NEUBRANDENBURG UNIVERSITY OF APPLIED SCIENCES

MASTER THESIS

Evaluation of the availability of communication channels for inland vessel navigation based on a long-term measurement campaign

---

Author:
Kateryna LUBYK

Supervisors:
Andreas WEHRENPFENNIG,
Ralf ZIEBOLD

A thesis submitted in fulfillment of the requirements for the degree of Master in Geodesy and Geoinformatics in the

Institute of Communications and Navigation
German Aerospace Center (DLR)

December 3, 2019

Declaration of Authorship

I, Kateryna LUBYK, declare that this thesis be titled, “Evaluation of the availability of communication channels for inland vessel navigation based on a long-term measurement campaign” and that the work presented in it is my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"Measure what can be measured, and make measurable what can not be measured."

Galileo

NEUBRANDENBURG UNIVERSITY OF APPLIED SCIENCES

Abstract

German Aerospace Center (DLR)

Master in Geodesy and Geoinformatics

Evaluation of the availability of communication channels for inland vessel navigation based on a long-term measurement campaign

by Kateryna LUBYK

This study investigates the capability of GSM and AIS communication channels for the provision of GNSS correction data, which are required for highly accurate positioning of inland vessels. The evaluation is based on a 1-year measurement campaign in the most important inland European waterway, the Rhine-Main-Danube Corridor. This research provides a map of the investigated internet availability in the main waterway path from Linz to Antwerp. Additionally, the capabilities of five AIS base stations broadcasting AIS binary messages in the VDES backup frequency in the test waterway area are estimated in this research.

The evaluated results indicate AIS communication channels can satisfy the established requirements. GSM would be able to satisfy the requirements but currently it is not reliably available on all parts of the inland waterways. Therefore, further research is needed to investigate VDES (new AIS generation) and further measurement campaigns need to be conducted. The analysis presented in this work is based on the manipulation big data, the integration of GIS technologies, and the development of custom algorithms.

Acknowledgements

This master thesis is based on my research work in the Institute of Communications and Navigation in the German Aerospace Center. Without sincere support from many people, it would not have been written. I would like to express my gratitude to all the people who made this thesis possible. Because of you, my research experience in Germany has become a story that I will always remember. First, I owe my far deepest gratitude to my supervisor Prof. Dr.-Ing. Andreas Wehrenpfennig, who gave me this opportunity for my master's thesis research and guided me into the research world. I would like to thank Dr. Ralf Ziebold for the steady support and guidance to complete this study. Also, I would like to thank him for the given topic, in which I have learned and developed myself as a future Geoscience specialist. Additionally, I thank my colleagues Martin Kriegel and Alvaro Caceres for helping me get into programming. I especially extend special gratitude to Martin Kriegel for helping me every day to get to DLR. All the colleagues at DLR are acknowledged for their contribution to an enjoyable working environment. Furthermore, I am very grateful to DLR to provide a scholarship for my research. Finally, I would like to express my gratitude to my family. This master thesis is dedicated to my parents, Maryna Lubyk and Volodymyr Lubyk, who supported and encourage me on my way.

Contents

List of Figures

List of Tables

List of Abbreviations

AIS	**A**utomatic **I**dentification **S**ystem
AOS	**A**utomated **O**peration **S**ystems
CCNR	**C**entral **C**ommission for the **N**avigation of the **R**hine
CSV	**C**omma **S**eparated **V**alues
CRS	**C**yclic **R**edundancy **C**heck
DGPS	**D**ifferential **G**lobal **P**ositioning **S**ystem
ETRS89	**E**uropean **T**errestrial **R**eference **S**ystem 1989
GIS	**G**eographic **I**nformation **S**ystem
GNSS	**G**lobal **N**avigation **S**atellite **S**ystem
GPX	**G**PS **E**xchange **F**ormat
GSM	**G**lobal **S**ystem for Mobile **C**ommunications
HTTP	**H**yper **T**ext **T**ransfer **P**rotocol
IMO	**I**nternational **M**aritime **O**rganization
NMEA	**N**ational **M**arine **E**lectronics **A**ssociation
NTRIP	**N**etworked **T**ransport of **R**TCM via **I**nternet **P**rotocol
PNT	**P**ositioning, **N**avigation and **T**iming
RADAR	**R**adio **D**etection and **R**anging
RINEX	**R**eceiver **I**ndependent **E**xchange **F**ormat
RKM	**R**iver **K**ilometers
RMC	**R**oot **M**ean **S**quare
RTCM	**R**adio **T**echnical **C**ommission for **M**aritime Services
RTK	**R**eal **T**ime **K**inematic
SAPOS	German **Sa**tellite **Po**sitioning **S**ervice
SPP	**S**ingle **P**oint **P**ositioning
SQL	**S**tructured **Q**uery **L**anguage
UTM	**U**niversal **T**ransverse **M**ercator coordinate system
VDES	**V**HF **D**ata **E**xchange **S**ystem
VHF	**V**ery **H**igh **F**requency
VRS	**V**irtual **R**eference **S**tation
VTS	**V**essel **T**raffic **S**ervice
WGS84	**W**orld **G**eodetic **S**ystem of 1984

Chapter 1

Introduction

Global Navigation Satellite Systems (GNSS) are the main information source of absolute Positioning, Navigation, and Timing (PNT) data for maritime and inland water navigation systems. This is supported by simultaneously adopting different processing channels to produce PNT data according to the special requirements of the various aspects of ship navigation.

The growing global transport of goods has become a vital factor increasing the size of ships. The dimensions of inland waterway vessels have some limitations, since they should be able to pass under the bridges and the through waterway locks. Another challenging task is the safety of passing the number of vessels yielding the same routes at the same time. Today, inland navigation is more challenging than maritime navigation and has higher requirements for accuracy and positioning.

According to the ”Annual Report 2018 Inland Navigation In Europe”, the number of accidents on inland waterways has been significantly reduced in the last 20 years. Specifically, the number of collisions between ships follows an obvious downward trend, which is due to the installation of electronic devices and equipment [28].

Nowadays, however, a significant number of inland accidents and casualties still exist. The rate of the accidents varies from one river to another. For instance, if you compare the Rhine to the Danube, the accident rate is lower. It depends on the quality of the navigation infrastructure and the specification of the inland waterway. According to accident statistics up to 2013, the accidents were determined as grounding, ship gets stuck, collision between ships , collision with infrastructure and bridges, pounding of waves, and other accidents. The most frequent cause of accident was the collision with infrastructure and bridges [28]. This kind causes 38-40% of all accidents. The next most frequent variety of accident was the collision between ships (18-19%) [28]. In last few years this type of accident is reducing. Nevertheless, collisions with infrastructure and bridges still shows an increasing tendency.This is why it is important to design and implement a safe navigation system to help skippers pass through difficult waterway areas.

Another aspect is the upcoming future of autonomous navigation. Due to the unpopular and complex profession of officer of the watch, it is necessary to reduce the number of people on board and try to automate transportation along inland waterways. Future technologies supporting autonomous navigation will require reliable and accurate positioning and communication between ship and shore.

This study focuses on the analysis of the high-precision phase-based positioning with a Real-Time Kinematic algorithm (RTK), using the correction data received by the GSM or AIS communication channel. The objective of this study is to evaluate GSM communication channel and the alternative AIS communication channel, investigate the current status of the availability of the communication. The data for the analysis was obtained from the one-year measurement campaign in the Rhine-Main-Danube waterway corridor.

1.1 Motivation

Inland waterway transport plays an important role for the transport of cargo in Europe. More than 37,000 kilometers of waterways connect hundreds of cities and industrial regions. 21 out of 28 Member States have inland waterways, 13 of which have networks which are interconnected. Within the data of the European Union, the share of road transport is constantly at a level of roughly 50%, whereas inland waterways represent less than 5% of the total transport amount [25]. But inland waterway transport has significant potential to increase the modal share. In comparison with other modes of transport which are often facing traffic jams and capacity problems, inland waterway transport is marked by its reliability, energy efficiency and sighificant capacity for increased exploitation. The European Commission aims to promote and strengthen the competitive position of inland waterways in the transport system and to facilitate its integration into the intermodal logistics chain. Inland waterway transport is a competitive alternative to road and rail transport. Specifically, it is an environmentally-friendly option in terms of energy consumption and noise emissions. It has the lowest amount of CO2 emissions in comparison to other modes of transport. Energy consumption per km/ton of transported goods constitutes approximately 17% of road transport and 50% of rail transport [25]. In addition, inland waterway transport provides a high degree of safety, in particular when it comes to the transportation of dangerous cargo. External costs for inland navigation such as costs deriving from climate gases, air pollutants, accidents and noise, are the lowest in contrast with other transport modes. Ultimately, it helps to deal with overloaded road networks in densely populated regions.

Presently, the development of safe and reliable inland waterway navigation is an important objective, especially considering the possible shift of transportation from the road and train transport to the inland waterway. To secure safe and effective ships' operation some modern equipment was produced, which assists the deck officers in ship control and maneuvering. The global economy develops quickly and establishes new waterway transport requirements, such as the optimization of the number of crew on board, energy savings, and cost reductions in the ship operations. In this respect, the automation of ship operations should be developed and focused on automatic guidance and advance driver assistance functions. This is a worthwhile solution for reducing the amount of crew and operating costs.

High precision positioning and reliable information about the surroundings and obstacles in the waterway are essential for these systems. These requirements can be satisfied via highly accurate and stable positioning, which depends on a reliable communication link. This study is considered a basis for the improvement and development of automatic machines and systems in inland waterway transport applications through the identification of areas with poor Internet access and analyzing the impact of obstacles on it. Moreover it can be used to develop connection availability towards the future of autonomous navigation.

1.2 Objectives

The aim of this thesis is to analyze the one year measurement campaign in the most important inland waterway route - the Rhine-Main-Danube Corridor and detect the parts of the path where there is no internet connection for autonomous navigation. This work is composed of the following tasks:

1. Extract relevant information from a real measurement database for evaluation of availability of connection.
2. Define the reference trajectory.
3. Filter the trajectory data from the disturbances data (maneuvering, mooring, change locations movements).
4. Calculate river kilometers in the waterway path from Linz to Antwerp.
5. Represent the measurement data in this waterway path and detect the position and river kilometers according to the reference trajectory .
6. Identify the vessel routes and status of the vessel in the route.
7. Evaluate the availability of the internet connection for the 2 Hz data in the waterway path.
8. Identify and explore the most frequently used the river kilometers.
9. Analyze river kilometers with a bad internet connection.
10. Analyze the influence of the connection of different environmental scenarios, and evaluate the influence of bridges and waterway locks on the availability of connection.
11. Evaluate the AIS communication link and simulate the error message rate from received RTCM-3 and AIS messages in the test bed area.

The development of this work has been financed and supported by the department of Nautical Systems, which is a part of the Institute of Communications and Navigation of the German Aerospace Center (DLR) for the period of eight months.

The algorithms have been built with Python, a modern, high-level language suitable for a wide variety of programming tasks. The Python language is an open-source language and allows for quick and easy solution development. The Python Standard Libraries make programming even more effective. These libraries make it easy to do things such as converting date and time values, manipulating strings, performing complex calculations, encoding and decoding data, file manipulation, working with databases, visualizing the data.

This work was done with the help of GIS technologies and the QGIS open-source platform.

1.3 Thesis Outline

The rest of this work is organized as follows: the next chapter provides a discussion of the current communication systems and state of the art related works and projects. In chapter 3, GNSS architecture and RTK measurement principles are briefly described. Then, chapter 4 gives an outlook of the measurement campaign and vessel along with equipment characteristics. In this chapter the database and the methodology for extracting relevant data are also described. Chapter 5 extensively explains all essential methods and tools that were used for the evaluation of the measurement campaign. Chapters 6 and 7 present the evaluation of the availability of GSM and AIS connections. These chapters, also contain the description and analysis of the influence of waterway infrastructures on the disturbance of communication and present the map of connection availability. Finally, in chapters, 8 and 9, the conclusions and future work are presented.

Chapter 2

Review of Navigation Communication Systems for Vessels

2.1 Vessel Navigation Historical Aspects

Inland water transportation is aimed at transporting goods and passengers by ship on inland waterways, both natural (rivers, lakes) and artificial (canals, reservoirs) between inland harbors or docks and wharves. The history of inland transport dates back more than one thousand years old. It is believed the first large river vessels were built in Ancient Egypt in the fourth millennium BC. The first sailors used redundant simple charts and observations of the Sun and stars for navigation. Maps, compasses, astrolabes, and calipers constituted the early instruments which were used for navigation. In the modern era, these devices have been significantly substituted by electronic and technological equivalents.

The twentieth-century technological revolution produced important advancements for marine and inland navigation, including radio beacons, radar, the GNSS compass, electronic river charts and the Global Positioning System (GPS).

Presently, positioning, electronic charts, and radar signals are overlapped when the nautical traffic situation is shown on a vessel display. In addition to this nautical information, ship parameters pertaining to specific ships are retrievable with the help of Inland AIS (Automatic Identification System) based on the other ships' radar echoes which can be accounted for during ship handling.

2.2 Automatic Identification System

AIS has been used as a communication platform to exchanges data between vessels since 2004 [16]. This system is implemented in different types of vessels. AIS information supplements marine radar, which is maintained as the first method of collision avoidance for water transportation. AIS uses marine radio frequencies for transmission between vessels instead of the mobile Internet. An AIS-like system could be established on the mobile Internet, but for navigation purposes this is currently very hazardous as the mobile Internet is not consistently available on the vessel route. Even the current AIS system is affected by safety issues caused by both technical and human errors [11].

The International Maritime Organisation (IMO) Resolution A.917(22) defines the objectives of AIS as [13]:

"AIS is intended to enhance: safety of life at sea; the safety and efficiency of navigation; and the protection of the marine environment. SOLAS regulation V/19 requires that AIS exchange data ship-to-ship and with shore-based facilities. Therefore, the purpose of AIS is to help identify vessels; assist in target tracking; simplify information exchange (e.g. reduce verbal mandatory ship reporting); and provide additional information to assist situation awareness".

