Performance of precise marine positioning using future modernised global satellite positioning systems and a novel partial ambiguity resolution technique

Alexander James Parkins

Department of Civil, Environmental and Geomatic Engineering

UCL

August 2009

Supervisors:
Professor Paul Cross
Professor Marek Ziebart

Thesis submitted for the degree of Doctor of Philosophy

I, Alexander James Parkins, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The International Maritime Organisation (IMO) established a set of positioning requirements for future Global Navigation Satellite System (GNSS) constellations in IMO resolution A.915. It is important to be able to determine if these requirements can be met, and what shore infrastructure would be required. This thesis describes the collection of data in a marine environment and the analysis of these data with regards to the requirements.

The data collection exercise was held at the beginning of May 2008 and saw *THV Alert* navigate into Harwich Harbour whilst Global Positioning System (GPS) observation data were recorded from onboard the vessel and from shore-based reference stations. Additional data were obtained from nearby Ordnance Survey reference stations, and two total stations were used to track the vessel's passage to provide a truth model. Several modernised GPS satellites were tracked. The data were processed under different scenarios, using software developed at UCL, and the positioning performance was analysed in the context of the IMO requirements. Potential performance improvements from modernised GPS and Galileo were then discussed.

Providing integrity through single-epoch real-time kinematic positioning, required to meet the strictest IMO requirements, is particularly difficult. The identification of phase observation outliers is not possible before the integer ambiguities are resolved, but an undetected outlier could prevent successful ambiguity resolution. It will not always be necessary to fix all the ambiguities to achieve the required positioning precision, particularly with a multi-GNSS constellation. This thesis introduces a new algorithm for partial ambiguity resolution in the presence of measurement bias. Although computationally intensive, this algorithm significantly improves the ambiguity resolution success rate, increasing the maximum baseline length over which the highest requirements are met with dual-frequency GPS from 1 km to 66 km.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms

ADOP Ambiguity Dilution of Precision

AFM Ambiguity Function Method

AIS Automatic Identification System

AltBOC Alternative Binary Offset Carrier

ARNS Aeronautical Radionavigation Service

ARP Antenna Reference Point

AtoN Aids to Navigation

BOC Binary Offset Carrier

BPSK Binary Phase-Shift Keying

CBOC Composite Binary Offset Carrier

CDMA Code Division Multiple Access

CRAIM Carrier Receiver Autonomous Integrity Monitoring

DGNSS Differential GNSS

DIA Detection, Identification and Adaptation

dwt Deadweight tonnage

ECDIS Electronic Chart Display and Information System

ECEF Earth Centred Earth Fixed

EGNOS European Geostationary Navigation Overlay Service

ESA European Space Agency

FDMA Frequency Division Multiple Access

FEC Forward Error Correction

GAGAN GPS Aided Geo Augmented Navigation

GLA General Lighthouse Authority

GLONASS GLObal'naya NAvigatsionnaya Sputnikovaya Sistema

GMF Global Mapping Function

GNSS Global Satellite Navigation System

GPS Global Positioning System

GPST GPS Time

GST Galileo System Time

GTRF Galileo Terrestrial Reference Frame

IALA International Association of marine aids to navigation and Lighthouse Authorities

IGS International GNSS Service

IGS International GNSS Service

IMO International Maritime Organisation

IOV In-Orbit Validation

IOV In-Orbit Validation

ITRF International Terrestrial Reference Frame

ITRS International Terrestrial Reference System

LAMBDA Least-squares AMBiguity Decorrelation Adjustment

MBOC Multiplexed Binary Offset Carrier

MDB Marginally Detectable Bias

MDB Marginally Detectable Bias

MSAS Multi-functional Satellite Augmentation System

OS Ordnance Survey

PRN Pseudo-Random Noise

PSD Power Spectral Density

QPSK Quadrature Phase-Shift Keying

RAIM Receiver Autonomous Integrity Monitoring

RTIA Ratio Test Integer Aperture estimator

RTK Real-Time Kinematic

SBAS Satellite-Based Augmentation System

SNR Signal to Noise Ratio

SoL Safety of Life

SOLAS Safety Of Life At Sea

SS Direct Sequence-Spread Spectrum

TAI International Atomic Time

TDM Time-Division Multiplexed

TEU Twenty-foot Equivalent Unit

THV Trinity House Vessel

TMBOC Time Multiplexed Binary Offset Carrier

UTC Coordinated Universal Time

VTS Vessel Traffic Services

WAAS Wide Area Augmentation System

WWRNS World-wide radionavigation system

Symbols

α Probability of making a type II error

β Probability of making a type II error

$\check{\cdot}$ Integer estimate

χ^2 Chi-squared statistic

Δ Gross error

Δ_i^u Marginally detectable error

ϵ Residual error

Γ Gamma function

γ Power of the outlier detection test

$\hat{\cdot}$ Least-squares estimate

$\hat{v}$ Least-squares residuals

λ Wavelength

μ Aperture parameter

Ω Aperture space

Φ	Phase measurement
$\Phi(x)$	Cumulative standard normal distribution
ρ	Geometric range
σ	Standard deviation
τ	GNSS signal travel time
A	Design matrix
c	Ratio test critical value
c	Speed of light
c	Vector specifying model error
dt^s	Satellite clock offset
dt_r	Receiver clock offset
$E\{\cdot\}$	Expected value
f	Frequency
f	Frequency
H_0	Null hypothesis
H_a	Alternative hypothesis
I	Ionospheric error
N	Phase ambiguity
P	Code observation
P_f	Ambiguity resolution fail rate
P_s	Ambiguity resolution success rate
Q	Covariance matrix
S_z	Pull-in region around z
T	Tropospheric error
W	Weight matrix
w_i	Outlier test statistic
x	Least-squares parameters
x_0	Initial values of parameters

y	Observed - computed
Z	Decorrelation matrix
a	Float ambiguities
MP_1	Multipath observable
z	Fixed ambiguities

Acknowledgements

I would like to thank my supervisors for their support over the course of my research, particularly Professor Paul Cross his guidance throughout. I would also like to thank Dr Sally Basker and Dr Alan Grant of the GLAs for their advice and assistance.

I am very grateful to Topcon UK for the loan of GPS receivers and total stations for the data collection, and in particular to Mark Francis and Stuart Minney for their practical assistance and time. I would also like to thank the crew of *THV Alert*, Nikos Zinas and Liz Jones for their assistance on that day; also Chris Danezis for writing the total station interface software.

Reference station data were generously supplied by the Ordnance Survey.

This work is funded through an EPSRC CASE studentship with Trinity House.

Chapter 1

Introduction

1.1 Background

International seaborne trade has grown by an average of $3{\cdot}1\%$ p.a. over the last 30 years, primarily driven by the growth in exports from Asia. There have been significant trends towards increased containerisation of goods and a "hub-feeder" model, where long-distance shipping takes place with large ships between busy hubs, and goods are subsequently transshipped with smaller ships to feeder ports; this allows more efficient large ships to make the long-distance journeys. The average ship size has grown by about 2% p.a. during this period, as shipping companies take advantage of economies of scale; the rate is more than double this for container ships, which make up an increasing proportion of the world fleet due to their flexibility. This size increase has been accompanied by a corresponding increase in average speed; larger ships are generally faster. Ports and shipping lanes are increasingly busy, and this congestion, combined with reduced crew sizes due to increased automation, has a negative impact on safety and efficiency.

In order to improve navigational safety and efficiency, the International Maritime Organisation (IMO) and the International Association of marine aids to navigation and Lighthouse Authorities (IALA) are promoting the strategy of "eNavigation", which is defined as the "harmonised collection, integration, exchange, presentation and analysis of maritime information onboard and ashore by electronic means" (IMO, 2008). Many of the technologies required for this are already available, and include Electronic Chart Display and Information Systems (ECDIS), which allow the real-time display of ship positions on electronic charts, and Automatic Identification Systems (AIS), which allow the broadcast of positional information between ships and shore stations. eNavigation requires robust and precise ship positions, for which Global Navigation Satellite Systems (GNSSs) will be relied upon as the primary positioning system.

Currently the US Global Positioning System (GPS) is the only fully-operational GNSS. This system was first declared operational in 1995 and nominally consists of 24 satellites orbiting the earth at a radius of 26,660 km; there are currently 31 operational satellites. Signals are broadcast on two frequencies: L1, at $1575{\cdot}42$ MHz, carries a civil and military signal, whereas L2, at $1227{\cdot}60$ MHz, only carries the military signal, which makes it difficult to track for civilian receivers. The US is currently in the process of modernising the GPS constellation to transmit improved signals on additional frequencies. There are currently seven operational satellites from the first stage of modernisation, Block IIR-Ms which introduce a new civil signal on L2. One of these satellite also transmits the first GPS signal on the new frequency, L5 at $1176{\cdot}45$ MHz,

which will feature in the next stage of modernisation. L5, like L1, is in the Aeronautical Radionavigation Service (ARNS) protected frequency band, and so is protected against interference and is suitable for safety-of-life uses such as automatic docking. The first of the next set of modernised satellites broadcasting L2C and L5, Block IIF, is expected to be launched in 2010. The final stage of the GPS modernisation program, which will start in 2014 at the earliest, will see a modernised civil signal on L1, with greatly improved tracking capabilities compared to the current code.

The EU are developing their own GNSS, called Galileo. There are two test satellites in orbit; the current timeline calls for four In-Orbit-Validation (IOV) satellites to be launched by the end of 2010, and the complete system by 2014. The Galileo satellites will broadcast similar signals to the GPS satellites: the Galileo Open Service, which is intended for general use, will feature similar signals to modernised GPS on L1 and L5 (called E1 and E5a), and will transmit a signal at $1207{\cdot}14$ MHz, E5b, instead of the L2 signal. Due to the similarities between the signals at L1/E1 and L5/E5a, it will be very easy for a receiver to track both systems simultaneously. Russia's GLONASS system, which was declared completed in 1995 but recently has suffered from a lack of satellites, is currently being replenished and may see a change in signal structure to be more compatible with GPS and Galileo; China is also launching its own system, Compass, that may be compatible with the other systems to some extent. The new signals from these modernised systems will facilitate more robust and precise positioning.

In order to assess whether the future performance of modernised GNSSs is sufficient to support eNavigation and marine navigation in general, the IMO has produced a series of requirements on the positioning performance of a future GNSS. The requirements are broken down into four different specific requirements: accuracy, integrity, continuity and availability. Accuracy places a 95% confidence bound on the difference between the observed positions and the truth, which is an important measure of the suitability of a positioning system for a particular requirement. Integrity specifies an alarm threshold and a maximum probability of the position being in error outside this threshold for more than 10 s, the time-to-alarm; this limits the probability of a dangerous error occurring. The continuity requirement makes sure that it is possible to achieve a position over the course of a manoeuvre with the specified probability, and the availability requirement specifies the overall availability of the system for navigation. The requirements are specified for many different marine positioning scenarios; this thesis focuses on port approach, port navigation, and automatic docking. The primary objective of this project is to investigate to what extent it may be possible to meet these requirements with current GPS, and to determine how the anticipated improvements in positioning capability with future GNSSs are likely to affect the ability to meet the IMO requirements.

There are three main GNSS positioning techniques: in order of increasing precision and complexity they are point positioning, Differential GNSS (DGNSS) and Real-Time Kinematic (RTK). Point positioning uses the code measurements at the roving receiver to obtain a position that is accurate to several metres; this is the simplest technique because it does not rely on shore infrastructure. DGNSS uses a shore-based receiver at a known location to estimate the errors in the code measurements, which are then used to correct the measurements at the rover; this can produce sub-metre accuracy, but the validity of the correction decreases with the distance

from the reference station due to spatial decorrelation of atmospheric and satellite orbit errors. RTK uses the phase measurements, which are much more precise than the code measurements but require an ambiguity resolution step to determine the number of wavelengths of the carrier wave between the receiver and the satellite. Centimetric accuracy can be achieved, but a nearby reference station is required and the procedure is much less robust than code positioning. This study will assess the strengths and weaknesses of these different techniques, how they may be used to meet the various IMO requirements, and what shore-based infrastructure and models would be appropriate to use.

During the course of the project, two particular areas requiring deeper investigation were identified. Demonstrating that the integrity requirements are met is a particular problem, due to the very low allowed probability of failure, and the low rate of occurrence of failure events. It is therefore necessary to investigate this topic further, and to develop a method for analysing the robustness to error of different positioning techniques. The second problem is the difficulty of high success rate ambiguity resolution for RTK positioning in difficult environments such as ports. RTK relies upon measurements of the sinusoidal carrier wave of the GPS signal, which are biased by an unknown integer number of cycles, the *integer ambiguity*; it is necessary to determine the value of this ambiguity before RTK positioning can be performed. However, this process is made much harder in difficult environments, and the IMO requirements have high specifications for continuity and availability, so this topic is looked at in greater detail in order to maximise the ambiguity resolution success rate and thereby increase the availability and continuity of RTK positioning in port environments. In particular, a novel algorithm is developed to allow a subset of the full set of ambiguities from all frequencies and satellites to be fixed, with improved robustness at the cost of computation time.

1.2 Objectives

The objective of this thesis is to investigate how the IMO positioning requirements for a future GNSS may be achieved. The performance of currently-operational GPS is assessed in a marine environment with different processing techniques and models, and the future improvements with modernised GNSS is assessed in light of GPS performance. The two critical areas of positioning, integrity and ambiguity resolution, will be studied in greater depth, and algorithms will be developed to study bias robustness and to improve ambiguity resolution in difficult environments. The specific objectives of this thesis are to determine:

- Which GNSS processing techniques may be appropriate to meet different IMO requirements;
- What shore-based infrastructure would be required to support this positioning, which is important for the viability of future positioning system implementations attempting to meet the requirements;
- How robust the different positioning techniques are to biased observations, which will give an idea of what positioning performance may be achieved using different techniques in very difficult environments;

- The performance improvement achieved using the modernised GPS L2C signal. This is the only modernised signal currently available, and an analysis of its performance will allow some assessment of the likely benefits of the other modernised signals;
- How the improvements from modernised GNSS are likely to affect the ability of the different positioning techniques to meet the IMO requirements;
- How to maximise the ambiguity resolution performance of RTK in difficult environments: good RTK ambiguity resolution performance will be essential to meet the IMO requirements with the strictest accuracy values.

Successfully achieving these objectives will give a good indication of which GNSS processing techniques will be appropriate to meet the future IMO requirements, and how easily they may be met.

1.3 Methodology

In order to achieve these objectives, it is necessary to collect GPS measurements and process them with a variety of different techniques and models, so that the results can be compared to the IMO requirements. It is important that these data be collected in a realistic environment so that the results of the processing reflect the results that would be achievable when using modernised GNSS in the conditions covered by the IMO requirements. Therefore the ship *THV Alert* and a jetty at Harwich harbour were made available by the GLAs for a day of data collection. A data collection exercise was planned with a GPS receiver mounted on *THV Alert* to recreate the conditions involved in port navigation and docking. The Ordnance Survey (OS), the national mapping agency of the UK, has a nationwide network of GPS reference stations and made 1 Hz GPS data available for this trial. Nearby OS reference stations with distances from Harwich ranging from 31 km to 111 km were used. In addition, two temporary reference stations were set up at a distance of 1 km and 15 km in order to provide short-baseline positioning. These reference stations allowed GPS processing over a range of baselines in order to determine what kind of shore-based infrastructure might be necessary to meet the IMO requirements with different processing techniques.

It is important to have an independent positioning system as a truth model, so that the quality of the GPS positions can be verified. The IMO requirements are too tight to be met by another radio-navigation system such as eLoran, so it is necessary to use a shore-based tracking system. There are several possibilities, such as a laser tracker or photogrammetric technique, but these are too short-ranged to allow operation while port navigation was carried out. It was therefore decided to use automatic tracking total stations which, when set up with known position and orientation, track a prism mounted on the ship and record the angle and distance several times each second. The range of these is likely to be in the region of 1 km, and preliminary trials were conducted to assess the range and tracking capability of these devices; the results of these trials informed the design of the data collection exercise.

Pre-existing GPS processing software packages provide limited scope for experimentation with different processing techniques and models. It was therefore necessary to develop GPS

processing software to enable processing of the collected data in a variety of ways and with different models. This allowed greater scope for experimentation and analysis of the particular difficulties of processing GPS data in the marine environment, and the development of novel processing and analysis techniques, particularly with regards to studying integrity and improving ambiguity resolution performance. This processing software is then used to process the data collected in Harwich harbour and the results compared to the truth model.

At the time of the experiment the only modernised signals available were 6 GPS satellites broadcasting the modernised L2C signal; receivers capable of tracking this were not common, but it was possible to borrow several with experimental firmware from Topcon UK for use in the trial. Analysis of the L2C signal will allow a quantitative assessment of the benefits of this particular modernised signal. It was not possible to directly analyse the benefits of the other modernised GPS signals and the Galileo system due to their unavailability; instead the literature on the subject is reviewed, and the expected benefits assessed in light of the results from the GPS data processing.

As mentioned in Section 1.1, two particular areas of GPS processing that require further analysis in the context of meeting the IMO requirements are integrity and ambiguity resolution. The data collected in Harwich and the processing software which will be developed are used to study these areas in greater detail. The algorithms used in the receiver to guarantee integrity are studied, and a comparative assessment of the robustness of different positioning techniques performed. A review is made of current ambiguity resolution algorithms, and further techniques are developed in order to increase the robustness of the ambiguity resolution, which is critical for meeting the strictest IMO requirements of automatic docking; the data collected in Harwich are used to test the effectiveness of any techniques developed.

1.4 Outline of thesis

Chapter 1 gives an introduction to the thesis and describes the objectives and the methodology for meeting these objectives.

Chapter 2 describes the IMO requirements for a future GNSS, and gives some background to the necessity of the requirements by describing long-term trends in world shipping and the development of eNavigation, dependent on GNSS for primary position input, to improve safety and efficiency in marine navigation.

Chapter 3 reviews in detail how the current GPS signal is processed. It describes the development of modernised GPS and Galileo, the signals that these new systems will transmit, and the improvements in the signal structure that will be implemented. Interoperability and compatibility issues are discussed, as are other future GNSSs, A review is undertaken of how improvements of the modernised signals and additional satellites will translate into improved positioning performance; this will facilitate the analysis of the possible positioning performance of modernised GNSS in light of the results of the data collection exercise.

Chapter 4 reviews the theory of ambiguity resolution and validation, which is essential to meet the most strict IMO requirements. Partial ambiguity resolution is discussed, and a novel technique for partial ambiguity resolution in the presence of biased observations is developed;

the benefits of this technique are assessed in the processing of the GPS data.

Chapter 5 describes techniques for the detection and exclusion of outliers in the phase and code observations, which is important for robust positioning in safety-of-life applications and to meet the IMO integrity requirements. This chapter shows how to compute the position error that is protected against with a given probability, and how to compute the size of outlier required to cause a position error of a given magnitude and the probability of detecting this outlier. This theory is used in the GPS data processing to make a comparison of the robustness of each positioning technique.

Chapter 6 describes the data collection exercises. The overall data collection plan is discussed, and the development of this plan through several preliminary exercises that were carried out to test the equipment and techniques is described. The main data collection exercise at Harwich harbour is described in detail. The method of generating the truth model from total station observations is given, and the multipath environment at each reference station is analysed.

Chapter 7 describes the software that was developed to allow the collected GPS data to be processed; the different processing techniques and models used are also discussed.

Chapter 8 describes the analysis of the data using the techniques and software developed in Chapter 7; it is determined which techniques would not appropriate to meet certain requirements, the robustness to outliers of each technique is analysed, and the residuals of the L2 and L2C observations are compared. The effectiveness of the ambiguity resolution techniques developed in Chapter 4 is assessed, and the conclusions about the performance benefits of modernised GPS that were made in Chapter 3 are used to predict the effect of modernised GNSS on the ability of different techniques to meet the different IMO requirements.

Chapter 9 gives conclusions and recommendations for further work.

Chapter 2

Positioning requirements and background

The world economy has grown rapidly over the last few years, driven by the double-digit growth of China and India. Over 90% of global trade is carried by sea (IMO, 2005), so this has resulted in significant expansion in seaborne trade, which has on average grown by $3{\cdot}1\%$ p.a. for the last 30 years (UNCTAD, 2008). There have been corresponding increases in commercial maritime activity and merchant fleet capacity; ports and seaways have become busier and more congested. Newer ships are larger and faster to take advantage of economies of scale, so the average ship size and speed increases each year. Although the current economic downturn will temporarily reverse these trends, it is likely that they will continue in the future when international trade recovers. The increased size and number of ships, together with smaller crew sizes due to increased automation, has an impact on navigational safety and efficiency. In order to improve safety and efficiency, the IMO is promoting the concept of eNavigation, which uses sophisticated techniques to collect, transfer and display navigational data. This requires a precise and reliable positional input, which will be primarily derived from GNSS. It is essential that GNSS positioning can provide sufficient performance to support eNavigation, and so the IMO has produced a set of requirements for a future GNSS for positioning in different scenarios. The background to these requirements is described in Section 2.1, some of the different regions where requirements are defined are described in Section 2.2, and some of the more general requirements are described in detail in Section 2.3. Section 2.4 describes the marine background and the eNavigation concept.

2.1 Positioning requirements

IMO resolution A.666(16) (IMO, 1989) describes the IMO policy and procedures that must be followed in order to adopt a radionavigation system for international use, in the case that the SOLAS (Safety Of Life At Sea) 1974 convention was modified to mandate the carriage of receiving equipment for the radionavigation system, which subsequently occurred. In particular, reference is made to the positioning requirements in IMO resolution A.529(13), which are split into two phases of the voyage: harbour entrances etc, and other waters; the required accuracy of the positioning system is dependent upon the minimum distance from danger. The SOLAS convention was first introduced in 1914 after the sinking of the *RMS Titanic* and subsequently updated several times, and currently mandates that ships over 3,000 tonnes carry an IMO-approved World-wide radionavigation system (WWRNS). Resolution A.666(16) was updated by resolution A.815(19) (IMO, 1995) and later by resolution A.953(23) IMO (2003), which revokes both

A.815(19) and A.529(13), and specifies both the procedure for approving a radionavigation system as a WWRNS and the positioning requirements that must be met as part of this adoption process.

Resolution A.860(20) (IMO, 1997) is the first resolution that sets out the procedure for adopting a future GNSS as a WWRNS, including a set of basic positioning requirements for all conditions. This was updated by resolution A.915(22) (IMO, 2001), which revokes A.860(20) and provides a much more comprehensive list of requirements across a broad spectrum of marine applications, such as automatic docking, icebreakers, buoy deployment and hydrographic surveying. It is the requirements specified in this document that are of interest in this project, as the aim is to assess the performance of a future GNSS. This is summarised in (UN MSC, 2004): resolution A.953(23) gives the formal requirements and procedures for accepting new systems as components of the WWRNS, and resolution A.915(22) sets the minimum requirements for a future GNSS to be considered in the framework of A.953(23). It is therefore important that a future GNSS can be shown to meet the requirements set out in A.915(22), in order that it can be included in the WWRNS and adopted for SOLAS vessels. It may be that carriage of Galileo receivers would be mandated by the IMO, but this is unlikely due to the costs of this to users, already on tight margins (Henderson *et al.*, 2009) .

The values of the requirements do not appear to have a particular scientific basis, but rather represent the agreed parameters of the maritime community. They are also to some extent chosen so that they may be met with the desired systems, rather than exclusively based upon the needs of ship positioning for a given application. The values are not static and are under continual review and development: (UN MSC, 2009) is a proposal to review and amend A.953(23) to better reflect the actual requirements of the different scenarios, for example reducing the time over which continuity must be maintained from 3 hours to 15 minutes in harbours, harbour approaches and restricted waters. The aim of this reduction is to make the requirements more achievable for current and future radionavigation systems; currently, continuity is particularly hard to achieve.

2.2 Phases of navigation

(IMO, 2001) gives many different applications, but does not define them. It is common to give different stages of navigation without a strict definition, or any definition at all; in general they are ill-defined. However IALA (2001) gives definitions of some general requirements, as set out below. The main difference in the scenarios is the navigation environment of the ship: how deep the water is, how close it is to obstructions and what navigational aids may be used. It is likely that the environment from the point of view of a GPS receiver will remain relatively similar throughout the stages, until it is very close to other ships and shore infrastructure.

2.2.1 Ocean navigation

- Beyond the continental shelf and more than 50 nm from land;
- In waters where position fixing by visual reference to land or fixed or floating aids to navigation is not possible;

- Sufficiently far from land masses such that hazards from shallow water or collisions are small.

2.2.2 Coastal navigation

- Within 50 nm from shore or on the limit of the continental shelf;
- In waters contiguous to major land masses or islands groups where sea lanes converge;
- Within range coastal vessel traffic systems;
- Increased activity from other vessels;
- Feasible to navigate by visual observation of shore or radar.

2.2.3 Harbour approach

- Transition from coastal navigation to harbour navigation;
- Ship moves to more restricted waters;
- Requires more frequent position fixing and increased possibility for collision with other vessels;
- Within coverage areas of aids to navigation, pilotage areas and vessel traffic services.

2.2.4 Restricted waters / harbour navigation

- Restricted room to manoeuvre;
- Relatively high probability of collision;
- Requires accurate position almost continuously;
- Potential grounding risk.

2.3 Requirement specification

The IMO requirements are specified in terms of accurac1y, integrity, continuity and availability, as described below. Table 2.1 shows a few of the more general requirements that are examined in this thesis.

2.3.1 Accuracy

> "The degree of conformance between the estimated or measured parameter of a craft at a given time and its true parameter at that time." (IMO, 2001)

The IMO requirements do not place a probability on the accuracy value, but for the purpose of this study it is assumed to be 2σ, as used in the ICAO requirements and suggested in other IMO documents.

Table 2.1: IMO requirements for a future GNSS

	Accuracy	Integrity			Availability	Continuity
	Horizontal (m)	Alert limit (m)	Time to alarm (s)	Integrity risk (per 3 hours)	% per 30 days	% over 3 hours
Ocean and coastal	10	25	10	10^{-5}	99·8	N/A
Port approach	10	25	10	10^{-5}	99·8	99.97
Port	1	2·5	10	10^{-5}	99·8	99.97
Automatic docking	0·1	0·25	10	10^{-5}	99·8	99.97

2.3.2 Integrity

> "The ability to provide users with warnings within a specified time when the system should not be used for navigation." (IMO, 2001)

There are three values here: the alert limit (threshold value), the time to alarm, and the integrity risk. The integrity risk is defined as:

> "The probability that a user will experience a position error larger than the threshold value without an alarm being raised within the specified time to alarm at any instant of time at any location in the coverage area." (IMO, 2001)

This requirement is well defined, but in practice requires knowledge of the probability of occurrence of an error of a magnitude sufficient to cause a position error larger than the threshold value. It is hard to determine this due to the rarity of such events.

2.3.3 Continuity

> "The probability that, assuming a fault-free receiver, a user will be able to determine a position with specified accuracy and is able to monitor the integrity of the determined position over the (short) time interval applicable for a particular operation within a limited part of the coverage area." (IMO, 2001)

The IMO requirements are specified over 3 hours. During this time period, 95% of positions must be less than the accuracy requirement away from the truth. Although it is not explicitly stated in the IMO requirements, in this project it is assumed that if the accuracy is sufficient, the continuity is only broken if a period greater than the integrity time-to-alarm passes during which there is either no position or no integrity monitoring available. This is justified because the integrity requirement allows the position to be an unlimited distance from truth for this period, and it is preferable to have no position output than an erroneous one.

2.3.4 Availability

> "The percentage of time that an aid, or system of aids, is performing a required function under stated conditions. Non-availability can be caused by scheduled and / or unscheduled interruptions." (IMO, 2001)

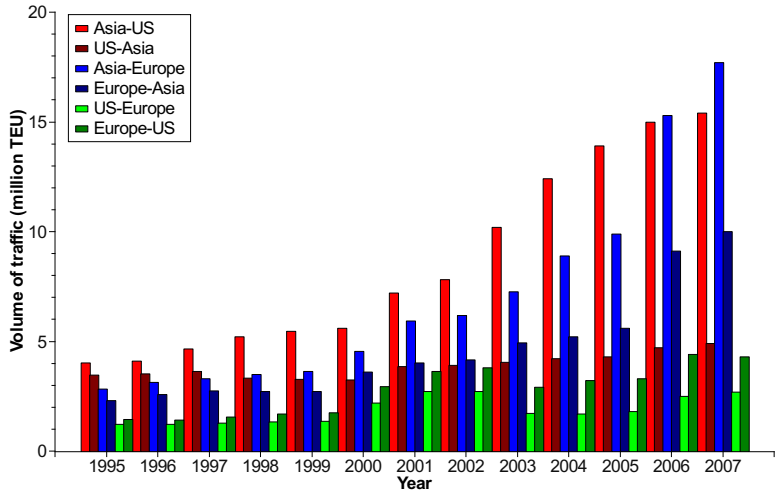


Figure 2.1: Cargo volumes along major trade routes, 1995 - 2007 (UNCTAD, various issues)

This is similar to the continuity requirement, but is is generally defined over a much longer time period, with a lower probability requirement.

2.4 Shipping background

This section gives an overview of recent trends in world seaborne trade and in ship design. These trends have lead to the need to development of the concept of eNavigation, described in Section 2.4.3, to improve navigational safety and efficiency.

2.4.1 Development of seaborne trade

Total maritime commercial activity increased by 63% in the 12 years to 2007, reaching $3{\cdot}3 \times 10^9$ tonne-miles (UNCTAD, various issues). This growth has been fuelled by an increase in Asian exports, with the proportion of goods loaded in Asian ports rising from 27% to 38% during this period. Figure 2.1 shows the volume of cargo carried along the major containerised trade routes and illustrates the rapid growth of exports from Asia: the volume of Asian exports has increased by 384% from 1995 to 2006, compared to an increase of 160% for all other routes.

The world merchant fleet is expanding: capacity has increased by 56% since 1995 to $1{\cdot}124$ billion deadweight tonnes (dwt) at the beginning of 2008. Figure 2.2 shows how this increase is divided amongst the major types of ship. Although the capacity of both the bulk carrier and tanker fleets has increased, the container fleet has expanded at an even greater rate, as container ships have taken over trade from other ship types. In 2007 the capacity of the international container fleet was predicted to grow by 45% by 2010, with most of this capacity concentrated in larger ships (House of Commons Transport Committee, 2007): the current economic downturn may extend this date, but the trend is clear.

The use of containers is a relatively new phenomenon, with the first dedicated ship built in the 1960s. Throughout the 1960s and 1970s containerisation took over all the major trade routes as the versatility of this mode of transport reduced loading times and shipping costs (Cullinane

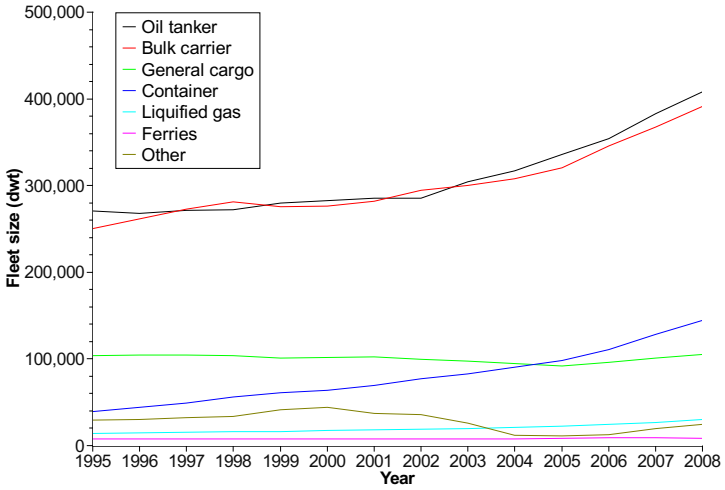


Figure 2.2: Composition of the world fleet 1995 - 2007 (UNCTAD, various issues)

& Khanna, 2000). This sector continues to increase its market share as a greater variety of goods are switched to containerised transport. One example of this is the transport of cars: this is already done on a small scale, and some car manufacturers are studying the potential for transporting four cars in a 45-foot container rather than traditional roll-on/roll-off car carrier (MAN Diesel, 2004).

The increase in traffic means that ports and shipping lanes are becoming ever busier. Global container port throughput has increased by 260% since 1995 (UNCTAD, *various issues*). Much of this increase has been concentrated in the largest ports due to the emergence of the "hub-feeder" concept. Larger ships provide economies of scale, as described in Section 2.4.2. The limitations to this economy are inflexibility (large ships are restricted to large ports) and time in port (cargo handling time increases with ship size). The first problem is overcome by global shipping alliances that can guarantee cargo for their large ships: in 2000 the top 20 operators controlled 98% of ships over 3500 TEU[1] (Cullinane & Khanna, 2000). The second is overcome by reducing the number of ports at which a ship calls. Therefore efficient long distance trade is carried out between large regional hubs, with cargo subsequently transshipped to feeder ports.

The UK follows the global trends. Annual tonnage handled by UK ports grew by around 30% between 1980 and 2000, with significantly stronger growth in imports than exports. Much of this growth is in containerised traffic, which increased by 112% between 1990 and 2006 (Department for Transport, 2006). Forecast growth rates to 2030 are in the region of 1% p.a. for all traffic, with containerised growth at 4% p.a. (House of Commons Transport Committee, 2007). The hub-feeder concept has influenced the development of UK ports: 95% of calls by container ships of over $1,500$ TEU were received at Felixstowe, Southampton, Medway and London (Department for Transport, 2006). This has generated a strong demand for transhipment services both from UK ports and European hubs such as Rotterdam (House of Commons Transport Committee, 2007).

[1] Twenty foot Equivalent Unit (TEU), the size of a small container, is commonly used to measure the capacity of containerised cargo vessels

Table 2.2: Current and predicted size of large ships (Buxton, 2004)

Ship Type	Typical Size (2004)	Growth Rate p.a. since 1975	Future Growth Rate (p.a.)	Typical Size (2020)
Large Bulk Carrier	180,000 dwt	$+2\%$	$+0{\cdot}5\% - +1\%$	210,000 dwt
Large Tanker	320,000 dwt	$+1\%$ (since 1985)	$+1\% - +2\%$	400,000 dwt
Large Container Ship	8,000 TEU	$+5\%$	$0\% - +5\%$	10,000 TEU – 20,000 TEU

The current economic downturn has, however, had a severe impact on global trade and hence the shipping business. The preceding boom resulted in many ships under construction at the start of the downturn, which has seen a drop in charter rates which for a 150,000 dwt Cape-class ship have dropped from $300,000 per day in autumn 2008 to $15,000 per day in March 2009 (The Economist, 2009). This is due to the current over-capacity, which has seen over 10% of the global capacity of container ships anchored empty outside the harbours of South-East Asia. This has led to the scrapping of many older ships, and will affect the size of the global fleet for may years to come.

2.4.2 Ship design trends

In general there is a cost reduction per TEU for operating a larger ship, so a shipping operator can derive a competitive advantage by acquiring larger ships (Cullinane & Khanna, 2000; MAN Diesel, 2004; Buxton, 2004). Buxton (2004) describes the past trends in ship sizes and predicts how this might continue: these trends are summarised in Table 2.2. Bulker and tanker technology is mature and the markets stable, so recent growth has been slow and steady. The size of a typical large bulk carrier has increased by about 2% p.a. since 1975, but this could slow as technical limits on hull structure or cargo operations are reached. Tanker sizes were dramatically reduced following the 1980s oil slump, when ships of up to $500,000$ dwt were shown to be too inflexible in the face of changing economic conditions (Buxton, 2004). However, tanker sizes have increased steadily since 1985, and this trend looks likely to continue into the future. Container ship sizes have shown a dramatic increase in the last 30 years, spurred by rapid trade development and increasing market share. There are currently twelve ships afloat with an official capacity of more than $10,000$ TEU, including eight ships owned and operated by Mærsk with a capacity of $12,508$ TEU. Studies suggest that "Malacca-max" ships of up to $18,000$ TEU are technically possible, this being the largest ship size able to pass through the Straits of Malacca (the second busiest shipping lane in the world after the Straits of Dover). However, an increase in the size of the largest container ships does not necessarily reduce demand for smaller ones (MAN Diesel, 2004). Indeed, greater economies may be achieved by increasing the size of smaller ships (Stopford, 2002).

Larger ships require more time loading and unloading in port. To compensate for this they often sail at a higher operational speed (MAN Diesel, 2004), as shown in Figure 2.3. While the smallest vessels might have an operational speed in the region of 19 knots, larger vessels travel at over 25 knots. This increased speed reduces time available for avoidance manoeuvres (Lamb

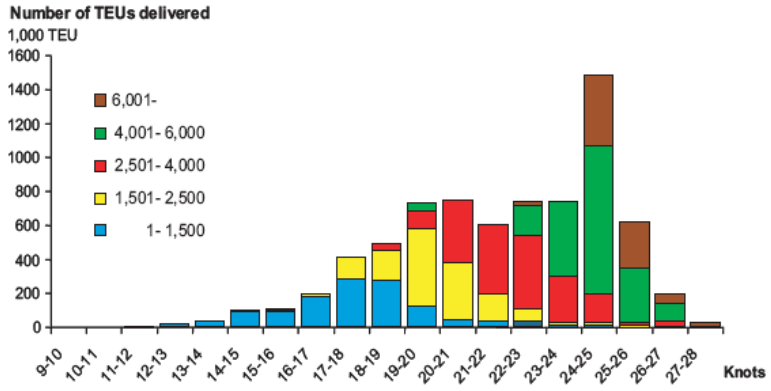


Figure 2.3: Average speed of container ships, 1960 - 2004, by ship size (MAN Diesel, 2004)

& Hunt, 2000) and increases the stress on the crew (Inoue, 2000), increasing both the likelihood and potential severity of an accident. Increased automation onboard ship has led to a reduction in crew size, which can result in increased crew fatigue and poor watch-keeping. Improved navigation technology is being developed to compensate for these problems and hence improve safety.

2.4.3 GNSS to assist navigation

In order to cope with these rapid increases in the volume of shipping and average size and speed of ships, the IMO, supported by IALA, is developing an "eNavigation" strategy to promote, coordinate, regulate and standardise the use of modern technologies for the improvement of navigational safety:

> "eNavigation is the harmonised collection, integration, exchange, presentation and analysis of maritime information onboard and ashore by electronic means to enhance berth-to-berth navigation and related services, for safety and security at sea and protection of the marine environment." (IMO, 2008)

Many of the technologies required for this are already available and in use. Ship-borne Automatic Identification Systems (AIS) transmit positional information to other ships and to shore stations; Electronic Chart Display and Information Systems (ECDIS) allow this information to be displayed on up-to-date electronic charts; and Vessel Traffic Services (VTS) allow port and harbour authorities to monitor vessel movements. IALA (2008) provides more information about the need for eNavigation and its development.

eNavigation can be expected to bring about substantial benefits to marine safety and efficiency. However, the system cannot work without a precise and robust positioning system. Although back-up systems such as eLoran, which allows positioning at the 10 m level by ranging from shore-based reference stations, might be used, the primary positioning system will be GNSS (de Halpert *et al.*, 2006). Therefore it is important that modernised GNSS can be shown to meet the positioning requirements for different scenarios, as discussed in Section 2.3.

2.5 Summary

The rapid increase in the size and quantity of ships in recent years has created a strong need for techniques to improve navigational safety. eNavigation is an initiative supported by the IMO that combines many modern techniques to enable the harmonised collection, integration, exchange, presentation and analysis of maritime information to improve the safety and efficiency of marine navigation. This relies upon GNSS as the primary positioning technology, and so it is therefore important that the positioning performance is appropriate for the task. The IMO has therefore developed a set of requirements for a future GNSS in different scenarios, in terms of accuracy, integrity, continuity and availability.

Chapter 3

Modernised GNSS

Currently, GPS broadcasts on two frequencies, L1 and L2, of which L2 is not designed for civilian use. The US is currently in the process of modernising GPS: there are seven satellites broadcasting a new civil signal on L2, called L2C, and one of these satellites is broadcasting a test signal on a new frequency, L5. The EU is also developing a similar system that will also broadcast on 3 frequencies, of which L1 and L5 are the same as GPS, which will allow easy interoperability. Russia is modernising its GLONASS positioning system and China is launching its own COMPASS positioning system, both of which may eventually be compatible with GPS and Galileo to some extent.

In order to determine how positioning with modernised GNSS can meet the IMO requirements, it is necessary to study the signal structure of current GPS, the improvements to this signal structure that will be provided by the various modernised signals, and the effect that these improvements will have on the positioning performance of a modernised GNSS receiver. This will provide the background necessary to predict the positioning performance of modernised GNSS for marine applications based on the performance of current GPS in this environment.

Section 3.1 provides an overview of the structure of current GPS signals and the general technique used to obtain range measurements from the GNSS signals; Section 3.2 describes the modernised GPS signal structure and Section 3.3 describes the Galileo signal structure. Section 3.4 outlines the details of the future improvements in signal structure, Section 3.5 gives an overview of other future GNSSs, and Section 3.6 discusses the compatibility and interoperability of the new signals. Section 3.7 assesses the effect of the improvements of modernised GNSS on aspects of positioning, which will later be used to interpret the results of the data collection exercise.

3.1 Overview of GPS

The structure of the GPS signals is defined in IS-GPS-200D and described in detail in Spilker & Van Dierendonck (1999) and Spilker (1980). This section gives an overview of the GPS signal structure, with the aim of facilitating subsequent explanation of the improved characteristics of the new signals.

The GPS constellation nominally consists of 24 satellites arranged in 6 evenly-spaced orbital planes inclined at 55° with respect to the equator. The satellite orbits are circular, with a radius of 26,660 km and an orbital period of 11 hours 58 minutes. GPS satellites have a very stable clock with a fundamental frequency $f_0 = 10{\cdot}23$ MHz, which is used to coherently generate the

transmitted signals. The current GPS satellites transmit on two carriers, L1 and L2, with central frequencies:

$$\begin{aligned} f_{L1} &= 154 \times f_0 = 1575{\cdot}42 \text{ MHz} \\ f_{L2} &= 120 \times f_0 = 1227{\cdot}60 \text{ MHz} \end{aligned} \tag{3.1}$$

Each signal is composed of three parts: a sinusoidal carrier wave at the central frequency modulated by a binary code and a data message. The binary code allows the receiver to determine the signal travel time, which is the basis of the GPS positioning technique. The data message provides additional information such as satellite ephemerides and clock offsets, UTC-GPST offset, ionospheric model and satellite health data. It takes 30 seconds to receive all the data necessary for navigation and $12{\cdot}5$ minutes to receive the entire message.

GPS is based on *direct sequence-spread spectrum* (DS-SS) signalling. An information waveform of low bandwidth (the data message) is modulated by a waveform of high bandwidth (the code) to produce a high-bandwidth signal for transmission. The data message requires a bandwidth of 100 Hz; the spread GPS signal is transmitted with a bandwidth of 20 MHz. Bandwidth is proportional to the chip (bit) rate of the waveform, so the speed of the code must be much greater than the data message; the data message is transmitted at 50 bps (bits per second), whereas the code on L1 has a rate of $1{\cdot}023$ cps (chips per second). The bits of the code are termed *chips* because they do not carry information.

Two types of spreading codes are transmitted: C/A (Coarse / Acquisition) and P (Precision). These are both pseudorandom noise (PRN) codes, i.e. they are designed to look random when taken over their whole period. Each GPS satellite transmits a different version of these codes, which allows the receiver to distinguish between transmissions from different satellites. This technique is called *code division multiple access* (CDMA).

The C/A-codes are broadcast on L1 and are freely available to civilian users. They are drawn from the family of Gold codes, which have low cross-correlation between members, and are very short at 1023 chips: these properties make it easy for a receiver to acquire the signal. The rate of transmission is $1{\cdot}023$ cps, giving a 1 ms period and 293 m wavelength.

The P-code is modulated on both L1 and L2 but is usually encrypted to form the P(Y) code, which is only available to military users. Since the C/A-code is not modulated on L2 it is not possible to obtain a code measurement directly on this frequency. There are various techniques available to allow measurement to be taken on L2 without knowledge of the encryption code, with reduced measurement precision and robustness (see e.g. Tsui (2000)). The complete P-code has a 38-week period; each satellite is assigned a unique week-long segment. The rate of transmission is $10{\cdot}23$ cps, giving a wavelength of $29{\cdot}3$ m.

The code and data signals are modulated on the carrier using *binary phase shift keying* (BPSK). The binary 0's and 1's of the code are represented by multiplying the carrier wave by $+1$ or -1 respectively, which is the equivalent of leaving the phase unchanged or changing it by $180°$, as shown in Figure 3.1. The code and data are modulated together on the same carrier; within one data bit there are exactly 20 C/A-code periods. The data and code bit transitions coincide because the chipping rates of both of these binary codes are coherently related to the same fundamental clock frequency.

L1 is modulated by both the C/A-code for civilian users and the P-code for military use.

