New copy

CRANFIELD UNIVERSITY



K SANGPETCHSONG

THE APPLICATION OF RELATIVE NAVIGATION TO CIVIL AIR TRAFFIC MANAGEMENT

COLLEGE OF AERONAUTICS

PhD THESIS

2000

ProQuest Number: 10820972

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.



ProQuest 10820972

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by Cranfield University.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

CRANFIELD UNIVERSITY

COLLEGE OF AERONAUTICS

PhD THESIS

CRANFIELD UNIVERSITY LIBRARY

Academic Year 1996-1999

K SANGPETCHSONG

The application of relative navigation to civil air traffic management

Supervisor: D Allerton

August 2000

This thesis is submitted in partial fulfilment of the requirements for the degree of PhD

ABSTRACT

This thesis addresses navigation and guidance which will be required for air traffic management in Future Air Navigation Systems (FANS) and Free Flight. In particular, the thesis covers the issues of data fusion and integrity monitoring, to provide an adequate level of aircraft separation assurance, based on relative navigation (RELNAV).

The evolution of air navigation systems is described. The principles of Kalman filtering and Joint Tactical Information Distribution System (JTIDS) RELNAV are covered. Sensor models of strapdown Inertial Navigation System (INS), Global Positioning System (GPS), and Automatic Dependent Surveillance-Broadcast (ADS-B) are developed in Matlab and integrated to form a hybrid navigation system. RELNAV algorithms for centralised and decentralised Kalman filtering are formulated, and their respective performances are analysed using Monte Carlo simulations for an airspace containing several aircraft. It is shown that RELNAV, based on the integration of INS and ADS-B, can enable aircraft to maintain safe separation independent of GPS, where it is assumed that an ADS-B datalink provides accurate time synchronisation.

An alternative approach that integrates INS, GPS, and ADS-B is developed and analysed. It is shown that this approach is more applicable to civil aviation because it eliminates the needs to establish and manage several navigation communities simultaneously, in effect, exploiting GPS as the navigation controller. The source selection functions used for RELNAV are also developed, and the stability and performance of this technique is evaluated from simulation studies.

A failure detection algorithm that monitors the residuals of a Kalman filter is derived and evaluated using Monte Carlo simulations of GPS failures. It is shown that this algorithm combines the use of likelihood functions and chi-squared tests, allowing both a false alarm rate to be selected and a failed sensor to be identified.

Finally, an algorithm is developed for separation assurance to determine the probability that aircraft are closer than a pre-determined distance, taking into account flight path prediction errors. It is shown that this algorithm simplifies aircraft conflict detection in three dimensions and allows the conflict probability at a particular time to be determined. This approach is validated using Monte Carlo simulations of aircraft trajectories which include near-misses.

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisor, Prof. Allerton, for all his priceless advice and kind support in both academic and personal matters. His advice has always kept me from falling into mistakes. Without his guidance and encouragement, producing this thesis would have been a totally different story. Let it be sufficient to say that there is barely a result in this thesis which was not obtained at his suggestion, or improved by his review. Any remaining mistakes in this work are due to my own fault.

I would also like to express my gratitude to Mr. John Wyatt and Miss Helen Clinch for their assistance with Matlab which was truly needed to carry out this work. I am grateful to Mr. John Harrington who helped obtain the research information with his usual promptness and competence. I would also like to thank Dr. John Ellis, Mr. Richard Odic, and Mr. Huamin Jia for their precious time and useful advice in reviewing this thesis.

I am greatly beholden to the Royal Thai Navy for giving me the opportunity to study in the United Kingdom.

Heartfelt indebtedness belongs to my mother and father who have always been patient, understanding, and supportive. I also sincerely appreciate the support from Pum and Jim+ during all this time. Lastly, I thank you Lek for the encouragement and for consistently asking to read this thesis.

TABLE OF CONTENTS

ABSTRACT i

ACKNOWLEDGEMENTS ii

TABLE OF CONTENTS iii

LIST OF FIGURES vii

LIST OF TABLES xi

NOTATION xii

CHAPTER 1 INTRODUCTION 1

1.1 Conventional Air Navigation Systems 1
1.2 Navigation systems integration 7
1.3 Absolute and relative navigation 8
1.4 The Future Air Navigation Systems (FANS) 9
1.5 The limitations of the proposed FANS implementation 11
1.6 Research objectives 14
1.7 Scope of the thesis 14

CHAPTER 2 AIR NAVIGATION SYSTEMS 16

2.1 Introduction 16
2.2 Communication component of FANS 16
2.2.1 Mode S extended squitter *16*
2.2.2 VHF Digital Link Mode 4 ..... *19*
2.3 Navigation component of FANS 22
2.3.1 Global Positioning System *23*
2.3.2 Inertial Navigation System *25*
2.4 Surveillance component of FANS 26
2.5 Traffic Alert and Collision Avoidance System 28

CHAPTER 3 KALMAN FILTERING THEORY 30

3.1 Introduction 30
3.2 Stochastic linear system model 30
3.2.1 Continuous time stochastic linear system model *31*
3.2.2 Discrete time stochastic linear system model *33*
3.2.3 Discrete time linear measurement model *35*

3.2.4 Shaping filter and augmented system model ... 36
3.3 DISCRETE TIME (SAMPLED-DATA) OPTIMAL ESTIMATOR: THE KALMAN FILTER ... 37
3.4 NON-LINEAR ESTIMATOR: THE EXTENDED KALMAN FILTER ... 39
3.5 KALMAN FILTER PERFORMANCE ANALYSIS ... 42
3.5.1 Monte Carlo simulation ... 42
3.5.2 Kalman filter divergence ... 43
3.5.3 Decentralised Kalman filtering ... 46

CHAPTER 4 NAVIGATION SYSTEMS AND SYSTEMS INTEGRATION ... 47

4.1 INTRODUCTION ... 47
4.2 UNDERLYING MODELLING AND ASSUMPTIONS ... 47
4.2.1 Coordinate systems and transformations ... 48
4.2.2 Development of the airspace model ... 50
4.2.3 Kinematics of a particle utilising intermediate frame of reference ... 52
4.2.4 Development of the aircraft model ... 53
4.3 STRAPDOWN INS NAVIGATION ALGORITHMS ... 54
4.3.1 Development of strapdown INS equations of motion ... 54
4.3.2 System error analysis ... 55
4.4 GPS NAVIGATION ALGORITHMS ... 59
4.4.1 Development of a simplified model of GPS ... 59
4.4.2 Development of general pseudorange measurement model ... 62
4.4.3 Stand-alone GPS navigation algorithms: the single-point solution ... 64
4.4.4 Dilution of precision ... 65
4.5 STATE-SPACE FORMULATION ... 66
4.6 STAND-ALONE SYSTEM PERFORMANCE ANALYSIS ... 68
4.6.1 Stand-alone strapdown INS performance analysis ... 69
4.6.2 Stand-alone GPS performance analysis ... 73
4.7 NAVIGATION SYSTEMS INTEGRATION ... 76
4.7.1 Introduction ... 76
4.7.2 Indirect feedback filter design of the integrated INS/GPS system ... 77
4.7.3 Integrated INS/GPS system performance analysis ... 78

CHAPTER 5 RELATIVE NAVIGATION ... 85

5.1 INTRODUCTION ... 85
5.2 JOINT TACTICAL INFORMATION DISTRIBUTION SYSTEM ... 85
5.3 PRINCIPLES OF JTIDS RELATIVE NAVIGATION ... 87
5.4 RELNAV NAVIGATION ALGORITHMS ... 91
5.4.1 Development of the relative grid equations of motion ... 92
5.4.2 Development of the equations of motion in the relative grid ... 94

5.5 STATE-SPACE FORMULATION 98
5.6 CENTRALISED RELNAV PERFORMANCE ANALYSIS 101

CHAPTER 6 RELATIVE NAVIGATION FOR CIVIL APPLICATIONS 110

6.1 INTRODUCTION 110
6.2 DECENTRALISED RELNAV NAVIGATION ALGORITHMS 110
6.2.1 Development of ownstate decentralised RELNAV overall system model 111
6.2.2 Derivation of ownstate decentralised RELNAV state estimator 111
6.3 DECENTRALISED RELNAV PERFORMANCE ANALYSIS 114
6.4 PRINCIPLES OF CIVIL RELATIVE NAVIGATION 118
6.4.1 Civil RELNAV navigation mechanisation 119
6.4.2 Development of source selection functions 120
6.5 INTEGRATED INS/GPS/ADS-B SYSTEM PERFORMANCE ANALYSIS 124

CHAPTER 7 SENSOR FAILURE DETECTION AND ISOLATION 131

7.1 INTRODUCTION 131
7.2 MAXIMUM LIKELIHOOD ESTIMATOR 131
7.3 KALMAN FILTER RESIDUAL MONITORING 133
7.4 CHI-SQUARED DISTRIBUTION 139
7.5 CHI-SQUARED TEST FOR SENSOR FAILURE DETECTION AND IDENTIFICATION 141
7.5.1 Previous developments of chi-squared-based FDI algorithms 141
7.5.2 Application of likelihood functions to chi-squared-based FDI algorithms 143

CHAPTER 8 AIRCRAFT TRAJECTORY CONFLICT PREDICTION 148

8.1 INTRODUCTION 148
8.2 RECENT DEVELOPMENTS IN FREE FLIGHT CONFLICT DETECTION ALGORITHMS 148
8.3 CONFLICT PROBABILITY ESTIMATION UTILISING SEPARATION ESTIMATION ERRORS 150
8.3.1 Aircraft conflict detection utilising separation estimation errors 150
8.3.2 Statistical characterisation of separation estimation errors 154
8.3.3 Characterisation of aircraft conflict probability estimation algorithm 156
8.3.4 Aircraft conflict probability estimation based on predicted flight trajectory 160
8.4 THREE DIMENSIONAL AIRCRAFT CONFLICT DETECTION ALGORITHM 162
8.4.1 Vertical conflict probability estimation 163
8.4.2 Aircraft conflict detection in three dimensions 163

CHAPTER 9 CONCLUSION AND SUGGESTED FURTHER WORKS 166

9.1 OVERVIEW OF THE THESIS AND RESULTS 166
9.2 RELATIVE NAVIGATION UTILISING ADS-B 168
9.3 AIRCRAFT CONFLICT DETECTION 171
9.4 SUGGESTIONS FOR FURTHER RESEARCH 172

REFERENCES 174

APPENDIX A ACRONYMS AND ABBREVIATIONS 181

APPENDIX B VALIDATION OF THE SIMULATED NAVIGATION SYSTEM MODELS 183

APPENDIX C GPS SATELLITE TRAJECTORY SIMULATION PARAMETERS 188

APPENDIX D LEAST SQUARES ESTIMATION 189

APPENDIX E RULES FOR EXPECTATION AND VARIANCE 191

APPENDIX F RELNAV INITIAL ERROR VARIANCE CALCULATIONS 192

APPENDIX G DERIVATION OF THE OWNSTATE DECENTRALISED RELNAV STATE ESTIMATOR 195

APPENDIX H DISCRIPTION OF THE SIMULATION PROGRAM 198

LIST OF FIGURES

Figure 1-1 Conventional airway system 2
Figure 1-2 Area Navigation (RNAV) concept 3
Figure 1-3 Conventional Air Navigation Systems 5
Figure 1-4 Comparison of the world fleet in 1998 and 2018 6
Figure 1-5 Integrated INS/LORAN system concept 7
Figure 1-6 Future Air Navigation Systems 10
Figure 1-7 Scope of the thesis 15
Figure 2-1 Mode S extended squitter message format 17
Figure 2-2 TDMA operational concept 19
Figure 2-3 Periodic broadcast protocol 20
Figure 2-4 Co-channel interference conditions 21
Figure 3-1 Organisation of a physical system, overall system model and Kalman filter. 31
Figure 3-2 The summary of discrete Kalman filter algorithm 38
Figure 3-3 The extended Kalman filter process 41
Figure 3-4 Kalman filtering examples 45
Figure 4-1 Coordinate systems, notations and transformations 48
Figure 4-2 Terrestrial navigation coordinate systems 50
Figure 4-3 The simplified airspace model (the *i*-frame) 51
Figure 4-4 Kinematics analysis of a particle utilising intermediate frames of reference 52
Figure 4-5 Geometry of waypoint bypassing 54
Figure 4-6 Strapdown INS frames of reference 57
Figure 4-7 The navigation process of the simplified model of a strapdown INS 58
Figure 4-8 The geometry of GPS satellite circular orbital trajectory 59
Figure 4-9 General 2-state clock error model 61
Figure 4-10 Comparison of true and modelled errors of a low accuracy strapdown INS 70
Figure 4-11 Comparison of a stationary strapdown INS and clock performance 71
Figure 4-12 Flight trajectory of the simulated aircraft 72
Figure 4-13 The performance of the stand-alone INS with the high accuracy sensors 73
Figure 4-14 The performance of the stand-alone GPS with the high accuracy sensors 74
Figure 4-15 The effects of the measurement accuracy on the stand-alone GPS performance 75
Figure 4-16 The indirect feedback Kalman filter configuration 78
Figure 4-17 The performance of the integrated INS/GPS system with high accuracy sensors 79
Figure 4-18 Comparison of the stand-alone and integrated system theoretical rms errors 80

Figure 4-19 The effects of the measurement accuracy on the integrated INS/GPS system performance 80
Figure 4-20 The effects of sensor accuracy on the integrated INS/GPS system performance 81
Figure 4-21 The divergence of the integrated INS/GPS Kalman filter 82
Figure 5-1 JTIDS network and stacked net operational concept 86
Figure 5-2 JTIDS relative navigation (RELNAV) operational concept 88
Figure 5-3 Active synchronization by Round Trip Timing interrogation 89
Figure 5-4 Coordinate systems used in RELNAV 92
Figure 5-5 Centralised RELNAV extended Kalman filter configuration 100
Figure 5-6 Flight trajectories for the RELNAV performance analysis 101
Figure 5-7 Theoretical rms errors of the RELNAV system states 102
Figure 5-8 The RELNAV performance of the first member aircraft 102
Figure 5-9 The RELNAV performance of the second member aircraft 103
Figure 5-10 Comparison of the theoretical rms errors for stand-alone INS and centralised RELNAV for the first member aircraft 104
Figure 5-11 Comparison of the theoretical rms errors for stand-alone INS and centralised RELNAV for the second member aircraft 104
Figure 5-12 Flight trajectories for testing the observability 105
Figure 5-13 The effects of the source motion on the observability and performance of RELNAV 106
Figure 5-14 Divergence of the RELNAV system state estimation errors 107
Figure 5-15 Divergence of the first member aircraft estimation errors 108
Figure 5-16 Divergence of the second member aircraft estimation errors 108
Figure 6-1 Decentralised ownstate RELNAV system configuration 113
Figure 6-2 Comparison of the centralised and decentralised RELNAV EKF rms errors of the RELNAV system states 114
Figure 6-3 Comparison of the centralised and decentralised RELNAV EKF rms errors of the first member aircraft 115
Figure 6-4 Comparison of the centralised and decentralised RELNAV EKF rms errors of the second member aircraft 115
Figure 6-5 The performance of the decentralised RELNAV EKF of the second member aircraft when measurements from the NC alone are used 117
Figure 6-6 The effect of resampling measurements on the performance of the decentralised RELNAV EKF 117
Figure 6-7 JTIDS RELNAV operational concepts 119

Figure 6-8 Integrated INS/GPS/ADS-B system operational concepts 120
Figure 6-9 Covariance-based hierarchy rank reversals 121
Figure 6-10 Integrated INS/GPS/ADS-B system configuration 123
Figure 6-11 Organisation of the accuracy levels, quality indices and rank assignment in the 8-rank and 15-rank source selection functions 124
Figure 6-12 Comparison of the performance of the integrated INS/GPS system and INS/GPS/ADS-B system for one visible GPS satellite with the 8-rank source selection function before filter tuning 127
Figure 6-13 Comparison of the performance of the integrated INS/GPS system and INS/GPS/ADS-B system for one visible GPS satellite with the 8-rank source selection function after filter tuning 128
Figure 6-14 Comparison of the performance of the integrated INS/GPS system and INS/GPS/ADS-B system for no visible GPS satellites with the 15-rank source selection function 129
Figure 6-15 Comparison of the integrated INS/GPS/ADS-B system performance for different number of visible GPS satellites 130
Figure 7-1 Divergence of the likelihood function 133
Figure 7-2 Likelihood function and Kalman filter residual monitoring for GPS failure detection. 138
Figure 7-3 Cumulative chi-squared distribution 140
Figure 7-4 Chi-squared-based FDI algorithm based on likelihood function 143
Figure 7-5 Mean of the chi-squared test statistic with no GPS failure 144
Figure 7-6 Distribution of the chi-squared test statistic with no GPS failure 145
Figure 7-7 Mean of the chi-squared test-statistic with GPS (SV4) failures 146
Figure 7-8 Distribution of the chi-squared test statistic with GPS (SV4) failures 146
Figure 7-9 Limitations of the chi-squared test statistic 147
Figure 8-1 Separation between a pair of aircraft in horizontal plane 151
Figure 8-2 Conflict probability estimation utilising separation estimation errors 153
Figure 8-3 Aircraft conflict probability estimation process 155
Figure 8-4 Conflict probability as a function of estimated separations and separation estimation error variance 157
Figure 8-5 Effects of estimated separations and geometry on the determination of aircraft conflict probability 160
Figure 8-6 Flight trajectories for the evaluation of aircraft conflict probability 161
Figure 8-7 Predicted rms errors of the stand-alone INS system 162

Figure 8-8 Aircraft conflict probability calculations based on the Kalman filter prediction-equation 162
Figure 8-9 Aircraft conflict detection algorithm in three-dimensional space 164
Figure B-1 Inertial sensor noise samples 183
Figure B-2 The comparison of the true and the theoretical inertial sensor noise characteristics 184
Figure B-3 The clock noise samples and rms error bounds 184
Figure B-4 The simulated navigational accuracy of a stationary strapdown INS and onboard clock (36-hour period) 185
Figure B-5 The simulated GPS satellite constellation defined in the earth frame in a 24-hour period 186
Figure B-6 Ground traces of the simulated GPS satellites as viewed in the simplified airspace in a 24-hour period 186
Figure B-7 GDOP of the simulated GPS satellites in a 24-hour period 187

LIST OF TABLES

Table 2-1 Various elements of an extended squitter 18
Table 2-2 Comparison of technical characteristics of candidate ADS-B datalink systems 22
Table 2-3 Estimated SPS C/A-Code pseudorange error budget 24
Table 2-4 Navigation Uncertainty Category – Position 27
Table 2-5 Navigation Uncertainty Category – Velocity 28
Table 2-6 ADS-B model parameters 28
Table 3-1 Statistical properties of the first-order Gaussian Markov and the random walk process 37
Table 3-2 The summary of the extended Kalman filter algorithms 42
Table 4-1 Strapdown INS sensor parameters 58
Table 4-2 Typical power spectral density (PSD) coefficients for various timing standards 61
Table 4-3 Sensor parameters for navigation system simulation 69
Table 4-4 Characteristics of stand-alone INS and GPS 76
Table 7-1 Demonstration of testing of goodness-of-fit using chi-squared distribution 140
Table 8-1 Characteristics of the conflict probability as a function of the estimated separation and the separation standard 158

NOTATION

$E[X]$	Expectation of a random variable X
$\mathrm{Var}[X]$	Variance of a random variable X
$\mathbf{A}^{\mathrm{T}}$, $\mathbf{a}^{\mathrm{T}}$	Transpose of a matrix $\mathbf{A}$ or a vector **a**. A matrix is represented by boldfaced capital letters. A vector is represented by boldfaced small letters.
f-frame	f-frame of reference. All frames of reference are assumed to be orthogonal, right-handed, Cartesian rectilinear coordinate systems.
i-frame	The simplified airspace model frame of reference
b-frame	The body frame of reference
n-frame	The navigation frame of reference
$\mathbf{v}_e^f$	Vector $\mathbf{v}$ given with respect to the e-frame and expressed in the f-frame. If the f-frame has axes x, y and z, $\mathbf{v}_e^f = \{[v_x, v_y, v_z]_e^f\}^{\mathrm{T}}$ in component form.
$\|\mathbf{v}^f\|$	The magnitude of vector $\mathbf{v}^f$, i.e. $\|\mathbf{v}^f\| = (v_x^2 + v_y^2 + v_z^2)^{\frac{1}{2}}$
$\dot{\mathbf{v}}_e^f$	Time derivative of vector $\mathbf{v}$ with respect to the e- frame expressed in the f-frame.
$\boldsymbol{\omega}_{ef}^f$	Angular velocity vector of the f-frame with respect to the e-frame expressed in the f-frame.
$[\mathbf{v}^f\times]$	Skew-symmetric matrix form of vector $\mathbf{v}^f$ where $[\mathbf{v}^f\times] = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix}$.

$\mathbf{C}_f^e$	Direction cosine matrix that transforms a vector from the f-frame to the e-frame, i.e. $\mathbf{v}^e = \mathbf{C}_f^e \mathbf{v}^f$.
$\mathbf{r}$	A position vector
$\boldsymbol{\phi}_{A_1A_2}^{A_2}$	A rotation vector describing the rotation of the A_2-frame relative to the A_1-frame
$\beta(\cdot)$	Brownian random process
$\boldsymbol{\Phi}(t_2,t_1)$	The state transition matrix that transforms state vectors from time t_1 to t_2.
$\lvert\mathbf{A}\rvert$	Determinant of matrix **A**
$\mathbf{I}$	Identity matrix
$\mathbf{A}_k$ $\mathbf{a}_k$	Matrix **A** or vector **a** at time t_k
$\mathbf{Q}$	The input (process) noise covariance matrix
$\mathbf{R}$	The measurement noise covariance matrix
$\mathbf{H}$	The measurement matrix
$\mathbf{w}$	The input (process) white noise
$\mathbf{v}$	The measurement white noise
$\mathbf{P}$	The estimation error covariance matrix
$\mathbf{z}$	The measurement vector
$\mathbf{K}$	The Kalman gain
$\hat{\mathbf{x}}^+$	The estimated state vector
$\Delta\hat{\mathbf{x}}^+$	The estimation error vector
$\hat{\mathbf{x}}^-$	The predicted state vector
$\Delta\hat{\mathbf{x}}^-$	The prediction error vector
τ	A time constant
σ	A standard deviation
ρ	Pseudorange measurements
ψ	Aircraft heading
$\boldsymbol{\alpha}$	The angle vector of the misalignment from the relative grid

NC	The navigation controller
n	The n^{th} member of the navigation community
$[U,V,W]^T$	The relative grid origin
$L(\mu,\sigma)$	The likelihood function of a normal random variable with mean μ and variance σ^2
$\chi^2_{\gamma=N}$	A chi-squared random variable with N degree of freedom

CHAPTER 1 INTRODUCTION

1.1 CONVENTIONAL AIR NAVIGATION SYSTEMS

Safety and efficiency of operation have always been the foremost concerns for the civil aviation community. To ensure safe and efficient flight, pilots must be able to navigate precisely over any terrain, in all weather and ambient lighting conditions, during all phases of flight (takeoff, climb, enroute, approach and landing). Furthermore, it is necessary to be able to control the flow of traffic to ensure safe aircraft separation and minimise the risks of aircraft collision both on ground and in the air. Air Navigation Systems (ANS) utilises communication, navigation, surveillance (CNS), and air traffic management (ATM) systems to afford these primary functions – navigation and control.

For the past 50 years, line-of-sight navigation systems such as VHF Omni-directional Range (VOR) and Distance Measuring Equipment (DME) have been used for navigation over large regions of continents. Nevertheless, these conventional systems have limited capability. Traditionally, pilots need to fly directly over ground-based radio beacons, following published airways (Figure 1-1), which usually results in additional operating costs due to delays and extra fuel burn. To provide sufficient coverage, large networks of radio beacons are also required owing to the limited range of these beacons, i.e. 200 nautical miles (nm), leading to high costs needed for air traffic service providers to maintain these systems. Furthermore, these systems provide limited accuracy. For example, because VOR uses bearing measurements, the accuracy deteriorates as a function of the distance to a radio beacon, where a typical bearing error of 1° gives rise to 2nm positioning error at 120nm.

Guidance systems such as Instrument Landing System (ILS) were developed to assist pilots during approach and landing in limited visibility conditions. However, ILS provides limited coverage (20nm within 10° of the localiser signal).

Long-range navigation systems that operate at HF frequencies such as LORAN-C and OMEGA have been developed to enable navigation beyond the coverage of line-of-sight systems. Nevertheless, these long-range systems afford poor accuracy owing to the susceptibility of HF signals to propagation anomalies such as

interference and fading. In addition, the availability of these systems still depends on the location of ground-based navigation aids. For example, the range of LORAN-C is 1400nm with an absolute accuracy of 0.25nm.

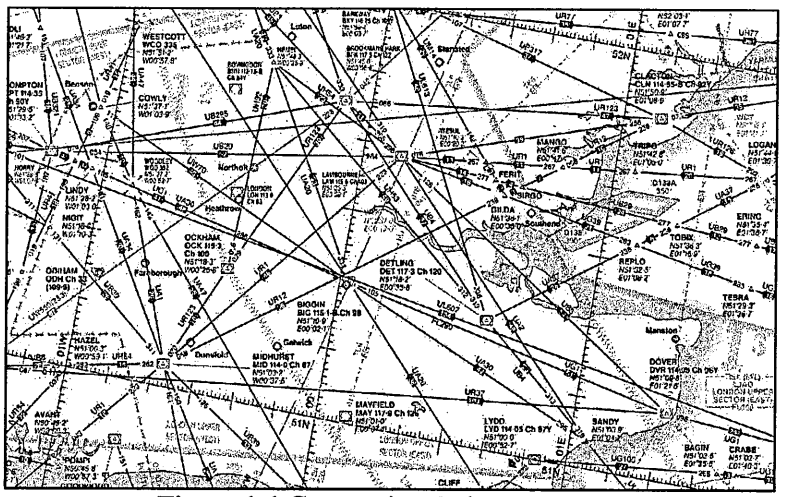


Figure 1-1 Conventional airway system

Self-contained systems, such as Inertial Navigation System (INS), have been developed to eliminate the reliance on ground-based navigation aids. INS is a form of dead reckoning navigation, based on the integration of sensor measurements to determine positions. Owing to error accumulation from the integration process, the performance of INS deteriorates with time. A typical INS installed in an airliner drifts at a rate of 1nm/hr.

In practice, several types of navigation systems must be installed to enable navigation in all phases of flight, which can lead to high costs of certification, installation and maintenance. Additionally, because these conventional systems afford poor accuracy, large values of aircraft spacing are mandatory to allow for the uncertainty in determining positions, which restricts the airspace capacity.

With the advances in digital computing technology, the need to optimise aircraft performance has led to the development of Flight Management Systems (FMS). A FMS enables flight crew to enter flight plans into flight management computers and modify these plans in-flight according to changes during the flight. A FMS monitors the flight envelope and automatically provides optimal guidance and control of an aircraft, based on knowledge of the performance characteristics of the aircraft and a combination of data from various navigation systems. A FMS affords highly advanced airborne

automation systems capable of utilising computerised flight planning and automatic flight control systems to reduce pilot workload. The ability to combine data from various sources introduces an alternative method of navigation known as Area Navigation (RNAV).

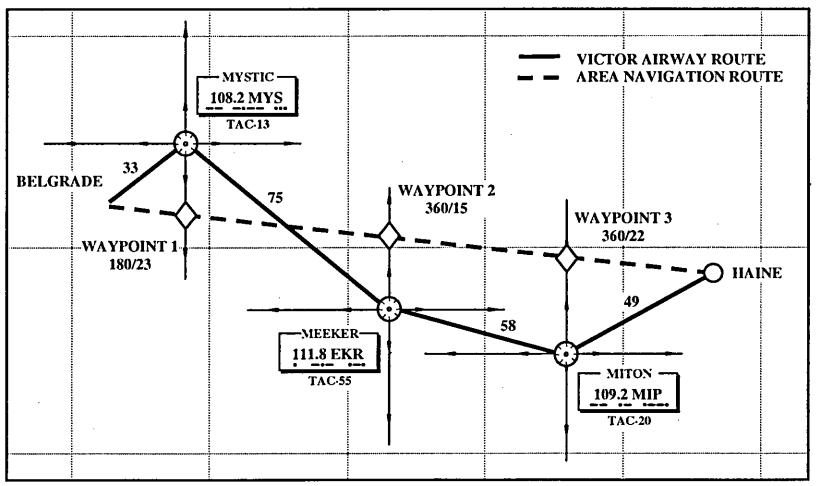


Figure 1-2 Area Navigation (RNAV) concept

RNAV automatically determines aircraft positions using data from one, or a combination, of navigation systems to enable an aircraft to navigate without the need to follow a fixed airway structure formed by networks of ground-based navigation aids, as shown in Figure 1-2. A RNAV system allows a flight crew to fly more direct and economical routes. Although VOR and DME can support RNAV, they can not support high levels of accuracy unless a large number of radio beacons are available. Self-contained systems such as INS can support RNAV, but these systems are expensive and require position updating by external systems to maintain high accuracy.

In addition to the requirement for accurate navigation, Air Traffic Control (ATC) systems are provided to ensure the safest and most efficient use of airspace. The present ATC systems have been built on conventional CNS systems and avionics equipment developed since 1945, when digital technology was still at an early stage of development. Subsequently, these ATC systems have not made full use of automation technology. Furthermore, air traffic controllers have to rely on VHF and HF radio (voice), and primary and secondary surveillance radar (PSR and SSR), which afford limited accuracy and coverage, as the primary means of communicating with flight crew and to manage air traffic flows.

In most cases, commercial aircraft need to operate under Instrument Flight Rules (IFR), to support all weather operations. IFR requires that pilots observe procedural controls and maintain radio contact with controllers during their entire flight. Procedural control starts with the pilots filing a flight plan and receiving a clearance for the flight, which involves extensive exchange of information (flight plans, meteorological forecasts, frequencies for radio and transponder use, notice to airmen, departure instructions, etc.) using radio. When the aircraft is ready for departure, a ground controller issues instructions for the aircraft to takeoff. Once airborne, the ground controller transfers (hand-off) the tracking of the aircraft to a departure controller, who provides heading and altitude instructions for the flight crew to climb to a specific altitude for the enroute phase of flight. Once enroute, the tracking is then transferred to enroute controllers.

During enroute flight, the controllers monitor the progress of aircraft in their airspace on radar displays, to separate aircraft and issue route or speed changes when needed to avoid bad weather or to maintain an efficient flow of traffic. As the aircraft flies out of range and toward another enroute control centre, the tracking controller hands off the aircraft to a controller at the next centre, who in turn monitors the aircraft as it continues on its flight path. This handing off process continues until the aircraft nears its destination, where the tracking is transferred to an approach controller, who is responsible for lining inbound aircraft up for the runway. Once the aircraft is properly spaced, a ground controller takes over and issues the landing instructions.

Normally, communications between controllers and flight crew depends on the use of radio, which is inefficient and restricts the capacity of the present ATC systems. The nature of voice communication limits the speed and volume of information transfer and also the opportunity to automate ATC systems using digital computers. Radio affords poor datalink management, and does not support automatic detection and correction of errors, which can easily be caused by noise on radio channels or problems of language difference. Long-range communication, which depends on HF radio, is also highly vulnerable to interference and fading.

Air traffic surveillance is inefficient because the accuracy and coverage of PSR and SSR are limited. PSR lacks the capability to identify unique aircraft because it uses

only range and bearing measurements from electromagnetic pulse echoes. Although SSR has been designed to overcome these limitations, the capacity of SSR can be affected, especially in high-traffic density airspace, owing to the use of an active transponder interrogation-reply scheme. The range of PSR and SSR is limited to 80nm and 250nm respectively. Beyond this coverage, controllers must depend on position reports by flight crew, derived from airborne long-range (and/or self-contained) navigation systems, using HF radio. However, this method is inefficient due to the limitations of HF radio and the inaccuracy of long-range navigation systems.

An enroute control centre is divided into several sectors, responsible for controlling traffic within specific airspace. Each sector may consist of as many as five controllers, each performing different tasks such as monitoring and separating traffic, handling flight data, co-ordinating with flight crew and controllers in other sectors, and managing the flow of traffic to minimise delays. Because the controllers perform these tasks with limited support of automation systems, the present ATC systems have limited capacity and are prone to human error. Owing to the limitations of conventional CNS systems and in order to limit the workload, controllers usually have to enforce procedural separation, which often leads to congestion, delays and reduction in airspace capacity. Automation systems are under development to help reduce controllers' workload including automatic conflict prediction and resolution systems and data link systems. Figure 1-3 illustrates the conventional ANS operational concept.

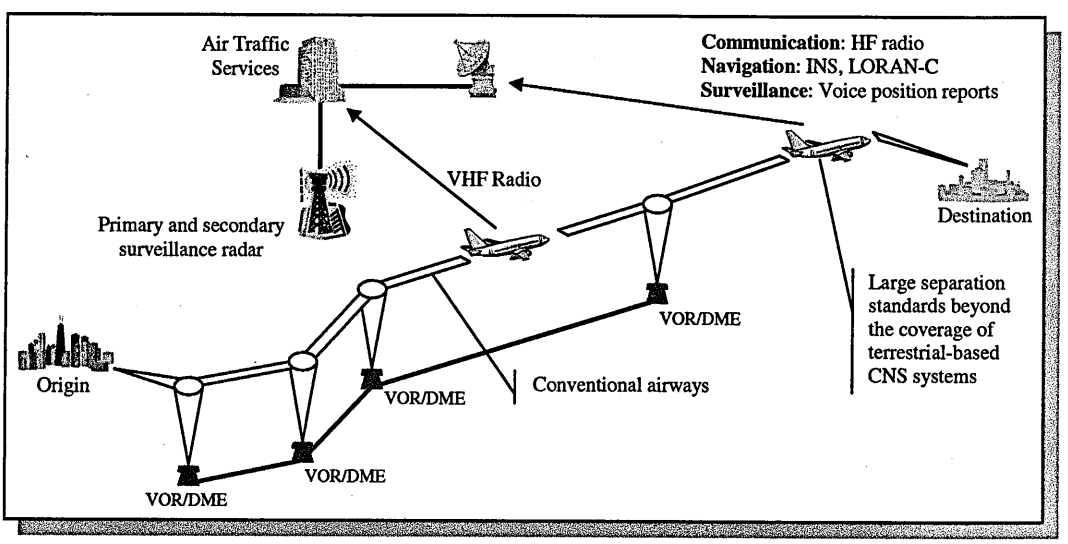


Figure 1-3 Conventional Air Navigation Systems

Despite air traffic controllers' efforts to ensure safe aircraft separation, midair collisions can occasionally occur. This has led to the development of Traffic Alert and Collision Avoidance System (TCAS). TCAS monitors the positions of nearby aircraft and determines if there is a likelihood for collision, independent of ATC system. If TCAS predicts a potential collision risk, it alerts the pilots automatically and issues conflict resolution advisories to avoid a collision. Nevertheless, because only range and range-rate are used in the calculation, TCAS is also prone to issuing unnecessary or false alarms and is limited to issuing vertical resolution advisories (altitude changes). Despite this problem, the introduction of TCAS fulfils an important role in terms of safety and redundancy (backup systems) in civil aviation.

In summary, insufficient use of automation and the limited coverage and accuracy of terrestrial-based CNS systems have limited the capacity of conventional ANS and prevented the capacity of advanced airborne systems such as FMS from being fully exploited. With increased air traffic density, this can lead to congestion and delays, and it has been estimated that the airline industry loses US$5 billions annually due to traffic delays alone.[1] It is widely accepted that world air travel will grow annually at 4-6% and, according to Boeing's report[2], over 20,000 new aircraft will be needed by the year 2018 to meet with the increased air travel demand (Figure 1-4). To accommodate this growth in air traffic and to maintain the competitiveness of the airline industry, it is crucial to increase the efficiency of ANS.

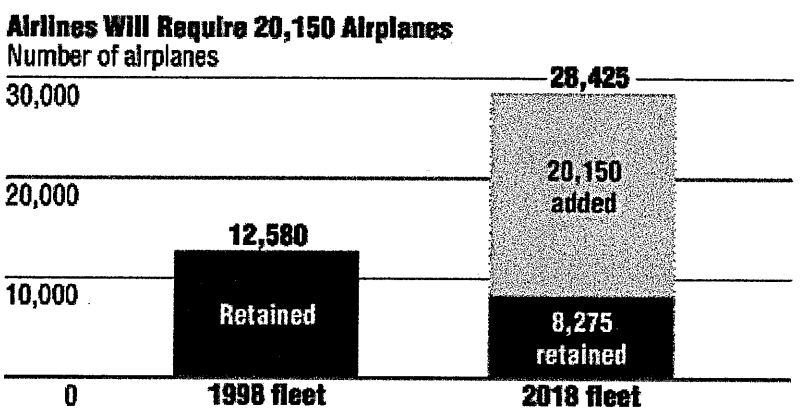


Figure 1-4 Comparison of the world fleet in 1998 and 2018[2]

1.2 NAVIGATION SYSTEMS INTEGRATION

An aircraft is usually equipped with several navigation systems to operate during different phases of flight and to provide redundancy. In practice, these systems will have different performance characteristics. Combining measurements from complementary navigation systems can make use of additional information to provide a more accurate and reliable system. For example, INS affords very accurate short-term navigation solutions at a high update rate. LORAN affords inferior accuracy to INS at a lower update rate but its performance does not deteriorate with time. In principle, measurements from LORAN sensors can be used to realign an INS at regular intervals, increasing the accuracy of navigation while eliminating the error growth, as shown in Figure 1-5. This concept is known as navigation systems integration.

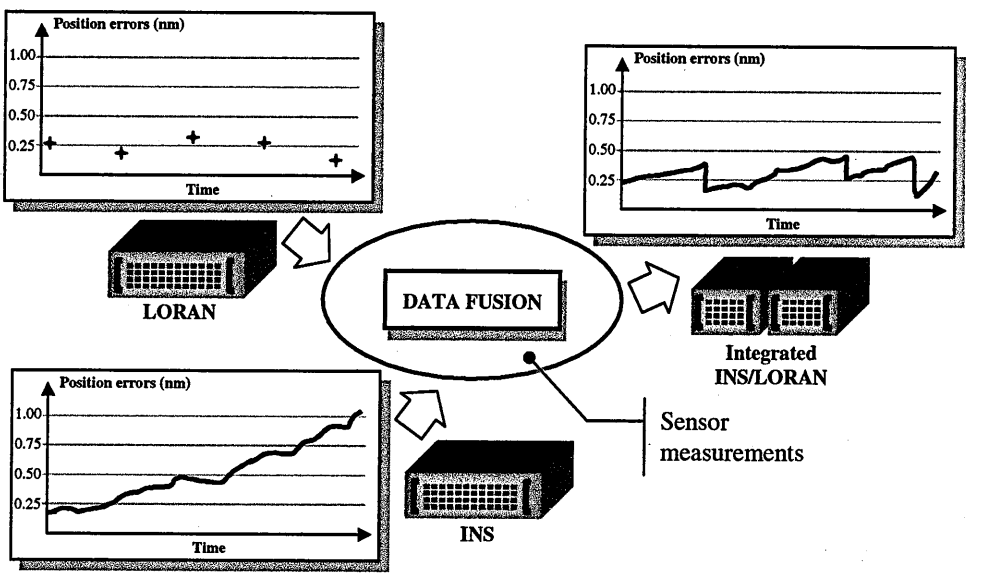


Figure 1-5 Integrated INS/LORAN system concept

The performance of every navigation system depends on the stochastic properties of the sensor signals. To combine measurements from these sensors optimally, it is necessary to apply data fusion techniques which take these stochastic properties into consideration. Kalman filtering has been extensively applied to integrate navigation systems for three reasons. Firstly, a Kalman filter is an optimal estimator. Secondly, it is based on time-domain analysis, which can accommodate a wide range of real-world problems. Thirdly, a Kalman filter is recursive in nature, which makes it straightforward to implement using digital computers. In the foreseeable future, Kalman

filtering will continue to serve as a principal data fusion algorithm used in navigation systems integration.

1.3 ABSOLUTE AND RELATIVE NAVIGATION

The ability to navigate accurately is as crucial to military organisations as it is to civil aviation. The US Department of Defense has developed the Global Positioning System (GPS), which is a satellite-based system, to overcome the limitations of conventional terrestrial-based systems. GPS uses range measurements derived by extracting timing information embodied in navigation signals to determine the position of a GPS receiver. GPS affords very high accuracy, in comparison to conventional radio navigation aids, and eliminates the needs to rely on ground-based beacons. GPS is capable of determining positions with the same high accuracy level, typically within 100 metres (m), anywhere on the earth's surface in all weather conditions. Because of these capabilities, GPS has had a significant impact on the evolution of civil ANS.

Besides accurate navigation, military organisations also need to be able to communicate efficiently. Link-16 provides this capability for NATO armed forces. This system is based on a communication architecture known as Joint Tactical Information Distribution System (JTIDS). JTIDS incorporates precise time synchronisation needed for datalink management, allowing precise relative ranges to be measured between members in a communication network. These relative range measurements are used in JTIDS to precisely determine the position of all members relative to each other, giving rise to a method of navigation known as Relative Navigation (RELNAV).

GPS and RELNAV are similar in principle. However, GPS is an "absolute" system since it determines positions with respect to the earth coordinate frame. On the other hand, RELNAV is a "relative" system since it determines positions with respect to a reference datum common to every member in a navigation community, without referring to external (absolute) sources. Armed forces need RELNAV to alleviate target registration problems such as two fighters reporting the same target at two different locations. In civil ATM, the ability to determine accurate relative positions can also be valuable, for example, in maintaining safe aircraft separation when there are disruptions to operations of absolute navigation systems such as GPS.

1.4 THE FUTURE AIR NAVIGATION SYSTEMS (FANS)

As already discussed, it is necessary to enhance the capabilities of CNS and ATM systems which form the basis of ANS to accommodate the expected increase in air traffic, improve safety, and reduce operating costs. This has led the International Civil Aviation Organisation (ICAO) to undertake the development of future ANS, known as FANS or CNS/ATM, based on the application of digital and satellite-based technologies.

In FANS, the Global Navigation Satellite System (GNSS), which will be based on similar basic principles to GPS, will be used as a principal navigation system. Airspace users can use relatively low cost receivers to determine their position very accurately (better than 10m) in all weather conditions anywhere in the world during all phases of flight. Conventional navigation systems such as VOR/DME will eventually be phased out, eliminating the overheads needed to maintain these radio beacons and to install several types of navigation systems in an aircraft. Another crucial advantage of GNSS is that it can be combined with INS to provide high performance RNAV, superior to traditional terrestrial-based navigation systems.

FANS will take advantage of digital datalink and satellite communication technologies to create a global communication network for the international aviation community, known as the Aeronautical Telecommunication Network (ATN). ATN will provide efficient information transfer between various ATC units and aircraft; for example, by providing controller-pilot data link communications (CPDLC). CPDLC enables standard ATC clearances and instructions, normally communicated by voice, to be carried out automatically using a digital datalink. Future communication systems, such as CPDLC, will be designed to minimise flight crew intervention during normal operations. In addition to reducing the workload of both pilots and controllers, a digital datalink is more efficient because it supports higher data rates, automation, and other datalink management capabilities, such as error detection and correction. For communication beyond the coverage of line-of-sight systems, satellites will be used to relay information, alleviating the problems associated with HF radio. The introduction of GNSS and digital datalinks also leads to a new concept of air traffic surveillance, known as Automatic Dependent Surveillance (ADS).

ADS is a surveillance function in which aircraft automatically transmit a state vector, containing position, velocity, and intention obtained from airborne systems (such as FMS and GNSS), to air traffic controllers on a regular basis in accordance with the established contract. ADS-Broadcast (ADS-B) is a special mode of ADS in which aircraft periodically broadcast their state vectors to controllers and nearby aircraft, without specific contracts. ADS permits enhanced ground and airborne situational awareness, especially, beyond the coverage of PSR and SSR. ADS can be used to provide accurate surveillance for many ATM applications, such as flight tracking, trajectory prediction and conflict detection and resolution. Figure 1-6 illustrates the FANS operational concept.

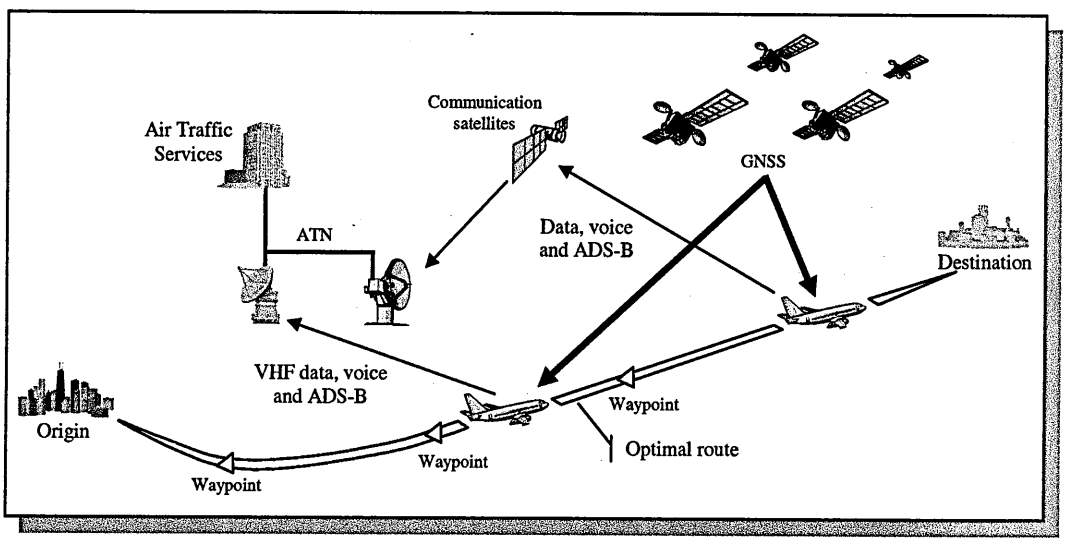


Figure 1-6 Future Air Navigation Systems

FANS offers enormous benefits to the international aviation community. For example, since 1995 Boeing has been installing "FANS-1" in aircraft, such as the B747 and the B777, to take advantage of digital and satellite-based technologies. FANS-1 makes primary use of GPS for navigation and VHF and satellite digital datalink to establish two-way datalinks between flight crew and controllers, thereby providing both ADS and CPDLC. Consequently, FANS-1 allows flight crew to adjust their course in-flight (based on obtained clearances) to fly more direct and fuel-efficient routes, especially over oceans, reducing flight time, delays, and operating costs. Furthermore, FANS-1 has already led to the reduction of mandatory separation between aircraft for oceanic flight from 100nm, 15min and 2000ft (lateral, longitudinal and vertical) to 30nm, 30nm and 1000ft, providing an increase of more than 50% in airspace capacity.[3]

Ultimately, airlines hope to gain total freedom for pilots to choose their own routes and to change these routes in-flight, safely and efficiently, with minimum constraints applied by ATC, which has led to the development of Free Flight concept.[88] Ensuring safe aircraft separation will be the most critical aspect of Free Flight. Under Free Flight, each aircraft will be surrounded by imaginary protective volumes, whose dimensions are derived from aircraft state vectors and CNS capabilities. Pilots will be allowed to manoeuvre freely provided that these protective volumes do not overlap. In flight conditions, when these volumes actually overlap, air traffic controllers will be required to intervene to ensure safe aircraft separation; for example, by issuing speed or course changes.

The controllers will monitor airspace usage, based on the use of decision support tools such as a conflict probe which can automatically predict flight paths, detect conflict, warn the controllers of collision risks, and provide resolution advisories to flight crew. An airborne Separation Assurance System (ASAS) will also be needed in aircraft to provide situation awareness and to assist flight crew in maintaining safe aircraft separation. ADS-B is likely to play a critical role in providing the information necessary for air traffic surveillance, conflict detection and conflict resolution co-ordination.

1.5 THE LIMITATIONS OF THE PROPOSED FANS IMPLEMENTATION

The needs and the benefits of FANS (and Free Flight) are apparent. Nevertheless, the implementation of FANS is complex because it covers diverse disciplines including air traffic management, communications, navigation, and surveillance. There are several institutional and technical issues in implementing FANS.

FANS will require changes in the rudiments of ATC from controlling to managing, based on the use of automation and decision support systems, and several ATC procedural changes. There are also several major implications for aircraft manufacturers and operators:

⇒ Attention must be given to the human factor issues to ensure that flight crew can operate in FANS environment.

⇒ The cost of development and installation of FANS equipment (both ground and airborne) is significant.

⇒ The certification of ATM system is complex, time consuming and costly.

⇒ Airlines will need to be confident that the investment in this new technology can be recovered within a few years.

This thesis focuses on solving one of the fundamental and critical technical aspects which underpin FANS - aircraft separation assurance and system integrity.

Based on concepts of FANS, GNSS constitutes an ideal navigation system. The ultimate goal is to develop GNSS, which is managed and controlled by the international aviation community, and certifiable as a sole-means of navigation for all phases of flight including precision approaches. GNSS is a long-term solution, which must overcome several political difficulties particularly agreements on who will pay for, develop, and control the system.

For the next few years, particularly before GNSS is fully implemented, GPS will serve as an initial building block for satellite-based navigation in FANS. Although the Russian GLONASS system also provides a satellite-based navigation system, it has been poorly maintained and, at present, only 9 of the 24 satellites are functioning properly.[84] For these reasons, GPS is a vital component for the development and implementation of FANS.

Nevertheless, it is commonly accepted that stand-alone GPS can not meet the stringent requirements of civil aviation operations for all phases of flight, in terms of the limited accuracy[i] and integrity[ii] of GPS.[6,77] For example, it has been shown that there can be periods of up to 40 minutes, twice per day, at certain geographic locations where there is insufficient satellite coverage to provide the geometry to maintain the required integrity.[78] At times, it can take more than 30 minutes before users are informed of GPS

---

[i] Accuracy is defined as the capability of a system to maintain aircraft position within error limits.

[ii] Integrity is defined as the capability of a system to provide a timely warning when the system should not be used for navigation.

errors.[79] The issue of the susceptibility of GPS to intentional and unintentional interference has also raised considerable concerns, especially, after interference problems from low cost portable GPS jammers have been reported.[77,80,81] In addition, GPS outages can be caused by the effects of solar activity on the ionosphere, which usually has unpredictable effects.[82,83]

The vulnerability of GPS to disruptions has raised considerable concern worldwide in using GPS (and other satellite-based systems) as a sole means of navigation. Momentary losses of GPS signals can jeopardise the safety of numerous aircraft simultaneously; a problem commonly referred to as a single point failure.[89] It is essential to ensure safe aircraft separation, when a disruption to GPS occurs, which is a critical aspect of FANS and Free Flight.

So far, the approach to improving the accuracy, integrity, continuity, and availability of GPS has been based on augmentation systems (differential techniques), where navigation signals from additional satellites and/or ground-based stations are transmitted to GPS users.[5] However, augmentation systems afford limited coverage, for example, within European or US mainland. Moreover, these augmentation systems have high costs of implementation. Consequently, augmentation systems are not generally considered as a viable long-term solution. Concerns about the high cost of acquiring augmentation systems and the vulnerabilities of GPS have led the FAA to re-examine available navigation alternatives.[7]

In Free Flight, aircraft will use ADS-B to broadcast their GNSS positions and intentions, which will inevitably contain a certain degree of uncertainty, caused by sensor errors and other navigation system errors. The effect of this erroneous information must be minimised in flight path prediction, conflict detection, and resolution generation. In other word, sensor data errors must be filtered in the computation of any navigation or guidance, which are based on ADS-B reports, used to detect flight path conflicts and ensure safe aircraft separation.

As already discussed, RELNAV is a method of navigation in which relative positions of every member in a navigation community are determined independent of external sources. This approach can be applied to ensure safe aircraft separation even

when there is a total loss of GPS. To promote safe and efficient flight under FANS and Free Flight, the research described in this thesis investigates these technical difficulties and proposes methods of solving them, based on a combination of data fusion techniques and RELNAV principles.

1.6 Research Objectives

The objectives of this thesis are as follows:

1. To develop a set of algorithms for system integration of aircraft navigation sensors for used in RELNAV applications.

2. To develop a suite of test programs in Matlab to test and validate the RELNAV principles and other algorithms developed in this thesis.

3. To investigate a method of improving the integrity based on the combination of GPS, INS, and ADS-B, in particular, to accommodate failure modes.

4. To develop a method for aircraft separation assurance based on data provided in ADS-B broadcasts.

1.7 Scope of the Thesis

This thesis covers two main areas. The first is to examine the feasibility of applying RELNAV principles to support the use of a satellite-based navigation system in civil ATM. The second is to develop a method to ensure safe aircraft separation based on the provision of ADS-B.

Chapters 2-4 review the background knowledge required in the subsequent analysis in this thesis. Chapter 2 examines CNS systems which form the basic building block of FANS and Free Flight. Chapter 3 covers the basis of Kalman filtering and implementation considerations in navigation systems integration. Chapter 4 reviews the modelling, performance analysis, and integration of navigation systems.

Chapters 5 and 6 contain an investigation into the application of RELNAV. In Chapter 5, an optimal (centralised) RELNAV algorithm is derived, based on JTIDS

RELNAV principles. In Chapter 6, a sub-optimal (decentralised) RELNAV algorithm appropriate to practical applications is developed and applied to civil ATM, based on the optimal solution derived in Chapter 5. It will be shown that RELNAV enables ADS-B to provide additional navigation sensors, allowing GPS to maintain its accuracy while degrading gracefully. Chapter 7 investigates sensor failure detection and isolation; a new method is developed, based on a combination of likelihood functions and chi-squared tests. Finally, in Chapter 8, a new conflict detection algorithm capable of considering the uncertainty in position reports is proposed. Monte Carlo simulation is used to validate and evaluate the algorithms developed in this thesis. Chapter 9 contains the summary of the findings and the contributions of this thesis to future development of civil ATM. Figure 1-7 depicts the scope of this thesis.

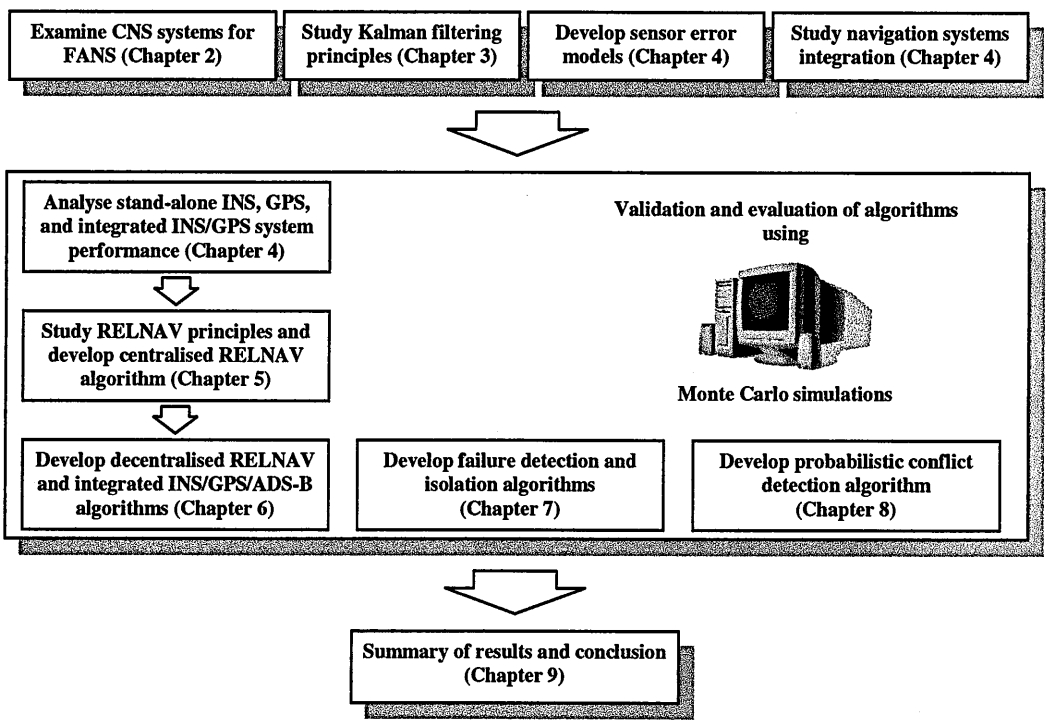


Figure 1-7 Scope of the thesis

CHAPTER 2 AIR NAVIGATION SYSTEMS

2.1 INTRODUCTION

As already discussed in Chapter 1, it is necessary to enhance the safety and efficiency of airspace to accommodate the expected increase in air traffic and reduce operating costs. These developments are made possible, based on advances in sensor, communication, digital, and satellite technologies. The advent of satellite-based navigation systems permits high-accuracy navigation worldwide. The introduction of advanced digital datalinks is expected to alleviate the problems associated with voice only communications. The provision of digital datalink systems allows navigation data to be passed between aircraft, affording new opportunities to increase navigation accuracy, by combining sensors and datalinks.

This chapter outlines the developments of enabling technologies for FANS and Free Flight covering communication, navigation, and surveillance (CNS) systems.

2.2 COMMUNICATION COMPONENT OF FANS

FANS aims to overcome the problems associated with voice communications, based on the application of advanced digital and satellite-based datalink systems. In addition, these digital datalinks will provide the basis of Automatic Dependent Surveillance-Broadcast (ADS-B), which is a vital element of air traffic surveillance in FANS. The following sections explain Mode S extended squitters and VHF Digital Link (VDL) Mode 4, the major candidate datalinks for ADS-B.

2.2.1 Mode S extended squitter

Secondary Surveillance Radar (SSR) has been originally designed to improve the efficiency of air traffic surveillance afforded by primary radar, utilising an active interrogate-reply surveillance scheme. The interrogation and reply messages, which are pulse trains of modulated signals, are transmitted at 1030 and 1090 MHz respectively. Several forms (modes) of SSR are provided, which can be determined from the preamble (timing pulses) embedded in the messages. Two most frequently used interrogations consist of Mode A (altitude) and Mode C (identity). Mode S (selective)

has evolved from the conventional SSR with an enhanced capability to identify a unique aircraft by referring to its 24-bit aircraft address.

A squitter is a broadcast of Mode S messages (56 bits) containing a 24-bit aircraft address and the transponder capability.[69,70] Mode S transponders automatically broadcast one *squitter* per second, which is primarily used for surveillance by Traffic Alert and Collision Avoidance Systems (TCAS). An extended squitter is an enhanced version of the squitter, allowing aircraft to broadcast longer messages (112 bits). The core function of extended squitters is to provide ADS-B using SSR Mode S. For reasons of compatibility, extended squitters use the same standard as SSR Mode S e.g. signal format, RF channel, modulation scheme, etc. Figure 2-1 shows the signal format of an extended squitter.

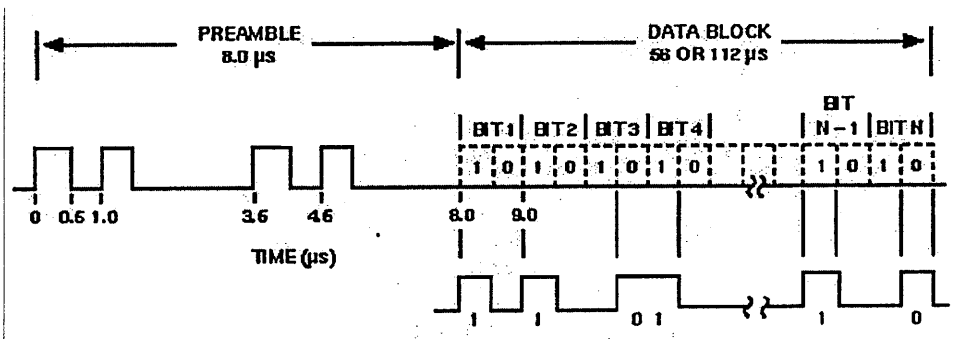


Figure 2-1 Mode S extended squitter message format

Mode S employs Pulse Position Modulation (PPM) technique to modulate the signal. A pulse ($0.5\mu s$) is transmitted either in the first half (indicating a 1) or the second half (indicating a 0) of the data bit period ($1\mu s$). Subsequently, the data rate within a message is 1M bits/sec. The 4-pulse preamble is used for synchronisation, which allows a receiver to detect the beginning of a message.

The position (in latitude and longitude) and velocity of aircraft are transmitted at a rate of 2Hz using position and velocity squitters. The identity of aircraft is transmitted at a rate of 0.4Hz using identity squitters. A data compression method known as Compact Position Reporting (CPR) is used to compress the 24-bits latitude and longitude data to 17 bits. Due to quantisation errors, the least-significant-bit (LSB) of the position reports is accurate to within 5m (1σ). Mode S extended squitter data structure is summarised in Table 2-1.

Table 2-1 Various elements of an extended squitter (the numbers given describe the number of bits required per each element) [76]

Report elements	Type of squitter		
	Position	Velocity	Identity
ICAO address	24	24	24
Latitude	17	-	-
Longitude	17	-	-
Attitude (geography)	12	8	-
Navigation Uncertainty Category[i] (position)	5	-	-
North velocity	-	11	-
East velocity	-	11	-
Altitude rate (geography)	-	10	-
Navigation Uncertainty Category (velocity)	-	3	-
Altitude (barometer)	12	-	-
Altitude rate (barometer)	-	10	-
Air speed	-	11	-
Ground speed, ground track	-	22	-
Turn indication	-	2	-
Call sign	-	-	48

To allow multiple access to the communication channel, Mode S extended squitter uses a pseudo-random transmission scheme, where a timing jitter is applied to the transmission of every message. The jitter, which is uniformly distributed within ±100ms, is randomly selected in such a way that the necessary transmission rates are ensured (2Hz for position-velocity squitters and 0.4Hz for identity squitters).

Because SSR Mode S was not originally designed as a datalink system, there are concerns over its long-term viability, especially, for surveillance in air traffic environments such as Free Flight. Mode S datalink is established based on Pulse Repetition Frequency mechanism, where a number of interrogation messages may have to be transmitted to identify other aircraft. This mechanism limits the capacity of Mode S owing to the signal interference (fruit[ii]) problems. Fruit is reception of the response one aircraft has given during the interrogation of another radar, and which may create an

[i] See Section 2.4 for the definition of the Navigation Uncertainty Category

[ii] False Replies from Unsynchronised Interrogator Transmissions (FRUIT)

erroneous target report. The limited capacity of Mode S is a major problem associated with extended squitters.

Although the range of extended squitters has been increased to 100nm, it has been shown that in high-density airspace it is possible to operate with fruit rates of 40000 fruits per second or higher.[71] Consequently, this has introduced difficulties with Mode S extended squitter receptions. At present, work is in progress to improve reception techniques for extended squitter operations in very high Mode A/C transponders environment. An alternative communication system that has been specifically designed to support FANS is described in the following section.

2.2.2 VHF Digital Link Mode 4

VHF Digital Link (VDL) Mode 4, also known as Self-Organising Time Division Multiple Access (STDMA), is a multi-purpose datalink system specifically designed for FANS.[72,73,74,75] VDL Mode 4 can operate on any of the 25kHz channels between 108-136.975MHz. Global Signalling Channel (GSC) refers to communication channels that provide worldwide standard communication channels and a means to access other datalink services. One VDL Mode 4 transponder can receive from up to four channels; for example, two GSCs and two local channels. The operational concept of VDL Mode 4 illustrated in Figure 2-2.

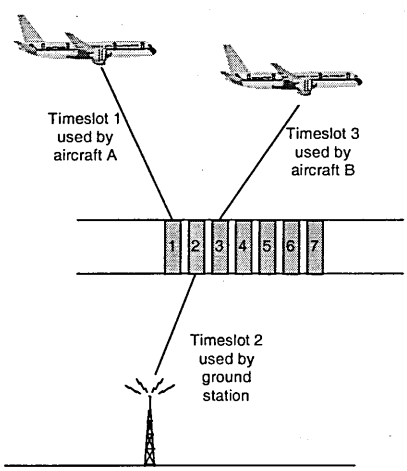


Figure 2-2 TDMA operational concept[74]

In TDMA, the available transmission time is divided into a number of timeslots. At the top level, a *superframe* spans a period of 60 seconds. Gaussian Filtered

Frequency Shift Keying (GFSK), the main modulation scheme for VDL Mode 4, affords 4500 timeslots per superframe with a data rate of 19200 bits/sec/channel within a message. VDL Mode 4 uses Universal Co-ordinated Time (UTC), derived from Global Navigation Satellite System (GNSS), as a primary source for synchronisation, where every 75th slot is synchronised to the UTC seconds. If UTC is not available, synchronisation can also be made with respect to other secondary time sources that can be related to UTC to maintain the continuity of the network.

VDL Mode 4 offers several protocols for establishing and managing the datalink. Periodic broadcast protocol is the most important protocol relating to ADS-B because it enables autonomous access to the network without requiring a master control station (self-organising). In operation, identity, position information, and network control parameters (time-out and offset parameters) can be transmitted completely in a single timeslot. That particular timeslot is then reserved for transmission by that aircraft for a specific number of consecutive superframes indicated by the time-out parameters. The offset parameters indicate which timeslot will be used after a time-out period. Aircraft change timeslots at regular intervals, between 3 and 8 minutes, to reduce the probability of co-channel interference.

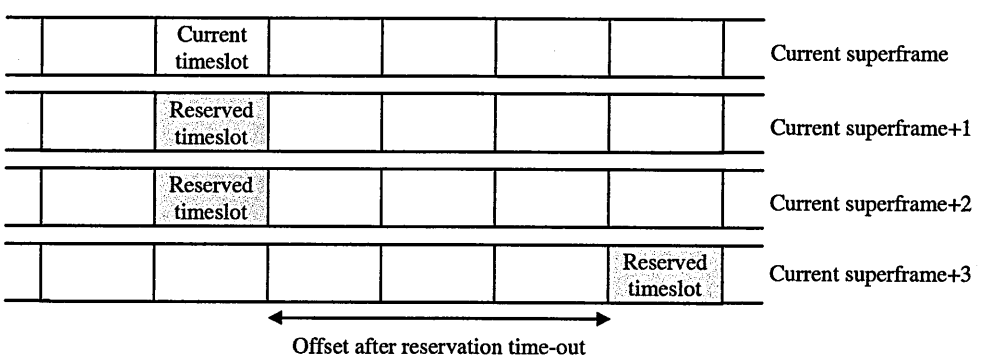


Figure 2-3 Periodic broadcast protocol

A new aircraft wishing to participate in a network can construct a reservation table by monitoring the usage of a channel over a period of 60 seconds (one superframe). First, an aircraft looks for unreserved timeslots. If there are no free timeslots available, timeslots reserved by other aircraft that are at a distance greater than 250nm can be used. This is feasible because of the favourable co-channel interference (CCI) conditions, which allows intended messages to be distinguishable (from messages

of other distance aircraft) when the signal strength differs by more than 10dB. This is equivalent to a factor of 3 in terms of distance. For example, as shown in Figure 2-4, aircraft A intends to transmit to aircraft B using the same timeslot as aircraft C. If the distance between aircraft A and B is *d*, aircraft A will successfully transmit to aircraft B, if the distance between aircraft C and B is greater than $3 \times d$. If more than one free timeslot is available, an aircraft will randomly select a specific timeslot from a range of available slots.

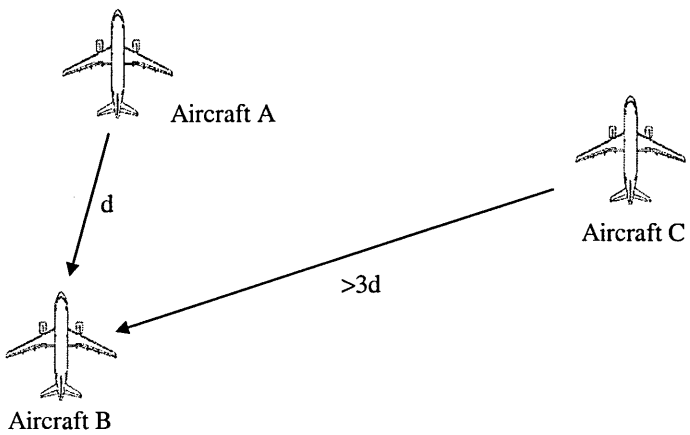


Figure 2-4 Co-channel interference conditions

VDL Mode 4 uses the *synchronisation burst format* to transmit ADS-B messages. One message consists of 256 bits. The first 40 bits are used for power stabilisation (16 bits) and synchronisation (24 bits). The next 192 bits are divided into fixed (56 bits) and variable (136 bits) data fields. These fields contain the information required for ADS-B. The last 24 bits of a message, which contain no useful information, are allocated to allow for propagation delays – affording an effective range of approximately 200nm.

Because VDL Mode 4 requires precise time synchronisation for datalink management, accurate timing information can be extracted from received massages to determine distances (pseudoranges) between aircraft. In VDL Mode 4, this capability gives rise to a concept known as *secondary navigation*, where these pseudorange measurements are used to determine the positions of aircraft. Although secondary navigation affords inferior accuracy to GNSS, it is expected to provide a useful back up of GNSS, especially, during short-term outages. Until now, there have been no technical developments in the field of the secondary navigation concept. In Section 6.4,

navigation algorithms are developed that afford this secondary navigation capability. Table 2-2 summarises the selected technical characteristics of Mode S extended squitter and VDL Mode 4.

Table 2-2 Comparison of technical characteristics of candidate ADS-B datalink systems[76]

Characteristics	Mode S extended squitter	VDL Mode 4
Frequency band	1090MHz	108-136.975MHz
Bit rate	1 Mega bits/sec	19200 bits/sec/channel
Modulation	PPM	GFSK
Synchronisation	4 pulse preamble	First 24 bits
Message length	112 bits	192 bits
Airborne longitude	CPR 17 bits, LSB ~5 metres	Compressed 18-22 bits even 16-20 bits odd, LSB ~1-18 metres
PVT segmentation (Position, Velocity and Time)	Yes: velocity in separate message	No: PVT in one message
Transmission rate for PVT	Position and velocity at 2Hz	PVT at 0.1Hz, En-route PVT at 0.2Hz, Terminal PVT at 0.667Hz with local channel
Transmission rate for identity	2.2 per second	Once every 5 minutes
Multiple access technique	Random message	TDMA
RF channels	One channel	2 (25kHz), Global Signalling Channels plus up to 2 local channels

The initial findings of Link Evaluation Team (LET)[76], a technical committee set up to evaluate datalink systems to provide ADS-B for Free Flight, has concluded that there are important remaining efforts to complete equipment development, certification, and installation that are concerned with widespread ADS-B deployment, independent of link choice. In addition, LET has not recommended any specific systems to be used for ADS-B and work is still in progress to evaluate these datalink systems.

2.3 NAVIGATION COMPONENT OF FANS

FANS aims to overcome the shortcomings of conventional terrestrial-based navigation aids, based on the use of satellite-based navigation systems. Sections 2.3.1 and 2.3.2 describe Global Positioning System (GPS) and Inertial Navigation System (INS), since they form the basis of navigation systems in FANS. The simulation models and the integration of these navigation systems are covered in detail in Chapter 4.

