Aalto University
School of Engineering

Global Navigation Satellite System (GNSS) Satellite and Pseudolite Signal Visibility simulation in Obstructed Environments

Master's Thesis, Department of Real Estate, Planning and Geoinformatics, School of Engineering, Aalto University

Espoo, October 3, 2012

Barchelor of Science in Technology
Esa-Pekka Sundell

Thesis supervisor: Professor Martin Vermeer
Thesis advisor: M.Sc. (Tech.) Heikki Laitinen

Aalto University
School of Engineering

P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of master's thesis

Author Esa-Pekka Sundell

Title of thesis Global Navigation Satellite System (GNSS) Satellite and Pseudolite Signal Visibility simulation in Obstructed Environments

Department Department of Real Estate, Planning and Geoinformatics

Professorship Geodesy **Code of Professorship** Maa-6

Thesis supervisor Professor Martin Vermeer

Thesis advisor M.Sc. (Tech.) Heikki Laitinen

Date October 3, 2012 **Number of pages** 68 **Language** English

Abstract Navigation with Global Navigation Satellite Systems (GNSS) works well in environments with clear sky, but navigation accuracy and availability is easily degraded when parts of the sky are blocked by terrain or solid structures. However, these kinds of obstructed environments sometimes require accurate navigation service to ensure safety and efficiency of operations, for example in an open-pit mine or a harbor. In these cases, land-based GNSS transmitters called pseudolites can provide additional navigation signals that can ensure the availability of navigation signals and enhance the accuracy back to good levels. Because of the many obstructions to signal propagation, there is a need for careful planning of the pseudolite locations. This research provides a useful tool for this purpose. The main task of this research is the development of new methodology and software for simulating the signal visibility of GNSS satellites and pseudolites, and for visualizing the *expected positioning accuracy* for a set of visible satellite and pseudolite signals. Techniques for simulating the direct signal propagation in complex three-dimensional environments are presented. The theory of GNSS navigation in general and the dilution of precision (DOP) values are explained and used as a basis for the definition for a good positioning accuracy. The result of this research is working software that simulates the signal visibilities and the quality of expected positioning accuracy in any environment, with and without pseudolites. Its performance is displayed with two test cases of environments where a pseudolite augmentation system would be feasible: a deep open-pit mine and a harbor pier with tall harbor cranes.

Keywords Global Navigation Satellite Systems, open-pit mine, harbor, pseudolites, signal propagation modeling, signal visibility, positioning accuracy, three-dimensional environment model

Aalto-yliopisto
Insinööritieteiden korkeakoulu

PL 11000, 00076 AALTO
www.aalto.fi
Diplomityön tiivistelmä

Tekijä Esa-Pekka Sundell

Työn nimi Global Navigation Satellite System (GNSS) satelliittien ja pseudoliittien signaalien kuuluvuuden mallinnus esteisessä ympäristössä

Laitos Maankäyttötieteiden laitos

Professuuri Geodesia **Professuurikoodi** Maa-6

Työn valvoja Professori Martin Vermeer

Työn ohjaaja Diplomi-insinööri Heikki Laitinen

Päivämäärä 3. lokakuuta 2012 **Sivumäärä** 68 **Kieli** Englanti

Tiivistelmä Navigointi ja paikannus Global Navigation Satellite System (GNSS) -järjestelmien avulla toimii hyvin ympäristöissä joissa taivaalle on hyvä näkyvyys, mutta navigoinnin toimivuus ja tarkkuus heikkenee jos ympärillä oleva maasto tai rakennukset peittävät osan taivaasta. Joskus tällaisissa peitteisissä ympäristöissä kuitenkin tarvitaan tarkkaa paikannusta työturvallisuuden ja -sujuvuuden varmistamiseksi, kuten esimerkiksi avolouhoksissa ja satamissa. Silloin navigoinnin toimivuutta ja tarkkuutta voidaan parantaa asentamalla alueelle maalle sijoitettuja GNSS-signaalin lähettimiä, joita kutsutaan pseudoliiteiksi. Koska peitteisillä alueilla on paljon esteitä navigointisignaalien etenemiselle, pseudoliittien sijainnit täytyy suunnitella tarkasti. Tämä tutkimus tarjoaa työkalun juuri tähän tarkoitukseen. Tutkimuksen tavoite on kehittää uusi menetelmä ja ohjelmisto, jolla voidaan simuloida GNSS-satelliittien ja pseudoliittien signaalien kuuluvuutta ja visualisoida *odotettavissa olevaa paikannustarkkuutta* kulloinkin käytettävissä olevista satelliitti- ja pseudoliittisignaaleista. Tutkimuksessa esitetään tekniikoita joilla mallinnetaan signaalin etenemistä monimutkaisessa kolmiulotteisessa ympäristössä. Lisäksi GNSS-paikannuksen ja dilution of precision (DOP) -arvojen teoriaa tarkastellaan ja käytetään perustelemaan määritelmä hyvälle paikannustarkkuudelle. Tutkimuksen tuloksena on toimiva ohjelmisto, joka simuloi signaalien kuuluvuutta ja paikannustarkkuuden laatua missä tahansa ympäristössä pseudoliittien kanssa tai ilman. Ohjelmiston toimivuus esitetään ajamalla se kahdelle kohteelle, joissa pseudoliittijärjestelmä olisi toteuttamiskelpoinen: syvä avolouhos ja satamalaituri jossa on korkeita nostureita.

Avainsanat Global Navigation Satellite Systems, avolouhos, satama, pseudoliitit, signaalin etenemisen mallinnus, signaalin kuuluvuus, paikannustarkkuus, kolmiulotteinen ympäristömalli

Acknowledgements

This thesis is the product of more than half a year of hard work. During that time, I received many useful insights and appropriate guidance from my advisor Heikki Laitinen and my supervisor Martin Vermeer. I want to thank them for always helping me when I was unsure of how to proceed.

This research was funded by Space Systems Finland (SSF) and the work was mostly done on their premises. I want to thank SSF for the possibility to conduct this research, and all of the people at SSF, who always supported me in my progress and made my working days brighter.

Writing this thesis would also not have been possible without the support of my family. My parents Pekka and Seija brought me up to be the man I am today, and my brother Jari and sister Heli always helped me with fellowship and advice. I also want to thank my fiancee Hangaram who was caring for me and cheering me up all along the way.

Thank you all!

Espoo, October 3, 2012

Esa-Pekka Sundell

Abbreviations and Acronyms

C/A code	Coarse/Acquisition code
C/A signal	Coarse/Acquisition signal
CDMA	Code division multiple access
DEM	Digital elevation model
DOP	Dilution of precision
EGNOS	European Geostationary Navigation Overlay Service
FDMA	Frequency division multiple access
GDOP	Geometric dilution of precision
GIS	Geoinformation system
GLONASS	Globalnaya Navigatsionnaya Sputnikovaya Sistema
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPSTk	GPS Toolkit
GTRF	Galileo Terrestrial Reference Frame
HDOP	Horizontal dilution of precision
IBLS	Integrity Beacon Landing System
IGS	International GNSS Service
ITRF	International Terrestrial Reference Frame
LiDAR	Light Detection and Ranging
LOS	Line of sight
MEO	Medium Earth Orbit
NAVSTAR	NAVigation System with Timing And Ranging
P code	Precision code
P signal	Precision signal
PDOP	Position dilution of precision
PE-90	Parameters of the Earth 1990
PREDISSAT	PREDIctive Software for Satellite Availability in the field of Transport
PRN code	Pseudorandom noise code
PZ-90	Parametry Zemli 1990
QoS	Quality of service
RF	Radio frequency
STK	Satellite Tool Kit
TAI	International Atomic Time

TDOP	Time dilution of precision
Utah AGRC	Utah Automated Geographic Reference Center
UTC	Coordinated Universal Time
VDOP	Vertical dilution of precision
WGS-84	World Geodetic System 1984

Contents

Chapter 1

Introduction

Imagine taking a GPS receiver and traveling around different places. Most of the time navigation will work without problems. Problems will occur maybe between tall buildings, when descending into a canyon, and certainly when entering a tunnel for example. It does not usually matter, since the navigation will return to work quickly when moving away from the problematic place. But in a situation where the GPS receiver is used for navigating for example a straddle carrier carrying a 50-ton container across a harbor, or a 100-ton haul truck driving in an open-pit mine, it can be very dangerous if the navigation suddenly stops working.

In environments that block a large part of the sky around the receiver, satellites cannot always provide continuous navigation service alone. One good solution to this problem is to place additional GPS transmitters, called pseudolites, on the ground around the critical area. When a few pseudolites are placed well enough, they can provide additional positioning signals to those corners and recesses where many satellite signals disappear from visibility.

The subject of this work is to create a methodology to simulate the satellite and pseudolite signal visibility across any three-dimensional environment, in order to detect the places where satellite visibility is significantly reduced, and to see how pseudolite signals will help there.

1.1 Motivation

Global Navigation Satellite Systems (GNSS), most prominently the Global Positioning System (GPS), have revolutionized the way to navigate and determine position on the Earth today. The applications of GNSS are not limited to traditional navigation and geodetic measurements, as they can be used for many other kinds of applications as well. And there are still many possibilities and potential in GNSS positioning that are not yet fully realized.

There are some restrictions to the utilization of GNSS in different applications, however. In order to successfully determine position with GNSS, the satellites must be visible. In other words, the signal must be able to properly travel from the satellites to the receiver. As the satellite signal travels through the Earth's atmosphere, the charged particles in the ionosphere and gases in the troposphere weaken and delay the signal to some extent. The signal still reaches the surface, and can be used for navigation. Even though the microwave signal passes through the atmosphere without major problems, it stops quite easily at solid obstacles. In practice all solid obstacles that block the visibility to the sky around the receiver limit the number of satellite signals that can be received. In environments where there are large structures against the sky, or the surrounding topography is very steep, this can prevent reliable positioning.

Pseudolites can be a solution to the limited visibility of satellites in a challenging environment. Pseudolites transmit signals similar to navigation satellites, but do not orbit the Earth on satellite vehicles. Pseudolite signals can be used together with satellite signals, which will increase positioning availability and expected positioning accuracy. When installed in obstructed environments, pseudolites will help with the problems in positioning that are caused by limited satellite visibility.

It is not a trivial task to choose pseudolite locations that will benefit the navigation accuracy in a meaningful way. A good pseudolite arrangement always depends on the environment, because ground shapes and structures block also pseudolite signals. This thesis will provide a means of modeling the effect on positioning accuracy of pseudolites in various environments, which will be useful in planning pseudolite setups for specific areas.

1.2 Contributions

The goal of this work is to develop a methodology and a piece of software that, firstly, simulates satellite and pseudolite signal visibility in a three-dimensional environment, and secondly, determines the expected positioning accuracy throughout the area. The software has the following requirements:

- The environment model must be able to include terrains with considerable altitude variations.
- The environment model must also allow structures that are located above the ground receiver locations, such as harbor cranes in a port environment. The simulation software must provide a method for modeling these kinds of terrains and structures.
- The software must be able to simulate the GNSS satellite signal visibility in a way that is physically realistic.

- In addition to satellite signals from different GNSS, the simulation must be able to include pseudolite signals. The software must provide a means for inserting pseudolites at any desired location inside the environment, and include them into the simulation in a realistic way.

- The software must be able to calculate proper values that can be used for determining whether the expected positioning accuracy is good or bad.

- The software must then be able to visualize the simulation results in a way that makes it easy to determine the most challenging locations for GNSS positioning.

The contribution of this thesis is the combination of the above functions in one piece of software. Prior research and existing software have demonstrated various methods to simulating GNSS satellite signal visibilities, but, to the knowledge of the author, the combination of complex environment models and the addition of pseudolites have not been implemented in a single simulation model before. This combination of functions will be a valuable tool in planning pseudolite arrangements for various environments.

1.3 Terms and Notations

This thesis contains various terms and notations that are used for the same concepts. The most commonly used terms like this will be covered here.

- *GNSS satellites* and *navigation satellites* are used analoguously to refer to the satellites of a GNSS.

- *GNSS signal visibility simulation system* or *tool*, *simulation program*, and in many places just *simulation software*, are all terms that are used to describe the kind of software system that is created in this thesis and similar systems. All of these mean a piece of software whose functions include GNSS signal visibility simulation.

There are also a few terms that are used when discussing the GNSS navigation performance or quality, mainly *visibility*, *availability*, and *accuracy*.

- *Visibility* or *signal visibility* is used in this thesis to indicate the visibility of a single GNSS satellite or a pseudolite. A satellite or a pseudolite is defined visible, if its signal can be acquired and used in positioning with a receiver in a specific location.

- *Availability* indicates the availability or possibility of GNSS positioning, i.e. GNSS positioning is available if at least four satellites or pseudolites are visible. Therefore, increased or decreased availability in an area means that there are, respectively, more or less locations where GNSS positioning will function properly.

- *Accuracy* of GNSS positioning in this thesis is defined as how close the positioning solutions are to the true receiver position. Accuracy is a very important quality value in this thesis, and it is described in more detail in Section 3.2.

1.4 Outline of the Thesis

This thesis begins with the introduction of basic concepts that need to be understood to analyze GNSS signal visibility. Chapter 2 describes the GNSS positioning systems and the concept of pseudolites. Chapter 3 then explains about the basic GNSS positioning algorithms and accuracy. Chapter 4 focuses on the physics of signal propagation and pseudolite properties that have to be considered when analyzing signal visibility. After this, Chapter 5 reviews some of the prior research that has been done on simulation systems similar to this work.

The remaining chapters present the implementation and utilization of the simulation software created in the course of this research. Chapter 6 introduces the implementation and algorithms of the software. Chapter 7 displays the results from two test cases of the simulation software to show how it works in practice. Chapter 8 discusses some observations about the performance of the simulation software, and finally, chapter 9 presents a review and the conclusions of this research.

Chapter 2

Global Navigation Satellite Systems and Pseudolites

Satellite-based positioning is the determination of positions on land, at sea, in air or in space by using artificial satellites that orbit the Earth. Global Navigation Satellite System (GNSS) is a term used for all satellite-based positioning systems, and also their combinations and augmentations (Hofmann-Wellenhof, Lichtenegger, & Wasle, 2008, p. 3). The beginning of this chapter describes the current GNSS in Section 2.1, after which the concept of pseudolites is explained in Section 2.2.

2.1 Global Navigation Satellite Systems

There are currently four modern Global Navigation Satellite Systems in operation or under deployment: GPS, GLONASS, Galileo, and Compass. This section will give a brief overview of all of these systems and their properties. For more detailed descriptions of these systems, refer to the book by Hofmann-Wellenhof et al. (2008, p. 309-430).

2.1.1 GPS

GPS is an abbreviation for the NAVigation System with Timing And Ranging (NAVSTAR) Global Positioning System. Development of GPS started in 1973 by the direction of the United States Department of Defense, and the task was given to the Joint Program Office, a component of the Space and Missile Center at El Seguendo, California. GPS full operational capability was declared in July 1995. It was designed to succeed Transit, the earlier and simpler satellite-based navigation system of the US.

GPS is a ranging system that determines unknown positions on land, at sea, in the air and in space using known positions of satellites. Its defined objectives include instantaneous position and velocity determination (i.e. navigation), and precise coordination of time (i.e. time transfer). The primary goals for the development of GPS were military ones, but later also the civilian use of the system was promoted by the direction of the US Congress.

The main properties of GPS are:

- Terrestrial reference system: World Geodetic System 1984 (WGS-84).
- Time system: GPS system time referenced to Coordinated Universal Time (UTC) as maintained in US Naval Observatory. GPS time was coincident with UTC in January 6th 1980 but does not take leap seconds, so it differs from UTC with an integer amount of seconds. GPS time has a constant offset of 19 seconds to International Atomic Time (TAI).
- Navigation services: Standard positioning service for civilian users and precise positioning service for authorized users. The separation of navigation services into two levels allows JPO to deny the full system accuracy to nonmilitary users.
- Signals: Carrier frequencies L1 (1575.42 MHz) and L2 (1227.60 MHz). Standard-accuracy C/A signal and high-accuracy P signal. There is also a new carrier frequency L5 (1176.45 MHz) that is currently only transmitted by the more recent GPS satellites launched since 2010, and is not yet in official use. L5 has a special safety-of-life code modulated on it.

There are two different ranging codes modulated on frequencies L1 and L2. The first is the coarse/acquisition (C/A) code, which is available for civilian use, and the second is the precision (P) code, which is reserved for the US military use only. C/A code is modulated only on the L1 frequency and P code on both L1 and L2. L5 is a new addition to the carrier frequencies, and it is designed to enhance the performance and reliability of GPS in safety-of-life applications, such as civilian aviation.

The GPS satellite constellation has the following properties:

- Constellation: 32 satellites in six orbit planes, each with four satellites.
- Orbits: Nearly circular orbits with an altitude of about 20,200 km. The satellites perform 2 full orbits per sidereal day, so the periods are approximately 11 hours 58 minutes. Orbit inclination is 55 degrees.
- 37 distinct pseudorandom noise (PRN) codes to separate signals from different satellites. This technique is called code division multiple access (CDMA). The first 32 codes are assigned to the satellites, the remaining 5 can be used by for example pseudolites.

2.1.2 GLONASS

GLONASS derives from the Russian "Globalnaya Navigatsionnaya Sputnikovaya Sistema", which translates to English as Global Navigation Satellite System. GLONASS development was initialized in the former Union of Soviet Socialist Republics in the mid-1970s. It was based on the experiences with the earlier Doppler satellite system called Tsikada. GLONASS reached its full operational capability in September 1996.

Similarly to GPS, GLONASS is a military system. Its purpose is to provide an unlimited number of air, marine, and any other type of users with all-weather three-dimensional positioning, velocity measuring and timing everywhere on the Earth or space close to the Earth at any time. Almost no information of the system was released outside the military until 1995, when the government of the Russian Federation officially announced that GLONASS will start to service also civilian users.

The main properties of GLONASS are:

- Terrestrial reference system: Parameters of the Earth 1990 (PE-90) or, as it is sometimes called with its Russian name, "Parametry Zemli 1990" (PZ-90).
- Time system: GLONASS time is maintained by a central synchronizer, and is closely related to UTC but has a constant offset of three hours, the time difference between Moscow and Greenwich. This means that GLONASS time has leap seconds. Apart from the constant integer offset, difference of GLONASS time to UTC is within 1 millisecond.
- Navigation services: Two levels of service, standard positioning service for civilian users and precise positioning service for the military.
- Signals: Carrier frequencies G1 (about 1602 MHz) and G2 (about 1246 MHz). Standard-accuracy C/A signal and high-accuracy P signal.

Similarly to GPS, the ranging codes are separated for the two service levels. The code for civilian use, called the C/A code or the S code, is modulated on satellite frequency G1. The code restricted for military use, the P code, is modulated on both frequencies G1 and G2. There is also a third frequency G3 allocated to GLONASS, but there is still little information available about its usage.

The GLONASS satellite constellation has the following properties:

- Constellation: 24 satellites in three orbit planes, each with eight satellites. 21 satellites are considered active satellites and the remaining three are "active on-orbit spares".
- Orbits: Circular orbits with an altitude of about 19,100 km, with periods of nominally 11 hours, 15 minutes, and 44 seconds. Orbit inclination 64.8 degrees, and the ascending nodes of orbit planes separated by 120 degrees.

- Implements the frequency division multiple access (FDMA) technique to differentiate between signals from different satellites. This means that the satellites have slightly different carrier frequencies from each other. GLONASS uses 15 bands that are spanned around the base carrier frequencies (1602 MHz and 1246 MHz) and separated by 0.5625 MHz for G1 and 0.4375 MHz for G2. To reduce the number of required frequencies, satellites on the opposite sides of the Earth (antipodes) transmit on the same frequency since they are never visible on the Earth at the same time.

Since the satellite signals are separated by carrier frequencies, GLONASS uses the same PRN codes for all satellites. The downside of the FDMA technology implemented in GLONASS is that it requires the receivers to be able to receive quite wide bandwidths. In the future, GLONASS will also start using the CDMA technique similarly to GPS, and FDMA will be supported only as a legacy technique. Currently, only the newest generation GLONASS satellites transmit signals that use CDMA.

2.1.3 Galileo

In 1994, by the request of the European Council, also the European Commission decided to contribute to satellite navigation. They set out to make progress in two steps: first create an augmentation system for existing GNSS (i.e. GPS and GLONASS), and then contribute to the development for a global navigation system for civil use.