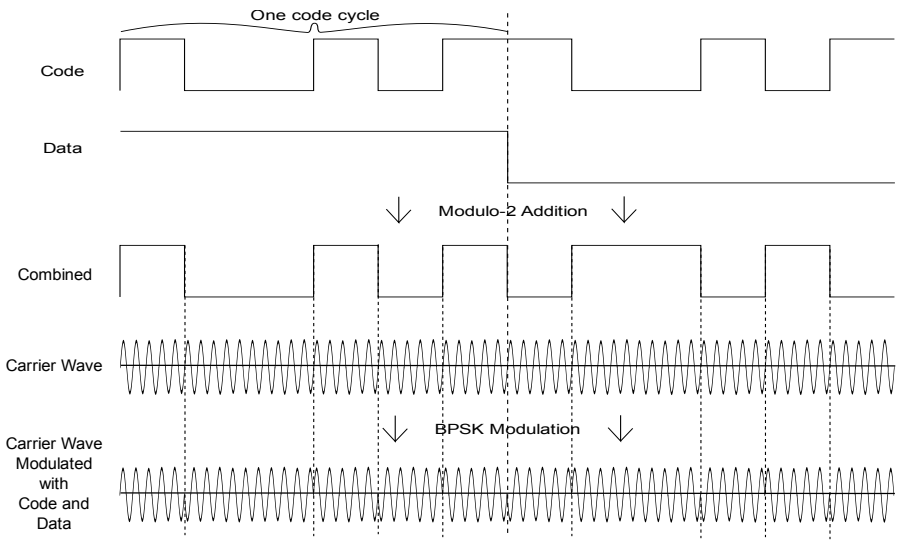


Figure 3.1: Binary phase shift keying modulation of code and data signals

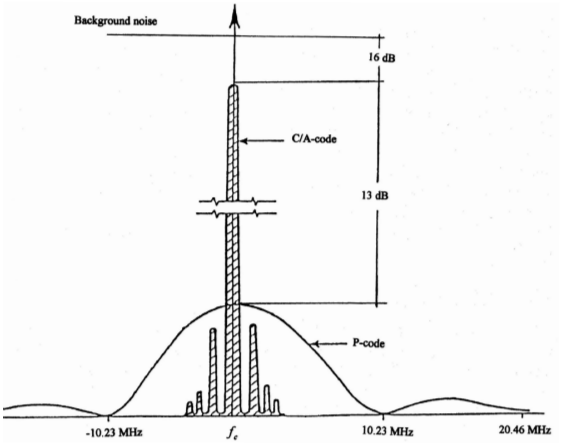


Figure 3.2: GPS power spectral density (Spilker, 1980)

These two codes are combined on the same carrier using *quadrature phase shift keying* (QPSK). The in-phase component of the carrier is BPSK modulated with the P-code and the quadrature component (90° out of phase) is modulated with the C/A-code.

The shape of these BPSK signals in the frequency domain can be described by the function $(\frac{\sin x}{x})^2$, referred to as $\text{sinc}^2(x)$. This function is symmetric about the central carrier frequency and zero at multiples of the chipping rate, as shown in Figure 3.2. The width of the spectrum is proportional to the chipping rate, as mentioned above. Due to the periodic nature of the modulating codes, this spectral power density is not continuous over the bandwidth, but rather is a line spectrum. The line spacing is $\frac{1}{NT_c}$ Hz, where N is the length of the code and T_c is the chip period. Thus, for the C/A-code the line spectrum spacing is 1 KHz, while the lines in the P-code spectrum are extremely dense and essentially continuous. The power at each line deviates slightly from the $\text{sinc}^2(x)$ envelope; the C/A-codes are particularly susceptible to narrowband interference at the frequencies where its line spectrum is the strongest.

The part of the spectrum outside the 20 MHz main lobe of the P-code is filtered out before transmission. Note that the peak in the C/A-code spectrum is 16 dB below the thermal noise

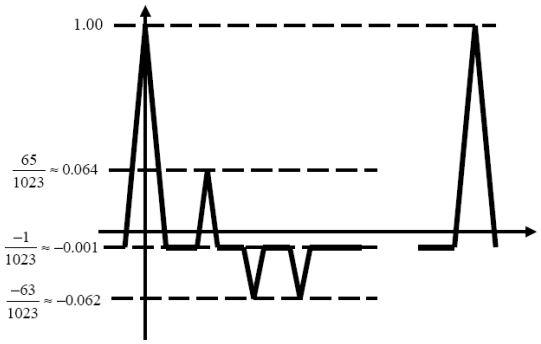


Figure 3.3: Autocorrelation of GPS C/A-code (Enge, 2003)

level, so GPS signals will not be visible if the spectrum is analysed with an oscilloscope. The means that they will not interfere with other systems at a similar frequency, such as line-of-sight microwave antennas.

When the signal reaches the GPS receiver it is modulo-2 added to a copy of the code which is generated by the receiver; this reforms the original signal (the data message). This is the essence of spread-spectrum signalling and greatly enhances robustness in the presence of narrowband interference. The *processing gain* of the system is the parameter $\frac{f_d}{f_c}$ that determines the proportion of narrowband interference that passes through to the reformed signal, where f_d is the frequency of the data message and f_c is the frequency of the spreading code.

The receiver generates a reference carrier and modulates it with the known PRN code. It then delays the reference signal until it matches up with the received signal: this time delay is the signal travel time used for positioning. The required time offset is computed by *autocorrelation*. For a given time offset Δt, received signal $s(n)$ and offset generated signal $s(n + \Delta t)$, the autocorrelation function is:

$$R(i) = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} s(t)s(t+\Delta t)dt = \begin{cases} 1 & \text{if } \Delta t = 0 \\ 1 - \frac{|\Delta t|}{T} & \text{if } |\Delta t| \leq T \\ \approx 0 & \text{if } |\Delta t| > T \end{cases} \tag{3.2}$$

where T is the chip length and τ the integration time. Longer integration times increase the probability of correctly determining the time offset, because the signal is accumulated but the random noise cancels out. The maximum possible integration time is limited by the period of the data bit, because the sign of the code bits is reversed at a data bit transition (see Figure 3.1). There are 20 C/A-code cycles for every data bit, so it is guaranteed that one of two contiguous sections of 10 code lengths will not contain a data bit transition. Therefore the presence of the data signal restricts the maximum integration time on the C/A-code to 10 ms.

Once the data message has been decoded it can be removed by modulo-2 addition in the same way as the spreading code, in a technique called *data-wiping*. This allows longer integration times until the message changes; nevertheless, the data message makes acquisition of the signal more difficult.

The autocorrelation function of the C/A-code is shown in Figure 3.3. The smaller peaks are called side-lobes, and the ratio of the height of the main peak to the height of the largest side-

lobe is a characteristic of the code which determines how robust the code is against incorrect correlations. In practice the peak will not be exactly 1, due to noise: how close it is depends on the signal to noise ratio, S/N_0, the ratio of the received GPS signal to the noise floor. The receiver adjusts the time offset Δt until $R(i)$ reaches this maximum. There are two corollaries to this process: firstly, longer codes are more robust against noise, because the random bits average to zero more reliably, and secondly a shorter chip duration will provide a sharper autocorrelation peak because the base of the autocorrelation triangle has a width of two chips. This will result in a more precise measurement.

The P-code is much longer than the C/A-code, which makes it very difficult to acquire directly, because there are many possible time offsets to evaluate. In the original GPS design it was intended that a receiver would initially acquire the C/A-code to determine the time offset and read the data message. One of the components of the data signal is the *hand over word*, which tells the receiver where in the week-long P-code to look. Once this is known, it is possible to lock on to the P-code and generate a more precise position using the longer, faster code.

Once the correct value of Δt has been determined, the generated PRN code is modulo-2 added to the received signal; the two signals are in phase and so cancel out. This collapses the waveform to a narrow-bandwidth signal, leaving the carrier modulated by the data message. This is then passed to a *phase-locked loop* (PLL), which generates a reference carrier at the same frequency and phase as the input signal. The PLL allows the receiver to measure the fractional carrier phase, which is the basis of RTK GPS positioning, and to recover the data message.

3.2 Modernised GPS

The first available modernised signal is a civil code on L2, currently transmitted by seven Block IIR-M satellites, with two more still to be launched. There will be a new encrypted code, the M-code, on L1 and L2 in addition to the P(Y)-code. There is currently one Block IIR-M transmitting a test signal signal on a new frequency, L5. The next block of satellites, Block IIF, will transmit the L2C and L5 codes; the first is expected to be launched in 2010. The final stage in GPS modernisation is the Block III satellites which will feature a new code on L1 and increased transmission power: these are anticipated in 2014 at the earliest. This section describes the new civil codes that will be transmitted by the modernised satellites.

3.2.1 L2

The first available modernised GPS signal is a civil signal on L2, called L2C. There are currently 6 satellites broadcasting L2C, with more being launched as old satellites are decommissioned. The structure of this new signal is specified in IS-GPS-200D. Fontana *et al.* (2001a,b) describe the design process and expected advantages, as summarised below. In addition to the new civil signal, the military P(Y)-code will be retained and the new military M-code will also be transmitted.

There will be two new sets of PRN codes, time division multiplexed (TDM) chip-by-chip on the quadrature component of the carrier. The first code, Civil Moderate (CM), has a length of 10,230 chips, while the second, Civil Long (CL) has a length of 767,250 chips. The in-

creased length over the C/A-code improves cross-correlation performance: the worst-case cross-correlation performance of the L2C codes is 45 dB, compared to 21 dB for C/A codes.

Both codes are of even length and perfectly balanced, i.e. sum to exactly zero. This allows the receiver to track each code separately by assuming the bits of the other code average to zero. The chip length is the same for each code, so the signal power is evenly divided; tracking only one code gives a 3 dB power reduction. The minimum received power from the from IIR-M and IIF satellites will be -160 dBW, with the civil signal $0{\cdot}4$ dB weaker than the military signal. Overall, L2C will be $2{\cdot}3$ dB weaker than the C/A-code on L1.

The chipping rate and modulation type of the L2C codes is the same as the C/A-code, producing a similar frequency power spectrum, as shown in Figure 3.4. The L2C codes are much longer than the C/A-codes, so the maximum lines have less power, which greatly increases robustness to narrow band interference.

The CM code is modulated by the new data message, CNAV, at a rate of 25 bps. FEC is applied to give a total transmitted bit rate of 50 cps. Compared to the conventional data message, FEC increases the data recovery threshold by 5 dB, the halved data transmission speed improves it by 3 dB and the half power (only modulated on one code) decreases it by 3 dB, for an overall improvement of 5 dB.

The CL code is a pilot component, which provides a 6 dB improvement to phase tracking threshold. The half power to the dataless component gives -3 dB, for a net performance increase of 3 dB when tracking the phase of the pilot component.

Each code is transmitted at a rate of $511{\cdot}5$ kcps, so the overall speed of the complete TDM L2C signal is the same as the C/A-code at $1{\cdot}023$ Mcps. The clock rate was limited to maintain spectral separation between the civil code and the new military M-code. The CM code has a duration of 20 ms and the CL code $1{\cdot}5$ s, so there are 75 repetitions of the CM code within every CL code cycle.

The length of the CL code is such that it would be very difficult to acquire it directly, so in normal operation the CM code would be acquired using a Costas loop. The receiver could then rapidly search the 75 possible time offsets between the two codes and lock on to the CL code. However, in a difficult environment where acquisition of the CM code is impossible, the CL code could be acquired directly due to its increased length and lack of data modulation. Once the receiver has acquired the CL code from one satellite, the range of possible offsets of this code from the other satellites is limited by the difference in signal travel time from the zenith to the horizon, around $18{\cdot}7$ ms. This gives a search range of around 19,130 chips, which is about twice the length of the search range for the CM code. It is therefore much easier to acquire the CL codes from subsequent satellites directly. There is little advantage to be gained by tracking both CM and CL codes (Fontana *et al.*, 2001b).

3.2.2 L5

The second modernised GPS signal that will be available is a completely new signal at L5, with a central frequency of $1176{\cdot}45$ MHz. This is within a protected Aeronautical Radionavigation Service (ARNS) band and is intended for use as a safety-of-life system. This frequency band is also used by aircraft navigation services, such as DME, TACAN and JTIDS, so the noise

floor will frequently be higher than the thermal noise. This is compensated for by the high received power of the L5 signal of -154 dBW. The signal structure is specified in IS-GPS-705 and described in Spilker & Van Dierendonck (1999); it is summarised below. The characteristics of this signal are very similar to those of Galileo E5a, with the same code length, code rate, chipping rate and use of tiered codes.

The signal is composed of two QPSK-modulated components with 50% power to each. These components are modulated by unique PRN codes which are broadcast at a rate of $10{\cdot}23$ Mcps, ten times the rate of the L1 C/A and L2C codes. This means that the null-to-null transmitted bandwidth is also ten times greater at 24 MHz, which improves tracking precision and robustness to multipath and RFI, at the expense of increased receiver cost and power usage.

The in-phase component is modulated by a PRN code of period 10,230 chips, so the duration is the same as the C/A-code at 1 ms. It is additionally overlaid by a 10-chip long Neumann-Hoffman (NH) code of duration 10 ms, where each chip of the NH code multiplies one whole length of the PRN code. This overlay code improves the cross-correlation properties of the combined code, but the tiered nature allows the receiver to acquire the shorter code before switching to the combined code for tracking. The PRN code is ten times the length of the C/A-code, but the better cross-correlation properties give a four-fold decrease in dwell-time at each potential offset, which mitigates the increase in acquisition time. The NH code also has the effect of reducing the spacing between the lines in the signal power spectrum from 1 kHz to 100 Hz, which reduces the effect of narrow band RFI. The NH codes also help to increase the robustness of the data bit synchronisation, because they are synchronised with the navigation bits. This component of the signal is also modulated by a data channel, at a rate of 50 bps. FEC($1/2$) is applied, so the data message is transmitted at 100 cps. The FEC provides a 5 dB improvement in the data demodulation threshold.

The quadrature component consists of a different PRN code of the same 10,230 chip length modulated by a 20-chip Neumann-Hoffman sequence. This is a pilot component, so is not modulated by a data message. The line spacing in the signal power spectrum for this component is even less at 50 Hz, which further improves the resistance to narrowband RFI.

Both of these codes have very low cross-correlation: the quadrature codes are -57 dB or less. Because the codes on each component of the carrier are different, there is a 2 dB improvement for tracking them both at the same time.

3.2.3 L1

The third modernised GPS signal to be available will be an additional modernised signal on L1. The L1 band is centred at $1575{\cdot}42$ MHz, which has the advantages of being the GNSS frequency least affected by ionospheric refraction and also within an ARNS protected band. On 26th June 2004, The US and EU signed the "Agreement on the promotion, provision and use of Galileo and GPS satellite-based navigation systems and related applications" to ensure the compatibility and interoperability of GPS and Galileo. Part of the agreement was that each system would adopt a signal on L1 with an identical power spectral density (PSD) when computed using all the components of the signal. Initially it was intended to use a BOC(1,1) modulation scheme, but a joint research group has subsequently identified and recommended another can-

didate modulation scheme, Multiplexed BOC (MBOC), with superior properties as outlined in Betz *et al.* (2006b,a). MBOC is defined in the frequency domain as the sum of $10/11$ of the normalised BOC(1,1) PSD and $1/11$ of the normalised BOC(6,1) PSD (Julien *et al.*, 2007); placing some of the signal power into a higher-frequency code improves signal tracking performance and multipath mitigation. This frequency-domain definition of the shared signal allows different implementations in the time domain: L1C will use Time Multiplexed BOC (TMBOC) to achieve the required PSD.

The C/A-code will continue to be broadcast on L1 for backward compatibility, so the GPS III satellites will modulate the C/A, L1C, P and M codes on the same carrier. The technique for doing this has not yet been determined, but will be flexible, with the phase relationship contained in the broadcast navigation message.

The L1C signal structure is specified in IS-GPS-800. The signal is split into two components, with 75% of the available power going to the pilot and 25% to the data channel; this ratio was determined after consultation with industry and experts (Hudnut & Titus, 2004). A TMBOC modulation scheme will be used for the pilot channel, where 4 out of every 33 symbols of the BOC(1,1) modulation are replaced by 6 faster symbols from a BOC(6,1) modulation. The data channel will be modulated by BOC(1,1) only; the faster BOC(6,1) chips are not modulated on the data component because data demodulation does not benefit from the higher-frequency contributions of the BOC(6,1) spreading symbols.

The modulating codes have a length of 10,230 chips and are based on Weil sequences; the chipping rate is $1{\cdot}023$ Mcps and the period is 10 ms. The codes with the best correlation properties were allocated the pilot component, as code acquisition and tracking will generally be performed on the pilot signal due to the power division. Each pair of pilot/data codes were chosen to have low correlation at zero offset. Each pilot component is additionally modulated by a unique 1,800 chip secondary code at a rate of 100 cps, with duration 18 seconds. This code is modulo-2 added to the spreading code, with each bit of the overlay code applied to a whole cycle of the spreading code; these codes reduce cross-correlation and aid in synchronisation to the data message boundary (Fontana *et al.*, 2001b).

3.3 Galileo

The Galileo system will consist of 30 satellites in three circular orbital planes with a radius of $29,601{\cdot}297$ km and an inclination of 56° with respect to the equator. There will be freely accessible civil signals overlaid on the GPS L1 and L5 frequencies, with very similar characteristics to the GPS signals, called E1 and E5a. In addition there will be another civil signal, E5b, adjacent to E5a at $1207{\cdot}14$ MHz, and an encrypted commercial signal, L6.

The Galileo system will provide four different services:

- The Open Service (OS) will use E1, E5a and E5b and will provide unencrypted data and ranging signals;
- The Safety of Life service (SoL) will transmit integrity and signal-in-space accuracy data on E1 and E5b;

- The Commercial Service (CS) will use E1, E5b and E6. Additional data will be transmitted to users; this will be encrypted and users will have to pay to access it;
- The Public Regulated Service (PRS) will be used by governments and will be transmitted on L1 and E6.

Galileo is designed to be compatible and interoperable with GPS. This section will describe the signal structure of the open service signals. These are specified in GJU (2006) and described in detail in Hein *et al.* (2001, 2002). There are currently two operational test satellites, GIOVE-A and GIOVE-B. It is anticipated that four In-Orbit-Validation (IOV) satellites will be launched by the end of 2010, and the full system by 2014, although these dates may be pushed back.

3.3.1 E1

The Galileo carrier at L1 is modified-hexaphase modulated by three signals. There is a military signal (L1a) and two civil components: a data component (L1b) and a pilot component (L1c). This section will give an overview of the two civil signals. As discussed in Section 3.2.3, due to the agreement between the EU and the US, Galileo E1 and GPS L1C will both use Multiplexed BOC (MBOC), which is defined in the frequency domain as the sum of $10/11$ of the normalised BOC(1,1) PSD and $1/11$ of the normalised BOC(6,1) PSD.

E1 will use a different modulation scheme to L1C to achieve the MBOC PSD: Composite BOC (CBOC). In this modulation scheme, the BOC(1,1) and BOC(6,1) sub-carriers are linearly combined, and both are present at all times. The use of a different modulation scheme to L1C may complicate receiver design (Julien *et al.*, 2007).

The modulating codes are the same length as L1C at 10,230 chips and the same speed at $1{\cdot}023$ Mcps, but are based on *random* or *memory* codes rather than Weil codes. These codes have been designed to achieve the best properties for a given length, and allow more flexibility than conventional codes. They are not generated onboard the satellite in the manner of conventional codes, but rather are stored in memory chips. This allows more flexibility in code design, and has been made possible because of the falling cost of such chips.

The main difference between the GPS and Galileo L1 signals, apart from the different modulation schemes, stems from the provision of integrity data and encrypted commercial data on the Galileo L1 signal. E1 has a greater proportion of power to the data component (50%, compared to 25% for GPS) and a much higher data rate (250 sps compared to 100 sps). GPS L1C also has a much longer overlay code, and hence superior correlation properties. Most receivers will acquire the E1 carrier using only the pilot component, because the high data rate limits the coherent integration time to only 4 ms on the data signal, compared to 10 ms for the C/A-code.

3.3.2 E5

The Galileo E5 codes are modulated on the carrier using AltBOC modulation, as described in Section 3.4.2. Each of the two side-lobes of the BOC modulation has a different code and so can be tracked separately. These two codes are called E5a and E5b, and taken individually are effectively a BPSK(10)-modulated code, with a primary code length of 10,230 chips and a chipping rate of $10{\cdot}23$ Mcps. These codes are overlaid by a longer secondary code and the power

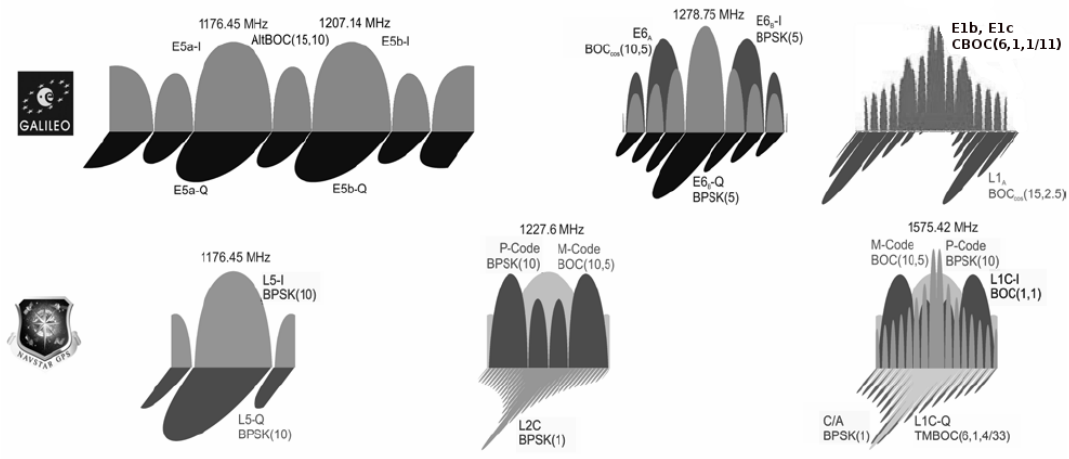


Figure 3.4: Power spectral density of the new signals (Wallner *et al.*, 2007; Avila-Rodríguez *et al.*, 2008)

is evenly divided between pilot and data components in phase quadrature. When tracked individually, they are very similar in structure and characteristics to GPS L5; the central frequency of E5a is identical to L5 and hence is interoperable without the need for additional front-end hardware. E5a supports the OS and transmits the F/NAV basic navigation message, while E5b supports the OS, CS and SoL service and transmits the I/NAV integrity message and encrypted commercial data; the E5b data message is transmitted at 250 cps, compared to 50 cps for E5a. Both of these signals are in the ARNS band, so will be suitable for use in safety-critical applications in conjunction with E1. The greatest advantage of Galileo E5 is realised when E5a and E5b are tracked coherently as a single wide-bandwidth signal. The wide bandwidth produces low code noise and has good potential for advanced multipath mitigation techniques (Simsky *et al.*, 2006a,b).

3.4 Technical improvements

This section describes some of the common design improvements of the new signals. Section 3.4.1 describes the characteristics of the new spreading codes, Section 3.4.2 describes alternative code modulation techniques, Section 3.4.3 describes the benefits of the new pilot signals and Section 3.4.4 describes the structure of the new navigation messages. In addition to these benefits, many of the new signals are transmitted with significantly higher power, which makes the signals easier to distinguish from the background noise and more resistant to interference.

This power spectral densities of the new signals are shown in Figure 3.4, and an overview of the structure of the new signals is given in Table A.1 and Table A.2 in Appendix A. For more information on the benefits of the improved signal structures see Enge (2003); Kaplan & Hegarty (2005); Parkinson & Spilker (1996).

3.4.1 Code characteristics

The worst-case cross-correlation of the C/A-code is $-21{\cdot}1$ dB, which has been deemed unacceptably low (Fontana *et al.*, 2001b). The modernised signals have improved upon this through the use of longer spreading codes, made feasible through advances in receiver technology. Modern receivers have many more correlators and can search a greater range of time offsets between the reference and received code rapidly; with few correlators it would take too long to acquire the modernised codes. As well as superior correlation, which improves sensitivity, longer codes produce a frequency spectrum with a greater number of more closely-spaced lines (see Section 3.1). The maximum lines in the spectrum have proportionally less power, which significantly reduces the vulnerability to narrowband RFI.

Tiered codes are used, which have the characteristics of a short code for fast acquisition and a long code for improved tracking and for acquisition in a low S/N_0 environment. Tiered codes generally consist of a medium-length fast primary code modulated by a slow, short secondary code. Each chip of the secondary code is multiplied by one whole length of the primary code. The composite code then has a duration equivalent to the long code and a chipping rate equivalent to the short code. Acquisition can be performed on the shorter primary code so that acquisition time is not increased, although this technique does restrict the integration time in the same way as the data message (Mattos, 2006). Once the primary code has been acquired there are relatively few possible time offsets to search in order to acquire the composite code, because the secondary code is short. Improved performance can then be obtained by tracking the combined code. In low S/N_0 environments, acquisition can be performed over the period of the full composite code, improving sensitivity at the cost of acquisition time. An additional advantage is derived when the secondary code on a pilot component is synchronised with the data message on the data component: the secondary code transitions can assist in data message recovery because the locations of the data bits are known.

Many of the new codes are transmitted with higher chipping rates than the C/A-code on L1. The base of the autocorrelation triangle of a BPSK-modulated signal (see Figure 3.3) has a width of two chips; increasing the chip rate reduces the chip length and so results in a sharper main peak in the autocorrelation function. This allows more accurate correlation of the code, which improves code precision. It also improves multipath mitigation: sharper peaks in the autocorrelation function reduce the distance that a signal needs to be delayed by in order to yield a distinct autocorrelation peak, which can be ignored.

A key function for the performance of a signal is its Gabor bandwidth, $\overline{\Delta\omega^2}$, also referred to as root-mean-square bandwidth. It is defined as the second derivative of the normalised autocorrelation function at the origin and is a measure of the spread of signal power within the bandwidth:

$$\overline{\Delta\omega^2} = (2\pi)^2 \frac{\int_{-\infty}^{\infty} f^2 |S(f)|^2 dt}{\int_{-\infty}^{\infty} |S(f)|^2 dt} \tag{3.3}$$

where f is the frequency and $S(f)$ is the normalised power spectral density of the spreading code. A code with a large Gabor bandwidth generally has a sharper autocorrelation peak, so offers a more precise code measurement and is more resistant to multipath. Gabor bandwidth is related to the code measurement precision by the Cramér-Rao lower bound, σ_ϵ^2, which is the

theoretical lower bound for the tracking variance, assuming perfect receiver implementation:

$$\sigma_\epsilon^2 = \frac{1}{2S/N_0 \overline{\Delta\omega^2}} \tag{3.4}$$

Equation 3.4 shows that increasing the Gabor bandwidth reduces the code noise; the f^2 term in Equation 3.3 implies that Gabor bandwidth may be increased by concentrating spectral power towards the edges of the frequency band. Null-to-null transmitted bandwidth is proportional to code chipping rate (the nulls is lie at $\pm$ the chipping rate from the central frequency), so faster codes spread the spectral power more and hence increase the precision. Improvements can also be obtained by using more sophisticated spreading codes that concentrate the power towards the edge of the bandwidth, as detailed in Section 3.4.2. For a more detailed description of Gabor bandwidth, see Spilker & Van Dierendonck (1999).

3.4.2 Modulation techniques

The GPS C/A-code is modulated on the L1 carrier using BPSK, as described in Section 3.1. The carrier wave is multiplied by a square subcarrier of frequency $n \times f_0$, where the fundamental satellite clock frequency $f_0 = 1{\cdot}023$ MHz, and the modulation scheme is termed BPSK(n). The C/A-code is modulated using BPSK(1)· The modernised signals in general use more sophisticated modulation schemes to enhance the signal characteristics.

The simplest of these alternate schemes is called Binary Offset Carrier (BOC) modulation, denoted BOC(n, m). The carrier is modulated by a spreading code at a chipping rate of $f_c = m \times f_0$ and additionally by a square sub-carrier at a frequency of $f_{sc} = n \times f_0$, where $n > m$. This sub-carrier splits the main lobe of the signal power spectrum into two lobes centred at $\pm f_{sc}$ from the central frequency. This gives greater code measurement precision and better multipath performance. The autocorrelation function of a BOC signal has three distinct peaks, so the receiver must be careful to track the correct one.

A version of BOC modulation will be used for the GPS and Galileo L1 civil signals. It is likely that the form adopted will be a Multiplexed BOC (MBOC) that places $^{10}/_{11}$ of the total signal power in BOC(1,1) and $^{1}/_{11}$ in a BOC(6,1) modulation. This increases the Gabor bandwidth of the signal and so reduces code measurement noise; it also provides scope for the use of advanced multipath mitigation techniques.

A variation on the BOC modulation scheme is Alternative-BOC (AltBOC(n, m)), which is similar to standard BOC modulation except that a different PRN code is used to modulate the upper and lower sidebands. It is then possible to acquire each sideband separately, or both together. Galileo will use AltBOC$(15, 10)$ to generate E5a and E5b, with power spectral density as shown in Figure 3.4. It will then be possible to acquire either of these independently, in which case they will each have the characteristics of a BPSK(10) signal such as GPS L5, or to track them together as a single wideband signal with improved performance characteristics.

3.4.3 Pilot signals

All of the modernised signals allow for a pilot component in some form. This is a component of a signal (e.g. the quadrature component) that is only modulated by the spreading code and not

by the data message. The reduced power to each signal component is more than made up for by the advantages of a dataless component.

The first advantage is a reduction in the minimum S/N_0 required to acquire a signal. The coherent integration time for carrier acquisition is not limited by the length of the data bit, as is the case for L1 C/A (see Section 3.1), so the receiver can integrate over a longer time period.

The second advantage is for carrier phase tracking. The data message is unknown, so it is not possible to remove it with modulo-2 addition, like the code (which is known by the receiver). In order to track the carrier phase, the receiver must use a special type of phase-locked loop, called a Costas PLL, which is insensitive to the phase transitions of the data message. The Costas loop is mathematically equivalent to a squaring loop, i.e. the received signal is multiplied by itself and the carrier is tracked at twice the original frequency. Therefore the phase transitions due to the BPSK-modulated data message are ignored, but phase measurement errors are doubled. This means that phase tracking errors greater than 90° cause a cycle slip with a Costas loop, whereas a standard PLL tolerates up to 180°. The optimal tracking loop bandwidth of the Costas loop is greater than for the standard PLL, which increases the noise in the phase measurement. A Costas loop results in a 6 dB reduction in tracking threshold; signal acquisition is not affected because this takes occurs before the use of the PLL.

Another disadvantage of the Costas loop is that the doubled frequency causes a 180° phase ambiguity, which must be resolved by looking at the known preamble of the navigation message. This is avoided with the standard PLL, which provides full-wavelength measurements of the carrier phase. Additionally, where there is a data component on the same carrier, the unambiguous carrier determined from the pilot component can be rotated in phase to provide unambiguous coherent data demodulation in the data component, reducing the data demodulation threshold.

A pilot component allows the receiver to track the carrier using a standard PLL, thus reducing the threshold for code acquisition and carrier tracking, and increasing the code measurement precision and cycle slip robustness. The pilot code also provides a better phase reference for data demodulation, improving the data recovery threshold. The data and pilot components are designed to be tracked together (Yang *et al.*, 2004); it is possible to track a component separately but this results in a reduction in the received power. The exception to this is the GPS L2C signal, where the two codes are chip-by-chip multiplexed and only one is tracked at a time (Fontana *et al.*, 2001b).

3.4.4 Navigation message

The modernised signals feature improved navigation message design, both in terms of the structure of the message and the method of transmission. The result is that the data message can be demodulated at lower S/N_0 levels and the receiver time-to-first-fix is reduced.

Forward error correction (FEC) is applied, which adds redundancy to the data transmission and allows the receiver to correct for lost data bits. The data rate is less than the the symbol transmission rate, so the data are transmitted more slowly but can be read at a 5 dB lower S/N_0 with the same success rate. Some of the new signals also use block interleaving of the message frames, where the order of data transmission is scrambled, so a burst error does not cause the loss of all data symbols relating to a data bit. However, this causes latency in the data transmission.

The data rate varies considerably across the new signals. Low data rates are used to improve reception and data demodulation threshold in low S/N_0 environments, while some signals carry high-rate data in order to guarantee the rapid reception of safety-of-life information or to provide encrypted commercial data.

The original GPS navigation message modulated on the L1 C/A-code is termed NAV. A new message structure has been designed for use on L2C and L5, called CNAV. This is more compact and flexible than the original NAV message, and the sequence and timing of each component can be specified by the control centre. FEC is applied, but not block interleaving.

The message structure for L1C is a further improved design and is termed CNAV-2. The data are divided into fixed (slowly changing) data, such as clock and ephemeris data, and variable data. The broadcast message then defines a period of time over which the fixed message data do not change, allowing the receiver to perform data-wiping on this portion of the message. About 60% of the data channel can then be used in the same manner as the pilot channel (Betz *et al.*, 2006a). CNAV-2 will feature more powerful FEC and block interleaving.

There will be three different types of Galileo navigation messages:

- F/NAV, freely accessible navigation message provided by the E5a signal for the Open Service;
- I/NAV, integrity navigation message provided by E5b and E1b signals, supporting Safety of Life Service and providing extended system integrity information;
- C/NAV, commercial navigation message type provided by the E6 signal.

The Galileo services are discussed in Section 3.3.

3.5 Other GNSSs

In addition to modernised GPS and Galileo there may soon be several other fully operational GNSSs that will, to some extent, be compatible with GPS and Galileo.

3.5.1 GLONASS

GLONASS is a Russian acronym for GLObal'naya NAvigatsionnaya Sputnikovaya Sistema (Global Navigation Satellite System). The first GLONASS satellite was launched by the Soviet Union in 1982, and a full constellation was achieved in 1995. However, due underfunding and a low satellite lifespan, the number of operational satellites declined to seven in 2001, when a modernisation program was announced. There are currently 20 operational satellites; the full modernised constellation will have 30 satellites (Inside GNSS, 2009) and there should be 24 satellites by the end of 2009 (Inside GNSS, 2008b).

3.5.2 Compass

Compass is a GNSS currently under development by China. The final constellation will consist of 5 geostationary (GEO) satellites and 30 Medium Earth Orbit (MEO) satellites; these will provide global coverage, with superior coverage over Asia. The first satellite was launched in

Name	Frequency (MHz)	Bandwidth (MHz)	Modulation
B1	1561·10	4·092	QPSK(2)
B1-2	1589·74	4·092	QPSK(2)
B2	1207·14	24	BPSK(10) + BPSK(2)
B3	1268·52	24	QPSK(10)

Table 3.1: Compass signal structure

2000 and the current system consists of 4 GEO and 1 MEO satellites; 10 more launches are planned in the next two years (Inside GNSS, 2008a) and full operational capability is expected to be achieved by 2012.

Compass will broadcast CDMA signals on 4 frequencies (Dong *et al.*, 2008), as shown in Table 3.1. All frequencies except for B1-2 are currently being transmitted (Grelier *et al.*, 2007; Gao *et al.*, 2008). There is some overlap with GPS and Galileo signals: B2 has the same central frequency as Galileo E5b, B3 will overlay Galileo E6, and B1 and B1-2 will overlay the GPS military and Galileo PRS codes at E1/L1. It is not clear how Compass will affect these signals, or how interoperable Compass will be with GPS and Galileo (GPS World, 2008b,a). Two services are planned, the Open Service, for all users, and the Authorised Service; it is not yet known which services will be supplied on which frequencies.

3.5.3 SBAS

In addition to the full GNSS, there are several Satellite-Based Augmentation Systems (SBAS), such as the US Wide Area Augmentation System (WAAS), the European Geostationary Navigation Overlay Service (EGNOS), the Japanese Multi-functional Satellite Augmentation System (MSAS), and India's future GPS Aided Geo Augmented Navigation (GAGAN). These use data from a series of ground stations with GNSS receivers to compute error models or corrections over a regional area. These corrections are then transmitted to users as a data message on a signal similar to the GPS L1 C/A code, where they can be applied the the measurements at the receiver in order to increase the accuracy to a level similar to that achieved with DGNSS. The ground stations also monitor the integrity of the received signals and alert users of any integrity errors. Because the corrections are transmitted on a ranging code, these satellites can be used for positioning. However, there is unlikely to be that many visible at one time; in Europe, the satellites transmitting the EGNOS corrections are in a geostationary orbit, which means that they have a low elevation and are easily obscured.

3.6 Compatibility and interoperability

3.6.1 Compatibility

Compatibility refers to the ability of GNSSs to be used together or independently without interfering with each other. The methodology for determining the GPS/Galileo radio frequency compatibility was specified as part of the US/EU agreement in 2004 ("Reference assumptions

for GPS/Galileo compatibility analyses"). The worst-case interference, of GPS L5 on Galileo E5a, is predicted to be 0·6 dB (Wallner *et al.*, 2005): inter-system interference will be almost undetectable under real conditions.

3.6.2 Interoperability

Interoperability refers to the ability of GNSSs to be used together to provide superior capabilities than either system alone. Interoperability is most important in processing areas dealt with by hardware, where differences significantly increase the cost of using multiple systems. GPS and Galileo will transmit signals with the same central frequency and similar signal structure at L1 and L5, so it will be easy to track both systems at this frequency; L2 and E5b are not interoperable, because there is no corresponding signal from the other system. The two systems use a different data message structure, but the decoding of this is done in software and so will have less impact on interoperability.

GLONASS currently transmits two signals using frequency division multiple access (FDMA), where every satellite transmits on a slightly different central frequency. This makes interoperability with GPS and Galileo more complicated and consequently more expensive, although many modern high-end receivers use both GPS and GLONASS. In 2006 the United States and Russia released a joint statement announcing that Russia was considering CDMA for GLONASS in the future, in order to increase interoperability with GPS. At the end of 2010 Russia will launch the first modernised GLONASS-K satellite, which will allow the evaluation of BOC(2,2) CDMA signals at L1 and possibly L3 (Inside GNSS, 2008b).

3.6.3 Time system

GPS time (GPST) is steered towards Coordinated Universal Time (UTC); Galileo System Time (GST) will be steered towards International Atomic Time (TAI). By definition, there is an integer number of seconds offset between UTC and TAI. However, even after this has been accounted for, there will still be a difference of the order of tens of nanoseconds between GPST and GST (Moudrak *et al.*, 2004), which will introduce a bias into the combined positioning solution. There are several ways to solve this problem:

- The time offset can be computed as an extra parameter in the receiver. This reduces the redundancy of the positioning solution and will therefore reduce the benefit derived from the second system in difficult environments where few additional satellites are visible. However, the time offset only changes slowly, so a previously computed time offset can be used during periods when few satellites are visible;
- It has been agreed that GPS and Galileo will broadcast the GPST-GST offset in the navigation message. Applying this correction will reduce the error, but a bias will still remain due to the uncertainty of the correction: this can be solved for as an additional parameter if sufficient satellites are visible;
- An external provider such as the IGS might compute the clock parameters for GPS and Galileo with respect to a common time scale, therefore eliminating the time offset problem (Moudrak *et al.*, 2004). These products could be available in real time.

3.6.4 Reference frame

The GPS reference frame is WGS84, which is a realisation of the International Terrestrial Reference System (ITRS). Galileo will use the Galileo Terrestrial Reference Frame (GTRF), which is an independent realisation of ITRS. The 2004 "Agreement on the promotion, provision and use of Galileo and GPS satellite-based navigation systems and related applications" specifies that the GPS and Galileo reference frames should be realised as closely as possible to the International Terrestrial Reference Frame (ITRF); the difference between the two will be of the order of 5 cm (Leonard *et al.*, 2002). The coordinate differences are known and can be applied in the user receiver; the offset is too small to affect the majority of users, and will disappear with differenced solutions.

There are many studies of the performance benefits of modernised GPS and Galileo, for a variety of applications. Section 3.7 describes the effect of the additional satellites and frequencies and the improved signal structure, as discussed in Chapter 3, on different aspects of GNSS performance. Section 8.7 describes the likely effect of these improvements on marine positioning, and in particular how future developments are likely to affect the ability of different processing techniques to meet the IMO requirements, as compared to the results achieved with current GPS described in Chapter 8.

3.7 Improvements

This section describes the anticipated performance improvements from modernised GPS and Galileo.

3.7.1 Satellite visibility

Tracking both the GPS and Galileo constellations will provide an approximate doubling of the number of visible satellites. The combined constellation will have greater positioning and integrity monitoring availability, particularly in areas such as congested ports with poor visibility of the sky (Merino *et al.*, 2001). The additional redundant observations will improve positioning accuracy and robustness. The combined constellation will also provide better satellite geometry, which will improve the positioning solution: O'Keefe (2001) shows that the 95th percentile of worldwide horizontal dilution of precision values with a 20° elevation cutoff will be improved from 3 with GPS alone to almost 1 with the combined constellation.

3.7.2 Code observations

Due to the increased signal bandwidth, faster codes and improved modulation schemes which put more signal power towards the edge of the broadcast spectrum, the code measurement noise on all modernised signals except GPS L2C will be significantly reduced compared to current L1 C/A code. In Avila-Rodríguez *et al.* (2004) the relative code measurement precisions for C/A, L5/E5a/E5b and E5a+b are given as 36 cm, 6 cm and 1 cm respectively. The precision of the old Galileo L1 modulation, BOC(1,1) is given as 11 cm; according to Avila-Rodríguez *et al.* (2008) the precision of the new MBOC(6,1,1/11) modulation scheme for GPS and Galileo L1

Code	Tracking noise (1 σ, cm)
GPS L1 C/A	15·0
GPS L2C	15·0
Galileo E1	6·1
Galileo E5a	3·9
Galileo E5b	3·7
Galileo E5a+b	0·9

Table 3.2: Measured GIOVE-A and L2C code precision with C/N_0 of 45 dB Hz (Simsky *et al.*, 2006b)

Baseline	GPS	GPS + Galileo
Short	$2{\cdot}4 \times 10^{-3}$	$< 10^{-8}$
Medium	$6{\cdot}3 \times 10^{-1}$	$6{\cdot}3 \times 10^{-2}$
Long	1·0	1·0

Table 3.3: Instantaneous dual-frequency ambiguity resolution fail rate (95th percentile over 24 hours, Munich) (Eissfeller *et al.*, 2001)

is about mid-way between BOC(1,1) and L5/E5a/E5b, or 8 cm. In Simsky *et al.* (2006b) the precision of the broadcast GIOVE-A and GPS L2C signals was measured: the results, given in Table 3.2, show the significant improvement in tracking precision of the Galileo signals. Note that the values are not directly comparable to the theoretical results of Avila-Rodríguez *et al.* (2004) due to the different C/N_0.

3.7.3 Ambiguity resolution

There are many studies of the effect of modernised GPS and Galileo on ambiguity resolution: the consensus is that the greatest improvement in ambiguity resolution success rate will be realised by the use of the additional satellites in a combined constellation, and there will be less benefit from the third frequency (Verhagen, 2005b). Exact numbers are difficult to quantify because of the dependency on the GNSS model used, the simulated observation precision and biases, and the baseline length. In Landau *et al.* (2004) the single-epoch geometry-based success rate is greater than 50% for GPS alone for baselines up to 10 km, with triple-frequency GPS up to 12·5 km, and with triple-frequency GPS and Galileo up to 17·5 km. The instantaneous single-epoch fail rates for different baseline lengths given in Eissfeller *et al.* (2001) are shown in Table 3.3: an order of magnitude increase is achieved over a medium baseline by using the combined constellation.

Single-epoch ambiguity resolution, which uses the code observations to obtain an estimate of the float ambiguities, will also benefit from the improvements in code accuracy, particularly in the presence of multipath (Milbert, 2005; Joosten *et al.*, 1999).

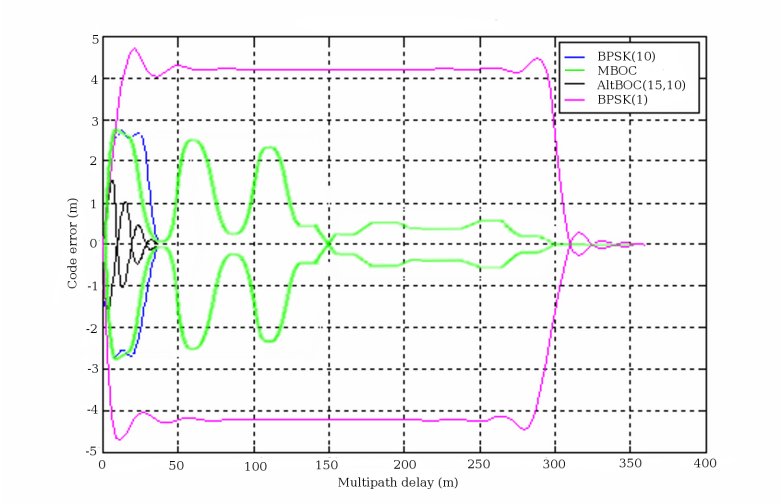


Figure 3.5: Code multipath error envelopes for signal-to-multipath ratio of 6 dB (Simsky *et al.*, 2006a, 2008)

3.7.4 Multipath mitigation

The superior modulation schemes and faster chipping rates of the new signals will provide increased robustness to multipath. Figure 3.5 shows the theoretical maximum multipath error for different modulation schemes. BPSK(1) is used for L1 C/A and L2C, MBOC is used for the modernised L1 signals, BPSK(10) is used by L5, E5a and E5b, and AltBOC is the coherent tracking of E5a+b. The maximum delay for BPSK(10) and AltBOC is $1/10$ of the maximum delay for BPSK(1) and MBOC, because the codes have 10 times the chipping rate and therefore $1/10$ the wavelength. MBOC shows greatly improved multipath resistance over BPSK(1), with a mean maximum multipath error of ~ 1 m compared to ~ 4 m. Coherent tracking of E5a+b results in a significant reduction in multipath compared to tracking each separately; the combined signal shows the greatest resistance to multipath. The new modulation schemes, in particular MBOC, have greater potential for the use of advanced multipath mitigation techniques (Hein *et al.*, 2006; Weill, 2002).

Simsky *et al.* (2006a) give multipath results measured from from GIOVE-A, and show that the Galileo codes are a significant improvement over GPS, particularly for long path-length multipath from low-elevation satellites. For example, the mean recorded multipath on one day was $0{\cdot}58$ m for C/A, $0{\cdot}38$ m for E1 (using BOC(1,1)), $0{\cdot}25$ m for E5a, $0{\cdot}27$ m for E5b and $0{\cdot}14$ m for E5a+b.