2.3.1 Global Positioning System

GPS has been developed and operated by the US Department of Defense, to provide accurate, worldwide, all-weather positioning system. GPS consists of a constellation of 24 satellites in six orbital planes, approximately 20,200km above the earth's surface at an inclination angle of 55°. Each GPS satellite continuously transmits navigation signals, which can be tracked by a GPS receiver. Based on the received signals, a receiver is able to determine the positions of the satellites that are being tracked and to measure the propagation time of these signals (*pseudorange* measurements). In principle, the user's position and time offset from GPS system time can be determined by solving a set of equations based on simultaneous measurements from at least four GPS satellites.

GPS provides two levels of services - *Precise Positioning Service* (PPS) and *Standard Positioning Service* (SPS). PPS affords a high accuracy navigation capability, but is only accessible by authorised users such as the US military. SPS, which provides inferior accuracy to PPS, is accessible by all other users including the civil aviation community. The inferior performance of SPS is caused mainly by the introduction of Selective Availability (SA) to preserve the US national security interests and the lack of additional signals that are available to the PPS's users to correct for atmospheric delays. SA is the deliberate degradation of satellite data for non-military users, which is achieved by "*dithering*" the satellite clock and introducing errors into the navigation message data. SA (in combination with other error sources, as shown in Table 2-3) leads to navigation accuracy (95%) of ≤100m horizontally, ≤156m vertically and time transfer accuracy of ≤340ns.

GPS satellite transmissions utilise spread spectrum modulation to provide the structure of ranging signals and essential navigation data such as satellite ephemerides and satellite health. The ranging signals for SPS users are 1023-bits Pseudorandom Noise (PRN) codes, known as coarse/acquisition (C/A)-code, that binary phase shift key (BPSK) modulate the satellite carrier L1 frequency (1575.42MHz). The 50 bits per second navigation data is also combined with the C/A-code prior to modulation with the L1 carrier, using the exclusive-or process. Spread spectrum modulation allows a GPS

receiver to detect and separate ranging signals from different satellites, which are transmitted on the same carrier frequency, using a technique called code division multiple access (CDMA). In principle, a replica of the C/A-code, which is unique to each satellite, is compared with a received signal. Once a match has been identified, the signal propagation time, which is needed to form pseudorange measurements, can be determined from the time shift of the replicated C/A-code.

The performance of GPS is predicated mainly on the satellites geometry in relation to the user, and the precision of pseudorange measurements, which can be corrupted by noise from several error sources, as shown in Table 2-3. In order to simulate GPS precisely, it is necessary to develop a simulation model that accurately models the GPS orbital trajectory and pseudorange measurement error sources. The development of the GPS simulation model used in this thesis is described in detail in Section 4.4.1.

Table 2-3 Estimated SPS C/A-Code pseudorange error budget[28]

Segment source	Error source	1σ error (m) with SA	1σ error (m) without SA
Space	Satellite clock stability	3.0	3.0
	Satellite perturbations	1.0	1.0
	Selective availability	32.3	-
	Other (thermal radiation, etc.)	0.5	0.5
Control	Ephemeris prediction error	4.2	4.2
	Other (thruster performance, etc.)	0.9	0.9
User	Ionospheric delay	5.0	5.0
	Troposheric delay	1.5	1.5
	Receiver noise and resolution	1.5	1.5
	Multipath	2.5	2.5
	Other (interchannel bias, etc.)	0.5	0.5
System UERE	Total (rms)	33.3	8.0

As already discussed in Section 1.5, the limited accuracy, integrity, and vulnerability (especially to signal interference) of the system are the major causes of concerns for utilising GPS as a sole means of navigation for civil aviation. Several augmentation systems have been developed to improve the accuracy and integrity of GPS. These augmentation systems operate on a principle known as *differential* GPS

(DGPS). In principle, DGPS estimates the errors of GPS measurements at a geographical location and broadcasts this information to all nearby users. These users are able to reduce their GPS measurement errors by applying differential corrections to their GPS measurements. Although these system augmentations can improve the performance of GPS significantly, they require the installation of new ground stations or satellites, with an effective coverage over specific regions such as the US continent or mainland Europe. The performance of GPS outside these areas is unaffected, e.g. over the oceans or in countries with less advanced ATC systems. To support the use of satellite-based navigation systems in FANS, it is of interests to examine methods, besides differential techniques, to alleviate the problems associated with these systems, which is one of the objectives of this work.

2.3.2 Inertial Navigation System

Inertial Navigation System (INS) was originally developed to provide an accurate, self-contained and unjammable navigation system for military applications. INS began to be installed onboard civil aircraft because of its capability to compute aircraft position continuously, especially for en-route navigation over the oceans where no ground-based navaids exist. Moreover, the outputs of INS can be used by automatic flight control and landing systems. Attitude and heading reference is also derived from an INS to display aircraft attitude on the Electronic Flight Instrument System (EFIS).

In principle, an INS is a form of dead reckoning navigation which determines the velocity, position and attitude of an aircraft by the integration of linear accelerations and angular velocities based on given initial conditions. INS employs a set of inertial sensors, i.e. accelerometers and gyroscopes, to measure accelerations and angular velocities respectively. The older version of INS uses a *stable* platform to support and maintain the orientation of these sensors. In the last 20 years *strapdown* systems have been used, where the inertial sensors are rigidly mounted on the airframe. Digital processors are used to replace the movement of mechanical gimbals to reduce the complexity of mechanical components and manufacturing costs. Nevertheless, both types of INS operate on the same basic principles.

The performance of an INS can be affected by several error sources such as initialisation errors and sensor errors. For example, the errors of a strapdown INS include initial misalignment of the computer platform, scale factor errors and limited word-length arithmetic. Since an INS operates by the integration of measurements, system errors accumulate resulting in deterioration of the system accuracy as a function of time. Based on present inertial sensor technology, a good quality INS used in aircraft navigation typically drifts at a rate of 1nm/hr.

Inertial sensors can be further classified according to the fabrication technology used in manufacture of these sensors i.e. *mechanical* sensors, and *solid-state* sensors (silicon, quartz, Micro-Electro-Mechanical Systems and Micro-Opto-Electro-Mechanical Systems).[85,86] Mechanical sensors refer to those sensors, which are implemented by spinning mass and force rebalance instruments. Because of maturity in this technology, these sensors afford extremely high performance with the disadvantage of high production costs. As advances in fabrication technology are made, it is predicted that the solid-state sensors will be used to replace mechanical inertial sensors in a wide range of applications with improved accuracy and reliability, reduced weight, size and costs. For example, an attempt is already being made to fabricate a set of inertial sensors on a single chip, with the goal of producing a low cost integrated INS/GPS system.[86]

This trend in developments will have significant impacts in all fields of navigation, including aircraft navigation. In this thesis, it is expected that INS will become an essential element of navigation systems for future aircraft as the progress is made in FANS systems and inertial sensor technology. This is because, in addition to its unique characteristics, INS can be integrated with other navigation systems (especially GPS) to enhance navigation performance in comparison with stand-alone systems. The development of INS sensor models used for the navigation simulation studies in this thesis, and the principles of navigation systems integration based on INS, are described further in Sections 4.3 and 4.7.

2.4 SURVEILLANCE COMPONENT OF FANS

ADS-B is an important element of the surveillance concept in FANS. ADS-B is a surveillance function where aircraft automatically broadcasts position and position-

related data derived from onboard navigation systems such as GPS, to other airborne or ground stations. Position-related data includes the information such as aircraft velocity vectors and trajectory change points. In principle, the controllers and pilots of nearby aircraft will be able to receive ADS-B messages and use the information to provide situation awareness, and to detect and resolve conflicting flight paths. ADS-B is expected to become an integral part of future navigation systems including situation awareness, conflict detection and separation assurance to improve the safety and effective use of airspace.

Included in every ADS-B message is the Navigation Uncertainty Category (NUC), which indicates the level of uncertainty of reported position and velocity. The NUC is transmitted so that surveillance applications can determine if the reported position and velocity have an acceptable level of integrity and accuracy for the intended use. The position uncertainty (NUC_P) and velocity uncertainty (NUC_R) can be determined from the horizontal errors, as shown in Table 2-4 and Table 2-5.[87] The horizontal position error is defined as the radius of a circle in the horizontal plane, with its centre being at the true position. This circle describes the region which is assured to contain the indicated horizontal position with 95% probability.

In operation, a number of aircraft can simultaneously broadcast information about their position, including the degree of uncertainty in the information (the NUC). The major difficulty, which is examined in this thesis, is to accommodate the uncertainties of predicted aircraft position and to enable this information to be used in conflict detection.

Table 2-4 Navigation Uncertainty Category – Position

NUC_P	Horizontal position error (95%)
0	Unknown
1	<10nm
2	<5nm
3	<1nm
4	<0.5nm
5	<0.25nm
6	<0.1nm
7	<0.05nm

Table 2-5 Navigation Uncertainty Category – Velocity

NUC_R	Horizontal velocity error (95%)
0	Unknown
1	<10m/s
2	<3m/s
3	<1m/s
4	<0.3m/s

At present, ADS-B has not been fully specified because the system is at an early stage of development. This situation poses several difficulties, particularly in the development of navigation algorithms based on ADS-B, which is the focus of this thesis. In order to develop a model of ADS-B in this thesis, the system parameters have been selected from the Minimum Aviation System Performance Standards (MASPS) for ADS-B developed by the RTCA. These parameters are summarised in Table 2-6.

Table 2-6 ADS-B model parameters

Range	40nm
Update period	3sec
Measurement noise (1σ)	20m

The range and update period values given in Table 2-6 have been chosen for ADS-B to support *both* conflict avoidance and separation assurance applications. (In Chapter 6, it will be shown that navigation based on ADS-B can operate with a longer update period.) The application of ADS-B proposed in this thesis requires the specification of the ADS-B pseudorange measurement noise, which is not available in the MASPS. In this thesis, ADS-B pseudorange measurement noise is modelled as Gaussian (normal) white (uncorrelated in time) noise with variance $(20m)^2$. This value has been selected because it is the permitted horizontal position error (1 dimensional) budgeted for ADS-B in the MASPS specification.

2.5 Traffic Alert and Collision Avoidance System

The primary objective of TCAS is to provide a backup of ATC separation procedures. TCAS has the capability to display relative positions and velocity of aircraft using colour-coded symbols and to warn flight crews of the risk of collision with other aircraft. TCAS operates by interrogating Mode S transponders of nearby aircraft. It

determines the location, speed, and course of each aircraft, based on range and range rate measurements from the transponders replies. Two types of TCAS are in use today – TCAS I and TCAS II. TCAS I is able to indicate the bearing and relative altitude of all aircraft within a selected range and to determine if an aircraft poses a potential collision threat. This constitutes a *traffic advisory* (TA). In addition to the TA, TCAS II is also capable of issuing a *resolution advisory* (RA), which provides advises to the flight crews to climb or descent, in order to avoid another aircraft.

The performance of TCAS is limited owing to the fact that only range-related measurements are used to determine potential threats. TCAS may issue a false alarm because the horizontal miss-distance can not be calculated precisely, based on range-related measurements alone. Moreover, TCAS is based on instantaneous measurements, whereas conflict detection for Free Flight will be based on predicted trajectories (for example, for 10 to 30 minutes in advance). For the above reasons, it is widely accepted that TCAS is not a suitable candidate airborne separation assurance system for Free Flight. Alternative systems and algorithms must be developed specifically for conflict detection and resolution that can make use of the information provided by ADS-B.

In this chapter, the major components of communication, navigation, and surveillance systems, which underpin FANS, have been described. In the following chapter, the basis of Kalman filtering, which is a data fusion technique widely applied to navigation systems integration, is explained.

CHAPTER 3 KALMAN FILTERING THEORY

3.1 INTRODUCTION

In 1960, R. E. Kalman developed a new approach to linear estimation and prediction which has become known universally as the Kalman filter[8]. Since then, it has become a fundamental tool for analysing and solving a broad class of problems where a system is modelled as a set of stochastic processes incorporating measurements corrupted by measurement noise. The Kalman filter has been widely accepted because it is the best filter (estimator) in the class of linear unbiased filters. It has been successfully applied to aircraft navigation where the signals measured by sensors are affected by noise and where modelling of the dynamic of navigation system is necessary. The Kalman filter has been widely used as a useful and practical tool for integrating and analysing the performance of navigation systems.

This chapter reviews the fundamental principles of Kalman filtering theory in order to understand its application for the design and analysis of navigation systems and systems integration. The discussion covers stochastic linear dynamics system modelling and Kalman filtering algorithms. The application of the Kalman filter to more practical problems such as non-linear estimation, filter performance analysis and decentralised filtering, particularly in aircraft navigation, is also explained.

3.2 STOCHASTIC LINEAR SYSTEM MODEL

The application of optimal estimation depends on the ability to accurately model a physical system under consideration. Although a stochastic model incorporates some forms of random processes, the development of the stochastic model is derived directly from deterministic system theories. This section first reviews a continuous time stochastic model for a linear dynamic system (plant). This is followed by the development of an equivalent discrete time plant model and discrete time measurement model which is suitable for the implementation of the Kalman filter algorithm using digital computers. Different methods for deriving discrete time parameters will be discussed. The term *overall* system model is used to refer to the combination of the plant and measurement model. Figure 3-1 illustrates the organisation of a physical

system, an overall system model and the Kalman filter.

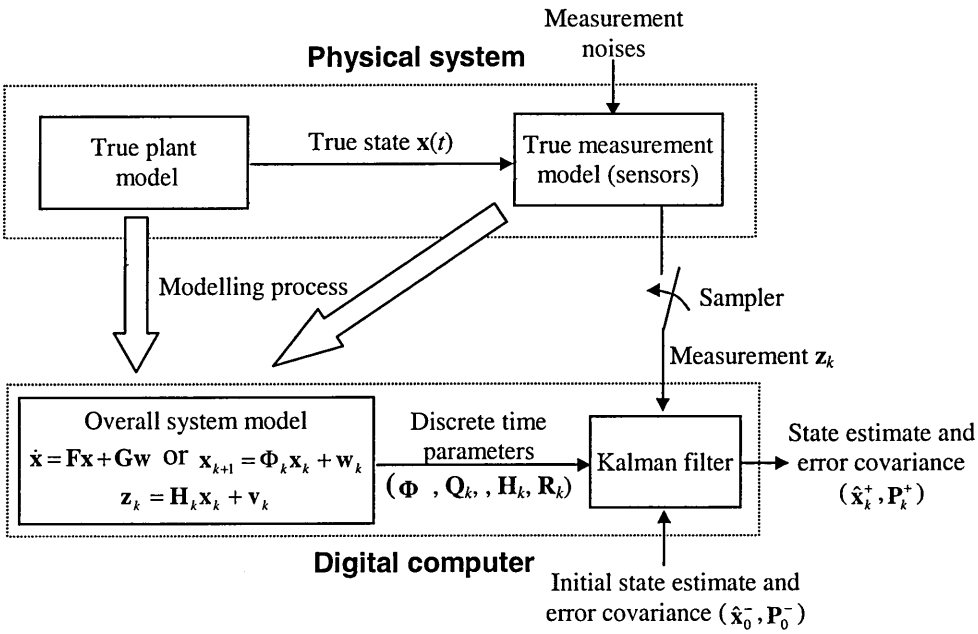


Figure 3-1 Organisation of a physical system, overall system model and Kalman filter.

3.2.1 Continuous time stochastic linear system model

A stochastic linear dynamic system can be represented in state-space notation by a set of linear differential equations as follows:

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{w}(t) \qquad \text{Eq. 3-1}$$

where $\mathbf{x}(t)$ is the system state vector consisting of the set of quantities which completely describe the unforced behaviour of the system. $\mathbf{w}(t)$ is Gaussian white noise associated with the plant with the following statistical properties[i]:

$$E[\mathbf{w}(t)] = \mathbf{0} \qquad \text{Eq. 3-2}$$

$$E[\mathbf{w}(t)\mathbf{w}^T(t')] = \mathbf{Q}(t)\delta(t - t') \qquad \text{Eq. 3-3}$$

where E and δ denote the expectation operator and delta function respectively. The justification for the assumption that the noise is Gaussian white noise will be discussed at the end of this section. To determine the system state at any time t, it can be shown that the analytical solution to Eq. 3-1 is given by: [9]

---

[i] Let ~N[**m**,**C**] also denotes a Gaussian (normal) random vector with mean **m** and covariance **C**.

$$\mathbf{x}(t) = \Phi(t,t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \Phi(t,\tau)\mathbf{G}(\tau)d\beta(\tau) \qquad \text{Eq. 3-4}$$

where β is the Brownian random process defined as the integral of Gaussian white noise with zero initial condition. The stochastic integral, $\int_{t_0}^{t} \Phi(t,\tau)\mathbf{G}(\tau)d\beta(\tau)$, is zero mean and consists of independent Gaussian increments. $\Phi(t,t_0)$ is the state transition matrix that transforms the state vector from time t_0 to t. It satisfies $\Phi(t_0,t_0) = \mathbf{I}$ and $\dot{\Phi}(t,t_0) = \mathbf{F}(t)\Phi(t,t_0)$ where $\mathbf{I}$ denotes the identity matrix. If $\mathbf{F}(t)$ is time invariant, it can be shown that $\Phi(t,t_0) = \Phi(t-t_0) = e^{\mathbf{F}\Delta t}$.

It can be seen from Eq. 3-4 that if $\mathbf{x}(t_0)$ is either a Gaussian random variable or a non-random variable, then $\mathbf{x}(t)$ is a Gaussian process, because the Gaussian property of the stochastic integral is preserved in linear operations. This property is particularly useful because the joint probability density function of $\mathbf{x}(t)$ at any given time can be described completely by the mean and covariance of $\mathbf{x}(t)$. The mean $\mathbf{m}(t)$ and covariance matrix $\mathbf{P}(t)$ of $\mathbf{x}(t)$ are defined as follows:

$$\mathbf{m}(t) = \mathrm{E}[\mathbf{x}(t)] \qquad \text{Eq. 3-5}$$

$$\mathbf{P}(t) = \mathrm{E}[(\mathbf{x}(t) - \mathbf{m}(t))(\mathbf{x}(t) - \mathbf{m}(t))^{\mathrm{T}}] \qquad \text{Eq. 3-6}$$

The superscript $^{\mathrm{T}}$ denotes the transpose operation. The joint probability density function of $\mathbf{x}(t)$ can be described as follows:

$$f_{\mathbf{x}(t)}(\zeta) = \frac{1}{(2\pi)^{\frac{n}{2}}|\mathbf{P}(t)|^{\frac{1}{2}}} \exp\{-\frac{1}{2}[(\zeta - \mathbf{m}(t))^{\mathrm{T}}\mathbf{P}^{-1}(t)(\zeta - \mathbf{m}(t))]\} \qquad \text{Eq. 3-7}$$

where $|\mathbf{P}(t)|$ is the determinant of $\mathbf{P}(t)$ and $\mathbf{P}^{-1}(t)$ and the inverse of $\mathbf{P}(t)$.

The statistical properties (mean and covariance) of $\mathbf{x}(t)$ can also be characterised as functions of time and system dynamics. It can be shown that, for a system defined by Eq. 3-1, the mean and the covariance matrix can be derived from the following differential equations:[9]

$$\dot{\mathbf{m}}(t) = \mathbf{F}(t)\mathbf{m}(t) \qquad \text{Eq. 3-8}$$

$$\dot{\mathbf{P}}(t) = \mathbf{F}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^{\mathrm{T}}(t) + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^{\mathrm{T}}(t) \qquad \text{Eq. 3-9}$$

Eq. 3-9 is known as the Riccati equation. Numerical integration is a direct method to solve this non-linear equation for a time varying system. However, if the system has constant coefficients or can be approximated as such, the solution can be reduced to solving a system of linear equations[10].

The justification for assuming that noise can be treated as Gaussian white noise will now be discussed. Whiteness implies that the noise is uncorrelated in time. Although white noise does not exist in practice, it provides a satisfactory substitution for wideband noise which affects any physical system within a limited bandwidth of that system. The assumption that noise is Gaussian is justified by using the central limit theorem. In practice, noise that affects any physical system is generated from a number of independent error sources. The theorem states that when a number of independent random variables are added together, the probability density function matches a Gaussian probability density function[10]. In addition, this assumption also simplifies the mathematics in the development of the Kalman filter algorithm.

In this section, the continuous time stochastic model of a linear dynamic system and its solution has been described using Eq. 3-1 and Eq. 3-4 based on the assumption that the noise is Gaussian white noise. The statistical properties of the system state vector can be characterised as functions of time using Eq. 3-8 and Eq. 3-9.

3.2.2 Discrete time stochastic linear system model

In practice, a digital computer is often used to implement the Kalman filter algorithms and data is processed at every sample time t_k. Thus, it is necessary to develop an equivalent discrete time stochastic model whose solution and statistical property is identical to the result of sampling the continuous time system of Eq. 3-1 at a discrete interval at time t_k. At sample time t_{k+1}, the solution to Eq. 3-1 can be written as follows:

$$\overbrace{\mathbf{x}(t_{k+1})}^{\mathbf{x}_{k+1}} = \overbrace{\Phi(t_{k+1},t_k)}^{\Phi_k}\overbrace{\mathbf{x}(t_k)}^{\mathbf{x}_k} + \overbrace{\int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)\mathbf{G}(\tau)d\beta(\tau)}^{\mathbf{w}_k} \qquad \text{Eq. 3-10}$$

Using an abbreviated notation replacing t_k with subscript k to distinguish between continuous and discrete time case, the equivalent discrete time model can be written as follows:

$$\mathbf{x}_{k+1} = \Phi_k \mathbf{x}_k + \mathbf{w}_k \qquad \text{Eq. 3-11}$$

where Φ_k is the transition matrix for the time step from t_k to t_{k+1} and $\mathbf{w}_k$ is the plant Gaussian white noise with the following statistical properties:

$$\mathrm{E}[\mathbf{w}_k] = \mathbf{0} \qquad \text{Eq. 3-12}$$

$$\mathrm{E}[\mathbf{w}_k \mathbf{w}_j^{\mathrm{T}}] = \begin{cases} \mathbf{Q}_k, & k = j \\ \mathbf{0}, & k \neq j \end{cases} \qquad \text{Eq. 3-13}$$

where $\mathbf{Q}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)\mathbf{G}(\tau)\mathbf{Q}(\tau)\mathbf{G}^{\mathrm{T}}(\tau)\Phi^{\mathrm{T}}(t_{k+1},\tau)d\tau$.

The discrete time parameters, Φ_k and $\mathbf{Q}_k$, completely describe the equivalent discrete time stochastic model corresponding to Eq. 3-1. Nonetheless, analytical solutions for Φ_k and $\mathbf{Q}_k$ are usually difficult to obtain. One method, which is commonly used, is numerical integration of the following differential equations to time $t = t_{k+1}$ (noting that $\Phi_k = \Phi(t_{k+1},t_k)$ and $\mathbf{Q}_k = \mathbf{Q}(t_{k+1},t_k)$):[9,10]

$$\dot{\Phi}(t,t_k) = \mathbf{F}(t)\Phi(t,t_k) \qquad \text{Eq. 3-14}$$

$$\dot{\mathbf{Q}}(t,t_k) = \mathbf{F}(t)\mathbf{Q}(t,t_k) + \mathbf{Q}(t,t_k)\mathbf{F}^{\mathrm{T}}(t) + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^{\mathrm{T}}(t) \qquad \text{Eq. 3-15}$$

subjected to the initial conditions $\Phi(t_k,t_k) = \mathbf{I}$ and $\mathbf{Q}(t_k,t_k) = 0$.

Alternatively, if the system is time invariant or slowly varying and the sample period ($\Delta t = t_{k+1} - t_k$) is short compared to the system natural transients, the following first-order approximation is often used:

$$\Phi_k \cong \mathbf{I} + \mathbf{F}_k \Delta t \qquad \text{Eq. 3-16}$$

$$\mathbf{Q}_k \cong \mathbf{G}_k \mathbf{Q}(t_k)\mathbf{G}_k^{\mathrm{T}} \Delta t \qquad \text{Eq. 3-17}$$

If the system parameters are constant (or it is reasonable to assume they are constant), another numerical method by van Loan can be used to calculate Φ_k and

$\mathbf{Q}_k$ by forming a matrix $\mathbf{A}$ as follows: [11]

$$\mathbf{A} = \begin{bmatrix} -\mathbf{F}_k & \mathbf{G}_k\mathbf{Q}_k\mathbf{G}_k^{\mathrm{T}} \\ \mathbf{0} & \mathbf{F}_k^{\mathrm{T}} \end{bmatrix} \Delta t \qquad \text{Eq. 3-18}$$

Let

$$\mathbf{B} = e^{\mathbf{A}} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \cdots & \Phi_k^{-1}\mathbf{Q}_k \\ \mathbf{0} & \Phi_k^{\mathrm{T}} \end{bmatrix} \qquad \text{Eq. 3-19}$$

Then, Φ_k is $\mathbf{B}_{22}^{\mathrm{T}}$ and $\mathbf{Q}_k$ is $\Phi_k \times \mathbf{B}_{12}$.

Applying an approach which is similar to the method used for the continuous time model, it can be shown that the mean and covariance of the system can be propagated as follows:

$$\mathbf{m}_{k+1} = \Phi_k \mathbf{m}_k \qquad \text{Eq. 3-20}$$

$$\mathbf{P}_{k+1} = \Phi_k \mathbf{P}_k \Phi_k^{\mathrm{T}} + \mathbf{Q}_k \qquad \text{Eq. 3-21}$$

In this section, the equivalent discrete time stochastic model has been derived and various methods for calculating the discrete time model parameters have been outlined. These techniques are commonly used in practical implementations of the Kalman filter algorithm, using a digital computer.

3.2.3 Discrete time linear measurement model

Some forms of measurements are always required to identify the state of a system. These measurements may be available either continuously or in discrete time in the form of sampled data. The latter is the most natural form for any system implemented using a digital computer. It is important to note that the sample interval does not have to be fixed. Assuming that the relationship between the system states and the measurements at time t_k can be described linearly by a measurement matrix $\mathbf{H}_k$, the resultant measurement model can be described as follows:

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \qquad \text{Eq. 3-22}$$

where $\mathbf{z}_k$ denotes a measurement in the present of Gaussian white noise, $\mathbf{v}_k$, with the following statistical properties:

$$E[\mathbf{v}_k] = \mathbf{0} \qquad \text{Eq. 3-23}$$

$$E[\mathbf{v}_k \mathbf{v}_j^{\mathrm{T}}] = \begin{cases} \mathbf{R}_k, & k = j \\ \mathbf{0}, & k \neq j \end{cases} \qquad \text{Eq. 3-24}$$

$$E[\mathbf{w}_k \mathbf{v}_j^{\mathrm{T}}] = 0, \quad \text{for all } k \text{ and } j \qquad \text{Eq. 3-25}$$

The next section will discuss a technique used to augment the overall system model to account for Gaussian non-white random processes.

3.2.4 Shaping filter and augmented system model

Despite several justifications made in assuming that the noise is Gaussian white noise, there are cases where non-white random processes (such as time correlated noise) are more suitable models. For such conditions, it is straightforward to augment the basic overall system model by using shaping filters to transform Gaussian white noise to other Gaussian random processes with the desired statistical properties[9,10,11]. The augmented model can be expressed as follows:

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \dot{\mathbf{x}}_p(t) \\ \dot{\mathbf{x}}_m(t) \end{bmatrix} = \begin{bmatrix} \mathbf{F}(t) & \mathbf{A}_p(t) & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_p(t) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_m(t) \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}_p(t) \\ \mathbf{x}_m(t) \end{bmatrix} + \begin{bmatrix} \mathbf{G}(t) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_p(t) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_m(t) \end{bmatrix} \mathbf{w}(t) \qquad \text{Eq. 3-26}$$

$$\mathbf{z}(t) = \begin{bmatrix} \mathbf{H}(t) & \mathbf{0} & \mathbf{H}_m(t) \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}_p(t) \\ \mathbf{x}_m(t) \end{bmatrix} + \mathbf{v}(t) \qquad \text{Eq. 3-27}$$

The augmented system makes use of the property of superposition of linear system theory. In Eq. 3-26 and Eq. 3-27, $\mathbf{x}_p(t)$ and $\mathbf{x}_m(t)$ are the augmented state vector for the plant and measurement model respectively. The relationship between the original plant and measurement model, and the augmented state vectors is described by the matrices $\mathbf{A}_p(t)$ and $\mathbf{H}_m(t)$. The ability to augment the system model to include a broad class of Gaussian random processes allows a Kalman filter to be applied to a wider range of problems.

In terms of navigation system analysis, the aircraft equations of motion, sensor characteristics and sensor measurements are used to derive the overall system model where the state vector consists of quantities such as aircraft attitude, position and

velocity. Gaussian non-white random processes such as random walk and first-order Gauss-Markov processes may also be used to augment the overall system model to provide more realistic sensor error models. It should be noted that a random walk process is obtained by integrating a Gaussian white noise. The statistical properties of the random walk and first-order Gaussian Markov process are summarised in Table 3-1. These two processes will be used extensively as sensor noise models for navigation system simulation and analysis in this thesis.

Table 3-1 Statistical properties of the first-order Gaussian Markov and the random walk process[9,10,11]

	Random walk process	**First-order Gaussian Markov process**
Continuous time model	$\dot{x}(t) = w(t)$	$\dot{x}(t) = -\beta x(t) + w(t)$
White noise PSD[i]	q	$2\beta\sigma^2$
Discrete time model	$x_{k+1} = x_k + w_k$	$x_{k+1} = e^{-\beta(t_{k+1}-t_k)}x_k + w_k$
Driving noise variance	$q_k = q(t_{k+1} - t_k)$	$\sigma^2(1-e^{-2\beta(t_{k+1}-t_k)})$
State variable mean	$\mathrm{E}[x] = 0$	$\mathrm{E}[x] = 0$
State variable variance	$\mathrm{E}[x^2] = qt$	$\mathrm{E}[x^2] = \sigma^2$
Autocorrelation function	-	$\vartheta(\tau) = \sigma^2 e^{-\beta\lvert\tau\rvert}$
Remarks	-	Time constant, $\tau = 1/\beta$

3.3 DISCRETE TIME (SAMPLED-DATA) OPTIMAL ESTIMATOR: THE KALMAN FILTER

The overall system model of a stochastic linear dynamic system was described in Section 3.2. This section reviews a discrete time optimal state estimator, the Kalman filter, used to estimate the state vector of a system based on such model. In this thesis, the term Kalman filter is used to refer to the *discrete* Kalman filter, which implies that the measurements are sampled and processed at every sample time t_k. However, the plant can be either a continuous time or an equivalent discrete time model. The discrete Kalman filter is the most common implementation of the set of Kalman filter algorithms. The discrete Kalman filter algorithm is summarized in Figure 3-2.

[i] Power Spectral Density

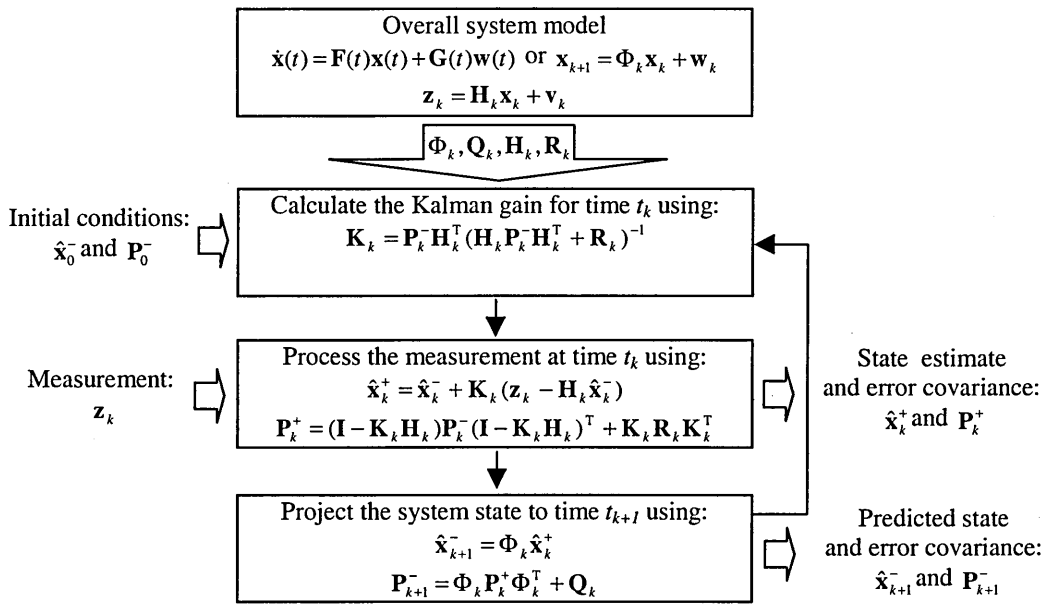


Figure 3-2 The summary of discrete Kalman filter algorithm

Assume that a system can be adequately modelled using Eq. 3-1 (or Eq. 3-11), and Eq. 3-22. The Kalman filter algorithm begins by estimating (predicting) the state vector $\hat{\mathbf{x}}_k^-$ and the associated prediction error covariance matrix $\mathbf{P}_k^-$ at an arbitrary initial time $t_{k=0}$ where $\hat{\mathbf{x}}_k^-=\mathrm{E}[\mathbf{x}_k]$ and $\mathbf{P}_k^-=\mathrm{E}[(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)^T]=\mathrm{E}[(\Delta\mathbf{x}_k^-)(\Delta\mathbf{x}_k^-)^T]$. The superscripts $\hat{}$ and $^-$ denote the state estimate and the prediction step (the estimation *before* incorporating any measurements) respectively. The relationship of the true state vector ($\mathbf{x}$), the estimated state vector ($\hat{\mathbf{x}}$) and the estimation error ($\Delta\mathbf{x}$) is defined as follows:

$$\mathbf{x}=\hat{\mathbf{x}}+\Delta\mathbf{x} \qquad \text{Eq. 3-28}$$

When a measurement $\mathbf{z}_k$ becomes available at time t_k, the Kalman filter uses it to estimate the state vector ($\hat{\mathbf{x}}_k^+$) according to:

$$\hat{\mathbf{x}}_k^+=\hat{\mathbf{x}}_k^-+\mathbf{K}_k(\mathbf{z}_k-\mathbf{H}_k\hat{\mathbf{x}}_k^-) \qquad \text{Eq. 3-29}$$

where the superscript $^+$ denotes the estimation step (the estimation *after* incorporating the measurements). $\mathbf{H}_k\hat{\mathbf{x}}_k^-$ can also be denoted as the predicted measurement $\hat{\mathbf{z}}_k^-$. The Kalman gain, $\mathbf{K}_k$, is defined as follows:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad \text{Eq. 3-30}$$

$\mathbf{K}_k$ in Eq. 3-30 is optimal because it satisfies the conditions that the mean-squared estimation error (the trace of the estimation error covariance matrix $\mathbf{P}_k^+$) is minimized. $\mathbf{P}_k^+$ is defined as $\mathbf{P}_k^+ = \mathrm{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_k^+)(\mathbf{x}_k - \hat{\mathbf{x}}_k^+)^T] = \mathrm{E}[(\Delta\mathbf{x}_k^+)(\Delta\mathbf{x}_k^+)^T]$ and can be calculated from:

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\mathbf{P}_k^- (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \quad \text{Eq. 3-31}$$

The recursive Kalman filter loop is then repeated, predicting the system state vector at the next sample time t_{k+1} according to:

$$\hat{\mathbf{x}}_{k+1}^- = \Phi_k \hat{\mathbf{x}}_k^+ \quad \text{Eq. 3-32}$$

with the corresponding prediction error covariance expressed as follows:

$$\mathbf{P}_{k+1}^- = \Phi_k \mathbf{P}_k^+ \Phi_k^T + \mathbf{Q}_k \quad \text{Eq. 3-33}$$

Recall that in terms of navigation system analysis, the state vector consists of aircraft states such as attitude, position and velocity and it assumes that the system model consists of Gaussian stochastic processes. Therefore, the mean $\mathbf{m}_k$ and the error covariance matrix $\mathbf{P}_k$ can be used to determine the error bounds of the aircraft states at any time t_k. The performance of a navigation system can be analysed by propagating $\mathbf{P}_k$ as a function of time – this is also known as *error covariance analysis*. It should be noted that the error covariance analysis can be carried out without the actual realisation of the measurements because the determination of $\mathbf{P}_k$ is independent of $\mathbf{z}_k$.

3.4 NON-LINEAR ESTIMATOR: THE EXTENDED KALMAN FILTER

The principles of stochastic system modelling and Kalman filter algorithm have been explained in sections 3.2 and 3.3 based on the assumption that the system model is linear. In practice, there are cases when the plant and/or the measurement models are non-linear, for instance, a pseudorange measurement model of a satellite navigation system. This section reviews a method known as the *extended* Kalman filter (EKF), which is used to modify the non-linear overall system model so that the Kalman filter algorithm can be implemented.[10,11]

Let a stochastic model of a non-linear system and non-linear measurement be described as follows (assuming a continuous plant and discrete measurements):

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t) + \mathbf{G}(t)\mathbf{w}(t) \qquad \text{Eq. 3-34}$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, t_k) + \mathbf{v}_k \qquad \text{Eq. 3-35}$$

where $\mathbf{f}$ and $\mathbf{h}$ are non-linear functions of the state vector. $\mathbf{w}(t)$ and $\mathbf{v}_k$ are the plant and measurement noise defined as before. The overall system model suitable for Kalman filtering can be obtained by linearizing Eq. 3-34 and Eq. 3-35 about a reference trajectory using Taylor series expansion[12]. The reference trajectory is defined as follows:

$$\mathbf{x}(t) = \mathbf{x}^*(t) + \Delta\mathbf{x}(t) \qquad \text{Eq. 3-36}$$

where $\mathbf{x}(t)$ is the true trajectory, $\mathbf{x}^*(t)$ is the reference trajectory and $\Delta\mathbf{x}(t)$ is a small perturbation. Substituting Eq. 3-36 in Eq. 3-34 and Eq. 3-35 yields:

$$\dot{\mathbf{x}}^*(t) + \Delta\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}^*(t) + \Delta\mathbf{x}(t), t) + \mathbf{G}(t)\mathbf{w}(t) \qquad \text{Eq. 3-37}$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k^* + \Delta\mathbf{x}_k, t_k) + \mathbf{v}_k \qquad \text{Eq. 3-38}$$

Expanding Eq. 3-37 and Eq. 3-38 using a Taylor series expansion and discarding terms higher than first-order, and noting that $\dot{\mathbf{x}}^*(t) = \mathbf{f}(\mathbf{x}^*(t), t)$, yields:

$$\begin{aligned}\Delta\dot{\mathbf{x}}(t) &= \left[\frac{\partial \mathbf{f}(\mathbf{x}, t)}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}^*(t)} \cdot \Delta\mathbf{x}(t) + \mathbf{G}(t)\mathbf{w}(t) \\ &= \mathbf{F}(\mathbf{x}^*(t), t)\Delta\mathbf{x}(t) + \mathbf{G}(t)\mathbf{w}(t) \qquad \text{Eq. 3-39}\end{aligned}$$

$$\begin{aligned}\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^*, t_k) &= \left[\frac{\partial \mathbf{h}(\mathbf{x}, t_k)}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}_k^*} \cdot \Delta\mathbf{x}_k + \mathbf{v}_k \\ &= \mathbf{H}_k(\mathbf{x}_k^*, t_k)\Delta\mathbf{x}_k + \mathbf{v}_k \qquad \text{Eq. 3-40}\end{aligned}$$

where the partial derivative of a matrix function is defined as follows:

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots \\ \vdots & \vdots & \end{bmatrix} \qquad \text{Eq. 3-41}$$

The choice of the reference trajectory used by the EKF is the filter's predicted trajectory where $\mathbf{x}^*(t) = \hat{\mathbf{x}}_k^-$. In a statistical sense, this is the best reference trajectory available. However, analysis of the EKF becomes more difficult because the Kalman gain now depends on the filter state estimates and hence, the actual measurement sequence. Therefore, the analysis of the EKF requires the use of Monte Carlo simulation. The EKF process is depicted in Figure 3-3.

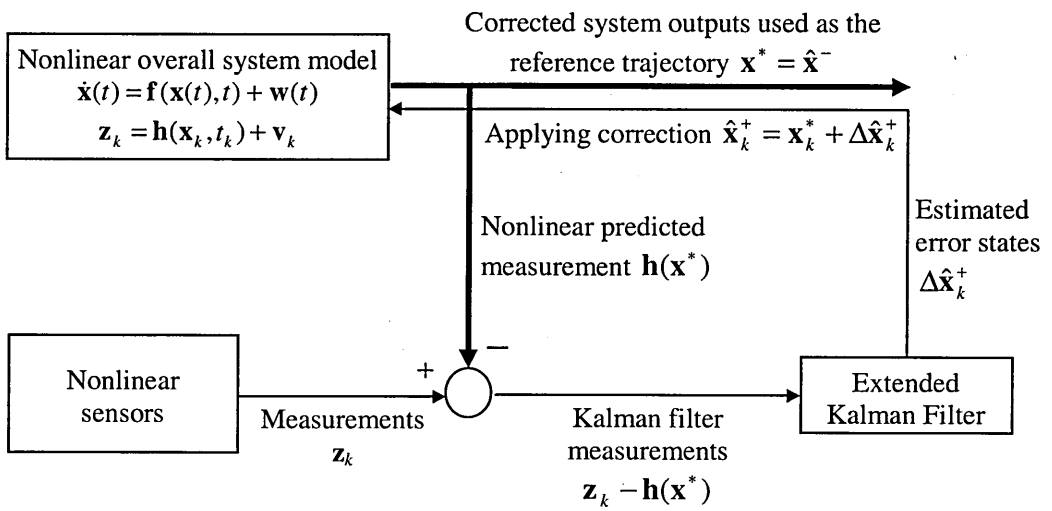


Figure 3-3 The extended Kalman filter process

The EKF differs from the Kalman filter as follows. The EKF uses Eq. 3-39 and Eq. 3-40 (with partial derivation matrices) as the overall system model with the *incremental* quantity $\Delta\mathbf{x}(t)$ as the system state vector. The measurement presented to the filter is the *difference* between the actual measurement and the predicted measurement based on the reference trajectory i.e. $\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^*)$. The EKF uses the incremental state estimate (the correction term) $\Delta\hat{\mathbf{x}}_k^+$ to update the reference trajectory and to obtain the corrected *total* state estimate (by the definition of Eq. 3-36) as follows:

$$\hat{\mathbf{x}}_k^+ = \mathbf{x}_k^* + \Delta\hat{\mathbf{x}}_k^+ \qquad \text{Eq. 3-42}$$

The components of $\Delta\hat{\mathbf{x}}_k^+$, which have been used for updating, are then reset to zero because they no longer contain any useful information. The non-trivial projection of the reset system state to the next sample time t_{k+1} is performed by using $\hat{\mathbf{x}}_k^+$ to solve the deterministic non-linear overall system model equations, $\mathbf{f}$ and $\mathbf{h}$, for $\hat{\mathbf{x}}_{k+1}^-$. The EKF

equations for the state estimate update, non-reset state prediction, filter gain and error covariance propagation remain the same as the standard Kalman filter. Table 3-2 summarises the EKF equations.

Table 3-2 The summary of the extended Kalman filter algorithms

Non-linear plant model Non-linear measurement model	$\dot{\mathbf{x}}(t)=\mathbf{f}(\mathbf{x}(t),t)+\mathbf{G}(t)\mathbf{w}(t)\,;\ \mathbf{w}(t)\sim \mathrm{N}[\mathbf{0},\mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^{\mathrm{T}}(t)]$ $\mathbf{z}_k=\mathbf{h}(\mathbf{x}_k,t_k)+\mathbf{v}_k\,;\ \mathbf{v}_k\sim \mathrm{N}[\mathbf{0},\mathbf{R}_k]$ and $\mathrm{E}[\mathbf{w}(t)\mathbf{v}_k^{\mathrm{T}}]=0$, for all k and t	
Reference trajectory	$\mathbf{x}^*(t_k)=\hat{\mathbf{x}}_k^-$	
EKF plant model EKF measurement model	$\Delta\dot{\mathbf{x}}(t)=\mathbf{F}(\mathbf{x}^*(t),t)\Delta\mathbf{x}(t)+\mathbf{G}(t)\mathbf{w}(t)\,;\ \mathbf{F}(\mathbf{x}^*(t),t)=\left[\dfrac{\partial\mathbf{f}(\mathbf{x},t)}{\partial\mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}^*(t)}$ $\mathbf{z}_k-\mathbf{h}(\mathbf{x}_k^*,t_k)=\mathbf{H}_k\Delta\mathbf{x}_k+\mathbf{v}_k\,;\ \mathbf{H}_k=\left[\dfrac{\partial\mathbf{h}(\mathbf{x},t_k)}{\partial\mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}_k^*}$	
Initial conditions	$\mathbf{x}(0)\sim \mathrm{N}(\hat{\mathbf{x}}_0^-,\mathbf{P}_0^-)$	
System state prediction	Reset states	Non-reset states
	$\dot{\mathbf{x}}(t)=\mathbf{f}(\mathbf{x}(t),t)$	$\Delta\hat{\mathbf{x}}_{k+1}^-=\Phi_k\Delta\hat{\mathbf{x}}_k^+\,;\ \Phi_k=e^{\mathbf{F}(\mathbf{x}_k^*,t_k)\Delta t}$
Prediction error covariance	$\mathbf{P}_{k+1}^-=\Phi_k\mathbf{P}_k^+\Phi_k^{\mathrm{T}}+\mathbf{Q}_k$	
State estimate update	$\Delta\hat{\mathbf{x}}_k^+=\Delta\hat{\mathbf{x}}_k^-+\mathbf{K}_k[\mathbf{z}_k-\mathbf{h}(\mathbf{x}_k^*,t_k)-\mathbf{H}_k\Delta\hat{\mathbf{x}}_k^-]$	
Estimation error covariance	$\mathbf{P}_k^+=(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}}+\mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^{\mathrm{T}}$	
Gain matrix	$\mathbf{K}_k=\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}}+\mathbf{R}_k)^{-1}$	

3.5 KALMAN FILTER PERFORMANCE ANALYSIS

The fundamental principles of Kalman filtering theory have been explained in the previous sections. This section covers the performance analysis and implementation considerations, which apply to practical implementation of Kalman filter algorithms.

3.5.1 Monte Carlo simulation

Monte Carlo simulation refers to a simulation method, which is used in applications where it is difficult or impossible to study the behaviour of a system analytically. This class of problems includes the extended Kalman filter. The behaviour

of a system which models the set of Gaussian random processes described by Eq. 3-11 can be analysed statistically, based on sequences of randomly generated noise with specific stochastic properties. For example, the plant noise $\mathbf{w}_k \sim N[\mathbf{0},\mathbf{Q}_k]$ can be derived from the transformation of standard normal (Gaussian) white noise sequences $\mathbf{u}_k \sim N[\mathbf{0},\mathbf{I}]$, which are available as standard normal random number generators in most dynamic simulation packages, as follows:[11]

$$\mathbf{w}_k = \mathbf{C}_k \mathbf{u}_k \qquad \text{Eq. 3-43}$$

Because $E[\mathbf{w}_k \mathbf{w}_k^T] = \mathbf{Q}_k = E[(\mathbf{C}_k \mathbf{u}_k)(\mathbf{C}_k \mathbf{u}_k)^T] = \mathbf{C}_k E[\mathbf{u}_k \mathbf{u}_k^T] \mathbf{C}_k^T$ and $E[\mathbf{u}_k \mathbf{u}_k^T] = \mathbf{I}$, $\mathbf{Q}_k$ reduces to:

$$\mathbf{Q}_k = \mathbf{C}_k \mathbf{C}_k^T \qquad \text{Eq. 3-44}$$

where $\mathbf{C}_k$ can be found by Cholesky factorisation.[11,38] Generally, the initial values of the state variable $\mathbf{x}$ are determined randomly from the population of random numbers $\mathbf{x}_0 \sim N[\mathbf{0},\mathbf{P}_0]$, where $\mathbf{P}_0 = E[\mathbf{x}_0 \mathbf{x}_0^T]$.

3.5.2 Kalman filter divergence

Kalman filter divergence refers to the discrepancy between the estimation errors predicted by the filter, the *theoretical* root-mean-square (rms) errors, and the *true* rms errors. The theoretical rms errors are determined by taking the square root of the diagonal terms of the estimation error covariance matrix **P,** based on the assumption that the errors are zero mean and uncorrelated in time. The true rms and mean errors used in this thesis have been determined from Monte Carlo simulations. Kalman filter divergence may be caused by a stochastically unstable system model, numerical (round off) errors or modelling errors.

A system model is stochastically stable if it is stochastically controllable and stochastically observable. A Kalman filter which is based on stochastically stable models is uniformly asymptotically globally stable.[9,49] This implies that the estimation error bounds (the covariance matrices) converge to zero (asymptotic), independent of the initial conditions (global) and absolute time (uniform). The actual system may be unstable, where the system states grow without bounds, while the filter can be stable, where the estimation errors remain bounded. This is a highly desirable filter

characteristic. Although, in practice, the stochastically stable system model requirement is quite restrictive and is not usually fulfilled, the Kalman filter affords some degree of natural stability.[10] In fact, divergence is usually caused by numerical errors and modelling errors (mismodelling) rather than system instability.[9,10]

In theory, the recursive loops of the Kalman filter can be repeated indefinitely. However, numerical errors may occur as the number of steps becomes large, which is common in any numerical procedure. A number of techniques can be applied to prevent divergence due to round off errors.[10,11,32] In addition to, high-precision arithmetic, the algorithm can be modified, for example, U-D factorisation is the equivalent of the regular Kalman filter algorithm; it propagates factors of the **P** matrix rather than **P** itself. A more numerically stable equation can be used to update the **P** matrix such as Eq. 3-31. In addition, the **P** matrix can also be symmetrised during each recursive step. It is also a good practice to avoid including deterministic processes (processes with zero-PSD input noise) in the system model because this causes the **P** matrix to approach a semidefinite condition as the number of steps becomes large. A small error may lead to non-positive definite **P** matrix and subsequent divergence. Therefore, it is common practice to replace the terms on the diagonal of the **Q** matrix which are zeros with small non-zero values. This is equivalent to adding small fictitious plant noise to the system and although it may lead to sub-optimality, it is preferable to have sub-optimal solutions than filter divergence.

In practice, a physical system can not be represented precisely by a mathematical model owing to either limited knowledge of the system or computing constraints which preclude incorporating all of the state variables in the model.[10,32] For example, an Inertial Navigation System (INS) can be affected by over 100 error sources whereas in practice even a modern processor is restricted to approximately 20 state variables.[36] In many cases, an overall system model used in a Kalman filter is a trade off between the performance of the processor and practical limitations. A divergence may occur if the model differs significantly from the actual system, for instance if the dominant system states are omitted from the model. This is equivalent to requesting the Kalman filter to fit data to a wrong curve. To avoid divergence due to modelling errors usually requires extensive system simulation and analysis, known as *sensitivity analysis*.

In practice, system parameters may be adjusted until the filter performance meets the requirements, known as filter *tuning*. One practical method that has been used widely to prevent divergence due to system modelling error is also to add fictitious plant noise to the system model.[10]

Figure 3-4 illustrates the Kalman filter performance in estimating the states of the 1st order Gaussian Markov process described in Table 3-1. Figure 3-4 (a) shows the true states, noisy measurements and the state estimates. It can be seen that the Kalman filter processes noisy measurements effectively, producing accurate state estimates that follow the true states closely. Figure 3-4 (b) illustrates the theoretical rms errors and the true rms and mean errors derived by means of Monte Carlo simulation. It shows that the theoretical and the true values compare closely. Moreover, the error samples are also bounded inside the 1σ-boundaries (68.3% for a Gaussian random variable). Figure 3-4 (c) illustrates the filter divergence due to modelling error. The actual system has $\beta = 0.05\,\mathrm{sec}^{-1}$ but the system model assumes $\beta = 0.005\,\mathrm{sec}^{-1}$.

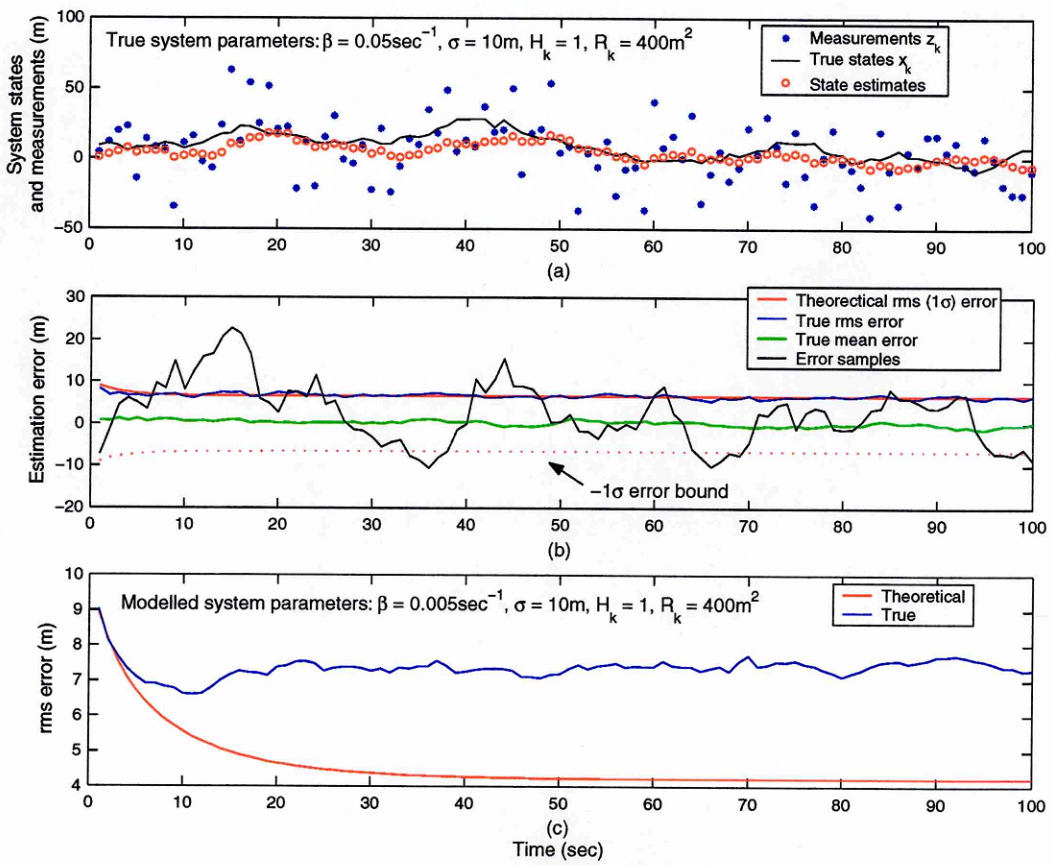


Figure 3-4 Kalman filtering examples

3.5.3 Decentralised Kalman filtering

In a design process, it is usually assumed that a single Kalman filter handles all the state variables and measurements simultaneously. This configuration is known as *centralised* Kalman filtering. In practice, in systems such as integrated navigation systems, data from various sensors may be pre-processed, for instance, by local Kalman filters, before being passed to the main Kalman filter. These pre-filtered measurements are correlated in time, and hence the basic assumptions used for the standard Kalman filter algorithms are invalid. This situation gives rise to the concept of "*decentralised*" Kalman filtering. The solutions used in decentralised filter can be based on either ad hoc methods or decentralised Kalman filtering theory.[11,13,14,15] In this thesis, the algorithms used for decentralised Kalman filtering, especially for relative navigation described in Chapter 6, are based on ad hoc methods. This approach has been widely accepted in many applications and it will be shown that the concept can be successfully applied to relative navigation for civil applications, which is the central theme of this thesis.

CHAPTER 4 NAVIGATION SYSTEMS AND SYSTEMS INTEGRATION

4.1 INTRODUCTION

In this chapter, the Kalman filtering theory covered in Chapter 3, will be applied to aircraft navigation system analysis and integration. The chapter begins by revising the underlying mathematics and assumptions which provide the basis for the development of aircraft guidance and navigation system models. The primary purpose in making these assumptions is to simplify the analysis without the loss in generality, allowing the same approaches to be applied to more complex systems.

The navigation systems under consideration include INS and GPS because they form the basis of navigation systems for FANS. A simplified model of a strapdown INS and GPS will be developed specifically for the computer simulation described in this thesis. For these simplified models, error characteristics of actual systems, for example, *Schuler oscillation* of a strapdown INS, are not modelled because the contributions made by these errors are not significant to the analysis of the navigation performance developed in this thesis. Moreover, it is straightforward to incorporate these errors in the overall system model.

The performance and limitations of stand-alone and integrated systems will be examined and analysed by means of Monte Carlo simulation which is required specifically for the analysis of the extended Kalman filter (EKF). The results obtained in this chapter subsequently provide the basis for the analysis of relative navigation to be developed in Chapter 5 and Chapter 6. Unlike absolute navigation systems such as INS and GPS, in relative navigation aircraft positions are determined relative to each other irrespective of their absolute positions (which are given with respect to the earth).

4.2 UNDERLYING MODELLING AND ASSUMPTIONS

The equations used in aircraft inertial navigation are formulated in a specific set of coordinate systems (frames of reference). The kinematics analysis is defined with respect to the chosen coordinate systems which can be derived from coordinate transformations, taking account of the *Coriolis* effect. The Coriolis terms arise in the

equations of motion when analysing the motion of a particle in several intermediate rotating frames of reference. This section reviews the coordinate transformations, the *differential operator* used to derive the Coriolis terms and aircraft and airspace modelling.

4.2.1 Coordinate systems and transformations

All rectilinear coordinate systems defined in this thesis are assumed to be orthogonal, right-handed, Cartesian frames of reference. A vector will be denoted using a boldface small letter with a superscript denoting the coordinate system in which the vector is defined (in component form). It will be assumed that all vectors are time varying but the implicit term *(t)* will be omitted in the notation. In Figure 4-1, $\mathbf{r}^{A_1}$ is the position vector of a particle P defined in the A_1-frame. In component form, a vector will be expressed using small letters in italic with subscripts denoting the respective axes of the coordinate system i.e. $\mathbf{r}^{A_1} = [r_x, r_y, r_z]^T$. For the analysis in this section, it is assumed that rotation angles are small, and hence small angle approximations are used.

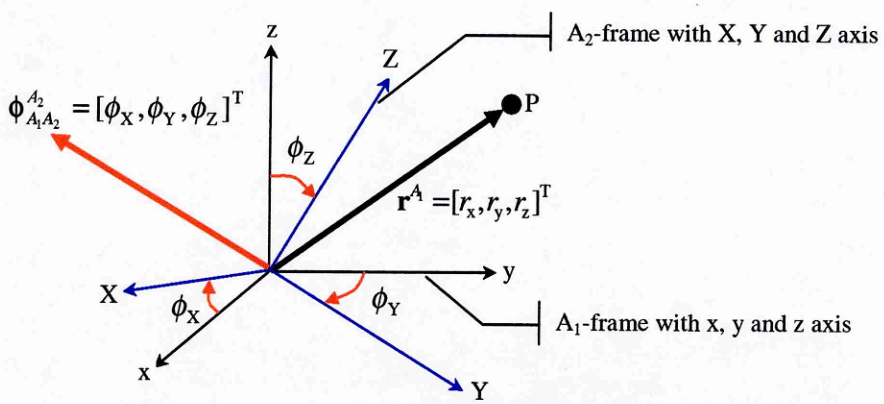


Figure 4-1 Coordinate systems, notations and transformations

In Figure 4-1, $\phi^{A_2}_{A_1A_2}$ is the *rotation vector* which describes the rotation of the A_2-frame relative to the A_1-frame where $\phi = \left\|\phi^{A_2}_{A_1A_2}\right\| = (\phi_X^2 + \phi_Y^2 + \phi_Z^2)^{\frac{1}{2}}$ is the magnitude of $\phi^{A_2}_{A_1A_2}$. The A_2-frame represents the rotation of the A_1-frame about $\phi^{A_2}_{A_1A_2}$ by an angle ϕ. It should be noted that $\phi^{A_2}_{A_1A_2} = \phi^{A_1}_{A_1A_2}$ is the unique property of the rotation vector but $\phi^{A_2}_{A_1A_2} = -\phi^{A_2}_{A_2A_1}$.[16] The Direction Cosine Matrix (DCM) [16,17,18] which transforms a vector defined in the A_2-frame to the A_1-frame is determined as follows:

$$\mathbf{C}_{A_2}^{A_1} = \mathbf{I} + \frac{\sin\phi}{\phi}[\boldsymbol{\phi}_{A_1A_2}^{A_2}\times] + \frac{(1-\cos\phi)}{\phi^2}[\boldsymbol{\phi}_{A_1A_2}^{A_2}\times]^2 \qquad \text{Eq. 4-1}$$

where $[\boldsymbol{\phi}_{A_1A_2}^{A_2}\times]$ is the skew-symmetric matrix of $\boldsymbol{\phi}_{A_1A_2}^{A_2}$ defined as follows:

$$[\boldsymbol{\phi}_{A_1A_2}^{A_2}\times] = \begin{bmatrix} 0 & -\phi_Z & \phi_Y \\ \phi_Z & 0 & -\phi_X \\ -\phi_Y & \phi_X & 0 \end{bmatrix} \qquad \text{Eq. 4-2}$$

The skew-symmetric matrix notation is used extensively in the derivation of strapdown INS navigation algorithms. It is also the vector cross product written in matrix form. For small rotation angles, Eq. 4-1 can be reduced as follows:

$$\mathbf{C}_{A_2}^{A_1} = \mathbf{I} + [\boldsymbol{\phi}_{A_1A_2}^{A_2}\times] \qquad \text{Eq. 4-3}$$

The DCM satisfies the following equations:

$$\mathbf{r}^{A_1} = \mathbf{C}_{A_2}^{A_1}\mathbf{r}^{A_2} \qquad \text{Eq. 4-4}$$

$$\mathbf{C}_{A_1}^{A_2} = \mathbf{C}_{A_2}^{A_1\,\mathrm{T}} = \mathbf{C}_{A_2}^{A_1\,-1} \cong (\mathbf{I} - [\boldsymbol{\phi}_{A_1A_2}^{A_2}\times]) \qquad \text{Eq. 4-5}$$

$$\mathbf{C}_{A_1}^{A_3} = \mathbf{C}_{A_2}^{A_3}\mathbf{C}_{A_1}^{A_2} \qquad \text{Eq. 4-6}$$

The DCM can also be propagated as a function of time based on the rotational rate of the coordinate system. Let $\boldsymbol{\omega}_{A_1A_2}^{A_2} = \dot{\boldsymbol{\phi}}_{A_1A_2}^{A_2}$ denote the rotational rate of the A_2-frame with respect to the A_1-frame. The time derivative of the DCM is given by:[17]

$$\dot{\mathbf{C}}_{A_2}^{A_1} = \mathbf{C}_{A_2}^{A_1}[\boldsymbol{\omega}_{A_1A_2}^{A_2}\times] \qquad \text{Eq. 4-7}$$

The solution to Eq. 4-7 can be obtained as follows:[17]

$$\mathbf{C}_{A_2\,k+1}^{A_1} = \mathbf{C}_{A_2\,k}^{A_1}\exp\left(\int_{t_k}^{t_{k+1}}[\boldsymbol{\omega}_{A_1A_2}^{A_2}\times]dt\right) \qquad \text{Eq. 4-8}$$

By assuming that the rotational rate $\boldsymbol{\omega}_{A_1A_2}^{A_2}$ remains fixed in space over the update interval, Eq. 4-8 can be written as follows:

$$\mathbf{C}_{A_2\,k+1}^{A_1} = \mathbf{C}_{A_2\,k}^{A_1}\exp([\boldsymbol{\sigma}_{A_1A_2}^{A_2}\times]) = \mathbf{C}_{A_2\,k}^{A_1}\mathbf{A}_k \qquad \text{Eq. 4-9}$$

where the rotation vector $\sigma_{A_1A_2}^{A_2} = \int_{t_k}^{t_{k+1}} \omega_{A_1A_2}^{A_2} dt$ can be used to find $\mathbf{A}_k$ by applying Eq. 4-1.

4.2.2 Development of the airspace model

This section describes three coordinate systems[17,19] that are generally used in terrestrial navigation – the *inertial* frame, the *earth* frame and the *local geographic* frame. The inertial frame is defined as a non-rotating and non-translating coordinate system with its origin at the centre of the earth. The accelerometer and gyroscope measurements are obtained with respect to this reference frame. The earth frame is defined as the coordinate system with its origin at the centre of the earth and its axes fixed with respect to the earth axes. The local geographic frame is defined as the coordinate system with its origin at the location of the navigation system and its axes aligned with the directions of north, east and the local vertical (up). Figure 4-2 illustrates these three coordinate systems.

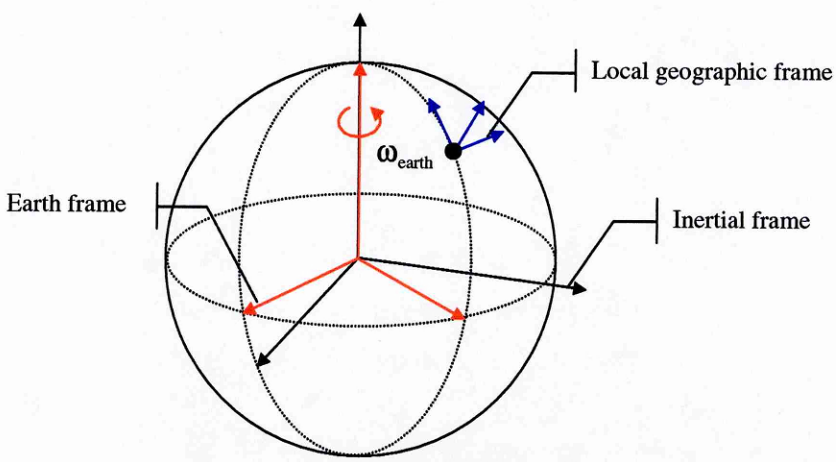


Figure 4-2 Terrestrial navigation coordinate systems

In practice, all these coordinate systems and the gravity vector will be included in the derivation of the inertial navigation algorithms and the earth will be modelled using the coordinate system defined in the World Geodetic System in 1984, known as WGS-84. The equations of motion are typically *mechanised* in the local geographic frame, which is moving with respect to the earth frame, which, in turn, moves with respect to the inertial frame. Nevertheless, to model all these coordinate systems and the gravity can add unnecessary complexity to the analysis because the kinematics analysis can be performed with respect to any number of moving (translating and/or rotating)

reference frames by applying a differential operator to take account of the Coriolis effect. Therefore, a simplified model of the airspace (denoted as the *i*-frame) has been developed in this thesis. The *i*-frame is a "*flat earth*" model derived from an angular cylindrical projection (Mercator projection)[20]. It also assumes that the earth is a perfect sphere with a radius of 6380000 metres[11]. Although the earth can be modelled more accurately using an oblate spheroid with an elliptical cross section[19], the spherical earth model provides an adequate reference earth model for the purposes of this thesis. The transformation between the *i*-frame and the spherical earth model is given by:

$$x = k\Lambda \qquad \text{Eq. 4-10}$$

$$y = k \ln\left|\tan\left(\frac{\Phi}{2} + \frac{\pi}{4}\right)\right| \qquad \text{Eq. 4-11}$$

$$z = Z \qquad \text{Eq. 4-12}$$

where Z is the aircraft altitude, k = Z + the earth radius, Λ is the geodetic longitude and Φ is the geodetic latitude. Figure 4-3 illustrates the simplified airspace model.

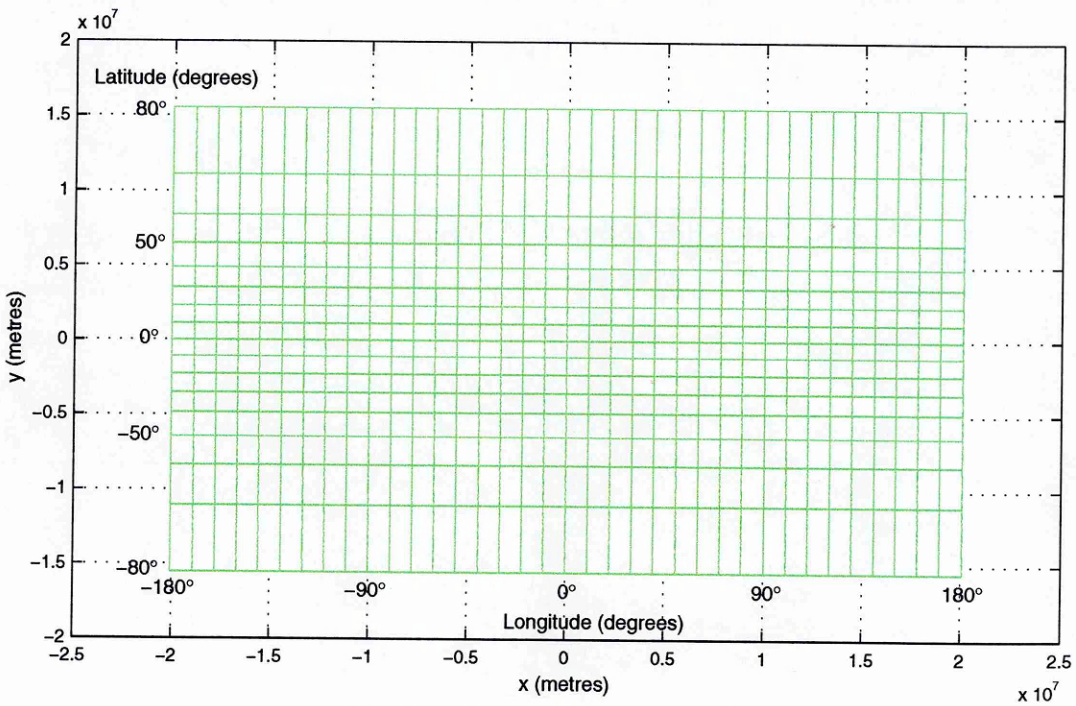


Figure 4-3 The simplified airspace model (the *i*-frame)

It will be shown in Appendix B that the characteristics of the simulated GPS (ground trace, observability and DOP) are preserved in this simplified airspace model.

4.2.3 Kinematics of a particle utilising intermediate frame of reference

This section reviews the kinematics analysis of a particle utilising intermediate reference frames, which form the basis of the derivation of strapdown INS navigation algorithms.