The first step resulted in the development and deployment of the European geostationary navigation overlay service (EGNOS). The second step started with cooperation with the US in the development of next-generation GPS. But, because GPS has always been considered a safety-critical system in the US, foreign countries were not allowed to participate in its definition and control. Europe also considered cooperation with Russia, but finally decided to develop its own GNSS.

Galileo was the name provisionally given to the European contribution to satellite navigation and it has since become the name for the European GNSS. The name is not an acronym, but it was chosen to give tribute to the Italian scientist and astronomer Galileo Galilei (1564-1642). Galileo is not intended to replace or to compete with other GNSS, but to be interoperable with them in a way that benefits civilian users.

At the moment of writing, Galileo has deployed two test satellites and two operational satellites, and the system in-orbit validation is in progress. Full operational capability is expected to be reached by the end of this decade. (European Space Agency, 2012)

The main properties of Galileo are:

- Terrestrial reference system: Galileo Terrestrial Reference Frame (GTRF), maintained by Galileo Geodetic Service Provider. GTRF will be related to the International Terrestrial Reference Frame (ITRF), and is specified to differ from the latest version of ITRF by no more than 3 centimeters.

- Time system: Galileo system time is a separately maintained continuous time scale that has a nominal offset of a constant amount of seconds to TAI.

- Navigation services: Satellite-only service, Open service, Commercial service, Safety-of-life service, and Public regulated service.

- Signals: Carrier frequencies E1 (1575.420 MHz), E6 (1278.750), E5 (1191.795 MHz; also denoted as E5a+E5b), E5a (1176.450 MHz), and E5b (1207.140 MHz). An open-access ranging code, ranging codes with commercial encryption, and ranging codes with governmental encryption.

Two of the frequency bands of Galileo, E5a and E1 are chosen purposely to be common with GPS to make interoperability easier. Also, the frequency band E5b will overlay with GLONASS carrier G3, for the same reason. There are different levels of ranging codes implemented in Galileo, but they are specified to be flexible in order to allow uploading of new settings into satellite memory.

The Galileo satellite constellation, once fully deployed, will have the following properties:

- Constellation: 30 satellites, of which 27 operational and three active spares.

- Orbits: Three circular orbit planes at an altitude of 23,222 km. Orbit inclination 56 degrees.

- Similarly to GPS, Galileo satellite signals are differentiated with the CDMA technique.

2.1.4 Compass

Compass, also known by the name BeiDou-2, is a GNSS currently under development by the People's Republic of China. The name of the system comes from the Chinese name for the constellation Big Dipper, "beidou", which means compass in English. The development of Compass will extend the already deployed and working BeiDou Satellite Navigation Experimental System (BeoDou-1), which is a two-satellite regional navigation system. (Montenbruck et al., 2012)

Compass was initialized in 2007 with the launch of the first navigation satellite into Medium altitude Earth Orbit (MEO). The system is expected to reach a regionally working constellation of 12 satellites before the end of 2012, which will provide navigation services in Asia and Australia. The system will be further developed to provide global service by 2020 with a total constellation of 30 MEO satellites and five geostationary satellites. (Montenbruck et al., 2012)

The content of the Compass navigation message has not yet been publicly released, and other technical details of the system are also not yet available to the public. This

makes real-time navigation with the system still impossible at this time, but it can already be used for geodetic positioning with post-processed orbit and clock solutions. (Montenbruck et al., 2012)

2.2 Pseudolites

This section will briefly describe what pseudolites are and how they can augment GNSS navigation. The use of pseudolites in the augmentation of GPS and other GNSS has been a widely studied subject since the beginning of these positioning systems. The information presented here is based on a prominent early study considering GPS augmentation with pseudolites conducted by Cobb (1997, p. 17-26).

Even though the official documentations of GNSS (Global Positioning Systems Directorate, 2011; European Space Agency, 2012) only describe the use of satellite signals, ground-based transmitters have been used together with satellites from the beginning of the GPS concept. These ground-based transmitters were first called pseudo-satellites, which was quickly shortened to pseudolites.

It has to be noted that some receivers cannot track pseudolites at all, or cannot track satellites in the presence of pseudolites. These practical difficulties of using pseudolites are covered later in Section 4.3. This section will assume ideal pseudolites and receivers that have no such difficulties.

2.2.1 Direct Ranging Pseudolite

The earliest pseudolite can be thought of as a complete ground-based satellite. It transmits the same code signals, on the same carrier frequencies, using the same timing with the satellite signals. Ideally the pseudolite will appear to a receiver as just an additional satellite signal. The only difference is that the pseudolite position must be described in geographic coordinates instead of orbital elements used by satellites.

This kind of pseudolites can be advantageous in areas where the GNSS satellite signal availability is limited. Each pseudolite will allow a receiver to navigate with one less satellite signal, so navigation becomes possible even when there are less than four satellites visible. See Figure 2.1 for a truck navigating with three satellite signals and a pseudolite signal. Pseudolites will also add to the fault safety of the navigation in the case of unexpected satellite signal failures. Because satellites and pseudolites are entirely separate systems, their common failure is highly unlikely.

If pseudolites are placed in places where their signals enhance the satellite-pseudolite geometry, they will not only improve the navigation signal availability, but also positioning accuracy. The effect of geometry of the satellites and pseudolites on the accuracy of positioning is further explained in Section 3.2. Because pseudolites can be

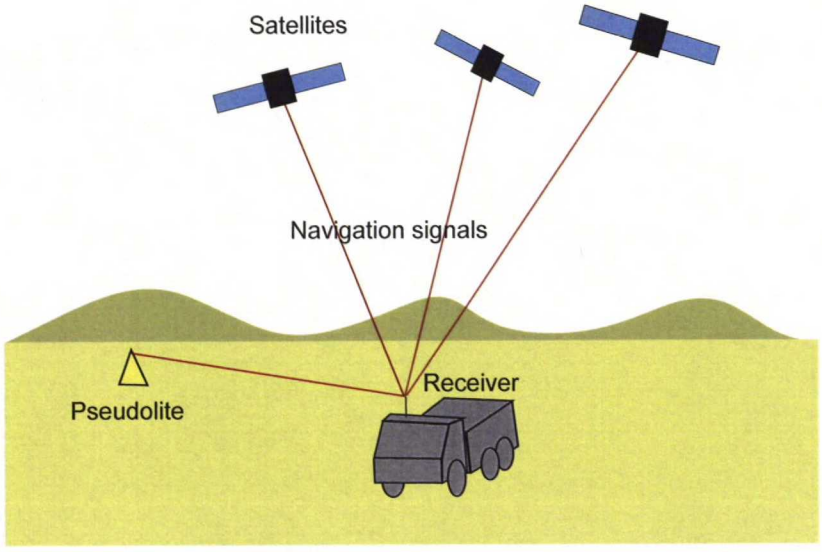

Figure 2.1: A navigation system with a pseudolite and GNSS satellites working together.

installed almost anywhere, they can potentially increase the positioning accuracy much more than additional satellite signals. Pseudolites are especially good at enhancing the vertical accuracy of GNSS positioning, which is usually worse than horizontal accuracy when positioning only with satellites.

2.2.2 Pseudolites in Carrier Phase Positioning

Most GNSS receivers use the modulated codes in the satellite signals for range measurements (see Section 3.1.1). The signal carrier phases can also be observed for this purpose instead of the modulated codes, which will result in much better accuracy. Using phase ranges is not as straightforward as using code ranges, but pseudolite signals in the navigation solution can help considerably.

For GPS satellites the modulated C/A code is approximately 293 meters long, and each cycle of the L1 carrier frequency is about 19 cm long. If a good receiver measures either feature with a precision of one percent, the range precision would be about 0.5 meters for C/A code and about 1 mm for carrier phase. Using phase ranges is clearly superior to code ranges in accuracy, but it has the problem of ambiguity that makes its use difficult in most real-time applications. The C/A code is designed to be unambiguous, so that each chip cannot be confused with its neighbors.[1] This means that the C/A code measurements give their ranges directly. Carrier cycles are not unique, however, and each cycle looks just like any other. The receiver can easily measure the fraction of the phase, but the range observation still contains an unknown, arbitrary number of whole cycles. This number of cycles is called the integer cycle ambiguity.

[1]The C/A code has an ambiguity at intervals of one epoch that corresponds to about 300 km. This ambiguity is easily resolved by examining the data modulation.

The ambiguities must be determined by different means than direct observations. One method to determine the ambiguities is to combine a set of simultaneous code and carrier phase measurements into a single large matrix, and solve it to determine all integer ambiguities at once. The sets must be sufficiently different from each other to provide accurate results. In practice, the sets must be separated enough in time for the satellite-receiver geometry to change. Usually the pseudolites are relatively close to the receiver, so if the receiver is moving, the satellite-pseudolite-receiver geometry changes very quickly. This change can make the ambiguity determination technique provide reliable results much faster.

Carrier phase ranges are used almost exclusively in relative positioning techniques, which are explained in more detail in Section 3.1.3. The integer ambiguity determination is not further considered in this thesis, but it was mentioned here because it is an important way in which pseudolites help to provide more reliable navigation service.

Chapter 3

Positioning Algorithms and Accuracy

The simulation model created in this thesis aims to model the expected accuracy of GNSS positioning. To properly understand how positioning accuracy can be modeled, basic GNSS positioning algorithms are presented in this chapter.

The principle of positioning in all GNSS is the same. They presume that the locations of the satellites are known at any time and the satellites transmit a signal whose travel time to the receiver can be observed. A detailed explanation of GNSS positioning algorithms is given in the following Section 3.1. After that, Section 3.2 concentrates on how GNSS positioning accuracy is determined in different positioning techniques and introduces the dilution of precision values.

3.1 Positioning Algorithms

The positioning algorithms used in GNSS are based on ranges or range differences measured to satellites. During position determination of a receiver, its distances to multiple satellites are measured at the same time. A distance to a satellite is called a range, which emphasizes the fact that a single distance observation actually represents a sphere in space. All points on the sphere share the observed distance to the satellite. (Hofmann-Wellenhof et al., 2008, p. 105-106)

The ranges to the satellites can be determined by observing the code modulated on the satellite signal or the signal phase. Also the position determination from the ranges can be done in a few different ways. This section explains these techniques as they are presented in the book of Hofmann-Wellenhof et al. (2008, p. 105-108, 161-178, 238-239, 250-266).

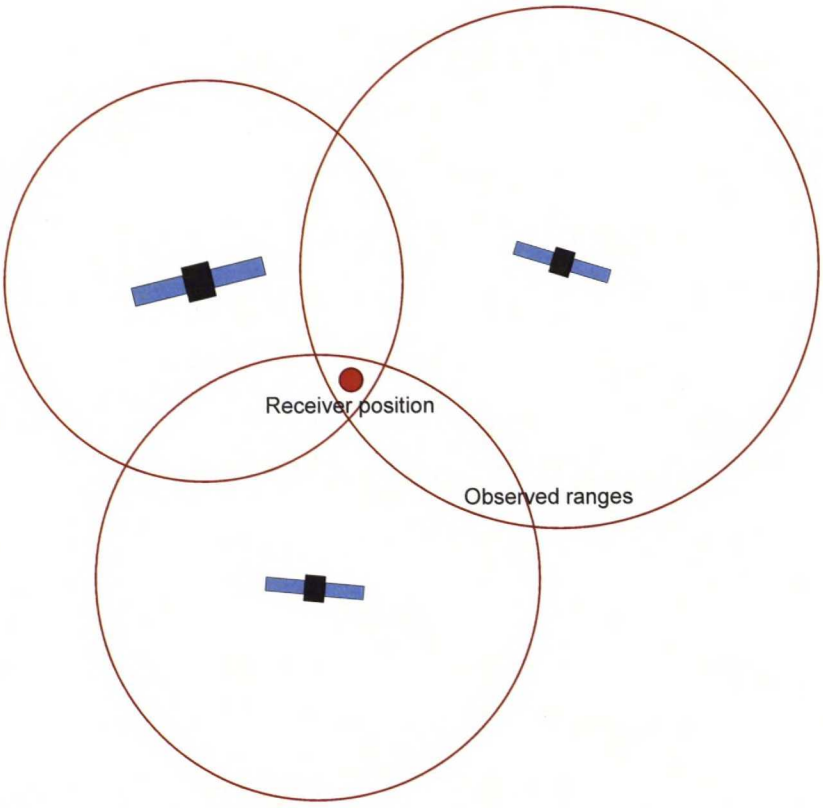

Figure 3.1: The receiver location lies in the intersection of the spheres that have the radius of the observed ranges and the satellites as the centers.

3.1.1 Point Positioning

The simplest way to determine position with GNSS observation is based on measuring the travel time to the satellites with just one receiver. The general way to do this is with code ranges. There is a specific pseudo-random code modulated on the carrier wave of the satellite signal. The receiver produces the same code, and then searches for the time delay to the satellite clock that produces exact correlation between the two codes. This time delay describes how long the signal has taken to travel from the satellite to the receiver.

In Cartesian coordinates, there are three unknowns, x, y, and z, so if the receiver clock and the satellite clock would run in the exactly same time, three range measurements would be enough to determine the receiver position. Each range provides a sphere that the receiver must lie on. Two spheres intersect on a circle in space. The circle and the third sphere intersect on two points: one point is the correct receiver position near the surface of the Earth, and the other point is usually somewhere far in space. See Figure 3.1 for an illustration of this concept.

In reality, however, the receiver clock is usually inexpensive and inaccurate. If the receiver clock differs from the satellite clock by $\Delta\delta_r^s$, all distance measurements will be wrong by $c\Delta\delta_r^s$, where c is the speed of light. This is why the receiver clock error δ_r is

also needed as an unknown in the solution. Satellites broadcast their clock errors with good accuracy, so the satellite clock error δ^s can usually be considered a known value.

Pseudoranges

Because of various errors and biases, which are discussed in Section 3.2.3, the observed ranges are not true geometrical distances. This is why they are called pseudoranges instead of just ranges. Following Hofmann-Wellenhof et al. (2008, p. 161), the code pseudorange can be modeled as

$$\begin{aligned} R_r^s(t) &= \varrho_r^s(t) + c\Delta\delta_r^s(t) \\ &= \varrho_r^s(t) + c\delta_r(t) - c\delta^s(t). \end{aligned}$$

$R_r^s(t)$ is the pseudorange between a satellite s and a receiver r on a given epoch t. It consists of $\varrho_r^s(t)$, the real geometrical distance between the receiver and the satellite, and $c\Delta\delta_r^s$, the clock error between the receiver and the satellite. The clock error was further divided into the satellite part $\delta^s(t)$ and the receiver part $\delta_r(t)$.

Then, separating the known and unknown quantities yields

$$R_r^s(t) + c\delta^s(t) = \varrho_r^s(t) + c\delta_r(t), \tag{3.1}$$

where the quantities on the left side are known and the quantities on the right side are unknown. This is a good form to be used in the following calculations. The true geometrical distance $\varrho_r^s(t)$ in the equation can be expressed as

$$\varrho_r^s(t) = \sqrt{(X^s(t) - X_r)^2 + (Y^s(t) - Y_r)^2 + (Z^s(t) - Z_r)^2}, \tag{3.2}$$

where $X^s(t)$, $Y^s(t)$ and $Z^s(t)$ are the coordinates of the satellite on epoch t, and X_r, Y_r and Z_r are the coordinates of the observing receiver.

Linear Model for Point Positioning

The computation of the receiver position is based on the solution of a system of range equations (Eq. (3.1)) to the satellites. Following Hofmann-Wellenhof et al. (2008, p. 161), a standard technique to solve the system of equations is by linearization. This simplifies the calculation considerably, because it allows implementing adjustment algorithms instead of complex nonlinear methods.

The only part in the observation equation that has unknowns in nonlinear form is the geometrical distance Eq. (3.2). The linearization starts with introducing an approximate receiver position (X_{r0}, Y_{r0}, Z_{r0}), so that the receiver coordinates can be expressed as the approximate receiver position and the difference to that position

$$\begin{aligned} X_r &= X_{r0} + \Delta X_r \\ Y_r &= Y_{r0} + \Delta Y_r \\ Z_r &= Z_{r0} + \Delta Z_r. \end{aligned}$$

The approximate distance from the satellite to the receiver, i.e. the true distance to the approximate receiver position, is a known value. It is given by

$$\varrho^s_{r0}(t) = \sqrt{(X^s(t) - X_{r0})^2 + (Y^s(t) - Y_{r0})^2 + (Z^s(t) - Z_{r0})^2}.$$

The differences ΔX_r, ΔY_r, ΔX_r are unknown, whereas the approximate coordinates X_{r0}, Y_{r0}, Z_{r0} are known. This split-up makes it possible to replace the distance equation Eq. (3.2) with a Taylor series with respect to the approximate position, which then becomes

$$\begin{aligned} \varrho^s_r(t) &= \varrho^s_{r0}(t) + \left.\frac{\partial \varrho^s_r(t)}{\partial X_r}\right|_{X_r=X_0} \Delta X_r + \left.\frac{\partial \varrho^s_r(t)}{\partial Y_r}\right|_{Y_r=Y_0} \Delta Y_r + \left.\frac{\partial \varrho^s_r(t)}{\partial Z_r}\right|_{Z_r=Z_0} \Delta Z_r \\ &= \varrho^s_{r0}(t) + \frac{X^s(t) - X_{r0}}{\varrho^s_{r0}(t)} \Delta X_r - \frac{Y^s(t) - Y_{r0}}{\varrho^s_{r0}(t)} \Delta Y_r - \frac{Z^s(t) - Z_{r0}}{\varrho^s_{r0}(t)} \Delta Z_r. \end{aligned} \tag{3.3}$$

With n observed satellites, there is a set of n linear equations. By combining Eq. (3.1) with Eq. (3.3), the observation equation for a satellite s becomes

$$\begin{aligned} R^s_r(t) - \varrho^s_{r0}(t) + c\delta^s(t) &= \frac{X^s(t) - X_{r0}}{\varrho^s_{r0}(t)} \Delta X_r - \frac{Y^s(t) - Y_{r0}}{\varrho^s_{r0}(t)} \Delta Y_r \\ &\quad - \frac{Z^s(t) - Z_{r0}}{\varrho^s_{r0}(t)} \Delta Z_r + c\delta_r(t), \end{aligned} \tag{3.4}$$

where known terms are written on the left, and unknowns on the right.

Solving the Linear Model for Point Positioning

The actual receiver position is determined with a least squares model. The set of linear equations can be written in a matrix-vector form

$$\mathbf{A}\mathbf{x} = \mathbf{l},$$

where $\mathbf{A}$ is the design matrix for point positioning, $\mathbf{x}$ contains the unknowns, and $\mathbf{l}$ contains the observations. Using the observation equations from Eq. (3.4), the matrices become

$$\mathbf{A} = \begin{bmatrix} -\frac{X^1 - X_{r0}}{\varrho^1_{r0}} & -\frac{Y^1 - Y_{r0}}{\varrho^1_{r0}} & -\frac{Z^1 - Z_{r0}}{\varrho^1_{r0}} & c \\ -\frac{X^2 - X_{r0}}{\varrho^2_{r0}} & -\frac{Y^2 - Y_{r0}}{\varrho^2_{r0}} & -\frac{Z^2 - Z_{r0}}{\varrho^2_{r0}} & c \\ \vdots & & & \\ -\frac{X^n - X_{r0}}{\varrho^n_{r0}} & -\frac{Y^n - Y_{r0}}{\varrho^n_{r0}} & -\frac{Z^n - Z_{r0}}{\varrho^n_{r0}} & c \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} \Delta X_r \\ \Delta Y_r \\ \Delta Z_r \\ \delta_r(t) \end{bmatrix} \quad \mathbf{l} = \begin{bmatrix} l^1 \\ l^2 \\ \vdots \\ l^n \end{bmatrix}, \tag{3.5}$$

where the observations $l^s = R^s_r(t) - \varrho^s_{r0}(t) + c\delta^s(t)$.

