Acceptable 56.7
Excessive Speed Vignette
Excessive 28.0
Acceptable 72.0
Speed Difference Vignette
Excessive 45.3
Acceptable 54.7
other True
Estimation
UAS RoTA
NCT RoTA
Rigidity
Lost NCT
Delay
Arbitration
No Manoeuvre
Wrong Manoeuvre
Early Manoeuvre Disenga...
Late Manoeuvre Engagem...
Insufficient Manoeuvre
Excessive Manoeuvre
Conflicted Manoeuvre
Control Flaw – Inadequate Control Action Node Structure : neticaLinks
Scenario Vignette Nodes : neticaLinks
ICAxx MonteCarlo : neticaCaseData
CFyy MonteCarlo : neticaCaseData
BN Defects MonteCarlo : neticaCaseData
Test Case Observed Flaw – Deviation Associations : neticaInference

NB: for Level 2 Intent Spec., “Systems Engineering View”



NB: for Level 1 Intent Spec., “Customer Usage & Goals View”



pie [testResult] SAA Scenario Perceived Range Estimate Vignette [inference]

pie [testResult] SAA Scenario Excessive NCT Speed Vignette [inference]

pie [testResult] SAA Scenario Minimum Warning Time Vignette [inference]

Figure A.10 Control Flaw Relative Likelihoods for Test Cases 6, 1 and 5 from Bayesian Learning

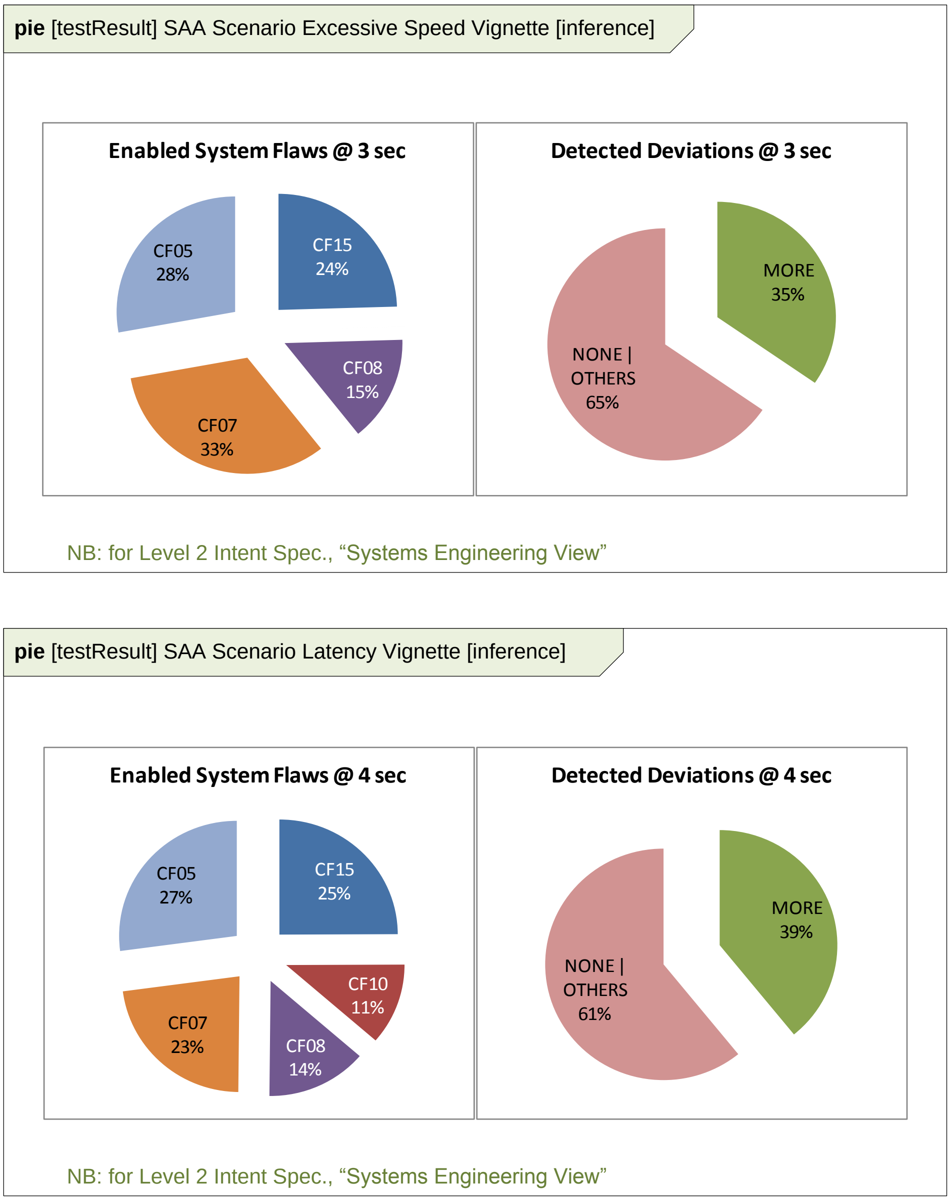
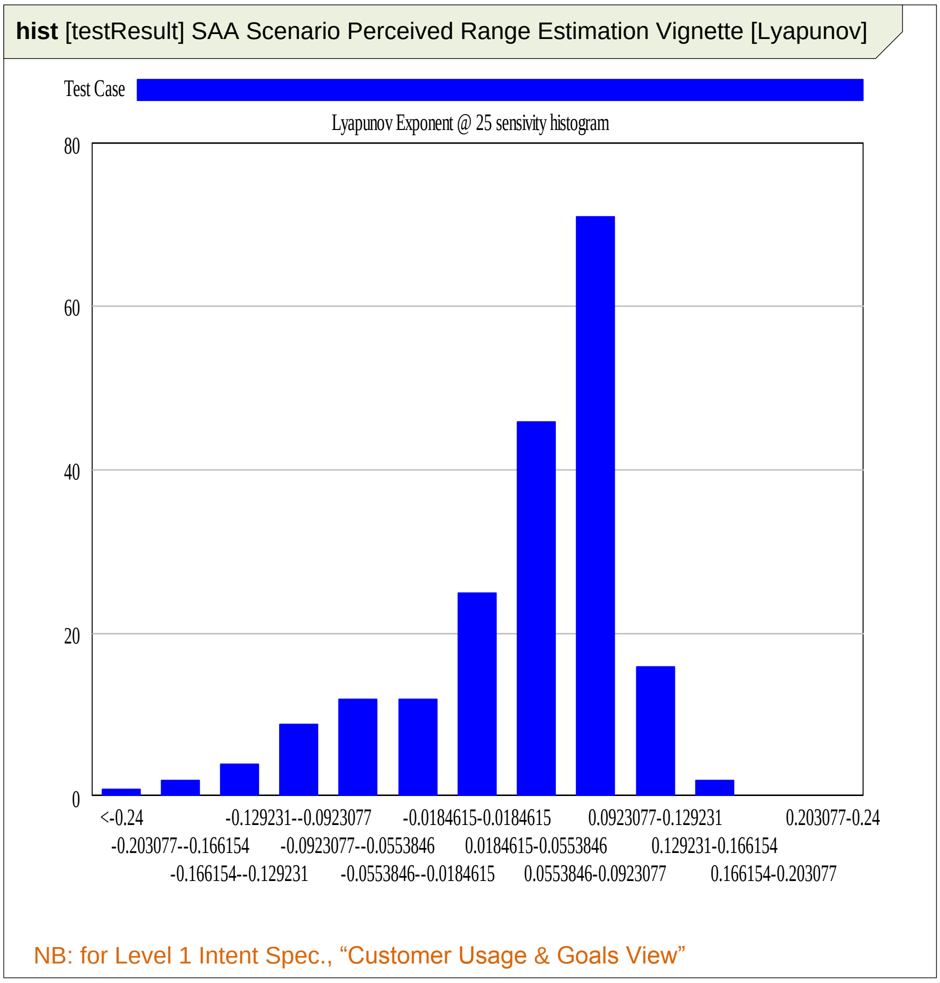
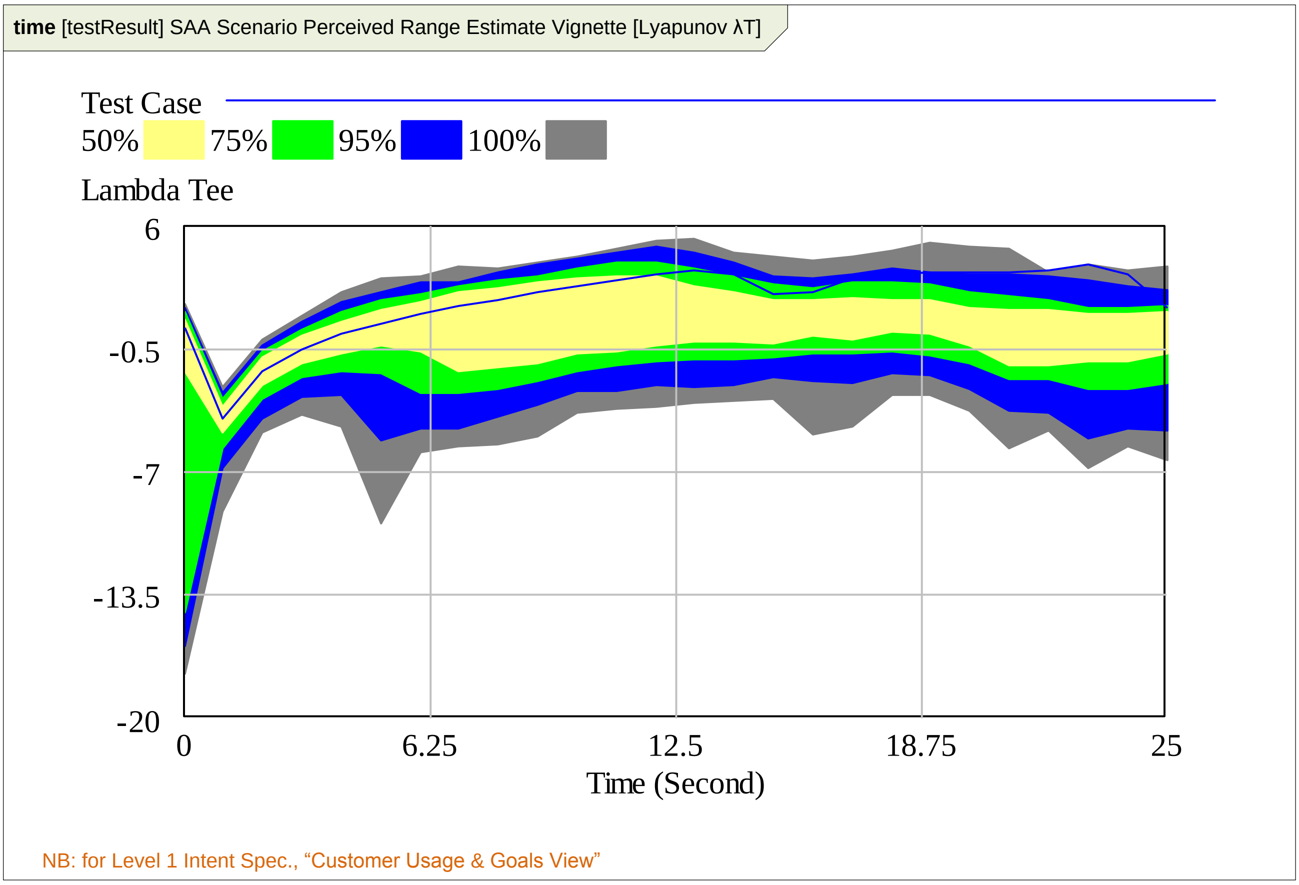


Figure A.11 Control Flaw Relative Likelihoods for Test Cases 3 and 4 from Bayesian Learning

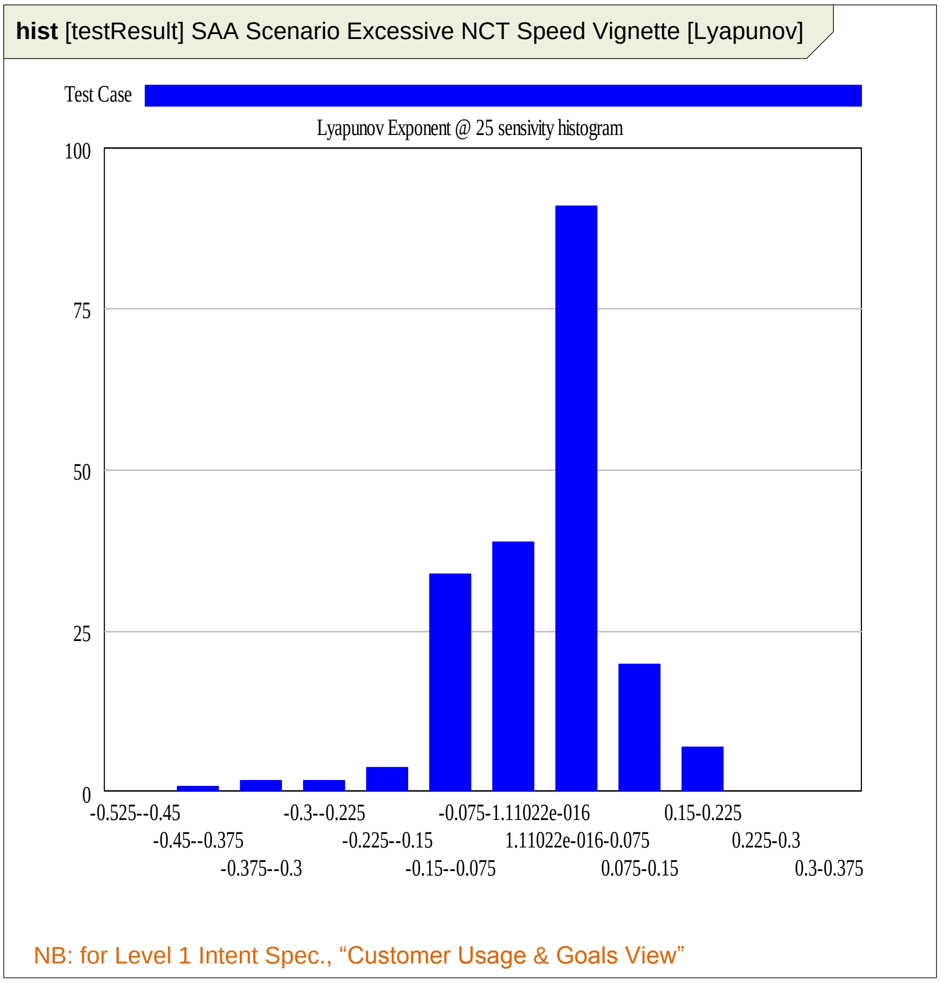


NB: for Level 1 Intent Spec., "Customer Usage & Goals View"

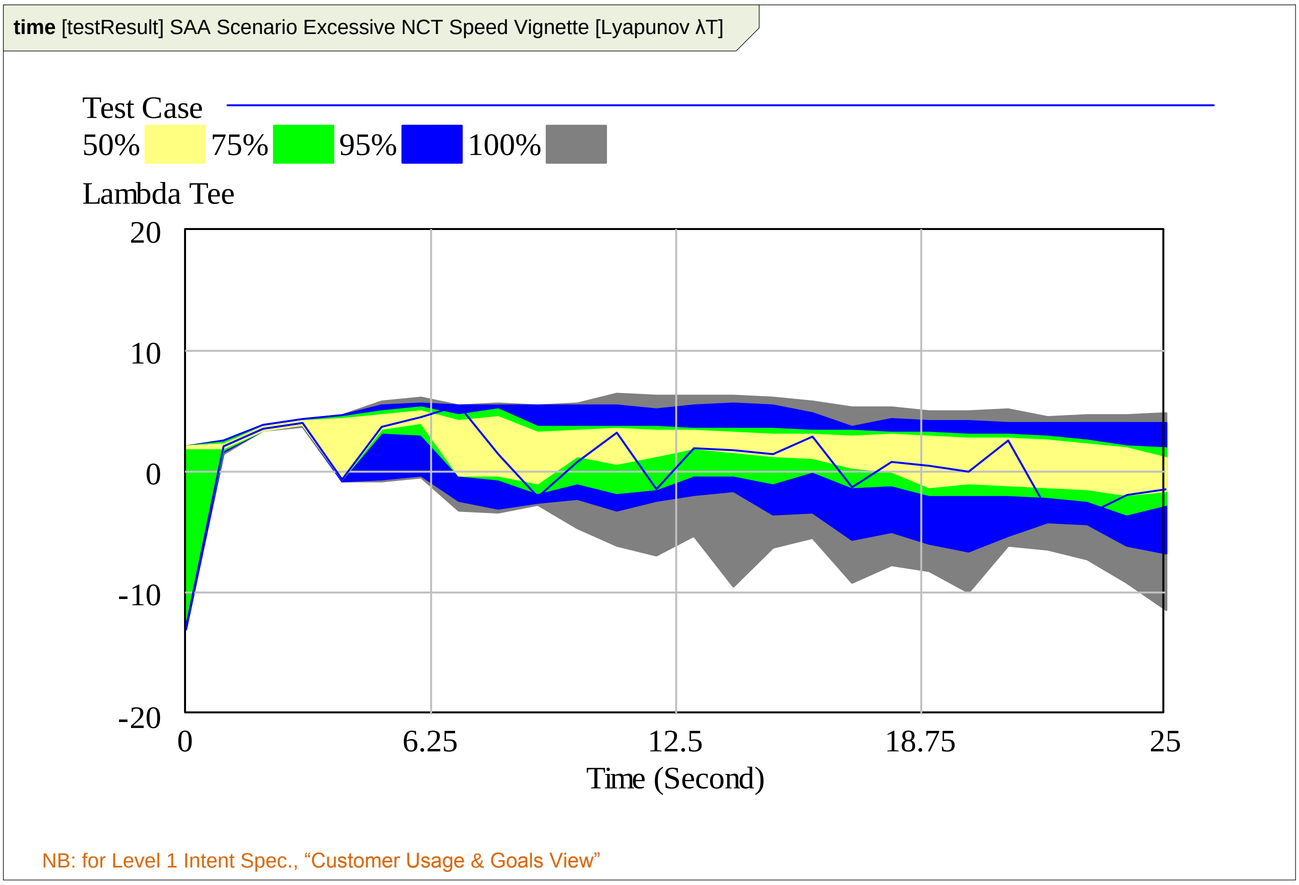


NB: for Level 1 Intent Spec., "Customer Usage & Goals View"

Figure A.12 Lyapunov Exponent Distributions at 25 seconds for TC 6

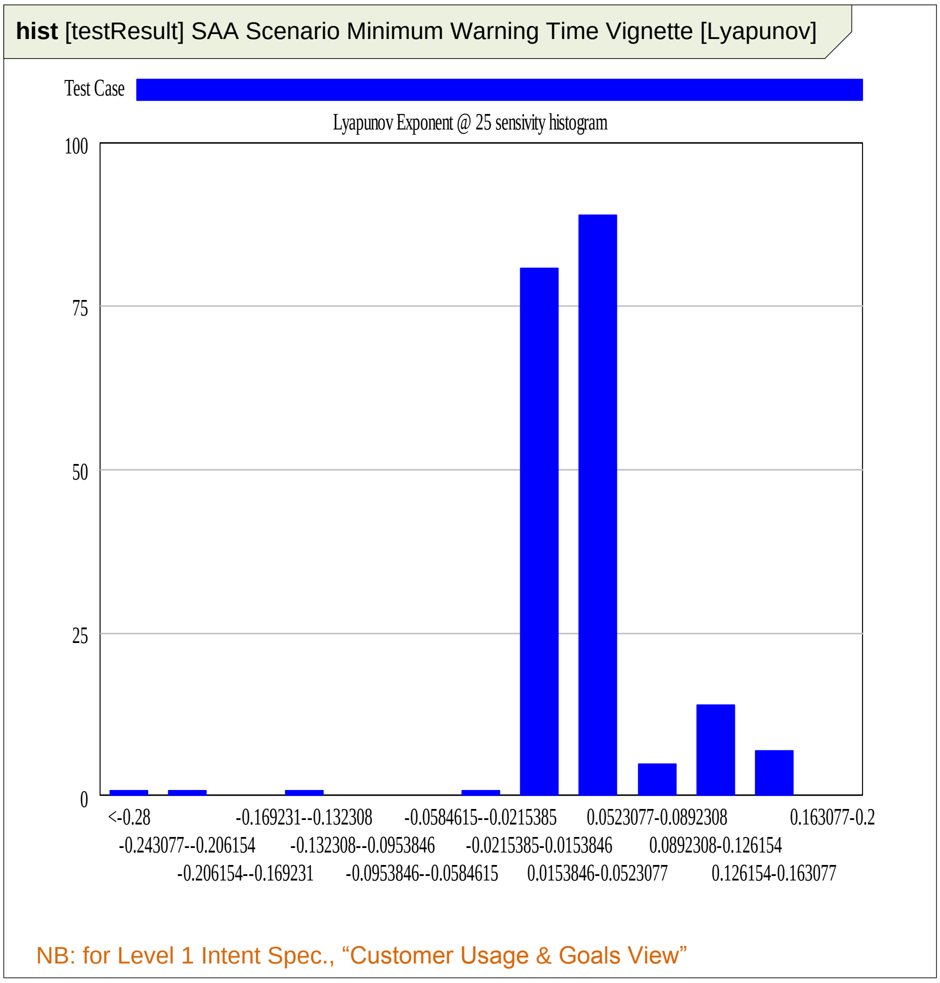


NB: for Level 1 Intent Spec., "Customer Usage & Goals View"



NB: for Level 1 Intent Spec., "Customer Usage & Goals View"