Collecting sensor data from vessels has numerous purposes. Data received by various vessels allow for new applications such as in the field of navigation. Moreover, collecting data, even from an individual boat, can give data which can be used to detect different problems quickly. An example of implementing marine data is using it as a support for navigation. Every vessel collects its own GPS location, heading, and velocity data which can then be transmitted to other vessels. This data can be used by the officer of the watch to assist navigation by having more information about the situation of other vessels in the waterway. This is particularly beneficial in bad weather conditions or on tight shipping routes when a visual observation or radar monitoring might not be enough to identify other vessels.

2.2.1 AIS Work Principle

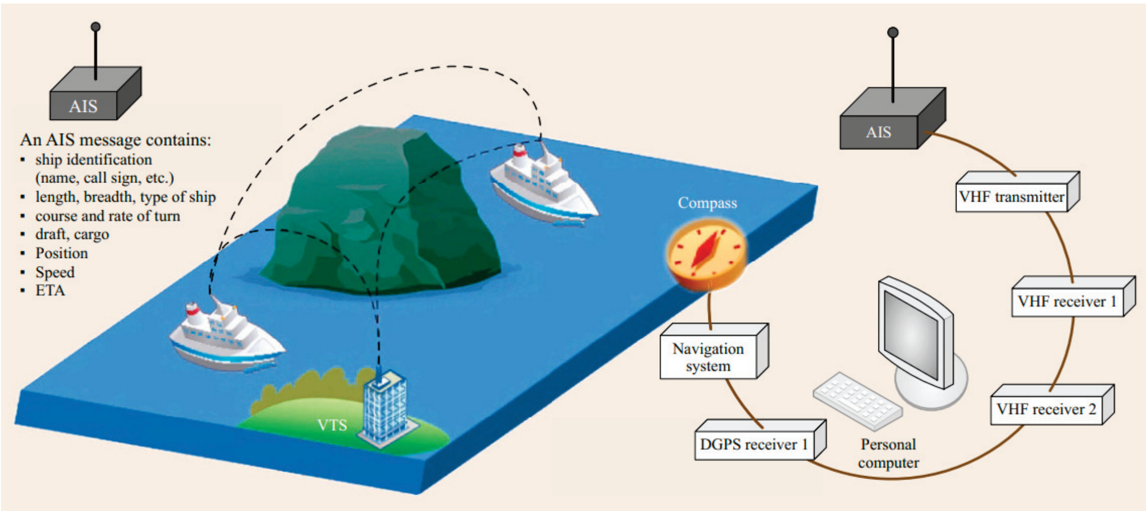


FIGURE 2.1: AIS Architecture [33]

AIS is established on ships and by vessel traffic services (VTS). This system allows for the identification and discovery of vessel locations by electronically exchanging data with other nearby ships, AIS base stations, and satellites. The approach when the detection of AIS signatures satellites is used is called Satellite-AIS (S-AIS). AIS transceivers automatically transmit data about the position, speed, and navigational status of the vessel. The data is reported from the ship's navigational sensors, such as from Global Navigation Satellite System (GNSS) receiver and from the GNSS compass. Additional information, such as the vessel's name and call sign, is also transmitted constantly and is programmed when installing the equipment.

AIS systems work on the scenarios for Self-Organizing Time Division Multiple Access (SOTDMA). This approach can be adjusted to all users and meets the likely future requirements for an effective use of the spectrum [32]. This means each vessel has its own unique time to transmit the data. The transmission is organized in slots, where each slot is the time period in which the transponder is able to transmit the data. The data is transmitted constantly with the same time intervals. These intervals can be every 2 –10 s, most quickly a vessel's speed while underway and every 3 min while vessels are at anchored.

The data is broadcast through a VHF (Very High Frequency) transmitter built into the transceiver. VHF is the range of radio frequency electromagnetic waves (radio waves) from 30 to 300 megahertz (MHz), with corresponding wavelengths of ten meters to one meter. AIS platform exchange data between ships and between ship and shore with two VHF communication channels (AIS 1: 161.975 MHz; AIS 2: 162.025 MHz) [36].

The coverage of this system depends on the height of the antenna. The propagation of the signal is better than radars because VHF has longer wavelengths. The large wavelength allows for good visibility in even the most difficult areas, such as river bends, behind islands, or in a hilly areas (see Fig. 2.1).

VTS is a traffic monitoring system established by harbor or port authorities. The delivered information on the vessel can be displayed on a screen or chart plotter showing the other vessels' positions in much the same manner as a radar display. AIS stations regularly synchronize themselves to each other, which helps avoid the overlaying of the slot transmissions [20].

2.2.2 Message Types

There exist 27 various kinds of top-level messages which can be broadcast by AIS transceivers [4]. AIS messages 6, 8, 25, and 26 supply "Application Specific Messages" (ASM), which allow "competent authorities" to define additional AIS message subtypes [32]. In this work Message type 8 will be used, which is an AIS binary broadcast message. This message will be changeable in length, depending on the quantity of binary data. The length can be shifted between 1 and 5 slots. Since the data content of this binary message is defined by the application, Message 8 is an Application Specific Message. In this work the transmitted length of the message will be 5 slots and will contain 1008 bits.

2.2.3 Message Format

An AIS slot is 26.667 ms long. This means that a slot that starts at 0.00 ends at 26.667 ms, which the beginning of the next slot [32]. The data modulation is 9600 bit/s, so each slot has a maximum capacity of 256 bits. When data is output on the VHF data link it should be grouped in bytes of 8 bits from the top to the bottom of the table associated with each message, in accordance with ISO/IEC 13239:2002 [32].

Each slot has the following structure:

```
<8 bit ramp up><24 bit preamble><8 bit start flag><168 bit payload>
<16 bit CRC><8 bit stop flag><24 bit buffer>
```

where [36],

- 24 bit preamble: this is a sequence of 0101...
- Start flag: this is ’0x7e’
- 168 bit payload, this is the body of an AIS message. For messages requiring more data, several slots (Maximum of 5) must be used.
- 16 bit CRC-16-CCITT(Cyclic redundancy check): this is a 16-bit polynomial to calculate the checksum.
- Stop flag: this is ’0x7e’
- 24 bit buffer: this is used for bit stuffing, synchronization jitter and distance delay.

2.3 Sensors for Collision Avoidance

Safe inland waterway navigation requires reliable and accurate information which should simultaneously present position and movement data of the vessel relative to the location and movement of other traffic members. This must also be done while taking into account the available space for movement in the waterway.

The navigational system on board most ships is currently based on RADAR and AIS. RADAR is an on-board device, which uses a rotating antenna to emit electromagnetic waves of the X-band (8-12 GHz) or S-band (2-4 GHz). The direction and distance of other vessels are determined by receiving and evaluating the reflected signals.

AIS exchanges including static and dynamic navigation data, as well as data related to the overall voyage. Dynamic data, in this case, is developed from built-in PNT sensors (among them GNSS, compass) and represent the information about the speed, course and etc.

AIS and RADAR systems have their strengths and weaknesses. The benefit of RADAR is the detection of other vessels is accomplished by an independent system on-board the ship. For this, the IMO declares RADAR as the main system for collision avoidance. However, RADAR also identifies more or less all objects that scatter electromagnetic waves. So in addition to other users of the traffic system it picks up everything from quaysides and wave crests to flocks of birds. On the other hand, objects located in the radio shadow of other objects cannot be identified with RADAR.

The positions provided by AIS, compared to RADAR, are presented as absolute values, which most of the time present higher accuracy than relative RADAR positions. Positions are identified using radio navigation systems, such as GPS, in which the GNSS receiver is either integrated into the AIS device or attached externally. Accordingly, the security discussion about the lack of resilience in GNSS systems is also reliant on the reliable provision of AIS data content.

The correct entry of static as well as voyage-related AIS data and the careful configuration of all navigation-related sensors which serve as AIS data sources must be handled by inland waterway technicians to ensure that AIS data content is reliable. This is why, despite the performance limitations of the RADAR, it is still the primary collision avoidance system. Thus, the current status of the vessel inland navigation is the combination of lookout, AIS and RADAR.

2.4 Autonomous navigation

Future developments in technology will transform and enchance the way in which the inland waterway sector operates. This will drive performance enhancements and create opportunities for inland waterway businesses to make better decisions. One such forward steps is autonomous shipping. Autonomy can be defined as a situation, where the vehicle can have freedom from external controls or influences. Auto-pilot is already established in most wheelhouses and the machine room is more and more automated. In the near future, inland navigation will use technologies for remote(shore) operations onboard the vessel. This is leading for applyiong fully automated operation systems (AOS). Systems are currently being investigated and developed that will implement automated functionality but with the possibility of human intervention on-board and on-shore. Companies like Massterly, Seafar, Rolls-Royce, Wilhelmsen, KONIGSBERG, . . . etc. are involved in designing the first automated cargo transportation vessel, with the goal of being available within a few years, both in maritime and inland navigation. [31][18][15].

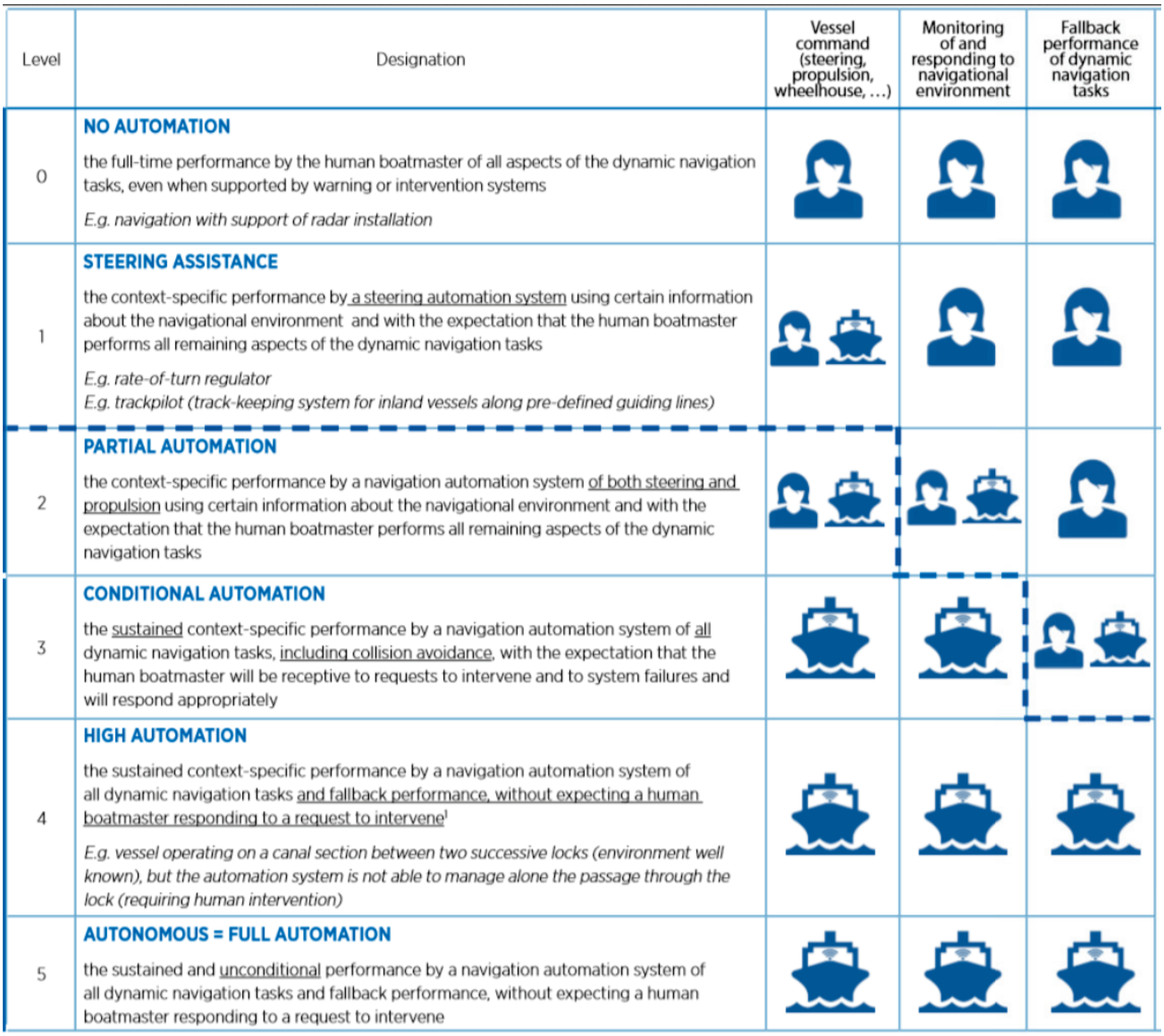
However, an autonomous ship also causes a lot of difficulties and challenges, technical and communication problems, which should be properly identified and solved. The autonomous vessels should have sufficient intelligence to make appropriate decisions in relation to possible internal and external variations. Such self-operating systems should include advanced decision-making facilities.Therefore the research centers try to develop various technologies to support the vessels future automation.

Recent technological progressions in autonomous systems, i.e. self-driving cars, robots, etc., are examples of systems with sufficient decision-making abilities. Moreover, these autonomous systems are also supported by internal and external IoT (i.e. internet of things), big data and communication infrastructure to overcome the respective challenges. Nevertheless, the success of decision-making innovations in these autonomous systems is yet to be evaluated comprehensively by the respective authorities.

The first international definition of levels of automation in inland navigation was provided by CCNR(Central Commission for Navigation on the Rhine)[26].

The figure 2.2 is describing five levels of automation, where the fifth is full automation and the first level is steering assistance. Autonomous functionality will allow the vessel to make decisions on her own, operate fully independently even without remote control, and adapt to all given situations and external variables, while upholding safety standards. Automation also combines a highly-developed form of artificial intelligence.