There will also be a reduction in the effect of phase multipath. The additional frequencies have different wavelengths and therefore a different multipath error at a given time, so multipath biases will be averaged out in the least-squares adjustment. In Lau & Cross (2003) it is shown that this effect, along with a phase multipath outlier rejection strategy, could improve the accuracy of phase positioning by 50% compared to current dual-frequency positioning. It is shown in Simsky (2006) that with three frequencies it will be possible to estimate the phase multipath from the measurement data from a single station by forming the geometry-free ionosphere-free linear combination, which contains only the combined tracking errors and multipath.

3.7.5 Ionospheric modelling

The Galileo broadcast ionospheric model, NeQuick, will be superior to the currently-used Klobuchar model, eliminating around 70% of the ionospheric error compared to 50% for Klobuchar (Somieski *et al.*, 2007); this will particularly benefit single-frequency point positioning. The noise multiplication factor for the L1/L5 ionosphere-free observable is $2{\cdot}59$, compared to $2{\cdot}98$ for the currently-used L1/L2 combination, due to the increased frequency separation. This, combined with the reduced code measurement noise and multipath, will improve the precision of the ionosphere-free observables and the estimation of the ionospheric error.

A three-frequency system will allow the estimation of higher-order ionospheric effects (Wang *et al.*, 2005; Julien *et al.*, 2004), which will improve ionospheric modelling. There will also be a greater choice of ionosphere-free observables for phase positioning (Odijk *et al.*, 2002).

3.8 Influence of GNSS on maritime activity

The advent of GPS has had a significant impact on marine navigation. As discussed in Chapter 2, carriage of a GNSS receiver is mandated for vessels covered by the Safety Of Life At Sea (SOLAS) convention, which covers commercial vessels over 300 dwt. The use of GNSS as a navigation aid is being developed further with Automatic Identification Systems (AIS), whereby ships transmit their GNSS-derived position to other vessels and to shore stations in order to improve safety and situational awareness. This positional information may then also be used by port authorities to strictly control the course of ships in congested areas to improve safety and shipping throughput. GNSS positions are also transmitted as part of distress signals, so that the position of a ship in trouble can easily be determined even far out to sea. GNSS is also used extensively in hydrographic surveying, dredging and fishing, and for pleasure yachting. The improvements in positioning performance provided by modernised GPS and Galileo will increase the utility of GNSS for these applications, and will allow further expansion into fields such as automatic docking, where GPS is currently not sufficiently reliable for general use (although some ferries do currently operate automatic docking systems). As discussed in Chapter 2, the IMO maintains a set of requirements on GNSS positioning; as the achievable positioning performance and the number of applications increases, these requirements are revised to be more specialised and more strict.

3.9 Summary

This chapter describes the structure and advantages of the new signals. For modernised GPS, there will be a new civil signal on L2, a new signal at L5 and an improved signal structure at L1. Galileo will bring 30 satellites transmitting Open Service signals that are interoperable with GPS at L1 (E1) and L5 (E5a), and an additional signal adjacent to E5a, called E5b, that can be combined with E5a as a single wide-bandwidth signal. Many of the new signals will feature longer, faster, tiered codes for improved acquisition threshold, tracking precision and robustness, and multipath mitigation, in addition to pilot signals which are not modulated by the data message and will increase the carrier tracking robustness. A theoretical analysis is made of

the areas where the additional satellites and improved signal structure may result in improved positioning performance. In Chapter 8 the effect of this improvement on the ability of GNSS to meet the IMO requirements is assessed in light of the results of the data collected in Harwich harbour, described in Chapter 8.

Chapter 4

Integer ambiguity resolution and validation

Ambiguity resolution is an essential step in RTK positioning, which is required to meet the most precise IMO requirements. RTK relies on measurements of phase of the carrier wave underlying GNSS signals. The carrier has a much shorter wavelength than the code, so measurements are more precise, but the receiver cannot measure the range directly like it does with the code.

At the epoch of transmission, the satellite broadcasts a sine wave carrier signal with a certain phase. As shown in Figure 4.1, this signal then travels to the receiver, going through a number of whole cycles, and arrives at the antenna at reception time with another phase. The receiver generates a copy of the signal identical to that generated in the satellite, and compares the phase of this copy to the phase of the received signal to obtain the phase observation. However, this does not tell the receiver how many whole cycles the signal passed through during the travel from the satellite to the receiver, and this information is essential to determine the the range to the satellite and thereby produce a position. This whole number of cycles is known as the *integer ambiguity*, and the value must be be determined before the phase observations can be used as precise ranges in RTK positioning: this process is known as *ambiguity resolution.* If the ambiguities are resolved incorrectly then large errors in the position may result, potentially causing an integrity error.

Section 4.1 gives an overview of the theory of ambiguity resolution, and Section 4.2 gives equations for computing the probability that the ambiguities are resolved to the correct values. Section 4.3 discusses ambiguity validation techniques, which are used to determine if there is sufficient confidence in the fixed ambiguities to use them for positioning: this is very important because using incorrect values for the ambiguities can result in integrity errors. Section 4.4 describes partial ambiguity resolution, which is a technique for resolving a subset of the ambiguities with a higher probability of resolving them correctly; this will allow more reliable positioning in difficult environments such as ports, at the expense of reduced accuracy. However, the current partial ambiguity resolution techniques choose the subset of the ambiguities to fix based on their theoretical precision rather than the actual observations, so partial ambiguity resolution can still fail due to biased observations, for example due to phase multipath (reflected GNSS signals reaching the receiver and combining with the direct signal to cause a range bias). Therefore, in Section 4.5 a new technique is developed to perform partial ambiguity resolution based upon the actual measurements; this should allow more reliable positioning in difficult environments and so should help in meeting the strictest IMO requirements, such as automatic

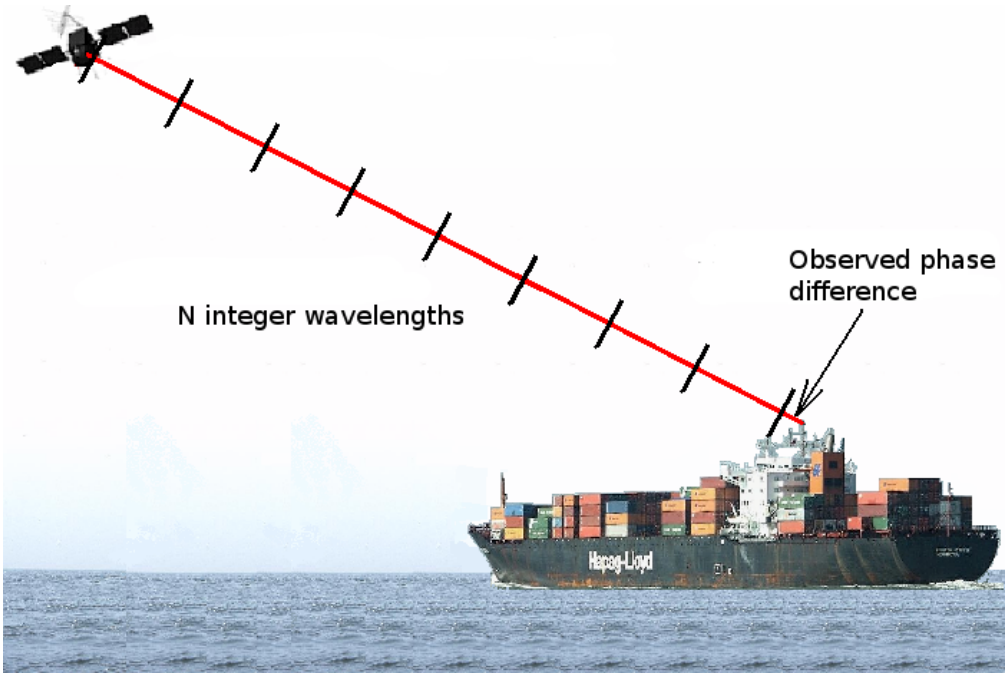


Figure 4.1: Phase measurement

docking. The benefits of this technique are assessed in Chapter 8 using real data.

4.1 Ambiguity resolution

RTK is the most accurate GNSS positioning technique, and is essential if a future GNSS is to meet the IMO requirements for automatic docking. Unlike point positioning and DGNSS, RTK relies upon the measurements of the underlying carrier phase of the GNSS signal. The observation equation is:

$$\Phi_r^s = \rho_r^s + c\left(dt_r\left(t\right) - dt^s\left(t - \tau_r^s\right)\right) + T_r^s - I_r^s + \lambda N_r^s + \epsilon_r^s \tag{4.1}$$

where:

Φ_r^s	is the phase observation in metres
ρ_r^s	is the geometric range
$dt_r\left(t\right)$	is the receiver clock offset at reception time
$dt^s\left(t - \tau_r^s\right)$	is the satellite clock offset at transit time
T_r^s	is the tropospheric error
I_r^s	is the ionospheric error
N_r^s	is the unknown number of cycles between the receiver and the satellite
ϵ_r^s	is the remaining error

In a single epoch, this does not in itself contain any information about the range to the satellite, due to the addition of the ambiguity term. However, if the value of this term can be determined, then the phase observations can be used as precise ranges. In the undifferenced measurement of Equation 4.1 there are too many error sources to enable the ambiguity to be estimated; in particular, the clock errors and hardware delays. Therefore the usual technique for ambiguity

resolution uses the combined observations from two receivers to two satellites are combined to form the double-differenced observation, which cancels the clock errors and hardware delays:

$$\Phi_{r_1,r_2}^{s_1,s_2} = \rho_{r_1,r_2}^{s_1,s_2} + T_{r_1,r_2}^{s_1,s_2} - I_{r_1,r_2}^{s_1,s_2} + \lambda N_{r_1,r_2}^{s_1,s_2} + \epsilon_{r_1,r_2}^{s_1,s_2} \tag{4.2}$$

where:

$\Phi_{r_1,r_2}^{s_1,s_2}$	is the double-differenced phase observation in metres
$\rho_{r_1,r_2}^{s_1,s_2}$	is the double-differenced geometric range
$T_{r_1,r_2}^{s_1,s_2}$	is the double-differenced tropospheric error
$I_{r_1,r_2}^{s_1,s_2}$	is the double-differenced ionospheric error
$N_{r_1,r_2}^{s_1,s_2}$	is the integer ambiguity
$\epsilon_{r_1,r_2}^{s_1,s_2}$	is the remaining error

The receiver and satellite clock offsets and the hardware delays are cancelled, and spatially-correlated error sources such as the tropospheric and ionospheric errors are reduced. If the integer value of the ambiguity can be determined, then the double-differenced phase observation can be used as a very precise range measurement.

Many different techniques for determining the integer value of the ambiguities have been described: see Kim & Langley (2000); Teunissen (2001c); Verhagen & Joosten (2004) for an overview. The search for the correct integer ambiguities can be carried out in either the coordinate domain or the ambiguity domain.

4.1.1 Coordinate domain

Coordinate domain techniques, such as the Ambiguity Function Method (AFM) (Counselman & Gourevitch, 1981), search around an approximate initial position for points in space that minimise the fractional part of the double-differenced phase measurement residuals. The idea is that the phase measurements will fit together best at the true position. The misfit between the phase measurements is quantified through the ambiguity function (Corbett & Cross, 1995):

$$A(X) = \left(\frac{1}{m \cdot (n-1) \cdot 2}\right) \sum_{k=1}^{m} \sum_{j=1}^{n-1} \sum_{l=1}^{2} \cos\left\{2\pi \left[\phi_{\text{obs}}^{kjl}(X_0) - \phi_{\text{cal}}^{kjl}(X)\right]\right\} \tag{4.3}$$

where:

$A(X)$	is the ambiguity function value at point X
k	is the fixed station count
j	is the satellite count
l	is the frequency count
X_0	is the correct position
X	is the trial position
$\phi_{\text{obs}}(X_0)$	is the double differenced observed phase
$\phi_{\text{cal}}(X)$	is the calculated double differenced phase at the test position
m	is the number of fixed stations
n	is the number of satellites

In some implementations the point with the maximum ambiguity function is taken as the ambiguity-fixed position (Mader, 1992); in others the best points are selected for further analysis. For example, Corbett & Cross (1995) perform an ambiguity-fixed solution with those sets of

fixed ambiguities that have a high ambiguity function value, and select the solution with the lowest residual.

The AFM has the advantage that it is immune to cycle slips and discontinuities of the data, and will take full advantage of additional phase measurements. However, it may be necessary to compute the ambiguity function for a large number of points, so computational efficiency is low; other techniques also have a higher success rate, so this technique is not widely used.

4.1.2 Ambiguity domain

The GNSS model can be described in the form of linear(ised) observations equations:

$$y = Aa + Bb + e \tag{4.4}$$

where A and B are the design matrices, $e \in \mathbb{R}^m$ is the noise vector, $y \in \mathbb{R}^m$ contains the observed-minus-computed double-difference observations, and the parameter vectors $a \in \mathbb{Z}^n$ and $b \in \mathbb{R}^p$ respectively contain the integer double-difference ambiguities in cycles, and the unknown rover coordinates (in the geometry-based model) or the unknown satellite ranges (in the geometry-free model), and any other real-valued parameters such as atmospheric parameters.

This model is then solved to obtain b, which contains the parameters of interest. The solution is obtained in three steps. In the first step, the float solution, the integer nature of a is disregarded and a standard least-squares adjustment performed. This results in real-valued estimates of a and b and the covariance matrix:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}; \begin{bmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{bmatrix} \tag{4.5}$$

where $\hat{\cdot}$ denotes the least-squares estimate. In the second step, ambiguity resolution, an ambiguity estimator is used to map the real-valued float ambiguities to integers. The ambiguity estimator is a mapping $M : \mathbb{R}^n \mapsto \mathbb{Z}^n$ such that:

$$\check{a} = M(\hat{a}), \ \check{a} \in \mathbb{Z}^n, \hat{a} \in \mathbb{R}^n \tag{4.6}$$

where $\check{\cdot}$ denotes the integer estimate. In the third step, the fixed solution, the fixed integer ambiguities are used to determine the parameter vector b to a higher precision. If b contains the final parameters of interest, it can be corrected to give:

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \check{a}) \tag{4.7}$$

Otherwise, the fixed ambiguities are used to correct the observed minus computed observations in y, and a new least-squares adjustment, without the ambiguity parameters, is performed to find the parameters of interest. This is necessary when the geometry-free model is used in the first stage and the parameters in b are the satellite ranges: a separate geometry-based ambiguity-fixed adjustment must be performed to determine the rover coordinates.

4.1.2.1 Admissible integer estimators

The ambiguity resolution problem is therefore to determine the mapping from the real-valued float ambiguities to the integer-valued fixed ambiguities. This procedure is treated rigorously by the theory of *admissible integer estimators* (Verhagen, 2005a; Teunissen, 2001c, 1999). The many-to-one mapping $M : \mathbb{R}^n \mapsto \mathbb{Z}^n$ defines a subset $S_z \subset \mathbb{R}^n$ around each integer vector $z \in \mathbb{Z}^n$ where each real-valued vector $x \in \mathbb{R}^n$ is mapped to z:

$$S_z = \{x \in \mathbb{R}^n | \, z = M(x)\}, \; z \in \mathbb{Z}^n \tag{4.8}$$

This subset is called the *pull-in region* of z (Teunissen, 1998c) and contains all real-valued float ambiguity vectors $\hat{a}$ that will be mapped to the integer vector z by the integer estimator, i.e. $\check{a} = z \iff \hat{a} \in S_z$. The integer estimator can be expressed as:

$$\check{a} = \sum_{z \in \mathbb{Z}^n} z s_z(\hat{a}) \tag{4.9}$$

where the indicator function:

$$s_z(x) = \begin{cases} 1 & \text{if } x \in S_z \\ 0 & \text{otherwise} \end{cases}$$

An integer estimator is *admissible* if $\forall z \in \mathbb{Z}^n$ the pull-in regions S_z satisfy (Teunissen, 1999):

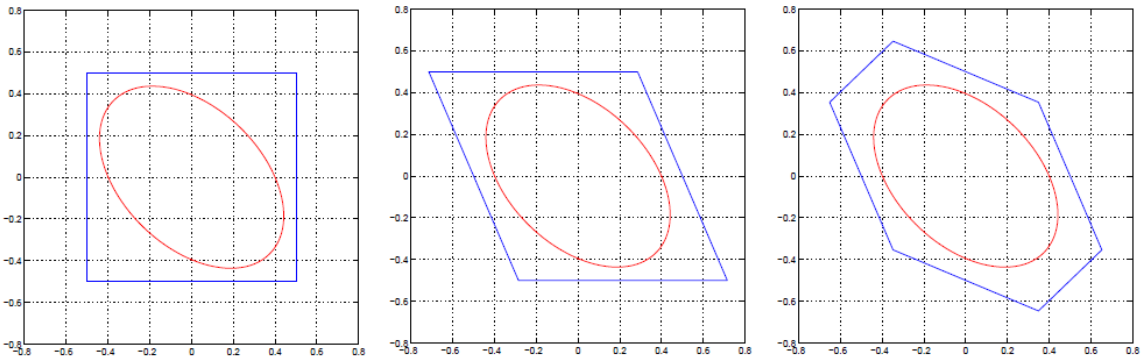
$$(i) \qquad \bigcup_{z \in \mathbb{Z}^n} S_z = \mathbb{R}^n \tag{4.10}$$

$$(ii) \qquad \text{Int}(S_u) \bigcap \text{Int}(S_z) = \emptyset, \; \forall u, z \in \mathbb{Z}^n, \; u \neq z \tag{4.11}$$

$$(iii) \qquad S_z = z + S_0, \; \forall z \in \mathbb{Z}^n \tag{4.12}$$

where 'Int' is the interior of the set. The first condition requires that the pull-in regions cover the whole of $\mathbb{R}^n$, so all float ambiguity vectors are mapped to an integer vector. The second condition prevents the pull-in regions from overlapping, so that each float ambiguity vector will only be mapped to a single integer vector. The third condition requires translational invariance, so every pull-in region is the same shape.

The three most commonly-used estimators are: integer rounding, integer bootstrapping and integer least-squares. Figure 4.2 shows the two-dimensional pull-in regions for these estimators. The red ellipse is a contour of the probability distribution of the float ambiguities, and the blue polygon is the pull-in region about $(0, 0)$. The more closely the pull-in region matches the shape of the float ambiguity probability distribution, the better the integer estimator. Integer rounding has a very low success rate and is not used in practice. Integer bootstrapping is usually applied with geometry-free model, and is fast to compute. Integer least-squares has the highest probability of correctly resolving the ambiguities, but requires the use of a search technique to find the correct solution and so is the most computationally intensive.



(a) Rounding (b) Bootstrapping (c) Integer least squares

Figure 4.2: Two-dimensional pull-in regions (Teunissen *et al.*, 2002)

4.1.2.2 Integer rounding

Integer rounding is the simplest technique: each real-valued ambiguity estimate is rounded to the nearest integer. Therefore if the set of float ambiguities is $\hat{a} = \begin{bmatrix} \hat{a}_1 & \cdots & \hat{a}_n \end{bmatrix}^T$, the set of rounded integer ambiguities is:

$$\check{a}_R = \begin{bmatrix} [\hat{a}_1] \\ \vdots \\ [\hat{a}_n] \end{bmatrix} \tag{4.13}$$

where $[\cdot]$ denotes rounding to the nearest integer. In practice this technique is not used due to the low success rate caused by the high correlation of the double-differenced ambiguities.

4.1.2.3 Integer bootstrapping

Integer bootstrapping is a sequential integer rounding technique, which takes account of the correlation between the ambiguities (Teunissen, 1993). The first ambiguity is rounded to the nearest integer; all subsequent integers are then corrected by virtue of their correlation with the first integer. The second ambiguity is then rounded and subsequent ambiguities corrected; this process of rounding and correction of subsequent ambiguities continues until all ambiguities have been fixed. In the case where the covariance matrix is diagonal, there is no correlation correction and bootstrapping is identical to integer rounding. The bootstrapped integer estimates are $\check{a}_B = [\check{a}_{B,1} \ldots \check{a}_{B,n}]^T$, where (Teunissen, 2001a):

$$\begin{array}{lcll} \check{a}_{B,1} & = & [\hat{a}_1] & \\ \check{a}_{B,2} & = & [\hat{a}_{2|1}] & = \quad \left[\hat{a}_2 - \sigma_{21}\sigma_1^{-2}\left(\hat{a}_1 - \check{a}_{B,1}\right)\right] \\ \vdots & & \vdots & \\ \check{a}_{B,n} & = & [\hat{a}_{n|N}] & = \quad \left[\hat{a}_n - \sum_{j=1}^{n-1} \sigma_{n,j|J}\sigma_{j|J}^{-2}\left(\hat{a}_{j|J} - \check{a}_{B,j}\right)\right] \end{array} \tag{4.14}$$

where $\hat{a}_{i|I}$ is the ith ambiguity estimate conditioned on the previous $I = \{1, \ldots, (i-1)\}$ sequentially rounded ambiguities. The results of integer bootstrapping are dependent on the order in which the ambiguities are fixed: the highest probability of successful ambiguity resolution is achieved when the ambiguities are fixed in decreasing order of precision.

4.1.2.4 Integer least squares

The integer least squares estimator is optimal, in the sense that it maximises the probability of correct ambiguity resolution (Teunissen, 1999). The fixed integer least squares solution is:

$$\check{a}_{LS} = \arg\min_{z \in \mathbb{Z}^n} \|\hat{a} - z\|^2_{Q_{\hat{a}}} \tag{4.15}$$

i.e. the chosen solution minimises the sum of the squares of the residuals in the metric of the covariance matrix of the float ambiguities. It is not possible to compute $\check{a}_{LS}$ directly as with bootstrapping and rounding: a search technique must be employed. Many different search techniques have been proposed (Kim & Langley, 2000), for example: Fast Ambiguity Resolution Approach (FARA), Fast Ambiguity Search Filter (FASF), Integrated Three-Carrier Ambiguity Resolution (ITCAR), (LAMBDA) (Teunissen, 1993), Least Squares Ambiguity Search Technique (LSAST), Modified Cholesky Decomposition, the Null-space method (Martin-Neira *et al.*, 1995) and Optimal Method for Estimating GPS Ambiguities (OMEGA) (Kim & Langley, 1999). The LAMBDA technique (de Jonge & Tiberius, 1996; Teunissen, 1995, 1993) is widely used, and provides an efficient search method which can be applied to any integer least-squares problem without modification of the algorithm.

4.1.2.5 Ambiguity decorrelation

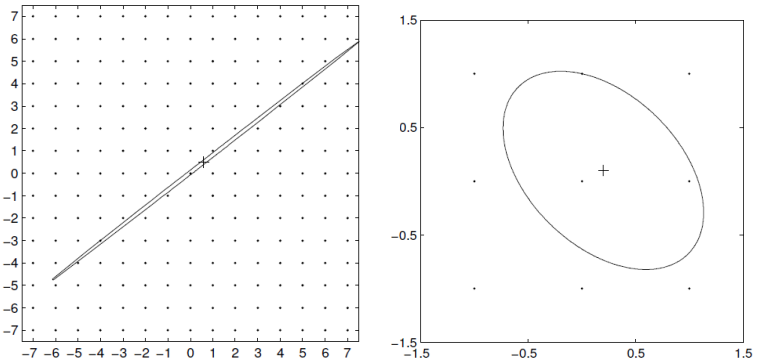
A common feature of many forms of bootstrapping or integer least squares search techniques is the application of a decorrelating linear transformation, Z, to obtain more precise ambiguities, $\hat{z}$:

$$\hat{z} = Z^T \hat{a}, \quad Q_{\hat{z}} = Z^T Q_{\hat{a}} Z, \quad \check{a} = Z^{-T} \check{z} \tag{4.16}$$

This transformation must be admissible, i.e. Z and Z^{-1} must have integer entries.

Such a transformation is computed as part of the LAMBDA search technique: the ambiguity search space is transformed from a highly elongated shape to a more compact one, as shown in Figure 4.3. This greatly increases the efficiency of the search because it is not necessary to examine points so far away from the float solution. However, due to the least-squares criterion, the final result is the same.

When used with the bootstrapping estimator, the decorrelation adjustment improves the success rate rather than computational efficiency. This is the basis of geometry-free ambiguity resolution techniques such as Cascade Integer Resolution (CIR) (Jung *et al.*, 2000) and Three-Carrier Ambiguity Resolution (TCAR) (Harris, 1997): a decorrelating adjustment, determined *a priori*, is applied to make the first ambiguities more precise and thereby increase the probability of fixing them successfully. With a dual-frequency system this may mean, for example, fixing the widelane combination for each satellite (wavelength $0{\cdot}86$ m) by integer rounding, then conditioning the L1 ambiguity by virtue of its correlation with the widelane and rounding it to the nearest integer. This allows the reconstruction of the L1 and L2 ambiguities, but due to its large wavelength the widelane is more precise (with the standard deviation in cycles) and so the probability of successful ambiguity resolution is increased. This approach is only suitable for the geometry-free case: it is not possible to specify linear combinations between the observations to different satellites *a priori*, due to the unknown satellite geometry.



(a) Before decorrelation adjustment (b) After decorrelation adjustment

Figure 4.3: Two-dimensional example of an ambiguity search space around the float solution (denoted +) (Verhagen, 2005a)

4.2 Computing ambiguity resolution success rate

It is possible to theoretically derive the probability of correctly resolving the ambiguities when using an ambiguity domain integer estimator. This can be useful for mission planning or for comparing different techniques, but the estimates are highly dependent on the float ambiguity covariance matrix and so their accuracy is limited by how well the covariance matrix models the true errors. The results in over-optimistic success rates in many situations. Conceptually, the success rate P_s is the integration of the probability density function f of the float ambiguities $\hat{a}$ over the pull-in region S_a of the true values a:

$$P_s = P\left(\check{a} = a\right) = \int_{S_a} f_{\hat{a}}\left(x\right) dx \tag{4.17}$$

4.2.1 Integer rounding

It is difficult to compute the success probability of integer rounding, $P\left(\check{a}_R = a\right)$, directly, but it is bounded by (Teunissen, 1998d):

$$\prod_{i=1}^{n}\left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_i}}\right)-1\right) \leq P\left(\check{a}_R = a\right) \leq \prod_{i=1}^{n}\left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{i|I}}}\right)-1\right) \tag{4.18}$$

where the $\Phi\left(x\right)$ is the function:

$$\Phi\left(x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{1}{2}z^2\right)dz \tag{4.19}$$

If the covariance matrix is diagonal, $\sigma_{\hat{a}_i} = \sigma_{\hat{a}_{i|I}}$ and the upper and lower bounds are identical.

4.2.2 Bootstrapping

The success rate of integer bootstrapping can be computed directly and is given in Teunissen (1998d) as:

$$P\left(\check{a}_B = a\right) = \prod_{i=1}^{n}\left(2\Phi\left(\frac{1}{2\sigma_{\hat{a}_{i|I}}}\right)-1\right) \tag{4.20}$$

This is the upper bound of the success rate of integer rounding as given in Equation 4.18, which implies that bootstrapping has at least as high a probability of successful ambiguity resolution as integer rounding. Unless the covariance matrix is diagonal, the presence of the conditional variances $\sigma_{\hat{a}_{i|I}}$ means that the success rate is dependent on the order in which the ambiguities are fixed.

4.2.3 Integer least squares

It is not possible to compute the success rate of integer least squares directly. Several techniques have been proposed for computing upper and lower bounds. It is suggested in Verhagen (2003) that, for a combination of ease of computing and accuracy, the bootstrapping success rate be used as a lower bound, and the upper bound be computed using the Ambiguity Dilution of Precision (ADOP)-based technique (Teunissen, 2000; Hassibi & Boyd, 1998):

$$P\left(\check{a}_B = a\right) \leq P\left(\check{a}_{LS} = a\right) \leq P\left(\chi^2 \leq \frac{c_n}{ADOP^2}\right) \tag{4.21}$$

where χ^2_n has a central Chi-squared distribution with n degrees of freedom, and

$$c_n = \frac{\left(\frac{n}{2}\Gamma\left(\frac{n}{2}\right)\right)^{\frac{2}{n}}}{\pi} \tag{4.22}$$

where Γ is the Gamma function. ADOP, a measure of the precision of the ambiguities (Teunissen *et al.*, 2000c), is calculated as:

$$ADOP = \sqrt{\det Q_{\hat{a}}}^{\frac{1}{n}} \tag{4.23}$$

The sharpness of the lower bound (i.e. the bootstrapping success rate) is reduced by the correlation of the ambiguities. It is therefore best to apply a decorrelating adjustment before carrying out the computation: when using the LAMBDA technique, such an adjustment is computed as part of the least-squares search algorithm. The upper bound is invariant to transformation of the ambiguities.

4.2.4 Success rate in the presence of bias

The preceding techniques for computing the ambiguity resolution success rate assume that the measurements are unbiased. The theory for computing the success rate in the presence of biased observations is similar, but the probability density function of the float ambiguities is no longer centred in the pull-in region. Teunissen (2001b); Teunissen *et al.* (2000a) give the probability of successful ambiguity resolution for integer bootstrapping in the presence of a bias as:

$$P_b\left(\check{a}_B = a\right) = \prod_{i=1}^{n}\left[\Phi\left(\frac{1-2\beta_i}{2\sigma_{\hat{a}_{i|I}}}\right) + \Phi\left(\frac{1+2\beta_i}{2\sigma_{\hat{a}_{i|I}}}\right) - 1\right] \tag{4.24}$$

where β_i is the ith element in the decorrelated bias vector b.

4.3 Ambiguity validation

The success rate as computed in Section 4.2 gives the probability of successful ambiguity resolution in general, based on the covariance matrix. It does not say anything about the specific case of the current epoch based on the float ambiguity vector, and is often over-optimistic due to the assumption of un-biased observations. Therefore it is necessary to apply a discrimination test after the best set of fixed ambiguities have been identified to determine if there is sufficient confidence in these values; if this test is failed then the float solution is maintained.

There are several tests currently used for this purpose, for example the ratio test, the difference test and the projector test. These tests compare the best and second-best sets of ambiguities to determine if they are sufficiently distinguishable. The most commonly used is the ratio test: the ratio of the squared norms of the difference of residuals from the float solution is computed, and the best fixed solution $\check{a}$ accepted iff:

$$\frac{\|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}}}{\|\hat{a} - \check{a}\|^2_{Q_{\hat{a}}}} > c \tag{4.25}$$

where $\check{a}_2$ is the second-best set of fixed ambiguities. The critical value c is usually chosen as $c = 2$, $2{\cdot}5$ or 3 (Verhagen, 2005a). However, the distribution of c is not known, so these values are empirically derived and have no sound theoretical basis. Note that c does not follow the f-distribution because $\|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}}}$ and $\|\hat{a} - \check{a}\|^2_{Q_{\hat{a}}}$ are not independent (Teunissen, 1998a).

4.3.1 Integer aperture estimation

4.3.1.1 Definition

The theory of integer aperture estimation (Teunissen, 2003a,c) can be used to provide a rigorous foundation for the currently used validation tests. This is an extension of the integer estimator theory described in Section 4.1.2 to provide an overall theory of integer estimation and validation. Requirement (i) for an integer estimator (Equation 4.10), that there are no gaps between the pull-in regions, is dropped, which allows three outcomes: *success* if the float ambiguities are resolved to the correct integers, *failure* if they are resolved to the incorrect integers, and *undecided* if they are not resolved, and the float solution is maintained. Therefore an integer aperture estimator must satisfy:

$$(i) \qquad \bigcup_{z \in \mathbb{Z}^n} \Omega_z = \Omega \tag{4.26}$$

$$(ii) \qquad \text{Int}\left(\Omega_u\right) \bigcap \text{Int}\left(\Omega_z\right) = \emptyset, \ \forall u, z \in \mathbb{Z}^n, u \neq z \tag{4.27}$$

$$(iii) \qquad \Omega_z = z + \Omega_0, \ \forall z \in \mathbb{Z}^n \tag{4.28}$$

where $\Omega \subset \mathbb{R}^n$ is the *aperture space* or acceptance region. As with the integer estimator, the acceptance regions are required to be disjunct and translationally invariant. Figure 4.4 shows a two-dimensional example of the acceptance regions for an integer aperture estimator compared to the pull-in regions for the integer least squares estimator: float solutions that fall outside the red ellipses are not fixed to integers.

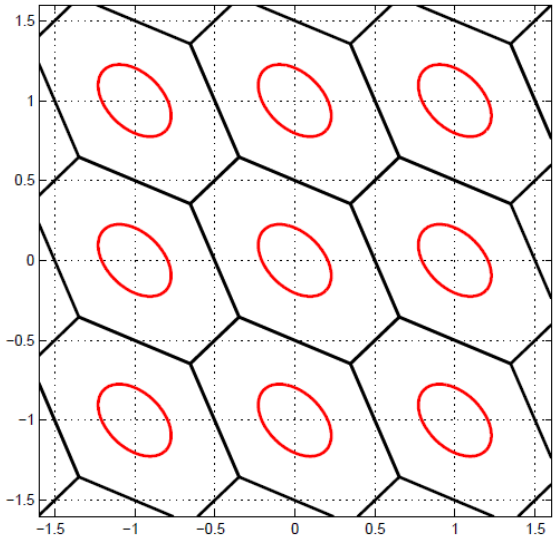


Figure 4.4: Two-dimensional example of aperture pull-in regions (red) and integer least squares pull-in regions (black) (Teunissen & Verhagen, 2004)

The integer aperture estimator, $\bar{a}$, is given by:

$$\bar{a} = \sum_{z \in \mathbb{Z}^n} z\omega_z(\hat{a}) + \hat{a}\left(1 - \sum_{z \in \mathbb{Z}^n} \omega_z(\hat{a})\right) \tag{4.29}$$

with the indicator function $\omega_z(x)$ defined as:

$$\omega_z(x) = \begin{cases} 1 & \text{if } x \in \Omega_z \\ 0 & \text{otherwise} \end{cases}$$

Therefore when $\hat{a} \in \Omega_z$ the ambiguity will be fixed to z, otherwise the float solution will be maintained.

There are two variables to be tuned: the size and shape of the integer aperture acceptance regions. The shape is determined by the estimator chosen, and the acceptance regions can be scaled to control the probabilities of success and failure by the choice of the aperture parameter μ. At one extreme, $\mu = 1$ and the integer aperture estimator is identical to the corresponding integer estimator; at the other, $\mu = 0$ and the acceptance region has no size and the ambiguities are never fixed.

It is proven in Teunissen & Verhagen (2004); Verhagen (2005a) that the ratio test, the difference test and the projector test are all examples of integer aperture estimators, which provides these these tests with a sound theoretical basis. There are many possible integer aperture estimators, such as the ellipsoid estimator shown in Figure 4.4, the bootstrapping and least squares estimators, which are based on a scaled-down pull-in region of the corresponding integer estimator, and the estimators corresponding to the ratio, difference and projector tests; see Verhagen (2005a) for an overview.

4.3.1.2 Optimal integer aperture estimator

The optimal integer aperture estimator is defined in Teunissen (2003c) and has the highest probability of successful ambiguity resolution for a given fail rate. The acceptance region is defined as:

$$\Omega_0 = \left\{ x \in S_0 | \sum_{z \in \mathbb{Z}^n} f_{\hat{a}}\left(x+z\right) \leq \mu f_{\hat{a}}\left(x+a\right) \right\} \tag{4.30}$$

whereS_0 is the integer least-squares pull-in region and $f_{\hat{a}}$ is the probability density function of the float ambiguities. $\check{a}$ is accepted iff

$$\frac{f_{\check{\epsilon}}\left(\hat{a}-\check{a}\right)}{f_{\hat{a}}\left(\hat{a}-\check{a}\right)} \leq c \tag{4.31}$$

where $f_{\check{\epsilon}}\left(x\right)$ is the probability density function of the ambiguity residuals $\check{\epsilon} = \hat{a} - \check{a}$; this depends on $\left\|\hat{a}-z\right\|^2_{Q_{\hat{a}}}$ $\forall z \in \mathbb{Z}^n$ (Verhagen, 2005a) and so is very computationally expensive to evaluate, which makes the optimal integer aperture estimator impractical to use.

4.3.1.3 Integer least squares with ratio test

It is possible to cast the ratio test in terms of an integer aperture estimator (Teunissen & Verhagen, 2004). The acceptance region is (Teunissen, 2003b):

$$\Omega = \left\{ x \in \mathbb{R}^n | \left\|x-\check{x}\right\|^2_{Q_{\hat{a}}} \leq \mu \left\|x-\check{x}_2\right\|^2_{Q_{\hat{a}}}, \, 0 < \mu \leq 1 \right\} \tag{4.32}$$

where $\check{x}$ and $\check{x}_2$ are the best and second-best integer-least-squares estimates of x respectively, and $\mu = {}^1\!/_c$ is the aperture parameter, which controls the size of the acceptance region.

4.3.1.4 Fixed fail rate estimation

In Verhagen (2005a) the concept of *fixed fail rate* integer aperture estimation is discussed; the aperture parameter μ is chosen with regards to the probability distribution of the float ambiguities, so that the probability of resolving the ambiguities to the incorrect integer value is guaranteed to be below a user-defined threshold (assuming unbiased observations). However, for many integer aperture estimators, including the Ratio Test Integer Aperture (RTIA) estimator, it is not possible to directly determine the appropriate value for μ. The following procedure for applying the RTIA estimator with a fixed fail rate is a practical implementation which circumvents the need to directly determine the aperture parameter corresponding to the fixed fail rate:

1. Choose fixed fail rate, P_F;
2. Collect observations and perform the least-squares adjustment;
3. Compute the ratio value $\mu' = \frac{\left\|\hat{a}-\check{a}\right\|^2_{Q_{\hat{a}}}}{\left\|\hat{a}-\check{a}_2\right\|^2_{Q_{\hat{a}}}}$;
4. Generate N samples of float ambiguities $\hat{x}_i \sim N\left(0, Q_{\hat{a}}\right)$;
5. For each sample, determine the integer least squares estimate of the integer values, $\check{x}_i$, and the ratio $\mu_i = \frac{\left\|\hat{x}_i-\check{x}_i\right\|^2_{Q_{\hat{a}}}}{\left\|\hat{x}_i-\check{x}_{i,2}\right\|^2_{Q_{\hat{a}}}}$;

6. Count the number of samples N_f for which $\mu_i \leq \mu'$ and $\check{x}_i \neq 0$;
7. Compute the fail rate with μ' as the critical value: $P_f(\mu') = \frac{N_f}{N}$;
8. If $P_f(\mu') \leq P_F$ then $\bar{a} = \check{a}$, otherwise $\bar{a} = \hat{a}$.

Good results have been found with N as small as a few thousand samples (Verhagen, 2005a); however each sample requires the running of the integer least squares algorithm, so this is still a significant computational burden. Verhagen (2007) describes the creation of a look-up table that allows an approximate value of the threshold parameter to be returned for a given number of observations and fixed fail rate. The RTIA estimator is close to optimal when the fixed fail rate approach is used to determine the aperture parameter (Verhagen, 2005a). This is a promising approach that seems to offer most of the improvement in ambiguity resolution success rate and the ability to choose a fixed fail rate, for minimal increase in computation time.

4.4 Partial ambiguity resolution

The concept of partial ambiguity resolution was first introduced in Teunissen *et al.* (1999). It is based on the idea that the ambiguity resolution success rate is increased when fewer integer constraints are placed upon the solution. This can result in a greater positional availability at the expense of reduced accuracy, and so can help RTK positioning to meet the IMO requirements. Partial ambiguity resolution can be applied to any model and ambiguity resolution technique, but the concept is most clearly demonstrated in the case of geometry-free bootstrapping. The bootstrapped success rate, as given by Equation 4.20, is the product of the probabilities of fixing each ambiguity:

$$P(\check{a}_B = a) = \prod_{i=1}^{n} \left(2\Phi \left(\frac{1}{2\sigma_{\hat{a}_{i|I}}} \right) - 1 \right) \tag{4.33}$$

Each term in the product is less than one, so each successive ambiguity fixed reduces the total success rate. If fixing the next ambiguity brings the success rate below the required limit then the bootstrapping algorithm can be stopped with fewer than the full set of ambiguities resolved. Fixing a subset of the ambiguities will result in a higher overall success rate, at the cost of reduced ambiguity-fixed precision (due to fewer fixed ambiguities).The effect of partial fixing on the precision of the parameters of interest (b) can be computed from the correlation between the ambiguities and the other parameters (Verhagen, 2005a):

$$Q_{\check{b}} = Q_{\hat{b}} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} Q_{\hat{a}\hat{b}} \tag{4.34}$$

Therefore, assuming the ambiguities are fixed correctly, partial fixing will always result in a more precise estimate of the parameters; if the fixed ambiguities have a high variance or low covariance with the parameters then the improvement could be small. As with normal bootstrapping, this procedure is applied to the decorrelated ambiguities; the most efficient way to do this is to use the Z-matrix computed as part of the LAMBDA algorithm. The fixed ambiguities are therefore linear combinations of the "real" ambiguities, and it is not possible to transform

back to obtain fixed values for the real ambiguities, because some of the required parameters remain unfixed.

An example of dual-frequency geometry-free partial fixing is the widelane-only solution. The L1 and L2 ambiguities are combined into widelane ambiguities, which are then fixed with a higher probability than L1 and L2 independently due to the increased precision. However, there is a smaller benefit to the fixed solution due to the long widelane wavelength.

The bootstrapping partial fixing algorithm is more powerful than the widelane because the Z-transformation is determined on-the-fly, based on the float ambiguity covariance matrix, and may achieve greater decorrelation. The LAMBDA Z-transformation will not necessarily choose the widelane linear combination as the most precise ambiguities and, depending on the success rate, it is possible that either the first decorrelated ambiguity (what would be the widelane) is not fixed for every satellite, or both decorrelated ambiguities are be fixed for some satellites. It is most common to analyse the effect of partial ambiguity resolution in the geometry-free case, see e.g. (Teunissen *et al.*, 2000b; Teunissen, 2002; de Jonge *et al.*, 2000).

The effect of this procedure with the geometry-based model is less intuitive, because the linear combinations formed through the LAMBDA Z-transformation are between all satellites and frequencies, rather than just those from the same satellite as is the case in the geometry-free model. Therefore the fixed decorrelated ambiguities are a linear combination of the "real" ambiguities from many satellites and frequencies. However, the technique is still valid: see e.g. (Petovello *et al.*, 2005; Cao *et al.*, 2007; O'Keefe *et al.*, 2006).

It is interesting to compare the effect of additional satellites and frequencies on ambiguity resolution for the geometry-free and geometry-based models. Although Equation 4.33 suggests that additional ambiguities will reduce the success rate, this effect can be mitigated by the benefits of the additional observations. Additional frequencies, and for the geometry-based case, additional satellites, allow greater decorrelation to be achieved and therefore improve the overall bootstrapping success rate. However, the geometry-free model does not take advantage of the additional satellites, and so the bootstrapping success rate is reduced in this case.

This partial ambiguity resolution technique is equally valid with LAMBDA. For integer least squares, additional observations increase the success rate, so the trade-off with partial ambiguity resolution is for a reduced number of more precise observations. The ambiguity subset chosen in this manner is optimal for sequential ambiguity resolution rather than the batch ambiguity resolution of integer least squares; the problem of determining the optimal batch subset is yet to be solved (Teunissen & Verhagen, 2007).

4.5 A new technique for partial ambiguity resolution in the presence of bias

The partial ambiguity resolution technique described in Section 4.4 selects a subset of ambiguities based on their unbiased precision, but real data are biased by residual atmospheric effects and multipath so the subset selected might not have the highest success rate. The success rates computed using the equations in Section 4.2 are in general over-optimistic, because they do not take account of these biases; it is not possible to use the equations in Section 4.2.4 because the

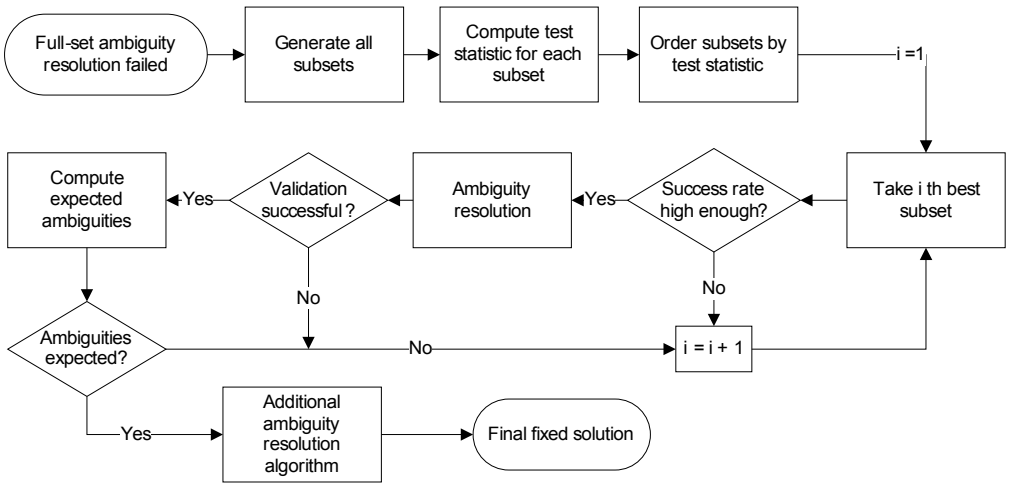


Figure 4.5: Flowchart for the subset ambiguity resolution algorithm

magnitude of the biases is unknown. When the Galileo satellites are operational, there may be a large number of satellites visible at one time, and therefore a large number of double-differenced phase observations. Given that an outlier in a single observation could cause a failure of the single-epoch ambiguity resolution, it seems reasonable to look for a technique that determines a subset of ambiguities to fix based on the real data. This will produce a greater ambiguity resolution success rate in the presence of biased observations, and therefore improved availability and continuity and hence the ability of RTK to meet the most precise IMO requirements for automatic docking. This section describes a new algorithm designed to improve the success rate of ambiguity resolution in the presence of biased observations.

