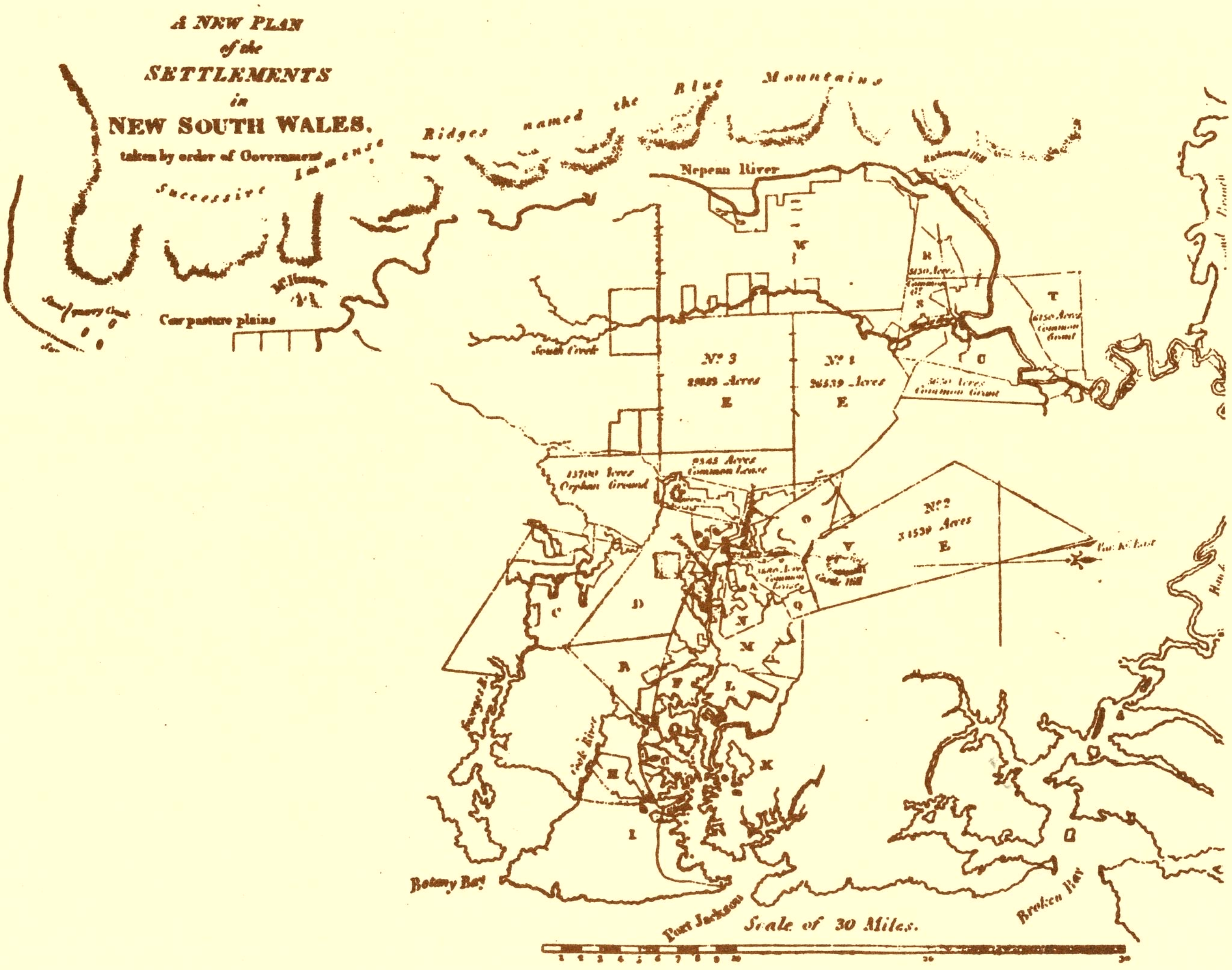
DEVELOPMENT OF A NAVIGATIONAL SYSTEM UTILIZING THE GLOBAL POSITIONING SYSTEM IN A REAL TIME, DIFFERENTIAL MODE

JOHN MICHAEL NOLAN



UNISURV S-40, 1990

Reports from

SCHOOL OF SURVEYING

THE UNIVERSITY OF NEW SOUTH WALES KENSINGTON, NSW AUSTRALIA 2033

UNISURV REPORT S-40, 1990

DEVELOPMENT OF A NAVIGATIONAL SYSTEM UTILIZING THE GLOBAL POSITIONING SYSTEM IN A REAL TIME, DIFFERENTIAL MODE

John Michael Nolan

Received: October, 1990

SCHOOL OF SURVEYING
UNIVERSITY OF NEW SOUTH WALES
P.O. BOX 1
KENSINGTON N.S.W. 2033
AUSTRALIA

National Library of Australia

Card No. and ISBN 0 85839 058 2

Editor's Note

The Unisurv-S series has been, since its inception in the mid-1960's, the vehicle by which work carried out within the School of Surveying, University of New South Wales, for research degrees and major projects has been published.

In recent years there has been an increase in research activity in other Australian Schools and Departments. In order to assist in the dissemination of this research, this School has decided that, where it is of sufficient quality and relevance, such research will be published within the Unisurv Series. This report, based upon research by John Nolan at Curtin University, Perth, W.A., is the first of these 'extra - UNSW' reports.

A.H.W. Kearsley
Publications Manager
School of Surveying
University of New South Wales
November, 1990

Abstract

This thesis documents the development of a navigational system that uses the Global Positioning System (GPS) to provide the user with knowledge of his position. The principal aim of this system is to provide metre level positioning, in real time, with a high degree of solution integrity.

To guarantee the integrity of the solution, steps have been taken in both algorithm design and implementation. Also, the user is given a variety of options with regard to monitoring the quality of the solution. A conservative estimate of the position accuracy is provided with every solution.

The system described uses two or more GPS receivers in a differential mode. That is, one receiver is sited on a known location and transmits corrections to one or more mobile receivers. This differential concept is essential if an improvement in the obtainable single receiver accuracy (between 20 and 100 metres) is required.

Corrections are passed by a suitable communication link. Most of the discussion in this paper focuses on the use of HF radio transmissions for this link.

The development of the system begins with a precise definition of the aims of the system and its proposed applications. Discussion is given to the hardware and software requirements and how these requirements are subsequently catered for.

A broad overview of the GPS system is offered, followed by a closer inspection of the characteristics of the raw data output by the various GPS receivers. The algorithms used in this system are then described in detail. This description is presented in the logical order of the development of the system.

Finally the results of field testing carried out to validate the system are presented.

Table of Contents

List of Figures

List of Tables

Special Note

All development described in this paper was sponsored by Associated Surveys International. The subject of this thesis specifically pertains to the design and implementation of a navigation system. This includes research into algorithms, implementation and testing. All source code remains the property of Associated Surveys International.

Examples of source code relevant to the explanation of the topics covered in this paper are contained both within the document and in Appendix A. The source code in this appendix is not for public release.

All research described in this paper was carried out solely by the author. All source code was written by the author with the exception of the code to interface the modem hardware. To expedite development, a private developer was contracted to implement this module.

Acknowledgements

My sincere appreciation is given to my employers, Associated Surveys International for providing me with the opportunity to develop this system. I have been fortunate during my time with this company to have had access to the valuable resources of this organisation. While this includes the latest in technical hardware and software, the most valuable resource has been the experience and insight of this large body of professional personnel.

Finally, I would like to thank my supervisor Dr Lawry White for his encouragement and help over the last few years.

1. OVERVIEW

1.1 Introduction

The [1]NAVSTAR Global Positioning System (GPS) is an American military based system designed to provide real time navigation anywhere in the world. It is an all weather system that when fully deployed will provide position information 24 hours per day. It accomplishes this by measuring distances to a number of satellites orbiting the earth. The proposed accuracy of the GPS system is 10m [2]($2\times d_{rms}$) for military users and 100m ($2\times d_{rms}$) for civil users. This level of accuracy is obtained using one suitable GPS receiver in a stand alone mode.

The first GPS satellite was launched in 1978 and since the early 1980's there have been sufficient satellites in orbit to test and evaluate the GPS system. During this time many methods have been developed to adapt the GPS system to suit the needs of different applications. Using GPS it is possible to measure distances of several hundred kilometres to an accuracy of less than a few parts per million. However, obtaining such results requires long observation periods where two receivers are stationary. Results are not available until several hours after the observation period.

One of the strengths of the GPS system is its versatility. Navigational position to better than one hundred metres can be determined with consummate ease. Alternatively, with different restrictions and operational procedures, geodetic accuracy (millimetre) solutions can be obtained.

The system described in this paper can be viewed as an in between case. The GPS is exploited to obtain metre level accuracy without the restrictions involved with

---

[1] Navigation System with Time And Ranging

[2] $2\times d_{rms} = 2\sqrt{[\sigma_{EE} + \sigma_{NN}]}$ (Ananda 1988)

geodetic accuracy measurements. This hybrid accuracy solution is obtained by using the two fundamental measurements provided by the GPS system. These are the carrier phase, used to obtain geodetic solutions and the code phase used to obtain the navigation solution.

The principal components of this real time navigation system presented are shown in figures 1 and 2. They consist of a base station installation and the rover or mobile installation. The hardware/software requirement at each station is identical. At the base station site corrections to the GPS satellite ranges are computed and passed via a communication link to the mobile site. At the mobile site these corrections are applied to the data being recorded by the mobile receiver. The corrected solution is then calculated and displayed to the user in various forms.

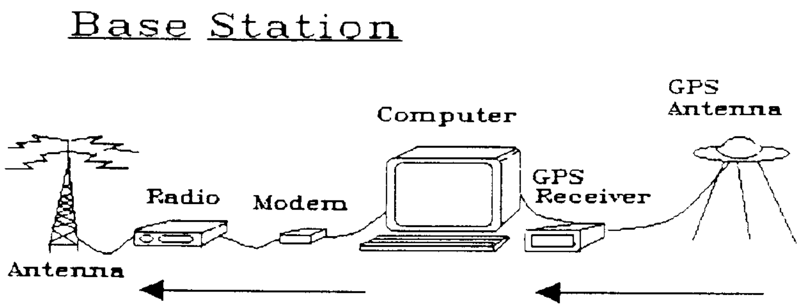


Figure 1: System Configuration (Base Station)

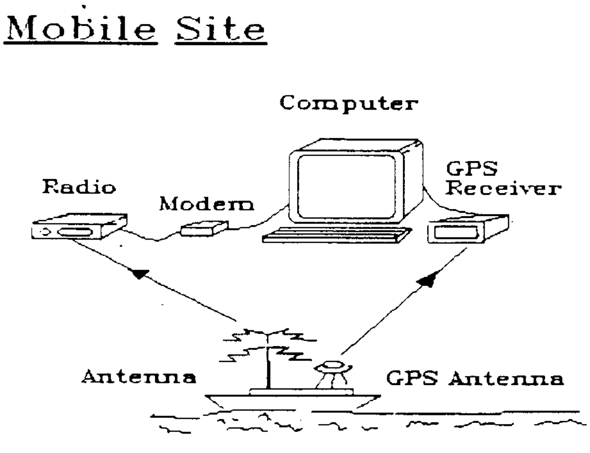


Figure 2: System Configuration (Rover Station)

This differential technique, whereby two GPS receivers are used together, is the key to obtaining increased accuracy results. However, while differencing certainly improves the accuracy of the system, many factors must be carefully controlled if the system integrity is to be maintained.

The chief motivation for the development of this system came from the commercial need for real time positioning in off-shore projects. Initially the requirement was to establish oil rig locations to an accuracy of approximately 5.0 metres [3](s.e.p.). Positioning errors in this type of work are very expensive.

Due to the high costs involved in the event of errors, the integrity of the solution is of absolute importance. If a solution is inferior, for any reason, then the operator must be warned. During my 1988 work commitment with Associated Surveys International, several commercial ventures were undertaken using GPS as the secondary system. During this period much was learned concerning the inherent problems caused by using GPS without strict quality control procedures. (Castalanelli et al 1988). The navigation system described in this paper is a direct result of these experiences.

---

[3] Spherical Error Probability. (Ananda 1988)

After the 1988 campaign it was proposed to develop a system that could be relied upon to fulfil off-shore navigation committments. It was also desirable that the system be adaptable to any other form of navigation requiring metre level accuracy.

The system will be used by a variety of operators in various conditions. Because of this, it needs to be simple to use and at the same time have the appropriate complexity to adapt to the needs of different applications.

The anticipated areas of application for this type of system are quite broad. Hence, if implemented well, the system could possibly have a life expectancy of between three and five years. Consequently, the software needs to be designed carefully to ensure that the system can evolve to adapt to future demands.

1.2 Aims

The following sections set out in more detail the goals defined for the development of the system.

1.2.1 Accuracy

The system is aimed at providing a horizontal accuracy of better than one metre ±10ppm ($2 \times d_{rms}$). However, it is natural that the higher the accuracy the better. In general, due to the nature of the GPS system, the horizontal component of position is typically twice the accuracy of the vertical component. Hence, to obtain an accuracy of better than one metre in the vertical necessitates obtaining an accuracy of better than 0.5m in the horizontal. This must be considered when specifying accuracies for this type of system.

1.2.2 Update Rate

The update rate of the system should be at least one second or better.

1.2.3 Integrity

The system should provide the operator with some form of assurance as to the integrity of the system. This can be accomplished by first designing the system so that all possible situations relating to degraded solution accuracy are handled by the software. Secondly, the user should be able to monitor the quality of the resultant solution.

1.2.4 Range

The system should be suitable for use at ranges of several thousand kilometres. There are several commercial positioning applications at this distance off-shore and so this type of range is essential. To facilitate maximum performance over a variety of ranges it is necessary to allow the system to be compatible with a variety of modems and radio types.

1.2.5 Receiver Independence

Considering the cost of this type of development it is unwise to be tied to one particular receiver type. Receiver dependence is one of the key drawbacks of existing GPS software as nearly all of the commercially available software is generated by the receiver manufacturers. If a system is reliant on one receiver type it may become redundant when the manufacturer discontinues a product or alters data formats. Also it may be inconvenient or uneconomical to use one particular receiver for all applications.

1.2.6 Hardware Independence

Where possible the system should not be tied down to a certain hardware configuration. Computer technology has advanced rapidly during the last decade and will continue to do so. However, the chief reason that software becomes redundant is not that the algorithms themselves become outdated. More often, the code is written for a certain type of hardware which has become obsolete or at least less than "state of the art". In an attempt to optimise the use of the present hardware, there is always the temptation to write code specific for a device. This temptation should be avoided. An important aim of the system should be that it can run (with only minor modifications) on many different hardware configurations. This hardware independence applies not only to the central processor, but to all components in the system (modems, radios and GPS receiver).

1.2.7 Multitasking

The ability of this type of system to be multi-tasking is very important. This ability, to run more than one program at the one time, has several advantages. First, it allows the user do other things (for example, reduce readings from other sensors, reduce GPS baselines or write a report) without interfering with the navigation facilities. More importantly, it allows other programs related to the system to be written separately and run concurrently with the navigation system (for instance, a program to reduce position using another system such as syledis). This type of separation of program development has the advantage that each programmer only has to concentrate on a smaller task. Naturally some tasks that require optimum use of shared data cannot be carried out efficiently in this manner.

Finally, in the future computer hardware will have the capacity to run many programs at once. Consequently, it would be a waste of resources to let one program use the entire processor by itself.

1.2.8 Multiple rover

In its simplest form, a differential GPS system requires one base station and one mobile or rover station. The base station determines the corrections to be passed to the rover station in real time. However, if several vessels require navigation in the same area, it is not economic to have one base station for each mobile station used. Also, each system would require a different communication frequency to avoid interference. As we can foresee that this use of multiple rovers is likely, it is prudent to plan the system so that multiple rover stations can be accommodated.

1.2.9 User Friendly

An important requirement of the system is that it should be "user friendly". There are two reasons for this. Firstly, if the program is easy to use, then the operator is less likely to make an error. Hence, making the program very simple and easy to use, increases the integrity of the system. It is futile having quality control monitoring options if the operator finds them difficult to use. Secondly, a system which requires minimal operator training, can be used by a greater number of people. This is a big advantage for any system. Put another way, it is a disadvantage having a good system if only a few highly trained operators can use it.

1.2.10 Customisation

Originally, the system was designed for use in the off-shore oil industry. It soon became apparent that a positioning system capable of providing sub metre accuracy would have other applications, both in land and airborne applications. It is inefficient to develop a system that can cater for one purpose only. However, a system that can cater for every possible task would be extremely difficult to develop and cumbersome to use. A more practical solution is to develop a system in such a manner that it can be customised to meet the needs of special applications. While the basis of the system is performing the same function, the user interface can be varied so the user sees only the information that is required for his particular task. Customisation improves the user friendliness of the system and hence also impacts on the integrity of the system.

1.2.11 Output Compatibility

In the field of off-shore navigation there are several proven commercial navigation systems. These systems take input from many different sensors, such as acoustics systems, syledis ranging systems, gyro headings etc. This input is processed and provided to the navigator in a sophisticated output form. This output includes graphics plots showing such things as runline offsets and position residuals. It is considered to be highly advantageous to be able to output data in a form that can be used by these types of commercially available navigation packages. Hence any system design should incorporate a provision for a flexible output of data that can be tailored to meet the needs of other packages.

In other areas, such as in land base pick-up, it may be required to output position information in the form required by a certain plotting package. While it is envisaged that the navigation system will have its own form of display for standalone use, where possible, full use of other existing packages should be exploited.

1.2.12 Future Growth

To justify spending the time and effort involved with this type of development it should be expected that the system will have a considerable life span. Achieving the set aims of hardware independence and customization naturally increases the life span of a system. However, software must also be written carefully to allow it to be modified without destroying its original structure. Also the system must be capable of being maintained and modified by a variety of people in the future. Hence care must be taken in the original design of the system to ensure that there will be no impediments to the growth of the system.

1.2.13 Economic maintainability

Software systems have to be maintained if they are to remain useful throughout their life. This need for maintenance arises because the software must fit in to a constantly changing environment. Consequently, when designing software every effort should be made to allow this essential function to be as simple as possible. In large systems more time is spent in reading code during modifications than in the initial writing of the program. Hence, in the long run, it is more economic to write well structured and planned code. This naturally requires an additional development overhead.

1.3 Overview

Section 2 (GPS System Overview) contains an introduction to the GPS system and its characteristics. The nature of the GPS signals is examined in detail as it is essential to the development of this type of system.

As can be seen from the proposed aims, there are many factors involved in designing a system. These factors have little to do with navigation but are of extreme importance when accessing the strengths of a system. No matter what algorithms are used and how they are implemented, if the basic system philosophy (hardware and software) does not provide the required amount of versatility the system's life will be short. Hence, before any development was commenced on the structure of the software itself, several key development issues needed to be addressed. Sections 3 (Programming Considerations) and 4 (Hardware Considerations) present the key issues concerning software and hardware respectively.

Section 5 (Algorithm Design) describes the algorithms used in the development of the system. These algorithms are each discussed in the logical order that the system was developed. Section 6 (Program Implementation) covers the methods used to implement these algorithms into computer code. Section 7 (Program Description) then describes the program as it looks to the user. This description of the program user interface offers a practical view of the functionality of the software in its present state.

Several tests were carried out to validate the operation of the software in a controlled situation. These tests are described and results analysed is section 8 (Test Results). Finally, strengths and weaknesses and overall conclusions regarding the system are discussed in section 9 (Summary and Conclusions)

1.4 Applications and requirements

To help understand the required function of the proposed navigation system, it is instructive to look at some of the proposed applications. The following briefly covers selected fields of use.

1.4.1 Rig Positioning

Due to the limited period of the day at present where GPS is operable, it is not possible to use GPS by itself for rig positioning. The high costs involved in the off-shore oil industry make it uneconomic to keep a drilling platform sitting idle while waiting for the GPS satellites to come into view to establish position. Hence until GPS becomes a 24 hour system, rig positioning will be done using other methods such as radio navigation or acoustic systems. However, even while this is the case, GPS navigation can still play an important role.

Radio navigation systems rely on distance measurements to several shore stations to determine the users position. The accuracy of these navigation systems degrades significantly as the distance from the shore stations increases (Chisholm 1987, p22). Due to logistic constraints it is often not possible to obtain ideal geometry for the shore stations. Hence, systematic errors caused by refractive effects are sometimes very difficult to detect.

At distances from 50km or more off-shore, differential GPS has the potential to be the more accurate system. Hence, during the appropriate window, GPS can be used as an independent check of the quality of the radio navigation system. As the range increases, GPS becomes a far more accurate alternative.

Acoustic systems can also be used for rig positioning in remote locations. These systems measure distances to several permanent beacons situated on the sea floor. Analogous to the GPS system, where the position of the GPS satellites must be known, the location of these beacons must also be established. GPS is an ideal positioning system for this.

Using either radio navigation or acoustic systems, GPS can be used as an excellent secondary system. In the future, when there is 24 hour satellite coverage, GPS will no doubt become the primary system. However, due to the high costs involved in the positioning of these expensive installations, there will still be need for a secondary system. This backup system provides an independent check for gross error and quality control.

In both its role as a back-up system and in the future as the primary system, several requirements are made of the GPS system. Of prime importance is the reliability of the position. Errors caused in rig positioning can result in greater losses than the cost of developing the entire system.

As the rigs themselves and the survey vessel are moving slowly, a solution update rate of once per second is adequate for this type of positioning. In all positioning tasks, the higher the accuracy achieved the better. However, in this case, the vessels themselves have a degree of motion associated with them. Hence an accuracy in the order of one or two metres is a more realistic goal.

1.4.2 Airborne Surveys

There are several methods of subsurface exploration that have been devised for airborne platforms. Using sophisticated measuring devices mounted on helicopters or aeroplanes, large surface areas can be examined in a relatively short time. Maps compiled from these surveys highlight possible areas for closer examination.

A key ingredient in this type of survey is in determining the position of the sensor during the course of the survey. This type of operation imposes several constraints. First, the update rate of the positions must be as high as possible. One second updates are barely adequate in this type of application.

Usually the accuracy requirement is such that errors cannot be observed at the map scale used. This may vary between two and ten metres depending on the map scale. Accurate results are not strictly required in real time. However, real time navigation to 100m is usually required to ensure that the aircraft is covering the area without leaving any significant gaps.

With some sensors, radio communications cannot be used during the survey as it may corrupt the recorded data. This affects how and if corrections can be passed to the aircraft in real time. One method that can be used is to record data at the base station and in the aircraft for post processing. The aircraft can navigate using the single receiver position.

When corrections are transmitted to the aircraft in real time, the radio/modem set-up upon the aircraft must be passive. Hence it must not acknowledge the receipt of the corrections.

1.4.3 Borehole setout and pick-up

The chief advantage of having a real time system is observed when we wish to place either ourselves or some object at a pre-specified location. In surveys where a monument exists and the requirement is to locate the coordinate position of this monument, usually real time results are not required.

Most exploration for mineralogy requires some form of sampling to establish the nature of the subsurface characteristics. The key method of carrying out this sampling is to drill test holes at certain locations and assay the samples obtained from these holes. To enable a thorough analysis of the results of this type of survey, it is important to know the coordinate position of these holes. Furthermore, the information is much more meaningful if these holes can be drilled in the appropriate positions as predefined by the geologist.

Real time differential GPS will become an ideal method for setting out the location of these bore holes. Many of these bore holes are set out over a broad and often featureless area. The accuracy required for this setout of position is usually at the metre level. In this type of operation, the ruggedness of the system is highly important as the equipment may need to travel by vehicle over very rough terrain. The system also needs to be able to be run totally from batteries. At present this imposes a problem when using portable computers with hard disks and backlit screens.

2. GPS SYSTEM OVERVIEW

Many good overviews of the GPS system can be found in the literature. The reader is referred to King et al (1985) and Scherrer (1985) for general overviews of GPS. A more thorough look at the GPS signal structure is given by Spilker (1978). The following overview introduces the basic concepts of the GPS system. This provides a background for a more detailed look at the basic GPS measurement types.

2.1 Introduction

The Global Positioning System (GPS) is a satellite based system used to determine a user's position and velocity anywhere on, or above, the earth's surface. This is accomplished by simultaneously measuring ranges to several GPS satellites orbiting the earth. Most descriptions break the system into three distinct segments. These are the space segment, the user segment and the control segment (see figure 3). A brief description of the function of each of these segments is given in the following.

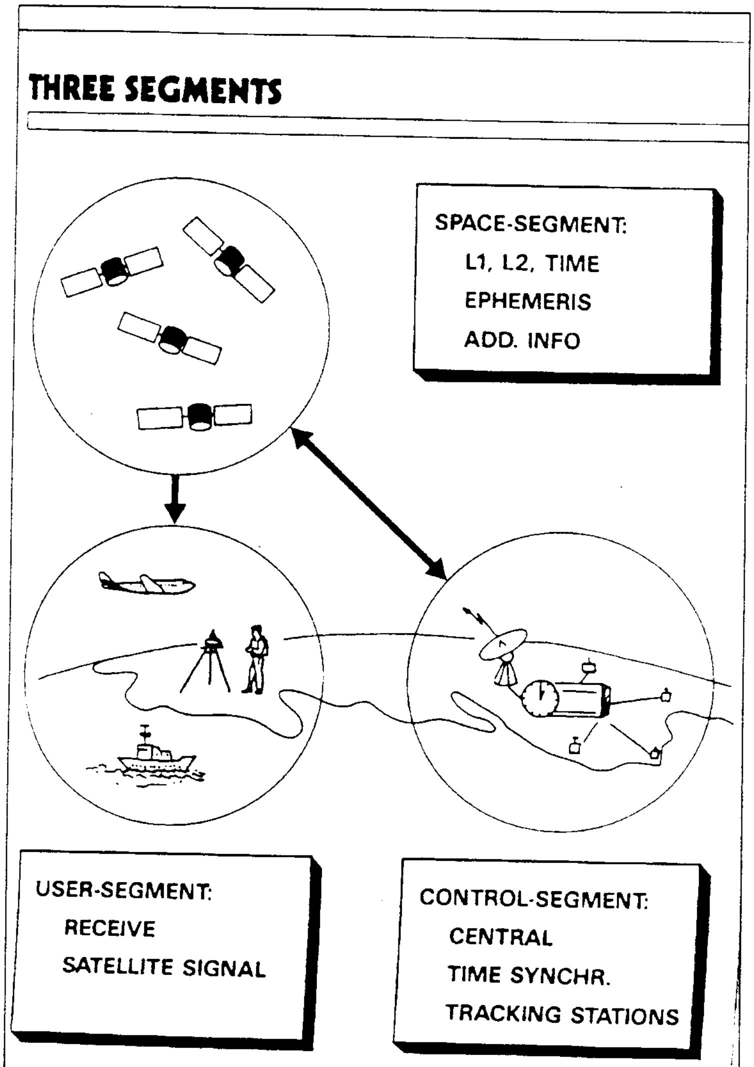


Figure 3: Three segments of GPS system (Scherrer 1985)

2.1.1 Space Segment

The final [4]Block II GPS constellation is expected to be in orbit by late 1994. It will consist of eighteen satellites in 6 orbital planes with 3 active spares also in orbit. Each orbital plane is inclined at the equator by 55 degrees and contains 3 equally spaced (120 degrees apart) satellites. The satellites are in near circular orbits, approximately 20,000km above the earth's surface. The period of the satellite orbits is 12 hours and consequently, to the user on the rotating earth, the constellation returns to the same location approximately every 24 hours. Figure 4 shows the final configuration of the Block II satellites.

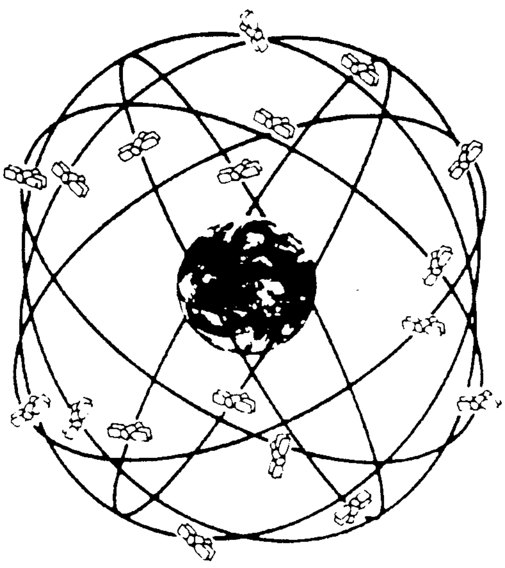


Figure 4: GPS Satellite Constellation (Sherrer 1985)

Each satellite is constantly transmitting two spread spectrum pseudo random noise (PRN) radio signals. A navigation message which contains satellite ephemeris and clock information is modulated onto this PRN signal. This information is uploaded to the satellites from the Control Segment (see section 2.1.2 below) and is needed to allow the GPS user to determine the position of the satellites and the errors in the satellite clocks. The navigation signals are transmitted at two frequencies, L1 (1575.42MHz) and L2 (1227.6MHz). Both these frequencies are derived from the same highly stable onboard atomic clock.

---

[4] The Block I phase consisted of the launching of 11 satellites of which 6 are still operational in early 1990. The Block I satellites are often referred to as the prototype constellation as their purpose was to test and validate the GPS system prior to the Block II launches.

2.1.2 Control Segment

In order for GPS to be a usable system it is necessary to know the position, in space, of the GPS satellites. Also the atomic clocks in each satellite must be constantly monitored and appropriate corrections for drift applied. This function is the task of the control segment.

The control segment consists of a master control station (located at Colorado Springs) and five tracking stations (at Hawaii, Colorado Springs, Ascension Island, Diego Garcia, and Kwajalein). These monitor stations are separated in longitude so that the maximum separation is less than 90 degrees. Data is collected at all tracking stations and sent to the master control station for processing. Here satellite clock parameters and ephemerides are predicted and sent to the appropriate ground stations for subsequent uploading to the satellites. Three ground stations are located with the monitor stations at Ascension Island, Diego Garcia and Kwajalein.

2.1.3 User Segment

The user segment is anyone with an appropriate receiver to track and decode the GPS signals. The main components required consist of the GPS antenna, receiver and data processor. Users can vary from mariners who need to know their location for navigation purposes, to land surveyors who may require very accurate positioning. The methodolgy and hardware employed are totally dependent on the user's application. This includes receiver type, measurements tracked and processing algorithms used.

2.1.4 System Overview

The basis of the GPS system is that the position of the user can be determined if accurate ranges can be measured to several points of known position. These points of known position are the GPS satellites. To make the system convenient to use, these satellites are constantly transmitting their position to the user in the form of an ephemeris message. This ephemeris contains Keplarian like elements from which the Earth Centered, Earth Fixed (ECEF) coordinates of each satellite can be calculated at any time.

The different methods of using the GPS system can be partly attributed to the methods used to determine the range measurements. Because GPS is a one way ranging system, unlike Electronic Distance Measuring Systems (EDM), accurate timing information is required to determine these ranges. To avoid the necessity for GPS receivers to have atomic clocks, receiver clock corrections are determined along with position in the navigation solution.

If the GPS receiver had a clock that could be exactly synchronized to GPS time, then measurements to as few as three satellites could be used to obtain three dimensional positioning. However, a fourth satellite is observed and this redundancy is used to determine the receiver clock offset.

Each GPS satellite transmits signals on two separate frequencies. This is specifically to facilitate the correction of certain refraction effects. Also the signal contains several modulations superimposed on these raw signals. Two types of timing codes are modulated onto the raw GPS carrier, namely the C/A (Clear Acquisition) code and the P (Precise) code. The P code provides a higher resolution by a factor of ten.

A second reason for having two codes is to allow the system to provide two distinct levels of accuracy. The C/A code was intended to provide coarse 100m positioning facility to civil users. Military users, with access to the more precise P code (or its encrypted Y code: see Selective Availability on page 23) will be able to obtain a much higher accuracy.

2.1.5 Satellite Window

When the final GPS constellation, consisiting of eighteen satellites plus three active spares, is in orbit there will be full twenty-four hour coverage. This means that at any point on the earths surface, a minimum of four satellites will be in view at all times.

However, until this time arrives, there is only a certain period of the day (the window) when sufficient GPS satellites are available for operation. A typical example of a GPS window can be seen in figures 5 and 6.

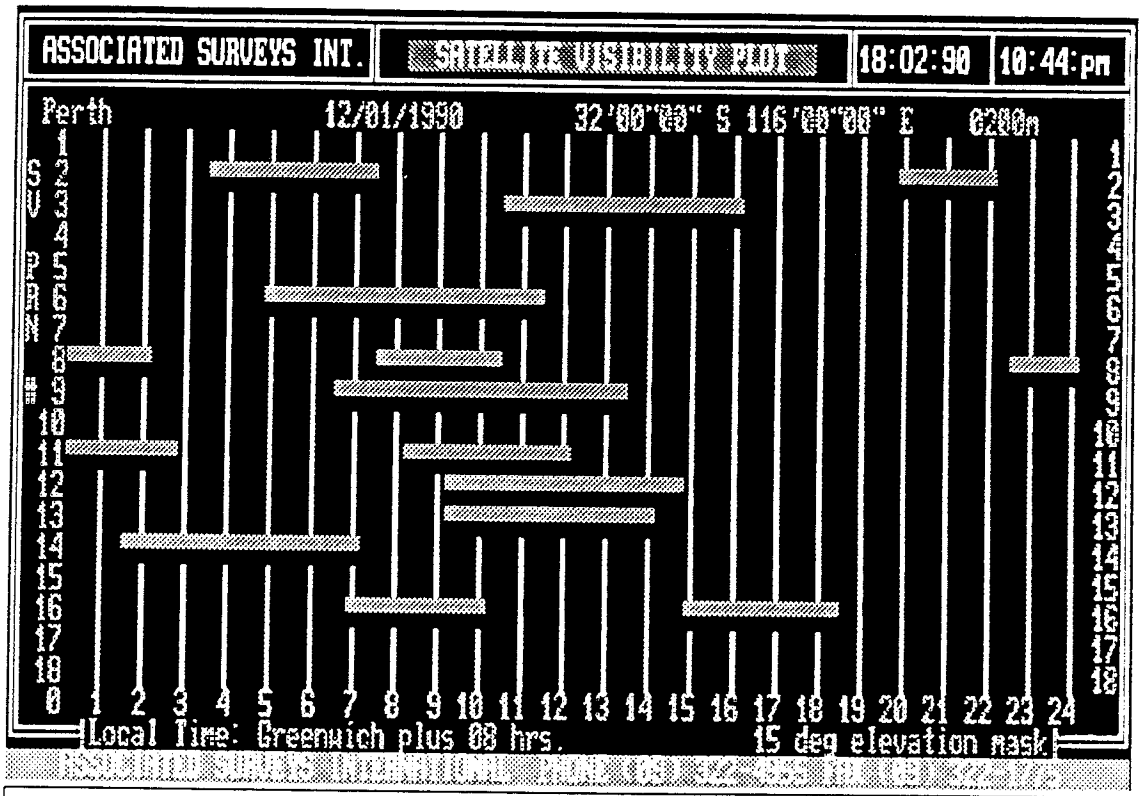


Figure 5: Typical Satellite Visibility Chart

Typically, a Satellite Visibility chart shows the period of the day when each SV is above the selected elevation mask. The Satellite Availability chart shows collectively, the number of satellites that are present for each period of the day. Hence the size of the GPS window can be seen more clearly from the Availability chart.

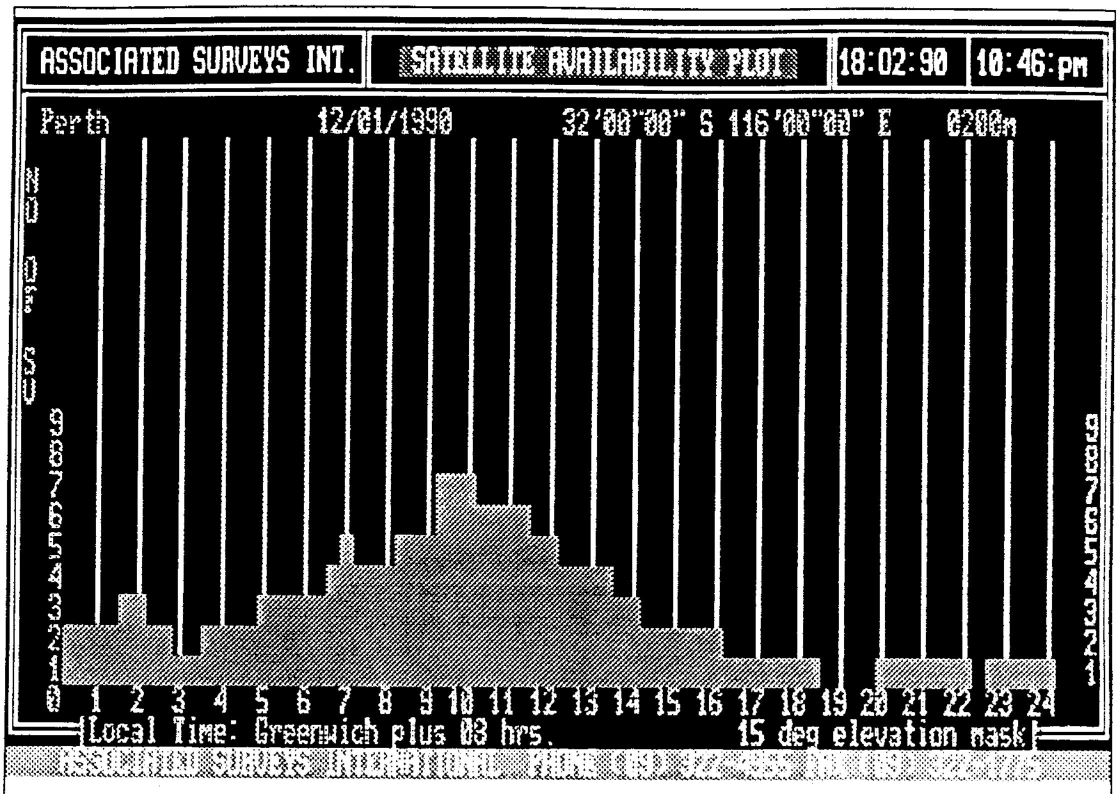


Figure 6: Typical Satellite Availability Chart

2.1.6 Dilution of Precision

Dilution of Precision (DOP) values provide a numerical means of measuring the strength of the geometry of a particular constellation. If there are four satellites in view and they are all in the same part of the sky, then most likely the DOP value will be very high. This indicates that the geometry is poor.

DOP values are derived by determining the error in position resulting from the assumption that each range has a standard deviation of unity. There are several types of DOP values, each relating to the component of interest. For instance, GDOP contains the sum of the variances in 3D position and clock offset. The PDOP value is obtained by only summing the position variances. Table 1 summarizes the derivation of DOP values. A full derivation of DOP values can be found in Jorgensen (1980) and Berveots (1987).