The shipping industry must reach several milestones to make autonomous ship navigation a reality. The state of "Remote Controlled Ship" can be an important milestone in this destination. Remote-controlled navigation facilities can always be a part of autonomous ship navigation. In general, each vessel navigation can have both autonomous and remote-controlled navigation sectors that are



Level	Designation	Vessel command (steering, propulsion, wheelhouse, ...)	Monitoring of and responding to navigational environment	Fallback performance of dynamic navigation tasks
0	**NO AUTOMATION** the full-time performance by the human boatmaster of all aspects of the dynamic navigation tasks, even when supported by warning or intervention systems *E.g. navigation with support of radar installation*			
1	**STEERING ASSISTANCE** the context-specific performance by a steering automation system using certain information about the navigational environment and with the expectation that the human boatmaster performs all remaining aspects of the dynamic navigation tasks *E.g. rate-of-turn regulator* *E.g. trackpilot (track-keeping system for inland vessels along pre-defined guiding lines)*			
2	**PARTIAL AUTOMATION** the context-specific performance by a navigation automation system of both steering and propulsion using certain information about the navigational environment and with the expectation that the human boatmaster performs all remaining aspects of the dynamic navigation tasks			
3	**CONDITIONAL AUTOMATION** the sustained context-specific performance by a navigation automation system of all dynamic navigation tasks, including collision avoidance, with the expectation that the human boatmaster will be receptive to requests to intervene and to system failures and will respond appropriately			
4	**HIGH AUTOMATION** the sustained context-specific performance by a navigation automation system of all dynamic navigation tasks and fallback performance, without expecting a human boatmaster responding to a request to intervene[1] *E.g. vessel operating on a canal section between two successive locks (environment well known), but the automation system is not able to manage alone the passage through the lock (requiring human intervention)*			
5	**AUTONOMOUS = FULL AUTOMATION** the sustained and unconditional performance by a navigation automation system of all dynamic navigation tasks and fallback performance, without expecting a human boatmaster responding to a request to intervene			

FIGURE 2.2: Levels of automation in inland navigation[26]

segmented via the operational requirements of future vessels. The required inland water way infrastructure to support both remote controlled and autonomous ship operations should be developed and is another important milestone in the autonomous destination. [22].

One project that intended to develop and test the concept of an autonomous ship was MUNIN (Maritime Unmanned Navigation through Intelligence in Networks).[23] Vessel navigation in this project is based on a combination of automated decision systems in addition to remote controlled actions via a shore-based station.

The MUNIN concept defined the following systems and entities:

- An Advanced Sensor Module that focuses on automated detection of objects and the recognition of sea phenomena. This information is gained by combining data provided by radar and AIS with daylight and infrared camera imagery.

- An Autonomous Navigation System (ANS) that enables navigation on a predefined voyage route within certain degrees of freedom and operates and makes decisions that avoid collisions and ensures the boat's stability in challenging weather conditions.

- A Shore Control Center (SCC) whose task is to permanently monitor and control the autonomously operated vessel. This is accomplished by the onboard crew of skilled nautical officers and engineers and by drones.

Due to safety concerns, especially in the initial phase, it is required that the system must intelligently operate at various autonomous levels without reducing the overall safety performance. Furthermore, if the remote control system fails, an unmanned ship needs reliable emergency methods to dock automatically and in a safe manner.

2.5 The concept of LAESSI

One of the important projects contributing to the development of inland ship navigation was LAESSI. Within the LAESSI (Control and Assistance Systems to Enhance the Safety of Navigation in Inland Waterways) project effective navigation assistance functions for inland waterway transport have been developed. The LAESSI project is aimed at making the environmentally-friendly mode of inland waterway transport safer.

As a part of the LAESSI project, a study to add detailed integrity information to the GNSS processing result was conducted. Providing highly accurate and reliable navigation data is a prerequisite for setting advanced driver assistance functions.

With this project, DLR has developed methods for the determination of highly accurate and reliable position, navigation and time information (PNT) for the ship forming an essential basis of the assistance functions. For this purpose, real-time kinematic (RTK)-based methods were developed for use under the difficult environmental conditions of inland navigation with shading and multipath propagation in the area of locks and under bridges.

LAESSI development concentrated on the development of the following four navigation assistance functionalities [5]:

- A Bridge Collision Warning system supplying a timely alert signal to the officer of the watch whenever the vessel, especially the wheelhouse or the mast, have no possibility safely pass under a bridge.

- A Mooring Assistance supplying an accurate display of the actual location. In particular, displaying highly accurate information about the distances to quay walls and other vessels..

- An Automatic Guidance System decreasing pressure on the skipper during route navigation. This is satisfied by highly accurate and tested position and heading information from the basis of this functionality, especially on narrow waterways.

- A Conning Display with clarity showing the motion of the ship. For this aim, it is also important to combine information from the propulsion systems as well as the impacts of wind and water currents.

The demand for project is evidenced by a large number of accidents caused by inland waterway vessels. The project LAESSI focuses, specifically, on the development of driver assistance functions in the areas of bridge approach warning, path guidance assistance and installation assistance, including the associated conning display.

The bridge approach warning should be able to check in sufficiently before a bridge passage to determine whether the ship and, in particular, its height-variable assemblies such as wheelhouse and Radarmast, can safely pass the next bridge. A possible warning message must be signaled several hundred meters before the bridge passage to have enough time for making a decision and successful maneuvering.

For the mooring assistant, the position orientation of the ship should be linked to the surroundings of the ship. The skipper is thus able to obtain an accurate representation of his situation. In particular, the current distances to quay walls or other ships and is thus assisted in maneuvering. With the help of the lane guidance assistant, the skipper should be relieved during the trip. Highly accurate, integrity checked position information is an important basis for this. The conning display represents the movement of the ship in a clear form. For this, it is necessary to be able to build on the position information by considering the drive equipment and the influence of the wind.

The work of the DLR is focused on the highly accurate determination of position, altitude, and pre-alignment of the ship with the help of Global Navigation Satellite Systems (GNSS). Anyone navigating with GNSS-based positioning must also be able to rely on it. Therefore, the topic of integrity was an important component of the overall system. PNT determination may not work in all situations (such as bridges). However it is crucial the skipper is reliably signaled when he can trust the measuring system. The development of integrity monitoring procedures, as well as the realization of an interface for the transfer of integrity information to the assistance functions, was an integral part of the work in the project.

In order to achieve the required accuracies, RTK methods were used for the PNT determination. For this purpose, carrier-phase-based correction data must be provided, which is derived from a network of land-based GNSS reference stations and transmitted to the barge. For transmission, a reliable data channel generally available along the inland waterways must be used. Since land stations for the Automatic Identification System (AIS) are already being operated on many federal waterways, furthering the use of this infrastructure is possible. However, the AIS communication channel does not provide enough bandwidth to transmit RTK correction data in correspondence to the standard AIS messages. For this reason in LAESSI project a spesial setup for transmittion AIS messages in backup frequencies were developed. This is initial investigations of transmission of RTK data using AIS technology (on VDES frequency channels) were carried out. The VDES is still in the process of standardization. Significant improvements can be expected with the future use of the VDES broadband channel (100kHz) and the development of appropriate coding schemes.

The measurement campaign investigated in this work was performed within the LAESSI project .The setup of the measurement system and the driver assistance function was also provided within this project (see overall system setup Fig.B.1). Also there were including the AIS/VDES setup, where in th one part of the waterway the RTK corrections where receiving from AIS stations.

2.6 Current Research Projects

Currently, several projects are underway to improve and develop inland navigation. In this chapter, we will briefly present some of them. All of these project implementations will be based on the AIS VDES technologies and require reliable communication channels.

- **AutonomSOW (2019-2033)**

The project AutonomSOW is an association of partners from industry, federal, and state science, which cooperates with the DLR. In July 2019, the project began with its first investigation. This project aims to make possible and demonstrate partially automated, automated and ultimately autonomous navigation in shipping. The systems, solutions, technologies, and regulations to be developed will actively support crew navigation, take on routine tasks, anticipate and efficiently manage challenges such as intercommunication, bottlenecks or locks, and increasing the safety of transport. The project's approach is based on a highly accurate satellite-based positioning of the ships and the detection of the environment by sensors. A communication system based on 5G and AIS VDES is possible with shore-based system partners that will enable the efficient management of waterway use according to principles yet to be determined. The aim is to test the technically necessary everyday equipment on board and integrate it into the digital management of the waterway. This also forms the basis for the basic concept for autonomous driving in inland navigation, which is to be developed. The beginning of 2020 should be started with the construction of a digital test field inland shipping for the automated and autonomous operation on the Spree-Oder-Waterway (SOW). In 2022, automated driving is placed within the scope of the project. In 2033, autonomous driving, based on the developed functions, is planned to be implemented [1].

- **NOVIMAR (2017-2021)**

The NOVIMAR (NOVel Iwt and MARitime transport concepts) project partners' goal is to develop the economic feasibility of waterborne transportation by introducing the concept of the vessel train [9]. The project's aim is optimization using existing short-sea, sea-river, and inland waterways. The NOVIMAR project runs in the period of 1st of June 2017 – 1st of June 2021. One of the challenges the developers met in this project were interlinking vessels within the train by a digital connection. Individual vessels will be able to join and leave the vessel train at places adjacent to their points of origin or destination at the seaside or inland. Partners of the NOVIMAR consortium are developing the appropriate tools for smart navigation, command, and control necessary to ensure safe operations. Certainly, for the purpose of navigating such a sophisticated system a reliable communication link and accurate positioning are needed.

In the current discussion of the project, using AIS was suggested to communicate the status and composition of the vessel train; a specific, designed AIS application message that contains information about the status of the vessel convoy; and a locking assistance system developed in the Scippper project that will interface with waterway locks.

- **SciPPPer (2018-2021)**

The abbreviation of project SciPPPer (SChleusenassistenzsystem basierend auf PPP und VDES für die BinnenschifffahRt) stands for "Inland Locking assistance navigation system based on PPP and VDES") [24].

The project started in 2018 and will be worked on until 2021. This project is a follow-up of the LAESSI project, since it left a lot of open questions which should be pursued. Within SciPPPer, the technology standard of the waterway infrastructure in Europe are developing because the improvement of inland and coastal shipping is an essential assistant measure to improve the efficiency of logistics.

Passing locks is the most common, but also the most demanding and critical in the inland navigation. The purpose of this project is to design an innovative assistance system which will automate the entering and exiting of a lock. This aim will be achieved by adopting new technologies: PPP (Precise Point Positioning) and using the new VDE (VHF Data Exchange) standard. The integration of the information and its presentation on a corresponding navigation display will be developed. Further, it will investigate the required short-range sensor technology within the framework of the project.

These projects are important to secure and expand Germany's position as a logistics hub in Europe. Consequently, the productive use of waterways and locks is an important strategy.

2.7 State of the Art

Today, precise navigation requirements have become crucial, principally for safety-critical applications such as in the aviation, maritime, railway, or automotive domains. Special requirements are established for inland waterways, but these are more challenging to develop due to the technological advancements in this branch. Inland waterways are exposed to various hazardous traffic areas such as narrow passages, bridges, waterway locks, and other vessels on the waterway. Moreover, inland navigation in a large ship reduces maneuverability. The aforementioned projects and future projects will improve the future of inland waterway navigation systems. They aim to supply driver assistance functions for inland vessel navigation, and upcoming techniques, now in trial status, combine AIS and GSM communication protocols.

Despite the fact GSM is determined to be a stable and low latency communication channel, its coverage is not sufficient in rural areas and inland waterway scenarios. However, AIS is an important information exchange system within the inland community, and its existing infrastructure can potentially cover the entire inland waterways.

The next-generation of AIS is now being developed and tested to fulfill the current need for data exchange capabilities. It is internationally called the VHF Data Exchange System (VDES) and will take into consideration the determined requirements for data exchange. Thus, the AIS radio channels (VHF Data Link VDL) will be guarded against overload as AIS populations increase. VDES will protect the original functions of AIS while supplying extra capacity for a broad range of applications in inland and maritime safety communication. VDES is designed to be a globally available digital data exchange system committed to maritime safety, security, efficiency and the protection of the environment. VDES has the ability to support inland data exchange such as River Information Services.

Most studies have only focused on the precision of the RTK measurement and the reliability of these measurements. One main issue which was not sufficiently investigated is the availability of communication channels to transmit this RTK correction. In the light of the recent need for autonomous and semi-autonomous navigation, there is a considerable concern about constant and uninterrupted communication channel. This paper focuses on analyzing the long, inland measurement campaign and the availability of GSM and AIS in the Rhine-Main-Danube waterway corridor.

Chapter 3

Satellite based Navigation

Defining a precise position has always been of considerable interest and the applications of such systems are almost unlimited.

Huge progress was made with the invention of the radio, which provided a way of determining location over large areas. For achieving good coverage over the whole world, radio transmitters were moved into space, now constituting the GNSS system.

Navigation Satellite Time and Ranging Global Positioning System (NAVSTAR GPS), or the GPS as it is more commonly called, is such a GNSS which is conducted by the United States Department of Defence. This was the first operational GNSS, but several other systems are now also implemented.

Other GNSS systems are: the Russian system Globalnaya Navigatsionnaya Sputnikovaya Sistema (GLONASS), the European Galileo system, and the Chinese Beidou Satellite Navigation System (BDS). They and others are either in use or are currently under development.

3.1 GNSS Architecture

Global Navigation Satellite System (GNSS) is a universal term designating a satellite navigation system (e.g. GPS, Glonass, Galileo and Beidou) which provides continuous positioning around the world [30]. Principally, GNSS system consists of three main segments: the space segment, which includes the satellites; the control segment (also referred to as the ground segment), which is responsible for the appropriate operation of the system; and the user segment, which includes the GNSS receivers, and provides positioning, velocity, and precise timing to users (see Fig.3.1).