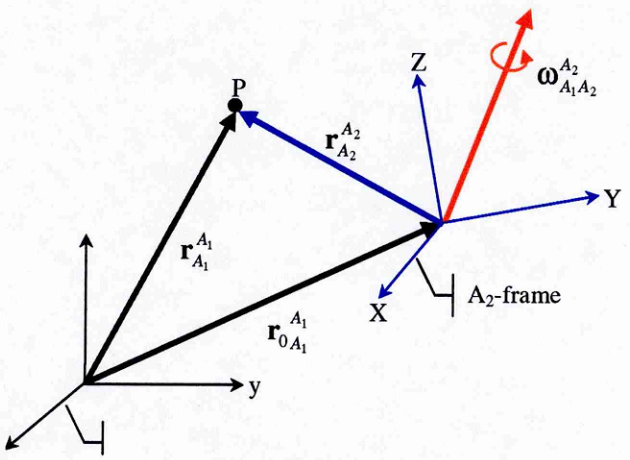


Figure 4-4 Kinematics analysis of a particle utilising intermediate frames of reference

In Figure 4-4, $\mathbf{r}_{A_1}^{A_1}$ and $\mathbf{r}_{A_2}^{A_2}$ are the position vectors of a particle P defined in the A_1-frame and the A_2-frame respectively. The superscript denotes the reference frame in which the vector is expressed (component form) and the subscript denotes the reference frame in which the vector is defined. For example, the vector which is *defined in the A_2-frame and expressed in the A_1-frame* can be obtained from the following transformation, $\mathbf{r}_{A_2}^{A_1} = \mathbf{C}_{A_2}^{A_1}\mathbf{r}_{A_2}^{A_2}$. The A_2-frame is the intermediate frame of reference, which is moving relative to the A_1-frame. The relationship between the position vectors of P defined in these two reference frames can be written as follows:

$$\mathbf{r}_{A_1}^{A_1} = \mathbf{r}_{0A_1}^{A_1} + \mathbf{r}_{A_2}^{A_1} = \mathbf{r}_{0A_1}^{A_1} + \mathbf{C}_{A_2}^{A_1}\mathbf{r}_{A_2}^{A_2} \qquad \text{Eq. 4-13}$$

Defining the motion of P in the A_2-frame, and the translation and the rotation of the A_2-frame relative to the A_1-frame as $\dot{\mathbf{r}}_{A_2}^{A_2}$, $\dot{\mathbf{r}}_{0A_1}^{A_1}$ and $\boldsymbol{\omega}_{A_1A_2}^{A_2}$ respectively, it can be shown that the "*differential operator*" which can be used to analyse the motion of P with respect to the A_1-frame is given by:[21,22]

$$\left.\frac{d}{dt}\right|_{A_1} = \left.\frac{\partial}{\partial t}\right|_{A_2} + [\boldsymbol{\omega}^{A_2}_{A_1A_2}\times] \qquad \text{Eq. 4-14}$$

By differentiating Eq. 4-13 and applying Eq. 4-14, the motion of P with respect to the A_1-frame can be determined based on its motion defined in the A_2-frame and the relative motion of the A_2-frame as follows:

$$\dot{\mathbf{r}}^{A_1}_{A_1} = \dot{\mathbf{r}}_{0\,A_1}^{A_1} + \mathbf{C}^{A_1}_{A_2}\left[\dot{\mathbf{r}}^{A_2}_{A_2} + [\boldsymbol{\omega}^{A_2}_{A_1A_2}\times]\mathbf{r}^{A_2}_{A_2}\right] \qquad \text{Eq. 4-15}$$

The last term in Eq. 4-15 is known as the Coriolis term. It arises as a result of the rotation of the intermediate frame of reference. In theory, the differential operator (Eq. 4-14) can be used to analyse the motion of a particle with respect to any number of intermediate reference frames. In terms of strapdown inertial navigation, it allows the equations of motion to be defined with respect to any reference frame, such as the local geographic frame, based on measurements obtained in the inertial frame.

4.2.4 Development of the aircraft model

The aircraft model used for simulation in this thesis is based on a database known as the Performance Table File (PTF) of the Base of Aircraft Data (BADA), revision 3.1.[23] The PTF is a collection of data files which contain aircraft performance data such as true air speed (TAS), rate of climb or descent and fuel consumption for different aircraft types, phases of flight (cruise, climb, descent) and flight levels. The data which has been included in the PTF files is based on the aircraft performance and operating procedure coefficients using a Total-Energy Model (TEM) to determine thrust, drag, aircraft speed and fuel flow.

In this thesis, the simulated aircraft is assumed to fly level flight at a constant TAS which has been extracted from the cruise TAS data of the PTF. The aircraft fly between waypoints using automatic first-order closed-loop heading control guidance. The commanded heading is initiated when the distance to the next waypoint is less than d_ψ, where d_ψ is calculated from a waypoint steering algorithm[24] (Figure 4-5) given by:

$$d_\psi = R\times\tan(\frac{\Delta\psi}{2}) \qquad \text{Eq. 4-16}$$

where $R = \text{TAS}/\dot{\psi}_{max}$ and $\Delta\psi$ is the change in heading. $\dot{\psi}_{max}$ is the maximum permissible rate of turn, which is calculated from the maximum allowable bank angle ϕ_{max} (45 degrees for civil flight during cruise) as follows:[25]

$$\dot{\psi}_{max} = \frac{g}{\text{TAS}} \tan(\phi_{max}) \qquad \text{Eq. 4-17}$$

where g = 9.81 m/s^2. Matlab source codes and details of the implementation of this heading control algorithm are given in Appendix H section H.4.

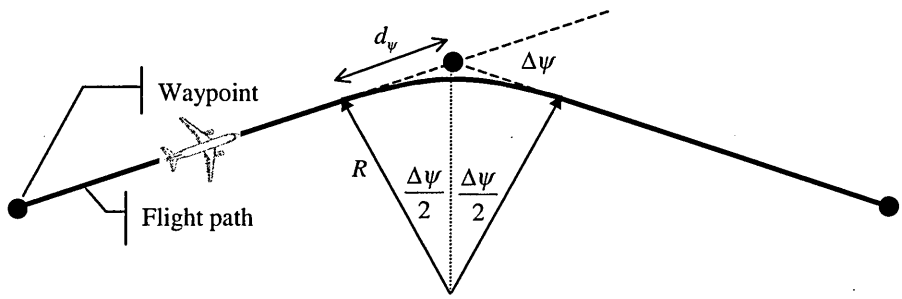


Figure 4-5 Geometry of waypoint bypassing[24]

4.3 STRAPDOWN INS NAVIGATION ALGORITHMS

In this section, the equations of motion for the simplified model of a strapdown INS are developed based on the underlying assumptions covered in Section 4.2. These equations will be used to derive the system error model used in Kalman filtering.

4.3.1 Development of strapdown INS equations of motion

In a real strapdown INS, a set of accelerometers and gyroscopes are used to measure the linear acceleration and the angular velocity of the vehicle's body frame of reference (denoted as the *b*-frame) with respect to the inertial frame. These measurements are integrated to derive the velocity, position, and attitude of the aircraft.

In this thesis, it will be assumed that the first stage of the integration process, including the measurement and integration of aircraft's accelerations, is perfect i.e. without errors. (In principle, this is equivalent to employing a set of *linear velocity sensors* to measure the linear velocity of the *b*-frame. Hence only the first-order terms, i.e. the *velocity*, need to be modelled.) In addition, it will also be assumed that the motion of the *b*-frame is measured with respect to the simplified airspace model (the *i*-

frame). These assumptions allow the order of the error model of strapdown INS to be reduced, simplifying the simulation and analysis significantly without affecting the validity of the results obtained in this thesis.

The reference frame in which the system is mechanised to navigate is called the navigation frame of reference (denoted as the *n*-frame). The attitude computer uses the measured angular velocity, $\boldsymbol{\omega}_{ib}^{b}$, to calculate the DCM for the transformations between the *b*-frame and the *n*-frame by solving Eq. 4-7. The DCM is used to transform the measured linear velocity, $\mathbf{v}_{i}^{b}$, defined in the *b*-frame and used in the *n*-frame. For simplicity, it will be assumed that the system is mechanised in the simplified airspace model i.e. the *i*-frame is also the *n*-frame. The differential equations that describe the motion of the navigation system can be written as follows:

$$\dot{\mathbf{r}}_{i}^{i} = \mathbf{v}_{i}^{i} = \mathbf{C}_{b}^{i}\mathbf{v}_{i}^{b} \qquad \text{Eq. 4-18}$$

$$\dot{\mathbf{C}}_{b}^{i} = \mathbf{C}_{b}^{i}[\boldsymbol{\omega}_{ib}^{b}\times] \qquad \text{Eq. 4-19}$$

Therefore, the position and the orientation of the navigation system with respect to the *i*-frame can be determined by solving the equations of motion (Eq. 4-18 and Eq. 4-19) based on given initial conditions.

4.3.2 System error analysis

In practice, the INS measurements will be corrupted by the measurement noise from various error sources such as sensor noise and initial misalignment. A first-order Gaussian Markov process and Gaussian white noise generally provide adequate sensor noise models of the gyroscope and the accelerometer.[11,26] Therefore, a *first-order Gaussian Markov process* will be used as the stochastic model of the *gyroscope error*, $\Delta\boldsymbol{\omega}_{ib}^{b}$. However, because the integration of the accelerometer Gaussian white noise produces random walk noise, a *random walk process* will be used to provide the stochastic model of the *linear velocity sensor error*, $\Delta\mathbf{v}_{ib}^{b}$. The statistical properties of these stochastic processes are given in Table 3-1. The sign convention for the true state, the measurement and the noise is given in Eq. 3-28. The estimated navigation equations can be written as follows:

$$\dot{\hat{\mathbf{r}}}_i^i = \hat{\mathbf{C}}_b^i \hat{\mathbf{v}}_i^b \qquad \text{Eq. 4-20}$$

$$\dot{\hat{\mathbf{C}}}_b^i = \hat{\mathbf{C}}_b^i [\hat{\boldsymbol{\omega}}_{ib}^b \times] \qquad \text{Eq. 4-21}$$

The attitude and position error equations will now be derived based on the approach described by Titterton[17].

Attitude error equations

For a perfect strapdown INS, the navigation computer will navigate with respect to the true navigation frame i.e. the *i*-frame. However, due to system errors, the navigation computer will actually navigate with respect to an estimated navigation frame – the computer frame of reference (denoted as the *c*-frame). The transformation between the body frame and the estimated navigation frame can be written as follows:

$$\hat{\mathbf{C}}_b^i = \mathbf{C}_b^c = \mathbf{C}_i^c \mathbf{C}_b^i \qquad \text{Eq. 4-22}$$

Let $\boldsymbol{\psi}$ be the misalignment angle vector of the true and the estimated navigation frame as shown in Figure 4-6. For small misalignment, we can write:

$$\mathbf{C}_i^c \cong \mathbf{I} - [\boldsymbol{\psi} \times] \qquad \text{Eq. 4-23}$$

Substituting Eq. 4-23 in Eq. 4-22 yields:

$$\hat{\mathbf{C}}_b^i = (\mathbf{I} - [\boldsymbol{\psi} \times]) \mathbf{C}_b^i \qquad \text{Eq. 4-24}$$

Rearranging Eq. 4-24 and differentiating it with respect to time, yields:

$$[\dot{\boldsymbol{\psi}} \times] = -\dot{\hat{\mathbf{C}}}_b^i \mathbf{C}_b^{i\,\mathrm{T}} - \hat{\mathbf{C}}_b^i \dot{\mathbf{C}}_b^{i\,\mathrm{T}} \qquad \text{Eq. 4-25}$$

Substituting Eq. 4-19 and Eq. 4-21 in Eq. 4-25, and using the property for a skew-symmetric matrix that, $[\boldsymbol{\omega}_{ib}^b \times]^{\mathrm{T}} = -[\boldsymbol{\omega}_{ib}^b \times]$, yields:

$$\begin{aligned}
[\dot{\boldsymbol{\psi}} \times] &= -\hat{\mathbf{C}}_b^i [\hat{\boldsymbol{\omega}}_{ib}^b \times] \mathbf{C}_b^{i\,\mathrm{T}} - \hat{\mathbf{C}}_b^i (\mathbf{C}_b^i [\boldsymbol{\omega}_{ib}^b \times])^{\mathrm{T}} \\
&= -\hat{\mathbf{C}}_b^i [\hat{\boldsymbol{\omega}}_{ib}^b \times] \mathbf{C}_b^{i\,\mathrm{T}} - \hat{\mathbf{C}}_b^i [\boldsymbol{\omega}_{ib}^b \times]^{\mathrm{T}} \mathbf{C}_b^{i\,\mathrm{T}} \\
&= \hat{\mathbf{C}}_b^i ([\boldsymbol{\omega}_{ib}^b \times] - [\hat{\boldsymbol{\omega}}_{ib}^b \times]) \mathbf{C}_b^{i\,\mathrm{T}} \qquad \text{Eq. 4-26}
\end{aligned}$$

Substituting $[\Delta\boldsymbol{\omega}_{ib}^b \times] = [\boldsymbol{\omega}_{ib}^b \times] - [\hat{\boldsymbol{\omega}}_{ib}^b \times]$ and Eq. 4-24 in Eq. 4-26, and ignoring the products of error terms yields:

$$[\dot{\psi}\times] = (\mathbf{I} - [\psi\times])\mathbf{C}_b^i[\Delta\omega_{ib}^b\times]\mathbf{C}_b^{i\,\mathrm{T}}$$

$$= \mathbf{C}_b^i[\Delta\omega_{ib}^b\times]\mathbf{C}_b^{i\,\mathrm{T}} \qquad \text{Eq. 4-27}$$

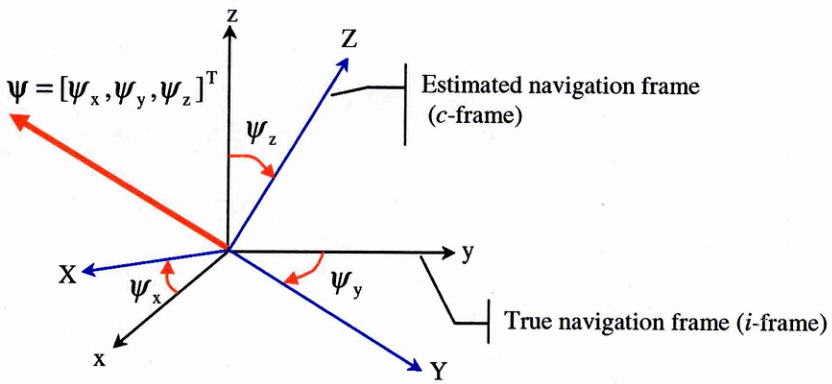


Figure 4-6 Strapdown INS frames of reference

From an element by element comparison and using the orthogonal property of the DCM, Eq. 4-27 can be written in vector form as follows:

$$\dot{\psi} = \mathbf{C}_b^i\Delta\omega_{ib}^b \qquad \text{Eq. 4-28}$$

Position error equations

The position error equation can be derived by subtracting Eq. 4-20 from Eq. 4-18 as follows:

$$\Delta\dot{\mathbf{r}}_i^i = \dot{\mathbf{r}}_i^i - \dot{\hat{\mathbf{r}}}_i^i = \mathbf{C}_b^i\mathbf{v}_i^b - \hat{\mathbf{C}}_b^i\hat{\mathbf{v}}_i^b \qquad \text{Eq. 4-29}$$

Substituting Eq. 4-24 and $\hat{\mathbf{v}}_i^b = \mathbf{v}_i^b - \Delta\mathbf{v}_i^b$ in Eq. 4-29, and ignoring the product of error terms, yields:

$$\Delta\dot{\mathbf{r}}_i^i = \mathbf{C}_b^i\mathbf{v}_i^b - (\mathbf{I} - [\psi\times])\mathbf{C}_b^i(\mathbf{v}_i^b - \Delta\mathbf{v}_i^b)$$

$$= [\psi\times]\mathbf{C}_b^i\mathbf{v}_i^b + \mathbf{C}_b^i\Delta\mathbf{v}_i^b \qquad \text{Eq. 4-30}$$

From an element by element comparison, Eq. 4-30 can be written in vector form as follows:

$$\Delta\dot{\mathbf{r}}_i^i = \mathbf{C}_b^i\Delta\mathbf{v}_i^b - \mathbf{v}_i^i\times\psi \qquad \text{Eq. 4-31}$$

Eq. 4-28 describes the rotation of the estimated navigation frame as a function of the gyroscope errors. Eq. 4-31 describes the position estimation errors as a function

of the linear velocity sensor errors and the misalignment angle vector. Both equations are significant because they model the strapdown INS error states (the misalignment and position estimation error) using a set of Gaussian stochastic processes which is applicable for Kalman filtering. They are also the simplified "*general error equations*" for strapdown INS (equation 11.17 and 11.22 given by Titterton[17]) where the system in this thesis consists of the first-order terms only and is mechanised in the simplified airspace model with opposite sign convention. Figure 4-7 outlines the inertial navigation process based on this simplified model of a strapdown INS.

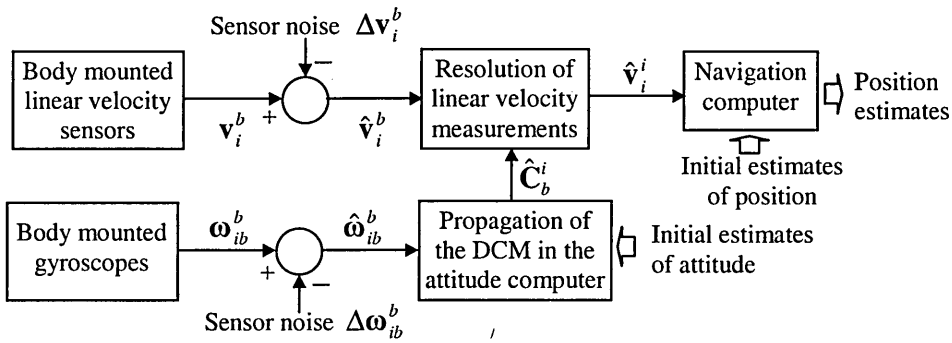


Figure 4-7 The navigation process of the simplified model of a strapdown INS

Three strapdown INS models will be simulated based on the levels of sensor accuracy given by the parameters in Table 4-1. The variance of the gyroscope error (denoted as σ_{gyro}^2) is proportional to the square of the gyroscope bias (Table 4-1), where the process time constant $1/\beta_{gyro}$ is between 10 to 15 hours[26]. The power spectral density (PSD) of the input white noise of the gyroscope error model is given by $2\sigma_{gyro}^2\beta_{gyro}$ as shown in Table 3-1. The PSD of the input white noise of the linear velocity sensor error model is given by q_{accel} (Table 4-1). These different INS models are representative of the range of inertial sensors used in civil aviation and provide the basis for the simulation of aircraft navigation and also for subsequent analysis of the Kalman filter performance.

Table 4-1 Strapdown INS sensor parameters[11]

	INS sensor accuracy		
	High	Medium	Low
Gyroscope bias (° per hour)	<0.01	0.1–1.0	10
Accelerometer white noise PSD , q_{accel} , $(\mu g/\sqrt{\text{Hz}})^2$	3^2-10^2	50^2	$>50^2$

4.4 GPS NAVIGATION ALGORITHMS

In this section, the simplified model of GPS and the GPS navigation algorithms are developed, based on the same assumptions made for the development of the strapdown INS navigation algorithms in Section 4.3.

4.4.1 Development of a simplified model of GPS

The precise description of the GPS satellite orbital trajectories is very complex. Consequently, a simplified model of the satellite orbital trajectory, based on *typical navigation satellite geometry*[11], is used for the simulation of GPS in this thesis. Although the simplified model assumes that the satellite orbital trajectory is circular, these assumptions have a minimal effect on the angular relationships between an observer and the simulated satellites. These angular relationships are vital for the simulation of GPS because they determine the satellite geometry, which dictates the time-varying characteristics of the system, including the observability and the dilution of precision. The simplified GPS satellite orbital trajectory is shown in Figure 4-8.

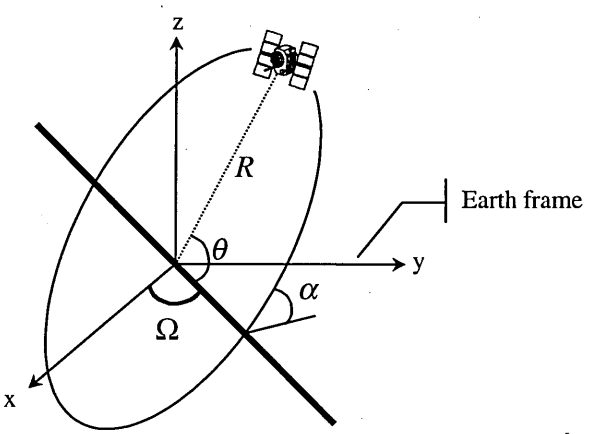


Figure 4-8 The geometry of GPS satellite circular orbital trajectory[11]

The satellite coordinate defined in the earth frame can be determined by solving the following set of equations:

$$x = R[\cos\theta\cos\Omega - \sin\theta\sin\Omega\cos\alpha] \quad \text{Eq. 4-32}$$

$$y = R[\cos\theta\sin\Omega + \sin\theta\cos\Omega\cos\alpha] \quad \text{Eq. 4-33}$$

$$z = R\sin\theta\sin\alpha \quad \text{Eq. 4-34}$$

where the inclination angle $\alpha = 55^\circ$ and the orbital radius $R = 26559800$ metres. The

right ascension Ω and the parameter θ are determined as follows:

$$\theta = \theta_0 + (t - t_0) \cdot \frac{360}{43082} \text{deg} \quad \text{Eq. 4-35}$$

$$\Omega = \Omega_0 - (t - t_0) \cdot \frac{360}{86164} \text{deg} \quad \text{Eq. 4-36}$$

In Eq. 4-35 and Eq. 4-36, t and t_0 are in second. The parameters Ω_0 and θ_0 for the current constellation of GPS satellites are given in Appendix C. Figure B-5 illustrates the simulated GPS satellite constellation for 24 satellites defined in the earth frame.

However, the simulated GPS satellite coordinates defined in the earth frame (x, y and z) must be transformed to the airspace model so that it can be used for navigation in the i-frame by the simulated aircraft. This can be achieved by transforming the satellite coordinate in the earth frame to the spherical earth frame (Λ, Φ and k) using Cartesian to Spherical transformation (Eq. 4-37, Eq. 4-38 and Eq. 4-39) and applying Eq. 4-10, Eq. 4-11 and Eq. 4-12.

$$k = \sqrt{x^2 + y^2 + z^2} \quad \text{Eq. 4-37}$$

$$\Phi = \tan^{-1}\left(\frac{k}{\sqrt{x^2 + y^2}}\right) \quad \text{Eq. 4-38}$$

$$\Lambda = \tan^{-1}\left(\frac{y}{x}\right) \quad \text{Eq. 4-39}$$

The GPS pseudorange measurements are corrupted by noise from several error sources, including Selective Availability (SA), ephemeris errors, satellite and receiver clock errors, atmospheric delays and multipath effects. The magnitude of these errors varies and each component can be modelled individually[28,29]. However, the effective pseudorange measurement errors can be considered as the statistical sum of the contributions from each of the error sources, which is termed the User Equivalent Range Error (UERE)[28]. Usually, the UERE is assumed to be a zero mean independent Gaussian noise and has the same value for all satellites. The variance of the UERE is determined from the sum of the variances of each of its components. Therefore, the pseudorange measurement error for each GPS satellite will be assumed to be Gaussian white noise with variance (denoted as σ_ρ^2) equal to $(33.3\text{m})^2$ (Table 2-3), which is the

estimated UERE for the Standard Positioning Service (SPS) with SA[28]. The SPS is accessible by most civilian and non-authorised users including the civil aviation.

One other source of error, which must be modelled, although in practice it can be eliminated with computation of the navigation equations, is the receiver clock bias. The GPS receiver clock error model is based on the *Allan variance parameters*[11]. Figure 4-9 illustrates the block diagram of this clock error model.

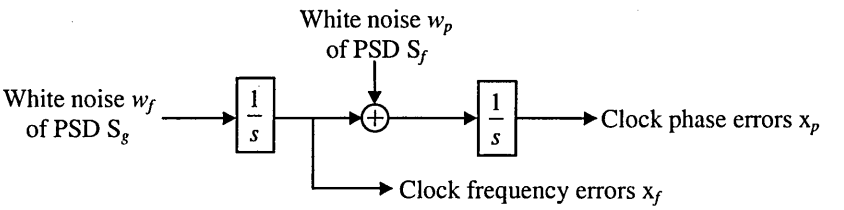


Figure 4-9 General 2-state clock error model[11]

The continuous-time clock error model is given by:

$$\dot{x}_p = x_f + w_p \qquad \text{Eq. 4-40}$$

$$\dot{x}_f = w_f \qquad \text{Eq. 4-41}$$

where x_p and x_f are the clock phase (seconds) and frequency (Hz) errors. w_p and w_f are the Gaussian white noise inputs with PSD S_f and S_g respectively.[i] S_f and S_g can be determined from the power spectral density coefficients given in Table 4-2 as follows:

$$S_f \cong \frac{h_0}{2} \qquad \text{Eq. 4-42}$$

$$S_g \cong 2\pi^2 h_2 \qquad \text{Eq. 4-43}$$

Table 4-2 Typical power spectral density (PSD) coefficients for various timing standards[11] (The units of h_0 and h_2 are omitted since they are not given in the reference[11])

Timing standard	Accuracy level	h_0	h_2
Compensated crystal	Low	2×10^{-19}	2×10^{-20}
Ovenized crystal	Medium	8×10^{-20}	4×10^{-23}
Rubidium	High	2×10^{-20}	4×10^{-29}

[i] Although the notation is confusing (the subscript f is used for both frequency and phase noise, i.e. $x_p \rightarrow w_p \rightarrow S_f$ and $x_f \rightarrow w_f \rightarrow S_g$), it is presented as such in order to be consistent with the reference[11].

4.4.2 Development of general pseudorange measurement model

This section reviews the derivation of the general pseudorange measurement model, which forms the basis of GPS navigation algorithms. The following notation is used in the derivation of the model, where the subscripts $_s$ and $_o$ denote the *source* and the *observer*. In terms of GPS navigation, the source and the observer refer to the GPS satellite and a GPS receiver respectively.

t_s	is the system time at which the signal is transmitted by the source.
t_o	is the system time at which the signal is received by the observer.
b_s	is the source clock offset from the system time.
b_o	is the observer clock offset from the system time.
$\hat{t}_s = t_s - b_s$	is the clock value at the time the signal is transmitted. The sign convention is defined in Eq. 3-28.
$\hat{t}_o = t_o - b_o$	is the clock value at the time the signal is received.
c	is the speed of light.
$\mathbf{r}_s = [r_{sx}, r_{sy}, r_{sz}]^T$	is the source position vector defined in the system reference frame given by the x, y and z axes.
$\mathbf{r}_o = [r_{ox}, r_{oy}, r_{oz}]^T$	is the observer position vector defined in the system reference frame.

The geometric range r between a source and an observer can be determined from the position vectors as follows:

$$r = \|\mathbf{r}_s - \mathbf{r}_o\| = \sqrt{(r_{sx} - r_{ox})^2 + (r_{sy} - r_{oy})^2 + (r_{sz} - r_{oz})^2} \qquad \text{Eq. 4-44}$$

Equivalently, r can be determined from the signal propagation time as follows:

$$r = c \cdot (t_o - t_s) \qquad \text{Eq. 4-45}$$

The pseudorange measurement ρ is given by:

$$\rho = c \cdot (\hat{t}_o - \hat{t}_s)$$

$$= c \cdot [t_o - b_o - (t_s - b_s)]$$

$$= c \cdot (t_o - t_s) + c \cdot (b_s - b_o) \qquad \text{Eq. 4-46}$$

Substituting Eq. 4-44 and Eq. 4-45 in Eq. 4-46 yields:

$$\rho = \sqrt{(r_{sx} - r_{ox})^2 + (r_{sy} - r_{oy})^2 + (r_{sz} - r_{oz})^2} + c \cdot (b_s - b_o) \qquad \text{Eq. 4-47}$$

The *general pseudorange measurement model*, which takes into account a measurement noise, can be derived from Eq. 4-47 as follows:

$$\rho = h(\mathbf{x}_s, \mathbf{x}_o) + v \qquad \text{Eq. 4-48}$$

where $\mathbf{x}_s = [r_{sx}, r_{sy}, r_{sz}, b_s]^{\mathrm{T}}$, $\mathbf{x}_o = [r_{ox}, r_{oy}, r_{oz}, b_o]^{\mathrm{T}}$ and v denote the source state vector, the observer state vector and the measurement noise respectively and the nonlinear function $h(\mathbf{x}_s, \mathbf{x}_o) = \sqrt{(r_{sx} - r_{ox})^2 + (r_{sy} - r_{oy})^2 + (r_{sz} - r_{oz})^2} + c \cdot (b_s - b_o)$.

In order to determine the observer state estimate $\hat{\mathbf{x}}_o$, a set of simultaneous measurements must be available from at least four different sources. Several analytical and numerical methods can be used to solve this set of non-linear measurement equations (Eq. 4-48).[28] One method is to linearise these non-linear equations about a reference trajectory ($\mathbf{x}_s^*$ and $\mathbf{x}_o^*$) using a Taylor series expansion, as explained in Section 3.4. The observer incremental state estimate $\Delta\hat{\mathbf{x}}_o$, which is obtained by solving the linearised equations, is then used to derive $\hat{\mathbf{x}}_o$ from the reference trajectory. This approach is widely used for stand-alone GPS navigation and is known as the *single-point solution*[30]. This method is important because it produces the *linearised pseudorange measurement model* in a form, which is applicable to Kalman filtering. It also models the source uncertainty, which is required for the derivation of relative navigation algorithms. Eq. 4-48 can be linearised as follows:

$$\rho = h(\mathbf{x}_s^* + \Delta\mathbf{x}_s, \mathbf{x}_o^* + \Delta\mathbf{x}_o) + v \qquad \text{Eq. 4-49}$$

where $\Delta\mathbf{x}_s = [\Delta r_{sx}, \Delta r_{sy}, \Delta r_{sz}, \Delta b_s]^{\mathrm{T}}$ and $\Delta\mathbf{x}_o = [\Delta r_{ox}, \Delta r_{oy}, \Delta r_{oz}, \Delta b_o]^{\mathrm{T}}$. Expanding Eq. 4-49 using Taylor series and ignoring terms higher than first-order, it can be shown that:

$$\rho \cong \rho^* + \left.\frac{\partial h}{\partial \mathbf{x}_s}\right|_{\substack{\mathbf{x}_s=\mathbf{x}_s^* \\ \mathbf{x}_o=\mathbf{x}_o^*}} \cdot \Delta\mathbf{x}_s + \left.\frac{\partial h}{\partial \mathbf{x}_o}\right|_{\substack{\mathbf{x}_s=\mathbf{x}_s^* \\ \mathbf{x}_o=\mathbf{x}_o^*}} \cdot \Delta\mathbf{x}_o + v \qquad \text{Eq. 4-50}$$

where $\rho^* = h(\mathbf{x}_s^*, \mathbf{x}_o^*)$ and the partial derivatives, which are evaluated at the reference trajectory, are given by:

$$\left.\frac{\partial h}{\partial \mathbf{x}_s}\right|_{\substack{\mathbf{x}_s=\mathbf{x}_s^* \\ \mathbf{x}_o=\mathbf{x}_o^*}} = \left[\frac{r_{sx}^* - r_{ox}^*}{\rho^*}, \frac{r_{sy}^* - r_{oy}^*}{\rho^*}, \frac{r_{sz}^* - r_{oz}^*}{\rho^*}, c\right] \qquad \text{Eq. 4-51}$$

$$\left.\frac{\partial h}{\partial \mathbf{x}_o}\right|_{\substack{\mathbf{x}_s=\mathbf{x}_s^* \\ \mathbf{x}_o=\mathbf{x}_o^*}} = -\left[\frac{r_{sx}^* - r_{ox}^*}{\rho^*}, \frac{r_{sy}^* - r_{oy}^*}{\rho^*}, \frac{r_{sz}^* - r_{oz}^*}{\rho^*}, c\right] = -\left.\frac{\partial h}{\partial \mathbf{x}_s}\right|_{\substack{\mathbf{x}_s=\mathbf{x}_s^* \\ \mathbf{x}_o=\mathbf{x}_o^*}} \qquad \text{Eq. 4-52}$$

Substituting $\Delta\rho = \rho - \rho^*$ and $\mathbf{h}_{so} = \left.\frac{\partial h}{\partial \mathbf{x}_o}\right|_{\substack{\mathbf{x}_s=\mathbf{x}_s^* \\ \mathbf{x}_o=\mathbf{x}_o^*}}$ in Eq. 4-50, the linearised pseudorange measurement model can be written as follows:

$$\Delta\rho = \mathbf{h}_{so}\Delta\mathbf{x}_o - \mathbf{h}_{so}\Delta\mathbf{x}_s + v \qquad \text{Eq. 4-53}$$

The next section discusses how to use Eq. 4-53 to determine the observer state estimate.

4.4.3 Stand-alone GPS navigation algorithms: the single-point solution

In practice, GPS navigation data processed by an observer are assumed to be perfect by definition, i.e. $\Delta\mathbf{x}_s = \mathbf{0}$. Therefore, Eq. 4-53 can be reduced and written in matrix form, which is the standard GPS linearised measurement model, as follows:

$$\Delta\boldsymbol{\rho} = \mathbf{H}\Delta\mathbf{x}_o + \mathbf{v} \qquad \text{Eq. 4-54}$$

where the i^{th} row of **H,** which takes the form of $\mathbf{h}_{so}$ defined in Eq. 4-52, represents the measurement from the i^{th} source as shown in Eq. 4-55.

$$\Delta\rho_i = -\left[\frac{r_{s_i x}^* - r_{ox}^*}{\rho_i^*}, \frac{r_{s_i y}^* - r_{oy}^*}{\rho_i^*}, \frac{r_{s_i z}^* - r_{oz}^*}{\rho_i^*}, c\right]\Delta\mathbf{x}_o + v_i \qquad \text{Eq. 4-55}$$

The **H** matrix is also referred to as the *geometry matrix* or the *direction cosine matrix* because it represents the source-observer angular relationships (cosine terms).

The geometry matrix is important because it can be used to determine the levels of accuracy of a navigation solution as a function of the observer-observer geometry. It can be shown that, by solving Eq. 4-54 based on *least squares estimation*, the estimate of $\Delta\mathbf{x}_o$ is given by:

$$\Delta\hat{\mathbf{x}}_o = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\Delta\boldsymbol{\rho} \qquad \text{Eq. 4-56}$$

where $\mathbf{R} = \mathrm{E}[\mathbf{v}\mathbf{v}^T]$ is the covariance matrix of the measurement noise $\mathbf{v}$ and the *estimation error covariance matrix* $\mathbf{P} = \mathrm{E}[(\Delta\mathbf{x}_o - \Delta\hat{\mathbf{x}}_o)(\Delta\mathbf{x}_o - \Delta\hat{\mathbf{x}}_o)^T]$, which identifies the error bounds of the state estimate, is given by:

$$\mathbf{P} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1} \qquad \text{Eq. 4-57}$$

The least squares estimation is widely used because it is a linear unbiased state estimator, which takes into account the stochastic properties of the measurement noise. The summary of the least squares estimation is given in Appendix D. $\Delta\hat{\mathbf{x}}_o$ can be used to derive the observer state estimate $\hat{\mathbf{x}}_o$ from the reference trajectory as follows:

$$\hat{\mathbf{x}}_o = \mathbf{x}_o^* + \Delta\hat{\mathbf{x}}_o \qquad \text{Eq. 4-58}$$

The single-point solution process can be iterated by using the most recent estimate $\hat{\mathbf{x}}_o$ as a new reference trajectory to refine the observer state estimate to the required accuracy levels. Nevertheless, it should be noted that, by assuming $\Delta\mathbf{x}_s = \mathbf{0}$, any deviations of the actual GPS states from the ideal system model will be included in the observer navigation solutions. This is equivalent to defining the states of the observer with respect to the *erroneous* GPS reference frame and system time.

4.4.4 Dilution of precision

This section reviews the use of the observer's estimation error covariance matrix $\mathbf{P}$ to determine the accuracy of a GPS navigation solution as a function of source-observer geometry i.e. from the *dilution of precision* (DOP). By assuming the measurement noise is uncorrelated and has equal variance σ^2, the measurement noise covariance matrix $\mathbf{R}$ can be expressed as $\mathbf{R} = \sigma^2\mathbf{I}$. Consequently, $\mathbf{P}$ in Eq. 4-57 can be reduced to:

$$\mathbf{P} = \sigma^2 (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \qquad \text{Eq. 4-59}$$

The DOP terms are determined, based on the contribution of geometry, as the appropriate sums of terms on the diagonal of the *normalised* **P** matrix, which is defined as follows:[27,30]

$$\text{normalized } \mathbf{P} = \mathbf{A} = \frac{\mathbf{P}}{\sigma^2} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \qquad \text{Eq. 4-60}$$

Therefore, for example, the geometrical DOP (GDOP) = $\sqrt{\text{trace}(\mathbf{A})}$ and the horizontal DOP (HDOP) = $\sqrt{\mathbf{A}_{11} + \mathbf{A}_{22}}$ where the subscript denotes the respective element of **A**.

Although the assumption about the measurement noise is unrealistic, the ability to determine the accuracy levels based on the geometry can be useful in practice. For example, when data from only a limited number of sources can be processed simultaneously, the comparison of the DOP values allow an observer to select the set of sources which will produce the best navigational performance. In this *"source selection"* process, the set of sources with the lowest DOP values is selected.

4.5 STATE-SPACE FORMULATION

In the previous sections, a set of differential equations were developed, which can be used to analyse the performance of the strapdown INS and GPS. This section summarises and presents these equations using the state-space notation described in Section 3.2. Let the respective axes of the *i*-frame be x, y and z axis such that, for example, the position vector, the linear and angular velocity vector can be written in component forms as $\mathbf{r}_i^i = [r_{\mathrm{x}}, r_{\mathrm{y}}, r_{\mathrm{z}}]^{\mathrm{T}}$, $\mathbf{v}_i^i = [v_{\mathrm{x}}, v_{\mathrm{y}}, v_{\mathrm{z}}]^{\mathrm{T}}$ and $\boldsymbol{\omega}_{ib}^b = [\omega_{\mathrm{x}}, \omega_{\mathrm{y}}, \omega_{\mathrm{z}}]^{\mathrm{T}}$. The simulated aircraft will be constrained to manoeuvre in the horizontal plane i.e. level flight, which is typical for en-route phase of flight, and, therefore, the following additional conditions are assumed to be satisfied:

$$v_{\mathrm{z}} = \Delta v_{\mathrm{z}} = \omega_{\mathrm{x}} = \omega_{\mathrm{y}} = \Delta\omega_{\mathrm{x}} = \Delta\omega_{\mathrm{y}} = \dot{\psi}_{\mathrm{x}} = \dot{\psi}_{\mathrm{y}} = 0 \qquad \text{Eq. 4-61}$$

$$\mathbf{C}_b^i = \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Eq. 4-62}$$

The navigation system errors can be described using an eight-state error model derived from the following set of differential equations:

$$\dot{\psi}_z = \Delta\omega_z \qquad \text{(from Eq. 4-28)} \qquad \text{Eq. 4-63}$$

$$\Delta\dot{r}_x = -v_y\psi_z + c_{11}\Delta v_x + c_{12}\Delta v_y \qquad \text{(from Eq. 4-31)} \qquad \text{Eq. 4-64}$$

$$\Delta\dot{r}_y = v_x\psi_z + c_{21}\Delta v_x + c_{22}\Delta v_y \qquad \text{(from Eq. 4-31)} \qquad \text{Eq. 4-65}$$

$$\dot{x}_p = x_f + w_p \qquad \text{(from Eq. 4-40)} \qquad \text{Eq. 4-66}$$

$$\dot{x}_f = w_f \qquad \text{(from Eq. 4-41)} \qquad \text{Eq. 4-67}$$

$$\Delta\dot{\omega}_z = -\beta_{gyro}\Delta\omega_z + w_{\Delta\omega_z} \qquad \text{(from Table 3-1)} \qquad \text{Eq. 4-68}$$

$$\Delta\dot{v}_x = w_{\Delta v_x} \qquad \text{(from Table 3-1)} \qquad \text{Eq. 4-69}$$

$$\Delta\dot{v}_y = w_{\Delta v_y} \qquad \text{(from Table 3-1)} \qquad \text{Eq. 4-70}$$

where w denotes the system input Gaussian white noise. ψ_z, Δr_x and Δr_y are the azimuth (heading), x-position and y-position errors. $\Delta\omega_z$, Δv_x and Δv_y are the vertical gyroscope, x- and y-linear velocity sensor noise. x_p and x_f are the clock phase and frequency errors. The augmented overall system model can be written in matrix form as follows:

$$\dot{\mathbf{x}}_o = \overbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -v_y & 0 & 0 & 0 & 0 & 0 & c_{11} & c_{12} \\ v_x & 0 & 0 & 0 & 0 & 0 & c_{21} & c_{22} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\beta_{gyro} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}^{\mathbf{F}} \mathbf{x}_o + \mathbf{G}\mathbf{w} \qquad \text{Eq. 4-71}$$

$$\begin{bmatrix} \Delta\rho_1 \\ \vdots \\ \Delta\rho_n \end{bmatrix} = \overbrace{\begin{bmatrix} 0 & \dfrac{r^*_{s_1x} - r^*_{ox}}{\rho^*_1} & \dfrac{r^*_{s_1y} - r^*_{oy}}{\rho^*_1} & c & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dfrac{r^*_{s_nx} - r^*_{ox}}{\rho^*_n} & \dfrac{r^*_{s_ny} - r^*_{oy}}{\rho^*_n} & c & 0 & 0 & 0 & 0 \end{bmatrix}}^{\mathbf{H}} \mathbf{x}_o + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \qquad \text{Eq. 4-72}$$

where $\mathbf{x}_o = [\psi_z, \Delta r_x, \Delta r_y, x_p, x_f, \Delta\omega_z, \Delta v_x, \Delta v_y]^T$ is the observer state vector. $\mathbf{G}$ is the 8-by-8 identity matrix and $\mathbf{w} = [0, 0, 0, w_p, w_f, w_{\Delta\omega_z}, w_{\Delta v_x}, w_{\Delta v_y}]^T$. The subscript $_n$ in the measurement model (Eq. 4-72) denotes the nth source. The PSD of the system input noise and the measurement noise error covariance matrix are given by the 8-by-8 diagonal matrix $\mathbf{GQG}^T$ and the matrix $\mathbf{R} = \mathrm{E}[\mathbf{vv}^T]$ in Eq. 4-73 and Eq. 4-74 respectively.

$$\mathbf{GQG}^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & S_f & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & S_g & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2\sigma^2_{gyro}\beta_{gyro} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & q_{accel} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{accel} \end{bmatrix} \qquad \text{Eq. 4-73}$$

$$\mathbf{R} = \begin{bmatrix} \sigma^2_{\rho_1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\rho_2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2_{\rho_n} \end{bmatrix} \qquad \text{Eq. 4-74}$$

In practice, the elements on the diagonal of $\mathbf{GQG}^T$ which are zeros are usually replaced with small non-zero terms to add fictitious input noise to the system, in order to improve the stability of the Kalman filter, as described in Section 3.5.2. The validation of the navigation system and sensor error models based on this augmented overall system model (Eq. 4-71 and Eq. 4-72) is described in Appendix B. The next section analyses the performance of the stand-alone navigation systems and discusses factors which influence the system performance based on the error model described in this section.

4.6 STAND-ALONE SYSTEM PERFORMANCE ANALYSIS

Accuracy is a critical attribute of a navigation system, and is widely used as a primary criterion for evaluating system performance. Theoretically, the accuracy can be determined from the root-mean-square (rms) errors, which are the square root of the

diagonal terms of the error covariance matrix (assuming that the errors are zero mean and uncorrelated in time). The *true* rms and mean errors used in this thesis have been derived from 50 Monte Carlo simulations and used to validate the error model by comparing these values with the theoretical values. Although 50 simulations do not produce precise values of the rms and mean errors, it adequately shows the trends, which is sufficient for the validation process required in this thesis. Three sensor accuracy levels based on the parameters in Table 4-3 (obtained from Table 4-1 and Table 4-2) are used in all subsequent analysis.

Table 4-3 Sensor parameters for navigation system simulation (The units of h_0 and h_2 are omitted since they are not given in the reference[11])

Sensor parameters	Accuracy levels		
	High	**Medium**	**Low**
Gyroscope bias ($^\circ/\text{hr}$)	0.01	1	10
Time constant (hours)	10	10	10
Accelerometer PSD ($\mu g/\sqrt{\text{Hz}}$)^2	3^2	50^2	200^2
Clock h_0	2×10^{-20}	8×10^{-20}	2×10^{-19}
Clock h_2	4×10^{-29}	4×10^{-23}	2×10^{-20}
Initial azimuth error (1σ) ($^\circ$)	0.01	0.1	1
Initial x and y position error (1σ) (m)	10	100	500
The initial errors are assumed zero mean Gaussian random variables.			

4.6.1 Stand-alone strapdown INS performance analysis

The effect of large misalignment errors on the system error model

The analysis of the strapdown INS is based on the assumption that system errors can be adequately modelled using Eq. 4-71. Figure 4-10 illustrates the *true* and the *modelled* errors obtained from one Monte Carlo simulation. The true errors are the differences between the true states of the simulated aircraft and the strapdown INS states. The modelled errors are determined by solving Eq. 4-71 using the true sensor errors, which are known precisely in the simulation. Figure 4-10b and Figure 4-10c show the divergence in the position errors as the azimuth error ψ_z accumulates. This divergence occurs because the magnitudes of higher order terms, which are ignored in Eq. 4-71 (on the assumption that the misalignment ψ is small), become significant.

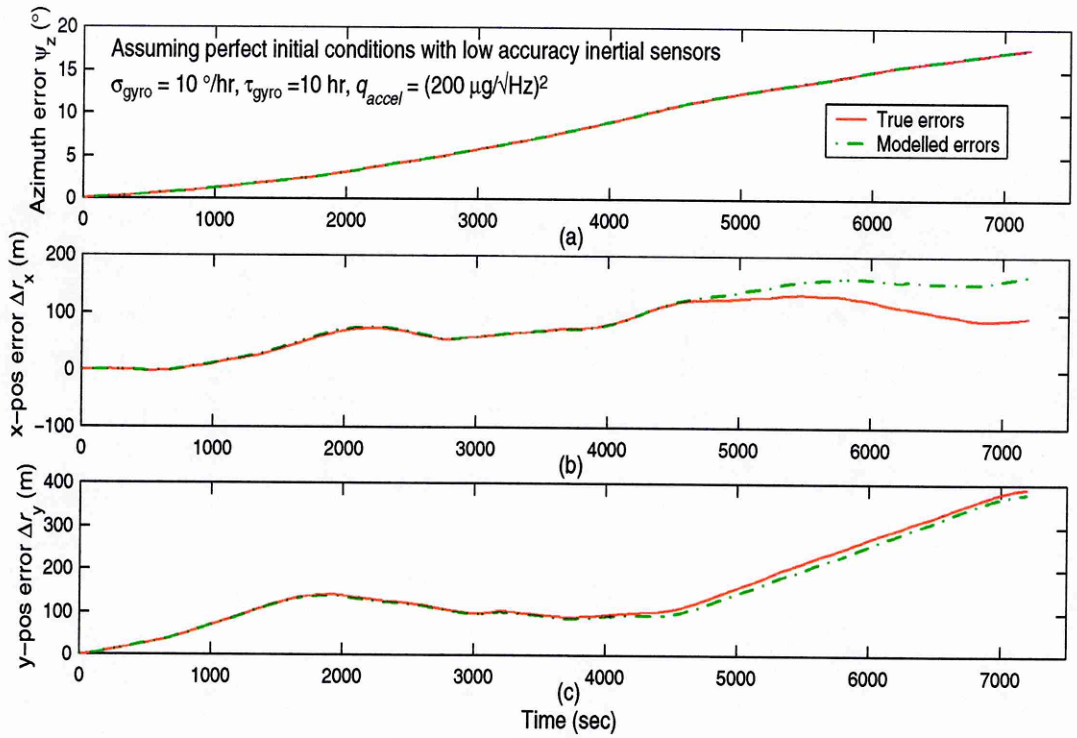


Figure 4-10 Comparison of true and modelled errors of a low accuracy strapdown INS

It is important to understand this effect because the azimuth error can become so large that the error model becomes invalid, affecting the performance of the Kalman filter and resulting in the diverge of the filter. Under these circumstances, it is necessary to develop an error model, which takes account of large misalignment errors.

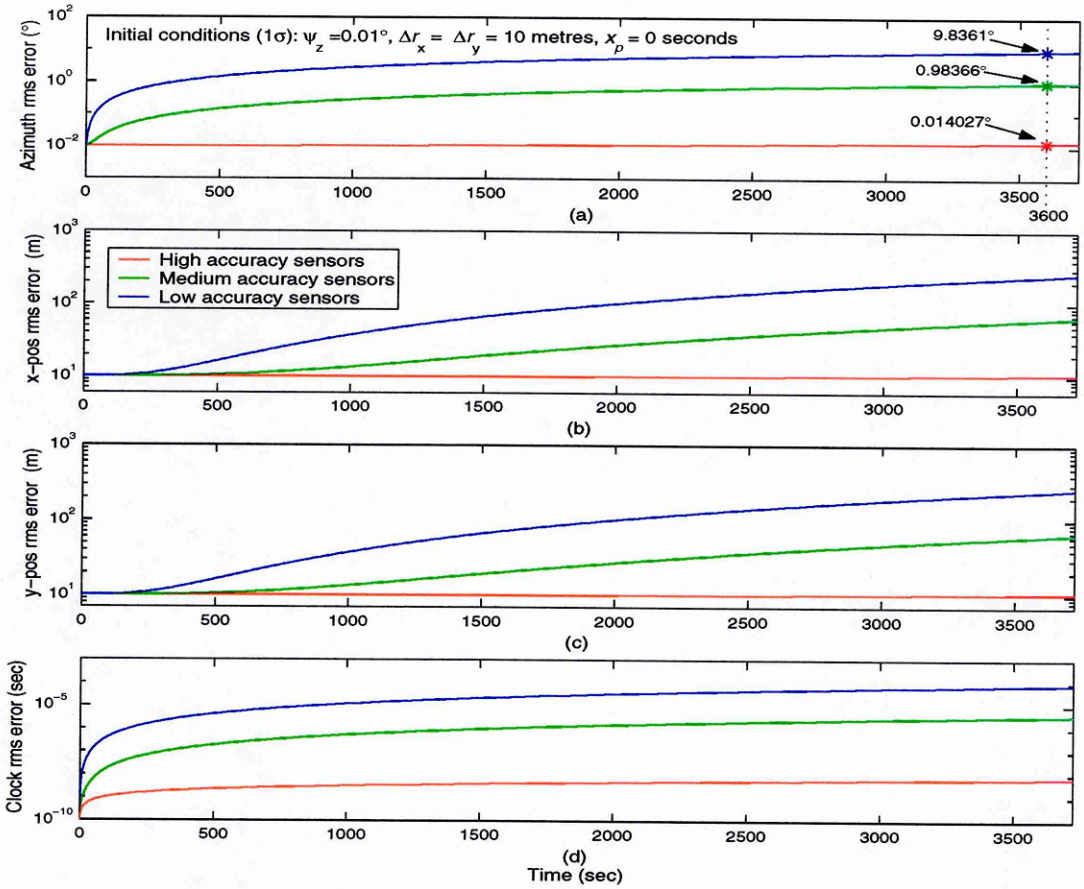


Figure 4-11 Comparison of a stationary strapdown INS and clock performance

The performance of an INS degrades with time because system errors accumulate from the integration of measurement noise.[33,34] This is illustrated in Figure 4-11 which shows the theoretical rms errors of a stationary strapdown INS for all three sensor accuracy levels. The error covariance matrix of the system is calculated by finding the equivalent discrete time model of Eq. 4-71 using the methods described in Section 3.2.2 and applying Eq. 3-21. In Figure 4-11a, the deviations of the azimuth rms errors after 3600 seconds from the three gyroscope bias values given in Table 4-3 are caused by truncating the higher order terms in the error model. In addition, Figure 4-11 illustrates that the system performance also depends on the sensor accuracy, where more accurate sensors reduce the rms errors. The rms errors of the simulated clock shown in Figure 4-11d also exhibit the same characteristics as the INS errors i.e. the errors grow with time and depend on the sensor quality.

The effect of aircraft manoeuvres on the stand-alone strapdown INS performance

This section analyses the effects of aircraft manoeuvres on the performance of the strapdown INS. The flight trajectories used for the analysis are shown in Figure 4-12. The simulated aircraft flies level at 20000 feet with a constant TAS of 387 knots, based on the BADA data for the Boeing 777 aircraft. This flight trajectory is also used for the subsequent analysis of the stand-alone GPS and integrated INS/GPS system. Although this trajectory has been chosen arbitrarily, it provides an adequate model of typical aircraft manoeuvres for the analysis in this thesis.

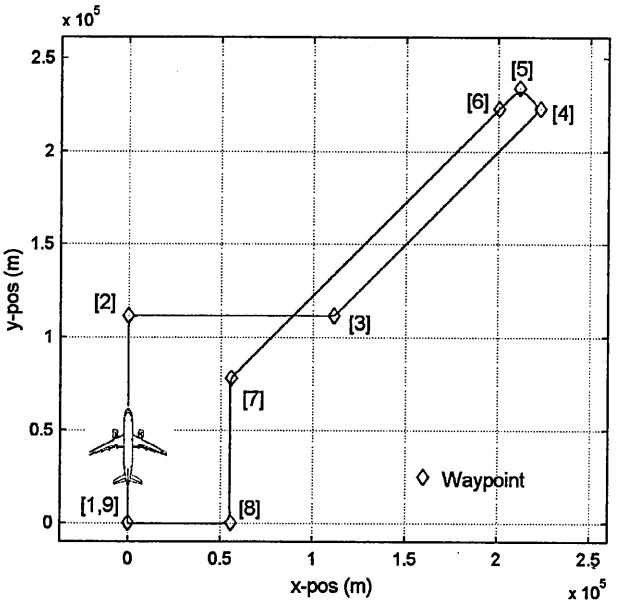


Figure 4-12 Flight trajectory of the simulated aircraft

The performance of the strapdown INS is affected by aircraft manoeuvres because the position error states are coupled with the velocity components in the error model (Eq. 4-64 and Eq. 4-65). The effects of the coupling can be seen in Figure 4-13 by comparing the change in the position rms errors with the change in the profile of velocity components. This implies that the position rms errors grow at a rate proportional to the aircraft velocity, which is confirmed by inspecting the true rms and mean errors. By observing the actual error samples in Figure 4-13, it should be noted that the strapdown INS navigation solution is very stable during relatively short time intervals because the error states do not change steeply between successive samples.

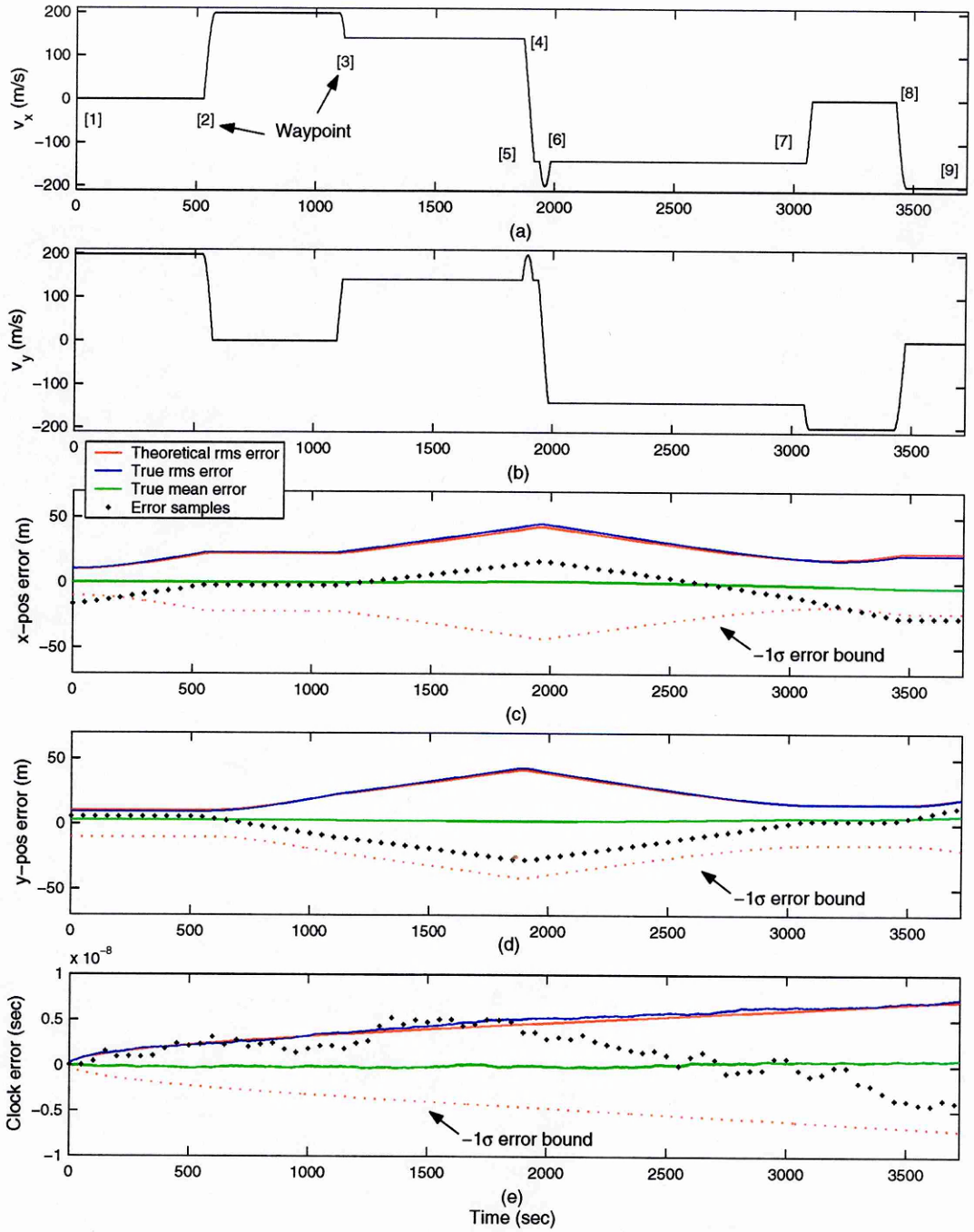


Figure 4-13 The performance of the stand-alone INS with the high accuracy sensors

4.6.2 Stand-alone GPS performance analysis

Although the performance of the stand-alone GPS single-point solution does not degrade with time, the system accuracy depends on the source-observer geometry and the pseudorange measurement accuracy.[28,35]

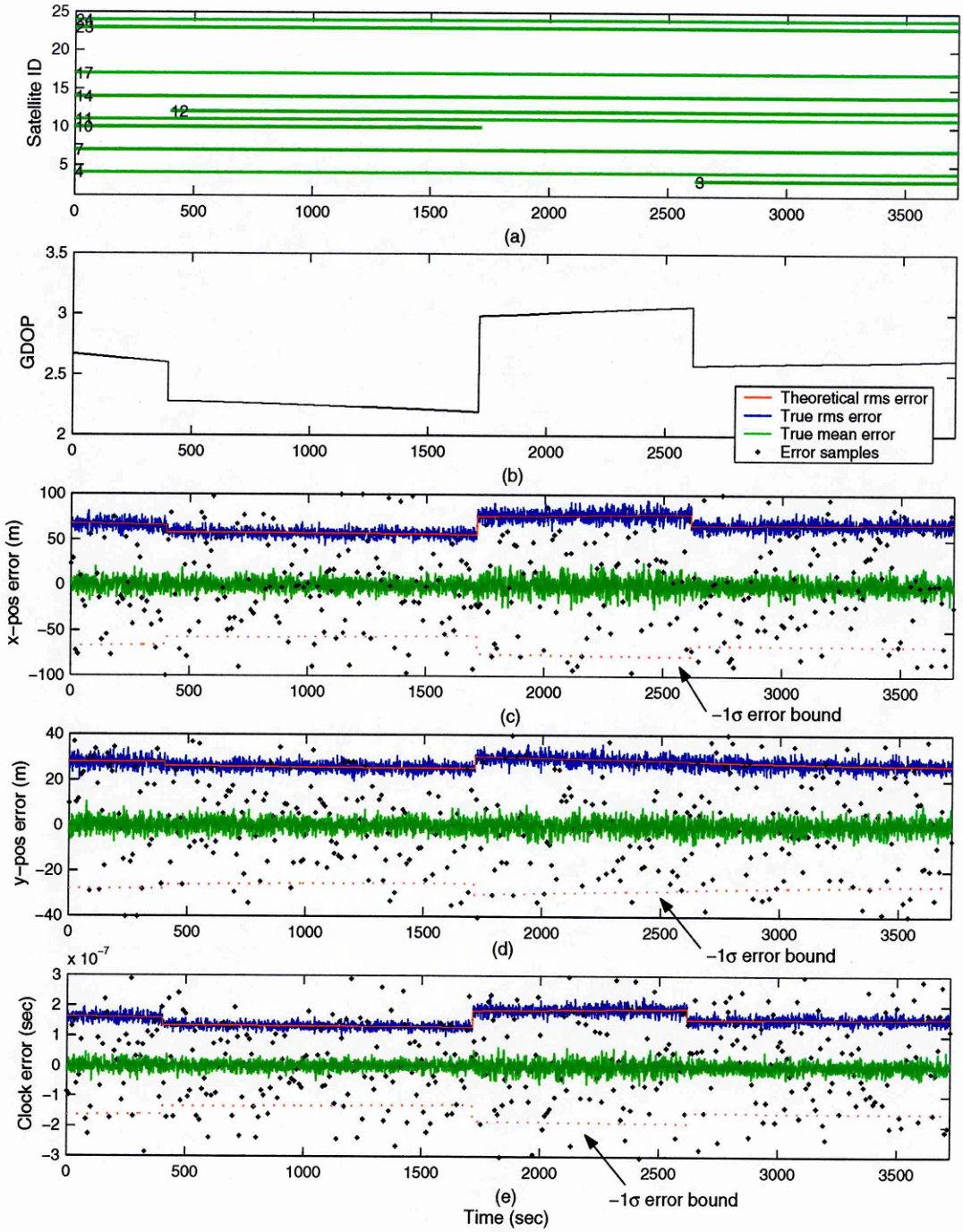


Figure 4-14 The performance of the stand-alone GPS with the high accuracy sensors

Figure 4-14a and Figure 4-14b show the available GPS satellites and the resulting GDOP for the simulated flight trajectory in Figure 4-12. The theoretical rms errors shown in Figure 4-14 are determined by solving Eq. 4-57. The effect of the source-observer geometry on the stand-alone GPS accuracy can be seen by comparing

the change in the rms errors with the change in the GDOP. The validity of the measurement model is confirmed by inspecting the true rms and mean errors. By observing the error samples in Figure 4-14, it should be noted that the stand-alone GPS errors change more steeply in comparison with the stand-alone INS. The effects of the measurement accuracy on the stand-alone GPS performance is illustrated in Figure 4-15, where the rms error is directly proportional to the magnitude of the measurement noise standard deviation. This implies that one method of improving the performance of GPS (or any systems that use pseudorange measurements for navigation such as relative navigation described in Chapter 5) can be improved directly by reducing the measurement noise.

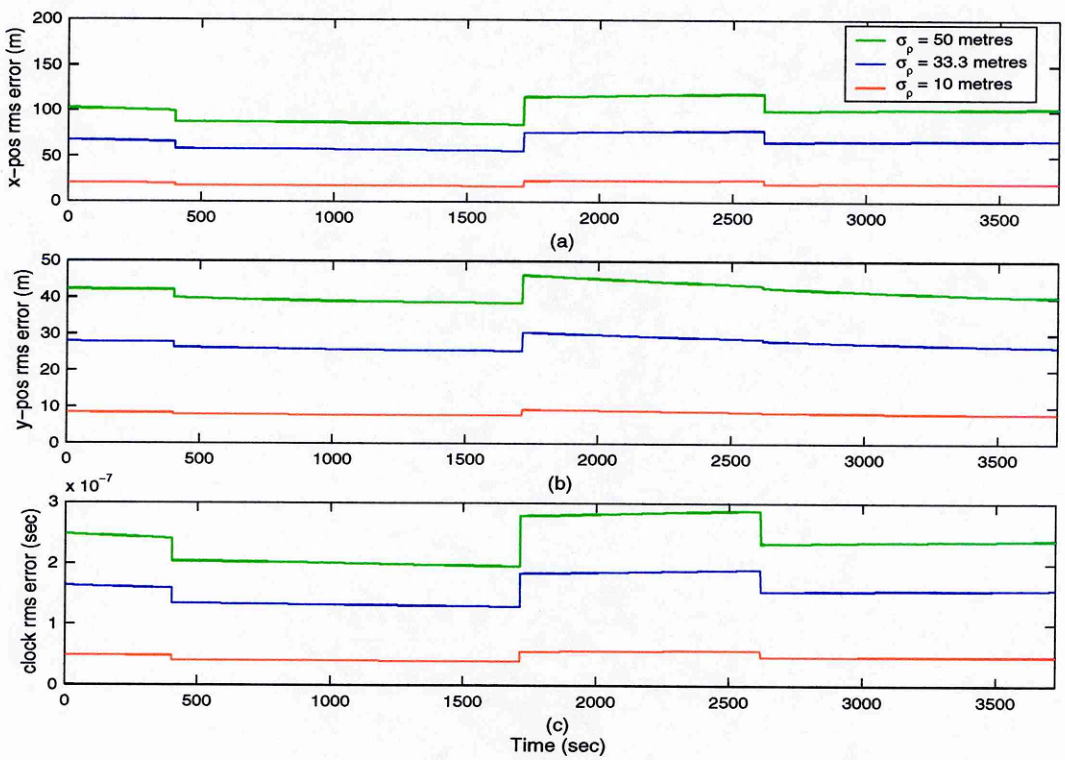


Figure 4-15 The effects of the measurement accuracy on the stand-alone GPS performance

In this section, the performance and error characteristics of the stand-alone strapdown INS and GPS system have been analysed based on error covariance analysis. Table 4-4 summarises the main characteristics of both systems.

Table 4-4 Characteristics of stand-alone INS and GPS[35,37]

Parameters	INS	GPS
Sensor output rate	High (50-200 Hz)	Low (1 Hz)
Continuity of outputs	Only interrupted by system failures	Susceptible to discontinuity owing to intentional and unintentional interference, and vehicle's manoeuvres
Error characteristics	Errors grow with time and are affected by vehicle manoeuvres Navigation solutions are stable during short time intervals and include attitude information	Errors do not grow with time but are affected by geometry Navigation solutions change steeply in comparison with the stand-alone INS

4.7 NAVIGATION SYSTEMS INTEGRATION

4.7.1 Introduction

Navigation systems integration is a method of combining information from different kinds of sensor systems, which are available in most aircraft, to produce navigation solutions which are superior to independent navigation systems. The integrated systems achieve improved performance by methodically combining data from various sensors, which have complementary error characteristics, and deriving the optimal navigation solutions by extracting the appropriate characteristics from each sensor. The Kalman filter is the primary state estimator applied for data fusion in integrated navigation systems. The acceptance of this filter comes from its ability to accommodate time varying systems and its recursive nature, which make it highly suitable for digital computer implementation. The Kalman filter is universally accepted as the *best* estimator of the class of linear estimators.[9,11] This section reviews and analyses the navigation systems integration based on INS and GPS which forms the basis of relative navigation investigated in this thesis.

It can be seen from Table 4-4 that INS and GPS have a number of complementary error characteristics and several benefits can be obtained from systems integration. The high sensor output rate of INS and its self-contained nature can be used to alleviate the problem of short-term GPS outages, which may be caused by signal interference, excessive vehicle manoeuvres and discontinuities caused by changes to the satellite constellation. The complete six degree of freedom (position and orientation)

navigation data provided by INS also augment the navigation states not available from the single-point solution of GPS. On the other hand, the manoeuvre and time independent properties of GPS can be used to suppress the error growth of INS i.e. to perform *in-flight alignment* of INS using GPS. Generally, the integration of INS/GPS results in a system with superior performance in comparison with stand-alone systems.

In practice, however, INS/GPS integration is a complex process involving several factors and considerations.[35,36] The degree of complexity in the design of the integrated system are usually determined by the system performance requirements. The next section discusses one integration configuration, which has been used extensively for integrating information from other sensors to aid INS, known as the *indirect feedback filter*[9] or the *feedback complementary filter*[11].

4.7.2 Indirect feedback filter design of the integrated INS/GPS system

The indirect feedback filter design is widely implemented in many real-time integrated systems because it offers several practical advantages including reduced computing load, superior reliability and performance.[9,11] The indirect feedback Kalman filter configuration for integrated INS/GPS system is illustrated in Figure 4-16. It can be seen that this configuration is very similar to the configuration used in the extended Kalman filter (EKF) described in Section 3.4. In integrated INS/GPS systems, the INS errors are modelled using the incremental state vector (Eq. 4-71) and a linearised measurement model, given by Eq. 4-72.

The primary advantage of this configuration is that the Kalman filter is situated outside the INS processing loop. Therefore, the Kalman filter can process the measurements at a much lower rate than the INS update rate, reducing the computational task while allowing the system to respond to the aircraft dynamics more accurately. The reliability of the system is also maintained because the operation of the INS is independent of the Kalman filter. The feedback configuration, which matches the level of accuracy of the reference trajectory, ensures continuous validity of the linearised model. Additionally, this configuration readily accepts diverse mixes of sensor aiding by modifications to the system software alone.

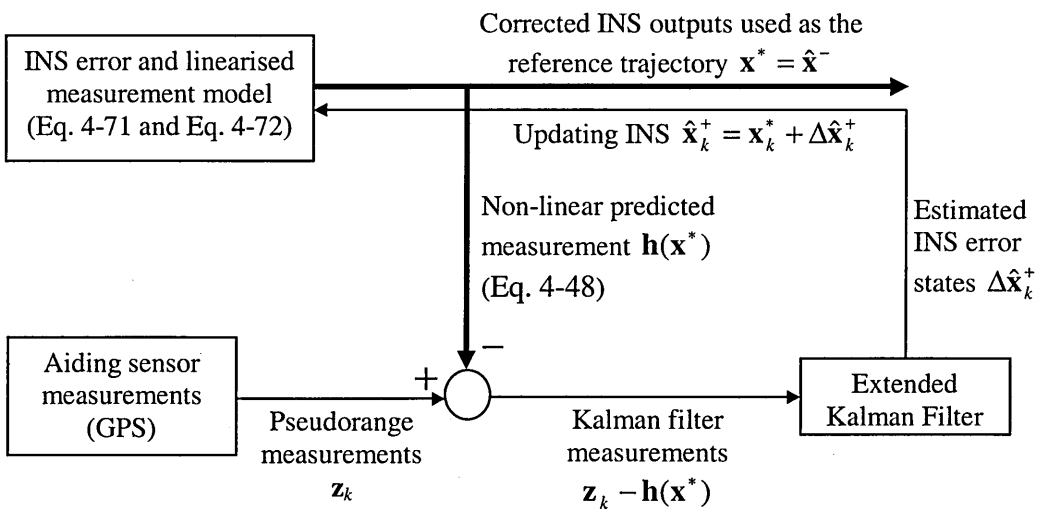


Figure 4-16 The indirect feedback Kalman filter configuration

The performance of the integrated INS/GPS system, based on the indirect feedback configuration, is analysed in the next section. It will be shown that the performance of the integrated system is superior to independent systems. In addition, some practical aspects of Kalman filter implementation, which are required for this application, will also be discussed.

