Integration of Satellite Positioning and a Track Database for Safety-Critical Railway Control Systems

by

Yuheng Zheng

A thesis submitted to the University College London in accordance with the requirements of degree of Doctor of Philosophy

Department of Civil, Environmental and Geomatic Engineering
University College London

October 2008

ABSTRACT

Although Global Navigation Satellite Systems (GNSS) have been widely used in aviation, vehicle and marine navigation, and have also found non-safety railway applications (e.g. for locating trains in order to provide passengers with arrival and departure information), they still cannot be used in a standalone mode for safety critical railway applications such as automatic train control, automatic door opening or train integrity monitoring. This is because GNSS suffers from the line-of-sight problem, namely, GNSS might be unavailable when trains run through the areas with low satellite visibility (e.g. in urban canyons, deep cutting sides and tunnels).

A potential solution is to integrate satellite navigation measurements with other sensors such as a track database, INS or an augmentation system. This thesis is concerned with the evaluation of the potential role of a track database for this purpose. A rigorous mathematical model for the integration of GNSS with the track database is developed. The key feature of this model is its ability to model errors in both GNSS measurements and the track database to achieve realistic performance statistics for the combined system. Knowledge of the position of the railway lines turns positioning, in principle, into a one dimensional problem. This thesis uses both simulated London area information and real railway satellite availability information from the Birmingham area to assess the improvements in Required Navigation Performance (RNP) parameters that might be obtained if railway authorities invest in a track database. The stimulation shows that the integration system improves the accuracy and increases the redundancy so that the system only needs as few as two satellites to calculate the position and accuracy, three satellites to compute the Receiver

Autonomous Integrity Monitoring (RAIM) and four satellites to do the Fault Detection and Exclusion (FDE). The cost-efficient accuracy of track database and suitable RNPs are also discussed for safety-critical railway requirements.

Dedications and Acknowledgements

This work is dedicated

- to my parents who lent me full support to my master's and doctoral studies;
- to my wife Dr. Huahui Zhao, who supported me and encouraged me when I was struggling with difficulties I had faced throughout my studies.

I especially want to thank my supervisor, Professor Paul Cross, for persevering with me as my supervisor throughout the time it took me to complete this project and write the dissertation.

I also want to thank my second supervisor Professor Marek Ziebart and my colleagues Dr. Lawrence Lau and Dr. Zenghong Li for their valuable suggestions for my project.

I would like to express my thanks to Rail Safety and Standards Board who supported this project.

I would also want to express my appreciation to Nottingham Scientific Ltd for providing me with the real time railway data in the Birmingham area.

Table of Contents

List of Tables

List of Figures

Acronyms

ABAS	Aircraft Based Augmentation System
ABS	Automatic Block Signaling
ASQF	Application Specific Qualification Facility
ATAL	Along Track Alarm Limit
BNSS	Beidou navigation satellite system
C/A-code	Coarse/Acquisition-code
CDMA	Code Division Multiple Access
CNSS	Compass navigation satellite system
CS	Commercial Service
CSNPC	China Satellite Navigation Project Center
CTAL	Cross Track Alarm Limit
DGPS	Differential Global Positioning System
DOP	Dilution of Precision
DOT/DOD	Department of Transport/Department of Defense
ECEF	Earth-Centered Earth-Fixed
ECONFIN	European Union's Economic and Financial Affairs Council
EGNOS	European Geostationary Navigation Overlay Service
ERTMS	European Railway Transport Management System
ESA	European Space Agency
ETCS	European Train Control System
EU	European Union
EUROCONTROL	European Organisation for the Safety of Air Navigation
EWAN	EGNOS wide area communication network
FAA	Federal Aviation Administration
FDE	Fault Detection and Exclusion
FDMA	Frequency Division Multiple Access
GAGAN	GEO Augmented Navigation
GBAS	Ground Based Augmentation System

GDOP	Geometric Dilution of Precision
GEO	Geostationary Earth Orbit
GLONASS	GLObal'naya NAvigatsionnaya Sputnikovaya Sistema
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GRAS	Ground Based Regional Augmentation System
GSM-R	Global System for Mobile Communications - Railway
GST	Galileo System Time
GTRF	Galileo Terrestrial Reference System
HAL	Height Alarm Limits
HDOP	Horizontal Dilution of Precision
HPL	Horizontal Protection Level
ICAO	International Civil Aviation Organization
IGS	International GNSS Service
IMO	International Maritime Organisation
INS	Inertial Navigation System
IRNSS	Indian Regional Navigational Satellite System
ITRS	International Terrestrial Reference System
ITU	International Telecommunication Union
LAAS	Local Area Augmentation System
LOS	Line-of-Sight
LS	Least Squares
LSP	Least Squares Parabola
MCC	Mission Control Centres
M-code	Military Code
MDE	Marginally Detectable Error
MEO	Medium Earth Orbit
MEDLL	Multipath Estimation Delay Lock Loop
MSAS	Multifunctional Satellite Augmentation System
NDGPS	Nationwide Differential Global Positioning System
NLES	Navigation Land Earth Stations

NMEA	National Marine Electronics Association
OS	Open Service
OTF	On-The-Fly
PACF	Performance Assessment and System Checkout Facility
P-code	Precision-code
PDOP	Position Dilution of Precision
PRS	Public Regulated Service
PPP	Public-Private Partnership
PZ-90	Earth Parameter System 1990
QZSS	Quasi-Zenith Satellite System
RAIM	Receiver Autonomous Integrity Monitoring
RNAV	RNP area navigation
RNP	Required Navigation Performance
RDSS	Radio Determination Satellite Service
RTK	Real-time Kinematic
SA	Selective Availability
SAR	Search and Rescue service
SBAS	Satellite Based Augmentation System
SIS	Signal in Space
SIL	Safety Integrity Level
SL	Standard Level
SOL	Safety of Life
SPS	Standard Positioning Service
SV	Space Vehicle
TAI	International Atomic Time
TATO	Timetable and Train Order
TDOP	Time Dilution of Precision
UERE	User Equivalent Range Error
UTC-C	Chinese Coordinated Universal Time
UV	UltraViolet
VDOP	Vertical Dilution of Precision

VHF	Very High Frequency
WAAS	Wide Area Augmentation System
WADGPS	Wide-area DGPS
WAGE	Wide Area GPS Enhancement
WGS'84	World Geodetic System of 198
WMS	Wide Area Master Stations
WRS	Wide Area Reference Stations
3D	Three-Dimensional

Chapter 1 Introduction

1.1 Background

Traditional railway signaling and control systems have problems such as difficulty in enhancing the capability of railway lines, train delay caused by signal failures, large investments and maintenances, incompatibility and non-interoperability between different systems. To improve the railway operational system, especially the future high speed and high density train control system, positioning systems plus the communication system might be applied to replace the current signaling and control system. According to the high navigation performances of Global Navigation Satellite System (GNSS), GNSS is a great choice for the train positioning system. However, since GNSS also suffers from the line-of-sight problem, it cannot be used alone for the railway positioning system. To compensate such a deficiency of GNSS, other sensors and technology should be used. This thesis is mainly concerned with integrating GNSS with a track database for safety-critical railway applications (e.g. automatic train control, automatic door openings at stations, train protection and warning system) .

1.1.1 Traditional Railway Signaling and Control System

In United Kingdom, the current railway signaling and control system is a block signaling system, called as the Automatic Block Signaling (ABS) control system, which was introduced over 100 years ago (Thomas et al., 2007). The train movement authorities of this system are delivered by the radio signals, colour line-side light signals, and remote signals. In this system, the train position is detected by track circuits which divide the railway tracks into many insulated rail blocks. At each block, a simple track circuit requires at least a feed and a relay. The feed is used to provide the electrical current at one end of the block whilst the relay is at the other end of the block. When the block is not occupied, the relay completes an electrical current and is energised by the current. In this situation, the clear (or unoccupied) signal for this block should be displayed. When a train enters into the block, the axle of the train creates a short-circuit and the relay is de-energised which means the display signal for this block should be the stop or occupied (Palmer, 2006). Therefore, trains are determined and separated by the blocks.

The ABS system can provide a high level of safety for railway control and operation systems and therefore is still serving the industry well and continues in revenue service operation in many railway lines around the world (Rumsey, 2006). However, this traditional technology has three major problems which would restrict its service for future high speed and high density railway operation.

Firstly, the ABS system requires lots of track-side transponder and other equipment installations. It costs railway companies large investments and high operational and maintenance fees. Therefore, the price of railway tickets keeps high and thus decreases its attractiveness among the transport options.

Secondly, the ABS system divides rail tracks into many fixed blocks. At a time, only one train can be allowed to enter any fixed blocks. Once the train travels into a block, the whole block is reported to be occupied regardless of the length or the speed of the train and the exact position of the train in the block is unknown. Therefore, the length of the fixed block has to be kept sufficiently long to allow the highest speed and longest trains to run within the block. When other trains which are not running at the highest speed travel in the fixed block, they are kept further apart in the minimum safe stopping distance. In this sense, the capability of railway lines does not reach its maximum.

Finally, the ABS system also faces the safety risk and delay problems. In the ABS system, if any equipment or wire was out of order, or the power supply was cut off, or the signal was failed to be shown, the train delay would be caused. For the safety problem, the contaminants on the rail and the vandalism would cause the risk for the ABS system (Palmer, 2006).

Further, in view of the strong competition with other public transport, the railway industry has to be concerned with the customer requirements. Specifically, the railway passenger requirements are simple and clear (Kenna, 2006; Alcouffe, 2001), summarised as follows:

1. Safe journey
2. Affordable price
3. Quick journey
4. More frequency
5. Punctuality
6. Information (pre-transport, during transport and post-transport)

Therefore, the future railway control system should at least have the following features:

1. High safety integrity level
2. Low Cost
3. High speed train operation
4. High density railway traffics
5. Accurate train position and speed

The existing ABS system could not completely fulfill the requirements above. To meet these goals and objectives, the GNSS-based railway control system is an option for the future high speed and high density railway operation.

1.1.2 Applications of GNSS in Railway

The GNSS has advantages such as 24 hour real time positioning, all weather conditions working, free of charge for users, and high navigation performances. Further, the high performances of accuracy, integrity and availability provided by the GNSS and GNSS augmentation systems have made the GNSS be widely applied in many applications.

Vehicles can use the GNSS to determine their location, speed, direction, all of which can be displayed on the moving maps. Boats and ships can also install the GNSS receivers to enable navigation in lakes, seas and oceans. GNSS and their augmentation systems are used in aviation for en-route navigation, approach, landing and departure of all flights. It can also be used in surveying, mobile phones, location-based services and mapping. The high accuracy performance of the GNSS makes it useful for all non-safety related applications. Additionally, the integrity information can be provided by the GNSS satellite itself (future), RAIM or GNSS augmentation systems; therefore, it brings about good opportunities for using GNSS in safety critical applications at all transport sectors. The successful applications of the GNSS in the safety-critical aviation navigation and maritime applications have suggested its prospect for railway safety-critical applications (Prasad 2005; Kiss 2000; Kaplan 2006).

Compared with the ABS system, GNSS has the benefits such as lower initial costs (e.g. all necessary equipments can be stored on the locomotive), less maintenance (e.g. transponders needed to be replaced owing to the vandalism), and potentiality of increasing the capability of railway lines for both freight and passenger trains due to its high navigation performances. However, using GNSS technology for the future railway control system in safety critical railway applications also faces challenges. Safety-of-life (SOL) in safety critical railway applications requires high integrity. For the GNSS, the more satellites can be tracked in view, the higher accuracy and integrity can be achieved. Even for the minimum requirement, GNSS technology still needs four visible satellites to calculate the position and the accuracy, and five visible satellites to conduct the integrity check. Unfortunately, the GNSS suffers from line-of-sight (LOS) problems, especially in the low satellite visibility environments such as deep cuttings, forests, urban canyons and tunnels. This will cause the loss of position and integrity, which is unacceptable for the safety-critical railway applications. Figure 1.1 gives an example of the signal blocking. When trains travel into the central

Therefore, the objective of this thesis aims to investigate the following questions:

1. What Required Navigation Performance (RNP) standards are needed for the safety-critical railway applications?
2. How does the GNSS standalone perform in an open area of railway environments?
3. How does the GNSS integrated with a track database perform in an open area of railway environments?
4. How does the GNSS standalone perform in low satellite visible railway environments?
5. How does the GNSS integrated with a track database perform in low satellite visible railway environments?
6. How does the GNSS standalone perform in a real railway line?
7. How does the GNSS integrated with a track database perform in a real railway line?
8. Has the integration system improved the GNSS performances?
9. What accuracy of the track database is cost efficient?
10. Have the performances of GNSS standalone and the GNSS/Track Database integration system achieved the RNPs of safety-critical railway applications?

Through exploration of the questions above, this research aims to achieve the following results:

- Present performances of the GNSS standalone in railway environments
- Present performances of the GNSS/Track Database integration system in railway environments
- Define the RNPs for the safety-critical railway applications

- Assess the improvements of the RNP parameters that might be obtained if railway authorities invest in a track database
- Assess the cost efficient accuracy of track database

This research aims to examine whether the track database is a cost-efficient compensation for satellite-based railway control systems rather than proving the necessity of integrated GNSS with track database for safety critical railway applications.

1.3 Thesis Outline

This thesis consists of eight chapters organised as follows:

Chapter 1 brief introduces the current and future railway control systems. The limitations of the current railway control system and the deficiency of the GNSS are summarised. This chapter concludes with the introduction of the GNSS/Track Database integration system and the objectives of this research.

Chapter 2 reviews the current and future GNSS in detail. The principles of the GNSS and their augmentation systems are described. The current applications of GNSS and the future possibility for railway applications are also reviewed.

Chapter 3 discusses the traditional railway control systems in terms of their advantages and disadvantages. This follows the rationale of choosing the satellite-based railway control system. The deficiencies of different complemented sensors for the GNSS and the advantages of integrating GNSS with a track database in railway applications are also presented.

Chapter 4 reports the four parameters of Required Navigation Performances (RNP), including accuracy, integrity, continuity and availability. The standards of RNPs for safety critical railway applications are also discussed.

Chapter 5 presents the application of GNSS/Track Database integrated system for the railway control system. The track database collection, the errors in track database, and the trajectory matching are described. It follows a detailed description of the mathematical model of the integration system for the linear and nonlinear cases of track lines developed in this research. Chapter 5 concludes with the review of the method of estimating the accuracy and the integrity of the integration system.

Chapter 6 compares the performances of the standalone GNSS and GNSS/Track Database integration system by using the developed mathematical model to analyze simulated data around London area in four scenarios, namely: the open areas of London regions, randomly reducing one satellite in view, randomly reducing two satellites in view, and using different track database accuracies in open areas.

Chapter 7 compares the performances of the GNSS alone and the integration system by using the mathematical model to analyze the real railway track and satellite availability information collected from Birmingham areas, in terms of the satellite visibility, estimated accuracy performance, estimated RAIM performance, and estimated availability of accuracy and integrity.

Chapter 8 summaries the research output of this thesis and provides implications for future research on the GNSS/track database integration system.

1.4 Contribution

The major contribution of this thesis is proving that the track database can not only be used for matching the GNSS position to the database or map, but also can be a cost-efficient compensation for the GNSS-based railway control system. With considering the relationship of train position and track data points and uncertainties in their measurements, the 3-D position problem reduces to the one dimension problem which means the requirement of satellite in view is reduced. Additionally, this thesis also discussed about the reference RNPs for the safety-critical railway applications, and suitable level of track database accuracy. The results from the simulated data suggest that the integrated system provides significant improvements in terms of accuracy, integrity and availability in all three directions such as along track, cross track and height. The performances of standalone GNSS and integration system give the idea for the railway authorities if they want to invest GNSS-based railway control system in the future.

1.5 Publication

Zheng, Y. (2007). *Improving Positioning Accuracy and Integrity in Rail Safety-Critical Applications through the Integration of GNSS with a Track Data Base.* Paper presented at the 20th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2007, Fort Worth Convention Center, Fort Worth, Texas. (Best Student Paper Award)

Zheng, Y. (2007). *The impact on Integrity of integrating a track database with GNSS for the High Safety Rail Applications.* Presentation at RIN New Navigators Seminar 2007, 20th June, 2007, Imperial College London.

Zheng, Y. (2006). *Using a Track Data Base to Improve the Availability of GNSS for Train.* Presentation at RIN New Navigators Seminar 2006, 21st June, 2006, Imperial College London.

Chapter 2 Fundamentals of GNSS

This chapter reviews the fundamental background of GNSS. The first section (2.1) reviews the current fully operation GNSS and the future operation GNSS. In section 2.2, the principles of GNSS operation and the current augmentation systems for GNSS are introduced. The chapter concludes with the review of the current applications of GNSS and the future possibility for railway applications (see 2.3). The review of the GNSS principles in this chapter is fundamental rather than comprehensive. The more comprehensive details are well described and discussed in the reference books (El-Rabbany, 2002; Hofmann-Wellenhof et al., 1997; Kaplan & Hegarty, 2006; Leick, 2004; Misra & Enge, 2001; Parkinson et al., 1996; Prasad & Ruggieri, 2005).

2.1 Global Navigation Satellite System

Global Navigation Satellite System (GNSS) is a general concept for the satellite navigation/positioning system which can provide the position information to receivers with global coverage. In view of the high speed development of the GNSS, it is now defined as a worldwide set of satellite navigation system (Kaplan & Hegarty, 2006). A GNSS receiver can get its three-dimensional positioning solutions (e.g. longitude, latitude, and altitude) with a few meter accuracy by using satellite signals transmitted from satellites to the receiver (Groves, 2008). Receivers can also calculate the precise time and velocity.

GNSS is becoming highly developed around the world. The United States NAVSTAR/GPS (Details described in Section 2.1.1) is the only one GNSS that is fully operational up till now. The Russian GLONASS is a GNSS which is expected to return to full operation by 2011 (Revnivykh, 2006). More GNSSs are under development. Galileo is the second generation GNSS which is developed by European Union (EU) and European Space Agency (ESA), expected to be in full operation around 2012 (Hein et al., 2007). China is also developing their new Compass (also called as Beidou-2) satellite system which will be another GNSS system and also expected to be in full operation by 2012 (Kaplan & Hegarty, 2006). Therefore, in the next five to ten years, more choices for GNSSs will emerge and there will be more than 120 satellites in orbit used for the satellite navigation.

The GNSS also includes other regional satellite navigation systems and augmentation systems. Chinese Beidou I, Japan's proposed Quasi-Zenith Satellite System and Indian Regional Navigational Satellite System are all the regional navigation systems which provide the navigation performances in the coverage areas. Additionally, in order to improve the performance of the GNSS system for different applications, many augmentation systems (e.g. LAAS, WAAS, EGNOS, DGPS, RTK) have been developed. Therefore, the GNSS system can be regarded as not only the global coverage satellite system but also the combination of all satellite navigation and related augmentation systems.

2.1.1 Global Positioning System (GPS) Overview

GPS is the first GNSS and was developed by the Department of Defense (DOD), United States. The full name of GPS is NAVigation Satellite Timing And Ranging/Global Positioning System (NAVSTAR/GPS). GPS was originally designed as a navigation system for U.S. military users. But it is also available for civilians, and therefore is a dual-use system for both military and civilian users. GPS is a one-way-ranging (passive) system which provides the worldwide, 24-hours real time, continuous, accurate, three-dimensional position, velocity and timing information to the users with appropriate receivers. GPS consists of three segments, namely, the Space Segment, the Control Segment, and the User Segment. The Space Segment deals with the launching of satellites. The Control Segment monitors and manages the satellites operation. The User Segment relates to the both military and civil receiver equipments development.

The baseline GPS constellation contains 24 satellites in six Earth-centered orbital planes with a radius of 26,560 km (i.e. about 20,163 km above the Earth). Each plane hosts four satellites. The orbits are nearly circular and at a 60^{o} spaced separation around the equator with a 55^{o} inclination relative to the equatorial plane. The GPS satellites have an orbit period of one-half of a sidereal day or 11 hours, 58 minutes. The first GPS satellite was launched on 22nd February 1978. It is one of the first generation of GPS satellites, indicated as Block I. Block I satellites were composed of 12 satellites: 11 were successfully launched and one was failed to be launched from 1978 to 1985. The purpose of the Block I was to build up the ground track network and test the GPS receiver performance and the possibility of the GPS operation.

The second generation of GPS satellites is known as Block II/IIA satellites which were developed for the first operational constellation. Block IIA ("A" denotes advance) is an advanced version of Block II with an enhancement of the navigation message storage. The first Block II satellite was successfully launched on 14th February 1989 and the first Block IIA satellite was launched on 26th November, 1990. From 1989 to 1997, a total of 28 Block II/IIA satellites were launched. The full operation constellation of GPS was declared in April 1995 (Hofmann-Wellenhof et al., 1997; Kaplan & Hegarty, 2006; Leick, 2004; Misra & Enge, 2001; Parkinson et al., 1996;

Prasad & Ruggieri, 2005). At the time of writing this thesis, some of these satellites were still in service.

Table 2.1 Summary of the GPS Satellites Launches

GPS Satellites Block	Years	Successful Launches
Block I	1978-1985	10
Block II	1989-1990	9
Block IIA	1990-1997	19
Block IIR	1997-2005	12
Block IIR-M	2005-2008	6

Because the life of the Block II/IIA was designed to be 7.5 years, the new generation of GPS Block IIR satellites were developed. The "R" denotes replenishment or replacement. The total plan of Block IIR is 21 satellites in orbit. The first Block IIR satellite was successfully launched on 23rd July 1997. Over the next seven years, 11 Block IIR satellites were launched. From 2005, a GPS modernization plan was introduced in order to improve the quality and protection for military and civil use. The rest of Block IIR satellites were converted to the Block IIR-M (Replenishment and Modernization) satellites. This new generation constellations employed a new military M-code (to enhance accuracy) and a second civil signal L2C. In addition, Block IIR-M will offer a modernized antenna panel that provides increased signal power to receivers on the ground, improving encryption and anti-jamming capabilities for the military (Parkinson et al., 1996). The first Block IIR-M1 or Block IIR-14 (M) was launched by a Delta II rocket on 25 September 2005. At time of this writing, six Block IIR-M satellites have been launched and the most recent one is Block IIR-19(M) which was launched on 15th March 2008 ((GPS World, 2008). Table 2.1 above gives a summary of the history of GPS satellites launch. The total number of GPS operation constellations has achieved 32 satellites up till now (U.S. Naval Observatory 2008), as shown in Table 2.2. Three notations are used to refer to the satellites. One is to assign a letter and a number to each satellite. The letter (i.e. A, B, C, D, E and F) represents the orbital plane of the satellite and the number (from 1 to 6) means the number of satellites on the plane. The second notation is Pseudorandom Noise (i.e. PRN) which means the PRN code generators on the satellite. The last notation is the space vehicle number (SVN).

The fourth generation of GPS satellites is Block IIF ("F" stands for Follow-on) and the future generation of GPS is called GPS III. Both of them will broadcast a new civil signal L5 to provide the safety-of-life service. The plan is to build up 12 satellites for the Block IIF and 32 satellites for GPS III (8 GPS IIIA, 8 GPS IIIB and 16 GPS IIIC). Unfortunately, the launch of the first GPS Block IIF satellites has been postponed to 2008 and GPS III plans to begin launches in 2013 (Kaplan & Hegarty, 2006). These programs are expected to improve position, navigation, and timing services for worldwide military and civil users and to provide advanced anti-jam capabilities yielding improved system security, accuracy and reliability.

The GPS satellite transmits navigation signal on two carrier frequencies (or sine waves) called L1 (1575.42 MHz), the primary frequency, and L2 (1227.60 MHz), the secondary frequency. These signals are generated synchronously, and if both signals are received by a user, the ionospheric delay can be calibrated. But for most civilian users like trains or cars, they only use one frequency (i.e. L1). The carrier frequencies are modulated by spread spectrum codes with a unique PRN sequences (or PRN codes) associated with each Space Vehicle (SV) and a navigation message data. All SVs use a CDMA (Code Division Multiple Access) technique.

Table 2.2 Current GPS configuration (On 27th May 2008)

Launch Order	PRN	SVN	Launch Date	Orbital Plane
IIA-10	32	23	26 November 1990	E5
IIA-11	24	24	04 July 1991	D5
IIA-12	25	25	23 February 1992	A5
IIA-14	26	26	07 July 1992	F5
IIA-15	27	27	09 September 1992	A4
IIA-21	09	39	26 June 1993	A1
IIA-22	05	35	30 August 1993	B5
IIA-23	04	34	26 October 1993	D4
IIA-24	06	36	10 March 1994	C1
IIA-25	03	33	28 March 1996	C2
IIA-26	10	40	16 July 1996	E3
IIA-27	30	30	12 September 1996	B2
IIA-28	08	38	06 November 1997	A3
IIR-2	13	43	23 July 1997	F3
IIR-3	11	46	07 October 1999	D2
IIR-4	20	51	11 May 2000	E1
IIR-5	28	44	16 July 2000	B3
IIR-6	14	41	10 November 2000	F1
IIR-7	18	54	30 January 2001	E4
IIR-8	16	56	29 January 2003	B1
IIR-9	21	45	31 March 2003	D3
IIR-10	22	47	21 December 2003	E2
IIR-11	19	59	20 March 2004	C3
IIR-12	23	60	23 June 2004	F4
IIR-13	02	61	06 November 2004	D1
IIR-14M	17	53	26 September 2005	C4
IIR-15M	31	52	25 September 2006	A2
IIR-16M	12	58	17 November 2006	B4
IIR-17M	15	55	17 October 2007	F2
IIR-18M	29	57	20 December 2007	C6
IIR-19M	48	07	15 March 2008	A6

There are two GPS codes, called as C/A-code (Coarse/Acquisition-code) and P-code (Precision-code). Both codes are transmitted on L1 frequency, but only P-code is modulated onto L2 frequency. The chipping rate for C/A-code broadcast is 1.023 MHz and for P-code, it is 10.23 MHz (ten times of C/A-code). Because P-code has a higher modulation bandwidth, the code ranging signal is more precise. The C/A code is available for civilian users. Since 1994, the P-code has been encrypted by the Y-code, normally indicated as P(Y) code. This feature is known as antispoofing (AS). The encrypted P(Y) code is only available for the U.S. military and other authorized users.

As for the GPS modernization program development (Block IIR-M, Block IIF and GPS III), the new civil code (C/A-code) is modulated on the L2 frequency (known as L2C). The new military code (M-code) is broadcast on both L1 and L2 frequencies. The availability of two C/A codes allows the stand-alone GPS receiver to calibrate the ionospheric delay. After the launch of GPS Block IIF satellites, a new third frequency (L5, 1176.45 MHz) would be operated for the safety-of-life service. Figure 2.1 illustrates the GNSS signal structure (Hein et al., 2007).

The GPS navigation data is also added in the navigation signal to transmit the navigation message. It is binary data stream which is transmitted at a low rate of 50 kbps. The navigation message contains keplerian elements which define the actual location of satellites, precise satellite clock parameters, the satellite health status, and the satellite almanac and ionoshperic data. Therefore, three components are in the satellite navigation signal: Carrier frequencies (L1, L2, L5 in the future), PRN codes (C/A, P(Y), M code) and Navigation data (Kaplan & Hegarty, 2006; Misra & Enge, 2001; Parkinson et al., 1996).

to provide position, velocity and timing determination. It was initially developed by the former Soviet Union and currently is operated by the Russian government. The development of GLONASS stared in 1976 and the first GLONASS satellite was launched on 12th October 1982. Originally, GLONASS was designed for the Soviet Union military users, but now it also defines as a dual-use system for both civil and military users. The nominal constellation is composed of 24 satellites (21 active satellites + 3 active spares) in three orbital planes separated by 120 degrees. The satellites operate in circular 19100-km orbits at an inclination of 64.8 degree to the Earth's surface. The GLONASS satellites have an orbit period of about 11 hours and 15 minutes. The spacing of satellites provides continuous and global coverage navigation service to the users near the Earth surface (Kaplan & Hegarty, 2006; Kazantsev, 1994; Leick, 2004).

The full operation constellation of GLONASS was declared in February 1996. But due to the financial problems of the Russian government, the system fell down rapidly in the following five years without sufficient constellation maintenance. By the end of 2001, it operated with only eight satellites. On 20th August 2001, the Russian government decided to rebuild the GLONASS. According to the declaration, the GLONASS aimed to restore its full operational constellation (FOC) in 2011. However, Russian space agency and India's space agency made the agreement to develop and maintain GLONASS together in 2004 and new funding was introduced into the GLONASS by the Russian Federation. Therefore, the GLONASS program appears to have been speeded up. Two government agencies will cooperate to restore the system to constant coverage of Russian and Indian area with 18 satellites in 2008, and be fully operational with all 24 satellites by 2011 (Groves, 2008; Hein et al., 2007; Kaplan & Hegarty, 2006; Revnivykh, 2006; Revnivykh et al., 2005).

Unlike other GNSSs (e.g. GPS or Galileo), GLONASS uses frequency division multiple access (FDMA) technology which GPS and Galileo use CDMA technology to recognize signals received by receivers. As the previous section introduced, each GPS satellite transmits different C/A and P(Y) codes on the same frequency (L1 and L2), but each GLONASS satellite transmits the same PRN code (C/A and P code) on different frequencies around L1 and L2. GLONASS satellite navigation signal comprises of three components: two L-band carriers (1602-1615.5 MHz for L1 and 1246-1256.5 MHz for L2), C/A and P code, and a navigation message. The carrier frequency is derived from the following equation:

$$f_L^k = (178.0 + k \times 0.0625) \cdot L \qquad (MHz) \tag{2.1}$$

where k was an integer which takes the value from 1 to 24, i.e. each satellite was assigned a number for the GLONASS channel. But after 2005, the Russians have modified two carrier frequencies to 1598.0625-1604.25 MHz for L1 and 1242.9375-1247.75 MHz for L2 (i.e. k=-7 to 4). Therefore, the k number only uses 12 values for all satellites and the satellites on the opposite side of the Earth need to share the same k number. The L is the factor number for the two L-band carrier frequencies,

$$L = \begin{cases} 9 & for \quad L1 \quad band \\ 7 & for \quad L2 \quad band \end{cases}. \tag{2.2}$$

Similar to GPS, there are also C/A code and P code for GLONASS, where C/A code for civil users is on L1 and P code for military users is on both L1 and L2. The chipping rate for C/A code and P code are 0.511 Mbps and 5.11 Mbps, respectively. The code length is 511 chips for C/A code and 33554432 chips for P code. The navigation message is a 50 bps data stream and it modulates on both C/A code and P code to generate two types of navigation messages (i.e. C/A navigation message and P code navigation message). The navigation message is to provide the major information of the satellite location, channel number and satellite health status. (Gibbons, 2007; Kaplan & Hegarty, 2006; Leick, 2004; Parkinson et al., 1996).

The time reference systems are also a different between GPS and GLONASS. The GPS time system is linked to Coordinated Universal Time, U.S. Naval Observatory (UTC (USNO)) whereas the GLONASS time system is linked to Coordinated Universal Time, Soviet Union (UTC (SU)). Another difference between these two systems is that the GPS and GLONASS use different coordinate frames to express the position of their satellites. GPS uses the World Geodetic System of 1984 (WGS'84) and GLONASS uses the Earth Parameter System 1990 (PZ-90). The maximum difference between these two systems could be 20m on the Earth surface. These two differences are the problems which need to be considered for the integration the GPS and the GLONASS. Table 2.3 compares of the GPS and the GLONASS.

At the time of this thesis writing, two generation GLONASS constellations have been launched. From 1982 to 2005, over 60 first generation GLONASS satellites were launched. They were defined as two blocks (Block I and Block II). The main difference between these two blocks is the lifetime of satellites. The lifetime for the GLONASS Block I satellites are around 14 months and for GLONASS Block IIc spacecrafts have been increased to three years. Like the GPS modification project, the Russian government also continues to improve their GNSS. The second generation of GLONASS constellation, called as GLONASS-M (where "M" stands for Modified), was started to be developed in 1990 and the first GLONASS-M spacecraft was launched on 10th December 2003. This new generation constellations possess a longer lifetime, improved navigation signals, improved navigation message and improved navigation performance. The latest designed generation, known as GLONASS-K, is the third generation. The first spacecraft of GLONASS-K is expected to be launched in 2009. The GLONASS-K satellites are designed with a longer lifetime of 10 to 12 years and a reduced weight about 800 kg than the GLONASS-M spacecrafts. A third L-band civil signal for safety-of-life applications, with the band of 1190-1212 MHz, will also be put in this series constellation.

On 25th September 2008, Russia successfully launched three GLONASS-M satellites which made the total operation constellation to 16 satellites. GLONASS positioning accuracy is claimed to be about 55 meters in the horizontal plane and about 70 meters in the vertical plane for civil users, and approximate 20 meters in the horizontal and 34 meters in the vertical for military users. It is believed that GLONASS could provide better accuracy for both civil and military users, especially after the time of next FOC.

The development of GLONASS not only depends on itself but also focuses on the interoperation with the GPS. In the market, some companies (e.g. Leica, Magellan, Septentrio) have already produced the receivers which could take both GPS and GLONASS signals. The integration GPS/GLONASS would be especially useful for civil applications because it would provide more satellites in the view and thus have better accuracy, continuity and availability. This would be extremely helpful when GNSS for railway applications is used because trains would travel through the tough environments wherein the performance of GNSS is suffered by the low satellite visibility (more details will be described in Chapters 4 and 5).

Table 2.3 Comparison of GPS and GLONASS

	GPS	GLONASS
Number of satellites	24	24
Orbital planes	6	3
Satellite per orbital plane	4	8
Inclination of orbital (deg)	55	64.8
Altitude (km)	20163	19100
Orbital period	11h58m	11h15m
Repeat ground path	1 sidereal days	1 sidereal day
Signal separation technique	CDMA	FDMA
Satellite coordinate frame	WGS'84	PZ-90
Time reference	UTC(USNO)	UTC(SU)
Carrier frequency	L1: 1575.42 MHz	L1: 1602.5625-1615.5 MHz
	L2: 1227.60 MHz	L2: 1246.4375-1256.5 MHz
C/A Code rate	1.023 MHz	0.511 MHz
P code rate	10.23 MHz	5.11 MHz

2.1.3 Galileo Overview

Galileo is the European own GNSS, initially built by the European Union and European Space Agency. It is an independent system from U.S. GPS and Russia GLONASS. Different from the GPS and the GLONASS, the Galileo system is specifically designed for civilian use by providing high accuracy and global coverage positioning services. The idea of Galileo began in the early 1990s, and the different concepts for Galileo were unified to one by the agreement of four EU countries (United Kingdom, German, Italy and France) at the end of 1999. Until the late of 2000, feasibility and definition phase of the Galileo system were finally completed. The EU and ESA confirmed that they would fund the Galileo program in March 2002. The development and validation phase started from 2001 and it is still ongoing. The phase consists of the consolidation of the space segment, ground-based infrastructure, and the validation of the system. Four prototype satellites will be launched to test the system in orbit in this phase. The constellation deployment phase was scheduled to be started from 2006 and completed in 2008, but it is now delayed to 2013 (Kaplan & Hegarty, 2006; Prasad & Ruggieri, 2005).

At the beginning stage of the Galileo program, it was only developed by the EU and ESA. However, as the time of Galileo project going on, many countries outside the EU showed their interest in this program. In September 2003, China claimed that they would invest about 2 billion CNY (155 million GBP) on the Galileo development. One year later, Israel also joined the Galileo project with the agreement of the EU. In the following two years, Ukraine, India, Morocco, Saudi Arabia and South Korea also joined this project and the Galileo project currently seems to be more like a multi-country project.

The progress of the Galileo project is very dramatic. Because the Galileo system is expected to have better accuracy than the GPS and to be available for all civil and military users, it might eliminate the influence of U.S. GPS. The potential to apply SA to the GPS could also be a challenge to the development of the GPS market. Therefore, EU has been under a big pressure by the U.S. government since the Galileo project started. Especially, after the terrorist attacks on 11th September 2001, the Galileo project was almost dead due to the pressure from the U.S. government. However, after a few months, the EU insisted on their development of the Galileo project, and the U.S. finally signed the agreement with the EU on the cooperation of GPS and Galileo in 2004. The political peril was not the only reason for the delay of the Galileo programme. The finance appeared to be another peril for the Galileo project. The original plan, called Public-Private Partnership (PPP) for which private investment would provide two-thirds of the investment needed to launch Galileo's infrastructure, seems to be impossible. In September 2007, the EU finance ministers submitted a proposal about all-public funding of the Galileo program which was approved by the European Union's Economic and Financial Affairs Council (ECONFIN) on 23rd November 2007. ECOFIN ministers also confirmed that the Galileo project should achieve FOC no later than 2013 (five years delay of 2008) (Ruiz, 2004). Till the time of this thesis writing, only two Galileo satellites were launched, one (designed as GIOVE-A (Galileo In-Orbit Validation Element)) was launched on 28th December 2005, and the other one launch of Galileo constellation (GIOVE-B) was on 27th April 2008.

The Galileo system will consist of 30 satellites, divided within three operational orbital planes at an altitude of 23616 km above the Earth's surface and with an inclination of 56 degrees. The orbit period for the Galileo satellite is about 14 hours and 22 minutes. Each orbital plane will be equally spread by nine operational satellites and one active spare satellite which is called as a Walker 27/3/1 constellation type. The geodetic coordinate reference frame for Galileo constellations is called as the Galileo Terrestrial Reference System (GTRF) which is also an independent realization of the International Terrestrial Reference System (ITRS). The GTRF only has a few centimetre difference from the GPS WGS-84 (which is also a realization of the ITRS). Therefore, the GTRF and WGS-84 are compatible for most users with this accuracy level. The time reference frame of Galileo system is also different from the GPS time system. Galileo will use the Galileo System Time (GST) based on the international atomic time (TAI) whereas GPS uses the UTC (USNO). Once Galileo is operational, the difference between this two time systems can be broadcast by Galileo satellites (Kaplan & Hegarty, 2006).

The Galileo system will provide five major services which are defined as: Open Service (OS), Safety of Life (SOL), Commercial Service (CS), Public Regulated Service (PRS), and the support to Search and Rescue service (SAR). These services will be provided worldwide and independently from other satellite navigation systems by using the signals broadcast by the Galileo satellites. The performance of the Galileo services is shown in Table 2.4. The OS is designed for mass-market users, to provide Position, Velocity and Time (PVT) information that can be accessed free of direct charge, suitable for such as in-car navigation and location system in mobile phones. The OS will be available for all the users which have receivers compatible with the Galileo signals. However, if the receivers are integrated with other GNSSs in the future, the navigation performances will be improved in severe environments, such as the urban canyon areas and the forests (Dutton et al., 2002; Galileo Brochure, March 2003; Hein et al., 2002; Onidi, 2002).

The SOL service is mainly used for users concerned with human safety. The transportations, which have stringent safety requirements, like the railway, aviation and maritime, are main users for this service. The SOL service will provide the same level of accuracy in position and timing as the OS service. However, the SOL service will offer the integrity information to assure users to received signals that are truly broadcast by the actual Galileo. The SOL service signals are carried on two frequency bands as the E5a+E5b and L1 bands.

The CS service aims at professional applications which require high accuracy positioning performances. It will provide enhanced navigation performances and added value data than that is offered by the OS. The foreseen applications will be based on: dissemination of data with a 500-bps rate, for value-added services; and broadcasting of two signals, separated in frequency from the OS signals to ease advanced applications. The CS service signals will be the OS signals and another two encrypted signals which are on the E6 frequency band.

The PRS service is a restricted service which is only available for the government-authorized users who require a high level of protection against the Galileo Signal in Space (SIS) jamming or interference. The PRS signals are encrypted and broadcast on separate frequencies. The access to the PRS will be controlled through a secure key management system by the member state governments. The main application for this service will be: the European police office, civil protection services, safety services and emergency services in EU; or law enforcement, customs, and intelligence services in member states.

In addition, the Galileo satellites will also support the humanitarian search and rescue service to fulfil the requirements and regulations of the International Maritime Organisation (IMO), and be backward compatible with the COSPAS-SARSAT (Cosmicheskaya Sistemma Poiska Avariynich Sudov-Search and Rescue Satellite-Aided Tracking) system.

The Galileo system will use similar CDMA technology to GPS. The differences are the ranging code types, data types and the carrier frequencies. The Galileo constellation satellite will provide ten navigation signals and one SAR signal in following frequency ranges E5 band (1164-1215 MHz), E6 band (1260-1300 MHz), E2-L1-E1 band (1559-1592 MHz), and L6 (1544-1545 MHz only for SAR signal). In the case of E5 carrier, the lower E5a and the upper E5b signals are modulated onto the single E5 band; therefore, the composite of the E5a and the E5b can be denoted as the E5 signal and be broadcast on a single extra-wide channel. Among these ten signals, six are used for the OS and the SOL service; two are specifically devoted to the CS; and two are specifically designed for the PRS. The ranging code is a sequence of +1 and -1 generated by highly stable, autonomous atomic clocks aboard Galileo satellites. In the Galileo navigation signal, five types of data can be delivered: navigation data, integrity data, commercial data, PRS data and SAR data. The navigation data are spread through the E5a, E5b and E2-L1-E1 carriers. The commercial data are transmitted on the E5b, E6 and E2-L1-E1 carriers. The integrity (described in Chapter 4) data are transmitted on the E5b and E2-L1-E1 carriers. PRS data are only transmitted on the E6 and L1 carriers. The carrier L6 is only used to transmit the SAR data. Table 2.5 gives the service and data plan of the Galileo program (Dinwiddy et al., 2004; Hein et al., 2002; Hein et al., 2004; Mattos, 2004; Pratt, 2005; Rodriguez et al., 2004). The Galileo signal structure is shown in Figure 2.1.

Table 2.4 Performance for the Galileo services

Galileo Services	Positioning accuracy	Availability	Integrity	Timing accuracy
OS	H: 15m V: 35m (single frequency) H: 4m V: 8m (dual frequency)	99.8%	None	30 ns
SOL	H: 4m V: 8m (dual frequency)	99.8%	Yes (Integrity risk: $3.5\times10^{-7}/150$ sec)	30 ns
CS	< 1m (dual frequency)	99.8%	None	30 ns
PRS	H: 15m V: 35m (single frequency) H: 6.5m V: 12m (dual frequency)	99-99.9%	Yes (Integrity risk: $3.5\times10^{-7}/150$ sec)	30 ns

Table 2.5 Galileo services and data plan

Carrier Frequencies	Service	Data type
E5a	OS/SOL	Navigation data
E5b	OS/SOL/CS	Navigation data/Commercial data/Integrity data
E6	PRS/CS	Commercial data/PRS data
E2-L1-E1	OS/SOL/CS	Navigation data/Commercial data/Integrity data/PRS data
L6	SAR	SAR data

2.1.4 Chinese Compass (Beidou) System

Compass navigation satellite system (CNSS) is the Chinese own developed GNSS. Beidou is the Chinese name for this system; therefore, it also can be called as Beidou navigation satellite system (BNSS). The CNSS is a multistage program operated by the China Satellite Navigation Project Center (CSNPC). The first stage, called as Beidou-1 navigation system, is the test navigation system. Three prototype Beidou-1 satellites were launched between October 2000 and May 2003. All of these three satellites are the Geostationary Earth Orbit (GEO) satellites which are located at high altitudes of 36000 km (about 22000 miles) above the Earth at 80^{o} E, 140^{o} E and 110.5^{o} E longitude, respectively. The Beidou-1 was fully operational at the beginning of 2004 and provided the services to customers over China and surrounding areas. Therefore, the Beidou-1 system is actually a regional satellite navigation system. Different from GPS, GLONASS and Galileo, which are the passive-systems employed one-way TOA measurements, the Beidou-1 system provides a radio determination satellite service (RDSS) which requires two-way range measurements to avoid synchronizing the receiver clock. With the estimated user altitude, the RDSS requires only two satellites to locate the two-dimensional user position at the operation centre. The Beidou-1 satellites transmit the signals at a frequency of 2491.75 ± 4.08 MHz (S-band). The geodetic coordinate reference frame for Beidou-1 system is the 1954 Beijing coordinate system which is based on the Krassovsky ellipsoid (Izotov. A., 1959) with $a = 6378245$ m, $\alpha = 1/298.3$, $e^2 = 0.00669342$. The Beidou-1 system also uses its own time reference system which is Chinese Coordinated Universal Time (UTC-C) observed by the atomic clocks in the Beijing control centre. The Beidou-1 system can provide positioning service with an accuracy of about 20 to 100 meters and a timing service with about 20 ns (Bian et al., 2005).

In October 2006, the China National Space Administration decided to upgrade and fully implement the Beidou-1 system to the next stage, which is designed as Beidou-2 system. Therefore, the Beidou-2 system is the fourth GNSS in the world. The current design for the Beidou-2 system consists of a constellation of 30 medium earth orbit (MEO) satellites and five GEO satellites with positions at 58.75° E, 80° E, 110.5° E, 140° E and 160° E. The 30 MEO constellations will be equally split to six orbital planes at an altitude of about 21500 km above the Earth's surface and with an inclination of 55 degrees. According to the Chinese government press, People's Daily Online, the Beidou-2 system will provide two services: an open civilian service, which is designed to provide positioning accuracy within 10 meters, 0.2 m/s for velocity accuracy, and timing accuracy within 50 ns and a safer and high precision service for government/military authorized users. According to International Telecommunication Union (ITU) filings by China in May 2004, each satellite of the Beidou-2 will broadcast signals in four carrier frequencies: 1561 MHz (E2), 1590 MHz (E1), 1268 MHz (E6), and 1207 MHz (E5b). The Beidou-2 navigation signals are also CDMA signals using binary or quadrature phase shift keying (BPSK, QPSK, respectively). The fourth GEO satellite was launched on 3rd February 2007 and the first MEO satellite (Compass-M1) was successfully launched on 14th April 2007. The CNSS is expected to achieve FOC in 2012. Because the CNSS is a government project for the military use, not much information about this system has been released in public. Therefore, the CNSS appears to be very mysterious for the world, but surely it will be unveiled in the near future (Forden, 2004; Gao et al., 2007; Grelier et al., 2007; ION Newsletter, 2006; Kaplan & Hegarty, 2006; Liu, 2006; People's Daily, 2007; Wilde et al., 2007).

2.2 GNSS Operation

This section discusses the operation of the GNSS. The position estimation by using standalone GNSS is introduced in 2.2.1. It is followed by the technology called as the Differential GPS (DGPS) (2.2.2). The real-time kinematic (RTK) is discussed in 2.2.3. The section ends with the augmentation system of the GNSS (2.2.4).