Figure A.13 Lyapunov Exponent Distributions at 25 seconds for TC 1



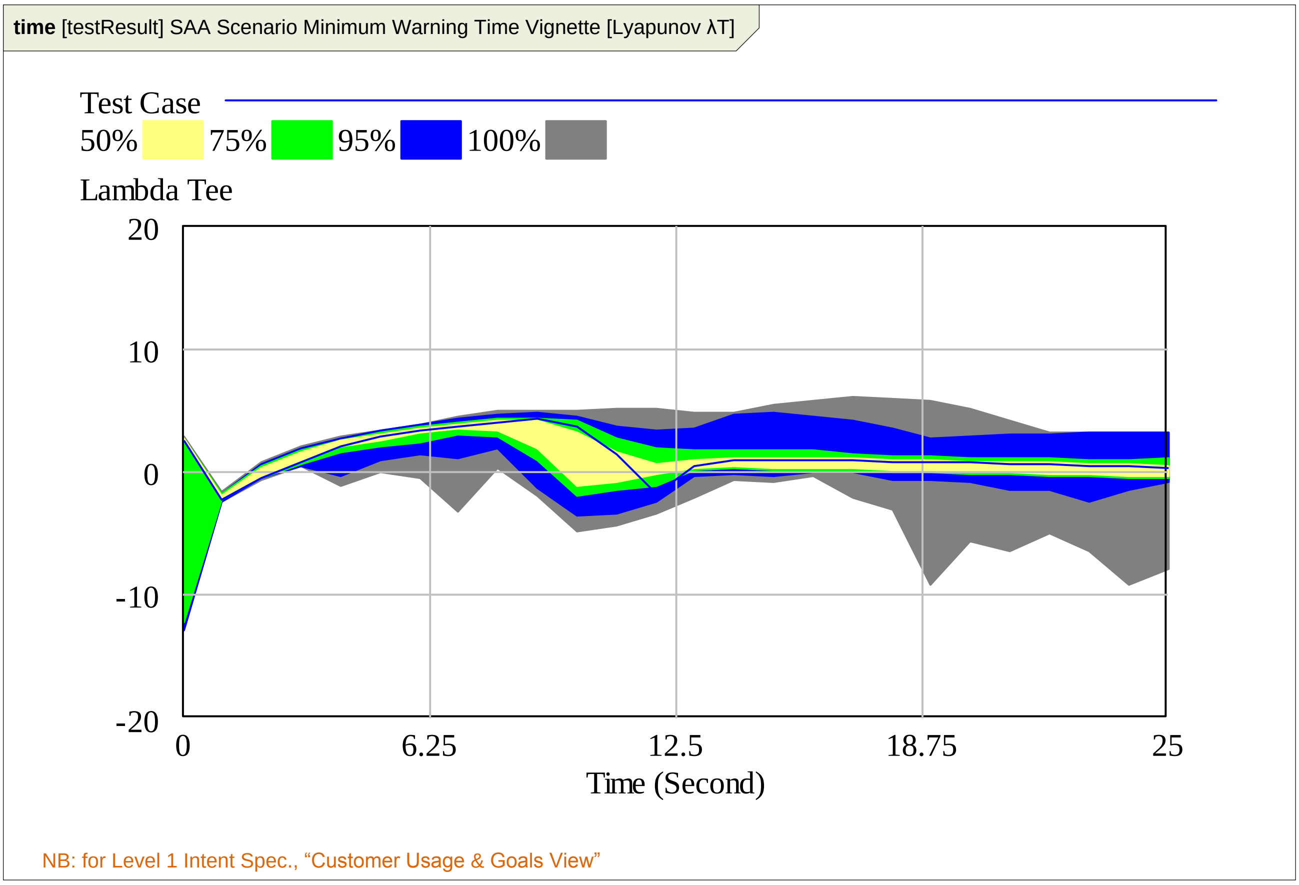
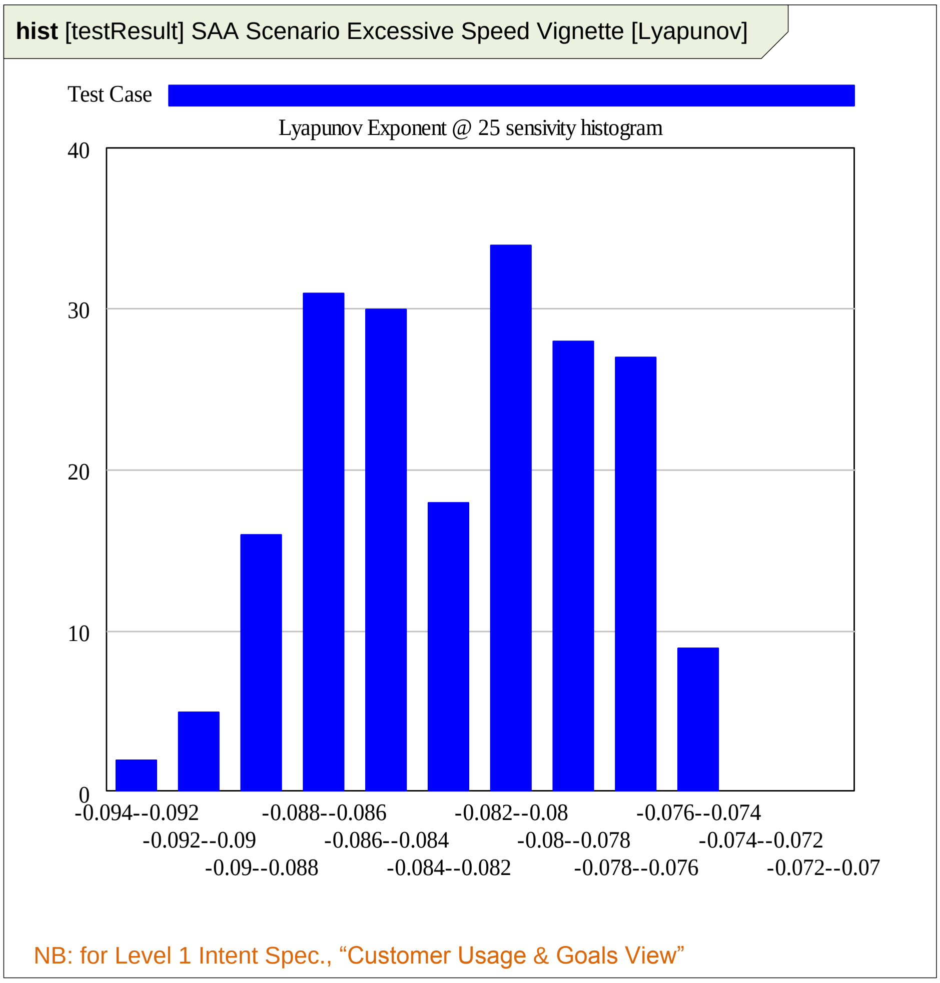
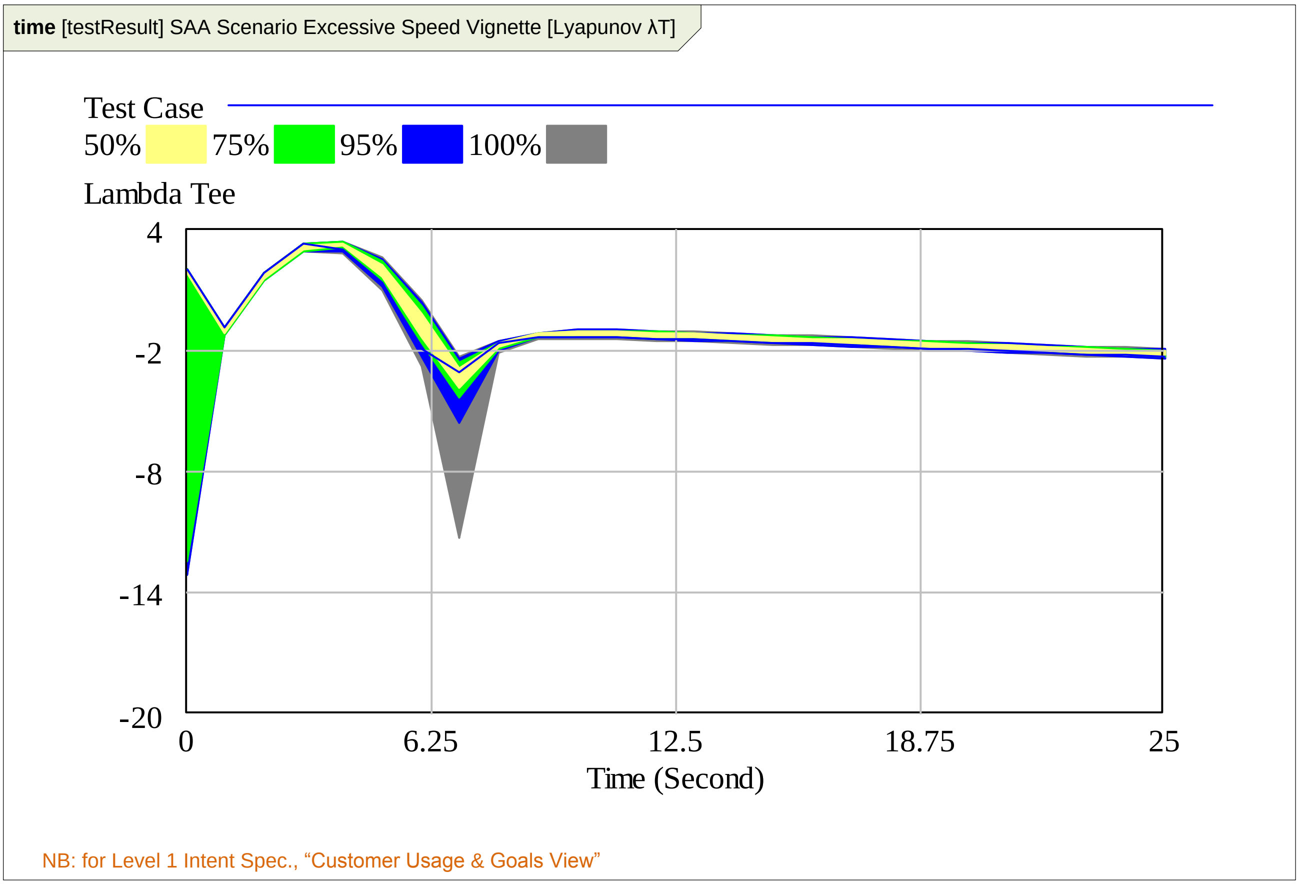


Figure A.14 Lyapunov Exponent Distributions at 25 seconds for TC 5

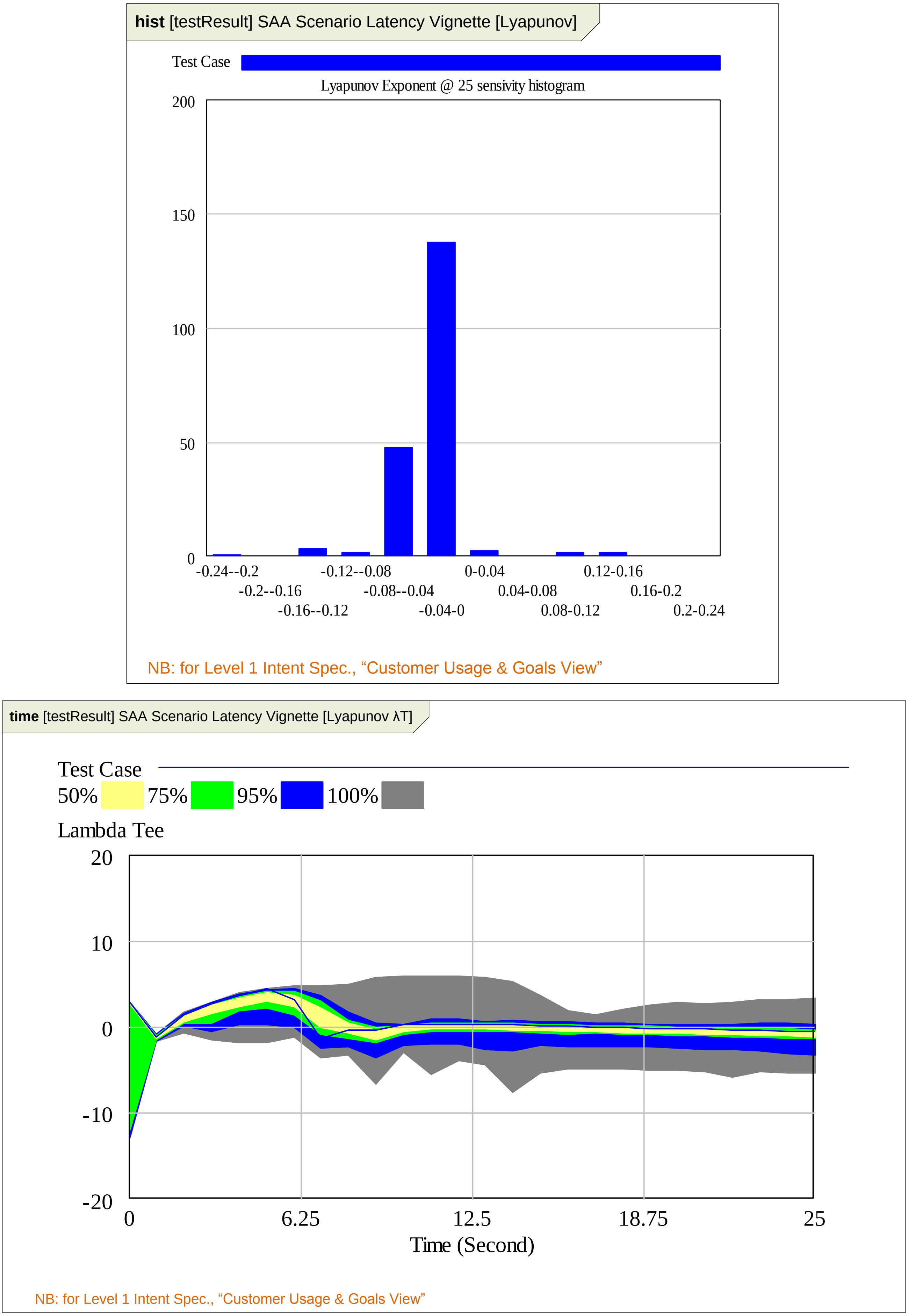

hist [testResult] SAA Scenario Excessive Speed Vignette [Lyapunov]

NB: for Level 1 Intent Spec., “Customer Usage & Goals View”


time [testResult] SAA Scenario Excessive Speed Vignette [Lyapunov λT]

NB: for Level 1 Intent Spec., “Customer Usage & Goals View”

Figure A.15 Lyapunov Exponent Distributions at 25 seconds for TC 3



hist [testResult] SAA Scenario Latency Vignette [Lyapunov]

NB: for Level 1 Intent Spec., “Customer Usage & Goals View”

time [testResult] SAA Scenario Latency Vignette [Lyapunov λT]

NB: for Level 1 Intent Spec., “Customer Usage & Goals View”

Figure A.16 Lyapunov Exponent Distributions at 25 seconds for TC 4

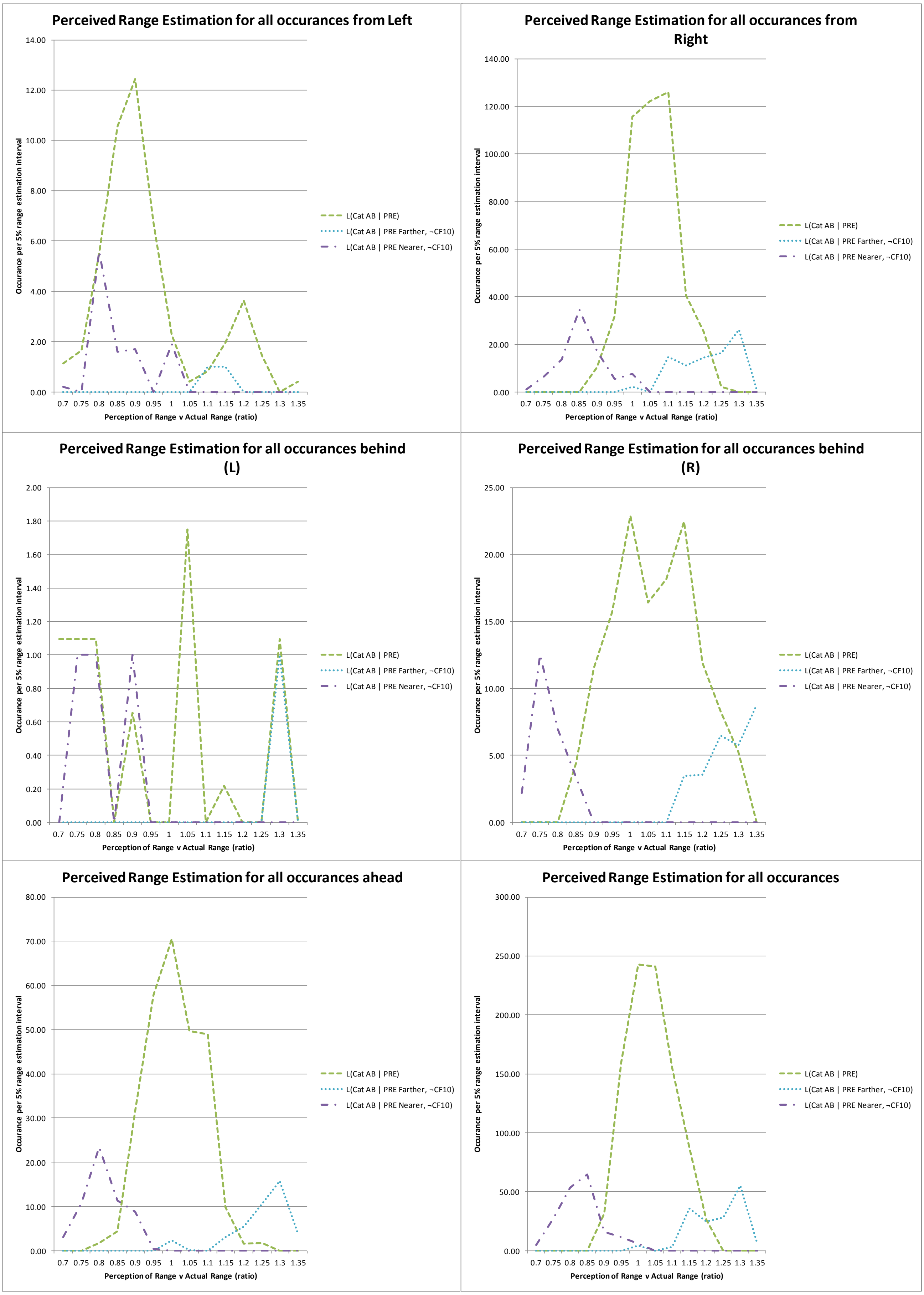
A.7. Bayesian Monte Carlo Responses

For the Bayesian Monte Carlo results, Figure A.17 and Figure A.18 present the observed distributions of sample averages formed with ten samples apiece, for varying combinations of just two of the deliberately injected control flaws – variations in the perceived range estimation (CF04) and in the latency or delay in commanding the emergency backup manoeuvre (CF10). Each of these distributions represents observations on the distribution of these parameters given the simultaneous presence of either a category A or B violation event. Therefore these distributions represent probability density functions for possible cause given category A or B events, *p(possible cause | CatAB),* with the overall rate representing the corresponding likelihood that a category A or B event occurs due to the possible cause, *L(CatAB | possible cause).*

By forming the distributions with sample averages, each distribution tends towards normal (Gaussian) as a consequence of the central limit theorem. Consequently, most of the larger distributions in these figures present an approximately normal shape, with the notable exception of the occurrences from “behind left” where there are very few *CatAB* events. Given the uniform distribution of values from which these causes are derived in the Monte Carlo simulations if we assume an absence of bias[7] the expectation would be that the mean and mode of the distribution tend towards the centre of the range, and be distributed evenly on either side of the mode. However, inspecting the top left plot (PRE for all occurrences from Left) reveals a very clear bi-modal distribution with the dip between the modes centred upon 1.05 (105% range estimation). Where these PRE values are then separated into “Farther” and “Nearer” groups, with the elimination of any potential “latency” control flaws, it would be expected that two distinct distributions are formed as can be seen, but again the modes for these are further apart than they should be for no bias.

Consequently, the behaviour of the model revealed through the Monte Carlo cases does provide a varying bias to the error value means responsible for defective behaviour; and in the case of the estimation error these values also fall naturally into Farther and Nearer Estimation Error defect groups. A particularly interesting feature of this is that the “Nearer” estimation errors, where the UAS estimates the NCT to be nearer than it actually is, produces a higher rate of *CatAB* events than “Farther” estimates – most apparent in the approaches from the left. This behaviour is also borne out in observations on the trained Bayesian network, where it was first observed. The implication of this is that in certain circumstances a “trigger-happy” approach to emergency behaviour is perhaps more likely to incur a constraint violation than deferred action. To some extent this is also borne out by the “latency” error behaviour (right-hand plot group) for the Left and Right approach cases each exhibiting a smaller peak at 1 second delay, and the main group modes occurring at between 3.4 and 3.8 seconds – again indicative of possible problems associated with a “trigger-happy” premature and overly responsive back-up behaviour.

[7] The distributions are not perfectly uniform, the latency population average is 2.469s in a 0-5 second window and the perceived range estimate population average is 1.044 in a window extending from 0.7 to 1.4.



Perceived Range Estimation for all occurances from Left

Perceived Range Estimation for all occurances from Right

Perceived Range Estimation for all occurances behind (L)

Perceived Range Estimation for all occurances behind (R)

Perceived Range Estimation for all occurances ahead

Perceived Range Estimation for all occurances

Figure A.17 Probability Density Functions for Sample Averages of Estimation

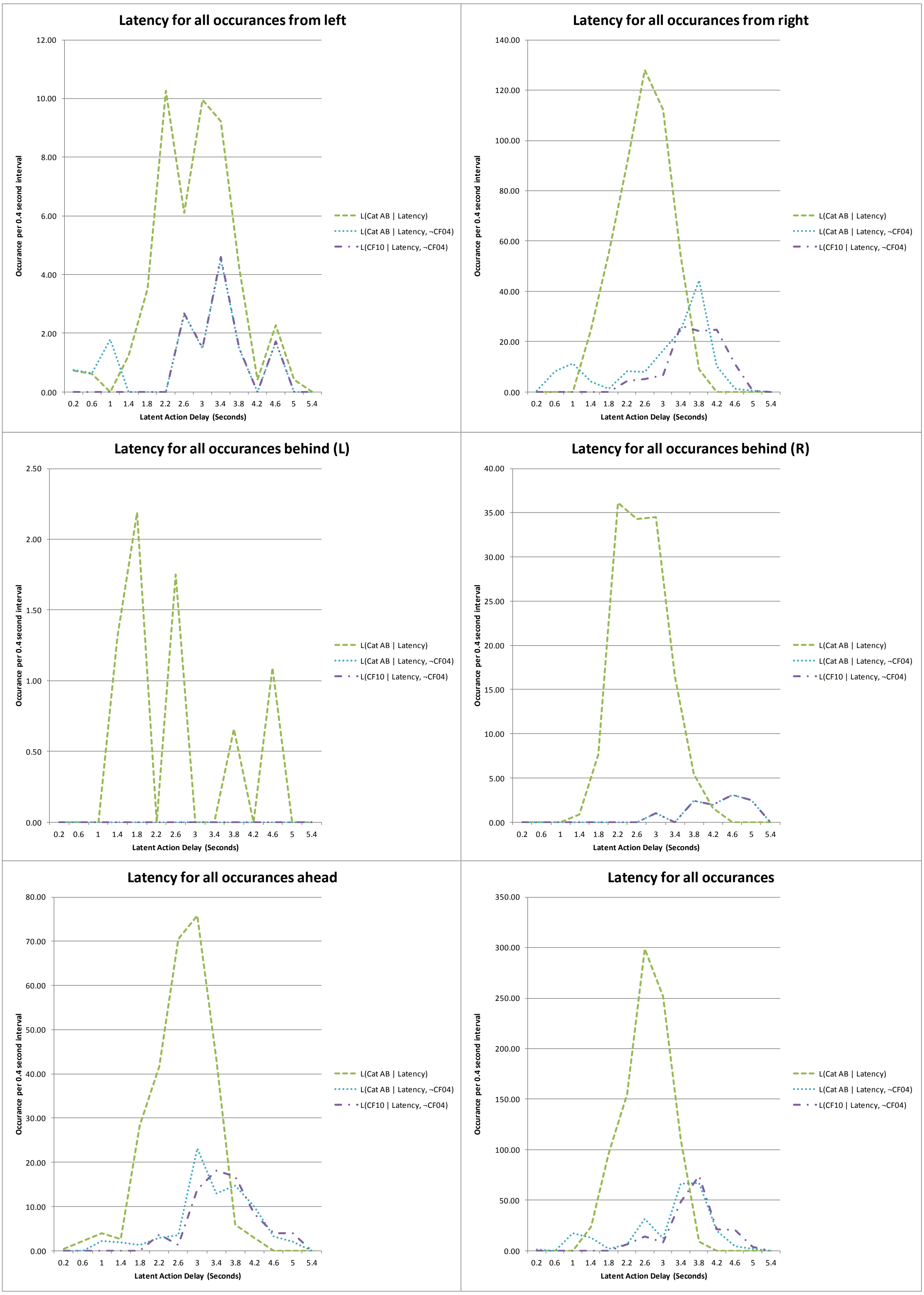


Figure A.18 Probability Density Functions for Sample Averages of Latency

The final step in this Bayesian Monte Carlo analysis is then to synthesise a probability density function surface correlating the likely probabilities that either a given estimation error (within CF04) and / or a given latency value (CF10) might cause a category A or B violation event. Unfortunately, as to whether a violation occurs as a consequence of only one of the causative factors, or in combination with the other, or indeed due to some other hidden cause, is unknown. Therefore the estimate has been composed as an average of two methods combining the maximum rates of occurrence of each potentially causative value with the other defect eliminated where detected (e.g. PRE selected only where the delay flaw CF10 is not detected), with the minimum rates accounting for where the two are believed to occur together – equations A.1 – A.4.

$$\boldsymbol{L_{XOR}} = \begin{bmatrix} \boldsymbol{L(PRE_{0.7}|\neg CF10) + L(Lat_{0.2}|\neg CF04)} & \cdots & \boldsymbol{L(PRE_{0.7}|\neg CF10) + L(Lat_{5.4}|\neg CF04)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{L(PRE_{1.35}|\neg CF10) + L(Lat_{0.2}|\neg CF04)} & \cdots & \boldsymbol{L(PRE_{1.35}|\neg CF10) + L(Lat_{5.4}|\neg CF04)} \end{bmatrix} \tag{A.1}$$

$$\boldsymbol{L_{AND}} = \begin{bmatrix} \sqrt{\boldsymbol{L(PRE_{0.7}|CF10) \cdot L(Lat_{0.2}|CF04)}} & \cdots & \sqrt{\boldsymbol{L(PRE_{0.7}|CF10) \cdot L(Lat_{5.4}|CF04)}} \\ \vdots & \ddots & \vdots \\ \sqrt{\boldsymbol{L(PRE_{1.35}|CF10) \cdot L(Lat_{0.2}|CF04)}} & \cdots & \sqrt{\boldsymbol{L(PRE_{1.35}|CF10) \cdot L(Lat_{5.4}|CF04)}} \end{bmatrix} \tag{A.2}$$

$$\boldsymbol{L_{MAX}} = \begin{bmatrix} \mathbf{max(L(PRE_{0.7}|\neg CF10), L(Lat_{0.2}|\neg CF04))} & \cdots & \mathbf{max(L(PRE_{0.7}|\neg CF10), L(Lat_{5.4}|\neg CF04))} \\ \vdots & \ddots & \vdots \\ \mathbf{max(L(PRE_{1.35}|\neg CF10), L(Lat_{0.2}|\neg CF04))} & \cdots & \mathbf{max(L(PRE_{1.35}|\neg CF10), L(Lat_{5.4}|\neg CF04))} \end{bmatrix}$$

(A.3)

$$\boldsymbol{L_{MIN}} = \begin{bmatrix} \boldsymbol{min(L(PRE_{0.7}|CF10), L(Lat_{0.2}|CF04))} & \cdots & \boldsymbol{min(L(PRE_{0.7}|CF10), L(Lat_{5.4}|CF04))} \\ \vdots & \ddots & \vdots \\ \boldsymbol{min(L(RE_{1.35}|CF10), L(Lat_{0.2}|CF04))} & \cdots & \boldsymbol{min(L(PRE_{1.35}|CF10), L(Lat_{5.4}|CF04))} \end{bmatrix} \tag{A.4}$$

Finally the likelihood surfaces for each direction of approach are calculated as shown in Figure A.19. These surfaces provide an estimation of the likelihood that category A or B violations occur as a consequence of a particular combination of perceived range estimation and latency values, within a rectangular interval (*ΔPRE* = 0.05, *ΔLat* = 0.4s); where *S* is the number of strips (14) across the plot and *CatAB* is the number of all category A or B violations within each set – equation A.5. In exploring this approach, it was observed that the numeric exclusive OR / AND method, and the MAX / MIN method, both produced very similar surfaces after renormalisation, where the option was then taken of blending the two, which together produces a slightly smoother and easier to interpret result.

$$\begin{aligned} &\boldsymbol{L(CatAB|PRE, Lat)} \\ &\qquad = \frac{\boldsymbol{S \cdot CatAB}}{\mathbf{2}} \\ &\qquad \cdot \left\{ \frac{\boldsymbol{L_{XOR} + L_{AND}}}{\sum_{\boldsymbol{PRE}=0.7,0.05}^{1.35} \sum_{\boldsymbol{Lat}=0.2,0.4}^{5.4} \left(\boldsymbol{L_{XOR_{PRE,Lat}} + L_{AND_{PRE,Lat}}}\right)} \right. \\ &\qquad + \left. \frac{\boldsymbol{L_{MAX} + L_{MIN}}}{\sum_{\boldsymbol{PRE}=0.7,0.05}^{1.35} \sum_{\boldsymbol{Lat}=0.2,0.4}^{5.4} \left(\boldsymbol{L_{MAX_{PRE,Lat}} + L_{MIN_{PRE,Lat}}}\right)} \right\} \end{aligned}$$

(A.5)