The principle of satellite-based positioning is based on the time-of-arrival of the signal received from the satellites.

A GNSS position solution is defined by passive ranging in three dimensions. The time of arrival signal, $t_(sa)$, is defined from the receiver clock, while the time of transmission, $t_(st)$, of each signal is received from its ranging code and data message. The range, ρ, from a satellite to the user, measured by GNSS user equipment, is gained by differencing the times of arrival and transmission and then multiplying by the speed of light, c. Accordingly,

$$\rho_j = (t_(sa), j - t_(st), j)c \tag{3.1}$$

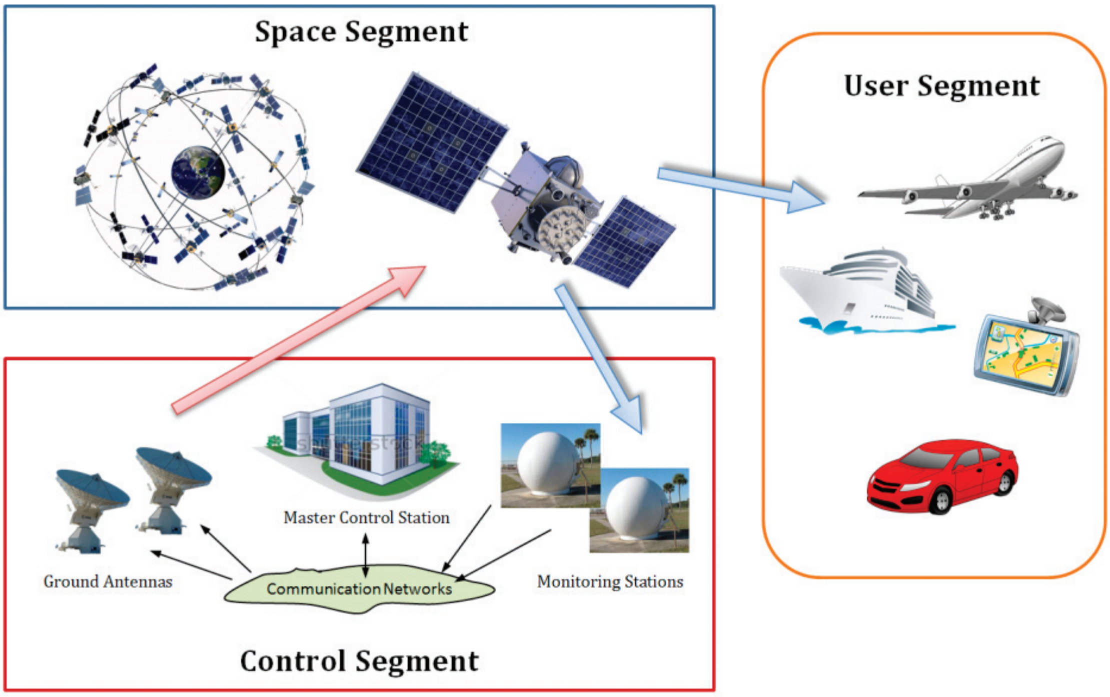


FIGURE 3.1: GNSS architecture [30]

where the index j is used to indicate the satellite number or receiver tracking channel and error sources have been neglected [10].

Because the position of the satellite constitutes part of the navigation message, finding the solution of the receiver positioning problem just consists of solving a geometric problem. It is necessary that the receiver is in range of at least four satellite signals, as the clock offset of the receiver must be evaluated synchronically with the position. The GNSS system ensures a basic radial positioning with meter accuracy in the horizontal plane and in the vertical axis. However, this greatly depends on the type of service, the quality of the receiver, and the geometry of the satellites with regard to the target. Furthermore, differential techniques can provide higher accuracy than a meter by making use of base stations at known locations to remove some of the errors. For example, carrier-phase positioning techniques can supply centimeter precision for real-time navigation and millimeter accuracy for surveying and geodetic applications [10].

3.1.1 GNSS Segments

Space Segment

The main purposes of the space segment are generation and broadcasting code and carrier phase signals and collecting and transmitting the navigation message uploaded by the control segment. The GNSS space segment is composed of satellite constellations, which have enough satellites to ensure visibility of at least four satellites at the same time from any point of the surface of the Earth. The transmission of the signal is regulated by highly stable atomic clocks onboard the satellites [30].

Control Segment

The control segment is accountable for the appropriate operation of the GNSS, and has the next basic functions [30]:

- to control and maintain the status and configuration of the satellite constellation;
- to predict ephemeris and satellite clock evolution;
- to keep the corresponding GNSS time scale (through atomic clocks);
- to update the navigation messages for all the satellites.

User Segment

The user segment consists of GNSS receivers, which have the following functions: receiving GNSS signals, determining pseudoranges (and other observables), and solving the navigation equations necessary to provide the coordinates and a very precise time. The standard GNSS receiver has the following elements: an antenna with preamplifier, a radio frequency section, a microprocessor, an intermediate-precision oscillator, a feeding source, some memory for data storage and an interface with the user. The antenna transforms the incoming GNSS radio signals to electrical signals, which are the inputs for the receiver. Next, the ranging processor uses acquisition and tracking algorithms to define the range from the antenna to each of the satellites used from the receiver outputs. It also controls the receiver and decodes the navigation messages. Lastly, the navigation processor uses the ranging measurements to compute a position, velocity, and time (PVT) solution. For additional information on the user's segment, refer to [10].

3.2 Error Sources

Satellite-based navigation is affected by several types of random errors and also affected by the influence on the passing signal through the atmosphere. Due to these errors, signals will deviate or the speed of the signal will be reduced while traveling from the satellite to the GPS receiver. According to the source of errors in satellite-based navigation can be distinguished in three main categories:

- satellite-based errors
- receiver-based errors
- signal propagation errors

3.2.1 Ionosphere Errors

Ionosphere delay occurs in the altitude from 50 km to 1000 km or even more. The speed of the GNSS signals which propagates in the ionosphere can be modified by the presence of an electrical charge in this layer of atmosphere [30]. Gas ionizes when the ultraviolet and the X-ray radiation interact with the atoms and the gas molecules.[8] The influence of the ionosphere varies with solar activity and the geomagnetic field. For GPS frequencies the resulting range error can vary from less than 1 m to more than 100 m. Carrier frequencies will be reflected by the passing the ionosphere when the signal is below 30 MHz. Just the higher frequencies can permeate through the ionosphere layer. This dependence on the signal frequency allows removing the first-order effects using two-frequency measurements. In this work two-frequency GNSS receivers were used, which eliminates the influence of the ionosphere.

3.2.2 Troposphere Errors

The troposphere is one of the atmospheric layers that is on the altitude of about 20 km above Earth's surface. GNSS signals are affected by the troposphere and have an extra delay in the measurement of the signal when passing through this layer from satellite to receiver. This layer is composed of dry gases and water vapor, occasionally causing the GNSS signals to refract and delay.

This delay is frequency-independent. However, it depends on the temperature, pressure, and humidity and on the transmitter and receiver antenna locations. Therefore, it is hard to model troposphere errors accurately using global general models [30].

3.2.3 Multipath Errors

The multipath error appears when a radio signal reaches the receiving antenna by two or more paths. The main cause is the location of the antenna close to the reflecting structures. For land scenarios, signals are generally reflected by the ground, buildings, or trees while, for ships, reflections off the host-vehicle body are more common.

Interference can also happen from diffracted signals. The reflected and diffracted signals are typically delayed with regard to the direct signals and have a lower amplitude unless the direct signals are attenuated (e.g., by a building or foliage). Low-elevation-angle signals are usually subject to the greatest multipath interference.

Minimization of the multipath error can be done by improving antenna directivity. For example, by attenuating the signal coming from certain low-elevation directions and removing the antenna away from reflecting objects [30].

3.3 GNSS Observation Techniques

In general, GNSS observables are pseudoranges obtained by code or phase measurements. These two approaches have different accuracy: code ranges have meter level precision, while the accuracy of the carrier phase is in the millimeter range.

The GNSS positioning can be performed by techniques such as Single Point Positioning (SPP), Differential Positioning with Corrections (DGPS), and Relative Positioning (Static, Rapid Static, Stop and Go, Kinematic and Real-Time Kinematic

(RTK)). GNSS SPP uses one GNSS receiver, while DGPS and relative GNSS positioning use two or more GNSS receivers, concurrently following the same satellites. Marine and inland navigation commonly use for navigation purposes the relative and differential positioning techniques. However, these techniques depend on corrections from the reference receiver.[8]

In this work RTK measurement technique is used since it can give centimeter-level accurate positioning.

3.3.1 RTK

Today, RTK-GNSS technique has been well established for many survey applications because there is no post-processing of the carrier phase data, and it has high accuracy and fast initialization. This approach has been used since the early 1990s.

The standard scenario for measurements in RTK technique requires two receivers: reference and 'roving' receiver and the available communication link to transfer the data. After setting up, the carrier phase measurements will continuously broadcast from the reference receiver to the roving receiver.

RTK approach allows reaching accuracies of up to centimeter-level almost immediately but for this a dense network of reference stations and a sufficient communication channel between the moving and the static receivers is needed.

The real-time positioning (RTK) approach allows to be fast and precise. However, this approach requires a service (in this work the Trimble Positioning Service was used) and it should be taken into account the most significant aspect to providing it is the availability of the communication link[3].

When the communication link is not available and the correction messages can not be transmitted from the reference station to the receiver, it will be automatically switched to the Single Point Positioning and calculate the position according to the satellite scenario. The accuracy of positioning with SPP can be in the meter range while, RTK is in the centimeter range.

3.3.2 Basic Principle of RTK

The Real-time kinematic (RTK) positioning method is based on the transmission of carrier-phase and pseudorange data over a radio link from GNSS reference station to a roving station. The main advantage of this method is that, with this technique the coordinates are determined in real-time.

The reference station (or base station) is a station where a GNSS receiver at a known location is established. This location of the reference station is pre-surveyed by either traditional methods or by GNSS observation for multiple days. Then the reference station can provide error data for every observation compared to its known position.

The signal's range can be affected by any obstructions along the propagation path. The ability to successfully use the received signal depends on various factors including the sensitivity of the receiver.

The data transmission can also be via the Internet using, e.g., the Network Transportation of RTCM Internet Protocol (NTRIP; [1.24]) and obtained by a hardwire or wireless link such as a mobile phone.[33]

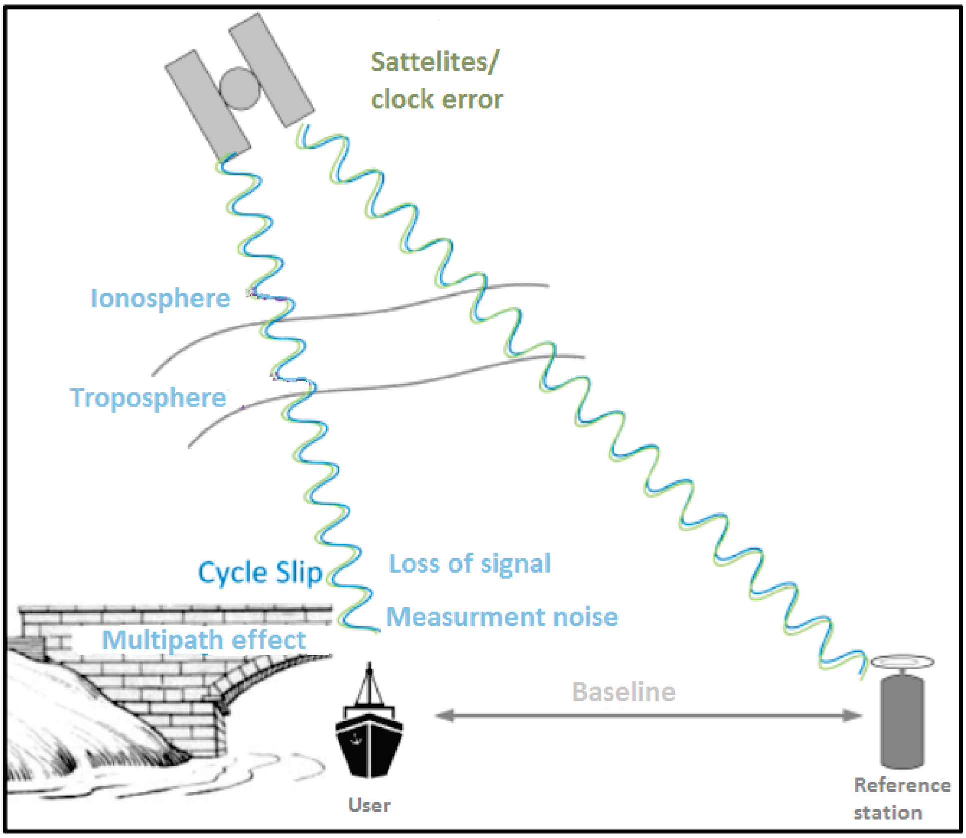


FIGURE 3.2: Principle of RTK

The working principle of RTK is a single-differencing GNSS observations between rover and base stations (the subscripts R and B will be used respectively) which helps to remove the satellite clock error and significantly diminish ionospheric and tropospheric influence on the signal propagation; then, a pivot satellite (referred to using the superscript r) which is determined for double-differencing, removing the clock offsets from the rover and base receivers.