4.5.1 Algorithm description

An algorithm for partial ambiguity resolution in the presence of biased observations, illustrated in Figure 4.5, is now described. The algorithm is only applied if the ambiguity validation test is failed when attempting to resolve the full set of ambiguities. All subsets of the full set above a certain size are generated and ordered according to some criterion. The normal ambiguity resolution and validation procedure is then applied to each subset in turn: the first subset that passes the ambiguity validation test is accepted. It is therefore highly likely that a set of fixed ambiguities will be obtained, and by trying the "best" subsets first it is more likely to be correct. The subsets can be ordered according to many criteria:

- Decreasing ratio $\frac{\|\hat{a}-\check{a}_2\|^2_{Q_{\hat{a}}}}{\|\hat{a}-\check{a}\|^2_{Q_{\hat{a}}}}$: increasing proportional distance of integers from the float values (this has the disadvantage that all subsets must have LAMBDA run on them);
- Decreasing ambiguity precision (determinant of the covariance matrix): start with the most precise observations, which are the most robust to bias;
- External information such as signal to noise ratio (SNR): start with the observations that are least likely to be be biased;

- Increasing ADOP (see Section 4.2.3): this combines precision with the number of observations;
- A combination of the above (e.g. $^{\text{mean SNR}}/_{\text{ADOP}}$: start with the observations that are least likely to contain a bias and are most robust against it).

The ratio criterion requires that LAMBDA must be run on all subsets before they are ordered, whereas with the other techniques the subsets can be ordered first and then LAMBDA run on each subset in turn, until one passes the criteria for acceptance. This results in significantly more processing, and would be impractical with large subsets. Using the SNR results in disproportionate down-weighting of low-elevation satellites, which results in poor geometry for resolving the ambiguities. ADOP combines the precision of the measurements with the satellite geometry, and so could be expected to be superior to using the precisions alone; therefore this technique is used in the processing.

An advantage of this method of partial ambiguity resolution is that the subset of ambiguities that are fixed are "real" ambiguities, unlike the decorrelated linear combination of real ambiguities that is fixed using the partial ambiguity resolution algorithm in Section 4.4. Therefore the accuracy of the ambiguity-fixed position is only affected by having fewer observations, rather than less precise observations. When many satellites are visible, this kind of partial ambiguity resolution may only have a small effect on the precision of the fixed solution.

The algorithm is only run if normal ambiguity resolution has failed, so can only improve the success rate. Every time the algorithm attempts to fix a subset, there are two types of error that can be made when applying the validation test. A Type I error, rejecting the correct ambiguities, generally has little effect, because there are many more such subsets to try to fix. However, a Type II error, accepting the ambiguities when they are false, is much more serious: position errors of several metres can be introduced in the fixed solution. This is therefore much worse than if the algorithm was not run and no position was returned. Tests have shown that Type II errors can occur in a significant proportion of epochs, so modification of the algorithm is necessary to make it viable.

4.5.2 Reducing Type II error probability

Several techniques are applied to reduce the probability of making a Type II error. If, for a given subset, the probability of successful ambiguity resolution (given by Equation 4.20) is below a given value then ambiguity resolution is not attempted; even if the validation test were passed, there would still be a high probability of making a Type II error with this subset. However, this technique alone is not sufficient to reduce the incidence of Type II error to an acceptable level.

The second technique is a fundamental change in concept: instead of allowing the fixed ambiguity set to take any value, it is only accepted if the values obtained are identical to those predicted from the previous epochs. Only epochs in which the entire set was fixed (i.e. not using this algorithm) are used for prediction. Therefore the algorithm can not be used to fix ambiguities independently, but can fill in gaps between fixed epochs, where a bias on one or more observations prevents successful ambiguity resolution, as long as the ambiguities do not change. The predicted values are generated by taking the weighted mode of the values for each

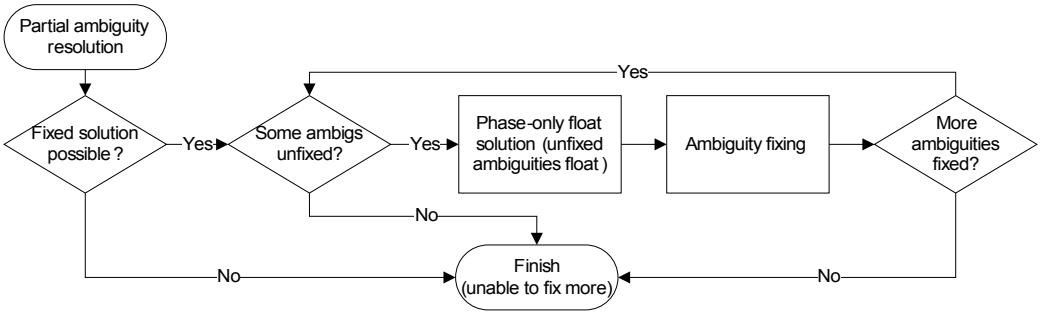


Figure 4.6: Flowchart of the additional ambiguity resolution algorithm

ambiguity over a fixed number of previous epochs. The weighting decays with time, so more recent values have more influence on the predicted value. If the ambiguities change then it will take a few epochs of correct fixing before the predicted values are correct.

With this modification, the algorithm is immune to Type II error except in the case where the expected values are incorrect, and the algorithm fixes an ambiguity subset to these values. The prediction will be incorrect in two circumstances: if in several previous epochs the ambiguities have been fixed to the same incorrect values, or if the ambiguities change and the algorithm fixes a subset to the previous values. The latter seems rather unlikely; the former is more likely but would be perpetuating the error from the standard ambiguity resolution, so is not a new source of error. However, this is still a risk to be aware of: in scenarios where the probability of incorrect ambiguity resolution is high then this technique could be risky. However, this risk could be mitigated by tightening the ambiguity validation criteria.

4.5.3 Improving positional precision

A potential problem with the subset ambiguity resolution algorithm is that there may be too few observations remaining to allow a sufficiently precise final position. A further algorithm is therefore used to fix ambiguities that were not fixed during the initial subset ambiguity resolution process. A least-squares adjustment is carried out, similar to the conventional float solution in that unknown ambiguities are included as parameters, but in this case the fixed phase observations provide the range information. These are more precise than the code observations used in the original float solution, so provide a greater probability of fixing the remaining ambiguities. The subset ambiguity resolution algorithm is then applied to the remaining float ambiguities: this is an iterative process which will end when either all the ambiguities have been fixed or the subset ambiguity resolution fails to fix any more. This is illustrated in Figure 4.6.

4.5.4 Computational efficiency

The primary drawback of the subset ambiguity resolution algorithm is the computation time, which is increased due to multiple repeated attempts to resolve the ambiguities: each subset that is tried requires the re-running of the ambiguity resolution algorithm. The subset ambiguity res-

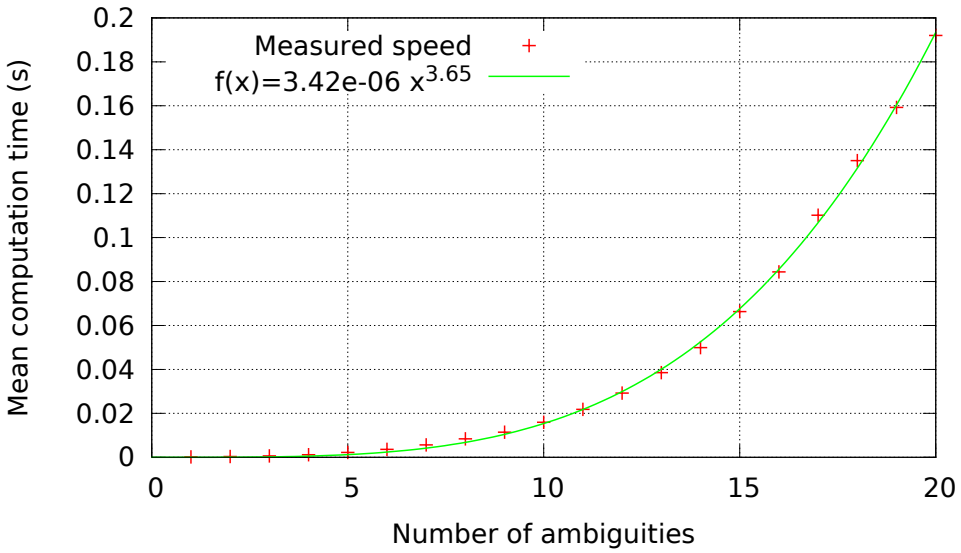


Figure 4.7: LAMBDA computational speed

olution algorithm is only run after the failure of normal ambiguity resolution, so is guaranteed to increase processing time: no ambiguities are fixed without additional processing. The current software implementation is far from optimal, but can be used to estimate how the processing speed scales with additional ambiguities; this will be the critical issue for the use with the additional satellites and frequencies of modernised GPS and Galileo. There are two main ways in which additional satellites increase the processing time: there are more subsets to try, and the LAMBDA algorithm takes longer when processing larger subsets.

Figure 4.7 shows how the LAMBDA algorithm scales with additional ambiguities, based on the average processing times over the data collected in Harwich harbour, as processed on a 1,800 MHz processor. Fitting a polynomial to the points shows the processing time increasing at $n^{3.65}$, which will result in long processing times for large subsets. Although LAMBDA, as a mechanisation of integer least-squares, provides the optimal result in the sense that the probability of successful ambiguity resolution is maximised, is not necessarily optimal for this situation if the processing time is taken into account: if there is limited processing time available to try to fix subsets, it may be more beneficial to try a greater number of subsets with a sub-optimal technique such as widelane-narrowlane linear combinations, as described in Section 4.1.2.5; due to the pre-defined decorrelation adjustment this technique is much more rapid to compute. It may be possible to use a hybrid technique, where the computational efficiency of LAMBDA is increased by storing information, such as the decorrelation adjustment, between epochs.

The number of subsets of a set of size n, not including the full or empty sets, is $2^n - 2$: Figure 4.8 shows how rapidly the number of subsets increases with the size of the full set. However, this does not mean that processing time will necessarily increase at this rate, because not all of these subsets will be processed. The larger number of subsets does give a greater chance for successful ambiguity resolution, but it will be necessary to use a time-based cutoff to stop the algorithm if the ambiguities are not resolved within a certain time. Further investigation would be necessary to determine the optimal subset determination strategy in this case: perhaps all the

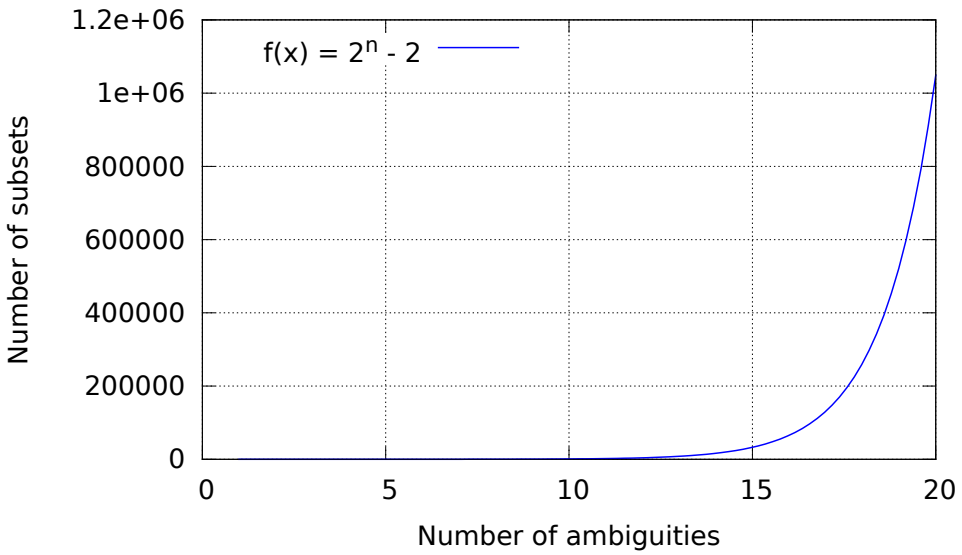


Figure 4.8: Number of subsets

	Total epochs	Success rate
LAMBDA only	2923	0.63%
Subset: 1 s cutoff	4063	0.88%
Subset: 2 s cutoff	4153	0.90%
Subset: 3 s cutoff	4346	0.94%
Subset: no cutoff	4545	0.98%

Table 4.1: Ambiguity resolution success rate from GORE with subset ambiguity resolution algorithm and time-based cutoff

observations from a single satellite should be excluded at once, and a minimum size should be put on the subset. Using the geometry-based algorithm, the probability of successful ambiguity resolution is increased with more satellites, so it is likely that that larger subsets would give better results.

The processing time with the subset ambiguity resolution algorithm is strongly dependent on how rapidly the ambiguities are fixed: if many subsets are tried then the processing time will be very long, but if the first subset is accepted then the additional processing time will be small. Figure 4.9 shows the time taken to perform ambiguity resolution using the subset ambiguity resolution algorithm on real data collected in Harwich harbour, as described in Chapter 6. Note that all epochs where the ambiguities were successfully resolved using LAMBDA alone took less than 0·2 s. No time-based cutoff was applied: the majority of additional epochs were processed within 1 s, and almost all within 3 s, so limiting the available processing time is clearly a viable strategy, and would be necessary for real-time processing. Table 4.1 shows rate of successful ambiguity resolution from GORE with LAMBDA alone, with the subset ambiguity resolution algorithm with a 1 s, 2 s and 3 s processing time cutoff, and with no time cutoff. 25% more epochs have ambiguities successfully resolved even with the 1 s cutoff; only a further 10% are resolved with no cutoff. This shows that most of the benefit of this algorithm

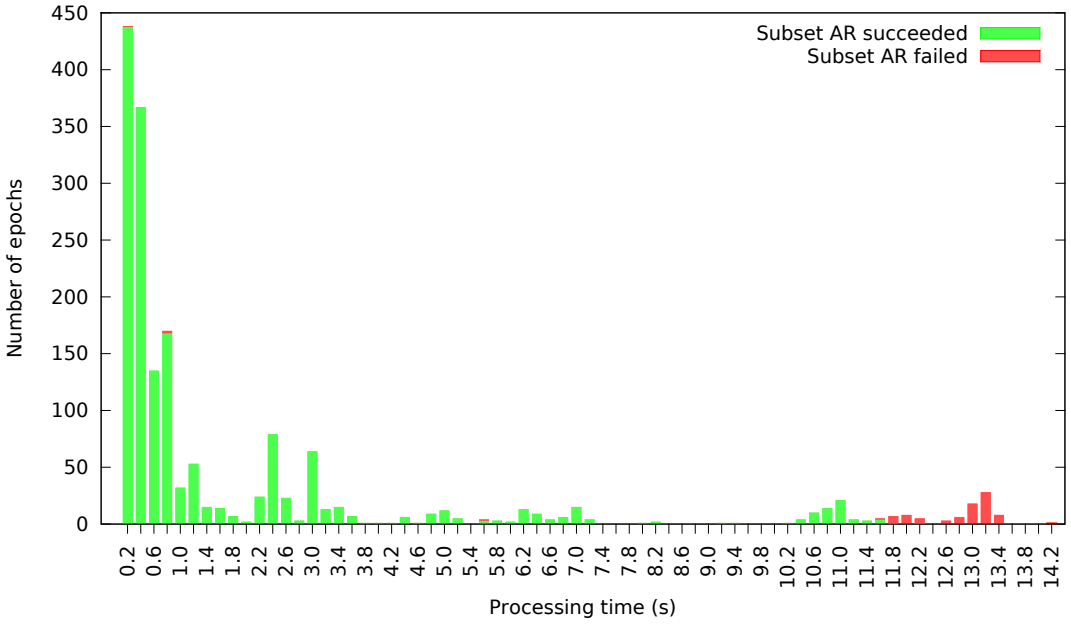


Figure 4.9: Subset AR algorithm processing time from GORE

can be derived with a relatively small increase in processing time. Note that these cutoff limits are very dependent on both the speed of the processor and the efficiency of the programming implementation. It is likely that a commercial receiver would have a slower processor, but a more efficiently-programmed implementation of the algorithm could greatly reduce the processing time. The algorithm is feasible for post-processing, and is well-suited to take advantage of multi-threading, since it relies on many runs of the same function. It could also be used in situations where the rover sends its observations to a base station for processing; cheap computation resources at the base station would make this technique viable.

4.6 Summary

This chapter studies ambiguity resolution and validation, which is a difficult aspect of RTK positioning, which is essential if the IMO requirements for automatic docking are to be met using GNSS. Integer least squares, as mechanised by the LAMBDA technique, is the optimal ambiguity resolution technique, although bootstrapping methods may be able to achieve good results in the geometry-free case with higher computational efficiency. For validating the ambiguities, it is best to use the fixed-fail-rate ratio-test technique, although the computational burden is currently too high for practical use. The ratio test with an experimentally-determined value of the critical value is most commonly used, but this is sub-optimal. A promising development is look-up tables of the optimal ratio test critical value for a given number of satellites and fixed fail rate.

Partial ambiguity resolution is discussed, where a subset of the (decorrelated) ambiguities is resolved, with a greater success rate. A novel technique for partial ambiguity resolution in the presence of biased observations is developed. If normal ambiguity resolution fails, then this algorithm generates subsets of the full sets of ambiguities, orders them according to some criteria and attempts to fix them in turn. In order to reduce the probability of incorrect fixing, each subset

is only attempted if the probability of successful ambiguity resolution is sufficiently high, and the fixed ambiguities are the values that are expected from the previous epochs. A further algorithm is then implemented in which the phase observations with fixed ambiguities are used to help to estimate the remaining unfixed ambiguities in an iterative process. This algorithm increases computational time, but is still viable, particularly if a time-based processing cutoff is applied; a 1 s cutoff still keeps the majority of the benefits. Results from using this algorithm will be demonstrated in Chapter 8.

Chapter 5

Integrity

The IMO requirements include a requirement for integrity, which places a limit on the probability that a biased observation causes a position error greater than the alert limit. This probability is the product of the probability of a bias occurring and the probability of the bias not being detected. Biased observations may be detected and excluded using a technique know as Receiver Autonomous Integrity Monitoring (RAIM), which allows receivers to independent monitor the integrity of received signals without input from an external source. Section 5.1 gives an overview of RAIM algorithms, and a commonly-used testing procedure is described in more detail in Section 5.2. For a bias of a given size in a specific observation, the size of the position error that is guarded against with a specified probability may be computed, as shown in Section 5.3. However, this is not sufficient to determine if the integrity requirements are met: the probability of a bias of the given size occurring must be known, and this is difficult to determine accurately due to the rare occurrence of such events and the corresponding low probability of missed detection of an error in the requirements $\left(10^{-5}\right)$, as discussed in Section 5.4.

There are two problems that further complicate the determination of integrity. The first is the 10 s time-to-alarm period: the bias must exist and be undetected for this whole period before the integrity requirements are breached, so assumptions must be made about how time-correlated biases are. The second is that larger biases are more easily detected, so the most dangerous bias from the perspective of breaching the requirements is one that causes a position error slightly larger than the alert limit. These problems make a direct analysis of whether a set of data meets the integrity requirements difficult to perform. Therefore, in order to avoid these problems, in Section 5.5 the algorithms in Section 5.4 are reversed so that, given a maximum allowable position bias (the alert limit), the size of bias in a given observation required to cause this error can be computed, along with the probability of detecting this bias. This will allow a comparative analysis to be made between the different GPS processing techniques when analysing the results of the data collection. An algorithm is also developed to quantify the relative robustness of ambiguity resolution using different techniques to biased phase observations.

5.1 Integrity algorithms

RAIM algorithms detect outliers by performing a consistency check between redundant observations; at least 5 visible satellites are required to detect an error. There are two main classes of integrity algorithms: snapshot algorithms, which use only the data from the current epoch, and recursive algorithms, which also use information from previous epochs. This project focuses on

single-epoch positioning so only snapshot techniques are considered here; recursive RAIM algorithms are generally used with a Kalman filter for positioning. Many RAIM techniques can be used for both snapshot and recursive solutions with little modification, but there is an additional need with recursive schemes to guard against slowly growing errors, which can be masked in the test statistics.

There have been many proposed code RAIM algorithms, many of which are equivalent (Brown, 1996). A commonly-used procedure is the Detection, Identification and Adaptation algorithm, described in Section 5.2, which is based on the theory of statistical hypothesis testing. An advantage of this technique is that it allows a simple evaluation of the position error that can be protected against with a specific probability, as described in Section 5.3.

Phase integrity is complicated by the ambiguity resolution step. Most carrier RAIM (CRAIM) techniques proceed by assuming that the ambiguities have been correctly fixed in a previous epoch; the data from the current epoch are then tested for the presence of a large outlier such as a cycle slip (Pervan *et al.*, 1998; Abousalem *et al.*, 1994), using similar algorithms to code RAIM. This is not compatible with the single-epoch approach, where the ambiguities are resolved every epoch. If the ambiguities are resolved correctly then biases in the ambiguity-fixed solution can be detected and excluded, but it is not possible to detect an outlier in single-epoch phase observations prior to fixing the ambiguities, because the bias propagates directly into the real-valued integer ambiguity estimate. This undetectable bias will then reduce the probability of successful ambiguity resolution.

5.2 Detection, identification and adaptation

The commonly used Detection, Identification and Adaptation (DIA) testing procedure (Teunissen, 1990, 1998b; Salzmann, 1995) uses the theory of statistical hypothesis testing to detect mis-specifications of the functional model. In the null hypothesis, the observations follow the least-squares model, $y = Ax$, and are normally distributed about the true value:

$$H_0 : E\{y\} = Ax;\ D\{y\} = Q_y \tag{5.1}$$

The three steps in the DIA algorithm are:

1. *Detection:* Perform an overall model test to determine if an unspecified model error has occurred. If this test is passed then no error is detected this epoch and the subsequent steps are not performed;

2. *Identification:* Determine the source of error by testing H_0 against various alternative hypotheses;

3. *Adaptation:* Adapt the null hypothesis to compensate for the detected error (either adopt the appropriate alternative hypothesis or exclude the faulty observation).

The algorithm excludes at most one outlier each time it is run, and is iterated until no further outliers are detected.

5.2.1 Detection

The first step determines if there is sufficient confidence to accept the null hypothesis, without the need for a specific alternative hypothesis; the null hypothesis is opposed with the most relaxed alternative hypothesis, which leaves the observables completely free:

$$H_0 : E\{y\} = Ax;\ H_a : E\{y\} \in \mathbb{R}^m \tag{5.2}$$

The test statistic for detection of a model error is then:

$$T_{m-n} = \frac{\hat{v}^T Q_y^{-1} \hat{v}}{m-n} \tag{5.3}$$

where $\hat{v}$ is the vector of least-squares residuals, m is the number of observations and n is the number of parameters: the test statistic is the sum of the weighted squares of the residuals divided by the redundancy. The test statistic has a central F-distribution with $m-n$ and ∞ degrees of freedom. The null hypothesis is rejected when:

$$T_{m-n} > F_{\alpha_{m-n}}(m-n, \infty, 0) \tag{5.4}$$

where α_{m-n} is the significance of the test.

5.2.2 Identification

If the detection test fails, the next step is to search for a specific model error. In the one-dimensional case, where the magnitude of the error can be specified by a scalar ∇, the alternative hypothesis is:

$$H_a : E\{y\} = Ax + c\nabla \tag{5.5}$$

where c is a vector that specifies the model error. The choice of c depends on the kind of model errors expected. To test for the presence of an outlier in the ith observation, c_i is taken as

$$c_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T \tag{5.6}$$

i.e. the null vector with the ith value unity. This test is usually performed for each $i = 1, \ldots m$, with corresponding test statistics w_i, to detect the presence of an outlier in any single observation. This process is known as *data snooping* (Baarda, 1968). The test statistics are then:

$$w_i = \frac{c_i^T Q_y^{-1} \hat{v}}{\sqrt{c_i^T Q_y^{-1} Q_{\hat{v}} Q_y^{-1} c_i}} \tag{5.7}$$

which have a standard normal distribution under the null hypothesis. The largest w_i is compared to the critical value to test for a gross error. Therefore a gross error on the jth observation is assumed when:

$$|w_j| \geq |w_i|\ \forall i \quad \text{and} \quad |w_j| > N_{\alpha/2}(0,1) \tag{5.8}$$

where α is the significance of the test. If the covariance matrix is diagonal, the test statistic simplifies to:

$$w_i = \frac{\hat{v}_i}{\sigma_{\hat{v}_i}} \tag{5.9}$$

which is the residual of the ith observation divided by the standard deviation of this residual.

This technique is valid when the covariance matrix of the observations Q_y reflects the population from which they are drawn, i.e. the unit variance $\sigma_0^2 = 1$. If this is not the case then a different test statistic is used (Cross, 1994):

$$\tau_i = \frac{\hat{v}_i}{\sigma_0 \sigma_{\hat{v}_i}} \tag{5.10}$$

which has the tau distribution:

$$\tau_\nu = \frac{\sqrt{\nu} t_{\nu-1}}{\sqrt{\nu - 1 + t_{\nu-1}^2}} \tag{5.11}$$

where ν is the number of degrees of freedom in the least-squares adjustment, and $t_{\nu-1}$ is a t-distribution with $\nu - 1$ degrees of freedom.

5.2.3 Adaptation

In principle, the alternative hypothesis becomes the new null hypothesis: another parameter is introduced into the least-squares solution to estimate the size of the bias. In practice a new adjustment is usually performed with the biased measurement excluded. The DIA procedure is then run again to enable the detection and exclusion of further biases.

5.3 Reliability

The DIA procedure described in Section 5.2 allows the possibility of the exclusion of biased observations. This section describes a technique for computing the probability that a given bias will be detected (internal reliability), and the effect that an undetected bias will have on the positioning solution (external reliability).

5.3.1 Internal reliability

When performing hypothesis testing there are four possible outcomes: the null hypothesis is correct and is accepted, the alternative hypothesis is correct and is accepted (i.e. bias correctly excluded), the null hypothesis is correct and is rejected (un-biased observation incorrectly excluded, a Type I error), and the alternative hypothesis is correct and is rejected (biased observation not excluded, a Type II error). α is the probability of making a Type I error and β is the probability of making a Type II error; $\gamma = 1 - \beta$ is the power of the test. The idea is to compute the size of the bias in a given observation that can be detected with probability γ, known as the Marginally Detectable Bias (MDB). Larger biases will have a higher probability of detection; smaller biases will have a lower probability of detection. The MDB can be computed for each model error.

The hypotheses for each measurement y_i is (Cross, 1994):

$$H_0 : y_i = \bar{y}_i + \epsilon_i;\ H_a : y_i = \bar{y}_i + \epsilon_i + \Delta_i \tag{5.12}$$

where ϵ_i is a normally distributed error and Δ_i is a gross error. If H_0 is true then the test statistic w_i for observation i will be distributed with a standard normal distribution; if H_a is true then w_i will have mean:

$$\delta_i = \frac{\Delta_i}{\sigma_{\hat{d}_i}} \tag{5.13}$$

where $\hat{d}_i = y_i - \hat{y}_i^C$. $\hat{y}_i^C$ is the ith observed quantity computed from the parameters derived from a least-squares computation of all the observations excluding y_i (Cross, 1994).

If values for α and β are specified then an upper bound for δ_i can be given as:

$$\delta_i^u = a + b \tag{5.14}$$

where a and b found from the inverse normal distribution: a from 2-tailed test with probability α; b from 1-tailed test with probability β, i.e. $a = \text{norminv}\,(\alpha/2)$, $b = \text{norminv}(\beta)$. Therefore the MDB is

$$\Delta_i^u = \delta_i^u \sigma_{\hat{d}_i} \tag{5.15}$$

It is shown in Cross (1994) that

$$\sigma_{\hat{d}_i} = \left(c_i^T Q_y^{-1} Q_{\hat{v}} Q_y^{-1} c_i\right)^{-\frac{1}{2}} \tag{5.16}$$

with c_i defined as in Equation 5.6. When Q_y is diagonal, this simplifies to

$$\sigma_{\hat{d}_i} = \frac{\sigma_i^2}{\sigma_{\hat{v}_i}} \tag{5.17}$$

so the MDB is

$$\Delta_i^u = \frac{\delta_i^u \sigma_i^2}{\sigma_{\hat{v}_i}} \tag{5.18}$$

5.3.2 External reliability

Internal reliability characterises the size of the MDB, that can be detected with specified probability. It is important to determine the effect that a bias, if undetected, would have on the positioning solution. This is known as *external reliability*. For each observation i, the bias vector:

$$p_i^{\Delta_i^u} = \begin{bmatrix} 0 \\ \vdots \\ \Delta_i^u \\ \vdots \\ 0 \end{bmatrix} \tag{5.19}$$

is the null vector with the corresponding MDB at the ith position. The effect that this has on the position solution is then:

$$\left(A^T W A\right)^{-1} A^T W p_k^{\Delta_i^u} = \begin{bmatrix} \delta X \\ \delta Y \\ \delta Z \\ \delta t \end{bmatrix} \tag{5.20}$$

The ECEF Cartesian coordinates are often rotated into the local topographic frame; the length of the $(\delta E, \delta N)$ vector is known as the Horizontal Protection Level (HPL), which is the level of integrity that can be guaranteed with the specified probabilities of making a Type I or Type II error. This is used to calculate RAIM availability.

5.4 Integrity availability

The IMO integrity requirements are specified in terms of the alert limit, integrity risk and time-to-alarm; the position error must only exceed the alert limit for a period greater than the time-to-alarm with a probability equal to the integrity risk. The integrity risk depends on both the probability of an outlier sufficiently large to cause a position error greater than the alert limit occurring, and the probability of it being detected (Ochieng *et al.*, 2002):

$$P\left(\text{Integrity error}\right) = P\left(\text{Sufficiently large error occurring}\right) \times P\left(\text{Missed detection}\right) \tag{5.21}$$

The IMO continuity requirements specify that integrity monitoring must be available (see Chapter 2); if the RAIM algorithm is insufficient to protect against an integrity error then a system cannot be used for safety-of-life positioning. In Section 5.3, it is shown how, given the probability of a false alarm, α, and the probability of missed detection, β, it is possible to compute the magnitude of the MDB in each observation guarded against, and the effect of this on the position. If this position error is less than the alert limit then RAIM is available; if it is larger then the RAIM algorithm is insufficiently robust and the position cannot be guaranteed free of errors of the given size with high enough probability. It therefore remains to determine appropriate values for α and β to make sure that the integrity requirements are met.

As shown in Equation 5.21, β depends upon $P\left(\text{Sufficiently large error occurring}\right)$. Usually, all biases are assumed to be sufficiently large, and a probability of occurrence is assigned. Events that cause an error of this magnitude are rare, so it is difficult to determine this value; the time-to-alarm also complicates matters because short-duration biases can be disregarded.

α is usually taken as a small value in the assumption that a false alarm prevents positioning; it would be possible to take this as a much larger value if the RAIM algorithm includes a fault exclusion step, because excluding a single observation generally has only a small effect on the solution.

After the selection of α and β, the MDB for each satellite and the error this MDB would cause on the positioning solution is then computed and compared to the alert limit.

5.5 Robustness to bias

In this section, the well-known theory described in the Section 5.3 is reversed so that, given a maximum allowable position bias (the alert limit), the size of bias in a given observation required to cause this error, and the probability of detecting this error, can be computed. This will allow a comparative analysis to be made between the different GPS processing techniques when analysing the results of the data collection. Therefore, although it will not be possible to determine if the integrity requirements are met (which would require knowledge of the probability of biases of specific size to be known), an easy comparison of the robustness of different techniques compared to given alert limits will be possible. A technique is also developed to allow an assessment of the robustness of ambiguity resolution to biased observations.

5.5.1 Code

In this section, the computation in Section 5.3 is performed in reverse: given the two-dimensional position error equal to the alert limit K, the magnitude of the MDB in a given observation required to produce this position error is calculated, and the probability of not detecting this MDB, β, is determined. The probability of erroneously rejecting an observation, α, is relatively less important (it matters less if a good observation is excluded than if a bad one is included) and is kept fixed at an arbitrary value, e.g. $\alpha = 0{\cdot}01$.

The Cartesian ECEF parameter vector is rotated to the local topocentric frame by the rotation matrix R, where

$$\begin{bmatrix} \delta_E \\ \delta_N \\ \delta_h \\ \delta_t \end{bmatrix} = R \begin{bmatrix} \delta_X \\ \delta_Y \\ \delta_Z \\ \delta_t \end{bmatrix} \tag{5.22}$$

Combining Equation 5.20 and Equation 5.22, the effect on the rotated parameters of a bias, Δ_i^ϵ, of size ϵ in the ith observation is:

$$R\left(A^T W A\right)^{-1} A^T W p_i^\epsilon = R \begin{bmatrix} \delta_X^\epsilon \\ \delta_Y^\epsilon \\ \delta_Z^\epsilon \\ \delta_t^\epsilon \end{bmatrix} = \begin{bmatrix} \delta_E^\epsilon \\ \delta_N^\epsilon \\ \delta_h^\epsilon \\ \delta_t^\epsilon \end{bmatrix} \tag{5.23}$$

where δ^ϵ is the error caused by a bias of size ϵ and p_i^ϵ is defined as in Equation 5.19. $p_i^\epsilon = \epsilon p_i^1$, so this is equivalent to

$$\epsilon R\left(A^T W A\right)^{-1} A^T W p_i^1 = \epsilon R \begin{bmatrix} \delta_X^1 \\ \delta_Y^1 \\ \delta_Z^1 \\ \delta_t^1 \end{bmatrix} = \epsilon \begin{bmatrix} \delta_E^1 \\ \delta_N^1 \\ \delta_h^1 \\ \delta_t^1 \end{bmatrix} \tag{5.24}$$

Define the plan error caused by a bias of size ϵ as $e^\epsilon := \sqrt{\left(\delta_E^\epsilon\right)^2 + \left(\delta_N^\epsilon\right)^2}$. From Equation 5.24 it follows that $e^\epsilon = \epsilon e^1$, where e^1 can be computed by taking $\epsilon = 1$. The requirement on

the bias ϵ is that the plan error must equal the alert limit, i.e. $e^{\epsilon} = \epsilon e^{1} = K$. Therefore the size of the MDB that produces a plan error equal to the alert limit is:

$$\Delta_i^u = \epsilon = \frac{K}{e^1} \tag{5.25}$$

It now remains to compute the probability of not detecting a bias of this size, β, for a given α. Assuming that the covariance matrix of the parameters is diagonal, re-arranging Equation 5.18 and substituting in Equation 5.14 gives:

$$b = \frac{\Delta_i^u \sigma_{v_i}}{\sigma_i^2} - a \tag{5.26}$$

where $a = \text{norminv}\,(\alpha/2)$, $b = \text{norminv}\,(\beta)$, σ_i is the precision of the ith observation and σ_{v_i} is the precision of its residual. σ_i and σ_{v_i} are available from the least-squares solution; this allows the computation of β for a given Δ_i^u.

5.5.2 Phase

As discussed in Section 5.1, in a single epoch it is not possible to detect phase biases prior to ambiguity resolution, but undetected biases can prevent successful ambiguity resolution. The following technique is designed to analyse the robustness of single-epoch ambiguity resolution to phase biases. For each epoch when the ambiguities are successfully resolved:

1. Compute the lower bound of the success rate in the presence of a bias (see Equation 4.24 in Chapter 4) with a bias of a given size (e.g. $1/8$ cycle) for each observation in turn;
2. Take the observation with the median success rate; this is representative of the average effect that a bias has on the overall success rate for all observations;
3. For this observation, use an iterative root-finding algorithm to determine the minimum size of bias required to cause the ambiguity resolution to be unsuccessful.

This gives a measure of the bias robustness of the positioning solution at each epoch. Taking the median value is required to reduce computation time. If this technique is performed on real data, biases will already be present in the observations due to atmospheric decorrelation, orbit errors and multipath, so this can only estimate the average *additional* bias.

5.6 Summary

Integrity is an important IMO requirements, and this chapter describes a commonly-used technique that can detect and exclude outliers in observations, and hence provide integrity. Equations are given for calculating the magnitude of the bias in a given observation that can be detected with a given probability, and the effect of this bias on the position. However, in order to determine if the integrity requirements are met, the probability occurrence of biases of different sizes must be known; this is not possible with any accuracy due to the infrequency of occurrence of such biases. Therefore equations are developed from the described algorithm to allow the

reverse computation: given a specific position error, the size of the bias in a given observation required to cause this error, and the probability of detecting this bias, can be computed. This allows a comparison to be made between different processing techniques when compared to the different IMO requirements. In addition, an algorithm is developed for assessing the robustness of ambiguity resolution to biased observations.

Chapter 6

Data collection

The data collection exercise was designed to obtain data in a port navigation and automatic docking environment, with an independent truth model, to allow an assessment of GPS positioning with regards to the IMO requirements. The truth model chosen was an automatic tracking total station which, when set up with known location and orientation, tracks and rapidly measures the angle and range to a prism mounted underneath the GPS antenna. Section 6.1 describes the plan for the main exercise. Several preliminary exercises were performed to assess the methodology and equipment prior to the main exercise, and are described in Section 6.2. Section 6.3 describes the main data collection exercise, and the generation of the truth model. Analysis of the data with respect to the IMO requirements is performed in Chapter 8.

6.1 Data collection plan

The aim of the data collection exercise is to obtain GPS data in a realistic environment corresponding to the IMO requirements, with an independent truth model to verify the positions obtained. This will allow the results generated from the processing of the data to be directly compared to the IMO requirements, and each GPS processing technique to be assessed to determine how it may be used to achieve the requirements. The three IMO requirements studied in this thesis are port approach, port navigation and automatic docking. The GLAs made a vessel, *THV Alert*, and a jetty available in Harwich for the purpose of the main experiment. It was possible to mount a GPS receiver on the stern of *Alert,* and to set up several GPS reference stations nearby; the Ordnance Survey, who run a national network of GPS reference stations, agreed to supply 1 Hz data from nearby reference stations. This variety of reference stations allows processing over a range of different baseline lengths. It was desirable to record data from the modernised GPS L2C signal, the only modernised signal available at the time of the experiment, so that the performance improvement of this new signal could be assessed: three GR-3 GNSS receivers, capable of tracking L2C, were loaned by Topcon UK for use in the experiment, and the data collection time was chosen to maximise the number of visible modernised satellites.

There were several possible options for a truth model. A laser tracker, when set up over a known point, can give very frequent and precise range and angle measurements to a special prism, but the range is limited to around 80 m, which would significantly limit the experiment. Another option is photogrammetric techniques, where targets on the ship are tracked by several video cameras and positions generated, but this is equally range-limited. The technique decided upon to generate the truth model is to use automatic tracking total stations. These are devices

that, when set up over a known point and with known orientation, measure the angle and range to a prism, and can be mounted directly under the GPS receiver. The automatic tracking total station automatically tracks the prism as it moves, and has a range of up to 1 km; this solution has also been used before to provide a truth model for GPS (Ueno, 1999). Topcon made two GPT-9000A automatic tracking total stations available for the main experiment. Several preliminary exercises were carried out in order to test and develop the technique for generating positions from a total station that correspond to the GPS positions. In Ueno (1999), a best-fit adjustment was made between the two sets of positions, which could potentially absorb a constant offset or rotation error in the GPS positions: the aim for this experiment is to obtain time-tagged total station positions synchronised to GPS time, and thereby make this adjustment unnecessary.

6.2 Preliminary data collection

6.2.1 London

The first experiment was carried out on the 23rd August 2007 in London, near UCL, as an initial assessment of the tracking capabilities of the total station. A 360° prism was mounted on a pole, with a GPS receiver mounted directly above. The total station was set up over a station that was subsequently surveyed with GPS to enable the prism positions to be put in the same reference frame as the GPS positions; a second surveyed station was used to determine the orientation of the total station. The total station then tracked the prism as it was moved about and the GPS receiver collected data. A second GPS receiver was set up nearby as a reference station; the position of this was determined by RTK positioning from a nearby Ordnance Survey reference station.

After the data had been collected and the positions of the stations established, an attempt was made to directly compare the prism and GPS positions. It had previously been determined that there were no total stations that could accept time input from a GPS receiver, so the prism positions were time-tagged with the total station clock time, and it was necessary to solve for the clock offset from GPS time. However, the total station rounded the time-tags on the prism positions to the nearest second. Whilst this is sufficient for the intended purpose of the automatic tracking feature - machine control or one-man surveying - significant time error was introduced in the time-stamped positions for a moving target, so a direct comparison between the prism and GPS positions was not possible.

6.2.2 Margate

The London trial established that direct recording by the total station would not produce prism positions with time-tags accurate enough to compare to the GPS positions, due to the time-tags being rounded to the nearest second. Therefore a new approach was developed whereby the output from the total station was streamed directly to a laptop via the serial port, where it was time-stamped with millisecond precision: this will produce a position error of a few millimetres, depending on the speed of the target. The initial plan was to have an additional GPS receiver streaming timing data to the laptop as well: software is freely-available that can be used to synchronise the laptop clock with this GPS output, thereby keeping the total station time-stamps

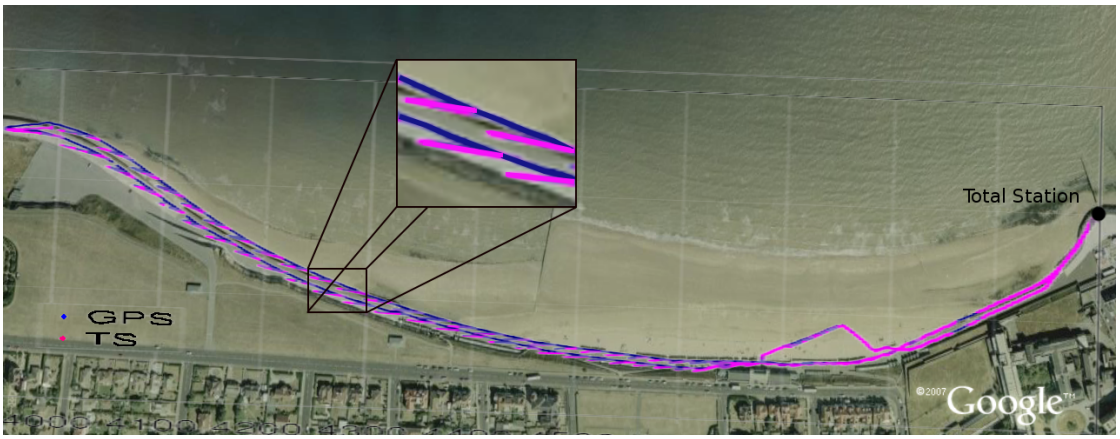


Figure 6.1: Plan view of Margate data collection

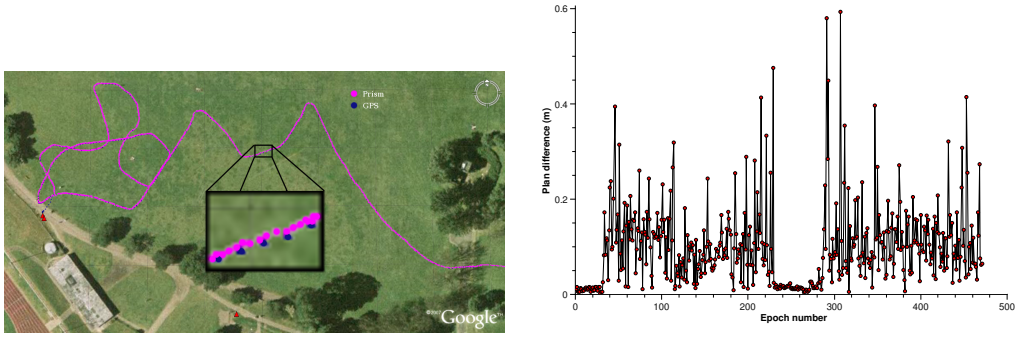
close to GPS time and eliminating the need to solve for the clock offset. However, this was not possible because both techniques require a serial port to stream the data, but the the available laptops had at most one serial port, and neither software worked through a serial to USB adaptor. Therefore the laptop clock was only synchronised with GPS time at the start of the experiment and thereafter left to drift.

The second preliminary data collection exercise was carried out on the 17th October 2007 in Margate, with the aim of testing this new technique for time-stamping the total station data near to GPS time using a laptop. Margate was chosen because there is a long shoreline, which provides a semi-marine environment for the test, and good visibility along the promenade for over 1 km, so the ability of the total station to track over long distances could be tested.

Figure 6.1 shows the promenade at Margate where the experiment was performed. The total station was set up at the East end as shown, next to a GPS reference receiver. The total station position and secondary station required for orientation were surveyed with GPS before the commencement of the experiment. The pole with the rover and the prism was carried along the promenade to a range of around 1 km, whilst being tracked by the total station. As the range increased over 400 m it became apparent that there was a problem with the tracking: the angular tracking was not smooth and the instrument cross-hairs lagged behind the prism and then "jumped" to catch up. The results of this can be seen in Figure 6.1, where the jumps in the prism positions (pink) are clearly visible (see the expanded inset). Attempts were made to rectify this problem at the time of the experiment, but these were unsuccessful. It was determined that the total station tracking capability and robustness degraded with range, so it was decided to limit future data collection to shorter ranges, of less than 400 m. Apart from the tracking problems, the new system of streaming the output positions to the total station for time-stamping with the total station clock time worked well.

6.2.3 St Albans

Investigation of the total station settings at Topcon after the Margate data collection exercise discovered certain settings on the instrument that were changed to provide smoother tracking. A further experiment was therefore carried out on the 22nd February 2008 in Verulamium Park,



(a) Plan view

(b) Plan difference between GPS and prism positions

Figure 6.2: St Albans data collection

St Albans, to test the new total station tracking. The setup is shown in Figure 6.2a: the GPS reference station and total station were set up as usual, and the data output to the laptop. Unfortunately, due to several problems during the setting up of the experiment, there was insufficient laptop battery life to record more than 10 minutes of data. However, it appeared that the new total station settings worked: the tracking remained smooth and the cross hair was fixed on the centre of the prism for the duration of the trial. It was therefore possible for the first time to directly compare the prism and GPS positions without the time-stamping problems encountered in the London trial or the severe jumps encountered in the Margate trial.