4.7.3 Integrated INS/GPS system performance analysis

Figure 4-17 illustrates the performance of the integrated INS/GPS system for the flight trajectory given in Figure 4-12 for an aircraft with high accuracy sensors[i]. By observing the error samples, it can be seen that the best characteristics of the INS and GPS sensors are retained; the error samples are smoothed and do not grow with time.

Figure 4-18 compares the theoretical rms errors for the stand-alone INS, GPS and the integrated system given in Figure 4-13, Figure 4-14 and Figure 4-17. It can be seen that the performance of the integrated system is superior to either system operating independently where the rms errors of the integrated system are reduced in comparison to the stand-alone systems. Moreover, the manoeuvre and time dependent errors of the INS and the geometry dependent errors of the GPS are clearly suppressed. Although the attitude information can not be determined by the GPS single-point solution, the GPS

[i] The terms high, medium and low, referring to sensor accuracy and initial condition accuracy, are given in Table 4-3.

pseudorange measurements can be used to perform in-flight alignment of the strapdown INS computer frame effectively, as shown in Figure 4-18a. In this case, the Kalman filter tries to align the computer frame of the strapdown INS with the GPS system reference frame. The effects of pseudorange measurement accuracy on the integrated system performance are illustrated in Figure 4-19 where the more accurate measurements result in better performance i.e. producing smaller rms errors, as expected.

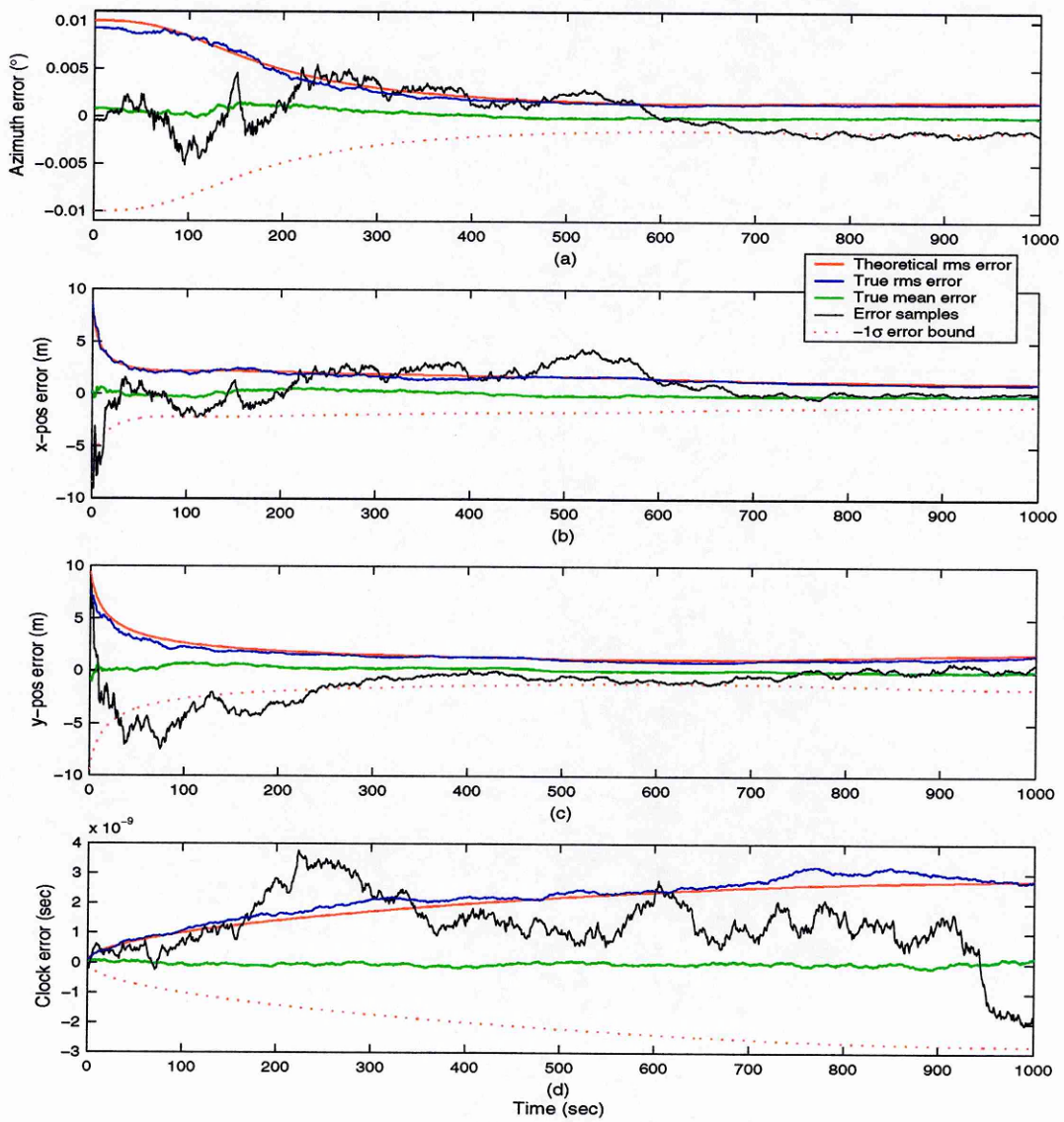


Figure 4-17 The performance of the integrated INS/GPS system with high accuracy sensors

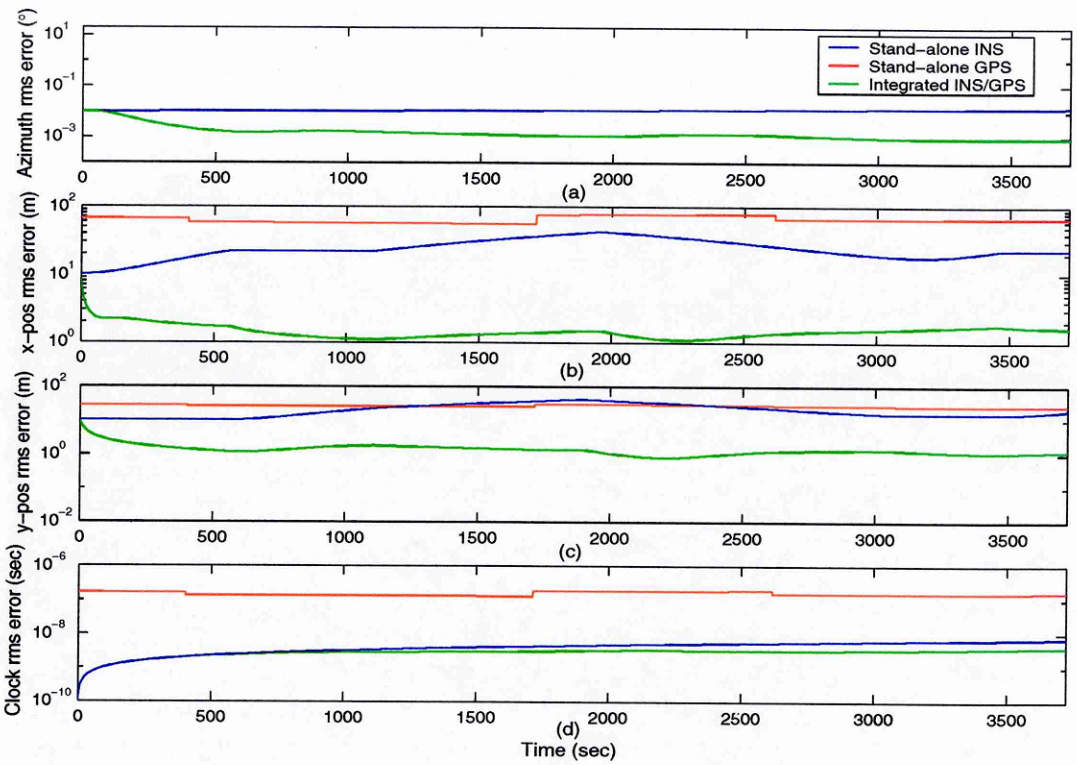


Figure 4-18 Comparison of the stand-alone and integrated system theoretical rms errors

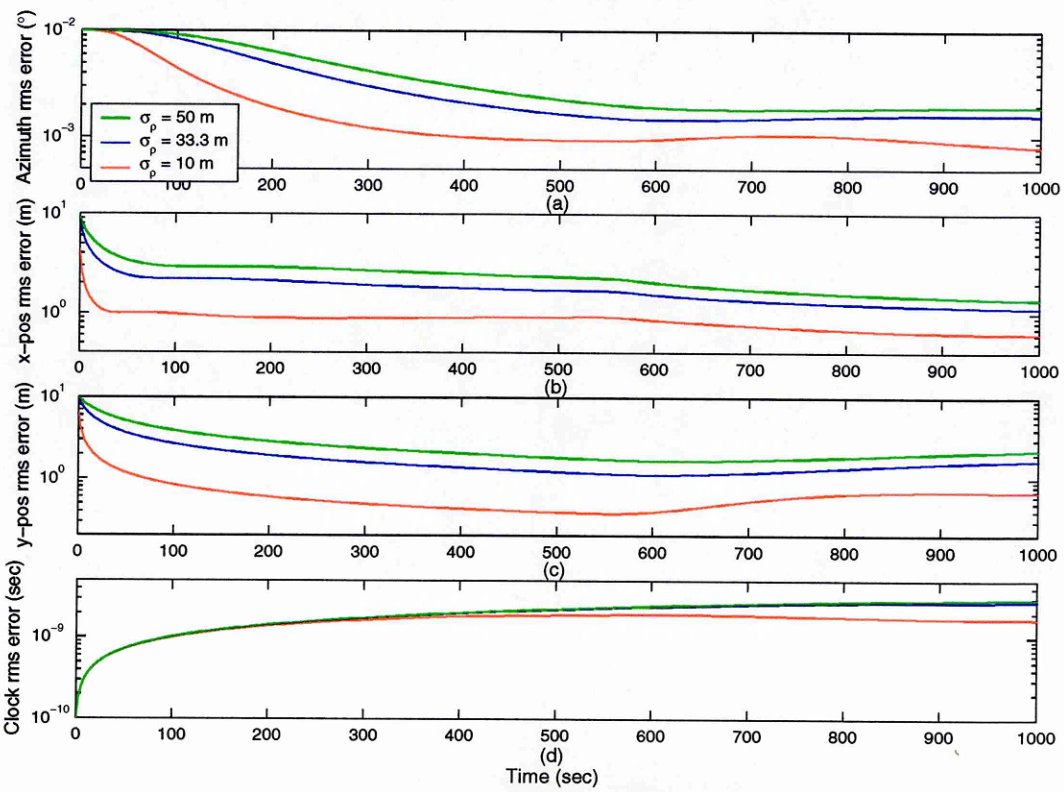


Figure 4-19 The effects of the measurement accuracy on the integrated INS/GPS system performance

Figure 4-20 illustrates the effects of the sensor and initial condition accuracy on the integrated system performance and an important characteristic of the Kalman filter, which is the convergence of the steady-state solutions. These results are shown for two levels of accuracy for the sensors and initial conditions. It can be seen that the theoretical rms errors of the integrated INS/GPS system converge to steady-state values, which are determined by the sensor accuracy and independent of the initial conditions. The convergence of the rms errors represents the ability of the Kalman filter to achieve the superior performance of an integrated navigation system, while retaining the desirable characteristics of all the available sensors and balancing the uncertainties in the system models.

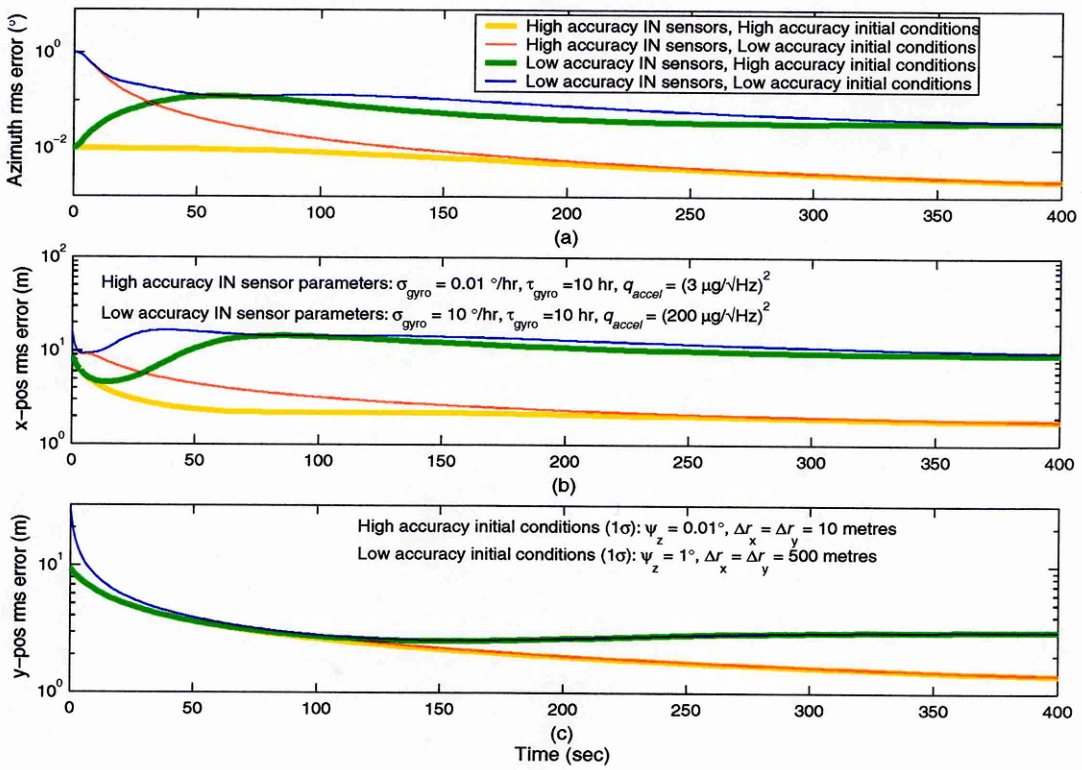


Figure 4-20 The effects of sensor accuracy on the integrated INS/GPS system performance

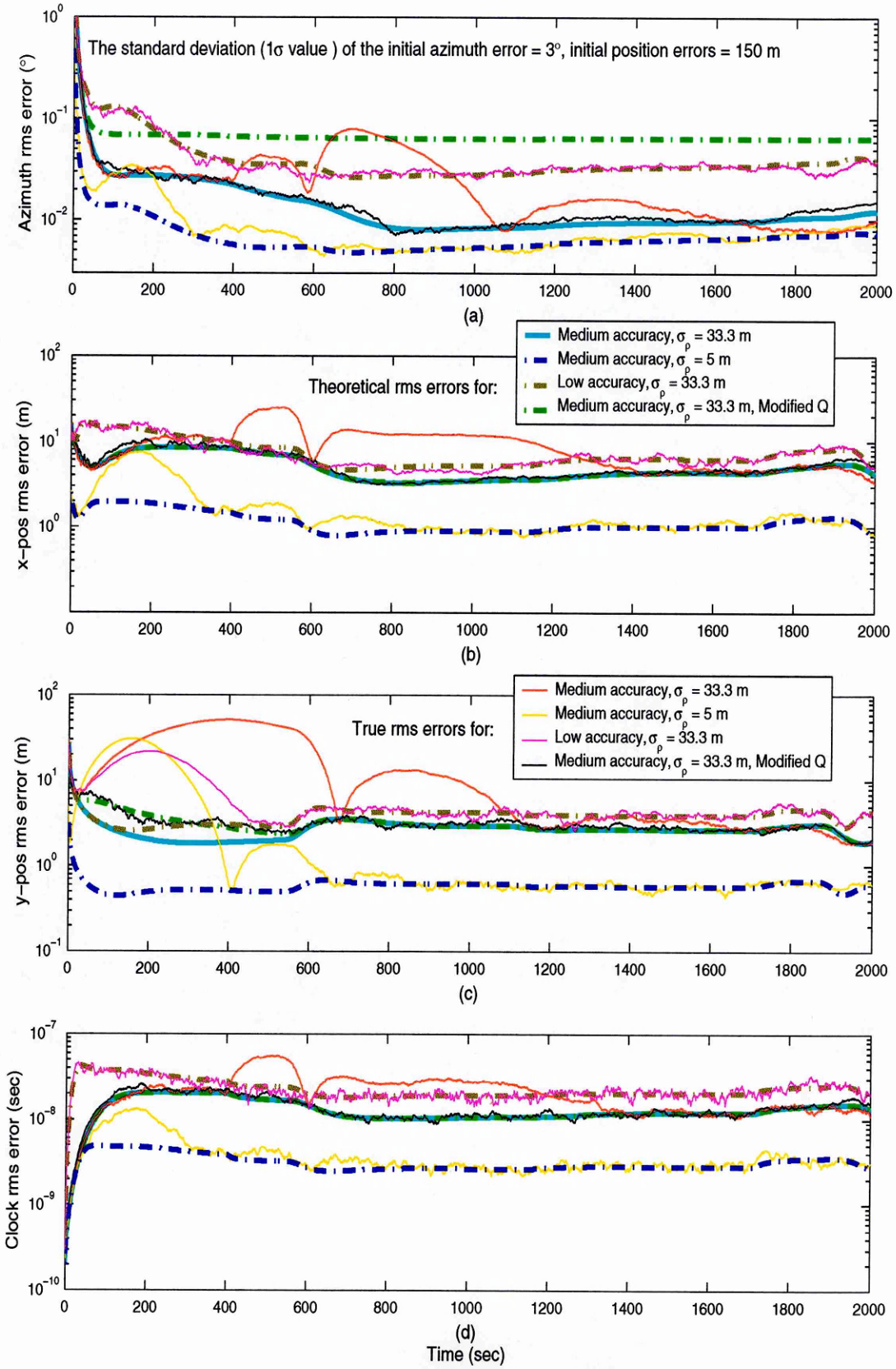


Figure 4-21 The divergence of the integrated INS/GPS Kalman filter

Figure 4-21 shows the theoretical and true rms errors for the simulation of the integrated INS/GPS system with both medium and low accuracy sensors. The standard deviation (1σ value) of the initial azimuth, position and pseudorange measurement error (σ_ρ) are 3°, 150m and 33.3m respectively. The purpose of this analysis is to illustrate the capability of the Kalman filter to derive optimal navigation solutions by weighting the uncertainties in the system, the system dynamics and available measurements. Although, the terms "medium" and "low" accuracy refer to the sensors of integrated systems, based on the parameter given in Table 4-3, the levels of the sensor accuracy (and the terms medium and low) should be viewed as qualitative rather than absolute measures.

By comparing the theoretical (═) and true (—) rms errors of the medium accuracy system with the standard deviation of the pseudorange measurement errors $\sigma_\rho = 33.3\text{m}$, it can be seen that the Kalman filter diverges for a limited period. This divergence occurs due to the invalidity of the error model, which is caused by accumulation of the azimuth error and the divergence between the reference trajectory (used in the linearised model) and the true trajectory, as discussed in Section 4.6.1. The period of divergence is limited because the errors in the reference trajectory (which is constantly updated by the EKF), and the azimuth errors are continually reduced by the filtering process. The divergence can be regarded as the transient response of the Kalman filter to external disturbances. This is an advantage of the EKF over other Kalman filter implementations ("*linearised*" Kalman filter[11]) which use predetermined reference trajectory.

Figure 4-21 also shows the theoretical (▪▪▪) and true (—) rms errors of the medium accuracy system when σ_ρ is reduced from 33.3m to 5m. The theoretical rms error for σ_ρ = 5m is smaller than the theoretical rms error for σ_ρ = 33.3m as expected. Moreover, the divergence period is also reduced for σ_ρ = 5m. This is because the EKF applies more weight to the reliable measurements in the filtering process, eliminating the divergence. On the other hand, by inspecting the theoretical (▪▪▪) and true (—) rms errors of the low accuracy system, the effect of divergence is also less severe in comparison with the medium accuracy system, for the values of σ_ρ = 33.3m. This is

because the EKF relies less on the INS outputs and more on the GPS measurements, resulting in faster correction of the INS errors.

It was found by observation that the filter divergence can be reduced by replacing the PSD of the linear velocity sensor errors ($w_{\Delta v_x}$ and $w_{\Delta v_y}$ in Eq. 4-69 and Eq. 4-70) with arbitrary large values at the beginning of the filter cycles. The modified system "tricks" the Kalman filter into believing that the INS is less reliable so that more weights are given to the GPS measurements. The elimination of the divergence can be seen by comparing the theoretical (▪▪▪) and true (—) rms errors of the modified system. The results obtained in Figure 4-21 are derived by multiplying the PSD by 10^5 during the first two steps of the filter cycles. This value of design parameter was selected to achieve the best trade-off between the increase in the rms errors and the control of the divergence. This example demonstrates that to apply the Kalman filter algorithms to solve real-world problems requires both theoretical and practical considerations involving extensive system simulation and analysis, ranging from comprehensive system modelling to filter tuning.

In this chapter, the strapdown INS error model and the general pseudorange measurement model have been developed based on the simplified models of INS and GPS. A method of integrating INS and GPS using the indirect feedback filter configuration have been covered. The characteristics of each system have been analysed. It has been shown that the performance of the integrated INS/GPS systems is superior to either system operating independently. The capability of the extended Kalman filter to optimally derive the state estimate by balancing the uncertainties in the system, the system dynamics and measurements has been demonstrated. Examples of the filter divergence and a practical solution have also been developed. The understanding of these aspects of systems integration will be incorporated in the relative navigation algorithms described in the next chapter.

CHAPTER 5 RELATIVE NAVIGATION

5.1 INTRODUCTION

Communication is a significant component in military operations. Owing to the nature of these operations, military communication systems are required to be highly secure, survivable and enable identification of friendly forces. Based on advances in digital technology, Link-16 is a highly secure and jam-resistant integrated communications, navigation and identification (ICNI) system currently used by the joint forces (army, navy and air force) of the North Atlantic Treaty Organisation (NATO) countries. The basic building block of Link-16 is a communication system architecture known as Joint Tactical Information Distribution System (JTIDS). At the system level, Link-16 as a system consists of platforms (aircraft, ships etc.), human machine interface (HMI), Tactical Digital Information Link J (TADIL-J) and JTIDS terminals.

This section presents an overview of JTIDS and its use in relative navigation (RELNAV). The potential of RELNAV for civil aviation applications and its principle of operation are also described. The overall system model for RELNAV is based on the simplified model of strapdown INS, described in Section 4.3. The similarities between the integrated INS/GPS system and RELNAV are outlined and, the performance of *centralised* RELNAV is analysed.

5.2 JOINT TACTICAL INFORMATION DISTRIBUTION SYSTEM

JTIDS is the communication system architecture based on Time Division Multiple Access (TDMA), which is similar in concept but not compatible with VDL Mode 4. JTIDS divides time into 12.8-minute Epochs, which are further subdivided into 98,304 7.8125-millisecond (msec) time slots. Because JTIDS terminals transmit data by broadcasting (omni-directional), all members of a JTIDS network are assigned blocks of logically contiguous time slots (usually within a 12-second frame, which repeat 64 times per Epoch), within which messages must be transmitted to prevent unintentional interference.

JTIDS employs 51 frequencies, separated by 3 MHz intervals, within the range 960-1215 MHz (excluding two 40 MHz bands at 1030 MHz and 1090 MHz, which are reserved for Mode S transponders). Each message pulse may be transmitted on any one of the 51 frequencies allocated to JTIDS. The jumping of frequency for each data pulse is pre-determined in a pseudo-random manner called frequency hopping. Theoretically, frequency hopping allows up to 128 different networks to operate simultaneously in the same geographical area by sharing the same time slots and frequencies, known as a *stacked net.* The stacked net concept of JTIDS affords the capability to relay information between members from different nets which may be outside the JTIDS operational range. Figure 5-1 illustrates the JTIDS network and the stacked net operational concept.

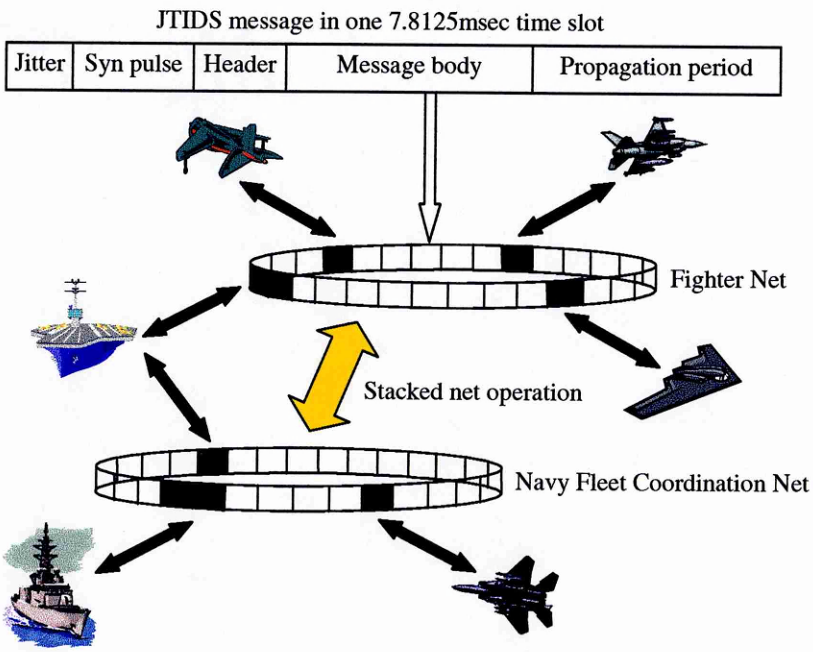


Figure 5-1 JTIDS network and stacked net operational concept

One JTIDS time slot is divided into a number of basic elements, as shown in Figure 5-1. The jitter is a pseudo-random start-of-slot time intended to reduce the susceptibility to enemy jamming. The synchronisation pulse (the Syn pulse in Figure 5-1) allows JTIDS terminals to detect arrival of a message and, more importantly, to perform the relative navigation calculations. The header contains message control information such as error correction coding instructions. The propagation period is included to allow for the propagation delay of transmitted signals.

RELNAV is a feature of JTIDS, which allows network members to perform dual-grid navigation in either the *relative grid* or the *absolute grid* (geodetic latitude longitude). This capability is facilitated by two elements of JTIDS: the precise time synchronisation and the ability to exchange data between network members.

One difficulty, which the original JTIDS RELNAV concept attempted to overcome, was the problem of *data registration*, where the same target appeared to be in different locations due to discrepancies in reported positions. JTIDS RELNAV solves this problem by establishing, in real-time, accurately correlated relative and absolute positions and tracks of all targets (friendly, unknown and hostile) in an operational area. Because the relative positions and tracks are determined very accurately without having to refer to external sources outside of the navigation community, it is possible for all JTIDS network members to operate purely in a relative grid. However, if any member knows its position in the absolute grid precisely, this information can be distributed to other members in the network allowing them to enhance their knowledge of their absolute positions. The ability to accurately determine relative positions and tracks, correlate position data, distribute positional accuracy and relay information between members constitute Link-16, providing a communication, surveillance and situational awareness system.[4] These assets can be highly valuable if they can be applied to civil air traffic management (ATM) as discussed earlier by Fried[42].

For example, even in the case of a complete failure of GPS, safe aircraft separations could still be maintained if accurate relative positions and tracks could be computed. The ability to correlate position data can also be used to process surveillance information such as ADS-B messages. This provides the motivation for investigation into the application of relative navigation to civil ATM. The JTIDS RELNAV principles are explained in the following section.

5.3 Principles of JTIDS Relative Navigation

Figure 5-2 illustrates the JTIDS RELNAV operational concept. In operation, JTIDS network members determine their relative positions and time offsets from the JTIDS system time through sequential passive time-of-arrival (TOA) measurements with respect to other members.

Usually, all active members periodically transmit Precise Participant Location and Identification (PPLI) messages within their allocated time slots containing their relative and absolute positions and also the corresponding quality of this information. Pseudorange measurements can be determined from the received PPLI messages by subtracting the known transmitting time from the measured TOA. Relative positions and time offsets can then be determined, for example, by solving the pseudorange measurement model based on the least-square estimation in Section 4.4.2 and 4.4.3. To enable members to keep track of their positions in the relative grid, some form of dead reckoning (DR) navigation such as INS is commonly used between pseudorange measurement updates.

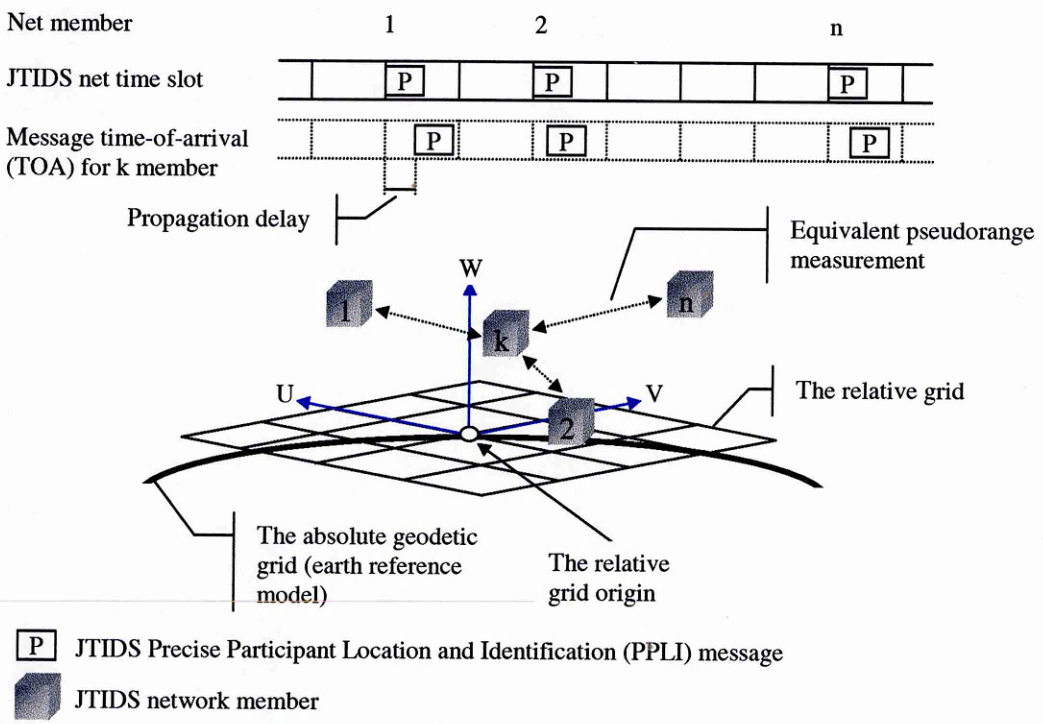
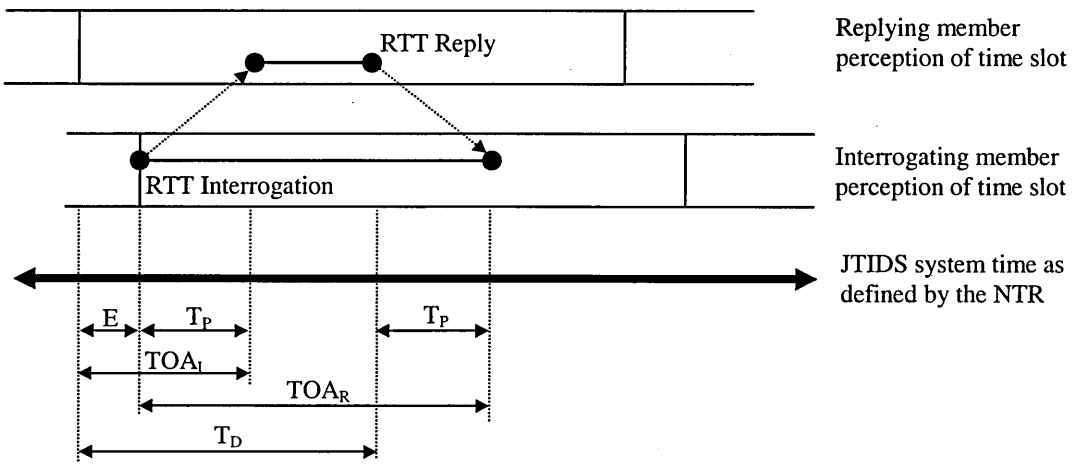


Figure 5-2 JTIDS relative navigation (RELNAV) operational concept

In JTIDS RELNAV, relative positions are determined with respect to a reference frame which is common to all members, known as the relative grid. The relative grid is defined as a Cartesian coordinate system tangential to the surface of the earth reference model at the grid origin which is usually a fixed coordinate close to the operational area. Typically, the relative grid is established by one member designated as the *Navigation Controller* (NC). The orientation of the relative grid is usually defined with respect to the orientation of the NC's attitude reference system such as the

computer frame of a strapdown INS. Subsequently, the relative grid will rotate owing to INS misalignment errors. Moreover, it will be shown in Section 5.4.1 that the relative grid origin will also translate as a function of INS sensor errors of the NC. Consequently, the member of a network with the highest accuracy INS is usually selected as the NC, in order to minimise the errors associated with the relative grid. Although, the INS of the NC constantly drifts, the reported relative positions are always assumed to be perfect (i.e. without errors). This assumption is critical and fundamental to the RELNAV operation.



E Interrogating member timing error
T_D Arbitrary time delay
T_P Propagation time
TOA_I Time-of-arrival of the RTT interrogation
TOA_R Time-of-arrival of the RTT reply

Figure 5-3 Active synchronization by Round Trip Timing interrogation[43]

The JTIDS system time is usually maintained by one member designated as the *Net Time Reference* (NTR). The reported time of the NTR is always assumed to be perfect by definition. In addition to the basic passive operation based on pseudorange observations, each member can also synchronise the time actively by means of the *Round Trip Timing* (RTT) protocol, illustrated in Figure 5-3. Members wishing to correct their clocks transmit the RTT interrogation messages to the sources with higher *time quality* (TQ) and wait for the RTT reply messages. From Figure 5-3, it can be shown that:

$$T_P = TOA_I - E \qquad \text{Eq. 5-1}$$

$$T_P = TOA_R + E - T_D \qquad \text{Eq. 5-2}$$

The time offset, E, from the source's clock is given by:

$$E = (TOA_I - TOA_R + T_D)/2 \qquad \text{Eq. 5-3}$$

If the source is the NTR, E is the time offset from JTIDS system time.

In practice, a member receives PPLI messages, which are transmitted from different sources, containing navigation data with different accuracy levels. It has been shown that arbitrary uses of pseudorange measurements can give rise to unstable navigation solutions leading to an unstable navigation community[46,47]. It has also been proven analytically that stable solutions can be obtained based on the *fixed-rank hierarchy*, provided certain assumptions are satisfied[45]. The fixed-rank hierarchy is a *source selection* process where a fixed-rank is assigned to each member which is not changed during a mission. Each member is restricted to pseudorange observations from members of superior ranks. The source selection function is a critical element of the successful operation of JTIDS RELNAV.

The source selection function in JTIDS RELNAV is based on a *covariance-based hierarchy*. The error covariance of the estimated *position* and *time* are converted to quality indices since practical limitations preclude the transmission of complete error covariance matrices. The position (PQ) and time (TQ) qualities range from 0 to 15 where the higher values represent the higher accuracy levels. These quality indices are included in every transmitted PPLI message and can be converted back to a full error covariance matrix by the receiving member. However, the reconstructed error covariance matrix differs from the actual error covariance matrix and this leads to sub-optimality. By definition, the NC always reports PQ values of 15, while the TQ values of the NTR is always 15. JTIDS RELNAV source selection function first screens all PPLI messages and rejects any messages of inferior quality. The remaining PPLI messages are ranked as a function of GDOP such that the optimal set of observations is selected. Otherwise, practical limitations would preclude the processing of all messages.

Another complexity, which arises in RELNAV, is how to determine the optimal navigation solutions taking into account the sequential nature of the observations, the dynamics of the platforms, the navigation system and measurement errors. JTIDS RELNAV uses an extended Kalman filter (EKF) as the principal data processing algorithm because it affords an optimal and versatile solution, as described in

Section 4.7. The JTIDS RELNAV system error model includes some 20 states, which describe the errors of the NC (the relative grid motion), the own platform INS and clock errors. The actual description of the JTIDS RELNAV system error model is classified and is not available in the public domain. Nevertheless, an overall system model for RELNAV has been developed in Section 5.4 based on the fundamental principles used in JTIDS RELNAV. It will be shown that this overall system model exhibits similar error characteristics to JTIDS RELNAV and provides the basis for the subsequent analysis of the RELNAV performance.

The principles of JTIDS RELNAV have been described. It can be seen that RELNAV is analogous to mechanising every member's INS to navigate in the relative grid established by the NC where *in-flight alignment* of each member's estimated navigation frame to the NC's estimated navigation frame is achieved using relative pseudorange observations. If all members are equipped with a strapdown INS, RELNAV tries to align their computer frames to the NC's computer frame. Similarly, time synchronisation is made with respect to the NTR's clock. Although the relative grid and the system time drift as functions of the sensor errors of the NC and the NTR, these errors are not observable by other members in the navigation community including, the NC and the NTR. Hence, the reported data of the NC and the NTR are always assumed to be perfect by definition; otherwise divergence would occur where the RELNAV EKF attempts to correct system states which are not observable. Highly accurate RELNAV capability can always be maintained while accurate relative pseudorange observations are available. Finally, an integrated INS/GPS system can be considered as a form of RELNAV where all GPS satellites act as both the NCs and the NTRs. GPS pseudorange measurements are used to perform in-flight alignment of the INS estimated navigation frame to the GPS reference frame and the receiver clock is synchronised to GPS system time. GPS system errors are not observable by any member of the integrated INS/GPS system.

5.4 RELNAV NAVIGATION ALGORITHMS

RELNAV differs from an integrated INS/GPS system in that the RELNAV reference frame (the relative grid) rotates and translates as a function of the NC's sensor

errors. Moreover, the RELNAV system time drifts as a function of the NTR's clock phase errors. These system errors must be modelled explicitly in RELNAV. The equations of motion for every member in the RELNAV navigation community must be developed with respect to this stochastically moving relative grid. RELNAV is also hierarchy-based where the sources have position and time errors of different qualities and these errors must be taken into consideration in the measurement model. This section develops the RELNAV algorithms suitable for EKF based on JTIDS RELNAV principles and the simplified models of strapdown INS, GPS and airspace described in Chapter 4. It will be assumed that every member uses a strapdown INS for computing relative positions. Figure 5-4 illustrates coordinate systems used in RELNAV.

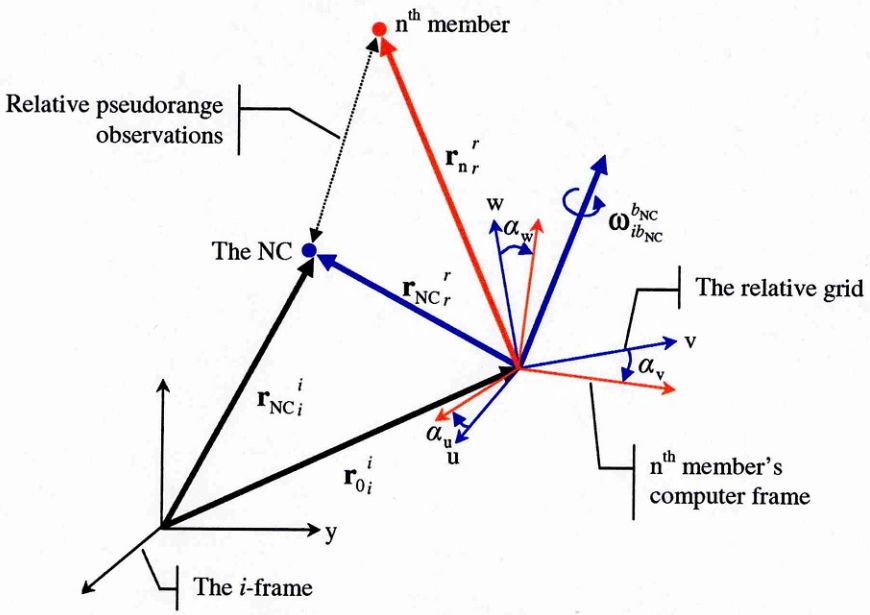


Figure 5-4 Coordinate systems used in RELNAV

5.4.1 Development of the relative grid equations of motion

The rotation of the relative grid

For simplicity, it will be assumed that the NC is mechanised to navigate in the true navigation frame (the *i*-frame); hence, the relative grid (the *r*-frame) is the *i*-frame as perceived by the NC. The relative grid will rotate at the same rate as the computer frame (the *c*-frame) of the NC's strapdown INS. Based on Eq. 4-28, the rotation of the relative grid can be modelled directly as follows:

$$\boldsymbol{\omega}_{ir}^{r} = \dot{\boldsymbol{\psi}}_{\mathrm{NC}} = \mathbf{C}_{b_{\mathrm{NC}}}^{r} \Delta\boldsymbol{\omega}_{ib_{\mathrm{NC}}}^{b_{\mathrm{NC}}} \qquad \text{Eq. 5-4}$$

where r and NC denote the relative grid and the Navigation Controller respectively. In effect, $\boldsymbol{\psi}_{\mathrm{NC}}$ is the misalignment of the relative grid from the i-frame. It should be noted that the r-frame is the c-frame of the NC's strapdown INS.

The translation of the relative grid

From Figure 5-4, the relationship between the relative grid origin and the NC position vector is defined as follows:

$$\mathbf{r}_{\mathrm{NC}i}^{\ i} = \mathbf{r}_{0i}^{\ i} + \mathbf{C}_{r}^{i}\mathbf{r}_{\mathrm{NC}r}^{\ r} \qquad \text{Eq. 5-5}$$

where $\mathbf{r}_{\mathrm{NC}}$ and $\mathbf{r}_{0}$ are the NC and relative grid origin position vector respectively. Differentiating Eq. 5-5 with respect to the i-frame and expressing all terms in the i-frame, taking into account the Coriolis effect, yields:

$$\dot{\mathbf{r}}_{\mathrm{NC}i}^{\ i} = \dot{\mathbf{r}}_{0i}^{\ i} + \mathbf{C}_{r}^{i}(\dot{\mathbf{r}}_{\mathrm{NC}r}^{\ r} + [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{\mathrm{NC}r}^{\ r}) \qquad \text{Eq. 5-6}$$

The motion of the relative grid origin with respect to the i-frame expressed in the i-frame can be obtained by rearranging Eq. 5-6 as follows:

$$\dot{\mathbf{r}}_{0i}^{\ i} = \dot{\mathbf{r}}_{\mathrm{NC}i}^{\ i} - \mathbf{C}_{r}^{i}(\dot{\mathbf{r}}_{\mathrm{NC}r}^{\ r} + [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{\mathrm{NC}r}^{\ r}) \qquad \text{Eq. 5-7}$$

where $\dot{\mathbf{r}}_{\mathrm{NC}i}^{\ i}$ is the true motion of the NC with respect to the i-frame and $\dot{\mathbf{r}}_{\mathrm{NC}r}^{\ r}$ is the motion of the NC as indicated by its strapdown INS. Therefore, from Eq. 4-18 and Eq. 4-20, $\dot{\mathbf{r}}_{\mathrm{NC}i}^{\ i}$ and $\dot{\mathbf{r}}_{\mathrm{NC}r}^{\ r}$ can be expressed as follows:

$$\dot{\mathbf{r}}_{\mathrm{NC}i}^{\ i} = \mathbf{C}_{b_{\mathrm{NC}}}^{i}\mathbf{v}_{i}^{b_{\mathrm{NC}}} \qquad \text{Eq. 5-8}$$

$$\dot{\mathbf{r}}_{\mathrm{NC}r}^{\ r} = \hat{\mathbf{C}}_{b_{\mathrm{NC}}}^{i}\hat{\mathbf{v}}_{i}^{b_{\mathrm{NC}}} \qquad \text{Eq. 5-9}$$

where $\mathbf{v}_{i}^{b_{\mathrm{NC}}}$ is the true velocity of the NC's body frame (the b-frame) with respect to the i-frame. From Eq. 4-23, the transformation between the i-frame and the r-frame can be expressed as follows:

$$\mathbf{C}_{i}^{r} \cong \mathbf{I} - [\boldsymbol{\psi}_{\mathrm{NC}}\times] \qquad \text{Eq. 5-10}$$

Substituting Eq. 5-8 and Eq. 5-9 in Eq. 5-7, noting that $\Delta\mathbf{v}_i^{b_{NC}} = \mathbf{v}_i^{b_{NC}} - \hat{\mathbf{v}}_i^{b_{NC}}$ and (from Eq. 4-22) that $\hat{\mathbf{C}}_{b_{NC}}^{i} = \mathbf{C}_{b_{NC}}^{c} = \mathbf{C}_{b_{NC}}^{r} = \mathbf{C}_i^r \mathbf{C}_{b_{NC}}^{i}$, yields:

$$\begin{aligned}
\dot{\mathbf{r}}_{0i}^{\ i} &= \mathbf{C}_{b_{NC}}^{i}\mathbf{v}_i^{b_{NC}} - \mathbf{C}_r^i(\hat{\mathbf{C}}_{b_{NC}}^{i}\hat{\mathbf{v}}_i^{b_{NC}} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{NC_r}^{\ r}) \\
&= \mathbf{C}_{b_{NC}}^{i}\mathbf{v}_i^{b_{NC}} - \mathbf{C}_r^i(\mathbf{C}_i^r\mathbf{C}_{b_{NC}}^{i}\hat{\mathbf{v}}_i^{b_{NC}} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{NC_r}^{\ r}) \\
&= \mathbf{C}_{b_{NC}}^{i}(\mathbf{v}_i^{b_{NC}} - \hat{\mathbf{v}}_i^{b_{NC}}) - \mathbf{C}_r^i[\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{NC_r}^{\ r} \\
&= \mathbf{C}_{b_{NC}}^{i}\Delta\mathbf{v}_i^{b_{NC}} - \mathbf{C}_r^i[\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{NC_r}^{\ r}
\end{aligned} \qquad \text{Eq. 5-11}$$

The motion of the relative grid origin can be determined as a function of the strapdown INS sensor errors of the NC by substituting Eq. 5-4 in Eq. 5-11 as follows:

$$\begin{aligned}
\dot{\mathbf{r}}_{0i}^{\ i} &= \mathbf{C}_{b_{NC}}^{i}\Delta\mathbf{v}_i^{b_{NC}} - \mathbf{C}_r^i\mathbf{C}_{b_{NC}}^{r}[\Delta\boldsymbol{\omega}_{ib_{NC}}^{b_{NC}}\times]\mathbf{r}_{NC_r}^{\ r} \\
&= \mathbf{C}_{b_{NC}}^{i}\Delta\mathbf{v}_i^{b_{NC}} - \mathbf{C}_{b_{NC}}^{i}[\Delta\boldsymbol{\omega}_{ib_{NC}}^{b_{NC}}\times]\mathbf{r}_{NC_r}^{\ r}
\end{aligned} \qquad \text{Eq. 5-12}$$

From an element-by-element comparison, Eq. 5-12 can be written in vector form as follows:

$$\dot{\mathbf{r}}_{0i}^{\ i} = \mathbf{C}_{b_{NC}}^{i}(\Delta\mathbf{v}_i^{b_{NC}} - \Delta\boldsymbol{\omega}_{ib_{NC}}^{b_{NC}} \times \mathbf{r}_{NC_r}^{\ r}) \qquad \text{Eq. 5-13}$$

It should be noted that, by applying Eq. 5-10 for small values of $\boldsymbol{\psi}_{NC}$ and $\dot{\mathbf{r}}_{0i}^{\ i}$, the motion of the relative grid origin expressed in the *r*-frame $\dot{\mathbf{r}}_{0i}^{\ r}$ can be approximated as follows:

$$\begin{aligned}
\dot{\mathbf{r}}_{0i}^{\ r} &= \mathbf{C}_i^r\dot{\mathbf{r}}_{0i}^{\ i} \\
&\cong (\mathbf{I} - [\boldsymbol{\psi}_{NC}\times])\dot{\mathbf{r}}_{0i}^{\ i} \\
&\cong \dot{\mathbf{r}}_{0i}^{\ i}
\end{aligned} \qquad \text{Eq. 5-14}$$

Eq. 5-14 is required in the derivation of the equations of motion in the relative grid.

5.4.2 Development of the equations of motion in the relative grid

This section develops the equations of motion in the relative grid for the n^{th} observing member (denoted by the subscript $_n$).

Relative attitude error equations

Let $\boldsymbol{\alpha}_n$ be the misalignment angle vector of the computer frame of the nth observing member from the relative grid, such that:

$$\mathbf{C}_{c_n}^{r} = \mathbf{C}_{i}^{r}\mathbf{C}_{c_n}^{i} \cong \mathbf{I} + [\boldsymbol{\alpha}_n \times] \qquad \text{Eq. 5-15}$$

From Eq. 4-23 and Eq. 4-5, the transformation between the *i*-frame and the *c*-frame of the nth member can be expressed as follows:

$$\mathbf{C}_{c_n}^{i} \cong \mathbf{I} + [\boldsymbol{\psi}_n \times] \qquad \text{Eq. 5-16}$$

Substituting Eq. 5-10 and Eq. 5-16 in Eq. 5-15 yields:

$$\mathbf{I} + [\boldsymbol{\alpha}_n \times] \cong (\mathbf{I} - [\boldsymbol{\psi}_{NC} \times])(\mathbf{I} + [\boldsymbol{\psi}_n \times]) \qquad \text{Eq. 5-17}$$

From an element-by-element comparison and (ignoring the products of error terms) Eq. 5-17 can be written in vector form as follows:

$$\boldsymbol{\alpha}_n = \boldsymbol{\psi}_n - \boldsymbol{\psi}_{NC} \qquad \text{Eq. 5-18}$$

Differentiating Eq. 5-18 with respect to time yields:

$$\dot{\boldsymbol{\alpha}}_n = \dot{\boldsymbol{\psi}}_n - \dot{\boldsymbol{\psi}}_{NC} \qquad \text{Eq. 5-19}$$

The misalignment of the *c*-frame of the INS of the nth member can be described as a function of the gyroscope errors by applying Eq. 4-28 and Eq. 5-4 to Eq. 5-19 as follows:

$$\dot{\boldsymbol{\alpha}}_n = \mathbf{C}_{b_n}^{r}\Delta\boldsymbol{\omega}_{ib_n}^{b_n} - \mathbf{C}_{b_{NC}}^{r}\Delta\boldsymbol{\omega}_{ib_{NC}}^{b_{NC}} \qquad \text{Eq. 5-20}$$

Relative position error equations

From Figure 5-4, the relationship between the position of the nth member in the relative grid and the *i*-frame can be described as follows:

$$\mathbf{r}_{n_i}^{\ i} = \mathbf{r}_{0_i}^{\ i} + \mathbf{C}_{r}^{i}\mathbf{r}_{n_r}^{\ r} \qquad \text{Eq. 5-21}$$

Eq. 5-21 is the most fundamental equation in the derivation of the RELNAV algorithms. The time differentiation of Eq. 5-21 observed from the *i*-frame and expressed in the *r*-frame, taking into account the Coriolis effect, is determined as follows:

$$\dot{\mathbf{r}}_{ni}^{\ r} = \dot{\mathbf{r}}_{0i}^{\ r} + (\dot{\mathbf{r}}_{nr}^{\ r} + [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r}) \qquad \text{Eq. 5-22}$$

where $\dot{\mathbf{r}}_{ni}^{\ r} = \mathbf{C}_i^r \dot{\mathbf{r}}_{ni}^{\ i}$, and $\dot{\mathbf{r}}_{ni}^{\ i} = \mathbf{C}_{b_n}^i \mathbf{v}_i^{b_n}$ is the true motion of the nth member with respect to and expressed in the *i*-frame. Eq. 5-22 can be rearranged to obtain the equation of motion in the *r*-frame as follows:

$$\dot{\mathbf{r}}_{nr}^{\ r} = \dot{\mathbf{r}}_{ni}^{\ r} - \dot{\mathbf{r}}_{0i}^{\ r} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r} \qquad \text{Eq. 5-23}$$

However, the estimated position of the nth member in the relative grid is determined from the INS, i.e.:

$$\hat{\dot{\mathbf{r}}}_{nr}^{\ r} = \hat{\mathbf{C}}_{b_n}^r \hat{\mathbf{v}}_i^{b_n} \qquad \text{Eq. 5-24}$$

Subtracting Eq. 5-24 from Eq. 5-23 and substituting $\dot{\mathbf{r}}_{ni}^{\ i} = \mathbf{C}_{b_n}^i \mathbf{v}_i^{b_n}$ yields:

$$\begin{aligned}
\Delta\dot{\mathbf{r}}_{nr}^{\ r} &= \dot{\mathbf{r}}_{nr}^{\ r} - \hat{\dot{\mathbf{r}}}_{nr}^{\ r} \\
&= \mathbf{C}_i^r \dot{\mathbf{r}}_{ni}^{\ i} - \dot{\mathbf{r}}_{0i}^{\ r} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r} - \hat{\mathbf{C}}_{b_n}^r \hat{\mathbf{v}}_i^{b_n} \\
&= \mathbf{C}_i^r \mathbf{C}_{b_n}^i \mathbf{v}_i^{b_n} - \dot{\mathbf{r}}_{0i}^{\ r} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r} - \hat{\mathbf{C}}_{b_n}^r \hat{\mathbf{v}}_i^{b_n} \\
&= \mathbf{C}_{b_n}^r \mathbf{v}_i^{b_n} - \dot{\mathbf{r}}_{0i}^{\ r} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r} - \hat{\mathbf{C}}_{b_n}^r \hat{\mathbf{v}}_i^{b_n}
\end{aligned} \qquad \text{Eq. 5-25}$$

Substituting Eq. 5-14 and Eq. 5-15 in Eq. 5-25, and noting that $\hat{\mathbf{C}}_{b_n}^r = \mathbf{C}_r^{c_n}\mathbf{C}_{b_n}^r$, Eq. 5-25 can be expressed as follows:

$$\begin{aligned}
\Delta\dot{\mathbf{r}}_{nr}^{\ r} &= \mathbf{C}_{b_n}^r \mathbf{v}_i^{b_n} - \hat{\mathbf{C}}_{b_n}^r \hat{\mathbf{v}}_i^{b_n} - \dot{\mathbf{r}}_{0i}^{\ i} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r} \\
&= \mathbf{C}_{b_n}^r \mathbf{v}_i^{b_n} - \mathbf{C}_r^{c_n}\mathbf{C}_{b_n}^r \hat{\mathbf{v}}_i^{b_n} - \dot{\mathbf{r}}_{0i}^{\ i} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r} \\
&= \mathbf{C}_{b_n}^r \mathbf{v}_i^{b_n} - (\mathbf{I} - [\boldsymbol{\alpha}_n\times])\mathbf{C}_{b_n}^r \hat{\mathbf{v}}_i^{b_n} - \dot{\mathbf{r}}_{0i}^{\ i} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r}
\end{aligned} \qquad \text{Eq. 5-26}$$

Substituting $\hat{\mathbf{v}}_i^{b_n} = \mathbf{v}_i^{b_n} - \Delta\mathbf{v}_i^{b_n}$ in Eq. 5-26 and ignoring the products of the error terms yields:

$$\begin{aligned}
\Delta\dot{\mathbf{r}}_{nr}^{\ r} &= \mathbf{C}_{b_n}^r \mathbf{v}_i^{b_n} - (\mathbf{I} - [\boldsymbol{\alpha}_n\times])\mathbf{C}_{b_n}^r (\mathbf{v}_i^{b_n} - \Delta\mathbf{v}_i^{b_n}) - \dot{\mathbf{r}}_{0i}^{\ i} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r} \\
&= \mathbf{C}_{b_n}^r \Delta\mathbf{v}_i^{b_n} + [\boldsymbol{\alpha}_n\times]\mathbf{C}_{b_n}^r \mathbf{v}_i^{b_n} - \dot{\mathbf{r}}_{0i}^{\ i} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r} \\
&= \mathbf{C}_{b_n}^r \Delta\mathbf{v}_i^{b_n} + [\boldsymbol{\alpha}_n\times]\mathbf{v}_i^{r} - \dot{\mathbf{r}}_{0i}^{\ i} - [\boldsymbol{\omega}_{ir}^{r}\times]\mathbf{r}_{nr}^{\ r}
\end{aligned} \qquad \text{Eq. 5-27}$$

From an element by element comparison, Eq. 5-27 can be written in vector form as follows:

$$\Delta\dot{\mathbf{r}}_{\mathrm{n}r}^{\ r} = \mathbf{C}_{b_{\mathrm{n}}}^{r}\,\Delta\mathbf{v}_{i}^{b_{\mathrm{n}}} - \mathbf{v}_{i}^{r}\times\boldsymbol{\alpha}_{\mathrm{n}} - \dot{\mathbf{r}}_{0i}^{\ i} - \boldsymbol{\omega}_{ir}^{r}\times\mathbf{r}_{\mathrm{n}r}^{\ r} \qquad \text{Eq. 5-28}$$

Timing error equations

Let t_{abs} denote an arbitrary absolute reference time. The relationship between t_{abs} and the RELNAV system time, defined by the NTR, can be described as follows:

$$t_{\mathrm{abs}} = \hat{t}_{\mathrm{NTR}} + \delta t_{\mathrm{NTR}} \qquad \text{Eq. 5-29}$$

where $\hat{t}_{\mathrm{NTR}}$ and δt_{NTR} denote the RELNAV system time (the NTR clock reading) and the NTR clock phase bias from the absolute reference time. Similarly, the relationship between t_{abs} and the clock of the n^{th} member can be expressed as follows:

$$t_{\mathrm{abs}} = \hat{t}_{\mathrm{n}} + \delta t_{\mathrm{n}} \qquad \text{Eq. 5-30}$$

where $\hat{t}_{\mathrm{n}}$ and δt_{n} denote the clock reading and the clock phase error of the n^{th} member. The time offset of the n^{th} member from the RELNAV system time can be determined as a function of the NTR and its own clock phase errors by subtracting Eq. 5-30 from Eq. 5-29 as follows:

$$\begin{aligned}\Delta t_{\mathrm{n}} &= \hat{t}_{\mathrm{NTR}} - \hat{t}_{\mathrm{n}} \\ &= \delta t_{\mathrm{n}} - \delta t_{\mathrm{NTR}}\end{aligned} \qquad \text{Eq. 5-31}$$

The time offset can be described as a function of time by differentiating Eq. 5-31 with respect to time as follows:

$$\Delta\dot{t}_{\mathrm{n}} = \delta\dot{t}_{\mathrm{n}} - \delta\dot{t}_{\mathrm{NTR}} \qquad \text{Eq. 5-32}$$

Relative pseudorange measurement model

The generalised pseudorange measurement model developed in Section 4.4.2 can be used directly as the measurement model for RELNAV because it already models the position and timing errors of the source. In order to apply the EKF, a linearised measurement model (Eq. 4-53) is used, which is repeated in scalar form in Eq. 5-33 for convenience.

$$\Delta\rho = \mathbf{h}_{so}\Delta\mathbf{x}_o - \mathbf{h}_{so}\Delta\mathbf{x}_s + v \qquad \text{Eq. 5-33}$$

In this section, the motion of the relative grid and the system time errors have been described as functions of the NC's strapdown INS sensor errors and the NTR's clock bias using Eq. 5-4, Eq. 5-13 and Eq. 5-32. The equations of motion of the n^{th} observing member in the relative grid have been derived in Eq. 5-20 and Eq. 5-28. The linearised measurement model is given by Eq. 5-33. These equations are significant because they represent an overall system model which is suitable for the RELNAV EKF. The comparison between RELNAV and the integrated INS/GPS system can now be made by comparing Eq. 4-28, Eq. 4-31 and Eq. 4-54 to Eq. 5-20, Eq. 5-28 and Eq. 5-33. It can be seen that the differences are that the RELNAV overall system model explicitly includes the stochastic motion of the navigation frame (the relative grid), the system time errors and the uncertainties of the sources of pseudorange observations. If the sensors of the NC and the NTR are perfect, i.e. the *r*-frame is the *i*-frame, the RELNAV equations (Eq. 5-20 and Eq. 5-28) reduce to integrated INS/GPS equations (Eq. 4-28 and Eq. 4-31).

5.5 STATE-SPACE FORMULATION

This section summarises the RELNAV algorithms using the state-space formulation notation described in Section 3.2, based on the same assumptions made in Section 4.5. In addition, for simplicity (but without a loss in generality), it will be assumed that the NC (denoted using the superscripts ^{NC}) also controls the RELNAV system time i.e. the NC is the NTR. Let the respective axes of the *r*-frame be given by the u, v and w axes, and $\mathbf{r}_{0_i}^{\ i} = [U,V,W]^{\mathrm{T}}$ denotes the relative grid origin in component form. The relative grid motion and the system time errors can be described in terms of the NC's sensor errors using the following set of differential equations:

$$\dot{\psi}_{\mathrm{w}}^{\mathrm{NC}} = \Delta\omega_{\mathrm{w}}^{\mathrm{NC}} \qquad \text{(from Eq. 5-4)} \qquad \text{Eq. 5-34}$$

$$\dot{U} = r_{\mathrm{v}}^{\mathrm{NC}}\Delta\omega_{\mathrm{w}}^{\mathrm{NC}} + c_{11}^{\mathrm{NC}}\Delta v_{\mathrm{u}}^{\mathrm{NC}} + c_{12}^{\mathrm{NC}}\Delta v_{\mathrm{v}}^{\mathrm{NC}} \qquad \text{(from Eq. 5-13)} \qquad \text{Eq. 5-35}$$

$$\dot{V} = -r_{\mathrm{u}}^{\mathrm{NC}}\Delta\omega_{\mathrm{w}}^{\mathrm{NC}} + c_{21}^{\mathrm{NC}}\Delta v_{\mathrm{u}}^{\mathrm{NC}} + c_{22}^{\mathrm{NC}}\Delta v_{\mathrm{v}}^{\mathrm{NC}} \qquad \text{(from Eq. 5-13)} \qquad \text{Eq. 5-36}$$

$$\dot{x}_p^{\mathrm{NC}} = x_f^{\mathrm{NC}} + w_p^{\mathrm{NC}} \qquad \text{(from Eq. 4-40)} \qquad \text{Eq. 5-37}$$

$$\dot{x}_f^{\mathrm{NC}} = w_f^{\mathrm{NC}} \qquad \text{(from Eq. 4-41)} \qquad \text{Eq. 5-38}$$

$$\Delta\dot{\omega}_{w}^{NC} = -\beta_{gyro}^{NC}\Delta\omega_{w}^{NC} + w_{\Delta\omega_{w}^{NC}} \quad \text{(from Table 3-1)} \quad \text{Eq. 5-39}$$

$$\Delta\dot{v}_{u}^{NC} = w_{\Delta v_{u}^{NC}} \quad \text{(from Table 3-1)} \quad \text{Eq. 5-40}$$

$$\Delta\dot{v}_{v}^{NC} = w_{\Delta v_{v}^{NC}} \quad \text{(from Table 3-1)} \quad \text{Eq. 5-41}$$

Similarly, the system errors model of the n^{th} member (denoted using the superscript ^{n}) can be described using the following set of differential equations:

$$\dot{\alpha}_{w}^{n} = \Delta\omega_{w}^{n} - \Delta\omega_{w}^{NC} \quad \text{(from Eq. 5-20)} \quad \text{Eq. 5-42}$$

$$\Delta\dot{r}_{u}^{n} = -v_{v}^{n}\alpha_{w}^{n} + c_{11}^{n}\Delta v_{u}^{n} + c_{12}^{n}\Delta v_{v}^{n} - c_{11}^{NC}\Delta v_{u}^{NC} - c_{12}^{NC}\Delta v_{v}^{NC} + (r_{v}^{n} - r_{v}^{NC})\Delta\omega_{w}^{NC} \quad \text{(from Eq. 5-28)} \quad \text{Eq. 5-43}$$

$$\Delta\dot{r}_{v}^{n} = v_{u}^{n}\alpha_{w}^{n} + c_{21}^{n}\Delta v_{u}^{n} + c_{22}^{n}\Delta v_{v}^{n} - c_{21}^{NC}\Delta v_{u}^{NC} - c_{22}^{NC}\Delta v_{v}^{NC} - (r_{u}^{n} - r_{u}^{NC})\Delta\omega_{w}^{NC} \quad \text{(from Eq. 5-28)} \quad \text{Eq. 5-44}$$

$$\Delta\dot{t}_{p}^{n} = (x_{f}^{n} + w_{p}^{n}) - (x_{f}^{NC} + w_{p}^{NC}) \quad \text{(from Eq. 5-32)} \quad \text{Eq. 5-45}$$

$$\dot{x}_{f}^{n} = w_{f}^{n} \quad \text{(from Eq. 4-41)} \quad \text{Eq. 5-46}$$

$$\Delta\dot{\omega}_{w}^{n} = -\beta_{gyro}^{n}\Delta\omega_{w}^{n} + w_{\Delta\omega_{w}^{n}} \quad \text{(from Table 3-1)} \quad \text{Eq. 5-47}$$

$$\Delta\dot{v}_{u}^{n} = w_{\Delta v_{u}^{n}} \quad \text{(from Table 3-1)} \quad \text{Eq. 5-48}$$

$$\Delta\dot{v}_{v}^{n} = w_{\Delta v_{v}^{n}} \quad \text{(from Table 3-1)} \quad \text{Eq. 5-49}$$

The plant model can be written in the standard matrix form $\dot{\mathbf{x}} = \mathbf{F}\mathbf{x} + \mathbf{G}\mathbf{w}$ based on Eq. 5-34 to Eq. 5-49, where the state vector $\mathbf{x} = [\mathbf{x}^{NC}, \mathbf{x}^{1}, \ldots, \mathbf{x}^{n}]^{T}$ and the input noise vector $\mathbf{w} = [\mathbf{w}^{NC}, \mathbf{w}^{1}, \ldots, \mathbf{w}^{n}]^{T}$ are defined as follows:

$$\mathbf{x}^{NC} = [\psi_{w}^{NC} \quad U \quad V \quad x_{p}^{NC} \quad x_{f}^{NC} \quad \Delta\omega_{z}^{NC} \quad \Delta v_{x}^{NC} \quad \Delta v_{y}^{NC}]^{T} \quad \text{Eq. 5-50}$$

$$\mathbf{x}^{n} = [\alpha_{w}^{n} \quad \Delta r_{u}^{n} \quad \Delta r_{v}^{n} \quad \Delta t_{p}^{n} \quad x_{f}^{n} \quad \Delta\omega_{z}^{n} \quad \Delta v_{x}^{n} \quad \Delta v_{y}^{n}]^{T} \quad \text{Eq. 5-51}$$

$$\mathbf{w}^{NC} = [0 \quad 0 \quad 0 \quad w_{p}^{NC} \quad w_{f}^{NC} \quad w_{\Delta\omega_{z}^{NC}} \quad w_{\Delta v_{x}^{NC}} \quad w_{\Delta v_{y}^{NC}}]^{T} \quad \text{Eq. 5-52}$$

$$\mathbf{w}^{n} = [0 \quad 0 \quad 0 \quad w_{p}^{n} \quad w_{f}^{n} \quad w_{\Delta\omega_{z}^{n}} \quad w_{\Delta v_{x}^{n}} \quad w_{\Delta v_{y}^{n}}]^{T} \quad \text{Eq. 5-53}$$

The measurement model is given by Eq. 5-33. In practice, the elements of $\mathbf{w}$ which are zeros are usually replaced with small non-zero terms to add fictitious noise to the system, to improve the stability of the Kalman filter, as described in Section 3.5.2.

It should be noted that the equations derived in Appendix A of Schneider[44] can be obtained by substituting $\Delta\dot{\theta}_a = -\Delta\omega_z$ and Eq. 5-54 in Eq. 5-34, Eq. 5-35, Eq. 5-36, Eq. 5-42, Eq. 5-43 and Eq. 5-44. This implies that Eq. 5-34 to Eq. 5-49 represent the RELNAV algorithms in the form of strapdown INS, including the error model of the system time.

$$\mathbf{C}_{b_n}^{r} = \begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Eq. 5-54}$$

The RELNAV algorithm developed in this section is centralised since all state variables and measurements are handled by a single master filter. Consequently, an enormous amount of data would have to be transferred from all members to the master EKF. In a practical system, each member is likely to have at least 20 state variables as discussed in Section 5.3. Moreover, all data transfers and processing has to be done in real-time, the limitations of computation and communication may make the centralised approach impractical. Nevertheless, the study of the centralised RELNAV is useful analytically because it produces the optimal solutions, providing a datum for performance comparison of other Kalman filtering implementations. The performance of the centralised RELNAV EKF is covered in the following section and the more practical implementation of the RELNAV algorithms based on decentralised filtering is developed in Chapter 6. Figure 5-5 depicts the centralised RELNAV extended Kalman filter configuration.

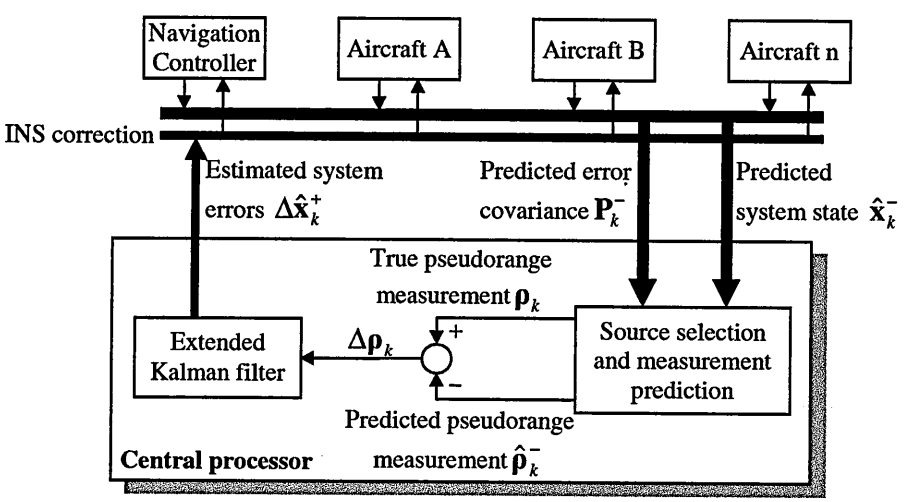


Figure 5-5 Centralised RELNAV extended Kalman filter configuration

5.6 CENTRALISED RELNAV PERFORMANCE ANALYSIS

Figure 5-6 illustrates the flight trajectories used in the RELNAV performance analysis studies. These flight trajectories have been chosen to represent typical civil air traffic manoeuvres, including parallel and crossing track sections. The NC, the first and the second member aircraft have high, medium and low accuracy systems respectively, based on the sensor parameters given in Table 4-3. The source selection function is based on the fixed-rank hierarchy where the NC, the first and second member aircraft are assigned 1st, 2nd and 3rd rank. Therefore, the first member aircraft can use the measurements from the NC and the second member aircraft can use the measurements from the NC and the first member aircraft. Figure 5-7 shows the theoretical rms errors of the RELNAV system states (the relative grid origin and orientation and the system time), which drift as functions of time. However, these errors are small because the NC has high accuracy sensors. Figure 5-8 and Figure 5-9 show the performance of the centralised RELNAV.