According to the [33], the code and carrier phase observations from the i^th satellite observed at the tracked vehicle are represented as:

$$\rho_R^i = ||p_i - p_R|| + I^i + T^i + c(dt_R - dt^i) + \varepsilon_R^i \tag{3.2}$$

$$\Phi_R^i = ||p_i - p_R|| + I^i + T^i + c(dt_R - dt^i) + \lambda N_R^i + \varepsilon_R^i \tag{3.3}$$

where:

ρ_R^i is the code observation [m],

Φ_R^i is the phase observation [m],

p_i is the position of the ith satellite,

p_R is the position of the rover,

I^i is the ionospheric error [m],

T^i is the tropospheric error [m],

c is the speed of the light [299 792 458 m/s],

dt^R is the receiver clock offset at reception time [s],

dt^i is the satellite clock offset at transit time [s],

λ is the satellite clock offset at transit time [s],

N^i is unknown number of cycles between the receiver and the satellite,

ε^i, ε^i are the remaining errors for the code and phase observations respectively [m],

When applying double-differences, the geometric distance between the satellite and rover is usually represented by the following linearised form:

$$||p_i - p_R|| = x^(b^l)p_R \tag{3.4}$$

where u^i is the line-of-sight unit vector from the receiver to the i^th satellite position.

After solving the ambiguities correctly, a final adjustment solution is performed to obtain the rover coordinates at cm accuracy level [14].

In the RTK technique, correction messages are transmitted to rovers in a certain format that every receiver manufacturer produces. For the prevention of the confusion different data formats can cause, the Radio Technical Commission for Maritime Services RTCM, Special Committee 104 (RTCM SC-104), has declared a standard format for transmitting correction messages between reference and rover stations and this has been named RTCM SC-104 [29].

Positioning accuracy increases the more reference and rover stations track satellites there are and, the faster the ambiguity resolution gets. The addition of GLONASS satellite signals and GPS signals can be used in this case.

The advantages of using RTK GNSS method are as following:

- the post-processing of RTK data is not required;
- the coordinates of measured points in the field can be transformed to local coordinate systems in real-time, with the requirement of a few points (at least three) whose coordinates are also known in the local coordinate system being available;
- the method supplies a reliable tool for positioning all points precisely. In standard kinematic surveys and in the case of (undetected) cycle slips or loss of lock in the reference station, kinematic positioning cannot be performed. In RTK, this is easily identified, and the survey continues with a new integer ambiguity introduced and fixed in real-time;
- with RTK it can be accurately surveyed (on a cm level) in the field in real-time.

Today, RTK GPS is broadly used all over the world in different fields and for different purposes. Nevertheless, the RTK method has some disadvantages which can be summarized as follows[17]:

- the radio distance from the reference to the rover receiver is limited. The practical distance is approximately 5-km from the radio transmitter;
- radio obstruction by topographical constraints like mountains or man-made obstacles like high-voltage towers can occurs, which lead to failure in obtaining the correction message in the rover receivers;
- this system fully depends on the reference receiver, and any failure in this will lead to a failure in the RTK technique which is based on the coordinates of the reference receiver only;
- corrections are accurately figuring the GPS errors in the territory observations ;
- accuracy degradation of the RTK occurs when the distance between reference and rover receivers increases or when the age of corrections increases.

3.4 Virtual Reference Station

To improve the precision of GPS positioning a few distinctive strategies exist. They are all based on data sent between the measurement equipment and other devices, such as permanent reference stations. These permanent reference stations are stations where a GNSS receiver is installed and users have the opportunity to use them as their base in relative positioning. Messages are sent from a reference station to a moving rover are called corrections. The commonly used format for corrections is RTCM. This format is a recommended standard developed by the Radio Technical Commission for Maritime Services. With a baseline processing, the effects of orbit errors, ionospheric and tropospheric refraction are decreased by forming differences of the observables, e.g., double differences. These effects increase with a growing baseline length. Consequently, it is essential to use short baselines between reference stations and the rover. Thus reference station networks like the German satellite positioning service (SAPOS) and several others have been designed to meet these requirements.

Nevertheless, the existing reference station network demands further decrease of the baseline length for differential positioning techniques. The concept is to generate "observation data" for a non-existing station from real observations of a multiple reference station network and to broadcast these data to the rover station. This is the basic principle of the virtual reference station (VRS) concept[12].

A VRS is an unoccupied reference station located in the vicinity of the meter from the RTK client. For this position, observation information is produced from the information of surrounding reference stations as if they had been used on that position by a GPS receiver. The principle is that the correction data for the rovers is obtained by interpolating the data of several reference stations, usually the data of three or more reference stations.

The VRS positioning service access proceeds as follows [27] :

- The data from the Reference Stations is transmitted to the processing and control center at a one-second interval.
- The RTK user transmits its rough position from a standalone GPS measurement in an NMEA GGA message to the control center
- The processing center determines in which triangle the user is stationed and computes a virtual reference station (as a function of the transmitted approximate position).
- The processing center transmits the data of the virtual reference station to the user in RTCM format.

This concept reduces the mistakes caused by the baseline distance between the rover and the reference station. Thus, the reliability of the positioning is heightened. The advantage of using VRS is, for example in case the reference station fails temporarily. The correction data will be computed with the surrounding reference stations. In addition, productivity is moved forward by clearly shorter initialization times. The precision of RTK positioning can be reached in centimeter-level by using VRS. However, this technique has some weaknesses. First of all, the VRS service dependences on a communication system such as the mobile phone network. Furthermore, VRS requires a bi-directional communication link between the receiver and the computation center because the rover has to send information about its current position and has to receive the VRS data. This telecommunication link must provide high bandwidth communication between all the elements of the system: the reference stations, the master control center, and the user receiver. The second challenge is errors can be produced by different tropospheric and stratospheric models applied between the computation center and the rover.

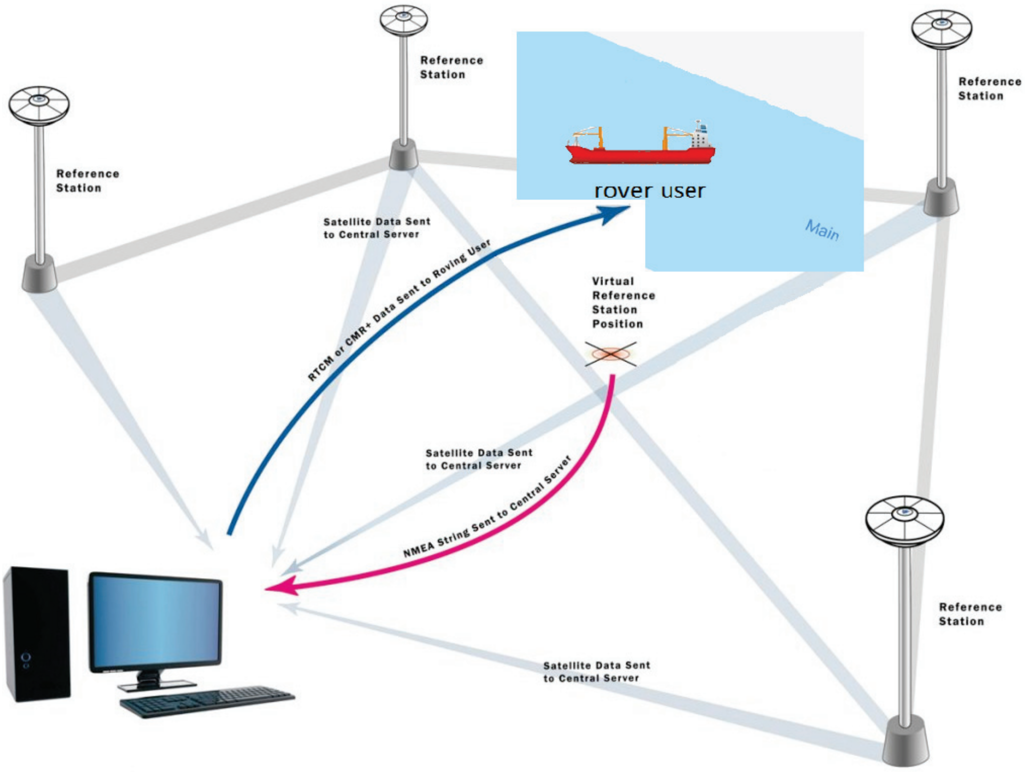


FIGURE 3.3: VRS communication [27]

Today, Network RTK is a broadly admitted technique, and, therefore, a growing number of organizations are establishing networks providing real-time corrections from a reference station network. In this work, Trimble VRS positioning service was used.

3.5 Coordinate Frames and Transformations

The science of navigation represents the position, orientation, and motion of objects. In this research work, an object is a vehicle with navigation equipment, such as a GNSS antenna. To specify the position and linear motion of an object we need to define an appropriate coordinate system. This system depends on different references such as the center of mass of that object, the geometrical center or an arbitrarily convenient point such as a corner and etc. In the case of radio positioning equipment, the phase center of the antenna is a suitable origin as this is the point at which the radio signals appear to arrive. The orientation and angular motion of an object should be described with a set of three axes. These axes need to be non-coplanar and also should be mutually perpendicular. The choice of the axis should fit the normal direction of motion for the object, the vertical direction when the object is at rest. For precise navigation, the relationship between the different coordinate frames must be properly defined. This section describes the coordinate systems used in this work.

In this work, the measurement campaign data followed the ETRS89 coordinate system. We are assuming for performing our further calculations, we can use WGS84 as a spatial reference system. As in table 3.1 the main parameters that are used for the ellipsoids are similar. And since we need to make a transformation between geographic coordinate system(longitude, latitude) to Cartesian coordinate system(x,y). The transformation between these systems was provided by Python library utm, which allows transforming the coordinates from WGS-84 to UTM and from UTM to WGS-84.

3.5.1 ETRS89

The European Datum ETRS89 is defined throughout Europe and is the nationwide uniform reference system for geospatial information of the Land Surveying and Real Estate Cadastre. It was derived directly from a global system (ITRF89)[19]. This coordinate system is an ECEF (Earth-Centered, Earth-Fixed) geodetic Cartesian reference frame, in which the Eurasian Plate represented is static. Coordinates in ETRS89 are based on the GRS80 ellipsoid. Coordinates and maps in Europe based on ETRS89 are not subject to change due to continental drift. The development of ETRS89 is based on the global ITRS geodetic datum, in which the presentation of the continental drift is adjusted so that the total apparent angular momentum of continental plates is near 0.

3.5.2 WGS84

The World Geodetic System 1984 (WGS84) is a three-dimensional Cartesian coordinate system with an earth fixed global reference frame. [12] This coordinate system is a standard for applications in cartography, geodesy, and satellite navigation, including GPS. A set of primary and secondary parameters describe an earth-model. The primary parameters determine the form of an earth ellipsoid, its angular velocity, and the mass of the earth. In this coordinate system, the GRS80 (Geodetic Reference System 1980) reference ellipsoid was originally used, but has sustained some minor refinements in later editions since its initial publication. Most of these refinements are essential for high-precision orbital calculations for satellites. However, they have little practical impact on typical topographical uses. The table 3.1 lists the main ellipsoid parameters.

A detailed gravity model for the earth is determined by the secondary parameters. These extra parameters are needed because WGS84 is used for determining the orbits of GPS satellites. The ellipsoid used in this coordinate system is composed to fit the geoid of the entire earth. Hence it usually does not fit the geoid in a specific country as well as the non-geocentric ellipsoid used for mapping the

Parameter	WGS84	GRS80
Semi-major axis of the ellipsoid (a)	6 378 137.0000 m	6 378 137.0000 m
Semi-major axis of the spheroid (b)	6 356 752.3142 m	6 356 752.3141 m
Flatting of the spheroid (f)	1/298.257223563	1/298.257222101

TABLE 3.1: Reference Ellipsoid Parameters

country. WGS84 is considered the most commonly used datum format with GPS receivers nowadays. The NMEA standard implements coordinates in this format.

3.5.3 UTM

The Universal Transverse Mercator (UTM) is a system for specifying coordinates to locations on the surface for the Earth. In this coordinate system, the earth is considered a perfect ellipsoid. The position is represented horizontally, which means it neglects altitude. Nevertheless, it varies from global latitude/longitude as this approach "cuts" the earth into 60 zones. Each zone projects on the plane as the basis for its coordinates. Defining a location indicates the zone and the x, y coordinate in that plane. Zones in UTM extent are frequently 6 degrees of longitude, and each has an appointed central meridian. The scale factor at the central meridian is defined to be 0.9996 of true scale for most UTM systems in use [12].

3.6 Data Formats

Various formats for the exchange of satellite navigation data are defined. The most widespread are RTCM, RINEX, and NMEA. They are internationally accepted and generally supported by all receiver manufacturers. The data format defined by the Radio Technical Commission for Maritime Services (RTCM) is adopted for the real-time broadcast of measurements and differential corrections. The receiver-independent exchange (RINEX) format is used for the transfer of raw data, especially for postprocessing applications. The National Marine Electronics Association (NMEA) determined an ASCII data format which is used for the transmission of position solutions.

3.6.1 RTCM Format

In spite of the fact that a few receiver producers have designed their own proprietary formats, the transmission of correction information between a reference and a remote receiver has been standardized in 1985 corresponding to the suggestions of the US Radio Specialized Commission for Oceanic Administrations, Uncommon Committee 104. The current version 3.1 is denoted as RTCM standard 10403.1.

Though the RTCM format was initially presented for differential operations in maritime applications, it is now utilized in all fields of applications for the broadcasting of all kinds of GNSS information.

A third version of RTCM has been established to enhance the capability of data broadcasting and to improve the integrity of the parity operation. The third version has been specially produced for large data transmissions and is adopted to RTK and network RTK operations. The message has the following structure: 8-bit preamble, 10- bit message length identifier, and 6 extra bits in the header reserved for future use. The data field has a maximum length of 1024 bytes followed by a 24-bit cyclic redundancy check (CRC) [12]. For the broadcasting of RTCM data over the Internet the networked transport of RTCM via Internet Protocol (NTRIP) format is used. This format was defined by the German Federal Agency for Cartography and Geodesy. NTRIP is based on the hypertext transfer protocol (HTTP).

3.6.2 NMEA Format

In 1983, US National Marine Electronics Association (NMEA) declared the NMEA-0183 interface specification. The NMEA 0183 standard represents an electrical interface and data protocol for communications within marine instrumentation. However, it became a voluntary industry standard interface for all GNSS receivers.