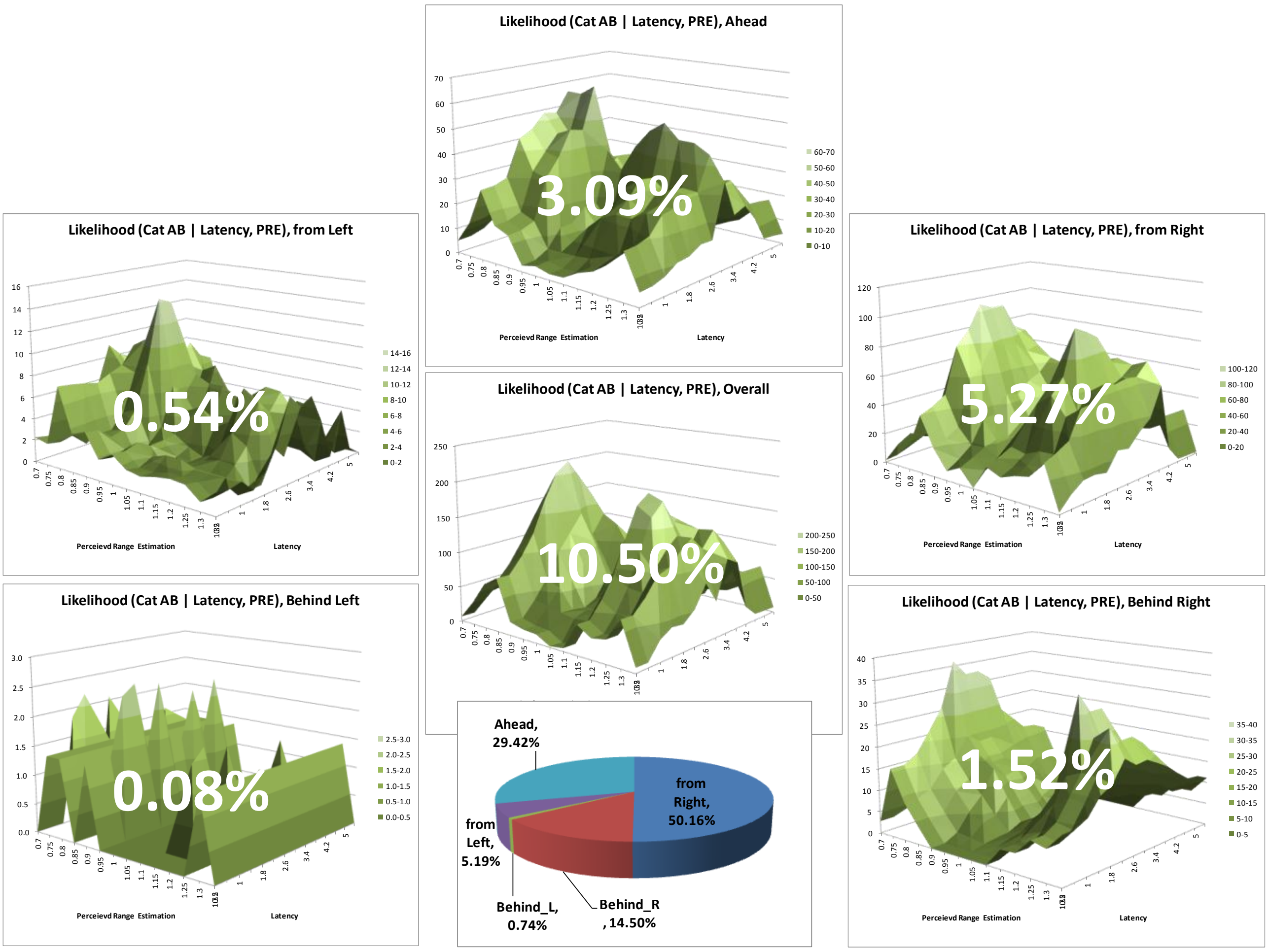


Figure A.19 Vignette Bayesian Monte Carlo parameter sensitivity surfaces

So as to make clear the relative scale of each surface, the prominent percentage values superimposed over the surface plots, in Figure A.19 above, represent the overall rate at which category A or B violations occur within the particular sector. As the only faults that are deliberately and knowingly injected into these vignettes are the variations in perceived estimation of range and latency, this violation rate might be taken as the overall likelihood that these violations occur due **only** to the defective values contained within the range of these variables. As the violation rate obviously varies also according to the direction of approach then clearly this is not true. The behaviour of the system is also affected by speed and direction, and perhaps embodies other hidden and unknown defects – certain speed ranges might themselves be considered defective. Given also that we do not know exactly where a particular boundary, in terms of speed or direction, lies within the model it may be very difficult to fully separate these effects. For the purposes of this study, the larger value is taken as the likelihood that over the entire range of speed, estimation and latency combinations, as explored by the Monte Carlo simulations, a violation event may occur within the given range of directions of approach for that sector. Whereas, assuming that the overall rate within any sector accommodates the speed and direction elements, then the surface under this is taken to represent variations in this rate due to estimation error and latency defects alone – which may not be entirely true as the exact location of the directional boundary is unknown, and even that assumes that there is a clear boundary and that there are no other unknown defects. Even so, a purely mechanical use of these surfaces might extract likelihood values for use in the calculation of risks related to the more narrowly defined combinations of defects as occur in the special cases (TC1 – 6).

A.8. Correlations between Lyapunov Exponent Values & Airprox Violations

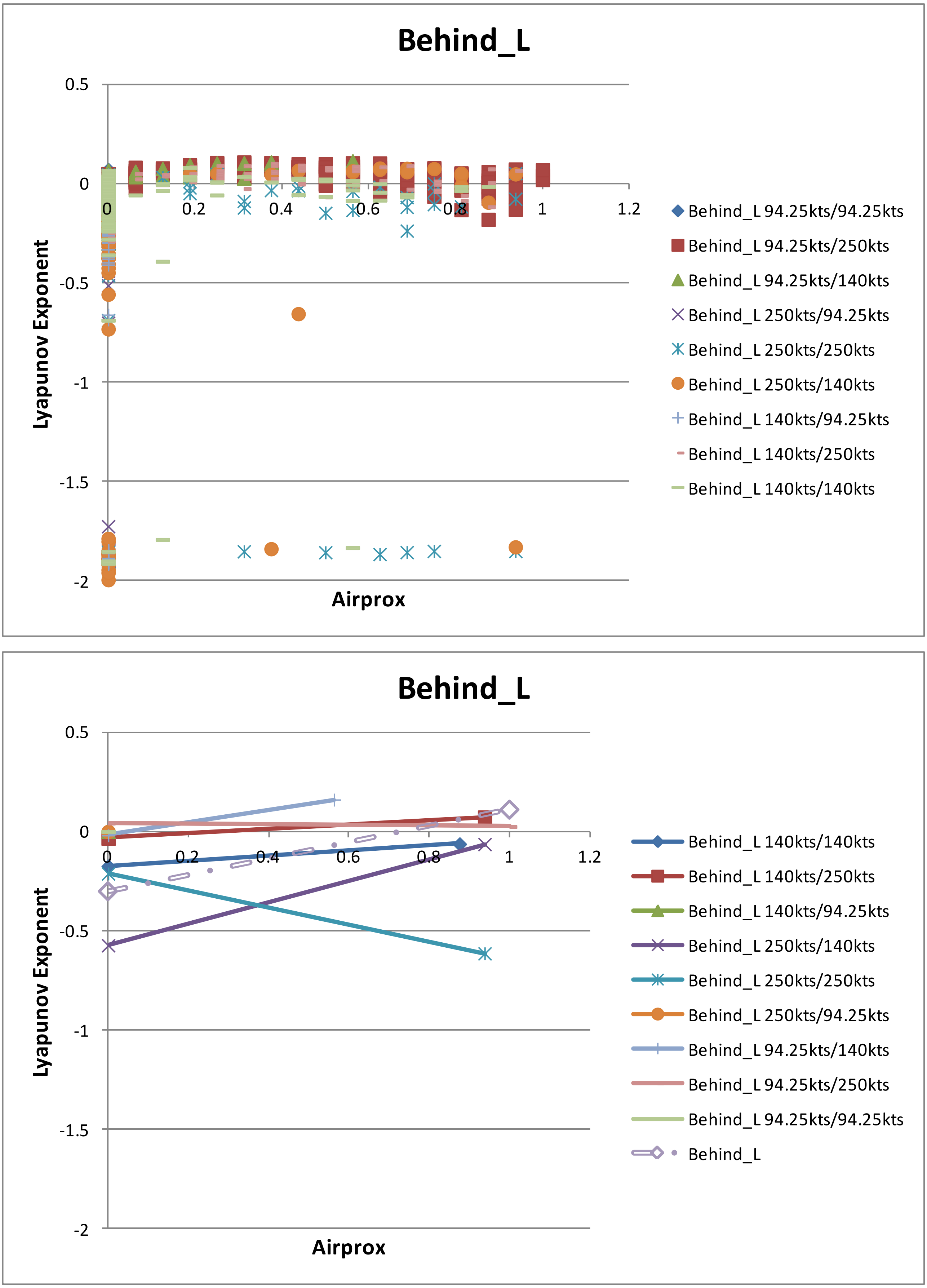
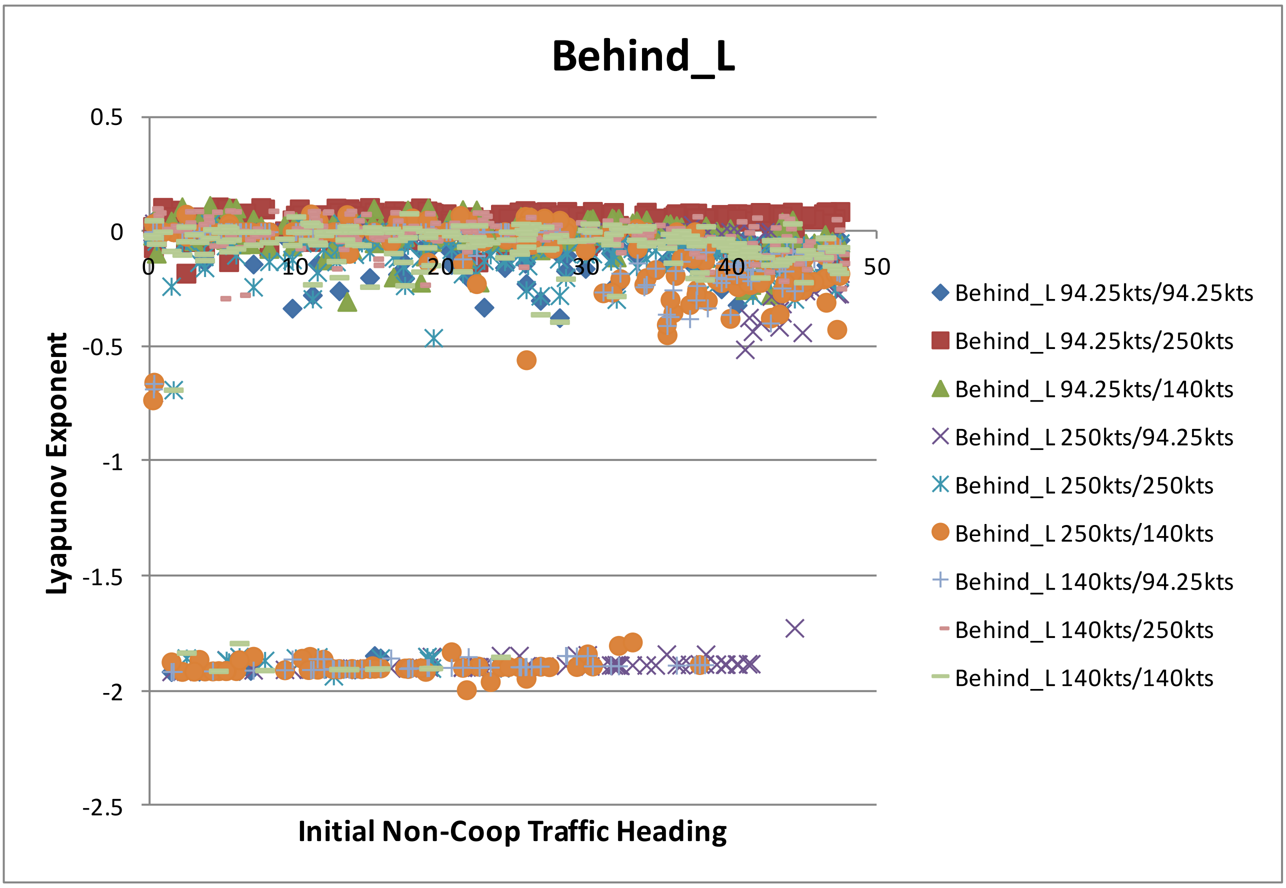


Figure A.20 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TC 9



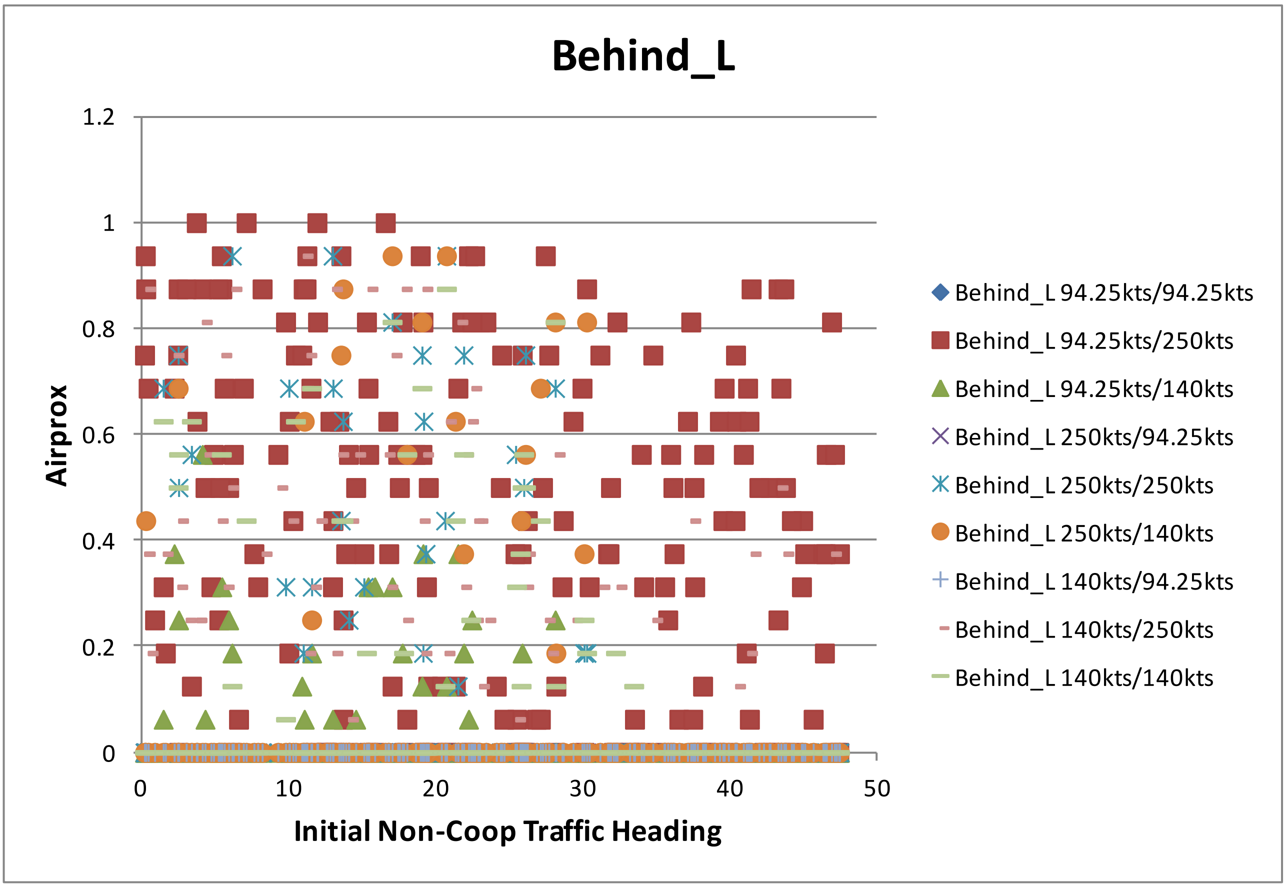
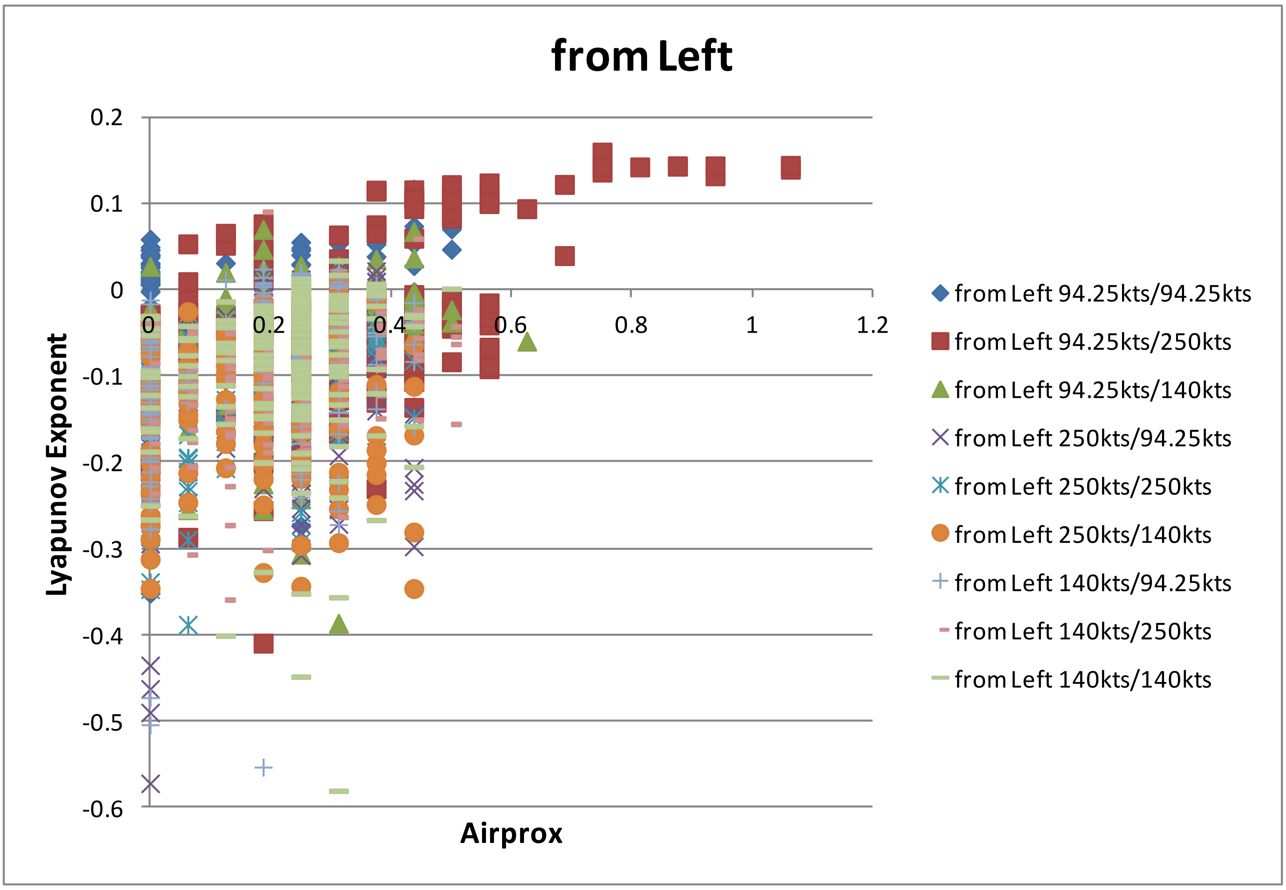


Figure A.21 Scatter Plots with respect to Initial Relative Heading including TC 9



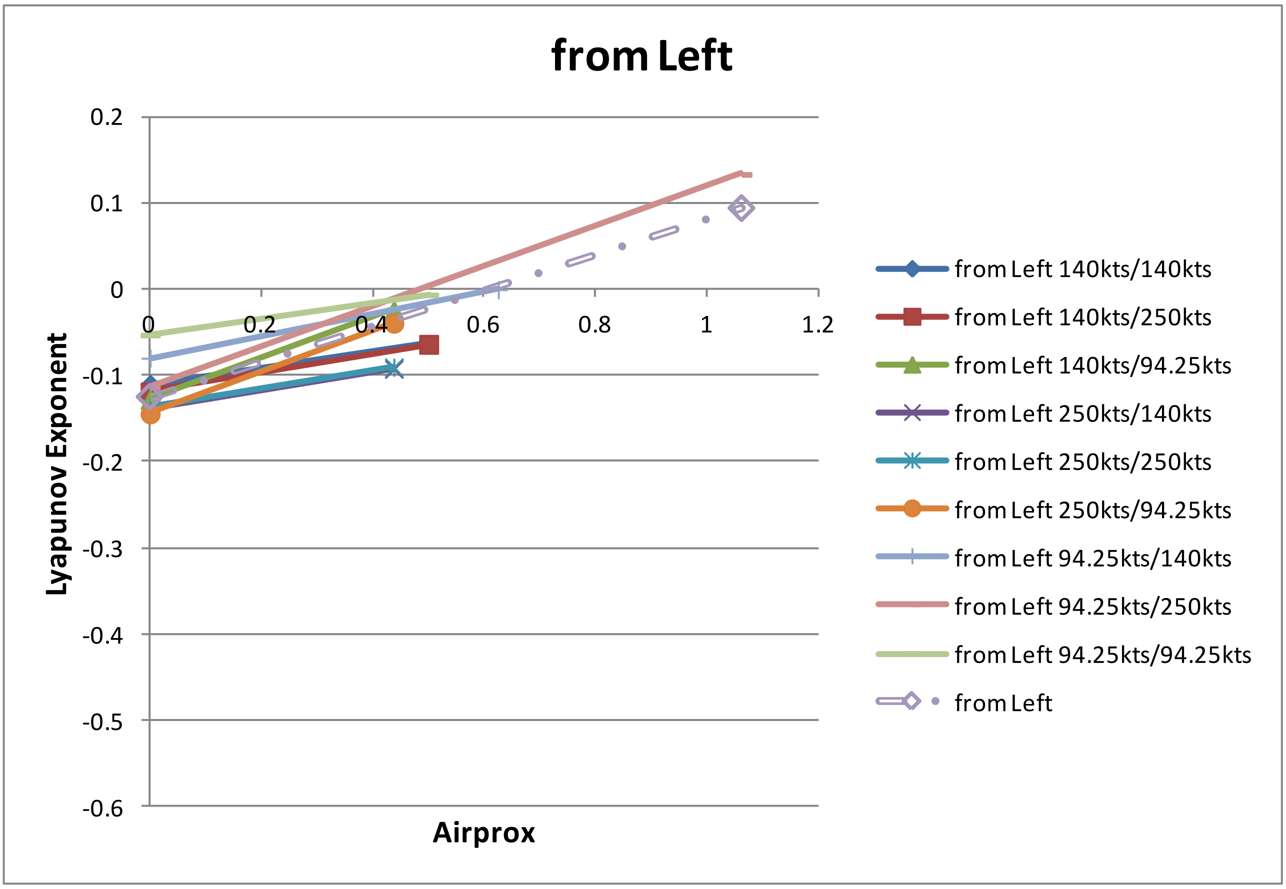
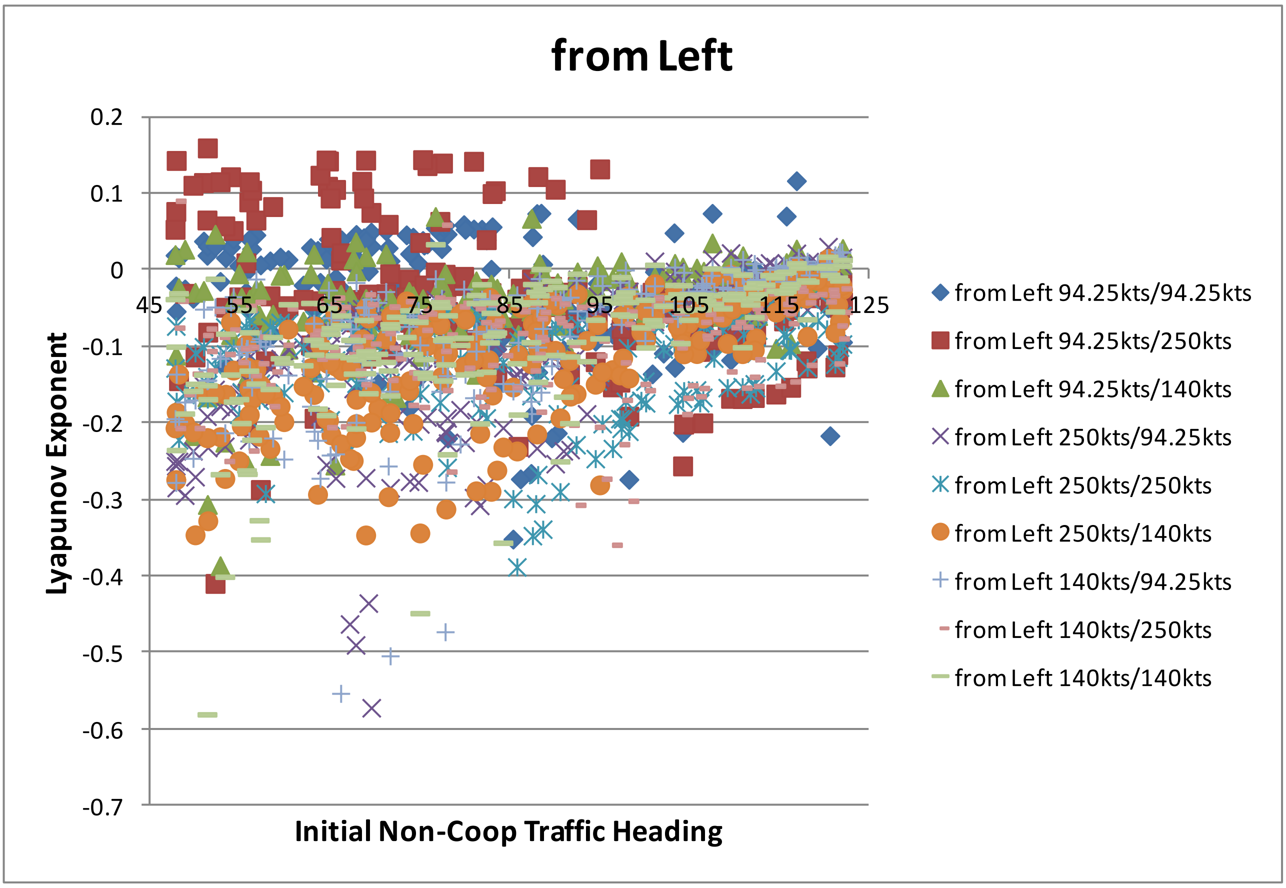