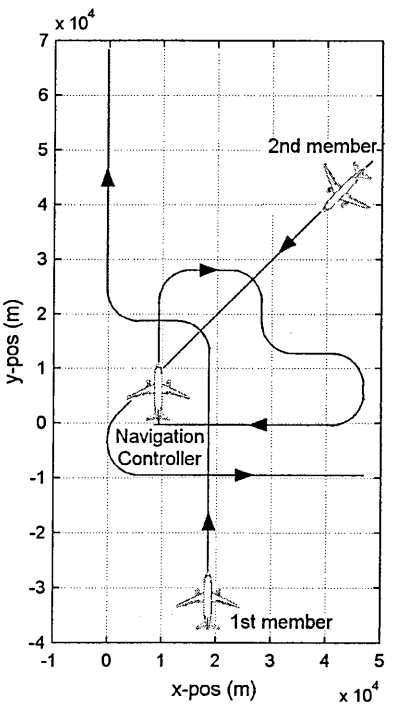


Figure 5-6 Flight trajectories for the RELNAV performance analysis

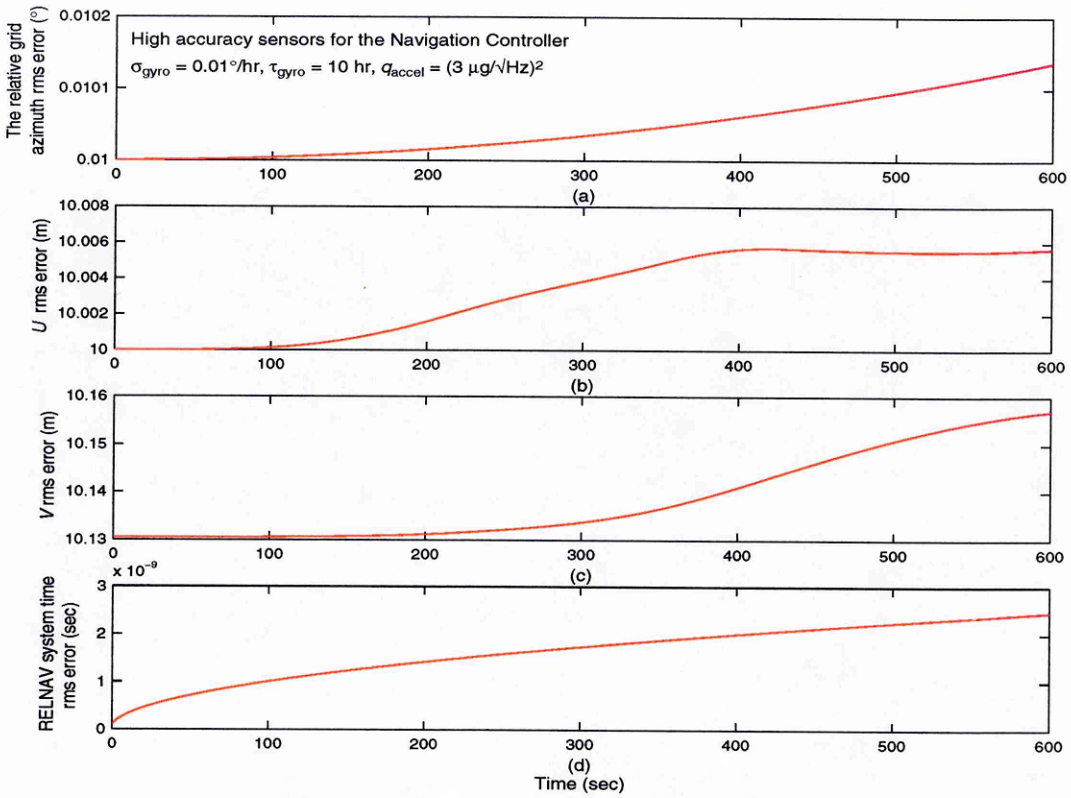


Figure 5-7 Theoretical rms errors of the RELNAV system states

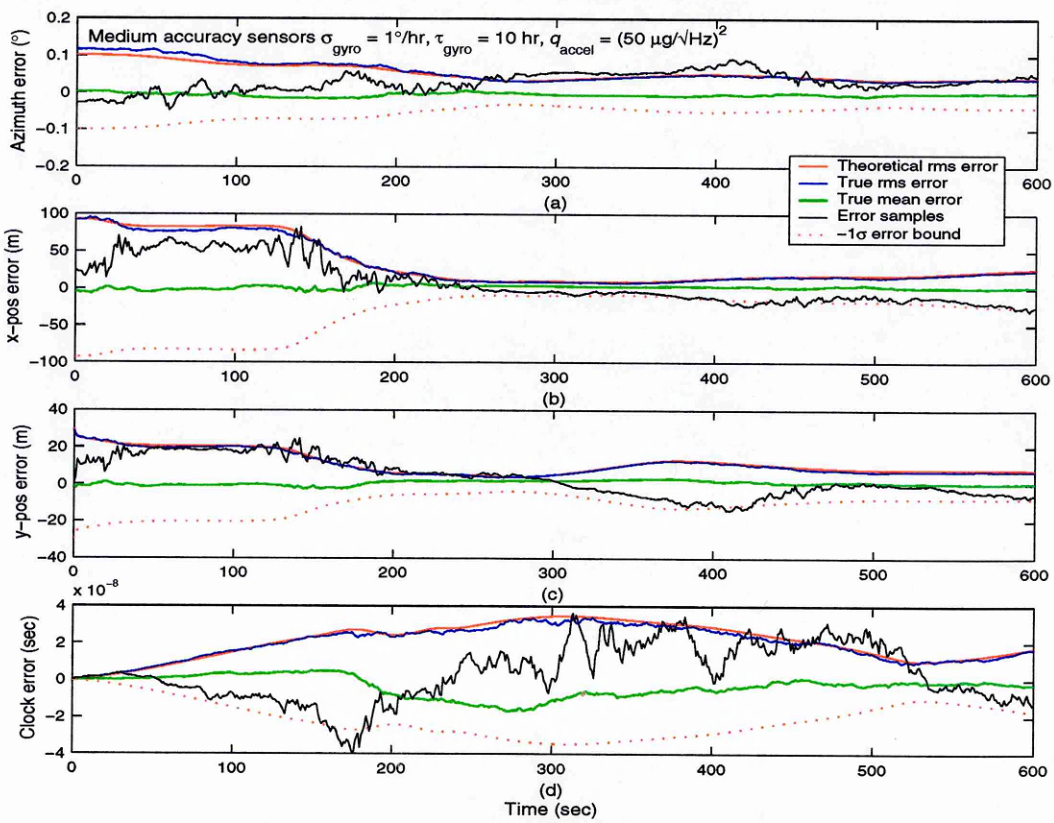


Figure 5-8 The RELNAV performance of the first member aircraft

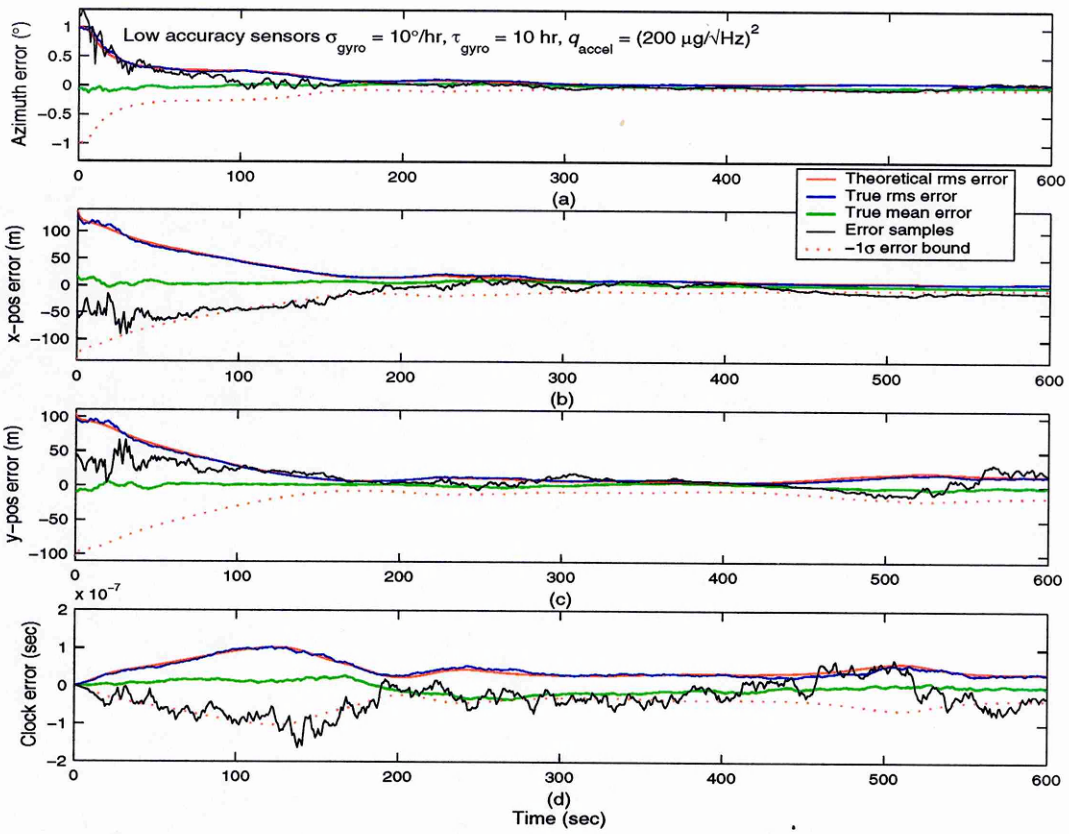


Figure 5-9 The RELNAV performance of the second member aircraft

Figure 5-10 and Figure 5-11 compare the theoretical rms error of the stand-alone INS and RELNAV for the first and second member aircraft. From inspection of the rms errors, it can be seen that the RELNAV performances for both aircraft are superior to the stand-alone INS. Because the reported positions of the NC are "*perfect*", the reduced rms errors of the other member aircraft imply that aircraft separations can be maintained (or even reduced) for the same level of confidence. These results are vital because they show that, theoretically, RELNAV can be used to accurately determine and correlate the relative positions of aircraft and, in addition, that it can be applied to civil ATM, even in the case of the absolute navigation system failures. Although the simulated flight trajectories in Figure 5-6 are chosen arbitrarily, it can be shown by simulation that the performance of RELNAV is superior in almost all situations except where the observability is poor or when an aircraft uses pseudorange observations from a source with inferior accuracy.

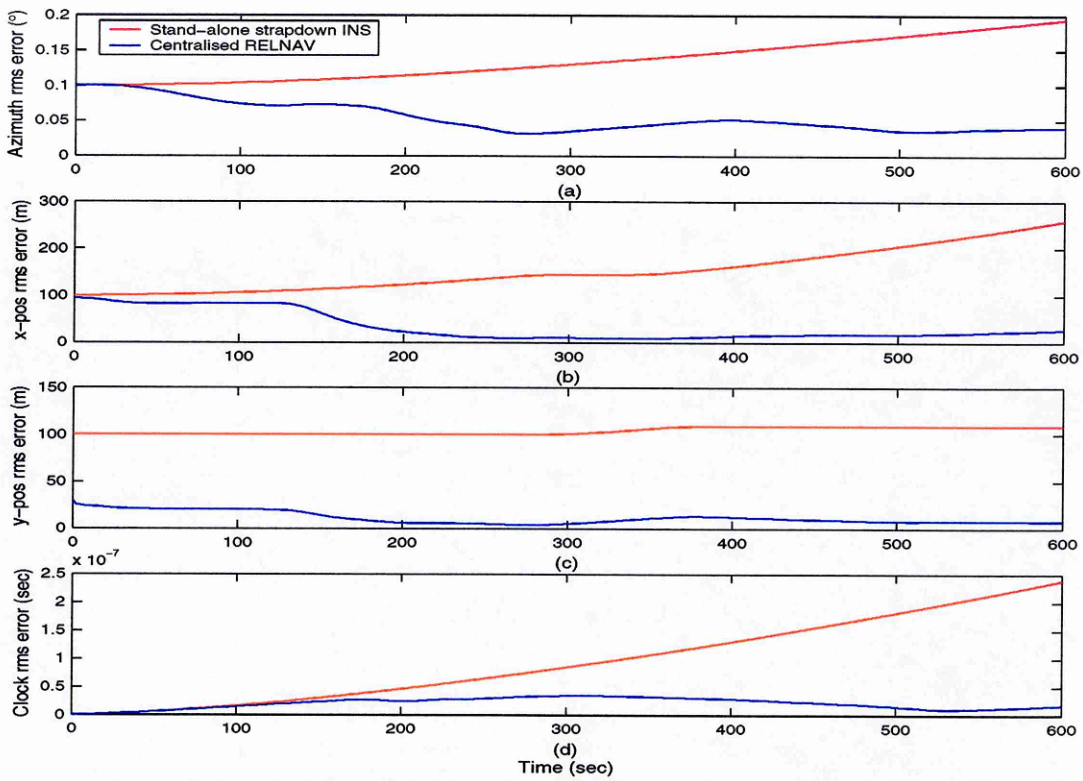


Figure 5-10 Comparison of the theoretical rms errors for stand-alone INS and centralised RELNAV for the first member aircraft

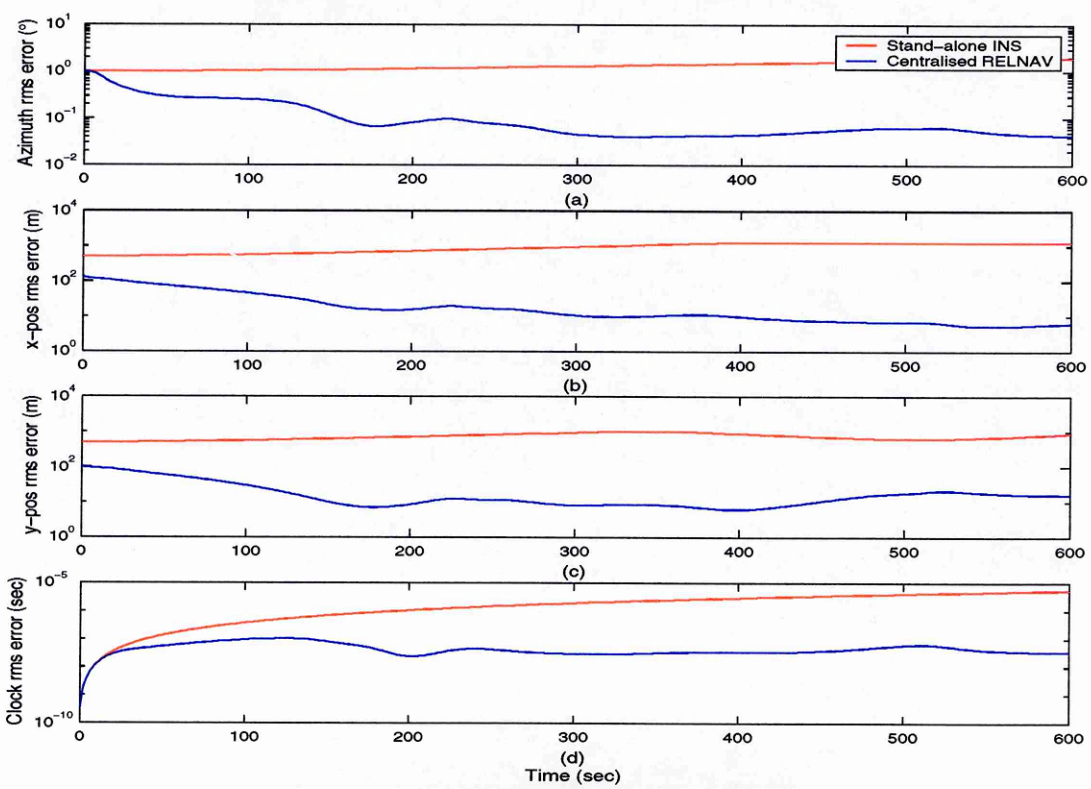


Figure 5-11 Comparison of the theoretical rms errors for stand-alone INS and centralised RELNAV for the second member aircraft

Observability plays a critical role in RELNAV because the performance of the EKF depends on it. Very accurate navigation solutions can be obtained, if the system is highly observable. However, poor observability can lead to inferior navigation solutions or divergence of the Kalman filter. Generally, the observability of RELNAV is related to the dilution of precision (DOP). It should be noted that, contrary to the single-point solution described in Section 4.4.3, which requires at least four simultaneous measurements, navigation by ranging to even a single source is possible based on the EKF approach. This is because the EKF makes use of the *memory* of the past measurements and known aircraft dynamics. However, the performance degrades as a function of the number of available sources and the DOP.[39] An example of a navigation system based on ranging to a single source is TRANSIT[40]. For the TRANSIT system, in order to improve the observability, pseudorange measurements are usually taken at relatively large sampling intervals (approximately every 2 minutes). For RELNAV, it is also known that the source(s) must change positions in order for the system to be observable and that the performance depends on the source-observer geometry.[41] The effect of source movements on the observability and performance of RELNAV is illustrated by the simulated flight trajectories shown in Figure 5-12.

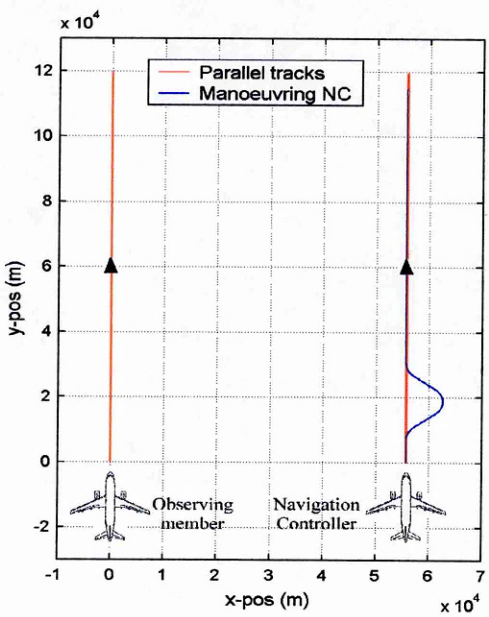


Figure 5-12 Flight trajectories for testing the observability

In Figure 5-12, the NC flies two trajectories –a parallel and a manoeuvring track. The comparison of the theoretical rms errors of stand-alone INS and RELNAV

for the observing member aircraft is shown in Figure 5-13. In the parallel track case, the effect of poor observability in the y direction is clearly seen from the magnitude of the y-position rms error. Even so, the y-position rms errors of the RELNAV parallel-track case degrade only slightly in comparison with the y-position rms errors of the stand-alone INS. This discrepancy is caused because the observing member in RELNAV has to also take into account the errors of the NC. More importantly, the azimuth, x-position and time of the observing member are still improved in comparison with the stand-alone INS, considering that the measurements are obtained from a single source. The effect of the source movement in improving the observability and performance of RELNAV can be seen from the y-position rms error of the RELNAV manoeuvring NC. This simulation demonstrates the effects of motion of the sources in improving the observability and performance of RELNAV. It also illustrates the ability of RELNAV to accurately determine relative positions even with just two members. For JTIDS RELNAV, it is feasible to ask a member (for example, a fast moving fighter aircraft in a fighter group) to manoeuvre to improve the observability. For civil applications, this problem may be more difficult to overcome and will be addressed further in Chapter 6.

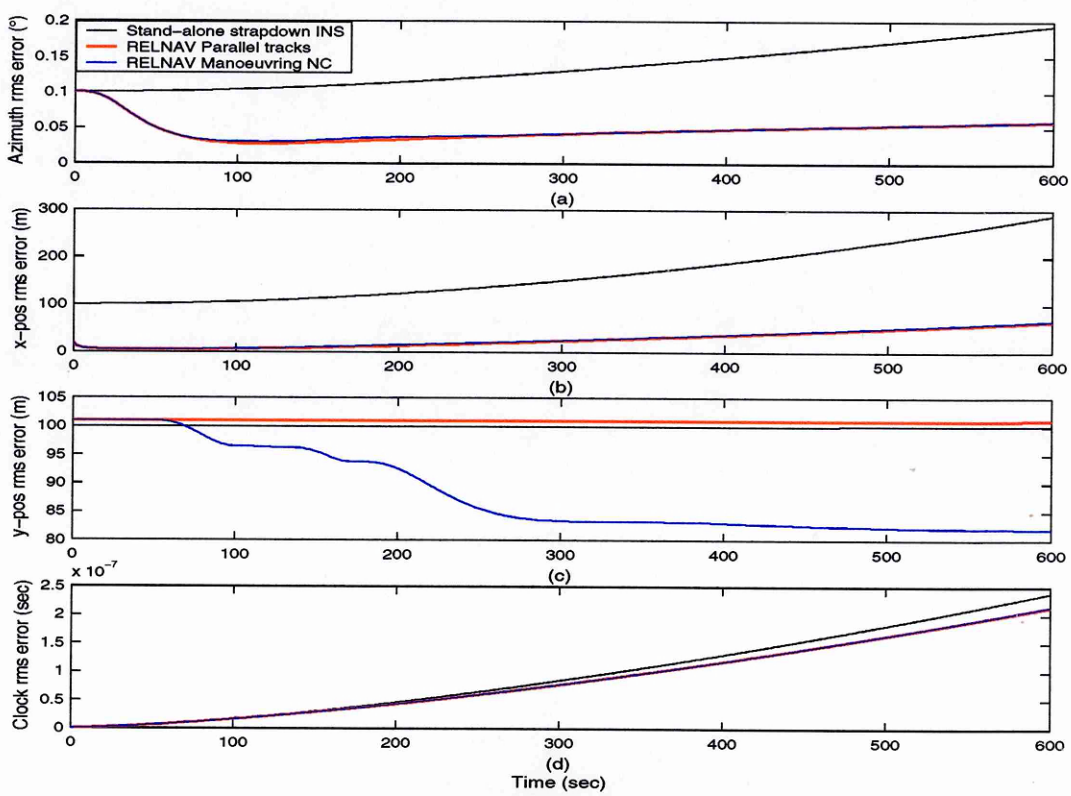


Figure 5-13 The effects of the source motion on the observability and performance of RELNAV

As already mentioned, the source selection function is crucial to the successful application of RELNAV. Figure 5-14, Figure 5-15 and Figure 5-16 demonstrate the divergence of the EKF and the unstable navigation community when the observers are allowed to use measurements from the sources with inferior accuracy. The simulations were based on the flight trajectories shown in Figure 5-6. In this case, the sensor accuracy levels are reversed. The NC, the first and the second member aircraft have low, medium and high accuracy systems respectively, but the ranking between these aircraft is retained. For example, the first member aircraft, which now has a medium accuracy system, is allowed to use measurements from the NC, which has a low accuracy system.

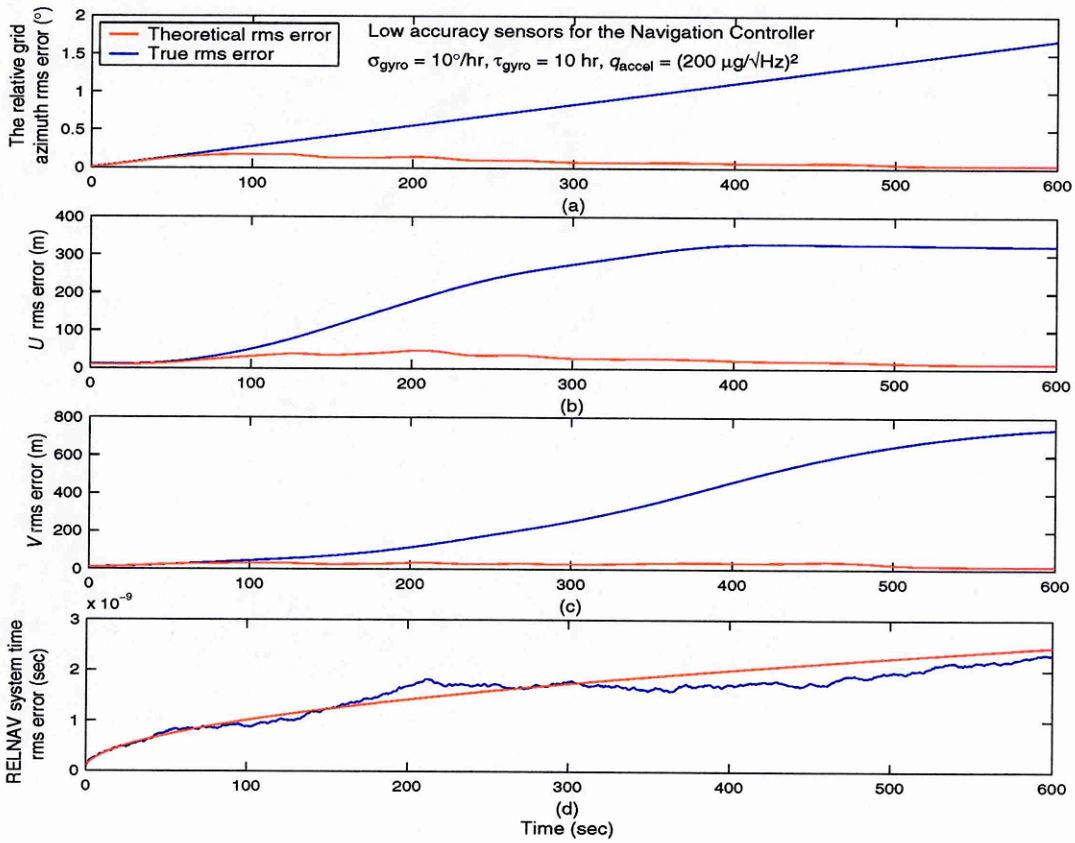


Figure 5-14 Divergence of the RELNAV system state estimation errors

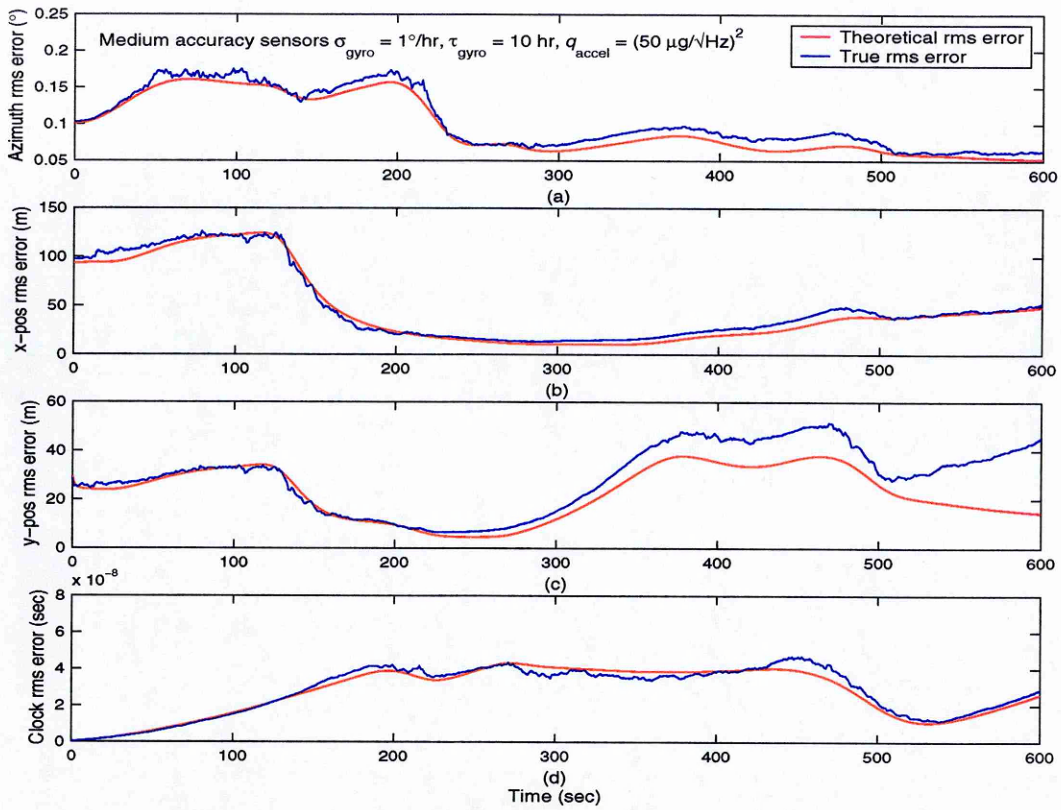


Figure 5-15 Divergence of the first member aircraft estimation errors

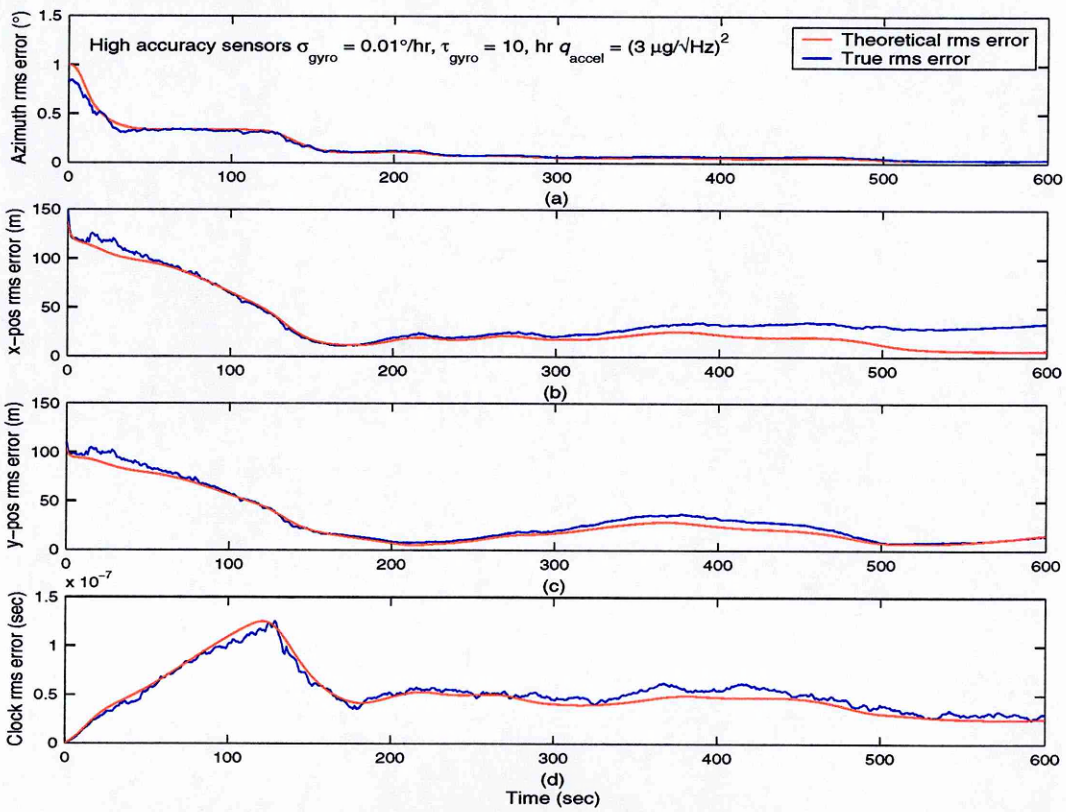


Figure 5-16 Divergence of the second member aircraft estimation errors

In forming a centralised RELNAV system model, the RELNAV system states are coupled with the state vectors of other members in the navigation community. If the other members have more accurate sensor than the NC (invalidating the fundamental assumption), the EKF tries to take into account these more accurate systems by reducing the errors of the RELNAV system states. However, because the relative grid system states are not observable to any member, this leads to divergence of the RELNAV system state estimation errors as shown in Figure 5-14. Subsequently, the distribution of absolute positional information will no longer be valid. Figure 5-15 and Figure 5-16 demonstrate the divergence of the estimation errors of the first and second member aircraft when the measurements from the source(s) with inferior accuracy are used in the EKF. Nevertheless, it should be noted that the EKF exhibits a certain degree of natural stability, and the growth of the divergence is limited.

In this section, the principles of JTIDS RELNAV have been explained. The centralised RELNAV navigation algorithms have been derived and the system performance has been analysed. The results obtained in this section confirm two importance aspects of RELNAV.

⇒ The source selection function is critical to the stability of the RELNAV EKF and, hence, the successful implementation of RELNAV; only observations from superior accuracy sources should be processed.

⇒ Secondly, the observability, which is dependent on aircraft manoeuvre, is a contributory factor to the performance of RELNAV.

The ability of RELNAV to accurately determine and correlate relative position data has been demonstrated and, it is clear that these assets can be highly valuable if they can be applied to civil ATM. However, from a practical viewpoint, the centralised RELNAV approach is impracticable, owing to practical limitations. The next section discusses the application of RELNAV to civil ATM based on the decentralised filtering, which is a more feasible approach.

CHAPTER 6 RELATIVE NAVIGATION FOR CIVIL APPLICATIONS

6.1 INTRODUCTION

The benefits of relative navigation (RELNAV) in civil aviation were covered in Chapter 5. It will be shown that it is necessary to implement the RELNAV algorithms based on decentralised filtering. This section covers the development of the ownstate decentralised RELNAV extended Kalman filter (EKF) and compares the system performance with the centralised RELNAV covered in Section 5.6. Factors, which influence the performance of the decentralised RELNAV, and practical methods, which can be used to alleviate the inherent problems, are discussed. The operational concept and implementation of the decentralised RELNAV algorithms for civil applications based on ADS-B is explained. Finally, the performance of a civil RELNAV system is analysed.

6.2 DECENTRALISED RELNAV NAVIGATION ALGORITHMS

The decentralised RELNAV EKF used in this thesis is derived from the *ownstate filter community* described by Gobbini[45,47]. Each member models only the errors of their own platform navigation systems and does not include the error states of other members in the overall system model. Measurements are processed by the EKF implemented in each individual platform. The main advantage of this approach is the elimination of a master filter and reduced computing requirements. It will be assumed that no sharing of information, other than that required in the ownstate RELNAV measurement model, is allowed. Although this assumption prevents the implementation of the optimal decentralised EKF[45], which produces the same optimal solutions as the centralised RELNAV using several distributed filters, it limits the amount of data transferred by the communication system such as ADS-B, which is particularly important from a practical viewpoint. However, for reason of simplicity, it will be assumed that full error covariance matrices of sources ($\mathbf{P}_{s_i}$ in Eq. 6-7) are known by an observer. The development of the ownstate decentralised RELNAV EKF is covered in the following sections.

6.2.1 Development of ownstate decentralised RELNAV overall system model

The equivalent discrete-time plant model of the ownstate decentralised RELNAV algorithms for the n^{th} observing member can be expressed as follows:

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k \mathbf{x}_k + \mathbf{w}_k \qquad \text{Eq. 6-1}$$

where $\mathbf{x}_k$ and $\mathbf{w}_k$ are the INS error states and the plant noise defined in Eq. 5-51 and Eq. 5-53. The measurement model is derived directly from Eq. 5-33 and can be written in matrix form as follows:

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k - \mathbf{u}_k + \mathbf{v}_k \qquad \text{Eq. 6-2}$$

where $\mathbf{v}_k$ is pseudorange measurement noise. $\mathbf{u}_k$ represents additional measurement noise resulting from uncertainties in the sources, such as the reported positions of received ADS-B messages. The i^{th} row of $\mathbf{u}_k$ is defined as follows:

$$u_i = \mathbf{h}_{s_i} \mathbf{x}_{s_i} \qquad \text{Eq. 6-3}$$

where $\mathbf{x}_{s_i}$ is the error state vector of the i^{th} source. The non-zero terms of $\mathbf{h}_{s_i}$ can be obtained from the respective direction cosine terms in Eq. 4-52. It will be assumed that $\mathbf{w}_k$, $\mathbf{u}_k$ and $\mathbf{v}_k$ are zero-mean and white. $\mathbf{u}_k$ is assumed to be white based on the assumption that the error $\mathbf{x}_{s_i}$ is white, i.e. uncorrelated between different sources and observers.

6.2.2 Derivation of ownstate decentralised RELNAV state estimator

The overall system model of the decentralised RELNAV is given by Eq. 6-1 and Eq. 6-2. This model is almost in a suitable form for the EKF except for the inclusion of $\mathbf{u}_k$. This section describes the modification to the standard EKF equations to take account of $\mathbf{u}_k$. The details of the derivation are given in Appendix G. Let $\mathbf{P}_k^- = \mathrm{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)^{\mathrm{T}}]$, where $\hat{\mathbf{x}}_k^-$ denotes the estimate of $\mathbf{x}_k$ before incorporating measurements. The standard Kalman filter update equation is given by:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \qquad \text{Eq. 6-4}$$

The estimation error covariance matrix $\mathbf{P}_k^+$ is given by:

$$\mathbf{P}_k^+ = \mathrm{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_k^+)(\mathbf{x}_k - \hat{\mathbf{x}}_k^+)^{\mathrm{T}}] \qquad \text{Eq. 6-5}$$

Substituting Eq. 6-2 and Eq. 6-4 in Eq. 6-5 and noting that the expectation of the product of uncorrelated noise is zero, it can be shown that $\mathbf{P}_k^+$ is given by:

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}} + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^{\mathrm{T}} \qquad \text{Eq. 6-6}$$

where $\mathbf{R}_k = \mathrm{E}[\mathbf{v}_k\mathbf{v}_k^{\mathrm{T}}]$ and $\mathbf{S}_k = \mathrm{E}[\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}}]$. Based on the assumption that source errors are white, the diagonal terms of $\mathbf{S}_k$ can be determined as follows:

$$\begin{aligned}\mathrm{E}[u_i u_i] &= \mathrm{E}[(\mathbf{h}_{s_i}\mathbf{x}_{s_i})(\mathbf{h}_{s_i}\mathbf{x}_{s_i})^{\mathrm{T}}] \\ &= \mathrm{E}[\mathbf{h}_{s_i}\mathbf{x}_{s_i}\mathbf{x}_{s_i}^{\mathrm{T}}\mathbf{h}_{s_i}^{\mathrm{T}}] \\ &= \mathbf{h}_{s_i}\mathbf{P}_{s_i}\mathbf{h}_{s_i}^{\mathrm{T}}\end{aligned} \qquad \text{Eq. 6-7}$$

where $\mathbf{P}_{s_i} = \mathrm{E}[\mathbf{x}_{s_i}\mathbf{x}_{s_i}^{\mathrm{T}}]$ is the error covariance matrix of the i$^{\text{th}}$ source, which can be used to determine the accuracy levels of the sources. $\mathbf{S}_k$ can be written in matrix form as follows:

$$\mathbf{S}_k = \begin{bmatrix} \mathbf{h}_{s_1}\mathbf{P}_{s_1}\mathbf{h}_{s_1}^{\mathrm{T}} & 0 & \cdots & 0 \\ 0 & \mathbf{h}_{s_2}\mathbf{P}_{s_2}\mathbf{h}_{s_2}^{\mathrm{T}} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{h}_{s_n}\mathbf{P}_{s_n}\mathbf{h}_{s_n}^{\mathrm{T}} \end{bmatrix} \qquad \text{Eq. 6-8}$$

By applying the same approach used in the derivation of the standard Kalman filter equations, the optimal (least-squares errors) gain $\mathbf{K}_k$ is determined as follows:

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k + \mathbf{S}_k)^{-1} \qquad \text{Eq. 6-9}$$

To project the system error states to time t_{k+1}, the standard EKF prediction equations given in Table 3-2 can be applied without modification. The difference between the ownstate decentralised RELNAV EKF and the standard EKF equations is the addition of $\mathbf{S}_k$ to the measurement noise covariance matrix $\mathbf{R}_k$. This is to be expected because the addition of $\mathbf{S}_k$ represents an increase in measurement noise levels caused by uncertainties in the sources. It should be noted that $\mathbf{S}_k$ also takes into account source-

observer geometry because of the presence of $\mathbf{h}_{s_i}$ in Eq. 6-7.

In the derivation of Eq. 6-6 and Eq. 6-8 (Appendix G), some error terms have been ignored based on the assumption that these errors are white. Therefore, the ownstate decentralised RELNAV EKF is sub-optimal. This assumption is not strictly valid because every member determines its distance from each other (denoted as "*ranging process*") and, hence, their errors are correlated. This assumption simplifies the derivation of the decentralised RELNAV EKF algorithms considerably and it will be shown that ignoring these error terms can cause instability of the decentralised RELNAV EKF. The levels of error correlation can invalidate these assumptions, leading to divergence of the decentralised EKF. However, it will be shown that these errors can be made to appear uncorrelated to the EKF (hence, eliminating the divergence) by the technique of measurement resampling. Figure 6-1 depicts the decentralised ownstate RELNAV system configuration.

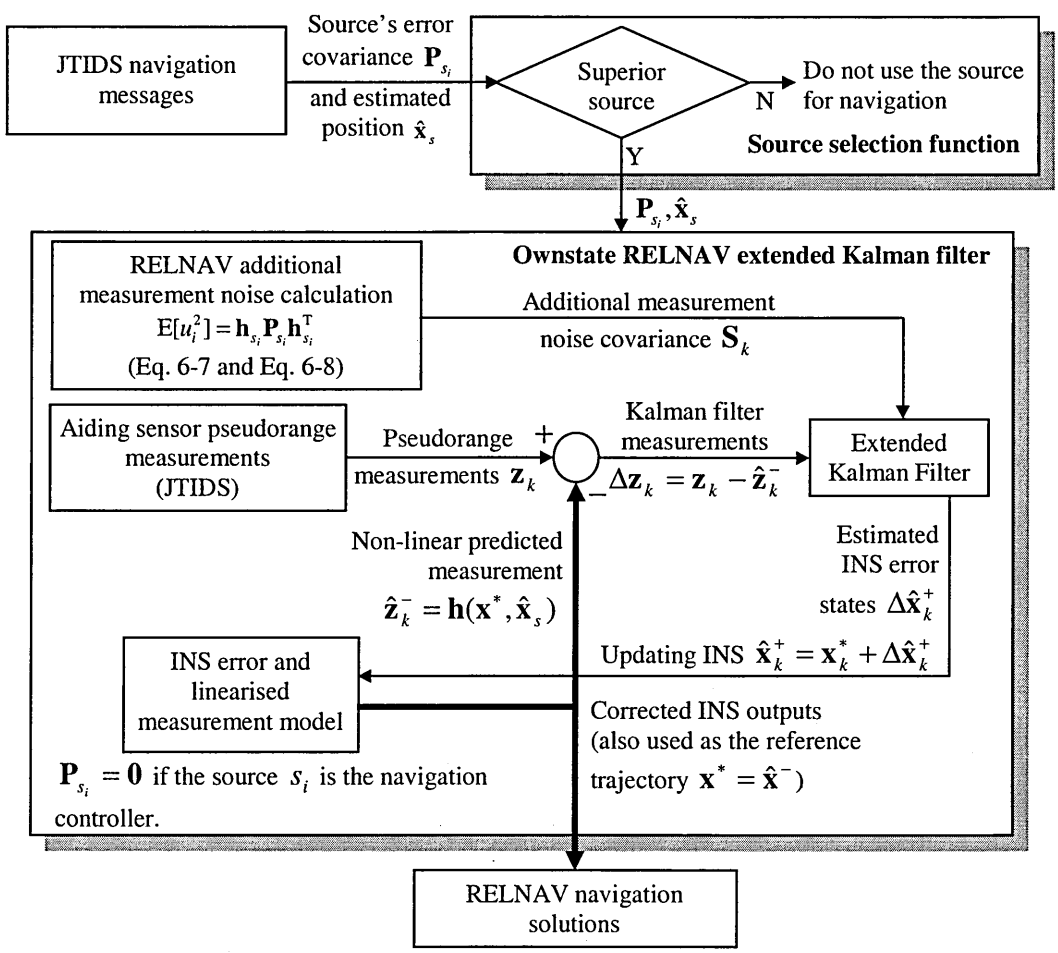


Figure 6-1 Decentralised ownstate RELNAV system configuration

6.3 DECENTRALISED RELNAV PERFORMANCE ANALYSIS

This section analyses the performance of the ownstate decentralised RELNAV EKF, based on the simulation configurations used in the analysis of the centralised RELNAV. Figure 6-2 shows that the rms errors of the RELNAV system states are the same for both the centralised and decentralised EKF as expected, because the error models of the RELNAV system states are identical for both filters.

The performance of the first member aircraft is shown in Figure 6-3. The sub-optimality of the decentralised EKF can be seen by comparing the theoretical rms errors, which is confirmed by the true rms errors. In the decentralised EKF, the RELNAV system states are not modelled by the first member aircraft. Consequently, the EKF put less weight on the plant model, leading to an increase in the computed rms errors. In this case, because the first member aircraft ranges to the Navigation Controller (NC) only, there are no modelling errors due to truncation of the correlated error terms. The analysis of the results shown in Figure 6-3 confirms that the EKF affords a certain degree of natural stability, even when several error terms (the RELNAV system states) have been excluded from the system model, providing that these errors are negligible.

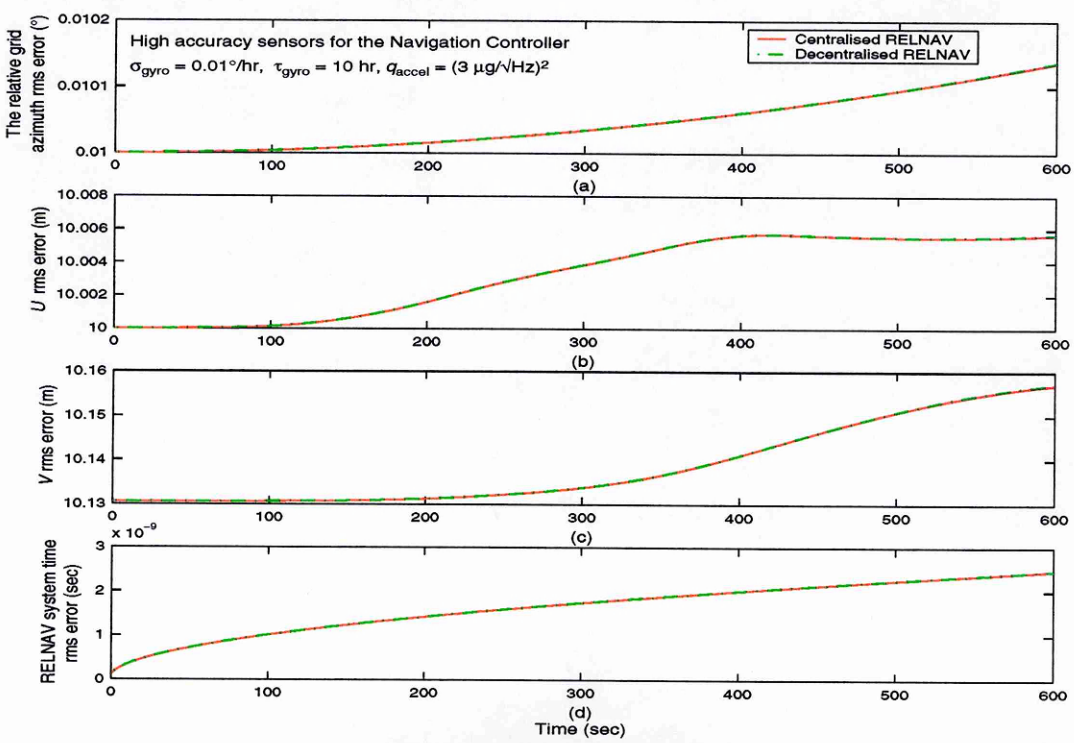


Figure 6-2 Comparison of the centralised and decentralised RELNAV EKF rms errors of the RELNAV system states

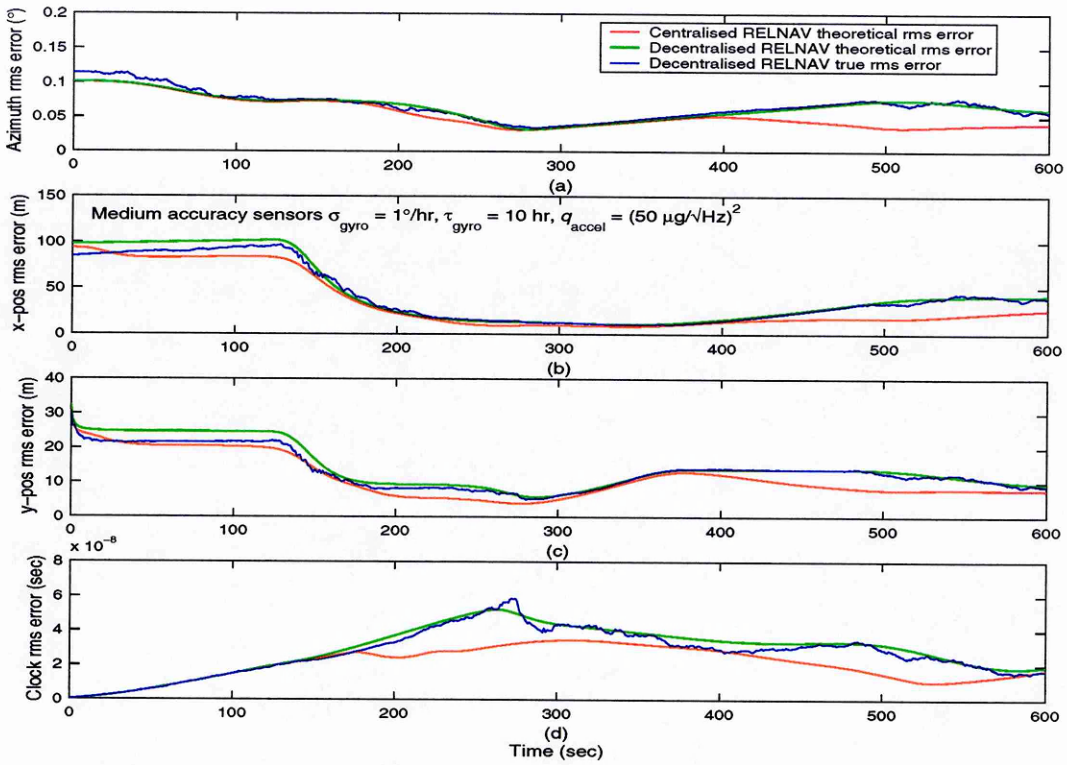


Figure 6-3 Comparison of the centralised and decentralised RELNAV EKF rms errors of the first member aircraft

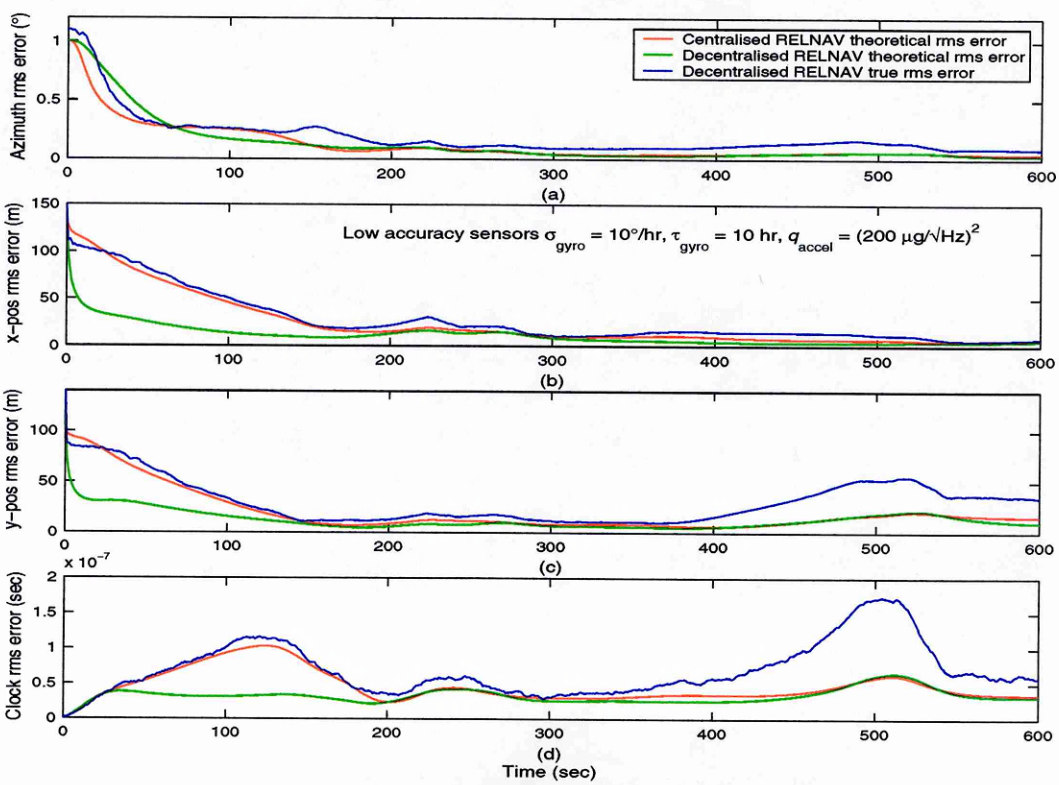


Figure 6-4 Comparison of the centralised and decentralised RELNAV EKF rms errors of the second member aircraft

Figure 6-4 shows the performance of the decentralised EKF of the second member aircraft. It can be seen that the theoretical rms errors of the decentralised EKF are less than the theoretical rms errors of the centralised EKF. This is because the RELNAV system states and the error states of the first member aircraft, have been excluded from the overall system model used by the second member aircraft. To the EKF, this makes a system appear easier to filter, than it actually is. Based on measurements with the same accuracy levels as those used by the centralised EKF, the decentralised EKF subsequently reduces the computed rms errors. However, the comparison of the theoretical and the true rms errors shows that the filter actually diverges. The divergence is attributed to the correlated error terms increasing to a level where they are no longer negligible. The errors of the first and the second member aircraft are correlated because both aircraft range to the same source - the NC. The effect of the error correlation on the divergence is confirmed by comparing the rms errors of the second member aircraft for the case where it ranges to the NC alone, as shown in Figure 6-5. In this case, the assumption that errors are uncorrelated is valid because the NC does not range to other members. Consequently, the filter performs correctly without diverging.

An ad hoc method that has been used successfully in decentralised filtering, to alleviate the problems of correlated sensor errors, is to use measurements taken from relatively large sampling intervals.[11] For example, if a sensor outputs measurements at 1-Hz rate, the measurements taken at every 20^{th} or 50^{th} sample may produce satisfactory performance. This method makes the measurements appear less correlated in time. However, an effective resampling rate usually depends on the actual levels of correlation of errors and may differ between applications.

Figure 6-6 illustrates the effect of measurement resampling on the performance of the decentralised EKF of the second member aircraft. In this case, the update rate of the observations from the NC remains at 1-Hz where every 60^{th} measurement from the first member aircraft is processed, i.e. one minute resampling rate. It can be seen that the divergence is eliminated and although, this method is sub-optimal, the performance is still superior to using measurements from the NC alone. This implies that measurements from the first member aircraft provide some useful information for filtering.

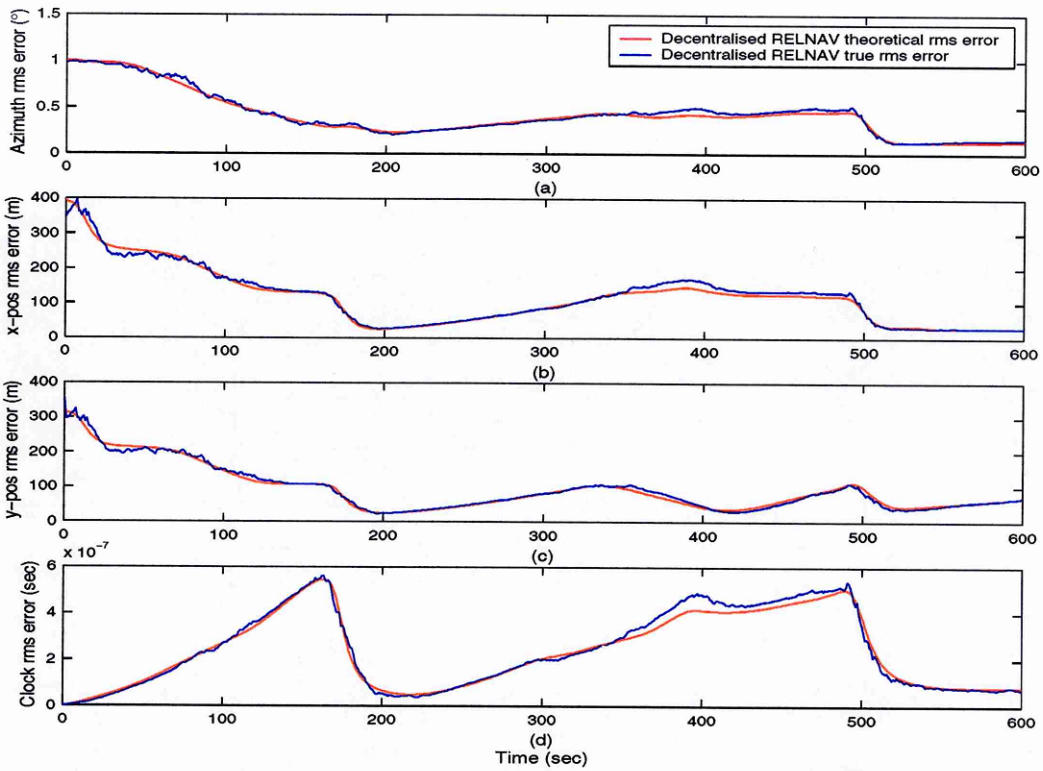


Figure 6-5 The performance of the decentralised RELNAV EKF of the second member aircraft when measurements from the NC alone are used

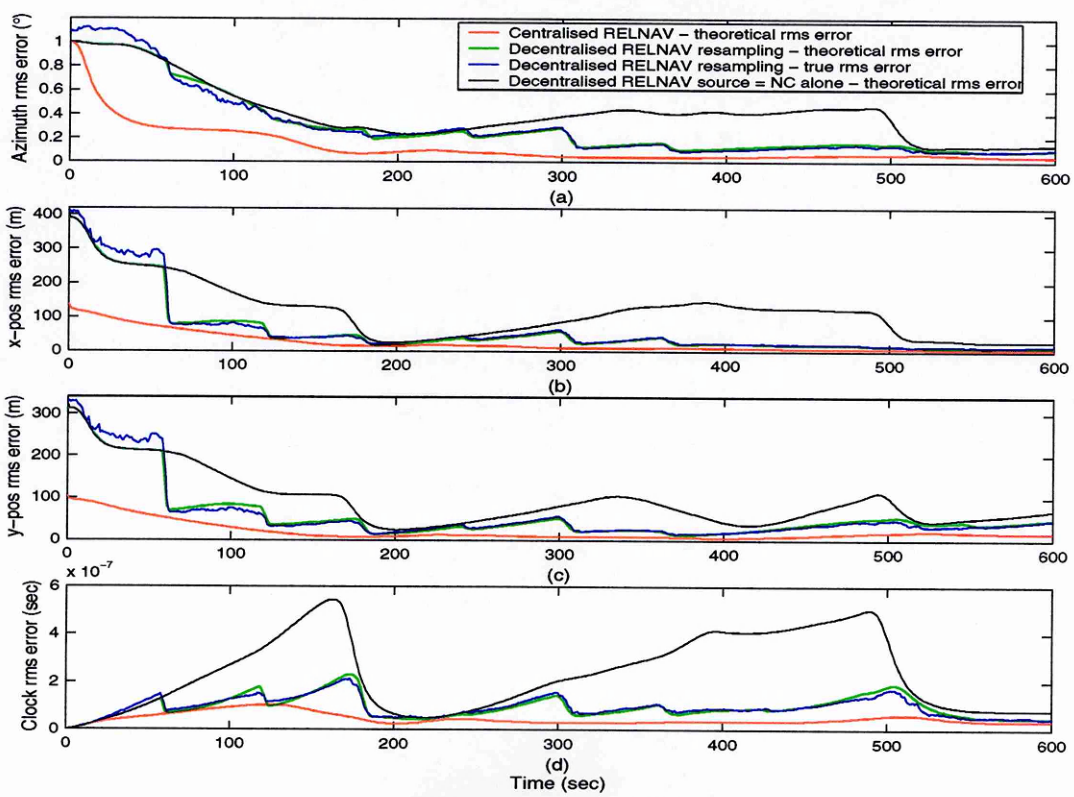


Figure 6-6 The effect of resampling measurements on the performance of the decentralised RELNAV EKF

In this section, the performance of the ownstate decentralised RELNAV EKF is analysed. It has been shown that the application of the ownstate decentralised EKF in RELNAV has the capability to accurately 1.) *determine and correlate relative position data* and 2.) *distribute navigation accuracy*. However, the risk of filter divergence is increased because of the assumption that the errors are uncorrelated, which is necessary to simplify the derivation of the algorithms. The method of measurement resampling has been shown to be successful in alleviating the problem of correlated errors. In conclusion, the analysis in this section confirms that the benefits of RELNAV can be obtained based on the application of the ownstate decentralised RELANV EKF, providing that the effects of ignoring the correlated error terms are taken into account. The following section discusses the application of the RELNAV principles to a civil air traffic environment using ADS-B.

6.4 PRINCIPLES OF CIVIL RELATIVE NAVIGATION

The principles of JTIDS RELNAV and the performance of the optimal centralised RELNAV have been covered in Chapter 5. The sub-optimal decentralised ownstate RELNAV algorithms, necessary for real-world applications, have been developed in Section 6.2 and the system performance has been analysed in Section 6.3. The usefulness of RELNAV, which is based on the ability to exchange data and determine precise relative pseudorange observations based on time-of-arrival (TOA) measurements of received messages, is highly desirable in the civil Air Traffic Management (ATM).

In civil ATM, ADS-B provides a fundamental capability that is similar to JTIDS, allowing JTIDS RELNAV principles to be applied. Although, ADS-B has not been fully specified, because the system is still under development, it is reasonable to assume that ADS-B will be able to provide accurate TOA and, hence, relative pseudorange measurements needed in RELNAV calculations.

Nevertheless, a major complexity in implementing RELNAV in a civil air traffic environment is establishing a navigation community. The designation of the NC and the Net Time Reference (NTR) and the establishment of a relative grid are difficult because the availability and capability of every member is not usually known in

advance. If several navigation communities are established due to limited coverage of ADS-B, further complexity is also introduced by interactions between members of the different communities. To avoid the problem of establishing and managing several local navigation communities, methods similar to those adopted for JTIDS RELNAV are proposed to provide an integrated INS/GPS/ADS-B system (civil RELNAV) for civil applications. The operational concept of a civil RELNAV is explained in the next section.

6.4.1 Civil RELNAV navigation mechanisation

In an integrated INS/GPS/ADS-B system, there is a single global navigation community, where GPS satellites are designated the NC and the NTR by definition. The relative grid is a reference frame common to all members, such as the earth frame. The problem of establishing and organising several local navigation communities is eliminated because all members navigate and synchronise with respect to a single common datum, which is the GPS system frame of reference and system time. In operation, ADS-B provides additional navigation sensors in the form of relative pseudorange measurements, which are combined with GPS pseudorange measurements, to aid INS, based on RELNAV principles. Figure 6-7 and Figure 6-8 illustrate the JTIDS RELNAV and integrated INS/GPS/ADS-B system operational concepts. Figure 6-10 shows the organisation of the ownstate decentralised RELNAV EKF.

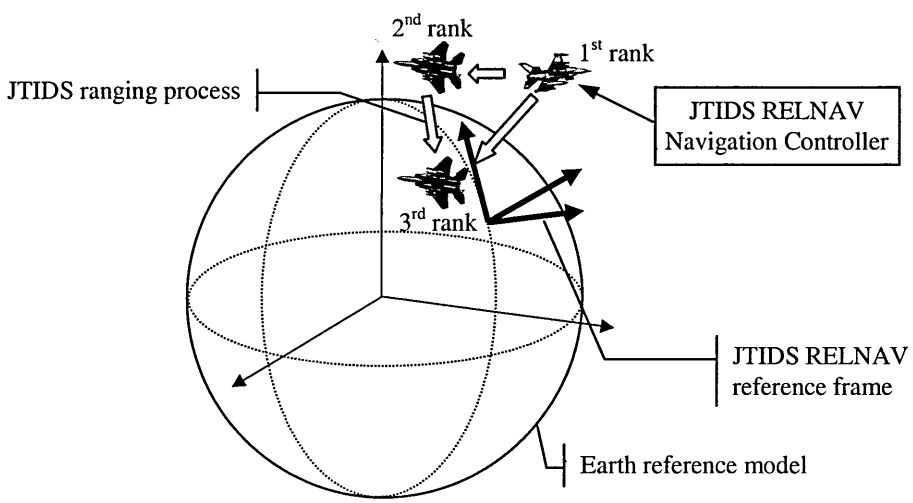


Figure 6-7 JTIDS RELNAV operational concepts

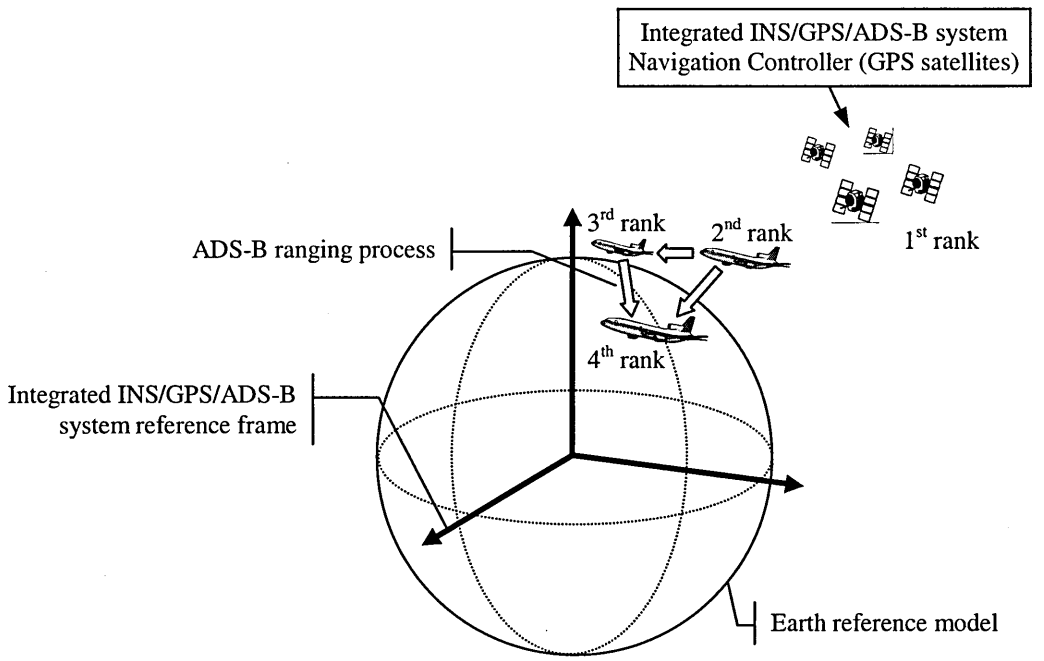


Figure 6-8 Integrated INS/GPS/ADS-B system operational concepts

6.4.2 Development of source selection functions

A source selection function is critical for RELNAV operations because it determines the accuracy and the stability of navigation solutions.[46,47,48] Navigation solutions are unstable when navigation filters diverge. It will be shown that, to ensure the stability of navigation solutions, the source selection function of the ownstate decentralised RELNAV EKF must establish observations from sources with superior accuracy. Additionally, the source selection function must also take into account the correlation of errors to prevent the divergence due to modelling error, as discussed in Section 6.3. In a civil air traffic environment, accurate navigation solutions and a stable navigation community are essential for safe aircraft operations.

In RELNAV, a navigation community can be categorised in three groups – democratic, fixed-rank hierarchy and covariance-based hierarchy. In a democratic community, a member can range to all other members. This type of community is unstable.[46,47] In a fixed-rank hierarchy, every member is assigned a fixed rank, which can not be changed during a mission. Higher rank[i] (less accuracy) members can range to

[i] In this thesis, the term "rank" is used specifically to refer to numerical values of parameter (see Figure 6-11). This convention is used so that the navigation controller is always assigned the rank value of 1.

other lower rank (more accuracy) members only. It has been proved that this type of community is stable providing that the system is stochastically stable.[45,47] However, because the ranks can not be altered during a mission, members in the fixed-rank hierarchy organisation are unable to take advantage of more accurate sources when they become available. Based on the previous discussion, the source selection function for the integrated INS/GPS/ADS-B system investigated in this thesis uses a covariance-based hierarchy, which is also the basic principle of JTIDS RELNAV source selection function.

In a covariance-based hierarchy community, rank assignment is based on the currently estimated accuracy levels of each member, which can be determined dynamically, for example, as the trace (sum of diagonal terms) of the estimation error covariance matrix. In theory, this approach assures that members always range to more accurate sources. In practice, however, filter divergence can occur because of rank reversals as depicted in Figure 6-9.

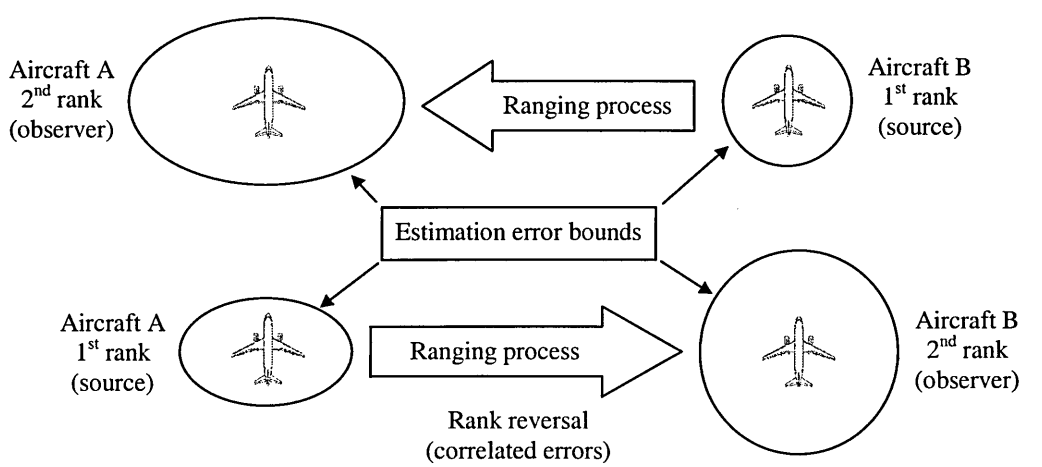


Figure 6-9 Covariance-based hierarchy rank reversals

In Figure 6-9, at the beginning of the ranging process aircraft B has superior accuracy to aircraft A, hence, aircraft A is allowed to range to aircraft B, leading to subsequent reduction in its estimation errors. Because aircraft B does not range to other members, its estimation errors will grow with time as functions of its INS sensor errors. As the ranging process continues, a situation is reached when the accuracy levels indicate that aircraft A is more accurate than aircraft B. Consequently, a rank reversal occurs, where aircraft B is then allowed to range to aircraft A. It may seem that the

accuracy of both aircraft can be maintained indefinitely by this process. Nevertheless, aircraft B's EKF will subsequently diverge owing to modelling errors. In the ownstate decentralised EKF formulation, the error states of aircraft A and aircraft B are assumed to be uncorrelated when, in fact, their errors are highly correlated because aircraft A ranged to aircraft B before the rank reversal.