The general least squares solution for the unknowns $\mathbf{x}$ in is given by

$$\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{P}\mathbf{l},$$

where $\mathbf{P}$ in is the weight matrix that contains weights for observed pseudoranges. Usually the weight matrix is derived from the covariance matrix of the observed pseudoranges $\mathbf{\Sigma}_l$ in the following way:

$$\mathbf{P} = \frac{\mathbf{\Sigma}_l^{-1}}{\sigma_0^2} = \frac{1}{\sigma_0^2}\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{bmatrix}^{-1}, \tag{3.6}$$

where σ_0^2 is the a priori variance of unit weight, which is often assumed to be 1. Usually the code observations are considered independent with equal variances. If this assumption is correct, the diagonal terms of $\mathbf{\Sigma}_l$ equal to σ_0^2, and the non-diagonal terms vanish. Consequently, the weight matrix $\mathbf{P}$ becomes an identity matrix and can be ignored, and the least squares solution is simplified to

$$\hat{\mathbf{x}} = \begin{bmatrix} \Delta\hat{X}_r \\ \Delta\hat{Y}_r \\ \Delta\hat{Z}_r \\ \hat{\delta}_r(t) \end{bmatrix} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{l}, \tag{3.7}$$

and the final receiver coordinates are

$$\begin{aligned} \hat{X}_r &= X_{r0} + \Delta\hat{X}_r \\ \hat{Y}_r &= Y_{r0} + \Delta\hat{Y}_r \\ \hat{Z}_r &= Z_{r0} + \Delta\hat{Z}_r. \end{aligned}$$

3.1.2 Differential Positioning

Differential positioning is a technique where multiple receivers are used at the same time relatively close to each other. The receivers observe their pseudoranges to the same satellites. This means that the paths that the signal travels through the atmosphere to these two receivers are approximately the same, and consequently the errors in the observed pseudoranges are similar. There is usually one receiver, called a reference or base station, which is set up in a known location. A base station's true geometric ranges to the satellites are known within the accuracy that the satellites' positions are known, and therefore the errors in the pseudoranges can be determined. These errors are then sent to the other receivers, which are called rovers, in unknown locations and are used to eliminate the errors in the observed pseudoranges. Then the point positioning technique is applied with the rovers' corrected pseudoranges.

There will still be some error remaining in the rover's pseudoranges due to the slight differences in the signal path to the base station and the rovers. Also receiver-related errors and errors from multipath propagation still remain in the solution. The errors in the satellite clock disappear completely, and also the ionospheric effects nearly completely. A little error from time synchronization, tropospheric effects and orbit errors also remains (more about observation errors in Section 3.2.1). This error is proportional to first order to the distance between the receivers, and gets larger as the receivers become more separated.

3.1.3 Relative Positioning

It is also possible to calculate differences between the observed pseudoranges of different receivers. This technique is called relative positioning, because the subtraction of the observed pseudoranges produces the relative position of the receivers with respect to one another. Since the paths from the receivers to the satellites are similar, the difference of the pseudoranges is free of most errors that the individual observations had, similarly to differential positioning. Relative positioning can be performed with either code-range or phase-range differences, but because the process of the positioning algorithm for code and phase ranges is similar, this section will only explain the technique with code-range. In practice, however, only phase-range differences are used in relative positioning, which is because of the significantly better accuracy.

Similarly to differential positioning, the technique requires a base station receiver with a known location, and a rover receiver with an unknown location. Let A be the location of the base station and B the location of the rover. Between these two points there is a vector $\mathbf{b}_{AB}$ which is called the *baseline* vector or just baseline. The components of the baseline are

$$\mathbf{b}_{AB} = \begin{bmatrix} X_B - X_A \\ Y_B - Y_A \\ Z_B - Z_A \end{bmatrix} = \begin{bmatrix} \Delta X_{AB} \\ \Delta Y_{AB} \\ \Delta Z_{AB} \end{bmatrix}.$$

When the coordinates of the base station are known, the rover's coordinates can easily be calculated from the baselines. The differencing can be made in three ways: between receivers, between satellites, and between epochs. These difference types are often combined to produce even better observations. Most relative positioning software use three combinations, which are single, double, and triple differences. Single differences correspond to differences between receivers, double differences correspond to differences between receivers and satellites together, and triple differences correspond to differences between receivers, satellites, and epochs.

Single differences

With code ranges, the basic pseudorange equation is Eq. (3.1). A single difference is formed by subtracting two pseudoranges from each other. The pseudoranges share a satellite j but lead to different receivers A and B. These pseudoranges become

$$R_A^j(t) + c\delta^j(t) = \varrho_A^j(t) + c\delta_A(t) \quad \text{and} \quad R_B^j(t) + c\delta^j(t) = \varrho_B^j(t) + c\delta_B(t).$$

The single difference is then

$$\begin{aligned} R_B^j(t) + c\delta^j(t) - R_A^j(t) - c\delta^j(t) &= \varrho_B^j(t) + c\delta_B(t) - \varrho_A^j(t) - c\delta_A(t) \\ R_B^j(t) - R_A^j(t) &= \varrho_B^j(t) - \varrho_A^j(t) + c(\delta_B(t) - \delta_A(t)). \end{aligned}$$

The result of the subtraction is that the satellite clock differences cancel on the left side. By introducing relative notations

$$\begin{aligned} R_{AB}^j(t) &= R_B^j(t) - R_A^j(t) \\ \varrho_{AB}^j(t) &= \varrho_B^j(t) - \varrho_A^j(t) \\ \delta_{AB}(t) &= \delta_B(t) - \delta_A(t)), \end{aligned}$$

the final single-difference equation is formed:

$$R^j_{AB}(t) = \varrho^j_{AB}(t) + c\delta_{AB}(t). \tag{3.8}$$

Double Differences

To obtain double differences, two single differences must be subtracted. Assuming two receivers A and B, and two satellites j and k, two single differences according to Eq. (3.8) can be formed:

$$R^j_{AB}(t) = \varrho^j_{AB}(t) + c\delta_{AB}(t) \quad \text{and} \quad R^k_{AB}(t) = \varrho^k_{AB}(t) + c\delta_{AB}(t).$$

The double difference is then

$$\begin{aligned} R^k_{AB}(t) - R^j_{AB}(t) &= \varrho^k_{AB}(t) + c\delta_{AB}(t) - \varrho^j_{AB}(t) - c\delta_{AB}(t) \\ &= \varrho^k_{AB}(t) - \varrho^j_{AB}(t). \end{aligned}$$

As can be seen from the above process, the receiver clock differences cancel on the right side. There is nothing left in the observation equation except the differences of the geometric ranges. A relative notation is used on the left side to form the final double-difference equation:

$$R^{jk}_{AB}(t) = \varrho^k_{AB}(t) - \varrho^j_{AB}(t). \tag{3.9}$$

Linear Model for Relative Positioning

A double-difference observation is composed of four ranges between the satellites and the receivers. This is depicted in Figure 3.2. The four ranges are

$$\begin{aligned} R^{jk}_{AB}(t) &= R^k_{AB}(t) - R^j_{AB}(t) \\ &= R^k_B(t) - R^j_B(t) - R^k_A(t) + R^j_A(t). \end{aligned} \tag{3.10}$$

The ranges are linearized similarly to the point positioning case. The equations are linearized about approximate receiver positions for both points A and B, and the linearization is done in the same way as in point positioning (Eq. (3.4)), which yields the four ranges

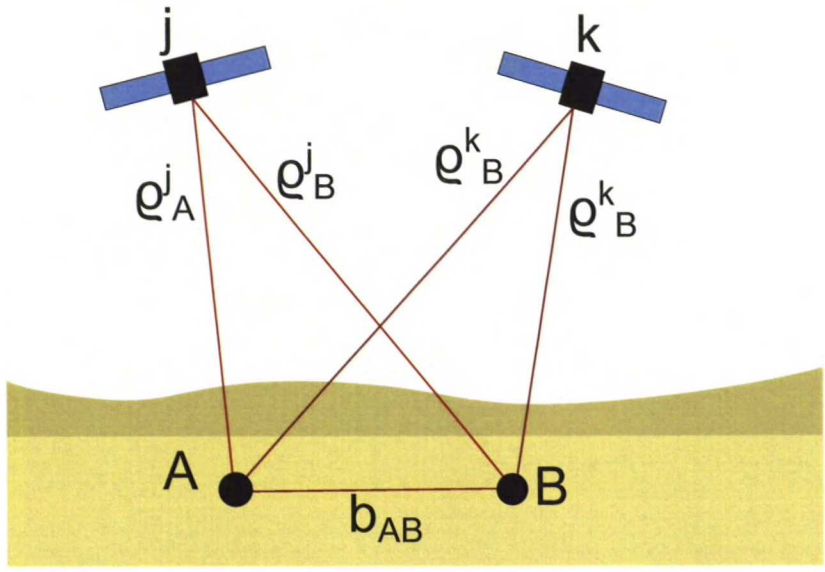

Figure 3.2: Relative positioning with double differences.

$$
\begin{aligned}
R^k_B(t) &= \varrho^k_{B0}(t) - \frac{X^k(t) - X_{B0}}{\varrho^k_{B0}(t)}\Delta X_B - \frac{Y^k(t) - Y_{B0}}{\varrho^k_{B0}(t)}\Delta Y_B - \frac{Z^k(t) - Z_{B0}}{\varrho^k_{B0}(t)}\Delta Z_B \\
&\quad + c\delta_B(t) - c\delta^k(t) \\
R^j_B(t) &= \varrho^j_{B0}(t) - \frac{X^j(t) - X_{B0}}{\varrho^j_{B0}(t)}\Delta X_B - \frac{Y^j(t) - Y_{B0}}{\varrho^j_{B0}(t)}\Delta Y_B - \frac{Z^j(t) - Z_{B0}}{\varrho^j_{B0}(t)}\Delta Z_B \\
&\quad + c\delta_B(t) - c\delta^j(t) \\
R^k_A(t) &= \varrho^k_{A0}(t) - \frac{A^k(t) - X_{A0}}{\varrho^k_{A0}(t)}\Delta X_A - \frac{Y^k(t) - Y_{A0}}{\varrho^k_{A0}(t)}\Delta Y_A - \frac{Z^k(t) - Z_{A0}}{\varrho^k_{A0}(t)}\Delta Z_A \\
&\quad + c\delta_A(t) - c\delta^k(t) \\
R^j_A(t) &= \varrho^j_{A0}(t) - \frac{A^j(t) - X_{A0}}{\varrho^j_{A0}(t)}\Delta X_A - \frac{Y^j(t) - Y_{A0}}{\varrho^j_{A0}(t)}\Delta Y_A - \frac{Z^j(t) - Z_{A0}}{\varrho^j_{A0}(t)}\Delta Z_A \\
&\quad + c\delta_A(t) - c\delta^j(t).
\end{aligned}
$$

In relative positioning, the coordinates of the base station must be known. If A is the base station, the approximate position is already the known position, so $X_{A0} = X_A$, $Y_{A0} = Y_A$, $Z_{A0} = Z_A$ and $\Delta X_A = \Delta Y_A = \Delta Z_A = 0$. This reduces the unknowns by three. Finally, the double-difference equation becomes

$$
\begin{aligned}
R^{jk}_{AB}(t) &= R^{k}_{AB}(t) - R^{j}_{AB}(t) = R^{k}_{B}(t) - R^{j}_{B}(t) - R^{k}_{A}(t) + R^{j}_{A}(t) \\
&= \varrho^{k}_{B0}(t) - \varrho^{j}_{B0}(t) - \varrho^{k}_{A}(t) + \varrho^{j}_{A}(t) \\
&\quad -\frac{X^k(t) - X_{B0}}{\varrho^{k}_{B0}(t)}\Delta X_B - \frac{Y^k(t) - Y_{B0}}{\varrho^{k}_{B0}(t)}\Delta Y_B - \frac{Z^k(t) - Z_{B0}}{\varrho^{k}_{B0}(t)}\Delta Z_B \\
&\quad +\frac{X^j(t) + X_{B0}}{\varrho^{j}_{B0}(t)}\Delta X_B + \frac{Y^j(t) - Y_{B0}}{\varrho^{j}_{B0}(t)}\Delta Y_B + \frac{Z^j(t) - Z_{B0}}{\varrho^{j}_{B0}(t)}\Delta Z_B.
\end{aligned}
$$

Notice that the clock errors have disappeared from the equation due to the double differencing. Gathering the known terms on the right and unknown terms on the left gives

$$
\begin{aligned}
&a^{jk}_{X_B}\Delta X_B + a^{jk}_{Y_B}\Delta Y_B + a^{jk}_{Z_B}\Delta Z_B = \\
&\qquad R^{jk}_{AB}(t) - \varrho^{k}_{B0}(t) + \varrho^{j}_{B0}(t) + \varrho^{k}_{A}(t) - \varrho^{j}_{A}(t),
\end{aligned}
$$

where the following abbreviations are used:

$$
\begin{aligned}
a^{jk}_{X_B}(t) &= \frac{X^k(t) - X_{B0}}{\varrho^{k}_{B0}(t)} - \frac{X^j(t) - X_{B0}}{\varrho^{j}_{B0}(t)} \\
a^{jk}_{Y_B}(t) &= \frac{Y^k(t) - Y_{B0}}{\varrho^{k}_{B0}(t)} - \frac{Y^j(t) + Y_{B0}}{\varrho^{j}_{B0}(t)} \\
a^{jk}_{Z_B}(t) &= \frac{Z^k(t) - Z_{B0}}{\varrho^{k}_{B0}(t)} - \frac{Z^j(t) - Z_{B0}}{\varrho^{j}_{B0}(t)}.
\end{aligned}
$$

Solving the Linear Model for Relative Positioning

With n observed satellites, there is a set of $n-1$ double difference equations. In matrix form

$$
\mathbf{Gx} = \mathbf{l},
$$

where $\mathbf{G}$ is the design matrix for relative positioning, in other words, the design matrix for a baseline vector. $\mathbf{x}$ contains the unknowns, and $\mathbf{l}$ contains the observations. The contents of these matrices are

$$
\mathbf{G} = \begin{bmatrix}
a^{12}_{X_B}(t) & a^{12}_{Y_B}(t) & a^{12}_{Z_B}(t) \\
a^{13}_{X_B}(t) & a^{13}_{Y_B}(t) & a^{13}_{Z_B}(t) \\
\vdots & & \\
a^{1n}_{X_B}(t) & a^{1n}_{Y_B}(t) & a^{1n}_{Z_B}(t)
\end{bmatrix} \quad
\mathbf{x} = \begin{bmatrix} \Delta X_B \\ \Delta Y_B \\ \Delta Z_B \end{bmatrix} \quad
\mathbf{l} = \begin{bmatrix} l^{12}_{AB} \\ l^{13}_{AB} \\ \vdots \\ l^{1n}_{AB} \end{bmatrix}, \tag{3.11}
$$

Table 3.1: Range biases (Hofmann-Wellenhof et al., 2008, p. 109)

Source	Effect
Satellite	Clock bias
	Orbital errors
Signal Propagation	Ionospheric refraction
	Tropospheric refraction
	Multipath
Receiver	Antenna phase center variation
	(depends on the direction of the signal)
	Clock bias

where the observations $l_{AB}^{jk} = R_{AB}^{jk}(t) - \varrho_{B0}^{k}(t) + \varrho_{B0}^{j}(t) + \varrho_{A}^{k}(t) - \varrho_{A}^{j}(t)$.

Similarly to the point positioning solution Eq. (3.7), the solution of this system is

$$\hat{\mathbf{x}} = \begin{bmatrix} \Delta\hat{X}_B \\ \Delta\hat{Y}_B \\ \Delta\hat{Z}_B \end{bmatrix} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{l},$$

and the final receiver coordinates are

$$\begin{aligned} \hat{X_B} &= X_{B0} + \Delta\hat{X}_B \\ \hat{Y_B} &= Y_{B0} + \Delta\hat{Y}_B \\ \hat{Z_B} &= Z_{B0} + \Delta\hat{Z}_B. \end{aligned}$$

3.2 Accuracy of the Positioning Solution

The positioning accuracy of GNSS is composed of many factors. The observed pseudoranges contain systematic and random errors, which then accumulate to the final positioning solution in a certain way. This section will explain what the error sources and magnitudes are in GNSS positioning, how the errors propagate, and how the accuracy of the positioning solution can be evaluated.

3.2.1 Measurement Errors

The code and phase pseudoranges are affected by systematic errors or biases as well as random noise. These errors can be classified in three groups: satellite-related errors, propagation-related errors, and receiver-related errors. The most significant of these errors have been listed in Table 3.1. (Hofmann-Wellenhof et al., 2008, p. 109)

Table 3.2: Typical magnitudes of range biases due to random noise (Hofmann-Wellenhof et al., 2008, p. 110)

Range	Bias
Code range (coarse code)	300 cm
Code range (precise code)	30 cm
Phase range	5 mm

Some of these biases can be modeled and their effect on the position accuracy can be reduced. Systematic errors can also be eliminated by combining observables appropriately, as was mentioned in the previous section. Differencing measurements of two receivers to the same satellite eliminates the satellite clock bias, and differencing measurements of one receiver to two satellites eliminates the receiver clock bias. This is why double-difference pseudoranges (see Section 3.1.3) are free of most systematic errors. (Hofmann-Wellenhof et al., 2008, p. 109)

The random noise mainly consists of actual observation noise and multipath effects. The range bias magnitudes due to noise are summarized in Table 3.2. The values assume a typical chip length of about 300 meters for a coarse (C/A) code and of about 30 meters for precise (P) code.

3.2.2 Satellite-Receiver Geometry

Satellite-receiver geometry has a significant influence on how the pseudorange errors cumulate into error in the positioning solution, as will be shown in this section. According to Hofmann-Wellenhof et al. (2008, p. 241), the covariance matrix of the estimated parameters reads

$$\Sigma_{\hat{x}} = \hat{\sigma}_0^2 \mathbf{Q}_{\hat{x}} = \hat{\sigma}_0^2 (\mathbf{A}^T \mathbf{Q}_l^{-1} \mathbf{A})^{-1}. \tag{3.12}$$

The diagonal elements of $\Sigma_{\hat{x}}$ contain the error of the positioning solution in the three coordinate directions. The off-diagonal elements express the correlations between the position errors. It is worthwhile to note that $\Sigma_{\hat{x}}$ is different from the covariance matrix of the range observations Σ_l that was mentioned earlier in Eq. (3.6). Instead of variances of the positioning solution, Σ_l contains the variances of the range measurements that were mentioned in the previous section.

This measurement uncertainty propagates to positioning solution covariances in $\Sigma_{\hat{x}}$ by the satellite geometry around the receiver. The combination of the measurement uncertainty and the satellite geometry determine the quality of the positioning solution together, as illustrated in Figure 3.3.

Pseudorange errors remain mostly random and uniform in scale for all pseudoranges when using same receivers and positioning techniques. Consequently, the changes in

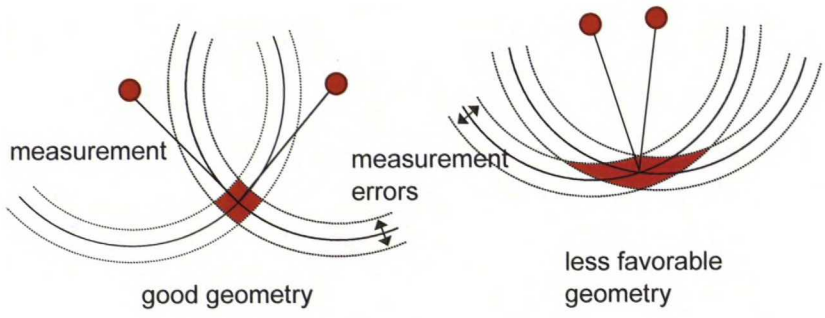

Figure 3.3: The effect of measurement error and observation geometry to the total error in position.

satellite-receiver geometry will account for most of the differences in the quality of positioning solutions, which is why a measure of this geometry is suitable for evaluating potential to provide accurate positioning.

3.2.3 Dilution of Precision

Measures of the quality of instantaneous satellite-receiver geometry are the dilution of precision (DOP) factors. In the simulation model that is used in this thesis, DOP factors are fundamental values that are used to evaluate the positioning quality for different points in the environment.