Figure A.22 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TC 11



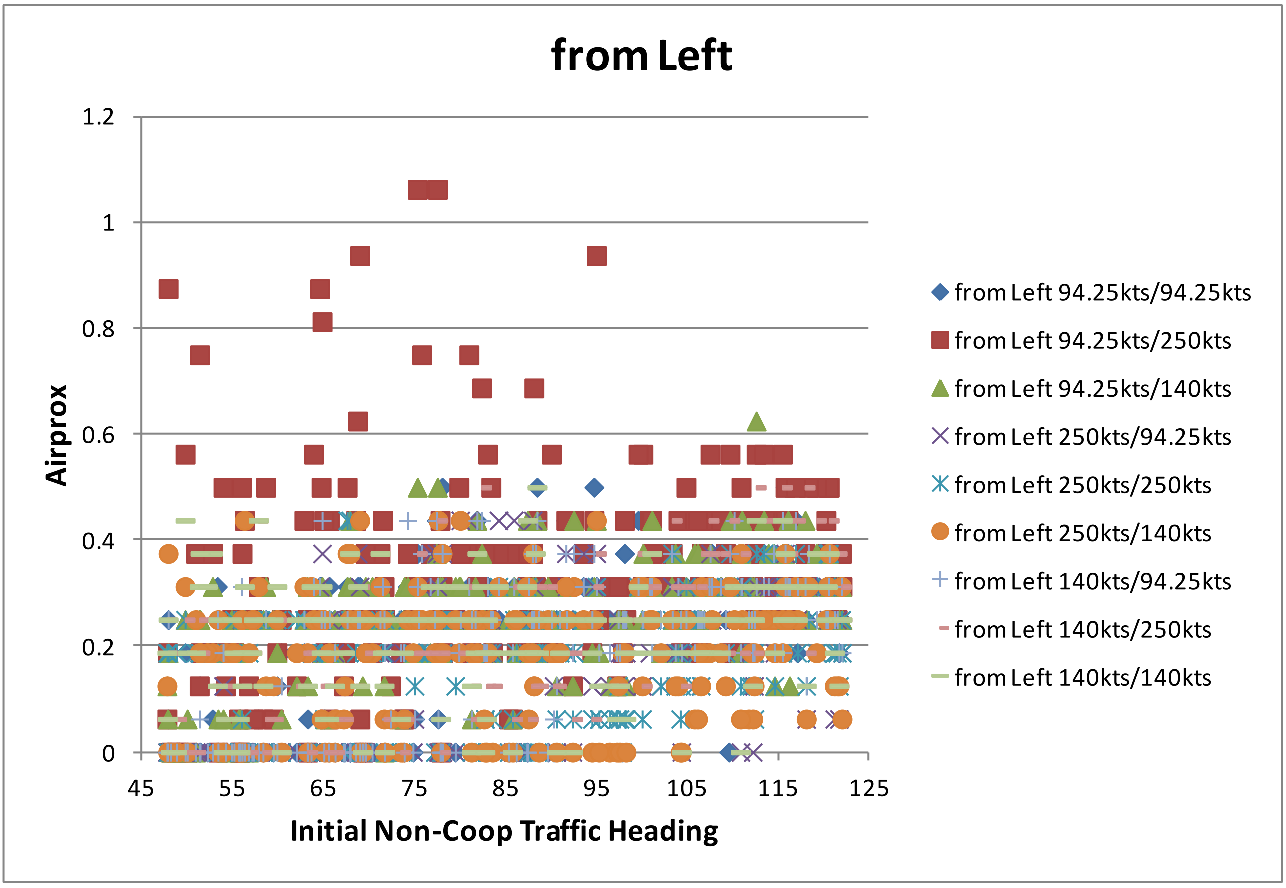
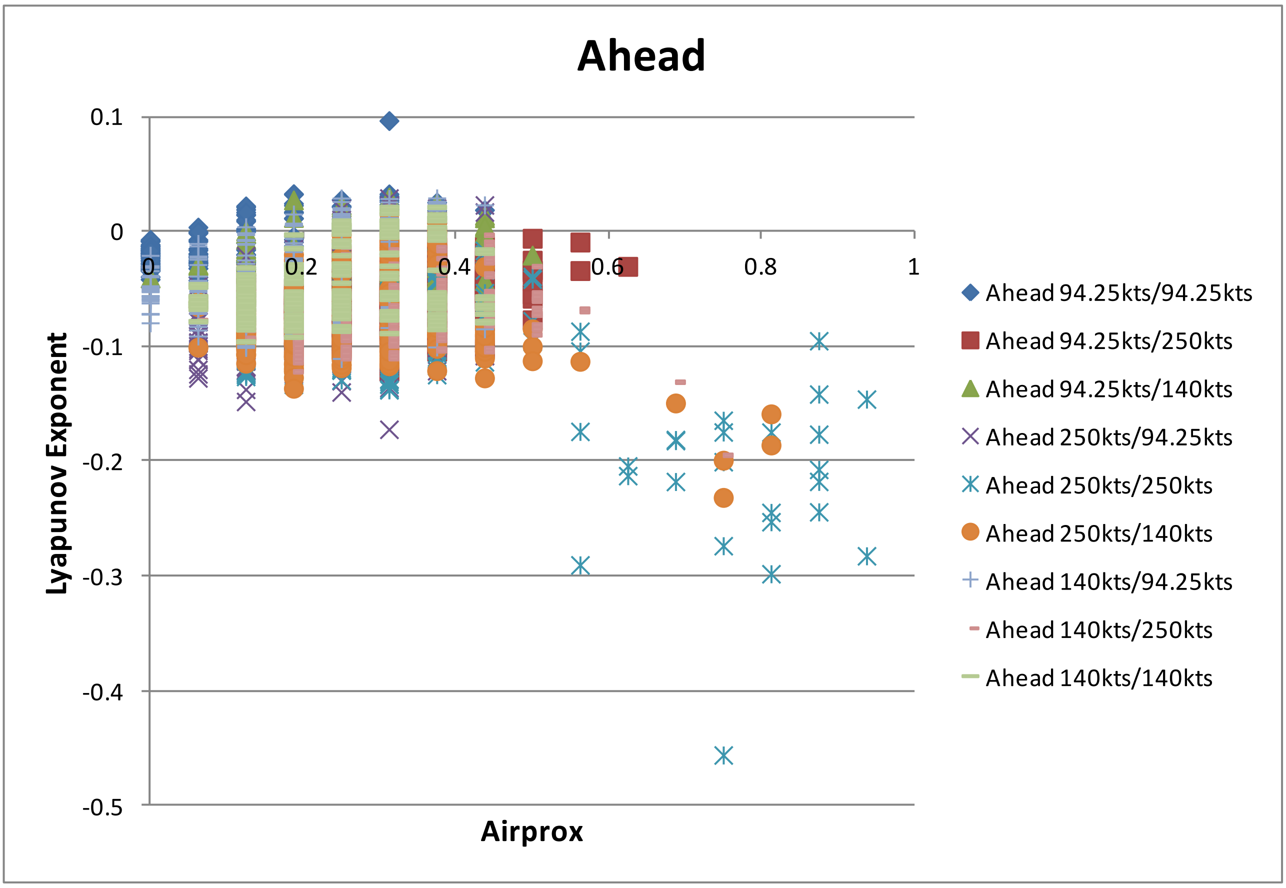


Figure A.23 Scatter Plots with respect to Initial Relative Heading including TC 9



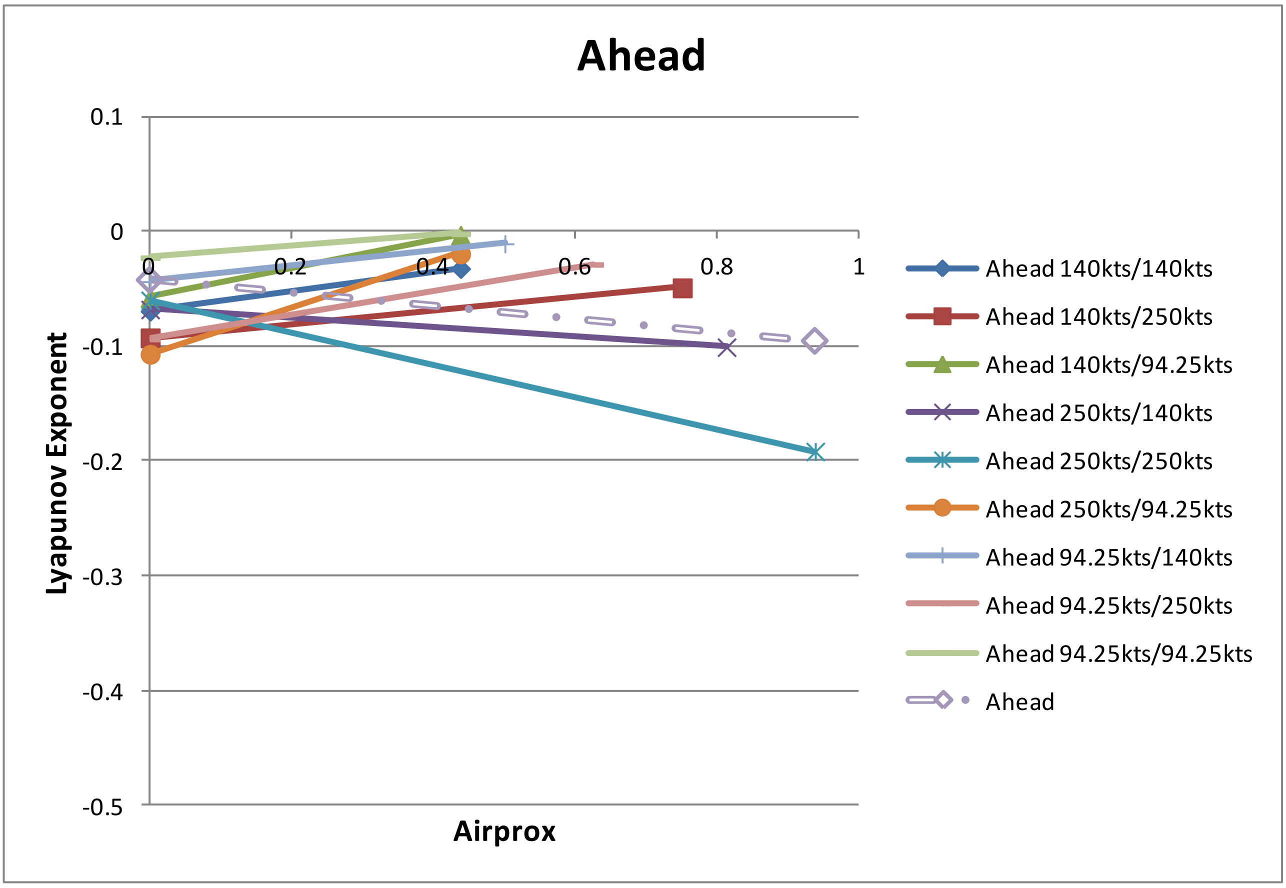
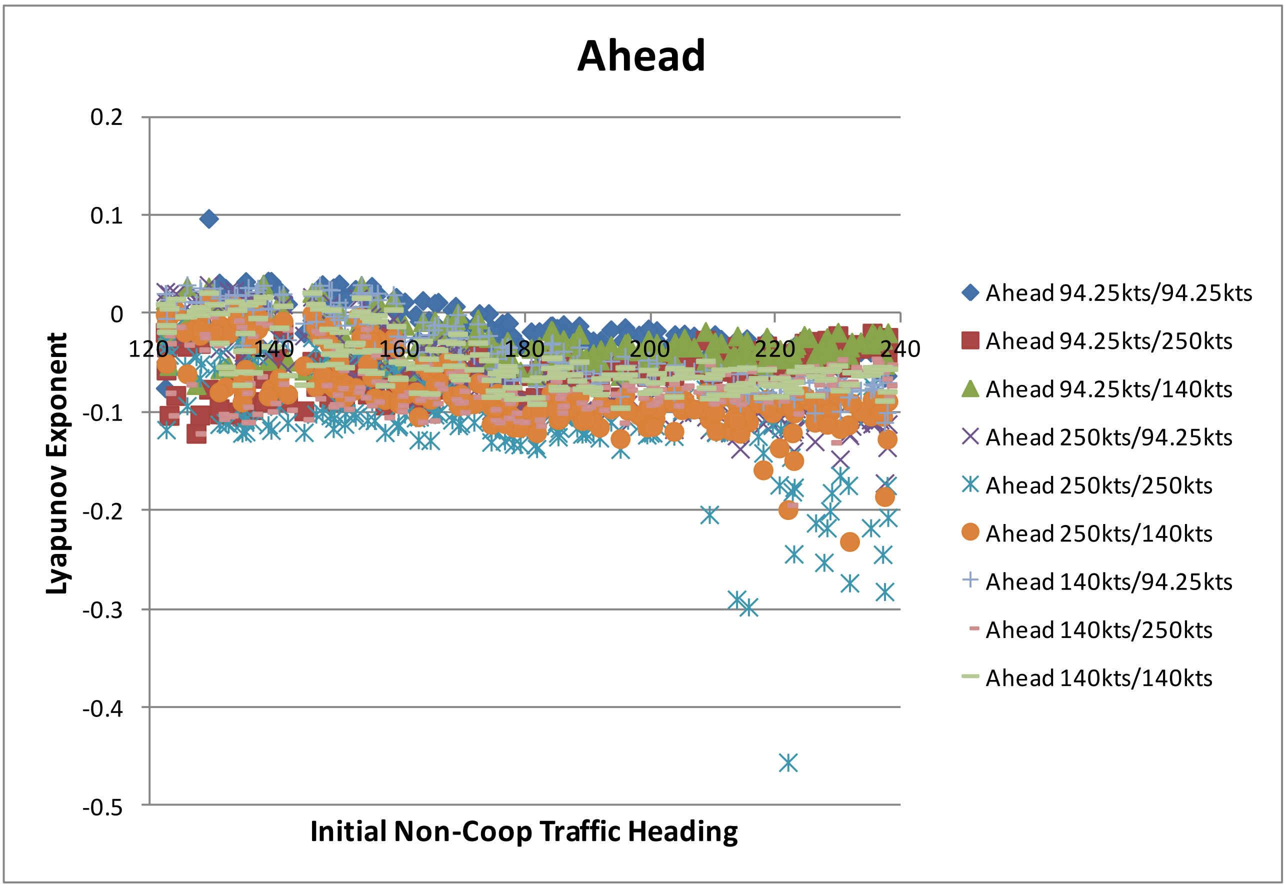


Figure A.24 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TCs 7 & 8



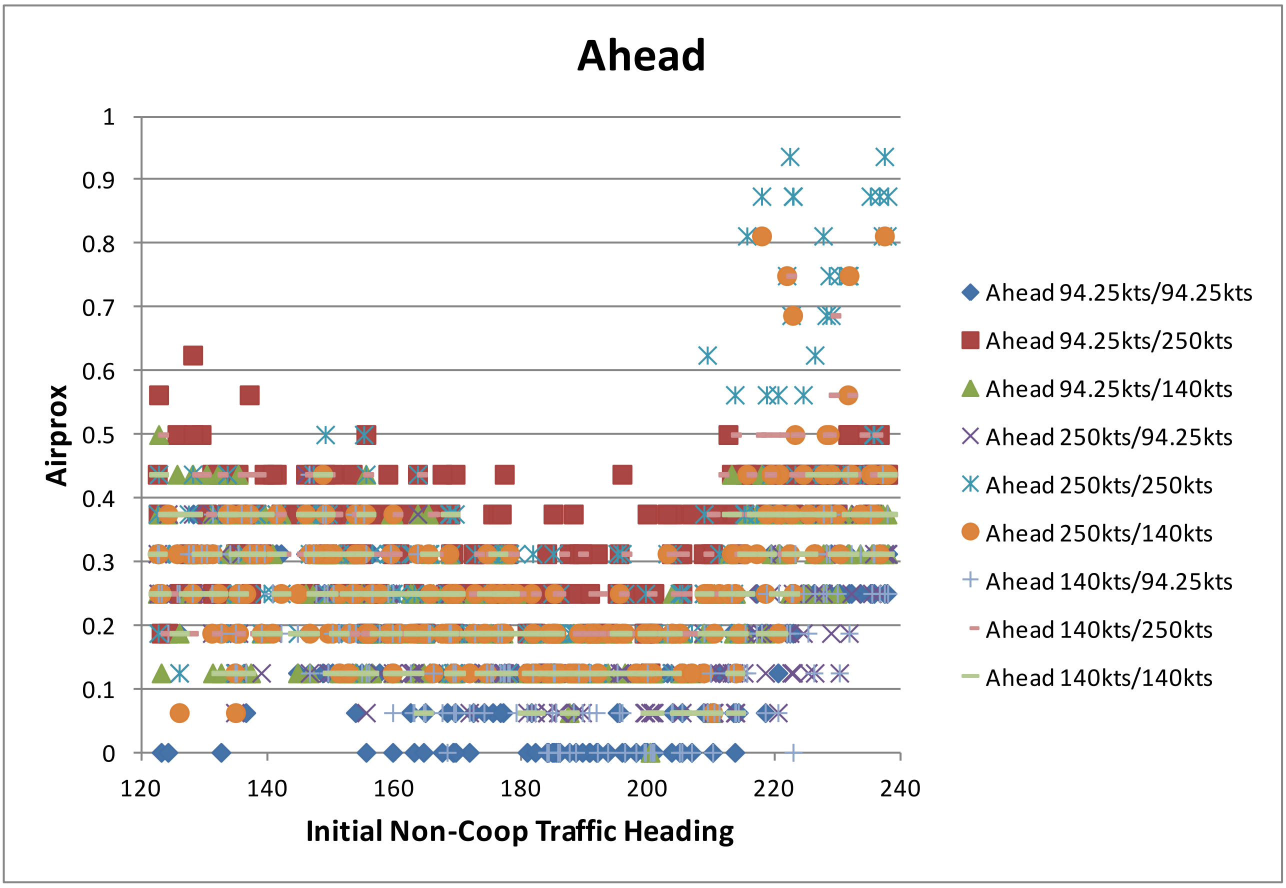
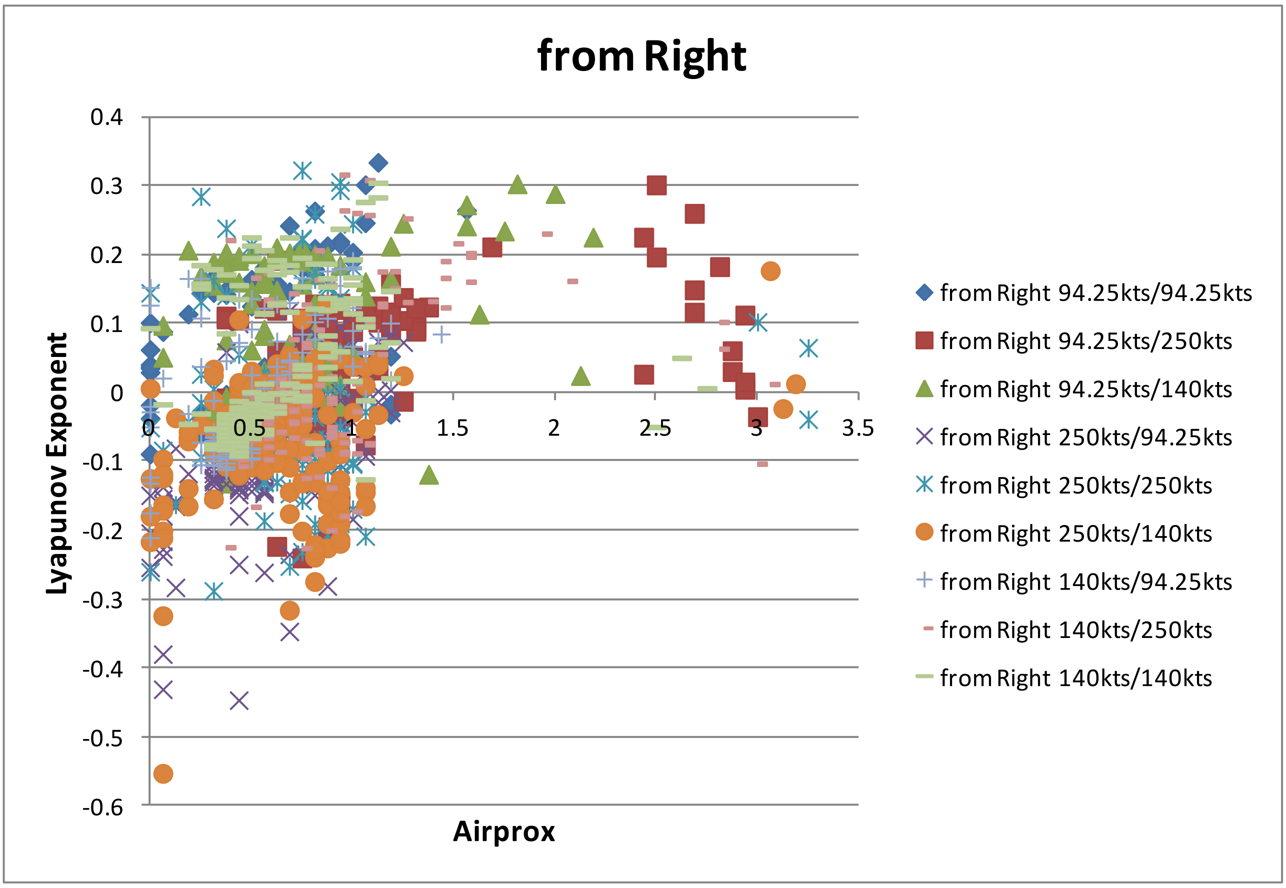


Figure A.25 Scatter Plots with respect to Initial Relative Heading including TCs 7 & 8