2.2.1 Standalone GNSS Operation Using PRN Codes

Positioning with GNSS is based on the Time-Of-Arrival (TOA) concept which calculates the range between the users and satellites. The range is derived from measured time difference by comparing the received PRN codes in the satellite signal and the receiver generating PRN code (Hofmann-Wellenhof et al., 1997). Figure 2.2 shows the satellite-to-user range. In Figure 2.2, the r_i represents the geometric range between the i^{th} satellite and the user. If the satellite system time, which the signal was left, is designed as t_s, and the receiver system time, which the signal was arrived, is designed as t_u; then

$$r = c \cdot (t_u - t_s) = c \cdot \Delta t, \tag{2.3}$$

where c stands for the speed of light. However, the receiver clock and the satellite atomic clock will generally have a bias error from the system time. The transmitting signal in the space will also be delayed by various other error components, such as ionospheric error, tropospheric error and multipath error. Therefore, the observed range, differed from the geometric range, is called as pseudoarange measurement, which is denoted as ρ and can be expressed as

$$\rho_i = r_i + c\cdot(t_u - t_s) + d\rho_i + I_{\rho_i} + T_{\rho_i} + \varepsilon_{\rho_i} \tag{2.4}$$

where

t_u is the receiver clock offset with respect to the system time

t_s is the satellite clock offset with respect to the system time

$d\rho_i$ is the ephemeris error

I_{ρ_i} is the ionospheric delay

T_{ρ_i} is the troposheric delay

ε_{ρ_i} is the errors including multipath, hardware bias, and receiver noise.

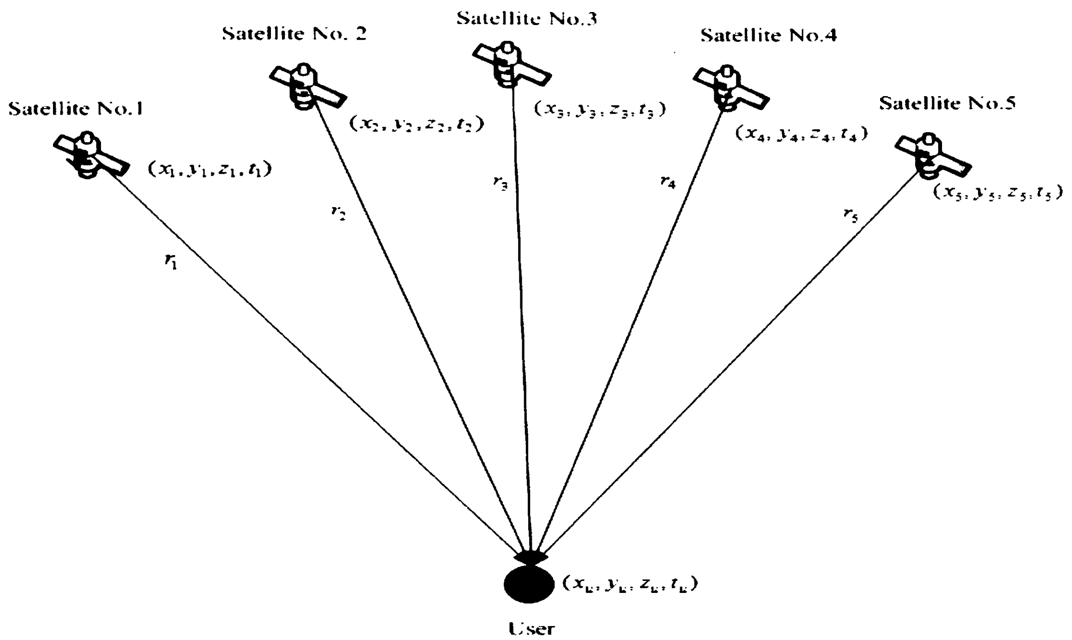


Figure 2.2 The satellite-to-user ranges

Satellite Clock Error

Although the satellite clock is made by a highly stable atomic clock which controls all timing operations including broadcast signal generation, the satellite clock error may deviate from the GNSS time due to the bias and drift. This offset can be up to approximately 1 ms, which could cause a 300 km error in the pseudorange (Kaplan & Hegarty, 2006). Fortunately, this error can be modified by the GNSS ground monitoring station and can be transmitted to the satellite to rebroadcast to users in the navigation message. Therefore, in this thesis, the effect of satellite clock error is assumed to be compensated and should not be considered as an unknown parameter anymore.

Ephemeris Error

The ephemeris is a couple of values that gives the positioning information of astronomical objects in the sky at a given time, and the ephemeris error is caused by the difference between the transmitted broadcast ephemeris in the navigation message and the true satellite location. The effect of satellite ephemeris is estimated by the ground monitor station. All satellite ephemeris effects are computed by the stations based on the previous measurements of satellite motion and uplinked to the satellite in the navigation data message for rebroadcast to the users. Since the satellites in the space are affected by various forces such as the Moon, the Sun and the gravity of Earth, the ground monitor stations are hard to model and predict precisely. Therefore, the ephemeris effect in the navigation message contains a residual error. The ephemeris error can be projected to three directions related to the satellite orbit: radial, along-track and across-track (see Figure 2.3) (Warren, 2002). The error in the radial direction is the smallest. Fortunately, the along-track and across track errors are also small along the line of sight direction which directly affects the pseudorange measurement error (Kaplan & Hegarty, 2006). The current estimate of the root mean square (rms) range error due to the ephemeris parameters is about 1.5 m (Misra & Enge, 2001).

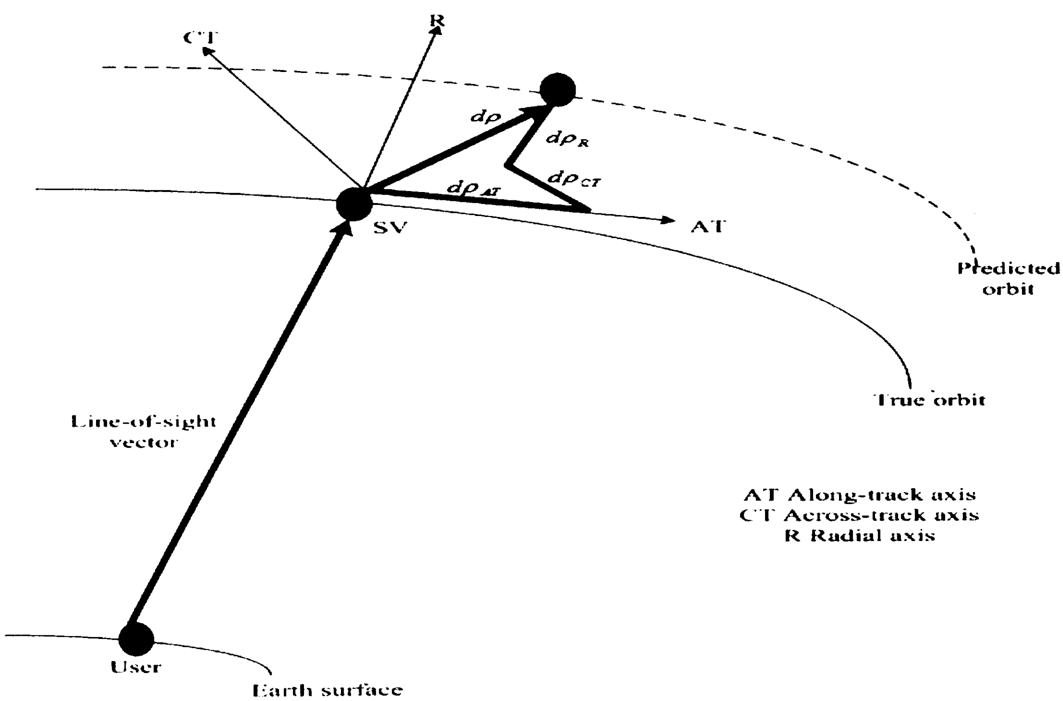


Figure 2.3 Ephemeris error components

Ionospheric Error

The GNSS signals are also affected by the medium they pass through when they travel from the satellites to the user receivers. GNSS signals do not travel at the vacuum speed of light when they are transmited through the atmosphere. The atmosphere changes the velocity of propagation of radio signals due to the refraction. The earth atmosphere is mainly divided into two regions: the upper atmosphere and the lower atmosphere. The upper atmosphere, normally referred as the Ionosphere, is the region located between about 50 km and about 1000 km above the earth. The ionosphere is the region of ionized gases which are characterized by the free electrons and ions. The ionization is caused by ultraviolet (UV) rays from the sun. When the sun rises, its UV ionizes a portion of gas molecules in the ionosphere and releases free electrons and ions which affect electromagnetic wave propagation and consequently the GNSS satellite signal broadcast (Kaplan & Hegarty, 2006; Misra & Enge, 2001). The signal is delayed due to the proportion of the total number of free electrons encountered and is also inversely proportional to the carrier frequency squared ($1/f^2$)(Parkinson et al., 1996). The code phase measurements are delayed by the same amount by which carrier phase measurements are advanced. Therefore, three ways can be used to correct the ionospheric effect. Firstly, it can be estimated by the internal broadcast ionospheric model. Secondly, it can be reduced by the dual-frequency receivers. Finally, it can be modulated by the real time updated from the multi-monitor stations (Klobuchar, 1991).

Tropospheric Error

Troposphere is the lower part of atmosphere from the Earth's surface up to 50 km. Unlike the ionosphere, the troposphere is a non-dispersive medium with respect to radio waves of frequencies up to 15 GHz. Consequently, it will delay all GNSS signals by the same amount. The signal delay depends on pressure, temperature, and relative humidity along the signal path. These components can be subdivided into the dry component and the wet component. Accordingly, the refractivity of troposphere, N_{trop}, can be expressed as follows:

$$N_{trop} = N_{dry} + N_{wet} \ . \tag{2.5}$$

The dry part typically contributes about 90% of the entire tropospheric delay, and it can be predicted with an accuracy of about 1%. The remaining 10% is the contribution from the wet component. The wet component is difficult to predict, since the water vapour density varies from time and space, and thus the prediction accuracy is only about 10%-20%. The troposheric effect reaches the minimum in the zenith direction (about two meters), and increases to about 25 m for a satellite at about 5^{o} of the elevation angle (Brunner & Welsch, 1993). Troposheric effect can be predicted by the tropospheric models such as the Hopfield model, Black and Eisner (B&E) model (Spilker, 1996) and Saastamoinen model (Saastamoinen, 1972). The tropospheric effect can also be mitigated by the differential technique and the residual effect with a baseline length of 100 km about 10-15 cm (Misra & Enge, 2001).

Multipath Error

In the errors of ε_{ρ_i}, the multipath error is a major error source, especially in the weak signal environment like the urban canyons or the forest. Multipath is the error caused by the reflection of a direct GNSS signal from such structures as buildings, vehicles and trees. Its impact on the measurements totally depends on the surroundings of the receiver antenna; thus, it is very difficult to predict and to compensate by the general model, and also cannot be mitigated by the differential processing. Therefore, the effect of multipath error can hugely depend on the reflecting geometry. However, multipath is repeated from day-to-day at a given location because of the periodic of satellite orbits which provide the repeated satellite-receiver geometry. Typically, the C/A-code multipath errors are in order of 20cm to several meters, depending on environmental conditions. Most multipath mitigation techniques are based on the design of suitable antenna site selection, receiver design, and error detection techniques for multipath minimization. For the code multipath mitigation, methods such as Narrow Correlator technique (van Dierendonck et al., 1992), Strobe Correlators, or Multipath Estimation Delay Lock Loop (MEDLL) technology (Ray, 2000), are well used in the receiver design.

Receiver Clock Error

The receiver clock error is the offset of the receiver clock from the reference GNSS time. Depending on the quality of the oscillator used in the GNSS receiver, the error could range from 200 ns up to a few ms, and the measurement ranges could therefore vary from a few metres to a few thousand kilometres. The error can be eliminated by single differencing between two satellites. However, in this thesis, the receiver clock error is treated as an unknown parameter in the position computations. Hence, the equation (2.4) can be expressed as

$$\rho_i = r_i + c \cdot t_u + \varepsilon_i \tag{2.6}$$

where ε_i is the composite of errors produced by, e.g. atmospheric delays, multipath, satellite ephemeris errors.

Calculating the User Position

In Figure 2.2, the satellite position is expressed as (x_i, y_i, z_i) and the user position is denoted as (x_u, y_u, z_u), for which x, y, z are the values in the Earth-Centered Earth-Fixed (ECEF) coordinate system. Therefore, the geometric range, r, can be computed by the following equation:

$$r_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2}\ . \tag{2.7}$$

Substitute equation (2.7) to equation (2.6), the pseudorange equation can be expressed as

$$\begin{cases} \sqrt{(x_1 - x_u)^2 + (y_1 - y_u)^2 + (z_1 - z_u)^2} + c \cdot t_u + \varepsilon_1 = \rho_1 \\ \sqrt{(x_2 - x_u)^2 + (y_2 - y_u)^2 + (z_2 - z_u)^2} + c \cdot t_u + \varepsilon_2 = \rho_2 \\ \sqrt{(x_3 - x_u)^2 + (y_3 - y_u)^2 + (z_3 - z_u)^2} + c \cdot t_u + \varepsilon_3 = \rho_3 \\ \vdots \\ \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} + c \cdot t_u + \varepsilon_i = \rho_i \end{cases} \tag{2.8}$$

Where i depends on the number of satellites which have been tracked in view. Therefore, the GNSS navigation systems can find a three-dimensional (3D) position only when the GNSS antenna can receive at least four different satellite signals (three satellites required with the height information added from the Digital Terrain Models) to solve four unknowns, including three coordinates of the user position (x_u, y_u, z_u) and one receiver clock offset (t_u). Normally, the GNSS solution can be solved by iteration techniques of the least squares (LS) method (Kaplan & Hegarty, 2006; Leick, 2004). To use LS method, the nonlinear mathematical model equation (2.8) can be denoted as the following matrix:

$$f(X) = l \qquad (2.9)$$

where $X = (x_u, y_u, z_u, ct_u)^T$ are the parameters and $l = (\rho_1, \rho_2, \cdots, \rho_i)^T$ are the observations. If $X_0 = (x_0, y_0, z_0, ct_0)^T$ is assumed as the approximate estimate coordinates for the user and the associated estimate predicted receiver clock offset, then:

$$X = X_0 + \Delta X \qquad (2.10)$$

where $\Delta X = (\delta x_u, \delta y_u, \delta z_u, c \cdot \delta t_u)^T$. Therefore,

$$f(X) = f(X_0 + \Delta X) \qquad (2.11)$$

The right hand function can be linearized around the approximate parameters X_0 by using Taylor series, giving:

$$A \Delta X = B + v \qquad (2.12)$$

where

$$A=\begin{bmatrix} \frac{\partial f_1}{\partial x_u} & \frac{\partial f_1}{\partial y_u} & \frac{\partial f_1}{\partial z_u} & \frac{\partial f_1}{\partial t_u} \\ \frac{\partial f_2}{\partial x_u} & \frac{\partial f_2}{\partial y_u} & \frac{\partial f_2}{\partial z_u} & \frac{\partial f_2}{\partial t_u} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_i}{\partial x_u} & \frac{\partial f_i}{\partial y_u} & \frac{\partial f_i}{\partial z_u} & \frac{\partial f_i}{\partial t_u} \end{bmatrix}; \tag{2.13}$$

$$\frac{\partial f_i}{\partial x_u}=\frac{x_i-x_u}{\sqrt{(x_i-x_u)^2+(y_i-y_u)^2+(z_i-z_u)^2}};$$

$$\frac{\partial f_i}{\partial y_u}=\frac{y_i-y_u}{\sqrt{(x_i-x_u)^2+(y_i-y_u)^2+(z_i-z_u)^2}};$$

$$\frac{\partial f_i}{\partial z_u}=\frac{z_i-z_u}{\sqrt{(x_i-x_u)^2+(y_i-y_u)^2+(z_i-z_u)^2}};$$

$$\frac{\partial f_i}{\partial t_u}=1.$$

A is referred to as the design matrix, which contains the direction vectors pointing from the approximate user position to the available satellites, and $B=[\rho_1-\hat{\rho}_1,\rho_2-\hat{\rho}_2,\cdots,\rho_i-\hat{\rho}_i]^T$. v is residual pseudorange errors and is assumed as normally distributed with zero mean and variance $Cov(l)$. $Cov(l)$ is the variance-covariance matrix of the observations. Generally, the $Cov(l)$ is diagonal, which means the observations are uncorrelated. Therefore, variance-covariance matrix of the pseudoranges can be shown as

$$Cov(l) = \begin{bmatrix} \sigma_{\rho_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{\rho_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{\rho_l}^2 \end{bmatrix}. \tag{2.14}$$

Because a small standard error associated with an observation means that a high weight is assigned to it, the weight matrix would be $W = Cov(l)^{-1}$.

The least squares method is to minimize the function $v^T Wv$. Take (2.12) to replace v

$$\begin{aligned} v^T Wv &= (A\Delta X - B)^T W(A\Delta X - B) \\ &= (\Delta X^T A^T - B^T)(WA\Delta X - WB) \\ &= \Delta X^T A^T WA\Delta X - B^T WA\Delta X - \Delta X^T A^T WB + B^T WB \end{aligned} \tag{2.15}$$

The minimum of $v^T Wv$ must occur at a value of ΔX that gives a zero for the gradient. Hence, set the gradient to be zero and seek a value that will minimize $v^T Wv$:

$$\frac{\partial v^T Wv}{\partial \Delta X} = 2A^T WA\Delta X - 2A^T WB = 0; \tag{2.16}$$

$$\Delta X = (A^T WA)^{-1} A^T WB \ . \tag{2.17}$$

Use the improved estimate ΔX to iterate until the change in the estimate is sufficiently small. Once the unknowns ΔX is obtained, the user coordinates and the receiver clock offset can be computed by equation (2.10). The navigation performances of GPS will be described in Chapter 4.

2.2.2 DGPS Overview

Differential Global Positioning System (DGPS) is an enhancement to the GPS. It is a method to improve the accuracy and integrity performance (details in Chapter 4) of the standalone GPS. The principal technique of DGPS is to use one or more reference stations with known locations to determine the biases in the satellite measurements; then these biases are provided as differential corrections to the rover/user receivers via a data link so that the correlated errors at the rover/user receivers can be mitigated (Kaplan & Hegarty, 2006; Misra & Enge, 2001). Figure 2.4 gives basic concepts of the local-area DGPS (LADGPS). The reference station T_0 is equipped with a GPS receiver, and the accurate position of this station in ECEF coordinates (x_0, y_0, z_0) is known by the previous surveying. According to the ephemeris data in the navigation message, the reference station can get all the satellite positions in view. Therefore, any geometric range between the station and the satellite, R_i, can be calculated by

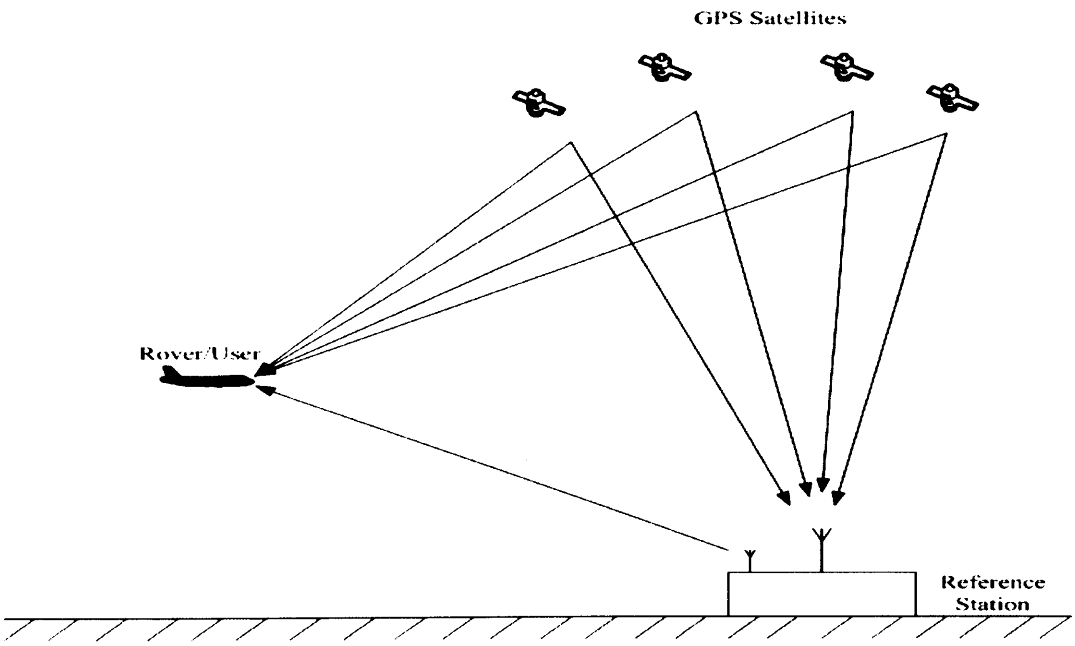


Figure 2.4 The concept of LADGPS

$$R_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} \; , \tag{2.18}$$

where (x_i, y_i, z_i) are the visible satellite positions in the navigation message. Additionally, the reference station could also receive the pseudorange measurements from each satellite. So the differential correction for each satellite can be computed by

$$\Delta\rho_i = R_i - \rho_i = -ct_0 - \varepsilon_i \tag{2.19}$$

where ct_0 represents the reference station clock offset from the GPS system time and ε_i is the i[th] satellite pseudorange error from the satellite to the reference station. Once the reference station gets the corrections, the corrections are broadcast to the rover/user receiver. Then the same satellite pseudorange measurements for the rover/user can be corrected by

$$\begin{aligned} \rho^i_{cor} &= \rho^i_u + \Delta\rho_i = R^i_u + ct_u + \varepsilon_u + \Delta\rho_i \\ &= \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} + cdt + d\varepsilon \end{aligned} \tag{2.20}$$

where $d\varepsilon$ is the residual pseudorange error and dt is the difference between the user and the reference station clock. Therefore, if more than four satellites are tracked, the position of the rover/user can be computed. Since the correction pseudoranges are used, a more accurate solution is attained (Kaplan & Hegarty, 2006).

Obviously, the differential correction can only correct the spatial and time correlated errors in the pseudorange, i.e. satellite ephemeris, satellite clock error, troposheric and ionospheric errors. That is to say, the errors, such as multipath, receiver noise and interference, still cannot be corrected by the DGPS. Further, the accuracy of DGPS depends on the closeness of the user to the reference station and also on the time delay of transmission of the corrections. In other words, the closer the user to the reference station and the closer the measurement epochs, the similar the correlated errors between the user and the reference station are. With the increasing distance of baseline and time latency between the user and the reference station, the correlation between errors is reduced.

To extend the coverage of LADGPS and reduce the spatial decorrelation impact, the concept of Wide-area DGPS (WADGPS) is introduced. The WADGPS includes a network of reference stations determines and continually updated the time and space varying components of the total error over the whole coverage region, and transfers the correction values to users within coverage region (Ashkenazi et al., 1992; Brown, 1989; Kaplan & Hegarty, 2006; Kee et al., 1991). A set of reference stations is deployed in a wide area, and each of them process the measurement differential corrections. The corrections are provided to the central processing site, and then are broadcast separately to the users for different error sources. Therefore, the users get a more accurate solution by using weighed corrections from stations. The weight only depends on the geometric of the user and reference stations, which means the largest weight is assigned to the closest reference station.

The accuracy delivered by DGPS can be sub-meters. It is very useful for many civil applications, such as guidance and approach situations in aviation navigation, safety critical train controls, as well as harbours and restricted waterways in marine navigation. All of these applications require high accuracy and integrity for which the standalone GPS is hard to offer. The DGPS technology is now widely used by navigation users in all GNSS areas.

2.2.3 Real-time Kinematic (RTK)

RTK is a technique based on the use of carrier phase measurements of the GPS for the high precision navigations and land surveying. RTK follows the same concept as the DGPS, but use the carrier phase measurements as its signal rather than the code pseudorange measurements in the DGPS. That is why sometimes the RTK method is denoted as carrier phase differential technique (Hofmann-Wellenhof et al., 1997).

The RTK employs a single reference station at a known location, measures the carrier phase measurements to form the differential corrections, then re-broadcasts to the rover receiver. Then the rover corrects its own carrier phase measurements to mitigate the correlated errors between the reference station and the rover. Carrier phase measurements are much more precise than the code pseudorange measurements, but they contain an unknown integer initialization constant, the so-called "integer ambiguity". Therefore, RTK positioning has to resolve integer ambiguities to achieve a high level of precision. This integer ambiguity problem can be solved by a technique called on-the-fly (OTF) ambiguity resolution. The success of OTF ambiguity resolution highly depends on the satellite geometry, and the ambiguities are solved faster when more satellites are tracked (Misra & Enge, 2001).

The expected positioning accuracy of the RTK method is about 2 to 5 cm within a baseline less than 15 km. In order to provide a reliable RTK service in a large area, multiple reference stations are needed to form a RTK network. The RTK and RTK network techniques are now widely used for a range of applications, such as land surveying, marine surveying, high precise navigation, auto-drive or auto-pilot system and precise farming (El-Rabbany, 2002; Langley, 1998).

2.2.4 Other Augmentation Systems

GNSS augmentation systems are designed to improve the navigation system performances of accuracy, integrity, and availability. According to different applications, the augmentation system can be divided into Ground Based Augmentation System (GBAS), Satellite Based Augmentation System (SBAS) and Aircraft Based Augmentation System (ABAS).

The above described DGPS and RTK systems are both GBAS systems. The U.S. Federal Aviation Administration (FAA) has also developed a GBAS system for the aircraft landing application, called Local Area Augmentation System (LAAS). The LAAS can support the precision approach, landing, and departure of all flights in the coverage area. The LAAS consists of local reference receivers and a central monitor station. The local reference receivers send the received pseudorange measurement data to the central monitor station. The central monitor station then processes the data to form the psuedorange corrections and correction rates, and broadcast to the airborne users via a Very High Frequency (VHF) data link. The aircraft receiver uses these corrections to correct its own GPS signals and to improve a position solution. The LAAS can obtain an accuracy of around 16m in the lateral position and about 4m in the vertical position (Lay et al., 2003; Prasad & Ruggieri, 2005). The other GBAS systems around world includes Australia's Ground Based Regional Augmentation System (GRAS), Russia's proposed differential correction and monitoring service, and the U.S. Nationwide Differential Global Positioning System (NDGPS).

The SBAS is a system which provides differential corrections and integrity information by using GEO satellites to broadcast messages. The DGPS data are broadcast through transponders on the GEO satellites on the same band as the GNSS constellations. A SBAS system contains a network ground reference station which monitors GNSS satellite signals, master stations which process reference station data and generate SBAS signals, and uplink stations to transmit the data to the GEO satellites in the space. At the time of this thesis writing, several SBAS systems are available or under development, such as the U.S. Wide Area Augmentation System (WAAS) and Wide Area GPS Enhancement (WAGE), the European Geostationary Navigation Overlay Service (EGNOS) system which is operated by ESA, the Japan's Multifunctional Satellite Augmentation System (MSAS), and the Indian proposed GPS and GEO Augmented Navigation (GAGAN) system. In this thesis, a summary of WAAS and EGNOS would be given asexamples of the SBAS system (see below).

The US FAA-developed WAAS (Enge et al., 1996; Skone et al., 2004) is an augmentation system for GPS to achieve the safety requirements of accuracy, integrity, and availability for aeronautical navigations. WAAS is also seen as a WADGPS system. As for 25 wide area reference stations (WRS) across the U.S., it composes the ground network to collect the measurements for the wide area master stations (WMS). The WMS processes the raw data to generate differential corrections and integrity data and send to the GEO satellites. The GEO satellites finally send back the differential corrections, which are coded in a navigation message of GPS/SPS-like signals transmitted at frequency L1 to the users. A WAAS-equipped receiver must be able to receive the additional ranging signal and to demodulate the navigation message for the differential corrections while computing their positions. The WAAS specification requires it to provide 7.6 meters or better position accuracy in both lateral and vertical directions, although the actual performance of WAAS could achieve about 1.0 meters laterally and 1.5 meters vertically at specific locations throughout most areas of United States. The WAAS also provides the integrity equivalent to or better than Receiver

Autonomous Integrity Monitoring (RAIM, described in chapter 4) and about 99.999% availability in coverage areas.

The EGNOS (Gauthier et al., 2003; Kirjner et al., 2003; Lyon et al., 2005; Soley et al., 2004) is an another SBAS system developed by the ESA, the EC and European Organisation for the Safety of Air Navigation (EUROCONTROL). Similar to the U.S. WAAS system, EGNOS provides the complementary information to the GPS, but also to the GLONASS and the future GALILEO system. The EGNOS system also consists of three segments: space segment, ground segment and user segment. The space segment includes three GEO satellites to provide a coverage area around Europe. The ground segment is composed of 34 ranging and integrity monitoring stations (RIMS), 4 mission control centres (MCC), 6 navigation land earth stations (NLES), the application specific qualification facility (ASQF), the performance assessment and system checkout facility (PACF) and the EGNOS wide area communication network (EWAN). The user segment consists of a GPS/GLONASS/GALILEO receiver and an EGNOS receiver. The EGNOS started its initial operation in July 2005 and gave about 2 meter accuracy and over 99% availability performances.

2.3 GNSS Railway Applications

As the above section summarised, that GNSS till now not only means GPS or GLONASS, it is a more comprehensive system. The fully GNSS indicates the concept including global satellite navigation systems (e.g. GPS, GLONASS, GALILEO, Beidou II), regional satellite navigation systems (e.g. QZSS, IRNSS, Beidou I), and augmentation system (e.g. GBAS, SBAS, ABAS). By 2013, more than a hundred of GNSS satellites will be available and many augmentation systems can be chosen. The high performance of accuracy, integrity and availability provided by the GNSS and

GNSS augmentation systems gives the GNSS the chance to be an important component in business, government and transport sectors.

In recent time, the GNSS has been used in many applications. Vehicles can use the GNSS to determine their locations, speeds, directions and also display them on the moving maps. Boats and ships can also install the GNSS receivers to navigate in lakes, seas and oceans. The GNSS and GNSS augmentation are used in the aviation for en-route navigation, approach, landing and departure of all flights. It can also been used in surveying and mapping. The high accuracy performance of the GNSS let it useful for all non-safety related applications. Additionally, the integrity information can also be provided by the GNSS satellite itself (future) and GNSS augmentation systems; therefore, it brings about good opportunities for using GNSS at safety critical applications at all transport sectors. The successful applications of GNSS on the safety-critical aviation navigation and maritime applications, suggests a prospect of applying it for railway safety-critical applications (Kaplan & Hegarty, 2006; Kiss, 2000; Prasad & Ruggieri, 2005).

However, using the GNSS for the safety critical railway application is more complicated than for the aviation and maritime applications. One significant problem is the line-of-sight (LOS) problem. For the GNSS, it requires the necessity for at least four (five for enable error detection) visible satellites to compute a position. In an open terrain, oceans or on aircraft, this is easily achievable, but for railways which often run through low satellite visibility areas such as deep cuttings, forests, urban canyons and tunnels, it will be hard to track sufficient satellites in these areas which will cause the loss of position and integrity. This is obviously unacceptable in a safety-critical application. Howbeit, this deficiency can be compensated by integrating with other sensors such as the track data base, odometer, INS, local area networks (wireless technology) and SBAS.

At the railways sector, the traditional train control, location and other automation technologies utilize the track-side equipments and transponders. It is a track-side block circuit system which allows the train to occupy a certain section of track between two block points. The whole section is reported to be occupied regardless of the length or speed of the train and the exact position of train in that section is not known. All these systems require high investment, high operational and maintenance cost, and they also face the risk of the vandalism. However, compared with traditional train location system, GNSS has the benefits such as lower initial cost (e.g. all necessary equipments can be stored on the locomotive), less maintenance (e.g. transponders needed to be replaced owing to the vandalism), and potentially increasing the capability of railway lines for both freight and passenger trains due to the high accuracy of GNSS. Some research projects have been demonstrated within Europe (details are described in Chapter 3), e.g. APOLO, LOCO, DemoOrt, GADEROS, RUNE, INTEGRAIL, SOCRATEC etc. This thesis is also to explore the use of GNSS to the railway applications with a new invented method to integrate the GNSS with the database (described in Chapter 5).

Chapter 3 Railway Control Systems

In this chapter, railway control systems are briefly introduced. In the first section (3.1), traditional railway control systems are introduced and compared in terms of their advantages and disadvantages. Section 3.2 discusses the reason for choosing the satellite-based railway control system for future. This chapter concludes with the description of flaws of different complemented sensors for GNSS and the necessary sensors for GNSS in the railway applications.

3.1 Traditional Railway Control System

Because trains are fixed on the railroad and are not allowed to change the railroad unless on the switch point, the possibility of the train collision is high. Furthermore, due to the high speed of train operation, trains cannot be stopped quickly and also cannot always be stopped within the sight distance of drivers. Therefore, a railway traffic control system needs to make efforts to avoid any possibility of collision. Under the traffic control, trains authorised movement is delivered from a responsible man (e.g. a signalman, a flagman or stationmaster) for each section of a rail network to the train crew. The different physical equipments are set for different rules. These operation rules are called differently in the different country such as method of working in U.K., method of operation in U.S., or safeworking in Australia.

The simplest railway traffic control system is the timetable operation. In this operation system, all trains ran according to a fixed timetable. It requires every train crew to be familiar with the fixed schedule (Bryan, 2006; Goverde, 2005). Trains needed to travel on each section of track at their scheduled time. No other trains were allowed to enter the same section at that time. For a single-track railroad, meeting points must be scheduled for the train running in the opposite direction. The train must wait for the other one at a passing pass and it was not allowed to move till the other arrived. Figure 3.1 below shows an example of passing pass.

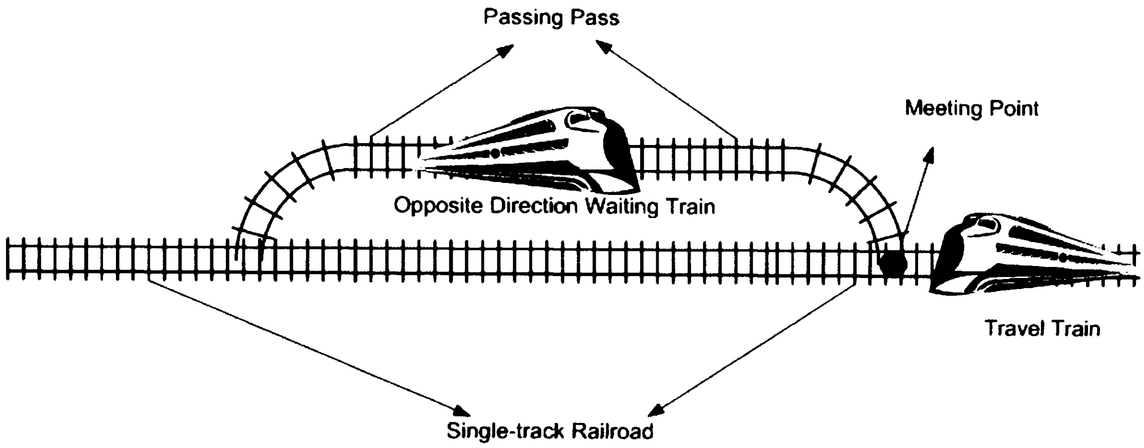


Figure 3.1 Example of Single-track Railroad

The timetable contained numerous types of important information of immense value to the driver. It listed different stations, the milepost of the station, the station number, and a list of scheduled trains. Usually all the first, second and third class regularly scheduled trains were listed. Figure 3.2 gives an example of the timetable.

St. Louis District—Kirkwood to Jefferson City									
Station Numbers	Miles from St. Louis	Time Table No. 44	TRAINS WESTWARD						
			FIRST CLASS			SECOND CLASS			THIRD
			15 Psgr	11 Psgr	19 Psgr	61 Red Ball Freight	71 Merchandise	77 Red Ball Freight	91 Local Freight
		STATIONS	Daily	Daily	Daily	Daily	Daily	Daily	Daily ex. Sun
—	---	St. Louis Union Station	9 00 A	2 02 P	6 40 P	11 00 A	7 30 P	10 30 P	12 01 A
13	13.48	CS KIRKWOOD WT	9 26	2 28	7 05	11 50	8 10	11 20	12 50
36	34.84	P PACIFIC JWY	9 52	....	7 32	....	....	....	2 45
54	54.93	CS WASHINGTON WCY	10 20	3 20	7 56	3 40 ·	9 10	12 12 A	4 00
125	125.33	N JEFFERSON CITY JWCTY	12 10 P	4 44 P	9 48 P	5 40 P	10 57 P	2 25 A	9 00 A

Figure 3.2 Example of a railway timetable (Robert J. Amsler, 2008)

The timetable system, however, had its disadvantages. Firstly, the track ahead was scheduled to be clear and there was no positive confirmation that it was clear. The system did not allow for engine failures and other problems alike. The timetable provided a sufficient time for the crew of a failed or delayed train to walk far enough to set up warning flags to warn other train crew. Secondly, the system was inflexible. Trains could not be added, delayed, or rescheduled without advance notice. Finally, as a corollary of the second deficiency, the system was inefficient. That is to say, to make the system flexible required the timetable to offer trains a broad allocation of time to allow for delays so that the line was in the possession of each train for longer than was necessary.

The above problems seemed to be mitigated with the invention of telegraph technology. Based upon telegraph technology, an advanced version of timetable operation was introduced, called as Timetable and Train Order (TATO). TATO system sent train orders to each train crew (Robert J. Amsler, 2008). Train orders allowed dispatchers to set up meeting point at sidings (passing pass) for trains. The train had to wait until the other passed and had to keep enough distances between trains if they were in the same direction. Therefore, the TATO operation allowed the cancellation, rescheduling, and addition of train services. The TATO operation was widely used on American railway traffic control until the 1960s, but it has been replaced by the block signalling system.

The block signalling system divides the railway lines into blocks, and only allows one train in each block at one time (GCORC, 2005). The train movement authorities are delivered by the radio signals, colour light signals and remote signals. After the Armagh rail disaster in 1889, the block signalling system becomes mandatory in

United Kingdom, and it is still used until now. Different from the timetable and TATO system for which blocks usually start and end at selected stations, blocks in the signalling-based system start and end at signals. The lengths of blocks are designed to allow trains to operate as frequently as necessary. Blocks can be kilometres long for low traffic railway lines, and can also be a few hundred metres long for the high traffic railway lines.

There are different types of block signalling systems. One is called manually-controlled block system for entering and leaving. This system needs a signalman to check the status of the block. The signalman firstly must be sure that the block has not been occupied by trains, and he then could authorize the next train to enter the block. To make sure the status of a block, a signalman needs two signals. One comes from the occupied train and the other comes from the signalman in the next block. When the train leaves the block, the driver must report the entry to the signalman of this block. At the same time, the signalman of the next block also needs to check the end-of-train marker before sending the message to the signalman of the previous block. This is to secure that there is no part of train remaining within the section. Only while these two messages are received, the signalman can let the next train to entry the block; otherwise, he has to stop the next train.

Another two systems called the permissive and absolute block control systems, which can be used in emergency and rescue situations. Both systems allow multiple trains to enter a block, but only at a slow safety speed (e.g. 20 mph or less) so that the trains can be stopped within the sight distance. The permissive block control system can be used in an emergency situation. For example, when the communication between the train driver and the signalman is disconnected, the system still permits trains to enter the occupied block at a low speed. However, the permissive block control system is not allowed to use in any poor visibility environments (e.g. fog, heavy rain, snowing). The absolute block control system can be used for rescuing failed trains. Under the authorized, the signalman must inform the driver of the precise information about the failed trains and permits the train to enter the occupied block to rescue the failed trains.

The automatic block signalling (ABS) control system is the most widely used block signalling system in United Kingdom. This system uses the automatic train detection technology to detect the train location so that it can indicate whether the block is occupied or not. According to the different technology of determination, it can be divided into two subsystems: track circuits and axle counters.

Track circuit is one of the most common approaches used to determine whether or not a block is occupied. The track at either end of each block is electrically isolated from the next block. An electrical current is fed to both running rails at one end of the block. A relay at the other end is connected to both rails. When the block is not occupied, the relay completes an electrical circuit, and is energized. When a train passes a signal and enters the block, it short-circuits the current in the block, and the relay is de-energized. Figure 3.3 presents the principle of this system operation.

power restoration after a power failure, an axle counted section cannot detect the occupied situation of the affected section only when a train has passed through it; nevertheless, a track circuited section is able to detect the presence of a train in section immediately.

3.2 Satellite-based Railway Control System

With regard to the railway control system, the railway industry mainly concerns issues such as, the safety-of-life, the flexibility, the capacity and the cost of the operational system. Although the ABS system guarantees a safe distance between trains and provides the flexibility to change the schedules, the capacity of that system is not enough and the cost to operate this system is high. Specifically, as the above section described, all the blocks in the ABS system are the “fixed block”, which means the length of the blocks are fixed. To let the high speed trains run, a long safe stopping distance is needed, and therefore, a long block is needed. Additionally, trains are kept further apart than the minimum safe stopping distance because their positions are not known precisely in the block, thus, the capacity of the ABS system is decreased. Further, the cost of the ABS system, including installing and maintenance cost, is also high. The vandalism is another big problem suffering the ABS system. The repair of vandalism is expensive and also causes safety problems. Therefore, the satellite-based railway control system can be a potential replacement for the current control system, at least be an attractive complementary system (Kiss, 2000; Juliette Marais et al., 2000; Prasad & Ruggieri, 2005; Thomas et al., 2007).

According to the high accuracy performance of GNSS, GNSS can determine the train position with an accuracy down to the meter level. Therefore, the trains can be separated by the minimum safe stop distance, and the capacity of the railway line is increased. Additionally, the lower initial cost (e.g. all necessary equipments can be installed on the locomotive) and less maintenance (e.g. no expense for the vandalism of track-side equipment) could give a big benefit for railway companies. However, GNSS suffers from the line-of-sight (LOS) problem. This is crucial in the railway environment, because trains run through the areas where there are low satellite visibility (e.g. the urban canyon areas, the deep cutting side, forests and tunnels). Any of above environments could block the satellite signals, thus leading to the insufficient visible satellites to determine the train position or to do the integrity test (J. Marais et al., 2004; Mazl & Preucil, 2003). Therefore, the future satellite-based railway control system must be an integration system which can determine its own position with great accuracy and integrity under all railway environments (Ahmad. Mirabadi et al., 2002; Ahmad Mirabadi et al., 2003; Mueller et al., 2003; Rome, 2003).

To compensate the deficiency of GNSS, GNSS can be integrated with other sensors such as odometer, INS (Inertial Navigation System), WAAS/EGNOS and the track database. Several research projects have been demonstrated within Europe, e.g. APOLO (Alcouffe et al., 2001), LOCOLOC/LOCOPROL (Mertens et al., 2003; Simsky et al., 2004), DEMOORT (Hartwig et al., 2005), GADEROS (Urech et al., 2002), RUNE (Genghi et al., 2003), ECORAIL, INTEGRAIL (Bedrich & Muncheberg, 2004), SATURN/SOCRATEC (Jose M Fraile, 1999; Jose M. Fraile & GmbH, 2000). These projects have designed different integration satellite-based train location systems. All projects are demonstrated with simulation data or the field test. The summaries of LOCOPROL and RUNE projects are given in the following part.

Rail User Navigation Equipment (RUNE) is a project developed and demonstrated by LABEN in Italy (Genghi et al., 2003). The purpose of this project is to demonstrate the improvement of the train satellite-based location system. The RUNE project introduces an autonomous satellite-based positioning system for trains in order to satisfy the European Railway Train Management System (ERTMS) requirements. The ERTMS requirements would be described in the next chapter. The RUNE system integrates three sensors, including the GPS/EGNOS receiver, Inertial Measurement Unit and the train odometer unit, and a database virtual balise map. The architecture of the RUNE project is shown in Figure 3.4 below. The project involves both the laboratory simulation and a three month on-board field-testing. All the results reach the requirements of ERTMS.

LOCOPROL/LOCOLOC is a project which aims to develop an innovative cost-effective satellite based fail-safe train location system as the core of a train protection, control and command system (Mertens et al., 2003; Simsky et al., 2004). LOCOPROL is especially designed for the low density traffic railway lines, and is proposed to achieve a significant reduction of the cost to allow their use on railway lines with the low and very low traffic density. The core of this project is to integrate the track trajectory information with the GNSS-based positioning system to reduce the three dimension train positions to one, namely, only along track is to be computed. This significantly increases availability and redundancy of GNSS satellites visibility. The output train position of this algorithm is not a single point but a segment of the track. The real position of a train is between the boundaries of this track segment. This system can also interlock with other sensors to reach the ERTMS standard. The brief architecture of LOCOPROL is presented in Figure 3.5 and the performances of this system are shown in Table 3.1.

Table 3.1 Performance of the LOCOPROL project

Accuracy	Integrity level	Availability
< 300 m (Along track)	10^{-9}-10^{-11}	< 275 m 50% (Along track) < 400 m 95% (Along track)

In sum, all above railway research shows that the requirements of railway applications (e.g. ERTMS) can be satisfied by the GNSS satellite navigation technology integrated with the track database and other sensors such as, INS, odometers, balises (electronic beacons or transponders placed between rails), etc. However, the railway industry still doubts about two issues: which sensors are the best complementary sensors for GNSS and which integration method is the most efficient technology. This is because that all supporting sensors have their benefits and deficiencies. For instance, the odometer is a cost effective sensor; however, it has random systematic errors caused by the slip and slide of the vehicle wheels or the creep phenomena. Position errors are increasing with time, and accurate velocity is hard to obtain during brakes when the weather condition is poor (e.g. rain, snow). The balise is the sensor which can correct the odometer error and track ambiguity solution, but the sensor installation and maintenance fees are high, and it also cannot modulate the velocity profile. The INS can compensate the GNSS LOS problem, but low-cost INS sensors will have large errors growth over the time. Although integrating with more expensive inertial sensors could improve the solution, it adds additional expenditure for commercial applications. The track database is cost efficienct but cannot fully compensate the LOS weakness of GNSS.

Nevertheless, the track database is a core element for GNSS because the topological positioning from GNSS must be performed with a track database. Therefore, in Chapter 5, a new rigorous mathematical model for the integration of GNSS with a track database will be introduced. A key feature of this model is its ability to model errors in both the GNSS observables and the track database in order to achieve realistic performance statistics for the combined system. Clearly, a track database railway positioning reduces, in principle, to a one dimensional positioning problem and the system only requires three satellites for Receiver Autonomous Integrity Monitoring (RAIM) and four satellites for Fault Detection and Exclusion (FDE). This new integration algorithm improves not only the accuracy and integrity performance but also the availability of RAIM on the ground. Therefore, this integration system is suitable for the low density traffic railway lines; it can also be adapted for medium traffic density lines or the future ERTMS/ETCS system dedicated to high speed and/or high-density traffic railway lines.

Chapter 4 Required Navigation Performance (RNP) for Safety Critical Railway Applications

The concept of Required Navigation Performance (RNP) was developed by the International Civil Aviation Organization (ICAO) in the early 1990s. It is regarded as a statement of the navigation performance necessary for operation within a defined airspace (Prasad & Ruggieri, 2005). RNP concepts could be used to define safety requirements for GNSS Signal-In-Space (SIS). It was originally designed for civil aviation applications, but it can also be applied to other safety related applications, such as railway safety control systems and safety harbour navigations. The RNP can be used to judge the suitability of a specific system for a particular application. As for civil aviation applications, it regulates that only a navigation system meeting the designated RNP level will be allowed to operate in the airspace. The RNP of a system entails four parameters, namely, accuracy, integrity, continuity and availability (W.Y. Ochieng et al., 1999; Prasad & Ruggieri, 2005). The four parameters are defined as follows:

- Accuracy refers to the closeness of the estimated position to the true value;
- Integrity is the ability of a system to provide timely warnings to users when the system should not be used for positioning. Integrity includes three parameters which are alarm limit, time-to-alarm and integrity risk. It requires the system to be able to detect the error which exceeds the alarm limit and delivers a warning to the user with given integrity risk in a given period time (time-to-alarm);
- Continuity is the capability of a system to provide required navigation accuracy and integrity during an intended period of operation;

- Availability stands for the percentage of time when the service of the navigation system is usable

The details of accuracy, integrity, continuity and availability will be described in Section 4.1 to Section 4.3, respectively. This chapter will also discuss the safety requirements for railway applications in the end.