The concept of this standard is data exchange between talkers and listeners. The principal is one talker sends data to one or more listeners through one-way serial communication channels. The NMEA data format is used to exchange information about the position including quality indicators, course over ground or speed over ground data.

The NMEA 0183 standard uses a simple ASCII - encoding standard for electronic communication [7]. Data is broadcasted from a talker to a listener by sentences, and the maximum amount of characters which can be in one sentence is 82. These sentences have a common structure. Each sentence starts with the delimiter «$» (dollar sign) and is ended with a <CR><LF> (carriage return, line feed) delimiter. A sentence consists of a number of fields containing an address field identifying the talker and the sentence formatter, zero or more data fields, and an optional (for most offers) checksum field to detect data transmission errors. Each field is followed by a comma deliminator, except for the address and checksum fields.

There are about 60 different approved sentence types. The beginning of the sentence shows GPS data with a "GP" indicator and for GLONASS information, the "GL" sign. The last data field is marked by an asterisk "*" and an optional checksum. The checksum is not mandatory for all message types. The checksum is given in hexadecimal format. Finally, it is followed by a carriage return and line-feed character [12].

The NMEA messages and their content used in this work are defined in table 3.2. The figure 3.4 shown the structure of the NMEA GGA sentence.

GNNS related NMEA sentances	
Sentence	Description of content
GGA	Global Positioning System Fix Data (time, ellipsoidal coordinates, number of satellites, dilution of precision, quality indicator, geoidal height, age of differential GPS data)
HDT	Heading true (heading true in degrees)
VTG	Navigation data (course over ground, speed over ground)

TABLE 3.2: NMEA sentences

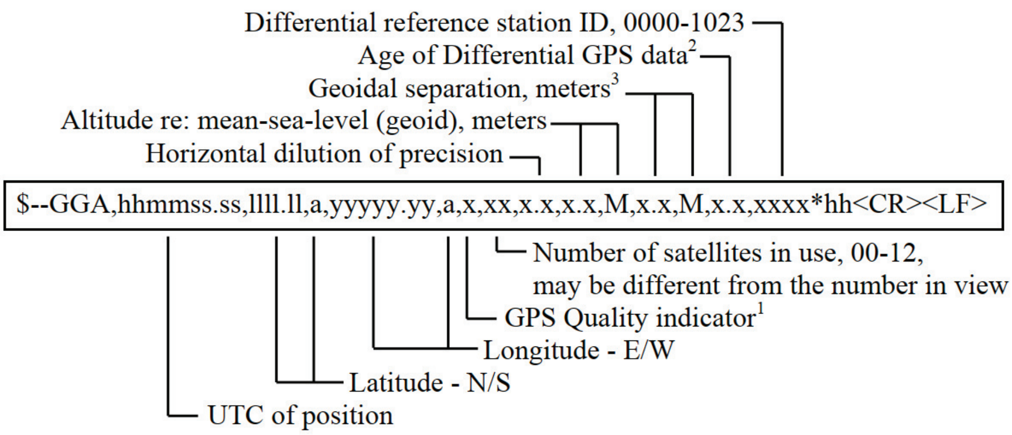


FIGURE 3.4: The structure of the NMEA GGA sentence

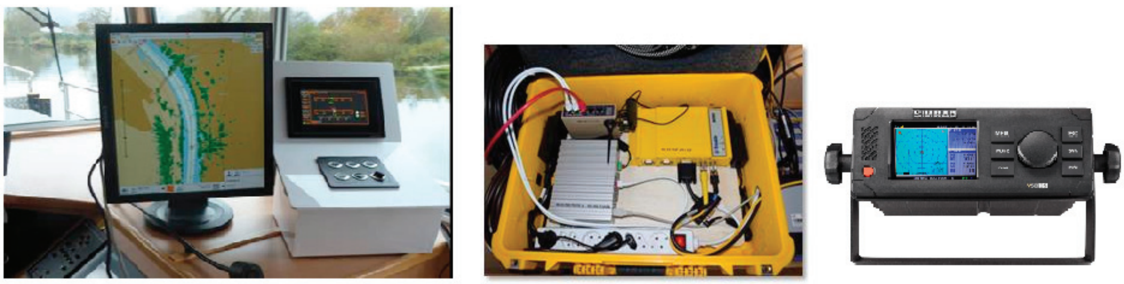
Chapter 4

Experimental Setup

This study was supported by the Department of Nautical Systems, which is part of the Institute of Communication and Navigation of the German Aerospace Center (DLR) [6]. The data for the evaluation was provided by DLR. In this chapter the measurement campaign and equipment used for this work will be described. Additionally the overall system measurement setup will be described. This chapter contains the description of the database structure and the method used for data visualization as well as the extraction of the appropriate data from the database.

4.1 Measurement Campaign

The measurement campaign under investigation was provided by a vessel in the LAESSI project. The data sources on board the EL-NINO provide their data in the form of NMEA data sets with update rates of 2 Hz, in accordance with the standard. Data was stored in a SQLite3 database. The duration of the measurement campaign was 1 year. The vessel was sailing in the inland waterway pathway of the Rhine-Main-Danube Corridor. It is an international waterway between the North Sea and the Black Sea. A general overview of the sailing trajectory by the vessel for the 1-year time frame is shown in figure4.2.



(a) Automatic guidance system (b) Trimble Receiver + BoxPC + switch (c) AIS / VDE Transceiver

FIGURE 4.1: Main components of the evaluated measurement campaign setup

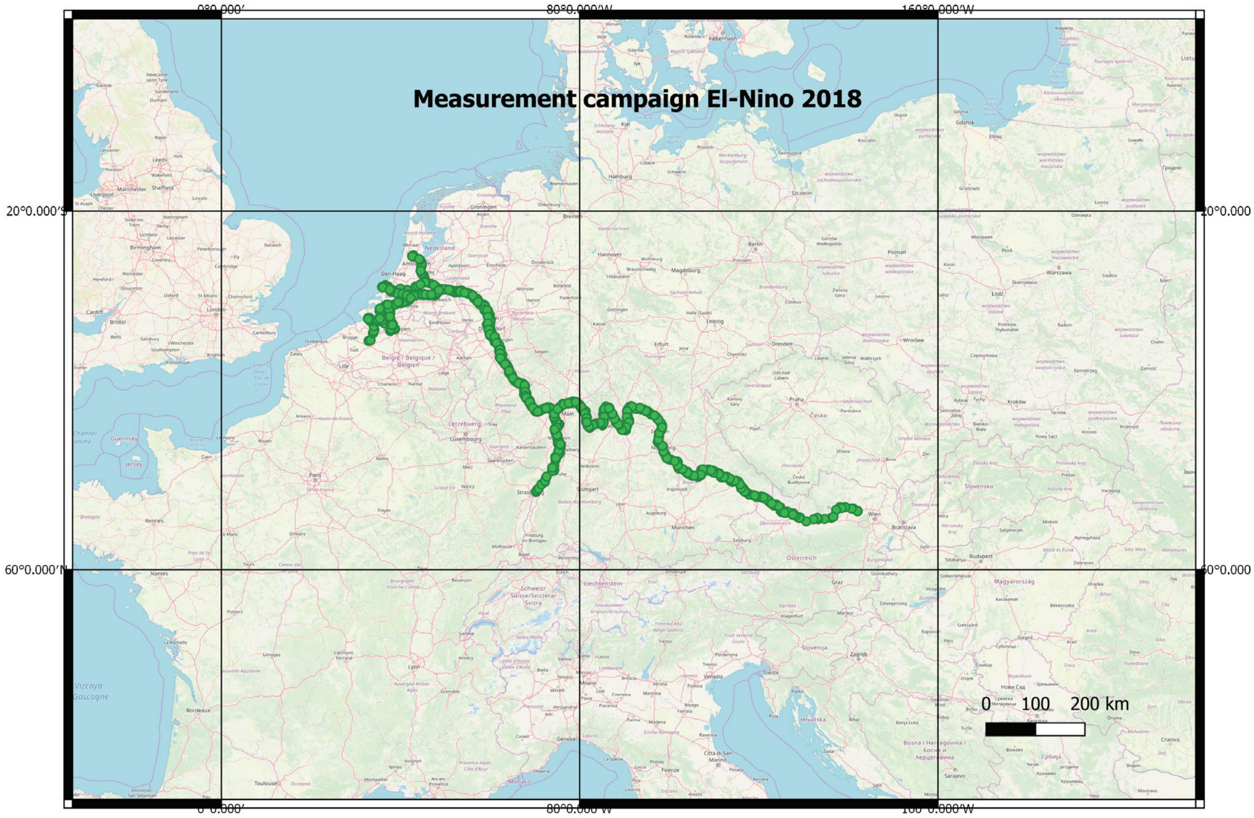


FIGURE 4.2: Measurement Campaign 2018

4.2 The Vessel and Equipment

In the context of this thesis, data was collected while the vessel travelled in the Rein-Main-Danube corridor.

For analysis, a cargo ship was used, having the characteristics represented in table 4.1.

Vessel EL-NINO	
Type	Cargo ship
Year of construction	1956
Length	185 m
Width	10 m
Dead Weight	4,322 t
Average Speed	10.5 km/h
Maximum Speed	20 km/h

TABLE 4.1: Vessel EL-NINO

The EL-NINO boat contains a NMEA 0183 network for communication between transducers and display instruments.

FIGURE 4.3: Vessel EL-NINO and the two established on-board GNSS antennas

Figure B.1 shows a schematic overview for the system setup of the measurement campaign. The main equipment components used in setting up the evaluation campaign are shown in the figure 4.1. For the evaluation of this work, the following were used: a Trimble GNSS receiver, Switch, PC, and Nauticast A2. Also, two GNSS antennas were established on the wheelhouse of the vessel (see Fig. 4.3). The overall measurement setup, for testing LAESSI assistance functionalities includes:

- Trimble BX982 GNSS receiver - receives RTK corrections from AIS/VDES or Mobile phone (Backup) and provides a highly accurate position solution;
- TP-Link TL-SG1008D (Gigabit Switch) - a device in a computer network that connects other devices together. Multiple data cables are plugged into a switch to enable communication between different networked devices;
- Alberding PC NISE105-E3845 (PC) - stores measurement information;
- Nauticast A2 - reception of AIS messages - VDE Channel;
- Argonics Trackpilot - automatically follows a predefined route;
- Radarpilot - designates the solution – a radar system including direct control of a track-keeping autopilot;
- Autopilot - allows the recording of specific turn rate data;
- LAESSI Assistance - Control and Assistance Systems to Enhance the Safety of Navigation in Inland Waterways including: the bridge collision warning system alerts, the mooring assistant, the automatic track control, and the conning indicator.

4.3 GNSS Receiver

During the measurement campaign, the Trimble BX982 GNSS receiver was used, which is commonly used for marine applications which require centimeter-level accuracy and heading. It is a multi-channel, multi-frequency GNSS receiver which allows OEM's and System Integrators to rapidly integrate centimeter-level positioning and precise heading into their application. The Trimble BX982 supports GPS L1/L2/L5, GLONASS L1/L2/L3, and BeiDou B1, B2 signals and Galileo signals. Code differential GPS positioning accuracy 3D: typically < 1 m. Accuracy and reliability may be subject to anomalies such as multipath, obstructions,satellite geometry, and atmospheric conditions. RTK positioning accuracy Horizontal: ±(8 mm + 1 ppm) RMS and Vertical: ±(15 mm + 1 ppm) RMS. Initialization time is less than 10 seconds [34].

4.4 Database

For this work a 1-year data set which was stored in a SQLite database was evaluated. SQLite is a relational database management system (RDBMS). In comparison to many other database management systems, SQLite is not a client-server database engine. Rather, it is embedded in the end program such as a receiver. SQLite performs most of the SQL standard, generally following PostgreSQL syntax. SQLite is a popular choice as an embedded database software for local/client storage in application software. In SQLite, there are a large number of bindings to other programming languages, including Python.

The database schema is represented in figure 4.4.

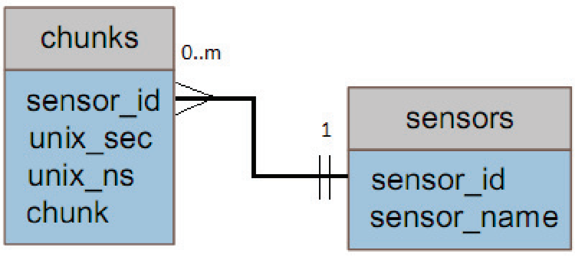


FIGURE 4.4: SQLite3 database schema

The table sensors contain the id and name of each sensor that was used in the measurement campaign (see Fig.4.5).

The table chunks represent information for each sensor according to time relevant message (see Fig.4.6).

	sensor id	sensor name
1	1	MSEN_ALB_Trimble2_RTCM3
2	2	MSEN_ALB_Trimble1_RTCM3
3	3	Trimble_NMEA2Hz
4	4	AIS_RTCM3
5	5	__dataCollector__
6	6	Backup_RTCM3
7	7	AIS_GSM_RTCM3

FIGURE 4.5: Table sensors in SQLite3 database

	sensor_id	unix_sec	unix_ns	chunk
1	3	1546732800	381277269	$GNGGA,000001.00,4947.89087671,N,00909.54433894,E,5,12,
2	4	1546732800	480451294	$GPVTG,,T,,M,0.099,N,0.184,K,A*2E
3	4	1546732800	520460688	$GPGGA,000001.00,4947.89026,N,00909.54516,E,1,08,1.19,12
4	4	1546732800	520721959	$GPGSA,A,3,27,29,26,21,16,20,10,31,,,,,1.88,1.19,1.46*06
5	4	1546732800	540418048	$GPZDA,000001.00,06,01,2019,00,00*6A
6	4	1546732800	540488673	$GPGBS,000001.00,1.0,0.7,1.3,,,,*44

FIGURE 4.6: Table chunks in SQLite3 database

For this work, the relevant sensor was the Trimble_NMEA2Hz. This sensor contains the NMEA messages sent out by the GNSS receiver.