Following Hofmann-Wellenhof et al. (2008, p. 262), dilution of precision values are generally calculated from the inverse of the so called normal equation matrix $\mathbf{N} = \mathbf{A}^T\mathbf{A}$. The design matrix of the linear equations $\mathbf{A}$ determines the elements of the matrix. As was shown in Eq. (3.5), the design matrix for point positioning is formed by

$$\mathbf{A} = \begin{bmatrix} -\frac{X^1 - X_{r0}}{\varrho^1_{r0}} & -\frac{Y^1 - Y_{r0}}{\varrho^1_{r0}} & -\frac{Z^1 - Z_{r0}}{\varrho^1_{r0}} & c \\ -\frac{X^2 - X_{r0}}{\varrho^2_{r0}} & -\frac{Y^2 - Y_{r0}}{\varrho^2_{r0}} & -\frac{Z^2 - Z_{r0}}{\varrho^2_{r0}} & c \\ \vdots & & & \\ -\frac{X^n - X_{r0}}{\varrho^n_{r0}} & -\frac{Y^n - Y_{r0}}{\varrho^n_{r0}} & -\frac{Z^n - Z_{r0}}{\varrho^n_{r0}} & c \end{bmatrix}.$$

In a constellation simulation model, the goal is to calculate the DOP values with a given receiver position and the satellite and pseudolite positions on a specific epoch. Therefore the receiver position can be considered accurate to the location that the DOP values are calculated for, and the observation constituents can be simplified to $X_{r0} = X_r$, $Y_{r0} = Y_r$, and $Z_{r0} = Z_r$.

The matrix for DOP calculations, which is sometimes called the cofactor matrix and denoted by $\mathbf{Q}$, can then be formed by

$$\begin{aligned} \mathbf{Q} &= \mathbf{N}^{-1} = (\mathbf{A}^T\mathbf{A})^{-1} \\ &= \begin{bmatrix} q_{XX} & q_{XY} & q_{XZ} & q_{Xt} \\ q_{XY} & q_{YY} & q_{YZ} & q_{Yt} \\ q_{XZ} & q_{YZ} & q_{ZZ} & q_{Zt} \\ q_{Xt} & q_{Yt} & q_{Zt} & q_{tt} \end{bmatrix}. \end{aligned} \tag{3.13}$$

This is a square matrix, where three rows and three columns correspond to the combinations of the receiver and satellite positions and one row and one column correspond to the receiver clock. The diagonal elements of this matrix are the basis for calculating DOP values.

There are three different kinds of dilution of precision values that can be combined from the matrix:

$$\begin{aligned} GDOP &= \sqrt{q_{XX} + q_{YY} + q_{ZZ} + q_{tt}} && \text{geometric dilution of precision} \\ PDOP &= \sqrt{q_{XX} + q_{YY} + q_{ZZ}} && \text{position dilution of precision} \\ TDOP &= \sqrt{q_{tt}} && \text{time dilution of precision.} \end{aligned} \tag{3.14}$$

Position dilution of precision is the most commonly used DOP value, since it represents how much the range error affects the position accuracy. If the design matrix has been calculated with receiver and satellite coordinates in a local coordinate system, i.e. the x- and y-coordinates are the local horizontal coordinates and the z-coordinates vertical, two more DOP values can be separated:

$$\begin{aligned} HDOP &= \sqrt{q_{XX} + q_{YY}} && \text{horizontal dilution of precision} \\ VDOP &= \sqrt{q_{ZZ}} && \text{vertical dilution of precision.} \end{aligned} \tag{3.15}$$

To get the horizontal and vertical DOP, the design matrix must be calculated with local coordinates. If geocentric coordinates are used for calculating the design matrix, as is usually the case, the matrix must be transformed into local coordinates before proceeding to the calculation of the cofactor matrix (Eq. (3.13)).

The DOP values depend on the available satellites and pseudolites, so they change in time as the satellites move on their orbits. A simulated GNSS system by Hofmann-Wellenhof et al. (2008, p. 264-265) with 27 satellites on 3 orbital planes at the height of 23,200 km would provide PDOP values between 1.1 and 2.9 at any time all across the Earth for clear-sky sites. The smaller the DOP values, the better the constellation of satellites is for positioning. Satellite geometry can be considered good when PDOP values are less than 3 and HDOP values less than 2. This kind of DOP values should be available anywhere on Earth if the sky is completely clear of obstacles. Any obstacles will block satellite signals and make the DOP values higher.

Dilution of Precision in Relative Positioning

Dilution of precision can also be calculated for relative positioning. These DOP values are considered relative DOP values. (Hofmann-Wellenhof et al., 2008, p. 265)

Relative DOP values are calculated with the same procedures as with the point positioning values. The only difference is that instead of the design matrix for point positioning $\mathbf{A}$, the design matrix for a baseline vector $\mathbf{G}$ (from Eq. (3.11)) is used in the calculation formulas. Following Nielsen (1997), the cofactor matrix $\mathbf{Q}$ is now formed by

$$\mathbf{Q} = (\mathbf{G}^T\mathbf{G})^{-1} = \begin{bmatrix} q_{XX} & q_{XY} & q_{XZ} \\ q_{XY} & q_{YY} & q_{YZ} \\ q_{XZ} & q_{YZ} & q_{ZZ} \end{bmatrix}.$$

The matrix contains no elements that come from the receiver clock, since the clock errors are eliminated in double differencing. The only DOP values for relative positioning are then PDOP, VDOP, and HDOP:

$PDOP_{DD} = \sqrt{q_{XX} + q_{YY} + q_{ZZ}}$	position dilution of precision for double differences
$HDOP_{DD} = \sqrt{q_{XX} + q_{YY}}$	horizontal dilution of precision for double differences
$VDOP_{DD} = \sqrt{q_{ZZ}}$	vertical dilution of precision for double differences.

These DOP calculation methods also apply to carrier phase observations. When the integer ambiguities are resolved, the resulting design matrix $\mathbf{G}$ will be exactly the same. (Nielsen, 1997)

An Upper Bound for Dilution of Precision in Relative positioning

The relationship between dilution of precision for point positioning and for relative positioning has been studied by Nielsen (1997). He discovered that the DOP factors for double-difference relative positioning are bounded from above by the corresponding DOP factors for point positioning. This was proven for the case of four satellites.

Teunissen (1998) studied Nielsen's theorem further, proving that it is also valid when there are more than four satellites involved in the positioning solution. In the special case when one of the satellites is located at the center of gravity of the receiver-satellite configuration, the DOP values for point positioning and double-difference positioning were found to be exactly the same.

Chapter 4

Considerations to Modeling Signal Visibility

The simplest way to model the visibility of GNSS signals is by assuming a so called *geometric definition of visibility*. Geometric visibility is defined by taking a straight *line of sight* (LOS) between the receiver and the satellite, and determining whether it is blocked by something in the environment. All possible refraction and diffraction effects on the signal are ignored. Two points in a terrain are considered to be mutually visible if the straight line segment between them does not go below the terrain at any point. Sometimes it is required that the line segment is strictly above the terrain and is allowed to touch the terrain only at the end points. (Stamm, 2001, p. 77)

In the case of electromagnetic radiation, however, the signal does not travel a straight path in reality, and signal propagation between two points does not occupy only a single line between them. This space for signal propagation is often modeled as a so called Fresnel zone, which is introduced in the following Section 4.1. This means that obstacles not directly on the signal path can have an effect on the signal. (Stamm, 2001, p. 66-71,77-78)

Signal strength also gets weaker as the receiver moves further away from the transmitter, due to free-space path loss. As the GNSS satellites are usually at a very long distance from the receiver, for example GPS satellites at about 20,000 km away, their signal strength stays mostly constant when moving on the Earth's surface. This is explained in a little more detail in Section 4.2. In the case of pseudolite signals, however, it is very important to take the distance to the pseudolite into account, as will be seen in Section 4.3.

4.1 Propagation of Radio Signals

As was stated above, if signal visibility is defined geometrically, the signal travels along a single line in space and the reflections and diffractions take place on certain points. In reality, the signal traveling from the GNSS satellite to the receiver occupies not

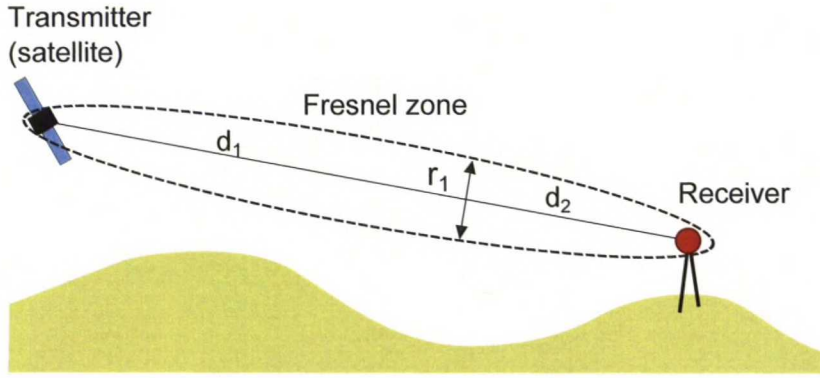

Figure 4.1: The first Fresnel zone ellipsoid between a receiver and a transmitter GNSS satellite. The locations of antennas of the receiver and transmitter are in the foci of the ellipsoid and the width of the zone is proportional to the distances d_1 and d_2 from these points.

only a line, but a certain space in between them. This is because electromagnetic radiation travels in wave fronts rather than single rays, according to the wave nature of electromagnetic radiation. Neglecting the effects caused by this phenomenon can lead to unrealistic results in the visibility analysis. (Stamm, 2001, p. 77)

This section will explain about Fresnel zones and the effect of signal refraction, but only to the extent that needs to be considered in this thesis. The information presented here follows work done by Stamm (2001), which should be referred to for more details about modeling signal propagation in electromagnetic radiation.

4.1.1 Fresnel Zones

The space that the signal mostly uses for propagation is called the first Fresnel zone. The first Fresnel zone is an ellipsoidal space between the transmitter and the receiver, whose dimensions depend on the distance between the transmitter and receiver and the wavelength of the radiation. This is illustrated in Figure 4.1. The first Fresnel zone radius r_1 is defined as

$$r_1 = \sqrt{\frac{\lambda d_1 d_2}{d_1 + d_2}}, \tag{4.1}$$

where λ is the wavelength and d_1 and d_2 are the projected distances from the end points to the point of interest. As the first Fresnel zone is where the signal will use to travel between two points, obstacles within it will cause path losses that reduce the signal power. If the first Fresnel zone is clear of all obstacles, the path losses are negligible. The points are then said to be *radio visible*. If there is an obstruction within the first Fresnel zone that is not larger than the zone, but larger in relation to the signal

wavelength, the signal will be diffracted and weakened. If there is an obstacle that completely blocks the first Fresnel zone, the signal transmission is blocked entirely.

Radio visibility is a very strict definition of visibility, and usually many points that are not radio visible with the transmitter still have enough signal strength to receive the signal. A little more loose definition is to say that the points are *θ-radio visible* if the power loss between them is less than a predefined threshold θ.

According to Stamm (2001, p. 93), a good approximation for signal visibility in most cases is to say that the signal has sufficient path clearance if the direct LOS is clear of the most outstanding obstruction for at least 60% of the radius of the first Fresnel zone r_1. So if the most outstanding obstruction is closer than 60% of the radius r_1 to the direct LOS, the signal has insufficient path clearance to be considered clearly visible.

Fresnel Zones of GNSS Signals

It can be seen from Eq. (4.1) that the radius of the Fresnel zone is proportional to the wavelength of the radiation. In the case of visible light, which has a very small wavelength around 860 nanometers, the zone becomes so thin that it is almost nothing but a simple line. GNSS carrier frequencies are generally in the microwave range and their wavelength is much longer, for example 19.05 centimeters (1575.42 MHz) for the GPS band L1, so the Fresnel zones are significantly wider and path losses are more probable.

Also, since the GNSS satellites are so far away, the distances d_1 can be considered infinitely large. Then $d_1 + d_2$ in the radius of the Fresnel zone (Eq. (4.1)) would be simplified to d_1, and the radius would become

$$r_1 = \sqrt{\frac{\lambda d_1 d_2}{d_1}} = \sqrt{\lambda d_2}$$

Fresnel loss can cause varying degrees of degradation to the GNSS signal by reducing the strength and quality of the signal. Ma et al. (2001) presented that GNSS signals with varying levels of degration can be classified in the following three categories:

1. Clear signal: The first Fresnel zone is clear of all obstacles, so the Fresnel path loss can be ignored. Signal quality is good and power loss is only caused by free-space loss and atmosphere absorption.

2. Shadowed signal: There are obstructions within the first Fresnel zone and the signal power and quality are degraded. The direct signal can still probably be used for positioning.

3. Blocked signal: The first Fresnel zone is completely obstructed and direct signal propagation is prevented. Signal may still reach the receiver through multipath effects: diffraction and reflection.

4.2 Satellite Signal Power Levels

Due to free-space path loss, the strength of an electromagnetic radiation signal at the receiver is an inverse square function of the range to the transmitter, so the signal grows weaker with more distance. The signals of GNSS satellites are transmitted at a constant power level of about 15 W, but they must then travel over 20,000 km and spread around to cover the whole Earth, weakening considerably in the process. (Cobb, 1997, p. 47)

Because the GNSS satellites are so far from the receivers, the range from a satellite to a receiver on Earth changes by about 20% at most. Also the satellites' antennas are designed in a way that their radiation patterns compensate for the change in distance to the Earth surface. Because of this, the strength of the signals from satellites remains relatively constant no matter how much the receiver moves. (Cobb, 1997, p. 49)

GPS signal specification (Global Positioning Systems Directorate, 2011) defines that when a navigation signal reaches the Earth surface, it will have a minimum strength of -160 dB less than one Watt, or -160 dBW. This same power level can also be expressed as 130 dB less than one milliwatt, or -130 dBm, in more convenient units. This means that, if there are no additional obstructions to the signal of a GNSS satellite, the signal will be visible wherever on the satellite's side of the Earth. It is not possible to move so far from the satellite that the signal would grow too weak to receive. For pseudolite signals, as the next section shows, this is not the case.

4.3 Pseudolite Signal Power Levels

Imagine a pseudolite installed to a ridge of an open-pit mine to help the heavy trucks and mine personnel navigate with better accuracy. The trucks might be driving at the other end of the mine 3 kilometers away, or as close as 10 meters from the pseudolite location. The difference in signal power at these two distances is considerable. The ratio of these two *far* and *near* ranges is 1:300. If a pseudolite transmits at a constant power level, the receiver will see the change in signal power level according to the inverse square of the distance separating them. In this case it would be about 50 dB. If the pseudolite transmit power is set in a way that the signal is still at receivable level of -130 dBm at the other end of the mine, the power level near the pseudolite would be about -80 dBm.

This is an overwhelmingly strong signal when compared to GNSS satellite signals of -130 dBm that the receiver is trying to receive at the same time. In the example situation, this is what would happen: in the far side of the mine, navigation with satellites and the pseudolite together would work well, but closer to the pseudolite, the pseudolite signal would begin to jam the satellite signals to an extent that the receiver will lose track of all but the pseudolite signal (Cobb, 1997, p. 49-50). This is called the near/far problem of pseudolites. See Figure 4.2 for an illustration of the issue.

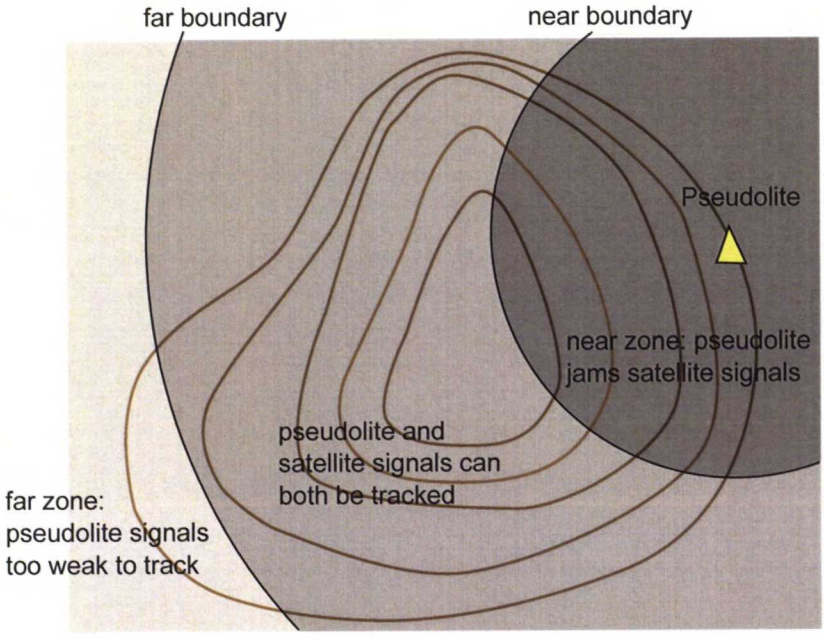

Figure 4.2: Zones of the near/far problem, illustrated for a pseudolite installed at the edge of an open-pit mine. The case depicted here would be highly unpractical, since the pseudolite would effectively prevent positioning in almost a third of the mine. It could still be a realistic case, if the receiver has a small tracking margin.

4.3.1 Near/Far Problem

The *far boundary* is a range where the pseudolite signal starts to be too weak to track. The *near boundary* is the range where the pseudolite signal starts to interfere with the satellite signals. The receiver must be between these boundaries to be able to use the signals together for positioning.

At the surface of the far boundary, the pseudolite signal strength would be equivalent to the minimum specified satellite signal strength of about -130 dBm. The distance to the far boundary is determined only by the pseudolite transmit power. The ratio of the distances to the far and near boundaries, however, is determined not only by absolute power level, but also the tracking margin of the receiver. (Cobb, 1997, p. 50-51)

The tracking margin of a receiver depends on the design of the receiver and how much interference it can handle compared to the strength of the tracked signal. A tracking margin of a typical GPS receiver sets the near/far ratio at about 6:1. For a functional system, it will be necessary to add safety factors to this value, so the system could be required to work with a near-far ratio of 3:1, for example. This could be crippling for many potential pseudolite systems. (Cobb, 1997, p. 50-54)

4.3.2 Working Around the Near/Far Problem

The research done on pseudolites by Cobb (1997, p. 54-62) introduces many ways that make it possible to build working systems with pseudolites despite the near-far problem. These include system designs that avoid the problem by using trajectory constraints, antenna patterns, and separate antennas for pseudolite and satellite signals. There are also special techniques that can be applied in the pseudolite signal transmission so that it would not interfere with satellite signal reception, like using transmission frequencies and codes different from the satellites or pulsing the transmission.

Trajectory Constraints and Antenna Patterns

Sometimes a pseudolite augmentation system can be designed in a way that the *near regions* are in no circumstances entered during navigation, and there is no need for pseudolite signals outside the far boundary. In this kind of a system the near-far problem would not be a problem at all.

The Integrity Beacon Landing System (IBLS) is an example of this kind of system. It is a relative carrier phase navigation system that provides airplanes with enhanced positioning accuracy when landing to airports. IBLS uses pseudolites to make it possible to initialize the integer ambiguities rapidly. After the ambiguities initialization, the receiver will not need the pseudolite signal for navigation. This kind of system is possible to design in a way that the trajectory of the airplanes goes over the pseudolite and never enters the near region.

Other systems can also be designed to make the near region inaccessible. For example a pseudolite used by ground vehicles could be installed at the top of a tower or a pylon whose height is larger than the near radius. Alternatively, if the movement of the navigating vehicles is limited to a specific area, the pseudolites could be installed so far from the edge of that area, that the vehicles would never come too close.

If the navigation trajectories of the vehicles are predictable, it is also possible to shape the radiation patterns of the transmitter and receiver antennas to deliver a weaker signal at the direction of the closest approach. It is possible to achieve up to 12 dB reduction in antenna gain from the far boundary to closest approach, which corresponds to 4 times better near-far range ratio.

Out-of-Band Transmissions

A pseudolite can also be built to transmit at different frequencies than the satellites. If the transmission frequency is sufficiently different, the pseudolite signals will be filtered out by the GNSS receiver and they will not hinder satellite signal reception.

The disadvantage of this solution is that a separate radio frequency (RF) tuner must be used to receive the pseudolite signals. A receiver with this kind of functionality would

be much more complex and expensive. In addition, the timing of the two different RF tuners can drift relative to each other, which will require consideration of more complicated positioning algorithms.

New Spreading Codes

It is also possible for a pseudolite to transmit their signals by using special ranging codes, which are different from the standard GNSS codes like the C/A code. This would require modifications to the receivers, since standard GNSS correlators could not understand the special codes. The upside is that non-standard codes have much less cross-correlation to the standard codes, and consequently the dynamic range of the receiver is better for the pseudolite signals.