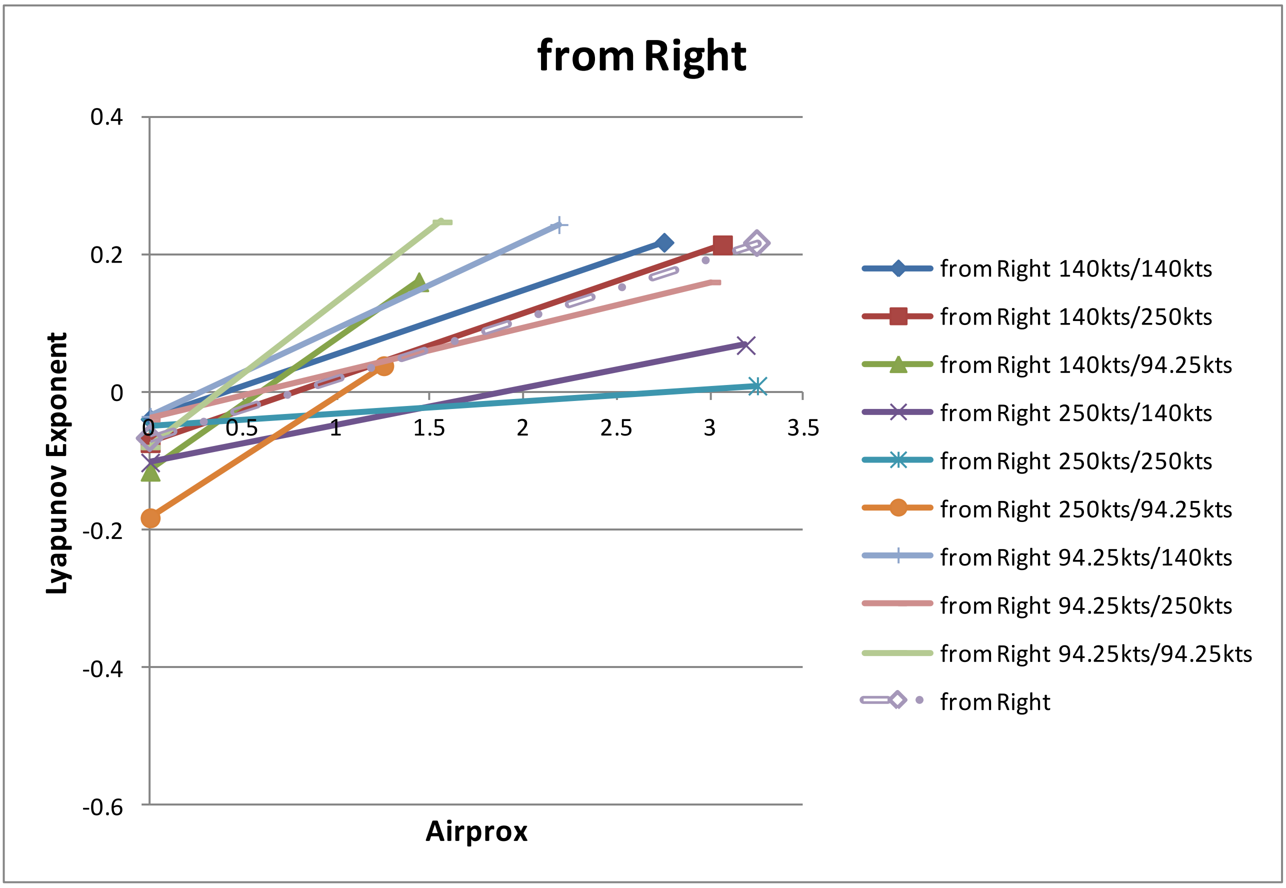


Figure A.26 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TC 12

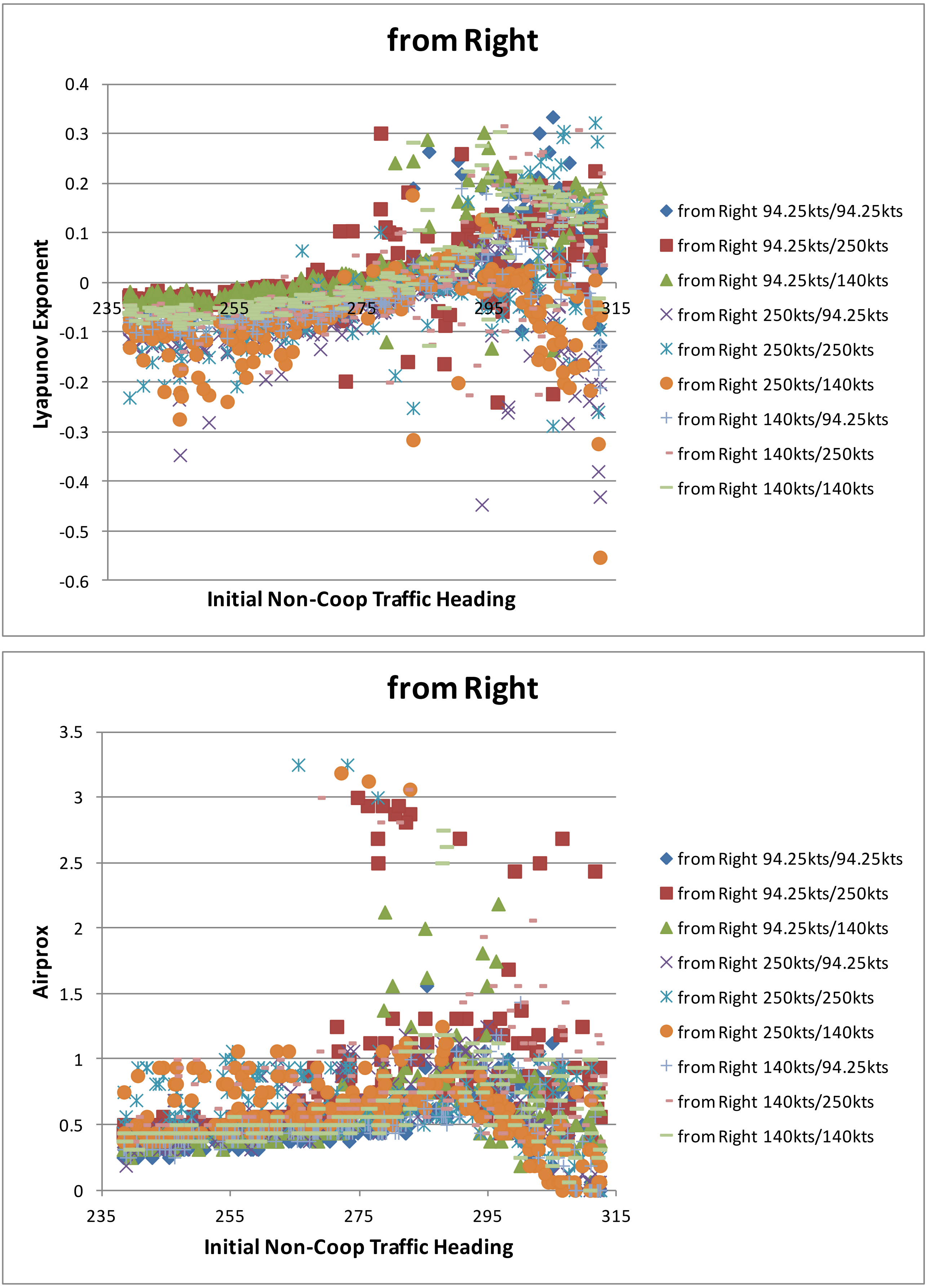
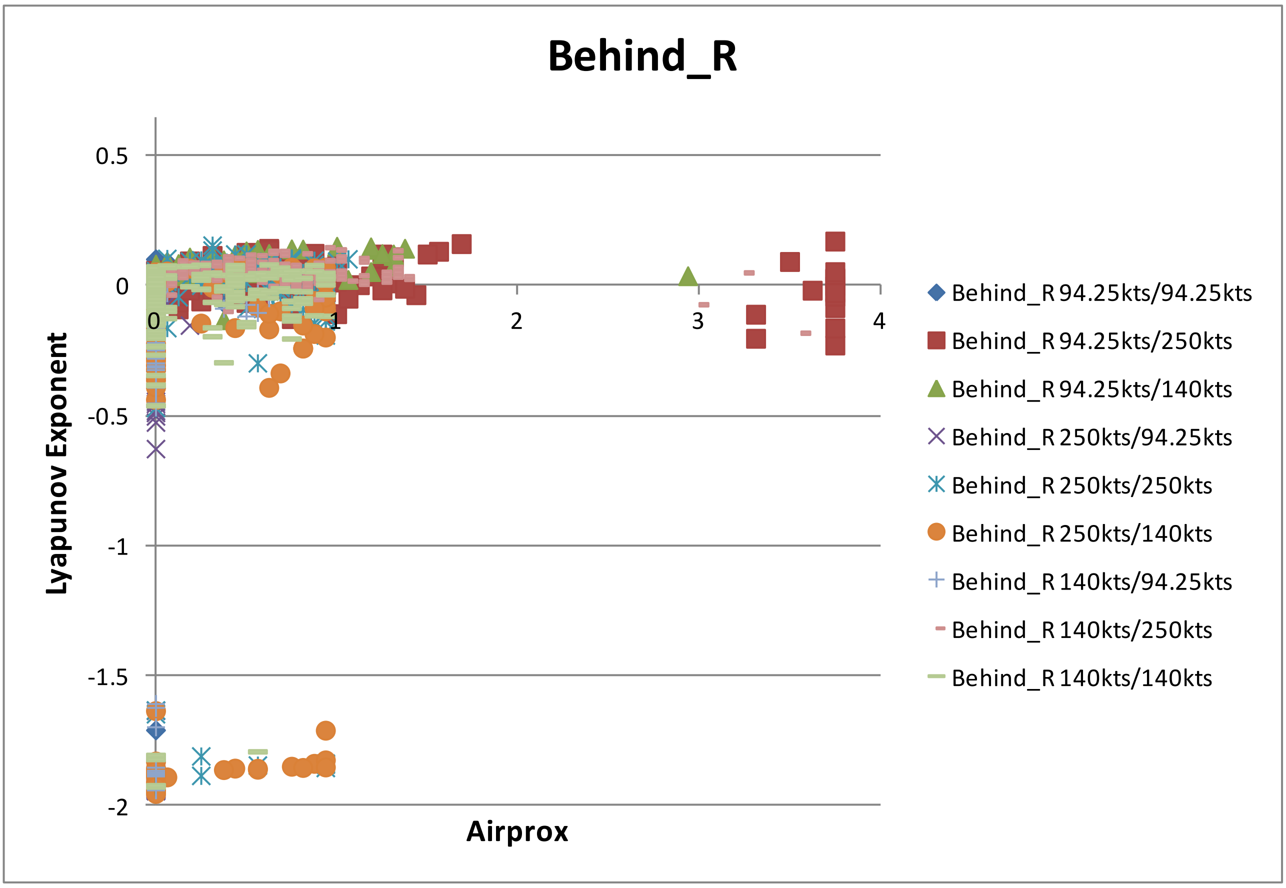


Figure A.27 Scatter Plots with respect to Initial Relative Heading including TC 12



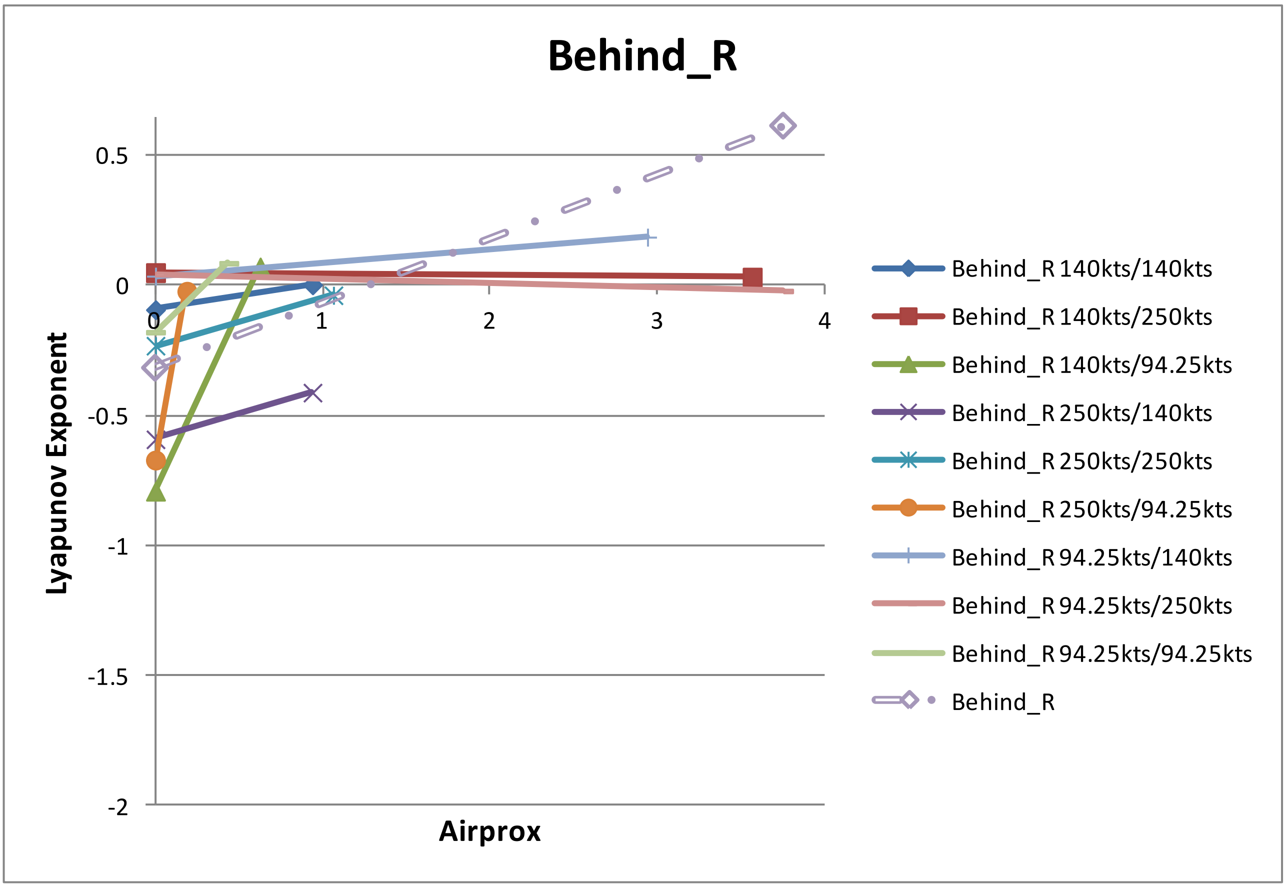
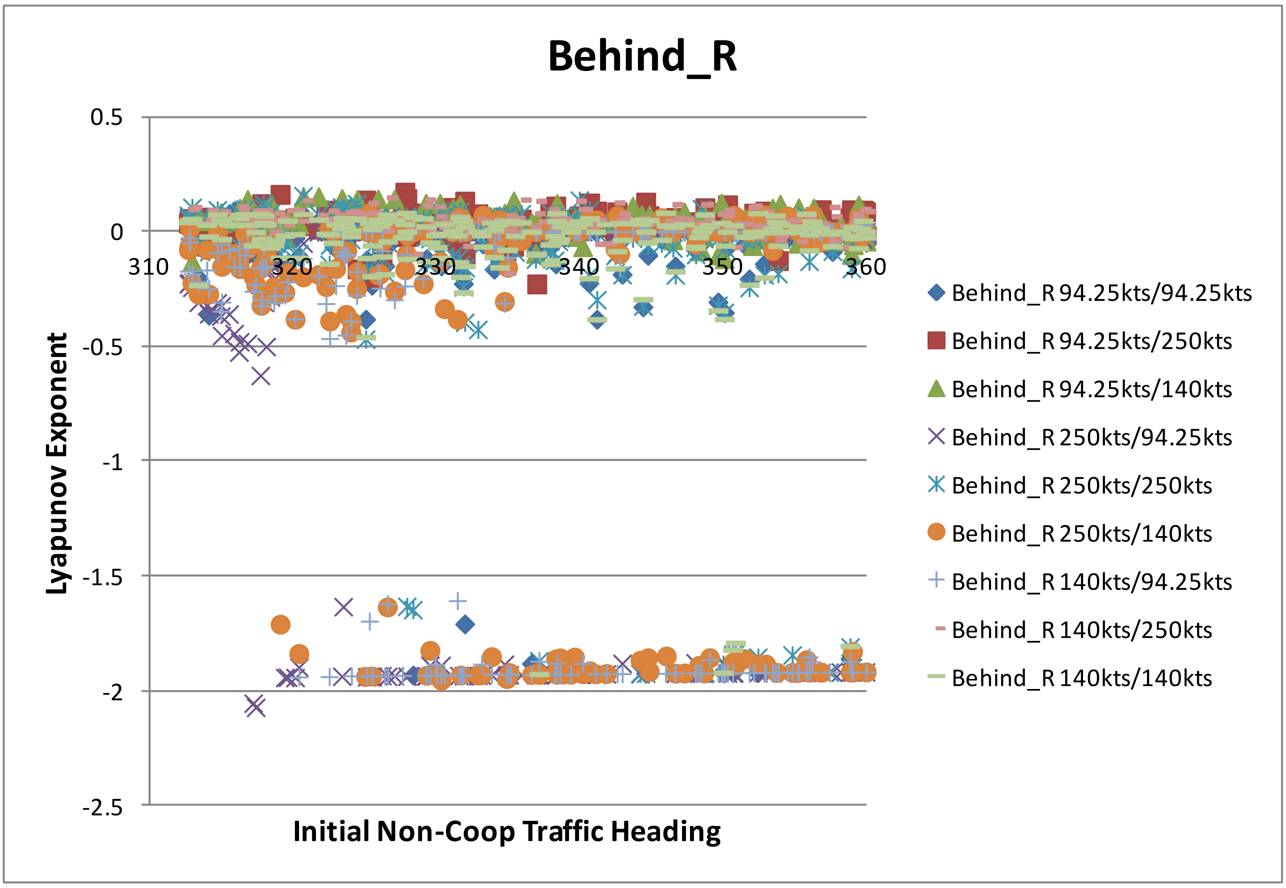


Figure A.28 Lyapunov v Airprox Correlation as Scatter Plot & Linear Regressions inc. TC 10



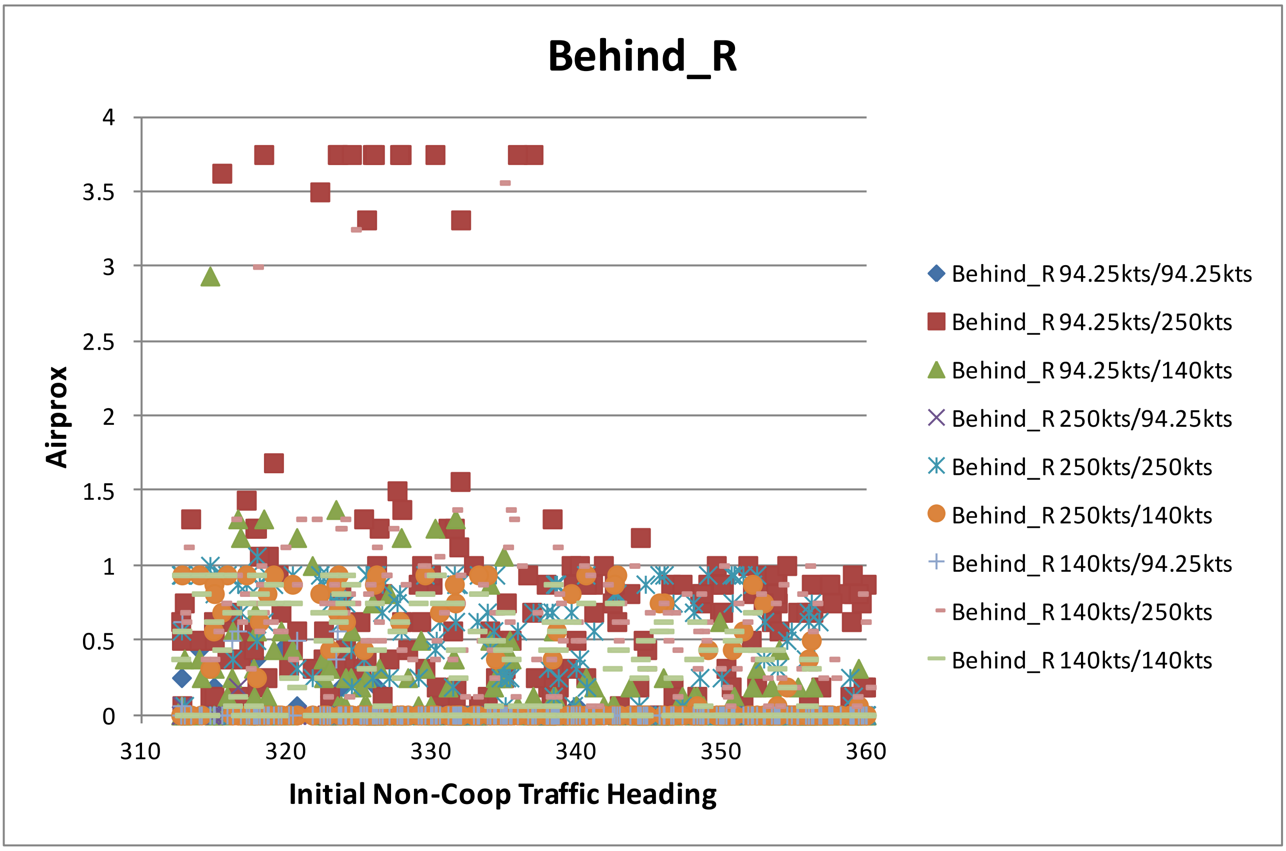


Figure A.29 Scatter Plots with respect to Initial Relative Heading including TC 10

This page is intentionally blank.

A.9. Hazardous Outcome Test Case Summaries

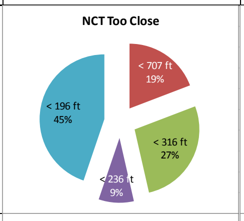
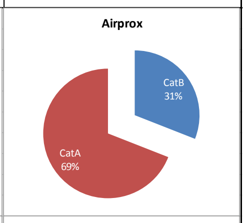
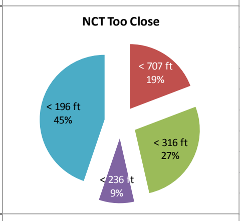
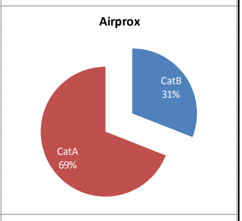
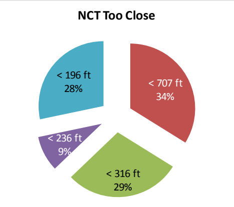
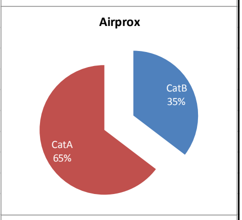
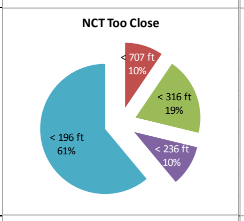
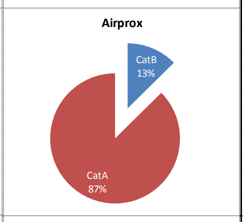
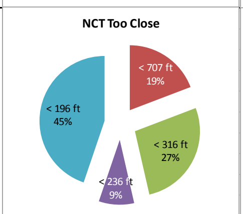
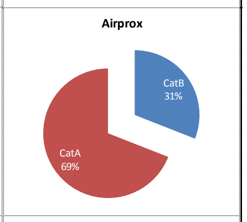
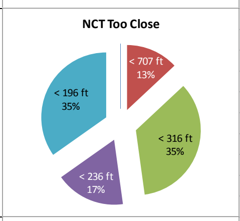
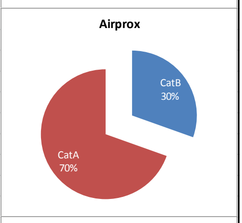
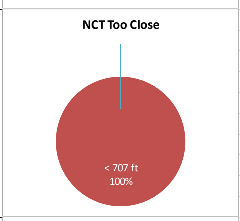
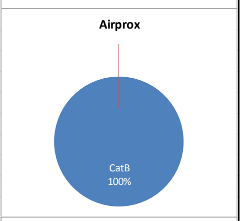
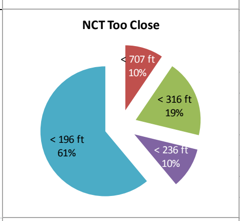
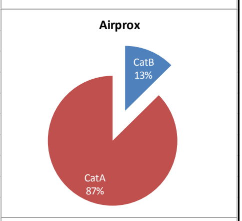
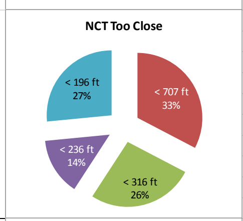
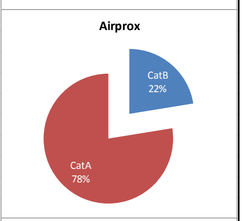
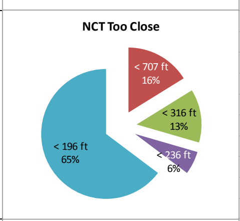
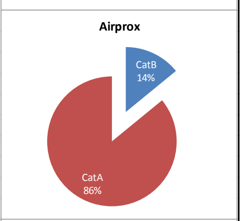
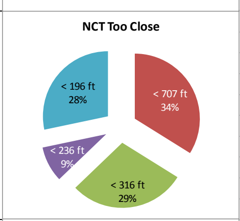
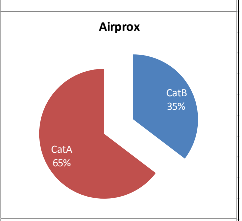
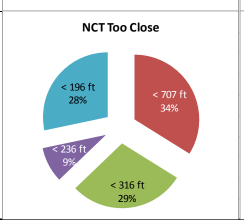
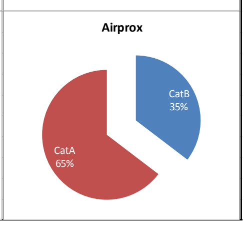
Hazard Context			System Flaws		Deviations			Test Cases			Diagnosed Defects			Inferred Airprox Event Likelihoods					
ID	Event	Special Conditions	Primary	Secondary	Avoidance Manoeuvre		Manoeuvring Airspeeds	No.	Trusted out of 12	Comment	Flawed	Fine	Unknown	UAS Surprised	UAS Threatened	NCT Surprised	NCT Panicked	NCT Too Close	UKAB Airprox Category (split between A & B)
HAZ 01.10	Approaching Visual Reporting Point or NAVAID	From Ahead with 0.5 NM separation	CF07	CF15 CF06 **CF04** CF05 CF15 CF06	LESS (ICA05)		DIFFERENT	8	12	PREL - S9 test case (speed difference).	0	5	1	🟢 1%	🟢 5%	🟢 15%	🟡 23%		
							HIGH	7	10	PREL - S3 test case (high speed).	2	1	3	🟢 1%	🟢 5%	🟢 15%	🟡 23%		
			CF07	CF05 CF15 **CF08**	MORE (ICA06)			3	7	ES - S8 test case.	1	2	3	🟢 8%	🟡 26%	🟢 0%	🟡 25%		
			CF15	CF07 CF06 CF08	CONFLICTING (ICA07)		DIFFERENT	1	1	**ENCTS (with offset defect) - S9 test case, and also S161 collision.**	1	5	0	🟢 3%	🟢 8%	🟡 30%	🔴 56%		
		From Right with 0.5 NM separation	**CF04**	CF15	LESS (ICA05)		LOW	12	5	PREL - S91 test case, prior to collision.	2	3	1	🟢 1%	🟢 5%	🟢 15%	🟡 23%		
						COMMISSION (ICA02)				PREL - S91 test case, at collision.				🟢 17%	🟢 0%	🟢 0%	🟢 0%		
			CF15	CF06 **CF08** CF07	EARLY (ICA03)		DIFFERENT	2	2	**ENCTS - S1 test case, precursor.**	1	4	1	🟢 0%	🟢 0%	🔴 62%	🔴 62%		
				CF06	CONFLICTING (ICA07)					**ENCTS - S1 test case at entry, and S85 at collision.**				🟢 3%	🟢 8%	🟡 30%	🔴 56%		
			CF07	**CF15**		COMMISSION (ICA02)				**ENCTS - S1 test case, exit.**				⚫ 86%	🟢 0%	🟢 0%	🟢 0%		
			CF15	**CF04**	LESS (ICA05)	CONFLICTING (ICA07)	LOW	6	4	PRE - S34 test case (MWT 5s).	1	4	1	🟢 1%	🟢 4%	🟢 11%	🟡 38%		
			CF04	**CF08** CF07	MORE (ICA06)		LOW	5	3	MWT - S3 test case, and also S27 at collision.	2	3	1	🟢 8%	🟡 26%	🟢 0%	🟡 25%		
			CF05	CF15 **CF08** CF10			HIGH	4	6	Latency - S5 test case, and also S67 at collision.	2	2	2	🟢 8%	🟡 26%	🟢 0%	🟡 25%		

Figure A.30 Hazardous Outcome Test Cases: Flaws, Deviations and Consequences for Approaches from Ahead and the Right