4.1 Accuracy

Accuracy is the degree of conformance of a measured position or an obtained solution to the true value. Precision is to describe the degree of closeness of repeated measurements to the sample mean (Leick, 2004). The accuracy of GNSS position estimate depends on two elements: the quality of the pseudoarange measurements and the user/satellite observation geometry. The major pseudoarange errors have been described in Chapter 2. Since these errors are reasonably independent, they can be allocated to individual satellite pseudoranges and considered as effectively resulting in the User Equivalent Range Error (UERE). As for a given satellite, the UERE is viewed as the statistical sum of each pseudorange error source (Kaplan & Hegarty, 2006; Kovach, 2000). Therefore, the UERE can be expressed as follows

$$\sigma_{UERE} = \sqrt{\sigma_{eph}^2 + \sigma_{sc}^2 + \sigma_{ion}^2 + \sigma_{tro}^2 + \sigma_{noise}^2} \tag{4.1}$$

where σ_{eph} is the standard deviation of ephemeris errors, σ_{sc} is the standard deviation of the error due to the satellite clock, σ_{ion} is the standard deviation of the ionospheric delay, σ_{tro} is the standard deviation of the troposheric delay, and σ_{noise} is the standard deviation of the errors due to the multipath, hardware bias, and receiver noise.

The Dilution of Precision (DOP) or Geometric Dilution of Precision (GDOP) is the concept to describe the geometric strength of user/satellite configuration on GNSS accuracy (Kaplan & Hegarty, 2006; Misra & Enge, 2001). When visible satellites are close together in the sky, the geometry is said to be weak whereas the GDOP value is high; when satellites are spread out in the sky, the geometry is strong whereas the GDOP value is low. In general, the more spread out the satellites are in the sky, the better the satellite geometry is.

Figure 4.1 shows a simple graphical explanation of the satellite geometry effect using two satellites. In this case, a user receiver obtains pseudorange measurements from a pair of satellites, S1 and S2. If pseudorange measurements are error free, the receiver will be located at the intersection of two middle circles centred at S1 and S2. However, because the measurements are imperfect, pseudorange distances will have an uncertainty region on both sides of the estimated distance. It can be seen that the user receiver will be located somewhere within the intersection of the uncertainty area (i.e. the shaded area). As shown in Figure 4.1, the quality of pseudorange measurements is the same in both case (a) and case (b). Clearly, in case (a), the size of the uncertainty area is small when the two satellites are spread out (i.e. far apart), so the quality of the position estimates is better which means good user/satellite geometry. Conversely, in case (b), when the two satellites are close to each other, the size of the uncertainty area is large. Therefore, the quality of position estimates is poor which means poor satellite geometry.

Generally, in the GNSS, accuracy can be measured by the statistical quantity (standard deviation), given that the GNSS measurements contain no systematic errors. The accuracy is expressed by

$$\sigma_p = GDOP \cdot \sigma_{UERE} \tag{4.2}$$

where σ_p is the standard deviation of the positioning accuracy, and σ_{UERE} is the standard deviation of the User Equivalent Range Error (UERE). GDOP is the geometric dilution of precision and is only associated with the satellite/user geometry. As described in Chapter 2, the least square method is commonly employed in the standard GNSS to get the relationship between the parameters and the measurements. Equation 2.17 gives the calculation of GNSS position

$$\Delta X = (A^T W A)^{-1} A^T W \Delta\rho \tag{4.3}$$

where $\Delta X = (dx_u, dy_u, dz_u, c \cdot dt_u)^T$, $\Delta\rho = (d\rho_1, \cdots d\rho_i)^T$, i is the number of satellites which are in view, A is the design matrix, and W is the weight matrix. In the weight matrix, the usual assumption is that the components of $\Delta\rho$ are identically distributed, independent, and have a variance equal to the square of the satellite UERE. Based on this assumption,

$$\sigma_{\rho_i} = \sigma_{UERE_i} \qquad (4.4).$$

To calculate the position accuracy, which means $Cov(\Delta X)$, the method of error propagation is used. According to Equation 4.3, the covariance of the ΔX could be derived as

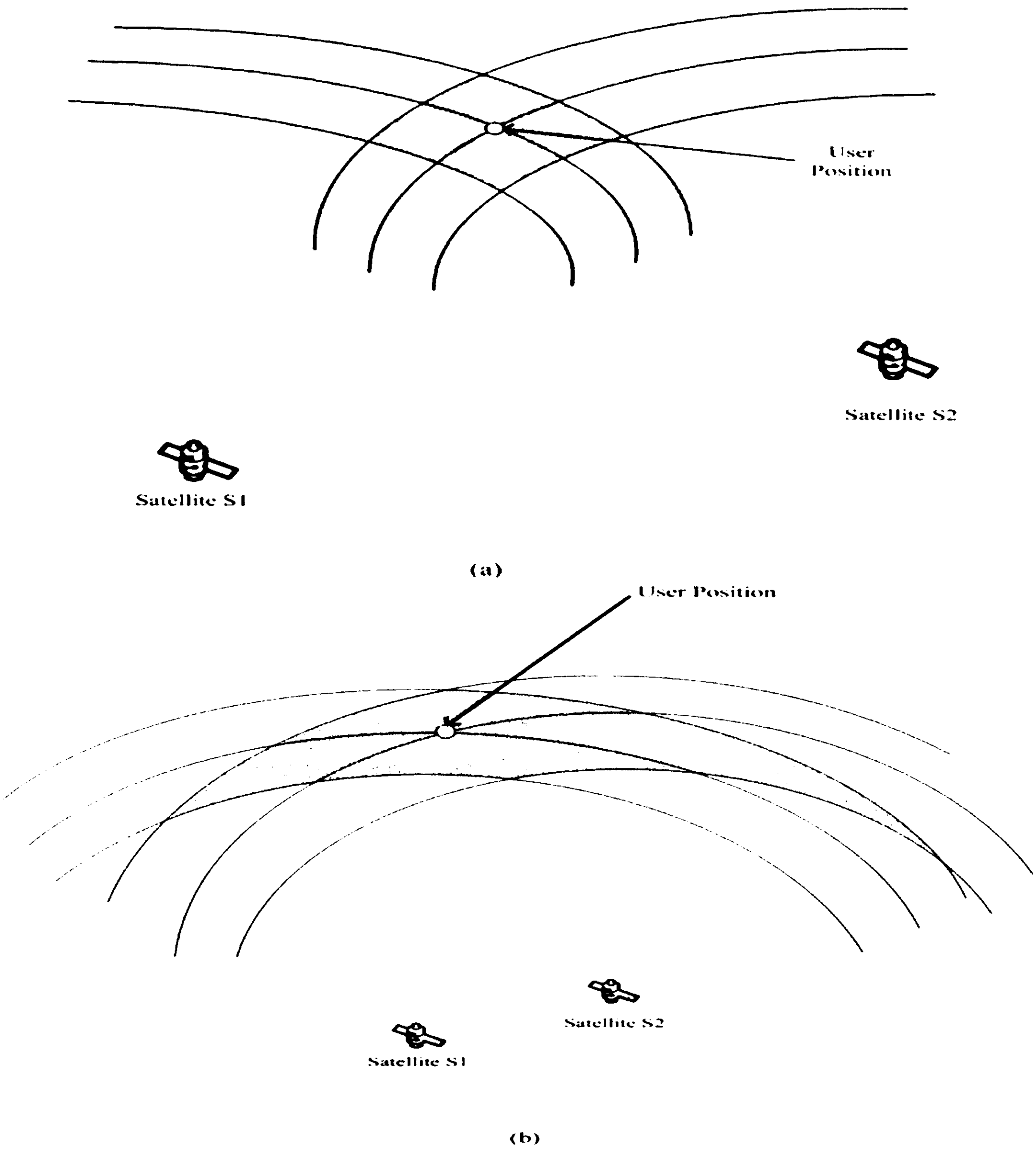


Figure 4.1 Relative Satellite Geometry and Dilution of Precision: (a) good GDOP; (b) bad GDOP

$$
\begin{aligned}
Cov(\Delta X) &= \{(A^T WA)^{-1} A^T W\} Cov(l) \{(A^T WA)^{-1} A^T W\}^T \\
&= (A^T WA)^{-1} A^T W \, Cov(l) W^T A (A^T WA)^{-1} \\
&= (A^T WA)^{-1} A^T W W^{-1} W^T A (A^T WA)^{-1} \\
&= (A^T WA)^{-1} A^T W^T A (A^T WA)^{-1}
\end{aligned}
\tag{4.5}
$$

The weight matrix W is symmetrical, so

$$Cov(\Delta X) = (A^T W A)^{-1} \tag{4.6}$$

The covariance matrices assigned to observations are called the stochastic model. Given that the covariance could come from the past experience, the manufacturer's specification, etc, it cannot be guaranteed that the stochastic model is always correct. Residuals are the biggest clue to check whether or not the stochastic model is correct. The term variance of unit weight is derived from (Leick, 2004)

$$\hat{\sigma}_0^2 = \frac{v^T W v}{n - m} \tag{4.7}$$

where n is the number of observations, and m is the number of parameters. The following results could be obtained:

- If the stochastic model is correct, we would expect to see a unit variance of about 1;
- If unit variance is greater than 1, it implies that on average, the residuals are too big, and therefore the standard errors are optimistic;
- If the unit variance is less than 1, it implies that on average, the standard errors are smaller than expected.

When the unit weight is not equal to 1, the stochastic model needs to be updated as follows

$$\sigma_{aposterior} = \hat{\sigma}_0 \sigma_{apriori} \tag{4.8}$$

where $\sigma_{apriori}$ is the standard error which is thought before the least squares process, and $\sigma_{aposterior}$ is the improved estimate standard error. Accordingly, the quality of the parameters could be updated to

$$Cov(\Delta X) = \hat{\sigma}_0^2 (A^T W A)^{-1} \tag{4.9}$$

In Equation (4.9), the accuracy is given by the Earth-centred Earth-Fixed (ECEF) coordinate system. This might not be especially useful for train location, because the railway industry is concerned with the accuracy of position in a local coordinate system which is based on the Along track, Across track and Height directions. This system is more easily visualised for train location than the ECEF system. To get the accuracy of the Along, Across and Height, two rotation transfers are required and can be achieved by following two steps: (1) transfer to the East, North and Height by the error propagation,

$$Cov_X^{(E,N,h)} = R Cov_X^{(x_u, y_u, z_u)} R^T \tag{4.10}$$

where R is the rotation matrix shown as

$$R = \begin{pmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\sin\phi \cdot \cos\lambda & -\sin\phi \cdot \sin\lambda & \cos\phi \\ \cos\phi \cdot \cos\lambda & \cos\phi \cdot \sin\lambda & \sin\phi \end{pmatrix} \quad (4.11),$$

Which λ, ϕ are the longitude and the latitude of the original point of the topographic coordinate system, respectively; (2) rotate to the Along, Across and Height directions,

$$Cov_X^{(Along, Across, h)} = R' Cov_X^{(E,N,h)} R'^T \tag{4.12}$$

where

$$R' = \begin{pmatrix} \sin\beta & \cos\beta & 0 \\ -\cos\beta & \sin\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (4.13),$$

which β is the azimuth of the track line cutting. So σ_{Along}, σ_{Across} and σ_h in the $Cov_X^{Along,Across,h}$ matrix are the position errors in the direction of the along track line, across track line and height, respectively.

From Equation 4.2, the GDOP can be defined in terms of ratio of combination of the components of $Cov(\Delta X)$ and σ_{UERE}, and is stated by the formula

$$GDOP = \frac{\sqrt{\sigma_{x_u}^2 + \sigma_{y_u}^2 + \sigma_{z_u}^2 + \sigma_{ct_u}^2}}{\sigma_{UERE}} \qquad (4.14)$$

According to Equation 4.6 and Equation 4.10, the GDOP value could be easily calculated as follows:

$$GDOP = \sqrt{a_{11} + a_{22} + a_{33} + a_{44}} \qquad (4.15)$$

where a_{ij} is the components of $(A^T A)^{-1}$ in component form

$$(A^T A)^{-1} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \qquad (4.16)$$

Commonly, the other used DOP parameters including Position Dilution of Precision (PDOP), Horizontal Dilution of Precision (HDOP), Vertical Dilution of Precision (VDOP), and Time Dilution of Precision (TDOP), can be obtained as follows:

$$PDOP = \frac{\sqrt{\sigma_{x_u}^2 + \sigma_{y_u}^2 + \sigma_{z_u}^2}}{\sigma_{UERE}} \tag{4.17}$$

$$HDOP = \frac{\sqrt{\sigma_{E_u}^2 + \sigma_{N_u}^2}}{\sigma_{UERE}} \tag{4.18}$$

$$VDOP = \frac{\sqrt{\sigma_{h_u}^2}}{\sigma_{UERE}} \tag{4.19}$$

$$TDOP = \frac{\sqrt{\sigma_{t_u}^2}}{\sigma_{UERE}} \tag{4.20}$$

Table 4.1 gives the GDOP values description (Person, 2007).

Table 4.1 Description of GDOP (Person, 2007)

GDOP Value	Rating	Description
0-1	Ideal	This is the highest level of GDOP.
1-3	Excellent	At this level, the accurate is good enough for most applications.
3-6	Good	At this level, position measurements could be used to make reliable en-route navigation.
6-10	Moderate	At this level, position measurements could be used for calculations, but to improve the quality, a more open view of the sky is recommended.
>10	Poor	At this level, the quality of position seems to be inaccurate.

4.2 Integrity

For the safety-critical railway application, in addition to the high accuracy of position, the concept of the integrity is also essential. Integrity relates to the trust that can be placed in the correctness of the information supplied by the navigation system. Its definition has been provided above. For the GNSS-based train location system, integrity is probably one of the hardest issues to be achieved due to the LOS problem. The integrity is another one which directly relates to the safety problem. A system can be called to have integrity as long as it never supplies a solution to a gross error, or has the ability to inform the user of a timely and valid warning message when the position error is out of tolerance (Ashkenazi et al., 1995).

The requirements of the integrity entail three parameters: alarm limit (including horizontal alarm limit and vertical alarm limit), time-to-alarm, and integrity risk. Three methods are used for GNSS integrity monitoring (Kaplan & Hegarty, 2006), embodying RAIM, SBAS, and GBAS. Due to different SOL requirements, methods vary.

RAIM is an approach for consistent check of the satellite measurements so as to estimate the quality of the resulting position (Hewitson & Wang, 2006; W. Y. Ochieng et al., 2001). If the consistency check fails, the receiver which has RAIM function can provide an alert to the pilot. It is essential in safety critical GNSS applications, such as in aviation, rail or marine navigation. RAIM uses the redundant GNSS pseudoarange measurements to detect the fault. Therefore, the more satellites are available, the more redundancies of pseudoranges will be available for RAIM. RAIM algorithms require a minimum of five visible satellites to detect an anomaly. The fault detection and exclusion (FDE, an enhanced version of RAIM) requires a minimum of six visible satellites to detect the faulty satellite and to remove it from the solution so that the navigation function could continue without interruption (Hatch et al., 2003). The major performance parameters of RAIM are alarm limit, time-to-alarm, false alarm rates, and probability of missed detection. The relationship between integrity and RAIM is given in Figure 4.2. Both alarm limit and time-to-alarm defined in integrity and RAIM are the same. The integrity risk is a product of the probability of missed detection multiplying the probability of a blunder occurring that will cause a position error exceeding the alarm limit. The statistic test used in RAIM is a function of the pseudoarange measurement residual and the amount of redundancy, by comparison with a threshold value which is determined in terms of the requirements for the probability of false alarm, the probability of missed detection, and the expected measurement noise.

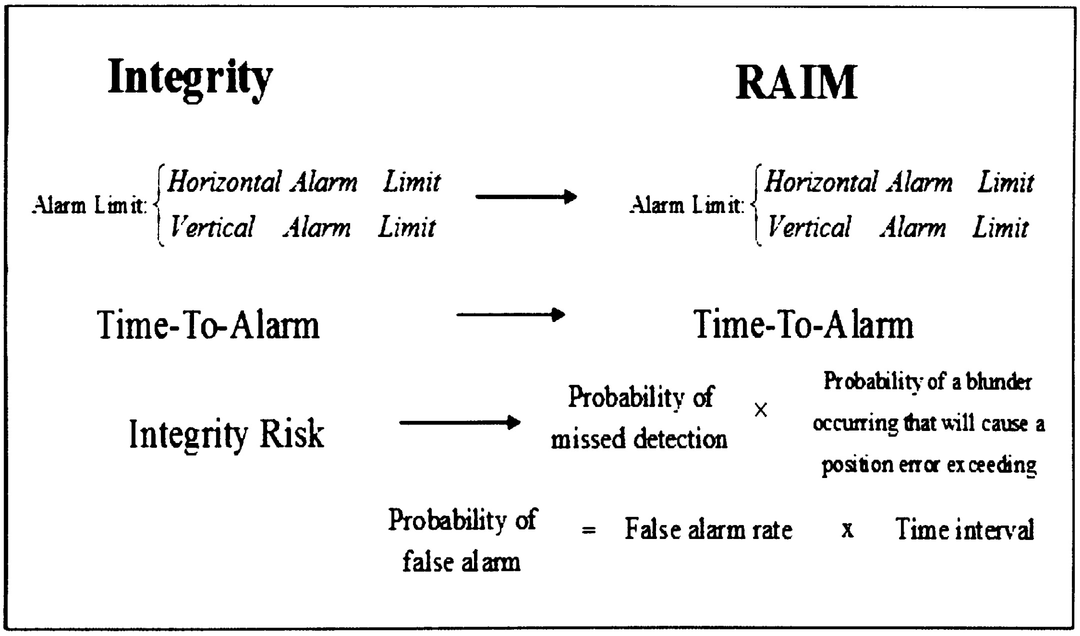


Figure 4.2 Relationship between Integrity and RAIM

The statistic method of calculation the RAIM is based on Marginally Detectable Error (MDE) algorithm. To calculate the RAIM, two major steps need to be followed (W.Y.

Ochieng et al., 2002). Firstly, with the given probability of false alarm and probability of missed detection, the minimum size of detectable gross error (i.e. a blunder or outlier) can be determined. Any large errors than the minimum size will be detected whilst smaller errors than the size will be seemed as random errors. In statistics, it is called as the internal reliability. Secondly, the impact of this smallest detectable error on the estimated receiver position can be obtained to determine whether or not the positioning error exceeds the alarm limit. This is also called as the external reliability in statistics. The whole MDE algorithm is well described by Cross (1994). As described in Chapter 2, the least squares method is used to compute the GNSS position in the current research. Therefore, only the MDE algorithm of analysing the results of a least squares computation is summarized below.

From Equation 2.9, the GNSS observations equation can be written as

$$f(X) = \ell \qquad (4.21).$$

Let ℓ_i be the i[th] observation of a vector of pseudorange observations ℓ , used in a least squares computation. To detect a gross error in observation by using a least squares method, the MDE algorithm depends on the least squares residual. If ℓ_i contains a gross error Δ_i when other pseudorange observations have only random normally distributed errors ε_i , the following null hypothesis can be set up

$$H_0 : \ell_i = \overline{\ell}_i + \varepsilon_i \qquad (4.22)$$

$$H_a : \ell_i = \overline{\ell}_i + \varepsilon_i + \Delta_i \qquad (4.23).$$

H_0 is the null hypothesis with no gross error in the observations. H_a is an alternative hypothesis against the null hypothesis, which means there is a gross error Δ_i in the observation ℓ_i. There are always two types of potential errors involved in the statistical test identified as type I error and type II error. The type I error is the error of rejecting the null hypothesis H_0 when H_0 is true. The probability of making a type I error is called the significance level of the test and is usually denoted by α. The probability of making a correct decision is called the confidence level, given by $(1-\alpha)$. The type II error is defined as the error of the null hypothesis which is accepted when ought to have been rejected. The probability of committing a type II error is denoted by β, and $(1-\beta)$ is called the power of test. The statistical test of null hypothesis against the alternative hypothesis is summarised in Table 4.2.

Table 4.2 Null hypothesis against alternative hypothesis

Decision	Actual Situation	
	H_0 is true (No gross error)	H_0 is false (There is a gross error)
Accept H_0	Correct (confidence level $1-\alpha$)	Type II error (Probability β)
Reject H_0 (Accept H_a)	Type I error (Significance level α)	Correct (Power of test $1-\beta$)

In the alternative hypothesis H_a, if ℓ_i contains a gross error Δ_i, it assumes that the observation ℓ_i still follows the normal distribution with a mean of $\overline{\ell}_i + \Delta_i$ rather than $\overline{\ell}_i$. From Equation 2.12 and the error propagation method, the covariance matrix of the estimated observed quantities, $C_{\hat{\ell}}$, can be obtained as follows

$$C_{\hat{\ell}} = AC_x^{(X,Y,Z,t)}A^T \tag{4.24}$$

where $C_{\hat{\ell}}$ is a $n \times n$ matrix and n is the number of visible satellites. The covariance matrix of the residuals, C_v, can be calculated by following

$$C_v = C_\ell - C_{\hat{\ell}} = \begin{bmatrix} \sigma_{v1}^2 & & & \\ & \sigma_{v2}^2 & & \cdots \\ \cdots & & \ddots & \\ & & & \sigma_{vn}^2 \end{bmatrix} \tag{4.25}$$

where C_v is still a $n \times n$ matrix and C_ℓ is the covariance matrix of the observations. It has been discussed in Baarda (1986) that in least squares estimation, the residuals are assumed as normally distributed population. If an observation contains a gross error, its residual will be biased, but will remain normally distributed population. To detect the gross error, the w statistic test is given by (Baarda, 1968; Cross et al., 1994; Hewitson, 2003; P.J.G., 1998)

$$\hat{w}_i = \frac{\hat{\ell}_{vi}}{\sigma_{\hat{\ell}_{vi}}} \tag{4.26}$$

where $\hat{\ell}_{vi} = \ell_i - \hat{\ell}_i$, and $\sigma_{\hat{\ell}_{vi}}$ is the standard error of $\hat{\ell}_{vi}$. If there is no gross error in the observations, $\hat{\omega}_i$ will be the standard normal distribution (i.e. zero mean and unity variance); otherwise, if ℓ_i contains a gross error Δ_i, the $\hat{w}_i$ will be still the normal distribution, but will have mean d_i where

$$d_i = \frac{\Delta_i}{\sigma_{\hat{\ell}_{vi}}} \tag{4.27}$$

and

$$\sigma_{\hat{\ell}_{vi}} = \frac{\sigma_i^2}{\sigma_{vi}} \tag{4.28}$$

where σ_i is the standard deviation of the observation i, and σ_{vi} comes from Equation 4.25. Therefore, if we specify the required significance level and required power of the test, the upper bond value of d_i, d_i^u, can be determined as follows

$$d_i^u = a + b \tag{4.29}$$

where a and b are found from the standard normal distribution table:

a from 2-tailed test with probability α;

b from 1-tailed test with probability β.

According to Equation 4.27-4.29, the upper limit on marginally undetectable errors for the i^th^ observation can be obtained by

$$\Delta_i^u = \frac{d_i^u \sigma_i^2}{\sigma_{vi}} \tag{4.30}$$

Figure 4.3 gives the example of unbiased and biased normal distribution with type I and type II errors. After the minimum detectable gross error (i.e. the maximum undetectable error) from each observation in the specified probability (i.e. internal reliability) is obtained, in order to calculate the effect on the solution of the minimum gross error detected with specified probability in the i^{th} observation (i.e. external reliability), evaluate

$$\left(A^T W A\right)^{-1} A^T W p_i = \begin{bmatrix} \delta X \\ \delta Y \\ \delta Z \\ \delta t \end{bmatrix} \tag{4.31}$$

where δ represents the effects of the minimum detectable gross error on the parameter and

$$p_i = \begin{bmatrix} 0 \\ \vdots \\ \Delta_i^u \\ \vdots \\ 0 \end{bmatrix} \tag{4.32}$$

is a null vector but with Δ^u_i on the ith element.

Similar to the accuracy, values in the along track, across track and height direction are easier to handle for the train controlling than in the ECEF coordinate system. In order to obtain along-track results, the position vector is initially rotated to the local topographic co-ordinate system by

$$\begin{aligned}
\delta_E &= -\delta_X \sin\lambda_p + \delta_Y \cos\lambda_p \\
\delta_N &= -\delta_X \sin\phi_p \cos\lambda_p - \delta_Y \sin\phi_p \sin\lambda_p + \delta_Z \cos\phi_p \\
\delta_h &= \delta_X \cos\phi_p \cos\lambda_p + \delta_Y \cos\phi_p \sin\lambda_p + \delta_Z \sin\phi_p
\end{aligned} \tag{4.33}$$

where ϕ_p is the latitude and λ_p is the longitude of the original of the local topographic coordinate. The external reliability vector is then rotated to align with the track direction to provide along-track results.

$$\delta_{Along} = \delta_E \sin\theta + \delta_N \cos\theta \qquad (4.34);$$

$$\delta_{Across} = -\delta_E \cos\theta + \delta_N \sin\theta \qquad (4.35).$$

where θ is the azimuth of the track line cutting.

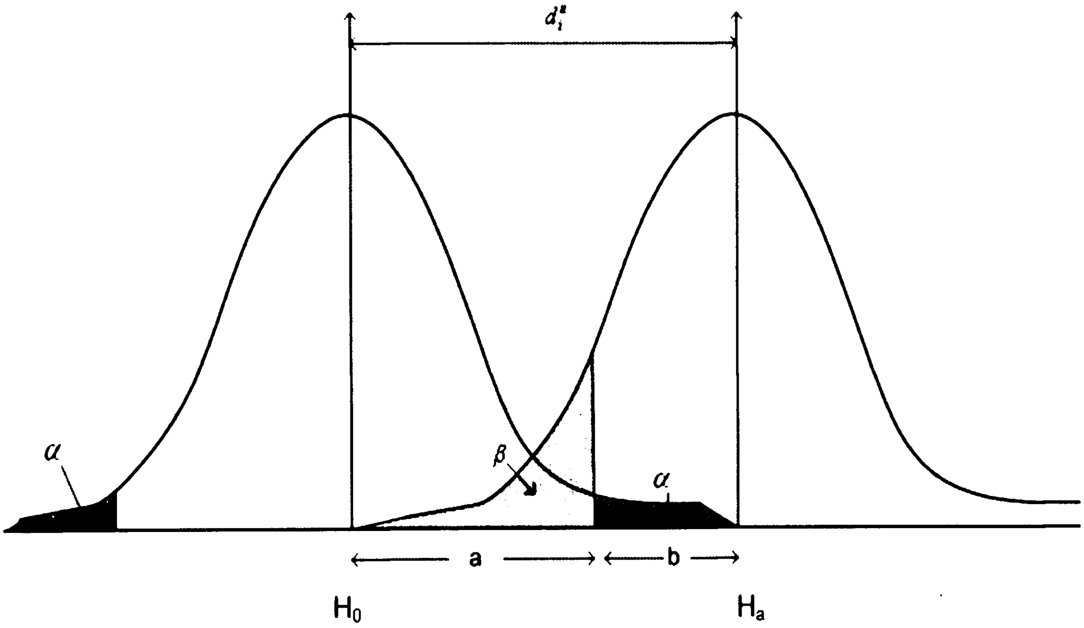


Figure 4.3 Unbiased and biased normal distributions with type I and type II errors

In the RAIM, the probability of false alarm is assigned as the type I error, α ; and probability of missed detection is assigned as the type II error, β. The output of the RAIM algorithm is the horizontal protection level (HPL), which is a circle centred at the true receiver position and assured to contain the indicated horizontal position with the given probability of false alarm and missed detection. It could also be the vertical protection level (VPL) (Kaplan & Hegarty, 2006). If the HPL does not exceed the horizontal alarm limit, RAIM is seen to be available for the intended operation;

otherwise, RAIM is seemed to be unavailable and the gross errors become a risk to the user.

4.3 Continuity and Availability

Continuity is one of the four main performances of RNP. It can be seen as the probability that the specified system performance will be maintained for the duration of a phase of operation, presuming that the system is available at the beginning of that phase of operation (DOD/DOT, 2002). Therefore, there is no uniform definition of the continuity with the specific performance requirements for any given application (Fernow & O'Laughlin, 2004). In this thesis, it is defined as the capability of GNSS signal-in-space (SIS) to provide required navigation accuracy and integrity performance without interruption during an intended period of operation. Continuity risk is a term to express the continuity. It is the maximum acceptable probability of the system which is interrupted and thus does not provide guidance information for the intended operation period. For continuity risk, only unscheduled interruptions are concerned for many applications because the scheduled maintenance activities (interruptions) can be advertised in advance. For the GPS satellite, the probability of unscheduled satellite failure is about 1×10^{-4} per hour (Kaplan & Hegarty, 2006). The Galileo SIS continuity risk is 8.0×10^{-6} in any 15 seconds for safety-of-life applications (Filip, 2006). For a combined GPS/Galileo system, continuity risk could reach better performance than for the individual GPS or Galileo system (Merino 2001). According to APOLO project (Alcouffe et al., 2001), continuity risk is suggested to be 8.0×10^{-6} in any 15 seconds.

Availability stands for the percentage of time when the services of the navigation system are usable. Due to the performance of accuracy and integrity, availability could be differed between the availability of accuracy and the availability of RAIM (Merino et al., 2001). The availability of accuracy is judged by GNSS accuracy outputs. If accuracy does not exceed threshold requirements, the system is set to be available and vice versa. As discussed in Section 4.2, the GNSS accuracy is expressed by

$$\sigma_p = GDOP \cdot \sigma_{UERE}$$

where σ_p is the standard deviation of the positioning accuracy, and σ_{UERE} is the standard deviation of the UERE. Accordingly, the availability of accuracy can also be expressed by the availability of GDOP, or the availability of HDOP, VDOP and PDOP, depending on the requirements. However, to determine the availability of accuracy for a specific location and time, the number of visible satellites, their positions and their UERE are required.

The availability of RAIM is a criterion to determine whether or not RAIM can be used. If the output HPL (or VPL) with the specific probability is less than HAL (or VAL), RAIM is said to be available; otherwise, it is not available. To assess the availability of RAIM, the HAL (or VAL), the probability of false alarm and the probability of miss detection need to be specified, and the HPL (or VPL) needs to be calculated.

4.4 Safety Critical Requirements for Railway Applications

As described above, the RNP concept was originally developed for civil aviation applications. The currently defined RNP standards are set for the operation with the necessary navigation performance in a defined airspace. Table 4.3 gives the four RNP area navigation (RNAV) parameters and their quantities (ICAO/SARPS, 2006).

Table 4.3 Required Navigation Performance [ICAO/SARPS, 2006] (H: horizontal; V: vertical)

Phase of Operation	Accuracy (95%)	Integrity			Continuity risk	Availability
		Integrity risk	Alert limit	Time to Alert		
En-route (oceanic)	2.0 nm (H)	10^{-7}/h	2.0 nm (H)	5 min	10^{-4} - 10^{-8}/h	0.99-0.99999
En-route (terminal)	0.4 nm (H)	10^{-7}/h	1.0 nm (H)	15 sec	10^{-4} - 10^{-8}/h	0.99-0.99999
NPA, Initial approach, intermediate approach, departure	220 m (H)	10^{-7}/h	556 m (H)	10 sec	10^{-4} - 10^{-8}/h	0.99-0.99999
APVI	16 m (H) 20 m(V)	2×10^{-7}/approach	40 m (H) 50 m (V)	10 sec	8×10^{-6} /15 sec	0.99-0.99999
APVII	16 m (H) 8 m (V)	2×10^{-7}/approach	40 m (H) 20 m (V)	6 sec	8×10^{-6} /15 sec	
Cat. I	16 m (H) 4-6.0 m (V)	2×10^{-7}/approach	40 m (H) 10-15 m (V)	6 sec	8×10^{-6} /15 sec	0.99-0.99999

Due to the complicated railway environment (e.g. trains not always run in open areas) and varied operation requirements, these standards cannot be used as reference standards for railway applications. New standards of RNP thereby need to be developed for railway control systems. The European Railway Transport Management System (ERTMS) has created a common Europe wide standard for different European railway system interoperation. The two main components of ERTMS are the European Train Control System (ETCS) which is a standard for on-board train control, and

GSM-R (Global System for Mobile communications - Railway) which is an international communication standard for railway communication and applications.

The ETCS is a signalling, control and train protection system designed to replace current different traditional incompatible safety railway control systems used by European Railways, especially on high-speed lines. It is used to control the movement of trains to ensure their performance to meet the safety requirements when running on different European railway networks. ETCS is classified into different equipment and functional levels, based on the equipment of the route and the approach to transmitting information to the train. There are four application levels (i.e. 0-3), detailed under.

ETCS Level 0

Level 0 refers to an ETCS vehicle which is not used on an ETCS route. The speed of the vehicle is monitored by its borne equipment so that it would not exceed the maximum speed of its type. The national trackside signals are under the driver's observation. The driver is confined to their national borders given that the meanings of signals vary from nations.

ETCS Level 1

ETCS Level 1 is a spot transmission based signalling system to overlay the existing signalling system. Because of the spot transmission of data, in order to obtain the next movement authority, the train first has to travel over the Eurobalise beacon wherein continuous trackside signals are collected and transmitted to the vehicle via cables. In addition, on this level, train speed, position and integrity are monitored by Track Circuit.

ETCS Level 2

ETCS Level 2 is a digital radio-based signal and train protection system. No more lineside signals are required, although Eurobalise is still used to locate training position. Train movements are monitored continually through GSM-R in the radio block centre. All trains automatically report their exact position and direction of travel to the Radio Block Centre at regular intervals, alongside speed information and route data.

ETCS Level 3

On Level 3, train integrity is determined on board to the very highest degree of reliability and the trackside balise is only for train location purpose. As with ETCS Level 2, movement authorities of the train on Level 3 are also monitored via GSM-R links in the radio black centre. Because it is always possible to determine which point on the route where the train has safely cleared in the radio block centre, ETCS Level 3 differs from classic operation with fixed intervals but allows trains to run on moving blocks. Table 4.4 summarizes the ETCS requirements (Genghi et al., 2003).

Table 4.4 ETCS requirements

Item	Value
Maximum distance inaccuracy	< 5 m + 5% of travelled space
Protected distance confidence level	99.7%
Maximum speed inaccuracy	2 km/h – 30 km/h
Maximum balise linking distance	2.5 km
Unavailability	10^{-7}

On ETCS level 3, the systems do not necessarily need the trackside equipment support. The train onboard system can compute its accurate position and integrity by itself. This may benefit from the use of GNSS technology. To be specific, the train control system is complicated, owing to various situations such as automatic trains control or manual trains control, freight trains control or passenger trains control, on-route control or switch-point control, low density railway lines or high density traffic railway lines. Accordingly, different RNP standards have to be applied for these varied situations. For example, we need very high integrity for both the automatic trains control and passenger trains control to reach the SOL requirements for the railway industry. We also need very high accuracy for both the switch-point control and busy traffic railway lines to facilitate tracking the train on right lines and controlling trains spacing. Basically, the applications of railway control systems can be divided into two major components: the safety related applications and non-safety related applications. The main requirements for railway non-safety related application are shown in Table 4.5. The non-safety applications can be used in such as passenger information systems, freight customers information, rolling stock maintenance etc. For safety related applications, the basic guidance of reference RNP standards are given by the GNSS Rail Advisory Forum in 2000 (GNSS Rail Advisory Forum, 2000) and APOLO project (Alcouffe et al., 2001). Table 4.6 summarises theses requirements.

Table 4.5 Main requirements for railway non-safety related applications (Alcouffe et al., 2001)

Accuracy	Integrity	Availability	Continuity	Real-time	Coverage
100 m	Non critical	98%	High	Non critical	Europe LM

Table 4.6 Safety related applications requirements

Application	Requirements					
Description	Accuracy	Integrity level	Time -to- alarm	Continuity risk	Availability	Coverage
Automatic train control on high density lines Train integrity monitoring	1 m (H)	3.3×10^{-9}	1 s	8.0×10^{-6} /15 sec	>99.98%	ELM
Train control on medium density lines	10 m (H)	3.3×10^{9}	1 s	8.0×10^{-6} /15 sec	>99.98%	ELM
Train control on low density lines	25 m (H)	3.3×10^{9}	1 s	8.0×10^{-6} /15 sec	>99.98%	ELM

According to "Railway Applications – Safety Related Electronic Systems" (CENELEC TC9x-SX9XA-WGA2, Draft 0.8 December 1994) (cited in Alcouffe et al., 2001), the safety integrity can be specified as the following four levels:

Table 4.7 Safety integrity level

Level	Safety Integrity
1	Low safety integrity
2	Medium safety integrity
3	High safety integrity
4	Very high safety integrity

The ETCS Level 3 safety integrity risk requirement is less than 2.5×10^{-10} /h, and it is equal to high safety integrity level. For the highest safety integrity level (level 4), the integrity risk is corresponded to about 4×10^{-12} /h. It should be noted that all figures in tables above are only recommended values. Further, according to the questionnaire results (Alcouffe et al., 2001) from the SNCF (French National Railway Company), some standards can be changed to more feasible values if the allocation of targets in system items is too high to demonstrate their achievement or its related cost.

However, in order to employ GNSS-based railway control system for railway safety critical applications, one unified RNP standard is insufficient. For specified systems and railway lines, it is important to define the RNP standards clearly. In this thesis, according to our integration GNSS/Track Database system, Table 4.8 introduces the reference RNP for safety critical railway applications in our demonstration. In this RNP standard, the horizontal plane requirements are divided to the along track and across track directions because the train position is located to these two directions. Since our integration system uses standard GPS receivers, the accuracy requirement is set to be 4 meters in the horizontal plane and it is enough for many applications. To obtain higher accuracy, the precise GPS receiver or DGPS could be employed but the cost will be higher. The availability requirement given here is also a recommended one, because there is not a general accepted figure for the sufficient availability. The more sensors are integrated, the higher availability is and the higher cost is too. Till now, continuity requirement is another open question for railway signalling, the figure given here is with referenced to Galileo SIS requirements. The test of continuity is beyond the scope of this thesis but is worth testing for future work. The integrity level is set from high safety integrity (level 3) to very high safety integrity (level 4). For GNSS-based railway control system, the higher integrity level is achieved, the higher possibility that the system can be used to replace the current control system or at least as a complementary system.

In the next chapter (Chapter 5), the integrated system (GNSS/Track Database) will be introduced.

Table 4.8 Reference RNP standards for safety critical railway applications

Requirement	Target Value
Accuracy	4 m (horizontal)
	10 m (vertical)
Integrity Risk	3.3×10^{-9} /h to 4×10^{-12} /h
Along track alarm limit (ATAL) Across track alarm limit (CTAL) Height Alarm limit (HAL)	10 m to 20 m 10 m to 20 m 25 m to 50 m
Time to alarm	1 s
Continuity	8.0×10^{-6} /15 sec
Availability	>99.98%

Chapter 5 GNSS/Track Database Integrated System

Global Navigation Satellite Systems (GNSS) have attracted increasing attention around the world and they have been successfully applied in aviation, vehicle and marine applications. Accordingly, applying GNSS as the primary positioning system in the railway application is a promising research area. GNSS has high accuracy, but it also suffers from line-of-sight (LOS) problems. This is crucial in the railway environment because trains often run through areas where there is low satellite visibility. Additionally, for applications such as signalling, train location systems have high safety-of-life (SOL) requirements and receiver autonomous integrity monitoring (RAIM) which is essential to satisfy SOL requirements, is especially poor in standalone GNSS when satellite visibility is restricted. In this chapter, a rigorous mathematical model for the integration of GNSS with a track database is developed. In Section 5.1, the information about the track database including the track database collection, the error in track database and the trajectory matching is given. It follows with a description of the mathematical model of the integration system, the Least Squares method to calculate the position and the strategy of system processing (Section 5.2). Section 5.3 presents the method to estimate the accuracy of the integrated system. Determination of the integrity of integrated system is shown in Section 5.4.

5.1 Track Database Information

In railway environments, rail tracks (railroads) are normally presented by the track lines which are centre lines of the tracks. The track lines themselves are not visible. Figure 5.1 shows the centre line of rail tracks. The track database is a database of rail tracks (track lines), consisting of a series of points (track data). The track data is stored as the three dimensional coordinate point in the track database. Unlike other applications such as aviation, vehicle and marine navigations, railway control system has a very special property that the trains must travel on the rail tracks. Therefore, when the train runs into any part of the rail track, it should be between two consecutive points (track data) related to the train position. The track database information is a key element in the GNSS-based railway control system. This is because that the three dimensional GNSS output positions need to be mapped to the control system and the discrete three dimensional track point database is the ideal choice for this purpose (Fraile 1999). Additionally, the track database information could be a cost-efficient complementary sensor for GNSS-based railway control system (Fraile, 1999; Simsky et al., 2004; Zheng, 2007). The track database could also help to generate the trajectory of trains travel so that the three dimensional train positions could be reduced to one dimensional train positions (Along track).

The track database collection is to obtain the coordinate information of the track points. Two ways can be used to determine the track database: the GPS technique or traditional surveying techniques.

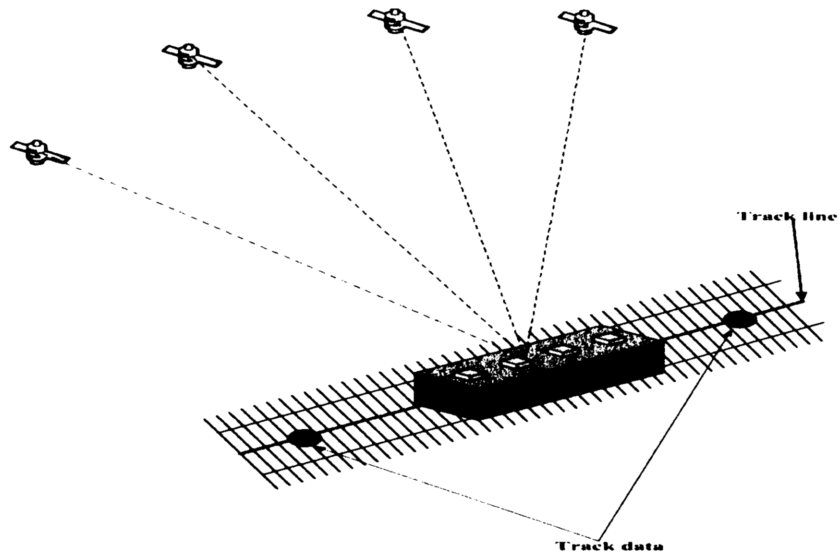


Figure 5.1 Centre line (track line) of rail tracks

In this thesis (Chapter 7), the track database was collected by the GPS equipment which was mounted on a test train around Birmingham. Because the GPS method is far more cost effective than traditional surveying methods (Euler et al., 1996). The traditional survey method can only collect about 1 to 1.5 km of rail tracks in a day. With the GPS technique, it can collect at least five times better than the traditional survey methods. Additionally, the GPS surveying can provide high accuracy track database. With the standard GPS receiver, the accuracy of track database can be on the meter level and interval between points is about 1.5 meters. If the DGPS is used, the accuracy can be down to center meter level (Euler et al., 1996).

In view of railway control system, since the train has to run along the track, the track line (trajectory function) can be used to demonstrate the movement of the train when it is on the rail track. In a three-dimensional space, the track line is a line of intersection between two surfaces. Accordingly, the trajectory function is virtually a group of functions for two crossed surfaces as follows:

$$\begin{cases} F_1(x,y,z)=0 \\ F_2(x,y,z)=0 \end{cases} \tag{5.1}$$

With the trajectory function, the relationship of three-dimensional coordinates of the train can be observed. Since the track line can be either a straight line or a curve, it needs to be analysed separately.

The case of straight line

In the case of a straight line, the straight track line in the three dimensional space is the line of intersection between two surfaces (see Figure 5.2).

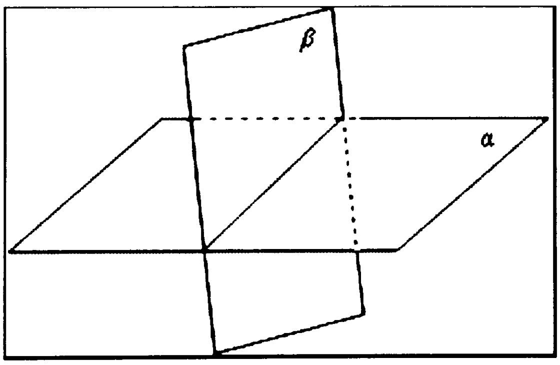


Figure 5.2 The common case of straight line in the 3D

The ordinary function of a straight line is defined as follows

$$\begin{cases} A_1x+B_1y+C_1z+D_1=0 \\ A_2x+B_2y+C_2z+D_2=0 \end{cases} \tag{5.2}$$

Figure 5.3 shows the projection of the straight line

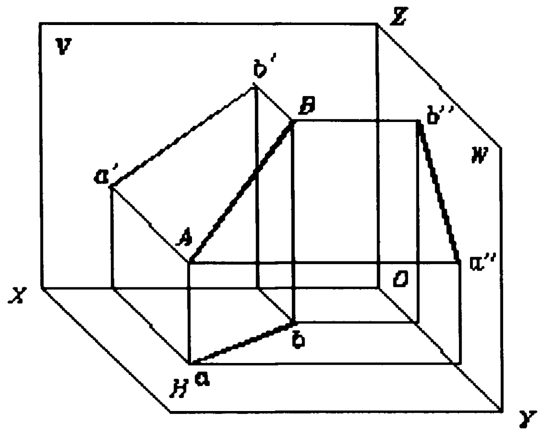


Figure 5.3 The projected case of straight line in the 3D space

Therefore, the straight line function can also be written as the project form by the following way:

$$\begin{cases} y = mx + g \\ z = nx + h \end{cases} \tag{5.3}$$

Where m,n and g,h are the slope and intersection of the line ab and line $a'b'$, respectively. In Equation (5.3), the line AB can be seen as the intersection line of plane $abBA$ and plane $a'b'BA$. In fact, it can be denoted as the intersection line of any two of the plane $abBA$, plane $a'b'BA$ and plane $a''b''BA$. With Equation (5.3), the relationship of coordinates x,y and z are defined. When the train travels on the track line AB, if x coordinate of the train and the trajectory function of line AB are known, the train coordinates y and z can be calculated by Equation (5.3).

The Trajectory Function of the Straight Line

There is no established trajectory function stored in the track database. All the track lines are instantiated by a series of continuous track points. As Figure 5.4 shows, the straight line l_1 is demonstrated by using track points X_1, X_2, X_3, X_4 and X_5.