Pulsed Pseudolite Signals

A pseudolite can only interfere with the satellite signals when it is transmitting. If the pseudolite transmits only 10% of the time, it will interfere only during that time. The remaining 90% of the time the receiver would only see the satellite signals. During these conditions, most receivers will be able to track satellites and pseudolites together. This kind of pulsed transmission technique is the most successful and widely used way to mitigate the signal interference problem of pseudolites.

There are various pulsing schemes that can be used for pseudolite signals, but a typical one uses short strong pulses of about 10% duty cycle. The pulse must be strong enough to get the pseudolite signal to the receiver despite the short duration. During the pulse, the signal from the pseudolite will generally be strong enough to exceed the dynamic range of the receiver. This will make the receiver *saturate* at a given power level, which the pulsing technique uses to its advantage.

Assume that a pseudolite uses a 10% duty cycle, alternating between for example 100 μs of transmission time and 900 μs off time. From a distance its signal might have a peak power level of -120 dBm, which will not yet saturate the receiver. Virtually all GNSS receivers integrate code signals over at least one C/A code epoch (1 ms), so the resulting average power level from the pseudolite would then become -130 dBm, which is the same that the satellites have. When the receiver approaches the pseudolite, the peak power level rises. When it exceeds a level of about -107 dBm, the receiver will start to saturate. After that point the peak power level seen by the receiver will remain constant, no matter how close it gets to the pseudolite. The receiver will still be able to track the pseudolite signal during the pulses, and desaturate to track satellite signals in between the pulses.

Ideally, the pulsing technique solves the near-far problem for one pseudolite. But in a system of multiple pseudolites, the pulsed signals can interfere with each other and the satellite reception. This sets some limitations for the number and locations of the

pseudolites: the total pulse cycle in any location must not exceed 20%, otherwise the satellite reception will be hindered. This means that only two pulsing pseudolites are allowed to be visible at once in any location. Also, a receiver in saturated condition can effectively receive only one signal, and only if it is significantly stronger than other signals. This can be solved for example by time synchronization of the pseudolite pulses, but the maximum duty cycle constraint still applies.

Chapter 5

Prior Research on GNSS Signal Visibility Simulation

GNSS signal visibility simulation systems have been developed for different purposes and with different properties. The purposes of GNSS simulation tools are generally one of (a) system analysis, (b) location and time planning, or (c) positioning quality improvement (Roongpiboonsopit, 2011). System analysis tries to understand the designs, factors and issues that affect the GNSS positioning performance, and uses the simulation results for this purpose. Location and time planning systems provide the means to decide where and at what time GNSS surveys should be conducted to get the most accurate and reliable results. Positioning quality improvement aims to reduce unnecessary errors in the GNSS positioning by simulating the positioning environment and signal propagation.

Previous research has studied, for example, how GNSS availability behaves in dense urban environments, where there are high buildings blocking the line of sight to the satellites, for example in Tokyo (Suh & Shibasaki, 2003, 2007), Seoul (Lee, Suh, & Shibasaki, 2008) or London (Taylor, Li, Kidner, Brunsdon, & Ware, 2007; Li, Taylor, Kidner, & Ware, 2008). Most of these systems only consider the visibility to GPS satellites, or satellites from other GNSS such as Galileo, to determine how well it is possible to navigate in the study area. One study also discusses the possibility of pseudolite augmentation (Suh & Shibasaki, 2003).

The studies have presented two different methods to model the environment. One approach is to use three-dimensional terrain data sets, which may include for example a digital elevation model and three-dimensional geoinformation system (GIS) data, and model the obstacles around the receiver (Suh & Shibasaki, 2003, 2007; Taylor et al., 2007). Another approach is to use photogrammetric methods on the receiver site, which effectively determine the dimensions and distances to the surrounding obstacles (Marais, Berbineau, & Heddebaut, 2005; Li et al., 2008).

Some of these prior studies have served as inspiration to the methodology chosen in this thesis, and the purpose of this chapter is to present those studies in further detail.

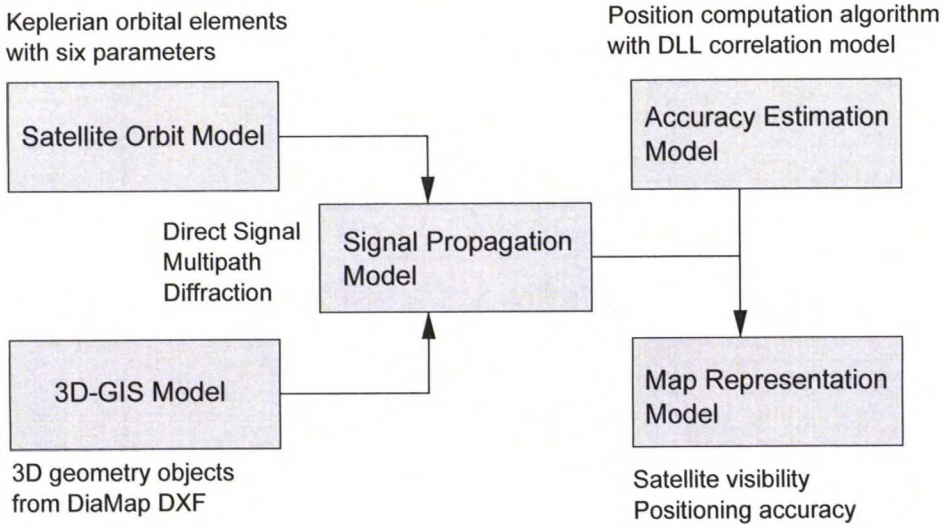

Figure 5.1: Organization of the simulation system by Suh and Shibasaki (2007).

Studies that are included in this chapter are: a simulation system created by Suh and Shibasaki (2003, 2007), a mobile GNSS visibility tool by Marais et al. (2005), and a GNSS quality of service simulation method by Roongpiboonsopit (2011). In addition, the last section of the chapter introduces some existing toolkits and software that can be used in signal visibility and navigation accuracy analysis.

5.1 Pseudolite and Satellite Visibility in Urban Environments

The GNSS satellite visibility simulation model created by Suh and Shibasaki (2003) was first created for testing optimal positions for pseudolites in urban environment. Their simulation model consists of a three-dimensional GIS data, a model of satellite orbits, and coordinates of installed pseudolites. The simulation model was implemented in the Java language, with GPS satellite and map data in individual classes. The structure of the Suh and Shibasaki's simulation model is depicted in Figure 5.1. A somewhat similar organization was implemented in the simulation software in this thesis.

The three dimensional GIS data used in the Suh and Shibasaki's simulation was the Mitsubishi Corporation DiaMap product, a 3D vector map of urban environments, and Shinjuku district in Tokyo was chosen as the focus area. They used YUMA almanacs from U.S. Coast Guard Navigation Center to model the satellite orbits. The simulation program worked in three stages: first it divided the study area into grid cells, then analyzed how many satellites are visible for each grid cell, and then finally calculated the corresponding DOP values.

The grid cells in the simulation were regular tetragons two meters wide, and they extended throughout the whole study area of Shinjuku. The satellite visibility and DOP values were calculated for each hour for the duration of 24 hours. The algorithm used in determining whether a satellite is visible or not, was a very simple one. It only

determines the straight line of sight from a grid cell to a satellite, and tests if there are any obstacles intersecting this line.

In their later study, Suh and Shibasaki (2007) developed the simulation program further, to take also the multi-path propagation and diffraction of the signal into account, and study their effect on the positioning accuracy. This was achieved by simulating the signal propagation in the 3D environment with a ray-tracing method. Testing results from the new simulation system confirmed that multipath signals from otherwise invisible satellites deteriorate the positioning accuracy significantly in highly built up areas such as Shinjuku.

5.2 GNSS Availability and Multipath Evaluation for Mobile Receivers

A somewhat similar GNSS signal visibility simulation tool has been developed by Marais et al. (2005). Their tool was created for analyzing GNSS availability in land transportation applications in urban environments. Instead of using a three-dimensional GIS model to model the surrounding environment, they used an optical approach to determine the visible sky on a specific trajectory. The simulation relied on Analytical Graphics, Inc. Satellite Tool Kit software to simulate satellite positions in the sky. After the simulation of satellite positions, the surrounding environment was analyzed from a video record that is shot driving on the chosen trajectory. Distances and heights of the main surrounding obstacles are defined for each epoch with a mono-camera stereovision process.

The first version of the tool only determined the satellites visible or blocked, and did not take the nature of the obstacles into account. This was done by creating a mask of the surrounding obstacles for each instant and analyzing whether the satellites are in the masked or unmasked area. The tool was then further developed to predict the multipath signals that are reflected from the obstacles and then received. The final simulation tool is called PREDIctive Software for Satellite Availability in the field of Transport (PREDISSAT) and is now applied for the development of a new railway control and command safe system for trains.

5.3 Methodology for Real-Time GNSS Quality of Service Prediction

In her doctoral dissertation, Roongpiboonsopit (2011) presented a detailed methodology for predicting the quality of service (QoS) of GNSS. The predicted measures were essentially visibility, availability, and accuracy for point positioning, but also more

general quality measures like average availability, average accuracy, continuity, and reliability for navigation along a route.

The techniques introduced in the dissertation utilize high-resolution 3D data sets and efficient line-of-sight algorithms to enable QoS prediction in real time. The work also introduced new satellite selection and route planning methods that take the QoS into account.

5.4 Existing Toolkits and Software

The software of all of the above research projects seems to have not been distributed to the public, so they could not be used as a base for the software of this thesis. Fortunately, there exist other software and toolkits that can help with GNSS signal visibility simulation.

One good software for GNSS observation planning is the Planning by Trimble Navigation, Inc., whose functions include for example determining satellite availability and calculating DOP values for specific epochs and time periods. There is no way of modeling individual obstacles in the program, but it is possible to set a simple elevation angle mask that changes according to the azimuth. Trimble Planning was used in this thesis to compare and evaluate the correctness of DOP calculations.

The Satellite Tool Kit (STK) by Analytical Graphics, Inc., which was used in the simulation system of Marais et al. (2005), is also a good candidate for helping in the simulation of GNSS signals. There exists a free version of the software, but it is quite limited in its functions. Some of the many commercial add-ons for STK seemed quite potent, but, due to their expensive cost, they were not used in this thesis.

Also many 3D-GIS software, such as ArcGIS 3D Analyst by Esri, have functions that could be used in GNSS signal visibility analysis. With 3D-GIS software the accommodation of a wide variety of environment data is very easy, and the software usually also have various visibility algorithms built in. However, it would be difficult to include GNSS satellite orbits into these programs.

After the search for readily available software tools, not one of them seems to be able to model pseudolites together with GNSS satellites. It was then decided that the software for this thesis had to be written mostly from scratch, and the search focus was turned to software libraries and toolkits that would help in the implementation of the required functions.

A few libraries and open source projects ended up as part of the simulation software: GPS Toolkit (Tolman et al., 2004), Mayavi (Ramachandran & Varoquaux, 2011), collada-interface (Scarpino, 2011), and TinyXML (Thomason, 2011). They will be introduced in more detail in Section 6.3.1 of the next chapter.

Chapter 6

Methodology

This chapter explains how the simulation software was put together. It begins by describing the main tasks of the software, then introduces the software's structure and design, and after that proceeds to explain the implementation and algorithms of the various functions.

6.1 Main Tasks of the Software

The main task of the software is to be able to determine what the positioning accuracy is like in a specified area and how it can be enhanced by adding pseudolites on specified locations. The functionality of the software can be broken down into smaller subtasks that perform this main task together. See Figure 6.1 for a simple use case diagram that shows the main tasks of the software.

As can be seen from the figure, the first of the software tasks is to load data files that contain 3D environment models, and store the geometry data within the program. The second task is to read GNSS ephemeris files to simulate the positions of satellites above the area at different times. The third task of the software is to include pseudolites into the simulation, specify their locations in the environment and their signal transmission range.

With these data, the software can proceed to the fourth task of calculating visibility of satellites and pseudolites within the environment, and then the fifth task, calculating dilution of precision values according to the visibilities. In order to simulate the effect on positioning accuracy when adding pseudolites to the area, the calculations can be conducted with only satellites, or together with pseudolites. The results of the simulation can then be visualized for example by showing locations with poor DOP values in the environment model.

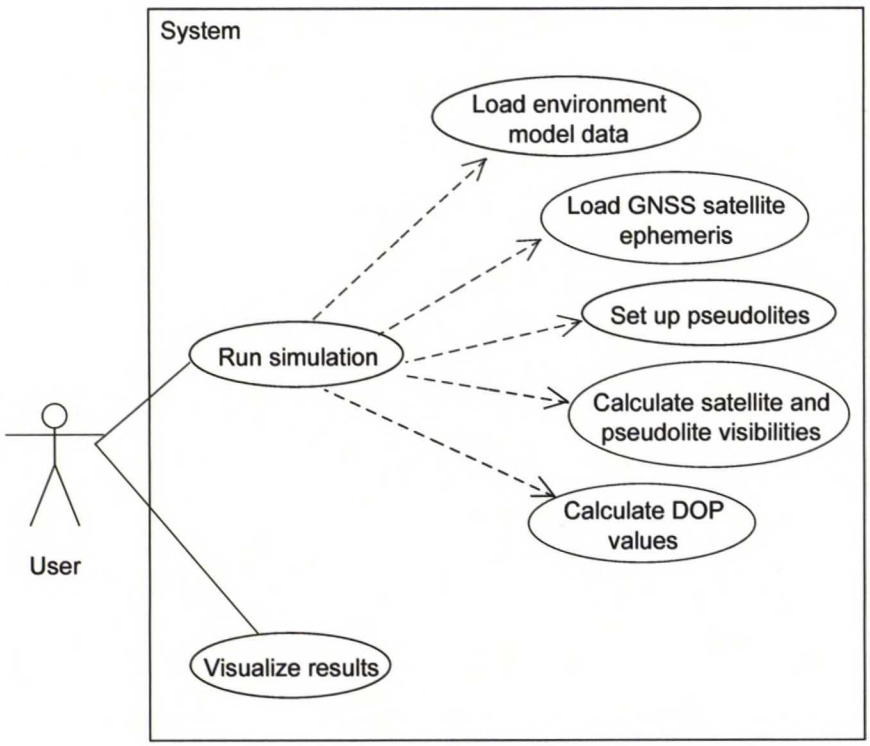

Figure 6.1: Main tasks of the simulation software, depicted as a simple use case diagram.

6.2 Design of the Simulation Software

To implement these functions, standard software engineering practices were followed, roughly according to the ICONIX process (Rosenberg & Scott, 2001). Starting with requirements analysis and use case modeling, then sketching the program structure with a domain model and robustness diagrams, finally a class diagram was produced. The final general structure of the software is depicted in Figure 6.2.

The modules of the program each perform a specific task for the main program. *Environment module* is responsible for reading 3D model files and storing the geometry data, *Satellites module* reads GNSS ephemeris or almanac files and can be used to retrieve satellite position information, and *Pseudolites module* handles all pseudolites included in the simulation. *Propagation analysis module* is used for calculating the visibilities, and *Accuracy analysis module* calculates dilution of precision values.

Receiver grid is a special class that consists of a 2D grid of points that cover the simulation area at a specified height from the environment ground. Visibilities are calculated for the locations specified by these grid points, and they are also used to store the calculated visibility and DOP results. *Scenario* is a class that represents a single complete simulation scenario, and it controls the other modules. Scenario class is also used for interfacing the rest of the program with simple calls to the various

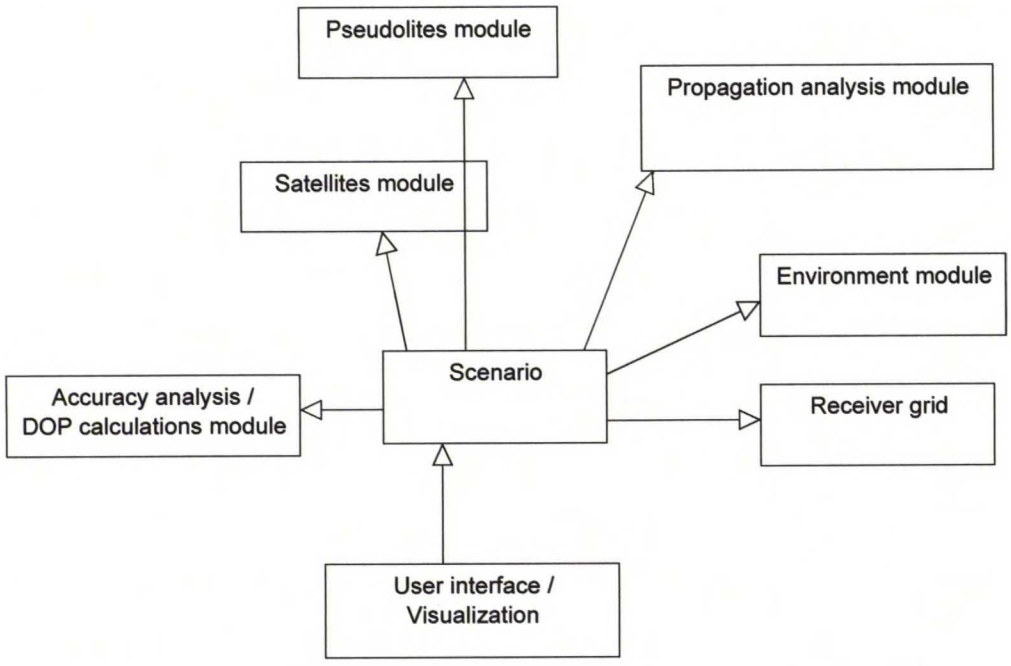

Figure 6.2: General structure of the simulation software with its main components and relations. The program is mainly interfaced through a central class called *Scenario*, which has ownership of all the significant high-level objects and assigns appropriate calls to the components for different functions.

tasks.

Data is transmitted in the software in a fashion described in Figure 6.3. First, environment data is loaded into the Environment module, satellite ephemeris data is loaded into the Satellites module, and pseudolites are created in the Pseudolites module. Then the Receiver grid is initialized, and visibilities are calculated in the Propagation analysis module using all of the data that was loaded earlier. The visibility information is in turn sent to the Accuracy analysis module that calculates the time series of DOP values. Finally, all necessary data is produced and it is used for visualization and interpretation of the simulation results.

6.3 Implementation of functions

The software is mostly implemented in the c++ programming language for calculation efficiency, easy portability of the code to other projects, and the possibility of using the GPSTk library (more about GPSTk in the next section). All functions that concern calculations are implemented in c++, and only interface and visualization parts are implemented in Python. Python was chosen here because of its easy scriptability and powerful visualization libraries.

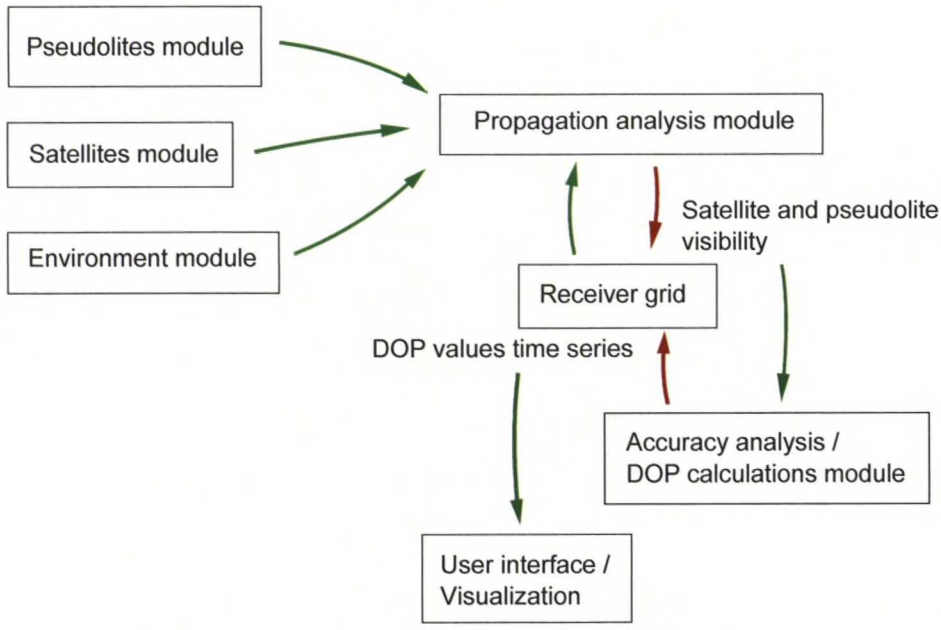

Figure 6.3: Data flow between the components of the simulation software. Visibility calculation in *Propagation analysis module* uses data from *Environment*, *Satellite*, and *Pseudolite* modules, as well as receiver position information from *Receiver grid*. The visibility information is used in turn to calculate dilution of precision values for different times. Final outputs of the program are DOP value time series for all Receiver grid points. These time series can then be visualized in different ways.