The vector of ranges R is related to the position vector (x) by the matrix of coefficients B

$$R = B\,x$$

if we assume that σ_{RR} (the variance of R) is a unit matrix then

$$x = (B^T B)^{-1} \times B^T R$$

$$\sigma_{XX} = (B^T B)^{-1} \times \sigma_{RR}$$

where

$$\sigma_{XX} = \begin{bmatrix} \sigma_{XX} & \sigma_{YX} & \sigma_{ZX} & \sigma_{TX} \\ \sigma_{XY} & \sigma_{YY} & \sigma_{ZY} & \sigma_{TY} \\ \sigma_{XZ} & \sigma_{YZ} & \sigma_{ZZ} & \sigma_{TZ} \\ \sigma_{XT} & \sigma_{YT} & \sigma_{ZT} & \sigma_{TT} \end{bmatrix}$$

The resultant dilution of precision values are given by

Geometric GDOP = $\sqrt{(\sigma_{XX} + \sigma_{YY} + \sigma_{ZZ} + \sigma_{TT})}$
Position PDOP = $\sqrt{(\sigma_{XX} + \sigma_{YY} + \sigma_{ZZ})}$
Horizontal HDOP = $\sqrt{(\sigma_{EE} + \sigma_{NN})}$
Vertical VDOP = $\sqrt{(\sigma_{HH})}$

where σEE, σNN and σHH are the components tranformed into local system.

Table 1: Formation of Dilution of Precision (DOP) values

2.2 Basic Measurement Types

The following section describes in some detail the characteristics of the raw GPS signal measurements. This understanding is essential in the development of our real time system as it employs a combination of the three basic receiver measurements. These three basic measurements are carrier phase, integrated Doppler and pseudo range or code phase. In this paper, the terms pseudo range and code phase both refer to the same quantity.

The majority of the following discussion on GPS measurements is generic, applying to all receiver types. However, all work carried out during this research was centred around the parallel channel type receivers, of which the Ashtech and Trimble are probably the best examples. Consequently, some of the more detailed examples of data output refer specifically to these particular types of GPS receivers.

2.2.1 GPS Signal Structure

Each satellite is constantly transmitting signals centred at the 2 frequencies of L1 (1575.42 MHz) and L2 (1227.6 MHz). Both these frequencies are derived from the "on board" atomic clock frequency of 10.23 MHz (Table 2).

CARRIER	L1	L2
WAVELENGTH	0.190m	0.244m
FREQUENCY	1575.42 MHz	1227.6 MHz
DERIVATION	154 × 10.23 MHz	120 × 10.23 MHz

Table 2: Summary of signal characteristics

The principal reason for having these two frequencies is to facilitate the removal of the ionospheric delay error.

The L1 signal is modulated with a 1.023 MHz C/A (Coarse Acquisition) pseudo random noise code. This code is a unique combination of 1023 bits and hence the code is repeated every millisecond. The modulation of the C/A code is achieved by changing the phase of the L1 carrier every 1540 cycles (see figure 7). An excellent description of this modulation technique is given in Talbot (1987).

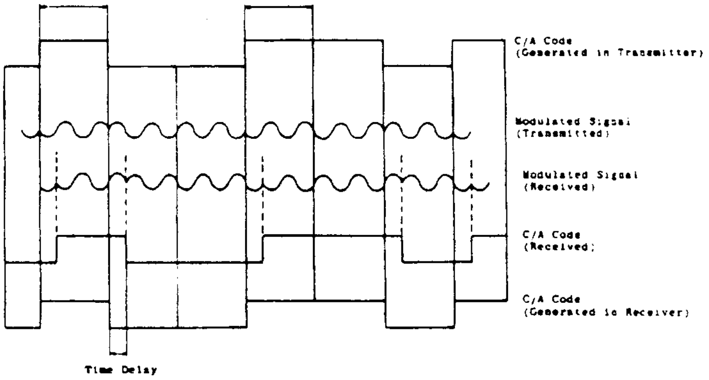


Figure 7: Code phase modulation (King et al 1985)

Similarly the L1 signal is also modulated with a 10.23 MHz P (Precise) code which has a period of exactly one week. This is a change of phase of the L1 carrier every 154 cycles. These modulations are carried out on two perpendicular components of the L1 carrier, the in-phase component and the quadrature carrier component. Both components are also modulated with a 50 bit per second navigation message. This message is superimposed as an additional change in phase every 31508400 cycles.

The full P code has a period of thirty-eight weeks, with each satellite being assigned a one week portion of this code. Hence each satellite emits a unique P code. Due to its long and rapid nature, it is difficult for a receiver to lock onto this code. To overcome this problem, the relatively simple C/A code is acquired first. This provides the receiver with timing information needed to facilitate acquiring the more complicated P code.

The L2 signal is bi-phase modulated with the P code and the navigation message. As the L2 carrier does not contain the C/A code, only users with access of the P code can acquire the L2 signal. Table 3 gives a summary of the characteristics of the C/A and P codes.

CODE	P(Precise)	C/A(Coarse/Acqusition)
FREQUENCY	10230000 Hz	1023000 Hz
PERIOD	1 week	1 millisecond
CODE CHIP	29.3m	293.3m

Table 3: Summary of code characteristics

The P code has ten times the resolution of the C/A code, hence pseudo range measurements made using the P code will naturally provide a higher accuracy. One of the advantages seen by the system designers for the two codes was that it allowed the system to provide two levels of accuracy. By design, the civil user with access to the C/A code would obtain position to an accuracy of 100m

($2 \times d_{rms}$). The military user with access to the P code would be able to obtain an accuracy of ten metres ($2 \times d_{rms}$).

2.2.2 Selective Availability

Following the launching of the prototype constellation it was found that positioning using the C/A code was capable of providing point positioning accuracy in the order of 20-40m. To overcome this, the Department of Defense (DoD) have implemented a policy of selective availability, whereby the satellite signals and broadcast ephemerides are deliberately degraded to the specified accuracy.

At present, the intentional degradation of the GPS signals is a controversial issue. The DoD argue that GPS is primarily a military system and that the 100m ($2 \times d_{rms}$) specification for civilian users is what was specified for the system. The civil users and manufacturers want the signals to be degraded only in times of national emergency.

The official DoD policy states that all Block II satellites will have selective availability (SA) turned on. When SA is on, all positioning services using the C/A or P code will be degraded to the 100 metre positioning level. Military users will have access to a Y code which contains a means of compensating for the orbit data and spoofing effects.

Anti-spoofing (AS) involves the encryption of the P code so the civilian user does not have access to it. DoD maintain that anti-spoofing may be on or off and can change at any time.

Despite criticism of these policies, at present the DoD shows no outward signs of relenting. However, Ashjaee (Ashjaee 1988) believes that although selective availability policy will not change, it will rarely be invoked. Table 4 gives a summary of the obtainable accuracies with and without SA.

CODE	CIVILIAN USER	MILITARY USER
C/A CODE	100m (2rms)	approx 30m
P CODE	100m (2rms)	approx 10m
Y CODE	NO ACCESS	approx 10m

Table 4: Summary of GPS accuracy with Selective Availability

2.2.3 Pseudo Range Measurements

A replica of the GPS PRN code is generated by the GPS receiver using the receiver's internal oscillator. This signal is then shifted in time to allow maximum correlation between this internally generated signal and the incoming GPS signal. (See figure 8)

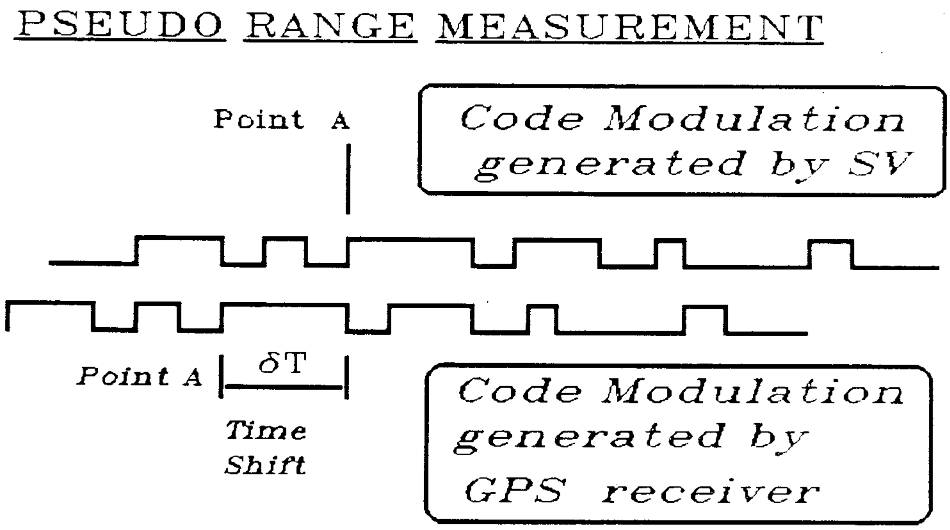


Figure 8: Pseudo range measurement

For the C/A code, which has a period of one millisecond, the maximum value for this time shift is one millisecond. The pseudo range is obtained by multiplying this time delay by the speed of light. Hence the pseudo range measurement derived from the C/A code measurement is only the fractional part (modulo 300km) of the true pseudo range distance. The P code, having a much longer period, provides the true pseudo range directly.

If there were no relative errors between the satellite and receiver clocks, then the pseudo range measurement would in fact be the true range. However, as the pseudo range measurements are scaled by the speed of light, a clock error of one millionth of a second would introduce a pseudo range error of some 300 metres. Due to the cost and mobility restrictions involved in using precise clocks, it is simpler to observe one extra satellite. This allows the errors present in the receiver clock to be solved. The pseudo range observation can be modelled as shown in table 5.

Pseudo Range Equation

$$\boxed{R = r + C + I + T + e} \quad [1]$$

where

R = Measured pseudo range
r = True range
C = Combined satellite/receiver clock errors
I = Ionospheric errors
T = Tropospheric errors
e = Unmodelled noise

Table 5: Pseudo range model

2.2.4 Carrier Phase

The carrier phase of the GPS signal has a wavelength of 19cm for L1 and 24cm for L2. By measuring the phase of the beat frequency it is possible to determine the carrier phase to less than a millimetre. Hence, the use of carrier phase measurements provides the most accurate method of utilizing the GPS signals. However, to extract this accuracy using the carrier phase requires a far greater amount of processing.

First, the code and navigation messages must be removed from the GPS signal. Code receivers use correlating techniques (Spilker 1978) to remove the code information from the GPS signal leaving a modulation free carrier. The following description of the carrier phase assumes that these code modulations have been removed.

For a constant position, the phase of a carrier signal can be viewed as the position of a cyclic oscillation at a certain time. For the signal shown in figure 9, the phase at time t = 1 is 0.25 of a cycle or 90 degrees. Over time the number of phase changes (0 to 360 degrees) can be counted and hence the continuous phase since a certain time can be accumulated. At time t = 13 the accumulated phase since t = 0 would be 3.25 cycles. The fractional phase at t = 13 would again be 0.25 of a cycle or 90 degrees.

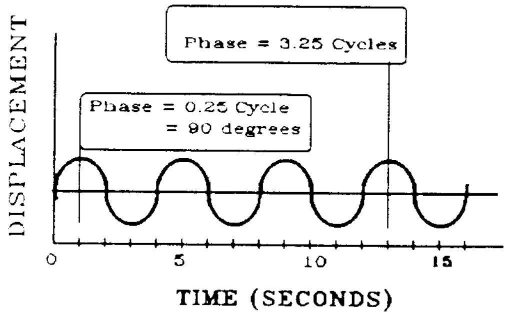


Figure 9: Carrier phase example

At the time of transmission the signal being generated by the satellite oscillator is at the nominal frequency f_S. As the signal is being transmitted both the satellite and the receiver are in motion. This relative motion between the satellite and receiver causes the wavelength of the propagating signal to be altered. This is the Doppler effect.

However, while the frequency and wavelength of the signal are altered, a point of constant phase travels at the same speed. Remondi (1985) offers the analogy of a signal being like a wave traversing the sea. The relative Doppler causes the waves to get closer together but the wave peak is still travelling at the same speed.

There are many possible methods for determining the phase of the incoming signal (Remondi 1985). However, the method most commonly used is to measure the phase of this incoming signal relative to the phase of the receiver's local oscillator. Among other factors, this method greatly reduces the physical complexity and hence cost of processing signals at very high frequencies.

The raw measurement can now be viewed as the difference between the Doppler shifted signal and the receiver generated signal. This, in effect, is a measure of the Doppler or relative motion between the satellite and receiver.

If there were no relative movement between the satellite and receiver then the beat frequency would be zero and the beat phase measured at any time would be constant. The sign of the beat frequency, and hence phase, changes depending on when the satellite is approaching or moving away from the receiver. The [5]RINEX format (Gurtner 1989) defines the phase change as being positive as the satellite moves away from the receiver.

At predetermined time intervals (usually measured by the receiver clock) the carrier beat phase of all channels of the GPS receiver are measured. This measurement does not need to be strictly simultaneous. However in this discussion it is assumed that measurements are simultaneous and hence that these measurements are made using parallel channel receivers.

Figure 10 shows an example of this measurement for a four channel receiver. At the time of measurement the fractional part of the carrier phase only is measured. That is, the fraction of a cycle since the last zero crossing of the carrier is determined. It is this measurement of the fractional carrier phase that the receiver hardware can measure most accurately (to a fraction of a millimetre). In essence, this measurement is made by determining the phase shift required to bring the receiver generated and satellite signal into maximum correlation.

---

[5] RINEX (Receiver Independent Exchange format)

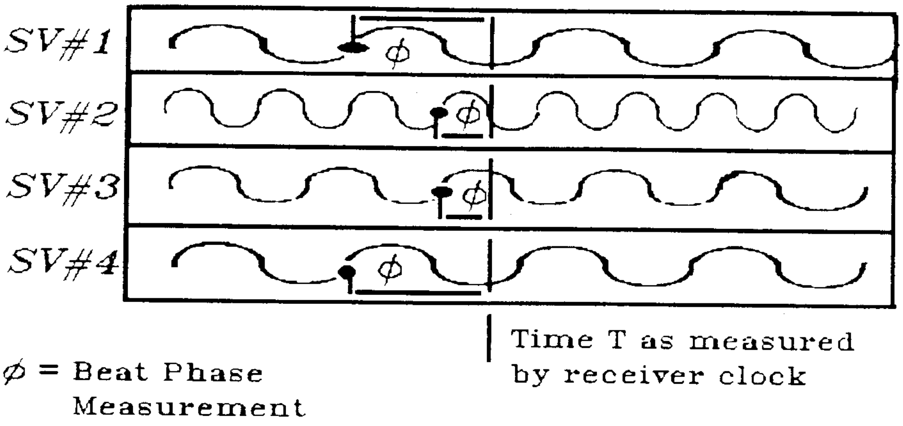


Figure 10 Example of carrier beat phase measurement

It should be noted that the receiver hardware cannot automatically determine whether this phase measurement is relative to the zero or 180 degrees phase crossover. The correct polarity of the phase measurement is determined from the code measurements. Hence upon initial lock of the GPS signals and after signal loss, the receiver sometimes has difficulty determining the correct signal polarity.

The fractional phase measurement can be made very accurately. However it is also very useful to know the integer number of wavelengths passing in between each of these instantaneous fractional phase measurements. If it is assumed that the receiver hardware has the ability to count the number or cycles passing in between each instantaneous phase measurement, then our phase measurements are continuous (at least since the initial lock of the satellite signal was obtained). These continuous phase measurements can then be viewed as a very accurate pseudo range measurement with the initial number of cycles missing. As long as the receiver maintains phase lock on this satellite, the missing number of cycles is the same for each measurement.

Much of the complexity in the processing of carrier phase measurements is in the determination of this missing cycle ambiguity. Excellent treatment of the modelling of the carrier phase for subsequent processing of these measurements is given in Remondi (1984), Goad (1985) and King et al (1985).

It should be reinforced that the carrier phase measurement actually made by the receiver is the fractional part of the carrier phase only. Hence, the receiver requires a method of determining the number of cycles passing between successive phase measurements in order to determine the continuous carrier phase reading.

The simplest method of concatenating these phase measurements is by using the measurements of the Doppler frequency (Remondi 1984) or the integrated Doppler as shown in section 2.2.6.

The relationship between the true range and the continuous and fractional carrier phase is shown in table table 6. It can be seen that in order for the fractional carrier phase measurement to be useful, two constants must be determined. The first is the value of inital distance between the satellite and receiver at the time that lock was obtained (N). The second is the number of whole cycles that have passed since lock was initially obtained (n0). The carrier phase ambiguity is split into these two components as n0 can be determined accurately by the receiver hardware. The accuracy of the determination of the inital bias N is the limiting factor in all positioning using carrier phase.

Fractional Carrier beat phase

$$\boxed{\phi_F = \phi_s(t_T) - \phi_r(t_R) + \phi_{noise}} \quad [2]$$

where

ϕ_F = fractional carrier beat phase
ϕ_s = carrier phase transmitted by satellite
ϕ_r = carrier phase generated by receiver oscillator
ϕ_{noise} = random measurement noise
t_T = GPS time of signal transmission
t_R = time of signal receipt measured by receiver clock

Continuous Carrier beat phase

ϕ_C = n + n0 + ϕ_F

where

ϕ_C = continuous carrier phase
N = number of cycles since initial lock of signal
n0 = unknown number of cycles at time of acquiring lock

Table 6: Continuous Carrier Phase

Table 7 shows the model relating the continuous carrier phase measurements to the true range.

Modelled Carrier Phase Equation

$$\boxed{\phi_C \times \lambda = r + c(\Delta tr - \Delta ts) + T + I + \epsilon} \quad [3]$$

where

ϕC = continuous carrier phase
λ = L1 wavelength
r = true range
c = speed of light
Δtr = receiver clock error at time of receipt
Δts = satellite clock error at transmission time
T = tropospheric delay
I = ionospheric delay
ϵ = measurement noise

Table 7: Modelled Carrier Phase

2.2.5 Doppler Frequency

It is possible to measure accurately the beat frequency of the incoming signal against the known reference oscillator. This Doppler frequency is actually the pseudo range rate or the combined change in position and clock offsets between the satellite and receiver.

This Doppler measurement can be integrated to obtain the integrated Doppler measurement described in section 2.2.6. The Doppler frequency also provides the user with the most accurate way of determining the velocity of the user in moving applications. Methods of using these Doppler measurements to strengthen the pseudorange solution and provide navigation with fewer satellites are also proposed (Ashjaee et al 1989).

Another use for the Doppler is when it is required to adjust the time tags of the GPS measurements. The Doppler can be used to interpolate the carrier phase measurements short distances to the new time tags.

2.2.6 Integrated Doppler

If the Doppler frequency is measured at small regular intervals (both the Trimble and Ashtech receivers use one millisecond intervals), then the rate of change of the Doppler frequency will be stable. Hence, the frequency measurements can be integrated by multiplying each measurement by the small time intervals and summating the result. Remondi (1984) shows that a small term due to relativity effects can be ignored for short intervals. Errors caused by this integration are related to the integration interval and the stability of the receiver clock.

This integration provides an accurate measure of the change in pseudo range between successive epoches. The integrated Doppler provides a more accurate measure of the change in range than simply taking the difference between two consecutive pseudo range measurements. However, the distance between the satellite and the receiver at the time that the integration started (the value N in table 6) is unknown.

To gain an appreciation for the accuracy of these integrated Doppler measurements, test data was recorded using an Ashtech XII receiver. The noise in the integrated Doppler was investigated by inspecting the phase residuals for this period. The plot of these residuals was very similar for all satellites. A typical plot can be seen in figure 11. From this plot it can be observed that the noise in the integrated doppler measurements is less than 0.03 cycle (one sigma). This amounts to less than 6 millimetres.

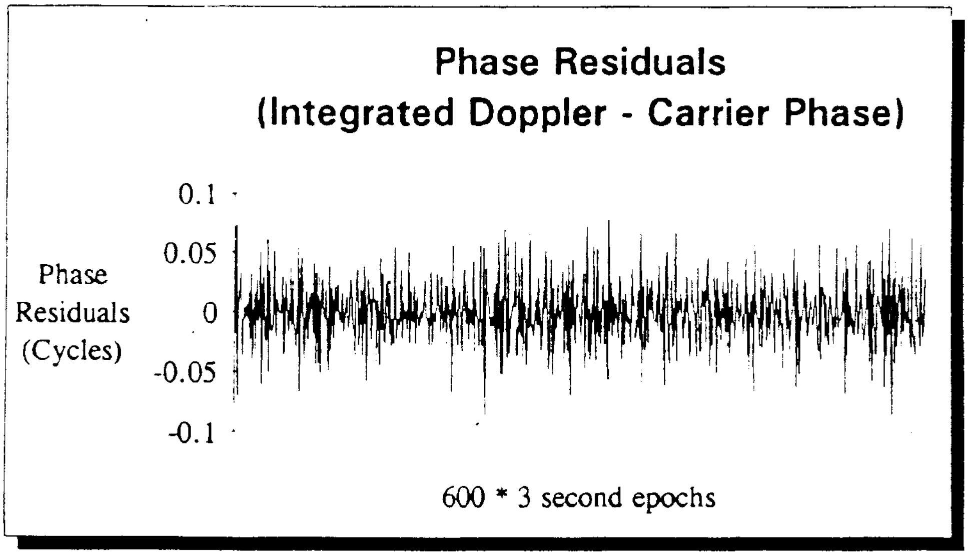


Figure 11: Noise in Integrated Doppler measurements

However, in addition to this random noise there is also an accumulative error due to the summation of the frequency measurement errors. Figure 12 shows a plot of the difference between the integrated Doppler values and the more accurate continious phase values. Over the 30 minute period of the plot, the maximum drift observed for any satellite was less than 0.7 of a cycle or 0.13 metre. Hence, if the integrated Doppler was used exclusively, it could cause range errors in the order of 0.5 metre in a two hour period.

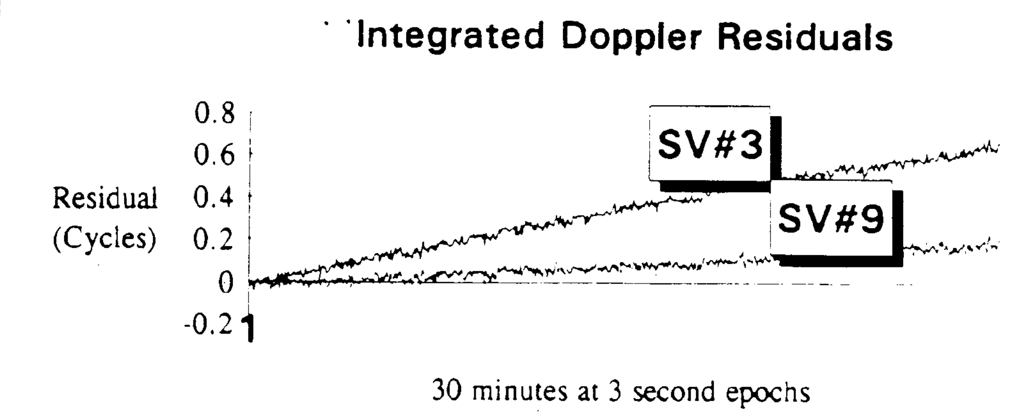


Figure 12: Drift of Integrated Doppler measurements with time

Figure 12 shows one small sample of data. However it is a fairly typical example and so provides a feeling for the accuracy of the integrated Doppler measurements.

The integrated Doppler measurement could be processed directly instead of using the more accurate carrier phase measurements. However, more importantly it is used to determine the number of cycles that have passed between consecutive carrier phase measurements. The carrier phase measurements are used to determine the fraction of a cycle, while the integrated Doppler is used to determine the number of cycles that have passed in between two measurement epochs. In this manner it is possible to determine a continuous phase measurement since the initial lock was obtained.

2.2.7 Determination of Continuous Carrier phase

The integrated Doppler measure of the change in range can be used to determine the integer number of cycles (n0) that have passed in between successive carrier phase measurements. This integer value can be determined as shown in table 8. (Ashjaee 1989).

The phase residual is the amount that n varies from an integer. It is predominantly due to noise in the integrated Doppler measurements. Hence the plot of this phase residual is a good indicator of the noise in the integrated Doppler measurements.

Relationship between carrier phase and Integrated Doppler (Ashtech measurements)

$$\boxed{\Delta n_{12} + C_2 - C_1 = D_2 - D_1} \qquad [4]$$

where

Δn_{12} = number of cycles passed between epoches 1 and 2
C_1 = fractional carrier phase measurement at t1
C_2 = fractional carrier phase measurement at t2
D_1 = integrated Doppler at time t1
D_2 = integrated Doppler at time t2

therefore

$$\Delta n_{12} = (D_2 - D_1) - (C_2 - C_1)$$

will be very close to an integer.

Phase Residual (ϕR)

$$\phi_R = \Delta n_{12} - \text{integer}(\Delta n_{12})$$

Integer number of cycles passed since lock acquired (n_0)

$$n0 = n0 + \Delta n_{12}$$

Continuous phase since lock obtained

$$\phi_C = n0 + C2 - C1$$

where

ϕ_C = Continuous phase at Time t2 (since time t_0)
n_0 = Integer number of cycles since lock time t_0

Table 8: Using Integrated Doppler to determine Continuous carrier phase

It should be noted that for the relationship shown in equation 4, table 8, to be correct, it is essential that the carrier phase and integrated Doppler measurements are made at the same time. Unfortunately, this is not the case as the carrier phase measurements are made at predetermined times of the receiver clock. The Doppler integration epoch may be up to one millisecond after this time.

However, this time difference can be easily corrected as it is a function of the raw pseudo range distance (see equation 5, table 9). The raw measurements output from the Trimble SX receiver needs to have this correction applied before they can be combined using the formula in table 8. The Ashtech receiver makes this correction to the integrated Doppler time tags internally. This ensures that both carrier phase and integrated Doppler are at the same time tag.

$$\boxed{D(tC) = D(tD) - p/c} \quad [5]$$

where

D = Integrated Doppler measurement
tC = time of carrier phase measurement
tD = time of Doppler integration

Table 9: Correction of integrated Doppler measurement to time of carrier phase measurement

2.3 Error Sources

The following describes the key error sources affecting any GPS measuring system. The aim of the real time navigation system is to minimize the effect of these errors. The residual effect of the remaining errors also must be estimated. Hence an understanding of its source is essential.

2.3.1 Satellite Orbit Error

A major source of error in the GPS system is due to the inaccurate coordinates of the GPS satellite positions and clock offsets. It is difficult to distinguish between a satellite position error and a clock offset as they both affect the solution in a similar manner. However, with differencing between receivers, satellite clock errors will cancel completely. Satellite position errors will only partially cancel depending on the relative geometry of the stations.

The broadcast ephemerides contain elements that are Keplarian in appearance. These elements have been computed by the control segment to match the true orbit only for the short period that the ephemeris is applicable. By design, this period of application for each ephemeris set is one hour. The ephemeris is still valid for an additional period of at least half an hour after this period. However, after this period, the degradation of the orbit is quite severe (Van Dierendonck et al 1982).

During normal operation periods it is expected that the Control segment will upload ephemeris data every eight hours. In this case the errors due to the prediction of the ephemeris is considered negligible when compared to the orbit estimation error (Ananda 1988). However, the clock error is totally due to the inability of the control segment to predict the instability of the satellite clock. Over an eight hour period this can be as much as 3.3 metres (one sigma)(Ananda 1988).

In the process of predicting these Keplarian elements for the period ahead, the Control segment also produces a statistic that represents the line of sight error vector. This statistic, the user range error, is contained in the ephemeris. It represents the projection of all system errors due to the inability of the control segment to estimate the spacecraft clock and ephemeris parameters.

The system has been designed so the user range error due to the space and Control segments is less than 6.0 metres (one sigma) (Ananda 1988). However, this may change with the enforcing of selective availability.

Remondi (1989) investigates the accuracy of the GPS broadcast ephemeris. This is achieved by a direct comparision with the precise ephemeris supplied by the Naval Surface Warfare Center (NSWC). One conclusion from this study was that the maximum difference between these ephemeris sources was less than 8 metres for satellites with cesium oscillators. Also it was determined that the precise ephemerides were good to approximately 5 metres.

For optimum performance it is important to use the ephemeris set pertinent to the time of operation. The major portion of this satellite orbit error is cancelled by the differencing process. However as the length of the baseline increases the residual error also increases. From similar triangles it can be seen that an error of x metres in satellite position can cause an error in baseline length of x multiplied by the ratio of the baseline distance divided by the orbital height (approximately 20,000km).

2.3.2 Satellite Clock Error

As stated in section 2.3.1, errors due to poor modelling of the satellite clocks can cause an error of up to 3.3 metres (one sigma). As this error will affect two sites identically, its effect will be cancelled entirely in the differencing process.

2.3.3 Tropospheric Delay

The Troposphere is the part of the atmosphere located between the earth's surface and extending to an altitude of approximately fifty kilometres. The effect of the troposphere on HF signals traversing through this medium is to refract the signal. Both pseudo range and carrier phase measurements are retarded by this tropospheric delay.

The magnitude of the delay caused by this tropospheric refraction typically varies from 2.3m at the Zenith to 25 metres at an elevation angle of five degrees. (Janes et al 1989) Some 90% of this delay is due to the dry gas component of the Troposphere. This can be modelled within 2-5%. The remaining 10% due to the water vapour content is much more difficult to model.

The magnitude of the tropospheric delay is a function of the elevation angle, temperature, pressure and humidity along the path of the signal. However, as these factors are difficult to measure, most models either assume standard atmosphere or make use of surface meteorological readings. There are several popular tropospheric correction models used in GPS data processing algorithms. Janes et al (1989) give a good summary of the Hopfield, Marini and Saastamoimen models. They concludes that such models can sufficiently determine the tropospheric delay to an accuracy level of approximately 10-20cm. Analysis of these models is beyond the scope of this thesis.

The major portion of the Tropospheric correction is eliminated in the differencing process. Consequently, including an accurate tropospheric model in real time software imposes a considerable overhead for a very minor accuracy improvement. A simple model relating the delay to the satellite elevation angle only, will eliminate the major part of the error (Ashjaee 1986b). This model, used in the Trimble receiver, is shown in table 10.

Tropospheric delay

$$\boxed{\text{Delay} = 2.4225/\ (\sin(E) + 0.025)\ \text{metres}} \qquad [6]$$

where

E = satellite elevation angle

Table 10: Simplified formula for Tropospheric Delay

Two sites, 100km apart, may have a different elevation angle to a satellite by as much as 0.29 degrees. Using the simplified formulae in table 10, this change in elevation would cause a maximum delta range correction of 0.15 metre.

2.3.4 Ionospheric Delay

The ionosphere is generally considered to be the region of atmosphere between 50km and 1000km above the earths surface. Ultraviolet radiation from the sun causes molecules of gas in this region to become ionized into positive ions and free electrons. Electromagnetic signals passing through this ionized medium are affected by its nonlinear dispersive properties.

When traversing this medium, the propagation of the phase of the GPS signal (phase velocity) is advanced with respect to the phase velocity of a wave in a vacuum. The effect of this is a small frequency shift of the incoming GPS signal. Hence, if an unambiguous phase measurement could be made, the observed measurement would be shorter than a measurement made in a vacuum. However, the propagation of the energy within the wave (the group velocity) is retarded. Hence the code information modulated on the signal takes longer to reach the receiver. This in turn can be seen as a slightly longer pseudo range measurement.

This delay can be as much as 15 metres in zenith and may increase by a factor of 2.5 at lower elevation angles (Remondi 1984). The magnitude of the delay is proportional to the Total Electron Content (TEC) and inversely proportional to the square of the frequency. Dual frequency receivers make use of this relationship with the signal frequency. By making observations on two separate frequencies, simultaneous equations can be solved to determine accurately this ionospheric delay (King et al 1985), (Hatch 1982). The selection of the two frequencies emitted by the GPS satellites was made to facilitate this correction.

Due to the complexities of the constituents of the ionosphere it is impossible to predict precisely this ionospheric delay. However, single frequency users can determine an ionospheric approximate correction through the use of an appropriate model. This model must determine the TEC along the path of the signal. The main factors that effect the TEC are:

- The time of the day
- The time of the year
- Period within the 11 year sunspot cycle
- Local sunspot activities

- Latitude of location.
- The elevation/azimuth of the satellite

Coefficients for a single frequency ionospheric model are contained in the GPS broadcast message. This model developed by Klobuchar can reduce the ionospheric correction by more than 50%. (ICD-GPS-200). More recent models (Finn et al 1989) require information such as a predicted value of the 10.7cm solar flux. Finn suggests that the broadcast message could be altered to include this information.

In general, high order accuracy ionospheric prediction models can be very computation intensive. Hence, in a real time, a trade-off must be made between accuracy and computational overhead.

2.3.5 Multipath

Multipath occurs when a reflected signal reaches the GPS antenna interfering with the direct signal. As GPS signals are right hand polarized, any reflected signal will be reversed and hence will be rejected by the GPS antenna. However, this rejection property of GPS antennas is usually poor at low angles. Hence, low elevation satellites are more subject to the effect of multipath errors.

Talbot (1988) argues that the noise level introduced by multipath effects can be reduced by weighting all observations by the signal to noise ratio of these measurements. In general, the affects of multipath can be minimized by careful choice of antenna sites. Where this is not possible, multipath effects are simply treated as noise in the data.

2.3.6 Receiver noise

One of the limiting factors concerning the accuracy that can be achieved using pseudo range measurements is due to the receiver hardware. With parallel channel receivers, inter channel biases must be determined by the receiver. This is done by tracking one satellite on all channels and determining the offsets due to electrical delays etc on each channel. Once this bias is determined it is applied to all subsequent measurements.

Any errors in this determination of the interchannel bias is peculiar to only one receiver and one satellite. Hence this error will not cancel but rather propagate in the differential process. It is assumed that all pseudo range measurements are made at the same instance. Ashtech state that the XII receiver provides synchronization of phase measurements to better that one nanosecond. This is excellent for phase measurements. However, a constant error in pseudo range measurements of 0.5 nanosecond would amount to a range error of 0.15 metre. Again, as this type of error would propagate in a differential system, it limits the obtainable accuracy.

Other receiver dependant factors such as the tracking bandwidth of the receiver are also important. The Ashtech receiver has a variable width tracking bandwidth. It automatically adjusts with the noise level of the incoming data. The dynamics of this automatic bandwidth can also affect the ultimate precision of the GPS data.

However, these types of factors cannot be intrinsically measured by the user. Instead their existence is acknowledged and modelled. From analysis of a number of results, the magnitude of the combination of these factors can be estimated for a particular receiver type.

2.4 GPS Methods

When the Global Positioning system was initially conceived it was to provide real time positioning through use of the code measurements it provided. However, since then both the civil and military users have developed ingenious methods of utilizing this system to extract greater accuracy and efficiency. The following sections give an overview of some of the basic methods of use of the GPS technology today. The terminology used is based on Remondi (1988).

2.4.1 Navigation

In navigation mode the GPS receiver is used utilizing the C/A or P codes. Usually this is for some form of navigational purpose where the receiver is in motion. The accuracy obtainable is 10 metres rms using the P code and 25 metres if using the C/A code. (This assumes selective availability is not being enforced.) If the satellite signals are being degraded then the accuracy of the system, to the civilian user, will be 100 metres.

All results are obtained in real time. The pseudo range solution may also be smoothed with the carrier phase to obtain a more precise if not more accurate solution.

2.4.2 Point Position

This method uses the same method as navigation but is used to obtain a good absolute position for a static point.

2.4.3 Static GPS

In static GPS, two or more receivers are used and the relative position between these receivers is determined very accurately. The carrier phase measurements are used, usually by forming double differences, solving for the integer biases and applying them to the measurements.

The observations require simultaneous occupancy of the stations for a period of 30 to 120 minutes. Results are obtained after processing of the data from both sites. With single frequency receivers, an accuracy of less than 2 ppm is quite obtainable.

Static GPS is the most common use of GPS for any case that requires an accuracy better than a few metres. It is used for conventional surveys, photogrammetric and seismic control, deformation monitoring and indeed anywhere where accuracy is required.

2.4.4 Kinematic

The kinematic method as introduced by Remondi (1985b) uses carrier phase in a double differencing manner similar to static GPS. However, the integer biases in the double difference solution are solved prior to the start of the survey. This can be achieved if the coordinates of the two starting stations are already known. Once the integer biases have been solved, they are held fixed as one receiver is taken to a number of other stations.

Providing this roving receiver retains lock on at least four satellites, then accurate coordinates for all stations visited can be determined. This method has the advantage that a large number of stations can be accurately determined very quickly. However, it also imposes several serious restrictions. First, the coordinates of the starting positions must be known accurately. Secondly, phase lock must be maintained on at least four satellites during the survey.

Because of the need to maintain lock, a practical minimum of five or six high satellites is required for kinematic GPS. This greatly reduces the window available for this type of work at present. It should be noted that, like static baselines, kinematic results are obtained after post processing. Passing the complete raw data required for forming double differences in real time would require an extremely good communication link.

2.4.5 Differential Dynamic Positioning

Remondi (1988) defines the term "dynamic positioning" to refer to any type of GPS positioning of a moving platform. It may use carrier phase, pseudo range or both. Similarly, he defines "differential" positioning in the broad sense to apply to any type of positioning where corrections are used from a reference site to improve the accuracy of a second site. Results can be obtained from post processing or obtained in real time.

In this light, the system described in this paper cannot be classified using just one of these names. Hence the system is referred to as utilizing "realtime, differential dynamic, GPS". However, even this description does not specify whether carrier phase smoothing techniques are being used.

2.4.6 Pseudo Kinematic

Another method of using GPS is the pseudo kinematic method. This method uses the carrier phase measurements in a similar manner to static GPS. Integer biases are determined during the survey using double differencing methods. However, rather than recording data continuously for a long period, two short periods of data (typically five minutes duration) are recorded with a large period in between.

Consequently, it is possible to visit many stations for a short period and by visiting each station twice, geodetic accuracy can be obtained. This is due to the fact that it is the change in geometry of the GPS satellites that gives the solution its strength in static baseline processing.

Pseudo kinematic may increase efficiency over static GPS methods for closely spaced points. However, in many cases the logistics of visiting points twice make the method unfavourable.

2.5 Receiver Technology

Since the introduction of the Macrometer V1000 receiver in 1983, GPS receivers have evolved rapidly. An understanding of some of the features of today's receivers can be seen in the evolution of GPS receiver hardware. The following discussion does not attempt to address every issue or receiver type. It is intended to provide a background knowledge only.

2.5.1 Macrometer

The interesting feature of the Macrometer, apart from its size, was that it did not track the GPS codes. Hence, it was necessary to bring the units together before and after the survey to synchronize the receiver clocks. Also as the receiver did not obtain the broadcast ephemeris, a suitable ephemeris had to be obtained from an external source prior to processing. At the time of its release, this feature was marketed as being an advantage in that the GPS codes were not required.

To obtain the phase measurements without knowledge of the GPS codes, the input signal is multiplied by itself to obtain a modulation free signal. This method of obtaining the carrier phase measurement is termed squaring the signal. Its disadvantage is that in the squaring process the measurement noise is also squared.

2.5.2 TI4100

Texas Instruments' TI4100 was the first receiver that made maximum use of both the pseudo range and carrier phase information.

The receiver has one single hardware channel, however up to four satellites on both L1 and L2 frequencies can be tracked. This is achieved by tracking each satellite in turn for a few milliseconds, returning to the first satellite every twenty milliseconds so the signals appears to be tracked continuously. This type of operation is termed "multiplexing". One of the advantages of multiplexing is that the GPS measurements are based on signals passing through identical receiver hardware. Hence errors due to electronic path delays and interchannel biases are virtually eliminated.

Another feature of the TI4100 is that it allowed the user to select from three tracking bandwidths depending on the amount of required dynamics of the application. Also it tracked both the P and C/A codes. The weight of the receiver without tape recording unit or antenna is 53 lbs.