A direct and optimal method to solve the divergence problem is to model the correlated error terms, based on the decentralised filtering theory.[45] From a practical viewpoint, however, this approach leads to significant increases in computation and communication. Because it is proposed that civil RELNAV will be based on ADS-B and that minimum demands should be placed on the system, practical methods of minimising the correlation of errors have been specifically investigated in this thesis. For example, the correlation of errors can be lessened by reducing the occurrences of rank reversals and measurement resampling. These methods can be implemented as additional rules in the source selection function by all members. Additionally, the measurement noise covariance matrix $\mathbf{S}_k$ can be increased (tuned) to account for the unmodelled correlated error terms too.

In this thesis, the accuracy levels of members in a covariance-based hierarchy community have been derived from traces of the estimation error covariance matrix **P** of the ownstate decentralised EKF. However, it can be shown that by assigning ranks to all existing members, based directly on the order of these accuracy levels, can lead to unstable navigation solutions because the error states are highly correlated as a consequence of frequent rank reversals. To avoid this instability, these accuracy levels are converted to *quality indices* (similar to the method used in JTIDS). In this thesis, accuracy level values between 0 to 20000 have been arbitrarily chosen and then divided into both 8 and 15 intervals, where each interval represents one quality index, as shown in Figure 6-11. The accuracy level values below 0 represent the highest quality reserved for 1st rank members such as the NC. Accuracy level values above 20000 represent the lowest accuracy members.

The grouping of members into the same quality index reduces the frequency of rank reversals caused by members with similar accuracy levels. In practice, if members have similar accuracy levels, it can be assumed that the performance of their navigation

systems is also similar and, in this case, ranging between these members offers little improvement. Therefore, these members are assigned the same quality index which is preferable to risking filter divergence. The selection of these accuracy levels and the parameters describe previously have been chosen arbitrarily to demonstrate the effects of source selection functions and its significance to the stability of a navigation community. It should be noted that this method does not guarantee filter stability. The organisation of the accuracy levels and quality indices is a trade-off between stability and filter performance. If the accuracy level intervals are small, members can take advantage of more accurate sources more frequently but with an increased sensitivity to instability. On the other hand, if the accuracy level intervals are large, the community will be more stable but the accuracy may not be distributed efficiently. The performance of the integrated INS/GPS/ADS-B system is analysed in the following section. Figure 6-10 depicts an integrated INS/GPS/ADS-B system configuration.

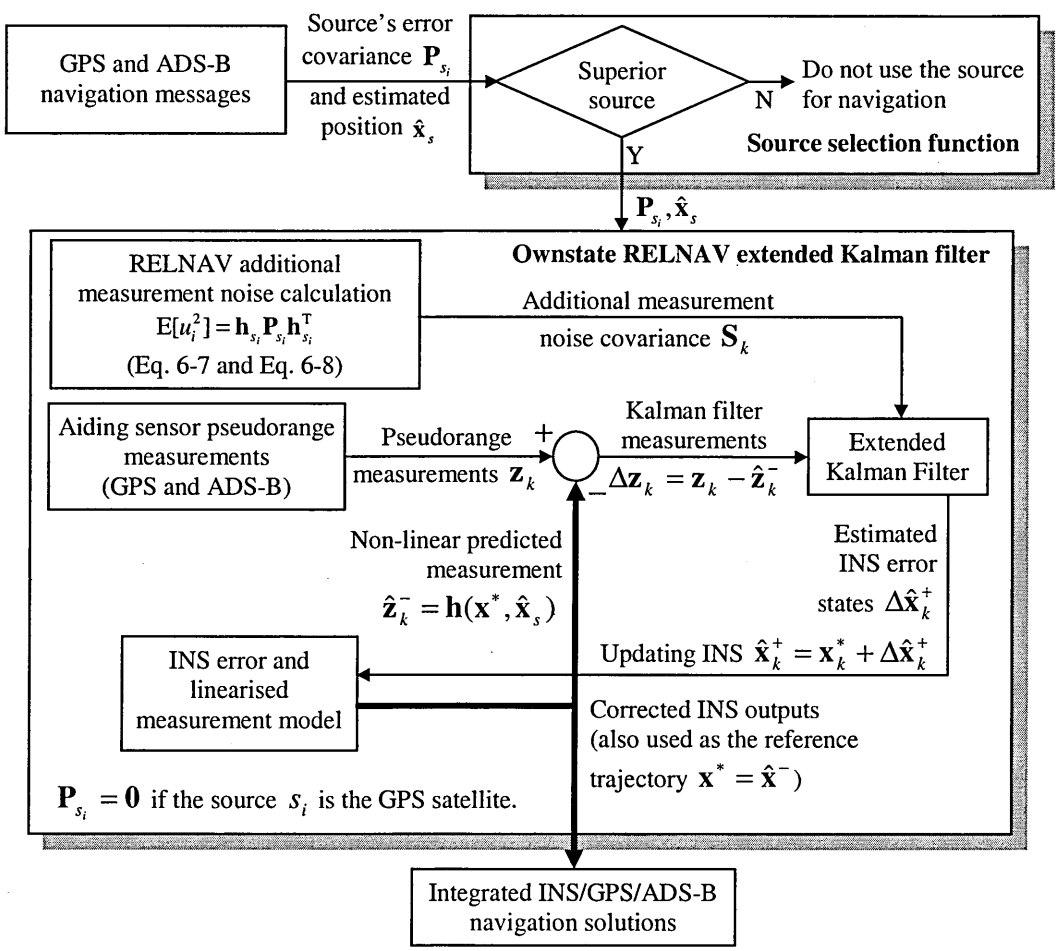


Figure 6-10 Integrated INS/GPS/ADS-B system configuration

6.5 INTEGRATED INS/GPS/ADS-B SYSTEM PERFORMANCE ANALYSIS

The simulated flight trajectories shown in Figure 5-6 have been used in the analysis of the integrated INS/GPS/ADS-B system performance. Because GPS satellites are designated as the NC in civil RELNAV, the NC, the first and the second member aircraft specified in Figure 5-6 are referred to as aircraft A, aircraft B and aircraft C respectively. All the simulated aircraft have medium accuracy sensors and medium accuracy initial conditions except for aircraft A, which has high accuracy initial conditions, based on the parameters given in Table 4-3. This configuration has been chosen to represent a typical civil air traffic environment, where most aircraft have similar sensor accuracy. Three case studies have been analysed when six, one and no GPS satellites are visible. In this thesis, it will be assumed that one ADS-B message from each member is received at least every 3-seconds, the variance of the pseudorange measurement noise is $(20\text{m})^2$ and the ADS-B range is 40 nautical miles, as discussed in Section 2.4.

When six GPS satellites are visible, all the simulated aircraft are assigned 2nd rank because GPS measurements provide very accurate navigation solutions. Therefore, ADS-B measurements are not used for navigation and the integrated INS/GPS/ADS-B system is effectively reduced to an integrated INS/GPS system.

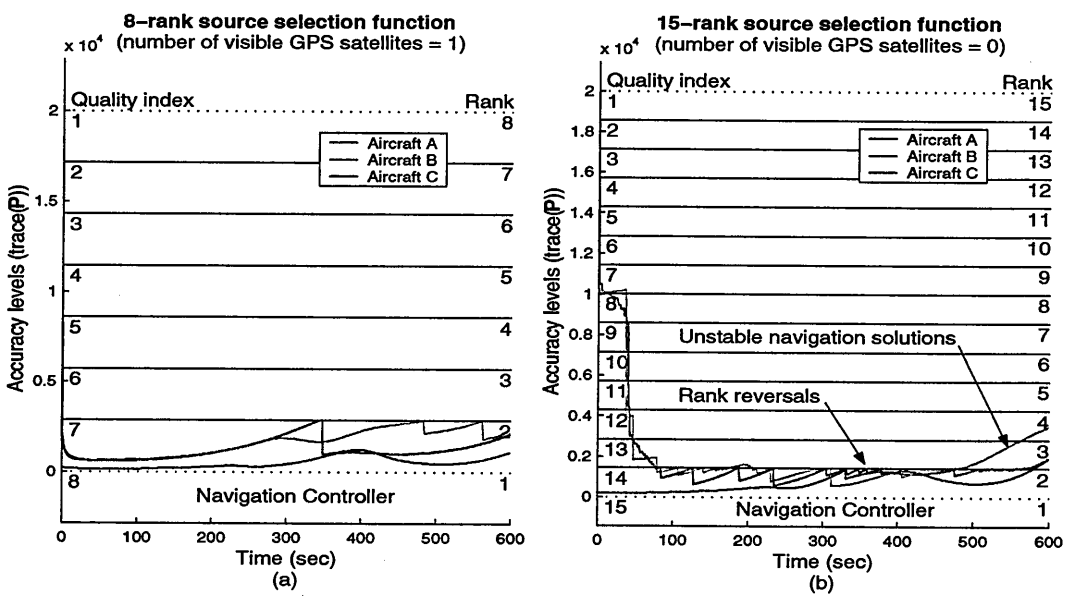


Figure 6-11 Organisation of the accuracy levels, quality indices and rank assignment in the 8-rank and 15-rank source selection functions

The case study when there is just one visible satellite is of special interest. Figure 6-11a shows the accuracy level profile for the 8-rank source selection function, which has been obtained from one Monte Carlo simulation. Aircraft A is not allowed to range to aircraft B or aircraft C because it is always assigned a lower rank (2^{nd} rank). However, the rms errors of aircraft B and aircraft C (Figure 6-12e to Figure 6-12l) indicate that the navigation systems of both aircraft have been improved by ranging to aircraft A.

Nevertheless, Figure 6-12g shows a constant divergence of the y-position rms errors of the aircraft B. This is the "*apparent*" divergence[10], which is caused by unmodelled error states. At the beginning of the filter cycles, aircraft B ranges to aircraft A and the visible GPS satellite. However, because aircraft A also ranged to the same visible GPS satellite before it was used in ranging by aircraft B, the errors of aircraft A will be correlated with the errors of aircraft B. These correlated errors are not modelled in the ownstate decentralised EKF model. Additionally the observability of aircraft B's y-position error state is poor, because the observability of the y-position error state depends only on the velocity in the x direction (Eq. 4-65), which is negligible for most of the time for the trajectory used by aircraft B. The effect of the unmodelled correlated errors, coupled with the poor observability, cause the y-position rms errors of aircraft B to diverge. However, aircraft C's errors do not diverge because it has superior observability and the EKF has a certain degree of natural stability. Figure 6-13 shows the rms errors of the simulated aircraft after the EKF has been tuned by deliberately adding $(15\text{m})^2$ to the ADS-B measurement noise covariance matrix $\mathbf{S}_k$. In this case study, the higher accuracy initial conditions of aircraft A has been used (distributed) to correct the navigation system of aircraft B and C, which have lower accuracy initial conditions. These results confirm that, based on the RELNAV principles, ADS-B sensors can be used to accurately correlate position data and to distribute the accuracy of superior members to inferior members, even when only one GPS satellite is visible.

The case study when there are no visible satellites is also of special interest. In operation, all members ensure that their errors are stable (bounded) by ranging to the NC based on the assumption that the navigation states of the NC are perfect (zero errors) and the system is stochastically stable. In the integrated INS/GPS/ADS-B

system, when there are no visible GPS satellites, a navigation community is established without the NC. Consequently, the errors of every member are allowed to grow as a function of time, unbounded. This situation can lead to an unstable community because of the rank reversal process described in Section 6.4.2. Figure 6-14 shows the rms errors for the simulation when there are no visible GPS satellites based on the 15-rank source selection function. It can be seen that, even without the NC, the accuracy of aircraft A is still distributed to aircraft B and C. However, as the errors of all the aircraft begin to grow, rank reversals (Figure 6-11b) occur resulting in instability of the navigation community as, shown in Figure 6-14. In the unstable community, although aircraft A is assigned a lower rank than aircraft B and C, measurements from aircraft A are not used to correct the navigation systems of these aircraft effectively. This is because navigation solutions of aircraft A are also unstable.

The above analysis suggests that a navigation community without an NC can become unstable. Additionally, the source selection function should be designed such that the accuracy level interval for the 2^{nd} rank includes only the accuracy level values for the steady-state solutions of the navigation system under consideration. This approach allows ADS-B measurements to be used to correct for the INS errors, which are not steady-state solutions, leading to faster convergence and minimising the risk of rank reversals and filter divergence. Finally, the theoretical rms errors of aircraft B and C for the integrated INS/GPS/ADS-B systems, based on the 8-rank source selection function, is shown in Figure 6-15 for different number of visible satellites. These results confirm that the integration of ADS-B measurements allows navigation systems to degrade gracefully when there are fewer than four GPS satellites.

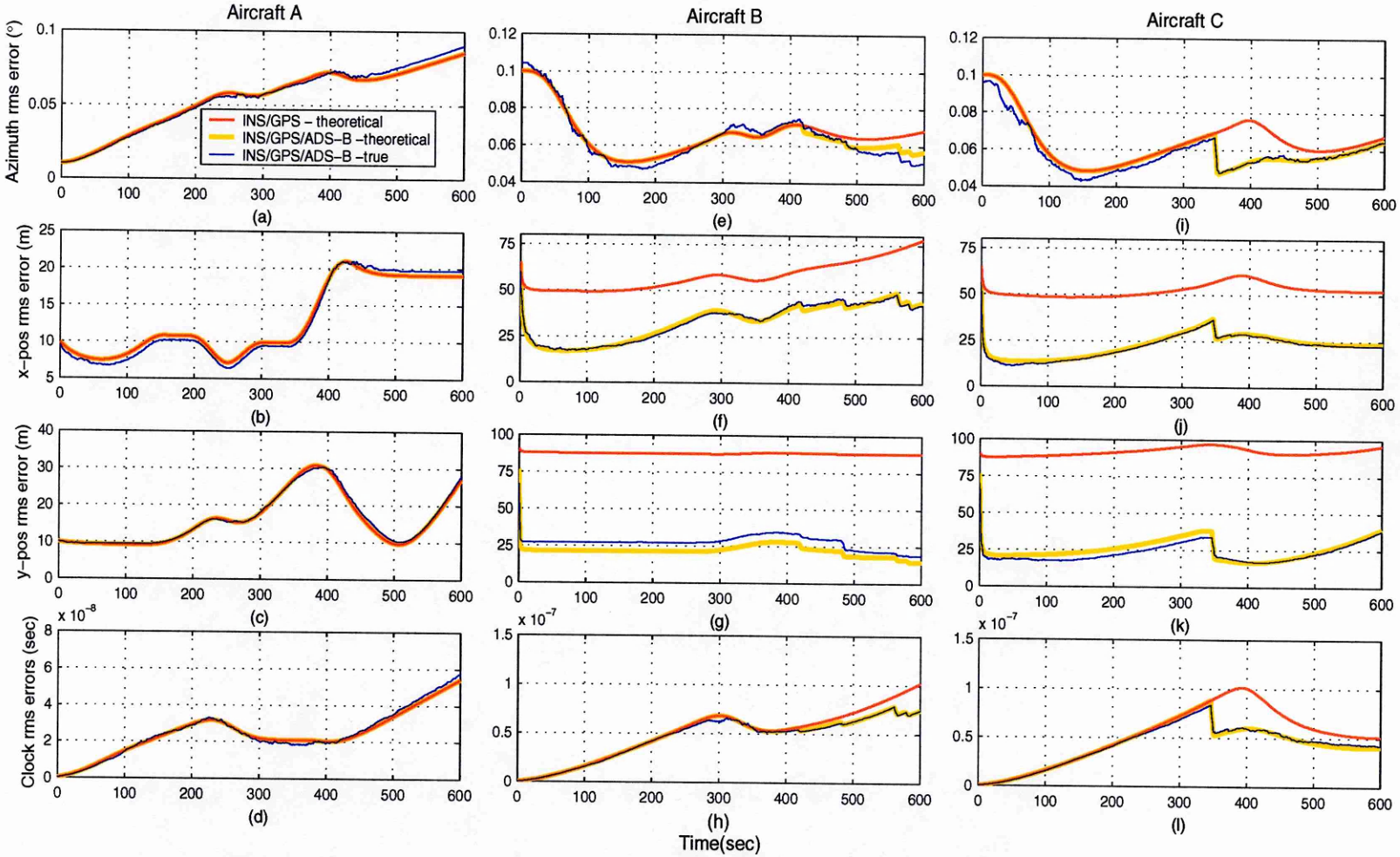


Figure 6-12 Comparison of the performance of the integrated INS/GPS system and INS/GPS/ADS-B system for one visible GPS satellite with the 8-rank source selection function before filter tuning

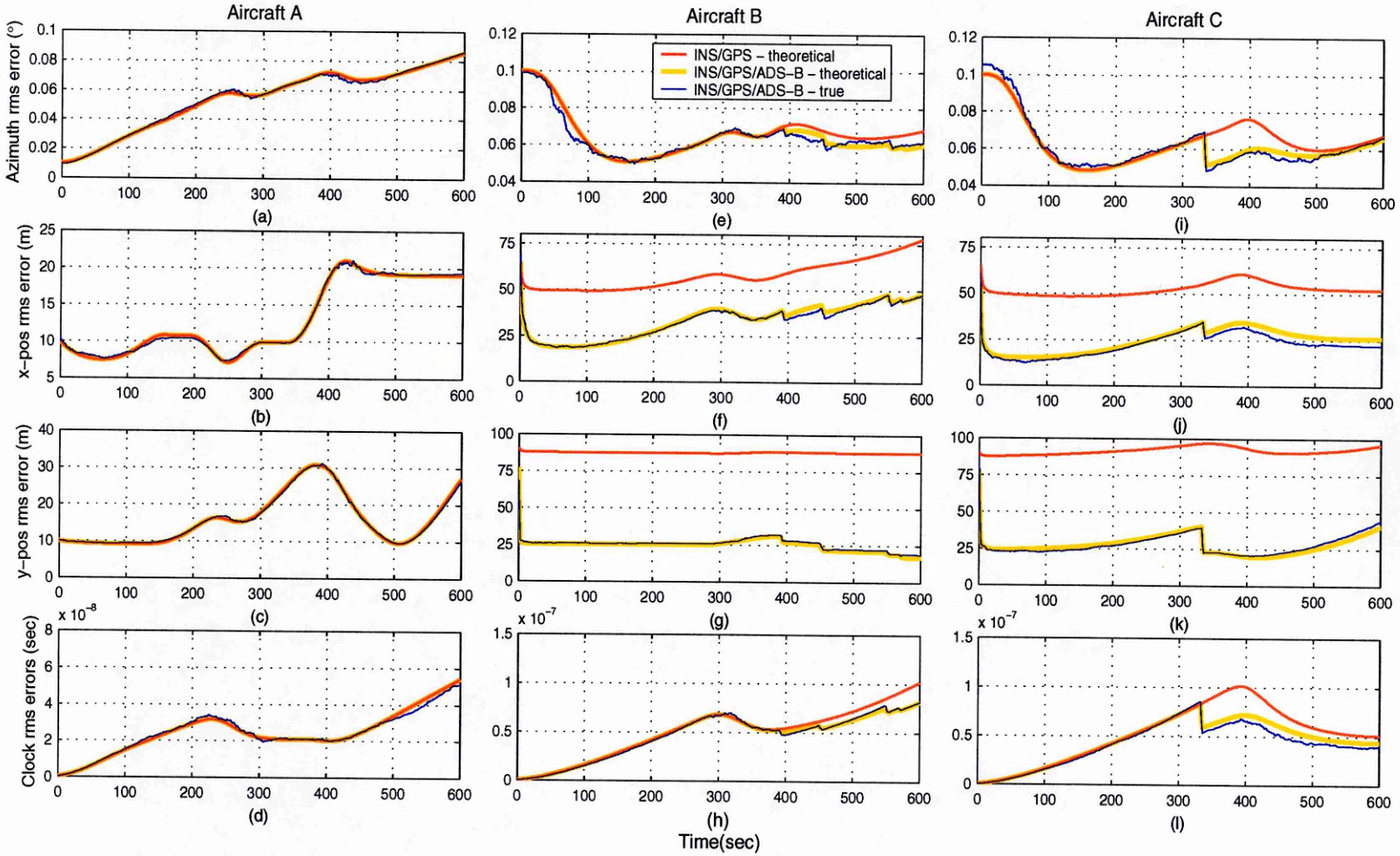


Figure 6-13 Comparison of the performance of the integrated INS/GPS system and INS/GPS/ADS-B system for one visible GPS satellite with the 8-rank source selection function after filter tuning

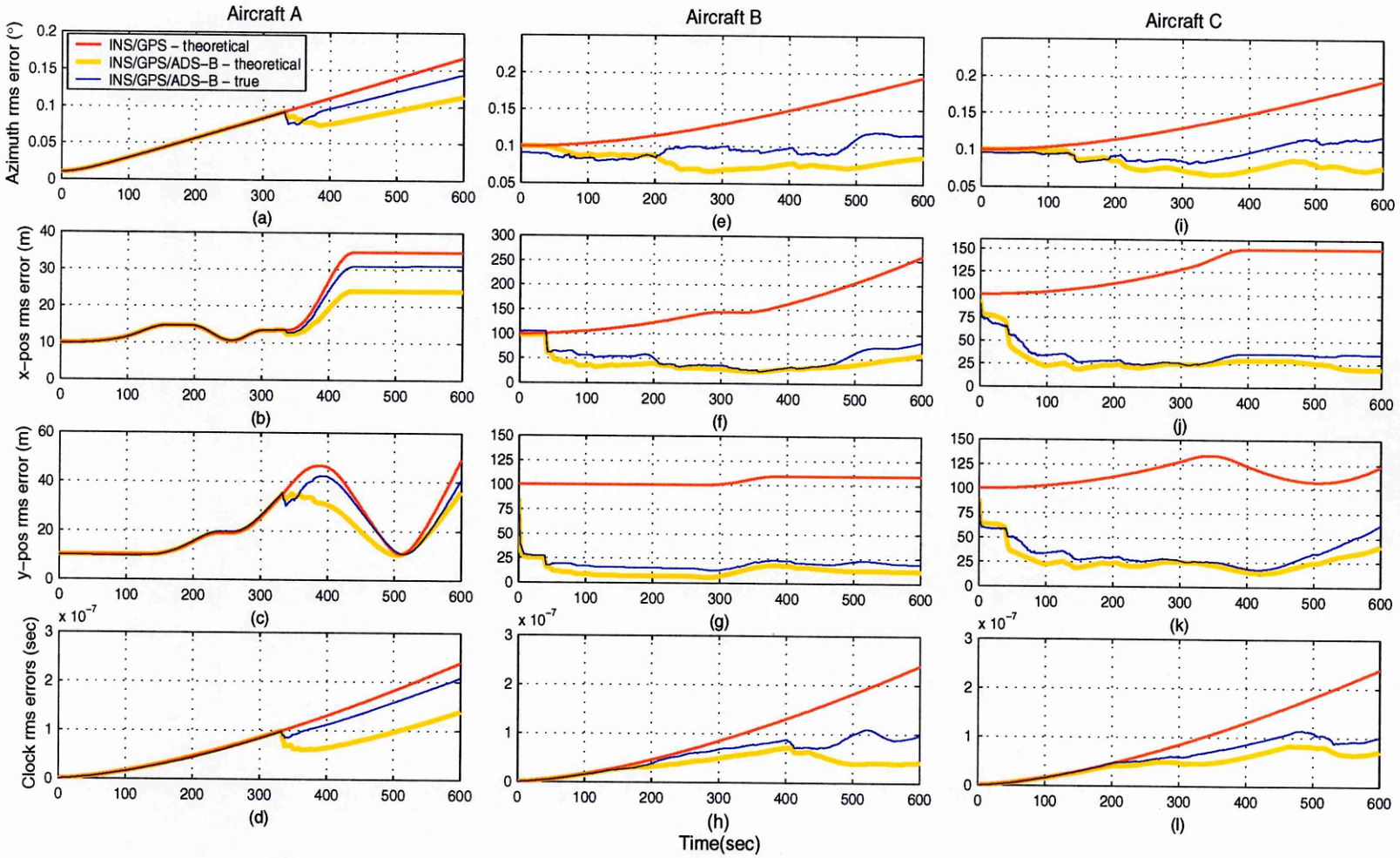


Figure 6-14 Comparison of the performance of the integrated INS/GPS system and INS/GPS/ADS-B system for no visible GPS satellites with the 15-rank source selection function

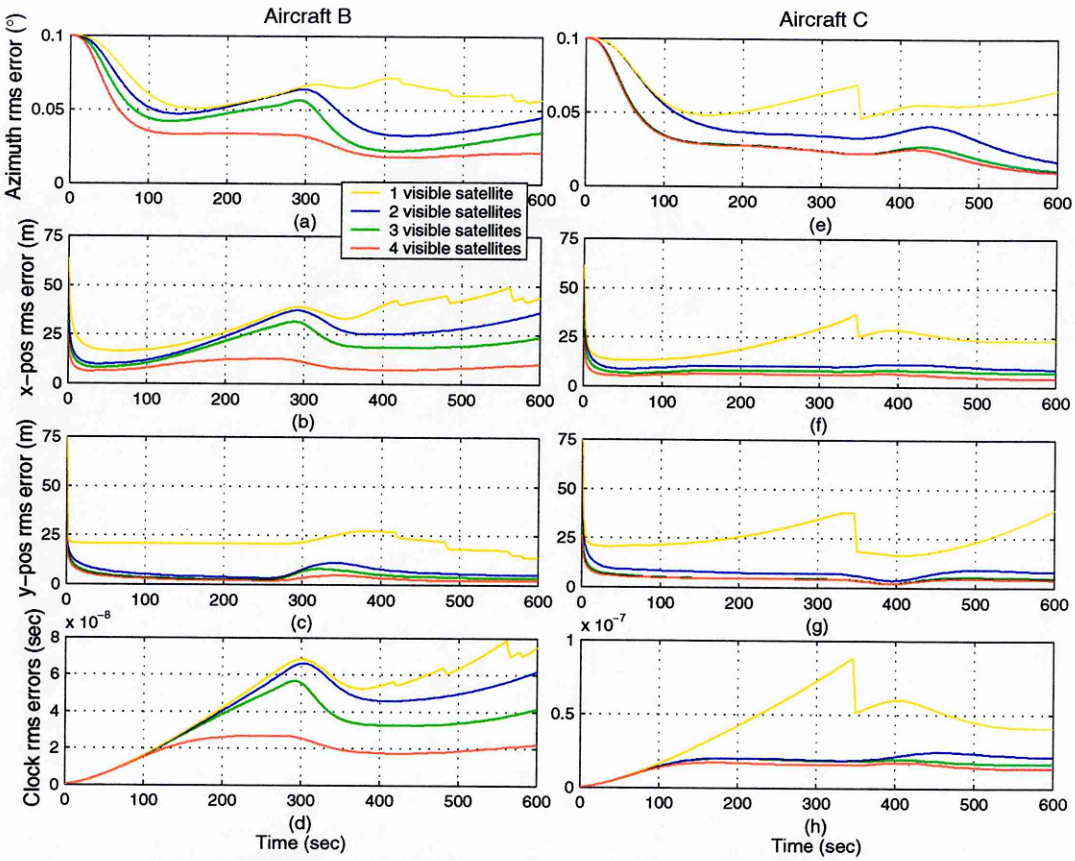


Figure 6-15 Comparison of the integrated INS/GPS/ADS-B system performance for different number of visible GPS satellites

In this chapter, it has been shown that the ADS-B can provide additional navigational sensors to form an integrated INS/GPS/ADS-B system based on the ownstate decentralised RELNAV EKF formulation. This integrated system offers the capability to accurately correlate position data and distribute navigation accuracy by processing ADS-B messages, allowing the performance of a navigation community to degrade gracefully when there are fewer than four GPS satellites. These capabilities are highly desirable in the civil ATM. However, the system stability is affected by two factors, namely, the modelling error (the truncation of the correlated error terms) and the observability. It has been shown that the problem of stability can be solved by applying carefully designed source selection functions, based on measurement resampling and rank assignment using quality indices, and filter tuning. Finally, it should be noted that the Kalman filter is merely a set of algorithms, which can be used to optimally estimate the state of a system. The success of applying the filter is predicated largely on the ability to understand the actual system and the system model under consideration.

CHAPTER 7 SENSOR FAILURE DETECTION AND ISOLATION

7.1 INTRODUCTION

Failure Detection and Isolation (FDI) is an essential function in aircraft navigation systems. Timely warning of sensor failures is critical to the safety of both civil and military aircraft operations. This chapter investigates a FDI algorithm, which monitors Kalman filter residuals based on the application of likelihood functions, and proposes the use of the chi-squared distribution to provide indices to be used in testing for failures. Monte Carlo simulations of an integrated INS/GPS system with an extended Kalman filter (EKF) are used to demonstrate and evaluate the performance of this algorithm.

7.2 MAXIMUM LIKELIHOOD ESTIMATOR

A maximum likelihood estimator is a method, which can be used to estimate the distribution parameters of a random variable from given data samples, for example, mean and variance of a Gaussian distribution. These parameters are determined from the calculation of a likelihood function, which is a quantitative measure of the goodness-of-fit of data samples to a model with a specific probability density function and a particular set of distribution parameters. For a Gaussian random variable X, the probability density function is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)\left[\frac{x-\mu}{\sigma}\right]^2} \qquad \text{Eq. 7-1}$$

The likelihood function for X, where μ and σ denotes the mean and variance of X, can be defined as the logarithm of the product of its probability density functions, based on N data samples, as follows:[52]

$$L(\mu,\sigma) = \ln\left[\prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)\left[\frac{x_i-\mu}{\sigma}\right]^2}\right]$$

$$= \ln\left[\left(\frac{1}{\sqrt{2\pi}}\right)^N \left(\frac{1}{\sigma}\right)^N e^{-(1/2\sigma^2)\sum_{i=1}^{N}(x_i-\mu)^2}\right]$$

$$= -N\ln(\sqrt{2\pi}) - N\ln\sigma - (1/2\sigma^2)\sum_{i=1}^{N}(x_i - \mu)^2 \qquad \text{Eq. 7-2}$$

In this case, the maximum likelihood estimator determines the values of μ and σ that maximise the likelihood function. This is achieved by differentiating Eq. 7-2 and solving for μ and σ when this function is zero. It can be shown that the mean and variance of *X*, which are estimated by the maximum likelihood estimator, are given by:[52]

$$\hat{\mu} = \frac{\left(\sum_{i=1}^{n} x_i\right)}{N} \qquad \text{Eq. 7-3}$$

$$\hat{\sigma}^2 = \frac{\left(\sum_{i=1}^{n} (x_i - \hat{\mu})^2\right)}{N} \qquad \text{Eq. 7-4}$$

If data samples $x_1, x_2, \ldots, x_n$ are actually samples of *X*, then $\sum_{i=1}^{N}(x_i - \mu)^2 = N\sigma^2$. Substituting this term in Eq. 7-2, the optimal likelihood function can be determined as follows:

$$L(\mu,\sigma)_{optimal} = -N\ln(\sqrt{2\pi}) - N\ln\sigma - (N/2)$$

$$= -\frac{N}{2}\left[\ln(2\pi) + \ln\sigma^2 + 1\right] \qquad \text{Eq. 7-5}$$

The likelihood function (Eq. 7-2) indicates the likelihood that data samples $x_1, x_2, \ldots, x_N$ are actually samples of a Gaussian random variable *X* with mean μ and variance σ^2. If the actual statistical properties (density function, mean or variance) of data samples vary from a presumed model, the computed likelihood function (based on these data samples) will diverge from the optimal values given by Eq. 7-5. Figure 7-1 illustrates examples of divergence of the likelihood function based on data samples, which have different statistical properties from a presumed model with $\mu = 0$ and $\sigma = 1$ for *N* = 10. This value of *N* has been chosen arbitrarily to demonstrate the concept. This property of the likelihood function has been used to monitor Kalman filter residuals[9] for sensor failure detection and isolation and will be discussed in the next section.

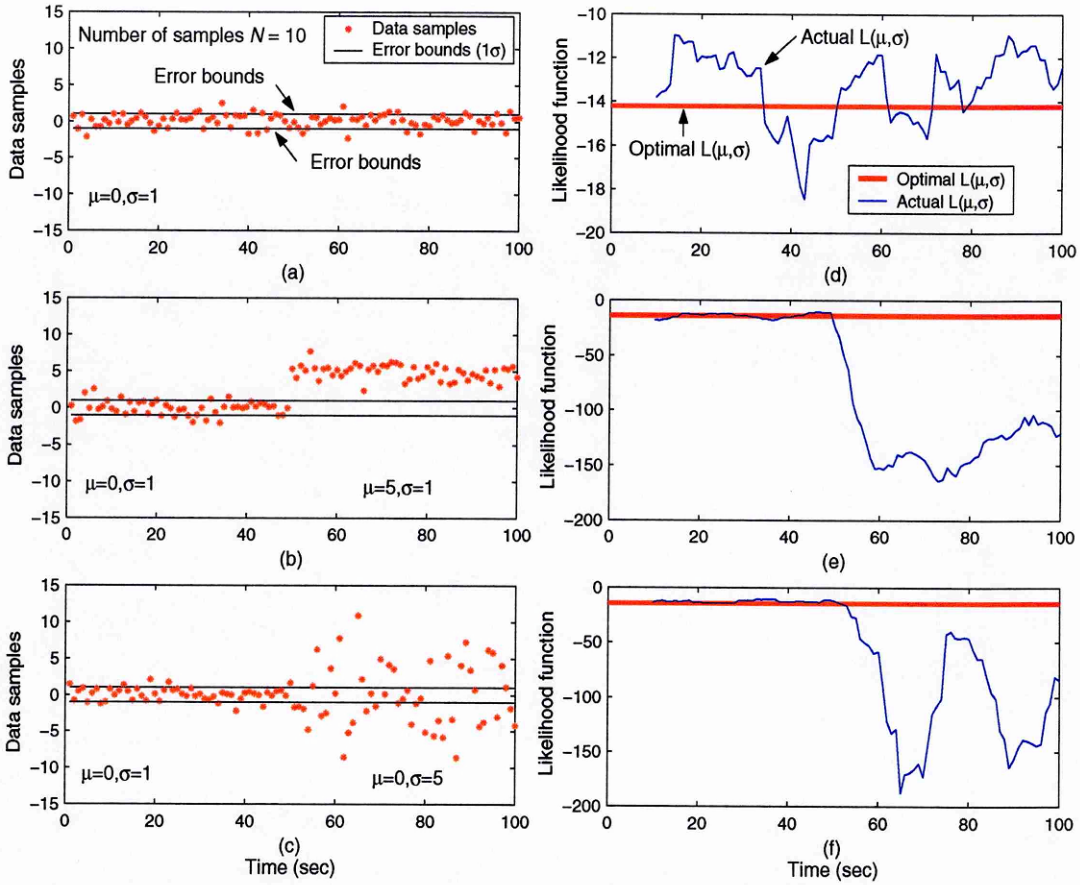


Figure 7-1 Divergence of the likelihood function

7.3 KALMAN FILTER RESIDUAL MONITORING

From the Kalman filter update equation (Eq. 3-29), the term $\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-$ is known as the *residual* (*innovation*) and is usually denoted as $\mathbf{r}_k$. Because the residuals describe the differences between actual and predicted measurements, they can be exploited to detect and isolate sensor failures. The following revision of Maybeck's method covers the derivation of the four statistical properties of a residual sequence (namely, the whiteness, the normality ("Gaussian-ness"), the zero-mean and the covariance $\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k$).[9] The derivation of the statistical properties of the residuals is conditional on the measurement at time *k*-1, i.e. $\mathbf{z}_{k-1}$. From Section 3.2, the overall system model of a linear stochastic system can be expressed as follows:

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k \mathbf{x}_k + \mathbf{w}_k \qquad \text{Eq. 7-6}$$

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \qquad \text{Eq. 7-7}$$

Noting that the Kalman filter update equations are given by $\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \mathbf{r}_k$ and $\hat{\mathbf{x}}_k^- = \mathbf{\Phi}_{k-1} \hat{\mathbf{x}}_{k-1}^+$, the residual $\mathbf{r}_k$ can be expressed as follows:

$$\begin{aligned}
\mathbf{r}_k &= \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^- \\
&= \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^- \\
&= \mathbf{H}_k (\mathbf{\Phi}_{k-1} \mathbf{x}_{k-1} + \mathbf{w}_{k-1}) + \mathbf{v}_k - \mathbf{H}_k \mathbf{\Phi}_{k-1} \hat{\mathbf{x}}_{k-1}^+ \\
&= \mathbf{H}_k \mathbf{\Phi}_{k-1} (\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}^+) + \mathbf{H}_k \mathbf{w}_{k-1} + \mathbf{v}_k \\
&= \mathbf{H}_k \mathbf{\Phi}_{k-1} \mathbf{e}_{k-1}^+ + \mathbf{H}_k \mathbf{w}_{k-1} + \mathbf{v}_k
\end{aligned} \qquad \text{Eq. 7-8}$$

Because $\mathbf{e}_{k-1}^+$, $\mathbf{w}_{k-1}$ and $\mathbf{v}_k$ are independent of the measurement $\mathbf{z}_{k-1}$, $\mathbf{r}_k$ is independent of $\mathbf{z}_{k-1}$. Based on the assumption that a linear combination of jointly Gaussian random variables is Gaussian, a residual sequence is *Gaussian*. Recalling the relationship between $\mathbf{r}_k$ and $\mathbf{z}_k$, $\mathbf{z}_{k-1}$ can be expressed in terms of $\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_{k-1}$ as follows:

$$\begin{aligned}
\mathbf{z}_{k-1} &= \mathbf{r}_{k-1} + \mathbf{H}_{k-1} \hat{\mathbf{x}}_{k-1}^- \\
&= \mathbf{r}_{k-1} + \mathbf{H}_{k-1} [\mathbf{\Phi}_{k-2} \hat{\mathbf{x}}_{k-2}^+] \\
&= \mathbf{r}_{k-1} + \mathbf{H}_{k-1} \{\mathbf{\Phi}_{k-2} [\hat{\mathbf{x}}_{k-2}^- + \mathbf{K}_{k-2} \mathbf{r}_{k-2}]\} \\
&= \mathbf{r}_{k-1} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} \mathbf{K}_{k-2} \mathbf{r}_{k-2} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} \hat{\mathbf{x}}_{k-2}^- \\
&= \mathbf{r}_{k-1} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} \mathbf{K}_{k-2} \mathbf{r}_{k-2} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} [\mathbf{\Phi}_{k-3} \hat{\mathbf{x}}_{k-3}^+] \\
&= \mathbf{r}_{k-1} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} \mathbf{K}_{k-2} \mathbf{r}_{k-2} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} \{\mathbf{\Phi}_{k-3} [\hat{\mathbf{x}}_{k-3}^- + \mathbf{K}_{k-3} \mathbf{r}_{k-3}]\} \\
&= \mathbf{r}_{k-1} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} \mathbf{K}_{k-2} \mathbf{r}_{k-2} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} \mathbf{\Phi}_{k-3} \mathbf{K}_{k-3} \mathbf{r}_{k-3} + \mathbf{H}_{k-1} \mathbf{\Phi}_{k-2} \mathbf{\Phi}_{k-3} \hat{\mathbf{x}}_{k-3}^-
\end{aligned}$$

Eq. 7-9

By substitution, Eq. 7-9 can be expanded to include the last term, $\mathbf{r}_1$. Because $\mathbf{r}_k$ is independent of $\mathbf{z}_{k-1}$, which is a linear function of $\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_{k-1}$, $\mathbf{r}_k$ is also independent of all previous $\mathbf{r}_k$'s, i.e. $\mathbf{r}_k$ is a *white* sequence. The *mean* and *covariance* of $\mathbf{r}_k$ are determined as follows:

$$\mathrm{E}[\mathbf{r}_k] = \mathrm{E}[\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-]$$

$$= \mathrm{E}[\mathbf{H}_k\mathbf{x} + \mathbf{v}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^-]$$

$$= \mathrm{E}[\mathbf{H}_k\mathbf{e}_k^- + \mathbf{v}_k] = \mathbf{0} \qquad \text{Eq. 7-10}$$

$$\mathrm{E}[\mathbf{r}_k\mathbf{r}_k^{\mathrm{T}}] = \mathrm{E}[\{\mathbf{H}_k\mathbf{e}_k^- + \mathbf{v}_k\}\{\mathbf{H}_k\mathbf{e}_k^- + \mathbf{v}_k\}^{\mathrm{T}}]$$

$$= \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k \qquad \text{Eq. 7-11}$$

These statistical properties are an ideal characterisation of a residual sequence, based on the assumption that the implemented filter model represents a real-world system accurately. If the implemented system model does not depict a real-world system accurately, correct characterisation of a residual sequence, based on these four statistical properties, can not be guaranteed. In operation, if a residual sequence exhibits correct properties up to a certain point in time and, subsequently, it exhibits different properties from those predicted by the filter, it can be assumed that an event has occurred, which causes the implemented model to deviate from a real-world system. If an event is triggered by a sensor failure, it can be detected by monitoring the statistical properties of the residual sequence. The measurement equation, described by Daly[50], which includes sensor failure models, can be modified to form a "*generalised*" measurement model, and expressed as follows:

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k + \boldsymbol{\varepsilon}_k \qquad \text{Eq. 7-12}$$

where $\boldsymbol{\varepsilon}_k$ represents sensor failures, modelled in terms of the mean or variance of the sensor noise.

The approach, to monitoring the statistical properties of the residuals in this thesis, is based on the application of the likelihood function, described in Section 7.2. If a residual sequence can be assumed to be a Gaussian white sequence, the mean and variance of the residual from the i^{th} sensor at time t_k, $r_i(t_k)$, can be determined as follows:[9]

$$\mu_i(t_k) = 0, \text{for all } t_k \text{ and all sensors} \qquad \text{Eq. 7-13}$$

$$\sigma_i^2(t_k) \text{ is given by the } i^{\text{th}} \text{ diagonal term of } \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k \qquad \text{Eq. 7-14}$$

By applying the same method used to derive Eq. 7-2, the likelihood function, for detecting a failure of the i^{th} sensor from N most recent measurements, can be determined as follows:

$$L_{N_i}(t_k) = \ln\left[\prod_{j=k-N+1}^{k} \frac{1}{\sigma_i(t_j)\sqrt{2\pi}} e^{-(1/2)\left[\frac{r_i(t_j)-\mu_i(t_j)}{\sigma_i(t_j)}\right]^2}\right]$$

$$= \sum_{j=k-N+1}^{k} \ln\left[\frac{1}{\sigma_i(t_j)\sqrt{2\pi}} e^{-(1/2)\left[\frac{r_i(t_j)}{\sigma_i(t_j)}\right]^2}\right]$$

$$= -\frac{N}{2}\ln(2\pi) - \sum_{j=k-N+1}^{k} \ln\left[\sigma_i(t_j)\right] - \frac{1}{2}\sum_{j=k-N+1}^{k} \frac{r_i^2(t_j)}{\sigma_i^2(t_j)}$$

$$= -\frac{1}{2}\left(N\ln(2\pi) + \sum_{j=k-N+1}^{k} \ln\left[\sigma_i^2(t_j)\right]\right) - \frac{1}{2}\sum_{j=k-N+1}^{k} \frac{r_i^2(t_j)}{\sigma_i^2(t_j)} \qquad \text{Eq. 7-15}$$

where the term $-\frac{1}{2}\left(N\ln(2\pi) + \sum_{j=k-N+1}^{k} \ln\left[\sigma_i^2(t_j)\right]\right)$ is $c_i(t_k)$.

Eq. 7-15 is the likelihood function suggested by Maybeck[9], where $c_i(t_k)$ is a slowly varying term independent of the actual measurements; therefore, it has no direct use for failure detection. $\sigma_i^2(t_j)$ is the predicted variance of the residuals assuming that no failures have occurred. The optimal likelihood function can be determined from the predicted distribution parameters, by applying the same approach used to derive Eq. 7-5, as follows:

$$L_{N_i}(t_k)_{optimal} = -\frac{1}{2}\left(N\ln(2\pi) + \sum_{j=k-N+1}^{k} \ln\left[\sigma_i^2(t_j)\right] + N\right) \qquad \text{Eq. 7-16}$$

The design parameter N is the number of the most recent residual samples to be used in the calculations of likelihood function. In effect, this approach is a moving window based on the last N samples. For small values of N detection algorithms can be too sensitive whereas large values of N may decrease the sensitivity of the algorithm to detect true failures. Maybeck suggests a number between 5 to 20 as a reasonable choice of N.

The likelihood function (Eq. 7-15) measures how well a residual sequence agrees with the statistical description, predicted by the filter. If the residuals have correct statistical properties, then the likelihood function (computed from the actual

residual sequence) will be close to the optimal values (Eq. 7-16). However, when a failure occurs (modelled as an *increase* in the mean or variance of the sensor noise), the residuals will exhibit different statistical properties and the likelihood function will become more *negative*. A failure can be declared when the likelihood function exceeds some predetermined threshold.

In almost all real-time applications, Kalman filters are based on simplified (sub-optimal) models, where some state variables are omitted due to practical limitations. If the real-world system is not represented accurately by the sub-optimal filter, then the statistical characterisation of the residuals may be imprecise, leading to degraded performance of the detection algorithms. In practice, Kalman filter design involves sensitivity analysis and tuning of the filters, which can be viewed as a process of attaining the best fit of the implemented model to the real-world system. In general, accurate characterisation of the residuals is maintained and failure detection can still be performed, based on residual monitoring. The effects of divergence of the statistical properties of the residuals, caused by gross errors in the model (such as sensor failures), are usually more severe than those exhibited by a properly designed sub-optimal filter.

Figure 7-2a to Figure 7-2d compare the actual (derived from Monte Carlo simulations) and the predicted mean and standard deviation of the residuals of the integrated INS/GPS system with an EKF. In these simulations, four GPS satellites (SV4, SV7, SV10 and SV17) are visible and the first satellite (SV4) fails after 150 sec. The failures are modelled as the addition of sensor noise with a mean of 50m and a standard deviation of 4m to all pseudorange measurements from SV4 after 150sec, in combination with the standard zero-mean measurement noise with 33.3m standard deviation.

GPS sensor failures can be modelled as an addition of pseudorange measurement noise with a standard deviation of 4m and a mean greater than 5m.[51] However, it should be noted that the actual requirements for GPS integrity monitoring is very stringent[59]; therefore the simulations in this chapter should be viewed as a benchmark for evaluating the FDI algorithm, being investigated in this thesis.

Figure 7-2e to Figure 7-2h compare the mean of the actual values of the likelihood functions with the optimal values (where N = 15 samples). In this demonstration, the failure detection threshold could be set arbitrarily at –85.

Two observations can be made from the analysis of the results shown in Figure 7-2. Firstly, although, it has been shown that a residual sequence is a white sequence, where r_k is uncorrelated with $r_{k-1}, r_{k-2}, \ldots, r_1$, there is some degree of correlation (noticeably after 150 sec) between the residuals of the failed and non-failed satellites. This is to be expected because the predicted error $\hat{\mathbf{x}}_k^-$ in the residual equation $\mathbf{r}_k = \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-$ is the same for all residual sequences, i.e. the residuals between different sensors are correlated. Even so, the performance of the algorithm is unaffected and the failed satellite (SV4) can be clearly identified. Secondly, the likelihood function of the failed satellite initially overshoots and then settles to a steady-state condition, which reflects the filter response to external disturbance (sensor failures).

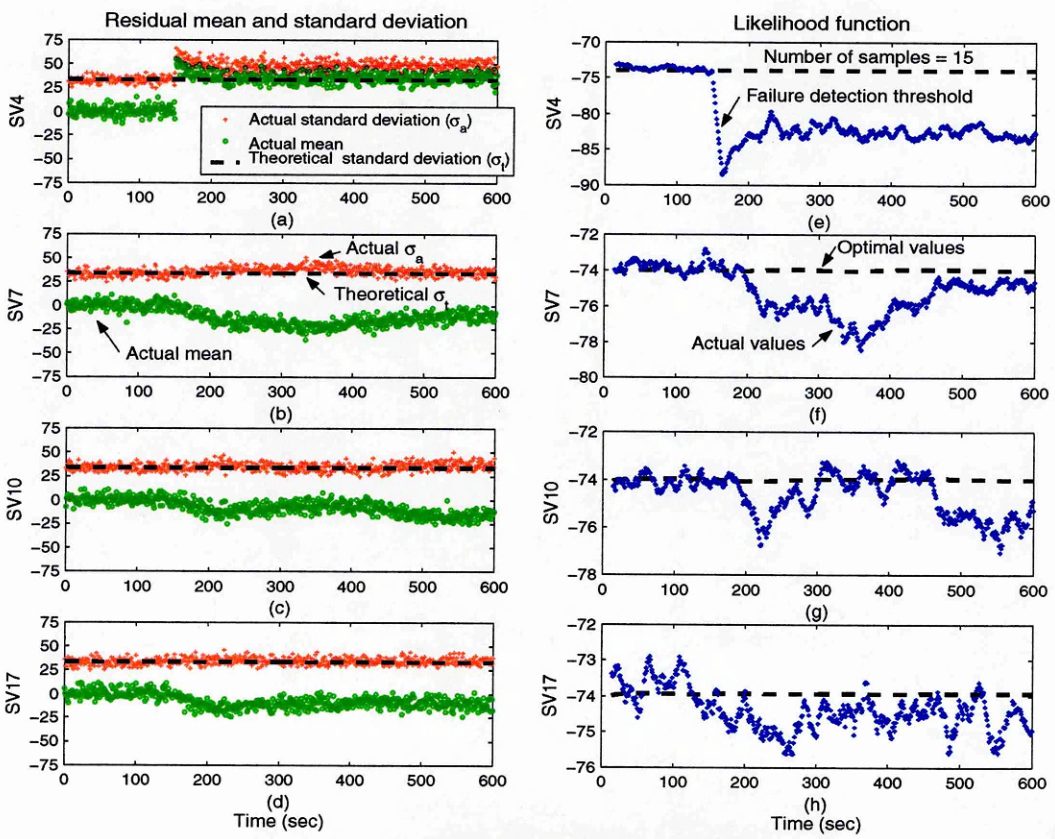


Figure 7-2 Likelihood function and Kalman filter residual monitoring for GPS failure detection. The y-axis of figures (a)-(d) is the pseudorange measurement error (m). The y-axis of figures (e)-(h) is the likelihood function value.

One difficulty associated with the detection of failure by residual monitoring, and directly applying the likelihood function, is in determining a meaningful detection threshold. The following section describes the chi-squared distribution, which can be used to select a meaningful detection threshold.

7.4 CHI-SQUARED DISTRIBUTION

A chi-squared random variable χ^2_γ is defined as a random variable, whose probability density function is given by:[52]

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{(\alpha-1)} e^{-(x/\beta)} \qquad \text{Eq. 7-17}$$

where the gamma function $\Gamma(\alpha) = \int_0^\infty z^{(\alpha-1)} e^{-z} dz$, $x > 0$, $\beta = 2$, $\alpha = \gamma/2$ and γ denotes the number of degrees of freedom. The mean and variance of a chi-squared random variable are given by γ and 2γ respectively. The cumulative distribution (probability) of a chi-squared random variable for a specific degree of freedom can be determined directly from tables[53] in most statistical textbooks or numerically[38] based on the incomplete gamma function. In this thesis, the chi-squared distribution is determined numerically. $P(T,\gamma)$ is defined as the probability that the observed chi-squared variable with γ degrees of freedom for a correct model should be less than a value T, i.e. $P(T,\gamma) \equiv \text{Prob}[\chi^2_\gamma < T]$, where $Q(T,\gamma) = 1 - P(T,\gamma) \equiv \text{Prob}[\chi^2_\gamma \geq T]$ is the complement of $P(T,\gamma)$. Figure 7-3 illustrates the cumulative distribution and the probability of a chi-squared random variable with 15 degrees of freedom for $T = 20$.

Let $X_1, X_2, \ldots, X_N$ be a random sample from a Gaussian distribution with mean μ_j and variance σ_j^2, it can be shown that:[38,52]

$$\chi^2_{\gamma=N} \equiv \sum_{j=1}^{N} \frac{(x_j - \mu_j)^2}{\sigma_j^2} \qquad \text{Eq. 7-18}$$

has a chi-squared distribution with N degrees of freedom, where the value of a chi-squared $\chi^2_{\gamma=N}$ can be computed from the data sample $x_1, x_2, \ldots, x_N$. The computed chi-

squared and the degrees of freedom (the number of available data samples N) can be used to determine the corresponding cumulative chi-squared probabilities. These probabilities indicate the likelihood that the data samples actually fit the statistical description (presumed model) of a Gaussian distribution with mean μ_j and variance σ_j^2.

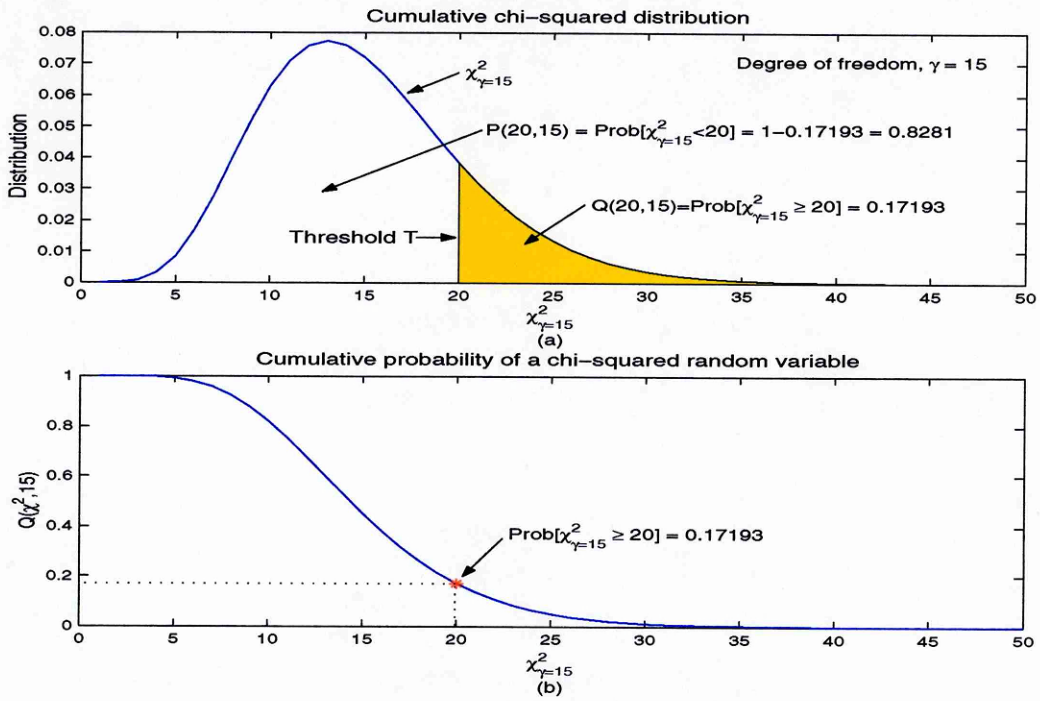


Figure 7-3 Cumulative chi-squared distribution

Table 7-1 Demonstration of testing of goodness-of-fit using chi-squared distribution

	Perfect model	**Incorrect model**
Mean μ	0	50
Variance σ^2	30^2	5^2
Data sample X_j ($N=\gamma=15$)	41.151, 7.3674, 3.5653, 11.541, -2.1046, -17.349, 14.079, 38.991, 49.044, 24.218, -30.824, 38.835, 0.44791, 6.5612, -21.084	58.566, 39.606, 50.565, 44.568, 42.208, 53.187, 47.977, 47.983, 50.42, 47.823, 47.187, 54.39, 45.927, 48.708, 52.466
Computed chi-squared values $\chi^2_{\gamma=15} = \sum_{j=1}^{15} \frac{(X_j-0)^2}{30^2}$	10.95	40.005
Cumulative probability $P[\chi^2_{\gamma=15} \geq T]$	0.11044	6.9524e-011

The numerical example given in Table 3-2 illustrates how a chi-squared distribution can be used to test the goodness-of-fit of data samples to a Gaussian distribution model. The chi-squared value T = 10.95 was calculated from the data samples, which have been generated from the *true* model (zero-mean Gaussian distribution with variance 30^2). The probability of obtaining chi-squared values of 10.95 or greater from data samples of the perfect model is given by $P[\chi^2_{\gamma=15} \geq 10.95] = 0.11044$. The chi-squared value of the "*incorrect*" data sample, which have been generated from the "*incorrect*" model (with distribution parameters $\mu = 50$ and $\sigma^2 = 5$), is 40.005. The probability of obtaining chi-squared values of 40.005 or greater from data samples of the true model is given by $P[\chi^2_{\gamma=15} \geq 40.005] = 6.9524 \times 10^{-11}$.

Consequently, it can be stated that it is unlikely that the "*wrong*" data samples are data samples of the true model, i.e. mismodeling has occurred. In sensor failure detection, discrepancies between the statistical properties of the presumed system model and the actual physical system may be caused by sensor failures, leading to chi-squared values with very low probability. In principle, when the predicted probability is below a certain threshold (or the computed chi-squared is greater than or equal to a specific threshold), a failure can be declared. The next section outlines the previous developments of FDI algorithms, which are based on the application of the chi-squared distribution, and derives another approach of a chi-squared-based FDI algorithm based on the likelihood function.

7.5 CHI-SQUARED TEST FOR SENSOR FAILURE DETECTION AND IDENTIFICATION

7.5.1 Previous developments of chi-squared-based FDI algorithms

In Section 7.4, it has been shown that a chi-squared $\chi^2_{\gamma=N}$ can be calculated from given data samples, based on the assumption that the data samples have been obtained from an assumed Gaussian model. The computed chi-squared (or the probability Q) can be used to indicate the agreement between the data samples and the presumed model. In order to detect a sensor failure, based on a hypothesis testing, two hypothesises (the *alternative hypothesis* H_1 and the *null hypothesis* H_0) can be stated as follows:

H_1: a failure has occurred ($\chi^2_{\gamma=N}(t_k) \geq \chi^2_{threshold}$) Eq. 7-19

H_0: no failure has occurred ($\chi^2_{\gamma=N}(t_k) < \chi^2_{threshold}$) Eq. 7-20

The detection threshold $\chi^2_{threshold}$ can be determined by specifying a constant probability of committing *Type I errors*[52], which are defined as situations in which the event H_0 has been rejected, although it is true. This is equivalent to declaring a failure, when no failure has actually occurred. In other words, the probability of committing Type I errors is the false alarm rate or probability of the failure detection algorithms. The primary advantage of constructing FDI algorithms, which are based on a chi-squared test statistic, is that a meaningful threshold can be selected by specifying a desired false alarm rate.

In the previous developments of chi-squared-based FDI algorithms, a chi-squared test statistic has been computed from a residual $\boldsymbol{\beta}_k$, which measure differences between the actual states of a system and the predicted states, based on a presumed model. This residual can be either the residual $\mathbf{r}_k$ of a single Kalman filter[55,57] or the residual, which has been generated, based specifically on a priori knowledge of a system[54,56] (for example, state estimates of a bank of auxiliary Kalman filters). In either case, a chi-squared test statistic is computed as follows:

$$\lambda = \boldsymbol{\beta}_k^{\mathrm{T}} \mathbf{B}_k^{-1} \boldsymbol{\beta}_k \qquad \text{Eq. 7-21}$$

where λ is a chi-squared test statistic with N degrees of freedom. N is the number of elements of the residual vector $\boldsymbol{\beta}_k$ and $\mathbf{B}_k$ is the error covariance matrix of the residual, which is defined as follows:

$$\mathbf{B}_k = \mathrm{E}[\boldsymbol{\beta}_k \boldsymbol{\beta}_k^{\mathrm{T}}] \qquad \text{Eq. 7-22}$$

These methods are a "*snap-shot*" approach[58] because a residual vector at a particular time t_k is used to compute a test statistic. One disadvantage of using a complete residual vector $\boldsymbol{\beta}_k$ to determine the test statistic according to Eq. 7-21, is that a failure of a particular sensor can not be identified. The next section discusses another approach to a chi-squared-based FDI algorithm, which affords the ability to identify a failed sensor, based on the application of the likelihood function and a single Kalman filter.

7.5.2 Application of likelihood functions to chi-squared-based FDI algorithms

By comparing the likelihood function (Eq. 7-15) with Eq. 7-18, and because the Kalman filter residuals are Gaussian, it can be seen that the last term of the likelihood function is a chi-squared random variable with N degrees of freedom. Moreover, $c_i(t_k)$ has no useful information for failure detection and, therefore, a test statistic for hypothesis testing of sensor failures (described in Section 7.5.1) can be formed using only the last term of Eq. 7-15. A chi-squared test statistic for monitoring residuals of a single Kalman filter can be expressed as follows:

$$\chi^2_{\gamma=N}(t_k) = \sum_{j=k-N+1}^{k} \frac{r_i^2(t_j)}{\sigma_i^2(t_j)} \qquad \text{Eq. 7-23}$$

The FDI algorithm, developed in this chapter, is illustrated schematically in Figure 7-4.

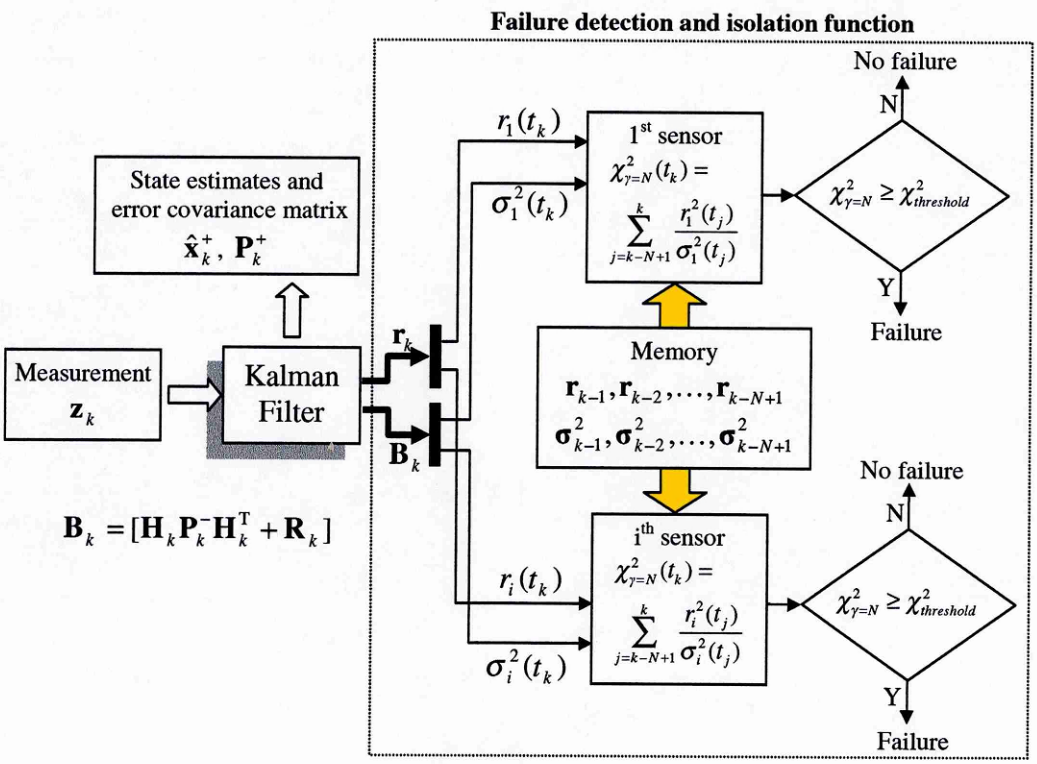


Figure 7-4 Chi-squared-based FDI algorithm based on likelihood function

To demonstrate the concept, the algorithm has been implemented to detect and identify failures of GPS satellites based on the case study used in Section 7.3. The algorithm has been set to monitor 15 consecutive residual samples, i.e. N = 15, with a constant 0.001 false alarm probability. Therefore, the threshold chi-squared can be determined from $Q[\chi^2_{threshold}, 15] = \text{Prob}[\chi^2_{\psi=15} \geq \chi^2_{threshold}] = 0.001$, which is 37.731. Figure 7-5 shows the mean of the chi-squared values when there are no GPS failures. The mean of the chi-squared test statistic is centred on the true mean, which is the number of degrees of freedom (15). The actual false alarm rate was determined to be 0.00097104 (which is close to 0.001). The comparison of the theoretical and the actual distribution of the chi-squared test statistic, when there are no failures, is also shown in Figure 7-6 for validation.

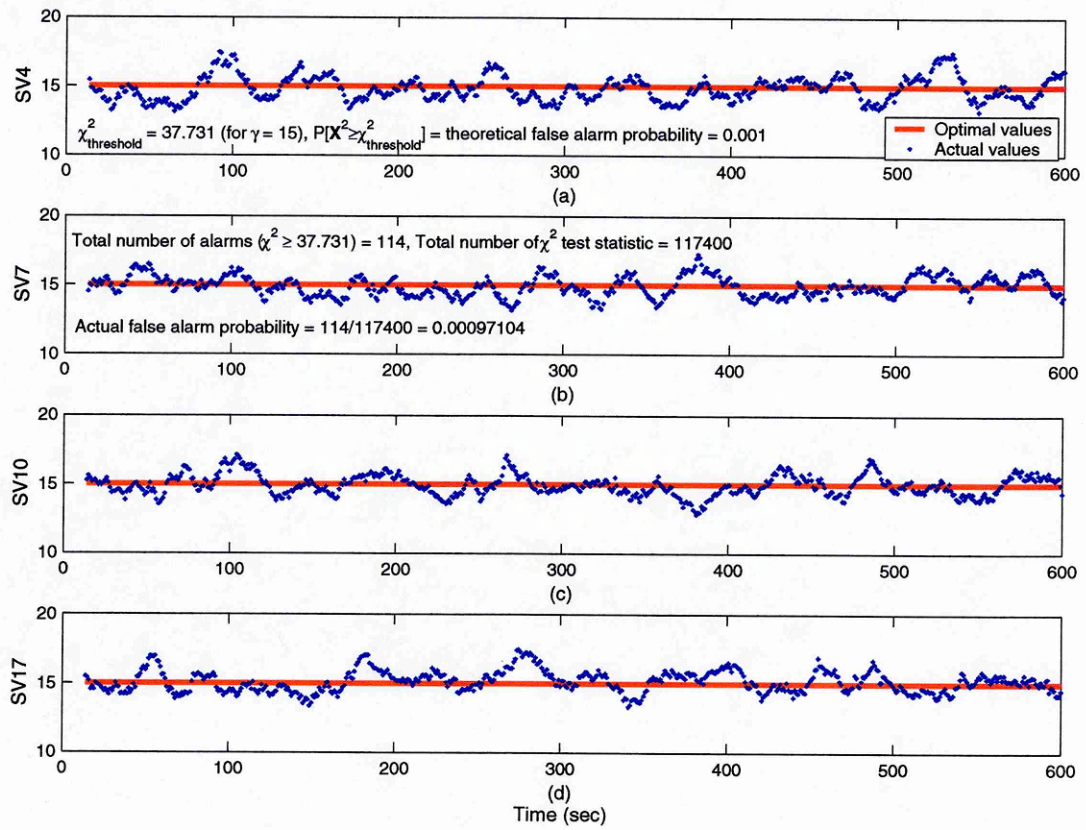


Figure 7-5 Mean of the chi-squared test statistic with no GPS failure

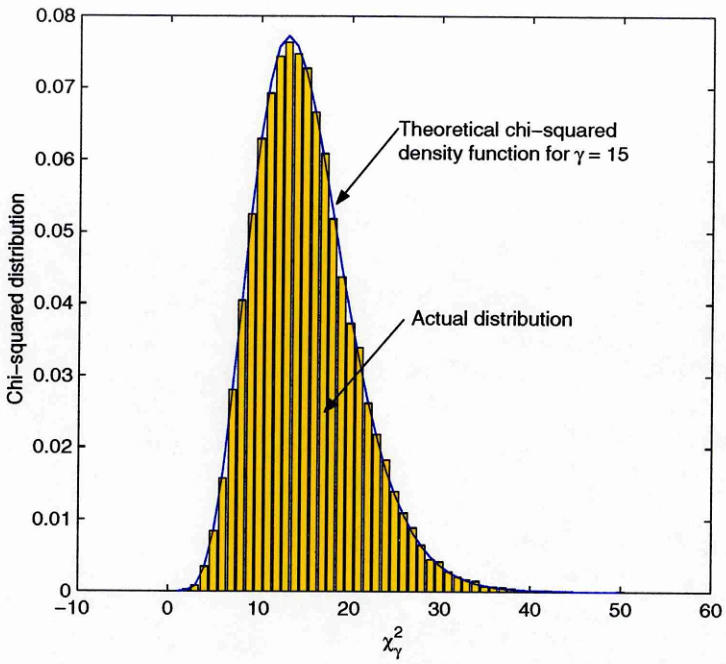


Figure 7-6 Distribution of the chi-squared test statistic with no GPS failure

Figure 7-7 shows the mean of the chi-squared test statistic for the simulations when GPS satellite (SV4) fails after 150 sec. The overshoot and the steady-state conditions can still be seen in the profile of the chi-squared test statistic of the failed satellite. Under such circumstances, failures can be detected and isolated whenever the test statistics are greater than the threshold value (37.731). The comparison of the actual and the theoretical distribution of the chi-squared test statistic is shown in Figure 7-8.

One drawback of this failure detection algorithm is that it is insensitive to sensor failures of small magnitude (relatively small mean or variance $\boldsymbol{\varepsilon}_k$ in Eq. 7-12). This is attributed to the inherent inability of any test statistic, including the chi-squared test, to distinguish between data samples that have very similar distribution parameters (mean and variance), as illustrated in Figure 7-9, where failures can not be detected for values of $\mu = 5$ but are clearly detectable for $\mu = 50$.