6.3.1 External Libraries Used in the Software

Even though the software was written from scratch for most parts, it uses external libraries for a few tasks. These external libraries will be briefly introduced here.

For reading and handling GNSS satellite data, a library called GPS Toolkit (GPSTk) (Tolman et al., 2004) is used. GPSTk is an open source library that has many solutions to GPS problems like reading and processing standard GNSS file formats such as RINEX. GPSTk was also used for various coordinate transformations.

The reading of KMZ files uses external code from the open source projects colladainterface (Scarpino, 2011) and TinyXML (Thomason, 2011). Their code was modified quite considerably during the development of the simulation software.

For displaying the results, a data visualization library for Python called Mayavi (Ramachandran & Varoquaux, 2011) is used. Mayavi can produce many kinds of visuals, but it was especially chosen for its proficiency at 3D plotting. It is capable of viewing the results of the analysis in 3D together with the environment 3D model, which makes the interpretation of the results easy.

6.3.2 Modeling the Environment

The main types of environment that the simulation will be conducted in will be harbors and open-pit mines. In harbors, the environment consists of mostly level ground with cranes and other structures above it. This is quite different from open-pit mines, where there are no structures but the ground has steep hills and large changes in elevation. Because the simulation must be able to include structures that occupy space above the actual ground, such as harbor cranes, modeling the environment with a simple digital elevation model would not work. Consequently, the environment in the software is modeled with three-dimensional polygons.

The software uses KMZ files created by Google SketchUp, a free three-dimensional modeling software, to load the environment model. KMZ files are actually zip packages that contain geographic information in a KML file (Wilson, 2008), and additional files that it refers to. Google SketchUp outputs KMZ-files that include two files, a COLLADA 3D-model file (Barnes & Finch, 2008), and a KML file that describes the model's geographic location and orientation. A file format like this was chosen for COLLADAs wide compatibility with various 3D modeling software, and because a KMZ file also conveniently stores a 3D-model's geographic reference information in the same package. In addition, both the KML and COLLADA files are open XML formats and therefore quite easy to read.

The environment models are created in SketchUp by importing digital elevation models (DEM), building geometries by hand, or combining these two tasks. Open-pit mine models were constructed in this work by importing an accurate digital elevation model of the mine area and aligning it with a background Google Maps image for geographic reference. SketchUp takes care of triangulating the model to polygons. See Figure 6.4 and 6.5 for an image of DEM data and an environment model created from it. The construction of harbor environment models did not need a DEM, but required manually building the structures (see Figures 7.6 and 7.7 in the next chapter).

In the simulation software, COLLADA files are parsed and the geometry information of the model is stored in a simple data structure. The geometry data is stored in two separate layers, one for geometries that represent the ground and one that represents all other structures. This is done in order to make sure that it is possible to explicitly define ground height at any location in the environment. See Figure 6.5 for a visualization of a loaded environment model of an open-pit mine and a grid of points representing receiver locations.

The geometries are stored as triangles and their vertex coordinates. The coordinates are stored in a local coordinate system specific to the model. The model's geographic location and orientation are stored as separate values, which are used to transform satellite locations into the local coordinate system (see Section 6.3.3). Since all of the geometries are simple triangles, determining their intersections with other geometries is quite straightforward, as can be seen in Section 6.3.4.

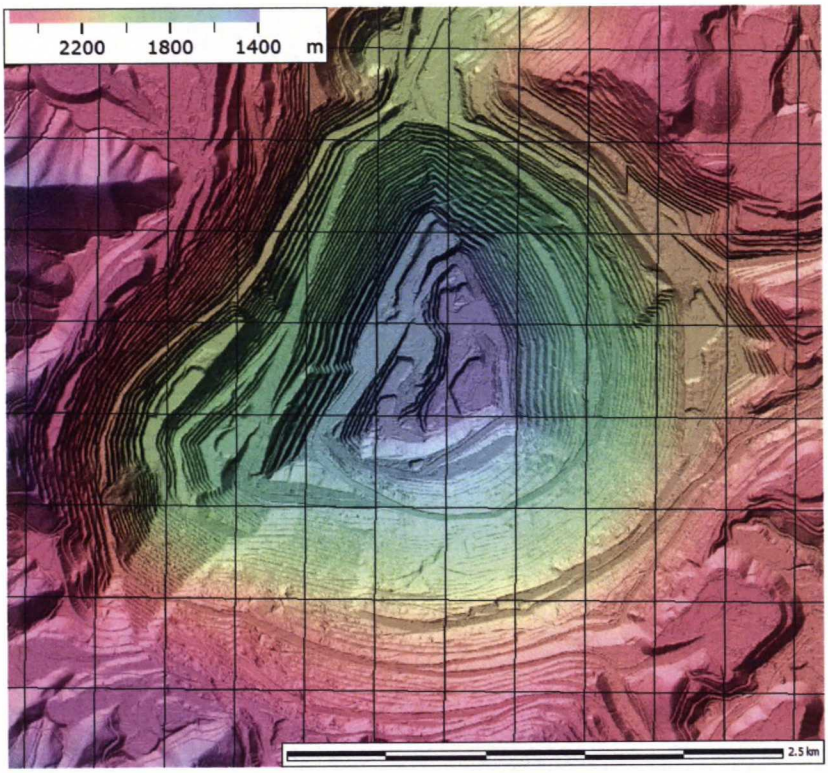

Figure 6.4: A digital elevation model data representing Bingham Canyon Mine, one of the largest open-pit mines in the world.

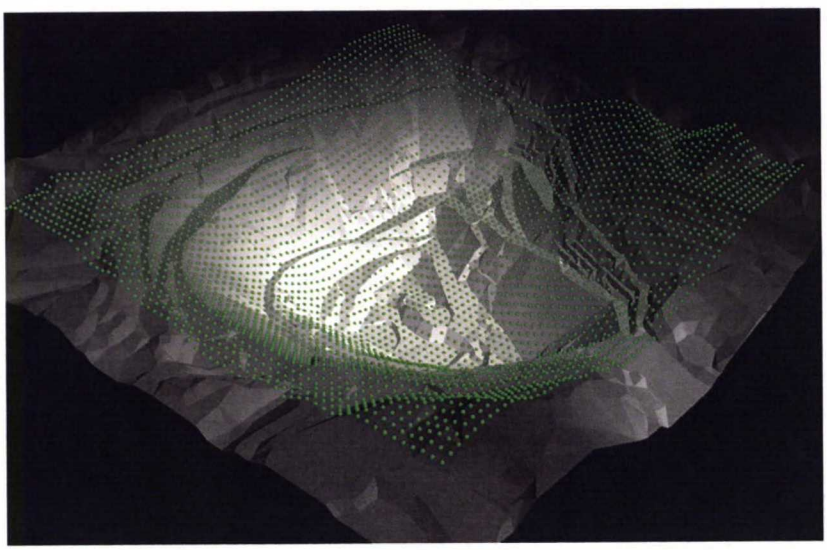

Figure 6.5: An environment model of Bingham Canyon Mine created with the digital elevation model data in Figure 6.4. There is a receiver grid initialized on the model.

6.3.3 Loading of Satellite Ephemeris

Loading of satellite ephemeris is implemented with the functions of the GPSTk library. The software supports reading all of the most common file types that are used for distributing satellite ephemeris, including RINEX nav, FIC, MDP, SP3, YUMA, and SEM (Tolman et al., 2004). The satellites' azimuth and elevation angles at different times can then be extracted by using the environment model's geographic location and orientation.

It is noteworthy to mention that the calculations in the software use the same satellite azimuth and elevation angles for all receiver locations in the simulation area, even though the locations are not exactly the same. Because the simulation areas are always relatively small, about 5 kilometers wide at most, the change in these angles would be negligible and there would be no benefit of calculating them separately for all points throughout the area.[1]

6.3.4 Propagation Modeling

According to Stamm (2001, p. 78), radio visibility is true as long as the first Fresnel zone between the transmitter and the receiver is free of any obstructions, as was stated in Section 4.1. The correct way to define radio visibilities is based on testing if there are intersections of the first Fresnel zones and the terrain. This can sometimes be a quite obstructing limitation and, in practice, the signal can often be visible even if the first Fresnel zone is partially obstructed.

In this simulation model, Fresnel zones intersections are tested to separate the completely unobstructed signals from partially obstructed, so called shadowed signals. Following the methodology of Ma et al. (2001) that was presented in Section 4.1, the visibility of a single signal is determined as one of the following three alternatives:

1. Clear signal: The first Fresnel zone is clear of all obstacles. Signal quality is good.

2. Shadowed signal: There are obstructions within the first Fresnel zone. Signal power and quality is degraded, but the direct signal can still probably be received and used for positioning.

3. Blocked signal: The direct line of sight is blocked. Direct signal propagation is prevented.

[1]Since the satellites are approximately 20,000 km away, the change in an angle across an area 5 kilometers wide would be at most about 0.014 degrees. This corresponds to 1.25 meter deviance to location at 5 kilometers distance, which is still well below the accuracy of the environment model and will not change the simulation results significantly.

Compared to the categories of Ma et al. (2001), the definition of a blocked signal is slightly stricter here, since it is easier to calculate and it ensures even higher quality signals as unblocked. The blocked signal with the above definition could in practice still reach the receiver through multipath effects: diffraction and reflection. This is, however, neglected in the model and it is presumed that these multipath signals can be determined and they are not included in the positioning calculations.

The visibility algorithm works in this order: First it checks whether there are obstructions blocking the direct line of sight. If the direct line of sight is blocked, the signal is considered blocked. If not, the first Fresnel zone is studied for intersections. If there is something within the first Fresnel zone, the signal is marked as a shadowed signal. Otherwise it is a clear signal.

The pseudolite signal visibility assumes pulsed transmissions, as described in Section 4.3.2. This means that there are no *near regions* where the pseudolites would jam satellite signals in the simulation. There is a specifically set distance to the far boundary, however, after which the pseudolite signal is no longer visible. The restrictions that pulsed transmissions set on the number and locations of pseudolites are not implemented in the simulation model, but they have to be considered when planning the pseudolites locations and signal ranges. Simply put, there can be no more than two pseudolites visible at any given location.

6.3.5 Calculation of the Dilution of Precision Factors

After the visibilities have been calculated, the dilution of precision values are calculated as described in Section 3.2.3, for each receiver point and epoch. This simulation model is designed to evaluate the availability and expected accuracy of GNSS positioning especially in systems that use relative positioning, but the DOP values for point positioning are used in the calculations for simplicity in implementation. Since relative positioning DOP values are bounded from above by point positioning DOP values (see Section 3.2.3), also point positioning values are well suited for this simulation model.

The actual DOP calculation is done in the Accuracy analysis module of the software. The module can be configured whether to use only signals with completely clear radio visibility, or all signals with line-of-sight visibility in the calculations. After the calculations the DOP values are stored as time series for each point in the receiver grid.

6.3.6 Visualizations

All visualizations are done in the Python environment by utilizing the Mayavi data visualization library (Ramachandran & Varoquaux, 2011). A basic visualization of the results is done by assigning colors to the different points in the receiver grid according

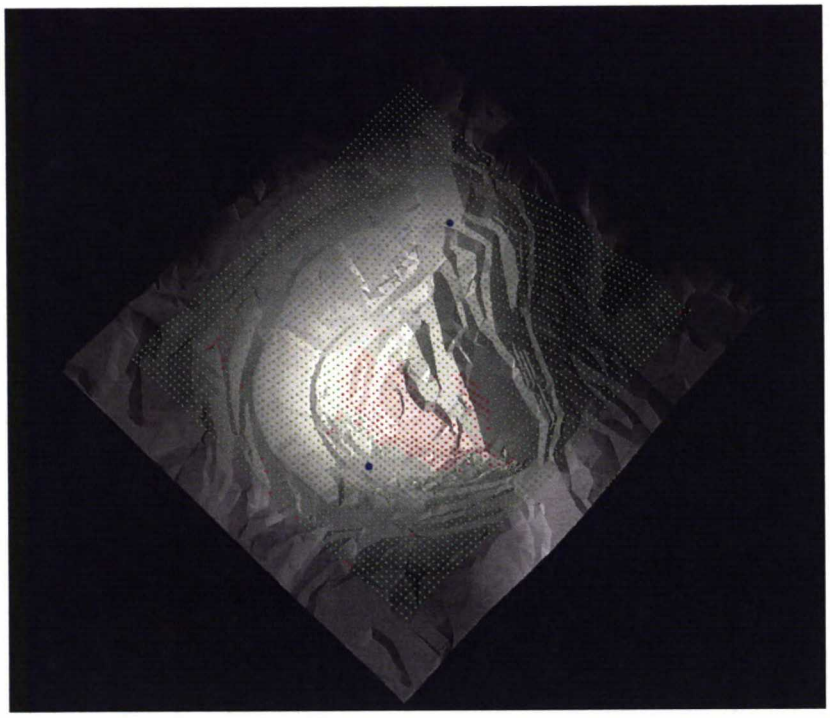

Figure 6.6: A visualization of the results of a visibility simulation. Points in green have at least six satellites or pseudolites visible at the time of the simulation (11/26/2002 at 12:00), and points in red have less than six visible. The two pseudolites in the area are depicted as large blue dots.

to the visible signals or DOP values of the points. A sample visualization can be seen in Figure 6.6.

Mayavi is a powerful and versatile tool for visualizations and there are also many other tools available in the Python environment. This is the main reason that Python was chosen as a secondary platform. In addition to these 3D visuals, with visualization libraries it would be easy to produce also for example 2D plots of the DOP time series or sky plots of visible satellites and pseudolites for individual points. In this thesis only this kind of 3D visualizations are presented, because the main interest is to see the general view of the GNSS performance.

6.4 Using the Software

The c++ source code of the program is compiled into a Python extension module, so that it can be used with Python scripts. This is achieved by using a tool called SWIG (Beazley, 1996) that wraps c++ interfaces and variable types into ones that Python understands.

The software is then run with a Python script. A sample script is listed in Figure 6.7. It can be seen that the script first creates a Scenario object, then defines the

```
print 'Importing pvamodule..'
import pvamodule
print 'Importing numpy..'
import numpy
print 'Importing mlab..'
from mayavi import mlab

# Load data for analysis scenario
sc = pvamodule.Scenario()
sc.loadEnvironment('bingham.kmz')
sc.loadSatellites('igr11942.sp3')

# Create the receiver point grid
sc.initializeReceiverGrid(50,5)

# Calculate satellite visibilities
stime = sc.satellites.getInitialTime()
etime = sc.satellites.getFinalTime()
print 'Start time ' ; stime.printTime()
print 'End time ' ; etime.printTime()
sc.time_interval = 15
sc.calculateSatVisibilities(stime, etime)

# Add a pseudolite
pl2_x = 1000.0
pl2_y = 482.0
pl2_z = sc.environment.groundHeight(pl2_x, pl2_y) + 50.0
pl2_loc = pvamodule.Point(pl2_x, pl2_y, pl2_z)
sc.pseudolites.addPseudolite(pl2_loc, 0.0)

# Calculate pseudolite visibilities
sc.calculatePlVisibilities()

# Calculate DOPs
sc.calculateDops()
```

Figure 6.7: A Python script that runs the simulation software.

files that make up the environment and satellite ephemeris, and initializes the Receiver grid. Then satellite and pseudolite visibilities are calculated. Notice how a pseudolite is added to the model before calculating pseudolite visibilities. Finally, the DOP values are calculated. Visualizations are done after this with separate scripts.

Chapter 7

Simulation Test Cases

The simulation software was tested on data from the Bingham Canyon Mine, which is located close to Salt Lake City in the United States. Bingham Canyon Mine, started in 1906 and still in full operation, is the largest copper mine in the USA and the world's deepest man-made excavation. The mine extends 4.5 kilometers from side to side and is about 1.2 kilometers deep. It will make a good example for satellite visibility simulations, since a lot of GNSS satellite signals are bound to become blocked for receivers in the bottom and ridges of a mine this deep.

The simulation software was also tested on a specifically constructed model of a harbor environment with crane structures. Harbor cranes are tall structures that effectively block many signals for receivers below and near them, reducing positioning accuracy and availability just where the straddle carriers or other container handling vehicles move most often. A case with two harbor cranes next to a large cargo ship is studied here.

This chapter introduces the data that was used for the environment models and satellite ephemeris, and shows the results of the visibility simulations with and without pseudolites. First, the case of the Bingham Canyon open-pit mine is explained thoroughly. After that, the harbor case is introduced with a little less detail.

7.1 Test Case: Open-Pit Mine

The environment model for the open-pit mine was created by processing a digital elevation model (DEM). The DEM data for the Bingham Canyon Mine was obtained from Utah Automated Geographic Reference Center (2012) (Utah AGRC). The DEM has been created from a Light Detection and Ranging (LiDAR) laser scan of the Salt Lake City area that was conducted in 2006-2007. The resolution of the downloaded DEM product is 2 meters, and the ground elevation accuracy is about 1.25 meters.

7.1.1 Processing the Digital Elevation Model

First, a region where the Bingham Canyon Mine is located was selected and downloaded from Utah AGRC. Then the mine area was cropped from the surroundings with a DEM editing software called MICRODEM (Guth, 2010). The cropped DEM that only covers the actual mine was shown in the previous chapter in Figure 6.4. In order to use the DEM in Google SketchUp, it was converted to USGS-DEM format (U.S. Geological Survey, National Mapping Division, 1998), an ASCII file format for storing digital elevation models designed by the United States Geological Survey.

The USGS-DEM file was loaded into SketchUp, which triangulates the elevation data to a specified amount of triangles in the importing process. Then a 3D polygon model of the environment is presented for editing. Additional structures like buildings could be added to the environment model in this stage of processing, but for an open-pit mine the plain ground geometry is sufficient. As a final step in creating the environment model, a background image is loaded into SketchUp from Google Maps, and the 3D model is carefully aligned to it in order to include correct geographic reference information. Then everything is ready, and a KMZ file can be exported to be used as the environment model in the simulation software.

7.1.2 Satellite Ephemeris Data

There are many different formats that are used for storing and distributing satellite ephemeris and almanacs, of which RINEX nav, FIC, MDP, SP3, YUMA, and SEM are supported in the simulation software created in this work. An almanac file only contains the orbital elements of the satellites and therefore only provides good estimates of the satellite locations. This would perfectly suffice for the case of a visibility simulation, but there is also no reason why not to use more accurate ephemeris files. This test case uses ephemeris data in the SP3 format, which is a widely used file format for storing accurate satellite ephemeris information.

The simulation test case used the International GNSS Service (IGS) Rapid ephemeris data for GPS satellites from November 22, 2002. The satellite positions in Rapid ephemeris data are slightly less accurate than in Precise ephemeris data, and were only chosen after experiencing some trouble with a Precise ephemeris file. The ephemeris file was downloaded from the IGS ftp servers (International GNSS Service, 2004), which host ephemeris data from GPS, GLONASS and other GNSS satellites, as well as Earth rotation and meteorological parameters. Only GPS satellites were included into the simulation because currently most GNSS receivers only support GPS.

This simulation test case was intended to display visibility of satellites and navigation accuracy generally. The date of the simulation data therefore did not matter as long as it is a day that represents typical satellite performance. November 22, 2002 was chosen in a quite random way.

The GPS satellites go around their orbits in exactly half a sidereal day, about 11 hours and 58 minutes, which means that the satellite constellation repeats its form twice a day. Also the Earth's rotation determines which satellites are visible at different times. Consequently, a full day's observations are required in order to test for all different geometries of visible satellites. This test simulation took a full 24 hours of observations from 00:00 to 23:45 with a time interval of 15 minutes.