The TI4100 was built to full military specifications. It is worth noting that a significant proportion of research carried out during the last 7 years was with these receivers. This is principally due to its reliability and dual frequency capability. Its shortcomings today are its size, and the limitation imposed in tracking only four satellites.

2.5.3 Wild Magnavox

The Wild Magnavox WM101 receiver has four physical hardware channels so four different satellites can be tracked simultaneously. In addition, each channel can operate sequentially, tracking up to two satellites on each of three channels. The fourth channel is used as a control channel sequentially tracking all available satellites.

Sequencing is a hardware/software method of tracking satellites for a short period (two seconds in the case of the WM101) and then switching to another satellite for a similar period. The principal advantage of sequencing is that it allows more satellites to be tracked for a certain number of channels. In static applications it is a simple matter for software to determine the change in carrier phase between successive measurements. However, in high dynamic modes the concatenation of the continuous carrier phase can become a problem.

In early 1989, Wild Magnavox released a dual frequency unit with eight physical channels available. A sequencing strategy is still required for dual frequency tracking of six or more satellites. The WM102 offers the ability to extract the L2 frequency using the P code. This is the more accurate method as squaring the frequency also increases the measurement noise. If the P code is encrypted, the WM102 can provide L2 measurements using a squaring technique.

2.5.4 Trimble Receivers

The Trimble 4000 series (S, SX, SL and SDL) offered parallel channels, one channel tracking each satellite. This receiver became very popular, partly due to this feature, but predominantly due to its lower pricing relative to the rest of the market. The original 4000S tracked only 4 satellites. Later models had 5 dedicated channels and this can be increased to 10 dedicated channels. A dual frequency unit was released in 1989, however, this implementation does not use the P code.

This dedicated channel technology combined with sophisticated internal software made the Trimble receivers an excellent choice for both geodetic and navigational work. The ST receivers released in early 1989 offer all the features of the Ashtech receivers described below.

2.5.5 Ashtech

The Ashtech receiver introduced in late 1988 can be viewed as the next step in the evolution of the Trimble style receivers. Through sophisticated custom hardware, 24 independent channels can be contained in a unit that weighs 10 lbs. This is one fifth the weight of the original TI4100.

The ability to track all the satellites in view at any time simplifies the requirements of survey planning. Another first with this receiver was that the user was no longer required to enter the initial position as part of the receiver set-up procedure. Previously all receivers required the user to enter an initial position to determine the ambiguity associated with the C/A code measurements.

The receiver also features a self adjusting tracking bandwidth that automatically adapts to the dynamics of the situation.

The new Trimble ST receiver is very similar to the Ashtech XII receiver in performance. It also has 'all in view' capability and does not require the entry of an initial position. At present both these receivers offer dual frequency models without the use of the P code.

2.5.6 Summary

In general the evolution of the GPS receiver has seen many changes. In hindsight it is apparent that the advantages obtained by tracking the codes make the use of squaring techniques an emergency option only. This trend is now becoming apparent for dual frequency implementation. Recently, public opinion has leaned towards the belief that selective availability will be enforced but that the P code will be rarely encrypted. Hence manufacturers that originally opted for the more reliable squaring techniques are now looking at utilizing the P code in their machines.

Similarly, the main advantage of sequencing and multiplexing technology was the reduced cost of receiver hardware. However, trends in hardware technology take away this advantage but do not alleviate the constraints imposed by sequencing strategies.

3. PROGRAMMING CONSIDERATIONS

After considerable research, it was decided to develop the navigation system to run under the "Microsoft Windows" environment. Windows is a subsystem that runs under the familiar MS-DOS operating system. It was also decided to develop all the source code in the 'C' programming language except for some small routines being written in assembler for higher performance. This section discusses some of the considerations behind these decisions.

Usually, a program is considered to be successful if it performs the task that it was proposed to do. However, considering the high cost of software development, it is imperative that a program is capable of adapting to a variety of future needs. It must be able to be maintained by other programmers less familiar with the overall system and should not require major structure changes to incorporate new functions.

As computer applications become bigger and more complex, the degree of structure and clarity must also increase. This is essential for the application to continue to grow. When writing small programs (typically less than 2000 lines) it is quite easy for the programmer to keep track of the program. If the program is unstructured or has a few idiosyncrasies, the programmer will no doubt be aware of these limitations.

As the program grows, many new problems arise. After a program reaches a certain size, it is impossible for the human mind to keep track of all of its complexities.

In Section 1, the aims of the project were defined. These goals included provision for future growth and economic maintainability. For this to be possible a program must be written with several things in mind.

The program source code must be well structured, modular, and clearly commented to facilitate easy maintenance of the program. This makes it feasible for other programmers to provide maintenance, support and enhancements to the program in the future.

At the time of writing this paper, the program source code consisted of over 10,000 lines and falls into the class of a medium sized program (between 2,000 and 100,000 lines of source code). It is anticipated that the program will have a life-span of three to five years and during this time will be maintained by many different people.

No language can ensure that the programmer will write well structured and commented code. However, the choice of language is a very important aspect of the development of this type of system. The same applies to the choice of operating system and software tools employed. These factors have a significant effect on the time taken to achieve the project goals and the resultant quality of the final product.

3.1 Languages

Since the development of FORTRAN, in 1954, programming languages have evolved rapidly. Amongst the many established languages available today, the best choice for a particular project is still subjective. In many organisations, upgraded versions of the FORTRAN language are still used for many scientific applications. Often this is simply because much of the organisation's existing software is already written in the FORTRAN language. Also their programmers are no doubt most comfortable with this language and hence it is simpler and more economic to continue using this older, less structured language.

In this case, the language is not used due to its own merits directly, but simply due to the extraneous factor of compatibility. Other extraneous factors that may affect the choice of language are the availability of programmers. For instance, it would be unwise to write a program in an obscure language (no matter how good the language was) if only very few people could maintain the code.

Similarly, the software tools available for a particular language affect its value. If programmers have already spent considerable time designing graphic libraries and printer interfaces for a particular language then this makes the use of this language much more attractive. The general popularity of a language in the industry is also important. If a large section of the industry is using a language it is highly likely that in the future, compilers, software libraries and programmers will be there to support the product.

Apart from these external factors, the merits of the programming language itself are also very important. First, the language must have the capacity to perform the specified task. It should facilitate the development of code that is easy to read and maintain. It should also encourage consistent error handling and modular design.

The following summarizes some of the languages available for use in the development of this project. It is beyond the scope of this paper to go into specific details of each of these languages. However, the following is intended to create awareness of the available options.

3.1.1 FORTRAN

Fortran was designed for scientific and numeric applications. The language was designed with the primary goal of execution efficiency. Program correctness was not a goal of the language and hence it contains many harmful features such as the GOTO instruction. The Fortran language has had several major upgrades (in particular in 1966 and 1977). The majority of scientific code written during the last three decades is in the FORTRAN language.

3.1.2 BASIC

Basic was introduced in 1965 and was the first truly interpretive language. This made it much easier to use for small programming applications. However, while this interpretive aspect is ideal for development and testing, the extra overhead required in running large programs inside an interpreter is a distinct disadvantage. In recent years BASIC has been upgraded to include many of the features of C and PASCAL in its interpretive mode. Modern BASIC can also be compiled outside of its interpreter making it a much more powerful high level language.

3.1.3 PASCAL

PASCAL was originally designed (in 1971) for teaching disciplined programming. It is a high level programming language, the chief strength of which lies in its structure. This highly structured approach to both code and data leads to improved clarity in programs which assists in writing and maintaining easy-to-read programs. PASCAL has influenced nearly all modern languages.

3.1.4 MODULA 2

MODULA 2 was developed in the late seventies by Nicholas Wirth who also developed the Pascal language. Wirth expanded Pascal with improved syntax, module structure, low level machine access and multi processing capabilities.

3.1.5 C

C was developed in 1974 for writing systems programs on the PDP11. It has been generalized and implemented on many computers large and small. C is a relatively low level language and hence is very efficient. This has the advantage that very little code needs to be optimised in assembler. The Unix operating system itself contains some 13000 lines of C source code. Only about 800 lines at the lowest level were required to be written in assembler. C is a very compact and portable language that has gained a lot of popularity, particularly in microcomputer applications.

Readability can be a problem with C because of the existence of too many ways to say the same thing. Also there is the capability of producing extremely terse code.

3.1.6 ADA

In the mid 1970's the United States Department of Defense (DoD) conducted studies that showed that enormous savings ($24 million between 1963-1999) could be achieved if DoD used one common language rather than the 450 different languages then used by its programmers. As a result of this study, the DoD instigated the development of a new programming language called ADA. ADA was designed to support numerical applications, systems programming and applications with real time and concurrent requirements. ADA is a large complex language which was developed using PASCAL as its starting point.

3.1.7 C++

Object oriented programming is the current fad in computer languages. Proclaimers of this "new way of looking at programming" say that programs written using the constructs of Object Orientated Programming System (OOPS) are much easier to maintain.

An OOPS program encapsulates the data types and their associated operations. By doing so it centralizes all access to the data type. Hence when a bug shows up, it can easily be associated with its data type and traced. C++ is an extension of the C language that allows the use of OOPS constructs. Other languages such as Smalltalk are totally OOPS languages.

3.2 Operating Systems

The choice of an operating platform for the development of the project had to be made carefully. An operating system is a translator that allows the user to communicate with the computer hardware. It controls such functions as loading and running software, reading and writing to disk, printers and display hardware. For good economic principles it was decided to develop the system for an IBM microcomputer environment. At the commencement of the project the following operating systems were available for this platform.

3.2.1 DOS

DOS (Disk Operating System) was initially developed in 1981. Since then it has evolved (through versions 1.0 to 4.0) along with the evolution of the personal computer from 8 bit to 32 bit machines. It is now the major operating system of the IBM compatible microcomputer industry.

It is a single tasking, single user system. This means that only one program can be run at any one time and only one user can use the processor at one time. With the earlier XT (8 bit) computers this limitation was quite reasonable as this was all that could be expected of the processor chip. However, with the introduction of the AT (16 bit) machines and the 386 (32 bit) machines, this limitation is now due only to the inadequacies of the DOS operating system.

In fact DOS does not take advantage of many of the features of the new Intel 80286 and 80386 chips, simply using them as very fast versions of the 8 bit 8086 chip.

DOS does not have a graphical user interface. All communication with the user is via a command line interface.

3.2.2 OS/2

In some ways OS/2 can be viewed as the next upgrade of the DOS operating system, albeit a very substantial upgrade. OS/2 uses the full power of the Intel 80286 and 80386 chips to provide a system capable of full multitasking, disk management, memory management and many other features. It comes with a standard user interface and provides device independent graphic drivers.

There is still some controversy regarding which system (OS/2 or a Unix derivative) will become the dominant operating system for microcomputers in the 1990's, Despite this, it seems fairly certain that OS/2 has a big future ahead of it.

However, while OS/2 is still evolving (having been officially released in 1988) there are several disadvantages in developing applications for it immediately. It requires a PC with at least 4 megabytes of RAM. This increases the hardware costs for both development and use of a system. Also, being such a new product, it is sure to evolve very rapidly during the next few years. Finally, at this stage there are very few applications available for OS/2. Hence it may be a case of constantly switching back and forward between DOS and OS/2 until OS/2 gains acceptance.

3.2.3 Microsoft Windows

Microsoft Windows is an extension of the DOS operating system. It gives DOS a graphical user interface and a simple method of multitasking. Applications written to run under Windows must be written specifically for this environment to take advantage of these functions.

There are many advantages to a developer in writing his application in the Windows environment. First, all Windows programs have a common interface and appearance. Hence a user who has used any other Windows program will easily be able to learn how to use another. Secondly, Windows allows the developer to write his application for a standard printer and standard monitor. Windows contains the appropriate device drivers to allow this standard output to be translated to the appropriate device. Hence, when a new printer is released, Microsoft will write a new driver to read the standard Windows output.

This advantage is highlighted by the fact that an application can be developed today and in two years time use a printer just released on the market. It makes more sense for Microsoft to develop this new driver once, than for each individual developer to have to modify code every time a new product is released.

The multitasking feature of Windows, although not as sophisticated as that of OS/2 does allow the running of several programs concurrently under DOS.

3.2.4 Xenix/Unix

Xenix is Microsoft's version of the Unix operating system for the PC. Unix was first introduced in 1969 and was the first interactive multitasking operating system. Since its introduction, it has gained widespread commercial acceptance (in main frame applications). The main hindrance in developing a Unix standard for the microcomputer is in defining a standard for the Unix system itself. At present there are several different strains of the Unix operating system.

Xenix has been available for the PC for some time. However, its drawback has been that the early microcomputers have not had the power to make use of many of the features of a Unix based system.

Also, until recently, Unix did not have a recognised graphical interface. This has been partly solved with the standardization of X-Windows and Motif. If these problems can be overcome, then a Unix based operating system may well rival OS/2 as the operating system of the 1990's.

3.3 Programming Decisions

A decision was made to develop the system under the Windows environment for the following reasons.

- It will run on present day 80286 and 80386 machines with less than 1 Mbyte of RAM (unlike OS/2 which requires a minimum of 4 Mbyte).
- It provides a consistent user interface with which users of other Windows products will be immediately familiar. Hence it is easier to learn and use.
- It provides a level of multitasking. This allows other Windows products to be used in conjunction with the navigation system.
- It provides a degree of hardware independence. Once the software is written for a standard device, it can be used with all devices supported by the Windows device drivers. This independence extends not only to video display and printers but also to mouse, keyboard, system timer and RS-232 ports.
- For developers, Windows provides a large library of built in routines that allow easy implementation of menus, dialogue boxes, scroll bars and other components of a user friendly interface.
- Windows provides an efficient system of memory management and use of RAM. To use more than 640K of RAM in a DOS program creates many problems. Windows provides this capability in a standardized manner.
- Windows development has many similarities with development under OS/2. Hence, the upgrade from Windows to OS/2 will not be as dramatic as coming directly from a DOS environment.

After deciding to use the Windows environment, the decision regarding the choice of language was limited by the fact that Windows supports only three languages at present. These three languages are C, Pascal and Assembler. Programming in Assembler was not considered feasible as even a medium size program would be very tedious to code and near impossible to maintain.

Both C and Pascal have excellent compilers available for microcomputers. However, C is considered the better choice due to its general popularity in the commercial microcomputer industry. The Windows libraries themselves are written in C. This popularity ensures that programmers will be available for future system maintenance.

The nature of the C language allows it to be very structured or very terse depending on how it is implemented. The language also contains efficient methods of handling exceptions or errors. As the C language has been standardised, subroutines written during the development of this system will be able to be re-used in other programs and on other machines in the future. This portability is one of the key advantages of the C language. Finally, as C is a relatively low level language, very little of the code will require optimizing in Assembler.

It should be stated that the C language is a relatively low level language and was not designed for developing extremely large systems. However, the use of the Windows library of functions gives C more of the power of a high level language.

Even so, languages such as Modula 2 and ADA have much to offer, particularly when systems will become large and require considerable maintenance. These programs were designed to meet the problems encountered in large systems maintained by many people. However, PC based compilers for ADA and Modula 2 do not provide the same level of hardware support as provided by the Windows environment.

Whether or not OOPS languages will live up to their promise is not known. However, it was considered somewhat premature to attempt to develop a system using one of the early C++ compilers available today.

4. HARDWARE CONSIDERATIONS

The aim of this section is to define the hardware requirements of the system and discuss the options available to meet these requirements.

4.1 Computer

The computer that will run the system is a key item in its success. The system requirements of this item are listed as follows:

Ruggedness:

The hardware must be well packaged and must be strong enough to endure the constant mobilization and demobilization to different sites. Also, it will often be in moving vessels such as ships and planes. Consequently, it must be resilient to a large degree of movement and vibration.

Portability:

The smaller and more portable the computer the better suited will it be to this type of system.

Economy:

The cost of the unit is an important consideration in the development of such a system. Particularly as, due to the vulnerability of a PC based system to environmental factors, it will generally be necessary to have a back-up computer available.

Communications Ports:

As a minimum, the system must receive data from the GPS receiver and transmit/receive data from the modem. Also the ability to output to an external log in real time is considered important. Hence, a minimum of three communication ports is considered essential for an effective system.

Graphics Capability:

In a modern system the use of a graphical interface is considered highly important. Hence, a screen with good graphics resolution and colour is essential. However, present technology does not provide colour screens in portable computers. Most laptop and portable computers contain some form of liquid crystal display (LCD) or gas plasma display. Advanced displays are usually back-lit for better visibility.

These displays can show shades of grey but not colour. However, if the display card in the computer supports colour, an external monitor can be connected to these machines when colour is required. Therefore, regardless of the monitor type, the type of graphics card is very important.

There are many graphics standards in the PC industry. The four major standards in the evolution of graphics capabilities are the MDA (Monochrome Display Adaptor), CGA (Colour Graphics Adaptor), EGA (Enhanced Graphics Adapter) and VGA (Video Graphics Adaptor). Each of these cards has a higher resolution and more colours available than the previous one. The evolution of display adaptors has gone hand in hand with the development of microprocessors. Naturally, faster CPU's (Central Processing Units) are required to control these more complex displays.

Power

It would be convenient if the system could be run from battery power (12-24V DC). However, the power requirements needed to run back-lit screens and hard disks necessitate the unit being run only on 120 or 240V AC power.

Expansion capability

The number of serial ports, memory cards or bus mouse cards that can be put into a computer is limited by the number of slots for these cards available on the motherboard. The amount of available slots is sometimes very limited, particularly with portable computers.

Microprocessor

The Central Processing Unit (CPU) dictates the ultimate capacity and speed of the computer. In fact most PC's can be classified by the types of processors they use. The first generation of PC's (XT's) used an Intel 8088, or similar processor. This was an 8 bit processor, meaning that it handled 8 data bits at a time.

The next generation introduced by IBM was the PC/AT which contained an Intel 80286 chip. This was a 16 bit processor, so it could handle twice as much information at a time. In 1987 Intel released the 80386 chip which is a 32 bit chip and hence faster again.

Another factor that affects the processing power of the computer is the operating speed of the chip. The original XT (8088 based) chips usually ran at 4.77 MHz. The original IBM AT machines ran at a clock speed of 6 MHz. However, manufacturers were able to increase this speed to around 12 MHz. At present it is possible to purchase 80386 machines that run at speeds of up to 33 MHz.

The next generation of silicon from Intel (the 80486) is now commercially available. This is still a 32 bit chip but has been redesigned with features such as RAM caching built into the chip. Hence, it offers a significant improvement in processing power over even the fastest 386 machines. It is interesting to note that the cost of these machines reflects their power, with the new 80486 machines being almost double the cost of the 80386 machines. This trend, more or less, continues down the ladder.

Another chip deserving mention is the 80386SX. This is an adaptation of the 80386 chip that has most of the functions of a full 80386. However, while the chip itself has 32 bit pathways, it fits into a 16 bit bus. Consequently, it has all the functionality of a 386 machine but is somewhat slower and considerably cheaper.

System Memory

Most IBM XT and AT type machines contain a minimum of 640KBytes of RAM. For programs running under DOS' this is all the memory that can be used and hence all that is required. Programs running under Windows can access any additional RAM present in the system. However, due to the current high cost of RAM (up to $1000 per megabyte) it was proposed to design the system so that it could take advantage of, but not require, additional RAM.

During the development of the code, it transpired that additional memory was required in order to use the debugging facilities provided with the Windows Development Kit.

Disk Storage

Due to typical high recording rates required in dynamic applications, a large amount of disk space is considered highly advantageous if logging is required.

4.2 Computer Selection

In mid 1989 when the hardware needed to be purchased, there were very few powerful portable computers available. The only full 386 machine available was manufactured by Compaq. However, this machine did not have a standard graphics card built in and therefore required a piggyback box to take an additional graphics EGA or VGA graphics card. This extra box detracted both from the machine's portability and its ruggedness.

Toshiba had a very good 286 portable available. However this machine was designed in the "laptop" mould and had very little room for expansion cards essential for extra I/O facilities.

Finally, the NEC 386SX portable was chosen as the one closest to fulfilling the requirements for the initial development and testing of the system. Table 11 contains the specifications for this unit. An additional Intel Above Board was also purchased to give the system an extra 2 megabytes of RAM. This extra RAM was required only for development of the system.

SPECIFICATIONS

CPU 80386SX

●	Word Length 16-bit	
●	Clock Rate 16MHz	
●	Wait States Zero	
●	Options	80387SX Maths Co-Processor

Random Access Memory

●	Standard	2Mb
●	Maximum	16Mb
●	Control	4 way page interleave

Read Only Memory

●	System	32Kb
●	BIOS	32Kb
●	Sockets	Two available for user definition

Keyboard

●	Keys	93
●	Function keys	12
●	Separate numeric keypad	
●	Separate cursor diamond	

I/O Facilities

●	Printer Port	
●	Serial Port RS232C asynchronous full/half duplex	
●	Keyboard	
●	External RGB monitor	

Expansion Slots

●	3 slots in total	3 x 8/16-bit full length

Display

●	Type	Plasma
●	Resolution	640 x 480
●	Grey Scales	16
●	Modes	VGA, EGA and CGA

External Monitors

●	Multisync II	Colour VGA compatible Maximum resolution 800 x 560
●	Multisync XL	Colour Maximum resolution 1024 x 768 19" diagonal

Table continued ......

....Continued from page 63

Hard Disc Drive	
• Formatted size	42Mb
• Access Time	28ms
• Physical Size	3.5" ST506 interface
Dimensions	
• Height	286mm
• Width	400mm
• Depth	195mm
Weight	
• Metric	9.6Kg
• Imperial	21 lbs
Power Supply	
• Voltage	240/120 volt switchable
• Wattage	130 watts

Table 11: Specifications for NEC 386SX Portable

These machines performed very well during field work in October/November 1989. However, when data is being logged to disk during a survey it was found to slow down the system significantly. Already, in 1990, full 386 portable machines are appearing on the market. The American manufacturer, Dolch, have now released a 486 portable computer. This highlights the fact that microcomputer hardware technology is improving at a very rapid rate reinforcing the need for software not to be tied to any one hardware configuration. All testing and implementation of the system were initially carried out using the NEC386SX portables. During 1990, Dolch units will be used where the extra processing speed is required.

4.3 Modems

A [6]Modem is required to decode the digital information to be sent into an analogue form that can be transmitted via a radio link.

---

[6] MODEM MOdulator DEModulator

The type of modem required is dependant on the medium used to send the information. Hence a different type of modem is required for sending data using a HF radio than for sending information through a telephone line. This is due to the different characteristics of these two forms of communication.

Some of the methods of data communication and their characteristics are summarized as follows:

4.3.1 HF Radio

Radio communication using high frequency radio signals offers a method of sending data very long distances. With good equipment it is possible to send data several thousand kilometres using HF signals. This transfer of data around the curved earth is achieved by bouncing the HF signals on the ionospheric layer. As the height of the layer varies considerably with time, this method of communication is somewhat unpredictable.

The ionosphere is constantly changing depending on the time of day, the time of year, location and local sunspot activity. Consequently, data communication between two points may be excellent at a certain time and some hours later be very much degraded. To overcome the problems associated with this dynamic atmosphere it is highly advantageous to have a wide variety of transmission frequencies available. Different frequencies have different refractive properties. It is therefore quite normal with HF communications for a certain frequency to be used quite well during the day time, but a much lower frequency being required to get the same communication at night.

It is possible to obtain radios that can transmit and receive on a vary large number of frequencies. However, the difficulty is not in obtaining such radios, but in obtaining permission for their use. An appropriate licence is required to be issued for each frequency to be used. Further, a separate license is required for transmission of data as opposed to normal voice transmission. Finally, it should be noted that obtaining appropriate licenses for transmission of data on a broad spectrum of frequencies can take several months to organise.

4.3.2 UHF/VHF Radio

Both UHF and VHF radios require line of sight for successful operation. As the transmission is direct, it is much more reliable than HF frequencies. Consequently, modems using UHF/VHF transmissions can be maintained at baud rates of up to 1200 baud. The same modem can be used for both HF frequency and UHF/VHF transmissions without any modifications. However, usually an EPROM (Erasable Program Read Only Memory) upgrade is required to change the baud rate to the maximum allowable for the transmission type.

4.3.3 Satellite

Communication satellites such as Aussat offer a very reliable method of transmitting GPS data between sites. The method is more reliable than HF radio as the transmissions, both to and from the satellite, are line of sight. At this stage, the principal disadvantage of this type of data link is simply the cost.

4.3.4 Hayes Compatible

During implementation and initial testing of the system, subroutines were also developed to allow the corrections to be sent between receivers, using a standard Hayes Compatible modem. Also code was written to allow the receivers to be connected back to back for testing of the system without the use of a modem.

4.3.5 Modem Selection

After inspection of several modems, it was found that the maximum obtainable baud rate, for HF radio transmission, was 300 baud. Two modem types were selected for comparison. These were from Transcom and Radtech, both Western Australian companies.

Both these modems have the ability to build up a data message from a number of transmissions. That is, if a signal is received in its entirety, an acknowledgement of this receipt is sent by the receiving modem. If only part of the signal is received, the valid part of this data is kept and from subsequent transmissions this data message is built up. Both modems perform this operation using their own proprietary parity checking algorithms.

It should be noted that this internal software in both modems is designed primarily for the transmission of ASCII 7 bit data. A control structure is used whereby certain bytes are recognized by the modem as control codes telling the modem to perform a particular function. Hence, pure binary data cannot be sent in a straight forward manner as some of the data would be misinterpreted as control codes.

To overcome this problem, all binary data is packed into 6 bit words prior to transmission. Data is then unpacked upon receipt by the receiving modem. This compaction of the binary information into 6 bit bytes increases the length of the data string by 33%. In other words, every 3 bytes becomes 4 bytes. This method is still more efficient than sending the data as pure ASCII numbers.

Transcom modems were eventually selected for the initial implementation of the system. This was primarily due to their smaller size but, in general, they appeared to be a more advanced product.

4.4 GPS Receiver

The overall philosophy of the system is to be independent of the receiver type and therefore support a number of receiver formats. This independence is achieved by writing any device dependant code in discrete modules. Hence, only this module need be changed or selected for use with another receiver type.

The type of receiver most suited to real time applications must be investigated with the following points in mind:

- The maximum number of satellites that the receiver can track is very important. If the receiver cannot track all the satellites in view, then the decision as to which combination of satellites is the best to track is very subjective. Software to handle this problem is a project in itself. Also, by not tracking all satellites in view , a receiver is wasting good data.

- The minimum recording rate that the receiver can output is critical, particularly for high dynamic applications such as in aircraft. Even in low dynamic applications such as at sea, a one second update rate is considered a minimal requirement.

- The amount of external control the operator has on the receiver is important. An operator should not have to learn both a software package and how to use a receiver. If the software can control all the functions of the receiver directly, then this makes the system much safer and easier to use.

- The resistance to dynamics of the receiver tracking loop must be considered. The receiver hardware must be capable of performing under high dynamic situations.

- The ruggedness, portability and operating specifications of the unit are also very important as the system will be mobilized quite frequently.

- It is highly advantageous if the receiver can log a large amount of data to solid state memory. This provides an alternative method of logging the raw data.

- It must be possible to obtain the complete raw data and ephemeris from the receiver in real time.

In the second half of 1988 when GPS receivers were being evaluated for this project, very few receivers offered the ability to track an optimum number of satellites. Consequently, the Trimble SL receivers were being used for all off-shore projects by ASI (Associated Surveys International) (Castalanelli et al 1988). The arrival on the market of the Ashtech receiver offered a marked improvement in receiver capability. Hence Ashtech receivers were purchased for the initial implementation of the navigation system.

At this stage, neither the Ashtech XII nor the Trimble ST had implemented a method of real time output. However, Ashtech were very helpful in providing the required output of raw data. A beta version of suitable EPROMs was obtained to allow this output facility in September of 1989. Prior to this, development was carried out using data downloaded from the receiver.

5. ALGORITHM DESIGN

5.1 Introduction

The following description of the algorithms used in the system are presented in the logical order of the growth of the system. When developing a system, such as the one described in this paper, it is prudent to break the development into clear stages. If possible it is also desirable to be able to access the success or otherwise of the development at each of these stages.

The following stages were proposed for the growth of the real time navigation system.

Stage 1: Pseudo range solution

Aim: To be able to take the raw data from a receiver and form the pseudo range solution in an identical manner to the internal processing of the GPS receiver.

Steps Required: In the formation of the pseudo range solution the following factors must be addressed.

- Satellite Clock Corrections
- Satellite Coordinate Calculation
- Receiver Clock Corrections
- Group Delay Correction
- Relativity Correction
- Millisecond Ambiguity
- Transmission Time.

Checks: This stage has an excellent check in that it should be possible to obtain an identical solution to that being calculated by the internal GPS software.

Stage 2: Carrier phase smoothing

Aim: To use the carrier phase/integrated Doppler measurements to smooth the pseudo range solution.

Steps required: Calculation of initial bias and checks for certain types of cycle slips.

Checks: The first and most obvious check is that the smoothed solution is in fact smooth and stable over short periods of time. The second check (more for the quality of the refraction corrections) is that the mean intial bias is reasonably stable.

Stage 3: Forming and passing corrections

Aim: To calculate corrections at the base site and pass them to the roving site for application.

Steps required: Form range corrections, Propogation of variances, Weightings and variance, Quality control statistics.

Checks: Back to back tests with two units over test baseline.

5.2 Pseudo range solution

The first stage of the development of the project was to extract the raw data from the GPS receiver and recalculate the point position solution using the raw code phase. This section highlights the algorithms used for this computation with emphasis on the corrections being applied.

5.2.1 Calculation of GPS Time

In the calculation of the position of the GPS satellite coordinates, the GPS time of the transmission of the GPS signal must be known. Usually, the time of the pseudo range measurement is made using the receiver oscillator. This time must then be corrected for a number of factors in order to obtain the true GPS time of transmission. Some of these corrections require an approximate knowledge of the satellite coordinates. Hence the operation is an iterative one whereby approximate values (for satellite coordinates, clocks and user positions) are used to generate corrections to obtain better approximate values.

The following formulae show the corrections applied to the code phase time to obtain GPS time. They include corrections for the satellite clock, relativistic effects and group delay. An explanation of each of these corrections is given below.

5.2.1.1 Transmission delay

In the reduction of the user position, the location of the satellite is required at the time that the signal leaves the satellite (Transmit Time) and not at the time it reaches the receiver (Received Time). Since it is usual for the time of the pseudo range measurement to be made using the receiver oscillator, this time must be corrected back to its time of transmission.

The GPS signal travels at the speed of light (c) for most of its journey and hence this time delay is quite simple to calculate (table 12).

$$\boxed{t_{TRANSMIT} = t_{RECEIVED} - p/c} \quad [7]$$

where

p = Pseudo range distance (metres)
c = Speed of light (metres / sec)
$t_{TRANSMIT}$ = Transmit time of signal
$t_{RECEIVED}$ = Receive time of signal

Table 12: Correction for transmission time

Errors in the receiver clock do not affect this calculation of transmit time as this error is also present in the pseudo range and therefore cancels in equation 7.

Errors in this calculation due to the atmospheric effects and noise in the pseudo range measurement do not significantly effect the subsequent calculation of satellite coordinates. This is because the maximum speed of a GPS satellite is less than 4000 metres/sec which is small compared to the speed of light (approximately

300,000,000 metres/sec). Hence, any error in the pseudo range measurement translates into an equivalent satellite coordinate error multiplied by (4000/300,000,000) or 1/75000.

5.2.1.2 Satellite Clock

The satellite clocks are offset from GPS time by a small amount. This error is monitored by the control segment and corrections are uploaded into the GPS satellites. The control segment ensures that this correction is always less than one millisecond. The correction can be calculated from the values contained in the broadcast message using the appropriate formula supplied in the ICD-GPS-200. (See table 13).

$$\boxed{\Delta t_{SV} = a_{f0} + a_{f1}\,(t-t_{oc}) + a_{f2}\,(t - t_{oc})^2} \quad [8]$$

where

Δt_{SV} = Satellite clock corrrection
t = GPS system time (seconds)
a_{f0} = 1st coefficient of clock offset
a_{f1} = 2nd coefficient of clock offset
a_{f2} = 3rd coefficient of clock offset
t_{oc} = reference time of clock parameters

Table 13: Satellite Clock Correction

5.2.1.3 Relativity

Apart from the clock errors caused by the satellite oscillator, there are also shifts caused by the effects of both special and general relativity. This is due to the difference in gravitational potential between the satellite and the user and also the difference in velocity of the users. These effects are compensated in two ways. The majority of the effect is corrected by setting the satellite clock frequency slightly lower by 4.45×10^{-3} Hz. The satellite signal speeds up as it approaches the earth so that if there were no apparent Doppler or atmospheric effects the observer would receive the signal at the design frequency (1575.42MHz for L1).

A second correction is made to account for the periodic components of the effects of relativity. The formula used to determine the magnitude of this correction is contained in the ICD-GPS-200 (1984), p68 and is shown in table 14. This correction accounts for the effects of the eccentricity of the satellite orbit.

$$\boxed{\Delta t_r = F\ e\ (A)^{\frac{1}{2}} \sin\ (E_k)} \quad [9]$$

where

t = $t_{SV} - \Delta t_r$
F = $-4.442809305 \times 10^{-10}$ sec/(metre)$^{\frac{1}{2}}$
e = eccentricity
$A^{\frac{1}{2}}$ = Square root of semi major axis
E_k = Eccentric anomaly

Table 14: Relativity Correction

5.2.1.4 Group Delay

The clock correction coefficients contained in the broadcast message were calculated using measurements on both L1 and L2 frequencies. Hence, they more specifically apply to measurements obtained using dual frequency measurements. To allow the single frequency user to corrrect for this assumption, a value is contained in the broadcast ephemeris to account for this slight difference. This term, (t_{GD}), is applied to all L1 only measurements. A function of this term is used to correct for measurements made using L2 signals only (ICD-GPS-200 1984 p69). Quite often this term is zero for all satellites. The correction is applied as shown in table 15.

$$\boxed{(\Delta t_{SV})_{L1} = \Delta t_{SV} - T_{GD}} \quad [10]$$

where

$\Delta(t_{SV})_{L1}$ = clock correction for L1 measurements

Δt_{SV} = satellite clock correction derived using coefficients in broadcast message

T_{GD} = Group Delay (secs) from Navigation message

Table 15: Correction for Group Delay

5.2.1.5 Millisecond Ambiguity

It was stated on page 24 that the C/A code has a period of one millisecond (approximately 300km). Hence, when dealing with measurements of pseudo range, using the C/A code, the missing number of one millisecond epochs needs to be determined. It should be emphasized that this problem does not exist when using the P code. The P code has a period of one week and hence provides the complete pseudo range measurement.

In order to solve this 300km ambiguity it is necessary to have an approximate knowledge of the coordinates of the user. If the position of the user is known to within 100km then this ambiguity can be solved using the method shown in table 16. In this algorithm it is assumed that approximate values for both the receiver position and receiver clock offset are known. The satellite coordinates and clock errors are initially estimated using the received time as a first estimate of transmit time.

An alternative to using the received time as the first estimate of transmit time is to use $t_{TRANSMIT} = t_{RECEIVED} - 0.075$ as the first approximation. (Remondi 1984) This value is chosen as the time of transmission for a GPS signal varies between 0.067 seconds and 0.090 seconds.

Approximate parameters.

Xr,Yr,Zr	Approximate Receiver Coordinates
Δt_r	Approximate Receiver Clock Correction
$t_{TRANSMIT}$	Time of Transmission of GPS signal. initially set to equal $t_{RECEIPT}$

Measurements

$t_{RECEIPT}$ = time of receipt of satellite signal

p = measured pseudorange (truncated to the nearest millisecond)

Method of Calculation

Xs,Ys,Zs = f(Ephem,$t_{TRANSMIT}$)

Δt_s = f(Ephem,$t_{TRANSMIT}$)

range = $\sqrt{[(X_s-X_r)^2+(Y_s-Y_r)^2+(Z_s-Z_r)^2]}$

frange = $p + \Delta t_s - \Delta t_r$

N = int((Range - frange)/one_millisec+0.5)

mrange = N * one_millisecond + frange

$t_{TRANSMIT}$ = $t_{RECEIPT} - (\text{mrange}+\Delta t_r)/c$

where

Δt_r	Receiver Clock Correction
Δt_s	Satellite Clock Correction
$t_{TRANSMIT}$	Time of Transmission of GPS signal.
$t_{RECEIVED}$	Time that Signal is received
X_s,Y_s,Z_s	Satellite coordinates
X_r,Y_r,Z_r	Receiver Coordinates
range	True range between SV and receiver
frange	measured range modulo one millisecond
mrange	measured range between SV and receiver
one_millisec	Distance travelled by light in one millisecond

Table 16: Determination of Millisecond Ambiguity

Because of this characteristic of the C/A code, most early GPS receivers required the user to enter an approximate location into the receiver prior to commencing a survey. However, modern receiver technology (the new Trimble ST and the Ashtech XII) use the modulation of the navigation message to resolve this ambiguity (Ashjaee 1988). The navigation message has a period of 30 seconds which translates into a range ambiguity of 9,000,000km. Consequently, the user need know his position only to the nearest 4,500,000km which more than covers anywhere on the earth.

This method of determining the millisecond ambiguity using the navigation message is contained in the internal software of the receiver. As a result the receivers mentioned do not require any user input prior to acquiring lock on to satellites and providing position output. These receivers compute and output the computed position along with the raw data. Consequently, this position can also be used as an initial position if required.

5.2.2 Calculation of Satellite Coordinates

The position of a GPS satellite can be calculated for a particular GPS time using the information contained in the broadcast ephemeris. The formulae, shown in table 17, (ICD-GPS-200 p78) are used to calculate the Earth Centred Earth Fixed (ECEF) coordinates of a GPS satellite.

$$
\begin{aligned}
\mu &= 3.986008 \times 10^{14}\ \text{metres}^3/\text{sec}^2 \\
\Omega &= 7.292115147 \times 10^{-5}\ \text{rad/sec} \\
A &= (\sqrt{A})^2 \\
n_o &= \sqrt{\mu/A^3} \\
t_k &= t - t_{oe} \\
n &= n_o + \Delta n \\
M_k &= E_k - e \sin E_k \\
\cos v_k &= (\cos E_k - e)/(1 - e \cos E_k) \\
\sin v_k &= \sqrt{(1 - e^2)} \sin E_k \;/\; (1 - e \cos E_k) \\
\phi_k &= v_k + w \\
\delta u_k &= C_{us} \sin 2\phi_k + C_{uc} \cos 2\phi_k \\
\delta r_k &= C_{rc} \cos 2\phi_k + C_{rs} \sin 2\phi_k \\
\delta i_k &= C_{ic} \cos 2\phi_k + C_{is} \sin 2\phi_k \\
u_k &= \phi_k + \delta u_k \\
r_k &= A\ (1 - e \cos E_k) + \delta r_k \\
i_k &= i_o + \delta i_k \\
x_k' &= r_k \cos u_k \\
y_k' &= r_k \sin u_k \\
\Omega_k &= \Omega_o + (\Omega\text{dot} - \Omega\text{dot}_e) - \Omega\text{dot}_e\, t_{oe} \\
x_k &= x_k' \cos \Omega_k - y_k' \cos i_k \sin \Omega_k \\
y_k &= x_k' \sin \Omega_k + y_k' \cos i_k \cos \Omega_k \\
z_k &= y_k' \sin i_k
\end{aligned}
$$

Table 17: Formulae for calculation of Satellite Coordinates

The Keplarian elements used in the above formulae are obtained from the broadcast ephemeris. A summary of these parameters is given in table 18. Table 19 contains the clock parameters contained in the broadcast message.