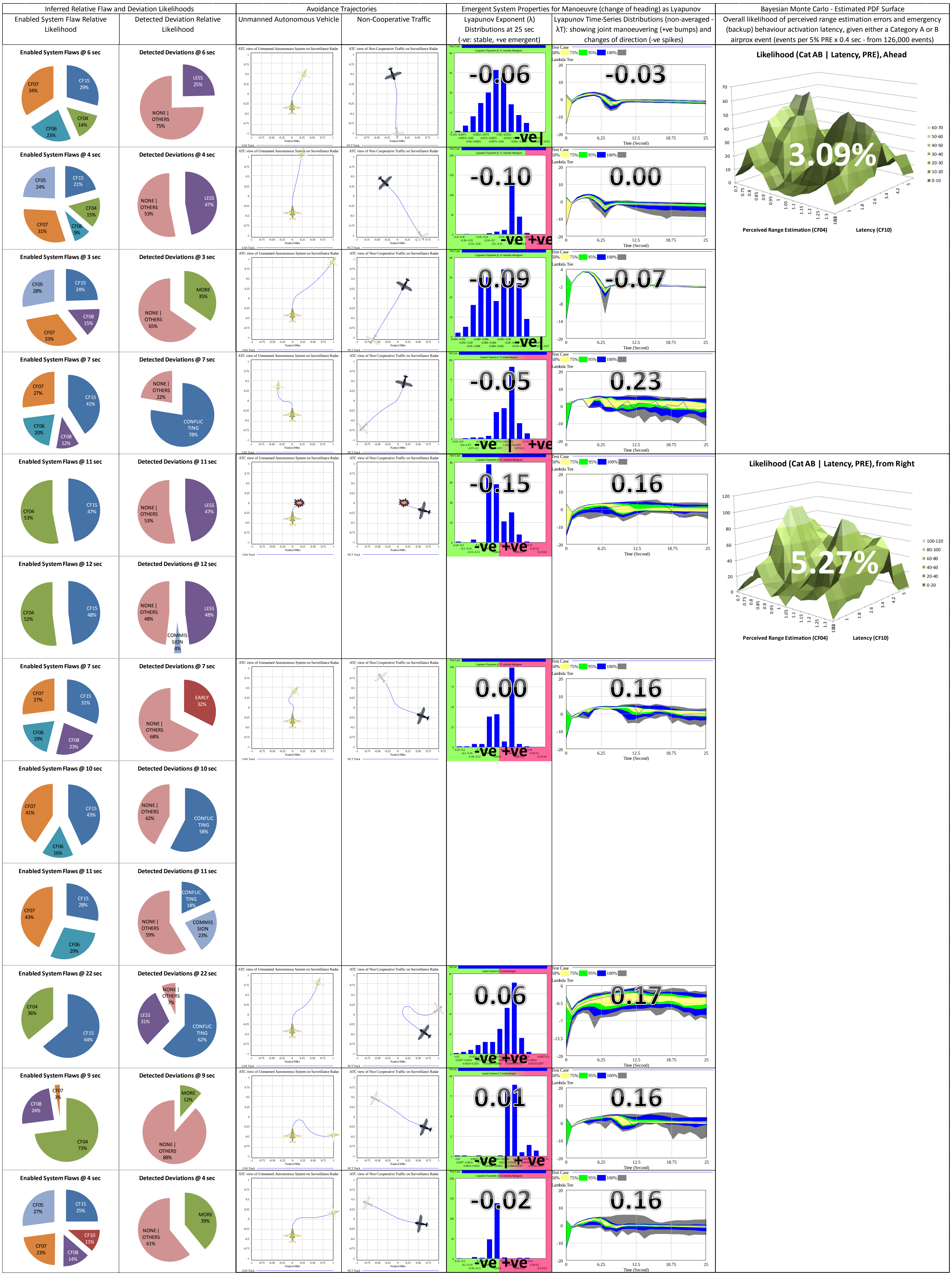


Figure A.31 Hazardous Outcome Test Cases: Flaw, Deviation, Perceived Estimation, and Latency Likelihoods, and Behaviours for Approaches from Ahead and the Right

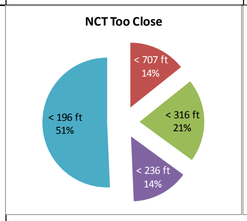
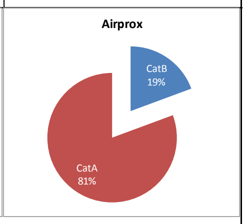
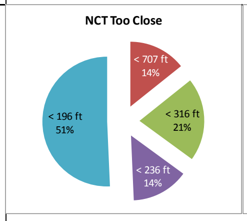
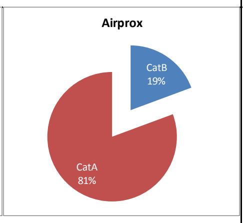
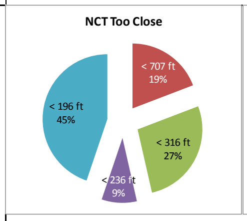
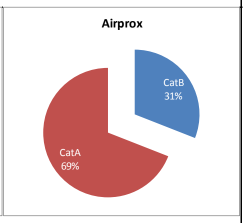
Hazard Context			System Flaws		Deviations				Test Cases			Diagnosed Defects			Inferred Airprox Event Likelihoods					
ID	Event	Special Conditions	Primary	Secondary	Avoidance Manoeuvre			Manoeuvring Airspeeds	No.	Trusted out of 12	Comment	Flawed	Fine	Unknown	UAS Surprised	UAS Threatened	NCT Surprised	NCT Panicked	NCT Too Close	UKAB Airprox Category (split between A & B)
HAZ 01.10	Approaching Visual Reporting Point or NAVAID	From Behind Right with 0.5 NM separation	CF07	CF15 **CF04** CF06	OMISSION (ICA01)	LATE (ICA04)	LESS (ICA05)	DIFFERENT	10	8	PREL - S200 test case, and also S95 collision.	1	3	2	🟢 2%	🟢 0%	🟢 16%	🟡 33%		
		From Behind Left with 0.5 NM separation	CF07	CF05 CF15 **CF04** CF06	OMISSION (ICA01)	LATE (ICA04)	LESS (ICA05)	HIGH	9	9	PREL - S164 test case.	2	3	1	🟢 2%	🟢 0%	🟢 16%	🟡 33%		
		From Left with 0.5 NM separation	CF05	CF07 CF15 **CF04** **CF10**	LESS (ICA05)			HIGH	11	11	PREL - S39 test case.	0	5	1	🟢 1%	🟢 5%	🟢 15%	🟡 23%		

Figure A.32 Hazardous Outcome Test Cases: Flaws, Deviations and Consequences for Approaches from Behind and the Left

Inferred Relative Flaw and Deviation Likelihoods		Avoidance Trajectories		Emergent System Properties for Manoeuvre (change of heading) as Lyapunov		Bayesian Monte Carlo - Estimated PDF Surface
Enabled System Flaw Relative Likelihood	Detected Deviation Relative Likelihood	Unmanned Autonomous Vehicle	Non-Cooperative Traffic	Lyapunov Exponent (λ) Distributions at 25 sec (-ve: stable, +ve emergent)	Lyapunov Time-Series Distributions (non-averaged - λT): showing joint manoeuvering (+ve bumps) and changes of direction (-ve spikes)	Overall likelihood of perceived range estimation errors and emergency (backup) behaviour activation latency, given either a Category A or B airprox event (events per 5% PRE x 0.4 sec - from 126,000 events)

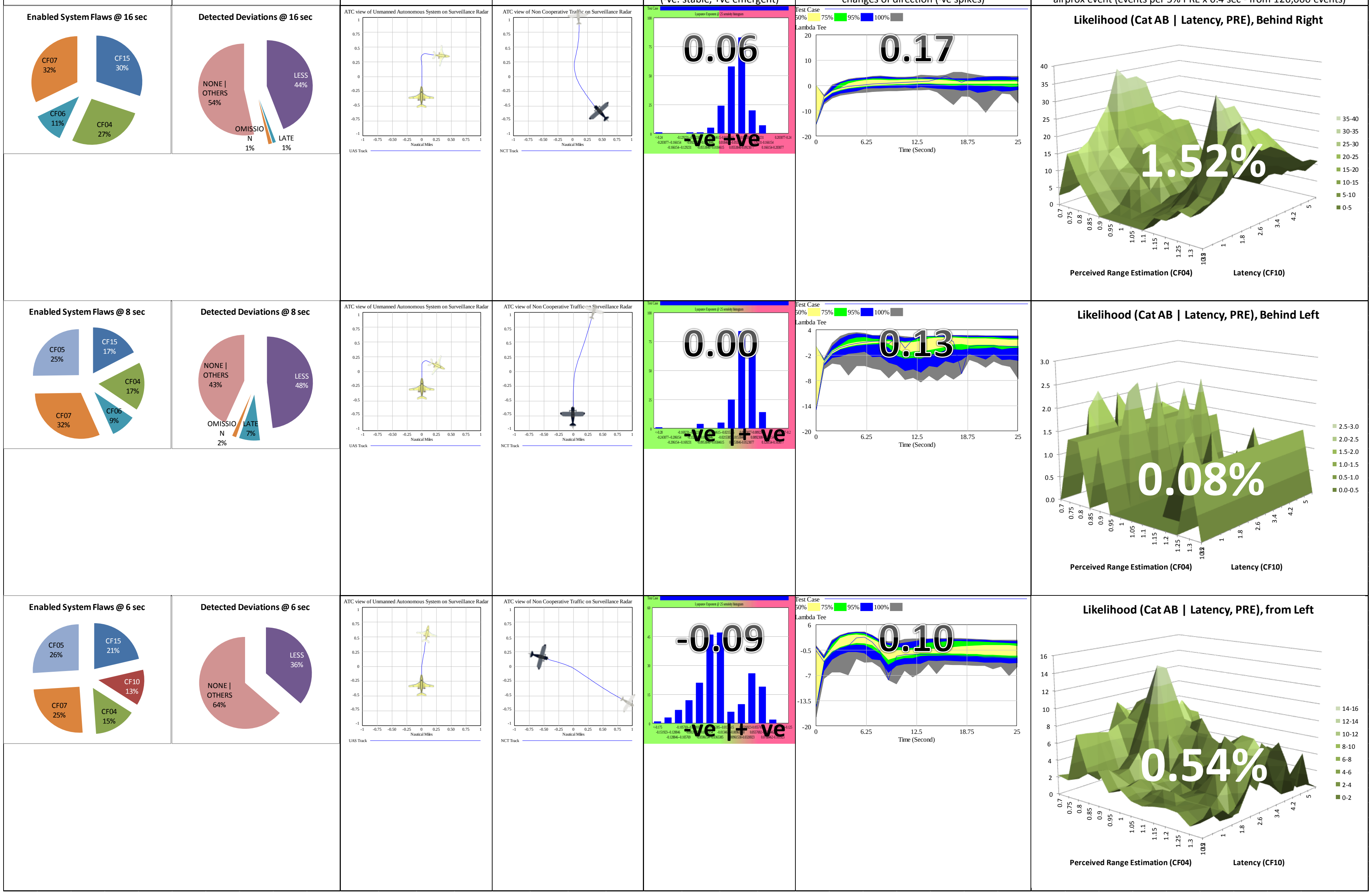


Figure A.33 Hazardous Outcome Test Cases: Flaw, Deviation, Perceived Estimation, and Latency Likelihoods, and Behaviours for Approaches from Behind and the Left

Appendix B. Hazards Advising Autonomy: Questionnaire

Motivation

This questionnaire arises from an on-going Systems Engineering Doctorate project. The purpose of this project is to investigate the development of a methodology and framework to determine autonomous system safety requirements relating to advisory functions in the context of system complexity and uncertainty. A variant of HAZOP (Hazard & Operability) study is proposed, along with the modelling of the likely system dynamics together with an inference mechanism to validate interpretations of predicted behaviour. These modelling approaches may even be used to represent a missing member of the HAZOP team – one with expertise accounting to some degree for the intended machine intelligence and its likely behaviour. It is a considerable challenge to adapt the HAZOP process to the fuzziness of systems with autonomous behaviour, including the dynamics of potentially dangerous interactions. Therefore the investigation is contextualised with a "toy-version" of a suitably representative problem. This questionnaire has been developed to assess to what extent dynamic behaviour for a given system might be intuitively grasped by specialists drawn from the field. With the outline system description given it is not the intention that you necessarily try to mentally emulate the model so much as to employ whatever experience, engineering judgement, and even "gut-feeling", as you might if considering the manoeuvre "deviations" as a member of a HAZOP team – the description serves as a guide only to intended behaviour. Please allow from an hour to 105 minutes to thoroughly read and complete this questionnaire – thank you.

Scenario Background

To this end a simplified "sense and avoid" system model has been created in order to facilitate an investigation of the hazard assessment concept, representing a two-dimensional interaction model with simplified rules and physics. The essential scenario is as presented alongside in Figure B.1. From this figure and for all of the vignettes within the scenario it is assumed that both vehicles are initially heading toward a common point of conflict (PoC).

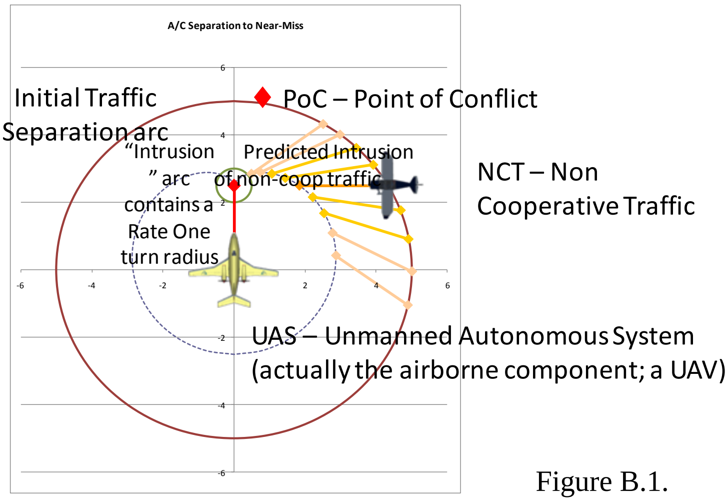

A/C Separation to Near-Miss
Initial Traffic Separation arc
PoC – Point of Conflict
"Intrusion" arc contains a Rate One turn radius
Predicted Intrusion of non-coop traffic
NCT – Non Cooperative Traffic
UAS – Unmanned Autonomous System (actually the airborne component; a UAV)
6
4
2
-2
-4
-6
-6
-4
-2
2
4
6

Figure B.1.

Both vehicles embody the Rules of the Air, as described overleaf. However the entity representing the UAV also embodies an additional emergency manoeuvre behaviour that overrides these rules of the air when threatened by the other intruding aircraft – which in general will cause the "UAV" to attempt to turn away from the direction of the threat. Also within these rules, each vehicle either takes the shortest turn towards its original heading where a rule allows a return to course or by default takes a turn to the left where the angle of turn required is approximately equal either way.

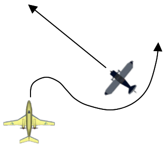
"Rules of the Air" – CAP 393 Section 4 General Flight Rules[8] (abridged)

Avoiding aerial collisions – 8

(5) … an aircraft which has the right-of-way under this rule shall maintain its course and speed.

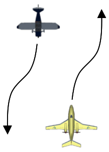
Converging – 9

(3) … when two aircraft are converging in the air at approximately the same altitude, the aircraft which has the other on its right shall give way.



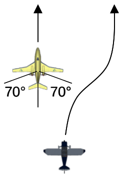
Approaching head-on – 10

When two aircraft are approaching head-on, or approximately so, in the air and there is a danger of collision, each shall alter its course to the right.



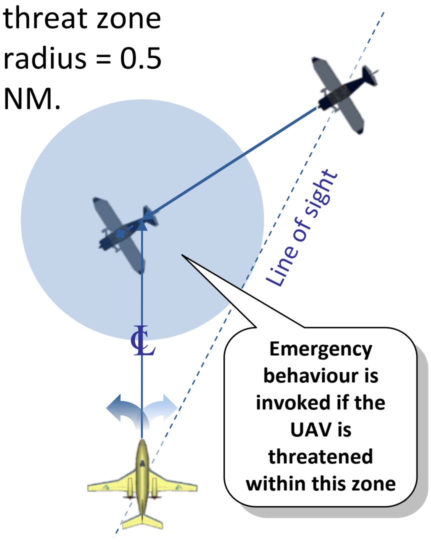
Overtaking – 11

(1) … an aircraft which is being overtaken in the air shall have the right-of-way and the overtaking aircraft, whether climbing, descending or in horizontal flight, shall keep out of the way of the other aircraft by altering course to the right.



For the emergency behaviour, the behavioural model for the UAV also presents a threat zone, of radius equivalent to that of a rate one turn at the current UAV airspeed, placed directly ahead, as illustrated alongside in Figure B.2. If the UAV model perceives that this threat exists, and with sufficient warning time, then an emergency manoeuvre is initiated, turning the UAV in the direction away from the intruder's current position with respect to the UAV's extended centre-line. The enactment of this behaviour is a discontinuity with respect to the "rules of the air" behaviour, and if this behaviour is incorrectly specified then the hazard is not mitigated but worsened. Otherwise, the UAV executes an avoidance manoeuvre consistent with the rules of the air and based only upon the line of sight to the intruding aircraft. When safe separation is achieved both model entities will then turn towards their original headings.

Figure B.2.


Collision threat zone radius = 0.5 NM.
Line of sight
℄
Emergency behaviour is invoked if the UAV is threatened within this zone

Three specific system defects have been incorporated into the model: errors in the perceived range estimation, variation of the warning time of threats, and variation of the latency in reacting to threats. Sensor models, fields and depth of view, explicit aerodynamic behaviour, inertial properties and delays in

[8] CAP 393 - Air Navigation: The Order and the Regulations. London, UK: TSO (The Stationery Office) on behalf of the UK Civil Aviation Authority, Safety Regulation Group, 2010

controller and actuator mechanisms are all either ignored, assumed to be “perfect”, or otherwise considerably simplified. Only dynamics essential for avoidance are represented.

Any separation less than a certain minimum is deemed a collision – freezing any further movement of the two vehicles and consequently freezing the simulated separation as effectively zero.

Pre-preliminary Safety Assessment / HAZOP Methodology

For the purposes of this study a variation of the STPA (Systems Theoretic Process Analysis), as developed by the Complex Systems Research Laboratory at MIT (http://sunnyday.mit.edu/), has been applied in order to develop a provisional HAZOP. This uses the abstract behavioural parameter “Manoeuvre” to which guidewords are applied, creating the seven deviations, or “Inadequate Control Action”, as detailed below:

ICA01 – Omission The UAV does not manoeuvre to avoid the other aircraft.

ICA02 – Commission The UAV manoeuvres in the wrong direction attempting to avoid the other aircraft.

ICA03 – Early The UAV prematurely ceases a manoeuvre to avoid the other aircraft and returns to the given heading.

ICA04 – Late The UAV belatedly initiates a manoeuvre to avoid the other aircraft.

ICA05 – Less The UAV rate of manoeuvre is insufficient to avoid the other aircraft.

ICA06 – More The UAV rate of manoeuvre is excessive in attempting to avoid the other aircraft.

ICA07 – Conflicting The UAV is unable to resolve the conflict between realising the System Goal and avoiding the other aircraft.

Associated with these deviations there are also the following fifteen potential control flaws, or system defects as identified through the STPA-HAZOP process. Together with the observed system dynamical behaviour, these deviations and system defects are to be also to be used within a system safety diagnostic inference model – aiding system hazard identification, inference and validation. However, only the highlighted control flaws are legitimate in the context of the current model – incorporating the specific defect representations as described in the scenario previously:

CF01 The UAV fails to detect the other aircraft altogether.

CF02 The UAV does not command a turning manoeuvre from the flight control system (FCS).

CF03 The UAV flight control surfaces do not respond to demands from the FCS.

CF04 The UAV incorrectly estimates the location and direction to the predicted point of conflict.

CF05 The UAV fails to follow established Rules of the Air.

CF06 The other aircraft fails to follow established Rules of the Air.

CF07 The UAV interprets the Rules of the Air too rigidly when closing within 500ft of the other aircraft.

CF08 The UAV loses detection of the other aircraft.

CF09 The UAV fails to detect the other aircraft at sufficient range.

CF10 The UAV delays the commanding of a turning manoeuvre from the flight control system (FCS).

CF11 The UAV flight control surfaces are slow to respond to demands from the FCS.

CF12 The UAV demanded rate of turn is insufficient for the closing angle, speed and range to the other aircraft.

CF13 The UAV maximum demanded rate of turn exceeds the maximum achievable rate of turn.

CF14 The UAV demanded rate of turn is excessive for the closing angle, speed and range to the other aircraft.

CF15 The UAV does not arbitrate the competing heading demands appropriately.

About Yourself

This questionnaire is anonymous. Please indicate the extent of your experience in the fields related to this study. This will help to cross-reference, correlate and benchmark the subsequent responses to the multiple-choice section that follows. You have the following years of experience in:

Operator, piloting, air crew and / or air traffic	none □	<1 yr □	1<5 yrs □	5<20 yrs □	>20 yrs □
System safety engineering	none □	<1 yr □	1<5 yrs □	5<20 yrs □	>20 yrs □
Robotics or autonomous systems engineering	none □	<1 yr □	1<5 yrs □	5<20 yrs □	>20 yrs □
Applying HAZOP to hazard identification	none □	<1 yr □	1<5 yrs □	5<20 yrs □	>20 yrs □
System simulation and / or modelling	none □	<1 yr □	1<5 yrs □	5<20 yrs □	>20 yrs □
System diagnostics	none □	<1 yr □	1<5 yrs □	5<20 yrs □	>20 yrs □

From the following list select **exactly four** of the terms or expressions that best qualifies your approach and perspective regarding the assessment of system safety. There are no right or wrong choices in this:

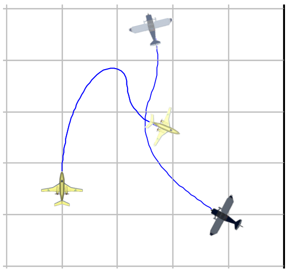
Component Failure Mode	□
Functional Failure Mode	□
Mission Phase Analysis	□
Classified Severity of Effect	□
Probabilistic Occurrence Rates	□
Release and Flow of Energy	□
System Operating State	□
Causal Propagation	□
Parameter Deviations	□
Human Reliability Error	□
Captured Expertise	□
Captured Socio-technical Behaviour	□
Captured Complex System Dynamics	□
Adaptation and Migration of Behaviour	□
Active Exploration	□
Safety Constraint Violation	□

The Instructions

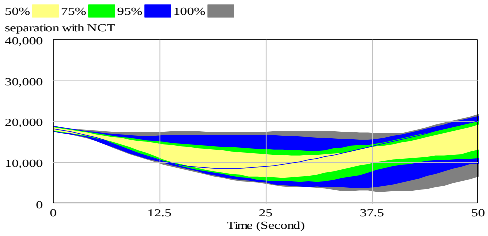
It is generally recognized that the validation of a hazard analysis is difficult to objectively accomplish in the absence of meaningful hazard statistics. Consequently, the hazard modelling method is to be validated, at least in part, by drawing upon the knowledge and experience of a group of subject matter experts. By executing these dynamical and inference models within a series of hazardous vignettes certain subtleties have been identified to be embodied in the following questions.

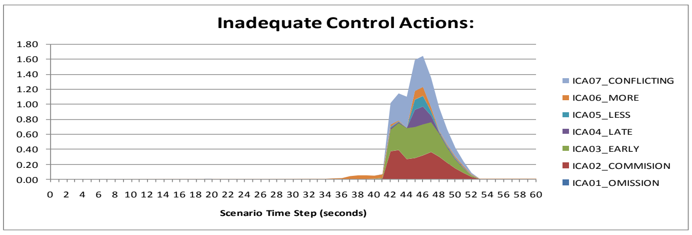
The main section of this questionnaire comprises eleven multiple choice questions contained within the following four sub-sections. These questions are designed to explore what is, or is not, intuitive in an understanding of the one-on-one collision avoidance scenario, based upon the simplified system model as described in the scenario background previously. Each question has four plausible options, of which only one is completely correct according to the model and its observed behaviour. The aim is to determine how close your specialist assessment of the situation is to matching the model. So as to provide wider scope in this assessment multiple selections are also allowed wherever you believe that the true answer may lie in more than one choice. Clearly, however, ticking all four options for a question reveals nothing, identical to having ticked none at all. For clarity, the UAV, where shown in the following, is at the centre or just below in each figure.

Some Notes on Graphical Notation



Consider the behaviour observed in the 50 second trajectories shown below. The UAV (starting in a northerly direction to the left) continuously misestimates the perceived range to the non-cooperative aircraft (coming in from the right). In this encounter the UAV attempts to manoeuvre behind the non-cooperative aircraft, being the aircraft which has the other on its right, and so shall give way. At a certain point in the encounter the non-cooperative aircraft then observes that the UAV is directly ahead, and so gives way to the right, but the UAV still attempts to manoeuvre behind. In the corresponding time-series plots, the distribution plot (below left) reveals purely statistical information in the encounter, whilst the plot to the right represents degrees of “belief” that a range of “Inadequate Control Actions” (ICA) are responsible for the behaviour. From the left-hand plot we can see that closest point of approach (<3000 ft at > 37.5 seconds) is an outlier in the distribution of separations near to the 100% limit – note also the slight double-dip. At the first closest approach, the initial peak in the “belief” plot, the inference is that the UAV is manoeuvring the wrong way and that it is attempting to return to heading with a behaviour that is “conflicted” – or unresolved. At the second peak it is also inferred that the manoeuvring is now late, less than required, but more than is appropriate.

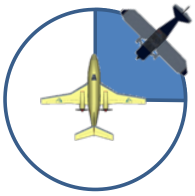




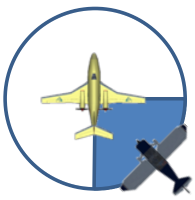
The Questions

Section 1: general situation awareness

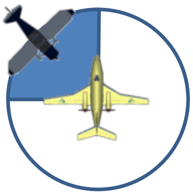
Q1.