Figure 6.2b shows the plan difference between the prism and GPS positions for each epoch, after processing with the technique described in Section 6.3.2.1. The receiver was stationary for epochs $0 - 30$ and $230 - 280$, during which time the two systems show very good correspondence (within 2 cm). However, the agreement is significantly worse, at the 10 cm level, when the rover is moving. This could be due to tracking errors, unknown delays in the total station measurements or time synchronisation errors. Another possible explanation is that the rover pole was not vertical whilst being carried over the uneven ground, producing an offset between the plan positions of the GPS antenna phase centre and the centre of the prism.

6.3 Main data collection exercise in Harwich Harbour

6.3.1 Description of data collection

The preliminary experiments described in Section 6.2 developed the technique of using a total station as a truth model, with time-stamped observations that can be directly compared to the GPS positions after solving for the clock offset and interpolating to the same time, as described in Section 6.3.2.1. In the last experiment, in St Albans, which unfortunately only had 10 minutes of data, agreement between prism and GPS positions was achieved at the 10 cm level when moving and the 2 cm level when stationary. Although this is still a significant error, the tracking and time synchronisation problems with the total station were been solved to the best level possible. It was decided to use two total stations in the final experiment, so that the errors could be mitigated by averaging between them, and to keep the total station ranges as short as possible.

Figure 6.3 shows the initial plan for the Harwich data collection. The data collection is split

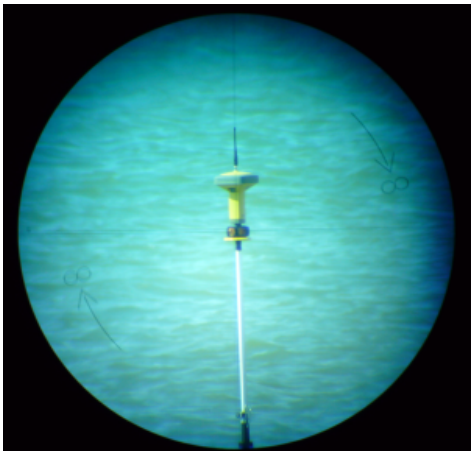
Figure 6.3: Initial design of Harwich data collection exercise, showing *Alert*'s *course and total* station locations with 400 m range circles

into two separate phases; each phase is monitored by both total stations, with the total stations being moved between phases. The first phase (yellow) recreates a port navigation environment, and the second phase (green) recreates a port navigation and automatic docking. The two tracks together give a complete approach and docking scenario. If time permitted then the pink and blue tracks could be used to collect further data.

The time of the main exercise, the evening of 1st June 2008, was chosen to allow the recording of the greatest possible number of L2C observations, as shown in Figure B.1 in Appendix B; the stations were set up on the previous day and a preliminary test using *Alert* was carried out in the morning. Three Topcon GR-3 GPS receivers, capable of tracking the modernised signal on L2, were used: the rover mounted on the stern of *THV Alert* (SHIP); a reference station on the roof of the nearby Trinity House building (BASE); and a second reference station in an open field 10 km away (WIX). These provided a range of nearby reference stations, and additional data were obtained from nearby Ordnance Survey reference stations: see Appendix B for a map (Figure B.2) and further details of the stations (Table B.1). A prism was mounted underneath the antenna on *THV Alert* and tracked by total stations from the shore, as shown in Figure 6.4, in the same manner as with the preliminary exercises. All GPS receivers were operated at a data rate of at least 1 Hz.

During the preliminary exercise in the morning, total station TS1 was set up to monitor the yellow path in Figure 6.3. However, due to shallow water, *Alert* was not able to approach closer than 700 m, which proved to be too far for the total station to track reliably in the rainy weather conditions and with the spray from the sea. It was therefore decided that this section would have to be abandoned and the date collection concentrated on the next phase, where the water was deeper closer to shore and the range to the total stations shorter.

(a) Total station tracking *THV Alert*



(b) GPS receiver and prism on *THV Alert*, viewed through the total station telescope

Figure 6.4: Photographs of Harwich data collection

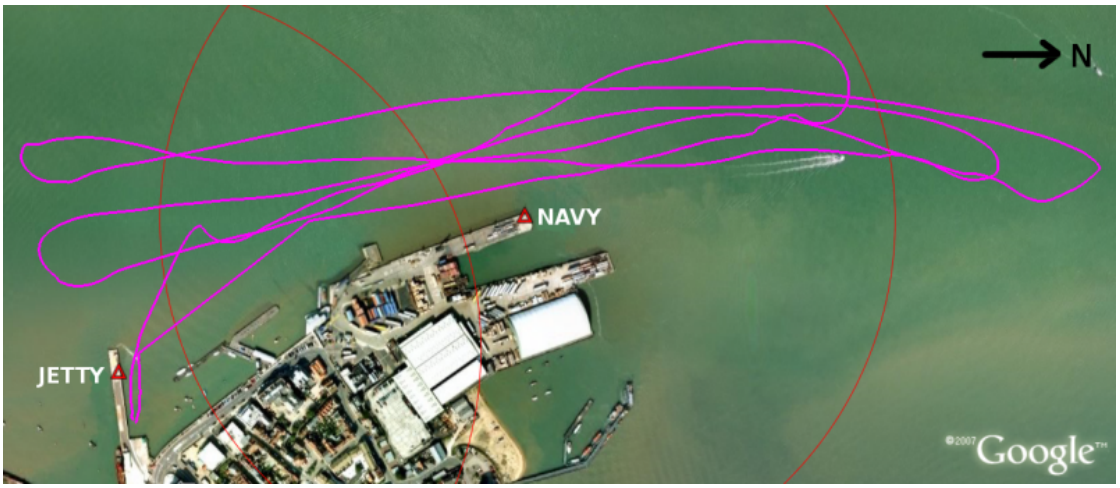


Figure 6.5: Plot of *THV Alert*'s course, showing total station locations with 400 m range circles

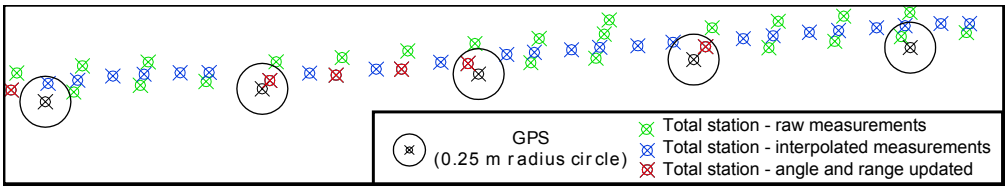


Figure 6.6: Plot of GPS and prism positions during Harwich data collection

Figure 6.5 shows the plotted true tracks of *Alert* during the final exercise in the evening. Automatic tracking total stations were set up at JETTY and NAVY. The pink track shows the course of *THV Alert* as she made several approaches to the jetty, recreating port approach and docking. *Alert* made several runs across in front of the total stations in order to allow the prism to be tracked at a reasonably close range. As she turned, it was necessary to move the pole with the GPS receiver and prism to the side of the ship closest to the shore to maximise the prism visibility: this resulted in short spans of data during the turns being excluded from the final processing. As many runs were performed as was possible in the time available.

6.3.2 Generation of truth model

6.3.2.1 Technique

A total station measures the angle and distance to the prism at a non-constant rate of $\sim$ 3 Hz: these observations were used to provide a truth model for the GPS positions, which were projected to a plane using a Transverse Mercator projection, and transformed to the local coordinate system. The positions obtained when the ship was turning and the rover was moved to the other side of the ship were excluded. The raw prism positions (green and red positions in Figure 6.6) are much less smooth than expected from the motion of the ship. Analysis of the observations showed that the total station did not always update the range or angle observations before output: of the 38,427 sets of observations from both total stations, 78% had updated range and 44% updated horizontal angle; both range and angle were updated in 24% of epochs, and neither were updated in 2% of epochs. Vertical angle was not considered due to the slow rate of change. The range is measured to the nearest centimetre, and the angle to 5" ($\approx$ 1 cm at 400 m); the target was moving at 1 m/s, so it is unlikely that the true observation was constant to this precision for any significant period of time. The raw prism positions therefore do not represent the true path of the ship, because of the substantial error introduced in the measurement update process. Those positions with both observations updated (red), show less noise but are too sparse to allow comparison to the 1 Hz GPS data.

In order to obtain a smooth track with enough positions for comparison with GPS, the un-updated measurements were estimated by linear interpolation between the nearest updated measurements: most positions are therefore determined from a combination of interpolated and observed parameters. Figure 6.6 shows that these positions (blue) exhibit much less noise than the raw positions. The subsequent analysis is performed on the union of the positions with both observations updated (red) and positions with at least one interpolated observation (blue).

The total stations did not measure at the same instant as the GPS receiver, so each set of total station positions were linearly interpolated in time to match the GPS epoch. If the GPS position

at epoch l is:

$$P\left(l\right)_{\text{GPS}} = \left(E\left(l\right)_{\text{GPS}}, N\left(l\right)_{\text{GPS}}\right) \tag{6.1}$$

then the interpolated prism position to match the GPS position is:

$$P(l)_{\text{TS}} = \left(E\left(k\right)_{\text{TS}} + a\left(E\left(m\right)_{\text{TS}} - E\left(k\right)_{\text{TS}}\right), N\left(k\right)_{\text{TS}} + a\left(N\left(m\right)_{\text{TS}} - N\left(k\right)_{\text{TS}}\right)\right) \tag{6.2}$$

where:

$$P\left(k\right)_{\text{TS}} = \left(E\left(k\right)_{\text{TS}}, N\left(k\right)_{\text{TS}}\right) \tag{6.3}$$

is the nearest prism position in the past;

$$P\left(m\right)_{\text{TS}} = \left(E\left(m\right)_{\text{TS}}, N\left(m\right)_{\text{TS}}\right) \tag{6.4}$$

is the nearest prism position in the future; and

$$a = \left(t\left(l\right) - t\left(k\right)\right) / \left(t\left(m\right) - t\left(k\right)\right) \tag{6.5}$$

The total station time-stamps are based on a local clock rather than GPS time, so a least-squares adjustment was performed to solve for the difference between the time systems. This was modelled as a constant offset δt and linear drift Δt of each local clock from GPS time:

$$t\left(i\right)_{\text{GPS}} = t\left(i\right)_{\text{TS}} + \delta t + \left(t\left(i\right)_{\text{TS}} - t\left(1\right)_{\text{TS}}\right)\Delta t \tag{6.6}$$

where $t\left(i\right)$ is the time of the current epoch and $t\left(1\right)$ is the time of the first epoch. δt was kept small by synchronising the local clocks to GPS time at the start of the exercise. For each epoch l the ship's velocity was estimated from the interpolated prism positions:

$$V\left(l\right) = \left(\frac{E\left(l+1\right) - E\left(l\right)}{t(l+1) - t(l)}, \frac{N\left(l+1\right) - N\left(l\right)}{t(l+1) - t(l)}\right) = \left(V_E\left(l\right), V_N\left(l\right)\right) \tag{6.7}$$

The least-squares adjustment is set up as follows:

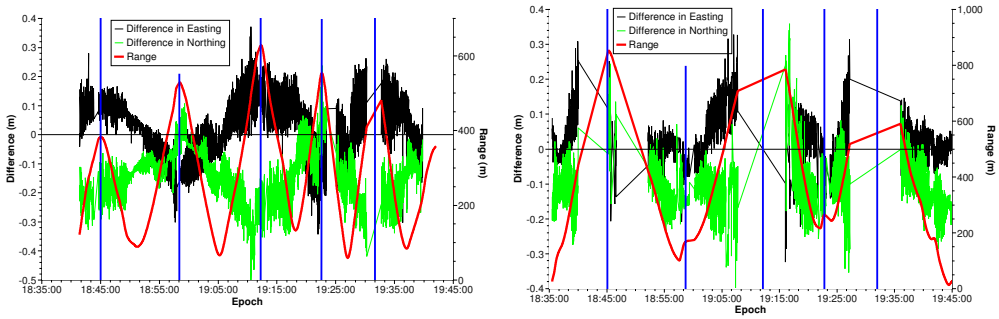
$$A = \begin{bmatrix} V_E\left(1\right) & V_E\left(1\right)\left(t\left(1\right) - t\left(1\right)\right) \\ V_N\left(1\right) & V_N\left(1\right)\left(t\left(1\right) - t\left(1\right)\right) \\ \vdots & \vdots \\ V_E\left(M\right) & V_E\left(M\right)\left(t\left(M\right) - t\left(1\right)\right) \\ V_N\left(M\right) & V_N\left(M\right)\left(t\left(M\right) - t\left(1\right)\right) \end{bmatrix} \tag{6.8}$$

$$x = \begin{bmatrix} \delta t \\ \Delta t \end{bmatrix};\ b = \begin{bmatrix} P^E_{\text{GPS}}\left(1\right) - P^E_{\text{TS}}\left(1\right) \\ P^N_{\text{GPS}}\left(1\right) - P^N_{\text{TS}}\left(1\right) \\ \vdots \\ P^E_{\text{GPS}}\left(M\right) - P^E_{\text{TS}}\left(M\right) \\ P^N_{\text{GPS}}\left(M\right) - P^N_{\text{TS}}\left(M\right) \end{bmatrix} \tag{6.9}$$

where M is the number of epochs with an interpolated prism position corresponding to a GPS

Table 6.1: Offset and drift of total station clocks from GPS time

Station	Clock offset (s)	Clock drift
NAVY	$-1{\cdot}04$	$8{\cdot}5 \times 10^{-6}$
JETTY	$-14{\cdot}94$	$1{\cdot}8 \times 10^{-5}$



(a) NAVY (b) JETTY

Figure 6.7: Graph of the difference between GPS and prism positions

position. The parameters are then solved as:

$$x = A^{-1}b \tag{6.10}$$

Equation 6.2 and Equation 6.10 are solved iteratively until convergence. Table 6.1 gives the computed clock offset and drift: it appears that the local clock at JETTY was not correctly synchronised to GPS time. Over the period of the data collection exercise, the clock at NAVY advanced by $0{\cdot}04$ s and the clock at JETTY advanced by $0{\cdot}08$ s. The ship is moving at > 1 m/s so, if not solved for, the clock drift would bias the final result by several centimetres.

The result of these adjustments is a set of prism positions from each total station with the same time-stamps as the GPS positions, in GPS time. These can then be directly compared to the GPS positions.

6.3.2.2 Results

Figure 6.7 shows the plan difference between the GPS and prism positions for each station, with the different "runs" (periods between turns) delimited by vertical blue lines. The gaps in the data are periods when the total station could not see the prism: there was worse visibility from JETTY than from NAVY. There was an unexpected offset between the GPS and prism positions, which is not constant but follows a distinct pattern that repeats with each run. There is an average offset in Northing of $14{\cdot}1$ cm from NAVY and $10{\cdot}1$ cm from JETTY; Easting offsets are much smaller at $0{\cdot}7$ cm and $-2{\cdot}4$ cm respectively. The ship's movement is mainly in the East-West direction, so some of the error in this direction may have been absorbed by the calculation of the clock offset in Equation 6.10.

Figure 6.8 shows the plan difference between prism positions after the iterative interpolation and time offset adjustments were applied to the two sets of prism positions (the positions from

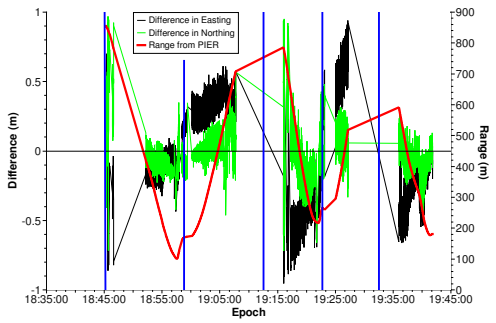
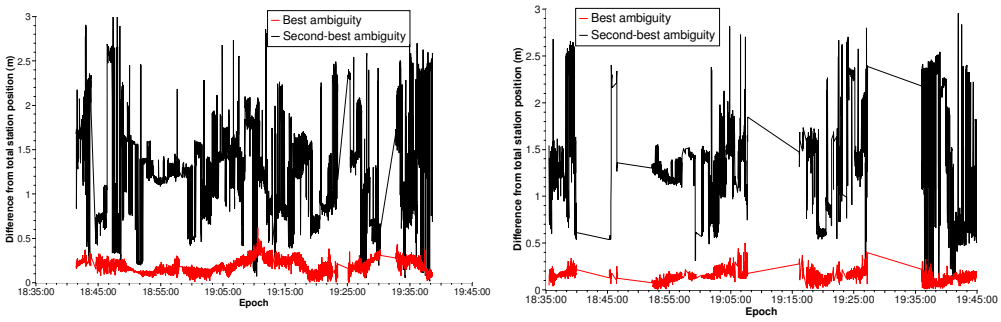


Figure 6.8: Graph of the difference between prism positions from NAVY and JETTY



(a) NAVY (b) JETTY

Figure 6.9: Graph of the difference between prism and GPS positions with best and second-best ambiguities

JETTY were adjusted to match the time-stamps and clock time of the positions from NAVY). The correspondence is worse between the two total stations ($\sim 0{\cdot}5$ m) than between either total station and GPS ($\sim 0{\cdot}2$ m). This implies that the poor agreement with GPS is substantially due to low total station measurement precision.

As discussed in Section 6.3.2.1, the total stations output positions at a rate of around 3 Hz, but the observations are not necessarily updated from the previous position. This implies that data output is not dependent upon obtaining an updated measurement. Therefore even when both measurements have been updated, the range and angle measurements are not necessarily synchronous, and the time-stamp records when the measurements were output rather than when they were obtained. These factors introduce significant error to the prism positions for a moving target. The interpolation and clock adjustment process will also introduce error, particularly during the periods when the ship is turning; interpolation errors should be smaller on the straight sections. The strictest IMO requirement on accuracy (95%) for a future GNSS is $0{\cdot}100$ m and the agreement between the total stations is $0{\cdot}283$ m: the total station measurements are not sufficiently precise to provide a truth model for these requirements.

6.3.2.3 Ambiguity validation

If the GPS ambiguities have been resolved correctly then the maximum error on any single GPS observation is $1/2$ cycle (10 cm for L1 and 12 cm for L2): this is therefore also the maximum position error. Over the very short baseline from BASE to SHIP (< 1 km), spatially correlated errors will almost completely cancel and the dominant error source will be phase multipath with a maximum magnitude of $1/4$ cycle (6 cm for L2). Therefore, assuming the ambiguities have been correctly resolved, the GPS positions should be more accurate than the prism positions.

The total station measurements can be used to validate the short-baseline ambiguity resolution. As discussed in Chapter 4, the LAMBDA method for ambiguity resolution selects the set of integers that minimise the sum of the squares of the distances to the (real valued) float ambiguities in the metric of the float ambiguity covariance matrix. If the sets of integers are ordered by increasing sum of the squares of the distances from the float values, then the first set is the most likely (in the least-squares sense) to be the correct set. If the first set is incorrect, the second set is subsequently the most likely to be correct. Therefore a study of how close the GPS positions generated by the first and second ambiguity sets are to the prism positions can be used to validate the ambiguity resolution.

Figure 6.9 shows the difference between the prism positions and the GPS positions with the first and second sets of ambiguities from each total station. In general, the first ambiguity position is closer to the prism position than the second. However, there are some epochs where the second ambiguity positions are the closest: this is particularly obvious around 19:10 from NAVY.

In the processed data, the first set of ambiguities only changes after a satellite has been lost and re-acquired, or when the reference satellite changes: this is in contrast to the second set, which often changes rapidly. For example, during the span 19:09:51 - 19:10:27, where there are 23 points where the second ambiguity positions are closest to the NAVY prism positions, the second ambiguity set varies between 3 different values but the first ambiguity set stays the same. This suggests that the first ambiguity set is indeed correct, and the difference from the prism positions is due to error in the total station measurement.

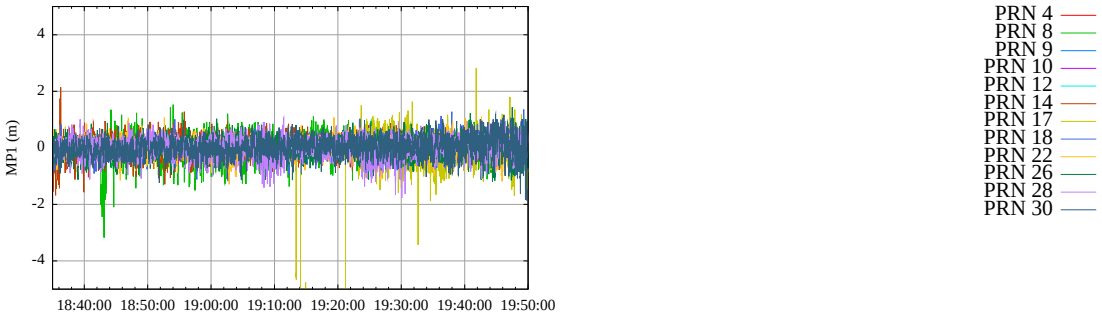
The second ambiguity position is closest in fewer than 1% of epochs, and the first set of ambiguities remains constant throughout the data collection exercise, apart from reference satellite changes and re-acquisitions. Therefore, although not in themselves sufficiently accurate to provide a truth model, the total station measurements provide confidence in the short-baseline ambiguity resolution, and these GPS positions are used as the truth in subsequent analyses.

6.3.3 Analysis of code multipath

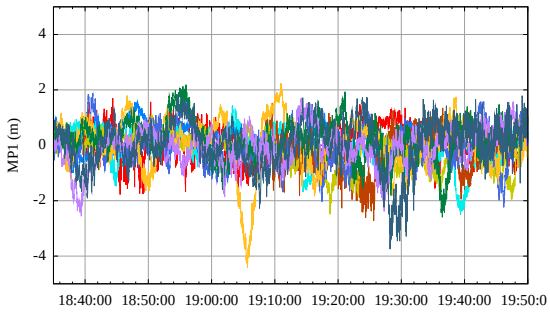
Analysis of the quality of the measurements at the receivers can help interpret the results. The MP_1 observable is used to estimate the magnitude of code multipath on L1:

$$\mathrm{MP}_1 = P_1 - 4{\cdot}0915\phi_1 + 3{\cdot}0915\phi_2 \tag{6.11}$$

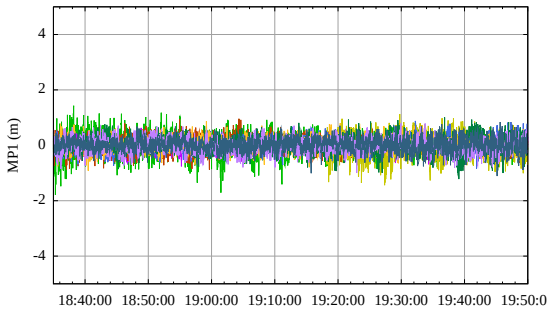
where P_1 and ϕ_1 are the code and phase observations on L1, and ϕ_2 the phase observation on L2. The noise level of this value gives an indication of code noise and multipath error.



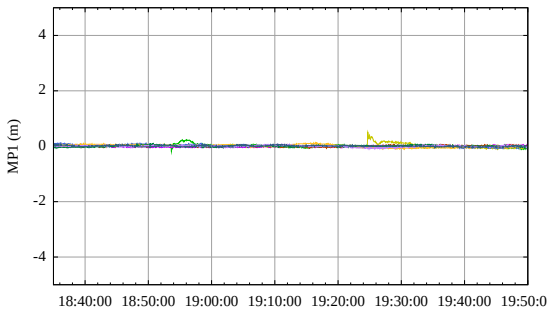
(a) SHIP



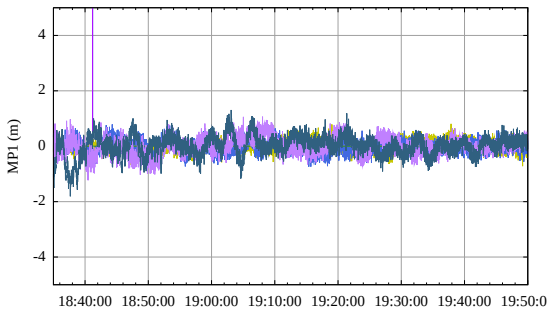
(b) BASE



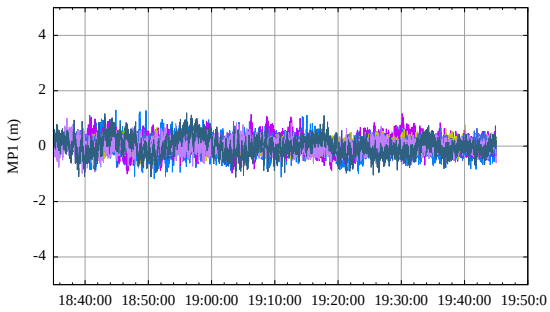
(c) WIX



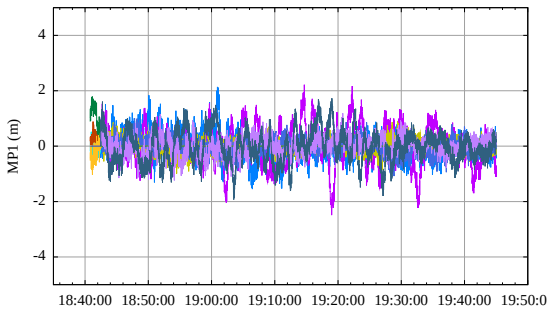
(d) ALDB



(e) SHOE



(f) MAID



(g) STRA

Figure 6.10: MP1 observables

Figure 6.10 shows the MP_1 values. The MP_1 observable from BASE is very noisy, indicating high levels of multipath: the receiver was located on the roof of the Trinity House building and substantial multipath would have been caused by nearby surfaces. In contrast, WIX, in an open field with a clear view of the sky, shows less multipath. SHIP shows relatively better performance than expected, given the location below the ship's superstructure, but is still worse than WIX. There is a clear cyclical multipath effect at MAID, and STRA is also relatively noisy. The noise level at SHOE is similar to that at WIX. The rest of the Ordnance Survey stations use a different type of receiver which phase-smooths the code observations, so there is minimal observable code noise and multipath; MP_1 graphs for these stations are the same as for ALDB. Phase-smoothing was disabled for the GR-3 receivers.

6.4 Summary

This chapter describes the main data collection exercise and the generation of the truth model, as well as the three preliminary exercises. In the first, in London, it was not possible to compare the GPS and prism positions because the total station time-tags were rounded to the nearest second, which introduced a large error for a moving prism. This led to the development of software to allow measurements to be streamed to a laptop and higher-precision time-tags applied. In the second, in Margate, the total station did not track smoothly; this was worse over longer ranges (> 400 m), but was subsequently improved by changing the total station settings. The third, in St Albans, demonstrated centimetric agreement between prism and GPS positions when the rover was stationary; this was an order of magnitude worse when it was moving. The lessons learnt from these exercises were applied to the main data collection exercise.

The main data collection exercise was carried out on *THV Alert* in Harwich Harbour. This was designed to allow data to be collected in a port navigation and automatic docking environment with an independent truth model, so that the results could be compared to the relevant IMO requirements. It was not possible to execute the original plan due to shallow water and range limitations for the total station tracking induced by the weather and sea spray. In the final experiment, *Alert* made several approaches to the GLA jetty whilst being tracked by two total stations. The prism positions from each total station were then interpolated to whole second positions and the clock offset solved for in an iterative procedure. The prism and GPS positions agreed at the decimetre level, and the two total stations only agreed to $\sim 0{\cdot}5$ m. It is likely that substantial errors in the prism positions were introduced by a fundamental problem with the total station tracking, which is the lack of synchronisation between the total station range measurements, angle measurements, and the time-stamp. The total station positions were therefore used to verify the ambiguity resolution of the short-baseline (1 km) GPS, and these are subsequently used as the truth model.

The code multipath was analysed at each receiver, and it was discovered that the short-range reference station at BASE had severe multipath errors. Several Ordnance Survey reference stations had phase smoothing enabled, which virtually nullified the code multipath.

Chapter 7

Positioning techniques and software

Once the GPS data have been collected in a realistic marine environment such as Harwich harbour, they must then be processed with a variety of different models in order to assess the positioning performance, and to determine which techniques and models may best enable future modernised GNSS to meet the IMO requirements. It is therefore necessary to have processing software that is flexible and allows easy changing of the specifics of each technique and model used. Unfortunately, currently-available software is designed for easy processing of survey data or scientific processing of large networks, and is not sufficiently flexible for use in this project. It was therefore necessary to write GNSS processing software, designed to be flexible and allow the easy choice of different processing techniques and models.

There are three possible GPS processing techniques that may be used to meet the IMO requirements. Point positioning uses the code measurements from the rover only, and does not require a shore station; however it is the least accurate. DGNSS also relies upon the code measurements, but uses a nearby reference station to cancel the satellite clock error and reduce spatially correlated errors such as the ionosphere, troposphere and satellite orbit error. The most precise technique, RTK, uses the phase observations to achieve very precise positioning, but the technique is more complex and less robust, and also requires a nearby reference station and a good-quality receiver. These different techniques are described in Section 7.1.

Each positioning technique is affected to a different extent by several error sources, such as atmospheric error, satellite orbit error or multipath. However, there are methods available to mitigate the effect of these errors: models of the atmosphere can be used to remove a proportion of the error, phase smoothing can reduce code multipath and the ionosphere-free code observable can be used to cancel the ionospheric error at the expense of increased multipath and measurement noise. These error sources and mitigating models or techniques are described in Section 7.2.

The software was developed in conjunction with Nicholas Zinas, also at UCL. It was designed to be as flexible as possible with regards to using different positioning techniques or models, with options easily changed via a configuration file. Although the only novel algorithm implemented was the subset ambiguity resolution algorithm, as described in Chapter 4, substantial work was required to deal with real data, which is often missing or containing errors. The software is described in Section 7.3, with flow diagrams showing the processing procedure for point poisoning, DGNSS and RTK.

7.1 Positioning techniques

7.1.1 Point positioning

Point positioning uses the code observations to produce a positioning solution. At a receiver r, for each satellite s and for each frequency, the observation equation at time t is:

$$P_r^s = \rho_r^s + c\left(dt_r\left(t\right) - dt^s\left(t - \tau_r^s\right)\right) + T_r^s + I_r^s + \epsilon_r^s\left(t\right) \tag{7.1}$$

where:

P_r^s	is the code observation
ρ_r^s	is the geometric range
$dt_r\left(t\right)$	is the receiver clock offset at reception time
$dt^s\left(t - \tau_r^s\right)$	is the satellite clock offset at transit time
I_r^s	is the ionospheric error
T_r^s	is the tropospheric error
ϵ_r^s	is the remaining error

$\epsilon_r^s\left(t\right)$ encompasses many residual error sources, such as errors in the atmospheric, satellite orbit and satellite clock models, and multipath and code measurement noise. The satellite clock offset at transmit time is obtained from the orbit files, and the other parameters are solved for in a weighted least-squares adjustment. The least-squares model is set up:

$$Ax = b + v \tag{7.2}$$

where A is the design matrix, x is the parameter vector, b is the observation vector and v is the vector of residuals. The geometric range in Equation 7.1 is expressed in terms of the known satellite transmit coordinates and the unknown rover coordinates:

$$\rho_r^s = \sqrt{\left(x^s - x_r\right)^2 + \left(y^s - y_r\right)^2 + \left(z^s - z_r\right)^2} \tag{7.3}$$

so the parameter vector $x = \left[x_r, y_r, z_r, dt_r\right]^T$. The observation is a non-linear function of the parameters, so non-linear least squares must be used. It is necessary to specify initial values of the parameters, x_0: the design matrix A is the Jacobian matrix of Equation 7.1 at x_0, b contains the observed-minus-computed values and x is the vector of updates to x_0. The satellite transmit positions are obtained from the satellite orbit file and depend on the signal travel time $\tau = \rho_r^s/c$, which is dependent on the results of the adjustment and the receiver clock offset, so the estimation procedure must be performed iteratively until convergence.

To increase the precision of the solution, the observations are weighted through the covariance matrix of the parameters, Q_l. A signal from a low-elevation satellite has passed through more atmosphere than a signal from a high-elevation satellite, and so contains greater error which is cancelled out less reliably by the atmospheric models. Therefore an elevation-dependent weighting of $1/\sin^2(\text{elevation})$ is applied to the observation precisions in Q_l. The least-squares

solution for the parameter vector is then:

$$x = \left(A^T Q_l^{-1} A\right)^{-1} A^T Q_l^{-1} b \tag{7.4}$$

7.1.2 DGNSS

In DGNSS, the errors in the observations at the rover are reduced or cancelled by differencing the code measurements with those from a nearby reference station. This is equivalent to estimating the error in the observations at a known point and then subtracting this error from the observations at the rover. For a receiver r_1, reference station r_2 and satellite s, the single-differenced observation equation at time t is:

$$\begin{aligned} P^s_{r_1,r_2}(t) &= P^s_{r_2}(t) - P^s_{r_1}(t) \\ &= \rho^s_{r_1,r_2}(t) + c\left(dt_{r_2}(t) - dt_{r_1}(t)\right) + \epsilon^s_{r_1,r_2}(t) \end{aligned} \tag{7.5}$$

The satellite clock error cancels completely; ionospheric, tropospheric and satellite orbit errors are spatially correlated and are reduced proportionally to the baseline length. It is necessary to perform point positioning for each receiver to obtain the satellite transmit positions. No elevation-dependent weighting is applied because the increased atmospheric errors are largely cancelled over the short baselines used; the main error sources are code measurement noise and multipath, which do not depend on elevation. The least-squares adjustment is then performed in the same manner as for point positioning.

7.1.3 RTK

RTK is the most accurate positioning technique. The single differences from each satellite are subtracted from the single differences from the reference (highest elevation) satellite to obtain double-differenced observations. The receiver only measures the current fractional phase, so phase observations are biased by an unknown number of whole cycles between the receiver and the satellite. For a receiver r_1, reference station r_2, satellite s_1 and reference satellite s_2, the double-differenced observation equation at time t is:

$$\Phi^{s_1,s_2}_{r_1,r_2} = \rho^{s_1,s_2}_{r_1,r_2} + T^{s_1,s_2}_{r_1,r_2} - I^{s_1,s_2}_{r_1,r_2} + \lambda N^{s_1,s_2}_{r_1,r_2} + \epsilon^{s_1,s_2}_{r_1,r_2} \tag{7.6}$$

where:

$\Phi^{s_1,s_2}_{r_1,r_2}$	is the double-differenced phase observation in metres
$\rho^{s_1,s_2}_{r_1,r_2}$	is the double-differenced geometric range
$T^{s_1,s_2}_{r_1,r_2}$	is the double-differenced tropospheric error
$I^{s_1,s_2}_{r_1,r_2}$	is the double-differenced ionospheric error
$N^{s_1,s_2}_{r_1,r_2}$	is the integer ambiguity
$\epsilon^{s_1,s_2}_{r_1,r_2}$	is the remaining error

RTK positioning relies on the estimation of the integer ambiguities in order to use the phase observations as precise ranges. The ambiguities are estimated and fixed in a four-step procedure, as shown in Figure 7.1; ambiguity estimation and validation is discussed in more detail in

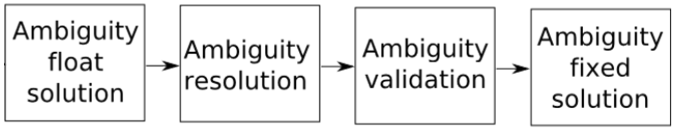


Figure 7.1: RTK positioning procedure

Chapter 4. The first step is to obtain an estimate of the real-valued float ambiguities and their covariance matrix. It is not possible to do this in a single-epoch adjustment using the phase data alone: with n satellites, there are $n-1$ double-differenced observations, $n-1$ unknown ambiguities and the unknown receiver position, so there are more unknowns than observations and the problem is rank-deficient. Conceptually, a single phase observation does not contain any range information due to the corresponding unknown real-valued ambiguity. This problem is solved in one of two ways: either data are collected over multiple epochs, so the changing satellite geometry provides range information, or the double-differenced code observations are used to obtain an initial position. The former approach has the advantage that data can be collected for as long as required to successfully resolve the ambiguities. However, this is harder to achieve if the receiver is in a challenging environment, where satellites may not be visible for the required period of time. The time taken to achieve ambiguity resolution could be a problem when attempting to meet strict continuity or availability requirements using this technique. The single-epoch approach, using the code observations, is more suited to a challenging environment as it can recover instantaneously from loss of satellite lock; this approach is therefore used in this project. After the ambiguities have been fixed to integers, it is possible (and necessary with the multi-epoch approach) to keep the same ambiguity values for subsequent epochs. However, this requires monitoring to eliminate cycle slips and incurs the risk of propagating incorrect ambiguity values; it is simpler to re-fix the ambiguities every epoch.

The next step in Figure 7.1, ambiguity resolution, determines the most likely set of integer values for the ambiguities, and the validation step determines if there is sufficient confidence in these to use them (using the incorrect set can result in a position error of several metres). If the validation test is passed then the ambiguity-fixed solution is performed, using the phase measurements to achieve a high-precision solution. The ambiguity-fixed step estimates the rover coordinates, so it is not necessary to have these as parameters in the float solution. This gives the option of a geometry-free float solution, where the parameters are the unknown double-differenced satellite ranges, rather than the rover coordinates. This model is commonly used because is linear, and it does not require knowledge of the satellite transmit coordinates. However, it has less redundancy than the geometry-based model, and is weaker due to the omission of the satellite geometry information. Therefore the geometry-based model, with the receiver coordinates as parameters, is used for this project.

7.2 Error sources and models

There are several sources of error that affect the GNSS observations, and these can often be mitigated against by the application of models of the error or a technique to reduce its effect.

The models applied in the software are the best available without input from external input; this criteria was established due to the potential difficulties of acquiring the external data, such as numerical weather models, with a ship-based rover.

7.2.1 Multipath

Signals from the satellite can be reflected by nearby objects; the reflected signal reaches the receiver slightly after the direct signal. If the reflection is delayed by a long time then the receiver can distinguish the direct and reflected signal, but short delays can cause a bias in the measurement. Multipath is a particularly serious error source because it is site-specific and so cannot be cancelled by a nearby reference station; it is also worse in built-up areas such as ports.

One method of mitigating code multipath is phase smoothing. This is a technique whereby the precisely measured difference in satellite-receiver range from epoch to epoch obtained from the phase observations is combined with the absolute range derived from the code observations, to obtain a code measurement with reduced noise from code measurement and multipath. The phase-smoothed code observable at epoch i is:

$$P_{sm}(i) = \frac{P(i)}{i} + \frac{i-1}{i}\left(P(i-1) + \Phi(i) - \Phi(i-1)\right) \tag{7.7}$$

where P is the code observable and Φ is the phase observable in metres. The phase-smoothed observable is initially weighted towards the code observation from the current epoch; over time the weight shifts towards deriving the inter-epoch difference from the phase observations. The ionosphere affects the code and the phase with the opposite sign, which causes P_{sm} to diverge from the true range. In a dual-frequency system a correction can be applied to compensate for this:

$$P'_{sm}(i) = P_{sm}(i) - 2\frac{\left(\Phi_{L1}(i) - \Phi_{L1}(i-1)\right) - \left(\Phi_{L2}(i) - \Phi_{L2}(i-1)\right)}{1 - \left(\frac{\lambda_{L2}}{\lambda_{L1}}\right)^2} \tag{7.8}$$

If only one frequency is available, the ionospheric divergence can be compensated for by restricting the maximum time window of the phase smoothing so the weighting never completely favours the phase, or by resetting the filter.

7.2.2 Troposphere

The troposphere stretches from the surface to about 10 km above the surface of the Earth, and delays the GNSS signal. The effect is the same on all GPS frequencies and the same for the code and the phase. The tropospheric delay is split into 'wet' and 'dry' components. The dry component, which depends upon the dry air, accounts for about 90% of the delay but can be accurately estimated from local meteorological data or a global model. The remaining 10%, the wet delay, is much harder to estimate due to uncertainties in the vertical distribution of water vapour; even measurements of the water vapour at ground level do not provide sufficient information to estimate the wet delay. Tropospheric delay is strongly dependent upon height.

Tropospheric delays are estimated or predicted as a zenith delay that is mapped to the specific observation by a mapping function. In our software, the tropospheric zenith delay was computed using the ESA model (Krueger *et al.*, 2004), which interpolates various parameters across

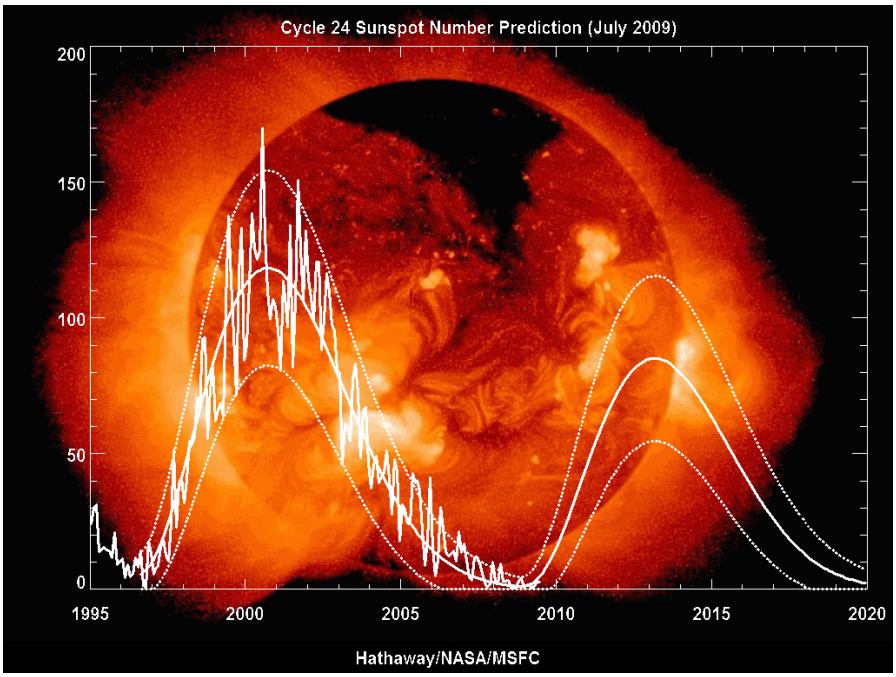


Figure 7.2: Sunspot cycle

a global grid, and depends on the latitude, longitude, time of day and time of year. The tropospheric mapping function, which maps this zenith delay to the elevation of a given satellite, was the Global Mapping Function (GMF) (Boehm *et al.*, 2006). These are the best models available without external input; if this constraint were relaxed then superior results could be achieved by integrating through numerical weather models. The tropospheric error is correlated over short baselines, so the error is reduced for differenced positioning techniques (DGNSS and RTK).

7.2.3 Ionosphere

The ionosphere is the highest layer of the atmosphere, stretching from about 70 km to over 1,000 km above the surface of the Earth. In this region, ultraviolet rays from the sun ionises gas molecules to release free electrons, which interfere with the GPS signals. Due to the dependence on the sun, the level of interference is affected by the time of day, with a maximum effect at about 14:00 and a weaker effect at night. The ionospheric error is also dependent on the 11-year sunspot cycle; the magnetic activity that accompanies the sunspots has a large effect on the ultraviolet emissions of the sun, and hence on the number of ions in the ionosphere. Figure 7.2 shows that the data collection exercise was carried out at a time of particularly few sunspots. The ionosphere delays the code measurement, and advances the phase measurement, unlike the troposphere which affects both equally. The error is frequency dependent, which is the main reason behind the multiple frequencies for GPS: the ionospheric error can be cancelled in a multi-frequency system by forming the ionosphere-free observable:

$$P^s_{r,IF} := \frac{f_1^2}{f_1^2 - f_2^2} P^s_{r,f1} - \frac{f_2^2}{f_1^2 - f_2^2} P^s_{r,f2} \tag{7.9}$$

where $P^s_{r,i}$ and f_i are the observation and frequency, respectively, of observable i. Ionospheric error is cancelled, but code noise and multipath is increased. With GPS L1 and L2, $P^s_{r,IF} = 2{\cdot}545P^s_{r,L1} - 1{\cdot}545P^s_{r,L2}$, and code noise and multipath are increased by a factor of 3. This multiplication of the noise limits the utility of the ionosphere-free observable to situations where the ionospheric error dominates the noise. In a single-frequency system this is not possible, so GPS includes an ionospheric model in the data message, called the Klobuchar model, (GPS ICD-200D). On average, this model cancels 50% of the ionospheric error; this can also be applied with dual-frequency data if the ionosphere-free combination is not formed. The tropospheric error is correlated over short-to-medium baselines, so the error is reduced for differenced positioning techniques (DGNSS and RTK).

7.2.4 Orbit error

The satellite orbit is derived either from the broadcast ephemeris contained in the GPS data message, or from a precise orbit file generated by an external agency such as the International GNSS Service (IGS) using observations from many GPS reference stations across the world. The precise orbits can be used for post-processing, or less-accurate predicted orbits can be used for real-time positioning, although these are still more accurate than the broadcast orbits. The error in the orbit directly translates into error in the ranges for point positioning; with differenced techniques the error is reduced in inverse proportion to the baseline length and therefore has less effect. The broadcast orbits are used in the processing of the data collected in Harwich, as these are easily available in real time.

7.2.5 Receiver antenna offset

The point of measurement of the GNSS signal in the antenna is offset from the physical reference point, known as the Antenna Reference Point (ARP); this offset depends on the elevation and azimuth of the incoming signal. Absolute antenna phase corrections were applied in our software. These corrections are particularly important when performing phase with two antennas of different types, because the errors will not cancel in the differencing and the resultant error can make correct ambiguity resolution very difficult.