For the selection of sensor data content, the following SQL query was used:

```
"SELECT chunk FROM chunks,sensors WHERE
chunks.sensor_id=sensors.sensor_id  AND
sensors.sensor_name=\"Trimble_NMEA2Hz\""
```

4.5 Data Extraction and Visualization

One of the important issues of this work is the handling of big data. For this purpose, SQL query relevant data were stored in CSV(Comma Separated Values) format since they are easy to read and post-process with Python, this format is the most common import and export format for spreadsheets and databases. Python enables easy-to-read CSV files by using the pandas library's DataFrame object. A DataFrame is two-dimensional, size-mutable, and potentially heterogeneous tabular data structure with labeled axes (rows and columns).

Another format used for data visualization is the GPX format. GPX, or GPS Exchange Format, is an XML schema designed as a standard GPS data format for the interchange of data between applications and Web services on the Internet. It can be used to describe waypoints, tracks, and routes. The format is open and can be used for free.

For extracting data from the database, sqlite3 and pynmea2 python libraries were used, which allow for the connection to the database, the identification of NMEA messages, and the parsing of them.

To enable us to do the evaluation, first we need to identify the main route of the vehicle. For the visualization of the one-year measurement campaign, QGIS Python Plugin - TimeManager and FFmpeg multimedia framework were used. With these tools, we can reproduce the vehicle position for each timestamp and store this in MP4 format.

The purpose of the Time Manager plugin for QGIS is to supply convenient browsing through temporal geodata. A dock widget gives a time slider and a configuration dialog for managing layers. Timestamps have to be in one of the appropriate formats. In our case, the time stamp was in the following format: YYYY-MM-DD HH:MM:SS.ssssss. This tool is essential when it is necessary to animate data and understand its behavior according to timestamps. To produce animations like the one below, QGIS Time Manager allows you to export several individual frames (map images) at a certain time interval from QGIS, which can then be combined into one animation. The example of a single-frame image is shown in the figure 4.7.

The first important step in this work was the creation of a tool to One of the first steps in this work was the creation of a tool to extract and decode NMEA messages and save it in a format that is easy to perform calculations on and to manipulate. The next important step was to understand the scale of the measurement campaign and the behavior of the vessel during the measurement campaign. Using these tools, we can manipulate the data from the measurement campaign and prepare the data for evaluation. These methods will be discussed in the next chapter.

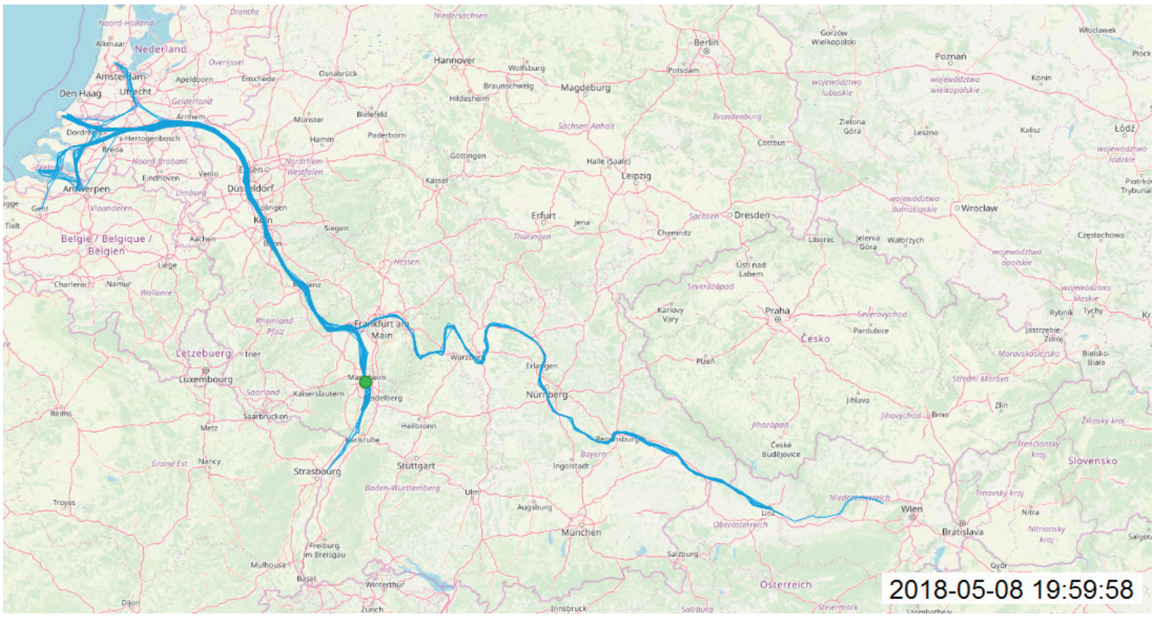


FIGURE 4.7: Vessel position according to time

Chapter 5

Methods and Tools for the Evaluation of the Measurement Campaign

Data analysis requires various pre-conditional operations. The first step of the process is to create a re-scaled reference trajectory according to which we can relate all other data. The next steps are to identify the vessel status and remove data points far from the reference trajectory, like paths in tributaries, or innactive data points such as long periods in harbors. Another important issue is the extraction and division of data into individual voyages. After finishing these steps, we can evaluate the availability of connections in the main waterway. In this chapter, data preparation and processing will be discussed.

5.1 Reference Trajectory

A vital issue for this work was to set data in one reference system - river kilometers, which enable the identification of each point in appropriate part of the waterway. For this task, it was important to have one reference trajectory according to which the calculations can be done. Considering that the vessel embarked on many short voyages, one continuous set of data from Linz to Antwerpen was extracted from the data (see Fig. 5.1). The time frame on this voyage was 15 days from 27.03.2018 to 10.04.2018. The total length constitutes 1392,68 km.

After analyzing the data of this voyage, numerous outliers were detected that needed to be eliminated because we only need the data on the main waterway path when the vehicle is not mooring, changing position or maneuvering.

We have developed a methodology that allows us to determine the status of the vehicle from the measurement data. AIS VDO messages already contain information about the vessel Navigational Status, but this data is not reliable according to various studies. For this reason, it was important to have this information accurate and reliable for further evaluations.

5.2 Determining the Navigational Status of the Vessel

As was mentioned in chapter 2, AIS systems provide information about the navigational status. So for this reason, the first approach in this work was trying to determine the reliable navigational status. For detecting maneuvering and mooring points two parameters were used: the speed and differences between the course

FIGURE 5.1: Reference trajectory

over the ground and the heading. It is known that the vessel uses different ranges of speeds for different movements in the waterway. It can be one of the parameters that can indicate the status of the vessel in the appropriate time frame.

Course Over Ground (COG) defines the direction of motion, with regard to the ground that a vessel has moved relative to the magnetic north pole or geographic north pole. Therefore, should a vessel be stationary, it has no COG. Under conditions where a vessel is encountering leeway (wind, current), a vessel's heading and COG may differ. Heading is used to represent the direction an object is pointing.(see fig.5.2)

For determining conditions for parameters we used a frequency distribution, which shows how often each different value in a set of data occurs. And from figure 5.3, we can make an assumption that the vessel was standing still when the speed is in the range of (0.0 - 0.5), maneuvering between (0.5 - 5.0), and sailing between (5.0 - 10.0+) km per hour.

The histogram 5.4 of differences between course over ground and heading shows that most frequently angles in range (0 - 50) degrees occur. In this range, the vessel was sailing. In this range, the ship sailed along the main waterway, when the angle increases and is out of the the investigated range we can suggest that it is doing maneuvering.

After applying these suggested conditions, we managed to eliminate a range of points where the vessel was standing still and maneuvering. However, the result was not satisfactory because there were still a significant number of points where the vehicle underwent maneuvers or were in harboring areas (see Fig.5.5

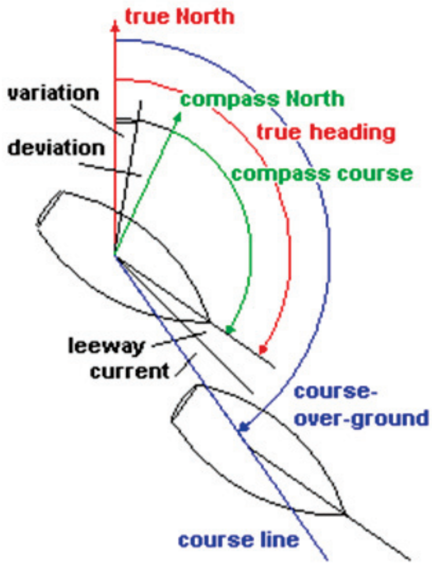


FIGURE 5.2: Directions of the vehicle

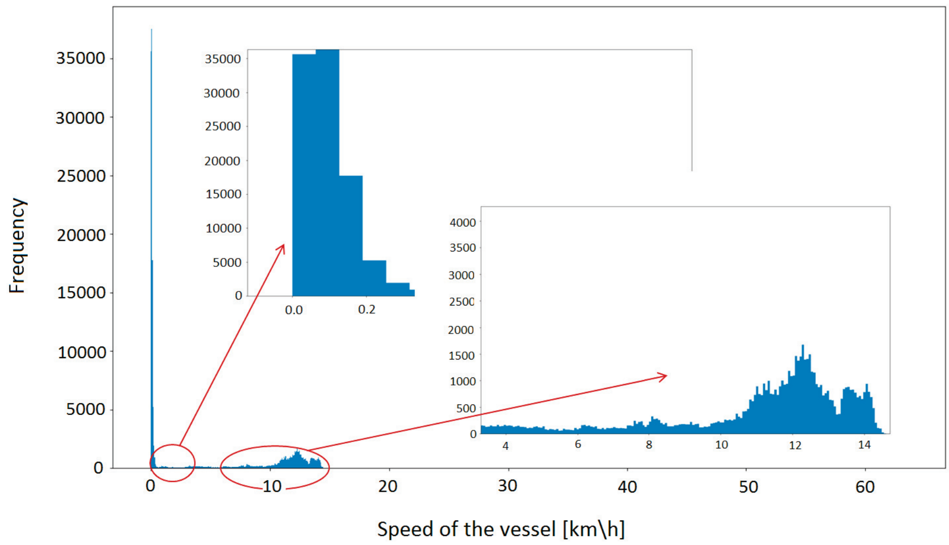


FIGURE 5.3: A frequency distribution of the speed of the vehicle

and Fig.5.6). In some cases, we eliminated some important points that caused us to not have the correct trajectory. And in order to achieve a smooth reference trajectory, this method could not be applied in this investigation.

The approach that has been applied here, checks the time when the vessel leaves the vessel-specific radius. For more accuracy, we checked 3 different radii: small - 5 meters, middle - 15 meters and large - 50 meters. For each point position, we check the time when the vessel leaves the area with a specific radius.(see fig.5.7)

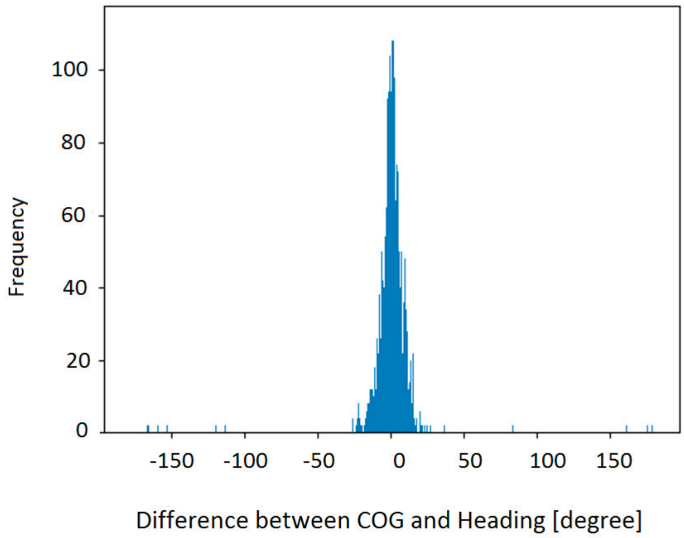


FIGURE 5.4: A frequency distribution of the angle between course over ground and heading

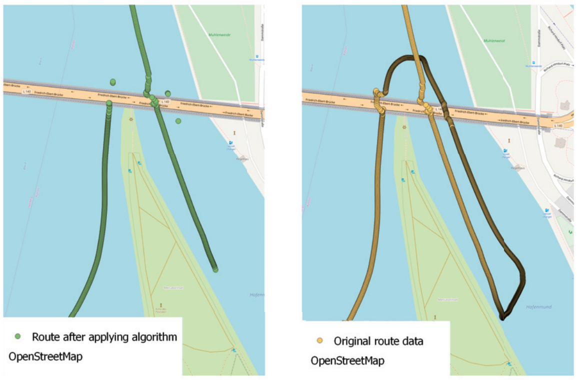


FIGURE 5.5: Before and after applying conditions in the harbour area

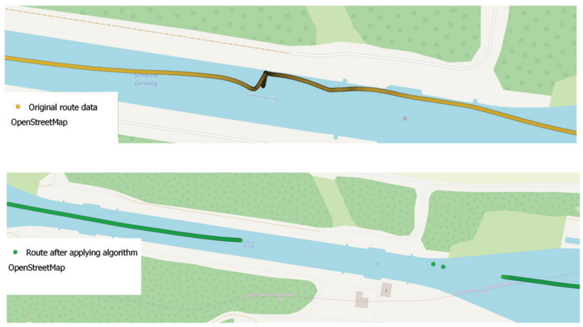


FIGURE 5.6: Before and after applying conditions in the waterway lock area

Name	Radius	Time	Speed
Small	5 m	30 sec	0,6 km/h
Middle	15 m	1 min	0,8 km/h
Large	50 m	1 min 30 sec	2 km/h

TABLE 5.1: Status conditions

For calculating the radius of departure from the great-circle, the distance between two points on a sphere was used, given their longitudes and latitudes (Haversine formula 5.1):

$$d = 2 * R * \arcsin(\sqrt{\Delta\varphi/2) + \cos\varphi_1 * \cos\varphi_2 * \sin(\Delta\lambda/2}) \quad (5.1)$$

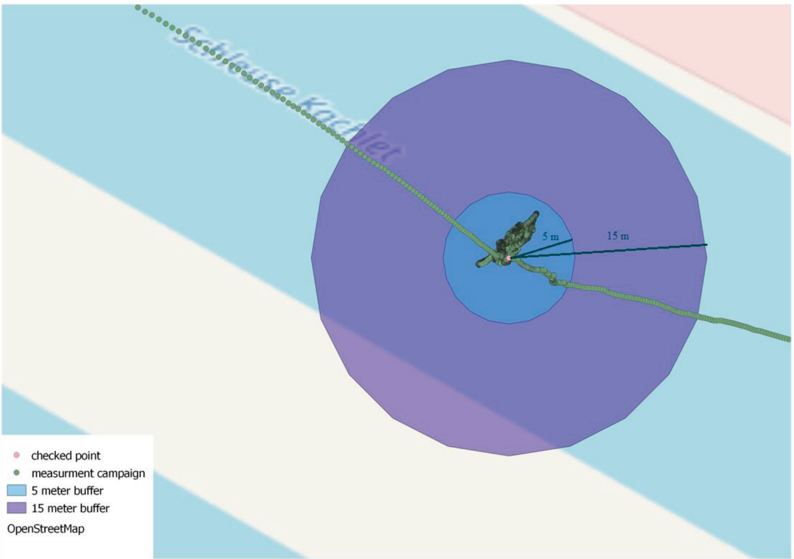


FIGURE 5.7: Use of different radii to detect manoeuvring and standing still areas

where,
φ is latitude in radians,
λ is longitude in radians,
R is earth's radius (mean radius = 6,371km).