7.1.3 Pseudolite Locations

DOP values depend on the geometry of the available satellites and pseudolites, so the pseudolite signals cannot help to improve DOP values if they don't provide an enhancement to this geometry. It is therefore important to choose the pseudolite locations not only in a way that makes their signals widely visible, but also in a way that provides the receivers additional strength to the positioning geometry.

The simulation model does not choose pseudolite locations automatically, so good locations have to be chosen with trial and error. Fortunately, it is quite easy to test if the locations are good or bad by examining the pseudolite visibilities separately. In this test case, the visibility simulation was tested with two different pseudolite arrangements, one that has a single pseudolite high on the edge of the mine, and a second that has two pseudolites a little bit lower on the hills. Illustrations of the chosen pseudolite locations can be seen in Figure 7.1. The visible receiver points from these pseudolites are depicted in Figure 7.2.

It can be seen that, in both cases, most of the receiver points inside the mine are provided with at least one pseudolite signal. Since the pseudolites are inside the mine in both cases, the pseudolite signals are no longer visible over the ridges of the mine. It is expected that the first pseudolite setting will already help to increase the dilution of precision values to some extent, but the second setting with two pseudolites will be even more beneficial.

It should be noted that the pseudolite signals in the system of two pseudolites are overlapping across most of the central mine, as can be seen in Figure 7.1. If this kind of system would be implemented in practice with pseudolites that are using pulsed transmissions, the signals would have to be time synchronized or the problem would have to be dealt with using some other technique.

7.1.4 Simulation Results

Like previously mentioned, the satellite visibilities were calculated for 24 hours at a time interval of 15 minutes. This resulted in a total of 96 time instances, or epochs. After the calculations it was noticed that there were nine epochs in which the software could not retrieve the satellite locations correctly, since on these epochs there were no visible satellites detected. These erroneous epochs were either in the very beginning

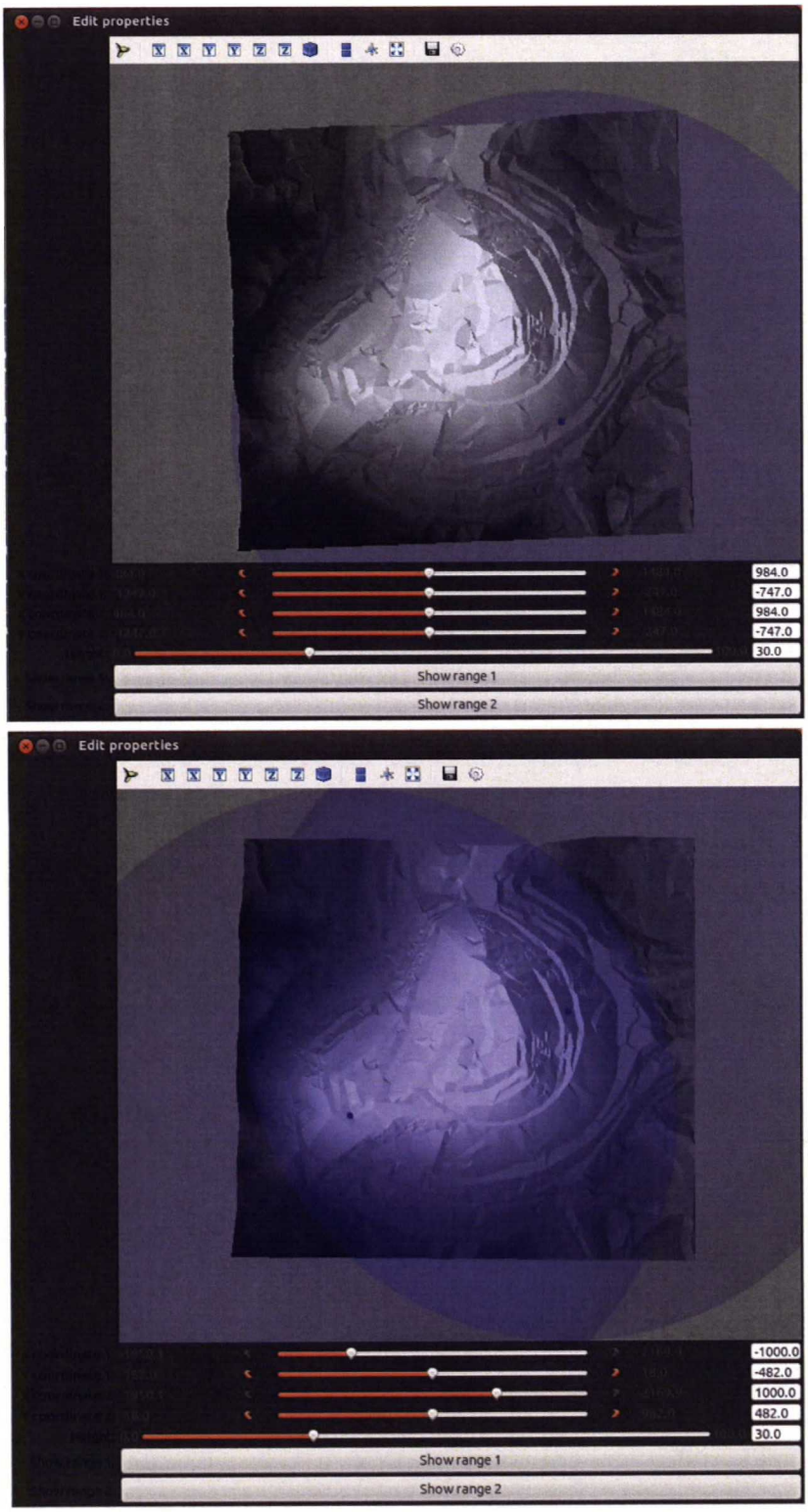

Figure 7.1: Planning of pseudolite locations for the simulation test case. The controls where the pseudolite locations can be modified can be seen below the visualization scenes. The far boundaries of the pseudolites are set to 3 km, and they are depicted here as spheres of that radius around the pseudolite locations.

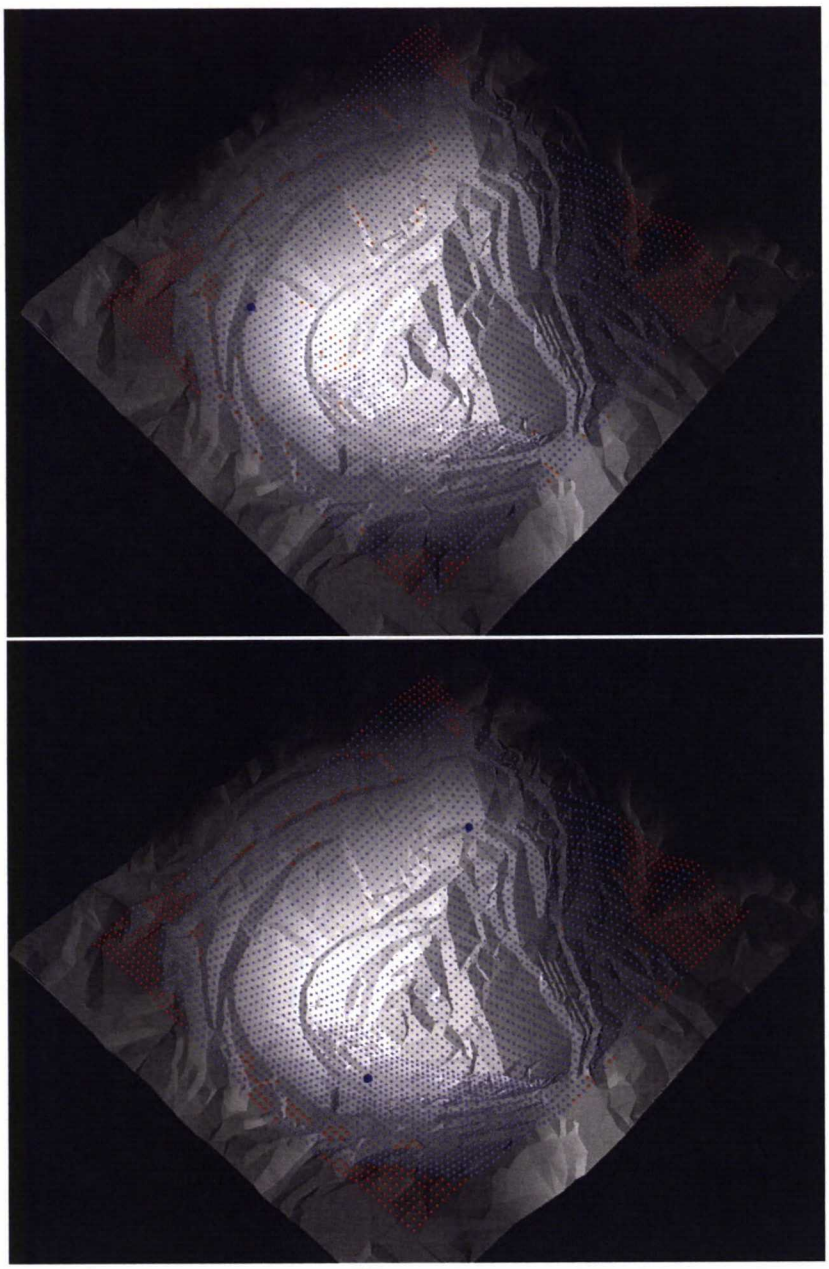

Figure 7.2: Visibility of pseudolites from the receiver points. The light blue points have at least one pseudolite visible, and red points have no pseudolites visible. The pseudolite locations are marked as larger blue dots.

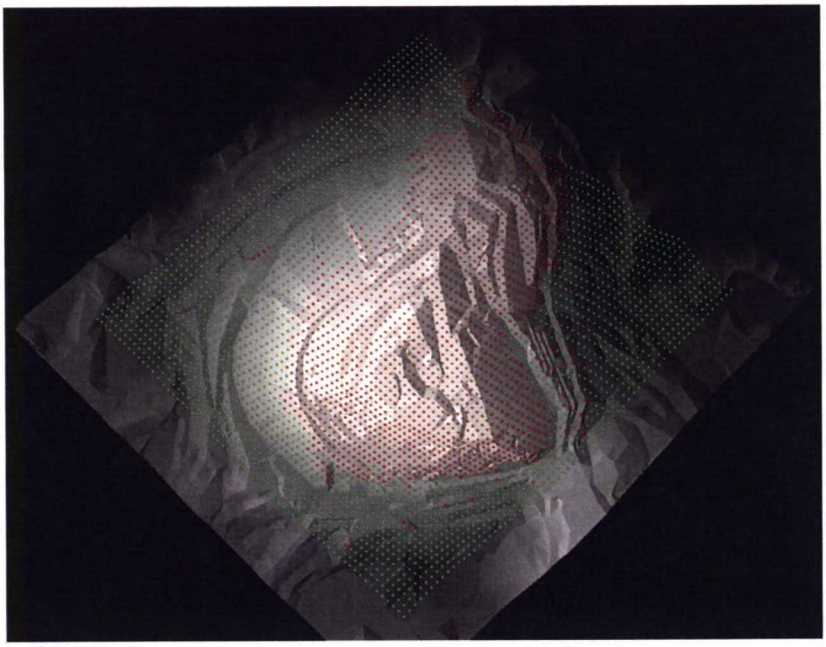

Figure 7.3: Results of the simulation without any pseudolites. Receiver points in green have good expected positioning accuracy. Receiver points in red have at least 30 epochs of 96 when DOP values are not good. Only about 42% of the receiver points are green.

or the end of the ephemeris file, so the errors must have been caused by the lack of interpolation data for GPSTk. Anyway, this issue had to be taken into account when visualizing the results.

The desired visualization in this test case was a one that displays the worst case expected positioning accuracy in the area, so it had to be a combination of the DOP values on all epochs in 24 hours. According to Hofmann-Wellenhof et al. (2008, p. 265), good geometry is available in a receiver point when its PDOP is less than 3 and HDOP is less than 2. This was set as the boundary for good and bad DOP in the simulation.

In the final visualization, a receiver point was considered to have good expected positioning accuracy if there is good DOP in at least 66 of the 96 epochs. This threshold of 30 epochs was set to eliminate the effect of the erroneous nine epochs, and to allow for bad DOP values in a few epochs. Because the requirements for a good DOP are quite strict, there would not have been good expected positioning accuracy anywhere on the model without this threshold.

The results of the simulation are displayed in Figures 7.3, 7.4, and 7.5. It can be seen from Figure 7.3 that the areas outside the mine have good positioning geometry even without the pseudolites, but the bottom and the hills have it much worse. Figures 7.4 and 7.5 display that the addition of pseudolites significantly improves the positioning geometry inside the mine. Two pseudolites at the lower hills remove the points with bad geometry almost completely, leaving only a few in the recesses of the bottom of the pit.

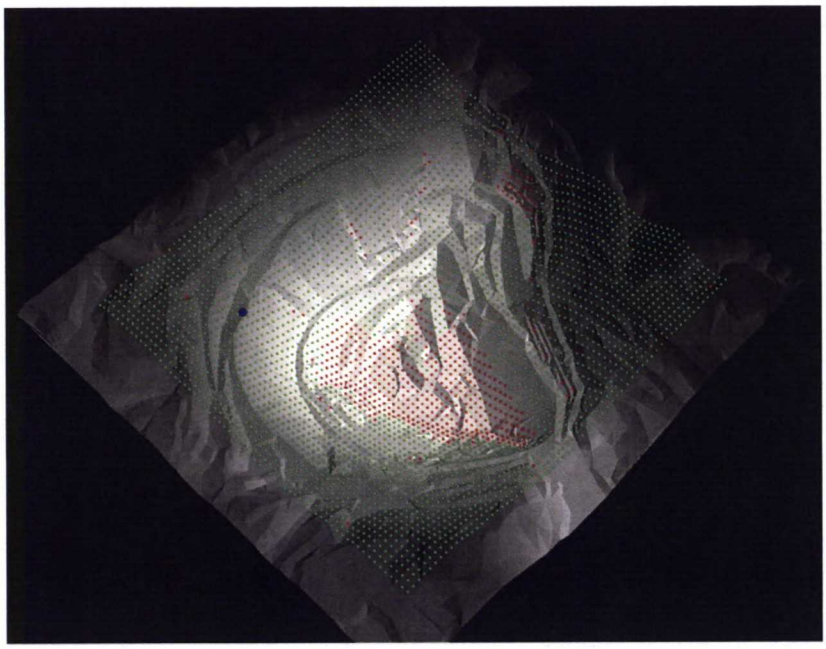

Figure 7.4: Results of the simulation with a single pseudolite on the ridge of the mine. Receiver points in green have good expected positioning accuracy. Receiver points in red have at least 30 epochs of 96 when DOP values are not good. The pseudolite is marked by a larger blue dot. About 87% of the receiver points are green.

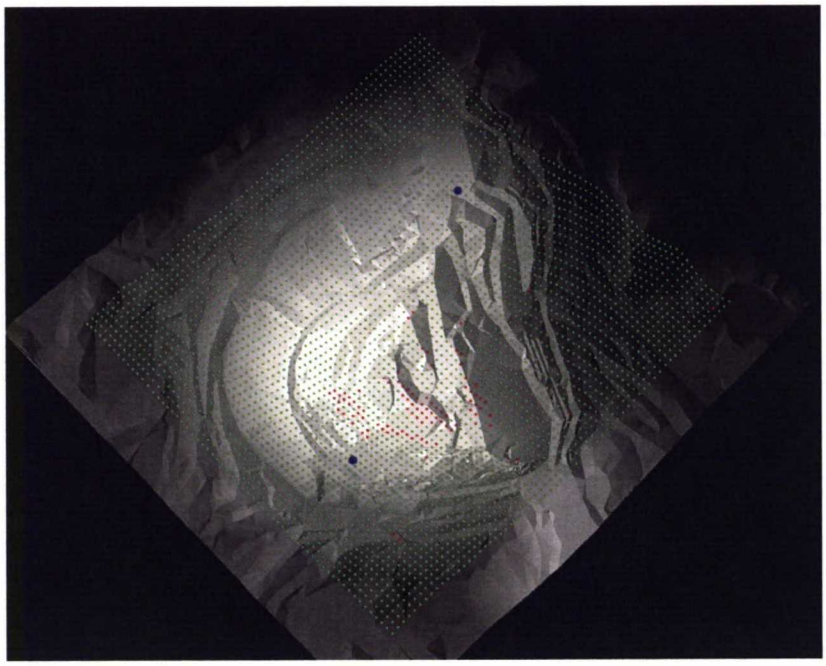

Figure 7.5: Results of the simulation with two pseudolites on the lower hills. Receiver points in green have good expected positioning accuracy. Receiver points in red have at least 30 epochs of 96 when DOP values are not good. Pseudolites are marked by larger blue dots. Now about 98% of the receiver points are green.

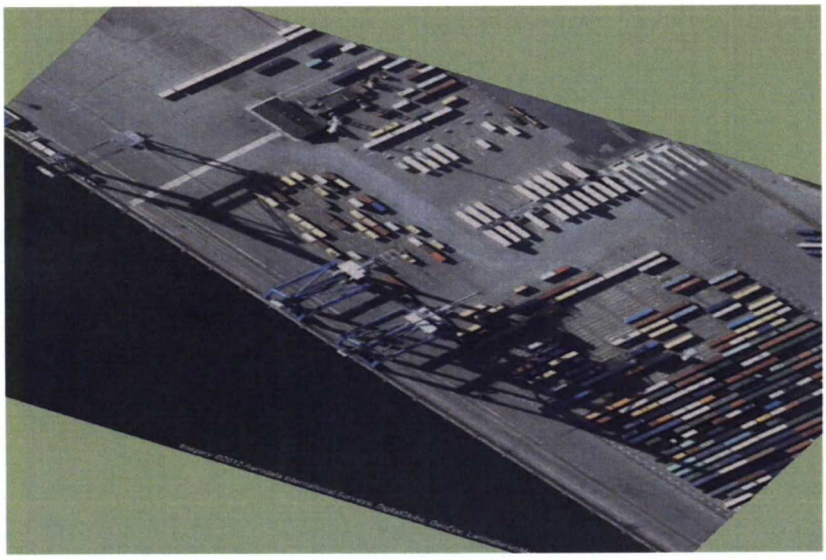

Figure 7.6: Modeling the harbor environment in SketchUp, step 1. A Google Maps background image of the Levantkaj pier (Copenhagen Malmö Port, Denmark) has been loaded. Two large harbor cranes can be seen on the pier next to the water line.

7.2 Test Case: Harbor

There was no input data for the harbor test case, since the structures of the model were all created manually. However, for the model to be realistic, the structures were created in the form of a pier in a real harbor. The harbor chosen for this test case was Copenhagen Malmö Port in Denmark, where the Levantkaj pier served as a model for the structures and their locations.

In contrast to the open-pit mine case, the creation of the environment started in SketchUp by loading the Google Maps image of the Levantkaj pier area. This stage can be seen in Figure 7.6, where the pier and two large harbor cranes can be seen. Then, simple models of the cranes were created, and also a model of a ship was constructed beside the cranes. The reason that the ship model was desired into the simulation, is that straddle carriers or other vehicles especially need to move around and below harbor cranes during loading and unloading a ship, and it is therefore a realistic case for a navigation simulation. The final harbor model created in SketchUp can be seen in Figure 7.7.

The simulation of the harbor case used the same date November 22, 2002 and ephemeris file that the open-pit mine case used. Also the simulation time was the same 24 hours, with a 15 minutes time interval between epochs.

The results of the harbor case simulation can be seen in Figures 7.8, 7.9, and 7.10. The simulation was conducted with exactly the same parameters as in the open-pit case, and also the visualization colors represent the expected positioning accuracy in the same way: green points have at least 66 epochs with a good DOP of the total 96 epochs, and red points have less.