7.3 Processing software

The software is a compilation of different modules that can be connected together to achieve the desired processing procedure. The exact execution of the software is controlled by a combination of changes in the code to enable specific modules, or settings in the settings file. This is best illustrated by a flow diagrams showing the procedure for obtaining a position with different positioning techniques. Figure 7.3 is the key for the flow diagrams. Black boxes represent modules that were not developed by the author; grey boxes were partially developed by the author.

Figure 7.4 shows the flow of point positioning. The observation data are contained in a RINEX file, which is read into the software one epoch at a time. The observations from the epoch are validated to ensure that they exist and have reasonable values, and phase smoothing and the

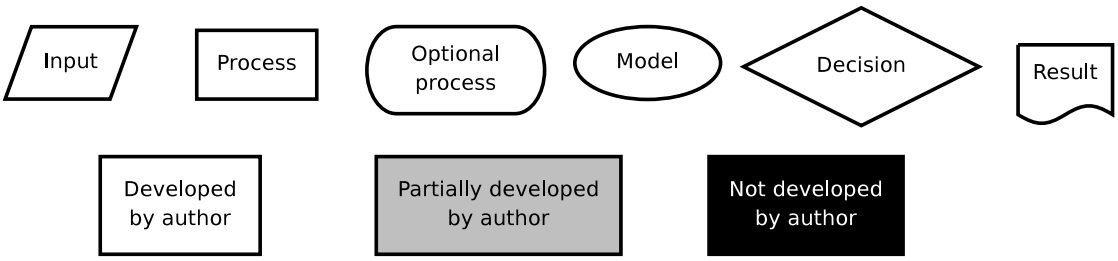


Figure 7.3: Flow diagram key

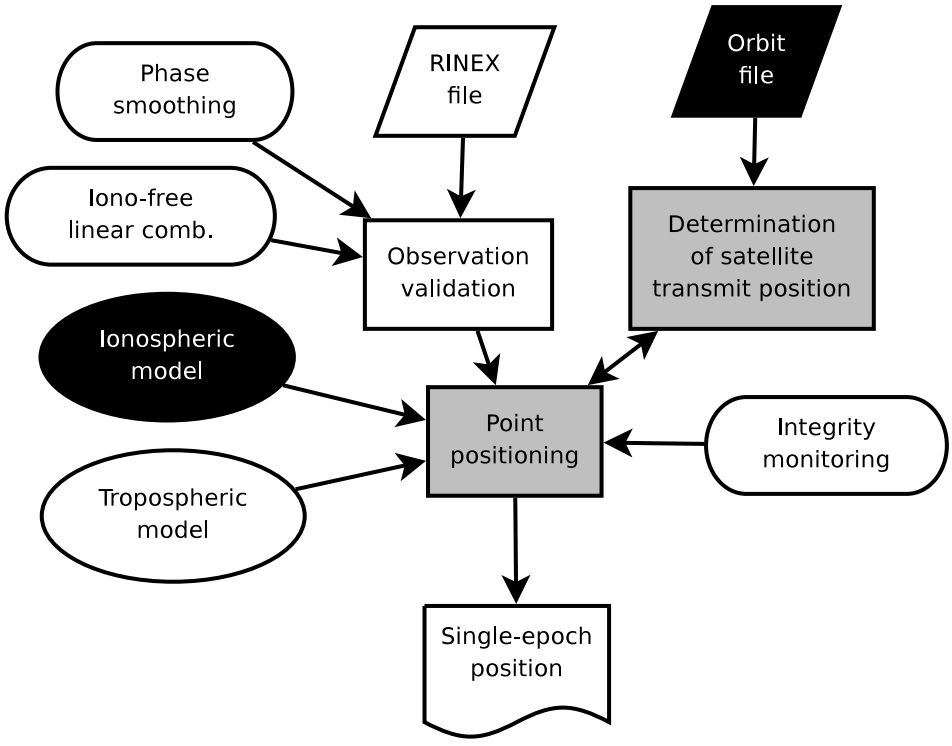


Figure 7.4: Point positioning flow diagram

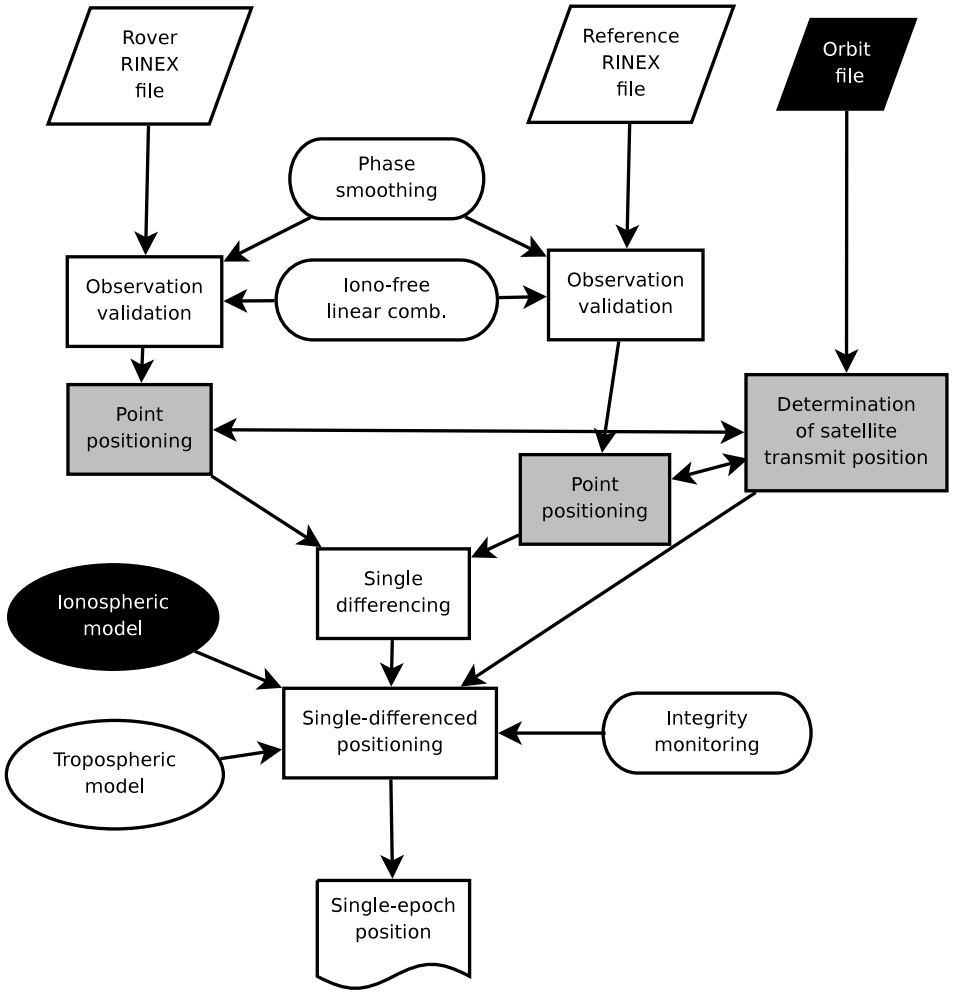


Figure 7.5: DGPS flow diagram

ionosphere-free linear combination may be applied, depending on the settings. Non-existent or obviously erroneous observations are flagged so that they are not used in the rest of the processing. The observations from the current epoch are combined with the orbit data read from a broadcast or precise ephemeris file in an iterative process to solve for the position, the receiver clock offset and the satellite positions and clock offsets at transmit time. At this stage, the atmospheric models are used to reduce the error from these sources. Integrity monitoring algorithms, as described in Chapter 5 may be applied to the residuals of the adjustment, in order to allow the exclusion of biased observations at this stage. After the position has been determined, the process repeats with the next epoch of observations from the RINEX file.

Figure 7.5 shows the flow of DGNSS positioning. The primary difference between DGNSS and point positioning is the inclusion of the observations from a reference station. A single epoch of observations from both receivers is read into the software and validated, and phase smoothing and ionosphere-free combination is optionally applied, in the same way as for point positioning; this is required to be performed separately for each receiver in order to determine the satellite transmit position and clock offset. At the next step the two set of observations are combined by single differencing, which cancels the receiver clock error and reduces the spatially correlated errors such as the atmosphere and the satellite orbit error. The single-differenced position is then

computed, and integrity monitoring applied if desired.

Figure 7.6 shows the flow of RTK positioning, which starts off in much the same way as DGNSS: the observations from both the rover and the reference station for a single epoch are read into the software, validation checks and phase-smoothing applied, and point positioning performed to determine the satellite transmit positions and clock offset. At the next stage, the observations from both receivers are double-differenced, and then the ambiguity float solution is performed, where the code and phase observations are combined with the rover position and double-differenced ambiguities as parameters. Ambiguity resolution and validation are then performed to obtain integer values for the ambiguities. If the validation test is failed then there is the option of using the subset ambiguity resolution algorithm developed in Chapter 4. If this is not used, or fails to resolve the ambiguities then no RTK position is achieved this epoch. If any fixed ambiguities are produced then the ambiguity fixed solution is performed, where the double-differenced phase observations with fixed ambiguities are used as precise range measurements to determine the final position.

7.4 Summary

There are several GNSS processing several that can be used to process the data collected in Harwich harbour, as described in Chapter 6. Point positioning does not require any shore infrastructure, but is the least accurate technique. DGNSS improves upon the accuracy of point positioning by using a nearby receiver to estimate the magnitude of the error on each observation; this value is then used to correct the observations at the rover, significantly reducing the magnitude of spatially correlated errors such as atmospheric or satellite orbit errors. RTK uses the phase observations to achieve very precise positions, but also requires a nearby reference station and is significantly less robust, due to the necessity of ambiguity resolution, as described in Chapter 4. This chapter describes the error sources that can affect these positioning techniques, and the models and algorithms that can be used to mitigate their impact. The final section describes the software written at UCL to allow flexible processing of the data collected in Harwich harbour, and gives examples of the flow of the processing for obtaining point positioning, DGNSS and RTK positions.

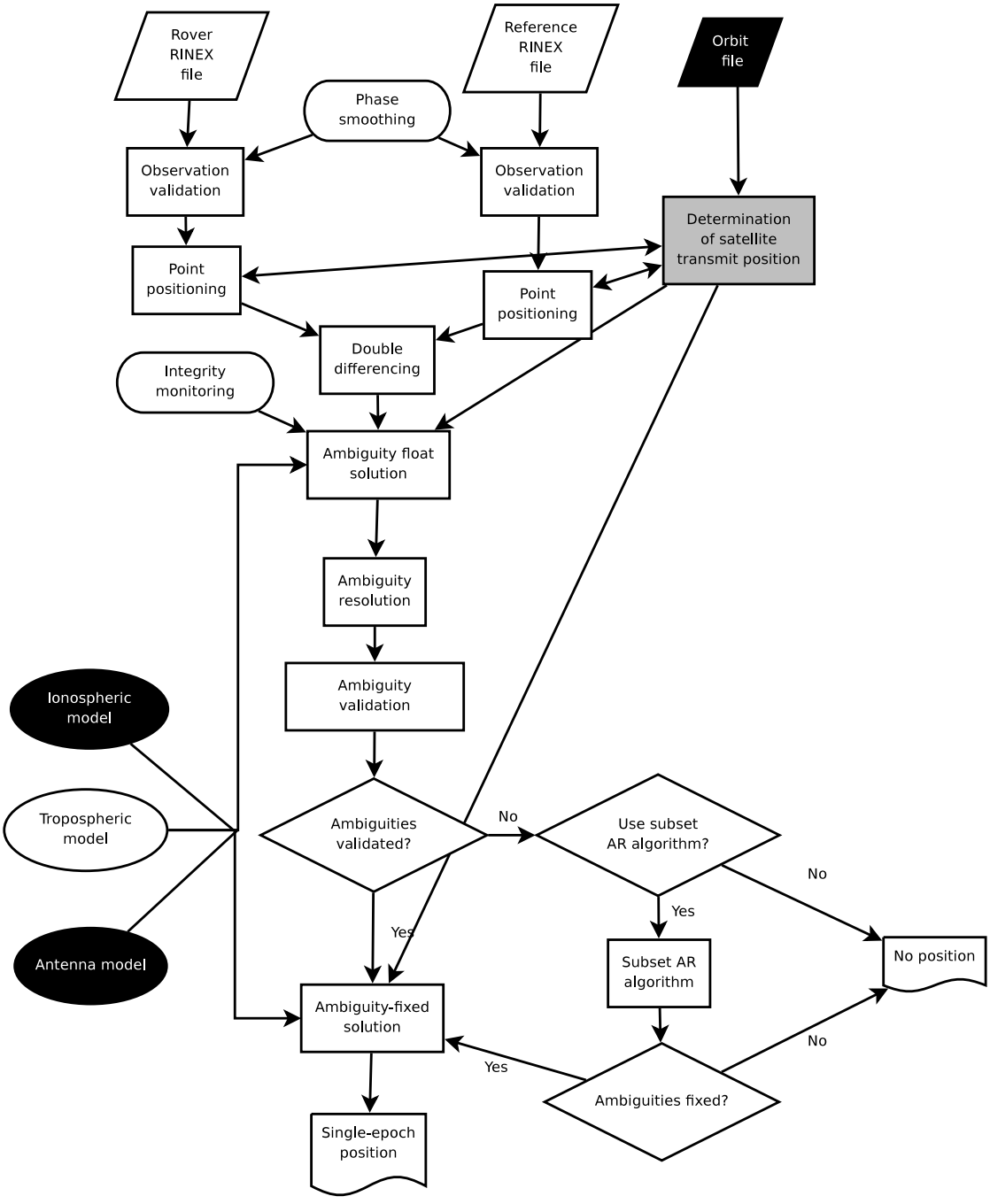


Figure 7.6: RTK flow diagram

Chapter 8

Data analysis

This chapter describes the analysis of the data obtained during the Harwich Harbour data collection exercise described in Chapter 6. The data were collected in an environment that provides a realistic example of port navigation and automatic docking, which are two of the IMO requirements for a future GNSS as are discussed in Chapter 2, and post-processed to simulate real-time processing. Due to the range limitations of the truth model, data were not collected in a port approach environment. This which differs from port navigation in two ways: there are fewer nearby obstructions and the receiver is further from the shore. Fewer nearby obstructions will result in improved positioning performance, as there will be more visible satellites and fewer reflectors to create multipath. The effect of a receiver further from shore can be simulated by processing the data with reference stations that are further away from the rover. Therefore the data are analysed with regards to the port approach, port navigation and automatic requirements, as set out in Table 8.1. The availability and continuity requirements are the same for all stages, but the 95% accuracy value and the alert limit are an order of magnitude smaller for each successive requirement. Automatic docking is the hardest to meet of all IMO requirements, with an accuracy requirement of $0{\cdot}1$ m.

It would require data over many years and in many different geographical locations and environments to to show that the requirements are met by any given technique; this analysis is beyond the scope of this thesis. However, what can be shown is the relative differences between processing techniques, and in particular which techniques cannot be used to meet specific requirements. The environment of the data collection was relatively benign; when approaching a large container terminal the multipath conditions at the receiver could be substantially worse, and the level of ionospheric activity was low. Therefore, if a technique does not meet the require-

Table 8.1: IMO requirements for a future GNSS

	Accuracy	Integrity			Availability	Continuity
	Horizontal (m)	Alert limit (m)	Time to alarm (s)	Integrity risk (per 3 hours)	% per 30 days	% over 3 hours
Port approach	10	25	10	10^{-5}	99·8	99·97
Port navigation	1	2·5	10	10^{-5}	99·8	99·97
Automatic docking	0·1	0·25	10	10^{-5}	99·8	99·97

ments in this situation, whether through low accuracy or an integrity, continuity or availability breach, then it is reasonable to assume that it will not be able to meet the requirement in the general case, over all times and locations. Therefore the results of this experiment are particularly useful to demonstrate which techniques can *not* be used to meet the requirements.

The novel partial ambiguity resolution algorithm described in Chapter 4 is applied to the real data collected in Harwich harbour, and the results compared to conventional RTK positioning. It is expected that this new technique will improve the robustness of the RTK positioning, and will improve the ability to meet the automatic docking requirement by reducing the occurrence of continuity and availability errors; the risk of propagating incorrect ambiguities to increase the occurrence of integrity errors will be assessed.

The IMO requirements are for a future GNSS, such as modernised GPS and Galileo. At the time of the experiment, the only modernised signals available were GPS L2C, and the time of the experiment was chosen to maximise the number of L2C signals that were visible. The observations from these modernised satellites are analysed to determine the performance improvement achieved over conventional codeless L2 tracking. It is expected that much better positioning performance will be achieved when the full modernised GPS and Galileo constellations are available; in this chapter the anticipated improvements and benefits as described in Chapter 3 are interpreted in light of the results of the data collection exercise.

Section 8.1 describes the technique used for analysing the data with respect to the requirements. This is followed by the results of different processing techniques: point positioning in Section 8.2, DGNSS in Section 8.3 and RTK in Section 8.4. In Section 8.5, techniques developed in Chapter 5 are used to analyse the relative robustness to outliers of the different processing techniques. In Section 8.6, a comparison is made between the modernised L2C and conventional L2 signals to determine the improvement in code tracking precision achieved with the modernised L2 code, and in Section 8.7 the effect of the anticipated performance improvements of modernised GNSS, as described in Chapter 3, on the ability of different techniques to meet the requirements is discussed.

8.1 Data processing technique

The data collected in Harwich Harbour are processed with regards to the IMO requirements for a future GNSS outlined in Table 8.1. As discussed in Chapter 6, the short-baseline RTK positions from BASE are used as the truth model: tracking total stations were used to provide confidence that the ambiguity resolution is correct. The dominant error source over this short baseline is phase multipath, with a maximum magnitude of about $1/4$ cycle (Lau & Cross, 2005): therefore the truth positions should be accurate to 6 cm.

The aim of the data analysis is to determine which processing techniques may be used to meet a given set of IMO requirements. The requirements are not specific enough to be directly applied to real data, so some assumptions are made to facilitate analysis. The most significant assumption is the extension of the integrity time-to-alarm period to the continuity and availability requirements, so that these requirements are only violated if no position is obtained for 10 s, in the same way that an integrity breach is not declared unless the position has been in error for

this period. This is justifiable because the requirements allow the position to be incorrect for the time-to-alarm period without the knowledge of the operator, and having no position is safer than having an erroneous one. This assumption is necessary because if it were not made, and a continuity and integrity breach was caused by the lack of a position for 1 s, the requirements would not be met by any technique: even the Ordnance Survey reference receivers occasionally have missing epochs of data. There are no requirements for height, so only the plan positions are considered. The requirements are defined as follows:

Accuracy

Accuracy is taken as the 95th percentile of the differences between the obtained positions and the truth. The IMO requirements do not specify which percentile is to be used for the accuracy, but 95% is the value in common usage (e.g. ICAO requirements) and is used when reporting the achievable accuracy of GPS in IMO (2003), so this is used in the subsequent analysis.

Integrity

Integrity is relatively straight-forward to interpret when analysing real data. If a position is obtained that is further than the alert limit from the truth, and no positions closer than the alert limit are obtained within the time-to-alarm, then an integrity breach has occurred. The integrity risk is taken as the proportion of 10 s periods over the course of the data collection that contain an integrity breach. The requirements specify the integrity risk statistic should be over 3 hours, and the data collection exercise only spans 75 minutes: however, a single integrity breach is sufficient to violate the integrity requirement over both time spans, so this will not have a significant effect on the results. Of course, there are more epochs in 3 hours during which an integrity error could occur, but this analysis is beyond the scope of this thesis due to the limitations on data collection time.

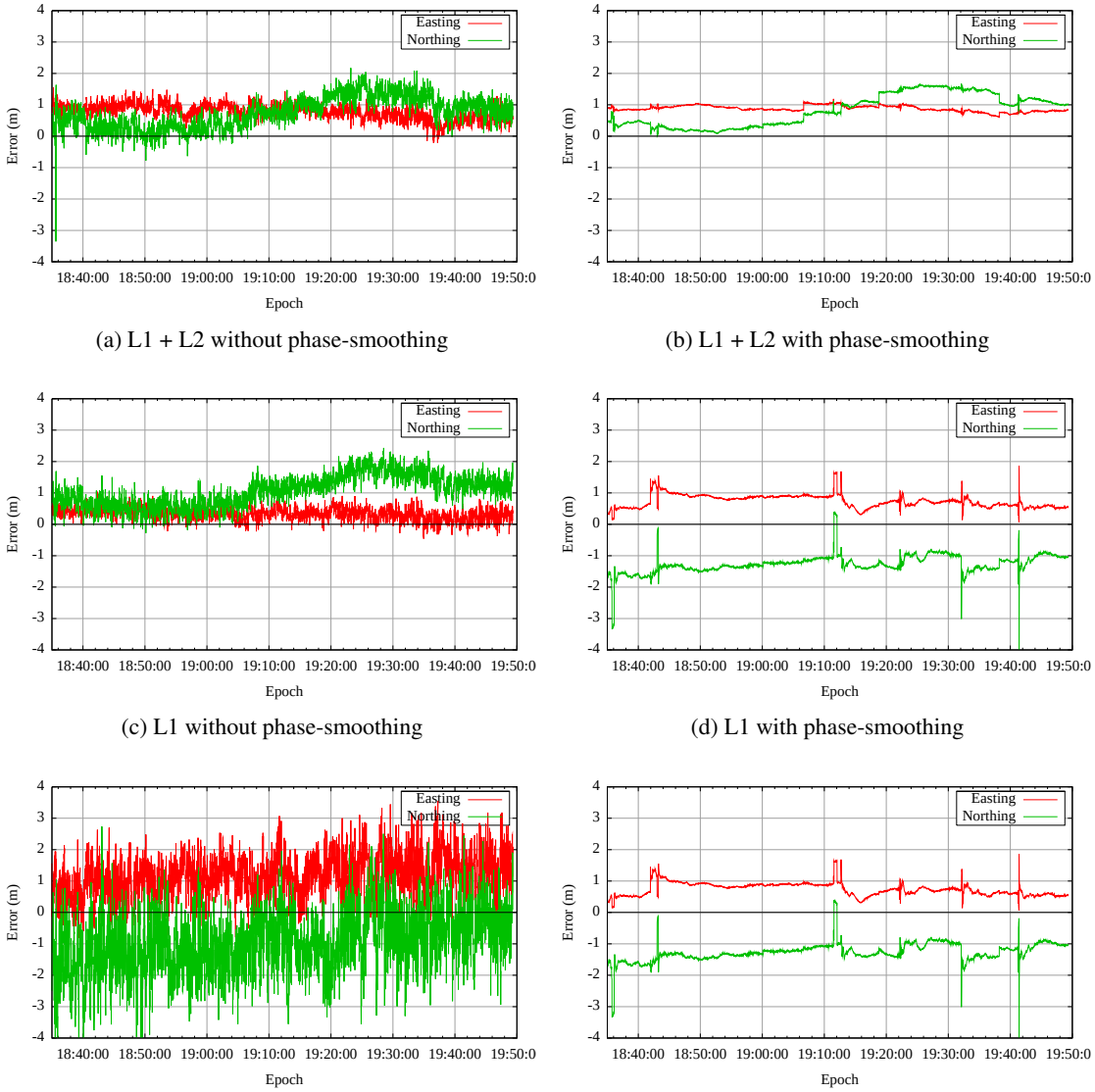
Availability

The availability requirement is over 30 days, so the short time span of the data collection exercise can only give an estimate of whether it would be met. The value used for the availability is the percentage of 10 s spans containing at least one valid position (i.e. the percentage of 10 s spans for which the system was available).

Continuity

The continuity value is the percentage of the total experiment time covered by the longest span with no 10 s gaps. This differs from the requirement, which is “the probability that ... a user will be able to determine a position with specified accuracy over the (short) time interval”. However, this is not possible to assess with a single span of data: the data are either continuous or not. This definition makes it possible to distinguish between different processing scenarios, while still making it possible to determine if the requirement was met.

The IMO requirements specify that integrity monitoring must be available over the time period. However, RAIM algorithms, as described in Chapter 5, strongly depend upon the value



(a) L1 + L2 without phase-smoothing

(b) L1 + L2 with phase-smoothing

(c) L1 without phase-smoothing

(d) L1 with phase-smoothing

(e) Iono-free without phase-smoothing

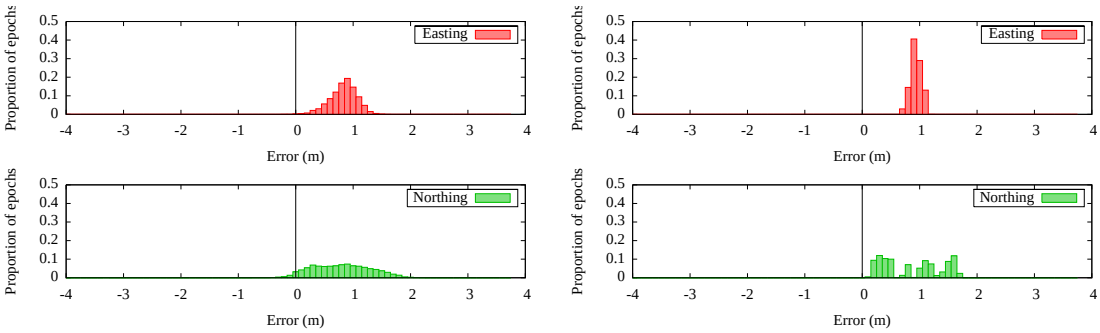
(f) Iono-free with phase-smoothing

Figure 8.1: Point positioning accuracy

chosen for the probability of missed detection of an error. This value must be chosen with reference to the integrity requirements, and depends on the probability of an error occurring. This probability is difficult to determine accurately due to the low rate of occurrence of such events, so in order to keep this analysis more general and based upon the data collected, this component of the requirement is ignored.

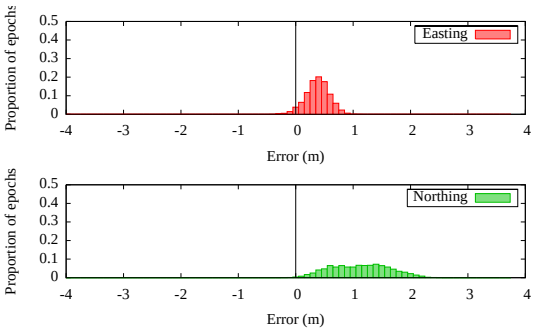
8.2 Point positioning

Single-frequency, dual-frequency and ionosphere-free point positioning were performed on the data from SHIP, both with and without phase-smoothing applied, for a total of six different processing techniques. The L1, L2 and L2C code observations were assumed to have standard deviations of 0·3 m, 0·6 m and 0·3 m respectively. Elevation-dependent weighting of $1/\sin^2(elevation)$ and an elevation cut-off of 15° were applied.

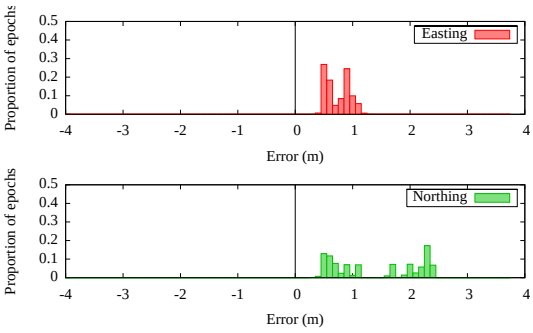


(a) L1 + L2 without phase-smoothing

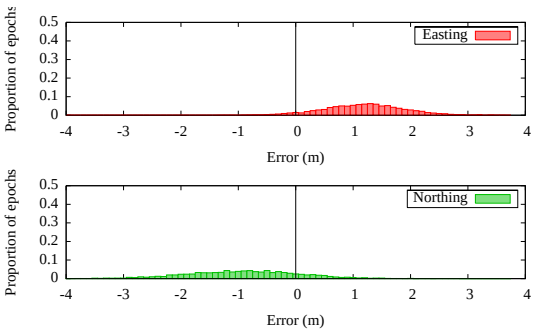
(b) L1 + L2 with phase-smoothing



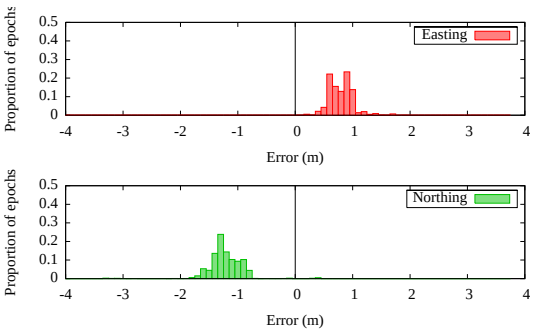
(c) L1 without phase-smoothing



(d) L1 with phase-smoothing



(e) Iono-free without phase-smoothing



(f) Iono-free with phase-smoothing

Figure 8.2: Point positioning histograms

Table 8.2: Point positioning accuracy (95%, m)

			Phase smoothed		
L1	L1+L2	Iono-free	L1	L1+L2	Iono-free
1·946	1·771	3·136	2·472	1·813	1·894

Table 8.2 shows the accuracy values for each technique: the highest achievable accuracy, 1·8 m for L1+L2, is not sufficient to meet the port navigation or automatic docking requirements of 1 m and 0·1 m respectively. Therefore the results are compared to the ocean, coastal and port approach requirements, where the 10 m accuracy was achieved using all processing techniques. With these requirements, there were no integrity, availability or continuity breaches for the duration of the trial for any of the processing techniques.

Figure 8.1 shows the position error against time for each positioning technique, and Figure 8.2 shows histograms of this error. As Figure 8.1 and Figure 8.2 show, the 95% position error statistic depends on both the noise and the offset from zero: if the positions were unbiased then the accuracy would be at the half-metre level if not using the ionosphere-free linear combination. The bias in Northing for the L1 and L1+L2 positions increases from $\sim$ 0·5 m to $\sim$ 1·5 m in the second half of the data processing period due to additional satellites rising at 19:12:36 and 19:26:23: these are low-elevation and therefore have high ionospheric delay, which increases the position bias. Figure 8.2 shows that the Northing results have about twice the standard deviation of the Eastings; this is due to the better East-West satellite geometry at higher latitudes (there are fewer satellites to the North).

Phase smoothing substantially reduces the noise, but does not reduce the mean offset from the truth; therefore the overall accuracy is not substantially improved. Indeed, the 95% accuracy is worse for L1 and L1+L2 with phase smoothing applied, as shown in Table 8.2.

The ionosphere-free linear combination eliminates ionospheric bias but multiplies the multipath and code measurement noise by a factor of 3, resulting in significantly increased noise in the position solution. The increase in position bias caused by the high ionospheric error on the additional low-elevation satellites visible for L1 and L2+L2 is not present. However, the offset from zero is not completely eliminated by the ionospheric-free combination, as this does not eliminate other errors such as the tropospheric delay or satellite orbit error. The higher measurement noise and multipath significantly reduces position noise: it is here that phase-smoothing has the greatest benefit, increasing the accuracy from 3·1 m to 1·9 m.

Table 8.2 shows that the L1+L2 observable is the most accurate at 1·8 m, and single-frequency positioning is 0·1 m worse. The ionosphere-free positioning is twice as poor, but phase-smoothing increases the accuracy to close to L1+L2. The phase-smoothed ionosphere-free processing technique also has the advantage that the accuracy should carry better to more difficult environments, as ionospheric delay and code multipath are cancelled and significantly reduced respectively.

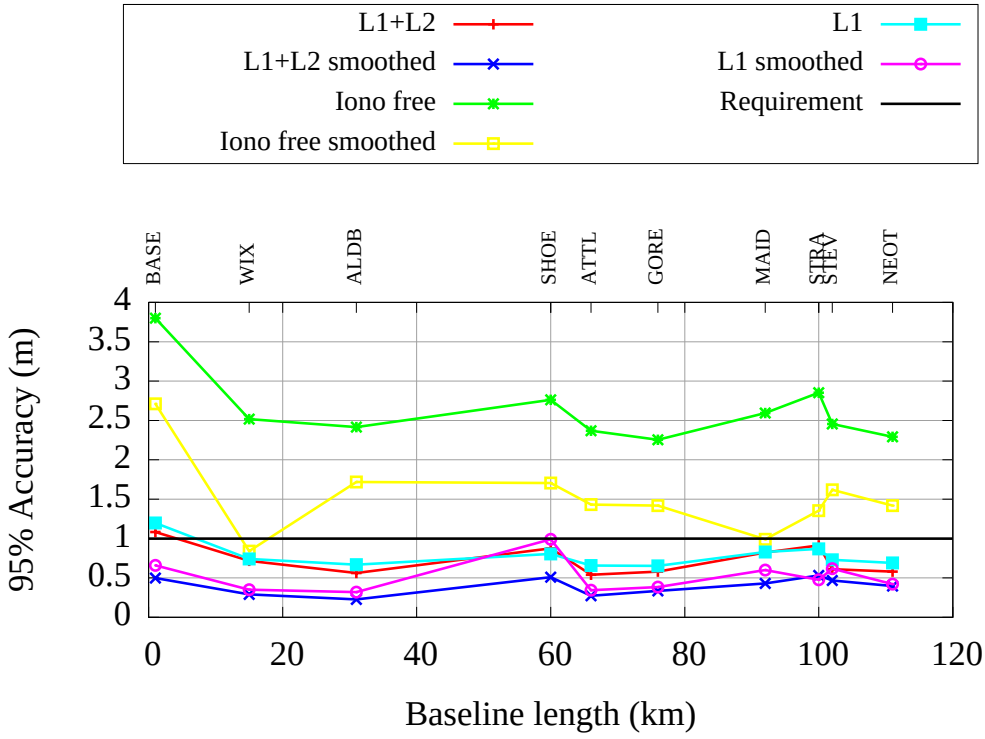


Figure 8.3: DGNSS accuracy

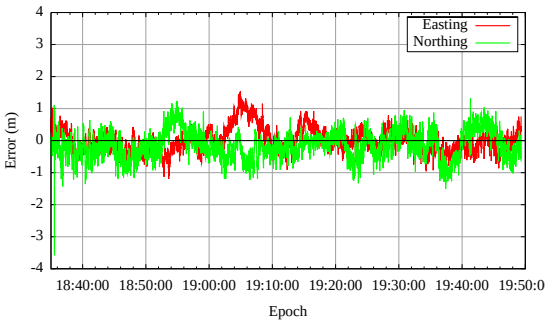
8.3 DGNSS

Differential GPS processing was performed using the L1, L1+L2 and ionosphere-free observables, both with and without phase smoothing applied, for each of the 10 different baseline lengths. The processing parameters were the same as for point positioning, except that no elevation-dependent weighting was applied.

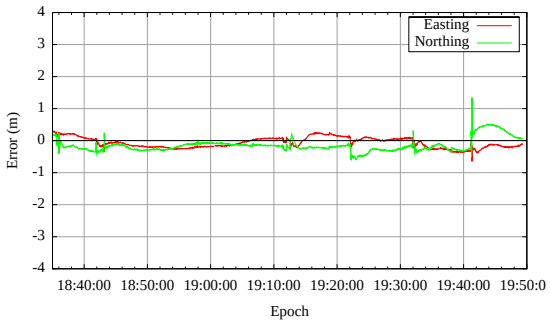
The accuracy results for all baselines and processing techniques are shown in Figure 8.3. The automatic docking accuracy requirements of 0·1 m are not met over any baseline length: the best achievable DGNSS accuracy was 0·23 m. All techniques apart from ionosphere-free meet the port navigation accuracy requirement (1 m) for most baselines, and the ocean, coastal and port approach accuracy requirement (10 m) is met by all techniques for all baselines. When each technique is analysed with respect to the requirements it is sufficiently accurate to meet, there were no integrity, continuity or availability breaches over any baseline.

There are 60 different combinations of processing technique and baseline: too many to analyse individually. Two cross-sections of the results are studied in depth: all processing techniques from BASE, and then a single processing technique from all reference stations.

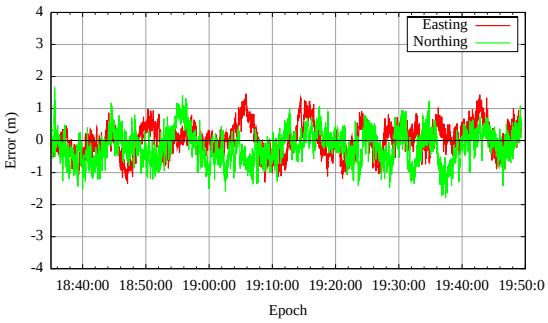
Figure 8.4 shows the error for each epoch of positions from BASE (1 km baseline) for each processing technique, and Figure 8.5 shows histograms of these errors. Unlike the point positioning errors in Section 8.2, the DGNSS errors are centred about zero: the single differencing has eliminated the systematic offset visible in Figure 8.1 and Figure 8.2. As noted in Chapter 6, the receiver at BASE was in a high-multipath environment: the characteristic periodic error caused by multipath is visible in the non-phase-smoothed data in Figure 8.4, and is particularly severe for the ionosphere-free results, with an amplitude of 3 m compared to 1 m for L1 and



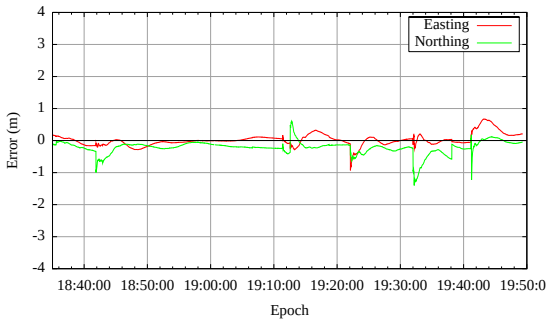
(a) L1 + L2 without phase-smoothing



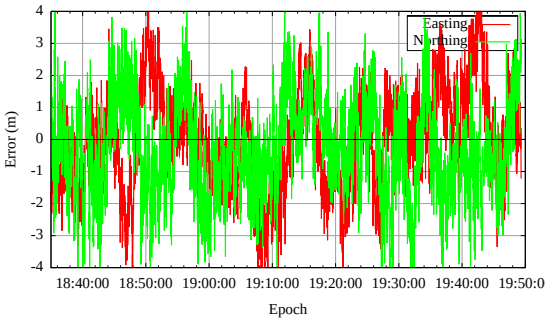
(b) L1 + L2 with phase-smoothing



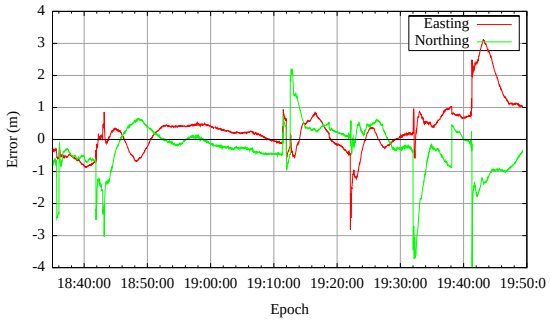
(c) L1 without phase-smoothing



(d) L1 with phase-smoothing

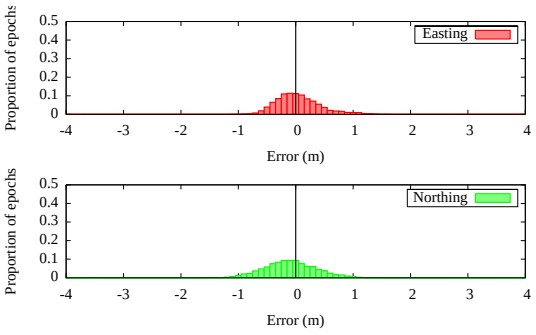


(e) Iono-free without phase-smoothing

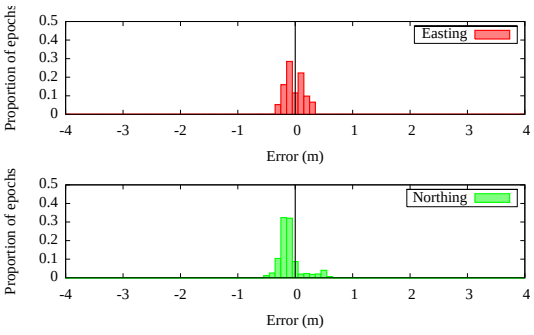


(f) Iono-free with phase-smoothing

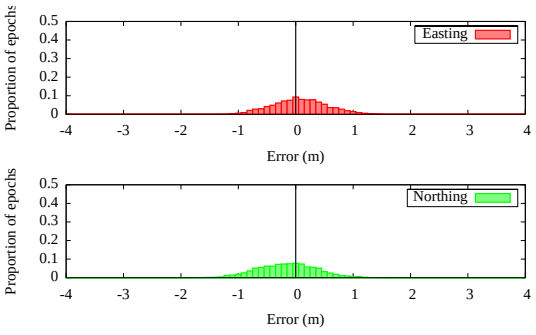
Figure 8.4: DGNSS accuracy from BASE



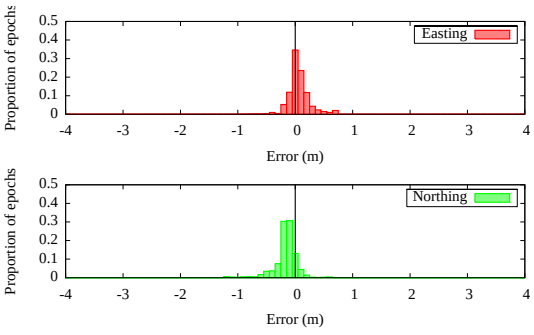
(a) L1 + L2 without phase-smoothing



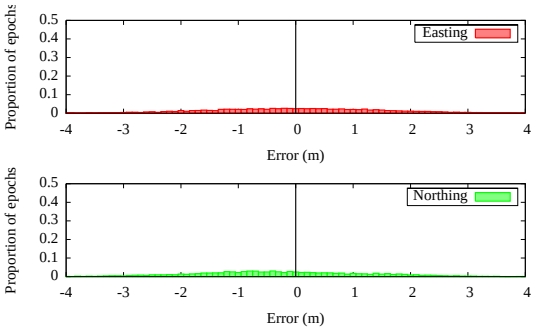
(b) L1 + L2 with phase-smoothing



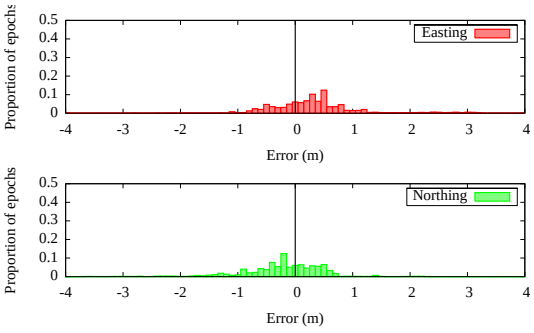
(c) L1 without phase-smoothing



(d) L1 with phase-smoothing

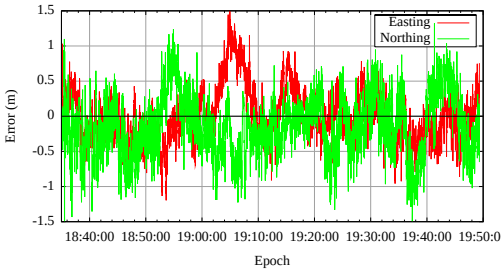


(e) Iono-free without phase-smoothing

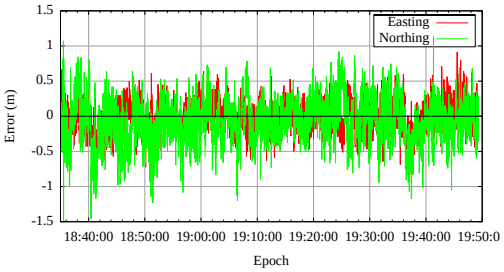


(f) Iono-free with phase-smoothing

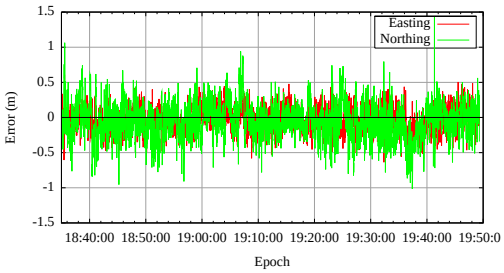
Figure 8.5: Histograms of DGNSS accuracy from BASE



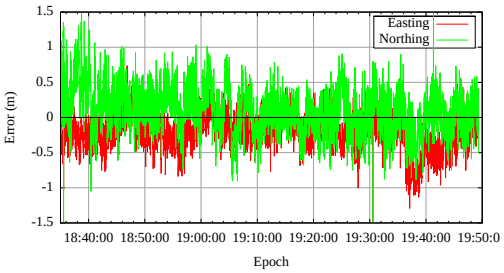
(a) BASE (1 km)



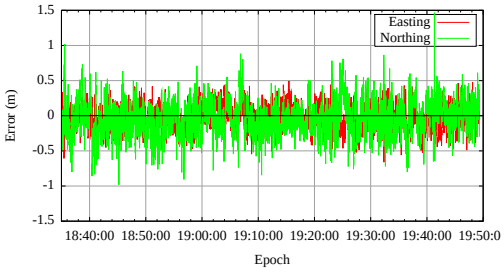
(b) WIX (15 km)



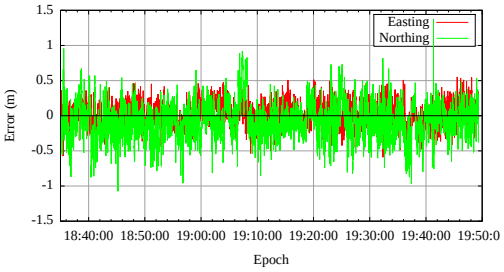
(c) ALDB (31 km)