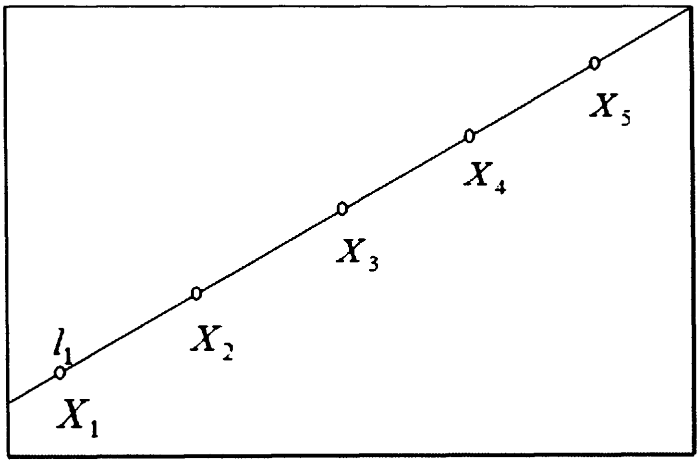


Figure 5.4 The straight line stored in the track database

Therefore, the function of l_1 can be obtained by using any two track point coordinates lying on the line.

$$\begin{cases} y = \dfrac{y_i - y_j}{x_i - x_j} \cdot x + y_j - \dfrac{y_i - y_j}{x_i - x_j} \cdot x_j \\ z = \dfrac{z_i - z_j}{x_i - x_j} \cdot x + z_j - \dfrac{z_i - z_j}{x_j - x_j} \cdot x_j \end{cases} \quad (where \quad i,j = 1,2,3,4,5, \quad i \neq j) \quad (5.4)$$

where (x_i, y_i, z_i) is the coordinate of X_i .If X_1 is the previous epoch position of the train in the track database, and the train is running along the straight line l_1, the function of l_1 can be written with point X_1 and point X_2 as the follow:

$$\begin{cases} y_u = \dfrac{y_2 - y_1}{x_2 - x_1} \cdot x_u + y_1 - \dfrac{y_2 - y_1}{x_2 - x_1} \cdot x_1 \\ z_u = \dfrac{z_2 - z_1}{x_2 - x_1} \cdot x_u + z_1 - \dfrac{z_2 - z_1}{x_2 - x_1} \cdot x_1 \end{cases} \quad (5.5)$$

where x_u, y_u, z_u is the coordinates of the train.

The case of the curve

The track line could also be a curve. Fortunately, according to the flatness of the earth surface, track line curvatures are reasonably small. The curve of track is seen to be very smooth. Actually, it is difficult to fit a curve in the three-dimensional space. Therefore, like the straight line case, the three-dimensional curve can be projected to two coordinate planes. This approach changes a three-dimensional curve fitting problem to two 2D curves, or a 2D curve and a 2D straight line fitting problem. Figure 5.5 presents the projection of curve in three-dimensional space.

The straight line trajectory function has been stated above. The process of fitting projected curve is similar in x-y or x-z coordinate plane, so only the case of x-y plane will be introduced in the following part. The Least Squares Parabola (LSP) fitting method is used to fit the curve in this thesis.

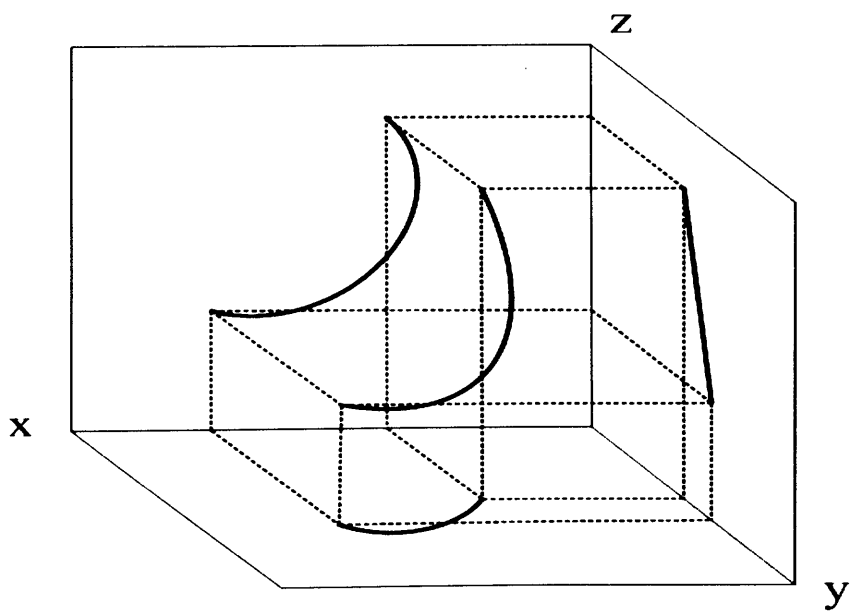


Figure 5.5 The projection of curve in 3D space.

The LSP uses a second degree curve, namely,

$$y = ax^2 + bx + c \tag{5.6}$$

to approximately fit the given set of data, $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$, where $n \geq 3$. Figure 5.6 shows the example of LSP fitting. The whole fitting system is

$$\begin{cases} ax_1^2 + bx_1 + c &= y_1 \\ ax_2^2 + bx_2 + c &= y_2 \\ \quad\vdots & \vdots \quad \vdots \\ ax_n^2 + bx_n + c &= y_n \end{cases} \tag{5.7}$$

Please note that a, b, and c are unknown coefficients while all x_i and y_i are given. The matrix form can be shown as

$$AX = B + v \tag{5.8}$$

where

$$A = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix}, \quad X = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad B = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

The best fitting curve has the least squares error, i.e. minimize $v^T v$. Please note, all the given set of track data, $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ are assumed to be equal quality. To obtain the least squares error, the minimum of $v^T v$ must occur at a value of X that gives a zero for the gradient. Hence,

$$\frac{\partial v^T v}{\partial X} = 0 \tag{5.9}$$

The unknown coefficient a, b, and c can be obtained via the following equation

$$X = (A^T A)^{-1} A^T B \tag{5.10}$$

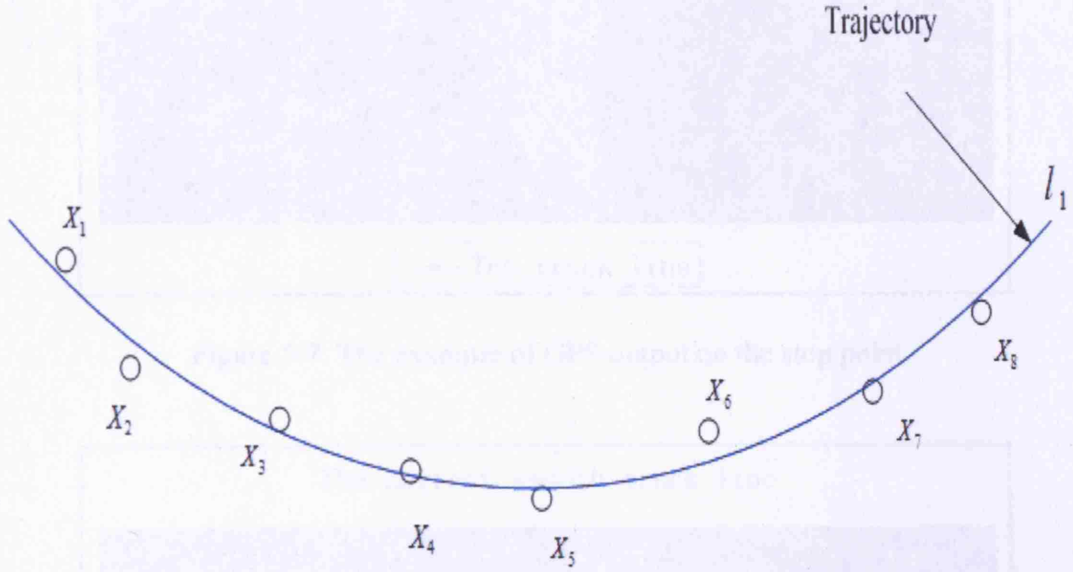


Figure 5.6 The example of Least Square Parabola fitting

Additionally, using GPS receivers to collect the track database needs to be very careful because when the train stops, the outputs of GPS are not fixed on the stop point. The positions from GPS are around the stop point during the stop time. Figure 5.7 gives an example of this situation. If GPS is used to generate the track line when the train does not move, it would obviously get the wrong track line. So the track line needs to be smooth at the stop point. Figure 5.8 shows the correct smooth track line by the manual correction.

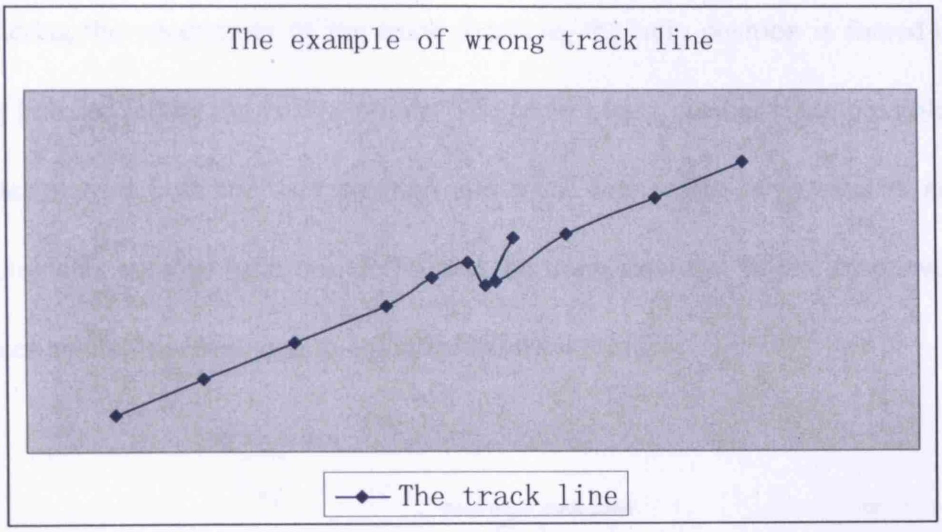


Figure 5.7. The example of GPS output on the stop point

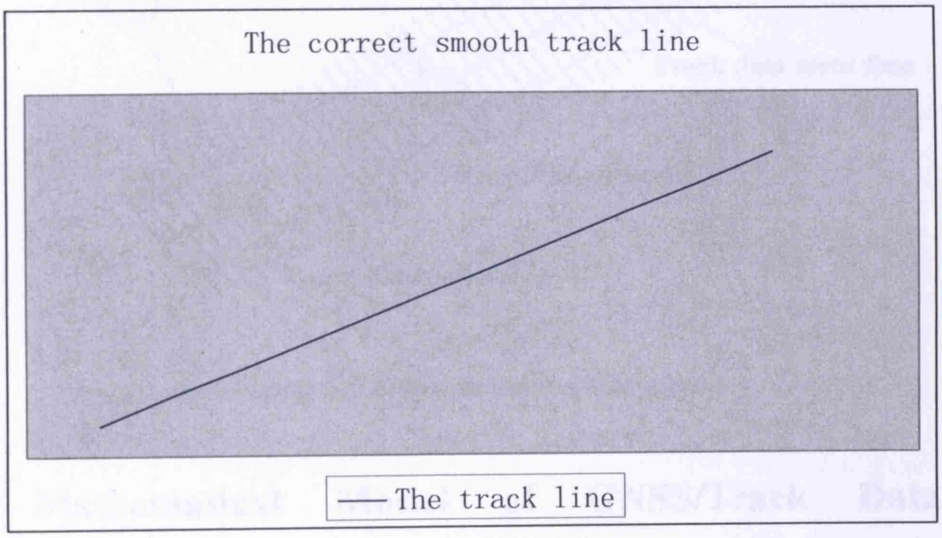


Figure 5.8 The smooth track line by the manual correction

Due to the instrument errors during the collection of the track data points, it is highly possible that the track data points do not exactly lie on the railway lines (real world). Figure 5.9 demonstrates the case of the track data error. As it reveals from Figure 5.9, the position of the train only lies on the track line which is derived by the prior and post error free track data points (red points). If the train position is calculated without considering the uncertainty of the track database, the train position is forced on the wrong line derived by the yellow points. The perfect track database is impossible, so it is better to treat both the user position and track data points as parameters and the measurements coming from the GNSS and the track database. In this case, even two satellites available are enough to calculate the train position.

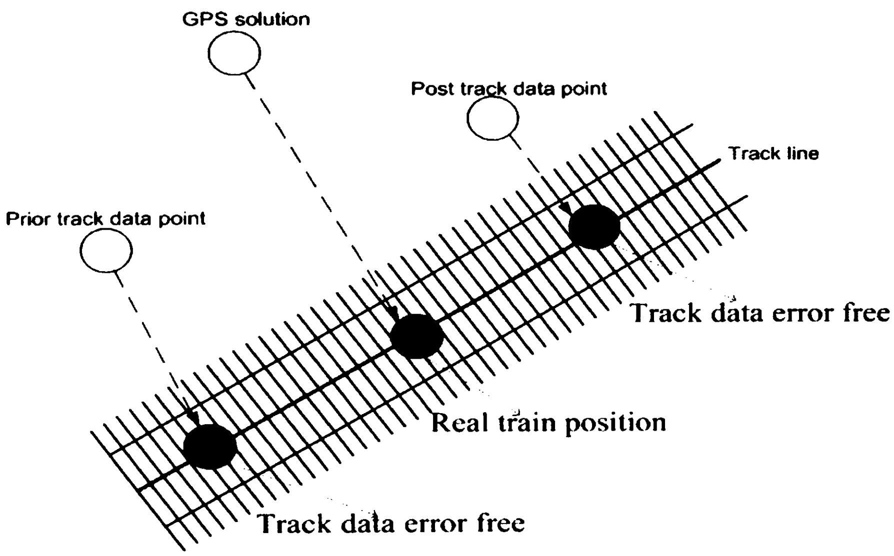


Figure 5.9 Errors in the track data base

5.2 Mathematical Model of GNSS/Track Database Integration System

The advantages of GNSS are obvious, but using GNSS for the railway applications is more complicated than other applications because rail tracks often run through the deep cuttings, the urban areas, forests or tunnels. Any of the above environments could block the satellite signals thus leading the insufficient visible satellites to calculate the GNSS position or the GNSS integrity test. To improve the performance of

GNSS-based railway control system, considering that trains must always travel on the track line, this section develops a new method of integrating the track database with GNSS to assist the train location system. A rigorous mathematical model for the GNSS/Track Database integration system is given in the following part. A key feature of this model is its ability to model errors in both GNSS observables and the track database in order to achieve realistic performance statistics for the combined system. Since the track line can be the straight line or the curve, the linear model and the nonlinear model are introduced in Section 5.2.1 and 5.2.2, respectively. In subsection 5.2.3, the strategy of processing above models is presented.

5.2.1 Linear Mathematical Model for GNSS/Track Data Integration System

When trains travel on the straight track line, if the prior and the post track points are known from the track database, the train position will satisfy the function of the straight line derived from these two points. That's to say, if the prior track point is X_1 (x_1, y_1, z_1) and the post track point is X_2 (x_2, y_2, z_2), the train position X (x_u, y_u, z_u) would satisfy

$$\begin{cases} y_u = \dfrac{y_2 - y_1}{x_2 - x_1} \cdot x_u + y_1 - \dfrac{y_2 - y_1}{x_2 - x_1} \cdot x_1 \\ z_u = \dfrac{z_2 - z_1}{x_2 - x_1} \cdot x_u + z_1 - \dfrac{z_2 - z_1}{x_2 - x_1} \cdot x_1 \end{cases} \quad (5.11)\ .$$

Equation (5.11) could, for instance, be used as an extra condition to help solve the y_u and the z_u coordinates of trains. Considering the user position and track data points as parameters, the linear mathematical model is presented as follows

$$
\begin{cases}
[(x_u-x_{s_1})^2+(\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_1})^2+(\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_1})^2]^{1/2}+c\cdot t_u & =\rho_1 \\
[(x_u-x_{s_2})^2+(\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_2})^2+(\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_2})^2]^{1/2}+c\cdot t_u & =\rho_2 \\
\vdots & \vdots \\
x_1 & =C_1 \\
y_1 & =C_2 \\
z_1 & =C_3 \\
x_2 & =C_4 \\
y_2 & =C_5 \\
z_2 & =C_6
\end{cases}
\qquad (5.12)
$$

where x_u, y_u, z_u and $x_{s_i}, y_{s_i}, z_{s_i}$ denote the user position and the i^{th} satellite position in three dimensions, respectively. Both of them are Cartesian coordinates in the Earth-centred Earth-fixed (ECEF) frame. t_u denotes the user receiver clock offset. c denotes the speed of the light. ρ_i denotes the pseudorange i^{th} satellite. C_1, C_2, C_3 are x, y, z values of the point X_1 in the track database, respectively. C_4, C_5, C_6 are x, y, z values of the point X_2, respectively.

Similar to solve GNSS alone, the linear mathematical model can also use the Least Squares method. Equation (5.12) can be denoted as the following matrix form:

$$f(X)=l \qquad (5.13)$$

where X denotes the parameters and l denotes the observations

$$X=(x_u, c\cdot t_u, x_1, y_1, z_1, x_2, y_2, z_2)^T,$$

$$l=(\rho_1, \rho_2, \cdots, \rho_n, C_1, C_2, C_3, C_4, C_5, C_6)^T.$$

If x_0 is the approximate x coordinate of the user and t_0 is the associated predicted receiver clock offset, then:

$$X = X_0 + \Delta X \tag{5.14}$$

where

$$X_0 = (x_0, t_0, x_1, y_1, z_1, x_2, y_2, z_2)^T,$$

$$\Delta X = (dx_u, c \cdot dt_u, dx_1, dy_1, dz_1, dx_2, dy_2, dz_2)^T.$$

Therefore,

$$f(X) = f(X_0 + \Delta X) \tag{5.15}$$

The right hand function can be linearized around the approximate parameters X_0 by using Taylor series, giving:

$$A\Delta X = B + v \tag{5.16}$$

where

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_u} & \frac{\partial f_1}{\partial t_u} & \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial z_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial y_2} & \frac{\partial f_1}{\partial z_2} \\ \frac{\partial f_2}{\partial x_u} & \frac{\partial f_2}{\partial t_u} & \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial z_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial y_2} & \frac{\partial f_2}{\partial z_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\frac{\partial f_i}{\partial x_u} = \frac{x_u - x_{s_i} + [\frac{y_2 - y_1}{x_2 - x_1} \cdot (x_u - x_1) + y_1 - y_{s_i}] \cdot \frac{y_2 - y_1}{x_2 - x_1} + [\frac{z_2 - z_1}{x_2 - x_1} \cdot (x_u - x_1) + z_1 - z_{s_i}] \cdot \frac{z_2 - z_1}{x_2 - x_1}}{[(x_u - x_{s_i})^2 + (\frac{y_2 - y_1}{x_2 - x_1} \cdot (x_u - x_1) + y_1 - y_{s_i})^2 + (\frac{z_2 - z_1}{x_2 - x_1} \cdot (x_u - x_1) + z_1 - z_{s_i})^2]^{1/2}}$$

$$\frac{\partial f_i}{\partial x_1}=\frac{[\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i}]\cdot\frac{(y_2-y_1)(x_u-x_2)}{(x_2-x_1)^2}+[\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i}]\cdot\frac{(z_2-z_1)(x_u-x_2)}{(x_2-x_1)^2}}{[(x_u-x_{s_i})^2+(\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i})^2+(\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i})^2]^{1/2}}$$

$$\frac{\partial f_i}{\partial t_u}=1$$

$$\frac{\partial f_i}{\partial y_1}=\frac{[\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i}]\cdot(-\frac{x_u-x_1}{x_2-x_1}+1)}{[(x_u-x_{s_i})^2+(\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i})^2+(\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i})^2]^{1/2}}$$

$$\frac{\partial f_i}{\partial z_1}=\frac{[\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i}]\cdot(-\frac{x_u-x_1}{x_2-x_1}+1)}{[(x_u-x_{s_i})^2+(\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i})^2+(\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i})^2]^{1/2}}$$

$$\frac{\partial f_i}{\partial x_2}=\frac{[\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i}]\cdot\frac{-(y_2-y_1)(x_u-x_1)}{(x_2-x_1)^2}+[\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i}]\cdot\frac{-(z_2-z_1)(x_u-x_1)}{(x_2-x_1)^2}}{[(x_u-x_{s_i})^2+(\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i})^2+(\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i})^2]^{1/2}}$$

$$\frac{\partial f_i}{\partial y_2}=\frac{[\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i}]\cdot(\frac{x_u-x_1}{x_2-x_1})}{[(x_u-x_{s_i})^2+(\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i})^2+(\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i})^2]^{1/2}}$$

$$\frac{\partial f_i}{\partial z_2}=\frac{[\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i}]\cdot(\frac{x_u-x_1}{x_2-x_1})}{[(x_u-x_{s_i})^2+(\frac{y_2-y_1}{x_2-x_1}\cdot(x_u-x_1)+y_1-y_{s_i})^2+(\frac{z_2-z_1}{x_2-x_1}\cdot(x_u-x_1)+z_1-z_{s_i})^2]^{1/2}}$$

$$B=\begin{bmatrix}\rho_1-\hat{\rho}_1\\ \rho_2-\hat{\rho}_2\\ \vdots\\ C_1-\hat{C}_1\\ C_2-\hat{C}_2\\ C_3-\hat{C}_3\\ C_4-\hat{C}_4\\ C_5-\hat{C}_5\\ C_6-\hat{C}_6\end{bmatrix}\qquad \Delta X=\begin{bmatrix}\Delta x_u\\ c\Delta t_u\\ \Delta x_1\\ \Delta y_1\\ \Delta z_1\\ \Delta x_2\\ \Delta y_2\\ \Delta z_2\end{bmatrix}$$

If $Cov(l)$ is the variance-covariance matrix of pseudoranges and track database, then the weight matrix would be $W = Cov(l)^{-1}$, defined as follows

$$W = C_l^{-1} = \begin{bmatrix} \sigma_{\rho_1}^2 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\rho_2}^2 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{x_1}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{y_1}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{z_1}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_2}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{y_2}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{z_2}^2 \end{bmatrix}^{-1}$$

where σ_{ρ_i} is the standard deviation of the ith pseudorange and σ_{x_i}, σ_{y_i}, σ_{z_i} are standard deviations of x, y, z coordinates of track data points in the track database, respectively. Therefore, the LS solution would be

$$\Delta X = (A^T WA)^{-1} A^T WB \tag{5.17}$$

After getting $x_u, x_1, y_1, z_1, x_2, y_2, z_2$, substitute them to the condition Equation 5.11, the y_u and the z_u coordinates of the user can also be obtained. In the linear mathematical model, only the x_u coordinate of the user needs to be calculated. In this sense, the three-dimensional train position problem reduces to one-dimensional train position problem. The minimum requirement to calculate the train position is now reduced to two satellites in view.

5.2.2 Nonlinear Mathematical Model for GNSS/Track Data Integration System

When a train runs on a track curve, it means the train will satisfy the function of the track curve. But unlike the straight line case, the trajectory function of track curve can not be perfectly obtained because there is no fix shape of the track curve. Therefore, The LSP fitting method is used to estimate the track curve. Figure 5.10 gives the example of track curve and estimated trajectory.

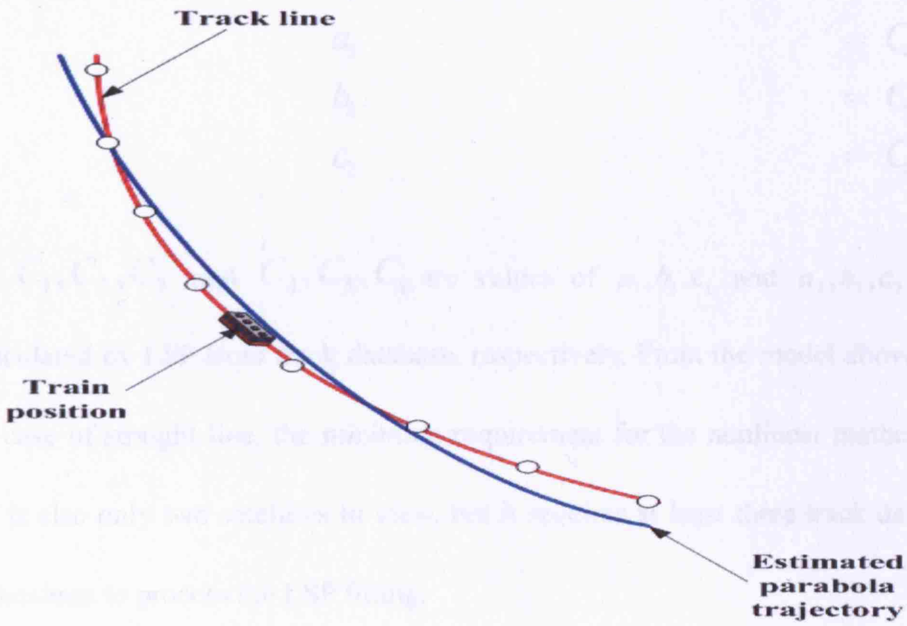


Figure 5.10 The track line and the estimated parabola trajectory

After finding the trajectory function of the track curve from the track database, the train position is now assumed to follow the estimated trajectory function. According to Equation 5.6, the constraint for the train position can be defined as

$$\begin{cases} y_u = a_1 \cdot x_u^2 + b_1 \cdot x_u + c_1 \\ z_u = a_2 \cdot x_u^2 + b_2 \cdot x_u + c_2 \end{cases} \tag{5.18}$$

where a_1, b_1, c_1 and a_2, b_2, c_2 are coefficients of estimated trajectory. Therefore, the nonlinear mathematical model for the GNSS/Track Database integration system can be described as follows

$$\begin{cases} [(x_u - x_{s_1})^2 + (a_1 \cdot x_u^2 + b_1 \cdot x_u + c_1 - y_{s_1})^2 + (a_2 \cdot x_u^2 + b_2 \cdot x_u + c_2 - z_{s_1})^2]^{1/2} + c \cdot t_u = \rho_1 \\ [(x_u - x_{s_2})^2 + (a_1 \cdot x_u^2 + b_1 \cdot x_u + c_1 - y_{s_2})^2 + (a_2 \cdot x_u^2 + b_2 \cdot x_u + c_2 - z_{s_2})^2]^{1/2} + c \cdot t_u = \rho_2 \\ \vdots \qquad \vdots \quad \vdots \quad \vdots \\ a_1 = C_1 \\ b_1 = C_2 \\ c_1 = C_3 \\ a_2 = C_4 \\ b_2 = C_5 \\ c_2 = C_6 \end{cases} \quad (5.19)$$

where C_1, C_2, C_3 and C_4, C_5, C_6 are values of a_1, b_1, c_1 and a_2, b_2, c_2 which are calculated by LSP from track database, respectively. From the model above, same as the case of straight line, the minimum requirement for the nonlinear mathematical model is also only two satellites in view, but it requires at least three track data from track database to process the LSP fitting.

LS method is also used to solve the nonlinear mathematical model. The process is similar to the above linear mathematical model. The matrix form of the model is

$$f(X) = l \qquad (5.20),$$

and the result is

$$\Delta X = (A^T W A)^{-1} A^T W B \qquad (5.21).$$

where

$$A=\begin{bmatrix} \frac{\partial f_1}{\partial x_u} & \frac{\partial f_1}{\partial t_u} & \frac{\partial f_1}{\partial a_1} & \frac{\partial f_1}{\partial b_1} & \frac{\partial f_1}{\partial c_1} & \frac{\partial f_1}{\partial a_2} & \frac{\partial f_1}{\partial b_2} & \frac{\partial f_1}{\partial c_2} \\ \frac{\partial f_2}{\partial x_u} & \frac{\partial f_2}{\partial t_u} & \frac{\partial f_2}{\partial a_1} & \frac{\partial f_2}{\partial b_1} & \frac{\partial f_2}{\partial c_1} & \frac{\partial f_2}{\partial a_2} & \frac{\partial f_2}{\partial b_2} & \frac{\partial f_2}{\partial c_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$B=\begin{bmatrix} \rho_1-\hat{\rho}_1 \\ \rho_2-\hat{\rho}_2 \\ \vdots \\ C_1-\hat{C}_1 \\ C_2-\hat{C}_2 \\ C_3-\hat{C}_3 \\ C_4-\hat{C}_4 \\ C_5-\hat{C}_5 \\ C_6-\hat{C}_6 \end{bmatrix}, \quad \Delta X = \begin{bmatrix} \Delta x_u \\ c\Delta t_u \\ \Delta a_1 \\ \Delta b_1 \\ \Delta c_1 \\ \Delta a_2 \\ \Delta b_2 \\ \Delta c_2 \end{bmatrix}$$

$$\frac{\partial f_i}{\partial x_u}=\frac{x_u-x_{s_i}+(a_1\cdot x_u^2+b_1\cdot x_u+c_1-y_{s_i})\cdot(2a_1x_u+b_1)+(a_2\cdot x_u^2+b_2\cdot x_u+c_2-z_{s_i})\cdot(2a_2x_u+b_2)}{[(x_u-x_{s_i})^2+(a_1\cdot x_u^2+b_1\cdot x_u+c_1-y_{s_i})^2+(a_2\cdot x_u^2+b_2\cdot x_u+c_2-z_{s_i})^2]^{1/2}}$$

$$\frac{\partial f_i}{\partial t_u}=1$$

$$\frac{\partial f_i}{\partial a_1}=0$$

$$\frac{\partial f_i}{\partial b_1}=0$$

$$\frac{\partial f_i}{\partial c_1}=0$$

$$\frac{\partial f_i}{\partial a_2} = 0$$

$$\frac{\partial f_i}{\partial b_2} = 0$$

$$\frac{\partial f_i}{\partial c_2} = 0$$

The covariance of observations is changed to

$$Cov(l) = \begin{bmatrix} \sigma_{\rho_1}^2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \sigma_{\rho_2}^2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{\rho_n}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{X_1} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{X_2} \end{bmatrix},$$

where

$$C_{X_1} = \begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_1,b_1} & \sigma_{a_1,c_1} \\ \sigma_{a_1,b_1} & \sigma_{b_1}^2 & \sigma_{b_1,c_1} \\ \sigma_{a_1,c_1} & \sigma_{b_1,c_1} & \sigma_{c_1}^2 \end{bmatrix} = (A_1^T A_1)^{-1} \ , C_{X_2} = \begin{bmatrix} \sigma_{a_2}^2 & \sigma_{a_2,b_2} & \sigma_{a_2,c_2} \\ \sigma_{a_2,b_2} & \sigma_{b_2}^2 & \sigma_{b_2,c_2} \\ \sigma_{a_2,c_2} & \sigma_{b_2,c_2} & \sigma_{c_2}^2 \end{bmatrix} = (A_2^T A_2)^{-1}.$$

The matrix A_1 and A_2 come from Equation 5.8. The weight matrix is still $W = Cov(l)^{-1}$. All coefficients a_1, b_1, c_1, a_2, b_2, c_2 are calculated by the LSP method from track database. According to Equation 5.9, all $\frac{\partial f_i}{\partial a_1}, \frac{\partial f_i}{\partial b_1}, \frac{\partial f_i}{\partial c_1}, \frac{\partial f_i}{\partial a_2}, \frac{\partial f_i}{\partial b_2}, \frac{\partial f_i}{\partial c_2}$ equal to zero.

5.2.3 Strategy of Processing the GNSS/Track Data Integration System

In above linear and nonlinear mathematical models, the key point of the integration system is that we use the track data to set a constraint for trains. So how to choose the track data becomes very important. From Figure 5.9, whenever the train runs into any part of rail tracks, it should between two consecutive track points related to the train position. If the track database is collected by the GPS surveying, the average distance between track data points can be very close, about 1.5 meters (Euler et al., 1996). According to the smoothness of rail tracks, the trajectory between any two consecutive track points within 1.5 meters interval can be seen as a straight line. Therefore, if the prior and post track points are known; only linear mathematical model is needed to find the train position. However, because the recent epoch train position is not known and the speed of train varies, it is hard to find the prior and post track points directly. So a strategy is needed to help the system to search the prior and post track points and to solve the train position automatically. The idea of the strategy is to find the final two consecutive track points which the train travels between them, then use the linear mathematical model to find the final position of the train. Figure 5.11 gives the example of strategy processing. The detail of the strategy is given as follows.

Step 1: Searching track points from the track database which guarantee the train travels between them. To do this, we call $a = \frac{v \cdot t}{d}$, where v is the velocity of the train at the previous epoch, t is the time interval between two consecutive epochs, and d is the average distant between two consecutive track points stored in the track database. If $a \leq 1$, go to **Step 6,** otherwise, go to **Step 2**.

Step 2: Define an interval which is from $x_{previous}$ to $x_{previous} + 2 \cdot a \cdot d$, where the $x_{previous}$ is the x coordinate of the previous user's position. Then go to **step 3**.

Step 3: Search the track database. If only two track points are found (i.e. $n_t = 2$), where n_t is the number of track point found, go to **Step 6**; otherwise, choose all track points in this interval, and continue to **Step 4.**

Step 4: Use track points to generate the trajectory function and use the nonlinear mathematical model to find the train position. Then go to **Step 5**.

Step 5: Reduce the interval. Set $a^1 = \frac{a}{2}$, if $a^1 \leq 1$, go to **Step 6**; otherwise, define the new interval by using $x_{calculate} \pm a^1 \cdot d$, where $x_{calculate}$ is the solution which is calculated by the nonlinear mathematical model, then go back to **Step 3**.

Step 6: Find the final two consecutive track points from the track database. If it is from Step 1, set the searching interval based on $x_{previous} \pm d$; if it is from Step 3, use two track points which is founded by Step 3; if it is from Step 5, set the searching interval based on $x_{calculate} \pm d$. After finding the prior and post track point, use the linear mathematical model to calculate the final train position.

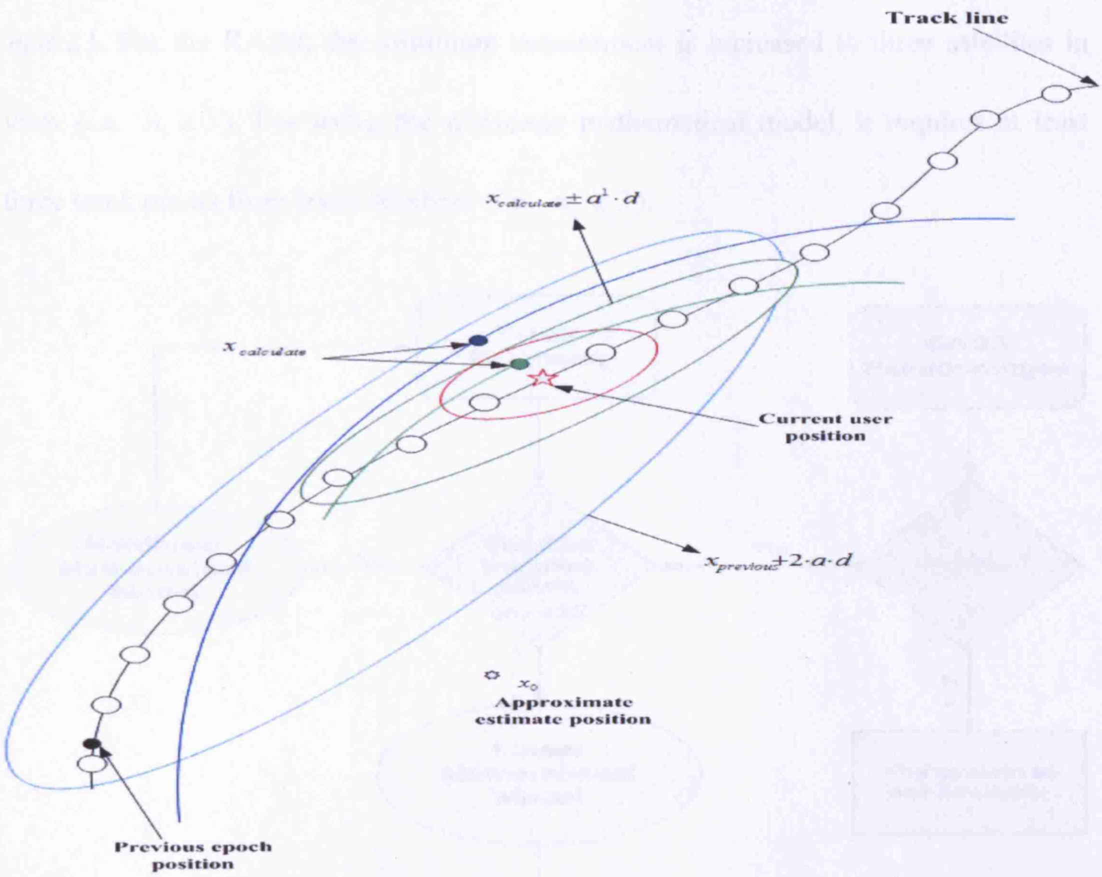


Figure 5.11 The example of strategy processing

With strategies above, the nonlinear mathematical model is only used to help search the final two track points. The reason is when a series consecutive track points are found, though it can be seen as a straight line between each two consecutive track points, it is still hard to guarantee the real trajectory between the whole series track points is the straight line. Therefore, it is better to generate the curve trajectory and use the nonlinear mathematical model. After only two track points left, the final train position is computed by the linear mathematical model. Figure 5.12 illustrates the principle of GNSS/Track Database integration system. For the integration system, the minimum requirement for calculating the train position is two satellites in view (i.e.

$n_1 \geq 2$). For the RAIM, the minimum requirement is increased to three satellites in view (i.e. $n_1 \geq 3$). For using the nonlinear mathematical model, it requires at least three track points from track database (i.e. $n_2 \geq 3$).

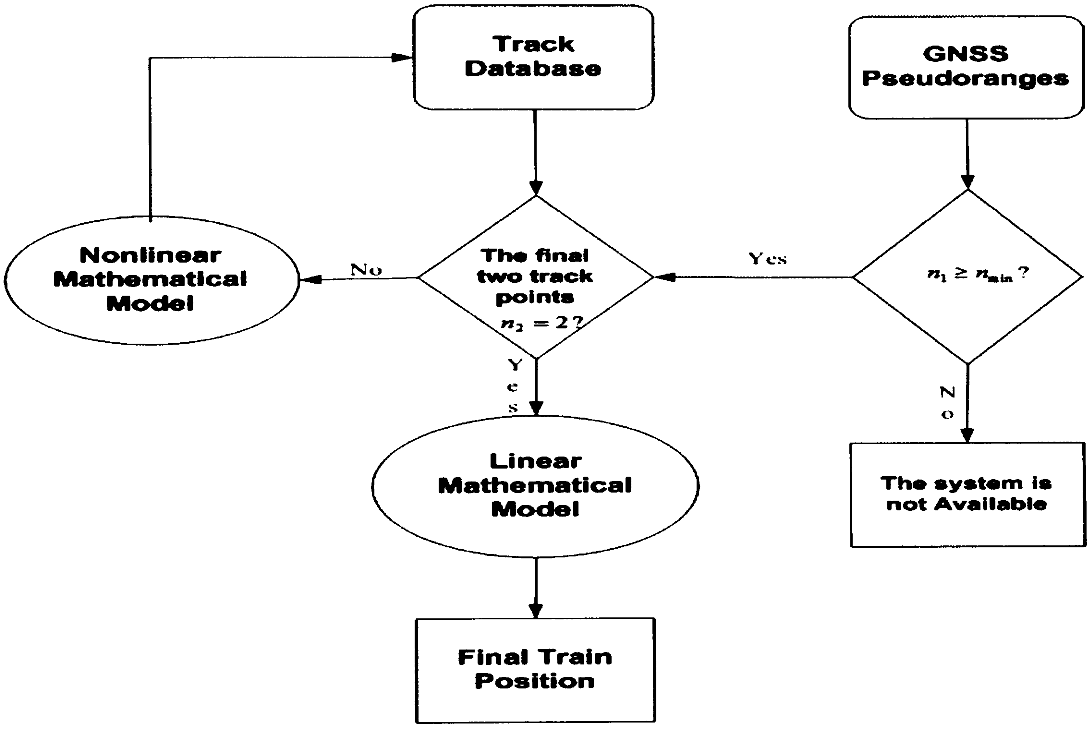


Figure 5.12 Principle of GNSS/Track Database integration system

5.3 Estimated Accuracy of Integration System

From Section 5.2, the final train position is calculated by the linear mathematical model. Therefore, according to Equation 5.17 and Equation 4.6, the covariance matrix of the parameters of the linear mathematical model can be computed by

$$Cov_X^{(x_u, t_u, x_1, y_1, z_1, x_2, y_2, z_2)} = (A^T W A)^{-1} \quad (5.22).$$

where

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_u} & \frac{\partial f_1}{\partial t_u} & \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial z_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial y_2} & \frac{\partial f_1}{\partial z_2} \\ \frac{\partial f_2}{\partial x_u} & \frac{\partial f_2}{\partial t_u} & \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial z_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial y_2} & \frac{\partial f_2}{\partial z_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$W = C_l^{-1} = \begin{bmatrix} \sigma_{\rho_1}^2 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\rho_2}^2 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{x_1}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{y_1}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{z_1}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_2}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{y_2}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{z_2}^2 \end{bmatrix}^{-1},$$

In the weight matrix, $\sigma_{\rho_i} = \sigma_{UERE_i}$, and σ_{C_i} are the standard deviations of track points in the track data base.

According Equation 5.11, the three dimensional solution of the train can be rewritten as

$$\begin{cases} x_u = & x_u \\ y_u = & y_1 \quad - \dfrac{y_2 - y_1}{x_2 - x_1} \cdot x_1 \quad + \dfrac{y_2 - y_1}{x_2 - x_1} \cdot x_u \\ z_u = & z_1 \quad - \dfrac{z_2 - z_1}{x_2 - x_1} \cdot x_1 \quad + \dfrac{z_2 - z_1}{x_2 - x_1} \cdot x_u \end{cases} \tag{5.23}$$

Therefore, the covariance matrix of the train position can be obtained by the error propagation

$$Cov_X^{(x_u,y_u,z_u)} = JCov_X^{(x_u,t_u,x_1,y_1,z_1,x_2,y_2,z_2)} J^T \tag{5.24}$$

where

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \dfrac{y_2 - y_1}{x_2 - x_1} & 0 & \dfrac{(y_2 - y_1)\cdot(x_u - x_2)}{(x_2 - x_1)^2} & \dfrac{x_2 - x_u}{x_2 - x_1} & 0 & \dfrac{(y_2 - y_1)\cdot(x_1 - x_u)}{(x_2 - x_1)^2} & \dfrac{x_u - x_1}{x_2 - x_1} & 0 \\ \dfrac{z_2 - z_1}{x_2 - x_1} & 0 & \dfrac{(z_2 - z_1)\cdot(x_u - x_2)}{(x_2 - x_1)^2} & 0 & \dfrac{x_2 - x_u}{x_2 - x_1} & \dfrac{(z_2 - z_1)\cdot(x_1 - x_u)}{(x_2 - x_1)^2} & 0 & \dfrac{x_u - x_1}{x_2 - x_1} \end{pmatrix}$$

The accuracy can also be presented by the along track, across track and height directions. To realise this, two rotations are needed. Firstly, transfer to the East, North and Height directions by the error propagation,

$$Cov_X^{(E,N,h)} = RCov_X^{(x_u,y_u,z_u)} R^T \tag{5.25}$$

where

$$R = \begin{pmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\sin\phi\cdot\cos\lambda & -\sin\phi\cdot\sin\lambda & \cos\phi \\ \cos\phi\cdot\cos\lambda & \cos\phi\cdot\sin\lambda & \sin\phi \end{pmatrix},$$

The λ, ϕ are the longitude and latitude of the original point of the topographic coordinate system, respectively. Then, rotate it to the Along track, Across track and Height directions,

$$Cov_X^{(Along, Across, h)} = R' Cov_X^{(E,N,h)} R'^T \tag{5.26}$$

where

$$R' = \begin{pmatrix} \sin\beta & \cos\beta & 0 \\ -\cos\beta & \sin\beta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

Which β is the Azimuth of the track line cutting.

Therefore, to estimate the accuracy of the integration system, the constellation configuration, the corresponding UERE budgets, the prior and post track points, and the accuracy of the track database are needed. The minimum requirement for estimating the accuracy is at least two satellites in view.

5.4 Estimated Integrity of Integration System

Again, the train position is obtained by the linear mathematical model. The nonlinear mathematical model is only used to help search the track points. Therefore, the estimated integrity is based on the linear mathematical model. Similar to the GNSS alone, the modified RAIM method can be used. The lower limit on MDE for the i^{th} observation is given by

$$\Delta_i^u = d_i^u \tau_i \sigma_i \tag{5.27}$$

where d_i^u is calculated as $d_i^u = a + b$, a and b are from the standard normal distribution table, and

$$\tau_i = \frac{\sigma_i}{\sigma_{vi}} \tag{5.28}$$

where σ_i is the standard deviation of the observation i, and σ_{vi} is the standard deviation of the residual v_i. The σ_{vi} comes from the diagonal of the covariance matrix of the residuals Cov_v, which is given by

$$Cov_v = Cov_l - Cov_{\hat{l}} = \begin{bmatrix} \sigma_{v1}^2 & & & \\ & \sigma_{v2}^2 & & \cdots \\ \cdots & & \ddots & \\ & & & \sigma_{vi}^2 \end{bmatrix} \tag{5.29}$$

where Cov_l is the covariance matrix of the observations, $Cov_{\hat{l}}$ is the covariance matrix of the observed parameters, calculated by

$$Cov_{\hat{l}} = ACov_X^{(x_u, ct_u, x_1, y_1, z_1, x_2, y_2, z_2)} A^T \tag{5.30}$$

The number of the observation i is from the total number of the visible satellites and the track points. $0 < i \le m + n$, where m is the number of the visible satellites and n equals to 6 entailing 3 coordinates of the prior track data point and 3 of the post track data point.

After the minimum detected gross error from each observation in the specified probability (i.e. internal reliability) is obtained, in order to calculate the effect on the solution of the minimum gross error detectable with specified probability in the i^{th} observation (i.e. external reliability), evaluate

$$\begin{aligned}&[\delta x_u, c\delta t_u, \delta x_1, \delta y_1, \delta z_1, \delta x_2, \delta y_2, \delta z_2]' \\ &= \left(A^T W A\right)^{-1} A^T W p_i\end{aligned} \quad (5.31)$$

where δ represents the effects of the minimum detectable gross error on the parameter and $p_i = (0, \cdots, \Delta_i^u, \cdots, 0)'$. Finally, the position shift is computed by

$$\begin{aligned}&[\delta x_u \quad \delta y_u \quad \delta z_u]' \\ &= J \cdot [\delta x_u, c\delta t_u, \delta x_1, \delta y_1, \delta z_1, \delta x_2, \delta y_2, \delta z_2]'\end{aligned} \quad (5.32)$$

where J is the Jacobian matrix from Equation 5.24. Similar to the accuracy, values in the along track, across track and height direction are easier to handle for the train controlling than in the ECEF coordinate system. In order to obtain along-track results, the position vector is initially rotated to the local topographic co-ordinate system by

$$\delta_E = -\delta_X \sin\lambda_p + \delta_Y \cos\lambda_p$$

$$\delta_N = -\delta_X \sin\phi_p \cos\lambda_p - \delta_Y \sin\phi_p \sin\lambda_p + \delta_Z \cos\phi_p \quad (5.33)$$

$$\delta_h = \delta_X \cos\phi_p \cos\lambda_p + \delta_Y \cos\phi_p \sin\lambda_p + \delta_Z \sin\phi_p$$

where ϕ_p is the latitude and λ_p is the longitude of the original of the local topographic coordinate. The external reliability vector is then rotated to align with the track direction to provide along-track results.

$$\delta_{Along} = \delta_E \sin\theta + \delta_N \cos\theta \qquad (5.34);$$

$$\delta_{Across} = -\delta_E \cos\theta + \delta_N \sin\theta \qquad (5.35).$$

where θ is the azimuth of the track line cutting.