M_0	Mean Anomaly at reference time
Δn	Mean Motion difference from computed value
e	Eccentricity
$(A)^{1/2}$	Square root of semi major axis of orbit
Ω_0	Right Ascension
i_0	Inclination angle at reference time
w	argument of perigee
Ωdot	Rate of right ascension
idot	Rate of inclination
C_{uc}	Correction term to argument of latitude
C_{us}	"
C_{rc}	Correction term to argument of orbit radius
C_{rs}	"
C_{ic}	Correction term to inclination of orbit
C_{is}	"
toe	Reference time of ephemeris
AODE	Age of data ephemeris

Table 18: Ephemeris parameters

Week Number	No of weeks since 6th January 1980
SV accuracy	
SV Health	Space Vehicle Health
T_{GD}	Group delay
AODC	Age of Clock data
t_{oc}	Time of Clock corrections
a_{f2}	Polynomial Coefficients for SV Clock
a_{f1}	
a_{f0}	

Table 19: Clock parameters

Table 20 gives an example of the program code used to calculate the satellite coordinates and the clock corrections.

```
/*
**    ______________________________________________________________
**    ##############################################################
**    earth_fixed(eph,sat,t)
**    Calculation of earth fixed coordinates from Broadcast
**    ephemeris parameters. (as per ICD-GPS-200 page 78)
**
**    Note: the SV clock correction is returned in the time
**    component of the satellite coordinate.
**
**    ##############################################################
**    Macro constants used:
**    """"""""""""""""""""""
**    rae    Earth Rotation Rate
**    Gme    Earths Universal Gravitation Constant
**    F      Relativistic constant
**
*/
earth_fixed(eph,sat,t)

struct nav     *eph;   /* EPHEMERIS USED FOR THIS CALCULATION        */
struct coord   *sat;   /* RETURNED EARTH FIXED COORDINATES           */
double         t       /* GPS TIME OF POSITION CALCULATION           */

{
/*
      *----------------*
      * Declarations
      *----------------*/
      double tolerence=5.0e-15;
      double A,n0,tk,n,Mk,Ek,corr,y,x,vk;
      double lk,duk,drk,dik,uk,rk,ik;
      double xko,yko,rak,zero=0.;
      double t_r, deltan,m0,omega0,i0,omega,idot;
      float omegadot;

      deltan        =eph->deltan       *pi;
      m0            =eph->m0           *pi;
      omega0        =eph->omega0       *pi;
      i0            =eph->i0           *pi;
      omega         =eph->omega        *pi;
      omegadot      =eph->omegadot     *pi;
      idot          =eph->idot         *pi;
/*
      *-----------------------------------------------------------*
      * Calculation of satellite coordinates as per p78 ICD200
      *-----------------------------------------------------------*/

      A      =eph->roota*eph->roota;      /* SEMI-MAJOR AXIS           */
      n0     =sqrt(Gme/(A*A*A));          /* COMPUTEDMEAN MOTION */
      tk     =t-(double)eph->toe;         /* TIME FROM EPH REF EPOCH*/
      n      =n0+(double)deltan;          /* CORRECTED MEAN MOTION */
      Mk     =(double)m0+n*tk;            /* MEAN ANOMALY              */
```

.....Table continued on next page

```
/*
        *-------------------------------------------------------------*
        * ITERATIVE SOLUTION OF KEPLER'S EQUATION
        * FOR ECCENTRIC ANOMALY
        *-------------------------------------------------------------*/
        Ek=Mk;
        do
        {
                corr=Ek-Mk-(eph->e)*sin(Ek);
                Ek=Ek-corr;
        } while ((rabs(corr))>tolerance);

                                        /* TRUE ANOMALY */
        y=(sqrt(1.0-eph->e*eph->e)*sin(Ek))/(1.0-eph->e*cos(Ek));
        x=(cos(Ek)-eph->e)/(1.0-eph->e*cos(Ek));
        vk=atan2(y,x);

        lk=vk+omega;                /* ARGUMENT OF LATITUDE */

                                    /* SECOND HARMONIC PERTURBATIONS*/
        duk=(double)eph->cus*sin(2*lk)+(double)eph->cuc*cos(2*lk);
        drk=(double)eph->crc*cos(2*lk)+(double)eph->crs*sin(2*lk);
        dik=(double)eph->cic*cos(2*lk)+(double)eph->cis*sin(2*lk);

        uk=lk+duk;                          /* CORRECTED ARGUMENT OF LATITUDE */
        rk=A*(1.0-eph->e*cos(Ek))+drk;          /* CORRECTED RADIUS */
        ik=i0+dik+(double)idot*tk;              /* CORRECTED INCLINATION*/

        xko=rk*cos(uk);                     /* POSITIONS IN ORBITAL PLANE */
        yko=rk*sin(uk);

                                    /* CORRECTED LONGITUDE OF ASCENDING NODE */
        rak=omega0+((double)omegadot-rae)*tk-rae*(double)eph->toe;

                                    /* EARTH FIXED COORDINATES */
        sat->x=xko*cos(rak)-yko*cos(ik)*sin(rak);
        sat->y=xko*sin(rak)+yko*cos(ik)*cos(rak);
        sat->z=yko*sin(ik);
/*
        *-------------------------------------------------------------*
        * Relativistic clock correction as per p68 ICD-GPS-200
        *-------------------------------------------------------------*/
        t_r=F*eph->e*eph->roota*sin(Ek);
/*
        *-------------------------------------------------------*
        * SV Clock offset as per p67 ICD-GPS-200
        *-------------------------------------------------------*/
        sat->t=eph->af0
                +eph->af1*(t-eph->toc)
                +eph->af2*(t-eph->toc)*(t-eph->toc)
                +t_r;

        return(0);
}
```

Table 20: Example of code to calculate ECEF Satellite Coordinates

5.2.3 Correction for Earth Rotation

The propagation time for the electromagnetic signal to travel from the GPS satellite to the receiver is in the range of 0.067 seconds to 0.090 seconds. During this time the earth has rotated approximately 36 metres at the equator. The satellite coordinates derived using the formulae in table 17 are earth centred and earth fixed. Hence, if the reference frame has moved, a correction must be applied to account for this movement.

A transformation must be applied to transform the ECEF coordinates of one epoch (at the time the signal left the satellite) to the epoch that the signal arrived at the receiver. As the Z axis of the ECEF system very nearly coincides with the rotational axis of the earth, this transformation is simply a rotation about this Z axis. (See Table 21.)

Rotation matrix

$$\begin{vmatrix} X \\ Y \\ Z \end{vmatrix} = \begin{vmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} \times \begin{vmatrix} X' \\ Y' \\ Z' \end{vmatrix}$$

therefore

$$dx = \cos\theta \times X' + \sin\theta \times Y' - X$$
$$dy = \sin\theta \times X' + \cos\theta \times Y' - Y$$

and

$$X = X' - dx$$
$$Y = Y' - dy$$
$$Z = Z'$$

where

θ = range/c × R_{ae} = Rotation angle
R_{ae} = Earth Rotation Rate = $7.2921151467*10^{-5}$m/s
range = true dist between satellite and receiver
XYZ = SV coordinates at transmission epoch
X'Y'Z' = SV coordinates at reception epoch
dx,dy = corrections for earth rotation

Table 21: Corrections for Earth Rotation

5.2.4 Position Solution

When simultaneous code phase measurements to four or more satellites are available, it is possible to determine the position of the GPS receiver. As it is required to solve for four parameters, three components of position coordinate and the receiver clock offset, then a minimum of four ranges is required to obtain a solution. When more than four ranges are available, the method of least squares is employed to determine the most likely position solution. The parametric method of least squares is described fully in White (1986) and Cross (1983).

In the following it is assumed that the position of the satellites can be calculated prior to the position computation. This is not strictly true as the calculation of satellite coordinates requires accurate GPS time which in turn requires a knowledge of the user position.

From the time of the observation it is possible to calculate the satellite coordinates. Each pseudo range measurement is an observation that can be related to the parameters required to be solved by equation 11 in table 22.

$$\boxed{R_i = \sqrt{[(X-X_i)^2 + (Y-Y_i)^2 + (Z-Z_i)^2]} + T} \qquad [11]$$

where

R_i	pseudorange to ith satellite
$x[X,Y,Z,T]$	the vector containing the receiver coordinates and clock offset.
X_i, Y_i, Z_i	the coordinates of the ith satellite
T	receiver clock offset

Table 22: Relationship between pseudorange and user position

However, the equation in this form is non-linear and difficult to solve directly. Hence linearization of the model is required. To linearize the model, it is necessary to have provisional values for our parameters of interest. These provisional (or approximate) values are related to the parameters as shown in vector form in table 23.

$$\boxed{x = x_0 + \Delta x} \quad [12]$$

where

$x = x\,[X,Y,Z,T]$ vector of parameters to be solved

$x_0 = x_0[X_0,Y_0,Z_0,T_0]$ vector of approximate coordinates

$\Delta x = \Delta x[\Delta x,\Delta y,\Delta z,\Delta t]$ vector of changes to parameters

Table 23: Provisional Coordinates

Equation 11 can be linearized using the first term of Taylors expansion as shown in table 24. Equation 13 can then be solved directly for the change in parameters $\Delta X[\Delta x,\Delta y,\Delta z]$.

Taylors Expansion

$$F[x_0 + \Delta x] = F[x_0] + dF/dx \;.\; \Delta x + \text{higher order terms}$$

Pseudo range Observation Equation

$$\boxed{R_i = RO_i + dF_i/dx.\Delta x + dF_i/dy.\Delta y + dF_i/dz.\Delta z + \Delta t} \quad [13]$$

where

$$RO_i = \surd[(X_0-X_i)^2 + (Y_0-Y_i)^2 + (Z_0-Z_i)^2] + T$$

$$dF_i/dx = (X_0-X_i)/(R_0 - T_0)$$
$$dF_i/dy = (Y_0-Y_i)/(R_0 - T_0)$$
$$dF_i/dz = (Z_0-Z_i)/(R_0 - T_0)$$

Note: Other variables as shown in tables 22 and 23

Table 24: Taylors Expansion of pseudo range model

For n pseudo range measurements these observations can be represented in the standard matrix form for observation equations, $v=Bx+T$ (White 1986), for subsequent solution using parametric least squares techniques. (Table 25)

$$\begin{array}{ccccccc} V & = & B & \times x & + & T \\ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ .. \\ v_n \end{bmatrix} & & \begin{bmatrix} F1/dx & F1/dy & F1/dz & 1 \\ F2/dx & F2/dy & F2/dz & 1 \\ F3/dx & F3/dy & F3/dz & 1 \\ ..... & ..... & ..... & . \\ Fn/dx & Fn/dy & Fn/dz & 1 \end{bmatrix} & \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta t \end{bmatrix} & & \begin{bmatrix} RO_1 - R_1 \\ RO_2 - R_2 \\ RO_3 - R_3 \\ .... \\ RO_n - R_n \end{bmatrix} \end{array}$$

$$P = \begin{bmatrix} W1 & 0 & 0 & .. & 0 \\ 0 & W2 & 0 & .. & 0 \\ 0 & 0 & W3 & .. & 0 \\ .. & .. & .. & .. & .. \\ 0 & 0 & 0 & 0 & Wn \end{bmatrix}$$

where

v = $n \times 1$ vector of residuals
B = $n \times 4$ matrix of coefficients
x = 4×1 matrix of parameters
T = $n \times 1$ matrix of corrections
P = $n \times n$ weight matrix

Table 25: Parametric Adjustment

Normal Equations in the form $(B^TPB)x + B^TPT = 0$ can then be solved to yield new corrections to the parameter vector x[X,Y,Z,T]. The solution is then iterated until the dot product of the vector Δx becomes zero. Table 26 provides a summary of this iterative least squares scheme.

Step 1: update estimate of parameters
$x = x_0 + \Delta x$

Step 2: Form observation equations
$v = Bx + T$

Step 3: Solve normal equations for Δx
$\Delta x = -(B^TPB)^{-1} \times B^TPT$

Step 4: Repeat steps 1 to 3 until $\Delta x . \Delta x = 0$

Step 5: Solution $x = x_0$
Variance $\sigma_{xx} = (B^TPB)-1$

Table 26: Iterative Least Squares Process

5.2.5 Altitude Hold

As shown in the previous section, observations to four satellites are required to obtain a position solution. This allows the determination of the 3 components of position and the receiver clock offset. If however, some of these parameters are already known, then observations are not required to as many satellites. The most common case is when the approximate altitude of the receiver is known. For instance, a vessel at sea would have a reasonable knowledge of its ellipsoidal height (perhaps to better than 10 metres).

When this is the case and it is desired to take advantage of this prior knowledge, it is possible to solve for only the horizontal position and clock offset using just three satellites. Similarly, if a precise oscillator is used with the receiver, it may not be required to determine the clock offset. Here again a solution can be obtained using only three satellites, this time solving for position parameters only.

If a precise clock and altitude hold were used it would also be possible to determine the horizontal position using only two satellites. It should be noted that both these methods usually produce inferior results when compared to the use of four or more satellites. However, the advantage of these types of methods is that it extends the usable satellite window. For economic reasons this can be quite beneficial during the prototype constellation configuration and prior to the completion of the block II satellite launches.

There are several methods that can be used to constrain the altitude of the solution, thus decreasing the degrees of freedom in the overall adjustment.

Adding Constraint Equation

This method involves adding the height information as a constraint equation into the adjustment (White 1986) . In this situation the equation is added as shown in table 27. The matrix of coefficients (B) and the matrix of corrections (T) are both repartitioned to include the constraint equation. Also an extra dummy parameter, (k), is solved for in the adjustment.

$-v + Bx + T = 0$ Observation Equation

$Cx + F = 0$ Constraint Equation

New normal equations with constraint

$$\left[\begin{array}{c|c} B^TPB & C^T \\ \hline C & 0 \end{array}\right] \times \left[\begin{array}{c} x \\ \hline -k \end{array}\right] + \left[\begin{array}{c} B^TPT \\ \hline F \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline 0 \end{array}\right]$$

Table 27: Constraint Equations

This method has the advantage that it would rigourously restrain the height in the adjustment regardless of the number of additional observations included. Its disadvantage is that it adds an extra degree of complexity to the initially simple parametric solution. Also there may be times when it is advantageous to be able to weight the fixed altitude (or semi fixed altitude) equation.

Parameter Elimination

This method involves transforming the parameters being solved so that the observation equations relate the range measurements to local system coordinates, ie Easting, Northing and Height. In this method, the height parameter could simply be removed from the solution when required and three satellites used. However, the simplicity offered by this approach is far outweighed by the inconvenience of working in a local system or transforming between systems.

Observation Equation

It was opted to implement the following scheme for fixing the altitude of position solution in the real time program.

The height is constrained by adding an additional observation equation with appropriate values and weighting to constrain the height to the desired value. The following shows the derivation of the appropriate observation equation for this purpose.

The relationship between the spheroidal height and the ECEF coordinates for a point is nonlinear as shown in table 28. (Bomford 1980).

$$\boxed{H = \sqrt{(X^2+Y^2)}/\cos\phi \quad - v} \quad [14]$$

where

$v = a(1-e^2\sin 2\phi)$
$\tan\phi = (Z + \epsilon\, b \sin^3 u)/(p - e^2\, a \cos^3 u)$
$\tan u = (Z/p)\,(a/b)$
$p = \sqrt{(X^2+Y^2)}$
$e2 = (a^2-b^2)/a^2$
$\epsilon = e^2\,(1-e^2)$
X, Y, Z = ECEF coordinates of position

Table 28: Non-linear height equation

In a similar fashion to the linearization of the observation equation for the satellite ranges, this equation can also be linearized using Taylor's theorem.

Using Taylors Theorem

$$H = f[X,Y,Z] = H_0 + dF/dx \,.\, \Delta x + dF/dy \,.\, \Delta y + dF/dz \,.\, \Delta z$$

where

H_0 is derived using the approximate coordinates of X,Y and Z with the formula in table 28.

dF/dx,dF/dy,dF/dz are the partial derivatives of the function with respect to X,Y and Z

Table 29: Linearization of equation

However, rather than obtain the values dF/dx,dF/dy and dF/dz by differentiating the complex formula in table 28, an approximation is used. The relationship between the height component and the parameters X,Y and Z can be approximated by the radius vector as shown in table 30.

Approximation

$$F[X,Y,Z] = R = \sqrt{(X^2 + Y^2 + Z^2)}$$

therefore

$$dF/dx = X/R$$
$$dF/dy = Y/R$$
$$dF/dz = Z/R$$

hence the observation equation becomes

$$H_{FIXED} = H_O + X/R \;.\; \Delta x + Y/R \;.\; \Delta y + Z/R \;.\; \Delta z$$

or in the form v = Bx + T

$$\boxed{v = X/R\ \Delta x + Y/R\ \Delta y + Z/R\ \Delta z + (H0 - H_{FIXED})} \quad [15]$$

Table 30: Observation Equation for Fixed Altitude

This type of observation equation is substituted for the fourth satellite when navigation is required using 3 satellites only. When only three satellites are being observed the weight given to this observation is inconsequential as the system of equations can be uniquely solved.

However, when four satellites are available and the known height information is included, the weighting given to this observation equation is important. Usually, this equation would only be used with four satellites if the geometry of these four satellites was poor. Hence in this case an appropriate weight must be given to the fixed altitude observation equation. The choice of the weighting of this additional observation is subjective when four or more observations are present.

5.2.6 Validation of algorithms

Following the implementation of the code to calculate the code phase solution it was necessary to test the performance of this code. The simplest means of testing this was to compare the calculated solution with that computed by the internal receiver hardware. Initially Trimble SX data were used for much of the development, with all validation carried out using Ashtech XIIL data. An agreement with the Ashtech receivers to the centimetre level was obtained.

Figure 13 shows a plot of the residuals obtained from subtracting the Ashtech receiver computed position and that computed with our own software. It can be seen that up to epoch 310 the two solutions agree only to the metre level. After epoch 310 the solutions agree to better than a centimetre. Figure 14 shows this agreement for the data after epoch 310 for the Easting residual. The Northing is very similar.

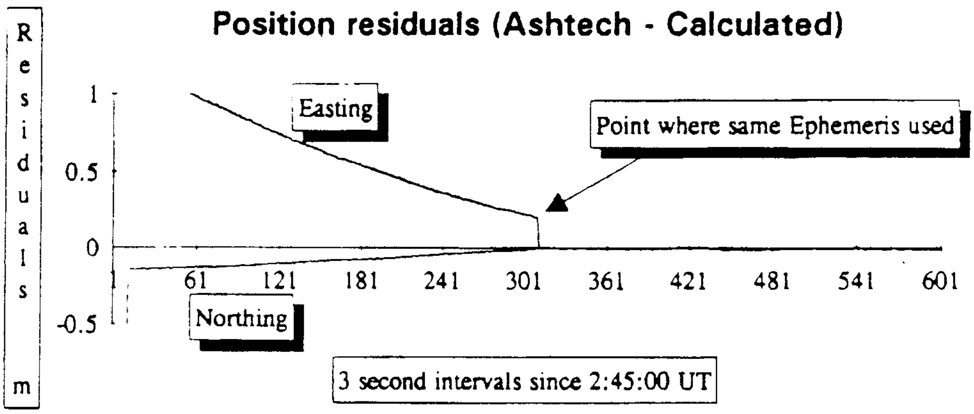


Figure 13: Position Residuals for Pseudo Range Position

Before epoch 310 the receiver is using a different ephemeris set in its position computation. After epoch 310 the same ephemeris set is being used in both calculations and hence the agreement is much better. The reason for the centimetre level difference is believed to be due to round off in the raw data measurements being output by the receiver.

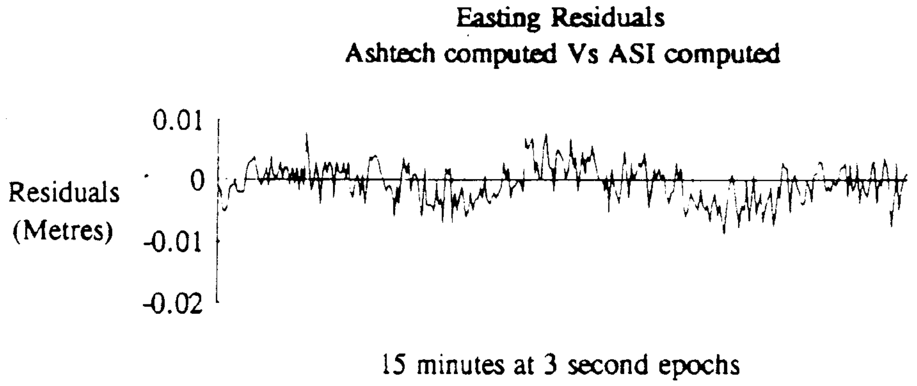


Figure 14: Position Residuals for Pseudo Range Solution

Consequently, this comparison with the Ashtech position validates the implementation of code used to obtain this solution. This is an important check as it effectively validates all routines and formulae for calculation of satellite coordinates, GPS time corrections and the resultant pseudo range solution.

5.3 Carrier phase smoothing

5.3.1 Introduction

The aim of the differencing technique is chiefly to cancel common biases. Errors caused by inaccurate orbit data and the effects of atmospheric refraction delays affect both stations in a similar way and hence can be cancelled by differencing. However, non correlated errors in these point position solutions will propagate.

In the raw pseudo range solution, the major random error is due to noise in the pseudo range measurements themselves. This noise varies between one and six metres, depending on the receiver type (Chisholm 1987). Assuming a pseudo range noise level of one metre and a PDOP of five, this could cause a position error in the order of five metres. The point position solution also may be biased by an orbit error of ten metres.

When the difference between the two stations is formed, the ten metre orbit error will effectively cancel. However, the five metre pseudo range noise will propagate and the differential solution may still be left with an error in the vicinity of five to ten metres. Hopefully, this highlights the need to eliminate the noise in the code phase measurements prior to differencing.

When carrier phase measurements are available, the simplest method to eliminate the code phase noise is to smooth it with the carrier phase measurements.

5.3.2 Initial Bias

The basis of this method, presented in Hatch (1982) is that the code phase measurements can be used to provide the initial bias missing from the carrier phase measurements. Once this initial bias has been determined, the more accurate carrier phase measurements can then be added to provide more precise range measurements.

The initial bias that needs to be determined is simply the true range between the satellite and the receiver at the commencement of the integration of the Doppler frequency. This bias can be estimated by subtracting the continuous carrier phase measurement from the code phase. As can be seen in equation 19 in table 31, the initial bias formed in this way is reasonably constant. It varies only with the change in ionospheric delay over time which is typically in the order of a few metres per hour. The initial bias also contains the clock errors and tropospheric errors present at the time integration started. However, just as the ranges are mapped back to this initial bias, so too are the clock and tropospheric errors. As these terms relate to a specific epoch, they are constant.

Hence, if this ionospheric error could be eliminated, the initial bias would be constant with a noise level equivalent to the noise level in the pseudo range measurements. As the code phase noise is some two orders of magnitude higher than the integrated Doppler, the noise contribution by the integrated Doppler is negligible.

$$\boxed{R = r + C + I + T}$$ Code Phase Equation [16]

where

R = Measured pseudo range
r = True range
C = Combined clock errors
I = Ionospheric errors
T = Tropospheric errors

$$\boxed{D = d + \Delta C - \Delta I + \Delta T}$$ Integrated Doppler [17]

where

D = measured integrated Doppler since t0
d = delta range measurement since t0
ΔC= Clock error since t0
ΔI= delta Ionospheric error since t0
ΔT= delta Tropospheric error since t0
t_0= time that Doppler integration started.

Initial Bias determination

$$B = R - D = (r - d) + (C - \Delta C) + (T - \Delta T) + (I + \Delta I)$$

therefore

$$\boxed{B = (b + C_0 + T_0 + I_0) + 2 \times \Delta I}$$ Initial Bias [18]

where

b = true range at time t0 $(b = r - d)$
C_0 = Clock bias at time t0 $(C_0 = C - \Delta C)$
T_0 = Tropospheric error at time t_0 $(T_0 = T - \Delta T)$
I_0 = Ionospheric error at time t_0 $(I_0 = i - \Delta I)$

Alternatively

$$\boxed{B = K + 2\Delta I}$$ Initial Bias [19]

where $K = b + C_0 + T_0 + I_0$ = constant

Table 31: Formation of Initial Bias Equation

If a set of n pseudo range and carrier phase measurements is taken to a satellite, at each of the n epochs a value for this initial bias can be obtained. By averaging these n values of initial bias a much more precise value can be determined. The variance of this new mean initial bias is inversely proportional to the number of observations averaged. (See table 33.)

Once an accurate value for this initial bias has been determined, a better estimate for the full range can be obtained by adding the integrated Doppler measurement to this initial bias at each epoch. This, in effect, provides an equation (equation 20) which is identical in form to the code phase equation (equation 16) except that is more precise.

$$S = B + D$$

$$S = (r - d + C - \Delta C + T - \Delta T + I + \Delta I) + (d + \Delta C - \Delta I + \Delta T)$$

$$\boxed{S = r + C + I + T} \quad \text{Smoothed Range} \quad [20]$$

All variables as defined in table 31

Table 32: Smoothed Range

This method of smoothing the code phase is elegant in its maximum use of both the carrier phase and the code phase measurements. In essence, the accurate carrier phase measurements are used to provide the change in range between epochs, while the code phase provides the absolute part of the range measurements.

Due to the nature of this method, the accuracy of the smoothed range increases with the number of measurements available for smoothing. Table 33 summarizes the propagation of the variances for the smoothed ranges under the assumption that there is no correlation between the two measurement types. While this assumption is incorrect, its effect is not considered consequential.

Variance of Initial Bias

From $B = R - D$

$$\sigma_{BB} = \sigma_{RR} + \sigma_{DD}$$

Variance of mean Initial Bias

From $B_{MEAN} = \sum B/n$

$$(\sigma_{BB})_{MEAN} = \sum \sigma_{BB}/n = \sum (\sigma_{DD} + \sigma_{RR})/n$$

Variance of Smoothed Range

From $S = B_{MEAN} + D$

$$\sigma_{SS} = \sum (\sigma_{DD} + \sigma_{RR})/n + \sigma_{DD}$$

Assuming σ_{RR} is significantly larger than σ_{DD}

$$\sigma_{SS} \approx \sum \sigma_{RR}/n$$

where

S = Smoothed Range
D = Integrated Doppler
R = Pseudo Range
B = Initial Bias
B_{MEAN} = Mean Initial Bias
σ_{SS} = variance of smoothed range
σ_{DD} = variance of Integrated Doppler measurement
σ_{RR} = variance of pseudorange
σ_{BB} = variance of initial bias
$(\sigma_{BB})_{MEAN}$ = variance of mean initial bias
n = number of measurements

Table 33: Variance of smoothed range

Figure 15 shows the initial bias determined for SV #03 over a twenty-five minute period between 10:45am and 11:40am on Julian day 320, 1989. From this graph, it can be seen that the spread of the initial bias is in the order of five metres. This directly reflects a noise level of similar magnitude in the original code phase measurements.

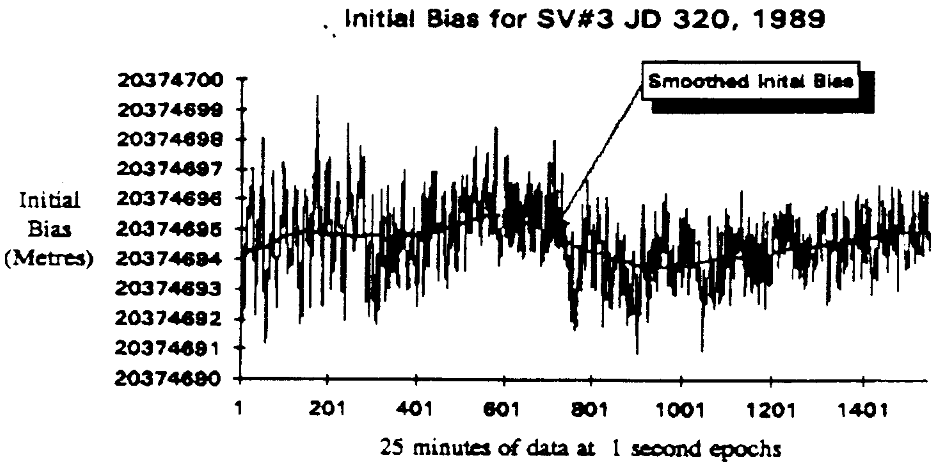


Figure 15: Intial Bias Plot

5.3.3 Effect of Ionosphere

The overall trend of the graph in figure 15 is due to the unmodelled effects of the ionosphere. As the signal travels through the ionosphere, its phase velocity is advanced thus decreasing the magnitude of the integrated Doppler and phase measurements. Simultaneously, the group phase, which contains the signal information is retarded which causes the code phase to become longer.

The ionosphere affects the two measurements differently and hence when the initial bias is formed by subtracting these two measurements, these ionospheric effects summate. (Equation 19)

During this 25 minute period the overall trend of the initial bias changes by several metres. Superimposed on figure 15 is a plot of the same initial bias but where each reading is obtained by taking the average of some three hundred continuous readings. It can be seen that this method smooths out the data without causing any significant bias.

Finally, figure 16 shows the smoothed initial bias plots for all the tracked satellites during this period. From this, it can be seen that the noise level of the smoothed initial bias is reduced to a few decimetres. However, the overall trend of the plot, caused by the ionospheric effects, can vary by several metres over a very short period (5-10 minutes).

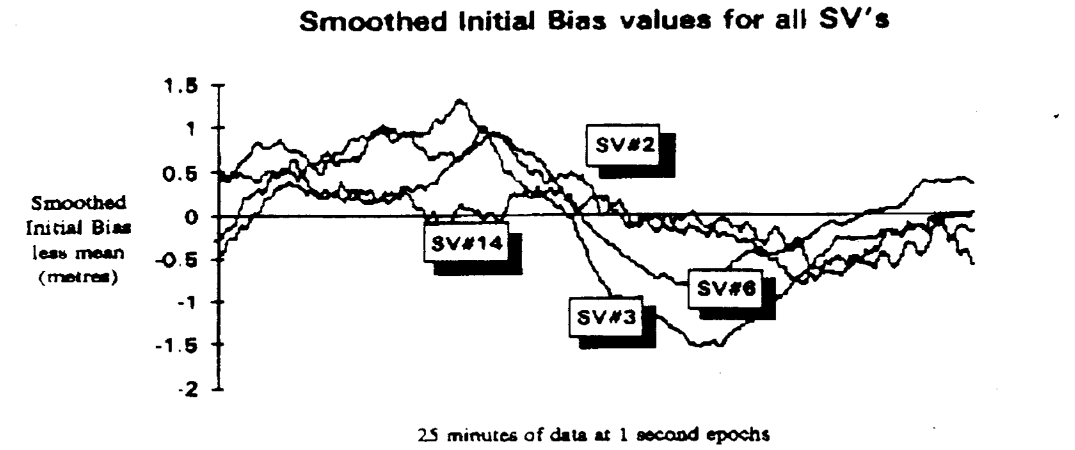


Figure 16: Smoothed Initial biases for all SVs

Smoothing the initial biases by averaging the previous n measurements is very effective in reducing the measurement noise in the pseudo range measurements. If dual frequency measurements were available, allowing the ionospheric correction to be removed, then all measurements could be used in the pseudo range smoothing.

However, with single frequency measurements the drift of the initial bias (caused by the effect of the ionosphere) limits the number of measurements that can be used in the smoothing process. At present the software uses the last 100 measurements for this smoothing. When recording at 1 second intervals this amounts to some 100 seconds of data. In the event of a loss of lock, the maximum jump in range due to the resetting of this smoothing counter would be always less than 0.3 metre and typically less than 0.1 metre.

This moving window technique is quite simple to implement, however it requires that each initial bias be stored in memory. If eight satellites are being tracked and the last 100 biases are being stored, then some seven KBytes of memory is required. This amount of memory is insignificant in a PC based system. However, to avoid this problem other methods have been proposed to perform this smoothing using a number of independent ramps. (Ashjee et al 1989) These methods do not increase the accuracy but rather are concerned with preserving system memory. Memory conservation is important when designing internal receiver software.

A method superior to the simple moving window is one which fits a regression line to the previous n measurements. From this linear fit, the appropriate value at the current epoch is estimated. This method provides a noise free value without any bias due to the ionospheric drift effect. The disadvantage of this method is that extra computational overhead is introduced.

5.3.4 Cycle Slip Detection

The continuous carrier phase is determined by using the integrated Doppler measurements (see Section 2.2.6 on page 31) to determine the number of cycles that have passed between two consecutive phase measurements. If there has been a loss of lock during this period, it is possible that the integrated Doppler measurement may be in error. This will cause a corresponding error in the continuous phase measurement. This type of an event is referred to as a cycle slip where the continuous phase count is in error by an unknown number of whole cycles.

It should be noted that the receiver hardware can usually detect such occurrences and resets the Doppler count to zero to indicate this. In this case the cycle slip is a controlled event and can be dealt with accordingly. However, if the receiver hardware does not detect the slip, then all the phase measurements from this time on will be biased by an unknown number of cycles. As this situation is highly undesirable it is important to monitor the solution for such undetected cycle slips.

Doppler count

There are several factors that indicate the presence of cycle slips. As mentioned above, if the receiver hardware detects a cycle slip it resets the Doppler count to zero. The integrated Doppler reading will also be reset to zero. In this case a new initial bias term must be calculated by averaging the initial bias terms collected from this time onwards.

Phase Residual

When the hardware has not indicated the presence of a cycle slip, several other indicators should still be monitored. The phase residual (discussed in section 2.2.7 on page 33) provides good assurance as to the quality of the phase observations. The phase residual is the difference between the continuous phase measurement and the integrated Doppler measurement. When measurements are of a good quality this phase residual is generally less that 0.02 of a cycle (for Ashtech XII data).

A high phase residual does not prove that a cycle slip is present. However, if the phase residuals are consistently low, then it is unlikely that any cycle slips have occurred.

Initial Bias Residual

A good indicator of the presence of large cycle slips is given by the initial bias residual. As discussed in section 5.3.2, the initial bias obtained for each epoch is averaged to obtain a more accurate value of this initial bias. A bias residual can be determined by subtracting this mean initial bias from the currently determined bias.

When no cycle slips are present, this residual should reflect the noise in the original pseudo range measurement. This was seen to be in the order of 1 metre for Ashtech measurements (figure 15). Consequently, any initial bias significantly greater than this value would indicate the presence of a cycle slip. In such cases the initial bias counter should be reset for this satellite.

Instantaneous Doppler

The Doppler measurements, made at each measurement epoch, can be used to check the validity of the integrated Doppler measurements. For a static receiver, the instantaneous Doppler readings can be used to determine the change in range between two consecutive epochs. The simplest method of determining the change in range is to mean the two Doppler measurements and multiply this mean by the epoch time interval. Ashjee (1986b) describes a method of using the Doppler measurements to repair cycle slips for static baseline measurements. In this method he employs the more elaborate scheme of fitting a polynomial to a series of ten Doppler measurements. An integration of this polynomial over the period of interest is performed to derive the missing cycle count.

However, for a static receiver collecting data every three seconds, the simple method of meaning the Doppler epochs is sufficient to determine the range difference to an accuracy of less than two cycles. Hence, this method can be used as a check for cycle slips larger than one cycle.

When 4 or more satellites are available

When four or more satellites are being tracked it is much easier to repair cycle slips. If data with no cycle slips is available from at least four satellites then it is a simple matter to determine the change in position of the GPS antenna for this epoch. Once this delta position has been accurately determined, cycle slips experienced on any other satellites can be repaired. This method, described fully in Allison (1989), is used in most kinematic processing where maintaining of lock is extremely important.

At present this type of implementation into the real time system has not been attempted. As it does present a significant processing overhead, strategies would need to be devised to ensure that this test is only applied when there is some reason to suspect a cycle slip. A routine to implement this checking facility could be called when phase residuals reach a certain limit or signal/noise ratio drop too low.

One problem with this method for cycle slip checking concerns the nature of cycle slips. If the cycle slip is due to a jump in the receiver clock, then it is likely that this slip will be present on all satellites. Hence the slip cannot be repaired in this manner.

Discussion

Using the initial bias residual, all cycle slips greater than a few metres can be trapped and appropriate action taken. Cycle slips greater than 0.2 metre can be detected using the instantaneous Doppler measurements for the base station receiver. However, this method is not reliable for a moving receiver.

If a cycle slip does pass through undetected, it will have two effects. First, the epoch containing the cycle slip will experience a significant jump in position. For a one metre cycle slip this may cause a position jump of five to ten metres depending on the satellite geometry. This erroneous measurement will also influence the initial bias calculation and hence cause a bias in the calculated position for the next few minutes. This effect would be diluted by the other measurements and would typically be less than a metre in magnitude.

The major concern of undetected cycle slips is that the position errors caused by their existence are not reflected in the calculated position accuracy statistics. Therefore, if position integrity is of importance, any possibility of a cycle slip should be noted and indicated in the solution accuracy statistics.

Possible cycle slips are indicated by several factors. First, the output of the Ashtech receiver contains a warning flag that is set if the receiver detects a possible cycle slip. A low signal noise ratio is another warning that data may be suspect. High phase residuals and bias residuals are strong indicators of the likelihood of cycle slips.

5.3.5 Validation and discussion

The principal check of the effectiveness of the carrier phase smoothing is that the solution for a single receiver is stable. Over a period of a few minutes, the solution will typically not drift more than a few decimetres. This is assuming that the same set of satellites are being tracked constantly.

To highlight the short term stability of the carrier phase smoothed solution, the antenna was moved in a figure eight with circle diameters of approximately five metres. The shape of the figure could be clearly seen on a screen plot.

Comparing the values of the point position solution obtained with known coordinates, it was found that the solution was usually within ten metres of the known WGS84 coordinates. This also provided assurance that all code was functioning correctly.

5.4 Forming and passing corrections

Once the smoothed ranges have been determined, a very stable position solution can be obtained. Typically, this position solution drifts slowly with time. When good satellite geometry is used, then a drift rate of some ten metres over an hour period is quite normal. However, the position will jump sharply (perhaps by twenty metres), if the number of satellites used in the position solution changes.

These jumps can be attributed to the effect of the biases being introduced/removed by the addition/removal of a satellite. It should be remembered that while the smoothed position is much more stable than the raw code phase derived position, it is no more accurate. In the smoothing process the random errors have been predominantly removed. However, the position is still being distorted by the influence of systematic errors or biases.

These biases are due to the effect of satellite position errors, residual satellite clock errors and unmodelled atmospheric effects. The position error introduced by use of the broadcast ephemeris alone can be in the order of twenty to thirty metres. This accuracy may be further degraded by the effects of selective availability.

The influence caused by the net effect of these biases is relatively constant over a short period of time. The drift in magnitude of these biases is the reason for the observed drift of the smoothed solution. However, when an extra satellite is introduced or removed from the solution, the net effect of this bias is altered considerably. This explains the large jumps in position experienced when the number of satellites tracked changes.

By correcting the solution with data derived from a known location, systematic biases present in the system can be reduced quite dramatically.