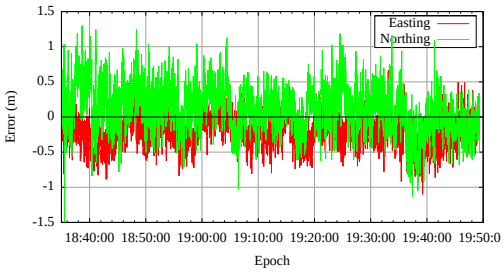
(d) SHOE (60 km)



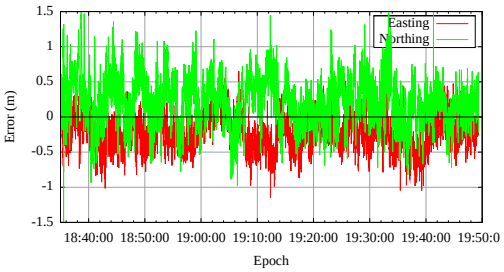
(e) ATTL (66 km)



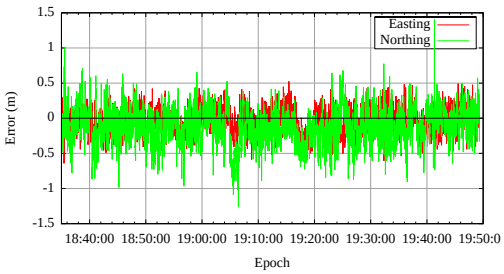
(f) GORE (76 km)



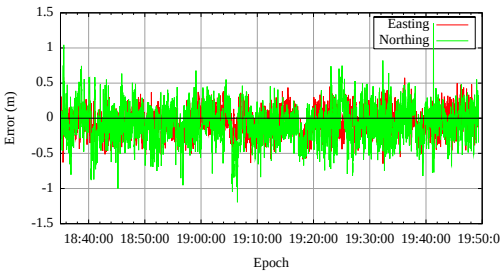
(g) MAID (92 km)



(h) STRA (100 km)

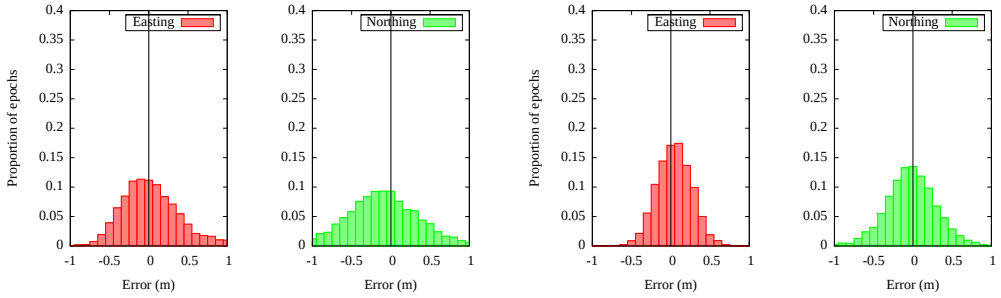


(i) STEV (102 km)



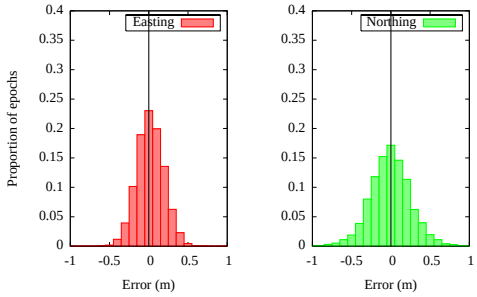
(j) NEOT (111 km)

Figure 8.6: Dual-frequency DGNSS accuracy

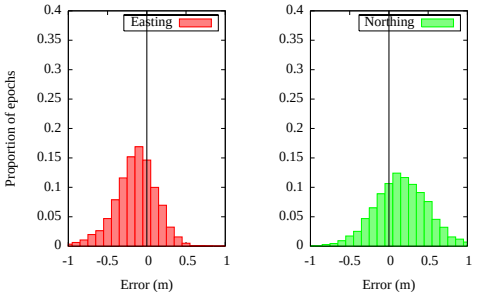


(a) BASE (1 km)

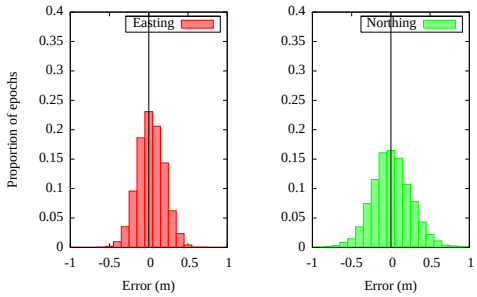
(b) WIX (15 km)



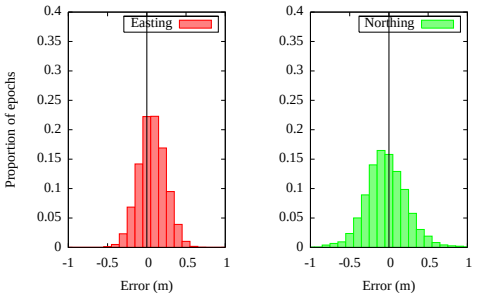
(c) ALDB (31 km)



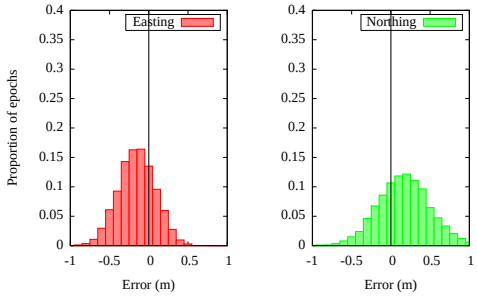
(d) SHOE (60 km)



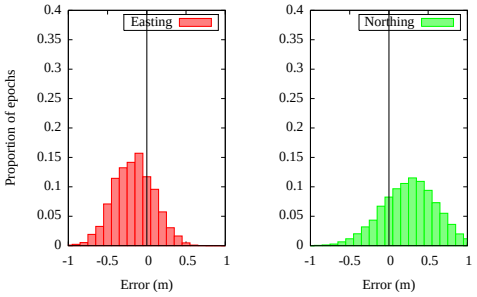
(e) ATTL (66 km)



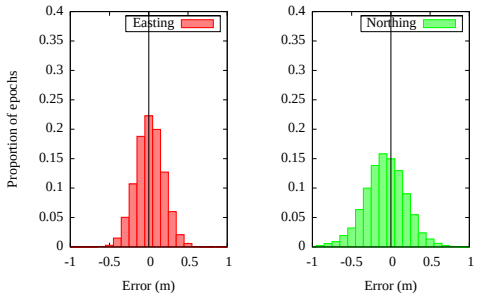
(f) GORE (76 km)



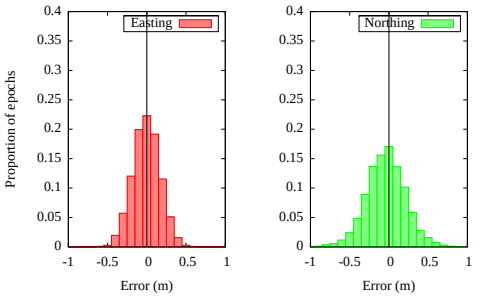
(g) MAID (92 km)



(h) STRA (100 km)



(i) STEV (102 km)



(j) NEOT (111 km)

Figure 8.7: Histograms of dual-frequency DGNSS accuracy

L1+L2. This is expected because the ionosphere-free combination multiplies the multipath and code noise error by 3. Phase-smoothing reduces the multipath error when using L1 or L1+L2 to give half-metre accuracy, but the phase-smoothed ionosphere-free results are still very poor; there is almost no ionospheric error to cancel over this short baseline, so the only effect of making this combination is to increase the noise and reduce the number of observables. However, over long baselines or at times of high ionospheric activity, this trade-off may be worthwhile in order to eliminate ionospheric noise. The single-frequency positioning is of similar quality to dual-frequency, and the accuracy in Easting is similar to the accuracy in Northing.

Figure 8.6 shows the dual-frequency positioning error for each epoch for each baseline. The least accurate positioning is from BASE (1·1 m), which is the shortest baseline, due to the high-multipath environment at the reference station. WIX, which uses the same receiver as BASE and is also a short baseline, is significantly better (0·7 m) due to the lower multipath at the reference station. The Ordnance Survey reference stations SHOE, MAID and STRA use a different type of receiver to ALDB, ATTL, GORE, STEV and NEOT: this latter group have phase-smoothing enabled at the reference station while the former do not. These two sets of stations are clearly distinguishable in the error histograms in Figure 8.7: the non-phase-smoothed group have a larger standard deviation and the mean error is offset from zero. The results from the phase-smoothed reference stations are still noisier than the phase-smoothed results in Figure 8.4, because the data at the rover are not phase-smoothed.

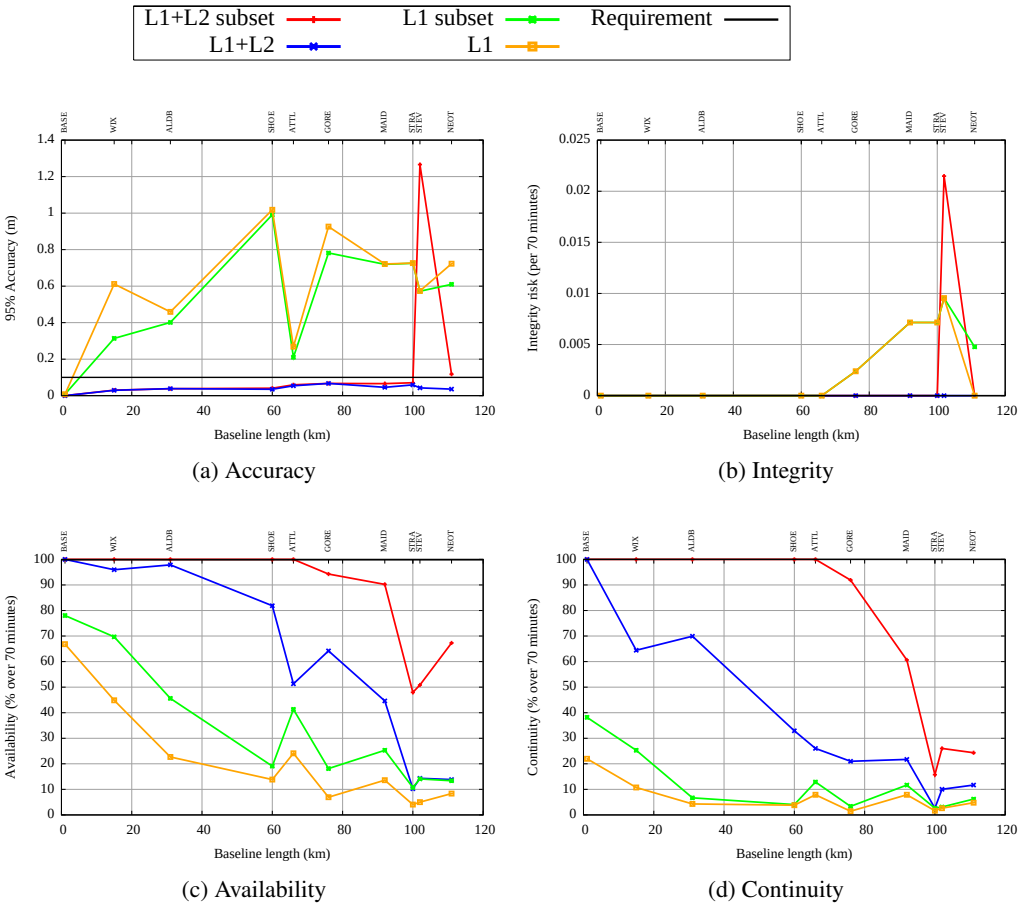
Figure 8.3 shows the accuracy results for every processing technique and baseline. The poor reference station environment at BASE is visible as reduced accuracy for all processing techniques from this reference station. There does not appear to be a correlation between baseline length and accuracy: the multipath error dominates the single-difference atmospheric and satellite orbit errors. The relatively high errors from STRA, MAID and SHOE are due to the receiver type.

The 1 m accuracy requirement was met using all techniques except ionosphere-free for all baselines except from BASE, where only the phase-smoothed results meet the requirement: this highlights the importance of correctly siting the reference station, and the benefits of phase-smoothing in removing multipath errors. The L1 results are slightly worse than the L1+L2, but L1 still met the accuracy requirement over all baselines. The correlation between accuracy and baseline length is low, so it is likely that the accuracy requirement could be met over even longer baselines. However, this may not be the case in times of higher ionospheric activity, and if it were necessary to use the ionosphere-free observable then it would not be possible to meet the port navigation accuracy requirements using DGNSS.

8.4 RTK

The RTK solution, with correct ambiguity resolution, is capable of meeting the strictest accuracy requirement of 0·1 m for automatic docking. The data are therefore processed with regard to this requirement. The subset ambiguity algorithm described in Chapter 4 is also applied to determine its effectiveness.

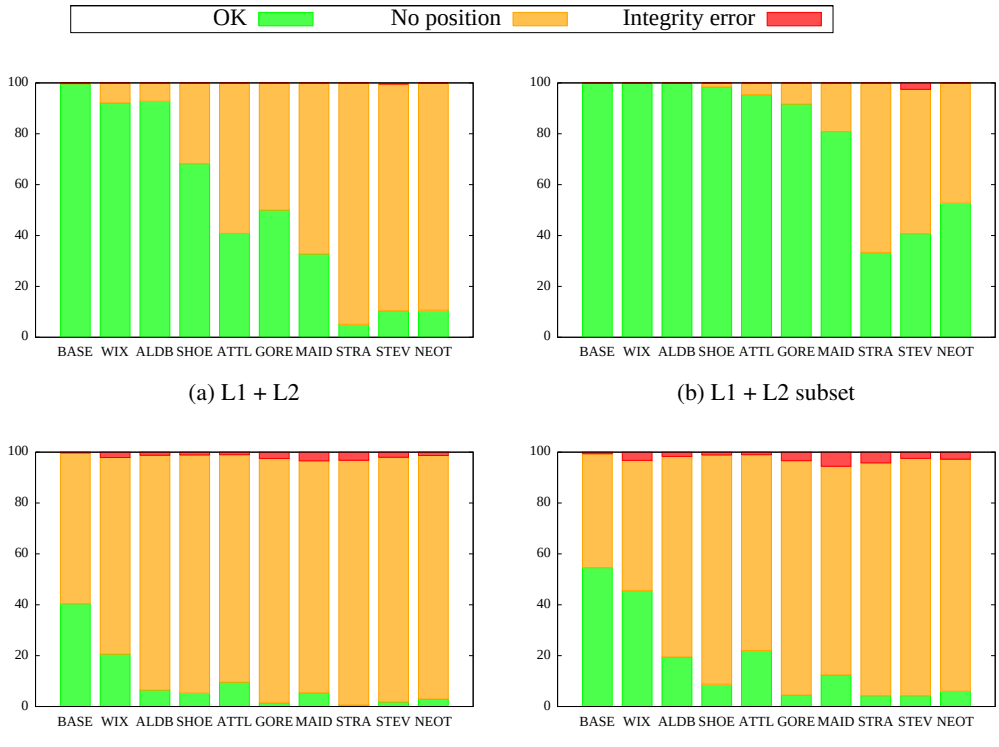
The ability of RTK to meet the requirements is dependent upon the single-epoch ambiguity



(a) Accuracy (b) Integrity

(c) Availability (d) Continuity

Figure 8.8: RTK positioning results



(a) L1 + L2 (b) L1 + L2 subset

(c) L1 (d) L1 subset

Figure 8.9: Ambiguity resolution success rate

resolution success rate: if the ambiguities are not resolved then the float solution must be used, which is equivalent to the DGNSS solution in Section 8.3 and is not sufficiently accurate to meet the automatic docking requirements. Figure 8.9 shows the ambiguity resolution success rate for each technique and baseline length. A success rate close to 100% was only achieved for dual-frequency processing over 1 km, or up to 30 km when using the subset ambiguity algorithm. The success rate decreases rapidly with baseline length due to atmospheric and orbit error decorrelation: the dual-frequency success rate is 81% over 76 km and 41% over 102 km. With L1+L2, incorrect ambiguity resolution only occurs at the longest baseline; the incorrect ambiguities are propagated with the subset ambiguity algorithm, causing some epochs that were not fixed to be fixed incorrectly; this increases the proportion of incorrectly fixed epochs from STEV from $0{\cdot}4\%$ to $2{\cdot}5\%$. Incorrect ambiguity resolution is common with single-frequency positioning, and the success rate is very low, only rising above 50% when using the subset ambiguity algorithm over the shortest baseline.

Figure 8.8a shows the accuracy results. L1+L2 met the requirements, achieving centimetric accuracy. As discussed in Chapter 6, the "truth" positions are the L1+L2 positions from BASE, with the ambiguity resolution validated by the total station observations: it is possible that these positions will will deviate from the true positions by up to $1/2$ cycle, or 12 cm. However, the major error source for the "truth" is phase multipath, with a maximum effect of $1/4$ cycle, as most other errors are cancelled over the short baseline; it is unlikely that the error often reaches this level. Therefore, these results give confidence that the dual-frequency positions are sufficiently accurate to meet the $0{\cdot}1$ m automatic docking accuracy requirement. The low L1 accuracy is caused by incorrect ambiguity resolution, which can cause a position error of several metres, as shown in Chapter 6. There was no incorrect ambiguity resolution for dual-frequency RTK except over the longest baselines. The peak at STEV for the dual-frequency subset ambiguity resolution is caused by the propagation of incorrectly resolved ambiguities. The short-baseline data are the truth model, and so have perfect accuracy. Except for the effects of incorrect ambiguity resolution, the accuracy is not significantly affected by baseline length.

Figure 8.8b shows the integrity errors for each baseline. If the ambiguities are correctly determined with the single-epoch technique then the maximum possible bias in any phase observation is $1/2$ cycle: if the bias were greater than this then the ambiguity would change so that the phase residual was less than $1/2$ cycle. The maximum position error is therefore $1/2$ cycle, or 12 cm for L2, which would occur if every observation had a $1/2$ cycle bias aligned in the same direction; this is less than the 25 cm alert limit requirement. Therefore an integrity error can only occur when the ambiguities have been incorrectly resolved. In general, only with single-frequency positioning are there enough epochs with incorrectly resolved ambiguities to cause an integrity breach, as shown in Figure 8.8b; the occasional incorrect ambiguity is absorbed by the time-to-alarm allowance. As seen in Figure 8.8a, the propagation of incorrect ambiguities by the subset ambiguity resolution algorithm has caused a high incidence of integrity breaches from STEV. Due to the small time-span of data analysed, even a single integrity breach is enough to cause an integrity risk much greater than the requirements allow. The probability of incorrectly resolving ambiguities increases proportionally to baseline length, which results in a reduction of system integrity over longer baselines.

Figure 8.8c shows the availability results. Availability is dependent on successful ambiguity resolution, and therefore drops off quickly as the baseline length increases. Single-frequency positioning does not achieve 70% availability, even over the 1 km baseline, and therefore does not meet the requirements over any baseline length. Dual-frequency positioning does not fail the availability requirements over the 1 km baseline from BASE, but the availability drops off rapidly (it is only 96% over 15 km), so the requirement is certainly not met over longer baselines. The subset ambiguity resolution algorithm extends the maximum baseline length to 66 km for L1+L2; the single-frequency availability is also increased, but still does not meet the requirements.

Figure 8.8d shows the continuity results. Continuity is similar to availability in that it is very dependent on successful ambiguity resolution and therefore decreases rapidly with increasing baseline length. The requirements are not breached by dual-frequency processing over 1 km, and up to 66 km when using the subset ambiguity resolution algorithm. Single-frequency processing does not provide sufficient continuity to meet the requirements.

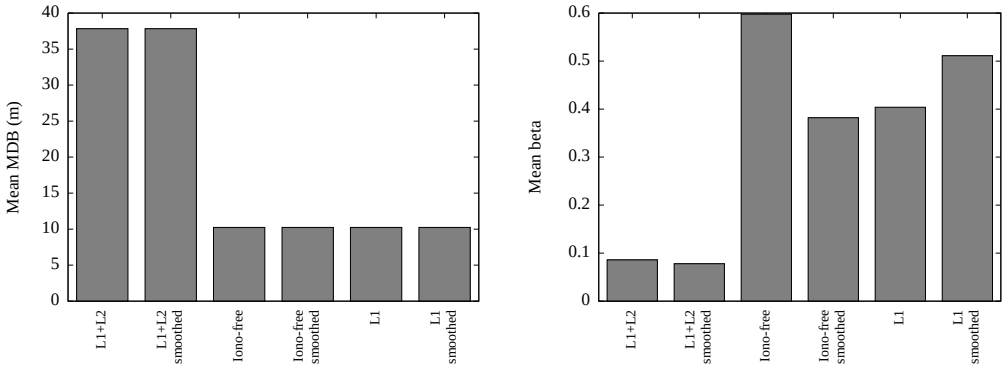
When all requirements are taken together, the only situations where the automatic docking requirements are not failed is when using dual-frequency processing over the 1 km baseline, and when using the subset ambiguity resolution algorithm over all baselines up to 66 km. The clearly shows the benefit of this algorithm. The single-frequency ambiguity resolution success rate is too low to allow any requirements to be met over any baseline length. Unlike with point positioning and DGNSS, availability and continuity, rather than accuracy, is the limiting factor with RTK. This also means that when RTK does not meet the automatic docking requirements, it also does not meet any other, less strict, requirements, because the continuity and availability requirements are the same. Also, unlike point positioning and DGNSS, using two frequencies greatly improves the positioning solution. The critical value of the ratio test could be increased for single-frequency positioning, in order to reduce the number of integrity errors and increase the 95% accuracy, but this will further reduce the availability and continuity.

8.5 Robustness to outliers

8.5.1 Code

During the data collection exercise there were no code outliers of sufficient magnitude to cause the position error to exceed the alert limit. However, due to the short period of data collected it is not possible to analyse how well the integrity requirements may be met over longer time periods, given the small allowable fail rate: 446×10 s periods of data were collected, and the integrity requirement is failed if greater than 1 in 10^5 of 3-hour periods contain an integrity breach.

As discussed in Chapter 5, it is possible to compute the minimum size of the bias (Marginally Detectable Bias, MDB) in each observation that will cause a position error equal to a given alert limit, and the probability (β) that this bias will not be detected by the RAIM algorithm. If the probability of this bias occurring is known then it is also possible to determine if the positioning solution meets the integrity requirements. However, this value is difficult to determine due to the low frequency of occurrence of such events and the dependence on the local environment; any value chosen would be somewhat arbitrary. Therefore, in this analysis, no attempt is made



(a) MDB (b) Beta

Figure 8.10: Point positioning robustness to outliers

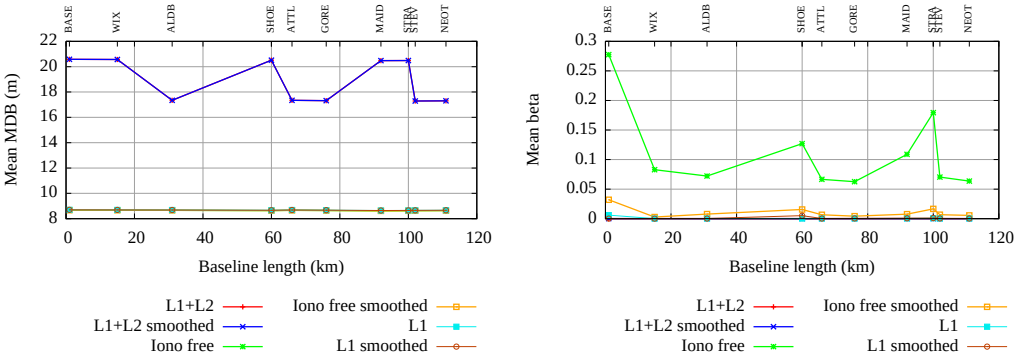
to tie the results to the requirements. The mean MDB and β over every satellite and epoch of data are compared across the different processing techniques and baselines to study the relative robustness to code outliers. Both the point positioning and DGNSS positions use the same alert limit of $2{\cdot}5$ m to facilitate comparison, although in practice point positioning is insufficiently accurate to meet this requirement: the actual point positioning alert limit would be 25 m, in which case the MDBs would be an order of magnitude larger, and correspondingly easier to detect (smaller β).

8.5.1.1 Point positioning

Figure 8.10a shows the mean MDB for point positioning for each different processing technique. Observations from the same satellite on different frequencies are assumed to have independent biases, so there are twice as many observations for L1+L2 than the other techniques; this has greatly increased the mean MDB from 10 m to 38 m. However, there is no difference between the techniques with the same number of observations; observation precision does not affect the MDB.

Figure 8.10b shows β, the mean probability of not detecting the MDB. The effect of the additional observations for L1+L2 is again the major factor, reducing β from $\sim 0{\cdot}5$ to $\sim 0{\cdot}08$, but there is also some difference between processing techniques: the probability of detecting the MDB improves with more precise observations, varying between $0{\cdot}60$ for ionosphere-free to $0{\cdot}38$ for phase-smoothed ionosphere-free.

Taken together these graphs show that the L1+L2 processing techniques, if errors on different frequencies from the same satellite are assumed to occur independently, have superior robustness to observation outliers: in order to cause a position error equal to the alert limit, a bias of $37{\cdot}8$ m would have to be present for L1+L2, with a $0{\cdot}09$ probability of not being detected; when using ionosphere-free this bias could be as small as $10{\cdot}24$ m with a probability of $0{\cdot}60$ of not being detected. This clearly shows the benefits of having more, precise observations; robustness to undetected outliers will greatly increase with the additional signals and satellites from modernised GPS and Galileo.



(a) MDB (b) Beta

Figure 8.11: DGNSS robustness to outliers for all baselines

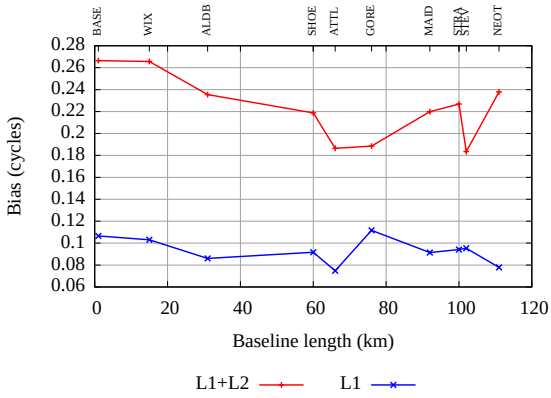


Figure 8.12: RTK phase bias robustness

8.5.1.2 DGNSS

Figure 8.11a shows the mean MDB for DGNSS for each different processing technique and baseline length. As with point positioning, the number of observations makes a significant difference, with the second frequency increasing the MDB from 9 m to 17 m or 20 m, depending on whether the reference station recorded the L2C observations. The processing technique again makes no difference.

Figure 8.11b shows the mean probabilities of not detecting the MDB for the different processing techniques and baseline lengths. As with point positioning, this is a combination of the number of observations and their precision. The ionosphere-free observable is significantly worse due to poor precision, with mean $\beta = 0{\cdot}11$ over all baselines. The mean over all baselines for L1+L2 is $\beta = 2 \times 10^{-4}$, compared to $\beta = 9 \times 10^{-4}$ for single-frequency: dual-frequency is an improvement, but the effect is not as strong as the size of the MDB.

The MDBs are smaller for a given technique than those for point positioning, but the mean β is also much smaller; smaller biases can cause the same position error, but these are much easier to detect.

Table 8.3: Mean L2 and L2C residuals (m)

	PRN 12		PRN 15		PRN 17	
Station	L2	L2C	L2	L2C	L2	L2C
SHIP	1·880	1·530	1·816	1·499	1·687	1·461
BASE	1·222	0·981	2·107	1·832	2·230	2·076
WIX	1·757	1·441	1·569	1·224	1·791	1·625
MAID	2·926	2·686	2·862	2·806	2·535	2·541
SHOE	2·745	2·518	2·543	2·468	2·377	2·330
STRA	2·722	2·537	2·741	2·687	2·388	2·427

8.5.2 Phase

As discussed in Chapter 5, it is only possible to perform the reliability analysis that was applied to the code on the ambiguity-fixed solution once the ambiguities have been successfully resolved, and the effect of biases at this stage is too small to significantly affect the solution. Instead, a technique to determine the average bias size that a positioning solution can tolerate before ambiguity resolution fails was used. Figure 8.12 shows this value averaged over all epochs for each baseline and for single- and dual- frequency. With single-frequency RTK, this value is around $0{\cdot}1$ cycle, whereas for dual-frequency positioning it is in the region of $0{\cdot}2 - 0{\cdot}25$ cycles: this shows that dual-frequency RTK is much more robust to biased observations. There is some difference between stations, and there may be a slight baseline-dependent effect, which could be due to the increased size of existing biases due to atmosphere over the longer baselines.

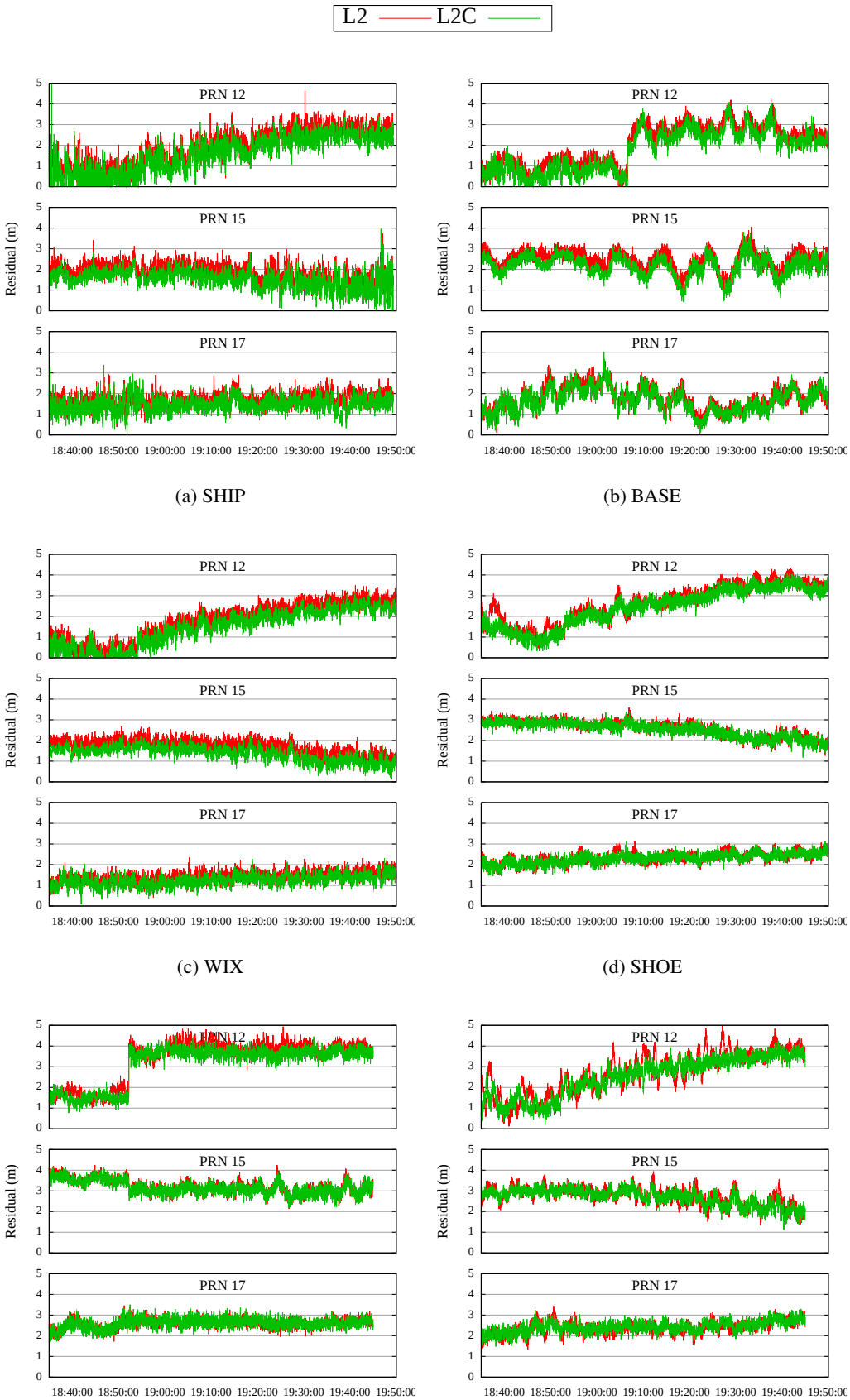
Single-epoch ambiguity resolution is affected by code bias as well as phase bias: a bias in a code observation will affect the float ambiguity values and will therefore also affect the ambiguity resolution. However, the relative robustness to this for each technique is the same as the robustness to code bias, so no further analysis is performed.

8.6 Analysis of L2C residuals

Several stations recorded both L2 and the new civil signal, L2C, from the three visible Block IIR-M satellites. Figure 8.13 shows the L2 and L2C point positioning residuals from an unweighted least-squares positioning solution from each station. The general trend of the residual for each satellite is the same for all stations and is caused by the change in satellite elevation over time.

There are two groups of receiver types: SHIP, BASE and WIX are Topcon GR-3's, while SHOE, MAID and STRA are Trimble NETR5's. The effect of the L2C signal is different for each group. For the GR-3's, the L2C residual follows the same pattern as the L2, but at a slightly reduced magnitude. So for example at BASE, in a high-multipath environment, the same multipath signal is displayed by both residuals. A different effect is displayed by the NETR5's: the means of the L2C residuals are not much smaller than for L2, but there is less multipath evident. This is particularly noticeable in PRN 12 from STRA and SHOE.

Table 8.3 shows a summary of the data, giving the mean point positioning residuals from L2 and L2C. The mean L2C residual was less than the mean L2 residual for all satellites and



(a) SHIP

(b) BASE

(c) WIX

(d) SHOE

(e) MAID

(f) STRA

Figure 8.13: L2 and L2C residuals

stations except for PRN 17 from MAID and STRA. The mean L2C residuals over all stations were 0·87, 0·90 and 0·95 times the mean L2 residuals for PRNs 12, 15 and 17 respectively: the L2C measurements were more accurate than L2.

8.7 Effect of improvements on positioning performance

This section describes how the performance enhancements described in Chapter 3 may affect the ability of the different GNSS processing techniques to meet the IMO requirements, in light of the results of the GPS-only positioning.

8.7.1 Point positioning

The best point positioning technique, using the phase-smoothed ionosphere-free observable, achieved an accuracy of 1·9 m: this is too low to meet the port navigation requirements of 1 m, but is well above the ocean and coastal navigation requirements of 10 m. The point positioning accuracy will be improved by the additional satellites, reduced code noise and multipath, and the lower-noise ionosphere-free observations; it is possible that the point positioning accuracy could reach the 1 m level, but this is unlikely to be achieved with sufficient availability. The improved Galileo broadcast ionospheric model will improve the accuracy of single-frequency positioning.

It was shown in Section 8.5 that the size of the observation bias that can be tolerated by the RAIM algorithm depends on the number of observations, and the probability of detecting the bias additionally depends on the accuracy of the code measurements; the additional satellites and frequencies should therefore greatly increase the bias robustness of point positioning, and point positioning should easily be capable of meeting the integrity requirements for ocean and coastal navigation.

8.7.2 DGNSS

Single- and dual- frequency DGPS met the 1 m accuracy requirements for port navigation over the longest baselines, except from reference station BASE which had high multipath. Atmospheric and orbit error are largely cancelled with differential positioning, so the improvement in multipath mitigation, code measurement precision and number of observations will have the most significant effect on DGNSS accuracy, particularly over shorter baselines; this should make sure that DGNSS is sufficiently accurate to meet the 1 m requirement in all conditions. The ionosphere-free GPS observable lacked the accuracy to meet this requirement, even when phase-smoothed, which will limit the maximum baseline length and could be a problem at times of high ionospheric activity. It is possible that the improved accuracy of the ionosphere-free observables with modernised GPS and Galileo could make 1 m accuracy achievable, which would allow DGNSS positioning over much longer baselines. Additional, more precise, observations will improve the DGNSS robustness in the same way as for point positioning.

Table 8.4: Processing techniques matched to requirements: failure type and baseline length

Technique	Point positioning	DGNSS	RTK	RTK+subset algorithm
Frequencies	L1	L1	L1+L2	L1+L2
Port approach	-	- up to 110 km	Continuity and availability failed > 1 km	Continuity and availability failed > 66 km
Port navigation	Accuracy failed	- up to 110 km	Continuity and availability failed > 1 km	Continuity and availability failed > 66 km
Automatic docking	Accuracy failed	Accuracy failed	Continuity and availability failed > 1 km	Continuity and availability failed > 66 km

8.7.3 RTK

RTK positioning had problems with availability and continuity, especially over longer baselines, due to the low ambiguity resolution success rate. In particular, the requirements were not met with a baseline of greater than 1 km with dual-frequency processing, and were not met at all for single-frequency. As discussed in Section 3.7.3, the ambiguity resolution success rate with modernised GPS and Galileo will be significantly increased. If the fail rate over medium baselines is decreased by a factor of 10 compared to dual-frequency GPS, as shown in Table 3.3, then the fail rate for the 15 km and 31 km baselines processed in Chapter 8 will be similar to the 1 km baseline with GPS alone, so it should be possible to meet the requirements up to this distance. Similarly, the maximum baseline length with the subset ambiguity resolution algorithm would be extended from 60 km to 92 km.

8.8 Summary

The GPS data collected in Harwich harbour were analysed using different processing techniques and compared to the IMO requirements as discussed in Chapter 2. The data were collected in a port navigation and automatic docking environment, but were also compared to the higher requirements for port approach, which is a more benign, lower multipath, environment. In order to make a positive statement that a requirement can be met with a specific technique, many years of data would have to be collected and analysed in different conditions and geographic locations, which is beyond the scope of this thesis. However, the results of the data collection exercise can be used to say which techniques can *not* be used to meet a given set of requirements: if the requirements are breached in this experiment then it is very unlikely that they would be met in the general case.

Table 8.4 shows the different processing techniques compared to the requirements, giving the failure type and baseline length. Accuracy was the limiting factor for point positioning, which could not meet the 1 m port navigation or the 0·1 m automatic docking requirement. DGNSS was also limited by accuracy, and could not meet the automatic docking requirement. In contrast to these code positioning techniques, RTK was limited by availability and continuity, due to the difficulty of achieving correct ambiguity resolution. However, this was mitigated to some extent by the novel subset ambiguity resolution algorithm developed in Chapter 4, which increased the baseline length over which the requirements were not failed from 1 km to 66 km. Single-

frequency positioning was as useful as dual-frequency for point positioning and for DGNSS for all baseline lengths, but was significantly worse for RTK, where the additional observations are essential to achieve successful ambiguity resolution.

Point positioning is the simplest technique to implement and does not require shore infrastructure. The accuracy is too low to meet the port navigation or automatic docking requirements, but is sufficient for port approach. The single-frequency results were almost as good as the dual frequency, and the ionospheric-free combination gave a similar accuracy, which would be important in periods of high ionospheric activity. There were no integrity, continuity or availability errors for the duration of the experiment.

DGNSS met the required accuracy for port navigation out to long baselines when using both single and dual frequency. However, during a period of high ionospheric activity it might be necessary to use the ionospheric-free observable to eliminate the ionosphere over longer baselines: for this experiment the resultant multiplication of code noise and multipath error would reduce the accuracy to below the port navigation requirements. High ionospheric activity could also reduce the range over which single-frequency DGNSS is practicable. There were no integrity, continuity or availability errors for the duration of the experiment.

Although RTK is more accurate than either point positioning or DGNSS, the ambiguity resolution step is not robust, particularly over longer baselines or with single-frequency data. This can result in the ambiguities not being fixed, affecting the continuity and availability, or being fixed incorrectly, causing an integrity error or reducing the accuracy. Because of this, no requirements are met over a baseline length greater than 1 km.

The subset ambiguity resolution algorithm extends the minimum baseline length over which the automatic docking integrity and continuity requirements are not failed using dual-frequency RTK from to 1 km to 66 km. However, over very long baselines the results were made worse by the propagation incorrectly fixed ambiguities. The algorithm is most useful to fill the gaps between successful ambiguity resolution over shorter baselines, where biases that prevent ambiguity resolution, such as multipath, affect different observations differently. Over long baselines, biases such as atmospheric decorrelation affect all observations and therefore provide less scope for successful subset ambiguity resolution. This experiment also demonstrates that it is dangerous to apply this algorithm in situations where the normal ambiguity resolution fixes the ambiguities incorrectly, as these errors are propagated to more epochs. However, the improvement in baseline length is significant, which shows the potential usefulness of this algorithm when attempting to meet the strict availability and continuity imposed by the IMO requirements.

In Section 8.5 it is shown that the number of available independent measurements strongly affects the bias robustness of both code- and phase- based positioning techniques. For RTK, the average size of the additional bias that can be tolerated in a single observation before ambiguity resolution is unsuccessful is more than doubled when using two frequencies, as is the mean MDB required to cause a position error equal to the alert limit for point positioning and DGNSS. The probability of detecting such a code bias is dependent on the precision of the code observations, and less advantage is derived from the additional frequency.

It was shown that the modernised L2C signal is not significantly more precise than the codeless L2; the greatest benefit may be for improved tracking robustness and signal acquisition in

difficult environments. In Section 8.7, the anticipated benefits of a future combined modernised GPS and Galileo system as discussed in Chapter 3 were interpreted in light of the ability of GPS to meet the IMO requirements. It is unlikely that the improvements will be sufficient to allow point positioning or DGNSS to meet the next-highest requirements. However, the improvements to accuracy and robustness will increase the likelihood of DGNSS meeting the 1 m port navigation requirement, even over long baselines and with high multipath and ionospheric activity. The greatest benefit will be when using RTK to meet the automatic docking requirements: as shown in Chapter 8, this is marginal with current dual-frequency GPS, but this may be possible over moderate-length baselines in the future.

It therefore seems likely that with modernised GPS and Galileo, point positioning would be sufficient to meet the port approach requirement, and DGNSS would be sufficient to meet the port navigation requirement over long baselines, except perhaps during periods of high ionospheric activity. RTK might be sufficient to meet the automatic docking requirements over short baselines; the subset ambiguity resolution algorithm would increase the robustness and may allow the requirements to be achieved over medium-length baselines. However, in order to demonstrate these conclusions, large quantities of data will have to be analysed from many different environments and geographical locations. In particular, the data collected at Harwich may have relatively low ionospheric error due to the state of the ionosphere at the time of data collection, and some port environments may exhibit greater multipath error due to the proximity of other ships, containers or cranes. However, multipath can be mitigated to some extent by phase-smoothing, and the superior characteristics of the modernised signals will result in a further reduction of the effect of multipath and the ionosphere, as discussed in Chapter 3.

Chapter 9

Conclusions and further work

9.1 Conclusions

The main aim of the thesis is to analyse the potential of different GPS positioning techniques to meet different IMO requirements, in particular port approach, port navigation and automatic docking, and to make a theoretical assessment of how this situation may change with the advent of modernised GPS and Galileo. The main way of achieving this objective is the analysis of real data collected in a realistic environment, with an independent positioning system as a truth model. An automatic tracking total station, which measures angles and distances to a prism mounted underneath the GPS antenna, was chosen to provide the truth model. Over the course of several preliminary data collection exercises, the technique for the truth model was developed, problems related to the tracking and the synchronisation of the observations were identified and resolved, and lessons learnt to apply to the final data collection. In particular, it was decided that the ideal range for the total station would be < 400 m, and it would be best to have two total stations tracking the prism at one time in order to improve robustness and precision.

The initial plan for the data collection was for the ship, *THV Alert,* to approach and enter Harwich harbour whilst being tracked by two total stations, to recreate a port approach environment. The two total stations would then be moved to allow the tracking of the *Alert* as she navigation through the port and then came to dock, covering the port navigation and automatic docking phases of the IMO requirements. However, initial tests in the morning of the data collection exercise showed that, due to adverse weather affecting the total station tracking and shallow water, *Alert* could not get sufficiently close to the total stations to allow tracking during the port approach phase; therefore the final data collection exercise was limited to the port navigation and automatic docking phases. However, the results are still compared to the port approach requirements due to the similarities of the environment. The final exercise saw *Alert* make several runs in front of the two total stations and come up alongside the jetty. GPS reference station data were collected from nearby Ordnance Survey reference stations and from two additional reference stations that were set up nearby to enable short-baseline processing. Data from the new modernised GPS signal, L2C, were collected at several stations and from *Alert* from three modernised GPS satellites in order to allow an analysis of the benefits of this signal.

The total station tracking was not sufficiently smooth to allow a direct comparison to the GPS positions. Lack of synchronisation between the distance measurements, the angle measurements, and the time-stamping of the data meant that after solving for the clock offset and interpolating the positions to the same time, the prism positions from each total station only

agreed at the $0{\cdot}5$ m level, which is too poor to enable analysis of the $0{\cdot}1$ m automatic docking requirements. The agreement between the the prism positions and the 1 km baseline GPS positions was at the $0{\cdot}2$ m level, which is sufficiently precise to give confidence in the correctness of the GPS ambiguity resolution (incorrect ambiguity resolution generally causes errors of several metres). This implies that these positions are accurate to within 12 cm, and therefore can be used as the truth model for all other positioning techniques.

Existing GNSS processing software is not sufficiently flexible to allow different processing techniques and models to be properly assessed and further techniques developed. Therefore GNSS processing software was developed to allow a flexible choice of models and techniques. In particular, the GPS data were processed using point positioning, DGPS and RTK over all baseline lengths and with the best models available without external input.

Point positioning achieved an accuracy of ~ 2 m, which is sufficient to meet the ocean and coastal navigation accuracy requirement of 10 m, but no stricter requirements. The ionosphere-free observable had an accuracy worse than 3 m, but phase-smoothing improved this to better than 2 m. This accuracy would be achievable in times of high ionospheric disturbance and in places with high code multipath, as these errors would be cancelled and significantly reduced, respectively. The single-frequency accuracy was similar to that for dual-frequency. There were no problems with availability, continuity or integrity.