However, to get the performance of RAIM, the requirements of Along Track Alarm Limit (ATAL), Cross Track Alarm Limit (CTAL), the probability of missed detection and false alarm, the time-to-alarm need to be known. Additionally, the prior and the post track points need to be known. For the GNSS/Track Database integration system, the RAIM requires at least three satellites in view and at least four satellites to do the FDE.

All the accuracy, the integrity and the availability of RAIM performances for the integration system will be tested in the following chapters. In Chapter 6, the data test is simulated around London. The track database is collected on a real railway line around Birmingham in Chapter 7.

Chapter 6 Data Simulation and Analysis around the London Area

According to the RNP requirements for safety-critical railway applications described in Chapter 4, using stand-alone GNSS for the train location system is not sufficient, especially for the integrity and the availability. This is because that a train often travels between deep cuttings, buildings or forests where only a small part of sky can be seen thus only a few satellite signals can be received. The insufficient visible satellites would cause bad quality positions; sometimes it would even cause no integrity and availability. Therefore, the track database is introduced to compensate GNSS for the high safety train location system in Chapter 5. The data test in the following two chapters (Chapter 6 and Chapter 7) would describe the performances of the GPS alone and the integrated GPS/Track Database around London (Chapter 6) and Birmingham (Chapter 7) areas, respectively. The accuracy, the integrity and the availability performances are compared in four different scenarios.

6.1 Data Test Description

As described in Chapter 4, the RNP requirements are used to judge the suitability of the GPS standalone or the GPS/Track Database integration system for specific applications. Therefore, to estimate the integrity and the availability of these systems, the RNP standards for safety-critical railway applications have to be defined before data simulations. According to Table 4.8, as required safety integrity levels and alarm limits will vary from different safety-critical railway applications, the RNP standard levels have been defined by four levels. Table 6.1 gives all four RNP standard levels for our data testing. The integrity and continuity are specified relative to a containment region which is a concept defined by RNP for Area Navigation (RNAV) in the aviation applications. The limit of the containment region is about 99.999% of positioning accuracy (between 4σ and 5σ). It means that the alarm limit for the integrity is about 2.5 times of the values of the accuracy (95%, 1.96σ). However, there is no specific restriction of this factor on alarm limits for railway applications. Due to the high accuracy requirement (e.g. 4 meters) for safety-critical railway applications, a relatively big alarm limit (e.g. five times of accuracy requirement) will only increase ten more meters for the minimum safety stopping distance. This is acceptable for medium density or low density traffic lines. As for high density traffic lines, if the position accuracy can be obtained by centimeter levels, the relatively big alarm limit might be acceptable depending on specific applications. In my data test, accuracy requirements are defined as 4 meters in the horizontal plane and 10 meters in the vertical plane. The data simulation is based on the standard GPS performance. Although the GPS compensated with RTK network could achieve centimeter level accuracy, it is not investigated in this thesis (beyond the scope and time span of this doctoral project). Therefore, the alarm limits are set to be 2.5 and 5 times of accuracy requirements (i.e. 10 meters in horizontal, 25 meters in vertical; and 20 meters in horizontal, 50 meters in vertical). The integrity risk is designed with two alternative levels (i.e. 3.3×10^{-9} per hour and 4×10^{-12} per hour). These two values stand for the range of values from high safety integrity level (Level 3) to very high safety integrity level (Level 4).

According to the relationship between the integrity and the RAIM in Figure 4.2, the relative RAIM requirements could also be defined. Table 6.2 illustrates the reference

Table 6.1 Referenced test RNP standard levels

<table>
<tr><th rowspan="3">Standard Level</th><th colspan="6">Requirements</th></tr>
<tr><th colspan="3">Integrity</th><th rowspan="2">Accuracy</th><th rowspan="2">Continuity</th><th rowspan="2">Availability</th></tr>
<tr><th>Integrity risk</th><th>Alarm Limits</th><th>Time to Alarm</th></tr>
<tr><td>1</td><td>3.3×10^{-9} /h</td><td>20 m (ATAL)
20 m (CTAL)
50 m (HAL)</td><td>1 s</td><td rowspan="4">4 m (horizontal)
10 m (vertical)</td><td rowspan="4">8.0×10^{-6} /15 sec</td><td rowspan="4">>99.98%</td></tr>
<tr><td>2</td><td>3.3×10^{-9} /h</td><td>10 m (ATAL)
10 m (CTAL)
25 m (HAL)</td><td>1 s</td></tr>
<tr><td>3</td><td>4×10^{-12} /h</td><td>20 m (ATAL)
20 m (CTAL)
50 m (HAL)</td><td>1 s</td></tr>
<tr><td>4</td><td>4×10^{-12} /h</td><td>10 m (ATAL)
10 m (CTAL)
25 m (HAL)</td><td>1 s</td></tr>
</table>

requirements of RAIM for safety-critical railway applications. The false alarm rate α

is fixed at 1×10^{-4} per hour based on the standard GPS false alarm rate. For the GPS satellites, the probability of a blunder occurring that will cause a position error exceeding the alarm limit is failures is about 1×10^{-4}, therefore the probability of missed detection β ranges from 3.3×10^{-5} to 4×10^{-8}. However, the standard levels have been set from the relatively relaxed safety level (Standard Level 1) to more stringently high safety level (Standard Level 4).

Table 6.2 Reference requirements of RAIM for safety-critical railway applications

Standard Level	RAIM Requirements					
	ATAL	CTAL	HAL	Probability of missed detection β	Probability of false alarm α	Mask Angle (^{0})
1	20 m	20 m	50 m	3.3×10^{-5}	1×10^{-4}	10
2	10 m	10 m	25 m			
3	20 m	20 m	50 m	4×10^{-8}		
4	10 m	10 m	25 m			

To estimate the accuracy and the RAIM performance, both the above integrity thresholds, the constellation configurations and the corresponding User Equivalent Range Error (UERE) budgets must be known. For the data simulation, the GPS constellation is extracted from the SP3 file (including satellite position and velocity information) of the International GNSS Service (IGS). A one day SP3 file is used in simulation to show the whole 24-hour performance of the system. The UERE budgets for GPS satellites were predicted by official United States Department of Transport/Department of Defense (DOT/DOD) and GALA (Galileo Overall Architecture Project) estimates (Ochieng et al., 2002), as shown in Table 6.3. During this test, the GALA estimates are used and the appropriate mapping functions have been utilized to model the effects of satellite elevation on the UERE.

Table 6.3. UERE budgets for GPS satellites

Elevation angles (°)	10	20	30	40	50	60	70	80	90
US DOD/DOT Estimates	2.35	2.13	2.07	2.06	2.06	2.05	2.05	2.05	2.05
GALA Estimates	5.23	3.10	2.77	2.45	2.28	2.19	2.25	2.29	2.27

In this chapter, two closed track points around the London region are selected from the map. They are assumed as the two final consecutive track points from the track database. The track line between these two points is treated as the straight line for the data simulation. The train position is in somewhere of the track line. Figure 6.1 shows the track line around London from the map. Figure 6.2 presents the direction of track line.

Figure 6.1 The test track line around London area

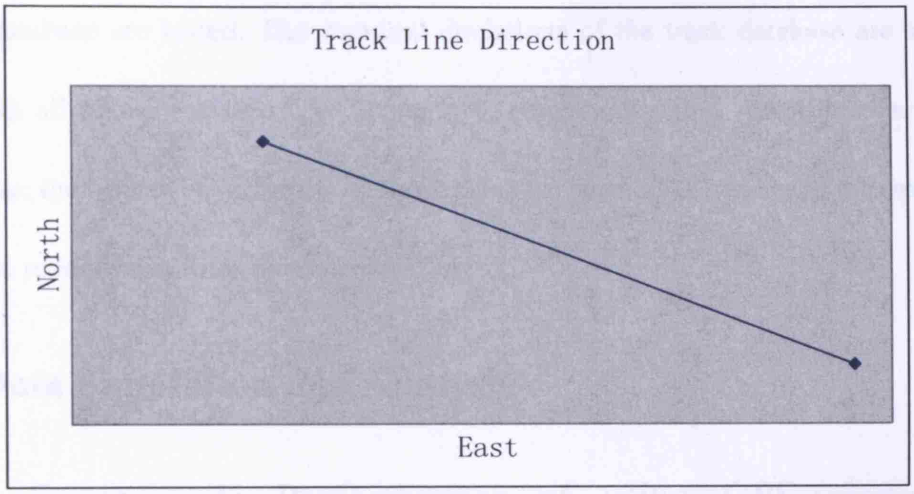


Figure 6.2 Direction of track line

Trains run through complex environments including open areas with good satellite visibility and forest or urban areas with poor satellite visibility. Furthermore, the track database, which is used in the integration system, contains measurement errors. Therefore, the data simulations are divided into four scenarios. In Scenario 1, the different performances between GPS alone and GPS/Track Database in the open area are discussed. In Scenario 2, one satellite is reduced randomly in view to make up slightly tough environment to check the accuracy, integrity and availability performance of using GPS only, or using GPS + Track Database. In Scenario 3, two satellites in view are randomly reduced to make up a low satellite visibility environment. Since the track database is used in the integration system, the accuracy of track database would have influences on the performance of the integration system. The more accurate track database is required, the more expense is cost on the collection of track database. The question is that which required accuracy levels of track database is cost-efficient. To answer this question, different accuracies of the track database are tested. The standard deviations of the track database are set to 1 meter in all above scenarios. In Scenario 4, compared with 1 meter accuracy track database, the improved accuracy of track database is set as 0.1 meter in the open area and two reduced satellites area, respectively.

6.2 Data Simulation and Analysis

6.2.1 Scenario 1: Performances of using GPS alone and GPS/Track Database in the open area of London region

In Figure 6.1, two particular track points in London region are selected as the track database measurements for the integration system. The train position is assumed in somewhere of this track line. The track line between two points is seemed as a straight line. The test date was on 21st May 2005 for the SP3 file. Figure 6.3 presents the satellite visibility of this particular track line in one day period. The mask angle is 10 degrees.

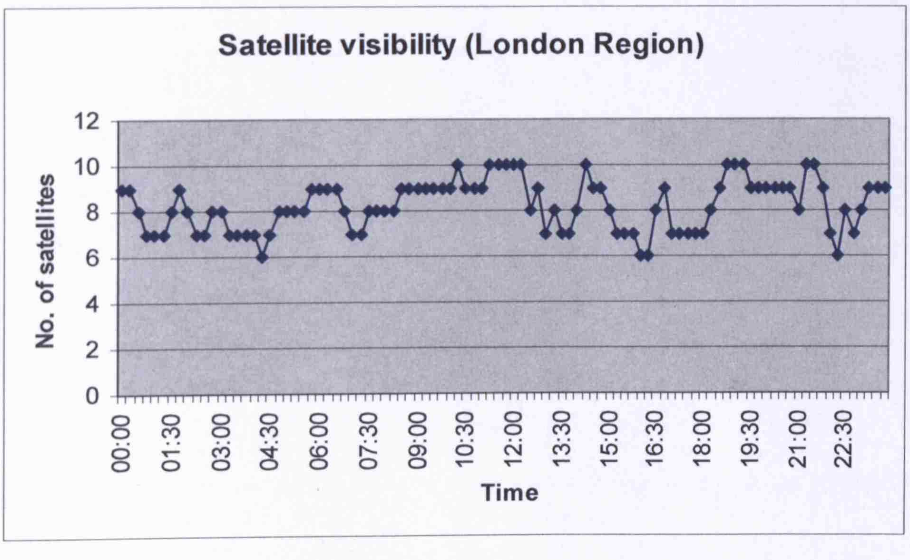


Figure 6.3 Visibility of satellites in the London open areas

In Figure 6.3, GPS has good satellite visibility in this open area. On 21st May 2005, six to ten satellites can always be seen during the whole day. To be specific, the total time of six satellites in view is about one hour (4.1%). The total time of seven, eight, nine and ten satellites in view are about 6.25 hours (26.2%), 5.75 hours (23.9%), 8 hours (33.3%) and 3 hours (12.5%), respectively. Therefore, in the open area, GPS has good satellite visibility. There are sufficient satellites for GPS alone or GPS/Track Database systems to obtain the RNP performances.

For GPS alone, the accuracy performance is shown in Figure 6.4. It has been illustrated separately in all three directions (i.e. along track, across track and height direction). It reveals that the GPS has good accuracy performances in open areas. In particular, in the along track direction, the accuracy of GPS is up to 2.99 meters. At most epochs, it is around 1.0 meters. In the across track direction, the accuracy is slightly more precise than the along track direction. All accuracies are less than 1.8 meters during the whole day.

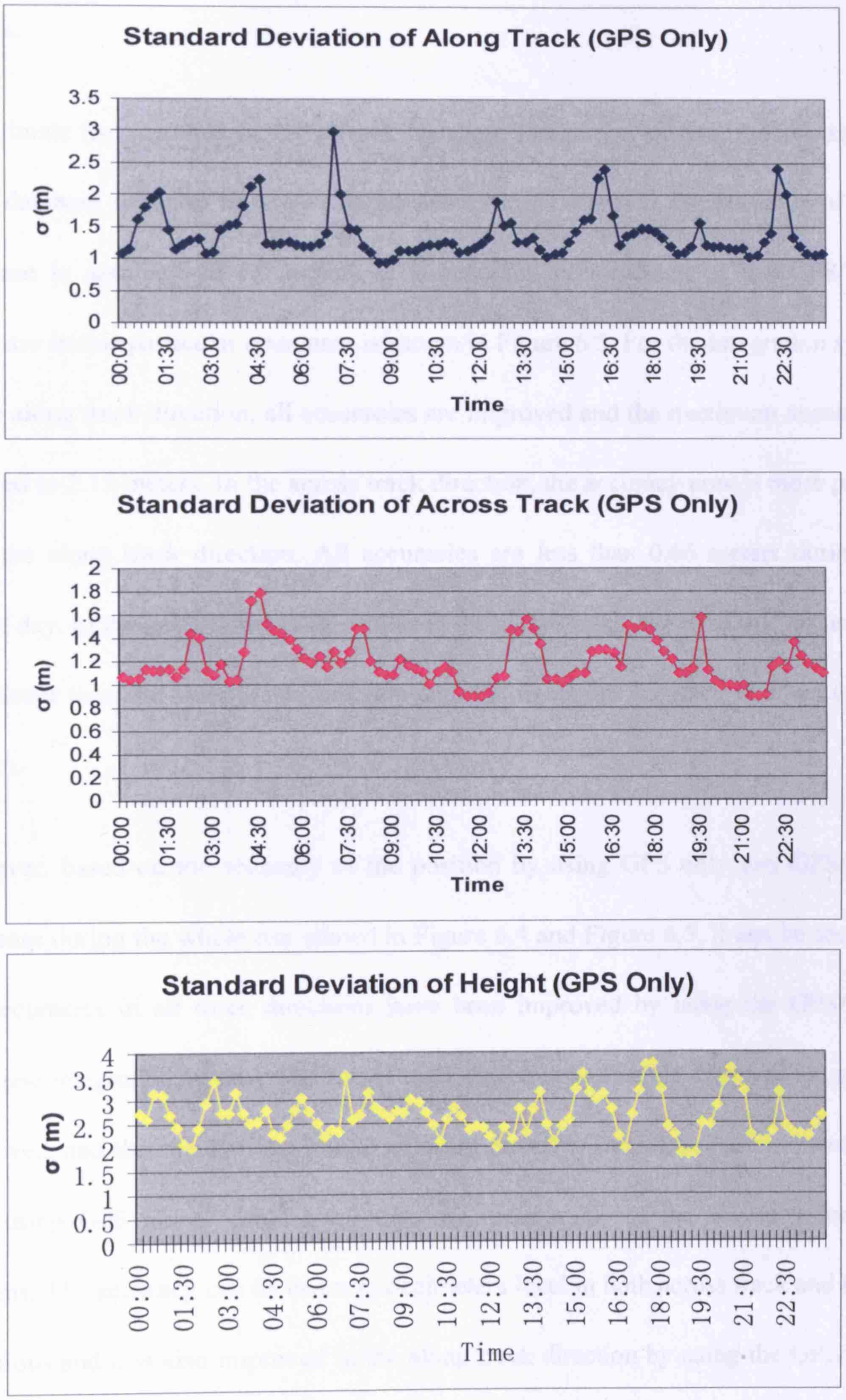


Figure 6.4 The simulated accuracy performance of the GPS only in open areas

In the height direction, the accuracy is worse than the other two directions, up to 3.8

meters.

To estimate the accuracy of GPS/Track Database integration system, the accuracy of track database needs to be known in advance. In this scenario, the accuracy of track database is assumed as 1.0 meters. The accuracy performance of the GPS/Track Database in this particular open area is shown in Figure 6.5. For the integration system, in the along track direction, all accuracies are improved and the maximum accuracy is reduced to 2.15 meters. In the across track direction, the accuracy now is more precise than the along track direction. All accuracies are less than 0.66 meters during the whole day. In the height direction, similar to the across track direction, the accuracy is also better than the along track direction and the maximum accuracy is down to 0.70 meters.

However, based on the accuracy of the position by using GPS only and GPS/Track Database during the whole day shown in Figure 6.4 and Figure 6.5, it can be seen that the accuracies in all three directions have been improved by using the GPS/Track Database integration system. The across track direction and height direction are greatly improved, and the standard deviations of using the GPS/Track Database are smoother than using GPS alone. Table 6.4 shows the mean value of the accuracy for both systems. The accuracy can be down to decimeters level in both across track and height directions and it is also improved in the along track direction by using the GPS/Track database integration system.

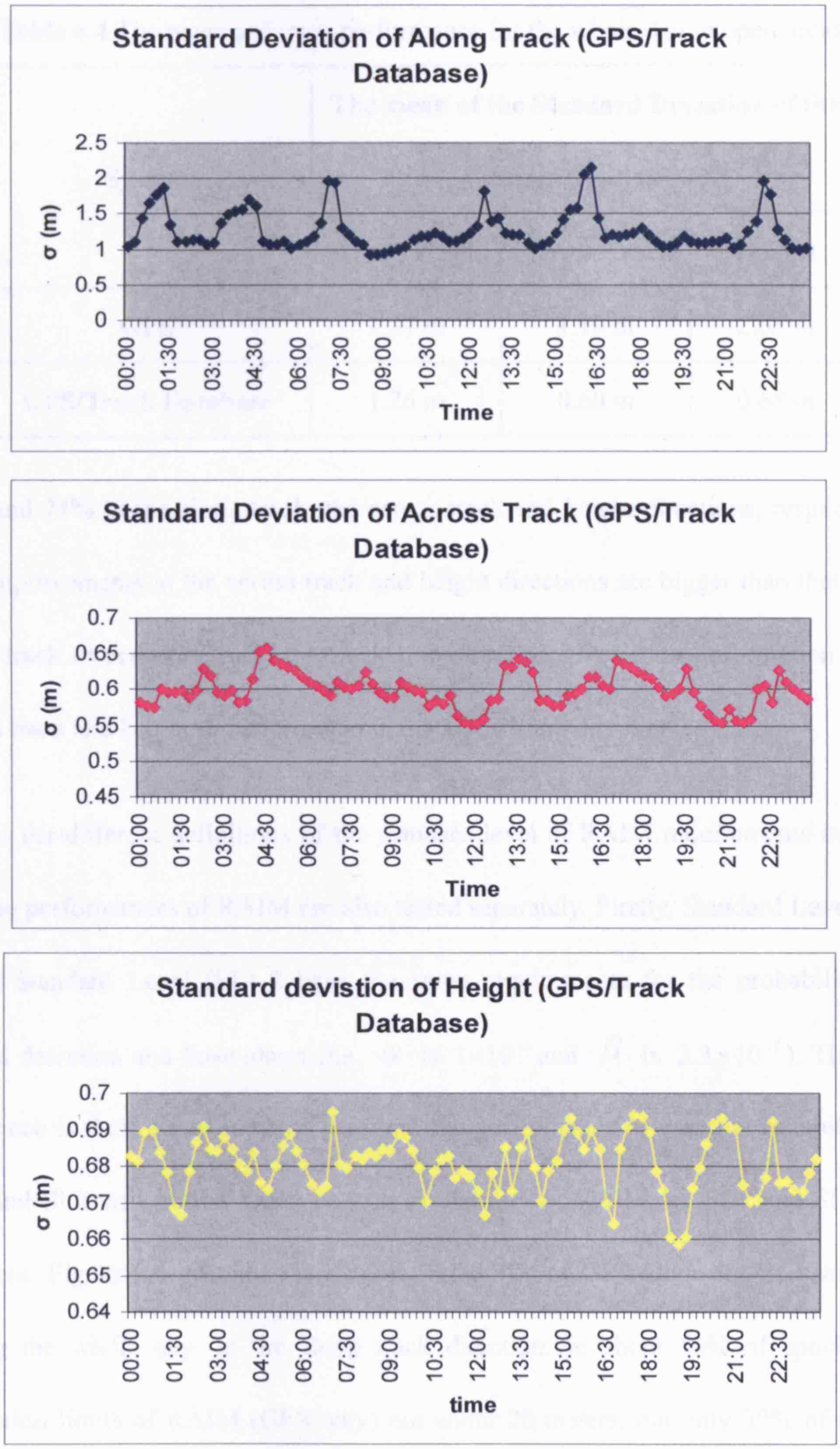


Figure 6.5 The simulated accuracy performance of the GPS/Track Database in open areas

The mean of the standard deviation of the user position has been enhanced about 9%,

Table 6.4 The mean accuracy performance for the whole day in open areas

System	The mean of the Standard Deviation of the user position		
	Along Track	Across Track	Height
GPS	1.37 m	1.19 m	2.65 m
GPS/Track Database	1.26 m	0.60 m	0.68 m

50% and 74% in the along track, the across track and height directions, respectively. The improvements in the across track and height directions are bigger than that in the along track direction because the track line direction gives more information in the across track and height directions than in the along track direction.

Due to the different definitions of the standard level of RAIM requirements in Table 6.2, the performances of RAIM are also tested separately. Firstly, Standard Level (SL) 1 and Standard Level (SL) 2 have the same requirements for the probabilities of missed detection and false alarm (i.e. α is 1×10^{-4} and β is 3.3×10^{-5}). The only difference is their alarm limits. Therefore, the performance of external reliability for SL1 and SL2 will be the same, but the availability of RAIM for SL1 and SL2 will different. Figure 6.6 presents the external reliability of GPS alone for SL1 and SL2 during the whole day. In the along track direction, in about 86% of epochs, the protection limits of RAIM (GPS only) are under 20 meters; but only 39% of epochs are less than 10 meters. The maximum position shift caused by MDEs is up to 48.01 meters during the whole day in this open area.

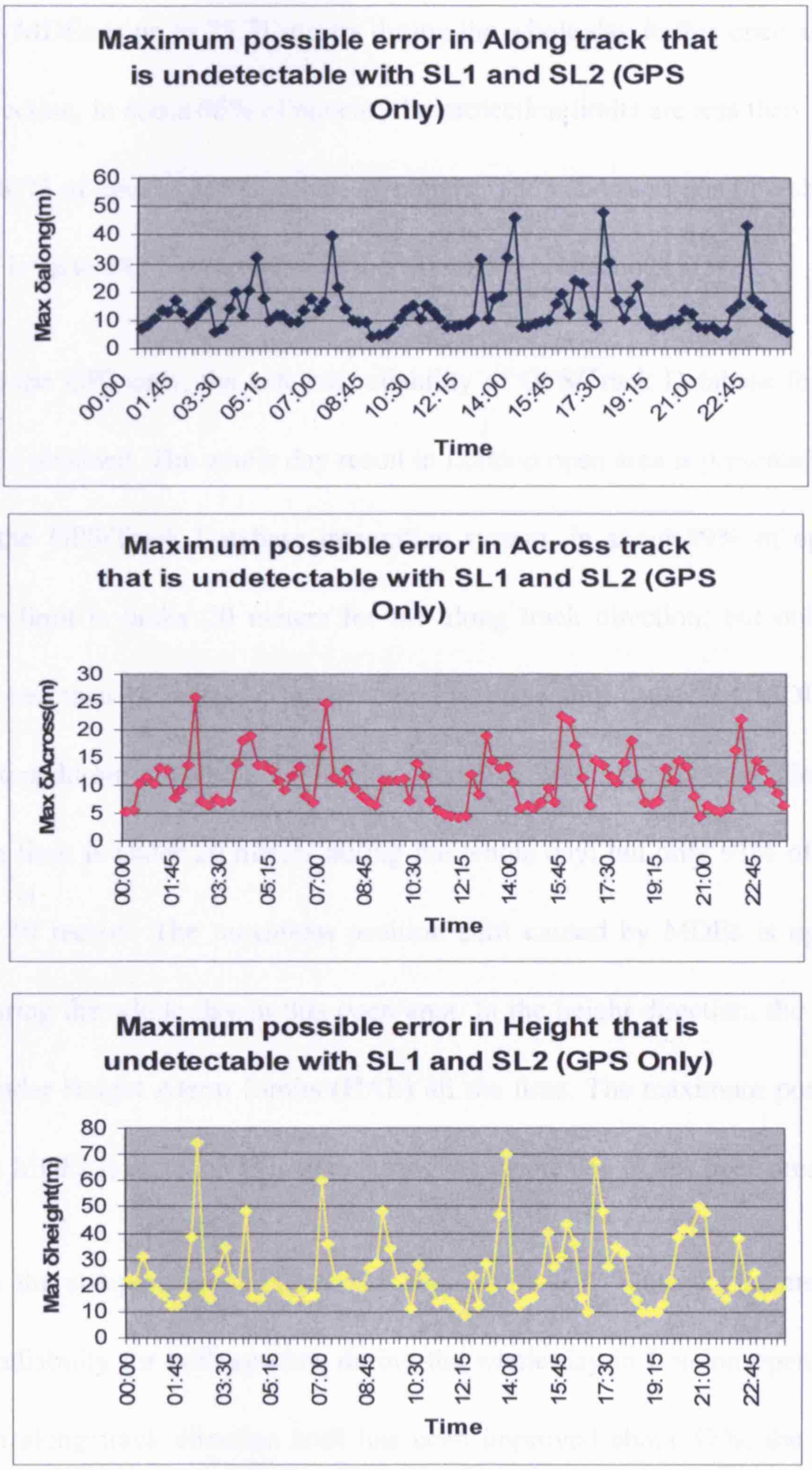


Figure 6.6 The external reliability of GPS only for SL1 and SL2 in London open area

In the across track direction, in about 95% of epochs, the protection limits are under 20 meters; but only 45% of epochs are less than 10 meters. The maximum position shift caused by MDEs is up to 25.71 meters during the whole day in this open area. In the height direction, in about 95% of epochs, the protection limits are less than 50 meters; but only 67% of epochs are less than 25 meters. The maximum position shift caused by MDEs is up to 74.58 meters during the whole day in this open area.

Similar to the GPS only, the external reliability of GPS/Track Database for SL1 and SL2 can be obtained. The whole day result in London open area is presented in Figure 6.7. For the GPS/Track Database integration system, in about 99% of epochs, the protection limit is under 20 meters for the along track direction; but only 73% of epochs is less than 10 meters. The maximum position shift caused by MDEs is up to 20.45 meters during the whole day in this open area. In the across track direction, the protection limit is under 20 meters during the whole day; but only 91% of epochs is less than 10 meters. The maximum position shift caused by MDEs is up to 12.96 meters during the whole day in this open area. In the height direction, the protection limit is under Height Alarm Limits (HAL) all the time. The maximum position shift caused by MDEs is up to 24.75 meters during the whole day in this open area.

Based on the comparison of Figure 6.6 and Figure 6.7, Table 6.5 summarises the external reliability for both systems during the whole day in London open area. The maximum along track direction shift has been improved about 57%, the maximum across track direction shift has been improved about 50%, and the height direction shift has been improved about 67% by the GPS/Track Database integration system, respectively.

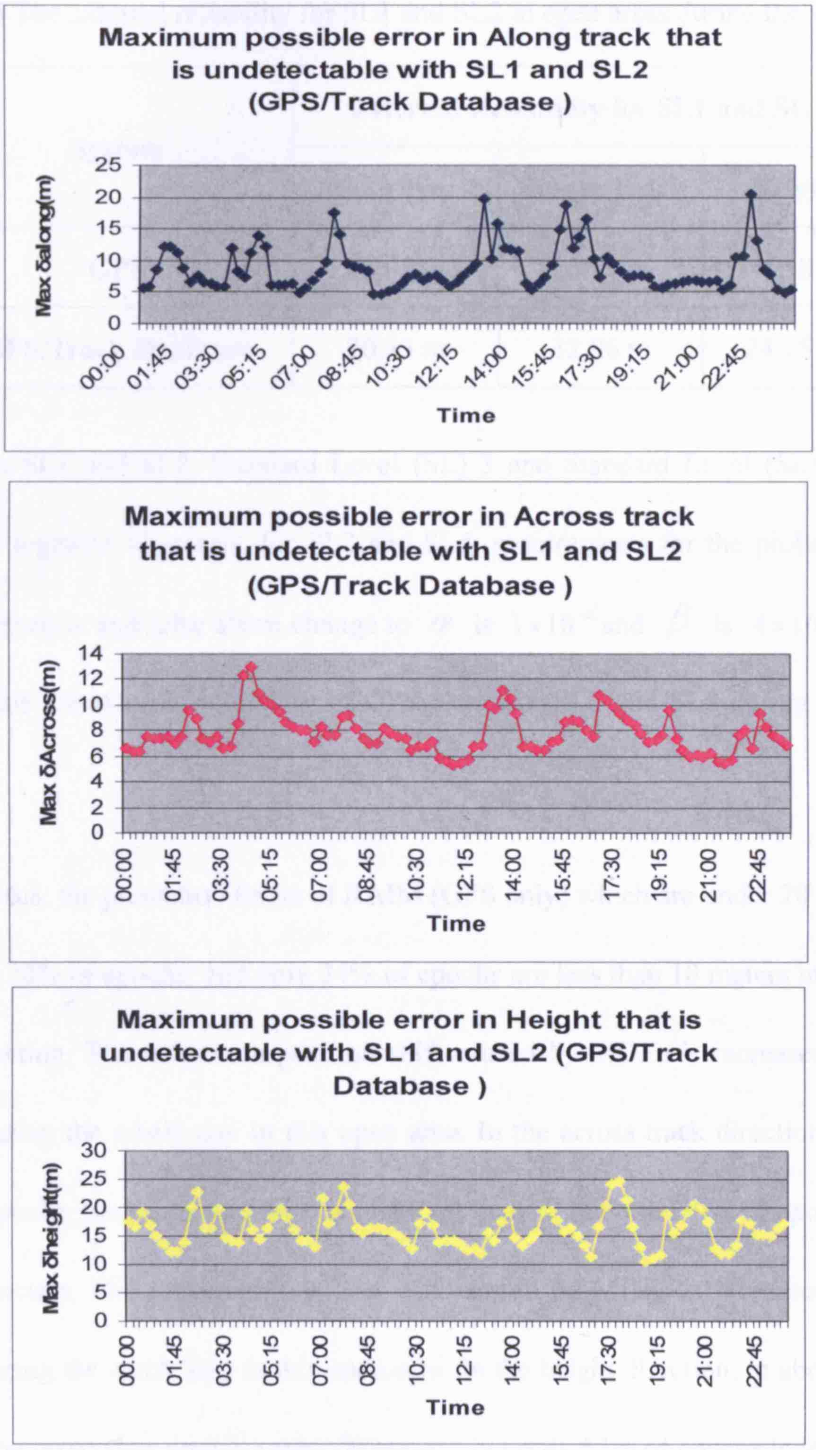


Figure 6.7 The external reliability of GPS/Track Database for SL1 and SL2 in London open area

Table 6.5 The external reliability for SL1 and SL2 in open areas during the whole day

System	External Reliability for SL1 and SL 2		
	Along Track	Across Track	Height
GPS	48.01 m	25.71 m	74.58 m
GPS/Track Database	20.45 m	12.96 m	24.75 m

Similar to SL1 and SL2, Standard Level (SL) 3 and Standard Level (SL) 4 can be simulated together. Currently, for SL3 and SL4, requirements for the probabilities of missed detection and false alarm change to α is 1×10^{-4} and β is 4×10^{-8}. Figure 6.8 presents the external reliability of GPS alone for SL3 and SL4 during the whole day.

In Figure 6.8, the protection limits of RAIM (GPS only) which are under 20 meters are reduce to 78% of epochs; and only 24% of epochs are less than 10 meters in the along track direction. The maximum position shift caused by MDEs is increased to 56.40 meters during the whole day in this open area. In the across track direction, in about 90% of epochs, the protection limit is under 20 meters; but only 39% of epochs is less than 10 meters. The maximum position shift caused by MDEs is increased to 30.20 meters during the whole day in this open area. In the height direction, in about 89% of epochs, the protection limit is under 20 meters; but only 53% of epochs is less than 10 meters. The maximum position shift caused by MDEs is increased to 87.62 meters during the whole day in this open area.

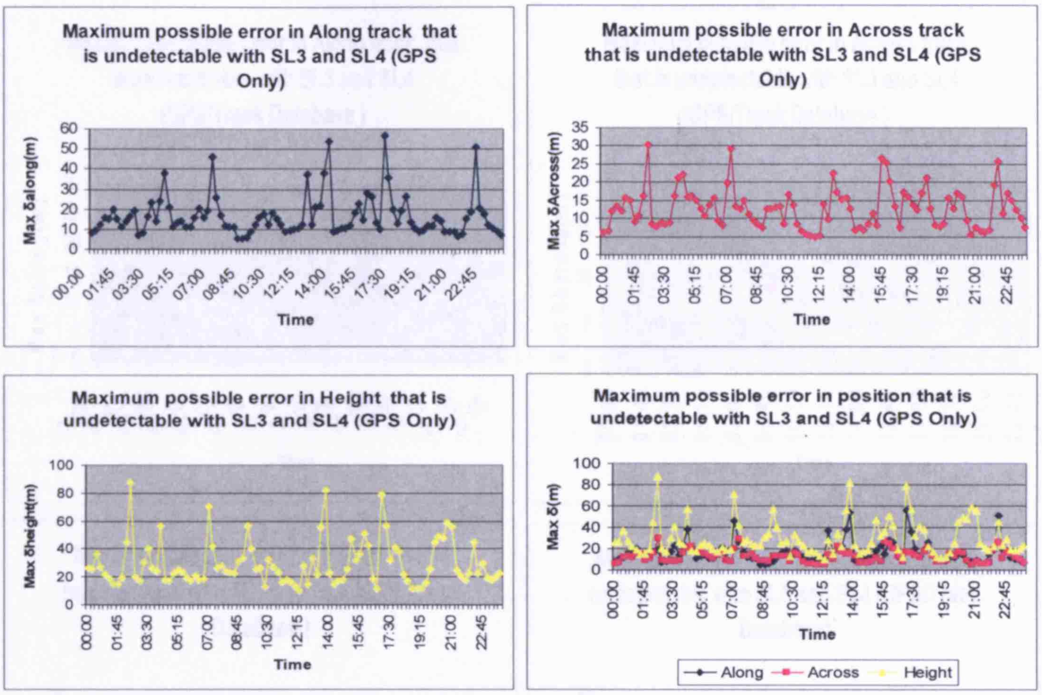


Figure 6.8 The external reliability of GPS only for SL3 and SL4 in London open area

The external reliability of GPS/Track Database for SL3 and SL4 can also be obtained. The whole day result in London open area is presented in Figure 6.9. For the GPS/Track Database integration system, in about 96% of epochs, the protection limits of RAIM are under 20 meters for the along track direction; but only 63% of epochs are less than 10 meters. The maximum position shift caused by MDEs is increased to 24.02 meters during the whole day in this open area. In the across track direction, the protection limits are still under 20 meters during the whole day; but only 72% of epochs are less than 10 meters. The maximum position shift caused by MDEs is increased to 15.23 meters during the whole day in this open area. In the height direction, the protection limits are less than 20 meters all the time, but only 94% of epochs are less than 10 meters now. The maximum position shift caused by MDEs is increased to 29.07 meters during the whole day in this open area.

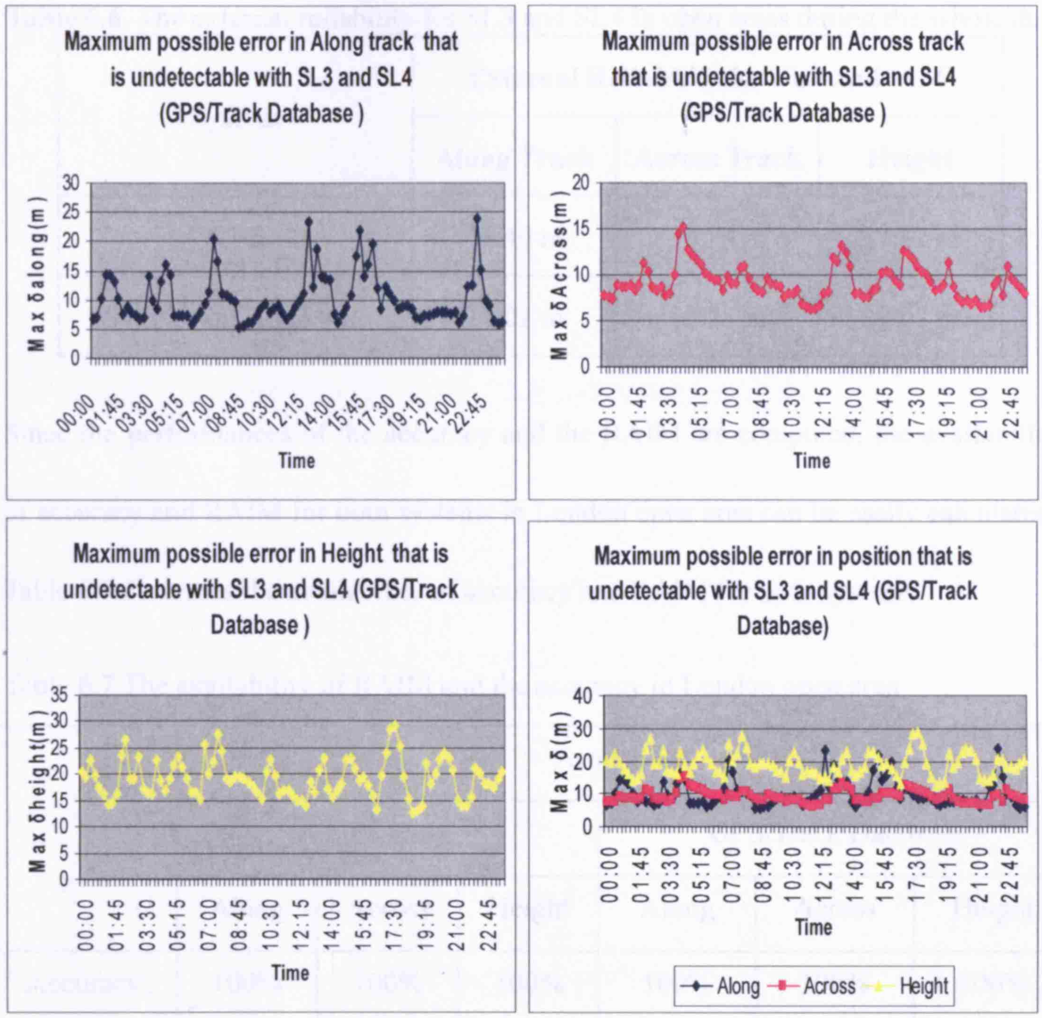


Figure 6.9 The external reliability of GPS/Track Database for SL3 and SL4 in London open area

Table 6.6 summarises the external reliability with SL3 and SL4 for both systems during the whole day in London open area. The maximum position shifts of the GPS only are 56.40 m (Along), 30.20 m (Across) and 87.62 m (height) during the whole journey, respectively. As for the integration system, all three directions are improved to 24.02 m (Along), 15.23 m (Across) and 29.07 m (height), respectively. The maximum along track direction shift has been improved about 57%, the maximum across track direction shift about 50% and the height direction shift about 67%.

Table 6.6. The external reliability for SL3 and SL4 in open areas during the whole day

System	**External Reliability for SL3 and SL 4**		
	Along Track	**Across Track**	**Height**
GPS	56.40 m	30.20 m	87.62 m
GPS/Track Database	24.02 m	15.23 m	29.07 m

Since the performances of the accuracy and the RAIM are computed, the availability of accuracy and RAIM for both systems in London open area can be easily calculated. Table 6.7 describes the availability of accuracy and RAIM for both systems.

Table 6.7 The availability of RAIM and the accuracy in London open area

	Availability					
	GPS Only			GPS/Track Database		
	Along	Across	Height	Along	Across	Height
Accuracy	100%	100%	100%	100%	100%	100%
SL1	86%	95%	95%	99%	100%	100%
SL2	39%	45%	67%	73%	91%	100%
SL3	78%	90%	89%	96%	100%	100%
SL4	24%	39%	53%	63%	72%	94%

However, the GPS standalone has good accuracy performances in open areas, but the availability of RAIM (GPS only) is really poor, especially with SL4. Fortunately, the integration system provides a significant improvement on the availability of RAIM for all standard levels. Especially, for the stringent high safety level (SL4), the availability can be improved about 40% in all three directions. Further, the availability of integrity can also be impacted by both alarm limits and integrity risk. Although the availabilities of accuracy are the same for both systems in open areas (see Table 6.4), the integration system improves the accuracy performance in all three directions.

6.2.2 Scenario 2: Performances of using GPS alone and GPS/Track Database by randomly reducing one satellite in view

Because trains always run through complex environments, it is hard to guarantee that trains are in the open areas all the time. The satellite signals can be obstructed by trees, buildings or deep cuttings so that the satellite visibility could be reduced. Therefore, the performances of system in reduced satellite visibility are also a vital issue in railway safety. In this scenario, all simulations are similar to Scenario 1, but one satellite is randomly reduced in view to test its influence on the system performance. That is to say, the number of satellites in view is reduced by one in each epoch. Figure 6.10 presents the satellite visibility in n-1 satellites environment during the whole day period. The mask angle is still 10 degrees.

From Figure 6.10, satellite visibility is reduced to five to nine satellites in view during the whole day. To be specific, the total time of five satellites in view is about one hour (4.1%); the total time of six, seven, eight and nine satellites in view are about 6.25 hours (26.2%), 5.75 hours (23.9%), 8 hours (33.3%) and 3 hours (12.5%), respectively. There are still enough satellites for the GPS alone system to calculate the position, accuracy and RAIM, but FDE would not be available when five satellites in view. For the integration system, all performance can be obtained during the whole day.

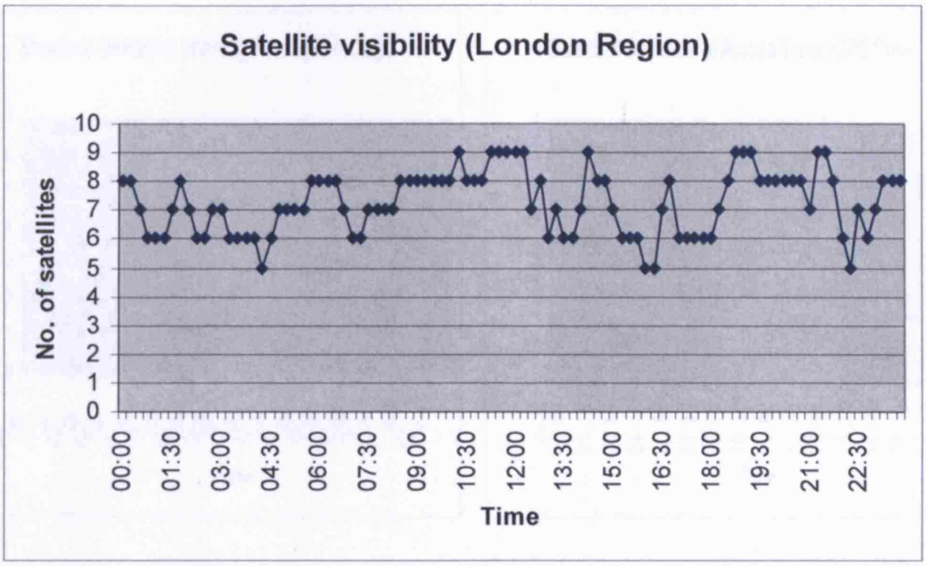


Figure 6.10 The visibility of satellites in n-1 satellites environment

For GPS alone, the accuracy performance in all three directions (i.e. along track, across track and height direction) is shown in Figure 6.11. Specifically, in the along track direction, the maximum accuracy of GPS alone is increased to 5.96 meters. At most epochs, the accuracies are from 1.00 meters to 3.00 meters. In the across track direction, the maximum accuracy jumps to 3.55 meters because the reducing satellite changes the GDOP. However, the accuracy in the across track direction is still slightly more precise than the along track direction. In the height direction, the accuracy is still worse than the other two directions and the maximum accuracy is increased to 10.05 meters.