5.4.1 Block Shift Correction

The difference between the GPS derived position and the known coordinates of a base station can be determined by simple vector subtraction. This difference represents the shift that is required to be added to the GPS derived position to make it coincide with the known position. If a second mobile GPS receiver is calculating its position in an identical manner, then the shifts determined at the base station is also relevant to this receiver.

However, it is very important that both receivers are calculating their position in an "identical manner". In this regard the following points (table 34) are extremely important.

- both sites must track identical satellites
- both sites must use identical ephemeris for all satellites.
- smoothing algorithms should be applied similarly at both sites.
- if corrections for ionospheric or tropospheric delays are being used then they must be applied similarly at both sites.

Table 34: Rules for Differencing

The main disadvantage of block shift corrections is that the user cannot ensure that both receivers will track the same satellites at all times. The following example highlights this problem.

A base station receiver is tracking five satellites and determining position and appropriate position corrections. These corrections are then sent via radio link to the mobile receiver. However, due to an obstruction of the mobile antenna, only four satellites are being tracked by the mobile receiver. Hence, these position corrections are of no use to the mobile unit and must be discarded.

The opposite may also occur. That is, the base station may be tracking four satellites and the remote station five. In this case, the position correction received from the base station can be used, providing that only the same four satellites are used to obtain the rover position solution.

To avoid such a problem, the method of range corrections is considered superior to using "block shift" position corrections. However, it should be noted that if the conditions set out in table 34 are met then both methods will yield identical results.

5.4.2 Range Corrections

Rather than determine the correction to be applied to the GPS derived position, the corrections to the individual ranges can be used. Range corrections can be determined by first calculating the vector distance between the satellite coordinate and the known ground coordinate. The measured range is then subtracted from this value as shown in table 35.

RANGE CORRECTION = CALCULATED RANGE - SMOOTHED RANGE

where

RANGE CORRECTION	Correction to be applied to a measured range to obtain the true range.
CALCULATED RANGE	Calculated distance between satellite coordinates and known receiver coordinate
SMOOTHED RANGE	Smoothed pseudo range corrected for receiver clock.

Table 35: Range Corrections

To obtain appropriate corrections, the receiver clock error must be determined so that it can be subtracted from the pseudo range to obtain the measured range. Hence, a position computation must be performed at the base site.

This position computation is not essential. However, without some estimate of the receiver clock correction, the size of the range corrections being passed would become extremely large. Also this position determination allows for a degree of quality control of the base station measurements.

If this calculation of receiver clock correction is in error, then it will be identical for each range correction. Hence, when the corrections are applied at the mobile station, this error will be partially taken up in the solution of the mobile receiver clock error. In essence, it is only the difference in receiver clock error that is of importance.

5.4.3 Form of Correction message

The form of the corrections sent via the communications link is very important. The first requirement is to keep the amount of data being transmitted as small as possible. The smaller the correction message, the quicker it can be sent and hence the more frequently corrections can be obtained. Frequent correction updates are essential, particularly with the proposed clock dithering methods being introduced by selective availability.

There has been considerable work on developing standards for these types of correction messages. Kalafus (Kalafus et al 1986) proposes the use of a set of standard forms packed into 30 bit words similar to the GPS navigation message. The same algorithm as the navigation message is also used for parity checking. This involves using the last 6 bits of every word to check the integrity of the data contained in the other 24 bits of the message. This type of checking is essential for any type of data transfer where data corruption is likely.

However, in our case, this parity would have been an unnecessary overhead, due to the built in intelligence of the modems being used for HF transmission. The internal modem software attaches its own parity bits prior to sending the message. If required, this parity is used to build up a complete message from a number of unsuccessful transmissions. Consequently, the addition of another 20% overhead in unnecessary parity was not considered worthwhile. However, to guarantee that the modem software is operating as intended, a check sum is included in the correction message.

The form of the correction message used is shown in table 36. It contains a correction header which includes the checksum, time of corrections and the number of SVs in the transmission. The body of the correction contains the SV prn number, the range correction for this satellite and an associated variance for this correction (see section 5.4.4). Also an identifier pertaining to the ephemeris data being used is contained in the message. This is so the software can ensure both sites are using identical ephemeris data.

	Number of bytes
Correction header	
Checksum	1 byte
ID	1 byte
No SVs	2 bytes
time	4 bytes
For each satellite correction	
SV prn number	1 byte
Correction Variance	1 byte
Correction	4 bytes
Ephemeris ID.	2 bytes

Table 36: Form of Range Correction message

5.4.4 Weighting of Corrected Solution

At the mobile site, the correction message is decoded and used to correct the observed ranges prior to the calculation of position. The accuracy of these new corrected ranges is now dependant on several factors. In order to obtain an accurate estimate of the final position, it is necessary to be able to estimate a variance factor for these corrected ranges. The following paragraphs describe the factors that influence this variance and how they are subsequently estimated.

Measurement noise

The influence of the code phase measurement noise on the solution is reduced by smoothing the code phase with the carrier phase (section 5.3). The variance of the smoothed measurement is inversely proportional to the number of measurements used in the smoothing. For this reason, the number of measurements used at the base station is passed from the base to mobile stations in the correction message. The resultant variance of the corrected range due to measurement noise is given by equation 21 in table 37.

$$\sigma_{NN} = \sigma_{RRb}/n_b + \sigma_{RRr}/n_r$$

if we assume that

$$\sigma_{RRb} = \sigma_{RRr} = \sigma_{RR}$$

then

$$\boxed{\sigma_{NN} = \sigma_{RR} \times [1/n_b + 1/n_r]} \quad [21]$$

where

σ_{NN} = variance due to measurement noise
σ_{RR} = variance due to code phase measurement noise
σ_{RRa} = variance due to code phase measurement noise at base
σ_{RRb} = variance due to code phase measurement noise at rover
n_b = number of measurements used in smoothing at base site
n_r = number of measurements used in smoothing at rover site

Table 37: Variance due to measurement noise

The value of σ_{RR} should be consistent for a particular make of receiver and hence can be estimated with reasonable confidence. Therefore, the appropriate weight for each corrected range is dependant only on the number of measurements used for smoothing at both the base and rover sites.

For this reason only the number of measurements used in the smoothing (n_b) is passed in the correction message to the rover site. This has a maximum value which relates to the maximum number of readings used for this smoothing. At present this is set to 100. From this, the variance of the corrected range can be formed by the rover using assumed values for σ_{RR}.

Interchannel bias

Due to careful design, the errors induced by the signals being processed through different hardware paths is minimal (Ashjaee 1986a). However, the presence of this source of error should still be recognised and represented in the error statistics. As biases from this source are receiver specific, their effect will summate in the differential process. The effect of this bias on the corrected range is equal to 2 ×

σ_{BB}, where σ_{BB} is the measurement variance due to the combined effects of multipath, interchannel bias, and electronic path delay.

The value of σ_{BB} should be constant for a particular type of receiver.

Orbit Decorrelation

Over short baselines the effect of errors in the satellite coordinates is significantly cancelled in the differencing process. As the baseline length increases, these orbit errors have a more significant effect. The magnitude of this error is proportional to the magnitude of the orbit error and the relative geometry of the baseline vector and the satellite.

If we assume that each satellite has a uniform variance of σ_{OO} in all three axes, then the effect this has on the differential range measurement can be readily determined. Figure 17 depicts the geometry of the situation. The variance of the corrected range measurement is equal to $\sigma_{OO} \times d$ where d is equal to the length of the vector between the two unit vectors R and B. This relationship can be obtained directly from similar triangles.

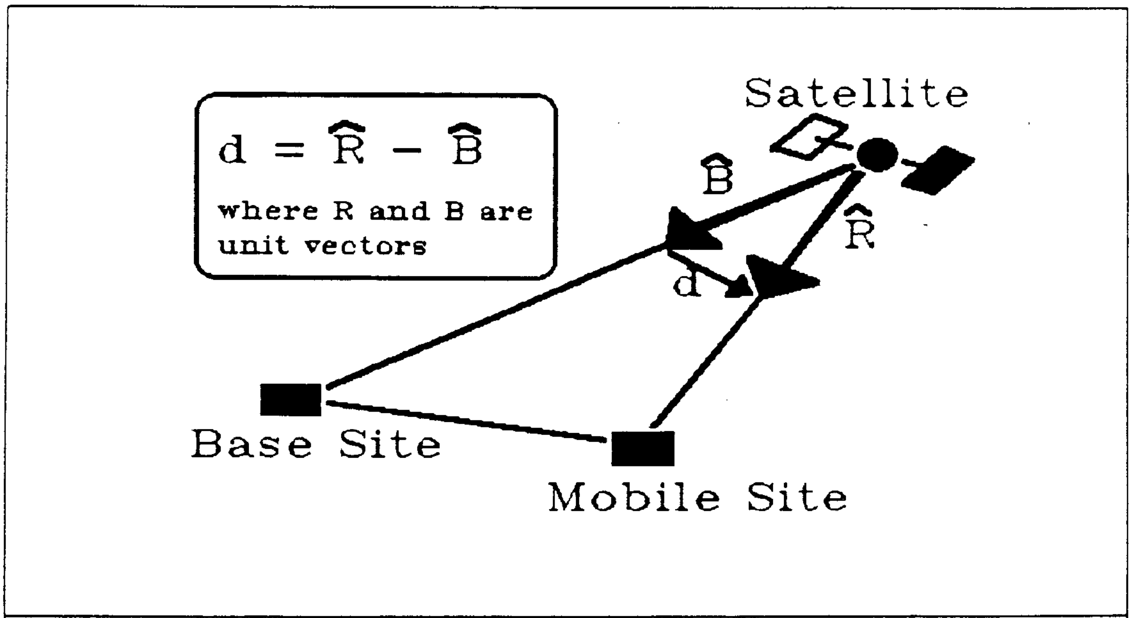


Figure 17: Geometry of orbit error propogation

Refraction effects

For the purposes of estimating solution variances the effect of the troposphere and ionosphere can be viewed simultaneously. As the distance between the base and mobile station increases, so does the error caused by the satellite signal passing through different refractive medium. Amongst other factors, this error is related to the change in elevation angle between the two sites. In an attempt to estimate the differential range errors caused by this error source, a very simple model is employed (Equation 22, Table 38).

$$\boxed{\sigma_{AA} = \sigma_{aa} \times (E_b - E_r)^2} \quad [22]$$

where

E_b = Elevation angle from base station to satellite

E_r = Elevation angle from rover station to satellite

σ_{aa} = variance of errors due to atmospheric effects

Table 38: Variance due to refraction effects

A better solution would be to include a variable to account for the diurnal variation. The errors due to the ionosphere are highly dependant on the time of day. However, more complicated models are not considered appropriate at this stage. Through careful selection of the value of σ_{aa} it should be possible to determine a variance that is appropriate for the worst part of the day. During night time and periods of low ionospheric activity this variance will simply be over conservative.

Correction Delay

Any correction determined by the base station is relevant only to the particular epoch of the determination. It can be readily seen that the corrections change by a few centimetres in a minute period.

Therefore errors introduced by using a correction some fifteen seconds old are minimal. This error is naturally a function of time and can be represented by the function shown in table 39.

$$\boxed{\sigma_{TT} = \sigma_{tt} \times (T_R - T_B)^2} \qquad [23]$$

where

σ_{TT} = is the variance due to the age of the correction
σ_{tt} = rate of change of corrections
T_R = Time of mobile station measurement.
T_B = Time of base station correction measurement.

Table 39: Variance due to Transmission delay

This error source could be significantly reduced by including range rate values in the correction message. This would allow the error to be corrected and a more accurate position obtained. The present system requires frequent updates to maintain the solution accuracy. If for some reason correction updates are not obtained, the term σ_{tt} ensures that this fact is reflected in the resultant position accuracy statistics.

Summary

The resultant variance for each range can be obtained by the summation of equations 21 to 23. The summation of these variances is shown in table 40. From experience gained in controlled tests it is possible to estimate appropriate variance factors for the variables σ_{RR}, σ_{BB}, σ_{AA}, σ_{OO} and σ_{TT}.

$$\sigma_{MM} = \sigma_{RR} \times (1/N_r + 1/N_b) + \sigma_{BB} \times 2 + \sigma_{AA} \times (E_b - E_r)^2 + \sigma_{OO} \times D + \sigma_{TT} \times (T_b - T_r) \} \quad [24]$$

where

σ_{MM} = variance of corrected ranges
σ_{RR} = variance due to code phase measurement noise
σ_{BB} = variance due to residual interchannel biases
σ_{AA} = variance due to orbit atmospheric errors
σ_{OO} = variance of satellite coordinates
σ_{TT} = variance due to age of correction data

N_b, N_r = number of measurements used for smoothing at rover and base site respectively
E_b, E_r = elevation angle of satellite at base and rover sites
T_b, T_r = Measurement times at base and rover sites
D = length of vector between the two unit vectors from the satellite to the base and rover sites

Table 40: Variance of corrected ranges

Once the range variances have been formed in this manner they can be used in the weight matrix for the least squares solution at the rover site. The resultant variance-covariance matrix now contains realistic values pertaining to the accuracy of the position solution. The trace of the variance-covariance matrix provides an estimate of the combined three dimensional position accuracy and clock offset.

A more convenient accuracy statistic is obtained by transforming the variance-covariance matrix into local system coordinates. From this the conventional statistic of $(2 \times d_{rms})$ can be formed. (See Table 41)

$$\begin{bmatrix} \sigma EE & \sigma EN & \sigma EH \\ \sigma NE & \sigma NN & \sigma NH \\ \sigma HE & \sigma HN & \sigma HH \end{bmatrix} = R^T \times \begin{bmatrix} \sigma XX & \sigma XY & \sigma XZ \\ \sigma YX & \sigma YY & \sigma YZ \\ \sigma ZX & \sigma ZY & \sigma ZZ \end{bmatrix} \times R$$

where

$$R = \begin{bmatrix} -\sin\phi\cos\lambda & -\sin\phi\sin\lambda & \cos\phi \\ -\sin\lambda & \cos\lambda & 0 \\ \cos\phi\cos\lambda & \cos\phi\sin\lambda & \sin\phi \end{bmatrix}$$

$$2 \times d_{rms} = 2 \times \sqrt{(\sigma EE + \sigma NN)}$$

Table 41: Formation of $2 \times d_{rms}$ error statistic

5.4.5 Validation of algorithms

Upon implementation of the algorithms for passing corrections, testing was carried out to determine the accuracy of the system. The results of these tests are presented in section 8.0.

6. PROGRAM IMPLEMENTATION

6.1 Introduction

After all the decisions have been made with regard to the system hardware/software and the program algorithms have been chosen, there still remains the task of putting the code together in a logical and structured manner. It is a fallacy to believe that if a program works then it is successful. Programs designed with this philosophy may work perfectly until they require modifications at a later date. A well designed system must be reliable, efficient, maintainable and provide an appropriate user interface.

The importance of the reliability of the software cannot be over stressed. The result of a serious system failure could cost more than the development of the whole system. However, less obvious is the need for the system to be easily maintainable. Software systems have to be maintained if they are to remain useful throughout their life. This need for maintenance arises because the software must fit in to a constantly changing environment.

The ability of the system to be maintained is a highly important factor. System maintenance can be broken into three types. (Sommerville 1989)

Perfective Maintenance: System improvement

An example of this would be rewriting or modifying a graphics routine to include a new feature, perhaps zooming. This might be an improvement found necessary by operators in field use.

Adaptive Maintenance: System evolution

A new interface may be required for a new receiver type. Alternatively, an existing receiver type may change it's data format.

Corrective Maintenance: System repair

This is usually in response to an bug or inaccuracy noticed during the field use of the program.

While every effort is made to reduce the need for corrective maintenance, in large programs, the odds are high that some bugs will exist. With care and good style these errors can only be minimized. Both adaptive and perfective maintenance are an expected part of a healthy system. Hence, it should be accepted that program maintenance is essential and the program should be designed to facilitate this efficiently.

The following section describes the goals of the program implementation intended make the system both maintainable and reliable.

6.2 Program Layout Standards

The layout of the program code affects the readability of the program. Good layout contains consistent indenting which highlights statements executed in the same loop. Header comments are consistent and contain all required information such as return values and functions called. The proper use of blank lines and paragraphing also greatly enhances the readability of a program.

However, the most important rule concerning program layout is consistency. Whatever conventions have been established, the same layout should be used throughout the program. (Sommerville 1989). With this in mind, the following conventions were chosen for use throughout this program.

- All indentation to be made in five space increments.

- Each function to have a header showing the function name, purpose, return values, any global variables used and all functions called. Each header is contained within a solid line to delineate it. All text starts in column 5 and columns 1 and 2 always contain asterisks. The function name is surrounded by hash marks. An example of the header style is shown in table 42.

```
/*
**     #####################
**     Least Squares Adjust
**     #####################
**
**     PURPOSE:       Form and Adjust Observations via iterative least squares
**
**
**     RETURN:        BOOL Returns TRUE if solution made or FALSE if not
**
**     GLOBALS:       Ash              NoAdj
**                    rec**            AltFix
**                    geog**           tolerence
**                    AdjustResid[]
**
**     FUNCTIONS CALLED:
**                    Wmalloc          Wmult_dimension
**                    Wmat_identity    Form_norm
**                    Parametric       Wfree
**                    Wmult_free
**
*/
```

Table 42: Example of function header

- All variable names are to be as descriptive as possible and use the [7]Hungarian notation. (Petzold 1988) In this notation the variable name is prefixed with several lower case variables indicating the data type. This is followed by a descriptive name consisting of lower and upper case letters. For example, in the variable wBufferSize, BufferSize is the descriptive name of the variable that contains the size of the buffer. The "w" prefixed to this variable informs the programmer that this variable is an unsigned integer.

Examples of other prefixes used are shown in table 43.

---

[7] Hungarian Notation: Method of naming variables named in honour of the Microsoft programmer Charles Simonyi.

c	char
by	BYTE (unsigned char)
n	short or int
i	int
b	BOOL (int)
w	WORD (unsigned int)
h	HANDLE (unsigned int)
l	LONG (long)
dw	DWORD (unsigned long)
fn	function
s	string
sz	string terminated with NULL (0)

Table 43: Hungarian Notation prefixes

- The descriptive part of all Global variables are shown in all upper case capitals.
- The descriptive part of all macro constants are shown in all lower case. This makes it simple to identify macro definitions.
- All comments and headings to be contained in a box as shown in table 44. One blank line should be left before each box.
- All local variables have descriptions given opposite the initial declaration. This practise also discourages the use of more than one declaration per line.
- All parameters to have description given opposite the initial declaration.

```
pvLeastSquares(NoSVs)

int NoSVs; /* Number of SV's present */

{
/*
     *------------------*
     * local variables
     *------------------*/
     double      **ppdBMatrix;           /* pointer to Obs Eq matrix                */
     double      *pdTMatrix;       /* pointer to Correction matrix       */
     double      *pdWeights;       /* pointer to weight matrix      */
     double      *pdResiduals;           /* pointer to vector of residuals          */
     double      pdParameters[4];        /* parameters to be solved        */
     double      pdCovariance[16]; /* Covariance matrix                  */
     double      dRadius;          /* Distance from point to ECEF origin */
     HANDLE      hB1,hB2;          /* Handles to multidimensioned matrix */
     HANDLE      hTMatrix;         /* Handle to memory for TMatrix       */
     HANDLE      hResiduals;       /* Handle to memory for Residuals     */
     HANDLE      hIdent;                 /* Handles to temp memory for Ident        */

/*
     *---------------------------------------------*
     * Make Adjustment for Alt hold case
     *---------------------------------------------*/
     NoAdj=NoSVs+AltFix;

/*
     *---------------------------------------------*
     * Create temp matrix space for adjustment
     *---------------------------------------------*/
     ppdBMatrix  =Wmult_dimension(NoAdj,4,&hB1,&hB2);
     pdTMatrix   =Wmalloc(NoAdj*sizeof(double),&hTMatrix);
     pdResiduals =Wmalloc(NoAdj*sizeof(double),&hResiduals);
     pdWeights   =Wmat_identity(NoAdj,&hIdent);

/*
     *---------------------------------------------*
     * Initialize parameters for adjustment
     *---------------------------------------------*/
     for(i=0;i<4;i++)
            Parameters[i]=0.0;
```

Table 44: Example of Program Layout

6.3 Programming methods

In designing a "maintainable" program it is important to maximize the cohesion of program components and to minimize the couples between components. (Sommerville 1989 p177). This should be the goal when developing large programs. One method of reducing couples is to minimize the use of global variables. In general, modules should be as self contained as possible. This is one of the chief advantages of object orientated or modular programming. A highly cohesive program is one where each unit performs a strong definite function.

In simple terms, a well written program is made of discrete functional modules, with a minimal number of couples between each one. The art of achieving this goal is part of the skill of programming. Some rules that assist in this are as follows

- All constants or variables that may be altered should be defined in a separate part of the program.
- No function should become too large. Typically 200 lines (with perhaps half this being comments) is large.
- Each module should have a clear function and purpose.
- Global variables should be used sparingly.
- Exceptions or errors should be anticipated and appropriately handled.
- Any device dependant code (such as code related to a certain receiver type) should be isolated into one particular area.

To facilitate efficient growth of the system it is convenient to develop process models and flow charts. This ensures that the program has an adequate structure or framework around which to build. These charts help the programmer ensure that the code is being implemented in the simplest manner possible. Also the links between elements can be seen more clearly. Nassi-Schneiderman diagrams

(Robertson 1989) provide a more structured alternative to these flow charts and were also used during the development.

6.4 Program Design: Level Zero

In top down programming, levels of process models can be used in a hierarchical system. A level zero flow chart showing the overall function of the software design can be seen in figure 18. This shows the basic difference in function of the software at the base and rover sites. It also defines the interaction that the system will have with the external devices (modem and GPS receiver).

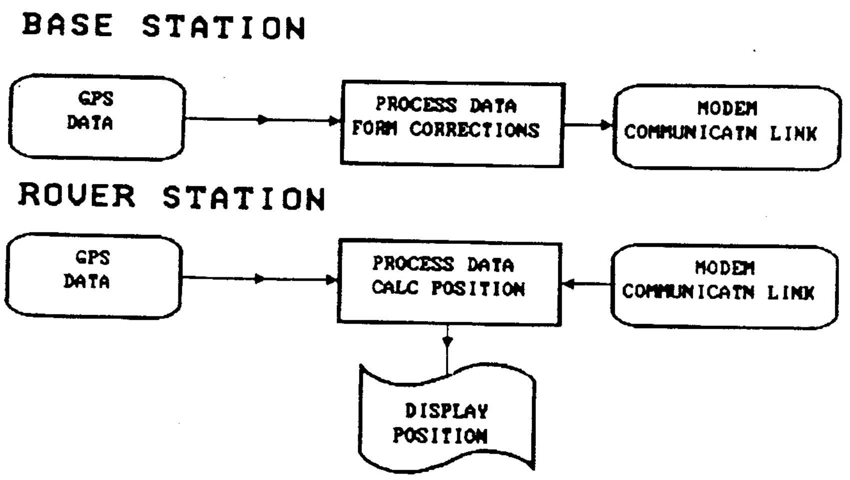


Figure 18: Overall system function

However, in practice, it is not efficient to have two separate programs, one for the base and one for the rover site. This type of separation causes problems in that two parallel systems must be developed with considerable duplication of code. Improvements to one system must then be continuously incorporated into the other.

A far better approach is to combine both modules so that the same system can be used as either a base or a rover site. The system then simply requires the operator to tell it which site it is. Figure 19 shows the same overall design but for the combined system. Depending on a conditional test with regard to the current operational mode, the program branches to perform the actual function. In practice, both systems follow a similar path with an appropriate branch occurring at many instances. These branches can be seen quite often in the lower level flow charts.

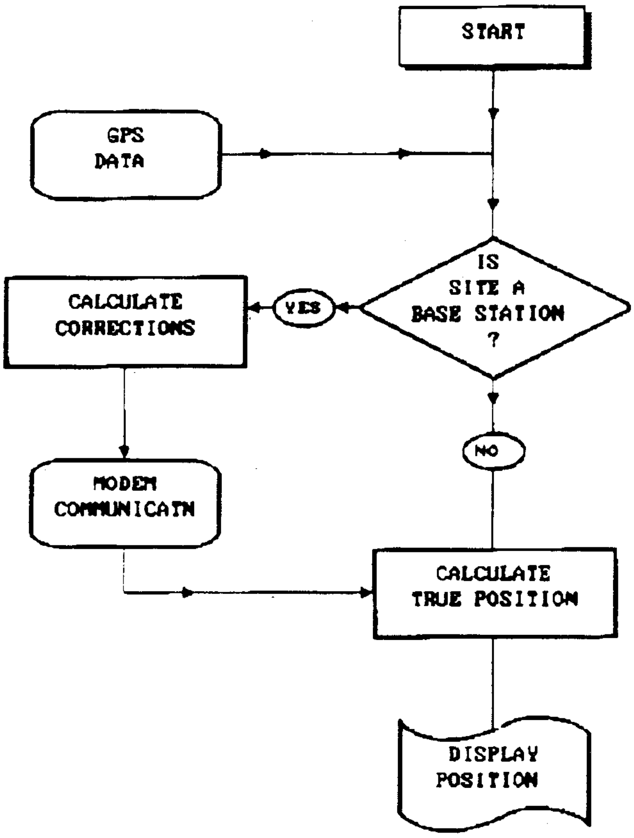


Figure 19: Overall System function showing combined modes

6.5 Program design

Level one

In the level one flow chart (figure 20) key program modules are defined. The functionality of these modules and the message flow between them are shown. In the NAVPROC module, GPS data is read from the GPS receiver and decoded. When a data set has been decoded a message is sent to the POSPROC module informing it that a new data set is available. Here, the position is calculated and, depending on whether the program is in base or rover mode, corrections are either sent or received from the MODPROC module.

This was the module design set out before the commencement of the coding of the program. During the development of the system, a great deal of change was seen in the lower level program design. However, this basic information flow is still evident.

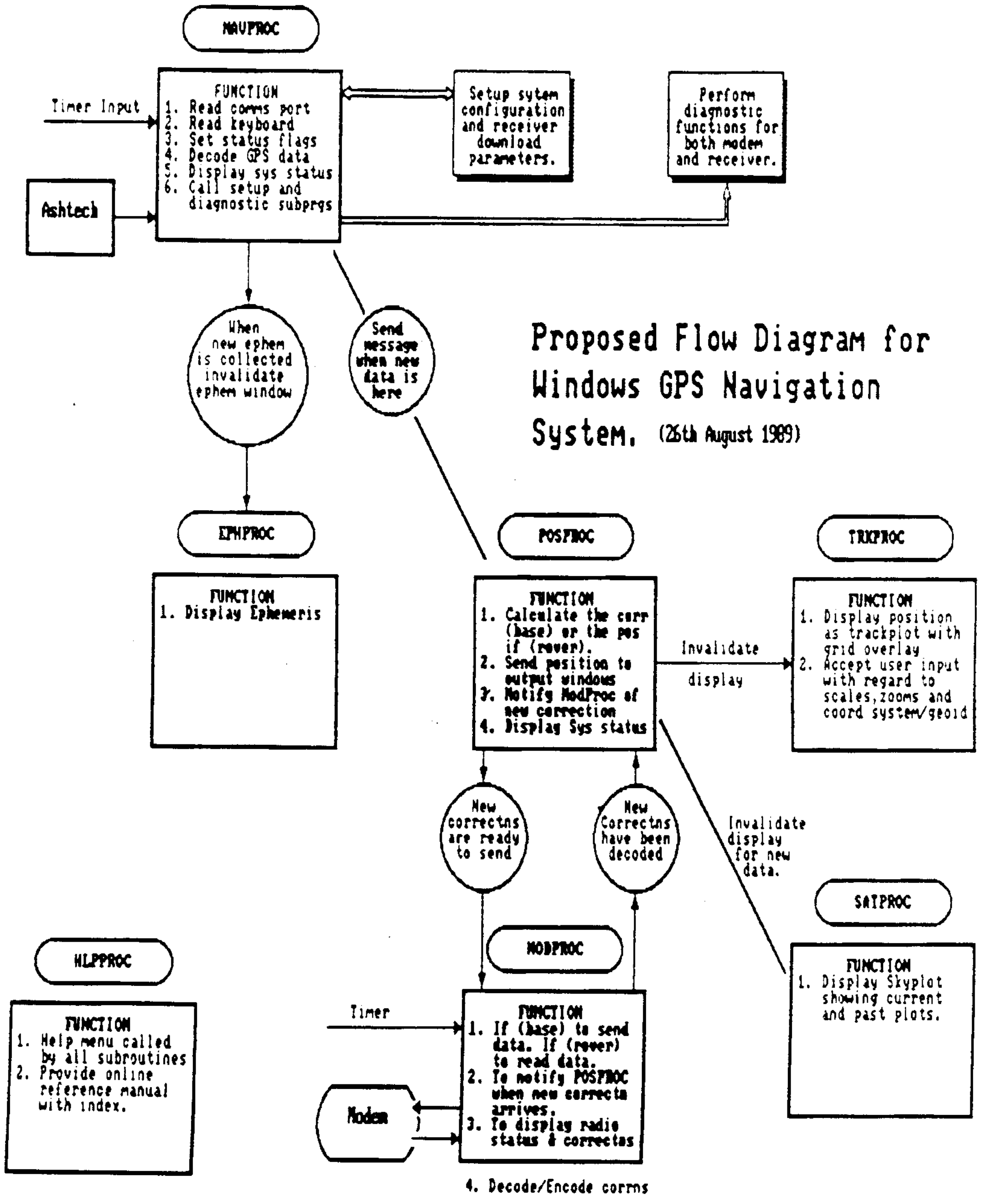


Figure 20 : Program Overview

Level 2

An example of a level 2 function is the process of determining the three dimensional position from the raw data. The code for this computation is performed by a group of functions in a discrete module. As this module encompasses the basic function of the program, its design is shown in the level two flow chart in figure 21. Source code for this module is contained in appendix A.

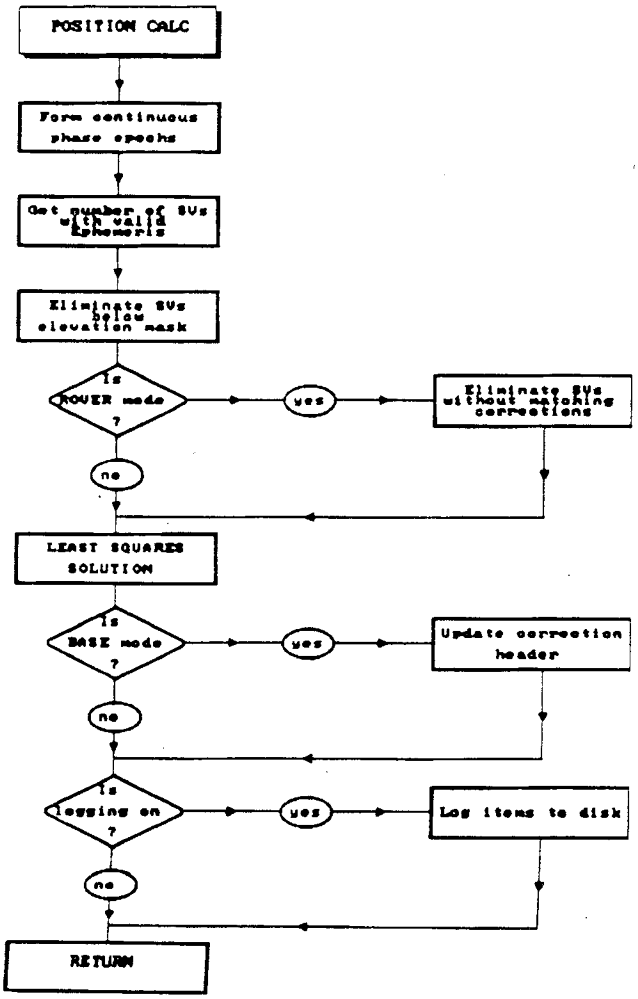


Figure 21: Flow diagram for Position Calculation Module

Level 3 functions

Each of the major steps in the Level 2 flow chart can be examined in further detail in lower level charts. Examples of these are shown in figures 22 and 23. Finally a hierarchy chart relating each of the functions in the module is shown in figure 24. With this type of planning it is much easier to view the logic of the code implementation. Links between functions can be easily seen. One aim of the module for position calculation was that it should be highly independent. However, there is a reasonable amount of coupling between the functions inside this module. Most of this coupling is through the use of common data in many of the functions.

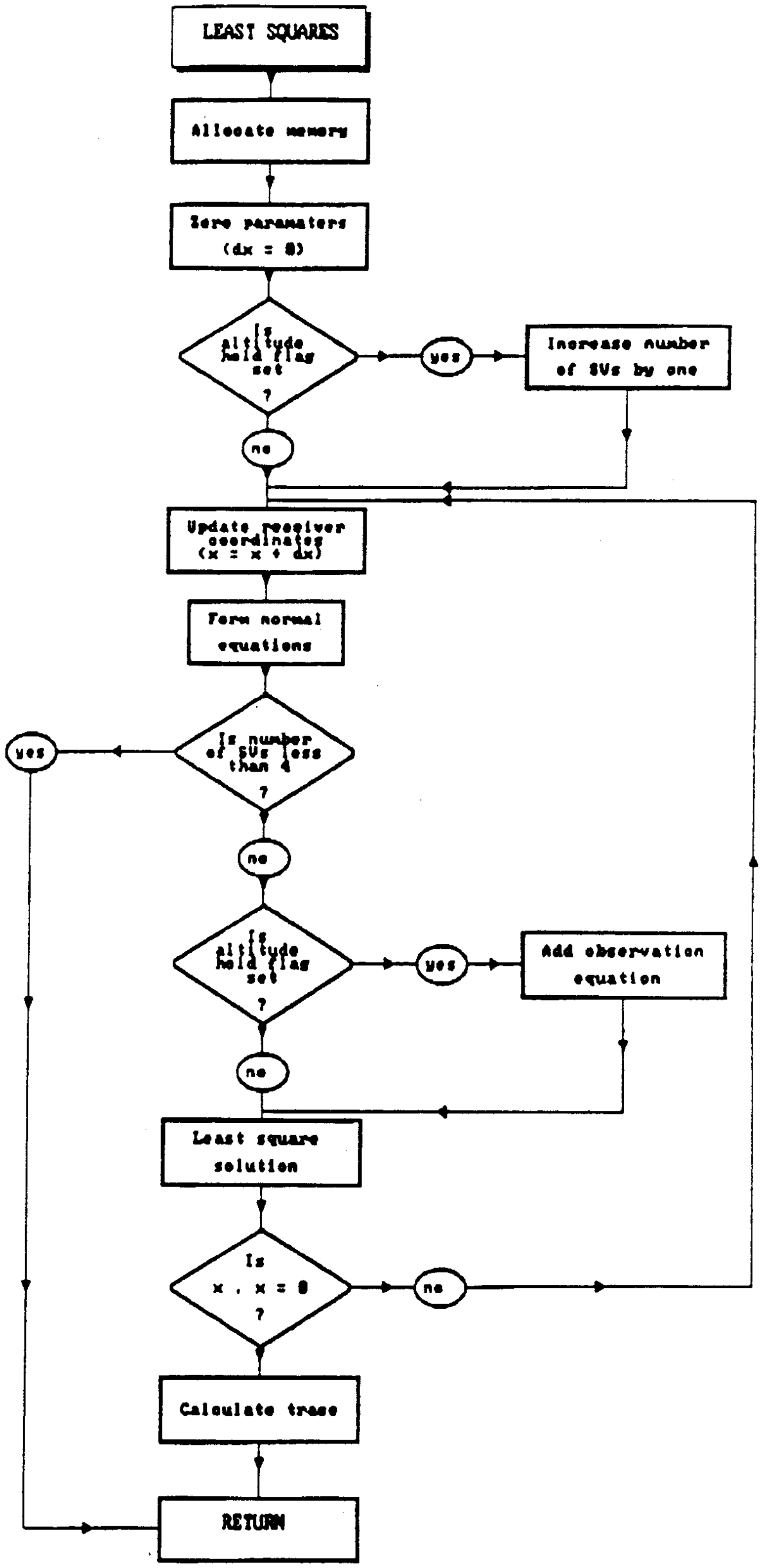


Figure 22: Least Squares Flow Chart

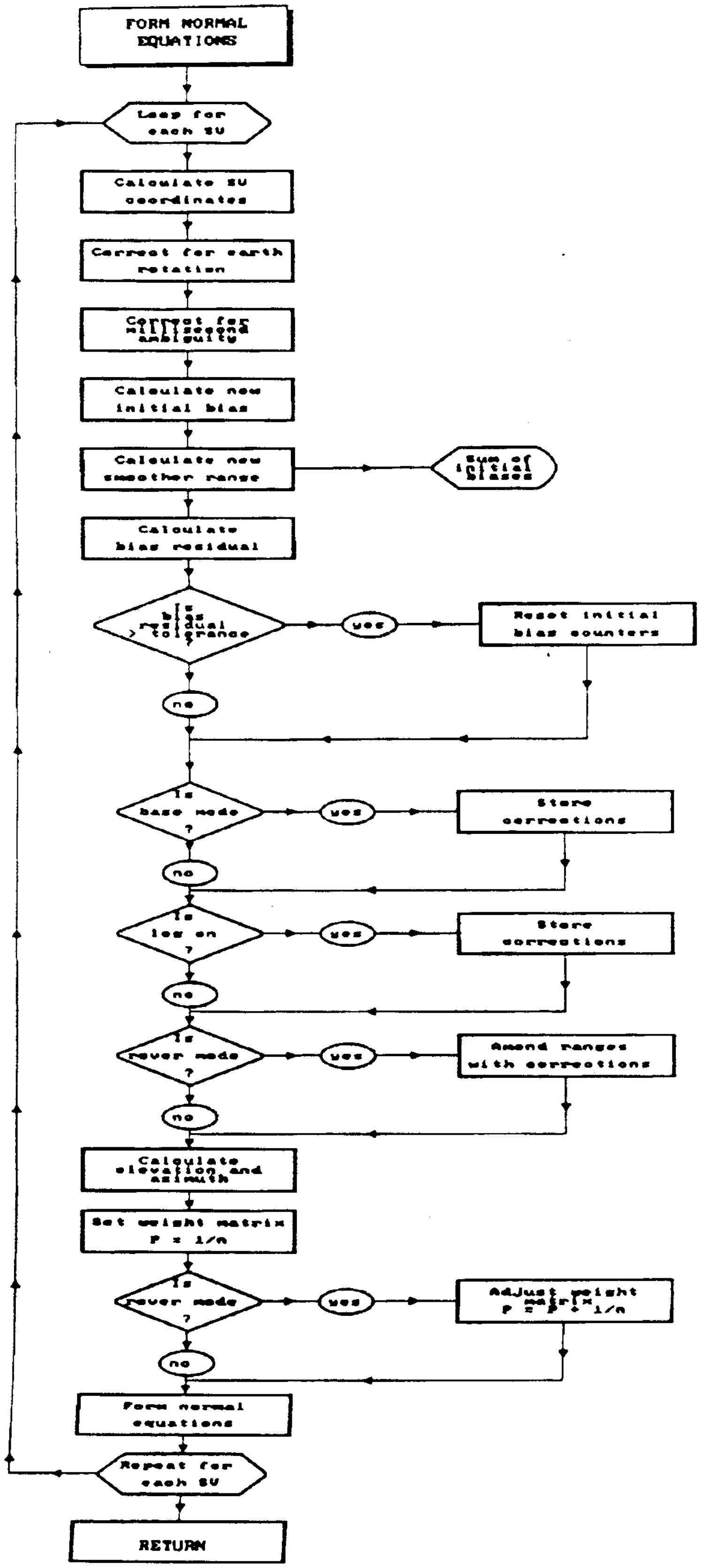


Figure 23: Formation of Normal Equations

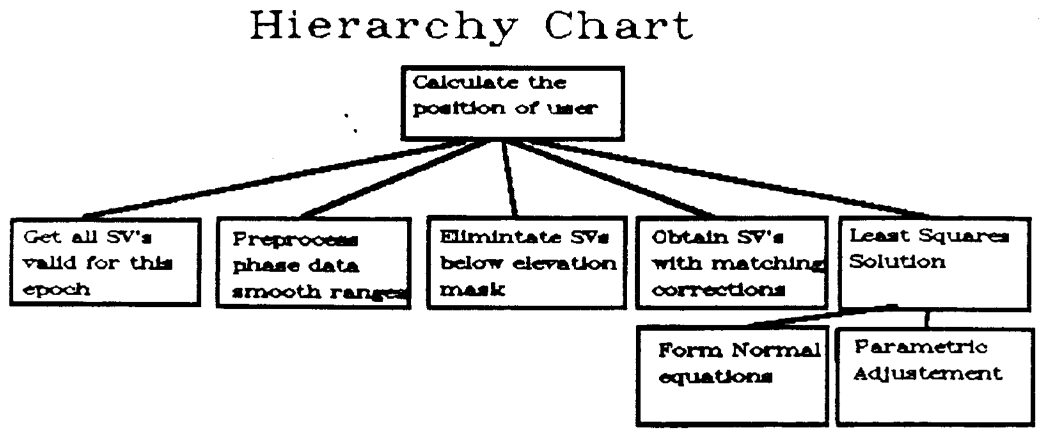


Figure 24: Hierarchy Chart for Position Calculation Module

6.6 Windows programming

Programming in the Windows environment has its own set of idiosyncrasies that must be learnt. To the user, each window is a display space on the computer screen. From the programmer's perspective, each window is an object that receives and processes messages. These messages are the key to the Windows system. All communication between modules and from the user (via keyboard and mouse input) is kept in a standard form and stored in the system message queue.