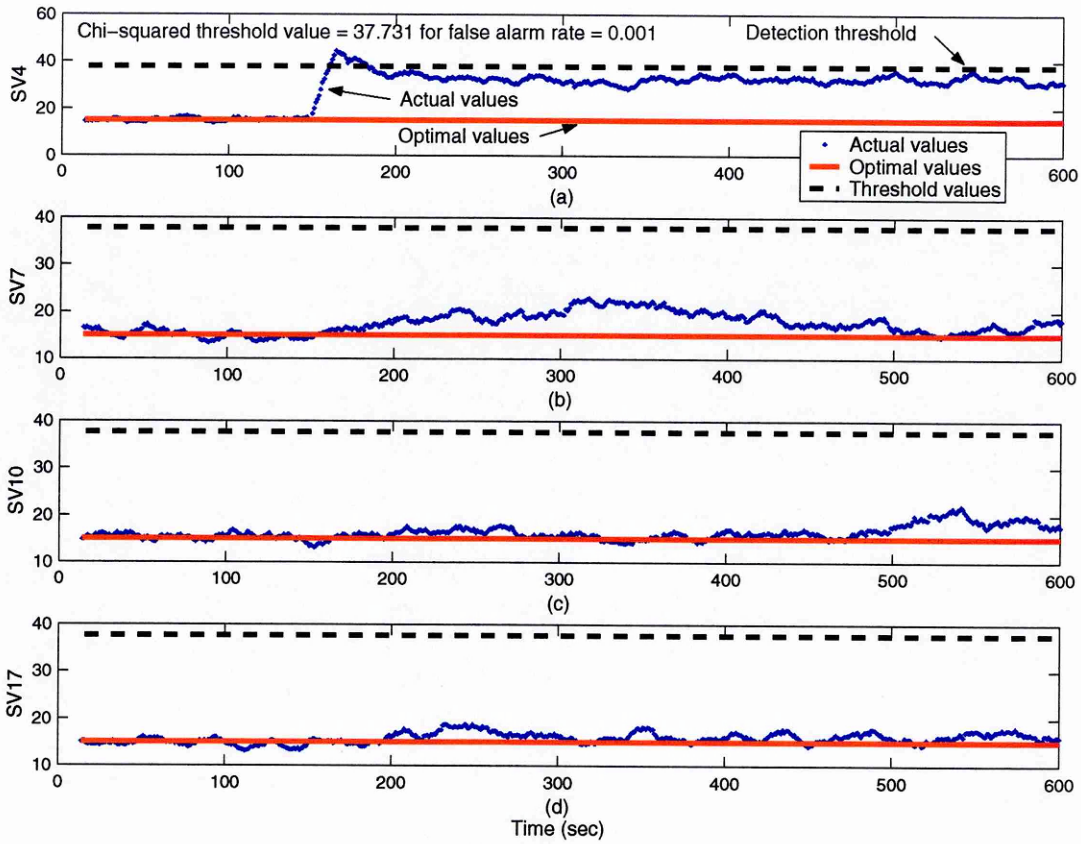


Figure 7-7 Mean of the chi-squared test-statistic with GPS (SV4) failures

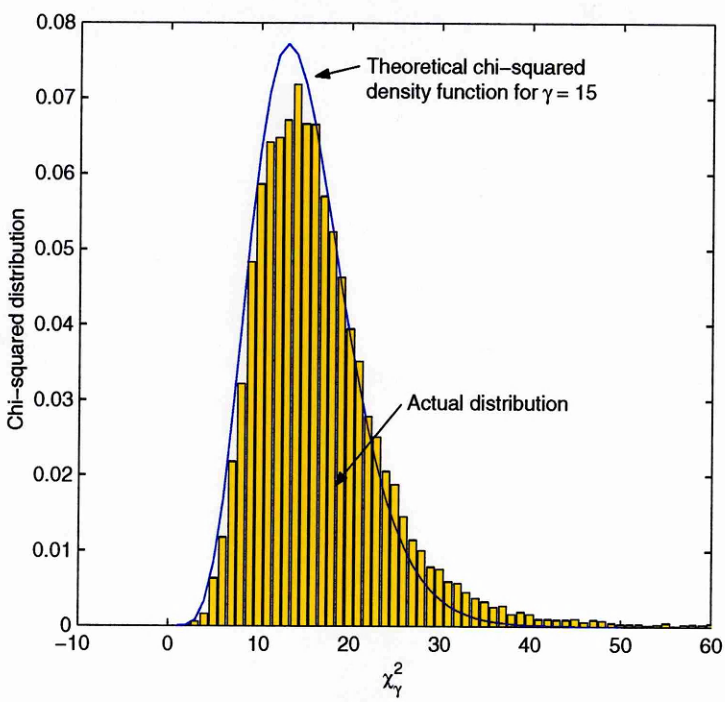


Figure 7-8 Distribution of the chi-squared test statistic with GPS (SV4) failures

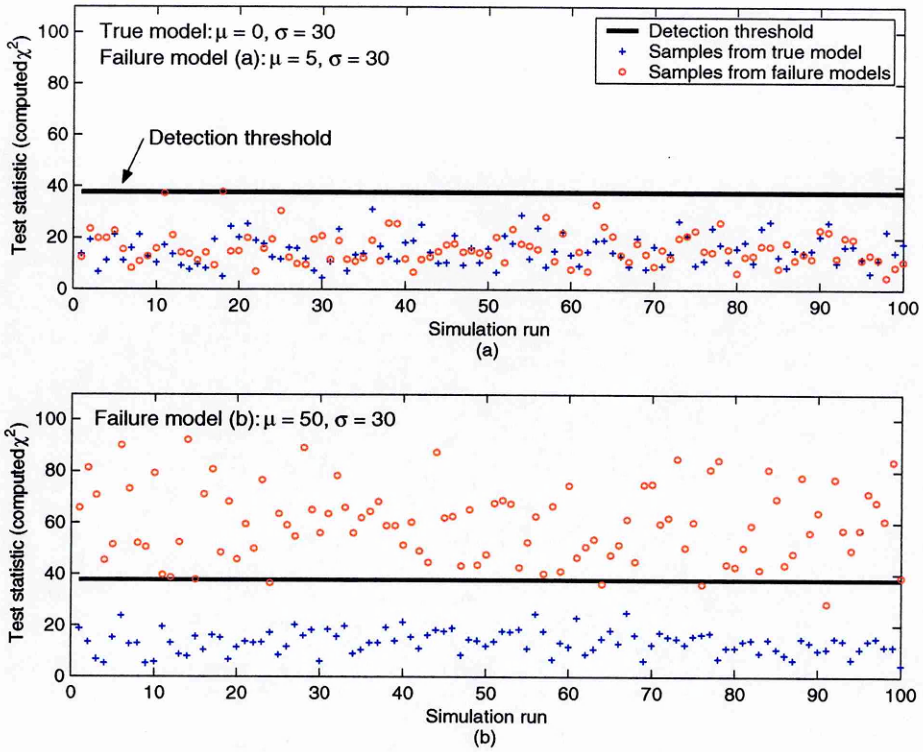


Figure 7-9 Limitations of the chi-squared test statistic

In this chapter, a likelihood function has been developed, which can be used for sensor failure detection and isolation, based on Kalman filter residuals monitoring. The technique is based on using a chi-squared test statistic, which is formed from the residuals. This approach allows a threshold to be selected by specifying a constant false alarm rate. The FDI algorithm, developed in this chapter, differs from other chi-squared-based algorithms in that the residual sequence from each sensor is monitored separately; thereby, enabling identification of a failed sensor. Moreover, the use of a moving window, which specifies the number of samples to be used to calculate the chi-squared test statistic, enables the sensitivity of the algorithm to be adjusted. The concept of this FDI algorithm has been demonstrated and the performance has been evaluated using Monte Carlo simulations of an integrated INS/GPS system with an EKF. It has been shown that although the performance of this algorithm degrades for failures with small magnitude, the algorithm proposed in this chapter can be used as a general FDI algorithm for detecting and isolating specific failure modes based on the implementation of a single Kalman filter.

CHAPTER 8 AIRCRAFT TRAJECTORY CONFLICT PREDICTION

8.1 INTRODUCTION

The ability to accurately predict conflicts of aircraft flight trajectories is a critical element of Air Traffic Management (ATM), particularly with increasing air traffic density and demands to increase airspace capacity. Additionally, automatic conflict prediction is essential in civil air traffic environments such as Free Flight, where the flight crews will be responsible for aircraft separation. In order to perform these tasks, flight crews also require advanced situational awareness and collision detection systems.

This chapter reviews recent developments in aircraft (trajectory) conflict detection algorithms (conflict probes), which are applicable to Free Flight. A new approach to aircraft conflict detection is then proposed. The derivation of this algorithm is supported by several numerical examples. Finally, analysis of characteristics and performance of the proposed algorithm is presented, based on Monte Carlo simulations.

8.2 RECENT DEVELOPMENTS IN FREE FLIGHT CONFLICT DETECTION ALGORITHMS

Free Flight refers to a concept, in which some of the responsibilities of aircraft separation tasks are transferred from air traffic controllers to the flight crews. The objective of Free Flight is to increase the flexibility and efficiency of airspace, by allowing pilots to fly user-preferred trajectories. It has been envisaged that Free Flight could lead to considerable reductions in operating costs for airlines. One critical issue for the implementation of Free Flight is the provision of flight trajectory prediction and conflict detection. The purpose of trajectory conflict prediction is to assist pilots in maintaining aircraft separation, while minimising the workload. Currently, the only airborne system that provides situational awareness and conflict detection capability is the Traffic Alert and Collision Avoidance System (TCAS). TCAS provides:

- ⇒ Displays of aircraft location and vertical speed.
- ⇒ Traffic advisories – warning of collision risk.
- ⇒ Resolution advisories – advice on manoeuvres to avoid collision.

TCAS operates by computing time-to-closest point of approach (CPA) based on measured ranges and range rates. When the computed time-to-CPA and the predicted vertical separation are less than predetermined thresholds, traffic advisories are generated. In addition to the traffic advisories, resolution advisories, which advice appropriate actions (climb or descent), are also generated. The primary objective of TCAS is the provision of a safety system, in support of the present Air Traffic Control systems and in this role, TCAS has been very successful. Even so, the capability of TCAS is limited because it only uses range and range rate measurements. Unnecessary advisories can also be generated because the horizontal miss distances to an intruder aircraft may not be known precisely. TCAS is based on instantaneous measurements, whereas conflict detection for Free Flight application will be based on predicted trajectories (for example, for 10 to 30 minutes in advance).

Aircraft conflict detection algorithms can be categorised in two groups – deterministic[63,64,65] and probabilistic[60,61,62,66]. (The term deterministic is specific to this thesis - Kremer[66] uses the term geometric). In principle, deterministic algorithms establish a protective volume surrounding an aircraft. If a distance between two protective volumes is less than a predetermined threshold, a conflict is said to have occurred. The construction of protective volumes is commonly based on predicted time-to-CPA or minimum miss distances. The shortcomings of these deterministic algorithms are that they tend to be conservative and insensitive to uncertainties, which exist in practical situations. For example, computed aircraft positions and predicted trajectories usually contain some errors, which are introduced by navigation sensor errors or discrepancies between predicted trajectory, based on prediction models, and the actual flight path. Generally, deterministic algorithms are deficient because they omit these stochastic factors.

Probabilistic conflict detection algorithms are designed to take these stochastic factors into consideration. One probabilistic conflict detection algorithm, developed by Paielli and Erzberger[60,61], is of special interest because it has been developed with the assumption that trajectory prediction errors are Gaussian, which is consistent with the implementation of Kalman filter algorithms. Additionally, this algorithm has already been implemented and demonstrated in a real-time application - the Center-TRACON Automation System (CTAS)[68].

Paielli's conflict probe operates by estimating the probability that a pair of aircraft will be in conflict. A conflict is defined as a situation in which aircraft are closer than a predetermined threshold, based on typical separation standards. The separation standards are the minimum allowable separations between aircraft. For en-route airspace, the horizontal separation is 5 nautical miles (nm) and the vertical separation standards are 2000 feet (ft) above 29000ft and 1000ft below 29000ft.[60] Effectively, if the estimated conflict probability is greater than a predetermined probability threshold, a conflict is declared and flight plans must be modified.

Paielli's conflict probe estimates the conflict probability by computing the volume under Gaussian distribution functions. This volume is based on transformations of aircraft position error covariance matrices and a constant relative velocity. Paielli's approach to conflict probability estimation is limited by the fact that it assumes a constant relative velocity. Although this assumption may be true for most of flight profiles, it leads to difficulties in the computation of the conflict probability in three-dimensions, especially when aircraft is climbing or descending. In addition, it should be noted that an analytical solution to conflict probability estimation at "*a particular time*" has not been found.[60]

This chapter presents an alternative approach to aircraft conflict probability estimation, which overcomes the shortcomings of Paielli's algorithm. The conflict probability estimation algorithm, developed in the following sections, is independent of both absolute and relative velocity and coordinate transformations. Moreover, the detection of aircraft conflicts in three-dimensional space is considerably simplified. The algorithm also enables the determination of a conflict probability at a particular time, using an analytical method, which is straightforward, in terms of computation.

8.3 CONFLICT PROBABILITY ESTIMATION UTILISING SEPARATION ESTIMATION ERRORS

8.3.1 Aircraft conflict detection utilising separation estimation errors

This section explains the derivation of an alternative approach to aircraft conflict probability estimation utilising separation estimation errors. The method

covered in this section is restricted to two dimensions – the horizontal plane. An extension to the method, which can be used to account for the altitude of an aircraft, will be covered in Section 8.4. Recall that a conflict between a pair of aircraft is defined as a situation in which the true separation between both aircraft is less than a predetermined threshold.

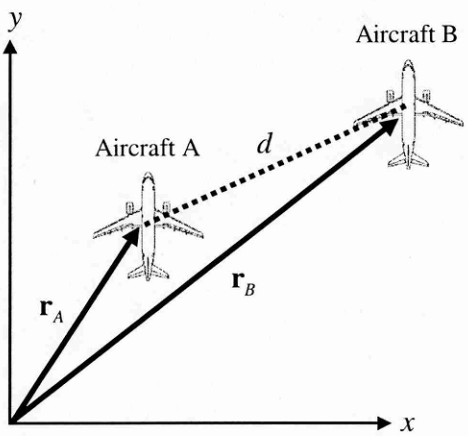


Figure 8-1 Separation between a pair of aircraft in horizontal plane

Consider an encounter between two aircraft, which are travelling at the same altitude, as shown in Figure 8-1. At a particular time, the true separation d between the aircraft A (denoted by the subscript $_A$) and aircraft B (denoted by the subscript $_B$) can be expressed as follows:

$$d = \|\mathbf{r}_A - \mathbf{r}_B\| \qquad \text{Eq. 8-1}$$

Eq. 8-1 can be expressed in component form as follows:

$$d = [(x_A - x_B)^2 + (y_A - y_B)^2]^{\frac{1}{2}} \qquad \text{Eq. 8-2}$$

where $\mathbf{r}_A = [x_A, y_A]^{\mathrm{T}}$ and $\mathbf{r}_B = [x_B, y_B]^{\mathrm{T}}$ are the true position vectors of aircraft A and B respectively. Let d_s denote a specific separation standard. By definition, an aircraft conflict occurs whenever Eq. 8-3 is satisfied.

$$d \le d_s \qquad \text{Eq. 8-3}$$

In practice, however, only the estimates of aircraft position $\hat{\mathbf{r}}$ are known. Defining the relationship between the true and the estimated position vectors as follows:

$$\mathbf{r} = \hat{\mathbf{r}} + \Delta\mathbf{r} \qquad \text{Eq. 8-4}$$

where $\Delta\mathbf{r} = [\Delta x, \Delta y]^T$ is the position estimation error. The estimated separation $\hat{d}$ can be determined as follows:

$$\hat{d} = \|\hat{\mathbf{r}}_A - \hat{\mathbf{r}}_B\| \qquad \text{Eq. 8-5}$$

Eq. 8-5 can be expressed in component form as follows:

$$\hat{d} = [(\hat{x}_A - \hat{x}_B)^2 + (\hat{y}_A - \hat{y}_B)^2]^{\frac{1}{2}} \qquad \text{Eq. 8-6}$$

The relationship between the true and the estimated separation can be described, using Taylor's expansion[12] as follows:

$$d = \hat{d} + \left[\left.\frac{\partial d}{\partial \mathbf{r}_B}\right|_{\substack{\mathbf{r}_A=\hat{\mathbf{r}}_A \\ \mathbf{r}_B=\hat{\mathbf{r}}_B}} \cdot \Delta\mathbf{r}_B\right] + \left[\left.\frac{\partial d}{\partial \mathbf{r}_A}\right|_{\substack{\mathbf{r}_A=\hat{\mathbf{r}}_A \\ \mathbf{r}_B=\hat{\mathbf{r}}_B}} \cdot \Delta\mathbf{r}_A\right] + \boldsymbol{\varepsilon} \qquad \text{Eq. 8-7}$$

where $\boldsymbol{\varepsilon}$ denotes higher order terms and the partial derivatives are defined as follows:

$$\left.\frac{\partial d}{\partial \mathbf{r}_B}\right|_{\substack{\mathbf{r}_A=\hat{\mathbf{r}}_A \\ \mathbf{r}_B=\hat{\mathbf{r}}_B}} = \left[\frac{\hat{x}_B - \hat{x}_A}{\hat{d}} \quad \frac{\hat{y}_B - \hat{y}_A}{\hat{d}}\right] \qquad \text{Eq. 8-8}$$

$$\left.\frac{\partial d}{\partial \mathbf{r}_A}\right|_{\substack{\mathbf{r}_A=\hat{\mathbf{r}}_A \\ \mathbf{r}_B=\hat{\mathbf{r}}_B}} = -\left[\frac{\hat{x}_B - \hat{x}_A}{\hat{d}} \quad \frac{\hat{y}_B - \hat{y}_A}{\hat{d}}\right] \qquad \text{Eq. 8-9}$$

Let $\mathbf{H} = \left.\frac{\partial d}{\partial \mathbf{r}_A}\right|_{\substack{\mathbf{r}_A=\hat{\mathbf{r}}_A \\ \mathbf{r}_B=\hat{\mathbf{r}}_B}}$ and ignoring higher order terms, Eq. 8-7 can be written as follows:

$$d \cong \hat{d} - \mathbf{H}\Delta\mathbf{r}_B + \mathbf{H}\Delta\mathbf{r}_A \qquad \text{Eq. 8-10}$$

Let the separation estimation error $\Delta d = d - \hat{d}$, then Eq. 8-10 can be written as follows:

$$\Delta d = \mathbf{H}\Delta\mathbf{r}_A - \mathbf{H}\Delta\mathbf{r}_B \qquad \text{Eq. 8-11}$$

In Eq. 8-3, the true separation d is used as a parameter for detecting aircraft conflict. By substituting $d = \hat{d} + \Delta d$ in Eq. 8-3, Δd can be used as an aircraft conflict detection parameter as follows:

$$\Delta d \leq \Delta d_{threshold} \qquad \text{Eq. 8-12}$$

where $\Delta d_{threshold}$ is an alternative conflict detection threshold, which is given by:

$$\Delta d_{threshold} = d_s - \hat{d} \qquad \text{Eq. 8-13}$$

Eq. 8-11, Eq. 8-12 and Eq. 8-13 are significant because they present a different approach to aircraft conflict detection utilising the separation estimation error Δd as a detection parameter.

Given the estimated positions of both aircraft ($\hat{\mathbf{r}}_A$ and $\hat{\mathbf{r}}_B$) at a particular time, the estimated separation $\hat{d}$ and the matrix $\mathbf{H}$ can be determined from Eq. 8-5 and Eq. 8-9 respectively. In addition, if the position estimation errors ($\Delta\mathbf{r}_A$ and $\Delta\mathbf{r}_B$) are known precisely, the separation estimation error Δd can be determined from Eq. 8-11. Subsequently, an aircraft conflict can be declared whenever Eq. 8-12 is satisfied. In practice, the position estimation error $\Delta\mathbf{r}$ is never known precisely. If the probability of aircraft conflict is defined as follows:

$$\textit{Aircraft conflict probability} = \text{Prob}[\Delta d \le \Delta d_{threshold}] \qquad \text{Eq. 8-14}$$

Then, the probability that a pair of aircraft are in conflict is given by the probability that Δd is less than $\Delta d_{threshold}$. If Δd is Gaussian with known mean and variance, this probability is simply the cumulative probability of a Gaussian (normal) distribution, i.e. $\text{Prob}[\Delta d \le \Delta d_{threshold}]$, as shown in Figure 8-2. The cumulative probability of a Gaussian distribution can be determined numerically using the *error function*[38,67]. The justification for assuming that Δd is Gaussian is covered in Section 8.3.3.

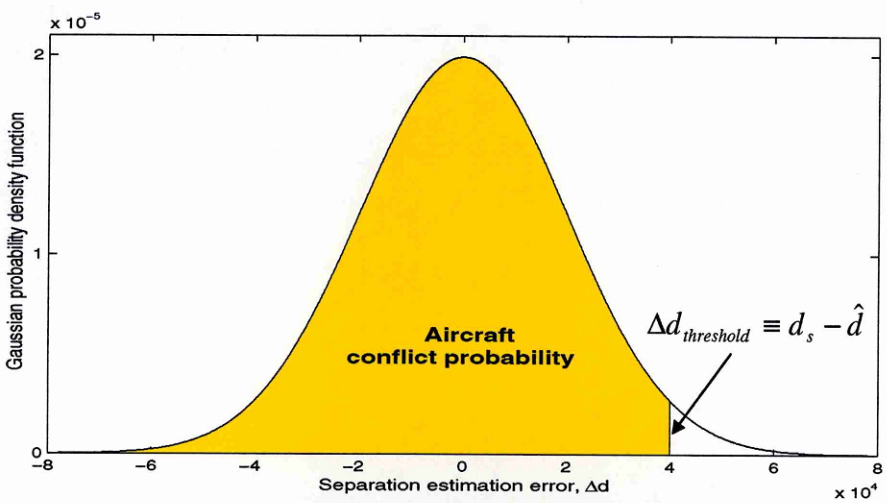


Figure 8-2 Conflict probability estimation utilising separation estimation errors

8.3.2 Statistical characterisation of separation estimation errors

In Section 8.3.1, it was shown that the probability of aircraft conflicts can be determined simply by computing the cumulative probability of a Gaussian distribution, if it can be assumed that the separation estimation error Δd is Gaussian with known mean and variance. From Eq. 8-11, if the position estimation errors $\Delta \mathbf{r}$ can be assumed to be Gaussian, Δd will also be Gaussian. From Eq. 8-11, the mean and variance of Δd can be determined as functions of the statistical properties of $\Delta \mathbf{r}$ as follows:

$$\mathrm{E}[\Delta d] = \mathrm{E}[\mathbf{H}\Delta \mathbf{r}_A - \mathbf{H}\Delta \mathbf{r}_B] = \mathbf{H}\mathrm{E}[\Delta \mathbf{r}_A] - \mathbf{H}\mathrm{E}[\Delta \mathbf{r}_B] \qquad \text{Eq. 8-15}$$

$$\begin{aligned}
\mathrm{Var}\,\Delta d &= \mathrm{E}[\Delta d^2] - \mathrm{E}[\Delta d]^2 \\
&= \mathrm{E}[(\mathbf{H}\Delta \mathbf{r}_A - \mathbf{H}\Delta \mathbf{r}_B)(\mathbf{H}\Delta \mathbf{r}_A - \mathbf{H}\Delta \mathbf{r}_B)^{\mathrm{T}}] - \mathrm{E}[\Delta d]^2 \\
&= \mathrm{E}[(\mathbf{H}\Delta \mathbf{r}_A - \mathbf{H}\Delta \mathbf{r}_B)(\Delta \mathbf{r}_A^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} - \Delta \mathbf{r}_B^{\mathrm{T}}\mathbf{H}^{\mathrm{T}})] - \mathrm{E}[\Delta d]^2 \\
&= \mathrm{E}[\mathbf{H}\Delta \mathbf{r}_A \Delta \mathbf{r}_A^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} - \mathbf{H}\Delta \mathbf{r}_B \Delta \mathbf{r}_A^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} - \mathbf{H}\Delta \mathbf{r}_A \Delta \mathbf{r}_B^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathbf{H}\Delta \mathbf{r}_B \Delta \mathbf{r}_B^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}] \\
&\quad - \mathrm{E}[\Delta d]^2
\end{aligned} \qquad \text{Eq. 8-16}$$

If $\Delta \mathbf{r}$ is zero-mean, i.e. $\mathrm{E}[\Delta \mathbf{r}_A] = \mathrm{E}[\Delta \mathbf{r}_B] = \mathbf{0}$, and there is no correlation between aircraft A and B, i.e. $\mathrm{E}[\Delta \mathbf{r}_A \Delta \mathbf{r}_B^{\mathrm{T}}] = \mathrm{E}[\Delta \mathbf{r}_B \Delta \mathbf{r}_A^{\mathrm{T}}] = \mathbf{0}$, then Eq. 8-15 and Eq. 8-16 can be expressed as follows:

$$\mathrm{E}[\Delta d] = 0 \qquad \text{Eq. 8-17}$$

$$\mathrm{Var}\,\Delta d = \mathrm{E}[\Delta d^2] = \mathbf{H}\mathrm{E}[\Delta \mathbf{r}_A \Delta \mathbf{r}_A^{\mathrm{T}}]\mathbf{H}^{\mathrm{T}} + \mathbf{H}\mathrm{E}[\Delta \mathbf{r}_B \Delta \mathbf{r}_B^{\mathrm{T}}]\mathbf{H}^{\mathrm{T}} \qquad \text{Eq. 8-18}$$

Eq. 8-17 and Eq. 8-18 characterise the mean and variance of Δd as functions of the statistical properties of $\Delta \mathbf{r}$. These equations are significant because they enable a direct determination of aircraft conflict probability. Let $\mathbf{P}_A = \mathrm{E}[\Delta \mathbf{r}_A \Delta \mathbf{r}_A^{\mathrm{T}}]$ and $\mathbf{P}_B = \mathrm{E}[\Delta \mathbf{r}_B \Delta \mathbf{r}_B^{\mathrm{T}}]$ denote the position error covariance matrices of aircraft A and B respectively. The process of determining the probability of aircraft conflict is shown schematically in Figure 8-3.

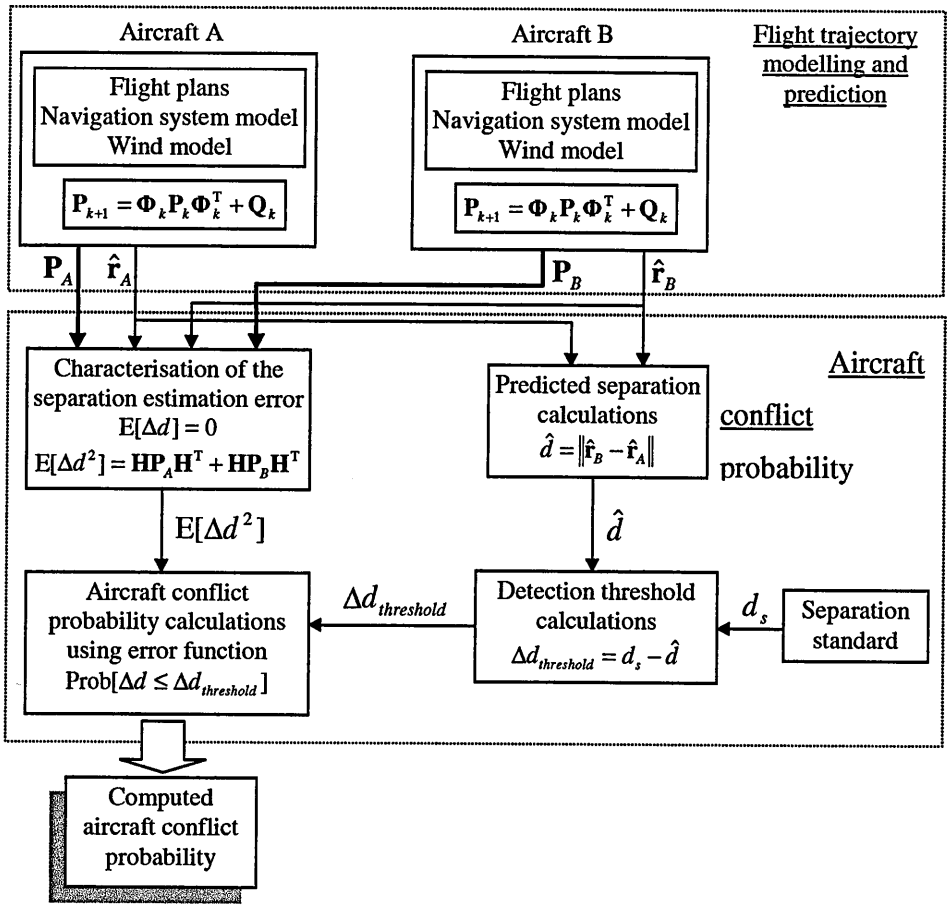


Figure 8-3 Aircraft conflict probability estimation process

Numerical example calculations

The following numerical example is used to illustrate the concept used in this conflict detection algorithm. All measurements are in metres. Let the estimated position vectors of aircraft A and B be $\hat{\mathbf{r}}_A = [0 \quad 0]^T$ and $\hat{\mathbf{r}}_B = [7202.6 \quad 7202.6]^T$ with the estimation error covariance matrices $\mathbf{P}_A = \begin{bmatrix} 200^2 & 0 \\ 0 & 200^2 \end{bmatrix}$ and $\mathbf{P}_B = \begin{bmatrix} 200^2 & 0 \\ 0 & 400^2 \end{bmatrix}$.

From Eq. 8-5 and Eq. 8-9, the estimated separation and the matrix $\mathbf{H}$ are given by $\hat{d} = 10186$ and $\mathbf{H} = [-0.70711 \quad -0.70711]$. The variance of the separation estimation error can be determined as follows:

$$\text{Var}\,\Delta d = \mathbf{H}\mathbf{P}_A\mathbf{H}^T + \mathbf{H}\mathbf{P}_B\mathbf{H}^T = 140000.4 \qquad \text{Eq. 8-19}$$

For a separation standard of 5nm ($d_s = 9260\,\text{m}$), the conflict detection threshold can be calculated as follows:

$$\Delta d_{threshold} = d_s - \hat{d} = 9260 - 10186 = -926 \qquad \text{Eq. 8-20}$$

Therefore, the probability of the aircraft conflict can be calculated numerically from the error function as $\text{Prob}[\Delta d \leq -926] = 0.0067$ (noting that Δd is a zero-mean Gaussian random variable with variance = 140000.4).

8.3.3 Characterisation of aircraft conflict probability estimation algorithm

In the previous section, it was shown that the probability of a conflict between a pair of aircraft can be determined as a probability that the separation estimation error Δd is less than a predetermined threshold. Although, it is not a condition that Δd must be Gaussian, the assumption that Δd is Gaussian simplifies the process considerably. The conflict probability can be determined directly from the error function, if the mean and variance of Δd are known. Δd is Gaussian based on the assumption that aircraft position prediction error $\Delta \mathbf{r}$ is Gaussian. It has been shown that the distribution of aircraft trajectory prediction errors, which have been generated specifically for ATM based on algorithms (such as those used in CTAS), is effectively Gaussian[60,61]. Moreover, the errors of navigation system state estimates, which are computed using algorithms such as the Kalman filter, are also Gaussian. These reasons justify the assumption that $\Delta \mathbf{r}$ and, hence, Δd are Gaussian in this thesis.

The conflict probability estimation algorithm derived in this thesis can be characterised as a function of the estimated separation $\hat{d}$ and the corresponding uncertainties (the variance of separation estimation errors $E[\Delta d^2]$). Figure 8-4 shows the predicted conflict probability for the separation standard d_s = 5nm, and a range of values of $\hat{d}$ and $E[\Delta d^2]$. (Although Figure 8-4 has been obtained from a set of parameters with specific numerical values, the results of this analysis can be applied in general.) Table 8-1 summarises the characteristics of the conflict probability as a function of the estimated separation and the separation standard.

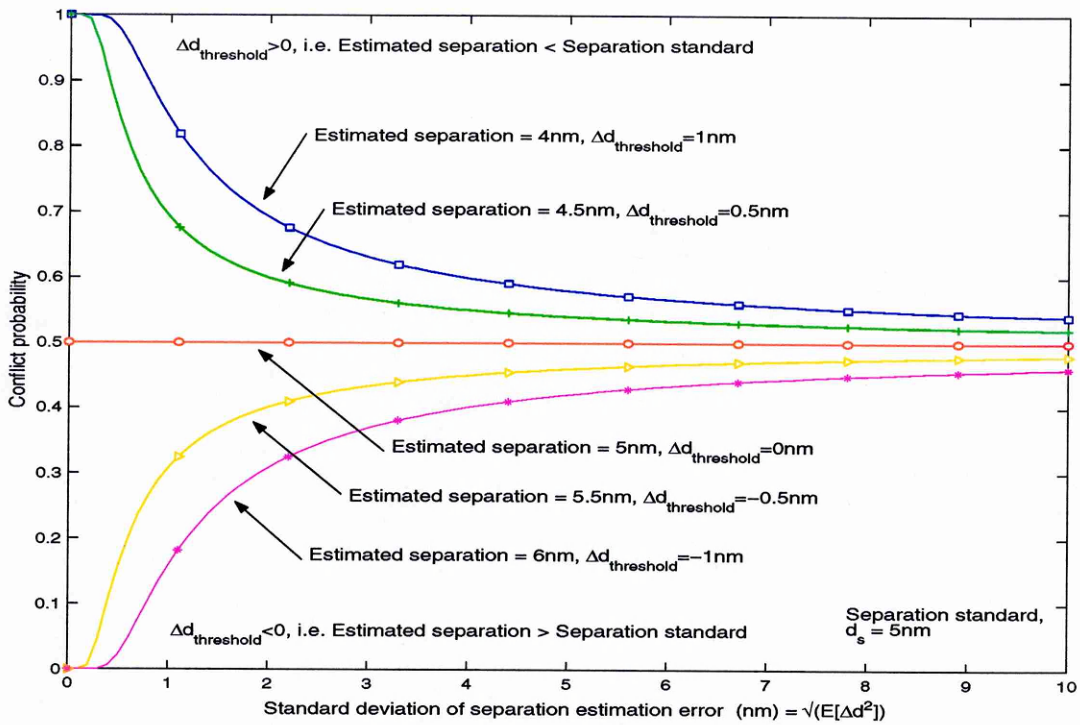


Figure 8-4 Conflict probability as a function of estimated separations and separation estimation error variance

From Figure 8-4, it can be seen that for the same levels of uncertainties in the estimated separation (constant $E[\Delta d^2]$), the aircraft conflict probability increases as the estimated separation decreases. This is to be expected because when aircraft are estimated to be closer together, there is a higher probability that they will be in conflict. When estimated separations equal a separation standard, the conflict detection threshold is zero and, consequently, the aircraft conflict probability is 0.5 for all values of $E[\Delta d^2]$. This is the cumulative probability that a zero-mean Gaussian random variable is less than or equal to zero.

When estimated separations are greater than a separation standard, the conflict detection thresholds are less than zero and, consequently, the conflict probability is always less than 0.5. Additionally, the conflict probability also increases as $E[\Delta d^2]$ increases. This is to be expected because increases in uncertainties of the estimated separations also increase the probability of aircraft conflict.

On the other hand, when estimated separations are less than a separation standard, the conflict detection thresholds are greater than zero and, consequently, the

conflict probability is always greater than 0.5. In addition, the conflict probability also decreases as $E[\Delta d^2]$ increases. This is to be expected because, if the probability of aircraft conflict is already high, increases in the uncertainties of estimated separations reduce the predicted probability of aircraft conflict.

Table 8-1 Characteristics of the conflict probability as a function of the estimated separation and the separation standard

	$\hat{d} > d_s$	$\hat{d} = d_s$	$\hat{d} > d_s$
$\Delta d_{threshold} = d_s - \hat{d}$	< 0	$= 0$	> 0
Conflict probability	< 0.5	$= 0.5$	> 0.5

In the derivation of this conflict probability estimation algorithm, it has been assumed that position estimation errors are uncorrelated between different aircraft (Eq. 8-16) and higher order terms can be neglected (Eq. 8-7). These assumptions are not strictly valid because position errors can be correlated, for example, if several aircraft are subjected to the same wind disturbances or using pseudorange observations from the same set of GPS satellites. It can be seen from Eq. 8-16 that taking the correlation error terms into account can increase or decrease the *true* variance of the separation estimation error (Δd), depending on the sign and magnitude of the correlation error terms. Additionally, the magnitude of higher order terms can become significant, especially, if aircraft position estimation errors are allowed to grow unbounded.

For example, the characteristics of the proposed conflict probability estimation algorithm is discussed for the case when the *true* variance ($E[\Delta d^2]$) of Δd is greater than the *computed* variance (which is based on the simplified form of the equation) as follows:

If estimated separations are less than a separation standard ($\Delta d_{threshold} > 0$), the computed conflict probability (based on the computed variance) will always be greater than the true conflict probability (based on the true variance). However, this characteristic of the conflict probability is preferable, even if the computed conflict probability is sub-optimal.

If estimated separations are greater than a separation standard ($\Delta d_{threshold} < 0$), the computed conflict probability will always be less than the true conflict probability. This situation is not desirable because it leads to an estimation of a conflict probability, which is lower than the true conflict probability. This problem may be corrected by increasing $E[\Delta d^2]$, for example, by including correlation error terms in the calculations. However, these correlation error terms may be difficult to find analytically. A more practical technique is to deliberately increase the covariance of position errors (c.f. Kalman filter tuning). This approach may require elaborate analysis of system models based on extensive Monte Carlo simulations.

These potential solutions and the effects of neglecting correlation error terms and high order terms require further investigations, but for reasons of time, they will not be pursued further in this thesis.

Figure 8-5 illustrates the effects of estimated separations and geometry on the determination of aircraft conflict probability. Aircraft A is at $\mathbf{r}_A = [0 \quad 0]^T$ with the error covariance matrix $\mathbf{P}_A = \begin{bmatrix} 200^2 & 0 \\ 0 & 200^2 \end{bmatrix}$. The separation standard d_s is 5nm. Aircraft B is positioned such that the estimated separations between these aircraft ranged between 5.5nm and 4.5nm, i.e. $\hat{d}$ was in the range $1.1 \times d_s$ to $0.9 \times d_s$, at three different crossing angles (0°, 45° and 90°). The position error covariance matrix of aircraft B is $\mathbf{P}_B = \begin{bmatrix} 200^2 & 0 \\ 0 & 400^2 \end{bmatrix}$. It can be shown that the variance of the separation estimation errors for 0°, 45° and 90° crossing angles are given by 80000, 140000 and 200000m^2 respectively.

From Figure 8-5, it can be seen that the conflict probability increases in all cases as the estimated separation decreases as expected. The comparison of the conflict probability for the three different crossing angles shown in Figure 8-5b, also illustrates that this conflict probability estimation algorithm is able to take into account the geometry of the aircraft because of the effects of the matrix **H** in Eq. 8-11.

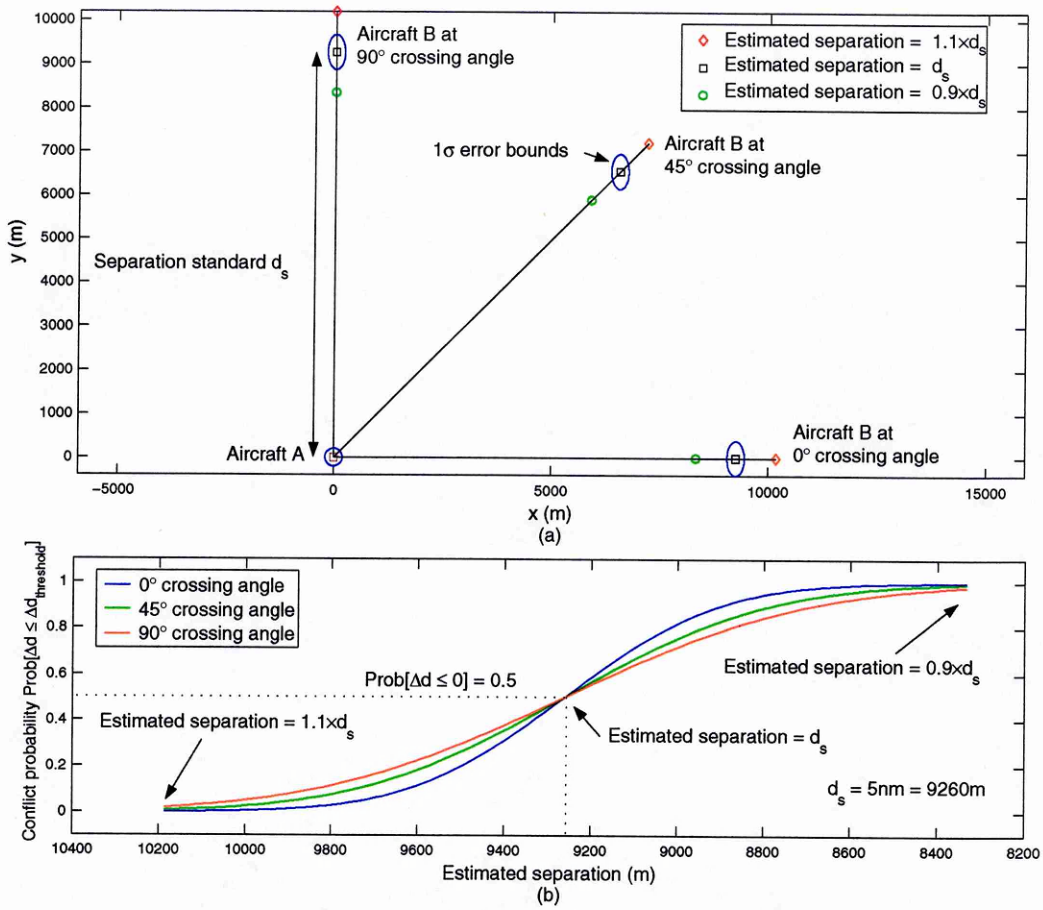


Figure 8-5 Effects of estimated separations and geometry on the determination of aircraft conflict probability

8.3.4 Aircraft conflict probability estimation based on predicted flight trajectory

In the previous sections, the fundamental principles and the characteristics of the aircraft conflict probability estimation algorithm have been described. In practice, it is likely that such algorithms will be used to estimate the probability based on predicted flight trajectories, providing the basis for conflict detection and flight plans modification, which underpins the safety and efficiency of ATM. This section briefly covers the modelling and prediction of flight trajectories, which can be applied to aircraft conflict probability estimation.

Aircraft flight trajectory prediction is complex because actual trajectories can be affected by diverse factors such as winds, navigation system errors, pilot and aircraft control errors. In order to predict flight trajectories accurately, these factors must be taken into consideration and modelled precisely. The flight trajectory prediction model,

used in this chapter, includes only the errors of navigation systems. This approach is adequate to demonstrate the concept of aircraft conflict prediction and to validate the conflict probability estimation algorithm developed in this thesis. The navigation system error model is based on the stand-alone INS error model developed in Chapter 4. Two conflicting flight trajectories, which have been chosen arbitrarily, are shown in Figure 8-6.

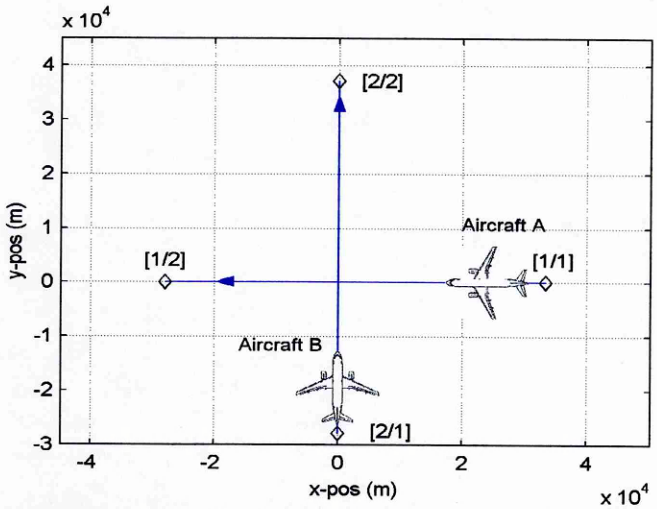


Figure 8-6 Flight trajectories for the evaluation of aircraft conflict probability

The rms errors of the navigation systems can be predicted by applying the standard Kalman filter prediction-equation (Eq. 3-33), which is replicated in Eq. 8-21.

$$\mathbf{P}_{k+1} = \mathbf{\Phi}_k \mathbf{P}_k \mathbf{\Phi}_k^{\mathrm{T}} + \mathbf{Q}_k \qquad \text{Eq. 8-21}$$

Figure 8-7 shows the predicted rms errors of the stand-alone INS systems of aircraft A and B for 300 seconds of simulated flight. To construct the necessary error covariance matrices ($\mathbf{P}_B$ and $\mathbf{P}_A$ in Figure 8-3), only the x and y position rms errors are required in the conflict probability estimation algorithm. It should be noted that the effects of heading errors are already incorporated in the INS error model. Figure 8-8a shows the predicted separations for the simulated flight trajectories. Figure 8-8b compares the computed and the actual conflict probability, where the actual conflict probability has been obtained from Monte Carlo simulations. It can be seen that the conflict probability estimation algorithm accurately predicts the probability that both aircraft will be closer than a predetermined separation standard. The conflict probability increases and decreases correctly as a function of the estimated separations and the estimated separation error variances.

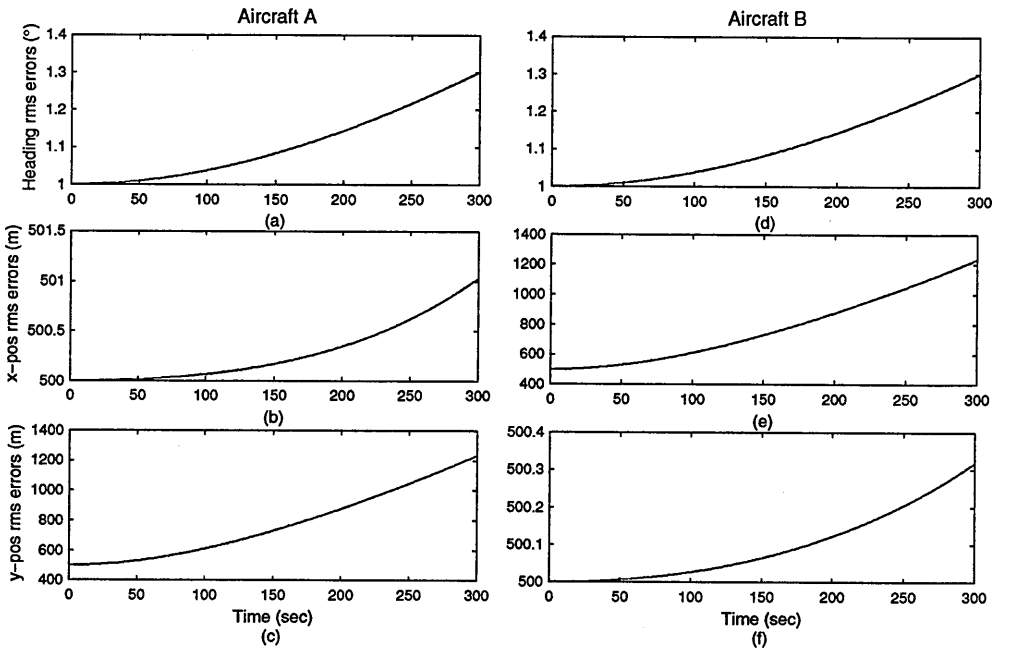


Figure 8-7 Predicted rms errors of the stand-alone INS system

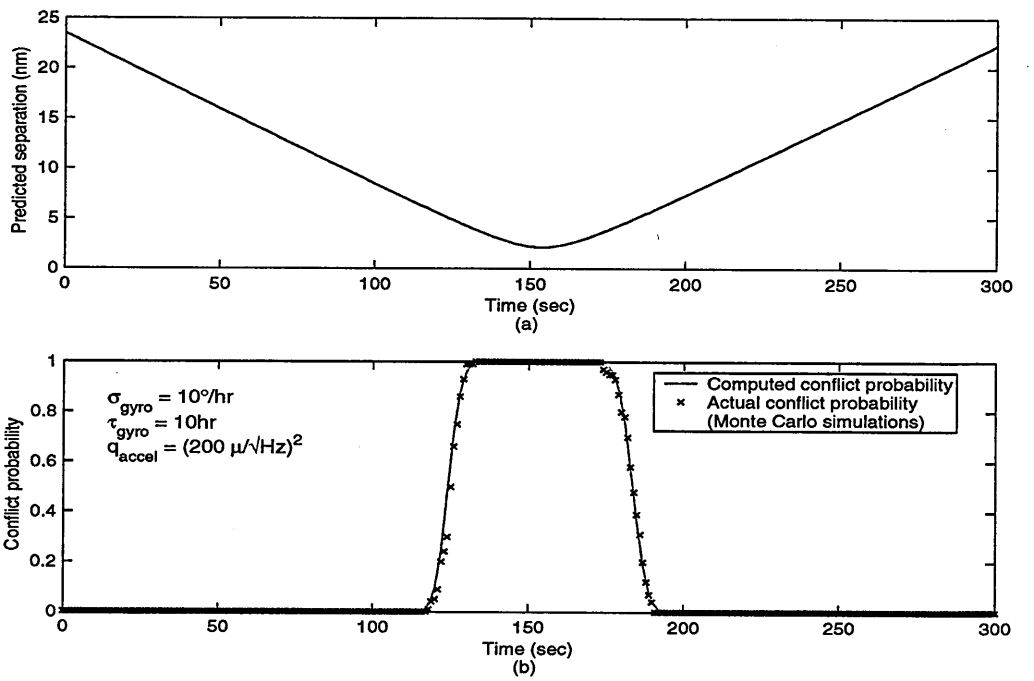


Figure 8-8 Aircraft conflict probability calculations based on the Kalman filter prediction-equation

8.4 THREE DIMENSIONAL AIR CRAFT CONFLICT DETECTION ALGORITHM

The estimation of aircraft conflict probability in the horizontal plane was described in Section 8.3. In practice, aircraft can also be in conflict in the vertical plane. This section outlines the modifications to the algorithm described in Section 8.3, to take

into account aircraft altitude, in order to estimate aircraft conflict probability in the vertical plane. In addition, a concept of aircraft conflict detection in a three-dimensional space is outlined.

8.4.1 Vertical conflict probability estimation

It is straightforward to derive the conflict probability in the vertical plane by directly substituting Eq. 8-22, Eq. 8-23 and Eq. 8-24 in the equations in Section 8.3.

$$y = \hat{y} = \Delta y = 0 \qquad \text{Eq. 8-22}$$

$$x = h,\ \Delta x = \Delta h,\ \hat{x} = \hat{h} \qquad \text{Eq. 8-23}$$

$$v = d,\ \Delta v = \Delta d,\ \hat{v} = \hat{d},\ v_s = d_s \qquad \text{Eq. 8-24}$$

where h and v denote aircraft altitude and vertical aircraft separation respectively. Consequently, the characterisation of the conflict probability described in Section 8.3.3 is also applicable to the vertical conflict probability. Monte Carlo analysis of vertical conflict probability estimation was not deemed to be necessary because it only requires changes of numerical values of the parameters.

8.4.2 Aircraft conflict detection in three dimensions

Recall that aircraft conflict is defined in terms of separation standards for the horizontal and vertical planes. An aircraft conflict is only declared when *both* separation standards (horizontal *and* vertical) have been exceeded. For example, when two aircraft are flying at the same altitude, their vertical separations are zero, giving a conflict in the vertical plane. However, both aircraft can still be safely separated while the horizontal separation is not exceeded. These aircraft will be in conflict when both the vertical and the horizontal separation standards have been exceeded. Subsequently, a conflict between two aircraft in three-dimensional space can be declared whenever the estimated horizontal and vertical conflict probability are above predetermined probability thresholds. Figure 8-9 shows the concept of the proposed conflict detection algorithm.

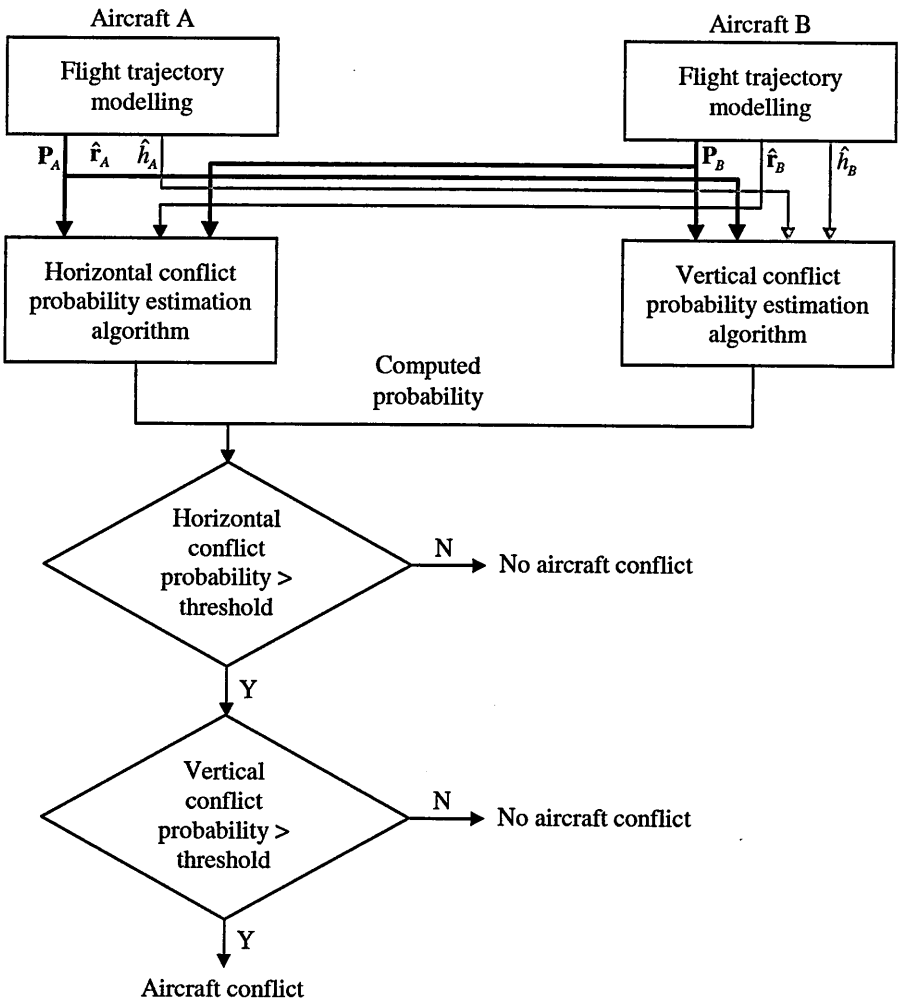


Figure 8-9 Aircraft conflict detection algorithm in three-dimensional space

The concept of conflict probability estimation can be extended to provide the basis of aircraft separation assurance. In principle, safe separation can be maintained by assuring that the predicted conflict probability is always below a certain limit. Furthermore, the conflict probability estimate can also be applied to the generation of conflict resolutions. For example, the conflict probability estimate can be used as an parameter to form a cost function, which is to be optimised, to generate the optimal conflict resolution. It will also be useful to investigate whether it is possible to express the conflict probability estimate analytically as a function of time, based on the error covariance propagation function i.e. the Riccati equation. Nevertheless, for reasons of time, these concepts have not been investigated in this thesis.

In this chapter, developments in aircraft trajectory conflict prediction algorithms have been investigated which are applicable to Free Flight. These algorithms can be categorised in two groups – deterministic and probabilistic. Deterministic algorithms are deficient because they do not take account of stochastic factors such as navigation sensor errors and modelling errors, whereas probabilistic algorithms can overcome this shortcoming.

In this thesis, the conflict probe developed by Paielli and Erzberger which is based on assumptions that are consistent with Kalman filter algorithms has been extended. Because Paielli's conflict probe operates by estimating the probability of aircraft conflict using transformations of position error covariance matrices and a constant relative velocity, it is inefficient in determining the probability of aircraft conflict at a particular time. Moreover, the determination of conflict probability in three-dimensional space is difficult, especially when aircraft are in climb or descent.

An alternative approach to conflict probability estimation, based on separation estimation errors, has been proposed in this chapter. This approach is computationally simple and overcomes the inherent shortcomings of Paielli's algorithm. The performance and characteristics of the proposed algorithm have been analysed from Monte Carlo simulations and it has been shown that this algorithm is able to estimate aircraft conflict probability accurately. Finally, the modifications to the equations for the horizontal plane have been outlined, extending the conflict detection algorithm to three-dimensional space.

CHAPTER 9 CONCLUSION AND SUGGESTED FURTHER WORKS

9.1 OVERVIEW OF THE THESIS AND RESULTS

This thesis addresses navigation and guidance which will be required for air traffic management (ATM) in FANS (and Free Flight). Innovative methods and algorithms for aircraft separation assurance were developed and validated using aircraft navigation simulation studies.

The thesis began with a review of the evolution of air navigation systems (Chapters 1 and 2). The necessity to improve the safety and efficiency of airspace, leading to the proposal for FANS, and the difficulties in implementing FANS was described, focusing particularly on the criticality of aircraft separation assurance. One major difficulty is centred on a single point failure problem, where momentary losses of GPS signals can jeopardise the safety of numerous aircraft simultaneously. Another difficulty is concerned with the use of predicted aircraft position (e.g. from ADS-B reports) in separation assurance.

The research described in this thesis focused on these technical difficulties. Firstly, an alternative method was developed to enable aircraft to precisely maintain separation, even during disruptions to GPS operations. Secondly, a method was devised to account for the uncertainties of predicted aircraft position, which is applicable to flight path conflict detection. The approach to these problems is based on a combination of data fusion techniques, relative navigation (RELNAV) principles, sensor failure detection and isolation (FDI), and probabilistic conflict detection.

The development of the underlying navigation simulation studies was the subject of Chapters 3 and 4. The basis of extended Kalman filters (EKF), which have been applied extensively to navigation systems integration, and analysis of their performance using Monte Carlo simulations was described (Chapter 3). The modelling of aircraft guidance, airspace, INS, and GPS used to develop a simulation program in Matlab, and the validation of these simulation models, was explained (Chapter 4 and Appendix B). The equations of motion and state-space system models suitable for the EKF were derived, based on strapdown INS notation and coordinate systems commonly

used in aircraft navigation (the first objective of the thesis, Section 1.6). The performance of a stand-alone INS, stand-alone GPS, and an integrated INS/GPS system was analysed using the simulation studies. These studies provided the foundation for investigating the application of relative navigation (RELNAV) to civil ATM (the second objective of the thesis).

Chapters 5 and 6 contained feasibility studies on the use of ADS-B to enable RELNAV. JTIDS and RELNAV principles were reviewed. The equations of relative motion and system models were derived. RELNAV algorithms for both centralised (Chapter 5) and decentralised (Chapter 6) EKF were formulated and validated using the simulation studies. It was shown that RELNAV, based on the integration of INS and ADS-B, enabled aircraft to precisely maintain safe separation, independent of GPS.

Principal implementation considerations of RELNAV were examined. The significance of the observability and selection of sources of available measurements (source selection functions) was highlighted. Source selection functions, which exploit error covariance matrices, and an alternative algorithm that integrates INS, GPS, and ADS-B were developed and evaluated (Sections 6.4 and 6.5). This algorithm is considered to be more applicable to civil ATM because it eliminates the needs to organise several RELNAV communities simultaneously. However, the algorithm requires that at least one satellite is available for navigation to ensure the stability.

Failure detection and isolation (FDI) algorithms were reviewed (Chapter 7). An alternative algorithm, which combines the use of likelihood functions and chi-squared tests, was derived and evaluated. It was shown that this algorithm allows both a false alarm rate to be selected and a failed sensor to be isolated in certain circumstances.

Finally, aircraft conflict detection algorithms were reviewed (Chapter 8). An alternative conflict probability estimation algorithm was derived and evaluated. It was shown that the proposed algorithm simplifies conflict detection in three dimensions, and allows the conflict probability to be determined independent of the aircraft velocity.

The results of this research and their implications are described in greater detail in the following sections.

9.2 RELATIVE NAVIGATION UTILISING ADS-B

An innovative method of assuring aircraft separation was proposed in this thesis, based on the use of ADS-B to provide RELNAV, introducing a new application of ADS-B in the form of an integrated communication and navigation system. The main attraction of this approach is the fact that ADS-B will be implemented as part of airspace modernisation. In principle, the benefits of this method can be gained by addition of data processors (software) which combine measurements of existing sensors.

The equations of motion in a relative grid were formulated (Section 5.4), based on the simplified model of a strapdown INS. The use of this simplified model of INS allowed all of the principal error terms (position, attitude, timing, and sensor errors) to be completely represented by a relatively small (8-states) error model (Section 5.5). This approach afforded straightforward system analysis, which could be extended to analysis of complex systems, based on the same principles. Furthermore, it was shown that this error model is in fact the same as a model developed by Schneider (using the transformation in Eq. 5-54), with an addition of the state variables for timing errors.

At first, the performance of RELNAV was analysed, based on a centralised approach, in which a single EKF processes all of available measurements. Although this approach is not practical, it does produce optimal solutions against which the performance of other algorithms can be compared.

It is known that, when fewer than four satellites are available, navigation using stand-alone GPS is not possible. The performance of a stand-alone INS can also degrade appreciably as a function of time. The analysis (Section 5.6) confirmed that RELNAV, based on the integration of INS and ADS-B, can enable aircraft to precisely maintain separation, independent of GPS. For example, results from the simulation studies showed that, after 10 minutes of a complete GPS failure, positioning errors (1σ) of a stand-alone INS[i] could be as large as 1000m whereas those of centralised RELNAV were maintained at 10m approximately (Figure 5-11). These results implied that, in principle, RELNAV can be used to improve the safety and integrity of FANS airspace,

---

[i] Low accuracy INS (Table 4-3): Gyroscope bias 10 °/hr and Accelerometer PSD 200^2 $(\mu g / \sqrt{\text{Hz}})^2$

especially, to accommodate GPS failures (the third objective of the thesis).

An alternative RELNAV algorithm was developed (Section 6.2 and Appendix G), based on the ownstate formulation proposed by Gobbini. Each aircraft models only the errors of its own navigation systems, while measurements are processed by the EKF implemented in each individual aircraft. This decentralised algorithm is necessary to reduce the computation and communication demands so that RELNAV can be implemented in practice. However, owing to the mismodelling (only the error states of the own aircraft are modelled and assumed to be white, i.e. uncorrelated in time), it was known that this algorithm was sub-optimal and susceptible to Kalman filter divergence.

The use of measurement resampling technique was examined to overcome this mismodelling problem (Section 6.3). It was shown that increasing the measurement sample intervals (in effect, making the measurements appear less correlated in time) can eliminate the divergence problem, allowing decentralised RELNAV to be used with slight deterioration in performance, compared with the centralised approach. The same simulation studies used to evaluate the centralised algorithm showed that, after 10 minutes of a complete GPS failure, positioning errors (1σ) of decentralised RELNAV were 50m approximately (Figure 6-5).

It was known that the use of error covariance matrices for source selection can result in the instability of a community of aircraft that is based on the ownstate formulation. In this thesis, one source selection method, which utilises the traces (sum of diagonal terms) of error covariance matrices to form performance indices, was developed and analysed using the simulation studies. The analysis (Section 6.4.2) showed that source selection functions can be designed in such a way that these performance indices are organised into sub-groups (ranks) to alleviate the mismodelling problem (due to frequent rank reversals), and hence improve the stability of the navigation community. Moreover, it was shown that the presence of the navigation controller is also crucial to ensure the stability of the community. These studies provided insightful knowledge about designing covariance-based source selection functions, which was not available previously in the public domain.

The concept of utilising ADS-B to afford RELNAV was proven to be feasible. However, it was considered that to enable aircraft to navigate purely in a relative grid,

i.e. as in JTIDS RELNAV, would require establishing and managing several local navigation communities concurrently, which could be too complex to be implemented in civil ATM. An alternative algorithm was developed (Section 6.4) to form an integrated INS/GPS/ADS-B system to eliminate this complication; in effect, exploiting GPS as the navigation controller of a single global navigation community. Although this implies that at least one GPS satellite (the controller) must be available for navigation to ensure the stability, this approach is justified because it eliminates the complication outlined above. In addition, it is considered extremely likely that civil ATM will operate with at least one GPS satellite available for navigation at all times.

The performance of the proposed integrated INS/GPS/ADS-B system was evaluated using the simulation studies. It was shown that addition of ADS-B sensors can improve the performance of an integrated INS/GPS system (Section 6.5). For example, it was shown that, when one GPS satellite is available for navigation, positioning errors of an integrated INS/GPS system were 100m whereas those of an integrated INS/GPS/ADS-B system were 20m approximately (Figure 6-12f-g).

An attempt was made to examine the use of RELNAV to improve the integrity (the provision of a timely warning of a failure) of GPS. For reasons of time, it was not possible to conclude whether RELNAV could be used to improve the integrity of GPS. However, a new FDI algorithm was developed, which extends the use of likelihood functions to monitor residual sequences of a Kalman filter, by combining it with chi-squared tests. The proposed algorithm allows the advantages of both techniques to be realised, i.e. both a false alarm rate can be selected and a failed sensor can be isolated. The simulation studies showed that this algorithm is suitable for detecting hard failures rather than soft failures (Section 7.5.2).

The studies on the application of RELNAV in this thesis is based on one crucial assumption that ADS-B datalinks will afford accurate time synchronisation – an equivalent value of $(20\text{m})^2$ in terms of a power spectral density of pseudorange measurement noise. It is possible that the present candidate datalinks for ADS-B will not be able to support this synchronisation capability. Nevertheless, this research has illustrated that potential benefits can be gained in terms of superior navigation performance by increasing the capability of these communication systems.

9.3 AIRCRAFT CONFLICT DETECTION

Existing probabilistic conflict detection algorithms, which consider the uncertainties of predicted flight paths in conflict detection, were examined (Section 8.2). A new conflict probability estimation algorithm, which determines the probability that two aircraft are closer than a pre-determined distance, was derived (Section 8.3). Since the proposed algorithm computes a conflict probability independent of the aircraft velocity, it affords the following advantages, which overcome the limitations of an algorithm developed by Paielli and Erzberger (which depends on an assumption of a constant relative velocity):

⇒ A conflict probability at a particular time can be determined analytically, based on the use of the error function.

⇒ Conflict detection in three dimensions is simplified because horizontal and vertical conflict probabilities are determined separately, independent of the aircraft velocity.

The simulation studies were used to validate and evaluate the performance of the proposed conflict probability estimation algorithm. It was pointed out that, owing to the truncation of higher-order terms in the derivation of this algorithm, large position estimation errors can affect the performance of the algorithm. Nevertheless, it is expected that the accuracy of the algorithm can be improved, by retaining these higher order terms.

The advantage of the proposed algorithm lies in the fact that, unlike conventional systems such as TCAS (which uses only range and range rate measurements), this algorithm exploits the information contained in ADS-B reports. It enables a conflict probability to be determined straightforwardly as a function of time and the uncertainties of predicted aircraft positions. Potentially, it provides the basis for future development of collision avoidance and separation assurance systems for FANS and Free Flight (the forth objective of the thesis).

9.4 SUGGESTIONS FOR FURTHER RESEARCH

Since this work focused mainly on initial feasibility studies, it was limited to simple simulations (based on a few typical flight paths and simplified models of ADS-B, INS, and GPS). More realistic simulations would be required to fully validate the concepts and algorithms proposed in this thesis. For example, further simulation studies should be based on standard aircraft equations of motion in three dimensions, realistic flight control and guidance algorithms, realistic models of aircraft navigation and communication systems, and typical Free Flight air traffic environment.

A considerable amount of research also remains to be done to address several practical issues, regarding the application of RELNAV principles. Firstly, RELNAV algorithms were formulated using strapdown INS notation, whereas, in practice, both strapdown and stable platform INS would be used. Therefore, a method of mixing different types of INS should be examined and it may be possible to derive a generic error model applicable to both types of INS. Secondly, Gobbini's method of decentralised reconstruction of the centralised optimal RELNAV algorithm is vulnerable to structural perturbations and costly in terms of communication demands. It would be interesting to develop another optimal decentralised algorithm which satisfies the practical limitations imposed by the current computational and communication technologies. Further research into these areas would provide knowledge to support the use of RELNAV in civil ATM.

The advantage of navigation purely in a relative grid is evident, particularly for aircraft separation assurance. It may be possible to develop an alternative approach, which does not rely on the use of GPS as the controller, to eliminate the complication of organising several local navigation communities concurrently. This may require drawing up a set of rules that permits several navigation communities to be established and managed simultaneously. Further research into this issue would be highly valuable because it would allow the benefits of RELNAV to be realised as originally envisaged.

It is worth recalling that RELNAV allows the position reports of nearby aircraft to be used for navigation, by accounting for the uncertainties of this information using stochastic models. This method expands the conventional approach to the solution

of the navigation equations, where knowledge about the sources is usually assumed to be perfect (without errors).

It is possible that the application of RELNAV principles could aid the development of other unrelated areas. For example, common methods of GPS integrity monitoring rely on using greater numbers of GPS satellites than those require solely for navigation. This provides redundant information needed for failure detection and has been one of the major reasons for the limited integrity of GPS. Further work might examine the use of position reports from other aircraft (e.g. from ADS-B) to provide the necessary redundant information. It may be possible to account for the uncertainties of these position reports using similar techniques to those employed for navigation.

Although much research still remains to be done, it is believed that potential benefits can be gained from the application of RELNAV principles to civil ATM. This is underpinned by the following considerations:

⇒ The analysis of RELNAV principles and performance described in this thesis is fundamental and should be equally applicable to complex systems.

⇒ A Kalman filter is merely a set of algorithms. The success of applying the filter is predicated largely on the ability to understand and accurately model the systems under consideration.

REFERENCES

1. Perry, Tekla. S. August 1997. In search of the future of air traffic control. IEEE Spetrum 34, no. 8: 19-35.
2. The Boeing Company. June 2000. Boeing Current Market Outlook 1999. [http://www.boeing.com/commercial/cmo/4wa06.html]
3. SITA. September 1996. Data Sheet FANS-1/A Communications. [http://www.sita.int/resources/if_fans1a.pdf]
4. Scott, W. B. 29 April 1996. JTIDS provides F-15Cs 'God's-eye view'. Aviation Week & Space Technology 144, no. 18: 63.
5. Corrigan, T. M and others. January 1999. GPS risk assessment study (final report) The Johns Hopkins University, Applied Physics Laboratory. [http://www.jhuapl.edu/transportation/aviation/gps/gps.pdf]
6. FAA. July 2000. GPS transition plan [http://www.faa.gov/asd/gpstrans.pdf]
7. MITRE. February 1999. Description of alternative navigation architectures. [http://www.faa.gov/asd/satnav/altnav1.pdf]
8. Kalman, R. E. March 1960. A new approach to linear filtering and prediction problems. Transaction of the ASME, Journal of Basic Engineering : 35-45.
9. Maybeck, Peter S. 1979. Stochastic models, estimation and control Volume 1. Academic Press, Inc.
10. Gelb, Arthur. 1974. Applied optimal estimation. The analytic Sciences Corporation.
11. Brown, Robert Grover and Patrick Y. C. Hwang. 1997. Introduction to random signals and applied Kalman filtering. John Wiley & Sons.
12. Jeffrey, Alan. 1996. Mathematics for engineers and scientists. Chapman & Hall.
13. Kerr, Thomas. January 1987. Decentralized filtering and redundancy management for multisensor navigation. IEEE Transactions on Aerospace and Electronic Systems AES-23, no. 1: 83-119.
14. Carlson, Neal A. May 1990. Federated square root filter for decentralized parallel processes. IEEE Transactions on Aerospace and Electronics Systems 26, no. 3: 517-25.