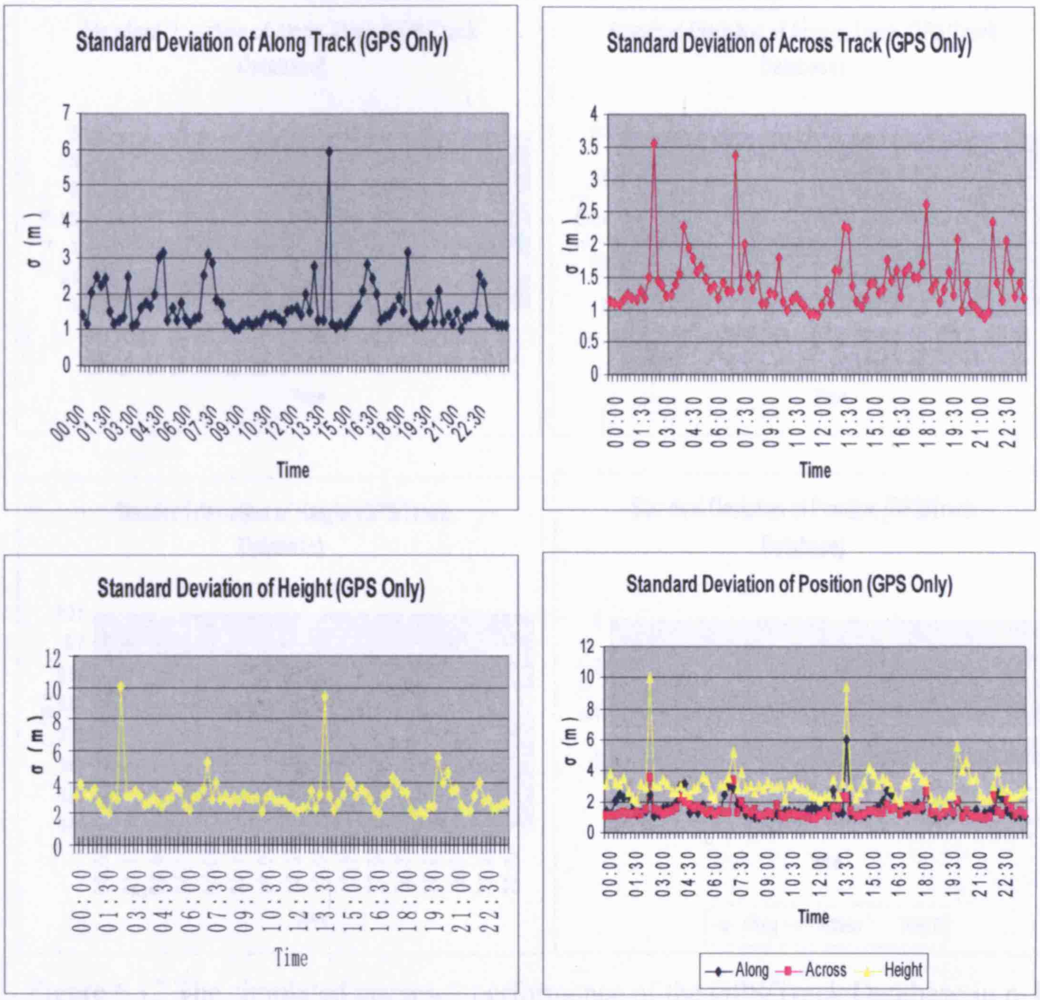


Figure 6.11 The simulated accuracy performance of the GPS only in n-1 satellites environment

To estimate the accuracy of the GPS/Track Database integration system in one reduced satellite environment, the accuracy of track database is still assumed as 1.0 meters in this scenario. The accuracy performance of the GPS/Track Database in this environment is presented in Figure 6.12. For the integration system, in the along track direction, all accuracies are improved and the maximum accuracy is reduced to 2.47 meters. In the across track direction, the accuracy now is more precise than the along track direction. All accuracies are less than 0.67 meters during the whole day. In the height direction, similar to the across track direction, the accuracy is also better than

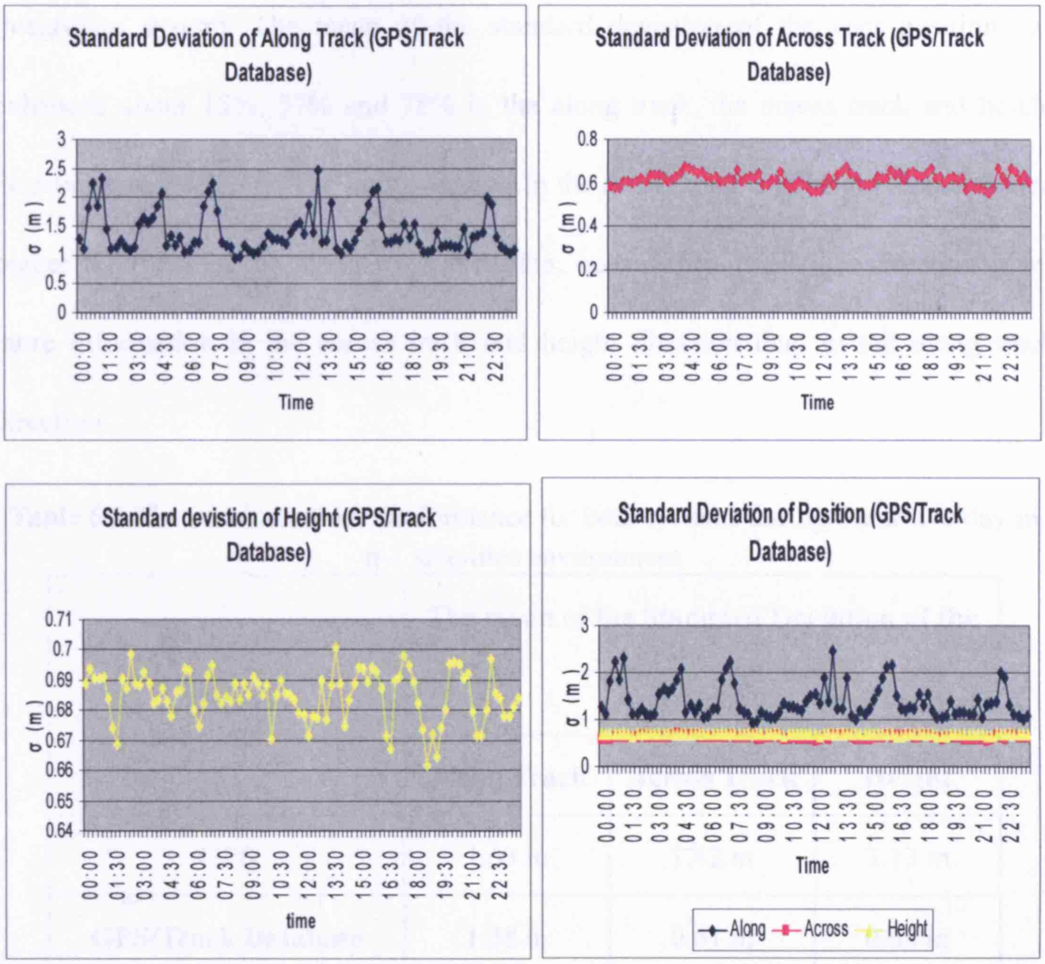


Figure 6.12 The simulated accuracy performance of the GPS/Track Database in n-1 satellites environment

the along track direction, and the maximum accuracy is down to 0.70 meters.

However, compared Figure 6.11 and Figure 6.12, it can be seen that the accuracies in all three directions have been improved by using the GPS/Track Database integration system. The across track direction and height direction are greatly improved, and the standard deviations of using the GPS/Track Database are smoother than using GPS only. Table 6.8 shows the mean value of the accuracy of both systems. The accuracy can still be down to decimeters levels in both across track and height directions and it is also improved in the along track direction by using the GPS/Track database integration system. The mean of the standard deviation of the user position has enhanced about 15%, 57% and 78% in the along track, the across track and height directions, respectively. The improvements in the across track and height directions are bigger than that in the along track direction, because the track line direction gives more information in the across track and height direction than in the along track direction.

Table 6.8 The mean accuracy performance for both systems during the whole day in n-1 satellites environment

System	The mean of the Standard Deviation of the user position		
	Along Track	Across Track	Height
GPS	1.63 m	1.42 m	3.13 m
GPS/Track Database	1.38 m	0.61 m	0.68 m

To estimate the performances of RAIM, similarly to Scenario 1, Standard Level (SL) 1 and Standard Level (SL) 2, which have the same requirements for the probabilities of missed detection and false alarm (i.e. α is 1×10^{-4} and β is 3.3×10^{-5}), can be tested together. Figure 6.13 presents the external reliability of GPS alone for SL1 and SL2 in n-1 satellites environment during the whole day. Obviously, the external reliability performance for the GPS alone system in this environment is worse than the performance in open areas. Specifically, in the along track direction, only about 65% of epochs, the protection limits of RAIM (GPS only) are under 20 meters; but protection limits which are less than 10 meters, are reduced to 20% of epochs.

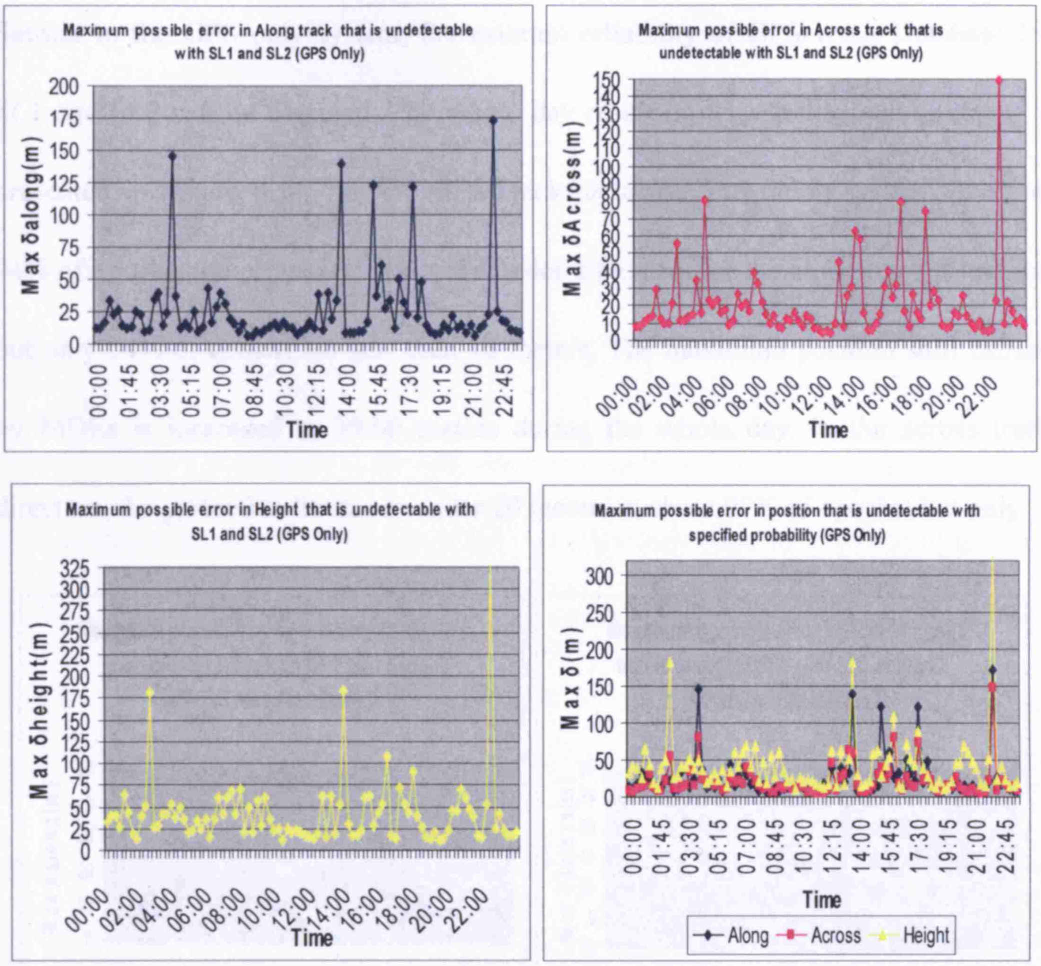


Figure 6.13 The external reliability of GPS only for SL1 and SL2 in n-1 satellites environment

The maximum position shift caused by MDEs is increased to 173.05 meters in n-1 satellites environment during the whole day. In the across track direction, in about 67% of epochs, the protection limits are under 20 meters; but only 31% of epochs are less than 10 meters. The maximum position shift caused by MDEs is increased to 148.87 meters during the whole day. In the height direction, in about 74% of epochs, the protection limits are less than 50 meters; but only 40% of epochs are less than 25 meters. The maximum position shift caused by MDEs is up to 326.03 meters during the whole day.

Similar to the GPS only system, the external reliability of GPS/Track Database for SL1 and SL2 can be obtained. The whole day result in n-1 satellites environment is presented in Figure 6.14. For the GPS/Track Database integration system, in about 94% of epochs, the protection limits are under 20 meters for the along track direction; but only 54% of epochs are less than 10 meters. The maximum position shift caused by MDEs is increased to 39.00 meters during the whole day. In the across track direction, the protection limits are under 20 meters in about 99% of epochs; but only

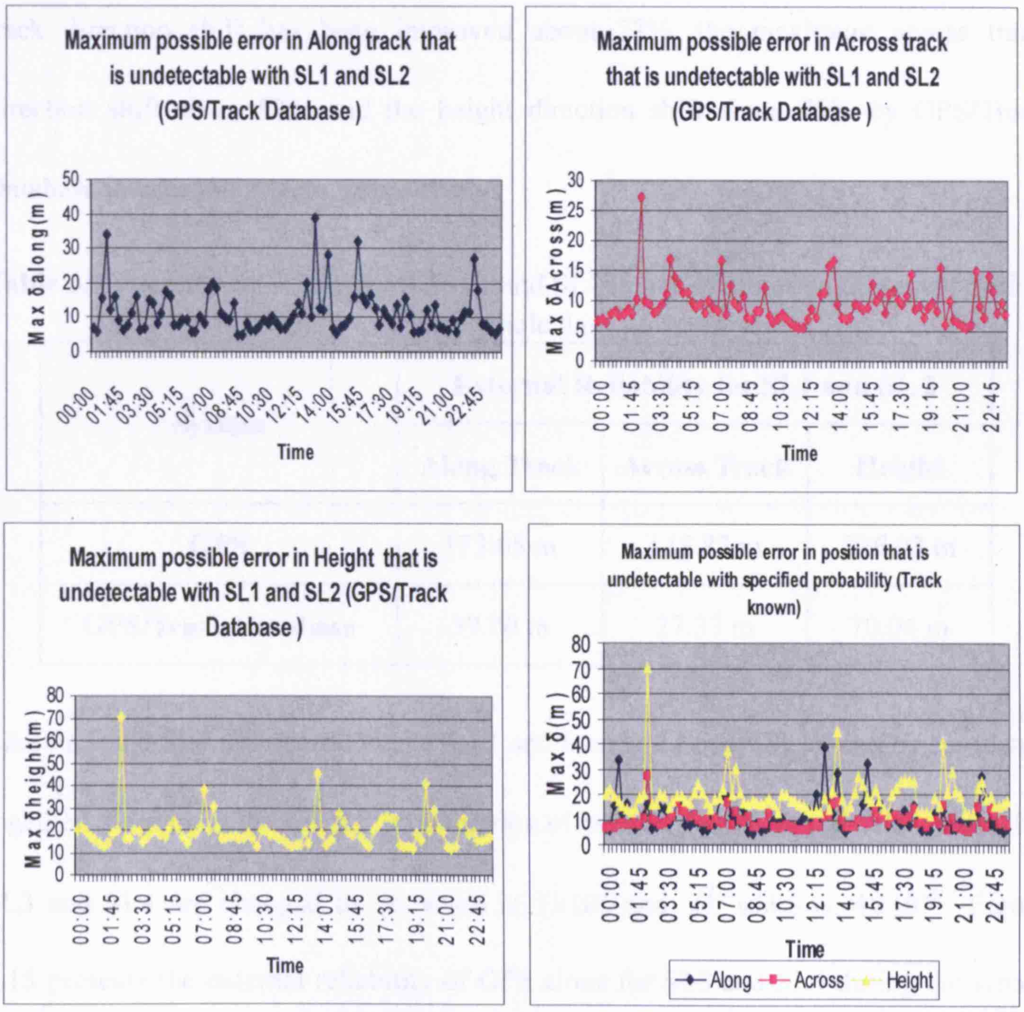


Figure 6.14 The external reliability of GPS/Track Database for SL1 and SL2 in n-1 satellites environment

69% of epochs are less than 10 meters. The maximum position shift caused by MDEs is increased to 27.33 meters during the whole day. In the height direction, the protection limits are less than 50 meters in about 99% of epochs; and the protection limits in 90% of epochs are less than 25 meters. The maximum position shift caused by MDEs is increased to 70.04 meters during the whole day.

Based on Figure 6.13 and Figure 6.14, Table 6.9 summarises the external reliability of both systems during the whole day in n-1 satellites environment. The maximum along track direction shift has been improved about 77%, the maximum across track direction shift about 82% and the height direction shift about 79% by GPS/Track Database integration system, respectively.

Table 6.9 The external reliability for SL1 and SL2 in n-1 satellites environment during the whole day

System	External Reliability for SL1 and SL 2		
	Along Track	**Across Track**	**Height**
GPS	173.05 m	148.87 m	326.03 m
GPS/Track Database	39.00 m	27.33 m	70.04 m

Like SL1 and SL2, Standard Level (SL) 3 and Standard Level (SL) 4 can be simulated together. Requirements for the probabilities of missed detection and false alarm for SL3 and SL4 are changed to α which is 1×10^{-4} and β which is 4×10^{-8}. Figure 6.15 presents the external reliability of GPS alone for SL3 and SL4 during the whole day.

In Figure 6.15, the protection limits of RAIM (GPS only) which are under 20 meters are reduced to 60% of epochs; and only 13% of epochs are less than 10 meters in the along track direction. The maximum position shift caused by MDEs is increased to 203.29 meters during the whole day. In the across track direction, in about 65% of epochs, the protection limits are less than 20 meters; but only 31% of epochs are less than 10 meters. The maximum position shift caused by MDEs is increased to 174.88 meters during the whole day. In the height direction, in about 66% of epochs, the protection limits are less than 50 meters; but only 27% of epochs are less than 25 meters. The maximum position shift caused by MDEs is increased to 382.99 meters during the whole day.

The external reliability of GPS/Track Database for SL3 and SL4 can also be obtained. The whole day result in n-1 satellites environment is presented in Figure 6.16. Now, for the GPS/Track Database integration system, in about 89% of epochs, protection limits of RAIM are under 20 meters in the along track direction; but only 44% of epochs are less than 10 meters. The maximum position shift caused by MDEs is increased to 45.82 meters during the whole day. In the across track direction, protection limits are less than 20 meters in 99% of epochs; but only 45% of epochs are less than 10 meters. The maximum position shift caused by MDEs is increased to 32.10 meters during the whole day. In the height direction, protection limits are less

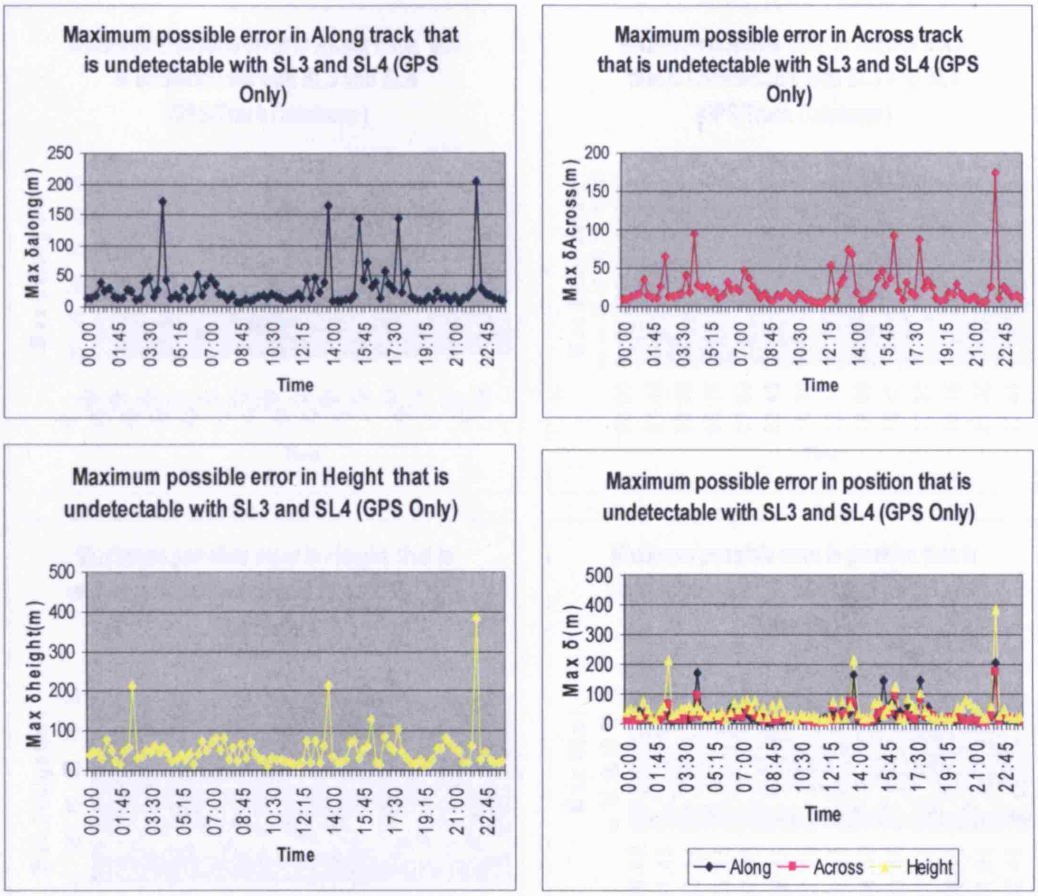


Figure 6.15 The external reliability of GPS only for SL3 and SL4 in n-1 satellites environment

than 50 meters in about 98% of epochs, but only 80% of epochs are less than 25 meters now. The maximum position shift caused by MDEs is increased to 82.28 meters during the whole day in n-1 satellites environment.

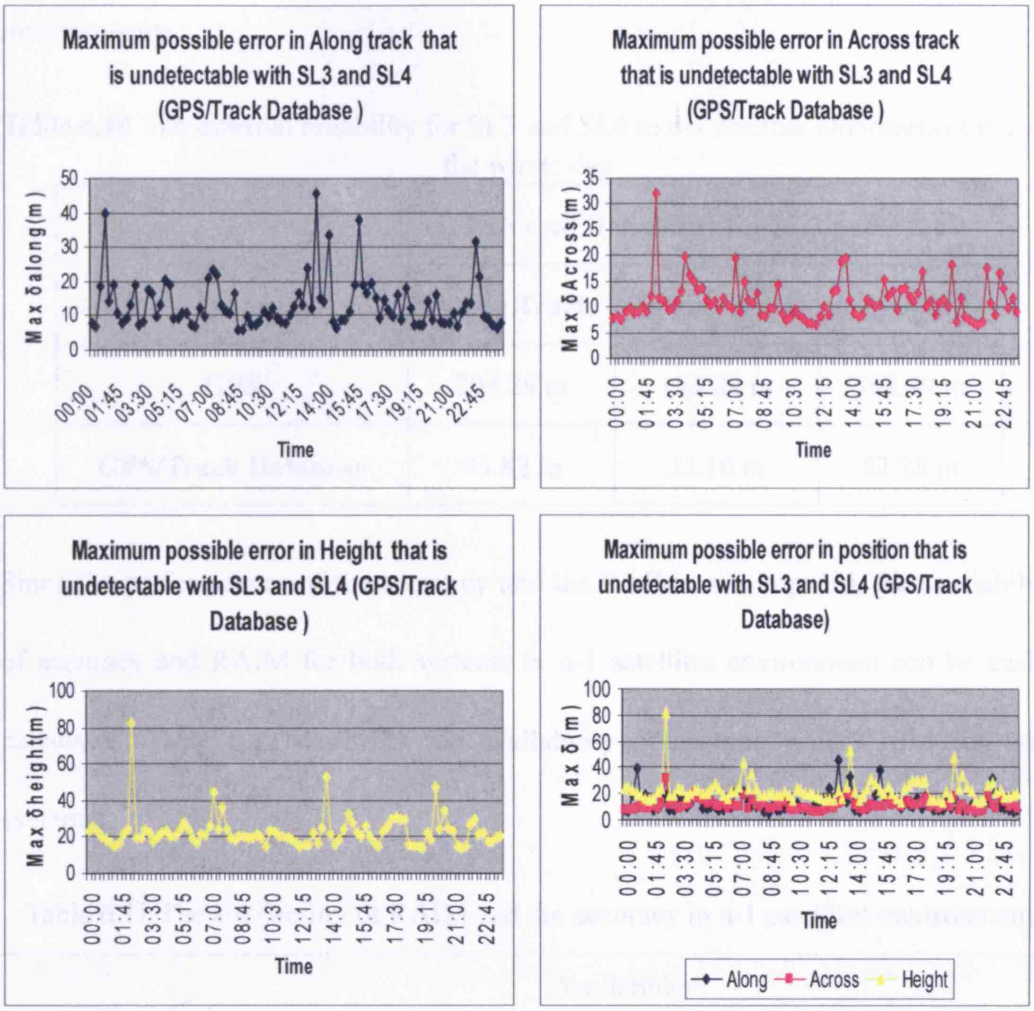


Figure 6.16 The external reliability of GPS/Track Database for SL3 and SL4 in n-1 satellites environment

Table 6.10 summarises the external reliability with SL3 and SL4 for both systems during the whole day in n-1 satellites environments. The maximum position shifts of the GPS only are 203.29 m (Along), 174.88 m (Across) and 382.99 m (height) during the whole journey, respectively. As for the integration system, the three directions are improved to 45.82 m (Along), 32.10 m (Across) and 82.28 m (height), respectively. The maximum along track direction shift has been improved about 77%, the maximum across track direction shift about 82% and the height direction shift about 79%. The reason caused the jump near 1:45am is because of the outlier of track database measurements.

Table 6.10 The external reliability for SL3 and SL4 in n-1 satellite environment during the whole day

System	External Reliability for SL3 and SL 4		
	Along Track	**Across Track**	**Height**
GPS	203.29 m	174.88 m	382.99 m
GPS/Track Database	45.82 m	32.10 m	82.28 m

Since the performances of the accuracy and the RAIM are computed, the availability of accuracy and RAIM for both systems in n-1 satellites environment can be easily calculated. Table 6.11 describes the availability of accuracy and RAIM for both systems.

Table 6.11 The availability of RAIM and the accuracy in n-1 satellites environment

	Availability					
	GPS Only			GPS/Track Database		
	Along	Across	Height	Along	Across	Height
Accuracy	99%	100%	99%	100%	100%	100%
SL1	65%	67%	74%	94%	99%	99%
SL2	20%	31%	40%	54%	69%	90%
SL3	60%	65%	66%	89%	99%	98%
SL4	13%	21%	27%	44%	45%	80%

However, when one satellite in view is randomly reduced, the GPS standalone still has good accuracy performance in most epochs, but it cannot guarantee the accuracy requirements during the whole day. The availability of RAIM (GPS only) is worse than that in open area; especially for SL4, it has only 13% in the along track direction. Fortunately, the integration system still brings about significant improvements on the availability of RAIM and accuracy for all standard levels. It still can guarantee all accuracy requirements and provides decimeter accuracy in the across track and height directions. Additionally, the integration system also has good performances on the external reliability. The protection limits are improved over 100 meters in all three directions. The availability of RAIM for the integration system is also improved about 30% on all standard levels.

6.2.3 Scenario 3: Performances of using GPS alone and GPS/Track Database by randomly reduces two satellites in view

When trains run through environments such as forests or urban canyons, more than one satellite signal is obstructed. To make up a low satellite visibility environment, in this scenario, two satellites in view are randomly reduced for all simulations. Therefore, the number of satellites in view is reduced by two in each epoch. Figure 6.17 presents the satellite visibility in n-2 satellites environment during the whole day period. The mask angle is still 10 degrees. In Figure 6.17, satellite visibility is reduced to four to eight satellites in view during the whole day. For GPS only, RAIM is not available when only four satellites are in view. When only five satellites are in view, FDE is also not available. However, there are still enough satellites for the integration system to calculate the position, accuracy and RAIM during the whole day.

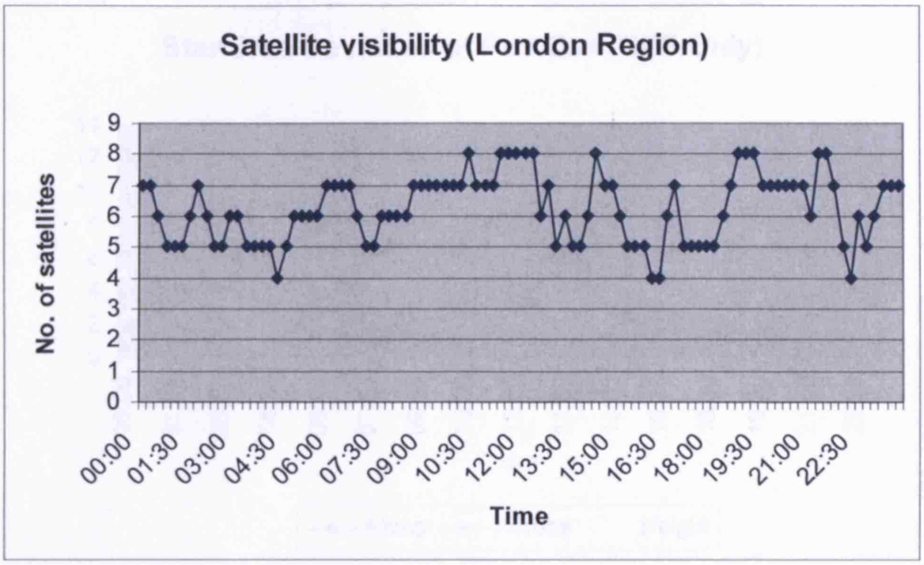


Figure 6.17 The visibility of satellites in n-2 satellites environment

Figure 6.18 and Figure 6.19 present the accuracy performance in three directions for the GPS only and the integration system during the whole day, respectively. Accuracies in all three directions have been improved by using the GPS/Track Database integration system. Accuracies in the across track direction and the height directions are improved greatly, and the standard deviations of using the GPS/Track Database are smoother than using GPS only. However, the accuracy performance in n-2 satellites environment are worse than that in open areas or n-1 satellites environment. Further, both systems cannot provide 100% availability of accuracy.

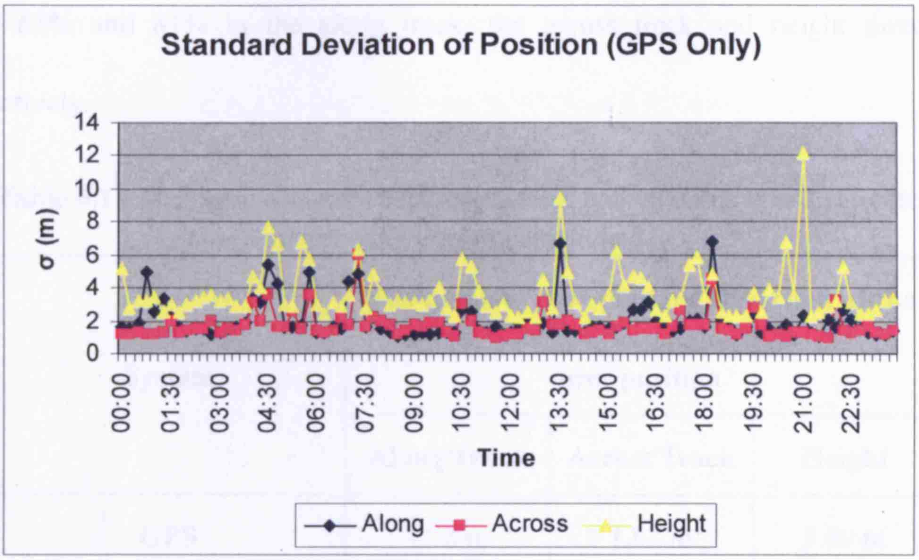


Figure 6.18 The simulated accuracy performance of the GPS only in n-2 satellite environment

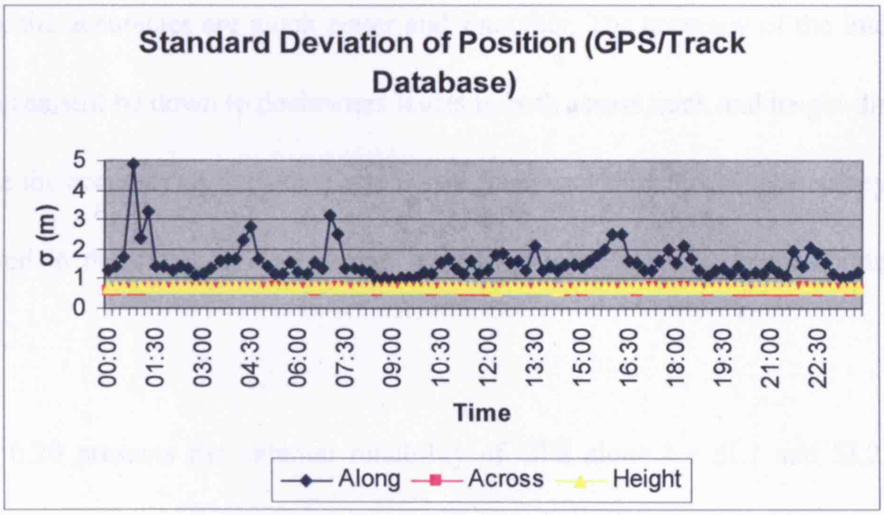


Figure 6.19 The simulated accuracy performance of the GPS/Track Database in n-2 satellites environment

The mean value of the accuracy for both systems is provided in Table 6.12. The accuracy is lower than that in open areas or n-1 satellites environment by using GPS only. For the GPS/Track database, the influence on the accuracy is smaller than GPS only. The mean of the standard deviation of the user position has been enhanced about

22%, 62% and 81% in the along track, the across track and height directions, respectively.

Table 6.12 The mean accuracy performance for both systems in n-2 satellites environment during the whole day

System	**The mean of the Standard Deviation of the user position**		
	Along Track	**Across Track**	**Height**
GPS	1.98 m	1.67 m	3.69 m
GPS/Track Database	1.54 m	0.63 m	0.69 m

For the GPS only, some epochs have rather poor accuracies, but for the integration system, the accuracies are much better and smoother. The accuracy of the integration system can still be down to decimeters levels in both across track and height directions, because the accuracy of track database is still fixed on 1.0 meters. The accuracy is also improved in the along track direction by using the GPS/Track database integration system.

Figure 6.20 presents the external reliability of GPS alone for SL1 and SL2 in n-2 satellites environment during the whole day. For the GPS only, the position shifts are largely worse than that in the open area or n-1 satellites environment. Some epochs are over 1000 meter shifts. Some epochs with only 4 satellites in view, do not have the RAIM outputs (i.e. infinity external reliability). They are obviously unacceptable for safety-critical railway applications. For the GPS/Track Database integration system, the external reliability for SL1 and SL2 is shown in Figure 6.21. In Figure 6.21, we can see that the position shifts also go worse than they are in the open area or n-1 satellites environment, but are not as worse as GPS only. The integration system still has RAIM performance in epochs wherein only four satellites are in view.

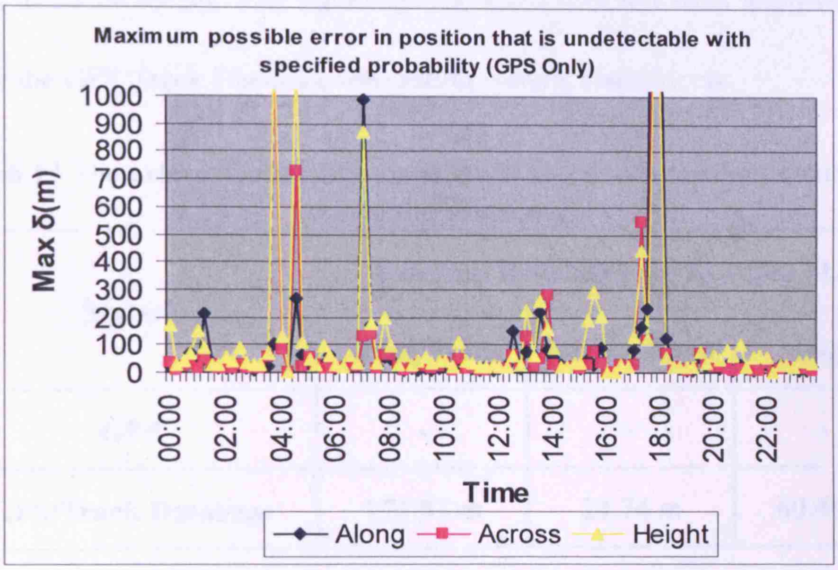


Figure 6.20 The external reliability of GPS only for SL1 and SL2 in n-2 satellites environment

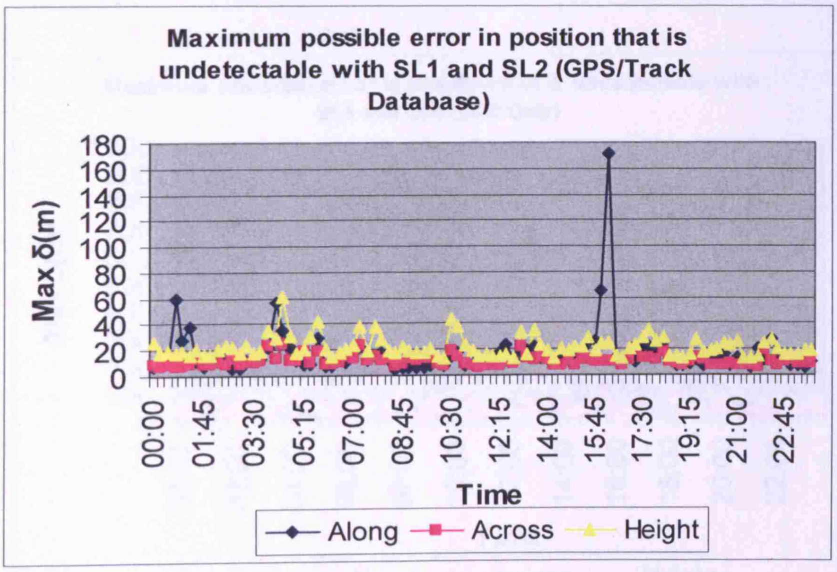


Figure 6.21 The external reliability of GPS/Track Database for SL1 and SL2 in n-2 satellites environment

Table 6.13 summarises the external reliability for both systems during the whole day in n-2 satellites environment. The maximum along track direction shift has been improved to 171.83 meters, the maximum across track direction shift has been improved to 24.74 meters, and the height direction shift has been improved to 60.44 meters by the GPS/Track Database integration system, respectively.

Table 6.13 The external reliability for SL1 and SL2 in n-2 satellites environment during the whole day

System	External Reliability for SL1 and SL 2		
	Along Track	Across Track	Height
GPS	-	-	-
GPS/Track Database	171.83 m	24.74 m	60.44 m

Similar to SL1 and SL2, SL 3 and SL 4 can be simulated together. Figure 6.22 and Figure 6.23 present the external reliability of GPS alone and GPS/Track Database for

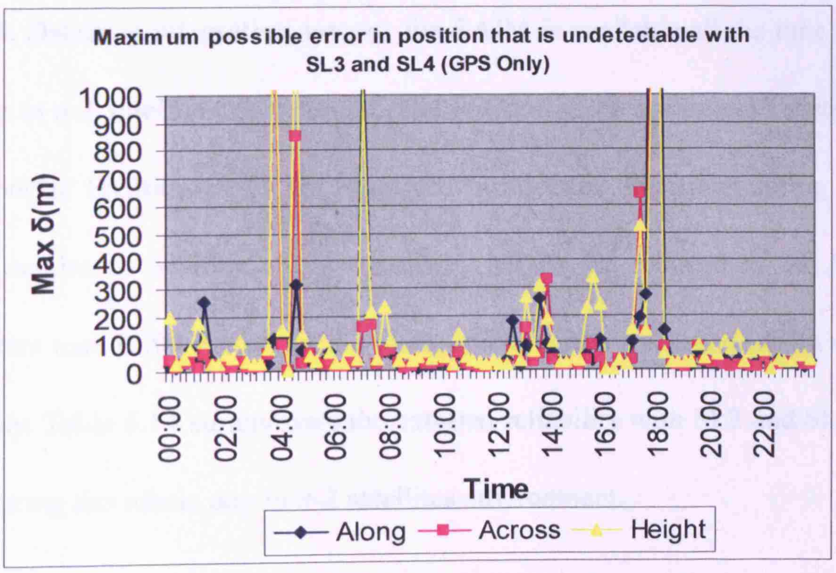


Figure 6.22 The external reliability of GPS only for SL3 and SL4 in n-2 satellites environment

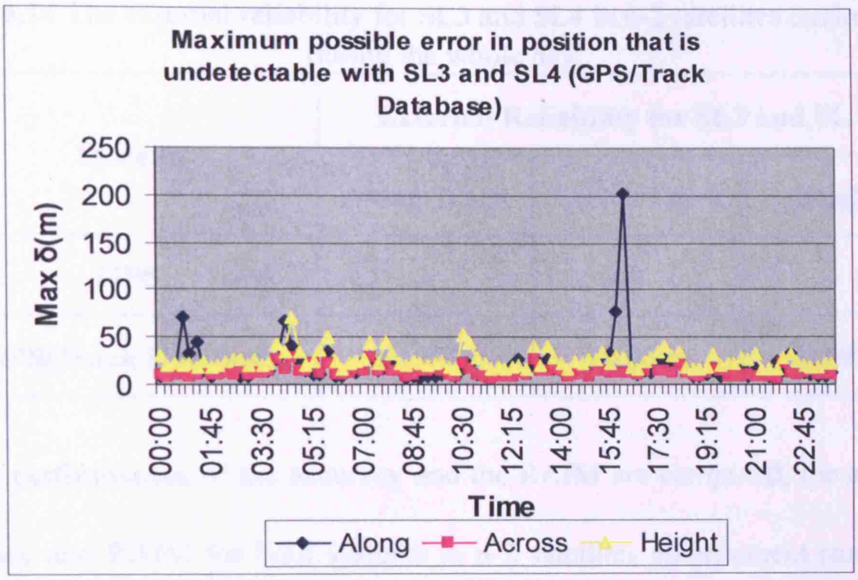


Figure 6.23 The external reliability of GPS/Track Database for SL3 and SL4 in n-2 satellites environment

SL3 and SL4 during the whole day, respectively.

For GPS only, about 4.1% of epochs, the RAIM is not available. Even though it has enough satellites in view, the position shifts are over 100 meters in many epochs in n-2 satellites environment with some epochs being over thousands meters. For the GPS/Track Database integration system, the RAIM is available all the time during the whole day in n-2 satellite environment. The position shifts are over 50 meters only in several epochs. Position shifts are improved in all three directions during the whole day. The maximum position shifts caused by MDEs are reduced to 201.09 meters, 29.06 meters and 70.99 meters in the along track, across track and height directions, respectively. Table 6.14 summarises the external reliability with SL3 and SL4 for both systems during the whole day in n-2 satellites environment.

Table 6.14 The external reliability for SL3 and SL4 in n-2 satellites environment during the whole day

System	External Reliability for SL3 and SL 4		
	Along Track	Across Track	Height
GPS	-	-	-
GPS/Track Database	201.09 m	29.06 m	70.99 m

Since the performances of the accuracy and the RAIM are computed, the availability of accuracy and RAIM for both systems in n-2 satellites environment can be easily calculated. Table 6.15 describes the availability of accuracy and RAIM for both systems.

Table 6.15 The availability of RAIM and the accuracy in n-2 satellites environment

	Availability					
	GPS Only			GPS/Track Database		
	Along	Across	Height	Along	Across	Height
Accuracy	91%	98%	99%	99%	100%	100%
SL1	46%	53%	54%	78%	96%	99%
SL2	16%	15%	21%	31%	49%	76%
SL3	39%	47%	51%	71%	90%	97%
SL4	8%	13%	10%	26%	30%	60%

However, when two satellites in view are randomly reduced, both the GPS standalone and GPS/Track Database system performance is worse than they are in the open area or n-1 satellites environment. The GPS alone system still has good accuracy performance in most epochs, but only 91% epochs of the accuracy reaches the requirements in the along track direction. For the integration system, it is improved to 99% of epochs. The availability of RAIM (GPS only) is really poor. Especially for SL4, it only has 8% in along track direction. This is obviously unacceptable for safety-critical railway applications. The integration system still results in significant improvements on the availability of RAIM and accuracy for all standard levels. Unfortunately, the performance of RAIM is also not acceptable for safety-critical applications. In this sense, other sensors are necessary to help the integration system to reach the requirement level of availability in this environment.

6.2.4 Scenario 4: Different Performances of GPS/Track Database by Using Different Track Database Accuracies in the Open Areas

In three scenarios above, the influences of satellite visibility on both systems are discussed. However, both the satellite visibility and the track database accuracy could impact the performances of the integration system. Because the cost of track database collections depends on the track database accuracy (i.e. higher track database accuracy, higher expense on the collection), the question is what level of track database accuracy is the most cost-efficient. To answer this question, the different track database accuracies are simulated in the open area (Scenario 4). The track database accuracy in the Scenario 1-3 is set to 1.0 meters (medium accuracy level). The alternative accuracies of track database are now set to be 0.1 meters (high accuracy level) and 10.0 meters (low accuracy level).

Similar to Scenario 1, all conditions are the same except the track database accuracy.

Because the track database accuracy will only affect the performance of the integration system, the satellite visibility, the performances of GPS only will be the same as Scenario 1. We only need to change the track database accuracy so that the performances of the integration system in this scenario can be obtained. Figure 6.24 and Figure 6.25 give the accuracy performances of the integration system with different accuracy levels of the track database.

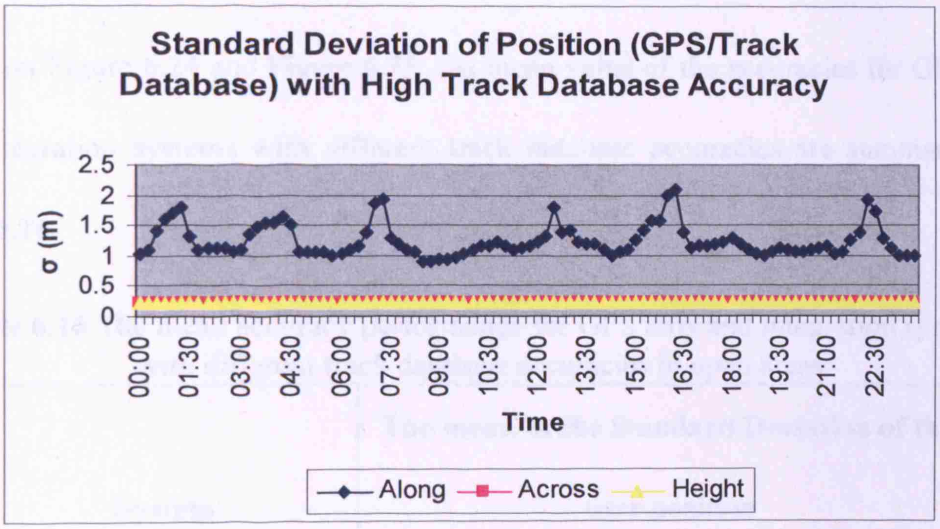


Figure 6.24 The simulated accuracy performance of the GPS/Track Database with 0.1 meters track database accuracy in open areas

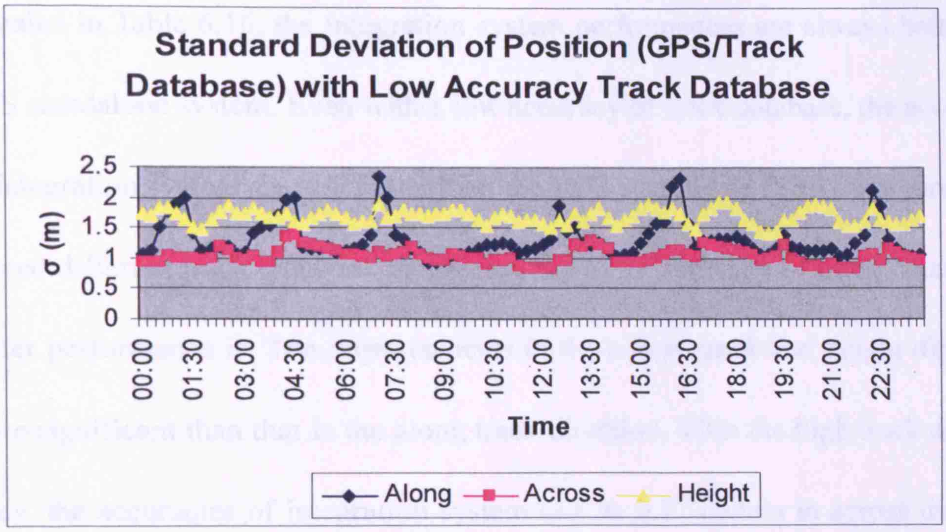


Figure 6.25 The simulated accuracy performance of the GPS/Track Database with 10.0 meters track database accuracy in open areas

Based on Figure 6.24 and Figure 6.25, the mean value of the accuracies for GPS only and integration systems with different track database accuracies are summarised in Table 6.16.