These messages are then processed in turn and sent to the appropriate program and window. For example, when a GPS data record is collected from the receiver (in the NAVPROC module), the following occurs. The data is stored in global memory and a message is put in the queue informing the POSPROC window of this event. When this message is received by the POSPROC module, the appropriate actions are taken and a position is calculated using this data.

In the meantime, other messages are also being sent into this queue. Such a message may be from the MODPROC module informing the NAVPROC module that a base station correction has been decoded. Alternatively, it may be informing a module that the user has double clicked on a particular selection in a dialogue box.

The use of this message structure is the means by which Windows programs are able to multitask. Consequently, it is important that code is written with this in mind. For instance, it is very important that no one module keeps control of the processor for too long. If the function is required to pause for any reason, perhaps waiting for input from a serial port, then it should return control so that other messages can be processed. Before returning control, the module can post a message to itself so that it will regain control after the message queue has been serviced. There are many references available on programming in the Windows environment (Petzold 1988). It is important that the Windows system is understood if programs are to be successfully developed for this environment.

6.7 Help

Both in-line help and external documentation are important for success with this type of system. User help should provide two kinds of information. One, when the user is in trouble, ie what do I do now? The second type of help is when the user simply requires information. In this early implementation of the system, provision has been made for full in line help. This can be called by pressing the F1 key at any time. However, at this stage, a comprehensive help manual for this on line facility has not been written. It is considered inefficient to attempt to write this sort of documentation until the program has evolved into a stable commercial product.

6.8 Implementation

Learning to use the functions of the Windows environment required a considerable length of time. This learning has not yet finished, more will be required in the future if printer/plotter output and more sophisticated displays are to be implemented. However, even at this stage of the development, the advantages offered by the Windows environment have been worthwhile. A great deal of time was spent in streamlining the logic of the code. The benefits of this extra work will be realized as the program grows on the basis of this work.

Several problems were uncovered in the developmental process. First, it was discovered that the Windows environment only provided support for two serial communication ports. This is a serious limitation as this is the minimum requirement for the system. To overcome this shortcoming during testing, data was sent to the navigational computer via one of the two parallel ports. A parallel/serial convertor was used to transform the data to the serial input required by the receiving computer.

Another, unexpected limitation of the Windows environment was that it does not support programming using a large memory model. In large models both code and data segments are referenced with far pointers. Due to Windows sophisticated memory swapping capabilities, addressing a data segment with far pointers can create problems when the address of a data segment is changed. This problem can be overcome through careful use of near and far pointers.

7. PROGRAM DESCRIPTION

7.1 Introduction

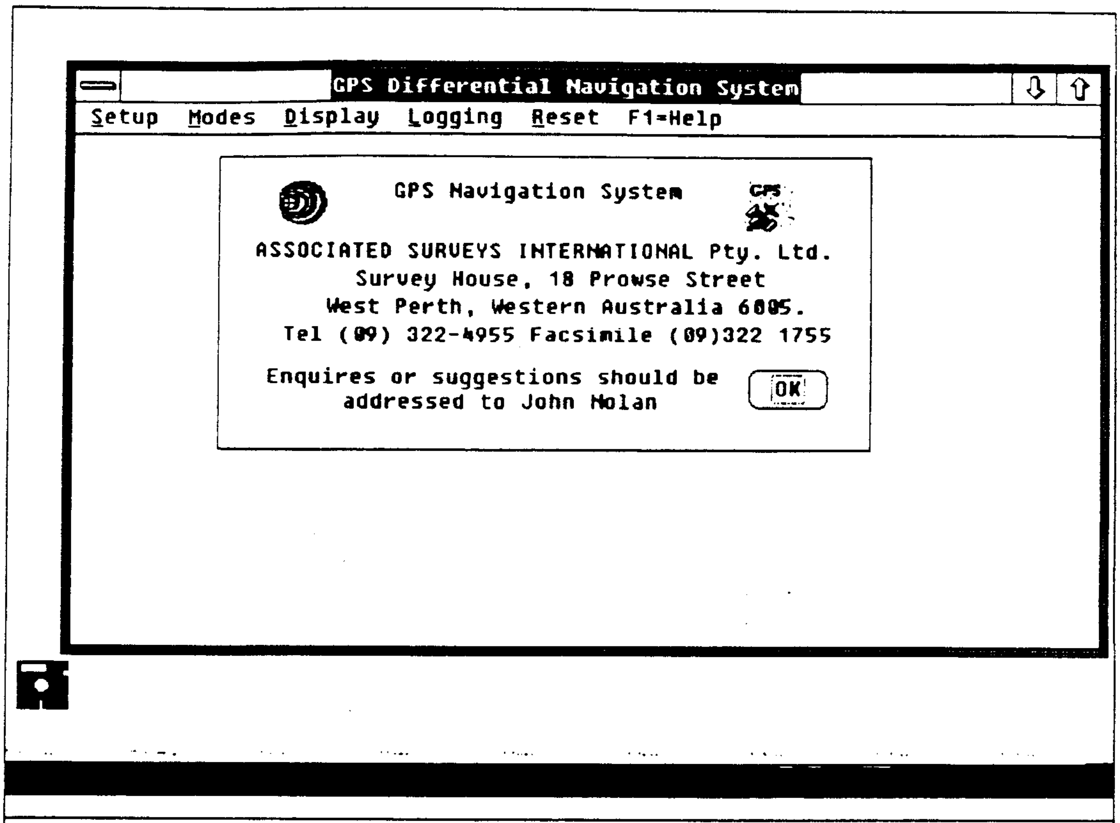
The following section describes the form of the current program interface and the options presented to the user. The program is currently undergoing modifications to customise it more for off-shore positioning. However, the program configuration presented here shows the software as it was during its test and validation period.

One of the strengths of the Microsoft® Windows user interface is the ease with which it can be altered. Menus and sub menus can be changed or have new items/options appended quite easily. Similarly, dialogue boxes can be added or restructured and thus alter the overall look of the program.

Despite the fact that the order of the menus may be changed quite drastically in the future, the functionality of the software should basically remain as described in the following sections.

7.2 Main Menu

Figure 25 shows the main menu options available on bootup of the program. Each of the menu options can be selected using the left mouse button. Alternatively, selections can be made by holding down the ALT key and pressing the appropriate letter in the menu item name. The appropriate letter for this selection is underlined in each case. This method of selection is consistent throughout the program and through all Windows programs.



GPS Navigation System

ASSOCIATED SURVEYS INTERNATIONAL Pty. Ltd.
Survey House, 18 Prowse Street
West Perth, Western Australia 6005.
Tel (09) 322-4955 Facsimile (09)322 1755

Enquires or suggestions should be addressed to John Nolan

Figure 25: Starting Menu

The menu items cover very broad areas. The SETUP menu contains all the functions for setting up the system for specific hardware and location requirements. The MODES options allows the user to select the appropriate mode of operation and the DISPLAY menu allows selection of the required information to be displayed. The LOGGING option controls what information is recorded and by what devices.

Upon selection of any of the items from the main menu a pull down menu is obtained. The selections available in each of these sub menus is described systematically, starting with the SETUP options.

7.3 Setup Options

The SETUP menu is shown in figure 26. The ellipses (...) seen after many of the menu options is a standard Windows convention. It informs the user that selection of this item will be followed by a dialogue box.

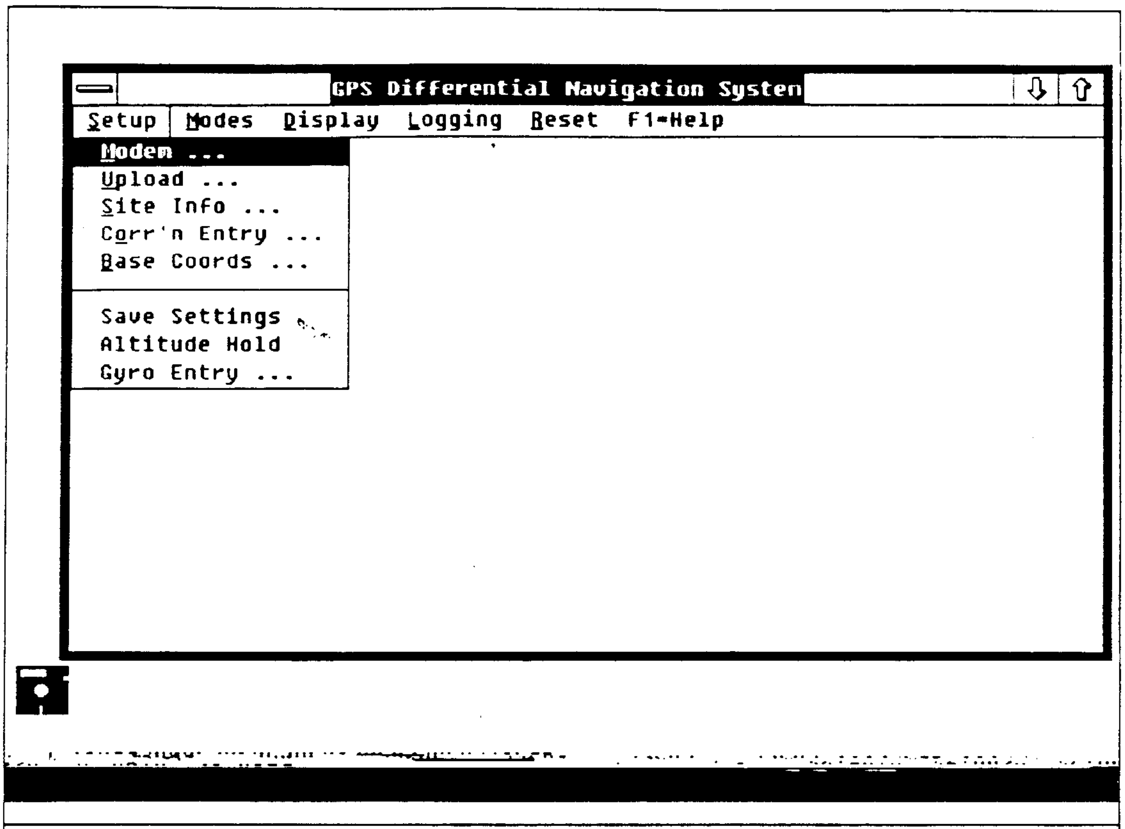


Figure 26: Setup Menu

Each of the items in this menu are discussed in order.

7.3.1 Modem

The modem dialogue box is shown in figure 27. This box allows selection of the parameters required to initialize the system for a particular modem.

Modem Type:

The modem type can be selected by clicking the mouse button in the modem type box. This toggles the available selection of modem types in the box to the right. The default modem selected in this figure is Transcom.

Transmit Delay:

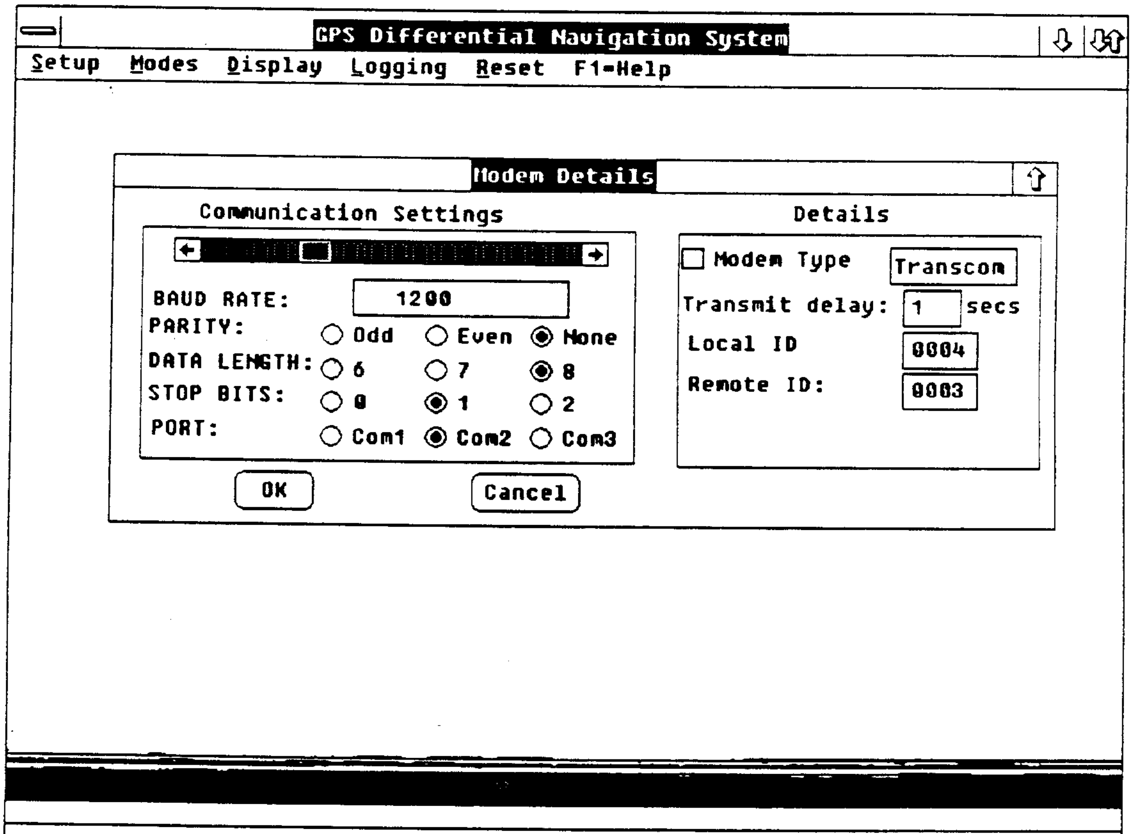
Certain radios may overheat if data is constantly transmitted without any delay. This field allows the user to nominate a delay (in seconds) between successive transmissions.

Local & Remote IDs:

Intelligent modems that require acknowledgment on the successful receipt of data usually require an ID for each modem. This is so the modem software can identify the modem from which the acknowledgment is coming. Where many mobile units are to be used, this is very important. These fields allow the user to enter appropriate ID numbers.

Communication settings

The baud rate and protocol settings can be set quite easily using this dialogue box. The baud rate is selected by simply moving the scroll bar thumb until the required baud is shown in the baud field.



BAUD RATE: 1200
PARITY: ○ Odd ○ Even ● None
DATA LENGTH: ○ 6 ○ 7 ● 8
STOP BITS: ○ 0 ● 1 ○ 2
PORT: ○ Com1 ● Com2 ○ Com3

☐ Modem Type: Transcom
Transmit delay: 1 secs
Local ID: 0004
Remote ID: 0003

Figure 27: Modem Setup Dialogue Box

Other parameters are selected using radio buttons. The selection of these is mutually exclusive. For example, if 2 stop bits are selected, then the highlight is automatically removed from 1 stop bit.

7.3.2 Upload

The UPLOAD menu allow the user to upload data into the memory of the receiver. This information is highly dependant on the receiver type used. Hence a unique dialogue box is displayed for a particular receiver type. The dialogue box for the Ashtech XII receiver is shown in figure 28. This box allows entry of the following information.

Enable/Disable satellites

This option allows the user to disable satellites. This may be required if it is known that certain SV's are not operating correctly.

Initial Position

The Ashtech receiver does not require an initial position for tracking purposes. However, a position is required if satellite elevations are to be computed when less than four satellites are present.

Elevation mask

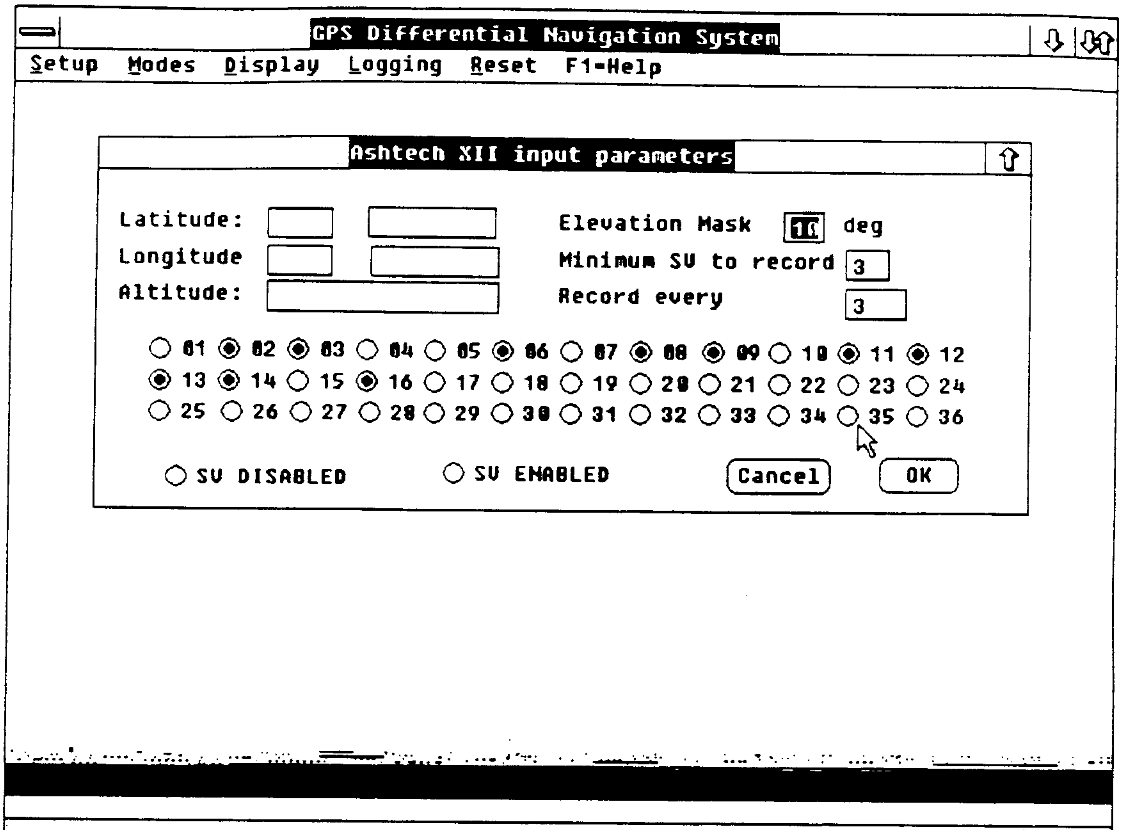
Satellites below this mask are not used in the position computation.

Minimum SVs

Data is recorded to the receiver memory when this minimum number of satellites are being tracked.

Recording Rate

The recording interval between measurement epochs.



Latitude:
Longitude
Altitude:
Elevation Mask 10 deg
Minimum SU to record 3
Record every 3

○ 01 ● 02 ● 03 ○ 04 ○ 05 ● 06 ○ 07 ● 08 ● 09 ○ 10 ● 11 ● 12
● 13 ● 14 ○ 15 ● 16 ○ 17 ○ 18 ○ 19 ○ 20 ○ 21 ○ 22 ○ 23 ○ 24
○ 25 ○ 26 ○ 27 ○ 28 ○ 29 ○ 30 ○ 31 ○ 32 ○ 33 ○ 34 ○ 35 ○ 36

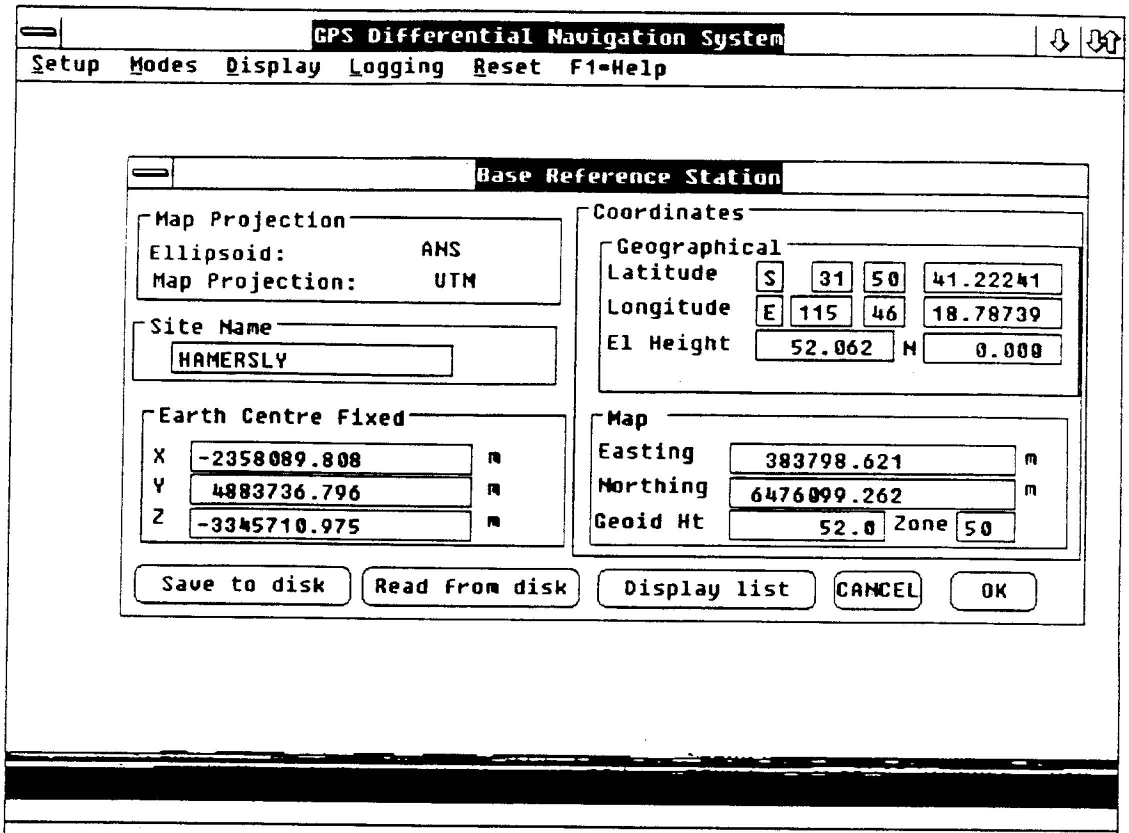
○ SU DISABLED ○ SU ENABLED

Figure 28: Ashtech Receiver Setup Dialogue Box

7.3.3 Base Coordinates

This dialogue box (figure 29) displays the name, map projection and spheroid of the reference station coordinates. These coordinates are displayed in geographical, map projection and Earth Centred Earth Fixed (ECEF) coordinates as can be seen in the example. The coordinates can be entered by changing any of these fields. Upon pressing ENTER (or the OK button), corresponding fields will be appropriately updated. It should be noted that the ECEF Cartesian coordinates are always shown in the WGS84 datum. Geographical coordinates entered on the ANS spheroid are transformed using an appropriate seven parameter transformation.

Base station entries can be saved to disk and recalled later, using the appropriate commands at the bottom of the dialogue box.



Map Projection
Ellipsoid: ANS
Map Projection: UTM

Site Name
HAMERSLY

Earth Centre Fixed
X -2358089.808 m
Y 4883736.796 m
Z -3345710.975 m

Coordinates
Geographical
Latitude S 31 50 41.22241
Longitude E 115 46 18.78739
El Height 52.062 M 0.000

Map
Easting 383798.621 m
Northing 6476099.262 m
Geoid Ht 52.0 Zone 50

Figure 29: Reference Station Dialogue Box

Save Settings

This option writes all current system settings to disk file. These options then become the default settings for the next invocation of the program. Settings are stored in the windows WIN.INI file. This file is in an ASCII text form and can be edited with a standard text editor. Hence these settings can be changed prior to starting the program, if required.

The advantage of this is that the settings for a specific hardware and location installation can be saved. Hence, if this survey is repeated at a later time, then it is a simple matter to read the required settings straight from disk file. This eliminates the need for operator input, thus reducing the likelihood of mistakes.

Altitude Hold

When this option is set a tick mark appears beside this menu item. When the altitude hold option is on, the height of the position is constrained to be the height value contained in the initial position entry. This constraint is primarily used when navigation is required with only three satellites. Alternatively, it is useful when there are more than three satellites but the geometry of the satellites is poor.

It should be noted that no matter how many satellites are present, having this option will affect the solution. Hence, when sufficient satellites are available (with good geometry), then this option should be disabled.

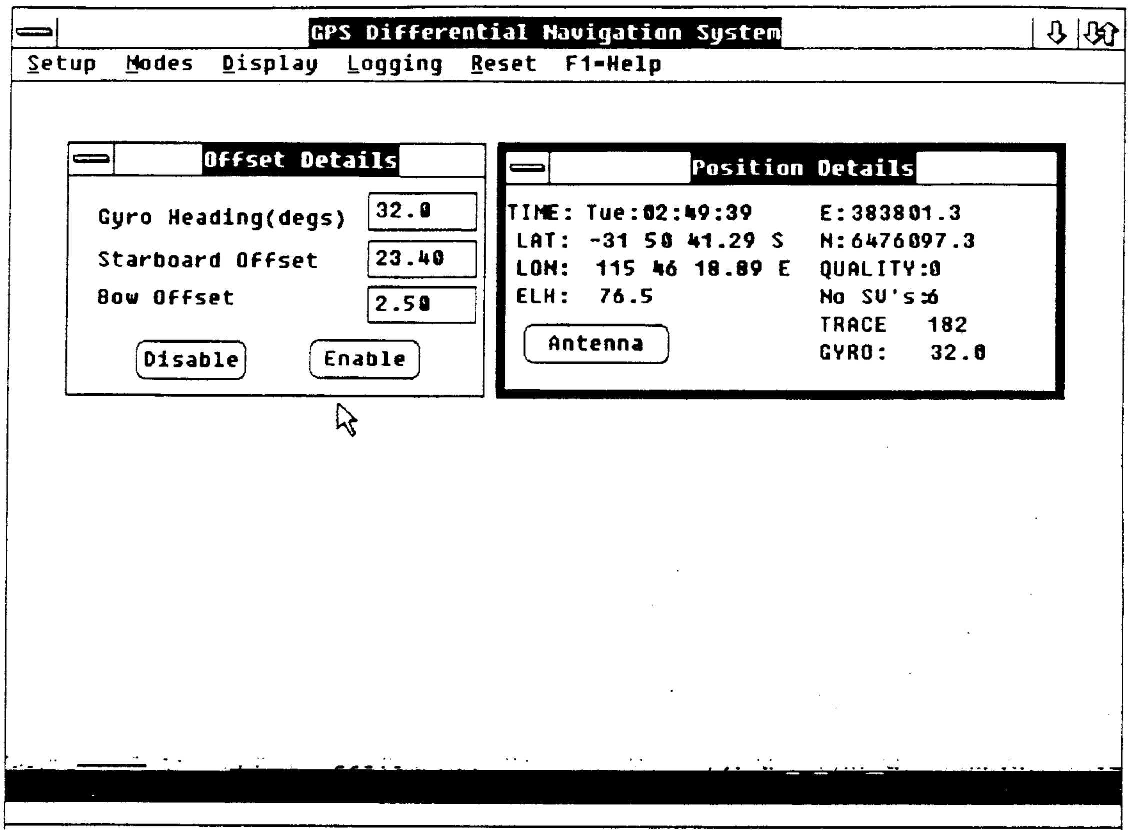
Gyro Entry

This option allows an offset position to be tracked instead of the position of the GPS antenna. An example of this is in locating an oil rig over a given location. In such cases, it is the position of the drill stem which is required to be placed over the location. Due to logistic constraints, it would be highly unlikely that the GPS antenna could be placed over the position of the drill stem. Hence the antenna is usually located at some distance away from the drill stem where it has good satellite visibility.

Consequently, once the GPS antenna position has been determined this offset position must be added to derive the position of interest, namely the drill stem. The offset between the drill stem and the antenna is usually a constant distance but may vary in azimuth depending on the orientation of the rig.

To facilitate the offset calculation, this dialogue box allows the user to enter the fixed offset relative to the orientation of the vessel. This offset is entered via a starboard and bow component which in effect defines the distance offset. The other information required is the heading of the vessel, usually obtained from a gyro reading or a compass.

Figure 30 shows the gyro dialogue box that allows the input of these parameters. When the ENABLE box is pressed then the position tracked in the position details box will change from Antenna to Offset. This indicates that the GPS antenna position is being corrected for the current offsets contained in the gyro dialogue box.



Gyro Heading(degs) 32.0
Starboard Offset 23.40
Bow Offset 2.50

TIME: Tue:02:49:39
LAT: -31 50 41.29 S
LON: 115 46 18.89 E
ELH: 76.5
E:383801.3
N:6476097.3
QUALITY:0
No SV's:6
TRACE 182
GYRO: 32.0

Figure 30: Gyro Entry and Position Dialogue Box

7.4 Modes Options

The modes menu (figure 31) allows the selection of the four basic modes of operation. The basic modes are as follows:

Base Station

In base mode, corrections to the ranges are determined and sent out via the communications link.

Rover Station

In rover mode, corrections are received from the base station and applied to the incoming measurements.

Stand Alone

This mode is intended for software use with one receiver only. Corrections are determined when the receiver is situated at a known location. These corrections are then used to correct observations made in a roving mode. The aim of this mode is to obtain maximum accuracy using one GPS receiver.

Base Monitor

The base monitor mode is identical to the base mode except that, as well as transmitting corrections, it also looks to receive a returned position from the mobile units. These positions can then be updated on the track plot display.

Due to the limited recording rates imposed with some communication forms, the update rate of the position of the mobile units will be poor. However, it does allow the position of a number of rovers to be tracked, which may be an advantage in some cases.

In the lower section of the modes menu are the selections that allow the user to track the appropriate information from the receiver. The data being received is marked with a tick mark. The data types currently supported are as follows:

Raw data:	Raw pseudo and carrier phase data
Ephemeris:	Ephemeris parameters for all satellites
Position:	Receiver calculated position
Navigate:	Both raw data and ephemeris updates

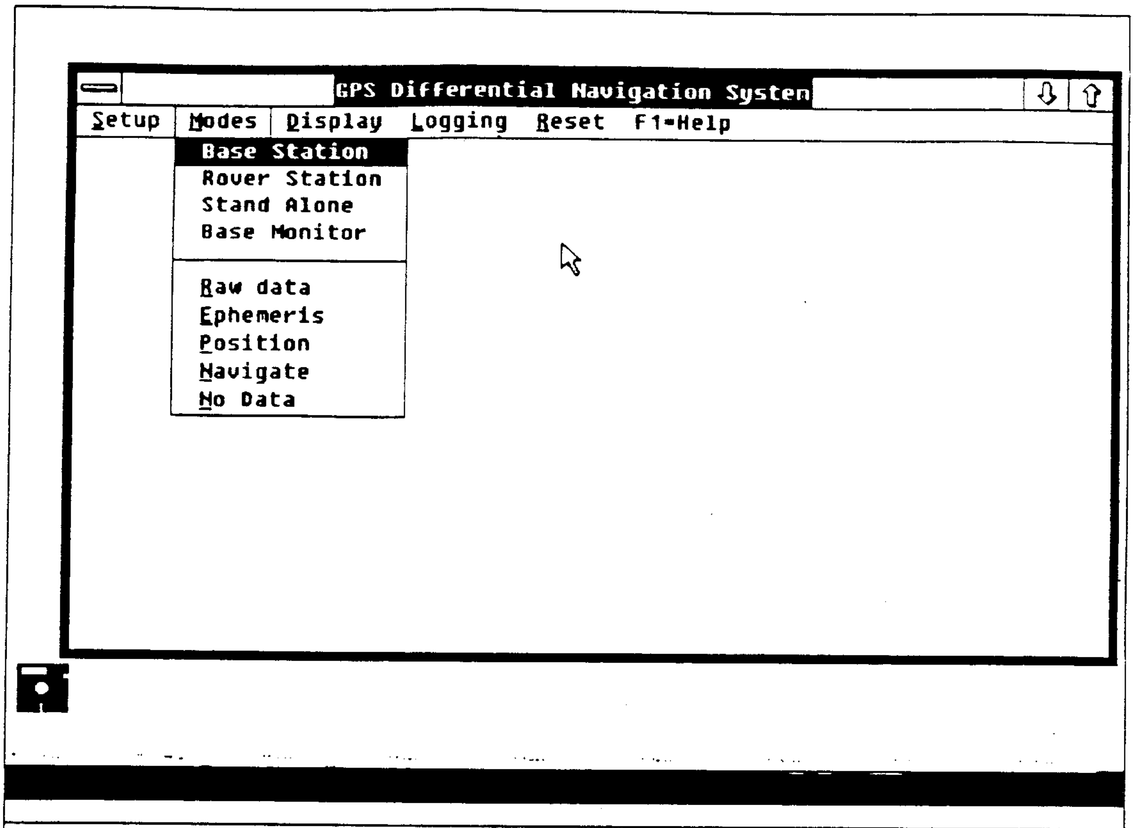
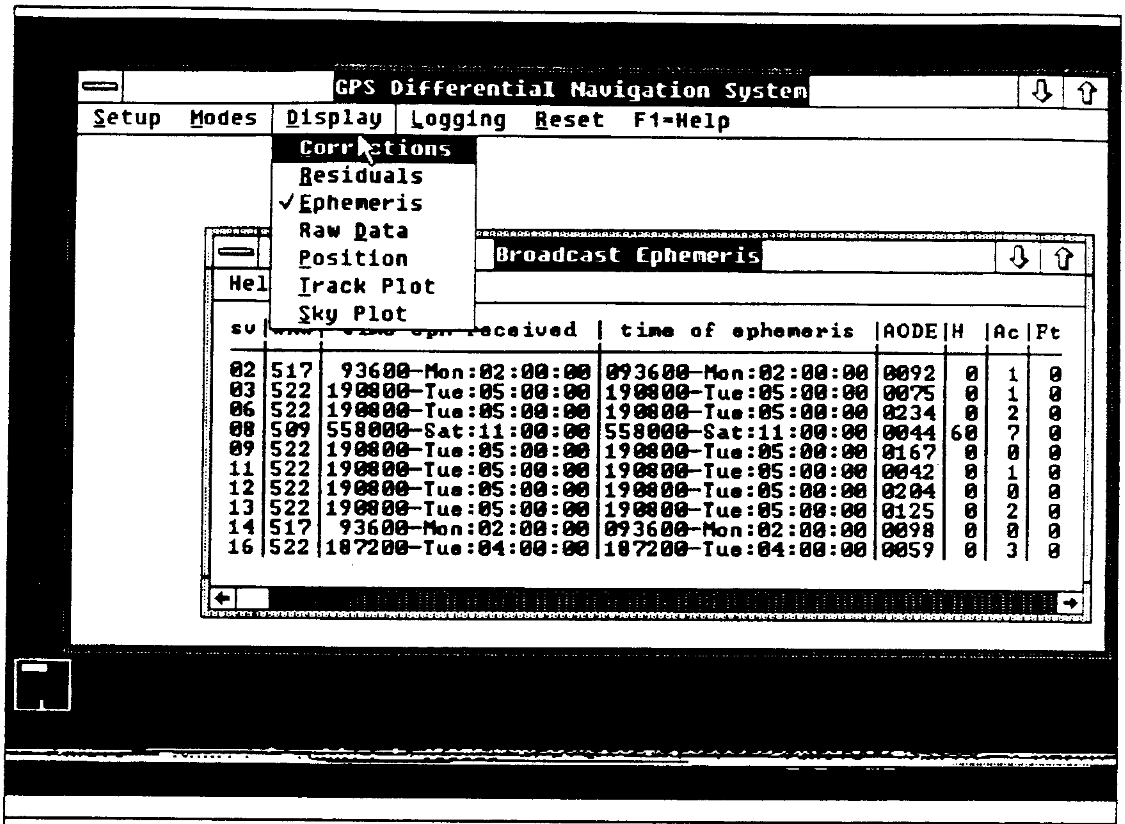


Figure 31: Modes Menu

7.5 Display Options

The display options are the key to the Windows user interface. Each display option shown in figure 32 creates its own window for the display of its option. In figure 32 the ephemeris display is also shown. Note that since this display is presently active, a tick mark appears in the main menu. Each of these windows can be displayed either by itself or in combination with other display windows.