15. Kerr, Thomas. November 1991. Comments on "Federated square root filter for decentralized parallel processes". IEEE Transactions on Aerospace and Electronics Systems 27, no. 6: 946-51.
16. Savage, Paul G. January 1998-28 February 1998. Strapdown inertial navigation integration algorithm design Part 1: Attitude algorithms. Journal of Guidance, Control and Dynamics 21, no. 1: 19-28.
17. Titterton, D. H. and J. L. Weston. 1997. Strapdown inertial navigation technology. Perter Peregrinus Ltd.
18. Shabana, Ahmed A. 1989. Dynamics of multibody systems. John Wiley & Sons, Inc.
19. Chatfield, Averil B. 1997. Fundamental of high accuracy inertial navigation. USA: American Institute of Aeronautics and Astronautics, Inc.
20. Forssell, Borje. 1991. Radionavigation systems. Great Britain: Prentice Hall International (UK) Ltd.
21. Williams, James H. 1997. Fundamentals of applied dynamics. John Wiley & Sons, Inc.
22. Maltbaek, J. C. 1975. Essential engineering dynamics. London: Crosby Lockwood Staples.
23. Eurocontrol. HOMEPAGE of BADA Aircraft Performance Database. December 1999. [http://www.eurocontrol.fr/projects/bada/]. 25 January 2000.
24. Siouris, George M. 1993. Aerospace avionics systems: a modern synthesis. The United State of America: Academic Press Limited.
25. Eurocontrol. 1998. User manual for the Base of Aircraft Data (BADA) Revision 3.0. EEC Note No. 6/98.
26. Erzberger, Heinz. 1967. Application of Kalman filtering to error correction of inertial navigators. National Aeronautics and Space Administration, NASA TN D-3874.
27. Spilker, James J. 1996. Satellite constellation and Geometric Dilution of Precision. in Global Positioning System: theory and applications Volume 1 Bradford W. Parkinson and James J Spilker, 177-208. USA: American Institute of Aeronautics and Astronautics, Inc.

28. Kaplan, Elliott D. 1996. Understanding GPS principles and applications. Artech House, INC.
29. Parkinson, Bradford W. and James J Spilker. 1996. Global Positioning System: theory and applications Volume I. USA: American Institute of Aeronautics and Astronautics, Inc.
30. Axelrad, P. and R. G. Brown. 1996. GPS navigation algorithms. in Global Positioning System: theory and applications Volume I Bradford W. Parkinson and James J Spilker, 409-33 American Institute of Aeronautics and Astronautics, Inc.
31. Strang, Gilbert. 1986. Introduction to applied mathematics. USA: Wellesley-Cambridge Press.
32. Joseph, P. D. November 1997 Joseph P. D. Homepage. [http://ourworld.compuserve.com/homepages/PDJoseph/intro.htm]. 14 February 2000.
33. Kayton, Myron and Walter R. Fried. 1969. Avionics navigation systems. USA: John Wiley & Sons, Inc.
34. Collinson, R P G. 1996. Introduction to Avionics. Great Britain: Chapman & Hall.
35. Gordon, G. S. Navigation systems integration. The GEC Journal of Technology. 1998; 15(2):80-90.
36. Greenspan, R. L. 1996. GPS and inertial integration. in Global Positioning System: theory and applications Volume II Bradford W. Parkinson and James J Spilker, 187-220. American Institute of Aeronautics and Astronautics, Inc.
37. Phillips, R. E. and Schmidt, G. T. GPS/INS Integration in. System implications and innovative applications of satellite navigation. France; 1996; AGARD-LS-207.
38. William H. Press and others. 1992. Numerical recipes in C: The art of scientific computing. Cambridge University Press.
39. Lewantowicz, Zdzislaw H. and Danny W. Keen. 1991. Graceful degradation of GPS/INS performance with fewer than four satellites. ION National technical meeting: Institute of Navigation, 269-76.
40. Stansell, Thomas A. 1979. TRANSIT - The current satellite navigation system, in AGARD-AG-245.

41. Schneider, Alan. M. July 1985. Observability of relative navigation using range-only measurements. IEEE Transactions on Aerospace and Electronics Systems AES-21, no. 4: 569-81.

42. Fried, W. R. January 1978. Principles and simulation of JTIDS Relative Navigation. IEEE Transactions on Aerospace and Electronic Systems AES-14, no. 1: 76-84.

43. Ranger, J. F. O. Principles of JTIDS Relative Navigation. Journal of Navigation 49, no. 1: 22-35.

44. Schneider, Alan. M. July 1985. Observability of relative navigation using range-only measurements. IEEE Transactions on Aerospace and Electronics Systems AES-21, no. 4: 569-81.

45. Gobbini, Giuseppe F. 1981. Relative navigation by means of passive rangings. Ph.D. Thesis, Massachusetts Institute of Technology.

46. Greenberg, M. S. and H. J. Rome. 1978. Stability analysis of relative navigation systems. Proceeding of the IEEE Position Location and Navigation Symposium, 335-44.

47. Widnall, W. S. and Giuseppe F. Gobbini. 1983. Stability of the decentralized estimation in the JTIDS relative navigation. IEEE Transaction on Aerospace and Electronic Systems AES-19, no. 2: 240-9.

48. Rome, H. J. and J. S. Stambaugh. 1976. Evaluation of the accuracy of community relative navigation organization concepts. Proceedings of the Institute of Navigation, 4-10.

49. Deyst, J. J. and C. F. Price. 1968. Conditions for the asymptotic stability of the discrete, minimum variance, linear estimator. IEEE Trans. Automatic Control AC-13: 702-5.

50. Daly, Kevin C., Eliezer Gai, and James V. Harrison. January 1979. Generalized likelihood test for FDI in redundant sensor configurations. Journal of Guidance and Control 2, no. 1: 9-17.

51. Parkinson, Bradford. W. and Penina Axelrad Techniques for autonomous GPS integrity monitoring. AGARD-AG-314.

52. Milton, J. S. and Jesse C. Arnold. 1990. Introduction to probability and statistics: Principles and applications for engineering and the computing sciences. 2ed. McGraw-Hill Publishing Company.

53. Staff of Research and Education Association. 1984. Handbook of mathematical, scientific, and engineering formulas, tables, functions, graphs, transforms. USA: Research and Education Association.

54. Da, Ren and Ching-Fang Lin. January 1995. A new failure detection approach and its application to GPS autonomous integrity monitoring. IEEE Transactions on Aerospace and Electronic Systems 31, no. 1: 499-506.

55. Hajiyev, Chingiz M. and Fikret Caliskan. 1998. Fault detection in flight control systems via innovation sequence of Kalman filter. UKACC International conference on CONTROL'98: IEE.

56. Chowdhury, Fahmida. 1998. A neuro-statistical method for fault detection in stochastic systems. Proceedings of the 1998 IEEE iternational conference on control applications : IEEE.

57. Brumback, B. D. and M. D. Srinath. June 1987. A chi-squared test for fault-detection in Kalman filters. IEEE Transactions Automatic Control AC-32, no. 6: 552-4.

58. Kerr, Thomas H. November 1990. Comments on "A chi-squared test for fault detection in Kalman filters". IEEE Transactions Automatic Control 35, no. 11: 1277-8.

59. Brown, R. G. 1996. Receiver autonomous integrity monitoring. Global Positioning System: theory and applications Volume II Bradford W. Parkinson and James J Spilker. USA: American Institute of Aeronautics and Astronautics, Inc.

60. Paielli, R. A. and H. Erzberger. 1997. Conflict probability estimation for Free Flight. Journal of Guidance, Control, and Dynamics 20, no. 3: 588-96.

61. Erzberger, Heinz and others. 1997. Conflict detection and resolution in the presence of prediction error. The 1st USA/Europe air traffic management R&D seminar. [http://atm-seminar-97.eurocontrol.fr/index.htm]

62. Bakker, G. J. Bert and Henk A. P. Blom. 1993. Air traffic collision risk modelling. Proceedings of the 32nd conference on decision and control: IEEE, 1464-9.

63. Zeghal, Karim. 1994. Towards the logic of an airborne collision avoidance system which ensures coordination with multiple cooperative intruders. 19th Congress of the ICAS : 2208-18.

64. Krozel, J, T Mueller, and G Hunter. 1996. Free flight conflict detection and resolution analysis. AIAA guidance navigation and control conference: AIAA.

65. Gazit, Ran Y. and J. David Powell. 1996. Aircraft collision avoidance based on GPS position broadcasts. Digital Avionics Systems Conference, 1996., 15th AIAA/IEEE : 393-9.

66. Kremer, Harmen, Bert Bakker, and Henk Blom. 1997. Probabilistic versus geometric conflict probing. The 1st USA/Europe air traffic management R&D seminar. [http://atm-seminar-97.eurocontrol.fr/index.htm].

67. Eric W. Weisstein. 1999. Gaussian Distribution. World of mathematics [http://mathworld.wolfram.com/GaussianDistribution.html]

68. Gregory L. Wong. 1999 The Center-TRACON Automation System. [http://www.ctas.arc.nasa.gov/]

69. Eurocontrol. 1997. Inventory and analysis of A/G applications and data networks. COM.ET2.ST15: Phase 1 Report.

70. Boisvert, R. E. and V. A. Orlando. ADS-Mode S system overview. Digital avionics systems conference, 1994 12th DASC AIAA/IEEE, 104-9.

71. Harman, W., J. Gertz, and A. Kaminsky. 1998. Techniques for improved reception of 1090MHz ADS-B signals. Digital avionics systems conference, 1998 17th DASC AIAA/IEEE/SAE, G25-1-G25/9.

72. Gustavsson, Niclas. 1996. VDL Mode 4/STDMA - a CNS data link. AIAA,IEEE Digital Avionics Systems Conference and Technical Display 15th : 111-6.

73. VDL Mode 4 Sub Group (SG/2). 1997. VDL Mode 4 standards and recommended practices. Version 5.4. [http://www.lfv.se/ans/card/]

74. VDL Mode 4 Sub Group (SG/2). 1997. VDL Mode 4 Manaul. Version 1.0. [http://www.lfv.se/ans/card/]

75. Gunnar, Frisk, Bengt Moberg, and Tim Katanik. 1999. Integrating DGNSS and CNS for seamless gate-to gate services: System concept and results from approach and landing trial. ION GPS Conference 99. [http://www.lfv.se/ans/card/]

76. Link Evaluation Team. Safe Flight 21 Link Evaluation Team Phase One Report. [http://www.faa.gov/safeflight21/mainlinks/Downloads/letreport1199.pdf]. 1 July 2000

77. Last, David. May 1997. Radionavigation at a time of decision. Journal of Navigation 50, no. 2: 224-31.

78. Senechal, C. G. January 1993. The value of GNSS to aircraft operators. Journal of Navigation 46, no. 1: 69-77.

79. Gerold, Adrian. September 1997. What's ahead for GPS. Avionics Magazine 21, no. 9: 32-8.

80. Carleton, Wes. December 1997. Satellite risk interference. Jane 's Airport Review 9, no. 10: 31.

81. Hazelwood, Ed. 2 February 1998. FAA considering backup for GPS. Aviation Week & Space Technology 148, no. 5: 58-9.

82. Nordwall, Bruce. D. December 1997. Solar storms threaten GPS reception. Aviation Week & Space Technology 147, no. 22: 61-2.

83. Sharkey, Sarah and Rolf Johannessen. May 1997. Reliability performance in GPS receivers and the nature of their failures. Journal of Navigation 50, no. 2: 172-83

84. Bretz, Elizabeth. A. April 2000. X marks the spot, maybe. IEEE Spectrum 37, no. 4.

85. Merhav, Shmuel. 1996. Aerospace sensor systems and applications. Springer-Verlag New York, Inc.

86. Barbour, Neil and Schmidt George. 1998. Inertial sensor technology trends. Proceedings of the 1998 workshop on autonomous underwater vhehicles, 55-62.

87. SC-186. 1998. Minimum Aviation System Performance Standards for Automatic Dependent Surveillance Broadcast (ADS-B). RTCA, RTCA/DO-242.

88. FAA. 2000 Free Flight [http://www.faa.gov/freeflight/]

89. Asbury, M. J. A. and Johannessen, R. 1995. Single point of failure in complex aviation systems of communication, navigation and surveillance. Journal of Navigation. 48 no. 2:192-203.

APPENDIX A ACRONYMS AND ABB REVIATIONS

ADS	Automatic Dependent Surveillance
ANS	Air Navigation Systems
ATC	Air Traffic Control
ATM	Air Traffic Management
ATN	Aeronautical Telecommunication Network
ATS	Air Traffic Services
CAA	Civil Aviation Authorities
CNS	Communication, Navigation, Surveillance
DCM	Direction Cosine Matrix
DGPS	Differential Global Positioning System
DME	Distance Measuring Equipment
DOD	Department of Defence
DR	Dead Reckoning
EKF	Extended Kalman Filter
FANS	Future Air Navigation Systems
FMS	Flight Management Systems
GLONASS	Global Orbiting Navigation Satellite Systems
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
ICAO	International Civil Aviation Organisation
INS	Inertial Navigation System
JTIDS	Joint Tactical Information Distribution System
LKF	Linearised Kalman Filter
NATS	National Air Traffic Service

NDB	Non-Directional Beacons
RAIM	Receiver Autonomous Integrity Monitoring
RELNAV	Relative Navigation
RNAV	Area Navigation
RNP	Required Navigation Performance.
RTCA	Radio Technical Commission for Aeronautics
SA	Selective Availability
TCAS	Traffic Alert and Collision Avoidance System
US	The United State
VHF	Very High Frequency
VOR	VHF Omni-Directional Range

APPENDIX B VALIDATION OF THE SIMULATED NAVIGATION SYSTEM MODELS

This appendix illustrates some sample plots of the sensor error characteristics, used for the validation of the simulated navigation system models described in this thesis.

The INS model

All results for the strapdown INS have been obtained from 36-hour period simulations of a stationary INS with gyroscope bias (σ_{gyro}) of 0.01 °/hr and 10-hour time constant (τ_{gyro}). The power spectral density (PSD) of system white noise of the linear velocity sensor error is given by $q_{accel} = (3\,\mu g/\sqrt{\text{Hz}})^2$. Figure B-1 shows the inertial sensor noise samples and the theoretical rms error bounds of the strapdown INS sensor error models (1st order Gaussian Markov and random walk processes), described in Section 4.3.

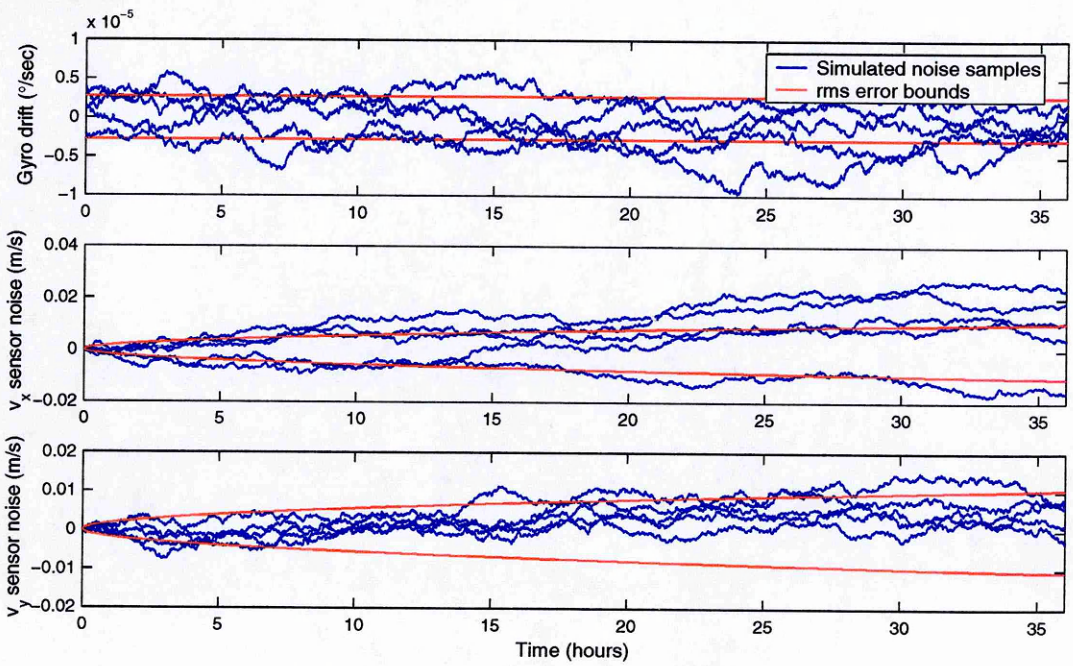


Figure B-1 Inertial sensor noise samples

Figure B-2 compares the true values (obtained from Monte Carlo simulations) and the theoretical values of the autocorrelation functions and the rms errors of the inertial sensor noise.

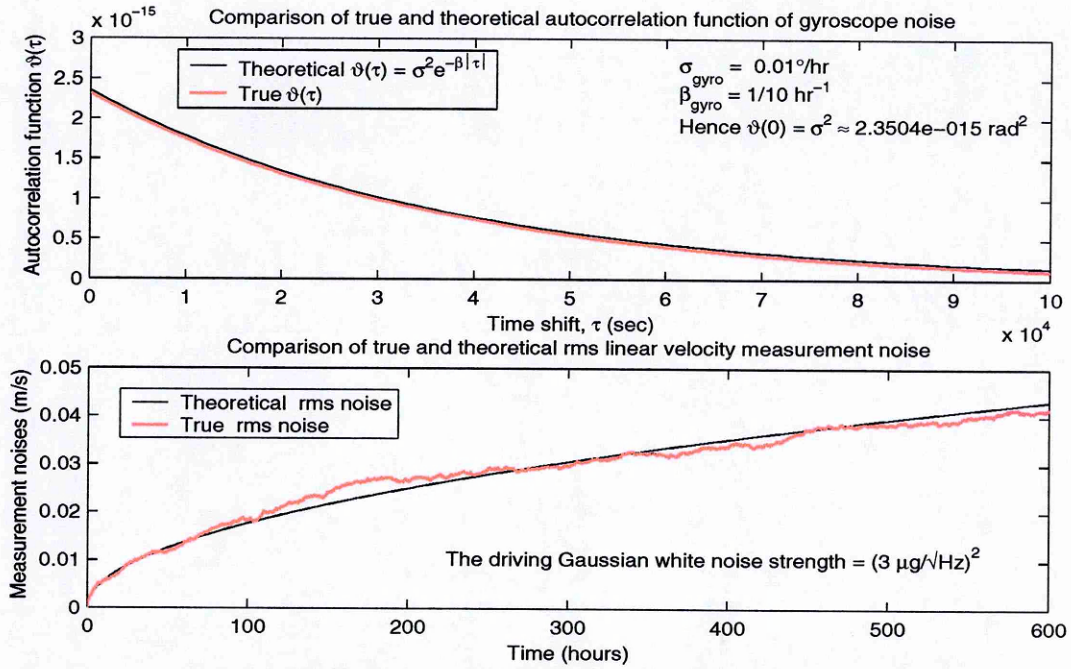


Figure B-2 The comparison of the true and the theoretical inertial sensor noise characteristics

Figure B-3 illustrates clock noise samples and rms error bounds of the clock error model, described in Section 4.4.1, for h_0 = 2e-20, h_{-1} = 7e-24 and h_{-2} = 4e-29.

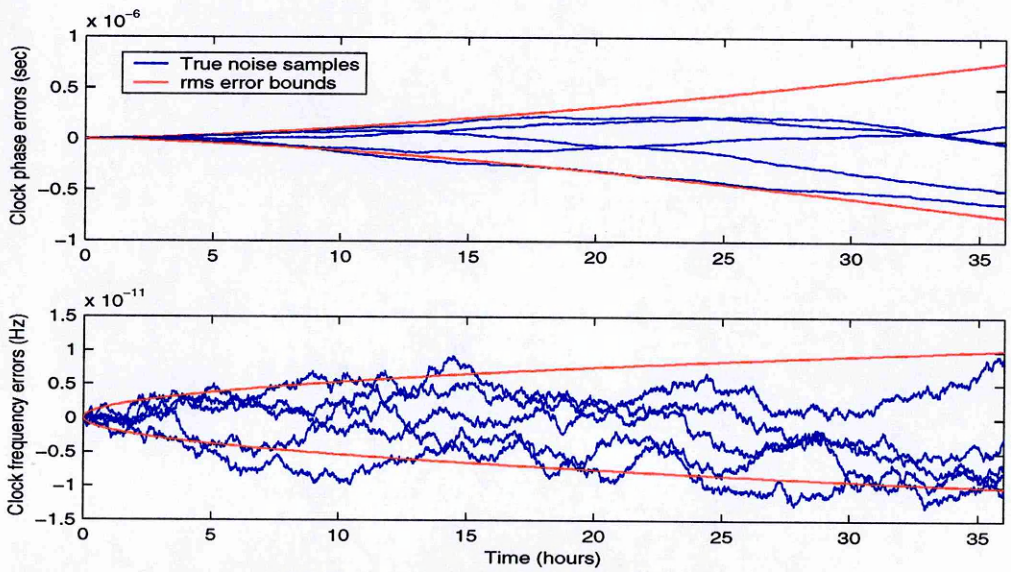


Figure B-3 The clock noise samples and rms error bounds

Based on the sensor error models simulated previously, the errors of a strapdown INS and an onboard clock are plotted as functions of time for 36-hour period as chown in Figure B-4.

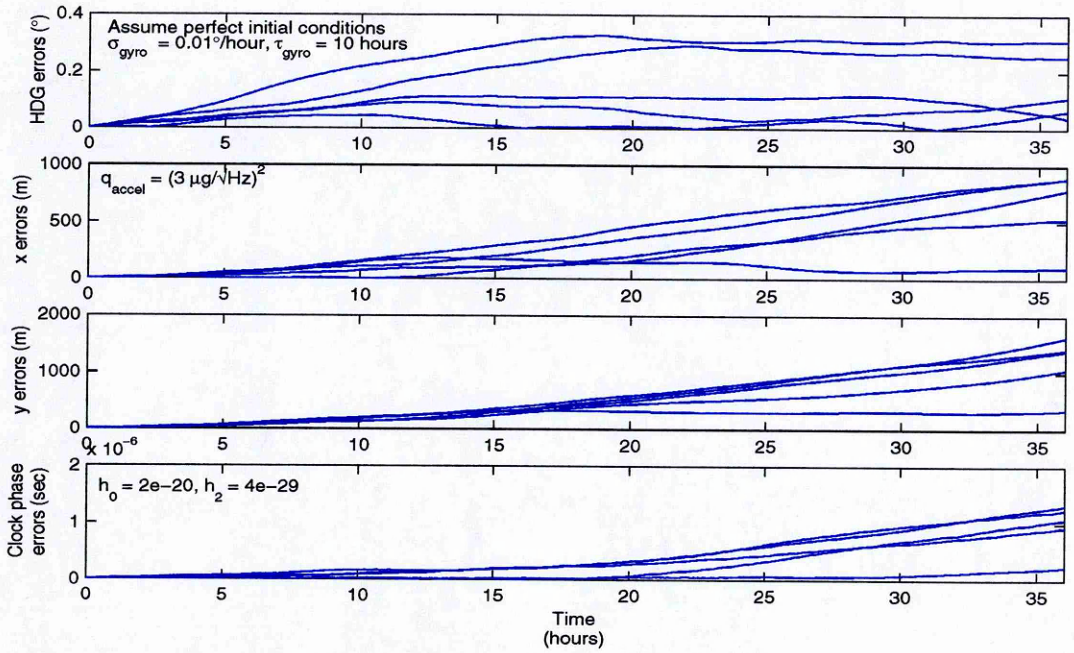


Figure B-4 The simulated navigational accuracy of a stationary strapdown INS and onboard clock (36-hour period)

The GPS model

The GPS model described in Section 4.4.1 has been validated by comparing the constellation, the ground traces and the GDOP of the simulated GPS satellites with the ideal cases described by Spilker[27] and Kaplan[28]. The constellation, ground traces and GDOP of the simulated GPS for a period of 24 hours are illustrated in Figure B-5, Figure B-6 and Figure B-7 respectively. It was found that the constellation, the ground traces and the GDOP of the simulated GPS satellites match very closely (qualitative) with the ideal cases.

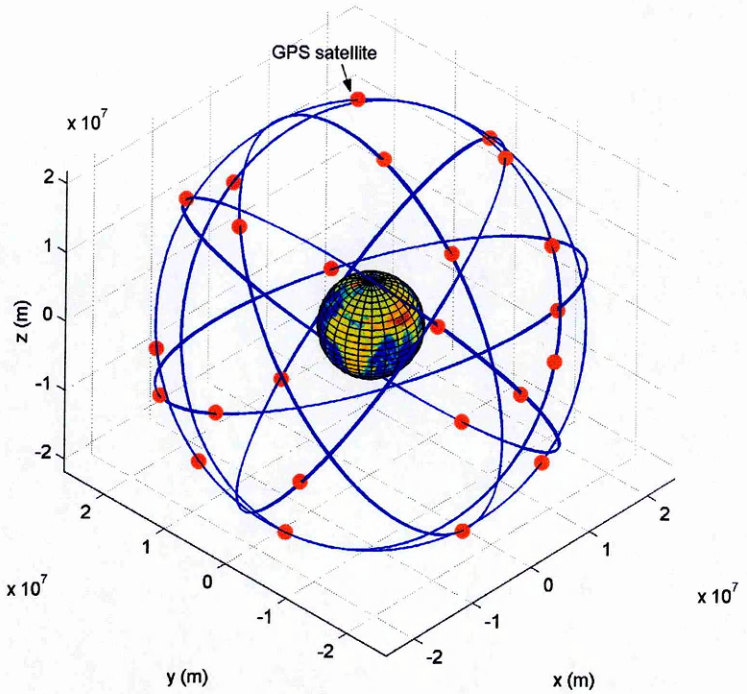


Figure B-5 The simulated GPS satellite constellation defined in the earth frame in a 24-hour period

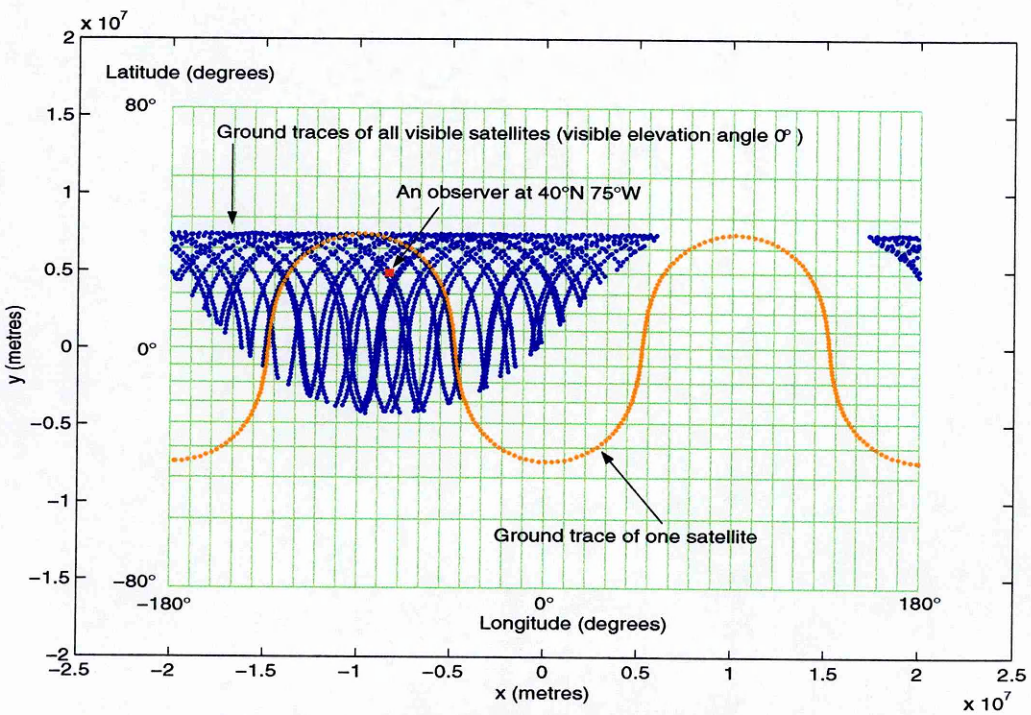


Figure B-6 Ground traces of the simulated GPS satellites as viewed in the simplified airspace in a 24-hour period

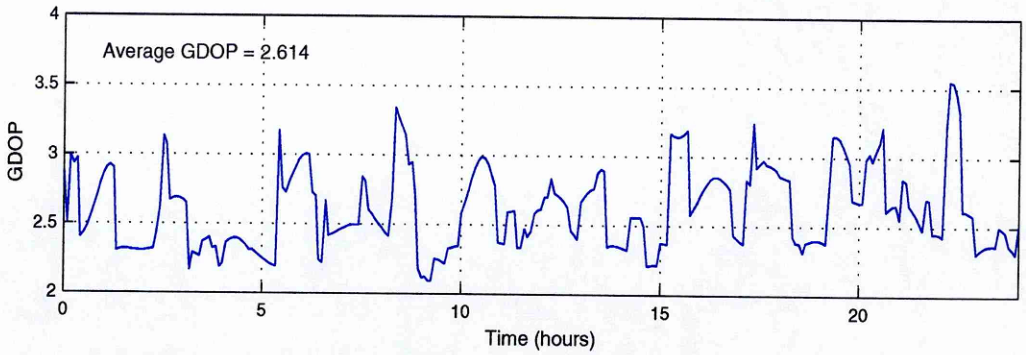


Figure B-7 GDOP of the simulated GPS satellites in a 24-hour period

APPENDIX C GPS SATELLITE TRAJ E CTORY SIMULATION PARAMETERS

This appendix lists the parameters for the simulation of the GPS satellite orbital trajectory based on typical navigation satellite geometry[11].

Satellite ID	$\Omega_0(^\circ)$	$\theta_0(^\circ)$
1	272.847	268.126
2	272.847	161.786
3	272.847	11.676
4	272.847	41.806
5	332.847	80.956
6	332.847	173.336
7	332.847	309.976
8	332.847	204.376
9	32.847	111.876
10	32.847	11.796
11	32.847	339.666
12	32.847	241.556
13	92.847	135.226
14	92.847	265.446
15	92.847	35.156
16	92.847	167.356
17	152.847	197.046
18	152.847	302.596
19	152.847	333.686
20	152.847	66.066
21	212.847	238.886
22	212.847	345.226
23	212.847	105.206
24	212.847	135.346

θ_0 and Ω_0 is the value of θ and Ω at the reference time t_0 which is midnight July 1, 1993.

APPENDIX D LEAST SQUARES EST IMATION

D.1 LEAST SQUARES SOLUTION OF $\mathbf{y} = \mathbf{Hx}$

This appendix reviews the least squares estimation[31]. Let a system be described by a set of linear equations $\mathbf{y} = \mathbf{Hx}$ where $\mathbf{y}$ is a m-by-1 vector containing m measurements. $\mathbf{x}$ is the n-by-1 system state vector to be determined from the measurements. $\mathbf{H}$ is a measurement matrix describing linear relationships between the system states and the measurements. For $m = n$, the basic solution for $\mathbf{x}$ is given by $\mathbf{x} = \mathbf{H}^{-1}\mathbf{y}$. However, in order to estimate the state vector $\hat{\mathbf{x}}$ for an overdetermined system; that is for $m > n$, let the estimation error $\mathbf{e}$ be defined as $\mathbf{e} = \mathbf{y} - \mathbf{H}\hat{\mathbf{x}}$. The *sum of the squares of the errors* $\|\mathbf{e}\|^2$ can be written as follows:

$$\|\mathbf{e}\|^2 = \mathbf{e}^{\mathrm{T}}\mathbf{e} = e_1^2 + e_2^2 + \ldots + e_n^2 \qquad \text{Eq. D-1}$$

The *least squares* solution is the solution that minimises $\|\mathbf{e}\|^2$ and can be obtained by substituting $\mathbf{e} = \mathbf{y} - \mathbf{H}\hat{\mathbf{x}}$ in Eq. D-1 as follows:

$$\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2 = \|\mathbf{H}\hat{\mathbf{x}} - \mathbf{y}\|^2 = (\mathbf{H}\hat{\mathbf{x}} - \mathbf{y})^{\mathrm{T}}(\mathbf{H}\hat{\mathbf{x}} - \mathbf{y}) \qquad \text{Eq. D-2}$$

Expanding Eq. D-2, noting that for real values of $\mathbf{y}$ and $\mathbf{H}\hat{\mathbf{x}}$, the order in the inner product is irrelevant, yields:

$$\begin{aligned}\|\mathbf{H}\hat{\mathbf{x}} - \mathbf{y}\|^2 &= \hat{\mathbf{x}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\hat{\mathbf{x}} - \hat{\mathbf{x}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{y} - \mathbf{y}^{\mathrm{T}}\mathbf{H}\hat{\mathbf{x}} + \mathbf{y}^{\mathrm{T}}\mathbf{y} \\ &= \hat{\mathbf{x}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\hat{\mathbf{x}} - 2\hat{\mathbf{x}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{y} + \mathbf{y}^{\mathrm{T}}\mathbf{y} \qquad \text{Eq. D-3}\end{aligned}$$

By differentiating Eq. D-3, setting the resultant equation to zero and solving it for $\hat{\mathbf{x}}$, it can be proved that the least squares estimate of the system state vector is given by:

$$\hat{\mathbf{x}} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{y} \qquad \text{Eq. D-4}$$

Algebraically, if the system $\mathbf{y} = \mathbf{Hx}$ is overdetermined, $\mathbf{y}$ can be derived by multiplying both sides by $\mathbf{H}^{\mathrm{T}}$ and solve the system $\mathbf{H}^{\mathrm{T}}\mathbf{y} = \mathbf{H}^{\mathrm{T}}\mathbf{Hx}$.

D.2 WEIGHTED LEAST SQUARES ESTIMATION

In the previous section, it was assumed that all measurements are equally reliable. However, there are cases when one measurement is more reliable than others. Under such circumstances, it is beneficial to rely on the more accurate measurements. This can be achieved by introducing a weighting matrix, **W**, where:

$$\|\mathbf{We}\|^2 = w_1^2 e_1^2 + w_2^2 e_2^2 + \ldots \qquad \text{Eq. D-5}$$

By multiplying both sides of $\mathbf{e} = \mathbf{y} - \mathbf{H}\hat{\mathbf{x}}$ by **W**, it can be seen that Eq. D-5 is the sum of squares of errors for the system $\mathbf{Wy} = \mathbf{WHx}$. By substituting **H** and **y** in Eq. D-4 with **WH** and **Wy**, the least squares solution for the system is given by:

$$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{CH})^{-1}\mathbf{H}^T\mathbf{Cy} \qquad \text{Eq. D-6}$$

where $\mathbf{C} = \mathbf{W}^T\mathbf{W}$.

It can be shown that, if the error **e** is zero mean i.e. $E[\mathbf{e}] = 0$ and has covariance matrix $\mathbf{R} = E[\mathbf{ee}^T]$, the best choice of **C** is $\mathbf{C} = \mathbf{R}^{-1}$. This results in a linear unbiased estimator given by:

$$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} \qquad \text{Eq. D-7}$$

When $\mathbf{C} = \mathbf{R}^{-1}$, the expected error in the estimate, measured by the covariance matrix $\mathbf{P} = E[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T]$, is also minimised and is given by:

$$\mathbf{P} = \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1} \qquad \text{Eq. D-8}$$

It should be noted that **P** is independent of the actual measurement **y**.

APPENDIX E RULES FOR EXPECTA TION AND VARIANCE

The rules for expectation and variance[52], which are used in the derivation of the initial error covariance matrices for the relative navigation Kalman filtering described in Appendix F, are summarised below. X, $\mathrm{E}[X]$, $\mathrm{Var}\ X$ and c denote a random variable X, the expectation of X, the variance of X and a constant respectively.

E.1 RULES FOR EXPECTATION

The expectation of any constant is that constant.

$$\mathrm{E}[c] = c \qquad \text{Eq. E-1}$$

Constants can be factored from expectations.

$$\mathrm{E}[cX] = c\mathrm{E}[X] \qquad \text{Eq. E-2}$$

The expectation of the sum of random variables is equal to the sum of the expectations whether or not the variables are independent (uncorrelated).

$$\mathrm{E}[X_1 + X_2 + \ldots + X_n] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \ldots + \mathrm{E}[X_n] \qquad \text{Eq. E-3}$$

The expectation of the product of random variables is equal to the product of the expectations, if the variables are independent.

$$\mathrm{E}[X_1 X_2 \ldots X_n] = \mathrm{E}[X_1] \cdot \mathrm{E}[X_2] \cdot \ldots \cdot \mathrm{E}[X_n] \qquad \text{Eq. E-4}$$

E.2 RULES FOR VARIANCE

1. The variance of any constant is zero.

$$\mathrm{Var}\ c = 0 \qquad \text{Eq. E-5}$$

Constants can be factored from variances.

$$\mathrm{Var}\ cX = c^2 \mathrm{Var}\ X \qquad \text{Eq. E-6}$$

The variance of the sum of independent random variables is equal to the sum of the variances.

$$\mathrm{Var}[X_1 + X_2 + \ldots + X_n] = \mathrm{Var}\ X_1 + \mathrm{Var}\ X_2 + \ldots + \mathrm{Var}\ X_n \qquad \text{Eq. E-7}$$

APPENDIX F RELNAV INITIAL ERROR VARIANCE CALCULATIONS

The equations, to derive the variances of system state vectors used as the diagonal terms of the initial error covariance matrix $\mathbf{P}_0$ for the relative navigation (RELNAV) Kalman filtering, are summarised below.

F.1 VARIANCES OF THE RELATIVE GRID STATE VECTORS

F.1.1 *Variance of the relative grid origin*

From Section 5.4.1, Eq. 5-5 can be written as follows:

$$\mathbf{r}_0^{\,i} = \mathbf{r}_{NC}^{\,i} - \mathbf{C}_r^i \mathbf{r}_{NC}^{\,r} \qquad \text{Eq. F-1}$$

The subscripts that denote the reference frames, which the vectors are defined with respect to, are omitted because they do not contribute to the analysis of stationary vector algebra. By substituting $\mathbf{r}_{NC}^{\,i} = \hat{\mathbf{r}}_{NC}^{\,i} + \Delta\mathbf{r}_{NC}^{\,i}$, where $\hat{\mathbf{r}}_{NC}^{\,i} = \mathbf{r}_{NC}^{\,r}$, and $\mathbf{C}_r^i \cong \mathbf{I} + [\boldsymbol{\psi}_{NC}\times]$ in Eq. F-1, yields:

$$\begin{aligned}\mathbf{r}_0^{\,i} &= \mathbf{r}_{NC}^{\,r} + \Delta\mathbf{r}_{NC}^{\,i} - (\mathbf{I} + [\boldsymbol{\psi}_{NC}\times])\mathbf{r}_{NC}^{\,r} \\ &= \Delta\mathbf{r}_{NC}^{\,i} - [\boldsymbol{\psi}_{NC}\times]\mathbf{r}_{NC}^{\,r}\end{aligned} \qquad \text{Eq. F-2}$$

Eq. F-2 can be written in component form as follows:[i]

$$U = \Delta r_u^{NC} + \psi_w^{NC} r_v^{NC} - \psi_v^{NC} r_w^{NC} \qquad \text{Eq. F-3}$$

$$V = \Delta r_v^{NC} + \psi_u^{NC} r_w^{NC} - \psi_w^{NC} r_u^{NC} \qquad \text{Eq. F-4}$$

$$W = \Delta r_w^{NC} + \psi_v^{NC} r_u^{NC} - \psi_u^{NC} r_v^{NC} \qquad \text{Eq. F-5}$$

Assuming that the error terms are uncorrelated and applying the rules of variances reviewed in Appendix E, Eq. F-3 can be used to find the variance of the relative grid origin in the u direction as follows:

$$\mathrm{Var}\,U = \mathrm{Var}\,[\Delta r_u^{NC} + \psi_w^{NC} r_v^{NC} - \psi_v^{NC} r_w^{NC}]$$

---

[i] U, V and W are the components of $\mathbf{r}_0^i$. The subscripts u, v and w denote the unit vectors of the i-frame.

$$= \mathrm{Var}\,\Delta r_u^{NC} + (r_v^{NC})^2 \mathrm{Var}\,\psi_w^{NC} + (r_w^{NC})^2 \mathrm{Var}\,\psi_v^{NC} \qquad \text{Eq. F-6}$$

Similarly, from Eq. F-4 and Eq. F-5, the variances of the relative grid origin in v and w direction can be expressed as follows:

$$\mathrm{Var}\,V = \mathrm{Var}\,\Delta r_v^{NC} + (r_w^{NC})^2 \mathrm{Var}\,\psi_u^{NC} + (r_u^{NC})^2 \mathrm{Var}\,\psi_w^{NC} \qquad \text{Eq. F-7}$$

$$\mathrm{Var}\,W = \mathrm{Var}\,\Delta r_w^{NC} + (r_u^{NC})^2 \mathrm{Var}\,\psi_v^{NC} + (r_v^{NC})^2 \mathrm{Var}\,\psi_u^{NC} \qquad \text{Eq. F-8}$$

F.1.2 *Variance of the relative grid alignment*

The relative grid alignment is a function of the Navigation Controller (NC)'s strapdown INS navigation frame orientation. Hence, the variance of the relative grid alignment is the variance of the NC's navigation frame misalignment, i.e.

$$\mathrm{Var}\,\psi_{NC} \qquad \text{Eq. F-9}$$

F.2 OBSERVING MEMBER ERROR VARIANCES

F.2.1 *Relative position error variance*

From Section 5.4.2, Eq. 5-21 can be written as follows:

$$\mathbf{r}_n^{\,r} = \mathbf{C}_i^r(\mathbf{r}_n^{\,i} - \mathbf{r}_0^{\,i}) \qquad \text{Eq. F-10}$$

where the subscripts that denote the reference frames, which the vectors are defined with respect to, are omitted. The estimated relative position is given by:

$$\hat{\mathbf{r}}_n^{\,r} = \hat{\mathbf{r}}_n^{\,i} \qquad \text{Eq. F-11}$$

Substituting Eq. F-10, Eq. F-11, $\mathbf{C}_i^r \cong \mathbf{I} - [\psi_{NC}\times]$ and $\hat{\mathbf{r}}_n^{\,i} = \mathbf{r}_n^{\,i} - \Delta\mathbf{r}_n^{\,i}$ in $\Delta\mathbf{r}_n^{\,r} = \mathbf{r}_n^{\,r} - \hat{\mathbf{r}}_n^{\,r}$, yields:

$$\Delta\mathbf{r}_n^{\,r} = (\mathbf{I} - [\psi_{NC}\times])(\mathbf{r}_n^{\,i} - \mathbf{r}_0^{\,i}) - (\mathbf{r}_n^{\,i} - \Delta\mathbf{r}_n^{\,i}) \qquad \text{Eq. F-12}$$

Ignoring the product of error terms, Eq. F-12 can be written as follows:

$$\Delta\mathbf{r}_n^{\,r} = -[\psi_{NC}\times]\mathbf{r}_n^{\,i} - \mathbf{r}_0^{\,i} + \Delta\mathbf{r}_n^{\,i} \qquad \text{Eq. F-13}$$

Eq. F-13 can be written in component form as follows:

$$\Delta \mathbf{r}_u^n = \psi_w^n \mathbf{r}_v^n - \psi_v^n \mathbf{r}_w^n - U + \Delta r_u^n \qquad \text{Eq. F-14}$$

$$\Delta \mathbf{r}_v^n = \psi_u^n \mathbf{r}_w^n - \psi_w^n \mathbf{r}_u^n - V + \Delta r_v^n \qquad \text{Eq. F-15}$$

$$\Delta \mathbf{r}_w^n = \psi_v^n \mathbf{r}_u^n - \psi_u^n \mathbf{r}_v^n - W + \Delta r_w^n \qquad \text{Eq. F-16}$$

Assuming that the error terms are uncorrelated, Eq. F-14 can be used to find the relative position error variance for the n^{th} observing member in the u direction as follows:

$$\begin{aligned} \operatorname{Var} \Delta \mathbf{r}_u^n &= \operatorname{Var} [\psi_w^n \mathbf{r}_v^n - \psi_v^n \mathbf{r}_w^n - U + \Delta r_u^n] \\ &= (\mathbf{r}_v^n)^2 \operatorname{Var} \psi_w^n + (\mathbf{r}_w^n)^2 \operatorname{Var} \psi_v^n + \operatorname{Var} U + \operatorname{Var} \Delta r_u^n \end{aligned} \qquad \text{Eq. F-17}$$

Similarly, from Eq. F-15 and Eq. F-16, the variances of the relative position error in the v and w direction can be expressed as follows:

$$\operatorname{Var} \Delta \mathbf{r}_v^n = (\mathbf{r}_u^n)^2 \operatorname{Var} \psi_w^n + (\mathbf{r}_w^n)^2 \operatorname{Var} \psi_u^n + \operatorname{Var} V + \operatorname{Var} \Delta r_v^n \qquad \text{Eq. F-18}$$

$$\operatorname{Var} \Delta \mathbf{r}_w^n = (\mathbf{r}_u^n)^2 \operatorname{Var} \psi_v^n + (\mathbf{r}_v^n)^2 \operatorname{Var} \psi_u^n + \operatorname{Var} W + \operatorname{Var} \Delta r_w^n \qquad \text{Eq. F-19}$$

F.2.2 *Relative misalignment error variance*

From Section 5.4.2, the misalignment error variance of the n^{th} observing member from the relative grid orientation can be determined directly from Eq. 5-18 as follows:

$$\operatorname{Var} \boldsymbol{\alpha}_n = \operatorname{Var} \boldsymbol{\Psi}_n - \operatorname{Var} \boldsymbol{\Psi}_{NC} \qquad \text{Eq. F-20}$$

APPENDIX G DERIVATION OF THE OWNSTATE DECENTRALISED RELNAV STATE ESTIMATOR

This appendix derives the least-squares state estimator for the ownstate decentralised relative navigation. The overall system model is given by:

$$\mathbf{x}_{ok+1} = \mathbf{\Phi}_k \mathbf{x}_{ok} + \mathbf{w}_k \quad \text{Eq. G-1}$$

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_{ok} - \mathbf{u}_k + \mathbf{v}_k \quad \text{Eq. G-2}$$

where $\mathbf{x}_o$ is the observer error state vector. It will be assumed that the plant and measurement noise is zero-mean, uncorrelated in time and are described as follows:

$$\mathrm{E}[\mathbf{w}_k \mathbf{w}_i^{\mathrm{T}}] = \begin{cases} \mathbf{Q}_k \text{ for } i = k \\ \mathbf{0} \text{ for } i \neq k \end{cases} \quad \text{Eq. G-3}$$

$$\mathrm{E}[\mathbf{v}_k \mathbf{v}_i^{\mathrm{T}}] = \begin{cases} \mathbf{R}_k \text{ for } i = k \\ \mathbf{0} \text{ for } i \neq k \end{cases} \quad \text{Eq. G-4}$$

$$\mathrm{E}[\mathbf{w}_k \mathbf{v}_i^{\mathrm{T}}] = \mathbf{0} \text{ for all } i \text{ and } k \quad \text{Eq. G-5}$$

$\mathbf{u}_k$ is the additional measurement noise defined in Eq. 6-3, which models the uncertainties in the sources. $\mathbf{u}_k$ is also assumed to be white and uncorrelated with $\mathbf{v}_k$ and $\mathbf{w}_k$. The estimation error covariance matrix of the observer is given by:

$$\mathbf{P}_{o\,k}^{+} = \mathrm{E}[(\mathbf{x}_{ok} - \hat{\mathbf{x}}_{ok}^{+})(\mathbf{x}_{ok} - \hat{\mathbf{x}}_{ok}^{+})^{\mathrm{T}}] \quad \text{Eq. G-6}$$

The standard Kalman filter update equation is given by:

$$\hat{\mathbf{x}}_k^{+} = \hat{\mathbf{x}}_k^{-} + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_k^{-}) \quad \text{Eq. G-7}$$

From equations Eq. G-2, based on the ownstate formulation, the predicted measurement can be expressed as follows:

$$\hat{\mathbf{z}}_k^{-} = \mathbf{H}_k \hat{\mathbf{x}}_{ok}^{-} \quad \text{Eq. G-8}$$

The estimation error states can be determined by substituting Eq. G-2 and Eq. G-8 in Eq. G-7 as follows:

$$\hat{\mathbf{x}}_{ok}^{+} = \hat{\mathbf{x}}_{ok}^{-} + \mathbf{K}_k[(\mathbf{H}_k \mathbf{x}_{ok} - \mathbf{u}_k + \mathbf{v}_k) - \mathbf{H}_k \hat{\mathbf{x}}_{ok}^{-}] \quad \text{Eq. G-9}$$

Substituting Eq. G-9 in Eq. G-6 yields:

$$\begin{aligned}\mathbf{P}_{o\,k}^{+} = \mathrm{E}[(\mathbf{x}_{ok} - \{\hat{\mathbf{x}}_{ok}^{-} + \mathbf{K}_k[(\mathbf{H}_k\mathbf{x}_{ok} - \mathbf{u}_k + \mathbf{v}_k) - \mathbf{H}_k\hat{\mathbf{x}}_{ok}^{-}]\}) \\ (\mathbf{x}_{ok} - \{\hat{\mathbf{x}}_{ok}^{-} + \mathbf{K}_k[(\mathbf{H}_k\mathbf{x}_{ok} - \mathbf{u}_k + \mathbf{v}_k) - \mathbf{H}_k\hat{\mathbf{x}}_{ok}^{-}]\})^{\mathrm{T}}]\end{aligned} \qquad \text{Eq. G-10}$$

Substituting $\mathbf{e}_{ok}^{-} = \mathbf{x}_{ok} - \hat{\mathbf{x}}_{ok}^{-}$ in Eq. G-10 yields:

$$\begin{aligned}\mathbf{P}_{o\,k}^{+} = \mathrm{E}[\{\mathbf{e}_{ok}^{-} - \mathbf{K}_k(\mathbf{H}_k\mathbf{e}_{ok}^{-} - \mathbf{u}_k + \mathbf{v}_k)\} \\ \{\mathbf{e}_{ok}^{-} - \mathbf{K}_k(\mathbf{H}_k\mathbf{e}_{ok}^{-} - \mathbf{u}_k + \mathbf{v}_k)\}^{\mathrm{T}}]\end{aligned} \qquad \text{Eq. G-11}$$

Eq. G-11 can be expanded as follows:

$$\begin{aligned}\mathbf{P}_{o\,k}^{+} = \mathrm{E}[&\mathbf{e}_{ok}^{-}\mathbf{e}_{ok}^{-\mathrm{T}} - \mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-}\mathbf{e}_{ok}^{-\mathrm{T}} + \mathbf{K}_k\mathbf{u}_k\mathbf{e}_{ok}^{-\mathrm{T}} - \mathbf{K}_k\mathbf{v}_k\mathbf{e}_{ok}^{-\mathrm{T}} \\ &- \mathbf{e}_{ok}^{-}(\mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-})^{\mathrm{T}} + \mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-}(\mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-})^{\mathrm{T}} \\ &- \mathbf{K}_k\mathbf{u}_k(\mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-})^{\mathrm{T}} + \mathbf{K}_k\mathbf{v}_k(\mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-})^{\mathrm{T}} \\ &+ \mathbf{e}_{ok}^{-}(\mathbf{K}_k\mathbf{u}_k)^{\mathrm{T}} - \mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-}(\mathbf{K}_k\mathbf{u}_k)^{\mathrm{T}} \\ &+ \mathbf{K}_k\mathbf{u}_k(\mathbf{K}_k\mathbf{u}_k)^{\mathrm{T}} - \mathbf{K}_k\mathbf{v}_k(\mathbf{K}_k\mathbf{u}_k)^{\mathrm{T}} \\ &- \mathbf{e}_{ok}^{-}(\mathbf{K}_k\mathbf{v}_k)^{\mathrm{T}} + \mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-}(\mathbf{K}_k\mathbf{v}_k)^{\mathrm{T}} \\ &- \mathbf{K}_k\mathbf{u}_k(\mathbf{K}_k\mathbf{v}_k)^{\mathrm{T}} + \mathbf{K}_k\mathbf{v}_k(\mathbf{K}_k\mathbf{v}_k)^{\mathrm{T}}]\end{aligned} \qquad \text{Eq. G-12}$$

Substituting $\mathbf{P}_{o\,k}^{-} = \mathrm{E}[\mathbf{e}_{ok}^{-}\mathbf{e}_{ok}^{-\mathrm{T}}]$, $\mathbf{S}_k = \mathrm{E}[\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}}]$ and $\mathbf{R}_k = \mathrm{E}[\mathbf{v}_k\mathbf{v}_k^{\mathrm{T}}]$ in Eq. G-12 yields:

$$\mathbf{P}_{o\,k}^{+} = \mathbf{P}_1 + \mathbf{P}_2 + \mathbf{P}_3 \qquad \text{Eq. G-13}$$

where

$$\begin{aligned}\mathbf{P}_1 = &\mathbf{P}_{o\,k}^{-} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{o\,k}^{-} - \mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} \\ &+ \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^{\mathrm{T}}\end{aligned} \qquad \text{Eq. G-14}$$

$$\begin{aligned}\mathbf{P}_2 = \mathrm{E}[&-\mathbf{K}_k\mathbf{v}_k\mathbf{e}_{ok}^{-\mathrm{T}} + \mathbf{K}_k\mathbf{v}_k\mathbf{e}_{ok}^{-\mathrm{T}}\mathbf{H}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} - \mathbf{K}_k\mathbf{v}_k\mathbf{u}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} - \mathbf{e}_{ok}^{-}\mathbf{v}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} \\ &+ \mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-}\mathbf{v}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} - \mathbf{K}_k\mathbf{u}_k\mathbf{v}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}}]\end{aligned} \qquad \text{Eq. G-15}$$

$$\begin{aligned}\mathbf{P}_3 = \mathrm{E}[&\mathbf{K}_k\mathbf{u}_k\mathbf{e}_{ok}^{-\mathrm{T}} - \mathbf{K}_k\mathbf{u}_k\mathbf{e}_{ok}^{-\mathrm{T}}\mathbf{H}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} + \mathbf{e}_{ok}^{-}\mathbf{u}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} \\ &- \mathbf{K}_k\mathbf{H}_k\mathbf{e}_{ok}^{-}\mathbf{u}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}}]\end{aligned} \qquad \text{Eq. G-16}$$

$\mathbf{P}_2$ is zero because the measurement noise $\mathbf{v}_k$ is assumed to be white. $\mathbf{P}_3$ is also zero based on the assumption that the errors of the source are white. Hence, the estimation error covariance matrix is expressed as follows:

$$\mathbf{P}_{o\,k}^{+} = \mathbf{P}_{o\,k}^{-} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{o\,k}^{-} - \mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^{\mathrm{T}} \qquad \text{Eq. G-17}$$

Eq. G-17 can be rearranged as follows:

$$\begin{aligned}\mathbf{P}_{o\,k}^{+} &= (\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{o\,k}^{-} - (\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^{\mathrm{T}} \\ &= (\mathbf{I}-\mathbf{K}_k\mathbf{h}_k)\mathbf{P}_{o\,k}^{-}(\mathbf{I}-\mathbf{h}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}}) + \mathbf{K}_k\mathbf{h}_k\mathbf{P}_{s\,k}^{-}\mathbf{h}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^{\mathrm{T}} \\ &= (\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{o\,k}^{-}(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}} + \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^{\mathrm{T}} + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^{\mathrm{T}} \qquad \text{Eq. G-18}\end{aligned}$$

The gain $\mathbf{K}_k$ that minimises the least-squares errors can be determined by solving the following differential equation:

$$\frac{\partial\,\mathrm{trace}(\mathbf{P}_{o\,k}^{+})}{\partial\mathbf{K}_k} = \mathbf{0} \qquad \text{Eq. G-19}$$

Additionally, the following equations[11] are applied in order to solve Eq. G-19.

$$\frac{\partial\,\mathrm{trace}(\mathbf{AB})}{\partial\mathbf{A}} = \mathbf{B}^{\mathrm{T}} \qquad \text{where } \mathbf{AB} \text{ must be square} \qquad \text{Eq. G-20}$$

$$\frac{\partial\,\mathrm{trace}(\mathbf{ACA}^{\mathrm{T}})}{\partial\mathbf{A}} = 2\mathbf{AC} \qquad \text{where } \mathbf{C} \text{ must be symmetric} \qquad \text{Eq. G-21}$$

$$\frac{d\,s}{d\mathbf{A}} = \begin{bmatrix} \dfrac{d\,s}{da_{11}} & \dfrac{d\,s}{da_{12}} & \cdots \\ \dfrac{d\,s}{da_{21}} & \ddots & \\ \vdots & & \end{bmatrix} \qquad \text{Eq. G-22}$$

Noting that trace($\mathbf{P}_k\mathbf{H}_k^{\mathrm{T}}\mathbf{K}_k^{\mathrm{T}}$) = trace($\mathbf{H}_k\mathbf{K}_k\mathbf{P}_k^{\mathrm{T}}$), and $\mathbf{P}_k$ is symmetric such that $\mathbf{P}_k = \mathbf{P}_k^{\mathrm{T}}$, it can be shown that, by using Eq. G-17, Eq. G-19 can be expressed as follows:

$$\frac{\partial\mathrm{trace}(\mathbf{P}_{o\,k}^{+})}{\partial\mathbf{K}_k} = -2\mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}} + 2\mathbf{K}_k[\mathbf{H}_k\mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}} + \mathbf{S}_k + \mathbf{R}_k] = \mathbf{0} \qquad \text{Eq. G-23}$$

Therefore, the gain $\mathbf{K}_k$ is given by:

$$\mathbf{K}_k = \mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}}[\mathbf{H}_k\mathbf{P}_{o\,k}^{-}\mathbf{H}_k^{\mathrm{T}} + \mathbf{S}_k + \mathbf{R}_k]^{-1} \qquad \text{Eq. G-24}$$

APPENDIX H DISCRIPTION OF TH E SIMULATION PROGRAM

H.1 INTRODUCTION

The simulation program was developed in Matlab Version 5.3.0.10183 (R11) and can be obtained by contacting:

Prof. David Allerton
College of Aeronautic
Cranfield University
Cranfield
Bedford MK43 0AL
United Kingdom
Email: dja@cranfield.ac.uk

or

Mr. Krisada Sangpetchsong
Email: aim12@mail.com

H.2 LIST OF PROGRAMS

Subdirectory **\ate**

autosim.m	Main simulation program
set_simu_path.m	Set Matlab paths to relevant subdirectories
startup.m	Initialise Matlab workspace for the simulation

Subdirectory **\ate\analysis_tb**

analyse_cenrn.m	Perform centralised RELNAV analysis
analyse_conflict_prediction.m	Perform conflict detection analysis
analyse_decrn.m	Perform decentralised RELNAV analysis
analyse_dr.m	Perform stand-alone INS analysis
analyse_drgps.m	Perform integrated INS/GPS analysis
analyse_gps.m	Perform stand-alone GPS analysis
analyse_rngps.m	Perform integrated INS/GPS/RELNAV analysis

Subdirectory **\ate\bada_tb**

init_bada.m	Load BADA data

bada_data.mat	BADA data file in Matlab format

Subdirectory **\ate\conflict_tb**

erf.m	Calculate the error function
nor_prob.m	Calculate the cumulative probability of the normal random variables
predict_trajectory.m	Predict flight path for conflict detection

Subdirectory **\ate\coor_tb**

ecef2geod.m	Transform a vector expressed in Earth-Centred Earth-Fixed coordinate system to Geodetic coordinate system (Latitude Longitude and Altitude)
geod2ecef.m	Transform a vector expressed in Geodetic coordinate system to Earth-Centred Earth-Fixed coordinate system
geod2grid.m	Transform a vector expressed in Geodetic coordinate system to map XYZ using Mercator projection
grid2geod.m	Transform a vector expressed in map XYZ using Mercator projection to Geodetic coordinate system

Subdirectory **\ate\dr_tb**

measure_dr.m	Model strapdown INS sensor measurement
update_dr.m	Strapdown INS attitude and position update algorithm

Subdirectory **\ate\gps_tb**

gps_ecef.m	Calculate GPS satellite Earth-Centred Earth-Fixed (ECEF) coordinates
gps_ranges.m	Determine the true GPS pseudorange measurements

init_gps.m	Initialise global variables for the use of GPS toolbox
point_solution.m	Determine GPS single point solutions

Subdirectory **\ate\integrity_tb**

find_lhfunc_ac.m	Calculate the likelihood function for independent (white) zero-mean Gaussian random variables

Subdirectory **\ate\kalman_tb**

dkalman.m	Discrete Kalman filtering algorithm
vanloan.m	Approximate the discrete-time parameters for discrete Kalman filtering

Subdirectory **\ate\plot_tb**

plot_ac_states.m	Plot simulated aircraft states
plot_iframe.m	Plot the simplified airspace

Subdirectory **\ate\relnav_tb**

determine_rank_band.m	Determine ranking intervals from covariance matrices
find_relnav_rank_vector.m	Determine the ranks for RELNAV
member_ranges.m	Determine true relative pseudorange measurements between each aircraft
measure_ads_b_ranges.m	Simulate relative pseudorange measurement process

Subdirectory **\ate\simulation_tb**

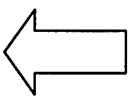
find_control.m	Determine command heading
init_ac.m	Initialise aircraft
init_simulation.m	Initialise simulation
simulation.m	Simulation routine
update_state.m	Update simulated aircraft states

H.3 QUICK REFERENCE TO RUNNING THE SIMULATION PROGRAM

H.3.1 *Running the simulation program*

To run the simulation program load Matlab, change the current working directory to **ate** subdirectory, run **startup.m** to initialise the simulation, and run **autosim.m**.

```
» cd \
» cd ate
» startup
» autosim
```



Matlab commands

H.3.2 *Initialising the simulation*

Before the simulation can be run successfully, **initialisation file**(s) must be defined in the **\ate\data\init_simulation** subdirectory. An initialisation file is a text file containing global simulation parameters such as simulation length, sample interval, the numbers of aircraft to be simulated, aircraft types, flight paths, sensor qualities, etc. Each initialisation file can be regarded as one specific simulation scenario. Several scenarios can be run as a batch processing by defining several initialisation files in the **\ate\data\init_simulation** subdirectory. This enables several scenarios to be run automatically, for example, overnight. The results (Matlab figure files, i.e. *.fig) for each simulation scenario will be stored in **\ate\data\results** subdirectory. An example initialisation file is shown below.

```
CC Initialisation file containing simulation global parameters
CC The order of appearance of these parameters in the initialisation file is important.
CC total simulation time (sec)
CD 10
CC simulation update rate (sec)
CD 1
CC [Total Monte Carlo run, reset GPS t0 flag]
CC reset GPS t0 flag = 1 resets the t0 so that GPS satellites geometry change for every Monte Carlo run
CD 1, 0
CC stand alone DR error covariance update rate (sec)
CD 1
CC stand alone GPS update rate (sec)
CD 1
CC Integrated DR/GPS update rate (sec)
CD 1
CC ADS-B update rate, Non NC update rate for decentralised RELNAV (sec)
CD 3, 1
CC Maximum ADS-B ranges (nautical miles)
CC set to 0 if do not want RELNAV capability
CD 40
CC Error covariance multiplying factor
CC [1sigma (60%), 2sigma (95%), 3sigma (99%)]
CD 1
CC GPS pseudorange measurement
CC white noises standard deviation (m)
```

```
CD 33.3
CC GPS failures parameters
CC assumed failure to be normal distribution additional pseudorange measurement errors
CC [failed SV ID, mean fail noise (m), 1 sigma fail noise, (m) mean fail time (s), 1 sigma fail time (s), NFS]
CC NFS = number of failed satellites allowed per simulation
CC setting NFS to zero if do NOT want satellite failure
CC If the provided failed SV ID is not in the range [1:24], it is selected by random
CC If failed SV = 0 failed SV is selected randomly from [1:24]
CC If failed SV = -1 failed SV is selected randomly from a list of visible satellites
CD 4, 50, 5, 150, 0, 0
CC Unavailable satellites ID [1:24]
CC These satellites will not be available to the observer. It is used to limit the number of visible satellites
CC 0
CC set to 1:24 if do not want GPS
CD 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24
CC 12, 24, 17, 23, 14, 11, 7, 10, 4
CC ADS-B pseudorange measurement
CC white noises standard deviation (m)
CD 20
CC Randomise ADS-B interval flag [0 = not random, 1 = random]
CD 1
CC Decentralised RELNAV Modify Wt mode [0 = no modification, 3 = modify by increasing all W]
CD 0
CC RELNAV/GPS/ADS-B Modify Wt mode [0 = no modification], Number of available ranks, Additional Rk (m^2)
CD 0, 1, 36
CC**********************************
CC Integrity monitoring parameters
CC**********************************
CC [mode option]
CC for mode: [0=no monitoring] [1=Likelihood function],  number of samples
CD 1, 15
CC**********************************
CC Plotting axis limits
CC**********************************
CC Axis limits to be used with
CC AXIS([XMIN XMAX YMIN YMAX])
CD -4e4, 6e4, -7e4, 9e4
CC**********************************
CC Conflict prediction parameters
CC**********************************
CC [minimum separation standard, conflict prediction update rate, conflict prediction look forward time, time k for
analyse conflict prediction]
CC distance in Nm, time in seconds
CD 5, 1, 600, 1
CC**********************************
CC Flight plans
CC**********************************
CD B777,ftp032,clockM,drL,initL
CD B777,ftp033,clockM,drL,initL
CC//
```

Additional simulation parameters which are contained in the **autosim.m** can also be used to control the simulation. Parameters of aircraft navigation systems are specified in data files (*.txt) in the following subdirectories:

\ate\data\clock	Clock accuracy levels
\ate\data\dr	INS accuracy levels
\ate\data\fltplan	Flight plans
\ate\data\init_simulation	Initialisation files
\ate\data\results	Simulation results

H.4 MATLAB SOURCE CODES FOR THE HEADING CONTROL ALGORITHM

The following source codes are extracted from function **find_control.m** in **\ate\simulation**. This function controls aircraft's manoeuvres between waypoints.

```
for ac_id = simu.ac_id{k} % process data for every simulated aircraft one by one

  ac(ac_id).tas(1,k) = ac(ac_id).tas(1,1); % true air speed, multiply by 0 to make stationary aircraft
  ac(ac_id).rocd(1,k) = ac(ac_id).rocd(1,1); % rate of climb or descent
  test_at_wpt(ac_id,k); % test if an aircraft arrives at a waypoint
  find_hdg_cmd(ac_id,k); % find heading command i.e. heading rate
  theta = pi/2-ac(ac_id).hdg(1,k);  % change from north 0 deg to east 0 deg
  ac(ac_id).Cbi(:,:,k) = [sin(theta) cos(theta) 0;-cos(theta) sin(theta) 0;0 0 1]; % direction cosine matrix
  ac(ac_id).v(:,k) = ac(ac_id).Cbi(:,:,k)*[0;ac(ac_id).tas(1,k);ac(ac_id).rocd(1,k)];

end


%-------------------------------------------------------
function test_at_wpt(ac_id,k)

% determine whether the aircraft is at the next WPT
% return the demanded heading (rad)

global ac simu

present_position = ac(ac_id).pos(:,k);

if  ac(ac_id).WPT.next < ac(ac_id).WPT.total % not travelling to the final WPT hence include distance to turn
  % Ref. Siouris, George M. Aerospace avionics systems: a modern synthesis.
  % The United State of America: Academic Press Limited; 1993; ISBN: 0-12-646890-7. p. 383
  % define 2 vectors (v01 and v02) between 3 waypoints in order to calculate required heading change
  v = diff(ac(ac_id).WPT.pos(:,[ac(ac_id).WPT.next-1:ac(ac_id).WPT.next+1]),1,2);  % v01 and v12
  true_track = find_hdg(v(2,:),v(1,:),'2pi'); % find true heading to waypoints
  d_track = true_track(2)-true_track(1); % calculate heading change
  distance_to_go = magnitude(ac(ac_id).WPT.pos(:,ac(ac_id).WPT.next)-present_position,1);
  distance_to_turn = ((ac(ac_id).tas(1,k)^2)/9.81)*abs(tan(d_track/2)); % tan(45*pi/180)*9.81 = 9.81
  % Check if the aircraft has passed the WPT
  p = [ac(ac_id).WPT.pos(:,ac(ac_id).WPT.next-1:ac(ac_id).WPT.next), present_position,...
      ac(ac_id).WPT.pos(:,ac(ac_id).WPT.next-1)];
  dp = diff(p,1,2);
  abc = magnitude(dp,1);
  C = acos((abc(1)^2+abc(2)^2-abc(3)^2)/(2*abc(1)*abc(2)));
  %end check
  if (distance_to_go <= distance_to_turn) | (C >= pi/2)
    ac(ac_id).WPT.next = ac(ac_id).WPT.next+1; % arrived at the turn point or already passed the next waypoint
    demanded_hdg = true_track(2); % the heading to the next waypoint+1 (by passing the next waypoint)
  else
    demanded_hdg = true_track(1); % the current heading to the next waypoint
  end
else % final waypoint
  % test by checking if the aircraft has passed the WPT i.e. 90 deg threshold
  p = [ac(ac_id).WPT.pos(:,ac(ac_id).WPT.next-1:ac(ac_id).WPT.next), present_position,...
      ac(ac_id).WPT.pos(:,ac(ac_id).WPT.next-1)];
  dp = diff(p,1,2);
  abc = magnitude(dp,1);
  C = acos((abc(1)^2+abc(2)^2-abc(3)^2)/(2*abc(1)*abc(2)));
  if C >= pi/2
    ac(ac_id).sim_end = k; % define the final k for ac(ac_id)
  end
  demanded_hdg = find_hdg(dp(2,1),dp(1,1),'2pi');
end
ac(ac_id).demanded_hdg(1,k) = demanded_hdg;

%-------------------------------------------------------
function find_hdg_cmd(ac_id,k)

% Calculate the heading command

global ac

present_hdg = ac(ac_id).hdg(1,k);

d_track = ac(ac_id).demanded_hdg(1,k) - present_hdg;
if abs(d_track)>pi
  d_track = d_track-sign(d_track)*2*pi;
end
```

```
if ac(ac_id).tas(1,k) ~= 0
  hdg_rate_max = 9.81/ac(ac_id).tas(1,k); % tan(45*pi/180)*9.81 = 9.81
else
  hdg_rate_max = inf; % stationary aircraft
end
if abs(d_track) > hdg_rate_max % @ bank 45 deg
  ac(ac_id).hdg_cmd(1,k) = sign(d_track)*hdg_rate_max;
  ac(ac_id).bank(1,k) = sign(ac(ac_id).hdg_cmd(1,k))*0. 7853; % 45*pi/180 = 0.7853 and +ve roll = right turn
else
  ac(ac_id).hdg_cmd(1,k) = d_track;
  ac(ac_id).bank(1,k) = sign(ac(ac_id).hdg_cmd(1,k))*atan(ac(ac_id).hdg_cmd(1,k)*ac(ac_id).tas(1,k)/9.81);
end
```