Table 6.16 The mean accuracy performance for GPS only and integration systems with different track database accuracies in open areas

System	The mean of the Standard Deviation of the user position		
	Along Track	Across Track	Height
GPS	1.37 m	1.19 m	2.65 m
GPS/Track Database ($\sigma_{track} = 0.1$)	1.24 m	0.22 m	0.22 m
GPS/Track Database ($\sigma_{track} = 1.0$)	1.26 m	0.60 m	0.68 m
GPS/Track Database ($\sigma_{track} = 10.0$)	1.32 m	1.03 m	1.69 m

As revealed in Table 6.16, the integration system performances are always better than the GPS standalone system. Even with a low accuracy of track database, the accuracies of the integration system are still better than the GPS standalone in all three directions. Compared different track database accuracies, the more accuracy of track database is, the better performance is. The improvements in the across track and height directions are more significant than that in the along track direction. With the high track database accuracy, the accuracies of integration system can be 0.22 meters in across track and height directions.

Figure 6.26 and Figure 6.27 present the external reliabilities of the integration system with different track database accuracies for SL1 and SL2 in open areas during the whole day. Compared with the GPS alone performance (Figure 6.6), position shifts are improved by the integration system with all high, medium and low accuracies of track database in all three directions.

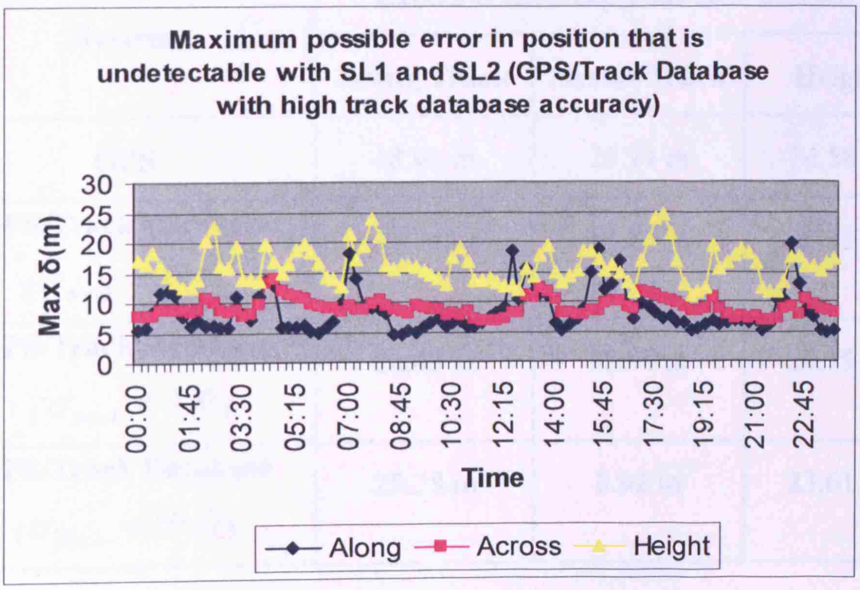


Figure 6.26 The external reliability of integration system with high accuracy of track database for SL1 and SL2 in open areas

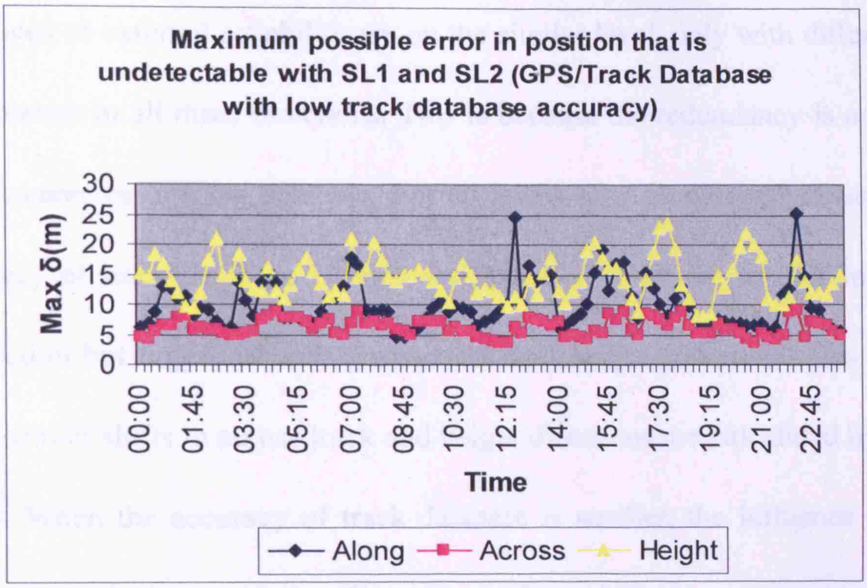


Figure 6.27 The external reliability of integration system with low accuracy of track database for SL1 and SL2 in open areas

According to Table 6.5, Table 6.17 shows the external reliability for the GPS only and integration systems with different accuracies of track database during the whole day in London open areas.

Table 6.17 The external reliability of integration systems with different accuracies of track database for SL1 and SL2 in open areas

System	External Reliability for SL1 and SL 2		
	Along Track	Across Track	Height
GPS	48.01 m	25.71 m	74.58 m
GPS/Track Database ($\sigma_{track} = 0.1$)	19.74 m	13.84 m	24.79 m
GPS/Track Database ($\sigma_{track} = 1.0$)	20.45 m	12.96 m	24.75 m
GPS/Track Database ($\sigma_{track} = 10.0$)	25.28 m	8.92 m	23.61 m

Compared the high accuracy with medium accuracy of track database, the performances of external reliability are on the similar level, only with differences less than 1.0 meters in all three directions. This is because the redundancy is not changed for both accuracies of track database. For all levels of track database accuracy, when the accuracy of track database is improved, position shifts are smaller in the along track direction but bigger in both across track and height directions. This is because that the position shifts in across track and height directions are calculated by Equation 5.31-5.35. When the accuracy of track database is smaller, the influence of internal reliability on track data is bigger. Thus, the position shifts in across track and height directions are bigger.

Similar to SL1 and SL2, SL 3 and SL 4 can be simulated together. Figure 6.28 and Figure 6.29 present the external reliabilities of the integration system with different track database accuracies for SL3 and SL4 in open areas during the whole day.

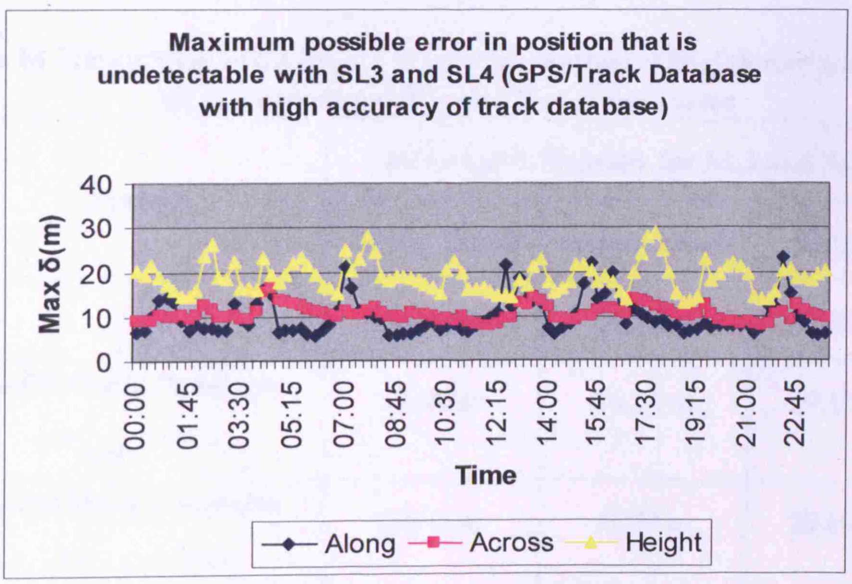


Figure 6.28 The external reliability of integration system with high accuracy of track database for SL3 and SL4 in open areas

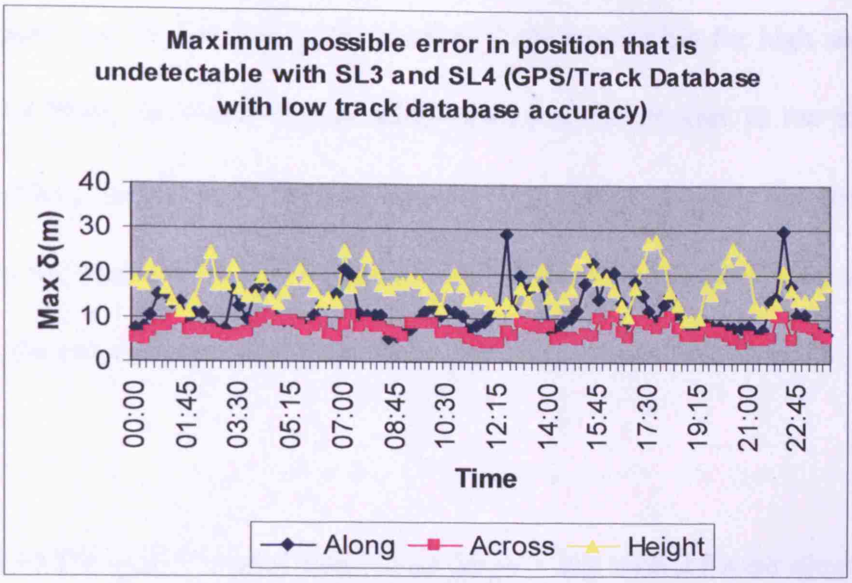


Figure 6.29 The external reliability of integration system with low accuracy of track database for SL3 and SL4 in open areas

For different accuracies of track database, according to Figure 6.8, Figure 6.9 and Figures 6.28-6.29, we can get the same conclusion as for SL1 and SL2. Table 6.18 summarises the external reliability for GPS only and integration systems with different accuracies of track database for SL3 and SL4 in London open areas.

Table 6.18 The external reliability of integration systems with different accuracies of track database for SL3 and SL4 in open areas

System	External Reliability for SL3 and SL 4		
	Along Track	**Across Track**	**Height**
GPS	56.40 m	30.20 m	87.62 m
GPS/Track Database ($\sigma_{track} = 0.1$)	23.20 m	16.26 m	29.15 m
GPS/Track Database ($\sigma_{track} = 1.0$)	24.02 m	15.23 m	29.07 m
GPS/Track Database ($\sigma_{track} = 10.0$)	29.70 m	10.47 m	27.73 m

It can be seen that the influences on position shifts are not big for high and medium accuracy of track database. The position shifts are the biggest in the along track direction when the track database with a low accuracy is used, but they are the smallest in the across track and height directions. Because the GPS is the test system of GNSS, the reference geoid model used in the simulation is WGS-84.

Table 6.19 The availability of RAIM and the accuracy with different accuracies of track database in London open area

Systems	Directions	Availability				
		Accuracy	SL1	SL2	SL3	SL4
GPS Only	Along	100%	86%	39%	78%	24%
	Across	100%	95%	45%	90%	39%
	Height	100%	95%	67%	89%	53%
GPS/Track Database ($\sigma_{track} = 0.1$)	Along	100%	100%	75%	96%	65%
	Across	100%	100%	80%	100%	42%
	Height	100%	100%	100%	100%	95%
GPS/Track Database ($\sigma_{track} = 1.0$)	Along	100%	99%	73%	96%	63%
	Across	100%	100%	91%	100%	72%
	Height	100%	100%	100%	100%	94%
GPS/Track Database ($\sigma_{track} = 10.0$)	Along	100%	98%	61%	95%	43%
	Across	100%	100%	100%	100%	94%
	Height	100%	100%	100%	100%	97%

Since the performances of the accuracy and the RAIM are computed, the availability of accuracy and RAIM for all systems in open areas can be easily calculated. Table 6.19 concludes the availability of accuracy and RAIM for all systems.

However, although the GPS standalone has good accuracy performance in open areas, the integration system with different accuracies of track database can still improve the accuracy performance in all directions. The improvements are significant in across track and height directions. The more accurate a track database is used, the bigger improvement can be obtained. Nevertheless, the improvements of RAIM performance by using different accuracies of track database are not exactly like the accuracy performance. Although, compared with the GPS standalone, the integration system improves the RAIM performance in all levels of track database accuracy; the availabilities of RAIM are not the highest in across track and height directions when the high accurate track database is used. As for the performances in all three directions, the medium accuracy of track database is suggested for the integration system unless there is a special requirement on the across track direction or the height direction.

In sum of all scenarios, either in open areas or in low satellite visibility environments, the GPS/Track Database integration system does result in significant improvements in all three directions. The integration system improves the accuracy and increases the redundancy so that the system only needs two satellites to calculate the position and accuracy, three satellites to compute the RAIM, and four satellites to do the FDE. The high accuracy of track database is not necessary. The most suitable accuracy of track database depends on the requirements of railway control systems. If high accuracy and high availability in the along track direction is required, a high accuracy of track database is suggested; otherwise, the medium accuracy of track database is more cost-efficient. Additionally, the required standard levels are most crucial for the RAIM performance. The integrity risk is the same important as the alarm limit. The reasonable integrity risk and alarm limits are required for safety-critical railway applications. When integrity risk is 3.3×10^{-9} /h, both GPS standalone and integration systems have better performance of RAIM for SL1 than for SL2. When integrity risk is 4×10^{-12} /h, both GPS standalone and integration system have better performance of RAIM for SL3 than for SL4. Compared SL3 with SL2, even with high integrity risk, if the alarm limits are tolerate big like SL3, the performances of the GPS standalone and the integration systems are better for SL3 than for SL2. The SL4 is obviously the most difficult level to achieve for both systems.

To sum up, the performances of integration system are assessed in different environments. Although the integration system has better performances, it still cannot fully compensate the weakness of GPS. For safety-critical railway control systems, more sensors are expected to be integrated with GPS/Track Database to improve the performance of the integration system.

Chapter 7 Data Simulation and Analysis on a Real Railway Line from the Birmingham Area

In Chapter 6, it has been shown that the accuracy and the integrity performance of the GNSS standalone or GNSS/Track Database integration system can be influence by the geometry of satellites in view. The more satellites can be tracked, the better accuracy and integrity can be achieved, given that the increasing number of satellites not only enhances the redundancy but also affords a better geometry of satellites. However, even with the same number of satellites in view, it can also cause huge differences in the accuracy and the integrity. The stand-alone GNSS works well in the open area due to a better geometry of the satellites, but it is not sufficient for the safety-critical railway application because satellite signals could be obstructed by trees, buildings or tunnels in low satellite visibility railway environments. As a result, a worse user/satellite geometry (i.e. poor GDOP) even no solution outputs was obtained because of the few satellites which are tracked. Fortunately, the GNSS/Track Database integration system could significantly enhance both performances of the accuracy and the integrity either in open areas or low satellite visibility environment. This is because the integration system not only increases the redundancy but also changes the use/satellite geometry to the track line direction and satellite geometry. As for the integration system, the geometric influence is now caused by both the track line direction and satellite positions. Therefore, when the track line direction and/or the satellite position are changed, the performances of the integration system will also be changed. In Chapter 6, the track line is fixed (i.e. static) and satellite positions are changed (i.e. dynamic) in the simulation. Therefore, the results only apply for the influence on the performances by the dynamic satellite geometry. However, in the real world, similar to the satellite position, the track lines are also constantly changed. Therefore, when a train travels on a real railway track, the performance of integration system is affected by the dynamic track line direction and satellite geometry. In this chapter, data simulations use real railway track and satellite availability information from Birmingham area to assess the influence of performances by dynamic track line directions and satellite geometry on both GPS standalone and GPS/Track Database integration systems for safety-critical railway applications.

7.1 Data Test Description

To assess the performances of GPS standalone or GPS/Track Database in the real railway environments, a real railway route was chosen between Lichfield and Redditch via Birmingham. The data were collected by the Nottingham Scientific Ltd. A GPS receiver was set on the roof of a train which was in normal revenue service rather than an experimental train. An integrated GPS/INS unit was used to collect positions of the train and also recorded the real time satellite availability information. All data were recorded in the NMEA data file. More details about equipment installation and data collection can be found in Thomas (2007). For the simulation of this research, only a part of data was chosen for the later analysis, from university (Birmingham) railway station to Redditch. Figure 7.1 shows the railway track via the Google Earth software.

This railway track was chosen as it contained most diverse railway environments such as single and multiple track sections, suburban and rural environments, symmetric and asymmetric cuttings, over-bridges and tunnels. The data were collected on 6th July 2006.

Figure 7.1 The test railway line between university (Birmingham) railway station and Redditch

The railway line was about 20 miles and the time of journey was about 30 minutes from the starting to the end point.

For the simulation, the outputs of GPS/INS positions were set as the track database for the GPS/Track Database integration system. It needs to be noted that it should be very careful to use GPS receivers to collect the track database because when the train stops, the outputs of GPS are not fixed on the stop point. The positions from GPS are around the stop point during the stop time. When GPS positions are used to generate the track line when the train does not move, it would obviously get the wrong track line.

Therefore, the track line needs to be smooth at the stop point. The whole track lines in this half-hour data have been smoothened via manual corrections.

The satellite availability information can also be extracted from the NMEA data during the whole journey. It means that the total number of satellites being tracked and PRN numbers of visible satellites are known in each epoch during the whole journey. Therefore, for the postpone processing, similar to Chapter 6, the information of constellation configurations can be extracted from SP3 file according to the satellite availability information. The constellation configurations are stored every fifteen minutes in SP3 file thus a linear interpolation method is required to obtain the configurations in every second for this particular journey.

However, to assess the performances of the GPS standalone and GPS/Track Database integration system, the UERE budgets and RNPs are also needed to be defined. In this chapter, we use the same values which were defined in Chapter 6 (see Table 6.1-6.3).

7.2 Satellite Visibility

The satellite availability information can be directly obtained from the real time NMEA data; therefore, it is easy to get the visibility of satellites in this particular journey. Figure 7.2 shows the satellite visibility through the whole travel journey. It can be seen that there are 20 epochs (1.02%) with no satellites being tracked during the whole journey. It means that there would be no solutions in these epochs whatever using GPS only or the integration GPS/Track database integration system. To be specific, these 20 epochs are composed of four intervals. Therefore, the satellite signals are blocked by four obstructers. Figure 7.3 presents these four obstructers between University Railway Station and Redditch.

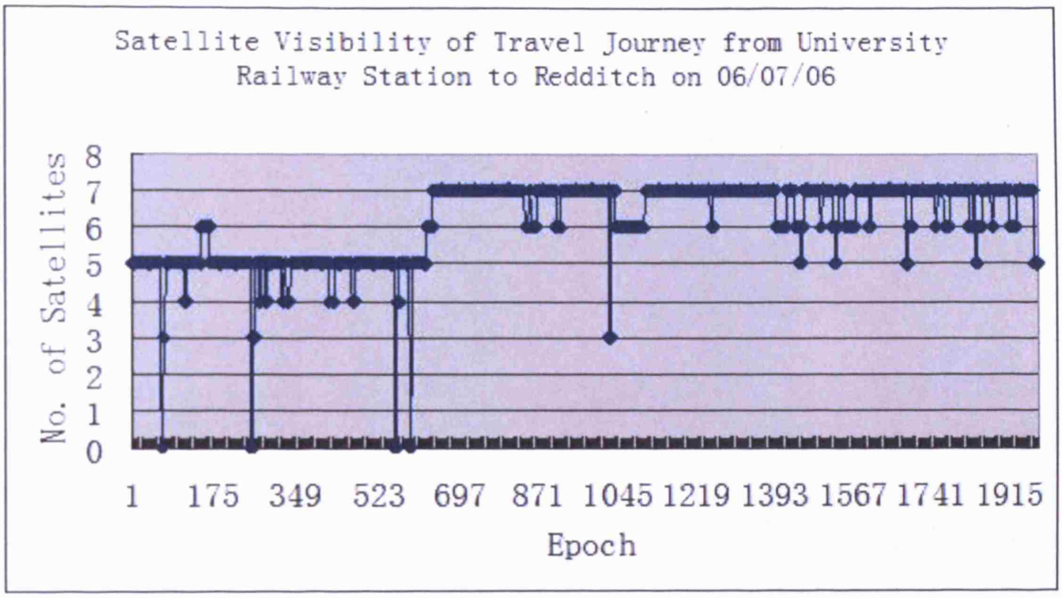


Figure 7.2 The satellite visibility of travel journey from University (Birmingham) Railway Station to Redditch on 06/07/06

The obstructers consist of three over-bridges and one small tunnel, as shown by red centre circles in Figure 7.3. When train travels through these obstructers, no satellite can be seen. In these epochs, the position of train can be calculated unless the system is integrated with other sensors such as the odometer or INS.

The total time of three satellites in view is 15 epochs (0.76%). For GPS standalone, there are not enough satellites to process the 3-D solution during these epochs. The GPS standalone system also cannot estimate the 3-D accuracy and check the integrity when only three satellites are tracked. However, the integration GPS/Track database system works during these epochs. There are enough satellites for the integration system to compute the 3-D position, estimate the accuracy, and calculate the RAIM. In these 15 epochs, satellite signals could be blocked by either the small over-bridges or deep cuttings. Figure 7.4 gives an example of a small over-bridge during the journey.

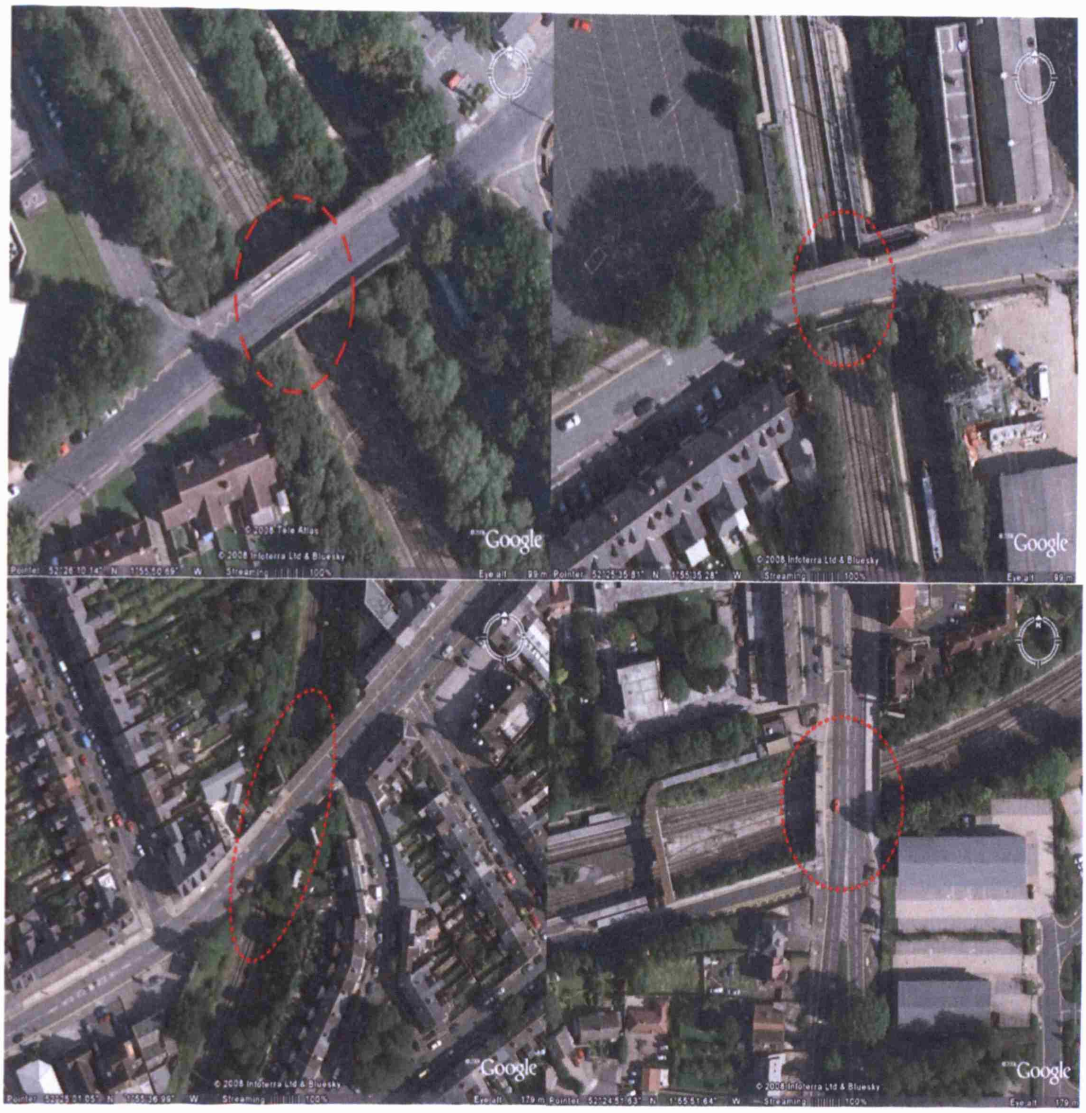


Figure 7.3 Four obstructers between from University (Birmingham) Railway Station and Redditch

The total time of four satellites in view is 50 epochs (2.54%). For GPS standalone, it can only calculate the 3-D position and estimate the accuracy, but the RAIM is not available in these epochs. For GPS/Track Database integration, there are still enough satellites to compute the position, accuracy, RAIM, and even FDE. The rest 95.68% of epochs, there always has five to seven satellites in view. Both the GPS only and integration system can compute the accuracy and RAIM in these epochs.

Figure 7.4 An example of small over-bridge during the journey

7.3 Estimated Accuracy Performance

After extracting the satellite availability information, the performance of accuracy for both systems can be estimated. For GPS only, the accuracy performance is presented in Figure 7.5. It has been illustrated separately in three directions (i.e. the along track, across track and height direction). There is no accuracy performance in total 35 epochs (i.e. 20 epochs of no satellite tracked and 15 epochs of three satellites tracked). They are set to be zero in Figure 7.5. To be specific, in the along track direction, when the accuracy of GPS only is available, it is up to 20.78 meters. The accuracies are over 10.0 meters in several epochs. At most epochs, they are between 1.0 meters and 3.0 meters. In the across track direction, the accuracies are worse than that in the along track direction. The accuracies are over 10.0 meters in several epochs. When the accuracy of GPS only is available, the maximum accuracy is increased to 29.16 meters during the journey. In the height direction, the accuracies are between 3.0 meters and 5.0 meters in most epochs. The maximum accuracy is about 7.38 meters.

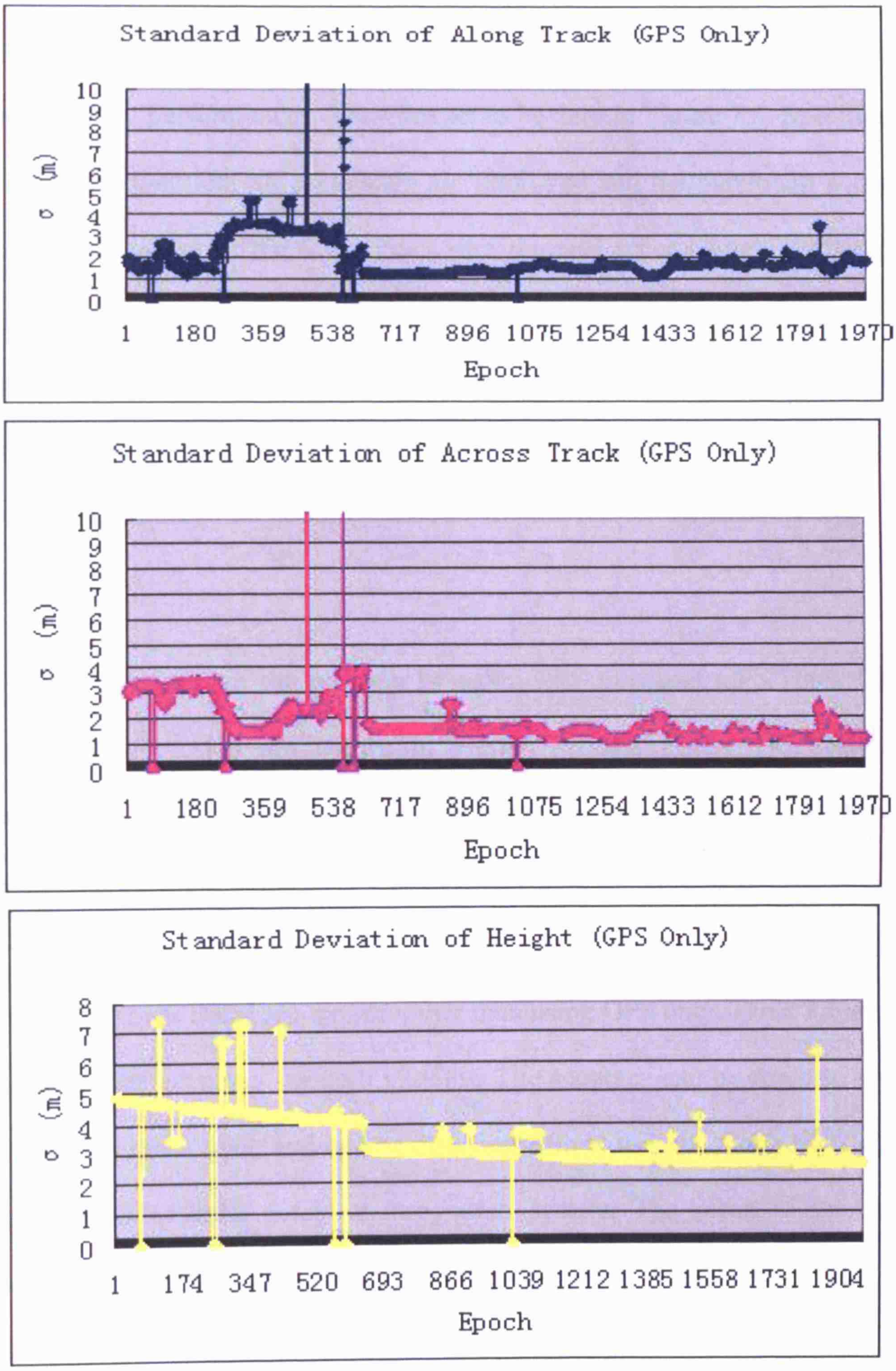


Figure 7.5 The accuracy performance of the GPS only in the travel journey from University (Birmingham) Railway Station to Redditch on 06/07/06

As Section 7.1 described, the real time GPS/INS positions are set as our reference track database. For the simulation, the accuracy of track database is still assumed as 1.0 meters. The accuracy performance of the GPS/Track Database for this particular journey is shown in Figure 7.6. For the integration system, only 20 epochs (1.02%) have no accuracy performances. They are set to be zero in Figure 7.6. Specifically, in the along track direction, all accuracies are improved and the maximum accuracy is reduced to 4.10 meters. In the across track direction, the accuracy now is more precise than that in the along track direction. All accuracies are less than 3.72 meters during the journey. At most epochs, they are less than 1.0 meters. In the height direction, similar to the across track direction, the accuracy is also better than the along track direction and the maximum accuracy is down to 3.80 meters. At most epochs, they are around 1.0 meters.

However, the accuracy of the position by using GPS only and GPS/Track Database during the travel period are shown in Figure 7.5 and Figure 7.6, respectively. Compared Figure 7.5-7.6, it can be seen that the accuracies in all three directions have been improved by using the GPS/Track Database integration system. The across track direction and height direction are greatly improved, and the standard deviations of using the GPS/Track Database are smoother than using GPS only. Table 7.1 shows the mean value of the accuracy for both systems. The accuracy can be down to about 1.0 meters in both across track and height directions. So is it in the along track direction by using the GPS/Track database integration system. The mean of the standard deviation of the user position has enhanced about 14.82%, 55.82% and 69.29% in the along track, the across track and height direction, respectively. The improvements in the across track and height directions are bigger than that in the along track direction because the track line direction gives more information in the across track and height direction than in the along track direction.

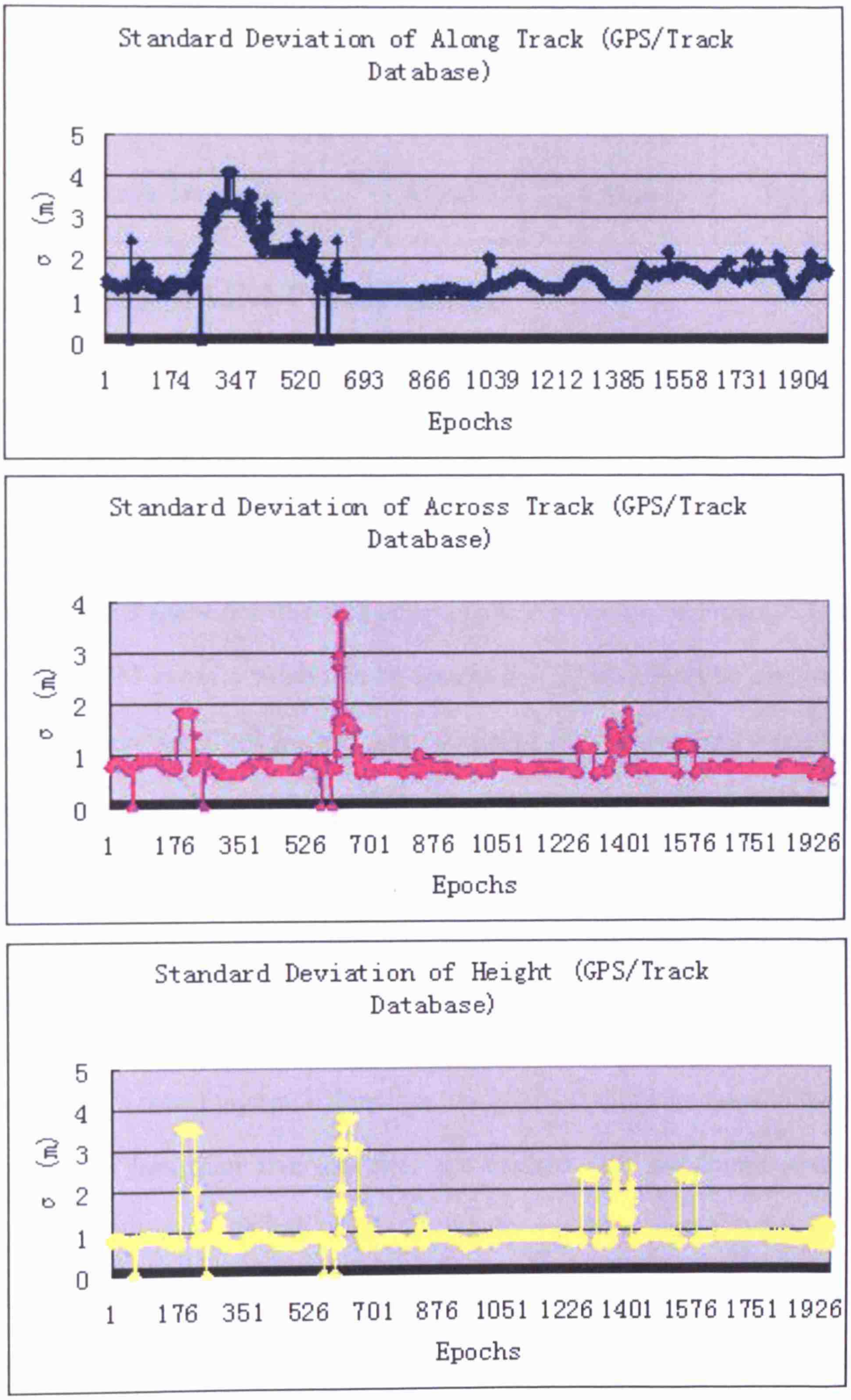


Figure 7.6 The accuracy performance of the GPS/Track Database in the travel journey from University (Birmingham) Railway Station to Redditch on 06/07/06

Table 7.1. The mean accuracy performance for both systems in the travel journey from University (Birmingham) Railway Station to Redditch on 06/07/06

System	The mean of the Standard Deviation of the user position		
	Along Track	Across Track	Height
GPS	1.89 m	1.87 m	3.36 m
GPS/Track Database	1.61 m	0.83 m	1.03 m

7.4 Estimated RAIM Performance

Due to different definitions of the standard level of RAIM requirements in Table 6.2, the performances of RAIM are also tested separately. Firstly, for Standard Level (SL) 1 and Standard Level (SL) 2 (i.e. α is 1×10^{-4} and β is 3.3×10^{-5}), the external reliability of GPS Only for this particular journey is shown in Figure 7.7. For GPS standalone, RAIM is not available in 85 epochs (i.e. 20 epochs of no satellite tracked, 15 epochs of three satellites tracked and 50 epochs of four satellites tracked). This is because there are not enough satellites for the GPS standalone system to compute the RAIM performance during these epochs. The 85 epochs are divided into 17 intervals thus set to be zero in Figure 7.7. In the rest epochs, the GPS standalone can calculate the RAIM performance.

To be specific, in the along track direction, the position shifts are huge in the first 640 epochs wherein less than five satellites are tracked. The maximum position shift caused by MDEs is up to 213.80 meters during the travelling period. When the number of visible satellites is increased to six or seven satellites, the position shifts are improved about twenty meters. Only about 27.72% of epochs, the protection limits of RAIM (GPS only) are less than 20 meters but no position shift is less than 10 meters.

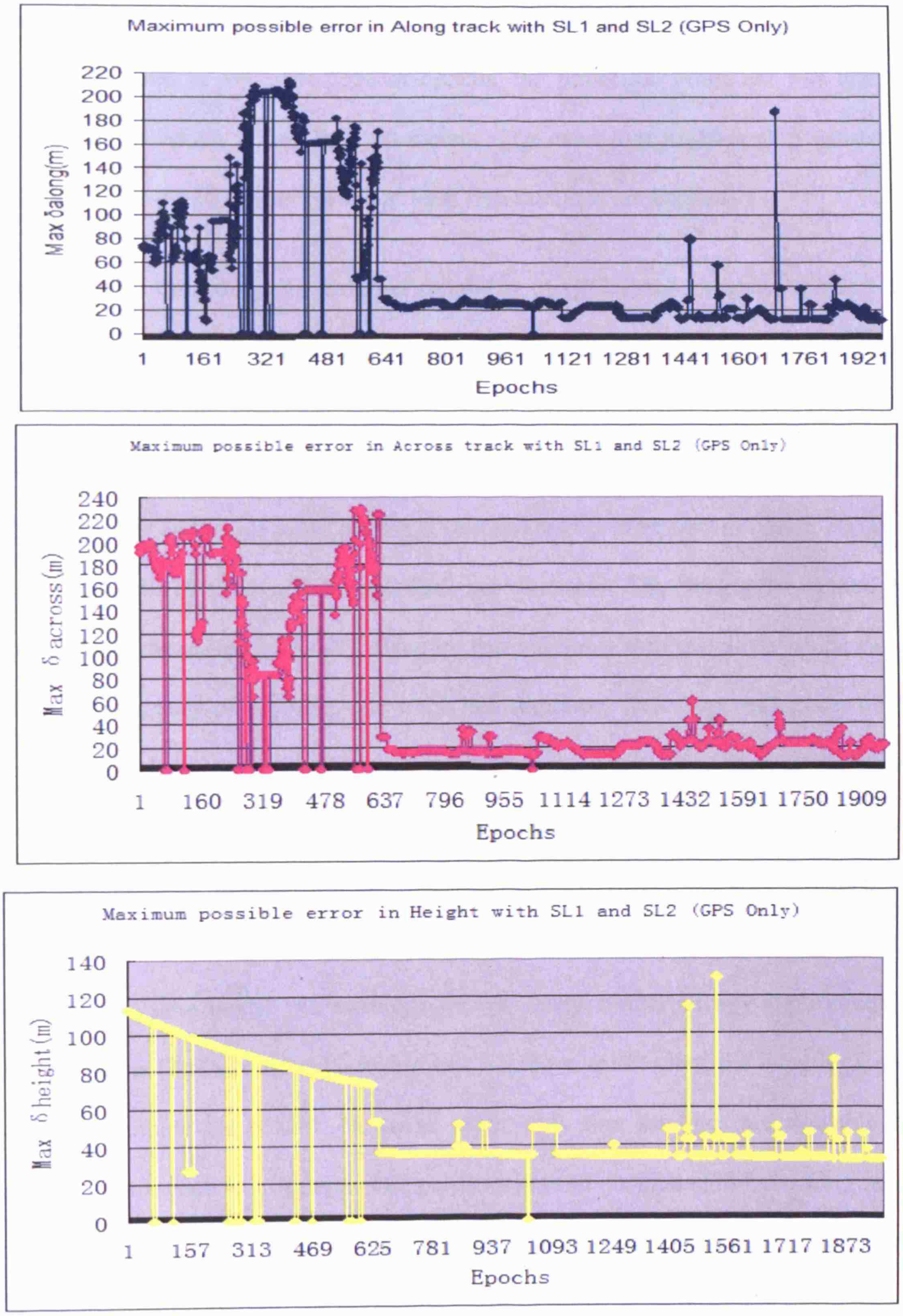


Figure 7.7 The external reliability of GPS only for SL1 and SL2 in the travel journey from University (Birmingham) Railway Station to Redditch on 06/07/06

In the across track direction, in about 36.19% of epochs, the protection limits are under 20 meters but no position shift is less than 10 meters. The maximum position shift caused by MDEs is up to 228.47 meters when at least five satellites are tracked. In the height direction, in about 66.75% of epochs, the protection limits are less than 50 meters but no epoch is less than 25 meters. The maximum position shift caused by MDEs is up to 130.47 meters when at least five satellites are tracked.

Similar to the GPS only, the external reliability of GPS/Track Database for SL1 and SL2 can be obtained. Figure 7.8 presents the external reliability of GPS/Track Database during the travelling period. We can see that for the GPS/Track Database integration system, RAIM is not available in only 20 epochs (i.e. no satellite tracked). The 20 epochs are divided into four intervals and are set to be zero in Figure 7.8. In 15 epochs (0.75% of times), three satellites are available. The integration system can obtain the performance of RAIM. Because there are only three satellites which can be seen, the geometry of satellites and track line directions have huge influences on the performance. In these 15 epochs, there are 10 epochs which have good geometry and the maximum position shifts are from 15 to 40 meters in the horizontal plane. 5 epochs have bad geometry, and the position shifts are up to 470 meters in the horizontal plane. Therefore, if there are three satellites available with bad geometry, the integrity of the integration system is still not available for the safety-critical railway applications. In the 50 epochs (2.5% of times) wherein four satellites are in view, the integrity system still can compute the RAIM. However, when only four satellites are available, the geometric influence is still strict. The performances of position shift have a big range, from 9 to 216 meters in the horizontal plane. When at least five satellites are tracked, for the integration system, the maximum position shifts are improved to 61.23 meters, 82.10 meters and 49.04 meters in the along track, across track and height direction, respectively.

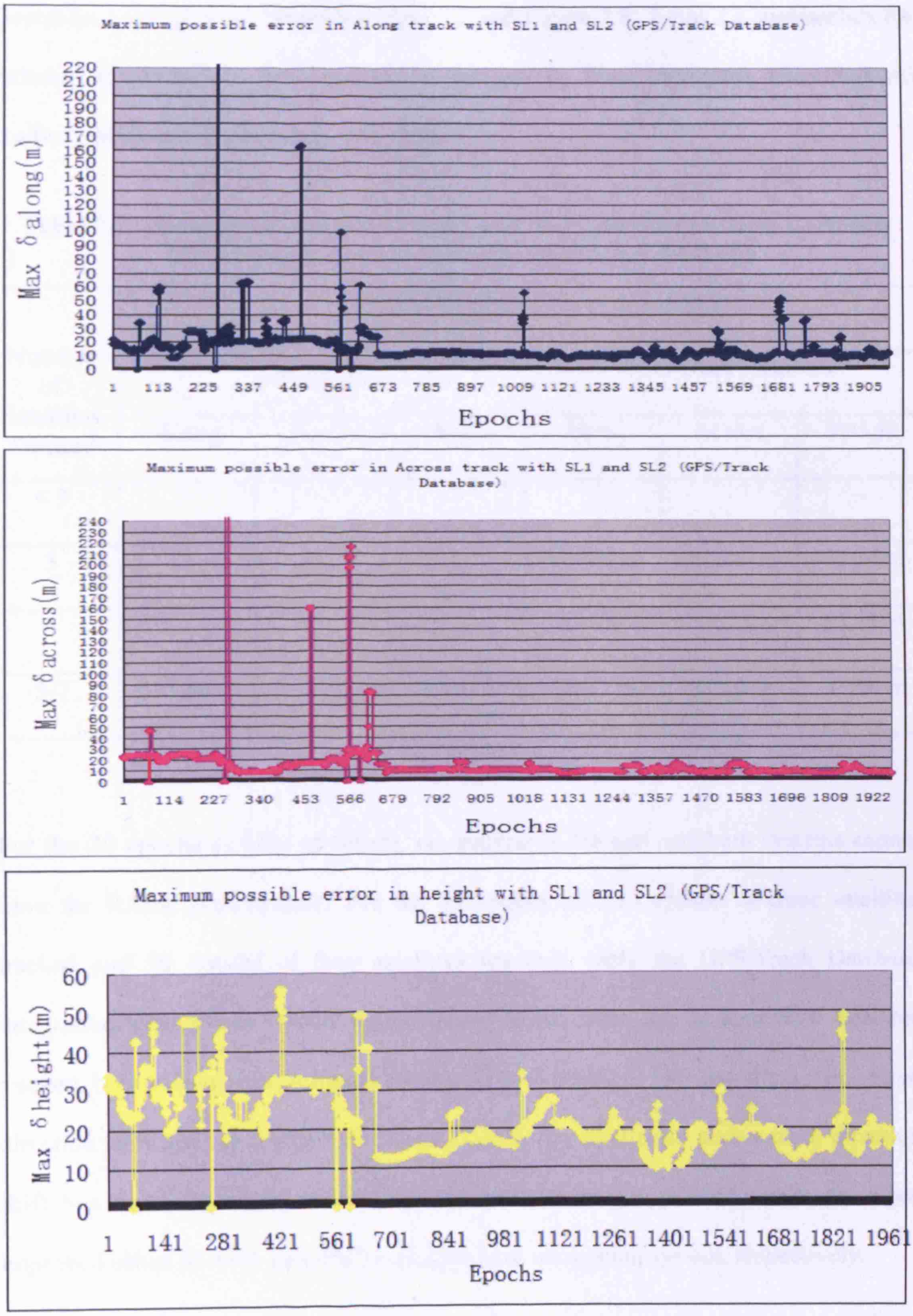


Figure 7.8 The external reliability of GPS/Track Database for SL1 and SL2 in the journey from University (Birmingham) Railway Station to Redditch on 06/07/06

Based on a comparison between Figure 7.7 and Figure 7.8, Table 7.2 summarises the external reliability for both systems in the journey from University (Birmingham) Railway Station to Redditch on 06/07/06.