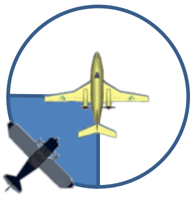
(a)



(b)



(c)



(d)

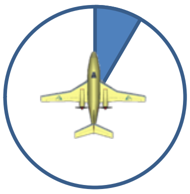
Where a UAV is positioned at the centre of a minimum separation zone, with a non-cooperative aircraft then approaching from an initial distance of 0.5 NM from the edge of this zone on any of the various directions as shown in Figure Q1, from which of these different direction segments do you believe that hazardous encounters are more likely to result in a collision?

Tick all that apply:

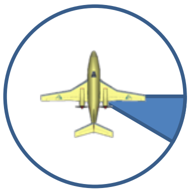
(a) □, (b) □, (c) □, (d) □

Optionally, add a reason here for your choice(s)

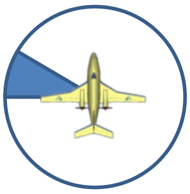
Q2.



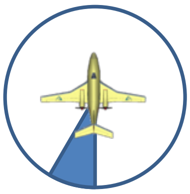
(a)



(b)



(c)



(d)

Where the UAV system **<u>does not</u>** exhibit internal defects with respect to estimating the location of the other aircraft, nor excessive latency in the emergency manoeuvre response, and has adequate warning time, from which of the different segments in Figure Q2 would you expect that hazardous encounters are more likely to result in a collision?

Tick all that apply:

(a) □, (b) □, (c) □, (d) □

Optionally, add a reason here for your choice(s)

Q3.

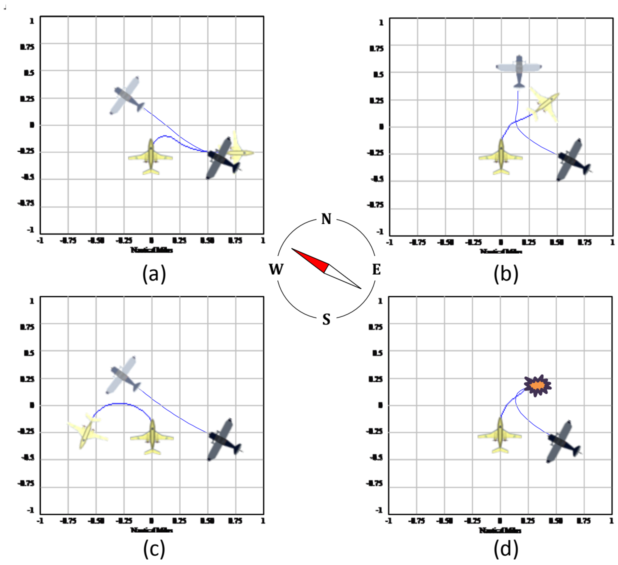
Where the UAV system **<u>does</u>** exhibit a specific defect with respect to either estimating the location of the non-cooperative aircraft, or excessive latency in the emergency manoeuvre response, or inadequate warning, from which of the segments in Figure Q2 would you expect that hazardous encounters are more likely to result in a collision?

Referring to Figure Q2, tick all that apply: (a) □, (b) □, (c) □, (d) □

Optionally, also add a reason here for your choice(s)

Section 2: anticipation of interactions

Q4.



(a) (b) (c) (d)

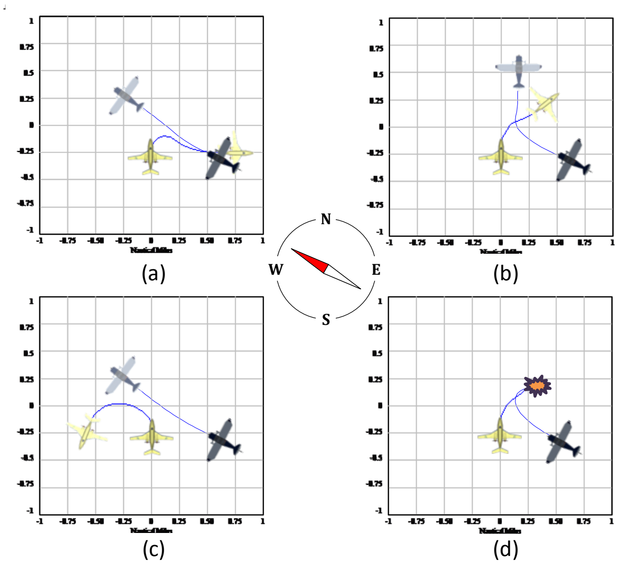
A non-cooperative aircraft breaches the 0.5 NM minimum separation constraint, flying with about 1½ times the airspeed of the UAV, on the approximate headings shown in Figure Q4. The UAV continuously incorrectly perceives the other aircraft to be somewhat further away than it is. Given that this produces a delayed switching of the UAV behaviour, which of the interactions represents the more likely outcome where initially the other aircraft is on a slightly more westerly heading, and consequently arriving more from the 3 o'clock position?

Tick all that apply:

(a) □, (b) □, (c) □, (d) □

Optionally, also add a reason here for your choice(s)

Q5.



(a) (b) (c) (d)

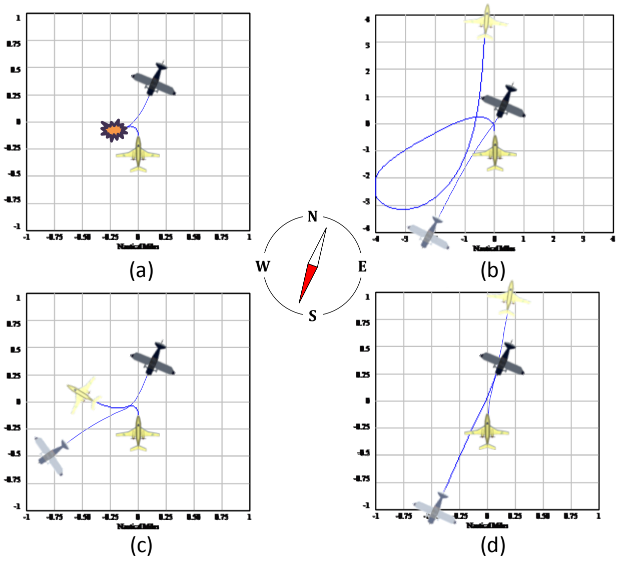
A non-cooperative aircraft breaches the 0.5 NM minimum separation constraint, flying at the same respective airspeeds as in Q4, and on the approximate headings as shown in Figure Q5. On this occasion the UAV system provides zero warning of the impending conflict and hence does not trigger an emergency manoeuvre and therefore its behaviour narrowly interprets the "rules of the air", as described on page 2. Otherwise, no other defect is present. Which of the interactions then represents the more likely outcome for this situation?

Tick all that apply:

(a) □, (b) □, (c) □, (d) □

Optionally, also add a reason here for your choice(s)

Q6.



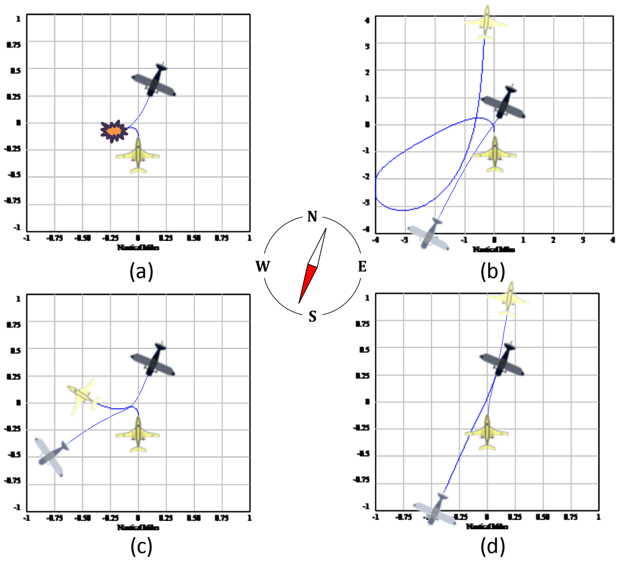
A non-cooperative aircraft breaches the 0.5 NM minimum separation constraint with both aircraft flying near to, or at, the 250 kts speed-limit, on the approximate headings shown in Figure Q6. Assuming that the emergency manoeuvre behaviour responds to threats observed to the right of the extended centre-line with an evasive turn to the left, which of the interactions represents a plausible outcome if initially the other aircraft is on a slightly more westerly heading, and consequently arriving more from the 1 o'clock position?

Tick all that apply:

(a) ☐, (b) ☐, (c) ☐, (d) ☐

Optionally, also add a reason here for your choice(s)

Q7.



A non-cooperative aircraft breaches the 0.5 NM minimum separation constraint, flying with slightly more than twice the airspeed of the UAV, on the approximate headings shown in Figure Q7. The emergency manoeuvre behaviour responds in the same manner as outlined in Q6, and in all other respects no specific defects are present. Which of the interactions might represent the outcome if initially the other aircraft is on a slightly more southerly heading, and therefore slightly more head-on with the UAV?

Tick all that apply:

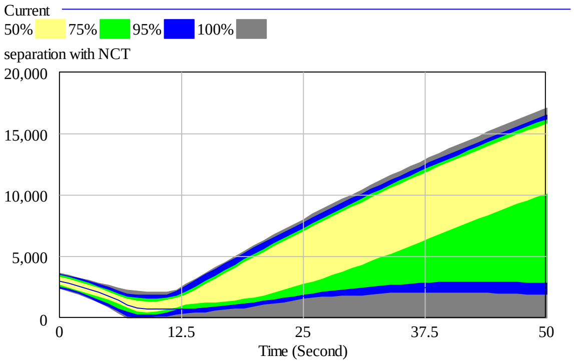
(a) ☐, (b) ☐, (c) ☐, (d) ☐

Optionally, also add a reason here for your choice(s)

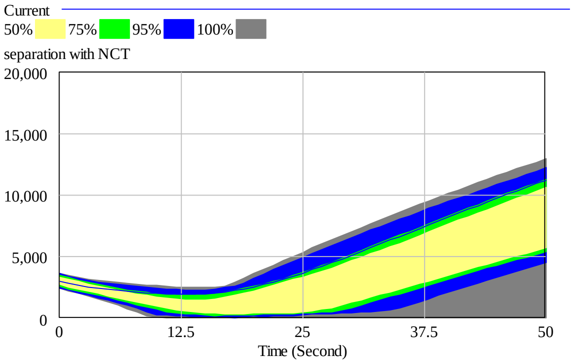
Section 3: Inferences

Each of the four vignettes shown in Figures Q8 and Q9 represents outcomes with respect to the “separation” variable in the results of four Monte Carlo simulations of 200 runs apiece. Each vignette embodies a variance of ±5° on the headings of both vehicles, ±5 kts on each airspeed, and an initial separation in the range of 0.4 – 0.6 NM. Some vignettes also embody variances of 0 to 2 seconds warning and -75% to +300% tolerance on the perceived range estimation, where appropriate.

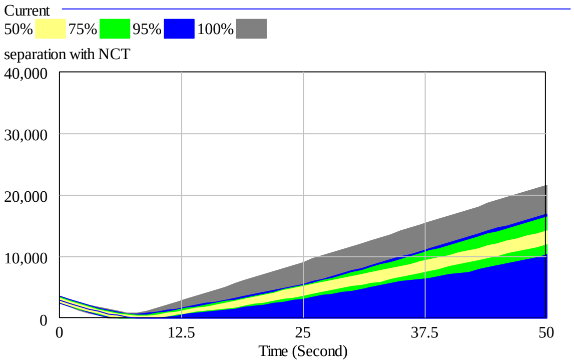
Q8.



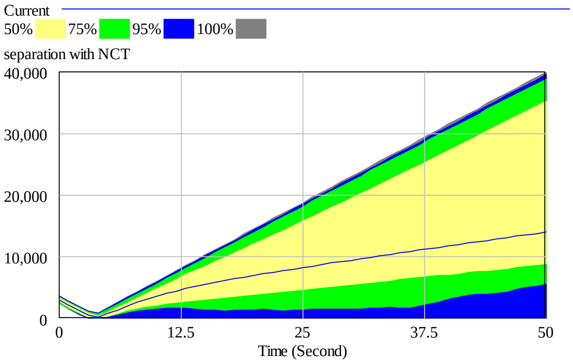
(a)



(b)



(c)



(d)

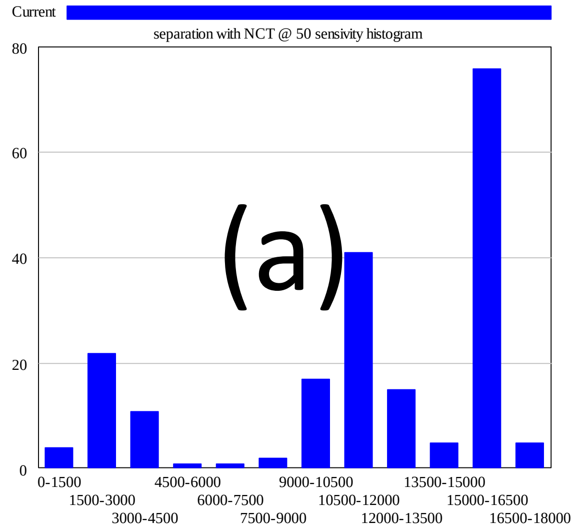
For the above varying spreads of separation trajectories, each over a 50 second period from the start of the encounter, and each showing the spread of separation distances in feet (vertical axis) across the 50%, 75%, 95% and 100% confidence intervals, which vignette would appear to represent the least risk of collision?

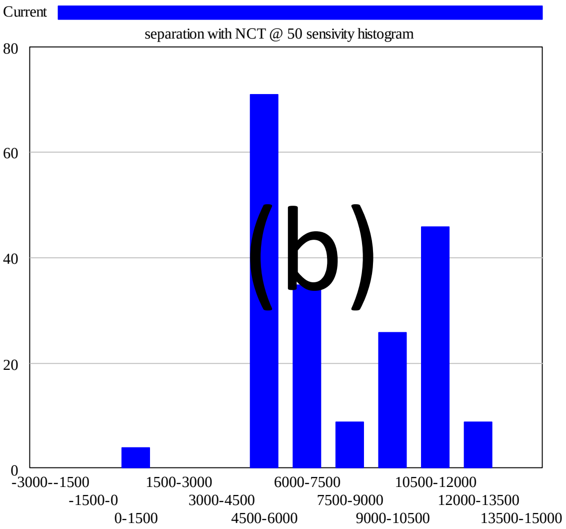
Tick the most appropriate:

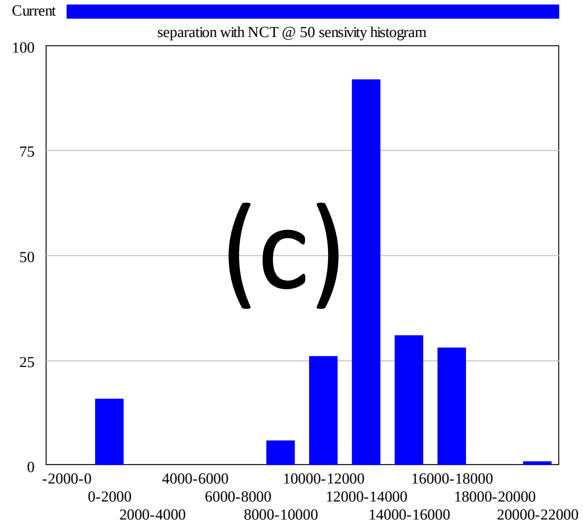
(a) □, (b) □, (c) □, (d) □

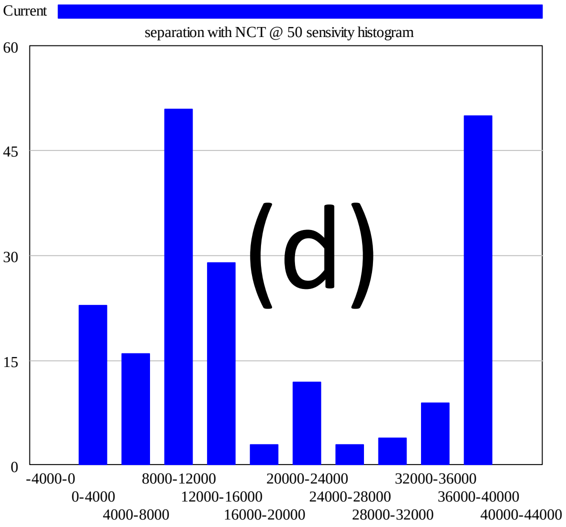
Optionally, also add a reason here for your choice(s)

Q9.









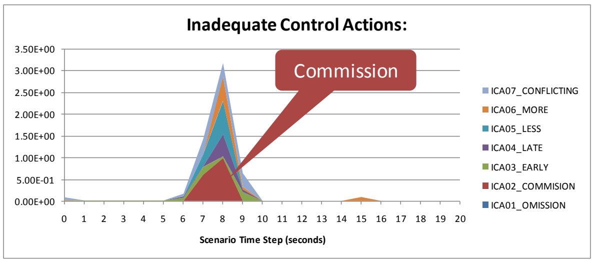
For the same Monte Carlo simulations as referred to previously in Q8, each of the above histograms represents the distribution of separation distances (horizontal axis) at 50 seconds into each vignette, which of these distributions appears to represent the greater risk of collision?

Tick the most appropriate:

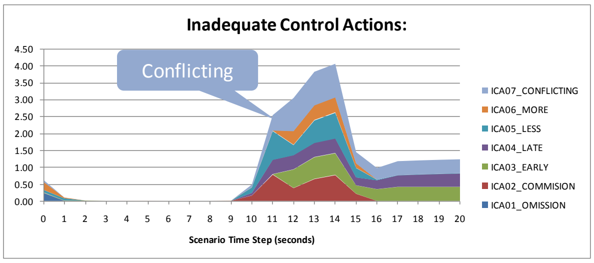
(a) □, (b) □, (c) □, (d) □

Optionally, also add a reason here for your choice(s)

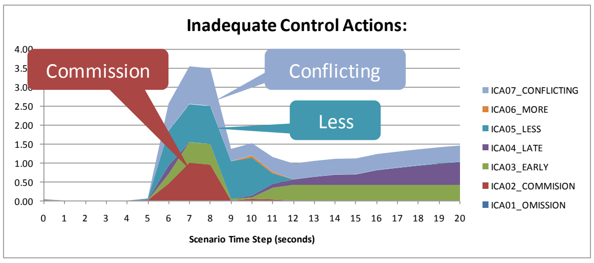
Q10.



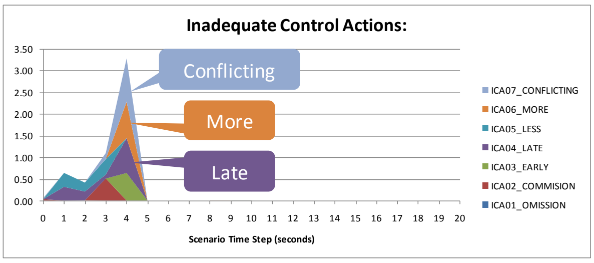
(a)



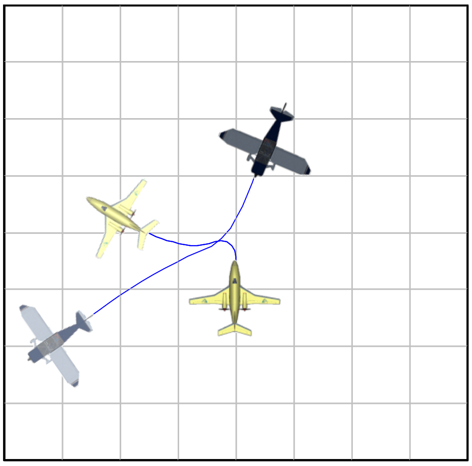
(b)



(c)



(d)



Referring also to the example given in the 'Notes on Graphical Notation', the above time-series plots in Figure Q10 represent four inferences taken from the previous vignettes. These represent inferences of the system defects, in terms of HAZOP deviations, or Inadequate Control Actions, with respect to an abstract parameter "Manoeuvre". Which of the above inference plots then represents the close encounter as depicted in the figure alongside where there is a failure of the UAV to anticipate the appropriate change in heading on the part of the other aircraft, and thereby then incorrectly to initiate a poorly resolved dangerous emergency manoeuvre taking it toward the other aircraft and also too close then to ensure that it might safely turn away?

Tick the most appropriate:

(a) □, (b) □, (c) □, (d) □

Optionally, also add a reason here for your choice(s)

Section 4: Prioritising risks

Q11.

Which of the following four item lists presents the apparent descending rank order of hazard risk associated with the four previously presented vignettes? Descending order – i.e. riskiest vignette at the top in each list.

a) A minimum warning time < 1 sec
Wider tolerance on perceived range estimation to 25% - 400%
Higher speeds of up to 250 kts in the case of both vehicles
A higher differential speed for the non-cooperative traffic

b) Wider tolerance on perceived range estimation to 25% - 400%
A minimum warning time < 1 sec
A higher differential speed for the non-cooperative traffic
Higher speeds of up to 250 kts in the case of both vehicles

c) Higher speeds of up to 250 kts in the case of both vehicles
A higher differential speed for the non-cooperative traffic
A minimum warning time < 1 sec
Wider tolerance on perceived range estimation to 25% - 400%

d) Wider tolerance on perceived range estimation to 25% - 400%
A higher differential speed for the non-cooperative traffic
Higher speeds of up to 250 kts in the case of both vehicles
A minimum warning time < 1 sec

Tick the most appropriate:

(a) ☐, (b) ☐, (c) ☐, (d) ☐

Optionally, add a reason here for your choice(s)

End of questionnaire.

Appendix C. Respondent Results and Analysis

Cohen’s Kappa [20] (κ) non-parametric statistical method for inter-rater agreement (α): where complete agreement = +1.

$$\alpha = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \equiv \begin{bmatrix} true + ve\ answer & false - ve\ answer \\ false + ve\ answer & true - ve\ answer \end{bmatrix} \equiv \begin{bmatrix} agree\ for & disagree\ against \\ disagree\ for & agree\ against \end{bmatrix} \tag{C.1}$$

$$\kappa = \frac{2 \cdot (\alpha_{11} \cdot \alpha_{22} - \alpha_{12} \cdot \alpha_{21})}{{\alpha_{12}}^2 + {\alpha_{21}}^2 + 2 \cdot \alpha_{11} \cdot \alpha_{22} + (\alpha_{11} + \alpha_{22}) \cdot (\alpha_{12} + \alpha_{21})} \tag{C.2}$$

Consider where $\alpha_{21} = \alpha_{12} = 0$; where every respondent agrees with the response(s) for the “correct” outcome whilst also agreeing with all of the posited responses that stand against the “correct” outcome, i.e. indicating perfect agreement (κ_A):

$$\kappa_A = \frac{2 \cdot \alpha_{11} \cdot \alpha_{22}}{2 \cdot \alpha_{11} \cdot \alpha_{22} + (\alpha_{11} + \alpha_{22}) \cdot (0)} = +1 \tag{C.3}$$

Alternatively consider where $\alpha_{11} = \alpha_{22} = 0$; where every respondent disagrees with the response(s) for the “correct” outcome whilst also disagreeing with all of the posited responses that stand against the “correct” outcome, i.e. indicating perfect disagreement (κ_D):

$$\kappa_D = \frac{2 \cdot (-\alpha_{12} \cdot \alpha_{21})}{{\alpha_{12}}^2 + {\alpha_{21}}^2 + (0) \cdot (\alpha_{12} + \alpha_{21})} = \frac{-2}{\frac{\alpha_{12}}{\alpha_{21}} + \frac{\alpha_{21}}{\alpha_{12}}} = \begin{cases} -1, & \alpha_{12} = \alpha_{21} \\ \rightarrow -0, & \alpha_{12} \gg \alpha_{21} \\ \rightarrow -0, & \alpha_{12} \ll \alpha_{21} \end{cases} \tag{C.4}$$

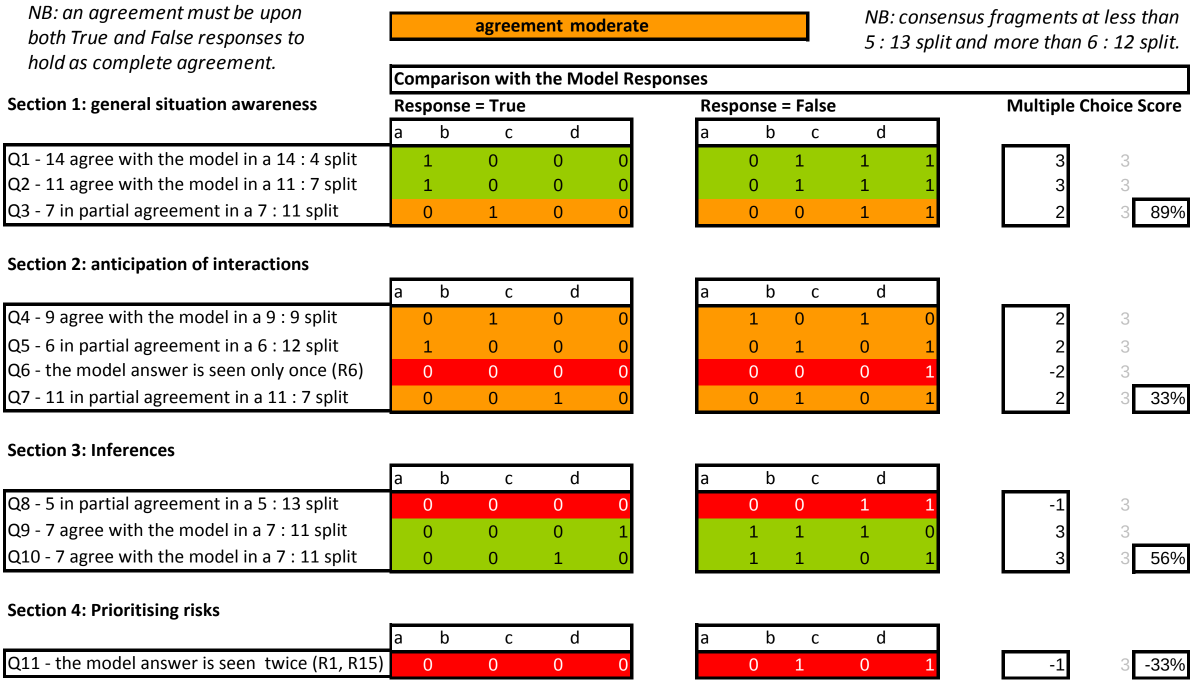
However if α11 = α_{12} = 0, or $\alpha_{21} = \alpha_{22} = 0$, or $\alpha_{11} = \alpha_{21} = 0$, or $\alpha_{12} = \alpha_{22} = 0$; where every respondent either disagrees with the response(s) for the “correct” outcome whilst also agreeing with all of the posited responses that stand against the “correct” outcome (disagreeing with the model where “correct” but accepting all other options as incorrect also – and therefore accepting none), or agrees with the response(s) for the “correct” outcome whilst also disagreeing with all of the posited responses that stand against the “correct” outcome (selecting all model responses as equally correct), or disagrees with the response(s) for the “correct” outcome whilst also agreeing with all of the posited responses that stand for the “correct” outcome (a straight contradiction), or agrees with the response(s) against the “correct” outcome whilst also disagreeing with all of the posited responses that stand against the “correct” outcome (another straight contradiction), i.e. all indicating an indeterminate measure of agreement (κ_R):

$$\mathrm{\kappa_R} = 0 \tag{C.5}$$

These measures of agreement have been categorised using a similar scale as proposed by Landis [64]:

$$C = \begin{cases} \text{"none"}, & \kappa \leq 0.0 \\ \text{"poor / slight"}, & 0.0 < \kappa \leq 0.2 \\ \text{"fair"}, & 0.2 < \kappa \leq 0.4 \\ \text{"moderate"}, & 0.4 < \kappa \leq 0.6 \\ \text{"substantial"}, & 0.6 < \kappa \leq 0.8 \\ \text{"almost perfect "}, & 0.8 < \kappa \leq 1.0 \end{cases} \tag{C.6}$$