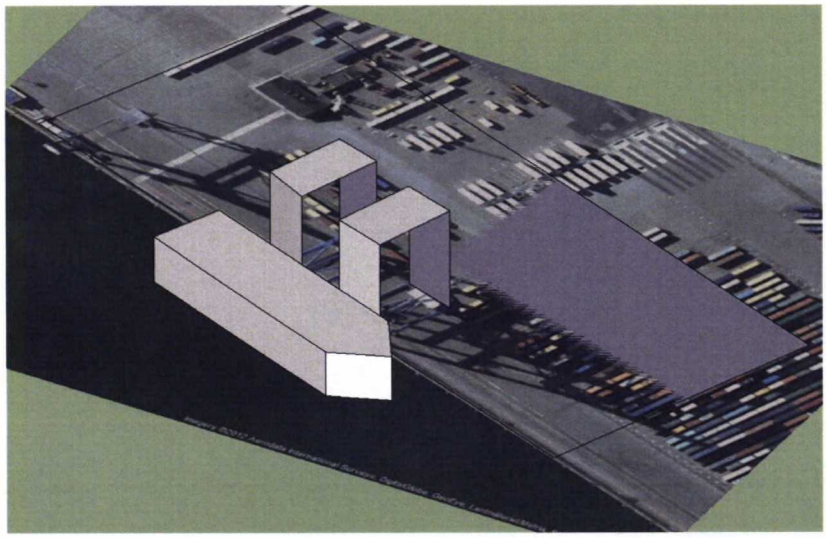

Figure 7.7: Modeling the harbor environment in SketchUp, step 2. Simple models of the harbor cranes and a ship beside them have been constructed. Notice also a rectangular geometry that lies on the pier. It is marked to be the ground layer in the simulation.

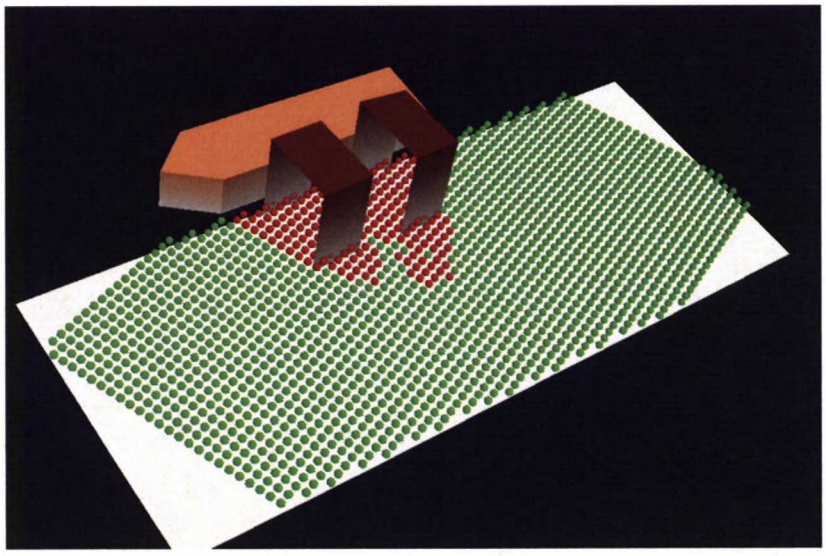

Figure 7.8: Results of the simulation of the harbor case without any pseudolites. Receiver points in green have good expected positioning accuracy. Receiver points in red have at least 30 epochs of 96 when DOP values are not good. Notice how the harbor cranes make the expected positioning accuracy worse around them.

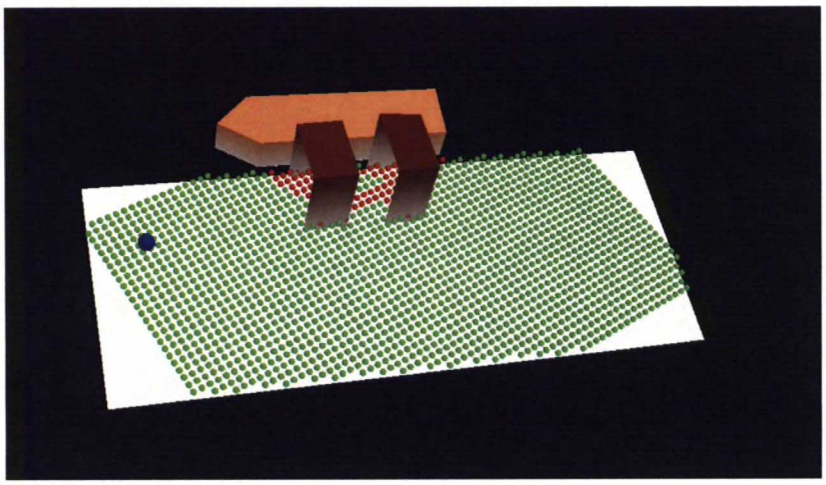

Figure 7.9: Results of the simulation of the harbor case with a single pseudolite. Receiver points in green have good expected positioning accuracy. Receiver points in red have at least 30 epochs of 96 when DOP values are not good. The pseudolite is marked by a larger blue dot.

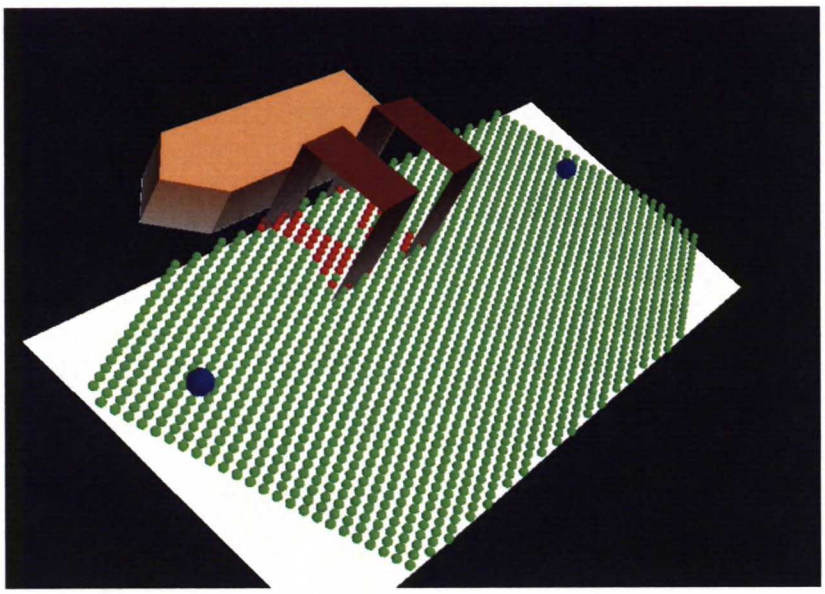

Figure 7.10: Results of the simulation of the harbor case with two pseudolites. Receiver points in green have good expected positioning accuracy. Receiver points in red have at least 30 epochs of 96 when DOP values are not good. The pseudolites are marked by larger blue dots. Notice how the expected positioning accuracy is still not good right below the harbor cranes.

Also the pseudolite additions were done in a similar way, first simulating a case with no pseudolites, and then with one or two pseudolites. It can be seen in figure 7.8, that the harbor cranes severely reduce the expected positioning accuracy in their vicinity. Pseudolites mitigate the problem to some extent, but the areas right below the cranes can have bad accuracy even with two pseudolites that provide additional signals from both sides. It must be considered, however, that the harbor crane models are created with very little detail, and in practice they might not block as many navigation signals as they do in this simulation.

It is also possible to install pseudolites to the bottoms of the upper parts of the harbor cranes, so that they would provide signals straight below them. This setting was not tested here, but it could have provided better expected positioning accuracy below the cranes than the separate pseudolites. If pseudolites were installed on the crane bottoms in reality, it would have to be considered that the pseudolites would then not have static coordinates, since they would be moving with the cranes.

Chapter 8

Discussion

It can be stated that the software implemented in this research works well and produces good results. The visualizations of the test cases presented in the previous chapter clearly show how the bottom of an open-pit mine and the areas below harbor cranes have bad DOP values, and consequently bad expected positioning accuracy.

There are still some observations that need to be pointed out before concluding the thesis. This chapter evaluates the results of the research, and presents some observations made by the author during the research process.

8.1 Efficiency of the Simulation Software

Environment modeling with 3D polygons, or triangles, produced good simulation results, but it made the calculations very heavy for large environment models. Because the simulation model was implemented without any proper optimizations, it had to calculate a huge number of triangle-line-segment intersections for one simulation. The number of calculations equals to the number of triangles in the model, times the number of receiver points in the grid, times the number of epochs, times the number of satellites in the constellation. The calculation of pseudolites visibilities was much faster, since they only have to be calculated for at most a few pseudolites and one epoch[1]. Also the DOP calculations were performed quite fast.

Using a laptop computer with average performance, the calculations for the open-pit mine test case took as long as about ten hours. The inefficiency of the visibility calculations can sometimes become a problem, since it restricts the use of very large and detailed environment models. Actually, also the open-pit mine model in the test case had to be restricted to about 5000 triangles to achieve the runtime of ten hours instead

[1]There is no difference in pseudolite visibilities between epochs, because the pseudolites are usually stationary.

of days. By contrast, the harbor test case, which had a much simpler environment model, needed only about ten minutes to run.

After the visibilities were calculated, however, they did not disappear from the memory, and they could be used with as many visualizations and analyses as was necessary, without having to run the calculations again. Also pseudolites could be installed and removed without having to calculate the satellite visibilities again, since only pseudolite visibilities were changed. In the end, after the long satellite visibility calculations, the flexibility of the software made it still quite comfortable to work with.

8.2 Evaluation of Methodology

The implementation of the software model was quite successful, and there is no doubt that the combination of c++ and Python works well here. Also the external libraries and toolkits, especially GPSTk and Mayavi, offered their functions to the simulation model without problems.

The methodology for determining satellite and pseudolite signal visibility did not settle for examining straight line segments, but it also modeled the signals' first Fresnel zones. This is a good method for modeling the direct propagation of the signal, but in reality there would be a lot of indirect propagation present. If the results of this simulation model were different from reality, the most significant reason would probably be the lack of multipath propagation modeling.

It can be discussed whether the dilution of precision value is the best value for modeling expected positioning accuracy. Some might argue that the observation errors in pseudoranges should also be included in the value. Indeed, if the simulation model could calculate for example the effect of multipath errors in the pseudorange, it would be wise to revise the value that describes expected positioning accuracy. The simulation model in this thesis has no way of modeling this kind of other factors, so it can only assume similar pseudorange error for all locations and satellites.

After the calculations of the satellite visibilities and DOP values, there was an unlimited number of ways to visualize the results. The test cases that were presented in this thesis used a combination of all calculated epochs and a definition of good and bad DOP, and assigned the receiver locations with appropriate colors. This thesis claims that this is a good way to visualize expected positioning accuracy with the data that was calculated, but no doubt also better ways exist.

8.3 Comparison to Prior Satellite Signal Visibility Systems

Some prior systems and software that were created for a similar purpose of GNSS signal visibility simulation than in this research were discussed in Chapter 5. Compared to

these systems, the simulation software in this thesis provides a new combination of functions. The focus of the simulation software in this thesis is also very unique, since its main goal is to help in planning good pseudolite locations.

The simulation system that most resembles the software created here is the one by Suh and Shibasaki (2003, 2007), which also provided functions to simulate pseudolite signals. The software created here, however, supports more complex and a larger variety of environments.

Chapter 9

Conclusions

This research started with the example of the GPS navigation of heavy machinery in open-pit mines and harbors, and how the availability and accuracy of navigation could be enhanced with pseudolites. The problem that was presented there concerned the location of the pseudolites: does a pseudolite placed to a given location actually improve the quality of GNSS positioning? This thesis has shown that it is possible to simulate and see how pseudolites improve the positioning accuracy around them. There is no longer need for guessing a good location for a pseudolite.

9.1 Results and Contributions

The beginning of this thesis introduced the Global Navigation Satellite Systems and pseudolites. It was discussed how pseudolites can enhance the positioning accuracy and reliability of GNSS navigation. Then the focus moved to the positioning algorithms and accuracy. It was concluded that the dilution of precision (DOP) values are well suitable for analyzing expected positioning accuracy, or what kind of positioning accuracy can be expected with a given set of satellite and pseudolite signals. After this background information, the thesis presented some important aspects of electromagnetic signal propagation and pseudolite properties. Without giving a thought to for example free-space path loss, wave-front propagation of the signal, or the near-far problem for pseudolites, the simulation model would not have produced very realistic results.

Then, a few prior research works that had been done on similar simulation models were introduced. Even though the actual software for any of those could not be used for the purpose of this thesis, they provided good insights and ideas on how to proceed with the implementation of this kind of a GNSS signal visibility simulation model. The details of the simulation software and specific implementation methods were then presented. The software has various separate functions, such as environment modeling,

signal propagation modeling, DOP calculation, and visualization of the results. New implementation strategies were presented for each of them in turn.

The software created in the process of this thesis works well and produces good results for evaluating the expected positioning accuracy. As can be seen from the test cases that were performed, the software worked equally well in two very different environments. The visualizations of the simulation results clearly reveal where the navigation accuracy is good, and where it is not so good. The simulation software is not very efficient and can run for long times with large environment models, but it does not fail to produce useful simulation results.

The results of the test cases done with the simulation model clearly showed, that it can be very difficult to maintain good GNSS positioning accuracy in obstructed environments like an open-pit mine or the vicinity of a harbor crane. The results also showed that pseudolites installed to appropriate locations can reduce these problems significantly, if not remove them completely. These were expected outcomes of the test cases, and they proved that the software works as intended.

9.2 Recommendations for Further Development

The result of this thesis, the simulation software, can be used for the planning of good pseudolite locations. This application has been mentioned in this thesis many times, since it was an important motivation for this work from the beginning. The best way to evaluate the performance of the software further, would therefore be to test it against real observations in a real pseudolite augmentation system. This would certainly reveal the weaknesses that the simulation software still has, and help it to become more realistic.

The software is currently run with various Python scripts, as was explained in Section 6.4. This is good because it allows the user to define the simulation settings very precisely, but the addition of a more user friendly interface would still be welcome. The development of a graphical interface for, for example, setting the Receiver grid, inserting pseudolites, and controlling the visualizations, would no doubt make the software much faster and easier to use.

The inefficiency of the software is also worth fixing. With some optimizations to the data structure and processing of the environment model, it would take much less time to calculate the signal visibilities. The software could then perform simulations for much larger and more detailed environments.

9.3 Summary

The goal of this research was to develop a methodology and a software tool to simulate GNSS satellite and pseudolite signal visibilities in obstructed environments and determine the expected positioning accuracy there. This goal has been achieved by, first, studying the background of GNSS positioning and accuracy and the nature of GNSS satellite and pseudolite signal propagation, and then, developing a piece of software that simulates signal visibilities according to the principles that were revealed in the studies. The simulation software works as intended and proved its functionality in two test cases. Now the simulation software is only waiting to prove itself in practical use: to help design a pseudolite system in the real world.

References and Bibliography

Barnes, M. & Finch, E. L. (2008). *COLLADA Digital Asset Schema Release 1.5.0 Specification.* Sony Computer Entertainment Inc. **retrieved from** `www.khronos.org/collada/`

Beazley, D. M. (1996, July). SWIG : An Easy to Use Tool for Integrating Scripting Languages with C and C++. In *4th Annual Tcl/Tk Workshop, Monterey, California.*

Cobb, H. S. (1997). *GPS Pseudolites: Theory, Design and Applications.* (PhD thesis, Stanford University).

European Space Agency. (2012). Galileo Navigation. [Online, Accessed July, 2012]. **retrieved from** `http://www.esa.int/esaNA/galileo.html`

Global Positioning Systems Directorate. (2011, Sept.). *IS-GPS-200F Navstar GPS Space Segment/Navigation User Segment Interfaces.* **retrieved from** `www.gps.gov/technical/icwg/`

Guth, P. (2010). MICRODEM Homepage. [Online, Accessed August, 2012]. **retrieved from** `http://www.usna.edu/Users/oceano/pguth/website/microdem/microdem.htm`

Hofmann-Wellenhof, B., Lichtenegger, H. & Wasle, E. (2008). *GNSS - Global Navigation Satellite Systems.* Wien, Austria: SpringerWienNewYork.

International GNSS Service. (2004). IGS Data & Products. [Online, Accessed April 2012]. **retrieved from** `http://igscb.jpl.nasa.gov/components/compindex.html`

Lee, Y.-W., Suh, Y.-C. & Shibasaki, R. (2008). A GIS-based Simulation to Predict GPS Availability along the Tehran Road in Seoul, Korea. *KSCE Journal of Civil Engineering, 12*(6), 401–408.

Li, J., Taylor, G., Kidner, D. & Ware, M. (2008). Prediction and Visualization of GPS Multipath Signals in Urban Areas Using LiDAR Digital Surface Models and Building Footprints. *International Journal of Geographical Information Science, 22*(11-12), 1197–1218. doi: `10.1080/13658810701851396`

Ma, C., Jee, G.-I., MacGougan, G., Lachapelle, G., Bloebaum, S., Cox, G. F., ... Shewfelt, J. L. (2001, Sept.). GPS Signal Degration Modeling. In *Proceedings of the Institute of Navigation ION GPS-2001.*

Marais, J., Berbineau, M. & Heddebaut, M. (2005). Land Mobile GNSS Availability and Multipath Evaluation Tool. *IEEE Transactions on Vehicular Technology, 54*(5), 1697–1704. doi: `10.1109/TVT.2005.853461`

Montenbruck, O., Hauschild, A., Steigenberger, P., Hugentobler, U., Teunissen, P. & Nakamura, S. (2012). Initial Assessment of the COMPASS/BeiDou-2 Regional Navigation Satellite System.

Nielsen, R. (1997, Jan.). Relationship Between Dilution of Precision for Point Positioning and for Relative Positioning with GPS. *IEEE Transactions on Aerospace and Electronic Systems*, *33*(1), 333–338. doi: 10.1109/7.570809

Ramachandran, P. & Varoquaux, G. (2011). Mayavi: 3D Visualization of Scientific Data. *Computing in Science & Engineering*, *13*(2), 40–51. doi: 10.1109/MCSE.2011.35

Roongpiboonsopit, D. (2011). *Navigation Recommender: Real-Time iGNSS QoS Prediction for Navigation Services.* (PhD thesis, School of Information Sciences, University of Pittsburgh).

Rosenberg, D. & Scott, K. (2001, Oct.). Introduction to the ICONIX Process of Software Modeling. *informIT.* **retrieved from** http://www.informit.com/articles/article.aspx?p=167902

Scarpino, M. (2011). collada-interface Project Page. [Online, Accessed August, 2012]. **retrieved from** http://code.google.com/p/collada-interface/

Stamm, C. (2001). *Algorithms and Software for Radio Signal Coverage Prediction in Terrains.* (PhD thesis, Swiss Federal Institute of Technology (ETH), Zurich).

Suh, Y.-C. & Shibasaki, R. (2003). Assessment of Pseudolite Layout Under Urban Environments Using a Simulation System for Seamless Positioning. *KSCE Journal of Civil Engineering*, *7*(3), 261–266. doi: 10.1007/BF02831777

Suh, Y.-C. & Shibasaki, R. (2007). Evaluation of Satellite-based Navigation Services in Complex Urban Environments Using a Three-dimensional GIS. *IEICE Transactions on Communications*, *E90B*(7), 1816–1825. doi: 10.1093/ietcom/e90-b.7.1816

Taylor, G., Li, J., Kidner, D., Brunsdon, C. & Ware, M. (2007). Modelling and prediction of GPS availability with digital photogrammetry and LiDAR. *International Journal of Geographical Information Science*, *21*(1), 1–20. doi: 10.1080/13658810600816540

Teunissen, P. (1998, Apr.). A Proof of Nielsen's Conjecture on the GPS Dilution of Precision. *IEEE Transactions on Aerospace and Electronic Systems*, *34*(2), 693–695. doi: 10.1109/7.670364

Thomason, L. (2011). TinyXml Project Page. [Online, Accessed August, 2012]. **retrieved from** http://sourceforge.net/projects/tinyxml/

Tolman, B., Harris, R. B., Gaussiran, T., Munton, D., Little, J., Mach, R., ... Renfro, B. (2004, Sept.). The GPS Toolkit: Open Source GPS Software. In *Proceedings of the Institute of Navigation ION GPS-2004.*

U.S. Geological Survey, National Mapping Division. (1998, Jan.). *National Mapping Program Technical Instructions: Standards for Digital Elevation Models.* **retrieved from** http://nationalmap.gov/standards/demstds.html

Utah Automated Geographic Reference Center. (2012). 2 Meter LiDAR Elevation Data. [Online, Accessed March, 2012]. **retrieved from** http://gis.utah.gov/data/elevation-terrain-data/2-meter-lidar/

Wilson, T. (2008). *OGC KML Version 2.2.0.* Open Geospatial Consortium Inc.