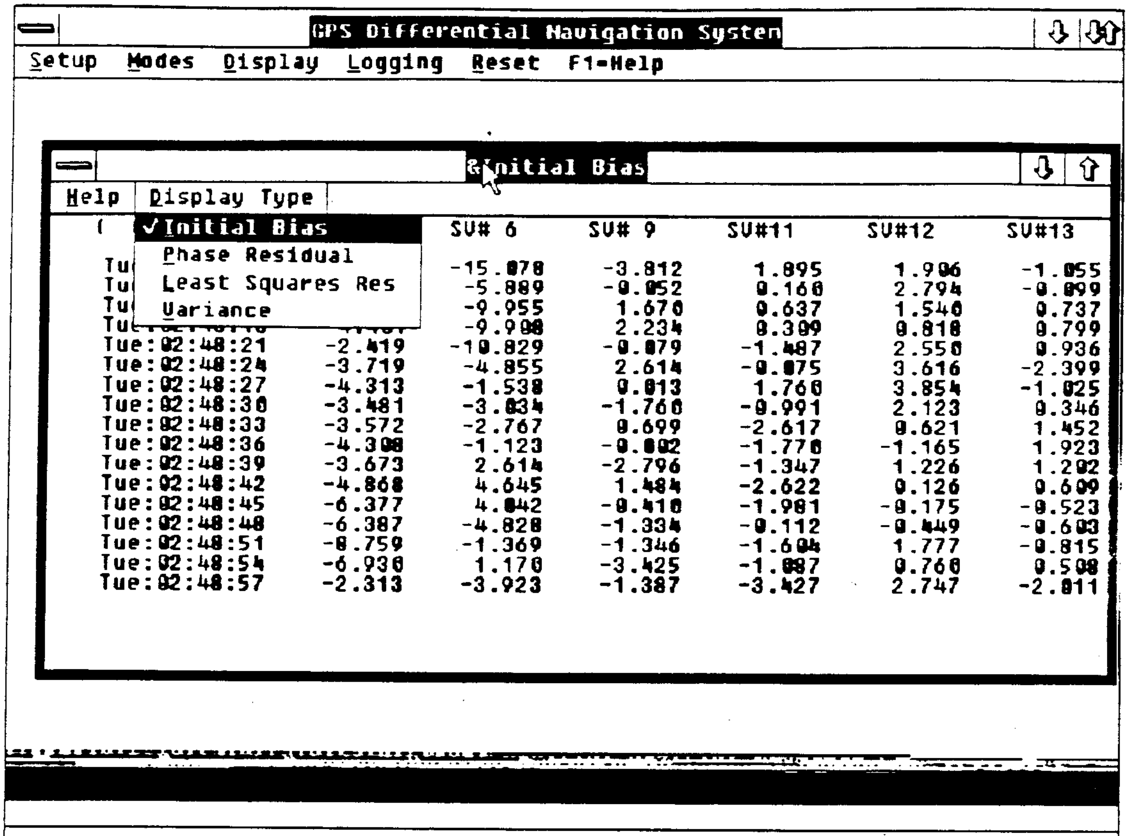
```
sv |                 eceived | time of ephemeris |AODE|H  |Ac|Pt

02 |517|  93600-Mon:02:00:00 |093600-Mon:02:00:00 |0092| 0 | 1| 0
03 |522|190800-Tue:05:00:00 |190800-Tue:05:00:00 |0075| 0 | 1| 0
06 |522|190800-Tue:05:00:00 |190800-Tue:05:00:00 |0234| 0 | 2| 0
08 |509|558000-Sat:11:00:00 |558000-Sat:11:00:00 |0044|60 | 7| 0
09 |522|190800-Tue:05:00:00 |190800-Tue:05:00:00 |0167| 0 | 0| 0
11 |522|190800-Tue:05:00:00 |190800-Tue:05:00:00 |0042| 0 | 1| 0
12 |522|190800-Tue:05:00:00 |190800-Tue:05:00:00 |0204| 0 | 0| 0
13 |522|190800-Tue:05:00:00 |190800-Tue:05:00:00 |0125| 0 | 2| 0
14 |517|  93600-Mon:02:00:00 |093600-Mon:02:00:00 |0098| 0 | 0| 0
16 |522|187200-Tue:04:00:00 |187200-Tue:04:00:00 |0059| 0 | 3| 0
```

Figure 32: Display Menu with Ephemeris Displayed

Examples of each display type are discussed in the following:

Residuals

The residuals display is the key display for quality control monitoring. It allows the user to monitor the initial biases, the phase residuals, the least squares residuals or the solution measurement variances. Each of these outputs is displayed in the form shown in figure 33. The sub-menu of "Display Type" can also be seen in this figure which allows selection of the appropriate residual. A description of each of these residual types can be found in section 5 on Algorithm design.



```
 (                                 SV# 6     SV# 9     SV#11     SV#12     SV#13

 Tu                               -15.078    -3.812     1.895     1.906    -1.055
 Tu                                -5.889    -0.052     0.160     2.794    -0.099
 Tu                                -9.955     1.670     0.637     1.540     0.737
 Tu                                -9.908     2.234     0.309     0.810     0.799
 Tue:02:48:21         -2.419      -10.829    -0.079    -1.487     2.550     0.936
 Tue:02:48:24         -3.719       -4.855     2.614    -0.075     3.616    -2.399
 Tue:02:48:27         -4.313       -1.538     0.013     1.760     3.854    -1.025
 Tue:02:48:30         -3.481       -3.034    -1.760    -0.991     2.123     0.346
 Tue:02:48:33         -3.572       -2.767     0.699    -2.617     0.621     1.452
 Tue:02:48:36         -4.308       -1.123    -0.002    -1.770    -1.165     1.923
 Tue:02:48:39         -3.673        2.614    -2.796    -1.347     1.226     1.202
 Tue:02:48:42         -4.868        4.645     1.484    -2.622     0.126     0.609
 Tue:02:48:45         -6.377        4.042    -0.410    -1.981    -0.175    -0.523
 Tue:02:48:48         -6.387       -4.828    -1.334    -0.112    -0.449    -0.603
 Tue:02:48:51         -8.759       -1.369    -1.346    -1.604     1.777    -0.815
 Tue:02:48:54         -6.930        1.170    -3.425    -1.087     0.760     0.508
 Tue:02:48:57         -2.313       -3.923    -1.387    -3.427     2.747    -2.011
```

Figure 33: Residuals Display Window

Ephemeris

The format of the ephemeris output is shown in figures 32 and 34. The Keplerian ephemeris parameters are not displayed but rather the relevant information identifying the ephemeris being used. The information contained in this ephemeris screen for each satellite is given in table 45.

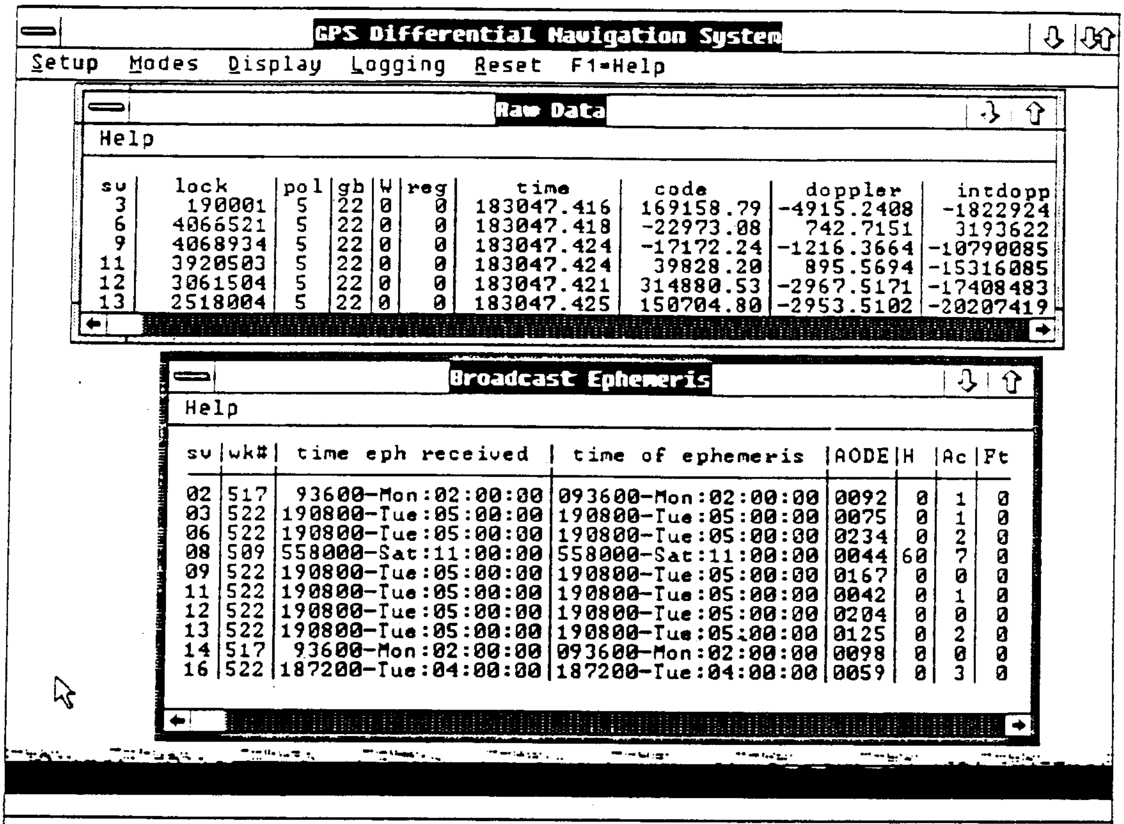
SV PRN number
GPS week number of ephemeris
Time that ephemeris was received
Reference time of ephemeris
Ephemeris set identification number (IODE)
Health flag
Accuracy flag
Fit Flag

Table 45: Information shown in Ephemeris display

Raw Data

The raw data being obtained from the GPS receiver is displayed as shown in figure 34. The raw data displayed has a horizontal scroll bar at the bottom of the window.

This scroll bar can be used to view information on the right hand side of the screen. This allows the window to be kept in a small part of the screen and only the relevant information viewed. The raw data screen displays all data contained in the Ashtech raw data structure. For a description of these outputs see Ashjee (1989a).



sv	lock	pol	gb	W	reg	time	code	doppler	intdopp
3	190001	5	22	0	0	183047.416	169158.79	-4915.2408	-1822924
6	4066521	5	22	0	0	183047.418	-22973.08	742.7151	3193622
9	4068934	5	22	0	0	183047.424	-17172.24	-1216.3664	-10790085
11	3920503	5	22	0	0	183047.424	39828.20	895.5694	-15316085
12	3061504	5	22	0	0	183047.421	314880.53	-2967.5171	-17408483
13	2518004	5	22	0	0	183047.425	150704.80	-2953.5102	-20207419

sv	wk#	time eph received	time of ephemeris	AODE	H	Ac	Ft
02	517	93600-Mon:02:00:00	093600-Mon:02:00:00	0092	0	1	0
03	522	190800-Tue:05:00:00	190800-Tue:05:00:00	0075	0	1	0
06	522	190800-Tue:05:00:00	190800-Tue:05:00:00	0234	0	2	0
08	509	558000-Sat:11:00:00	558000-Sat:11:00:00	0044	60	7	0
09	522	190800-Tue:05:00:00	190800-Tue:05:00:00	0167	0	0	0
11	522	190800-Tue:05:00:00	190800-Tue:05:00:00	0042	0	1	0
12	522	190800-Tue:05:00:00	190800-Tue:05:00:00	0204	0	0	0
13	522	190800-Tue:05:00:00	190800-Tue:05:00:00	0125	0	2	0
14	517	93600-Mon:02:00:00	093600-Mon:02:00:00	0098	0	0	0
16	522	187200-Tue:04:00:00	187200-Tue:04:00:00	0059	0	3	0

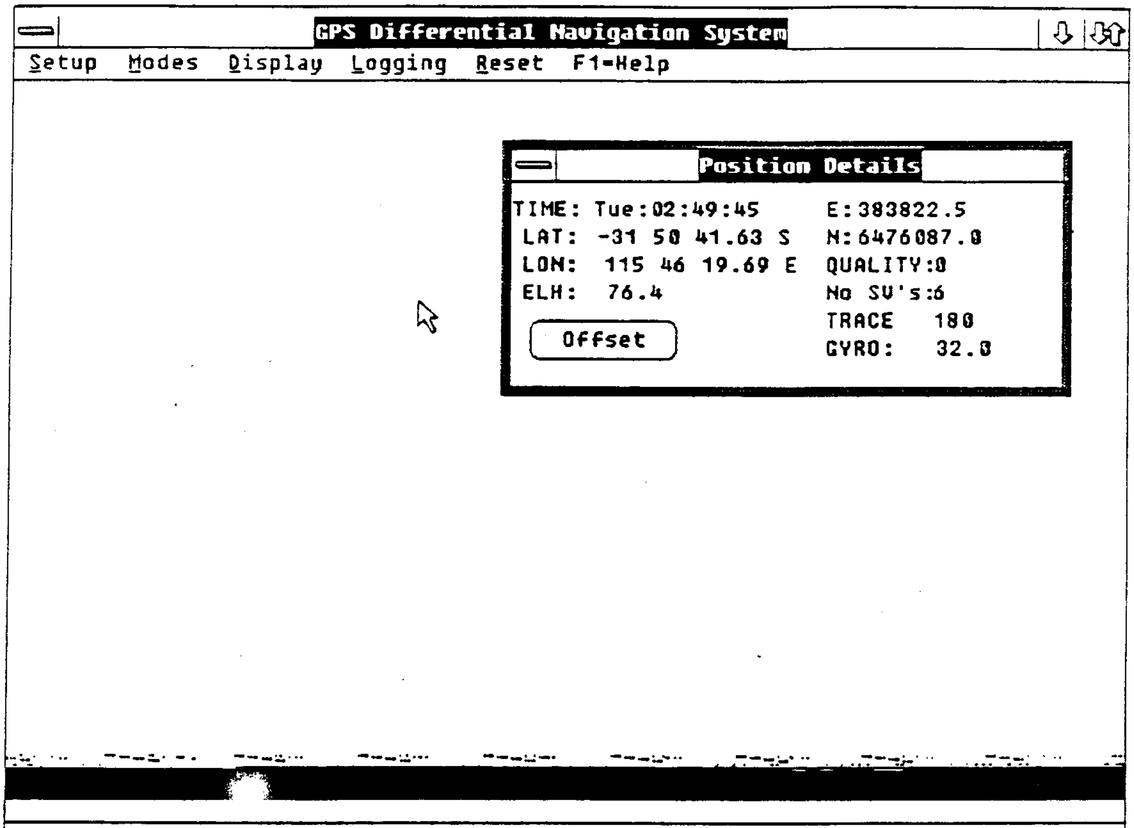
Figure 34: Display Windows Showing Ephemeris and Raw Data

Position

The position window (Figure 35) is the key window for text display of the users position. It is usually displayed with a combination of other windows. It contains in compact form the current position in both geographical and UTM coordinates, the number of satellites tracked as well as two factors concerning the quality of the solution, Quality and Trace.

The Quality parameter has now been replaced by a term called "Update". This parameter is simply the number of seconds that have elapsed since the last correction update. If the station is in base mode, this term is zero. The Trace term is the $2 \times d_{rms}$ discussed in section 5.4.4.

Also a button display highlights whether the current position being tracked is the GPS antenna or a predefined offset.



```
TIME: Tue:02:49:45          E:383822.5
 LAT:  -31 50 41.63 S       N:6476087.0
 LON:  115 46 19.69 E       QUALITY:0
 ELH:  76.4                 No SV's:6
                            TRACE    180
                            GYRO:    32.0
```

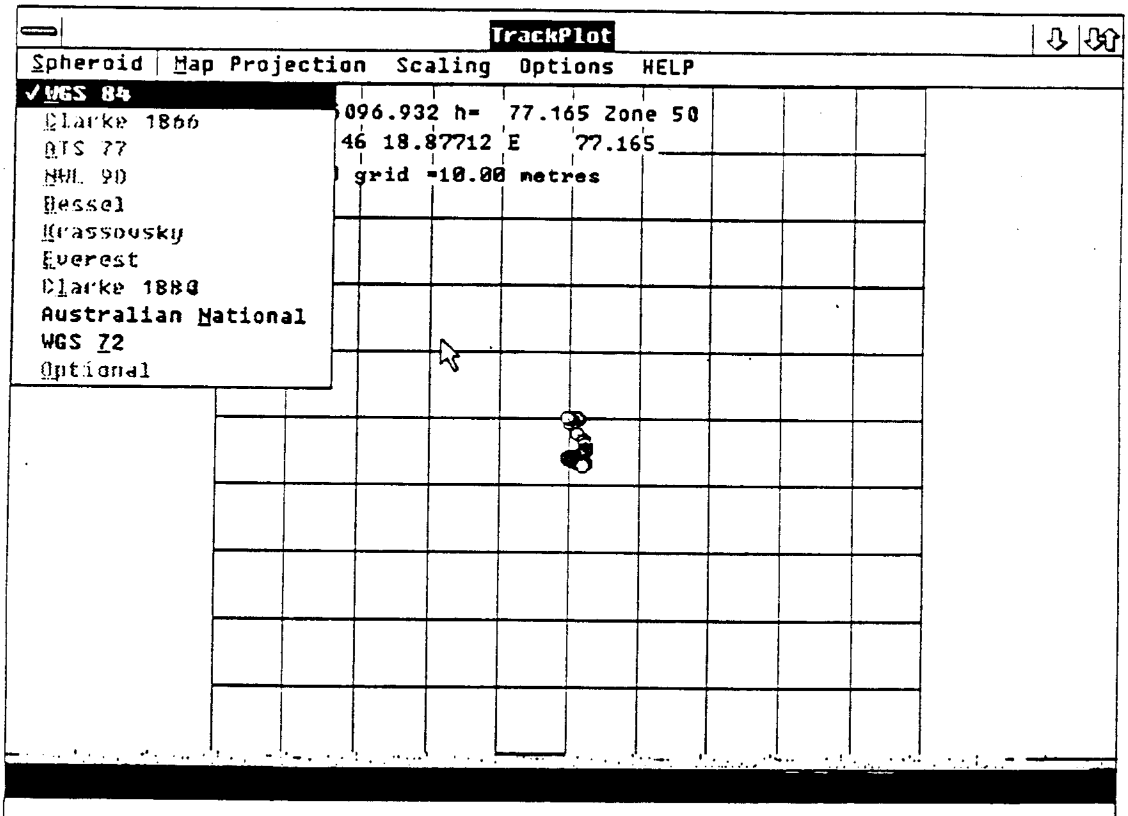
Figure 35: Position Window

TrackPlot

The track plot shown in figure 36 is a primitive version of the proposed track plot facility. This version simply plots the current position of the user on the screen at the currently selected scale. Every time the scale of the plot is changed the position history is lost. The text information shown on the screen is intended for debugging purposes only. This trackplot facility allows a much easier visual inspection of the smoothness of the position solutions. This is particularly true for the positions generated by the base station. By monitoring this plot the base station operator can obtain a good understanding as to the stability of his position solution.

For use as a navigational display this facility needs to be upgraded to include plotting of overlays, waypoints and runlines. These types of options can be implemented quite easily due to the graphics nature of the Windows environment.

The plot shown in figure 36 is the position history from one receiver in "Stand Alone" mode over a 10 minute period. It can be seen that the position solution skews some 10 metres during this time period. The spheroid menu selections are also shown in this figure. At present, both the spheroid type and map projection can be selected from the "Track Plot" menu. Whether this is the place to put this selection is currently being evaluated.



```
096.932 h=  77.165 Zone 50
46 18.87712 E     77.165
grid =10.00 metres
```

Figure 36: Track Plot Window

SkyPlot

The skyplot option shows a skychart where the satellite positions are plotted on a polar chart. The elevation is plotted as the radial distance against the satellite azimuths. These plots are very useful to access the effects of obstructions to the satellite signals. This also allows the user to monitor the strengths and weaknesses of the changing satellite constellation. Figure 37 shows a full screen sky chart. Other charts are shown in combinations with other data later in this section.

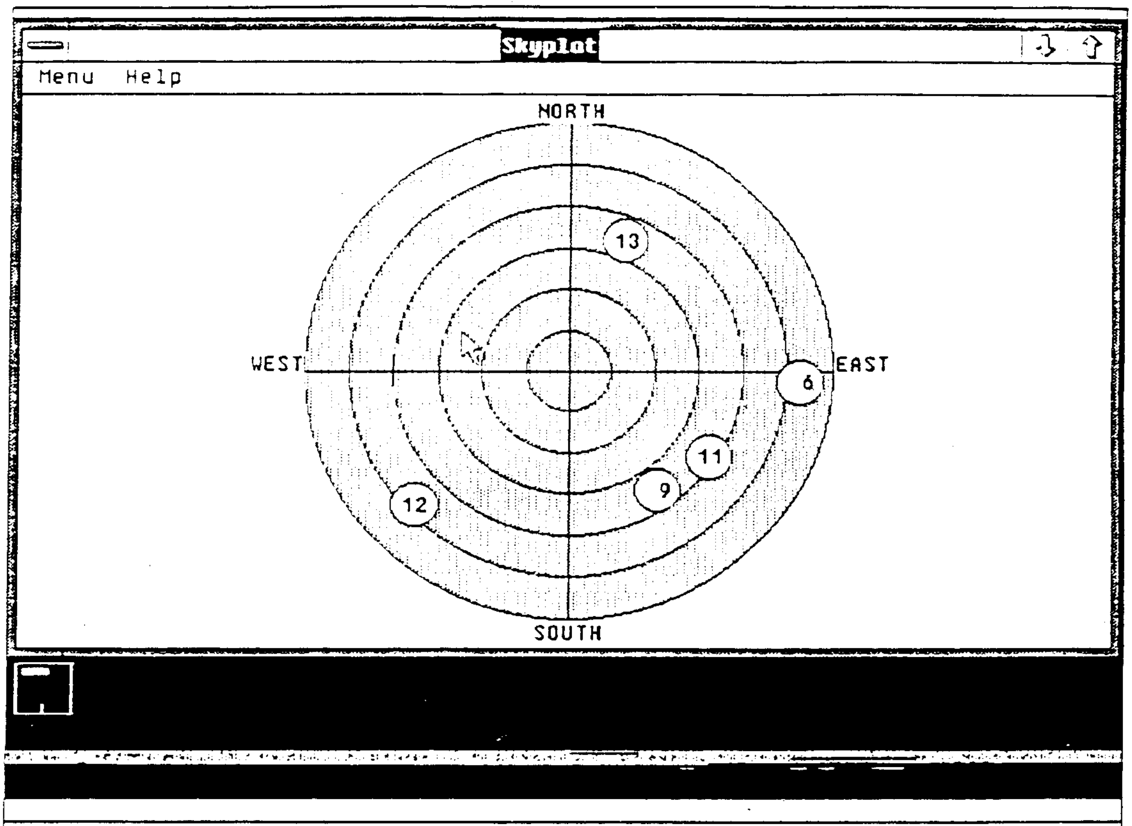
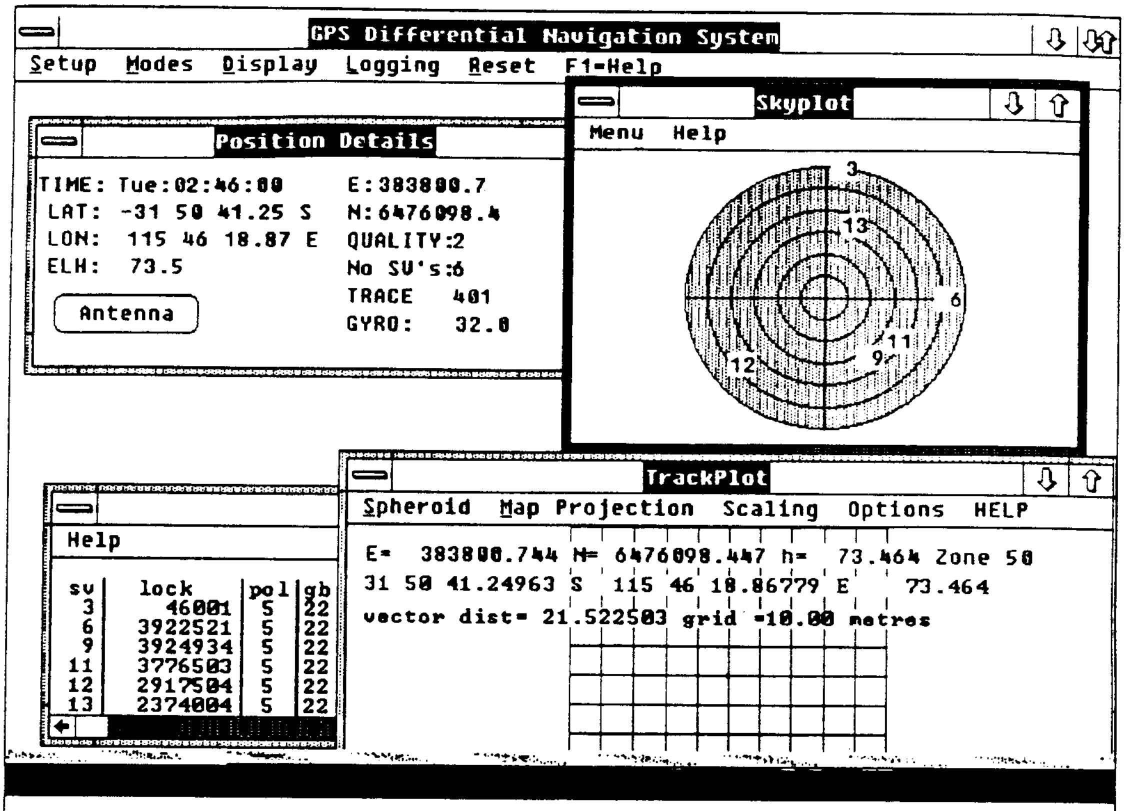


Figure 37: Sky Plot Window

Combinations

The advantage of having each of the above display types as separate windows is that each window can be displayed in combinations with other windows. Hence, the user can customise his screen to see exactly the information that he requires. Some examples of this are shown in figures 38,39 and 40.

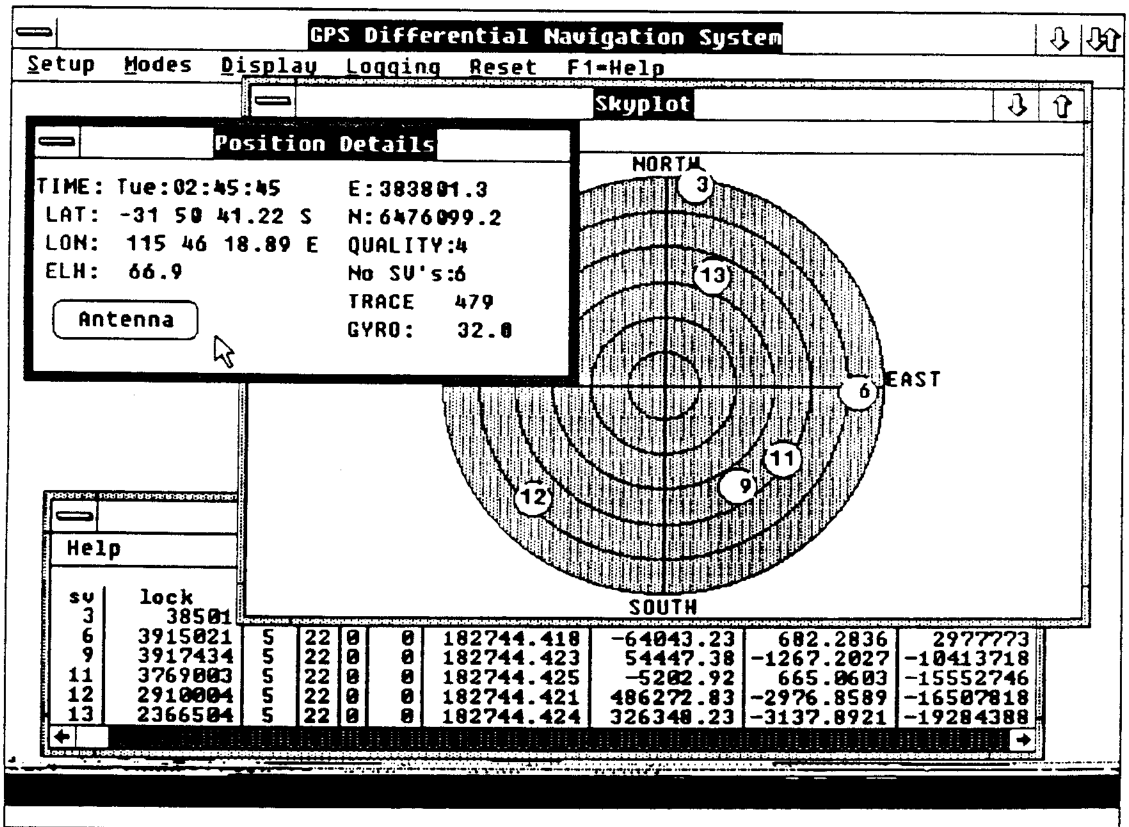


```
TIME: Tue:02:46:00          E:383800.7
 LAT:  -31 50 41.25 S       N:6476098.4
 LON:   115 46 18.87 E      QUALITY:2
 ELH:   73.5                No SU's:6
                            TRACE    401
                            GYRO:    32.0
```

```
 sv |    lock  |pol |gb
  3 |     46001|  5 |22
  6 |   3922521|  5 |22
  9 |   3924934|  5 |22
 11 |   3776503|  5 |22
 12 |   2917504|  5 |22
 13 |   2374004|  5 |22
```

```
E=  383800.744 N= 6476098.447 h=  73.464 Zone 50
31 50 41.24963 S  115 46 18.86779 E    73.464
vector dist= 21.522503 grid =10.00 metres
```

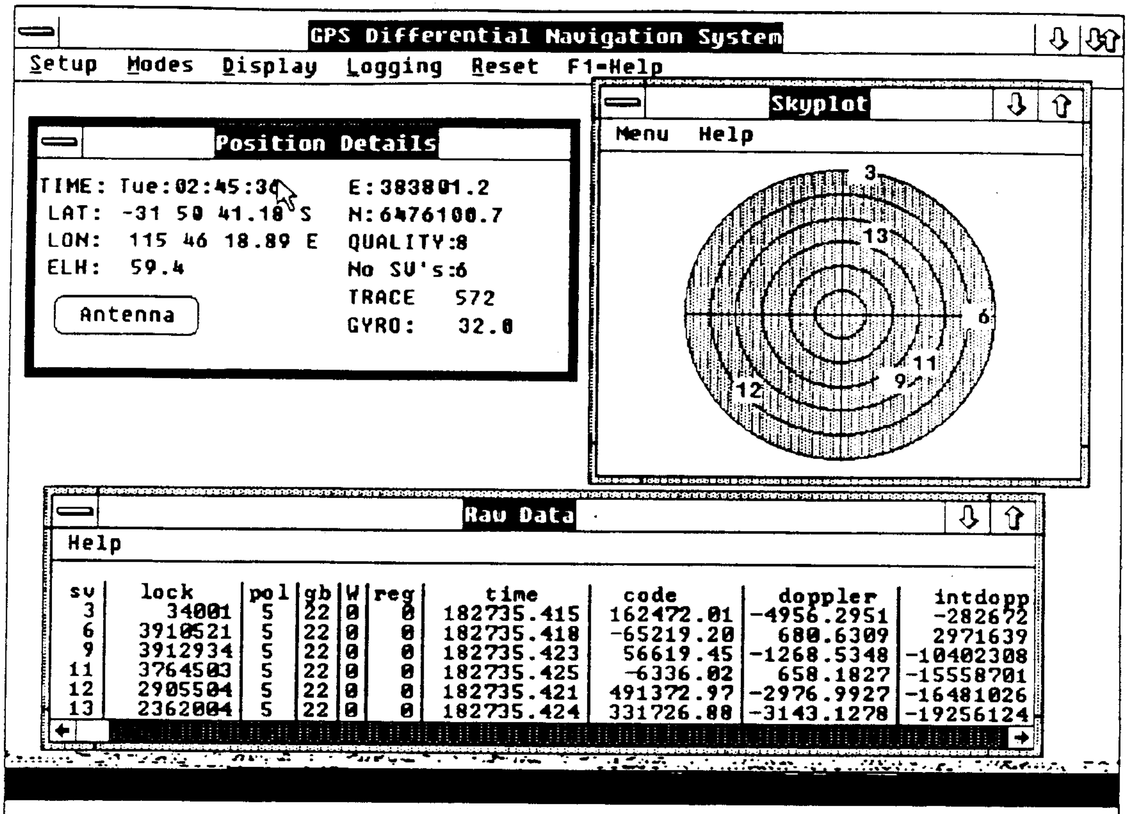
Figure 38: Example of Overlapping Windows



```
TIME: Tue:02:45:45          E:383801.3
 LAT:  -31 50 41.22 S       N:6476099.2
 LON:   115 46 18.89 E      QUALITY:4
 ELH:   66.9                No SU's:6
                            TRACE    479
                            GYRO:    32.0
```

```
 sv |    lock
  3 |     38501
  6 |   3915021| 5 |22|0|  0| 182744.418| -64043.23|  602.2836|   2977773
  9 |   3917434| 5 |22|0|  0| 182744.423|  54447.38|-1267.2027| -10413718
 11 |   3769003| 5 |22|0|  0| 182744.425|  -5202.92|  665.0603| -15552746
 12 |   2910004| 5 |22|0|  0| 182744.421| 486272.83|-2976.8589| -16507818
 13 |   2366504| 5 |22|0|  0| 182744.424| 326340.23|-3137.8921| -19284388
```

Figure 39: Example of Window Screen



TIME: Tue:02:45:36
LAT: -31 50 41.18 S
LON: 115 46 18.89 E
ELH: 59.4

E:383801.2
N:6476100.7
QUALITY:8
No SV's:6
TRACE 572
GYRO: 32.0

sv	lock	pol	gb	W	reg	time	code	doppler	intdopp
3	34001	5	22	0	0	182735.415	162472.01	-4956.2951	-282672
6	3910521	5	22	0	0	182735.418	-65219.20	680.6309	2971639
9	3912934	5	22	0	0	182735.423	56619.45	-1268.5348	-10402308
11	3764503	5	22	0	0	182735.425	-6336.02	658.1827	-15558701
12	2905504	5	22	0	0	182735.421	491372.97	-2976.9927	-16481026
13	2362004	5	22	0	0	182735.424	331726.88	-3143.1278	-19256124

Figure 40: Example of Window Screens

7.6 Logging Options

The logging menu is shown in figure 41. The three options in this menu are quite straight forward. When an option is enabled, a tick mark can be seen next to the option. The three options are as follows

TRAK

When this option is enabled, position output is sent through the printer port (LPT1) in a form suitable for input to the TRAKIV navigational computer. This option allows the position information to be combined with information from other sensors and output to be provided by the TRAKIV computer.

Printer

This option enables online output to a printer through serial port LPT2. This information is identical to the data being logged to disk using the Disk option. An example of the information being logged is shown in table 46. There are three types of information logged. The first type is the system setup information. This is logged whenever the user selects the "Save Settings" option in the "Setup" menu. The second type of information is the ephemeris information. This information is sent to the log whenever the "Display Ephemeris" option is selected. The third type of information is the position and range data which is logged whenever the position is being recorded.

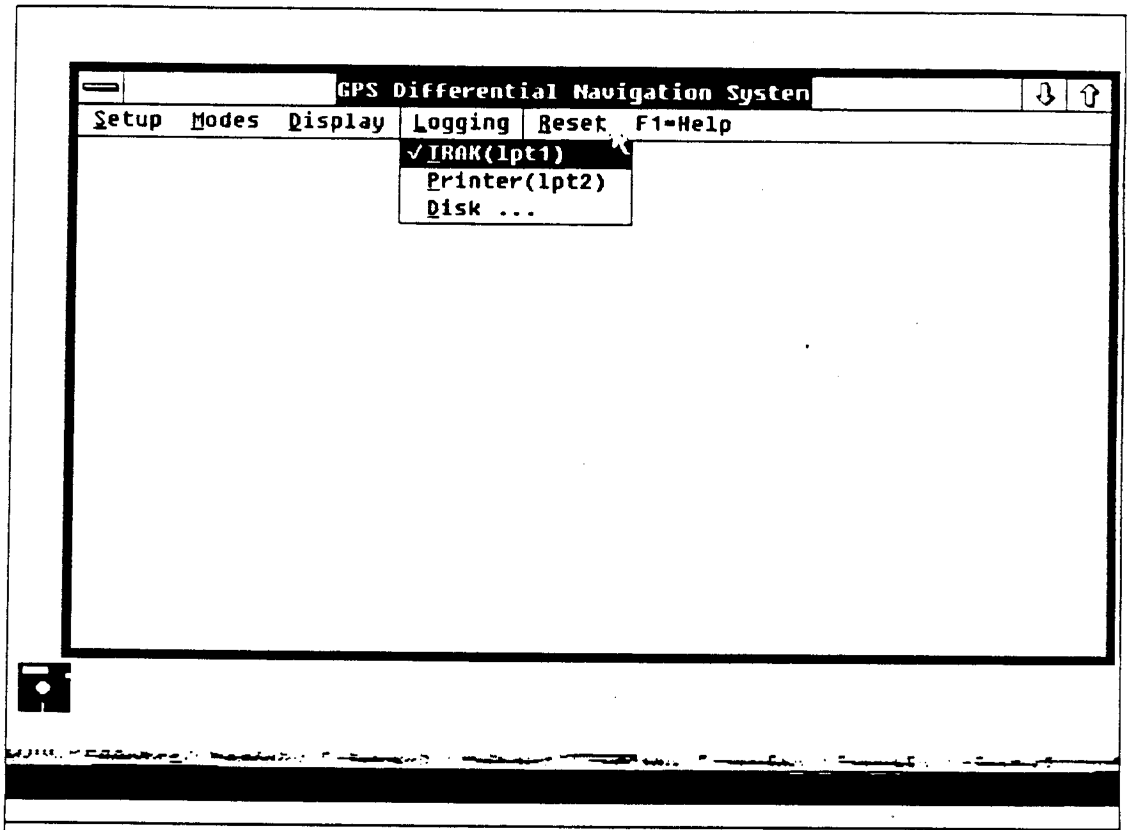


Figure 41: Logging Options

Disk

The Disk option invokes a dialogue box that prompts for the name of the disk file to which output is to be logged. Output in the identical form as described in the Printer Output setup is logged to this file. It should be noted that this logging to disk slows down the processing speed of the computer considerably. Hence, depending on the computer hardware, it may be necessary to increase the recording interval to accommodate disk logging.

```
Associated Surveys GPS Navigation Log
10th December 1989

Mode                     : Rover Station
Receiver Type         : Ashtech XIIL
Logging Interval     : 3 seconds
Elevation Mask       : 10 degrees
Min SV's Tracked     : 3
Operator name         : J.M.NOLAN
Modem type             : Transcom
Local ID                 : 0003
Remote ID               : 0004
Reference Spheroid : ANS
Map Projection       : UTM
Transformation parameters used
DX 116.000 DY 50.470 DZ -141.690 RX 0.23 RY 0.39 RZ 0.34 SC
0.0983
Initial Position (Altitude used in Altitude hold)
20 37 17.40076 S   116 44 51.27331 E      42.000
Reference Site       : Sale Office
20 37 17.40076 S   116 44 51.27331 E      42.000

sv|wk#| time eph received | time of ephemeris |AODE|H |Ac|Ft
--+---+-------------------+-------------------+----+--+--+--
03|513|456840-Fri:06:54:00|460800-Fri:08:00:00|0018| 0| 0| 0
06|513|456840-Fri:06:54:00|460800-Fri:08:00:00|0023| 0| 0| 0
09|513|456840-Fri:06:54:00|456704-Fri:06:51:44|0078| 0| 0| 0
11|513|456840-Fri:06:54:00|460800-Fri:08:00:00|0198| 0| 0| 0
12|513|456840-Fri:06:54:00|460800-Fri:08:00:00|0226| 0| 1| 0
13|513|456840-Fri:06:54:00|460800-Fri:08:00:00|0018| 0| 0| 0
16|513|456840-Fri:06:54:00|460800-Fri:08:00:00|0202|63| 7| 0


SV13 Range=20328423.6 Corr=2.0
SV12 Range=22941539.0 Corr=7.5
SV11 Range=21969992.4 Corr=32.6
SV09 Range=21992122.4 Corr=0.1
SV06 Range=23214646.1 Corr=-107.2
SV03 Range=22068268.3 Corr=86.4
Fri:07:00:21 E:447075.3 N:7839039.7 H:31.1 NoSVs 6 Trace 2.3 delta
12

SV13 Range=20327979.7 Corr=2.0
SV12 Range=22940251.3 Corr=7.5
SV11 Range=21971612.6 Corr=32.6
SV09 Range=21991444.5 Corr=0.1
SV06 Range=23214982.0 Corr=-107.2
SV03 Range=22066251.3 Corr=86.4
Fri:07:00:24 E:447080.9 N:7839039.6 H:30.7 NoSVs 6 Trace 2.3 delta
15
```

Table 46 Example of output log

8. TEST RESULTS

Validation and verification of a software package must be carefully monitored through each stage of the development process. The distinction between the terms, validation and verification is summarized by Boehm (1979):

- Validation: Are we building the right product?
- Verification: Are we building the product right?