DGNSS achieved an accuracy of $\sim 0{\cdot}5$ m over most baselines with single- or dual- frequency positioning, which is sufficient to meet the port navigation accuracy requirements. The exception was the shortest baseline (BASE, 1 km), where high multipath at the reference station reduced the accuracy to worse than 1 m. Phase-smoothing increased the accuracy from all stations, and in particular increased the accuracy of positioning from BASE to below 1 m, sufficient to meet the port accuracy requirements. There is not a strong correlation between accuracy and baseline length, so it is likely that these requirements would be achieved over significantly longer baselines. However, the ionosphere-free accuracy was at the 3 m level, and even when phase-smoothed only achieved, $\sim 1{\cdot}5$ m, so it would not be possible to meet the port navigation requirements over very long baselines or at times of high ionospheric activity. There were no problems with availability, continuity or integrity, and single-frequency positioning achieved similar results to dual-frequency.

RTK, with correct ambiguity resolution, was sufficiently accurate to meet the automatic docking requirements. However, low ambiguity resolution success rates caused continuity and availability problems, which meant that the requirements were failed except for over the 1 km baseline with dual frequency. Unlike the code-based positioning techniques, dual-frequency RTK was greatly superior to single-frequency; ambiguity resolution success rates over 1 km were $90{\cdot}6\%$ and 40% respectively. Success rate, and hence availability and continuity, decreased rapidly with baseline length: the dual-frequency success rate was 81% over 76 km and 41% over 102 km. Over the longest baselines, incorrect ambiguity resolution caused some integrity errors.

Given a bias in an observation of a specified size, conventional Receiver Autonomous Integrity Algorithms (RAIM) allow the computation of the probability that the bias will be detected, and the effect on the positioning solution if it is not detected. However, in order to relate this infor-

mation to the IMO integrity requirements, the probability of the error of a given size occurring must be known; this is hard to estimate due to the rarity of such events. The situation is further complicated by the 10 s time-to-alarm limit, which requires knowledge of how biases are correlated in time. Therefore, in order to make an assessment of the relative robustness to bias of different techniques, the algorithms were reversed so that, given the alert limit, the size of bias in a given observation required for the position to breach the alert limit can be computed, as well as the probability of detecting this bias. A technique was also developed to study the size of phase bias required to cause ambiguity resolution to fail.

It was shown that the size of observation bias that can be tolerated depends strongly upon the number of independent observations for both code- and phase- based positioning; this will increase greatly with additional satellites and frequencies (assuming that biases are not correlated between frequencies, which will be the case for some error sources). The probability of detecting code biases depends upon the precision of the code measurements; phase biases cannot be detected in a single epoch before ambiguity resolution occurs, and are not significant if the ambiguities have been correctly resolved.

The receivers that recorded L2C also recorded L2 from the same satellite. An unweighted point positioning adjustment was performed, and the L2 and L2C residuals compared. On average, the point positioning residuals for the L2C observations were found to be $0{\cdot}9$ times the L2 residuals for the same satellite, showing that the increase in measurement precision on L2C is relatively small. A different pattern was noted for different types of receiver. It may be that the greatest benefit of L2C is in increased tracking robustness and signal acquisition in difficult environments.

The improvements in the signal structure of the modernised signals were discussed. The modernised signals will feature longer, faster, tiered codes which will be more precise and robust, and more resistant to multipath. Pilot signals, transmitted without the data message, will allow more robust phase tracking, and the new signals will be transmitted with greater power. Modernised GPS and Galileo will be interoperable at two frequencies, and time system and reference frame differences can be overcome or ignored; a receiver will easily be able to track double the number of satellites. Ambiguity resolution success rate will be substantially increased, and the ionosphere-free observable will have less noise. It is unlikely that these benefits will allow any processing technique to meet stricter requirements than currently achievable with GPS alone, but will make it easier for each technique to meet its respective requirements. Point positioning was easily accurate enough to meet the 10 m accuracy requirements, so there will be little benefit from modernised GNSS. DGNSS may be able to meet the 1 m requirement more robustly, over longer baseline lengths and in worse ionospheric and multipath environments, or with greater occlusion of the sky. The greatest benefit will be for meeting the $0{\cdot}1$ m requirement with RTK; this was not possible over a baseline greater than 1 km in this experiment with current GPS alone; it may be possible to achieve this over a 30 km baseline with modernised GPS and Galileo.

It was determined that the problem of ambiguity resolution severely affected the ability of RTK to meet the IMO requirements for automatic docking: the difficult environment meant that the ambiguities were often not resolved, and therefore no position achieved, especially over longer baselines. Partial ambiguity techniques were discussed, which allow a subset of the am-

biguities to be resolved, providing a greater probability of success at the expense of reduced position accuracy. However, the current procedure assumes unbiased observations, and so is not necessarily suitable for a difficult environment such as a port, where there may be high multipath on some signals that prevents successful ambiguity resolution. Therefore, a new technique for partial ambiguity resolution in the presence of biased phase observations was developed. This technique is only applied when normal ambiguity resolution has failed. The expected value of the ambiguities for the current epoch is determined by taking the weighted mode of the values from previous epochs. All subsets of the full set of ambiguities are generated and ordered according to some criterion, such as ambiguity dilution of precision. Each subset is then fixed in turn, if the probability of success is sufficiently high. If a subset is fixed to the expected values then it is accepted; if no subsets are fixed correctly then the ambiguities are kept unfixed this epoch. Once some ambiguities have been fixed, a new adjustment is performed in order to increase the accuracy of the final solution, using the fixed ambiguities to provide precise phase range measurements with the unfixed ambiguities as unknown parameters. Experimental results show that this algorithm is effective at increasing the ambiguity resolution success rate and that it improves the ability of RTK to meet the IMO requirements, increasing the minimum baseline length over the requirements were not failed from 1 km to 66 km. Due to the need to perform the ambiguity resolution procedure many times, the algorithm is very computationally intensive; however, it was shown that applying a time-based cutoff to processing with this algorithm may allow many of the benefits to be obtained with a reasonable processing time, and the technique may be particularly suitable for processing where the observations from the rover are sent back to a reference station for processing. Several techniques were suggested to improve the processing speed of the algorithm.

9.2 Meeting the requirements

Although, with a small set of data, it is not possible to positively show that the requirements can be met in all locations and scenarios, it is possible to determine which techniques will not be suitable to meet a given requirement, and to get an idea about which techniques are likely to be able to meet the requirement. This section discusses how well the different requirements were met with current GPS over the course of the Harwich data collection, the likely effect of modernised GPS and Galileo, and what this will mean for Aids to Navigation (AtoN) service provision.

9.2.1 Port approach

Table 9.1: Port approach requirements

Accuracy	Integrity			Availability	Continuity
Horizontal (m)	Alert limit (m)	Time to alarm (s)	Integrity risk (per 3 hours)	% per 30 days	% over 3 hours
10	25	10	10^{-5}	99·8	99·97

The 10 m accuracy requirement means that port approach should be achievable with current GPS, as an accuracy of $\sim$ 2 m was achieved in this experiment using the phase-smoothed ionosphere-free observation, which would have been achieved even in times of high ionospheric activity or in situations with high multipath. The alert limit is also high and so the integrity requirements should be achievable, depending on the frequency of occurrence of outliers. Based on the results of this experiment, the availability and continuity did not present a problem, but a large quantity of data would need to be analysed in order to properly assess this. Modernised GPS and Galileo will improve the positioning performance, but given that the requirements are likely to be met with GPS alone, the benefit is not likely to be significant, although the availability and continuity will be increased. DGNSS, requiring a shore-based reference station, is more accurate than point positioning and would likely also meet the port approach requirements. However, RTK, except over very short baselines, does not meet the availability or continuity requirements, and so is not suitable for meeting the port approach requirements. It is likely that the port approach requirements may be met without modernised GNSS or AtoN service provision; however, either of these may be necessary to assist in meeting the integrity, availability, or continuity requirements.

9.2.2 Port navigation

Table 9.2: Port navigation requirements

Accuracy	Integrity			Availability	Continuity
Horizontal (m)	Alert limit (m)	Time to alarm (s)	Integrity risk (per 3 hours)	% per 30 days	% over 3 hours
1	2·5	10	10^{-5}	99·8	99·97

The port navigation accuracy requirement of 1 m precludes the use of point positioning, which achieved an accuracy of $\sim$ 2 m. RTK did not meet the availability and continuity requirements over baselines longer than 1 km, so it is unlikely that RTK could be used to meet the port navigation requirements for large ports. This leaves DGNSS as the remaining positioning technique. Phase-smoothed dual- or single- frequency DGNSS achieved an accuracy of $\sim$ 0·5 m over all baselines, with no integrity, availability or continuity errors, so it is likely that this technique could be used to meet the port navigation requirements over long baselines, i.e. requiring little shore infrastructure to cover a large area. However, the ionosphere-free accuracy was at the 1·5 m level, even when phase-smoothed, which is insufficient to meet these requirements. Therefore shorter baselines might be necessary to meet these requirements at times of high ionospheric activity. However, the additional observations and improved ionosphere mitigation from modernised GPS and Galileo will extend the baseline length.

9.2.3 Automatic docking

Table 9.3: Automatic docking requirements

Accuracy	Integrity			Availability	Continuity
Horizontal (m)	Alert limit (m)	Time to alarm (s)	Integrity risk (per 3 hours)	% per 30 days	% over 3 hours
0·1	0·25	10	10^{-5}	99·8	99·97

The 0·1 m accuracy requirement for automatic docking precludes the use of point positioning and DGNSS; only RTK is sufficiently accurate to meet this. However, due to the difficulty of ambiguity resolution, in the Harwich data collection exercise the availability and continuity requirements were failed over baselines greater than 1 km, which shows the difficulty of achieving these requirements. It is also essential to use dual-frequency data, as the single-frequency positioning had very low availability and continuity. It is in meeting this requirement that the additional frequencies and satellites of modernised GPS and Galileo may show most benefit by increasing the minimum baseline length over which ambiguity resolution is sufficiently reliable. The novel subset ambiguity resolution developed in this thesis also shows great benefit for meeting the requirements. However, it is not likely that the requirements could be achieved reliably over even medium-length baselines even with modernised GNSS and the subset ambiguity resolution algorithm; the ideal situation would be a reference station local to every port.

9.3 Original work

During the course of this project the following original work was produced:

- Collection of information about long-term trends in international shipping and ship design, in order to obtain a greater understanding of the background to the necessity of eNavigation and the IMO requirements;
- Collection of information about the signal structure of modernised GPS and Galileo and the likely improvements to positioning performance obtained from the use of these systems, to allow the benefits of modernised GNSS to meeting the IMO requirements to be assessed;
- Development of GNSS processing software to allow flexible processing of the collected data and development of further algorithms, with work required to combine pre-existing algorithms to achieve the required results, and to deal with real data;
- Development of an algorithm to compute, for a given alert limit, the minimum size of bias required to cause a position error and the probability of detecting this bias, in order allow a comparative analysis of the robustness to bias of code positioning techniques and thereby assess the abilities of different processing techniques to meet the integrity requirements, and the potential benefits of modernised GPS and Galileo;

- Development of an algorithm to determine the relative robustness to bias of phase positioning techniques, for the same purpose as the code robustness algorithm;
- Assessment of the real-world performance of different GPS positioning techniques in scenarios corresponding to different IMO requirements, with an independent truth model and different baseline lengths;
- Assessment of the performance benefits of the new L2C signal, as part of the assessment of the future benefits of modernised GNSS to meeting the IMO requirements;
- Theoretical analysis of the effect of the improvement in positioning performance from modernised GNSS on positioning in a marine environment, and in particular upon the ability of different processing techniques to meet the different IMO requirements;
- Development and testing of a new algorithm for additional ambiguity resolution. This subset ambiguity resolution algorithm allows some ambiguities to be resolved if normal ambiguity resolution fails, and can improve the robustness of ambiguity resolution to biased observations. Testing was performed on the real data collected in a marine environment, and clearly shows the benefits of this algorithm.

9.4 Further work

Further processing could be performed on the collected data. An interesting analysis would be to use the reference station data, with different configurations of reference stations, to determine the performance of network RTK. Atmospheric corrections could then be interpolated across the network, which should increase the accuracy of DGNSS, and in particular improve the ambiguity resolution success rate. This effect would be most significant over longer baselines: good RTK results might be achieved by a network of medium-baseline reference stations. It would also be interesting to study other ways to improve the ambiguity resolution success rate, for example using the conventional partial ambiguity algorithm or the fixed-fail-rate ratio test. GLONASS data were recorded from some reference stations; if these could be processed then the additional observations provided would give some insight into the benefits of the additional Galileo satellites.

Simulated modernised GPS and Galileo data could be used to analyse the effect on the positioning solutions of various combinations of additional frequencies and satellites. Ideally these signals would be generated and combined with the real data, but if this proved to be unfeasible then all the signals, including GPS, could be generated with the appropriate noise level, using the observed noise level as the baseline.

If the experiment were performed again, then it would be beneficial to use a more accurate technique to provide the truth model. The difficulty with this is that GPS provides a very easily-obtained and accurate solution: it is difficult to do better with another system. It is possible that a different type of total station could be obtained that would track more smoothly. Otherwise a laser tracker or a photogrammetric technique could be used, particularly over shorter ranges. It would be useful to have reference stations further away from the rover to analyse the effect of spatial decorrelation on DGNSS accuracy; EGNOS corrections could also be used.

In order to allow a more powerful analysis of whether the different positioning techniques can meet the IMO requirements, rather than where they cannot, a greater quantity of data from different times and locations would need to be collected. In particular, it would be desirable to collect data at times of higher ionospheric activity and in situations of higher multipath at the reference receiver.

Further work could be done to develop the partial ambiguity resolution algorithm. In particular, the computational efficiency may need to be improved before it is practically useful. The inefficiency stems from the fact that ambiguity resolution must be performed on many different subsets. Currently the LAMBDA algorithm is used, which is relatively computationally intensive. Efficiency might be increased by using the bootstrapping technique to fix the ambiguities: this would still require the initial part of the LAMBDA algorithm, the determination of the decorrelating Z-transformation, but would remove the need for the search process. The decorrelation adjustment chosen depends on the satellite geometry and measurement precision, which do not change rapidly; it might be possible to further improve computational efficiency by storing the transformation for each subset between epochs.

Appendix A

GNSS Signal Structure

Table A.1: GPS signal characteristics

Carrier	L1			L2		L5	
Central Frequency (MHz)	1575·42			1227·60		1176·45	
ARNS	Yes			No		Yes	
Null-to-Null Bandwidth (MHz)	2·046			20·46		20·46	
Min. Received Power (dBW)	−158·5	−157		−160		−154	
Multiplex Scheme	CASM			Chip-by-Chip		QPSK	
Signal Component	C/A	L1C-I	L1C-Q	L2C CM	L2C CL	L5-I	L5-Q
Power Distribution	100%	25%	75%	50%	50%	50%	50%
Modulation Type	BPSK(1)	TMBOC	TMBOC	BPSK(1)	BPSK(1)	BPSK(10)	BPSK(10)
Primary Code Length (chips)	1,023	10,230	10,230	10,230	767,250	10,230	10,230
Code Rate (Mcps)	1·023	1·023	1·023	0·5115	0·5115	10·23	10·23
Secondary Code Length (chips)	-	-	1800	-	-	10	20
Data Message Type	NAV	CNAV-2	N/A (Pilot)	CNAV	N/A (Pilot)	CNAV	N/A (Pilot)
Data Message Rate (bps)	50	50	-	25	-	50	-
Error Correction	-	FEC($1/2$); Interleave	-	FEC($1/2$)	-	FEC($1/2$)	-
Symbol Rate (sps)	-	100	-	50	-	100	-

Table A.2: Summary of Galileo signal characteristics

Carrier	E1		E5a		E5b	
Central Frequency (MHz)	1575·42		1176·45		1207·14	
ARNS	Yes		Yes		Yes	
Null-to-Null Bandwidth (MHz)	2·046		51 combined (20·46 individually)			
Min. Received Power (dBW)	-157		-155		-155	
Multiplex Scheme	CASM		QPSK		QPSK	
Signal Component	E1b	E1c	E5a-I	E5a-Q	E5b-I	E5b-Q
Power Distribution	50%	50%	50%	50%	50%	50%
Modulation Type	CBOC	CBOC	AltBOC(15,10) (BPSK(10) individually)			
Primary Code Length (chips)	4092	4092	10,230	10,230	10,230	10,230
Code Rate (Mcps)	1·023	1·023	10·230	10·230	10·230	10·230
Secondary Code Length (chips)	-	25	20	100	4	100
Data Message Type	I/NAV	N/A (Pilot)	F/NAV	N/A (Pilot)	I/NAV	N/A (Pilot)
Data Message Rate (bps)	125	-	25	-	125	-
Error Correction	FEC(1/2); Interleave	-	FEC(1/2); Interleave	-	FEC(1/2); Interleave	-
Symbol Rate (sps)	250	-	50	-	250	-

Appendix B

Data collection

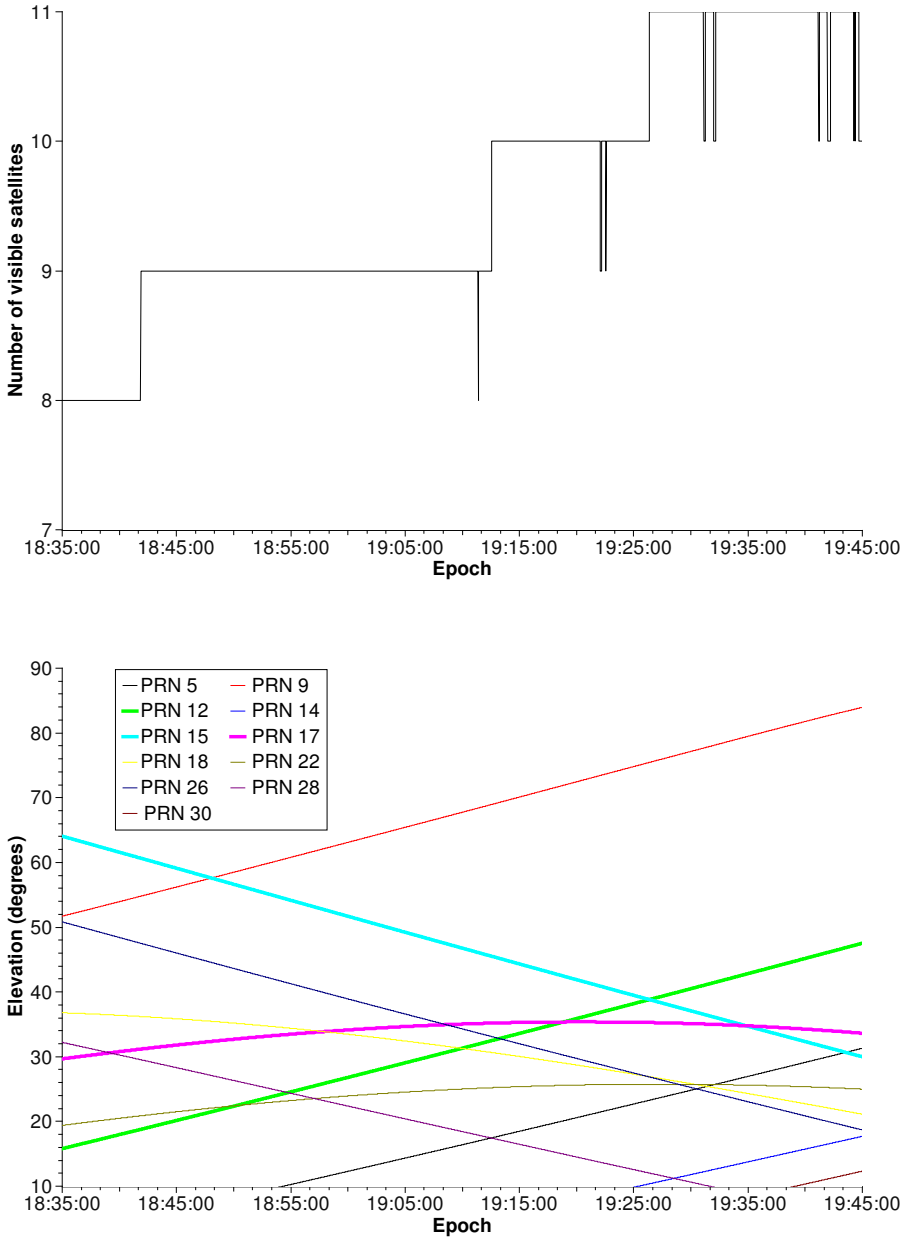


Figure B.1: Visibility of GPS satellites during the Harwich data collection (block IIR-M in bold)

Table B.1: Reference station details for the Harwich data collection

Name	Description	Type	Distance (km)
SHIP	GPS + L2C	Topcon GR-3	-
NAVY	Total station	Topcon GPT-9000A	0·5
JETTY	Total station	Topcon GPT-9000A	0·5
BASE	GPS + L2C	Topcon GR-3	1
WIX	GPS + L2C	Topcon GR-3	15
ALDB	GPS	Leica SR530	31
SHOE	GPS + L2C	Trimble NETR5	60
ATTL	GPS	Leica SR530	66
GORE	GPS	Leica SR530	76
MAID	GPS+ L2C	Trimble NETR5	92
STRA	GPS+ L2C	Trimble NETR5	100
STEV	GPS	Leica SR530	102
NEOT	GPS	Leica SR530	111

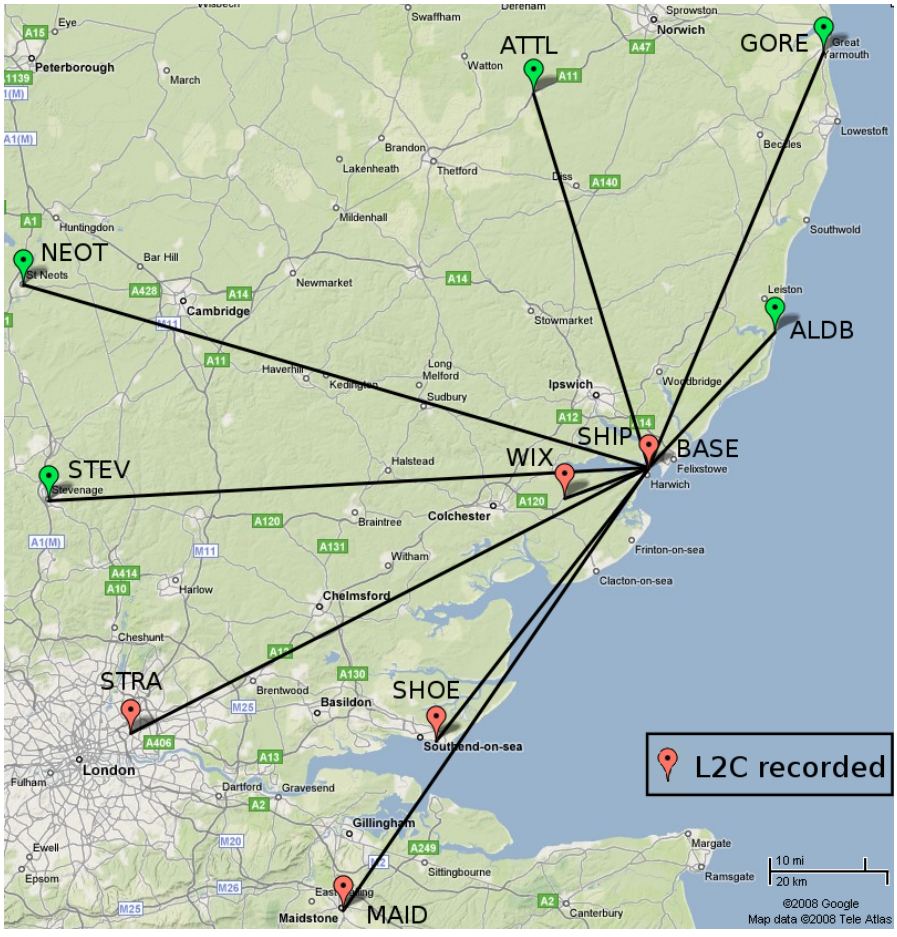


Figure B.2: Map of reference stations used in the Harwich data collection

Bibliography

ABOUSALEM, M.A., MCLELLAN, J.F. & KRAKIWSKY, E.J. (1994). A new technique for quality control in GPS kinematic positioning. In *IEEE Position Location and Navigation Symposium, 1994*, 621–628.

AVILA-RODRÍGUEZ, J.A., IRSIGLER, M., HEIN, G.W. & PANY, T. (2004). Combined Galileo/GPS frequency and signal performance analysis. In *Proc. ION GNSS 2004*, 362–649, Long Beach.

AVILA-RODRÍGUEZ, J.A., HEIN, G., WALLNER, S., ISSLER, J., RIES, L., LESTARQUIT, L., DE LATOUR, A., GODET, J., BASTIDE, F. & PRATT, T. (2008). The MBOC modulation: the final touch to the Galileo frequency and signal plan. *Navigation*, **55**, 15.

BAARDA, W. (1968). *A testing procedure for use in geodetic networks*, vol. 2. Delft, Kanaalweg 4, Rijkscommissie voor Geodesie, 1968.

BETZ, J., BLANCO, M., CAHN, C., DAFESH, P., HEGARTY, C., HUDNUT, K., KASEMSRI, V., KEEGAN, R., KOVACH, K., LENAHAN, L., MA, H., RUSHANAN, J., SKLAR, D., STANSELL, T., WANG, C. & YI, S. (2006a). Description of the L1C signal. In *ION GNSS 2006*, 2080–2091, Fort Worth.

BETZ, J.W., CAHN, C.R., DAFESH, P.A., HEGARTY, C.J., HUDNUT, K.W., JONES, C.A.J., KEEGAN, R., KOVACH, K., LENAHAN, L.L.S., RUSHANAN, J.J., STANSELL, T.A., WANG, C.C. & YI, S.K. (2006b). L1C signal design options, the MITRE Corporation.

BOEHM, J., NIELL, A., TREGONING, P. & SCHUH, H. (2006). Global mapping function (GMF): A new empirical mapping function based on numerical weather model data. *Geophys. Res. Lett*, **33**.

BROWN, R. (1996). Receiver autonomous integrity monitoring. *Global Positioning System: Theory and applications.*, **2**, 143–165.

BUXTON, I. (2004). Trends in ship sizes - will hulls always grow larger? *Naval Architect*, **April**, 22–26.

CAO, W., O'KEEFE, K. & CANNON, M.E. (2007). Partial ambiguity fixing within multiple frequencies and systems. In *ION GNSS 2007*, Fort Worth.

CORBETT, S.J. & CROSS, P.A. (1995). GPS single epoch ambiguity resolution. *Survey Review*, **33**, 149–160.

COUNSELMAN, C.C. & GOUREVITCH, S.A. (1981). Miniature interferometer terminals for earth surveying: Ambiguity and multipath with global positioning system. *Geoscience and Remote Sensing, IEEE Transactions on*, **GE-19**, 244–252.

CROSS, P.A. (1994). *Advanced Least Squares applied to position fixing*. Working Papers, University of East London.

CULLINANE, K. & KHANNA, M. (2000). Economies of scale in large container ships: optimal size and geographical implications. *Journal of Transport Geography*, **8**, 181–195.

DE HALPERT, J., BASKER, S. & PARKINS, A. (2006). Future trends in shipping and its demand on aids to navigation. In *Aids to Navigation in the Digital World, 16th conference of the International Association of Marine Aids to Navigation and Lighthouse Authorities*, Shanghai.

DE JONGE, P. & TIBERIUS, C. (1996). *The LAMBDA method for Integer Ambiguity Estimation: Implementation Aspects*. Delft Geodetic Computing Centre.

DE JONGE, P., TEUNISSEN, P., JONKMAN, N. & JOOSTEN, P. (2000). The distributional dependence of the range on triple frequency GPS ambiguity resolution. *2000: Navigating into the New Millennium*, 605–612.

DEPARTMENT FOR TRANSPORT (2006). *Focus on Ports*. Palgrave Macmillan.

DONG, S., WU, H., LI, X., GUO, S. & YANG, Q. (2008). The Compass and its time reference system. *Metrologia*, **45**, S47–S50.

EISSFELLER, B., TIBERIUS, C., PANY, T., BIBERGER, R. & HEINRICHS, G. (2001). Real-time kinematic in the light of GPS modernization and Galileo. In *Proc. of ION GPS 2001*, 650–662, Salt Lake City.

ENGE, P. (2003). GPS modernization: Capabilities of the new civil signals. In *Australian International Aerospace Congress*, Brisbane.

FONTANA, R., CHEUNG, W. & STANSELL, T. (2001a). The modernized L2 civil signal. *GPS World*, **12**, 28–34.

FONTANA, R.D., CHEUNG, W., NOVAK, P.M. & STANSELL, T.A. (2001b). The new L2 civil signal. In *ION GPS 2001*, 617–631, Salt Lake City.

GAO, G., CHEN, A., LO, S., DE LORENZO, D., WALTER, T. & ENGE, P. (2008). Compass-M1 broadcast codes and their application to acquisition and tracking. In *ION NTM 2008*.

GJU (2006). Galileo open service signal in space interface control document (OS SIS ICD).

GPS WORLD (2008a). Galileo, Compass on collision course.

GPS WORLD (2008b). GNSS interoperability in China: overrated? *GPS World*.

GRELIER, T., GHION, A., DANTEPAL, J., RIES, L., DELATOUR, A., ISSLER, J.L., AVILA-RODRIGUEZ, J., WALLNER, S. & HEIN, G. (2007). Compass signal structure and first measurements. In *ION GNSS 2007*, 3015–3024, Fort Worth.

HARRIS, R. (1997). Direct resolution of carrier-phase ambiguity by "bridging the wavelength gap". Tech. rep., ESATEC, tST/60107/RAH/Word.

HASSIBI, A. & BOYD, S. (1998). Integer parameter estimation in linear models with applications to GPS. *IEEE Transactions on Signal Processing*, **46**, 2938–2952.

HEIN, G., AVILA-RODRIGUEZ, J., WALLNER, S., PRATT, A., OWEN, J., ISSLER, J., BETZ, J., HEGARTY, C., LENAHAN, L., RUSHANAN, J. *et al.* (2006). MBOC: The new optimized spreading modulation recommended for Galileo L1 OS and GPS L1C. In *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, 883–892.

HEIN, G.W., GODET, J., ISSLER, J.L., MARTIN, J.C., LUCAS-RODRIGUEZ, R. & PRATT, T. (2001). The Galileo frequency structure and signal design. In *ION GPS 2001*, 1273–1282, Salt Lake City.

HEIN, G.W., GODET, J., ISSLER, J.L., MARTIN, J.C., ERHARD, P., LUCAS-RODRIGUEZ, R. & PRATT, T. (2002). Status of Galileo frequency and signal design. In *ION GPS 2002*, 266–277, Portland.

HENDERSON, C., SWANN, J., SAGE, A. & MITCHELL, S. (2009). Galileo downstream applications study. Report for the UK DfT.

HOUSE OF COMMONS TRANSPORT COMMITTEE (2007). The ports industry in England and Wales.

HUDNUT, K.W. & TITUS, B. (2004). GPS L1 civil signal modernisation (L1C), report to the interagency GPS executive board.

IALA (2001). Aids to navigation guide.

IALA (2008). eNavigation: Frequently asked questions, version 1.1.

IMO (1989). World-wide radionavigation system, resolution A.666(16).

IMO (1995). World-wide radionavigation system, resolution A.815(19).

IMO (1997). Maritime policy for a future global navigation satellite system (GNSS), resolution A.860(20).

IMO (2001). Revised maritime policy and requirements for a future global navigation satellite system (GNSS), resolution A.915(22).

IMO (2003). World-wide radionavigation system, resolution A.953 (23).

IMO (2005). International shipping - carrier of world trade, background paper, World Maritime Day 2005.

IMO (2008). Development of an e-Navigation strategy. COMSAR 12/11.

INOUE, K. (2000). Evaluation method of ship-handling difficulty for navigation in restricted and congested waterways. *The Journal of Navigation*, **53**, 167–180.

INSIDE GNSS (2008a). China to launch 10 Compass satellites over next two years.

INSIDE GNSS (2008b). International system providers meeting (ICG-3) reflects GNSS's competing interests, cooperative objectives.

INSIDE GNSS (2009). Russia launches three more GLONASS-M space vehicles.

JOOSTEN, P., TEUNISSEN, P. & JONKMAN, N. (1999). GNSS three carrier phase ambiguity resolution using the LAMBDA method. In *ION GNSS 1999*, vol. 99, 5–8.

JULIEN, O., ALVES, P., CANNON, M. & LACHAPELLE, G. (2004). Improved triple-frequency GPS/Galileo carrier phase ambiguity resolution using a stochastic ionosphere modeling. In *Proceedings of ION, NTM*, 01–26.

JULIEN, O., MACABIAU, C., RODRIGUEZ, J.A.A., WALLNER, S., PAONNI, M., HEIN, G.W., ISSLER, J.L. & RIES, L. (2007). On potential CBOC/TMBOC common receiver architectures. In *ION GNSS 2007*.

JUNG, J., ENGE, P. & PERVAN, B. (2000). Optimization of cascade integer resolution with three civil GPS frequencies. In *ION GPS 2000*, vol. 2200.

KAPLAN, E.D. & HEGARTY, C.J. (2005). *Understanding GPS: Principles and Applications*. Artech House, second edition edn.

KIM, D. & LANGLEY, R. (1999). An optimized least-squares technique for improving ambiguity resolution and computational efficiency. In *ION GPS 1999*, vol. 1588.

KIM, D. & LANGLEY, R. (2000). GPS ambiguity resolution and validation: Methodologies, trends and issues. In *7th GNSS Workshop-International Symposium on GPS/GNSS, Seoul, Korea, Nov*.

KRUEGER, E., SCHUELER, T., HEIN, G., MARTELLUCCI, A. & BLARZINO, G. (2004). Galileo tropospheric correction approaches developed within GSTB-V1. In *Proceedings of ION GNSS 2004*, 17–19.

LAMB, W.G.P. & HUNT, J.M. (2000). Multiple encounter avoidance manoeuvres. *The Journal of Navigation*, **53**, 181–186.

LANDAU, H., VOLLATH, U. & CHEN, X. (2004). Benefits of modernised GPS/Galileo to RTK positioning. In *GNSS 2004*, Sydney.

LAU, L. & CROSS, P. (2003). Impact of GPS modernization on precise carrier phase-based positioning in the presence of multipath. In *ION GPS 2003*, 2163–2172, Portland.

LAU, L. & CROSS, P. (2005). Use of signal-to-noise ratios for real-time GNSS phase multipath mitigation. In *NAV 05*, London.

LEONARD, A., BLOMENHOFER, H. & IZQUIERDO, I. (2002). GPS and Galileo interoperability and synergies. In *Proc. of ION GPS 2002*, 330–341, Portland.

MADER, G. (1992). Rapid static and kinematic global positioning system solutions using the ambiguity function technique. *Journal of Geophysical Research*, **97**, 3271–3283.

MAN DIESEL (2004). Propulsion trends in container vessels, copenhagen, Denmark.

MARTIN-NEIRA, M., TOLEDO, M. & PELAEZ, A. (1995). The null space method for GPS integer ambiguity resolution. In *Proceedings of DSNS'95*, 24–28.

MATTOS, P. (2006). The implications of simultaneous processing of the Galileo L1b/c signals. In *ION GNSS 2006*, 1861–1865, Fort Worth.

MERINO, M.M.R., ALARCON, A.J.G., VILLARES, I.J. & MONSECO, E.H. (2001). An integrated GNSS concept, Galileo & GPS, benefits in terms of accuracy, integrity, availability and continuity. In *ION GPS 2001*, 2114–2124, Salt Lake City.

MILBERT, D. (2005). Influence of pseudorange accuracy on phase ambiguity resolution in various GPS modernisation scenarios. *Navigation*, **52(1)**, 29–38.

MOUDRAK, A., KONOVALTSEV, A., FURTHNER, J., HAMMESFAHR, J., BAUCH, A., DEFRAIGNE, P. & BEDRICH, S. (2004). Timing aspects of GPS-Galileo interoperability: challenges and solutions. In *Proceedings of the 2004 Precise Time and Time Interval Meeting (PTTI)*, 279–292.

OCHIENG, W., SHERIDAN, K., SAUER, K., HAN, X., CROSS, P., LANNELONGUE, S., AMMOUR, N. & PETIT, K. (2002). An assessment of the RAIM performance of a combined Galileo/GPS navigation system using the marginally detectable errors (MDE) algorithm. *GPS Solutions*, **5**, 42–51.

ODIJK, D., TEUISSEN, P. & TIBERIUS, C. (2002). Triple-frequency ionosphere-free phase combinations for ambiguity resolution. In *Proceedings of ENC 2002*.

O'KEEFE, K. (2001). Availability and reliability advantages of GPS/Galileo integration. In *ION GPS 2001*, 2096–2104, Salt Lake City.

O'KEEFE, K., PETOVELLO, M., LACHAPELLE, G. & CANNON, M. (2006). Assessing probability of correct ambiguity resolution in the presence of time-correlated errors. *Navigation*, **53**, 269–282.

PARKINSON, B.W. & SPILKER, J.J. (1996). *Global Positioning System: Theory and Applications*, vol. I. American Institute of Aeronautics and Astronautics.

PERVAN, B., LAWRENCE, D. & PARKINSON, B. (1998). Autonomous fault detection and removal using GPS carrier phase. *Aerospace and Electronic Systems, IEEE Transactions on*, **34**, 897–906.

PETOVELLO, M., O'KEEFE, K., LACHAPELLE, G. & CANNON, M. (2005). Quantifying ambiguity resolution performance in the presence of time-correlated measurement errors using geometric-based techniques. In *Proceedings of ION AM*, 1073–1085.

SALZMANN, M. (1995). Real-time adaptation for model errors in dynamic systems. *Journal of Geodesy*, **69**, 81–91.

SIMSKY, A. (2006). Three's the charm: triple-frequency combinations in future GNSS. *Inside GNSS*, **1**, 38–41.

SIMSKY, A., SLEEWAEGEN, J.M., HOLLREISER, M. & CRISC, M. (2006a). Performance assessment of Galileo ranging signals transmitted by GSTB-V2 satellites. In *ION GNSS 2006*, 1547–1559, Fort Worth.

SIMSKY, A., SLEEWAEGEN, J.M. & NEMRY, P. (2006b). Early performance results for new Galileo and GPS signals-in-space. In *ENC GNSS 2006*, Manchester.

SIMSKY, A., MERTENS, D., SLEEWAEGEN, J.M., DE WILDE, W., HOLLREISER, M. & CRISCI, M. (2008). Mboc vs boc(1,1). *Inside GNSS*, 36–40.

SOMIESKI, A., BURGI, C. & FAVEY, E. (2007). Evaluation and comparison of different methods of ionospheric delay mitigation for future Galileo mass market receivers. In *ION NTM 2007*, Fort Worth.

SPILKER, J.J. (1980). GPS signal structure and performance characteristics. In *Global Positioning System: theory and applications*, vol. I, 29–54, The Institute of Navigation.

SPILKER, J.J. & VAN DIERENDONCK, A. (1999). Proposed new civil GPS signal at 1176.45 MHz. In *ION GPS 1999*, 1717–1725, Nashville.

STOPFORD, M. (2002). Is the drive for ever bigger containerships irresistible? In *Lloyds List Shipping Forecasting Conference*.

TEUNISSEN, P. (1990). Quality control in integrated navigation systems. *Aerospace and Electronic Systems Magazine, IEEE*, **5**, 35–41.

TEUNISSEN, P. (1993). Least-squares estimation of the integer GPS ambiguities. In *Section IV Theory and Methodology, IAG General Meeting, Beijing, China*.

TEUNISSEN, P. (1995). The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *Journal of Geodesy*, **70**, 65–82.

TEUNISSEN, P. (1998a). *GPS for Geodesy*, chap. 7, 271–318. Springer, 2nd edn.

TEUNISSEN, P. (1998b). Minimal detectable biases of GPS data. *Journal of Geodesy*, **72**, 236–244.

TEUNISSEN, P. (1998c). On the integer normal distribution of the GPS ambiguities. *Artificial Satellites*, **33**, 49–64.

TEUNISSEN, P. (1998d). Success probability of integer GPS ambiguity rounding and bootstrapping. *Journal of Geodesy*, **72**, 606–612.

TEUNISSEN, P. (1999). An optimality property of the integer least-squares estimator. *Journal of Geodesy*, **73**, 587–593.

TEUNISSEN, P. (2000). ADOP based upperbounds for the bootstrapped and the least-squares ambiguity success rates. *Artificial Satellites*, **35**, 171–179.

TEUNISSEN, P. (2001a). GNSS ambiguity bootstrapping: Theory and applications. In *Proceedings Int. Symp. on Kinematic Systems in Geodesy, Geomatics and Navigation*, 5–8.

TEUNISSEN, P. (2001b). Integer estimation in the presence of biases. *Journal of Geodesy*, **75**, 399–407.

TEUNISSEN, P. (2001c). Statistical GNSS carrier phase ambiguity resolution: a review. *Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing, 2001.*, 4–12.

TEUNISSEN, P. (2002). Galileo and ambiguity resolution. In *There is more than geometry, a collection of articles in honour of Prof. Theo Bogaerts*.

TEUNISSEN, P. (2003a). Integer aperture GNSS ambiguity resolution. *Artificial Satellites*, **38**, 79–88.

TEUNISSEN, P. (2003b). Theory of integer equivariant estimation with application to GNSS. *Journal of Geodesy*, **77**, 402–410.

TEUNISSEN, P. (2003c). Towards a unified theory of GNSS ambiguity resolution. *Journal of Global Positioning Systems*, **2**, 1–12.

TEUNISSEN, P. & VERHAGEN, S. (2004). On the foundation of the popular ratio test for GNSS ambiguity resolution. In *ION GNSS 2004*, Long Beach.

TEUNISSEN, P. & VERHAGEN, S. (2007). GNSS carrier phase ambiguity resolution: Challenges and open problems. In *Proceedings of the Scientific meetings of the IAG General Assembly 2007*, Perugia, Italy.

TEUNISSEN, P., JOOSTEN, P. & TIBERIUS, C. (1999). Geometry-free ambiguity success rates in case of partial fixing. In *Proceeding of National Technical Meeting & 19th Biennal Guidance Test Symposium, ION 1999*, 201–207.

TEUNISSEN, P., JOOSTEN, P. & TIBERIUS, C. (2000a). Bias robustness of GPS ambiguity resolution. In *ION GPS 2000*, 104–111, Salt Lake City.

TEUNISSEN, P., N.F. JONKMAN, P.J. & TIBERIUS, C. (2000b). Long baseline 3 frequency differential GNSS. In *Proc. of the IEEE Position Location and Navigation Symposium*, 7–14, San Diego.

TEUNISSEN, P., ODIJK, D. & JONG, C. (2000c). Ambiguity dilution of precision : an additional tool for GPS quality control. In *LGR-Series*, 21, 261–270, Delft Geodetic Computing Centre, Delft.

TEUNISSEN, P., JOOSTEN, P. & TIBERIUS, C. (2002). A comparison of TCAR, CIR and LAMBDA GNSS ambiguity resolution. In *ION GPS 2002: 15th International Technical Meeting of the Satellite Division of The Institute of Navigation*.

THE ECONOMIST (2009). Shipbuilding: Sink or swim. **March 26th**.

TSUI, J. (2000). *Fundamentals of global positioning system receivers*. Wiley New York.

UENO, M. (1999). *Use of GPS for a berthing guidance system*. Ph.D. thesis, Laval University.

UN MSC (2004). Safety of navigation: Update on the galileo program and IMO related activities.

UN MSC (2009). Work programme: world-wide radionavigation system (WWRNS).

UNCTAD (2008). Review of maritime transport, 2008, united Nations, New York and Geneva.

VERHAGEN, S. (2003). On the approximation of the integer least-sqaures success rate: which lower or upper bound to use? *Journal of Global Positioning Systems*, **2**, 117–124.

VERHAGEN, S. (2005a). *The GNSS integer ambiguities: estimation and validation*. Ph.D. thesis, Delft University of Technology.

VERHAGEN, S. (2005b). How will the new frequencies in GPS and Galileo affect carrier phase ambiguity resolution? *Inside GNSS*.

VERHAGEN, S. (2007). Improved performance of multi-carrier ambiguity resolution based on the LAMBDA method. In *Proceedings of the 3rd ESA Workshop on Satellite Navigation User Equipment Technologies, NAVITEC 2006*, Noordwijk, The Netherlands.

VERHAGEN, S. & JOOSTEN, P. (2004). Analysis of integer ambiguity resolution algorithms. In *ENC 2004*.

WALLNER, S., HEIN, G., PANY, T., AVILA-RODRIGUEZ, J. & PÓSFAY, A. (2005). Interference computations between GPS and Galileo. In *ION GNSS 2005*, 13–16.

WALLNER, S., HEIN, G.W. & AVILA-RODRIGUEZ, J.A. (2007). Interference computations between several GNSS systems. In *Proceedings of ESA Navitec 2006*.

WANG, Z., WU, Y., ZHANG, K. & MENG, Y. (2005). Triple-frequency method for high-order ionospheric refractive error modelling in GPS modernization. *Journal of Global Positioning Systems*, **4**, 291–295.

WEILL, L.R. (2002). Multipath mitigation using modernized GPS signals: How good can it get? In *ION GPS 2002*, 493–505, Portland.

YANG, C., HEGARTY, C. & TRAN, M. (2004). Acquisition of the GPS L5 signal using coherent combining of I5 and Q5. In *ION GNSS 2004*, 2184–2195, Long Beach.