Table 7.2. The external reliability for SL1 and SL2 in the journey from University (Birmingham) Railway Station to Redditch on 06/07/06

Number of Satellites Tracked	External Reliability for SL1 and SL2					
	GPS Only			GPS/Track Database		
	Along	Across	Height	Along	Across	Height
< 3	-	-	-	-	-	-
3	-	-	-	378.07 m	467.04 m	42.74 m
4	-	-	-	162.75 m	215.25 m	55.41 m
5-7	213.80 m	228.47 m	130.47 m	61.23 m	82.10 m	49.04 m

For the 20 epochs (1.02% of times), no satellite is tracked and both systems cannot have the RAIM performance. For the 55 epochs (i.e. 15 epochs of three satellites tracked and 50 epochs of four satellites tracked), Only the GPS/Track Database integration system has RAIM performance. When there are at least five satellites tracked, both systems can compute the RAIM performance. The maximum along track direction shift has been improved about 71.36%; the maximum across track direction shift has been improved about 64.07%; and the height direction shift has been improved about 62.41% by GPS/Track Database integration system, respectively.

Similar to SL1 and SL2, SL 3 and SL4 can be simulated together. Currently, for SL3 and SL4, requirements for the probabilities of missed detection and false alarm change to α as 1×10^{-4} and β as 4×10^{-8}. Figure 7.9 presents the external reliability of GPS alone for SL3 and SL4 during the travelling period.

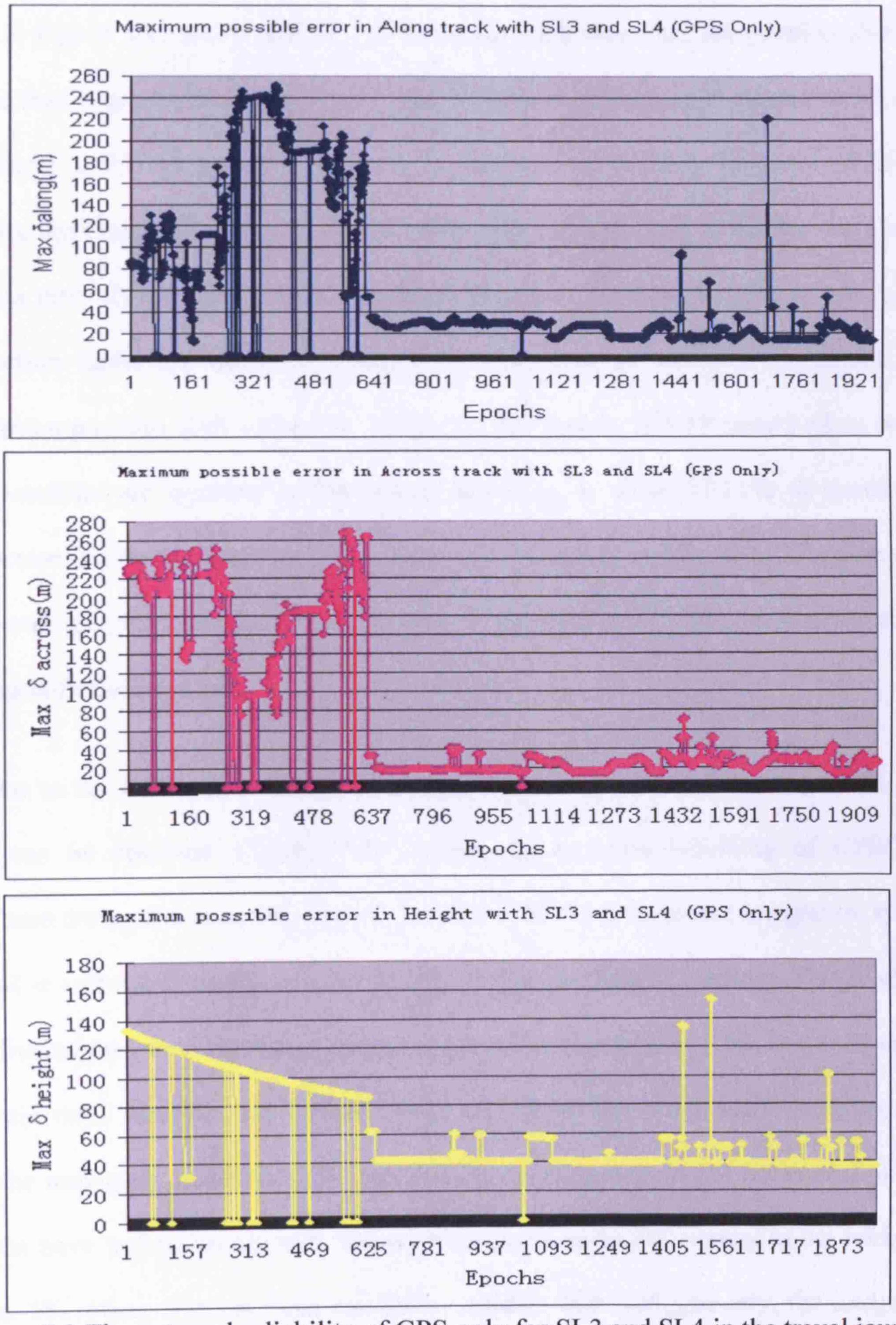
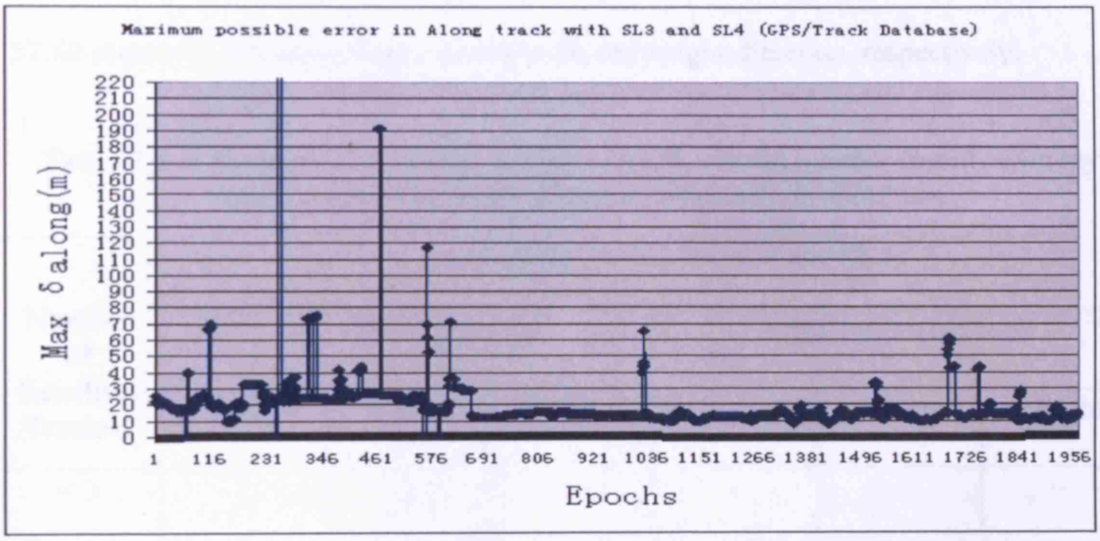
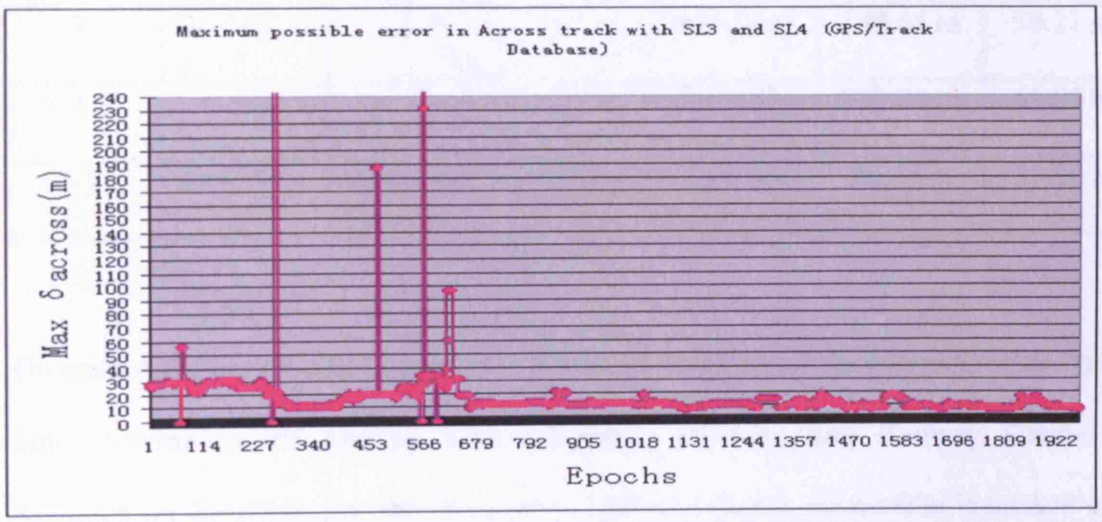


Figure 7.9 The external reliability of GPS only for SL3 and SL4 in the travel journey from University (Birmingham) Railway Station to Redditch on 06/07/06

For GPS standalone, RAIM is still not available in 85 epochs (i.e. 20 epochs with no satellite tracked, 15 epochs with three satellites tracked and 50 epochs with four satellites tracked). The 85 epochs are still divided into 17 intervals and are set to be zero in Figure 7.9. To be specific, in the along track direction, the position shifts are worse than they are for SL1 and SL2. The maximum position shift caused by MDEs is increased to 251.15 meters during the travelling period. Only in about 24.52% of epochs, the protection limits of RAIM (GPS only) are less than 20 meters but it is still no less than 10 meters. In the across track direction, in about 30.17% of epochs, the protection limits are under 20 meters but it is also no less than 10 meters. The maximum position shift caused by MDEs is increased to 268.39 meters when at least five satellites are tracked. In the height direction, in about 57.11% of epochs, the protection limits are less than 50 meters but no epoch is less than 25 meters. The maximum position shift caused by MDEs is increased to 153.26 meters when at least five satellites are tracked.

Similar to the GPS only, the external reliability of GPS/Track Database for SL1 and SL2 can be obtained. Figure 7.10 presents the external reliability of GPS/Track Database during the travelling period. For the GPS/Track Database integration system, RAIM is still not available in only 20 epochs (i.e. no satellite tracked). The 20 epochs are divided into four intervals and are set to be zero in Figure 7.10. In the 15 epochs wherein three satellites are tracked, there are 10 epochs which have good geometry and the maximum position shifts are from 17 to 65 meters in the horizontal plane. 5 epochs have bad geometry with the position shifts up to 550 meters in the horizontal plane. Therefore, if it has three satellites available with bad geometry, the integrity of the integration system is still not available for the safety-critical railway applications. In the 50 epochs (2.5% of times) wherein four satellites are in view, the integrity system still can compute the RAIM. However, when only four satellites are available, the geometric influence is still strict.





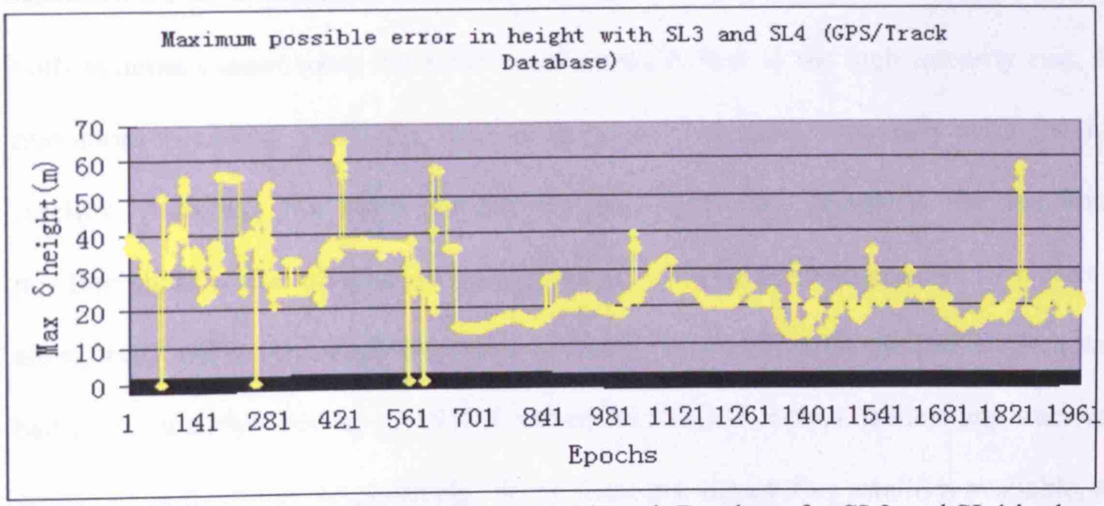


Figure 7.10 The external reliability of GPS/Track Database for SL3 and SL4 in the journey from University (Birmingham) Railway Station to Redditch on 06/07/06

The performances of position shifts have a big range, from 11 to 252 meters in the horizontal plane. When there are at least five satellites are tracked, for the integration system, the maximum position shifts are improved to 71.93 meters, 96.44 meters and 57.60 meters in the along track, across track and height direction, respectively.

Table 7.3. The external reliability for SL3 and SL4 in the journey from University (Birmingham) Railway Station to Redditch on 06/07/06

Number of Satellites Tracked	External Reliability for SL3 and SL4					
	GPS Only			GPS/Track Database		
	Along	Across	Height	Along	Across	Height
< 3	-	-	-	-	-	-
3	-	-	-	444.12 m	548.64 m	50.21 m
4	-	-	-	191.18 m	252.86 m	65.09 m
5-7	251.15 m	268.39 m	153.26 m	71.93 m	96.44 m	57.60 m

Compared Figure 7.9 and Figure 7.10, Table 7.3 summarises the external reliability of both systems in the journey from University (Birmingham) Railway Station to Redditch on 06/07/06. For the 20 epochs (1.02% of times), no satellite is tracked and both systems cannot have the RAIM performance. Due to the high integrity risk, the maximum positions shifts for both systems are increased especially with the low satellite visibility. For three visible satellites with bad geometry, the maximum position shifts of the integration system are up to 444.12 meters and 548.64 meters in along track and across track direction, respectively. As for four satellites tracked with bad geometry, they are up to 191.18 meters and 252.86 meters in the along track and across track direction, respectively. When there are at least five satellites available, the maximum position shifts of the GPS only are increased to 251.15 meters (Along), 268.93 meters (Across) and 153.26 meters (height), respectively. For the integration system, they are improved to 71.93 meters (Along), 96.44 meters (Across) and 57.60 meters (height). The maximum along track direction shift has been improved about 71.36%; the maximum across track direction shift has been improved about 64.07%; and the height direction shift has been improved about 62.41% by the GPS/Track Database integration system, respectively.

7.5 Availability

Since the performances of the accuracy and the RAIM are computed, the availability of accuracy and RAIM for both systems can be easily calculated. Table 7.4 summarises the availability of accuracy and RAIM for both systems in the journey from University (Birmingham) Railway Station to Redditch on 06/07/06.

Table 7.4 The availability of RAIM and the accuracy for both systems in the journey from University (Birmingham) Railway Station to Redditch on 06/07/06.

	Availability					
	GPS Only			GPS/Track Database		
	Along	Across	Height	Along	Across	Height
Accuracy	96.50%	97.72%	98.22%	98.22%	98.98%	98.98%
SL1	27.72%	36.19%	66.75%	78.27%	82.49%	98.58%
SL2	0.0%	0.0%	0.0%	18.43%	44.11%	72.54%
SL3	24.52%	30.71%	57.11%	71.73%	75.84%	95.03%
SL4	0.0%	0.0%	0.0%	4.57%	24.31%	60.81%

Both the GPS standalone and GPS/Track Database systems have good accuracy performance in this journey; however, according to the blocked satellite signals, neither of them can guarantee 100% availability of accuracy during the travelling period. However, only the integration system can provide the accuracy performance when only two or three satellites are tracked. Therefore, the integration system improves the availability of accuracy. Additionally, the integration system also improves availability of accuracy and provides better accuracy performance in all three directions.

The availabilities of RAIM (GPS only) are really poor for all levels of RNP standard. Especially for SL2 and SL4, the RAIM (GPS only) is not available during the whole travelling period. Fortunately, the integration system gives the significant improvement on the availability of RAIM for all standard levels. But for standard levels of small alarm limits (SL2 and SL4), the availability of RAIM (GPS/Track Database) is insufficient for safety-critical railway applications. For SL1 and SL3, the availability of RAIM is improved over 40% and 30% in the horizontal and vertical planes by the integration system, respectively.

To sum up, according to the different requirements for safety-critical railway applications, different thresholds of RNPs suggested in Table 6.1 is to assess the operation train performance from University (Birmingham) Railway Station to Redditch on 06/07/06 in Chapter 7. The results demonstrate significant improvements of performances by using track database to complement GPS for the safety-critical railway applications. When the GPS is integrated with track database, the accuracy, RAIM and availability are improved. Both GPS only and GPS/Track Database systems have high accuracy of the positions. The improvements of the accuracy are better in the across track and height directions rather than in the along track direction by using the integrated system. However, due to the LOS problem, the integration system can only improve the availability of accuracy to 98.22%, 98.98% and 98.98% in the along track, across track and height direction, respectively. The results also reveal the importance of the geometry of satellites and track line directions on the low satellite visibility. It means, for the integration system, the geometric influence could be caused not only by the reduced satellites and moving satellite positions but also by the change of the track line direction. The bad geometry of satellites and track line directions could cause the position shifts to several hundred meters which means the RAIM is not available for safety-critical railway applications. Furthermore, the integration system can calculate the position with as few as two satellites tracked and compute the RAIM with as few as three satellites available. Nevertheless, the integration system is not able to completely compensate the weakness of GNSS. Therefore, to get the extra higher availability, GNSS needs to be integrated with other sensors which would not be affected by the LOS problem.

Chapter 8 Conclusions

This chapter summarizes the performances of the GNSS standalone system and GNSS integrated with a track database system for the safety-critical railway applications in different railway environments. This chapter concludes with the implications for the future work on the GNSS-based railway control system.

8.1 Summary and Conclusions

In the traditional railway control systems (e.g. track circuits and axle counters), trains are allowed to occupy a certain section of track between two block points. The whole section which is indicated as "fixed block" is reported to be occupied regardless of the length or the speed of trains. Additionally, the exact position of the train is not known in the fix block. A long fixed block is thereby needed for the trains at a high speed, and trains should always be kept further apart than the minimum safe stopping distance. In this sense, the capability of the traditional railway control system does not achieve its maximum capability. Furthermore, these systems require large investments and high operational and maintenance costs; consequently, the price of railway tickets is high. In order to meet with the passenger's requirements on railway performance parameters (e.g. price, frequency, punctuality) and the strong competition with other public transport, although the traditional railway control systems can guarantee a high level of safety, the railway industry is interested in developing a new control system which is expected as a "moving block" system for the high speed train operation.

Compared with traditional railway control systems, the GNSS has benefits such as lower initial (e.g. all necessary equipments can be stored on the locomotive) and maintenance costs (e.g. no track side equipments are needed owing to the vandalism) as well as potentially improved capability of railway lines with the high accuracy performance of GNSS. However, in recent time, GNSS has also been found in non-safety railway applications (e.g. for locating trains in order to provide passengers with arrival and departure information); however, it still cannot be used in a standalone mode for safety critical railway applications such as automatic train control, automatic door opening, and train integrity monitoring. This is because GNSS suffers from the line-of-sight problem, namely: GNSS might be unavailable when trains run through the areas with low satellite visibility (e.g. in urban canyons, deep cutting sides and tunnels). This is obviously unacceptable for safety critical railway applications for which the integrity is as important as the accuracy.

To compensate the deficiency of GNSS, in this thesis, a new rigorous mathematical model for the integration of the GNSS with the track database is developed. A key feature of this model is its ability to model errors of both GNSS and the track database measurements in order to achieve realistic performance statistics for the integration system. According to the difference of track lines, linear and nonlinear models are built to solve the train position for the straight track line and the curve of track line. A strategy is also given to search the final two track points for the linear mathematical model. The stimulations show that the integration system reduces positioning, in principle, into a one dimensional problem so that the system only needs as few as two satellites to calculate the position and accuracy, three satellites to compute the RAIM and four satellites to do the FDE.

However, for the satellite-based railway control system, the relevant required navigation performances are needed to be defined for the related safety critical railway applications. Unfortunately, there is no unified definition of RNPs for safety critical railway applications in the world.

This thesis used both simulated London area information and real railway satellite availability information in the Birmingham area to obtain the performances of Required Navigation Performance (RNP) parameters for both GNSS standalone system and GNSS/Track Database integration system. The major findings are summarised as follows:

- **Reference RNPs in safety-critical railway applications:** The reference RNPs for the safety-critical applications is discussed in this thesis. According to the railway characteristic, the horizontal plane requirements being divided into the along and across track directions as the train position is located to these two directions. The accuracy requirement is set to be 4 meters in the along track and across track directions, and it is enough for many applications. The integrity level is set from 3.3×10^{-9}/h (high safety integrity level) to 4×10^{-12}/h (very high safety integrity level). The simulations show that the alarm limits and the integrity risk are the major influences on the RAIM performances of the GNSS standalone and integration system.

- **Performances of GNSS standalone and GNSS integrated with a track database in an open area of railway environments:** The GNSS standalone has good navigation performances in that the satellite visibility is good around London open area railway environments. The accuracy of GPS standalone is up to 2.99 meters and around 1.0 meters at most epochs. Nevertheless, the availability of GNSS standalone is not sufficient according to the high RNP requirements. For SL4, it is only 13% in the along track direction. Compared to the standalone GNSS, the integration system improves the accuracy about 9%, 50% and 74% in the along track, the across track and height directions, respectively. The improvements of the accuracy for the integrated system are more obvious in the across track and height directions than in the along track direction because the track line direction provides more information in the across track and height directions than in the along track direction. As for the RAIM performance and availability, the integration system also achieves a great improvement. In particular, for the stringent high safety level (SL4), the external reliability has been improved over 50% and the availability of RAIM has been improved about 40% in all three directions.
- **Performances of GNSS standalone and GNSS integrated with a track database in low satellite visibility railway environments:** In the tough environment where the visibility of satellites is low (i.e. n-2 satellites are in view), the GNSS standalone navigation performances are worse than they are in the open area as described above. The GNSS still has good accuracy performances in most

epochs; however, only 91% epochs of the accuracy reaches the requirements in the along track direction. The availability is rather poor for all standard levels. Especially for SL4, it is only 8% in the along track direction. Compared to the GNSS standalone, the improvements of accuracy, integrity and availability by the integration system are more apparent than they are in the open areas due to the poor performance of the standalone GNSS in the tough areas.

- **Performances of GNSS standalone and GNSS integrated with a track database in a real railway line:** Data tests used real railway track and satellite availability information from University (Birmingham) Railway Station to Redditch in the Birmingham area. The influence of performances by dynamic track line direction and satellite geometry in both GNSS standalone and GNSS/Track Database integration systems for safety-critical railway applications was tested. The results show significant improvements of performances by using track database to complement the GNSS. For the GNSS standalone, the accuracy is between 1.0 meters to 3.0 meters at most epochs. For the integration system, the accuracy is less than 1.0 meters at most epochs. The improvements of the accuracy are more apparent in the across track and height directions than in the along track direction by using the integrated system. However, due to the LOS problem, the integration system can only improve the availability of accuracy to 98.22%, 98.98% and 98.98% in the along track, across track and height directions, respectively. The results also reveal the importance of the geometry of satellites and track line directions with a low satellite visibility. It means, for the integration system, the geometric influence could be caused not only by the number and positions of satellites but also by the change of the track line direction. The bad geometry of satellites and track line directions could cause the position to shift up to several hundred meters. This means the RAIM is not available for safety-critical railway applications. However, when integrity risk is 3.3×10^{-9}/h, both GPS standalone and integration systems have better performances of RAIM for SL1 than for SL2. When integrity risk is 4×10^{-12}/h, both GPS standalone and integrated system have better performances of RAIM for SL3 than for SL4. Compared SL3 with SL2, if the alarm limits are tolerate big like SL3, the performances of the GPS standalone and the integration systems are better for SL3 than for SL2, even with a high integrity risk. The SL4 is obviously the most difficult level to achieve for both systems.

- **The cost efficient accuracy of the track database:** The different accuracies of track database are also tested in this research. The results show that the most suitable accuracy of track database depends on the requirements of railway control systems. If high accuracy and high availability in the along track direction are required, a high accuracy of track database (i.e. 0.1 meters) is suggested; otherwise, the medium accuracy of track database (i.e. 1.0 meters) is more cost-efficient.
- **Improvements by the integration system**: The GNSS/Track Database integration system provides significant improvements of accuracy, integrity and availability in all three directions (i.e. along track, across track and height

directions). The accuracy of the integration system can be improved about 14.82%, 55.82% and 69.29% in the along track, the across track and height directions, respectively. For SL1, the availability of RAIM by the integration system is improved about 50% in the horizontal plane. The integration system also increases the redundancy so that the system only needs two satellites to calculate the position and accuracy, and three satellites to compute the RAIM and four satellites to do the FDE.

- **Other sensors for the integration system to improve the performance for safety-critical railway applications**: Since the integration system is not able to completely compensate the weakness of the GNSS (i.e. no GNSS signals or less than two satellite signals), it cannot satisfy the high availability requirements of safety-critical railway applications. The potential solution is to integrate with other sensors such as INS and odometers so that the high accuracy, integrity and availability performances could be achieved.

The findings described above show that the current research has fulfilled the research objectives introduced in 1.2.

8.2 Future Work

For the safety-critical railway applications, the integrity is a key measurement for the railway control systems. Although the satellite-based railway control system can provide high accuracy, its integrity performance without any augmentation is poor, especially in the low satellite visibility environments. However, the GNSS/Track

Database integration system provides significant improvements on all RNP parameters. However, it cannot fully remedy the weakness of the GNSS. To satisfy the high level of RNPs, the following recommendations are made for future work.

Firstly, to fully compensate the LOS problem of GNSS, the GNSS/Track Database integration system needs to be integrated with other sensors which will not be affected by the LOS problem. One of the recommendations is to use the INS sensors such as gyroscopes, accelerometers and odometers. The balance between the cost and the performance is a key criterion for a cost efficient integration system.

Secondly, in order to achieve the interoperation with ECTS, the GNSS/Track Database integration system should be integrated with Eurobalise, GSM-R, interlocking and radio block centres. Therefore, in future ECTS Level 3 applications, the train can determine its position by itself. There is no need of the track-side equipments. The train movement can be authorized by the radio block center through the GSM-R. The train integrity can be calculated by the onboard equipments. The integration system will change the "fix block" to the "moving block" so that the maximum capability of railway lines can be reached. However, the future satellite-based railway control system will be a replacement or at least a compensation system for the current railway control systems.

Thirdly, to enhance the performance of GNSS, both ground and satellite augmentation systems (e.g. EGNOS/WAAS, DGPS, RTK) can be used in the satellite-based railway control system. The EGNOS/WAAS can provide the integrity information for the integration system. The DGPS and RTK technologies can provide the position accuracy from meter level to centimeter level. Integration with these augmentation systems can make the satellite-based integration system be available in wider safety-critical railway application areas.

Fourthly, for the GNSS/Track Database integration system, the RAIM method is developed to detect the fault in the measurements. In the GNSS/Track Database integration system, either the satellite measurements or track database measurements can have the fault. If we want to locate the fault and exclude it from the measurements, the FDE for the integration system is needed. To detect and exclude one fault, an additional requirement for the FDE is at least four satellites which can be tracked.

Fifthly, due to the limitation of the scope and time of this doctoral research, the demonstrations were focused on the post processing. The real time processing is a priority for the future work so that real impacts of urban canyons, suburban canyons, forests, tunnels, over-bridges, overhead powers and local multipaths on the integration system can be located. Additionally, the processing time is also crucial for the integration system. Therefore, different software and strategy will be test in the real time processing.

Finally, the multi-system GNSS receiver, which is compatible with GPS, Galileo, GLONASS and Beidou-2, is also an important research area for the GNSS-based railway control system in the future. Either additional satellites or additional signals would help to improve the RNP parameters performance for the GNSS-based train control system. Furthermore, because the safety-critical railway applications have high requirements for SOL, the worst case of the integration system should be considered, and the alternative method should be developed in the future work.

References

Alcouffe, F., Alcouffe, F., & Barbu, G. (2001). *Advanced Position Locator System: Telematics Applications Programme Transport: Final Report* (No. TR4003 APOLO).

Ashkenazi, V., Chen, W., Ochieng, W. Y., Hill, C. J., & Moore, T. (1995, 16 Nov. 1995). *Design tools for GNSS.* Paper presented at the Implementation of GNSS, IEE Colloquium on, London, UK.

Ashkenazi, V., Hill, C. J., & Nagle, J. (1992, September 16-18, 1992). *Wide Area Differential GPS: A Performance Study.* Paper presented at the Fifth International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS-92), Albuquerque, NM.

Baarda, W. (1968). A testing procedure for see in geodetic networks. *Netherlands Geodetic Commission, New Series, 2*(4).

Bedrich, S., & Muncheberg, S. (2004). *GNSS-Based Sensor Fusion for Telematic Applications in Railway Traffic.* Paper presented at the ION GNSS 17th International Technical Meeting of the Satellite Division, Long Beach, CA.

Bian, S., Jin, J., & Fang, Z. (2005). The Beidou Satellite Positioning System and Its Positioning Accuracy. *Navigation: Journal of the Institute of Navigation, 52*(3), 123-129.

BNSF Railway Company. (2005). *General Code of Operating Rules* (5th Edition ed.).

Brown, A. (1989). Extended Differential GPS. *Navigation: Journal of The Institute of Navigation, 36*(3).

Brunner, F. K., & Welsch, W. M. (1993). Effect of the Troposphere on GPS Measurements. *GPS World, 4*(1), 42-51.

Bryan, F. W. (2006, May 1, 2006). Railroad's traffic control systems. *Trains Magazine.*

Cross, P. (1994). *Advanced least squares applied to position fixing: Working paper.* London: University of East London, School of Surveying

Cross, P., Hawksbee, D. J., & Nicolai, R. (1994). Quality measures for differential GPS positioning. *The Hydrographic Journal, Hydrographic Society, 72*, 17-22.

Dinwiddy, S., Breeuwer, E., & Hahn, J. (2004, May 16-19, 2004). *The Galileo system.* Paper presented at the ENC GNSS 2004, Rotterdam, The Netherlands.

DOD/DOT. (2002). *2001 Federal Radionavigation Plan*: DOT-VNTSC-RSPA-01-3.1/DOD-4650.5.

Dutton, L., Rumens, D., & Ruiz, W. F. a. L. (2002, Sept. 24-27 2002). *Galileo's services.* Paper presented at the Proc. ION GPS 2002, Portland, OR.

El-Rabbany, A. (2002). *Introduction to GPS: the Global Positioning System.* Norwood, MA: ARTECH HOUSE, INC.

Enge, P., Walter, T., Pullen, S., Kee, C., Chao, Y.-C., & Tsai, Y.-J. (1996). Wide Area Augmentation of the Global Positioning System. *Proceedings of the IEEE, 84*, 1063-1088.

Euler, -. I. H.-J., Hill, C. D., AG, L., & Muller, D.-I. U. (1996). *Real-Time Precise GPS for Railroad Mapping.* Paper presented at the Position Location and Navigation Symposium, Atlanta, GA, USA.

Fernow, J., & O'Laughlin, D. (2004). *Estimating Continuity of GNSS.* Paper presented at the ION GNSS 17th International Technical Meeting of Satellite Division, Long Beach, CA.

Filip, A. (2006). Railway Safety Certification Requirements for the Galileo Signal-in-Space.

Fontana, R., & Latterman, D. (2000, Jun 26-28, 2000). *GPS modernization and the future.* Paper presented at the IAIN World Congress, in association with ION AM 2000, San Diego, CA.

Forden, G. (2004). The Military Capabilities and Implications of China's Indigenous Satellite-Based Navigation System. *Science and Global Security, 12*, 219-250.

Fraile, J. M. (1999). *Development of a GPS-Based Train Localisation System for Low Traffic Rail Lines.* Paper presented at the European Symposium on Transport Telematics, ETT'99, Potsdam, Germany.

Fraile, J. M., & GmbH, S. (2000). *Advanced GNSS-based Localisation System for Railway Applications.*

FRP. (1999). 1999 Federal Radionavigation Plan: Spring Field.

Galileo Brochure. (March 2003). *Galileo, The European Programme for Global Navigation Services*. Retrieved. from.

Gao, G., Chen, A., Lo, S., Lorenzo, D., & Enge, P. (2007, July/August 2007). The Compass MEO Satellite Codes. *InsideGNSS*, 36-43.

Gauthier, L., Michel, P., & Ventura-Traveset, J. (2003, April 22-25, 2003). *EGNOS - the First European Implementation of GNSS, Programme Development Status Overview.* Paper presented at the ENC GNSS 2003, Graz, Austria.

GCORC. (2005). *General Code of Operating Rules* (5th Edition ed.): General Code of Operating Rules Committee.

Genghi, A., Marradi, L., Martinelli, L., Campa, L., Labbiento, G., Cianci, J., et al. (2003). *The RUNE Project: Design and Demonstration of a GPS/EGNOS-Based Railway User Navigation Equipment.* Paper presented at the ION GPS/GNSS 2003, Portland, OR.

Gibbons, G. (2007, Spring 2007). GLONASS The Way Ahead. *Inside GNSS*, 20-23.

GNSS Rail Advisory Forum. (2000). *Requirements of Rail Applications*.

Goverde, R. M. P. (2005). *Punctuality of Railway Operations and Timetable Stability Analysis.* Unpublished PhD Thesis, The Netherlands TRAIL Research School, Delft, The Netherlands.

GPS World. (2008, Mar 17, 2008). IIR-19(M)/SVN48 Successfully Reaches Orbit. *GPS World*.

Grelier, T., Dantepal, J., Delatour, A., Ghion, A., & Ries, L. (2007, May/June 2007). Inital Observations and Analysis of Compass MEO Satellite Signals. *InsideGNSS*, 39-43.

Groves, P. D. (2008). *GNSS, INERTIAL, AND MULTISENSOR INTEGRATED NAVIGATION SYSTEMS*. London: Artech House.

Hartwig, K., M.Grimm, Horste, M. M. Z., & Lemmer, k. Requirements for Safety Relevant Positioning Applications in Rail Traffic - A Demonstrator for a Train Borne Navigation Platform Called 'DemoOrt'.

Hatch, R., Sharpe, T., & Yang, Y. (2003). *A Simple RAIM and Fault Isolation Scheme.* Paper presented at the ION GPS/GNSS 2003, Portland, OR.

Hein, G. W., Godet, J., Issler, J.-L., Martin, J. C., Erhard, P., Lucas-Rodriguez, R., et al. (2002, Sept. 24-27 2002). *Status of Galileo frequency and signal design.* Paper presented at the ION GPS 2002, Portland, OR.

Hein, G. W., Irsigler, M., Avila-Rodriguez, J., & Pany, T. (2004, May 16-19, 2004). *Performance of Galileo L1 signal candidates.* Paper presented at the ENC GNSS 2004, The European Navigation Conference, Rotterdam, The Netherlands.

Hein, G. W., Rodriguez, J., Wallner, S., Eissfeller, B., Pany, T., & Hartl, P. (2007, January/February 2007). Envisioning a Future GNSS System of Systems Part 1. *Inside GNSS,* 58-67.

Hewitson, S. (2003). *GNSS Receiver Autonomous Integrity Monitoring: A Separability Analysis.* Paper presented at the ION GPS/GNSS 2003, Portland, OR.

Hewitson, S., & Wang, J. (2006). GNSS Receiver Autonomous Integrity Monitoring (RAIM) Performance Analysis. *GPS Solution, 10,* 155-170.

Hofmann-Wellenhof, B., Lichtenegger, H., & Collins, J. (1997). *Global Positioning System: Theory and Practice Fourth, revised edition.* New York: Springer-Verlag Wien New York.

ION Newsletter. (2006). CHINA ANNOUNCES PLANS For Its Own GNSS. *ION Newsletter.*

Izotov, A.A. (1959). *Reference-ellipsoid and the standard geodetic datum adopted in the USSR.* Journal of Geodesy 33 (3), 1-6.

Kenna, P. M. (2006) *The Neeeds of the Business and the operator.* Paper presented at the The 11th IET Professional Development Course on Railway Signalling and Control Systems

Kaplan, E. D., & Hegarty, C. J. (2006). *Understanding GPS Principles and Applications Second Edition.* Norwood, MA: ARTECH HOUSE INC.

Kazantsev, V. N., et al.,. (1994, October 18-21, 1994). *Overview and Design of the GLONASS System.* Paper presented at the International Conference on Satellite Communications, Moscow, Russia.

Kee, C., Parkinson, B. W., & Axelrad, P. (1991). Wide Area Differential GPS. *Navigation: Journal of The Institute of Navigation, 38*(2), 123-145.

Kirjner, R., Lyon, A., & Westbrook, J. (2003, September 9-12, 2003). *Beyond the EGNOS system test bed providing EGNOS services.* Paper presented at the ION GPS 2003, Porland, OR.

Kiss, G. K. (2000). The Use of Modern Satellite Systems in the Railway Traffic. *Periodica Polytechnica Ser. Transp. Eng., 28*(1-2), 123-130.

Klobuchar, J. A. (1991). Ionospheric Effects on GPS. *GPS World, 2*(4), 48-51.

Kovach, K. (2000, January 2000). *New User Equivalent Range Error (UERE) Budget for the Modernized Navstar Global Positioning System (GPS).* Paper presented at the Institute of Navigation National Technical Meeting, Anaheim, CA.

Langley, R. B. (1998). RTK GPS. *GPS World, 9*(9), 70-76.

Lay, R., Mathur, N. G., & Shetty, R. (2003, April 2003 and November 2003). LAAS CAT. II/III Update. *SatNav News, 19 and 21,* 7-8 and 4-5.

Leick, A. (2004a). *GPS satellite surveying 3rd edition*: John Wiley.

Leick, A. (2004b). *GPS Satellite Surveying (3rd Edition).* Hoboken, NJ John Wiley.

Liu, J. (2006, September 2006). Development Update Navigation and Positioning in China. *InsideGNSS,* 46-50.

Lyon, A., Westbrook, J., & Guida, U. (2005, September 13-16, 2005). *Operating EGNOS.* Paper presented at the ION GNSS 18th International Technical Meeting of the Satellite Division, Long Beach, CA.

Maeda, H. (2005, September 2005). *QZSS Overview and Interoperability.* Paper presented at the ION GNSS 2005, Long Beach, CA.

Mangoe. (2006). track circuit. Retrieved http://en.wikipedia.org/wiki/Image:Clear_track_circuit.PNG

Marais, J., Lefebvre, S., & Berbineau, M. (2004). *Satellite Propagation Path Model along a Railway Track for GNSS Applications.* Paper presented at the IEEE

Vehicular Technology Conference, Los Angeles.

Marais, J., Meunier, B., & Berbineau, M. (2000). *Evaluation of GPS Availability for Train Positioning along a Railway Line.* Paper presented at the IEEE Vehicular Technology Conference Boston, MA, USA.

Mattos, P. (2004, May 2004). Acquiring sensitivity - to bring new signals indoors. *GPS World, 15,* 28-33.

Mazl, R., & Preucil, L. (2003). *Sensor Data Fusion for Inertial Navigation of Trains In GPS-dark areas.* Paper presented at the IEEE Intelligent Vehicles Symposium, 2003.

Mcdonald, K., & Hegarty, C. (2000, June 26-28, 2000). *Post-modernization performance capabilities.* Paper presented at the ION AM 2000, San Diego, CA.

Merino, M. M. R., Alarcon, A. G. G., Villares, I. J., & Monseco, E. H. (2001). *An Integrated GNSS Concept, Galileo & GPS, Benefits in terms of Accuracy, Integrity, Availability and Continuity.* Paper presented at the ION GPS, Salt Lake City, UT.

Mertens, P., Franckart, J.-P., & Starck, A. (2003). *LOCOPROL: A Low Cost Train Location and Signaling System for "Low Density" Lines.* Paper presented at the WCRR 2003, Cherryfield, Maine.

Mirabadi, A., Sandidzaden, M. A., Hosseingholian, & Schmid, F. (2002). *Fault Tolerant Train Navigation System via Integration of GPS, Balises, Tachomters and Doppler Radar.*

Mirabadi, A., Schmid, F., & Mort, N. (2003). Multisensor Integration Methods in the Development of a Fault-Tolerant Train Navigation System. *The Journal of Navigation, 56,* 385-398.

Misra, P., & Enge, P. (2001). *Global Position System: Signals, Measurements, and Performance.* Lincoln, Mass: Ganga-Jamuna Press.

Mueller, K. T., Dow, D., Brawner, J., Bortins, R., Alban, S., & Meek, F. (2003). *Field Test Results with a High Speed Rail Prototype GPS Locomotive Location System.* Paper presented at the ION GPS/GNSS 2003, Portland, OR.

Ochieng, W. Y., K.F.Sheridan, K.Sauer, & X.Han. (2002). An Assessment of the RAIM

Performance of a Combined Galileo/GPS Navigation System Using the Marginally Detectable Errors (MDE) Algorithm. *GPS Solution, 5*(3), 42-51.

Ochieng, W. Y., Sauer, K., P.A.Cross, K.F.Sheridan, J.Iliffe, S.Lannelongue, et al. (2001). Potential Performance Levels of a Combined Galileo/GPS Navigation System. *The Journal of Navigation, 54*(2), 185-197.

Ochieng, W. Y., Shardlow, P. J., & Johnston, G. (1999). Advanced transport telematics positioning requirements: An assessment of GPS performance in Great London. *The Journal of Navigation, 52*(3), 342-355.

Onidi, O. (2002, Sept 24-27 2002). *GALILEO is launched.* Paper presented at the ION GPS 2002, Portland, OR.

Palmer, J. (2006, 5-9 June, 2006) *The Need for Train Detection.* Paper presented at the The 11th IET Professional Development Course on Railway Signalling and Control Systems.

P.J.G., T. (1998). Quality Control and GPS. In P. J. G. Teunissen & A. Kleusberg (Eds.), *GPS for Geodesy* (2nd Edition ed.): Springer Verlag.

Parkinson, B. W., James, J., & Spilker, J. (1996). *Global Positioning System: Theory and Applications Volume I* (Vol. 163). Washington: American Institution of Aeronautics and Astronautics, Inc.

People's Daily. (2007, April 16, 2007). China to Build Global Satellite Navigation System. *The People's Daily Online.*

Person, J. (2007). Writing Your Own GPS Applications: Part 2 - Causes of Precision Error. Retrieved http://www.developerfusion.co.uk/show/4652/2/

Prasad, R., & Ruggieri, M. (2005). *Applied Satellite Navigation Using GPS, GALILEO, and Augmentation Systems.* London: Artech House.

Pratt, A. R. (2005, November 2005). *Galileo Signal Structure.* Paper presented at the NAV 05, London, U.K.

Rumsey, A. F. (2006) *Developments in Train Control Worldwide.* Paper presented at the The 11th IET Professional Development Course on Railway Signalling and Control Systems

Ray, J. (2000). *Mitigation of GPS Code and Carrier Phase Multipath Effects Using a*

Multi-antenna System. Unpublished PhD University of Calgary, Calgary, Canada.

Revnivykh, S. (2006). *GLONASS Status Update.* Paper presented at the Proceedings of 46th CGSIC Meeting, Fort Worth, Texas, USA.

Revnivykh, S., Klimov, V., Persev, V., Mitrikas, V., & Kossenko, V. (2005, September 13-16, 2005). *GLONASS Status, Performance and Perspectives.* Paper presented at the ION-GNSS-2005, Long Beach.

Robert J. Amsler, J. (2008). Timetable and Train Order Operation, A Primer. from http://www.gatewaynmra.org/tto.htm (retrieved on 19th February 2008)

Rodriguez, J., Irsigler, M., Hein, G. W., & Pany, T. (2004, Sept. 21-24, 2004). *Combined Galileo/GPS frequency and signal performance analysis.* Paper presented at the ION GNSS 2004, Long Beach, CA.

Rome, H. J. (September 9-12, 2003). *One Dimensional Integrated GPS-Odometer Navigation with Applications to Rail Track Analysis.* Paper presented at the ION GPS/GNSS 2003, Portland, OR.

Saastamoinen, J. (1972). Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites. *Geophysical Monograph 17, American Geophysical Union.*

Sandhoo, K., & Shaw, D. T. M. (2000, Sept. 19-22, 2000). *Modernization of the Global Positioning System.* Paper presented at the ION GPS 2000, Salt Lake City, UT.

Simsky, A., Wilms, F., & J-P.Franckart. (2004). *GNSS-based Failsafe Train Positioning System for Low-density Traffic Lines Based on One-dimensional Positioning Algorithm.* Paper presented at the 2nd ESA Workshop on Satellite Navigation User Equipment Technologics NAVITEC'2004, Noordwijk, the Netherlands.

Singh, A., & Saraswati, S. K. (2006). India Heading for a Regional Navigation Satellite System. *Coordinates,* 6-8.

Skone, S., Hoyle, V., & Coster, A. (2004, May 16-19, 2004). *WAAS Performance under Increased Ionospheric Conditions.* Paper presented at the ENC GNSS 2004, The European Navigation Conference, Rotterdam, The Netherlands.

Soley, S., Farnworth, R., van den Berg, A., Kremers, R., Sanz, J., Macabiau, C., et al. (2004, May 16-19, 2004). *The data collection network: EGNOS revealed.* Paper presented at the ENC GNSS 2004, Rotterdam, The Netherlands.

Spilker, J. (1996). Troposheric effects on GPS. In B. Parkinson, J. Spilker, P. Axelrad & P. Enge (Eds.), *Global Positioning System: Theory and Applications, Vol. I* (pp. 517-546): American Institute of Aeronautics and Astronautics.

Thomas, M., Ejimogu, L., Lowe, D., Dumville, M., W.Roberts, P.Cross, et al. (2007). *Dependability of GNSS on the UK Railways.* Paper presented at the Nav 07, London.

U.S. Coast Guard. (2008). GPS CONSTELLATION STATUS FOR 25 APR 2008: U.S. Coast Guard

US Naval Observatory. (2008) Current GPS Constellation. from http://tycho.usno.navy.mil/gpscurr.html (Retrieved 27th May, 2008)

Urech, A., Diestro, P. J., & Gonzalez, O. (May 13-16, 2002). *GADEROS, a Galileo Demonstrator for Railway Operation System.* Paper presented at the Proceedings of DASIA 2002, Dublin, Ireland.

Van Dierendonck, A. J., Fenton, P., & Ford, T. (1992). Theory and performance of narrow correlator spacing in a GPS receiver. *Navigation: Journal of The Institute of Navigation, 39*(3), 265-283.

Wilde, W. D., Boon, F., Sleewaegen, J.-M., & Wilms, F. (2007, July/August 2007). Tracking China's MEO Satellite on a Hardware Receiver. *InsideGNSS,* 44-48.

Zheng, Y. (2007). *Improving Positioning Accuracy and Integrity in Rail Safety-Critical Applications through the Integration of GNSS with a Track Data Base.* Paper presented at the 20th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2007, Fort Worth Convention Center, Fort Worth, Texas.