Indeed, both these factors required attention during the development stages. In section 5 (Algorithm Design) it was shown that the system development could be subdivided into three separate stages. At each of these stages several checks are performed to verify the performance of the system.

Upon completion of the system, more elaborate testing is required to ensure that the system is operating correctly. This aims of this testing are as follows:

- To detect the presence of any bugs in the system so they can be corrected.
- To determine the accuracy of the system under a variety of conditions.
- To establish appropriate variances for the parameters outlined is section 5.4.4. This is to enable our calculated error statistic of $2 \times d_{rms}$ to be realistic.

This initial testing of the system, sometimes referred to as acceptance testing or alpha testing was carried out to establish these factors. With regard to detecting the presence of bugs, it should be noted that testing can prove only the presence of errors. If no errors are found, it does not prove that errors are not present.

Despite careful planning, many coding and design errors were discovered during the alpha testing phase.

8.1 Alpha Testing

One metre baseline test

All initial testing was performed using 2 receivers connected via a serial cable. Modems were not used in this set-up. The two antennas were placed in the same area separated by a few metres. This short baseline test effectively eliminates any errors due to orbit de-correlation and refraction effects. Also, with units connected without modems, update rates of one second could be obtained. This eliminates the possibility of any induced error due to the age of the transmitted correction.

Position results were recorded to disk file for subsequent analysis. By adopting a coordinate for the base station receiver, an accurate coordinate was determined for the stationary roving receiver. This known coordinate was then subtracted from each of the GPS determined positions to obtain position residuals.

The chief component of interest in this analysis is the horizontal error. Therefore the distance residual ($\sqrt{[\Delta E^2 + \Delta N^2]}$) is formed for inspection. Figure 42 shows a plot of the distance residual for a period of 100 minutes. The thinner line on this plot shows the calculated $2 \times d_{rms}$ statistic. This data was chosen as it highlights many factors.

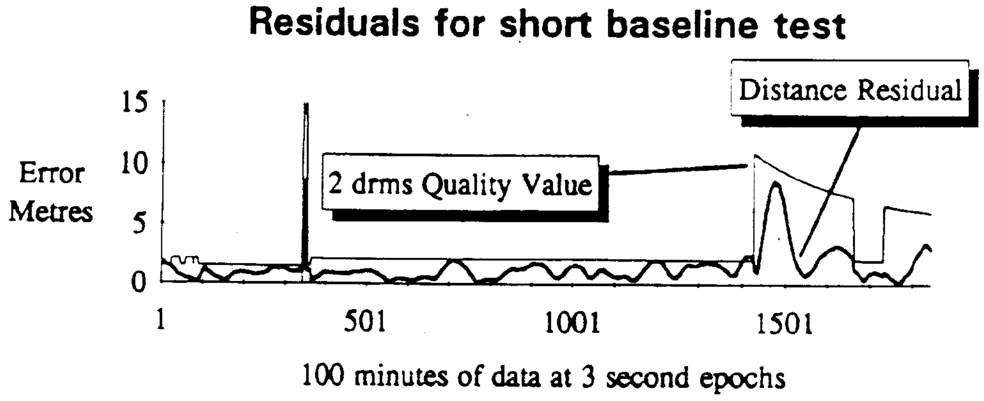


Figure 42: Short baseline test residuals

From this plot, it is evident that the distance residual is always less than the $2 \times d_{rms}$ statistic. At approximately epoch 350 a large spike can be seen. This is due to a loss of lock on one satellite. This left the navigation solution with only four satellites which evidently had poor geometry. Consequently, the distance residual for this period was seen to be as high as ten metres. However, this event was also recognised by the software and indicated by the $2 \times d_{rms}$ statistic. This statistic rose as high as thirty metres during this period. Hence, the user is adequately warned of the degraded solution during this period.

A less severe case, but equally important, occurs approximately at epoch 1430. (figure 42) At this time the number of satellites being tracked drops from five to four. It briefly returns to five satellites at epoch 1660 and back to four at epoch 1730. The solution obtained using the four satellite constellation at this time is still quite inferior to that of the five satellite constellation. Hence this change in the number of SVs used is clearly reflected in the plot of the $2 \times d_{rms}$ statistic shown in figure 42. The relationship between the $2 \times d_{rms}$ and the number of satellites tracked is shown more clearly in figure 43. This plot shows the both the quantities superimposed on an area chart. It can be easily seen that when the number of SVs drops to four, the $2 \times d_{rms}$ statistic increases.

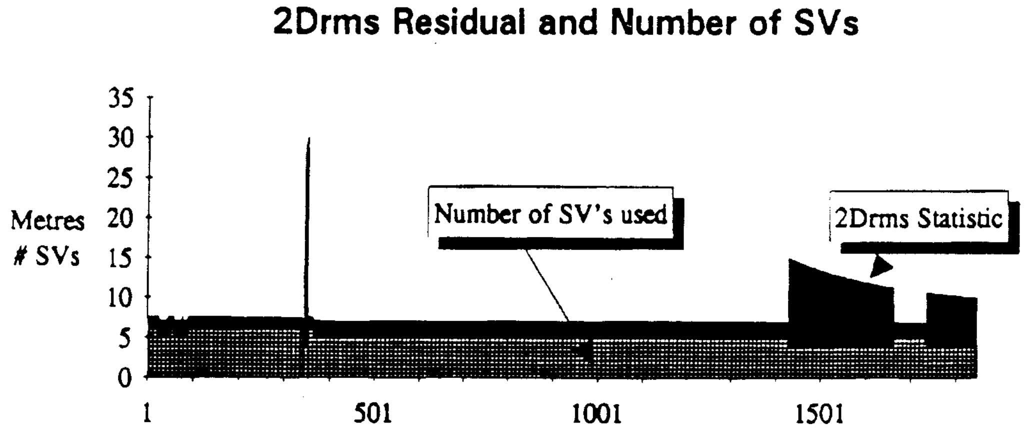


Figure 43: Area Plot showing # SVs and $2 \times d_{rms}$ statistic

Figure 44 shows essentially the same data as shown in figure 42 but this time in an area plot. In this plot the close relationship between the $2 \times d_{rms}$ statistic and the actual observed distance residuals can be seen.

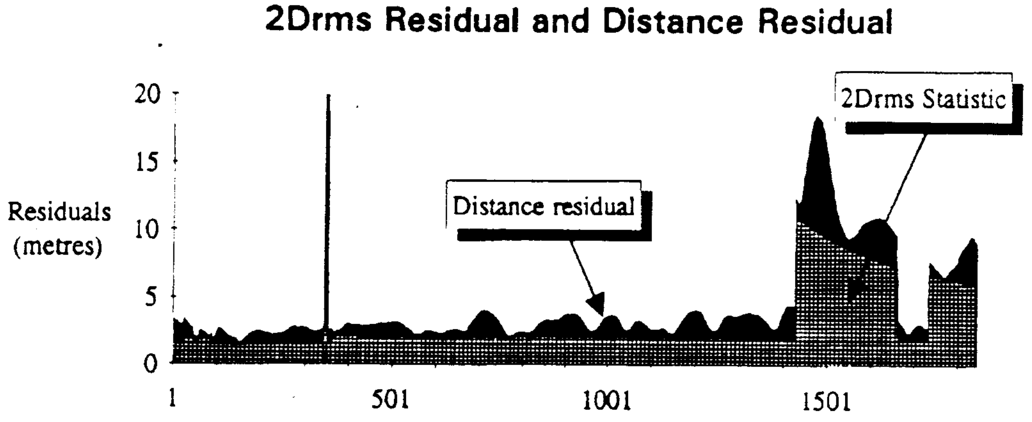


Figure 44: Area Plot showing observed residuals and $2 \times d_{rms}$ statistic

It is pleasing that the relationship between the observed and predicted accuracy is strong. However, it should be noted that the magnitude of the distance residuals is larger than expected. For the period of consistent five SV coverage, between epoch 400 and epoch 1400, the following statistics (shown in table 47) were obtained.

Mean Easting Residual = +0.094 m

Mean Northing Residual = -0.051 m

Variance Easting σEE = 0.650 m^2

Variance Northing σNN = 0.735 m^2

Std Deviation Easting σE = 0.806 m

Std Deviation Northing σN = 0.857 m

$2 \times D_{RMS} = 2 \times \sqrt{(\sigma_{EE} + \sigma_{NN})} = 2.354$ m

Circular Probability Error (50%) = 1.476 m

Note: CEP $\approx$ 1.774 σ (Ananda 1988)

Table 47: Summary of Test Statistics

Ananda (1988) shows the relationship between $2\times D_{RMS}$ and σ to be approximated by $2\times D_{RMS} = 2.828\ \sigma$. Hence the probability of any residual falling inside a circle of radius sphere of 2.354 metres is statistically less than one percent (2.8 standard deviation). There is a 50% probability that a residual will fall in the circle of radius 1.774 metres. This 50% probability statistic called Circular Error Probability (CEP) is widely used in GPS literature in quoting the accuracy of navigation systems.

This magnitude of error is quite acceptable for many types of navigation. However, it must be remembered that this is for a period of good satellite geometry and with all other error sources such as orbit and atmospheric effects eliminated.

In order to allow the error statistic to agree conservatively with the observed residuals, it was necessary to increase the magnitude of the apriori term σ_{BB} (see section 5.4.4). This term represents the variance of errors caused by interchannel biases and electronic path delays. From Ashjaee (1989), it is believed that the error due to this source (for Ashtech XII receivers) is in the order of a few centimetres. However, in order to obtain agreement between the observed and predicted residuals this term was increased to 0.04 centimetres^2.

Four Kilometre Baseline

Prior to field testing of the system, a full test was implemented over an accurately known 4 kilometre baseline. Effects due to refraction and orbit errors will still be minimal over this length baseline. However, routines for modem control and data transfer were evaluated. Also the effects of the delay in obtaining corrections could be directly accessed. VHF radios were used for this test with 300 baud modems. Although, 600 baud could be used with the VHF system it was considered more prudent to perform all tests with the slower baud rate. This was principally as, when using HF radios, 300 baud would be the maximum transmission rate.

Figure 45 shows a plot of the residuals for an hours data during this test. The data has been sorted in ascending order of the PDOP value. In figure 45 the PDOP is referred to as the Trace. This sorting clearly shows how the solution degrades with satellite geometry. During the period when the PDOP was low the solution accuracy was observed to be very similar to that shown in the initial short baseline tests. As the PDOP reaches 6 a very noisy patch of data can be seen. This was due to an error in the program logic which was highlighted by this test and the error rectified.

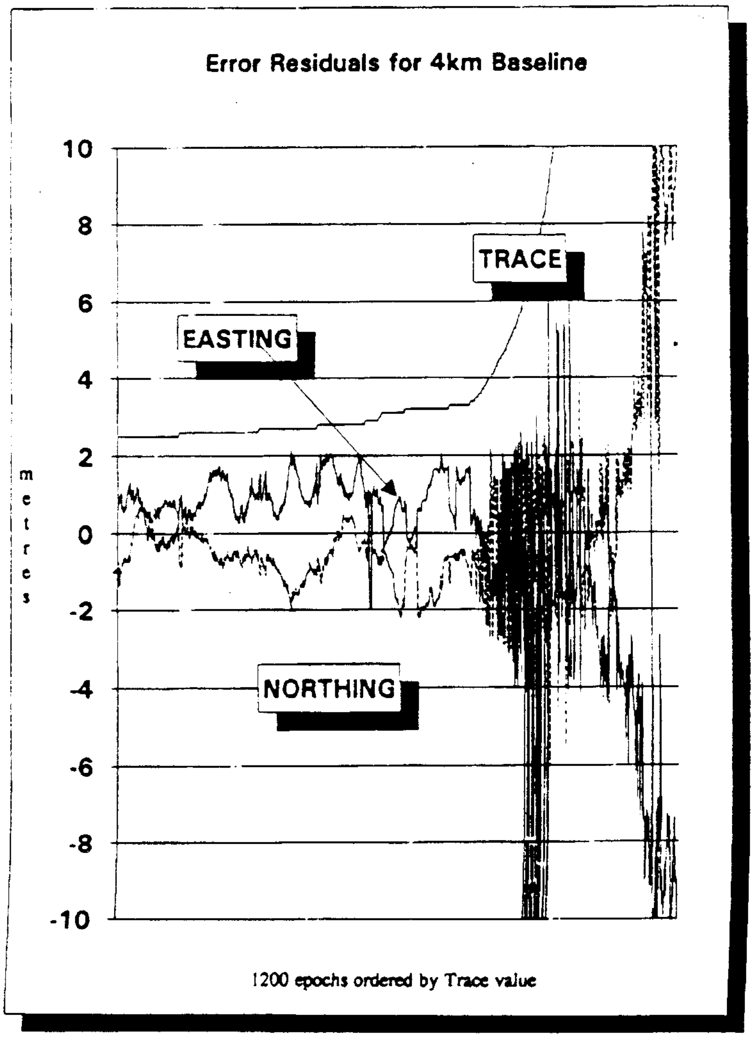


Figure 45: Residuals from 4km baseline trial

This 4km baseline test verified that all components of the program were functioning correctly. The system was providing position information to an accuracy of better than two metres $2 \times d_{rms}$ when the PDOP was below five. Consequently, the system was considered ready for beta testing.

8.2 Beta Testing

Polarfix Comparison

There are disadvantages in a developer performing all tests on his own system. This is due to the tendency for the developer not to design the test environment objectively. To ensure that all testing is impartial, beta testing is usually performed. In beta testing, the product is provided to a number of users who agree to use the system and report all system errors. It exposes the system to real use and problems that may not have been expected by the developer.

The system was tested during several in-house exercises by Associated Surveys International personnel. In one test the system was directly compared against a Polarfix navigation system. The Polarfix is a theodolite/EDM system that automatically tracks a set of prisms located at the mobile site. The base unit calculates the position of the moving vessel and transmits it to the mobile unit. Both the base GPS receiver and a Polarfix navigation system were set-up at the Freshwater Bay Yacht Club. During the exercise, both units logged data to a Qubit TRAKIV navigation computer. This allowed the position residual to be monitored at all times during the exercise.

This Comparison of the GPS navigation system and the Polarfix unit was carried out over two separate days and through a significant part of the GPS window. Observation of the residuals during this test confirmed that the system was functioning as intended. Unfortunately, no data was logged during these trials.

Transits

In off-shore work it is a normal calibration procedure to test the navigation system prior to commencing a survey. One common method used is to run transits between channel markers at accurately known positions. The vessel is steered along the line of two accurately coordinated channel markers. These markers are entered as a run-line into the navigation computer and the offset of the GPS antenna from the line of the channel markers is determined. When the vessel is on the line of the channel markers then it would be expected that this calculated offset distance be close to zero.

These transits provide a quick method of detecting if the navigation system is functioning correctly. Errors larger than five metres can be easily detected. In one case, an error of over one hundred metres was discovered. This was found to be due to corrupted parameters for the conversion between WGS84 and ANS84.

8.3 Conclusions

The tests conducted on the navigation system provide confidence that the system is working as intended. They highlight two important points. First that the system is quite stable when there is good satellite geometry. Secondly that the calculated error statistic of $2 \times D_{RMS}$ provides a reliable estimate of the accuracy of the system.

However, the observed accuracy of 2.17 metres ($2 \times D_{RMS}$) is higher than anticipated. This may be attributed to a larger than expected interchannel bias in the receivers. Alternatively, and perhaps more likely, it may be due to some deficiency in the current implementation of the system. The present system accuracy is more than adequate for most purposes. However, the reasons for this loss of accuracy is a concern that should be addressed. Also at present no serious analysis has been undertaken with regard to the degradation of the system accuracy with the increase in baseline length.

The current values adopted for the variances described in section 5.4.4 are shown in table 48.

σRR	=	$1.00m^2$	pseudo range measurement noise
sBB	=	$0.04m^2$	interchannel bias
sOO	=	$100.m^2$	orbit errors
sAA	=	$100.m^2$	refraction errors
sDD	=	$0.00001\ m^2/minute$	Correction drift

Table 48: Adopted Variances

Of these, σRR has been accurately determined from inspection of code phase residuals. The variance σBB was determined from analysis of results during the short baseline testing. This value was adjusted to ensure that the $2 \times d_{rms}$ statistic was conservative. All other values have been estimated.

9. SUMMARY AND CONCLUSIONS

In early 1990, six new block II satellites have been successfully deployed. It is the eve of the GPS becoming a 24 hour system. During the next two years it is anticipated that GPS will become a major navigational tool. It will virtually replace existing radio navigation and acoustic systems for many applications. However, in order for this to happen, the quality and quantity of available software for using the GPS system must increase. At present software development is lagging the rapid increases in hardware technology.

The system described in this paper introduces a software package that attempts to address this deficiency. The system uses the GPS system in a differential mode to provide accurate positioning. Initial test results show an accuracy of better than two metres ($2 \times d_{rms}$). The system has a range of several thousand kilometres. This is totally dependant on the quality of the communication link used. The maximum position update rate is one second. This rate is limited by the receiver hardware being used. Also when logging to disk, a 80386 processor is required to maintain this logging rate.

The system was written to run under the Microsoft Windows environment. This immediately provides several advantages. First, it provides a friendly graphical user interface. This interface will be immediately familiar to users of other Windows applications. Another important feature of the Windows environment is its multitasking capabilities. This allows other programs to be run concurrently. Windows also provides device independence and support for a large number of printers, plotters and video displays.

The goal of the system is to provide the user with accurate position. In using differential GPS, it is imperative that both units are using identical satellites and identical ephemeris. The design of this navigational system ensures that this is always the case.

One of the principal aims of the system is to provide the user with statistics concerning the quality of the solution. Solution quality is determined by the propagation of apriori variances. These include estimates of orbit errors, atmospheric errors, interchannel biases and code phase measurement noise.

A major source of error inherent in this type of system is caused by the presence of cycle slips. These slips occur when the receiver hardware cannot accurately determine the number of full carrier phase cycles that have passed between two consecutive epochs. The system detects all large cycle slips (greater than ten cycles) automatically. Smaller cycle slips are more difficult to detect. For the static base station receiver, cycle slips greater than one cycle can be detected. This is achieved using the instantaneous Doppler measurements.

The operator can monitor several statistics that indicate the presence of cycle slips. At present, these residual displays are purely numerical.

The system provides several implementation features which are briefly described below:

- All parameters can be altered at run time using the appropriate dialogue boxes. Alternatively, settings can be saved to an ASCII file and re-read later. These settings can naturally be altered with a text editor.
- Provision has been made for tracking of an offset position for times when the GPS antenna is not the point of interest.
- Corrections can be passed via voice links and entered manually. This facility is a back-up for times when data transmission is not possible.
- All receiver functions such as file names, recording intervals etc are controlled by the system. Hence, once set up, it is not necessary to touch the receiver.

- The altitude can be constrained to a fixed height in the solution. This allows navigation to continue with three satellites. Alternatively, this option may be required when four satellites are available but the satellite geometry is poor.

- The software can operate using one receiver in a stand alone mode. By setting the receiver over a known location, corrections can be determined. These corrections are then applied as the receiver visits other locations. These positions are initially quite accurate but degrade with time.

- The system provides a simple track plot facility. This allows the base station operator to monitor the stability of the base station corrections.

- Both raw data and abbreviated solution data can be logged on request.

It is anticipated that in the future there will be a demand for the type of software described in this paper. Consequently, every effort was made to allow the system to be capable of adapting to future needs. The structure of the program code was carefully designed with the aim of making it easy to maintain.

The modular structure of the program units also allow the device dependant code to be isolated. Hence it is a simple matter to include support for a wide number of receiver and modem types. All software was written in C, which is a very popular language for microprocessor development. A standard program layout was defined making the source code more readable.

At present the software is being Beta tested on commercial work. It has proven itself to be adequate for many positioning tasks. However, there are many features and improvements that must be addressed. A few of these features are as follows:

- Inclusion of range rate corrections in the correction message. This would reduce the reliance of the system on having recent correction updates.

- Support for a wider number of receivers and modems. At present, the software supports only the Ashtech XII receiver and two types of modems.
- Incorporation of a sophisticated graphics display with zooming and panning and position memory.
- Graphical representation of all residuals. This would allow the user to access the quality of the solution much more easily.
- More elaborate cycle slip checking method for the case when good data has been obtained on four satellites.
- Inclusion of full on-line help facilities and a reference manual.

It is recognised that the system is far from completed. At present it successfully performs its primary function in a reasonable friendly manner. However, the true success or failure of this type of system will lie in its ability to evolve. If it has been designed well, then it will be around long enough to eliminate its inefficiencies.

BIBLIOGRAPHY

Allison T. (1989). Real Time Cycle-Slip Fixing During Kinematic Surveys. Proceedings of 5th International Geodetic Symposium on Satellite Positioning. New Mexico State University. March 13-17 1989. Vol.1, pp 330-337.

Ananda M. (1988). The Global Positioning System (GPS) Accuracy, System Error Budget, Space and Control Segment Overview. International Symposia, Global Positioning Systems. Department of Geographic Information, Gold Coast, Queensland. October 17-19, 1988. pp I.2.1-I.2.19.

Ashjaee J. (1986a). GPS Doppler Processing for Precise Positioning in Dynamic Applications. Global Positioning System. Papers published in Navigation. Journal of the (U.S.) Institute of Navigation, Washington, D.C. Vol.III, pp 54-69.

Ashjaee J. (1986b). Real Time First Order Survey with GPS Single Frequency C/A Code. Proceedings of the Fourth International Geodetic Symposium on Satellite Positioning. Austin, Texas. April 28 - May 2, 1986. Vol.1 pp 489-502.

Ashjaee J. (1988), On Precision of GPS C/A Code. Based on a paper presented at PLANS 1986. IEEE AES Magazine, June 1988.

Ashjaee J. (1989). Precision Survey with Ashtech XII, The All-in-One, All-in-View. Proceedings of the Fifth International Geodetic Symposium on Satellite Positioning. New Mexico State University. March 13-17, 1989. Vol.I pp 316-329.

Ashjaee J., Lorenz R., Sutherland R., Dutilloy J., Minazio J-B, Abtahi R., Eichner J-M., Kosmalska J., Helky R. (1989). New GPS Developments and Ashtech M-XII. ION-GPS 89, Colorado Springs, September 1989.

Berveots S.G. (1987). Variance and Bias in GPS Positioning. Centenary GPS Conference, Melbourne, August 1987.

Boehm B.W. (1979). Software Engineering: R & D trends and defense needs, in Research Directions in Software Technology, Wegner, P. (ed.) Cambridge, Mas.: MIT Press.

Bomford, G. (1980). Geodesy, Fourth Edition. Clarendon Press, Oxford. p.679.

Castalanelli M.A., Nolan J.M. (1988). Testing and Evaluation of Real Time Positioning Using Differenced Pseudo Range Solutions. International Symposia, Global Positioning Systems. Department of Geographic Information, Gold Coast, Queensland. October 17-19, 1988. pp VIII.3.1-VIII.3.21.

Chisholm G.S. (1987). Integration of GPS into Hydrographic Survey Operations. Unisurv Report S20,1987. School of Surveying, University of New South Wales, Kensington. Received July 1987.

Cross P.A. (1983). Advanced Least Squares Applied to Position Fixing. Paper No.6, North East London Poly., UK.

Finn A., Matthewman J. (1989). A Single-Frequency Ionospheric Refraction-Correction Algorithm for TRANSIT and GPS. Proceedings of the 5th International Geodetic Symposium on Satellite Positioning, New Mexico State University. March 13-17, 1989. Vol.II, pp.737-756.

Goad C.C., (1985). Precise Relative Position Determination Using Global Positioning System Carrier Phase Measurements in a Nondifference Mode. Proceedings of the First International Symposium on Precise Positioning with the Global Positioning System. April 15-19, 1985. Rockville, Maryland. Vol.1, pp.347-356.

Gurtner W. (1989). A Common Exchange Format for GPS Data. Proceedings of the Fifth International Geodetic Symposium on Satellite Positioning, New Mexico State University. March 13-17, 1989. Vol.II, pp 920-931.

Hatch R. (1982). The Synergism of GPS Code and Carrier Measurements. Proceedings of the Third International Goedetic Symposium on Satellite Doppler Positioning. New Mexico State University. February 8-12, 1982. Vol. 2. pp 1213-1228.

ICD-GPS-200 (1984). Interface Control Document. Satellite Systems Division, Rockwell International Corporation.

Janes H.W., Langley R.B., Newby S.P. (1989). A Comparison of Several Models for the Prediction of Tropospheric Propagating Delay. Proceedings of the Fifth International Geodetic Symposium on Satellite Positioning, New Mexico State University. March 13-17, 1989. Vol.II, pp.777-788.

Jorgensen P.S. (1980). Navstar/Global Positioning System 18-Satellite Constellations. Global Positioning System. Papers published in Navigation. Journal of the (U.S.) Institute of Navigation, Washington, D.C. Vol.II, pp 1-12.

Kalafus R.M., Van Dierendonck A.J., Pealer N.A. (1986). Special Committee 104. Recommendations for Differential GPS Service. Global Positioning System. Papers published in Navigation. The Institute of Navigation, Washington. Vol.III, pp 101-116.

King R.W., Masters E.G., Rizos C, Stolz A, Collins J. (1985), Surveying with GPS, Monograph 9, School of Surveying, University of New South Wales, Kensington, NSW.

Petzold C. (1988). Programming Windows. Microsoft Press, a Division of Microsoft Corporation. Washington.

Remondi B.W. (1984). Using the Global Positioning System (GPS) phase observable for relative geodesy: modelling, processing and results. PhD Dissertation, University for Space Research, The University of Texas in Austin.

Remondi B.W. (1985a). Modeling the GPS Carrier Phase for Geodetic Applications. Proceedings of the First International Symposium on Precise Positioning with the Global Positioning System. Rockville, Maryland. April 15-19, 1985. Vol.1 pp.325-336.

Remondi B.W. (1985b). Performing Centimeter Accuracy Relative Surveys in Seconds Using GPS Carrier Phase. Proceedings of the First International Symposium on Precise Positioning with the Global Positioning System. Rockville, Maryland. April 15-19, 1985. Vol.2 pp.789-807.

Remondi B.W. (1988). Kinematic and Pseudo-Kinematic GPS. Internal paper, Ashtech Inc., Sunnyvale, Calif.

Remondi B.W. (1989). GPS Broadcast Orbits Versus Precise Orbits: A Comparative Study. GPS Bulletin, Coordination of Space Technifques for Geodesy and Geodynamics. November/December 1989. Vol.2 No.6.

Robertson L.A. (1989). Simple Program Design. Thomas Nelson Australia, Melbourne.

Scherrer R. (1985). The WM GPS Primer. WILD Publication, Heerbrugg, Switzerland.

Sommerville I. (1989) Software Engineering. Third Edition. Addison-Wesley Publishing Company, England.

Spilker J.J. (1978). GPS Signal Structure and Performance Characteristics, Navigation, Journal of the (U.S.) Institute of Navigation, Washington, D.C. Vol. 25, No.2, pp 121-146.

Talbot N.C. (1987). GPS Signal Structure and Receiver Operation. Centenary GPS Conference. Royal Melbourne Institute of Technology, Melbourne.

Talbot N. (1988). Optimal Weighting of GPS Carrier Phase Observations Based on the Signal-to-Noise Ratio. International Symposia, Global Positioning Systems. Department of Geographic Information, Gold Coast, Queensland. October 17-19, 1988. pp V4 1-17.

Van Dierendonck A.J., Russell S.S., Kopitzke E.R., Birnbaum M. (1978), The GPS Navigation Message. Global Positioning System. Papers published in Navigation, Journal of the (U.S.) Institute of Navigation, Washington, D.C. Vol.I pp 55-73

White L.A. (1986). Calculus of Observations 380/580. 1986 Edition. Department of Surveying and Mapping, Western Australian Institute of Technology, Perth.

Publications from

THE SCHOOL OF SURVEYING, THE UNIVERSITY OF NEW SOUTH WALES.

All prices include postage by surface mail. Air mail rates on application. (Effective Nov. 1990)

To order, write to Publications Officer, School of Surveying, The University of New South Wales,
P.O. Box 1, Kensington N.S.W., 2033 AUSTRALIA

NOTE: ALL ORDERS MUST BE PREPAID

UNISURV REPORTS - G SERIES

Price (including postage): $6.00

G14. A. Stolz, "The computation of three dimensional Cartesian coordinates of terrestrial networks by the use of local astronomic vector systems", Unisurv Rep. 18, 47 pp, 1970.

G16. R.S. Mather et al, "Communications from Australia to Section V, International Association of Geodesy, XV General Assembly, International Union of Geodesy and Geophysics, Moscow 1971", Unisurv Rep. 22, 72 pp, 1971.

G17. Papers by R.S. Mather, H.L. Mitchell & A. Stolz on the following topics:- Four-dimensional geodesy, Network adjustment and Sea surface topography, Unisurv G17, 73 pp, 1972.

G18. Papers by L. Berlin, G.J.F. Holden, P.V. Angus-Leppan, H.L. Mitchell & A.H. Campbell on the following topics:- Photogrammetry co-ordinate systems for surveying integration, Geopotential networks and Linear measurement, Unisurv G18, 80 pp, 1972.

G19. R.S. Mather, P.V. Angus-Leppan, A. Stolz & I. Lloyd, "Aspects of four-dimensional geodesy", Unisurv G19, 100 pp, 1973.

G20. Papers by J.S. Allman, R.C. Lister, J.C. Trinder & R.S. Mather on the following topics:- Network adjustments, Photogrammetry, and 4-Dimensional geodesy, Unisurv G20, 133 pp, 1974.

G21. Papers by E. Grafarend, R.S. Mather & P.V. Angus-Leppan on the following topics:- Mathematical geodesy, Coastal geodesy and Refraction, Unisurv G21, 100 pp, 1974.

G22. Papers by R.S. Mather, J.R. Gilliland, F.K. Brunner, J.C. Trinder, K. Bretreger & G. Halsey on the following topics:- Gravity, Levelling, Refraction, ERTS imagery, Tidal effects on satellite orbits and Photogrammetry, Unisurv G22, 96 pp, 1975.

G23. Papers by R.S. Mather, E.G. Anderson, C. Rizos, K. Bretreger, K. Leppert, B.V. Hamon & P.V. Angus-Leppan on the following topics:- Earth tides, Sea surface topography, Atmospheric effects in physical geodesy, Mean sea level and Systematic errors in levelling, Unisurv G23, 96 pp, 1975.

G24. Papers by R.C. Patterson, R.S. Mather, R. Coleman, O.L. Colombo, J.C. Trinder, S.U. Nasca, T.L. Duyet & K. Bretreger on the following topics:- Adjustment theory, Sea surface topography determinations, Applications of LANDSAT imagery, Ocean loading of Earth tides, Physical geodesy, Photogrammetry and Oceanographic applications of satellites, Unisurv G24, 151 pp, 1976.

G25. Papers by S.M. Nakiboglu, B. Ducarme, P. Melchior, R.S. Mather, B.C. Barlow, C. Rizos, B. Hirsch, K. Bretreger, F.K. Brunner & P.V. Angus-Leppan on the following topics:- Hydrostatic equilibrium figures of the Earth, Earth tides, Gravity anomaly data banks for Australia, Recovery of tidal signals from satellite altimetry, Meteorological parameters for modelling terrestrial refraction and Crustal motion studies in Australia, Unisurv G25, 124 pp, 1976.

G26. Papers by R.S. Mather, E.G. Masters, R. Coleman, C. Rizos, B. Hirsch, C.S. Fraser, F.K. Brunner, P.V. Angus-Leppan, A.J. McCarthy & C. Wardrop on the following topics:- Four-dimensional geodesy, GEOS-3 altimetry data analysis, analysis of meteorological measurements for microwave EDM and Meteorological data logging system for geodetic refraction research, Unisurv G26, 113 pp, 1977.

G27. Papers by F.K. Brunner, C.S. Fraser, S.U. Nasca, J.C. Trinder, L. Berlin, R.S. Mather, O.L. Colombo & P.V. Angus-Leppan on the following topics:- Micrometeorology in geodetic refraction, LANDSAT imagery in topographic mapping, adjustment of large systems, GEOS-3 data analysis, Kernel functions and EDM reductions over sea, Unisurv G27, 101 pp, 1977.

G29. Papers by F.L. Clarke, R.S. Mather, D.R. Larden & J.R. Gilliland on the following topics:- Three dimensional network adjustment incorporating ξ , η and N, Geoid determinations with satellite altimetry, Geodynamic information from secular gravity changes and Height and free-air anomaly correlation, Unisurv G29, 87 pp, 1978.

From June 1979 Unisurv G's name was changed to Australian Journal of Geodesy, Photogrammetry and Surveying. These can be ordered from The Managing Editor, Australian Journal of Geodesy, Photogrammetry and Surveying, Institution of Surveyors - Australia , Nos 27 - 29 Napier Close, Deakin, ACT 2600, AUSTRALIA.

UNISURV REPORTS - S SERIES

S8 - S20	Price (including postage):		$10.00
S27 onwards	Price (including postage):	Individuals	$20.00
		Institutions	$28.00

S8 A. Stolz, "Three-D Cartesian co-ordinates of part of the Australian geodetic network by the use of local astronomic vector systems", Unisurv Rep. S 8, 182 pp, 1972.

S9 H.L. Mitchell, "Relations between MSL & geodetic levelling in Australia", Unisurv Rep. S 9, 264 pp, 1973.

S10 A.J. Robinson, "Study of zero error & ground swing of the model MRA101 tellurometer", Unisurv Rep. S 10, 200 pp, 1973.

S12. G.J.F. Holden, "An evaluation of orthophotography in an integrated mapping system", Unisurv Rep. S 12, 232 pp, 1974.

S14. Edward G. Anderson, "The Effect of Topography on Solutions of Stokes' Problem", Unisurv Rep. S 14, 252 pp, 1976.

S15. A.H.W. Kearsley, "The Computation of Deflections of the Vertical from Gravity Anomalies", Unisurv Rep. S 15, 181 pp, 1976.

S16. K. Bretreger, "Earth Tide Effects on Geodetic Observations", Unisurv S 16, 173 pp, 1978.

S17. C. Rizos, "The role of the gravity field in sea surface topography studies", Unisurv S 17, 299 pp, 1980.

S18. B.C. Forster, "Some measures of urban residual quality from LANDSAT multi-spectral data", Unisurv S 18, 223 pp, 1981.

S19. Richard Coleman, "A Geodetic Basis for recovering Ocean Dynamic Information from Satellite Altimetry", Unisurv S 19,332 pp, 1981.

S20. Douglas R. Larden, "Monitoring the Earth's Rotation by Lunar Laser Ranging", Unisurv Report S 20, 280 pp, 1982.

S27. Bruce R. Harvey, "The Combination of VLBI and Ground Data for Geodesy and Geophysics", Unisurv Report S27, 239 pp, 1985.

S29. Gary S. Chisholm, "Integration of GPS into hydrographic survey operations", Unisurv S29, 190 pp, 1987.

S30. Gary Alan Jeffress, "An investigation of Doppler satellite positioning multi-station software", Unisurv S30, 118 pp, 1987.

S31. Jahja Soetandi, "A model for a cadastral land information system for Indonesia", Unisurv S31, 168 pp, 1988.

S32. D. B. Grant, "Combination of terrestrial and GPS data for earth deformation studies" Unisurv S32, 285 pp, 1990.

S33. R. D. Holloway, "The integration of GPS heights into the Australian Height Datum", Unisurv S33, 151 pp.,1988.

S34. Robin C. Mullin, "Data update in a Land Information Network", Unisurv S34, 168 pp. 1988.

S35. Bertrand Merminod, "The use of Kalman filters in GPS Navigation", Unisurv S35, 203 pp., 1989.

S36. Andrew R. Marshall; "Network design and optimisation in close range Photogrammetry", Unisurv S36, 249 pp., 1989.

S37. Wattana Jaroondhampinij, "A model of Computerised parcel-based Land Information System for the Department of Lands, Thailand," Unisurv S37, 281 pp., 1989.

S38. C. Rizos (Ed.), D.B. Grant, A. Stolz, B. Merminod, C.C. Mazur "Contributions to GPS Studies", Unisurv S38, 204 pp., 1990.

S39. C. Bosloper, "Multipath and GPS short periodic components of the time variation of the differential dispersive delay", Unisurv S39, 214 pp., 1990.

S40. John Michael Nolan, "Development of a Navigational System utilizing the Global Positioning System in a real time, differential mode", Unisurv S40, 163 pp., 1990.

S41. Roderick T. Macleod, "The resolution of Mean Sea Level anomalies along the NSW coastline using the Global Positioning System", 278 pp., 1990.

PROCEEDINGS

Prices include postage by surface mail

P1. P.V. Angus-Leppan (Editor), "Proceedings of conference on refraction effects in geodesy & electronic distance measurement", 264 pp., 1968. Price: $10.00

P2. R.S. Mather & P.V. Angus-Leppan (Eds), "Australian Academy of Science/International Association of Geodesy Symposium on Earth's Gravitational Field & Secular Variations in Position", 740 pp., 1973. Price $15.00

MONOGRAPHS

Prices include postage by surface mail

M1.	R.S. Mather, "The theory and geodetic use of some common projections", (2nd edition), 125 pp., 1978.	Price	$12.00
M2.	R.S. Mather, "The analysis of the earth's gravity field", 172 pp., 1971.	Price	$8.00
M3.	G.G. Bennett, "Tables for prediction of daylight stars", 24 pp.,1974.	Price	$5.00
M4.	G.G. Bennett, J.G. Freislich & M. Maughan, "Star prediction tables for the fixing of position", 200 pp., 1974.	Price	$8.00
M5.	M. Maughan, "Survey computations", 98 pp., 1975.	Price	$11.00
M7.	J.M. Rueger, "Introduction to Electronic Distance Measurement", (2nd Edition), 140 pp., 1988.	Price	$17.00
M8.	A.H.W. Kearsley, "Geodetic Surveying". 77pp., 1988.	Price	$11.00
M10.	W. Faig, "Aerial Triangulation and Digital Mapping", 102. pp., 1986.	Price	$16.00
M11.	W.F. Caspary, "Concepts of Network and Deformation Analysis", 183 pp., 1988.	Price	$25.00
M12.	F.K. Brunner, "Atmospheric Effects on Geodetic Space Measurements", 110 pp., 1988.	Price	$16.00
M13	Bruce R. Harvey, "Practical Least Squares and Statistics for Surveyors", 229 pp., 1990.	Price	$23.00