C.1. Cohort Results and Question-by-Question Analysis



NB: an agreement must be upon both True and False responses to hold as complete agreement.

agreement moderate

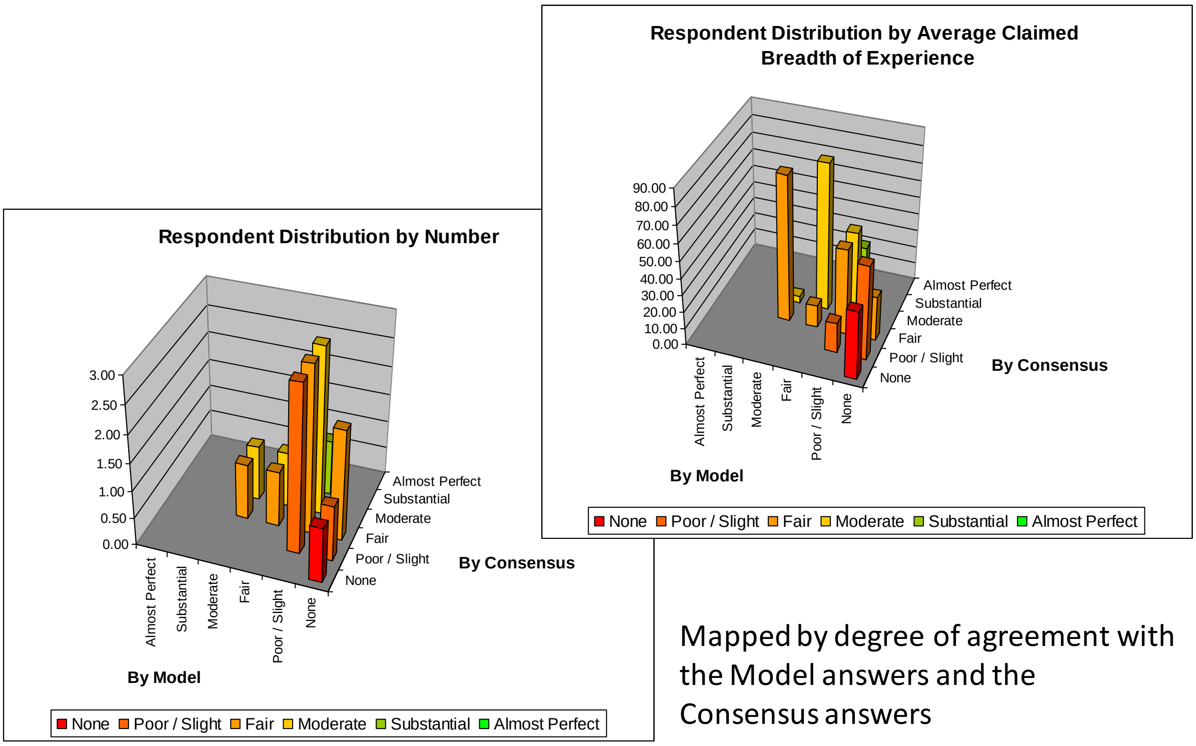
NB: consensus fragments at less than 5 : 13 split and more than 6 : 12 split.

Comparison with the Model Responses

	Response = True				Response = False				Multiple Choice Score		
Section 1: general situation awareness	a	b	c	d	a	b	c	d			
Q1 - 14 agree with the model in a 14 : 4 split	1	0	0	0	0	1	1	1	3	3	
Q2 - 11 agree with the model in a 11 : 7 split	1	0	0	0	0	1	1	1	3	3	
Q3 - 7 in partial agreement in a 7 : 11 split	0	1	0	0	0	0	1	1	2	3	89%
Section 2: anticipation of interactions	a	b	c	d	a	b	c	d			
Q4 - 9 agree with the model in a 9 : 9 split	0	1	0	0	1	0	1	0	2	3	
Q5 - 6 in partial agreement in a 6 : 12 split	1	0	0	0	0	1	0	1	2	3	
Q6 - the model answer is seen only once (R6)	0	0	0	0	0	0	0	1	-2	3	
Q7 - 11 in partial agreement in a 11 : 7 split	0	0	1	0	0	1	0	1	2	3	33%
Section 3: Inferences	a	b	c	d	a	b	c	d			
Q8 - 5 in partial agreement in a 5 : 13 split	0	0	0	0	0	0	1	1	-1	3	
Q9 - 7 agree with the model in a 7 : 11 split	0	0	0	1	1	1	1	0	3	3	
Q10 - 7 agree with the model in a 7 : 11 split	0	0	1	0	1	1	0	1	3	3	56%
Section 4: Prioritising risks	a	b	c	d	a	b	c	d			
Q11 - the model answer is seen twice (R1, R15)	0	0	0	0	0	1	0	1	-1	3	-33%

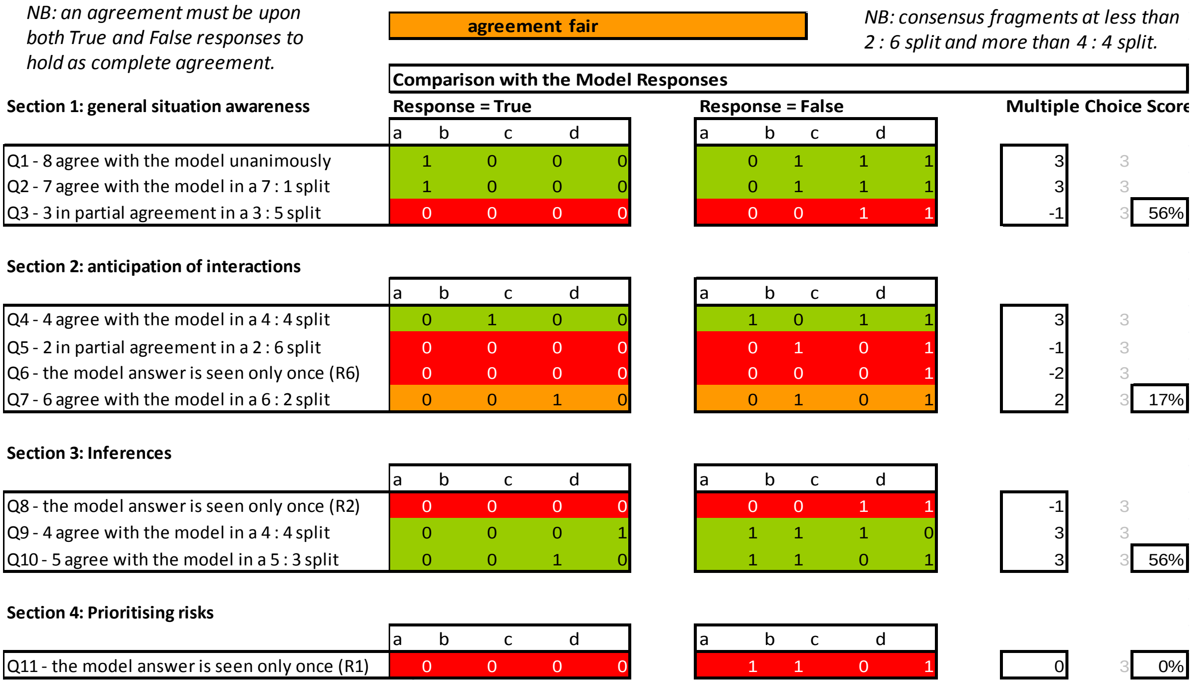
Figure C.1 Questionnaire Responses for 6 : 12 split (any 6 agreements from 18 respondents).

In Figure C.1 and Figure C.3, “green” scores represent (aggregate) agreements categorised as “substantial” or “almost perfect”, “orange” scores for “fair” to “moderate” categories, and “red” for “none” to “poor / slight” agreement.



Mapped by degree of agreement with the Model answers and the Consensus answers

Figure C.2 Questionnaire Respondent Agreement Distributions (for 18 respondents).



NB: an agreement must be upon both True and False responses to hold as complete agreement.

agreement fair

NB: consensus fragments at less than 2 : 6 split and more than 4 : 4 split.

Comparison with the Model Responses

	Response = True				Response = False				Multiple Choice Score		
Section 1: general situation awareness	a	b	c	d	a	b	c	d			
Q1 - 8 agree with the model unanimously	1	0	0	0	0	1	1	1	3	3	
Q2 - 7 agree with the model in a 7 : 1 split	1	0	0	0	0	1	1	1	3	3	
Q3 - 3 in partial agreement in a 3 : 5 split	0	0	0	0	0	0	1	1	-1	3	56%
Section 2: anticipation of interactions	a	b	c	d	a	b	c	d			
Q4 - 4 agree with the model in a 4 : 4 split	0	1	0	0	1	0	1	1	3	3	
Q5 - 2 in partial agreement in a 2 : 6 split	0	0	0	0	0	1	0	1	-1	3	
Q6 - the model answer is seen only once (R6)	0	0	0	0	0	0	0	1	-2	3	
Q7 - 6 agree with the model in a 6 : 2 split	0	0	1	0	0	1	0	1	2	3	17%
Section 3: Inferences	a	b	c	d	a	b	c	d			
Q8 - the model answer is seen only once (R2)	0	0	0	0	0	0	1	1	-1	3	
Q9 - 4 agree with the model in a 4 : 4 split	0	0	0	1	1	1	1	0	3	3	
Q10 - 5 agree with the model in a 5 : 3 split	0	0	1	0	1	1	0	1	3	3	56%
Section 4: Prioritising risks	a	b	c	d	a	b	c	d			
Q11 - the model answer is seen only once (R1)	0	0	0	0	1	1	0	1	0	3	0%

Figure C.3 Questionnaire Responses for 4 : 4 split (any 4 agreements from best 8 respondents).

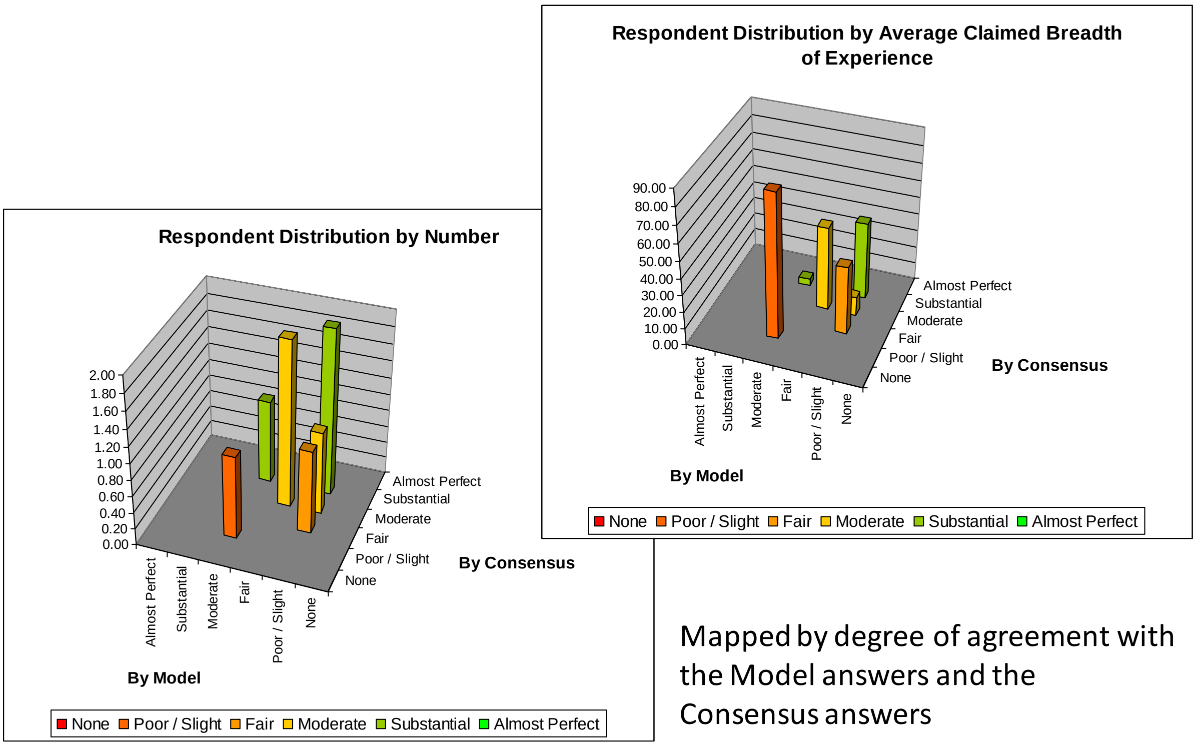
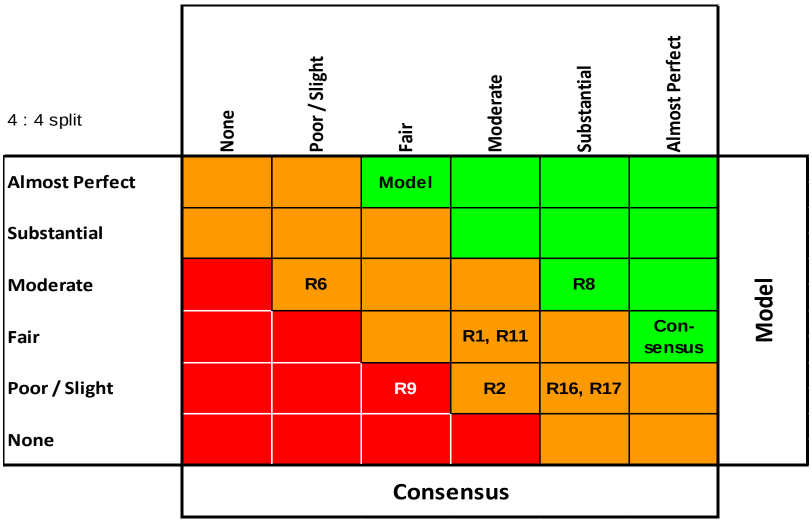


Figure C.4 Questionnaire Respondent Agreement Distributions (for best 8 respondents).

C.2. Cohort Groupings and Authority (for best 8 respondents)



4 : 4 split	None	Poor / Slight	Fair	Moderate	Substantial	Almost Perfect
Almost Perfect	🟧	🟧	🟩 Model	🟩	🟩	🟩
Substantial	🟧	🟧	🟧	🟩	🟩	🟩
Moderate	🟥	🟧 R6	🟧	🟧	🟩 R8	🟩
Fair	🟥	🟥	🟧	🟧 R1, R11	🟧	🟩 Con-sensus
Poor / Slight	🟥	🟥	🟥 R9	🟧 R2	🟧 R16, R17	🟧
None	🟥	🟥	🟥	🟥	🟧	🟧

Model
Consensus

Figure C.5 Model v Consensus – respondent agreement matrix (including R9 in best 8).

In Figure C.5 respondent R9 claims substantial system simulation expertise, but degrades the degree of agreement between the consensus perspective and the "correct" behaviour indicated by the model.

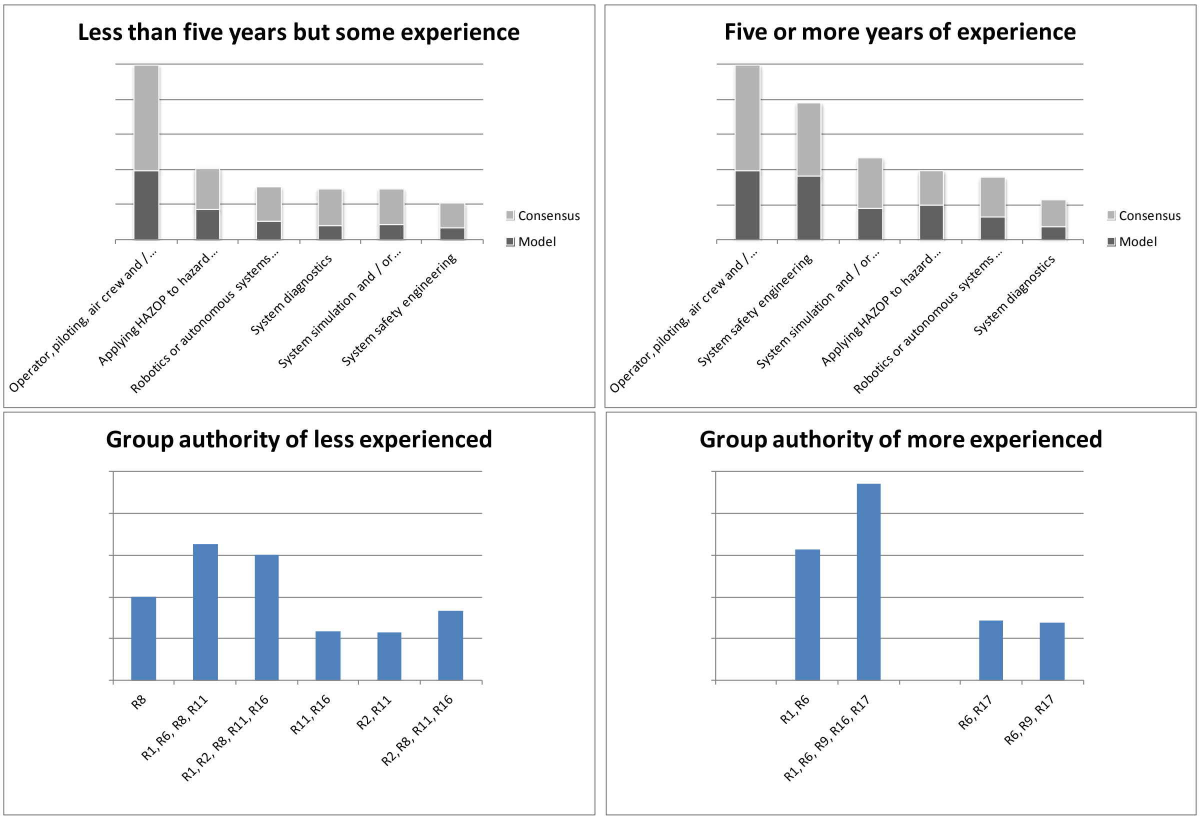
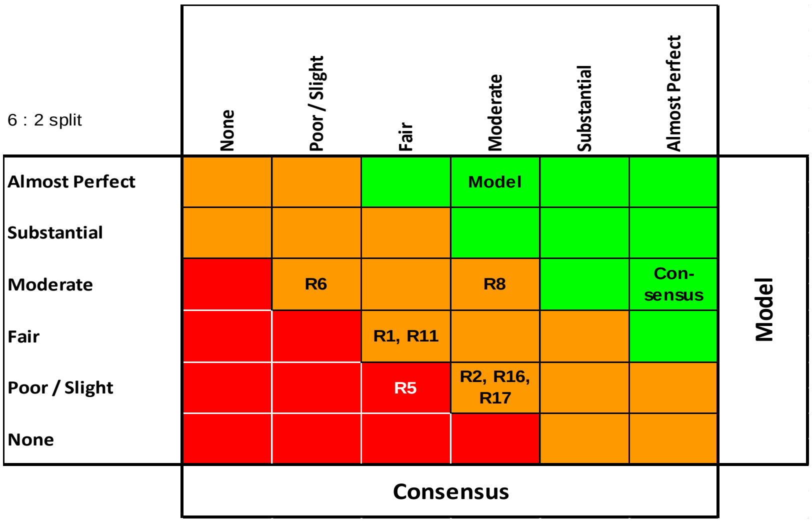


Figure C.6 Normalised average agreement & authority weighted by claimed experience (Experience < 5 years – left, > 5 years – right).

A weighted average by experience is applied to these aggregate scores, for 1, 5, 20 and 40 or more years:

$$\kappa_{more} = \frac{\kappa_{E1} + 3 \times \kappa_{E5} + 9 \times \kappa_{E20} + 27 \times \kappa_{E40}}{N_{E1} + 3 \times N_{E5} + 9 \times N_{E20} + 27 \times N_{E40}} \qquad (C.7)$$

$$\kappa_{less} = \frac{27 \times \kappa_{E1} + 9 \times \kappa_{E5} + 3 \times \kappa_{E20} + \kappa_{E40}}{27 \times N_{E1} + 9 \times N_{E5} + 3 \times N_{E20} + N_{E40}} \qquad (C.8)$$



6 : 2 split

Model \ Consensus	None	Poor / Slight	Fair	Moderate	Substantial	Almost Perfect
Almost Perfect	🟧	🟧	🟩	🟩 Model	🟩	🟩
Substantial	🟧	🟧	🟧	🟩	🟩	🟩
Moderate	🟥	🟧 R6	🟧	🟧 R8	🟩	🟩 Con-sensus
Fair	🟥	🟥	🟧 R1, R11	🟧	🟧	🟩
Poor / Slight	🟥	🟥	🟥 R5	🟧 R2, R16, R17	🟧	🟧
None	🟥	🟥	🟥	🟥	🟧	🟧

Figure C.7 Model v Consensus – respondent agreement matrix (including R5 in best 8).

In Figure C.7, by including respondent R5 claiming some piloting expertise, a better alignment is obtained between model's "correct" behaviour and the belief held in the consensus concerning the expected behaviour.

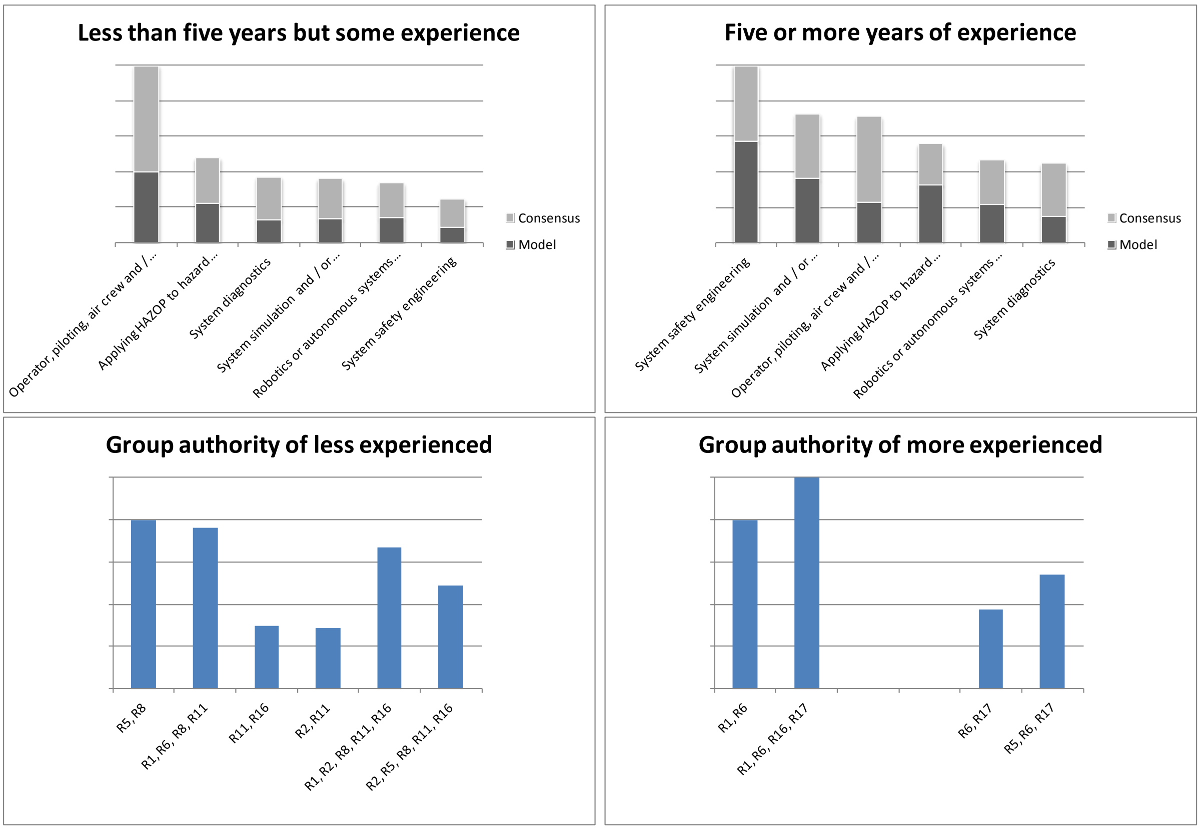


Figure C.8 Normalised average agreement & authority weighted by claimed experience (Experience < 5 years – left, > 5 years – right).

Likewise, a weighted average by experience is applied to the aggregate scores, for 1, 5, 20 and 40 or more years' experience (equations C.7 and C.8).

This page is intentionally blank.

(end of thesis)