After applying this algorithm for 1 continuous route, which was 15 days long, sailing from Linz to Antwerpen, we achieved the following result (Fig.5.8):

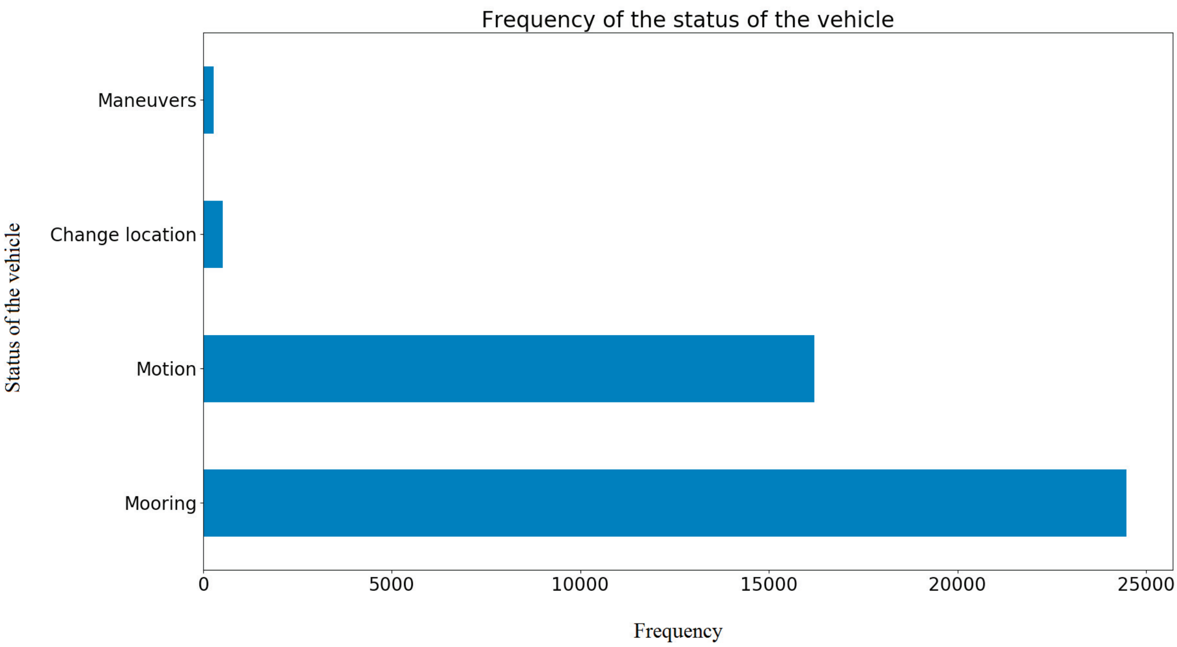


FIGURE 5.8: A frequency of the status of the vehicle

Most of the time - 59%, the vehicle was in a harbor area. To achieve the reference trajectory, points with the status of mooring, changing location, or maneuvering were skipped. From this data set, only 39% of the data where the vehicle was moving were taken. The result was checked in the QGIS platform and has shown a smooth trajectory that is situated only in the main waterway. This trajectory we will use as a reference trajectory. This reference trajectory will be used in the next steps for recalculating river kilometers and the distances to this trajectory.

5.3 Defining River Kilometres

River kilometres (RKM) is a measure of distance in kilometres along a river starting from its mouth. River kilometers begin at zero and increase further upstream. RKM is similar to vehicle roadway kilometers markers, except that river kilometers are not marked on the physical river; instead, they are marked on navigation charts and topographical maps. It should be noted that the total of the river kilometers is not the same as the length of the river. Rather, it is a method of determining location of any feature along the river relative to its distance from the mouth when measured along the course (or navigable channel) of the river.

In cases when some ambiguity exists as to where the river mouth begins, the zero river kilometer point may not be exactly at the mouth. For example, when a river has more than one stream, a series of river mile strings referring to the distance to the ocean or the sea may be used [37].

In this work, the evaluated measurement campaign river kilometers approach was used. Since we evaluated several waterways, all rivers were assigned to a one-river kilometer system. For better evaluation of the measurement campaign, the calculation of river kilometers starts from the beginning of the vessel route. The river kilometer value identifies the location of each point from the measurement campaign relative to the reference trajectory. The 0 kilometer point was assigned in the area of Linz (Austria) and the end kilometer in Antwerpen (Belgium). In between this waterway range, the river passed through big cities like Düsseldorf, Frankfurt am Main, and Nuremberg. The total distance of the investigated trajectory is 1392,68 km.

In an attempt to determine the kilometers of the river, the spline function was first used. A spline is a special function defined piecewise by polynomials. This approach can give us the approximate river curve, by using curve fitting, and an interactive curve design for the complex investigated shape. This approach works well with the small trajectories and when in this trajectory, there exists only one value of y for each value x, i.e. the trajectory should be unambiguous. In our case, the river has many curves that don't satisfy the above preconditions. Therefore, long and complex trajectory needs to be handled differently.

5.4 River Kilometers Algorithm

For calculating river kilometers we will use a reference trajectory that were defined in previous chapter. In this trajectory, we identify for each point the kilometer which it has in the waterway path. The next task we attempt to solve is how to apply this system to all our points from the measurement campaign.

The calculation of the river kilometers is based on the 3 points approach at which one point from the measurement campaign intersects two nearest points from the reference trajectory. This approach will be discussed in more detail in the next chapter. To calculate river kilometers, the following parameters were used:

- reference trajectory
- the transformation from WGS 84 to UTM coordinate system
- the distance to the trajectory altitude were calculated in terms of the sides theorem
- nearest points algorithm

After the kilometers and the distance to the reference path were calculated for all points of the measurement campaign, all data were intersected with the reference path. Data that were more than 150 m from the reference path were discarded.

5.5 Background for River Kilometers Calculations

At first, we will consider a three-point task, where two points are from our reference trajectory and the third point is the point for which we want to calculate kilometer and the distance to the trajectory. We already know the kilometers value for two points from trajectory. And we need to find this value for the third point. This task is to consider a line-segment AB (reference trajectory) and a point C (measurement campaign) that forms a triangle (see Fig. 5.9).

For this task, the altitude in terms of the sides theorem were used. For any triangle with sides a, b, c, and semiperimeter s:

$$s = \frac{(\mathrm{a} + \mathrm{b} + \mathrm{c})}{2} \tag{5.2}$$

the altitude from side a is given by h_a:

$$h_a = \frac{2\sqrt{\mathrm{s}(\mathrm{s}-\mathrm{a})(\mathrm{s}-\mathrm{b})(\mathrm{s}-\mathrm{c})}}{\mathrm{a}} \tag{5.3}$$

To achieve the distance between the point that is a projection of point C onto the line segment AB and the point A, the Pythagorean theorem was used.

$$a^2 + b^2 = c^2 \tag{5.4}$$

To set the appropriate river kilometer value for each point from the measurement campaign from 5.4 we use the next equation:

$$x = \sqrt{b^2 - h_a^2} \tag{5.5}$$

As a result, for each point, the river kilometer according to the reference trajectory and the distance to the trajectory segment were determined.

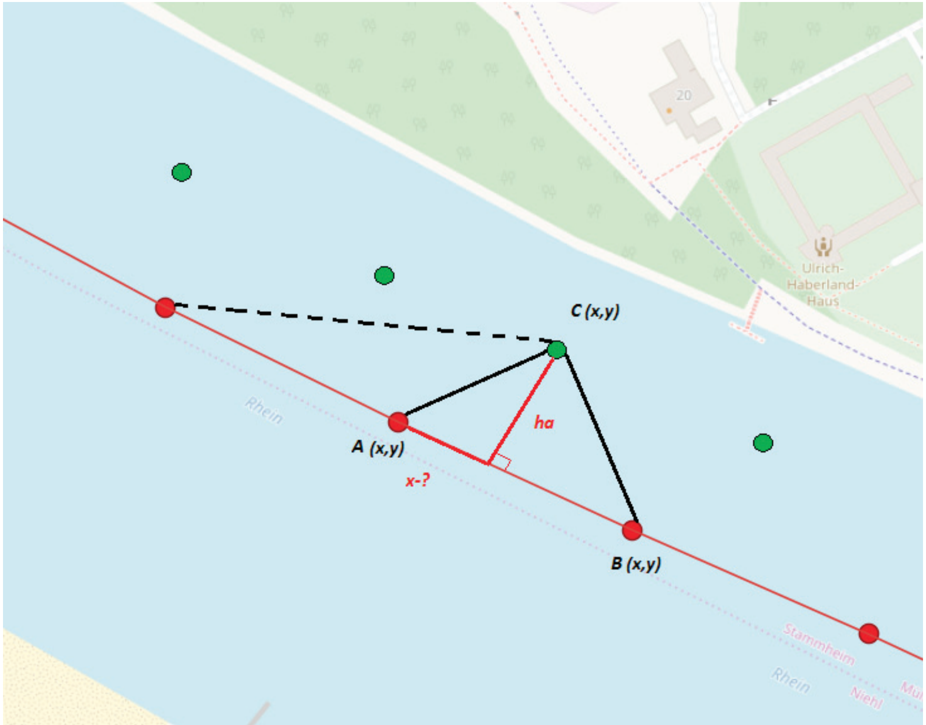


FIGURE 5.9: Line-segment AB (reference trajectory), and a point C (measurement campaign)

The x value represents a distance between a point A and a projection point C onto line AB. The last step is to find the river kilometer for point C which is calculated by A - x. Thus, we achieve the river kilometer according to the reference trajectory and the distance to the trajectory segment.

The next step is to implement this approach for all the points from the measurement campaign. For every point from the measurement campaign, we need to find two appropriate the closest points from the reference trajectory. This task can be solved by finding the nearest neighbour geometry. This approach is one of the most commonly used GIS tasks. In this study, we have a number of point objects representing the reference trajectory of the vehicle, where each point represents a specific river kilometer, and then another set of locations representing a 1-year ship track on the inland waterways. Then, the task was to determine which points from the reference trajectory is the closest one to the individual measurement campaign points. This task is a typical nearest neighbour analysis, where the purpose is to obtain the closest geometry to another geometry.

In Python, this kind of analysis can be done with a shapely function termed $nearest\_points()$ that returns a tuple of the nearest points in the input geometries. The $object.distance()$ returns the minimum distance to the other geometric object. After we have found for a specific measurement point the one nearest river kilometer from the trajectory, we need to find the second nearest point. For this purpose, we check the previous and the next points on the reference trajectory and take the point which has the shortest distance (see Fig. 5.9).

After selecting the point which has the shortest distance to the measurement point, the task reduces to the triangle with the known coordinates in apexes. Then it is required to do the above-represented calculations in the Cartesian coordinate system. For this coordinates from WGS 84 to UTM coordinate system were transferred. This task was done with the help of python library UTM. The next step is to calculate the perpendicular distance to the reference trajectory and define in which kilometer this point is in the reference trajectory.

The algorithm for the river kilometers system was implemented with Python. For computations the following libraries were used:

- shapely
- pandas
- geopandas
- utm
- time
- pdb
- ctypes

A flowchart that represents the whole process of calculating river kilometers to all measurement data according to the reference trajectory is represented in the figure 5.10.

5.6 Detecting and Removing Outliers from the Data

The important step before evaluation of the data is data preparation. At first, the data were set in one river kilometer system. After this, all data were thoroughly checked for continuous behavior. For this an algorithm flagged rows where the river kilometer increases more than the common average speed permits. To set all the data in terms of the reference trajectory, a path buffer with a radius of 150 meters was used. This value was chosen according to the distribution of distances to the reference trajectory (see Fig. 5.11).

With this buffer of 150 meters around the reference trajectory, the data that were not within the buffer and that were when the vehicle was on shore or in a harbor were removed. In this case, the connection links with appropriate data could be evaluated. The figure 5.12 shows the result of implementing a buffer in one part of the waterway path.

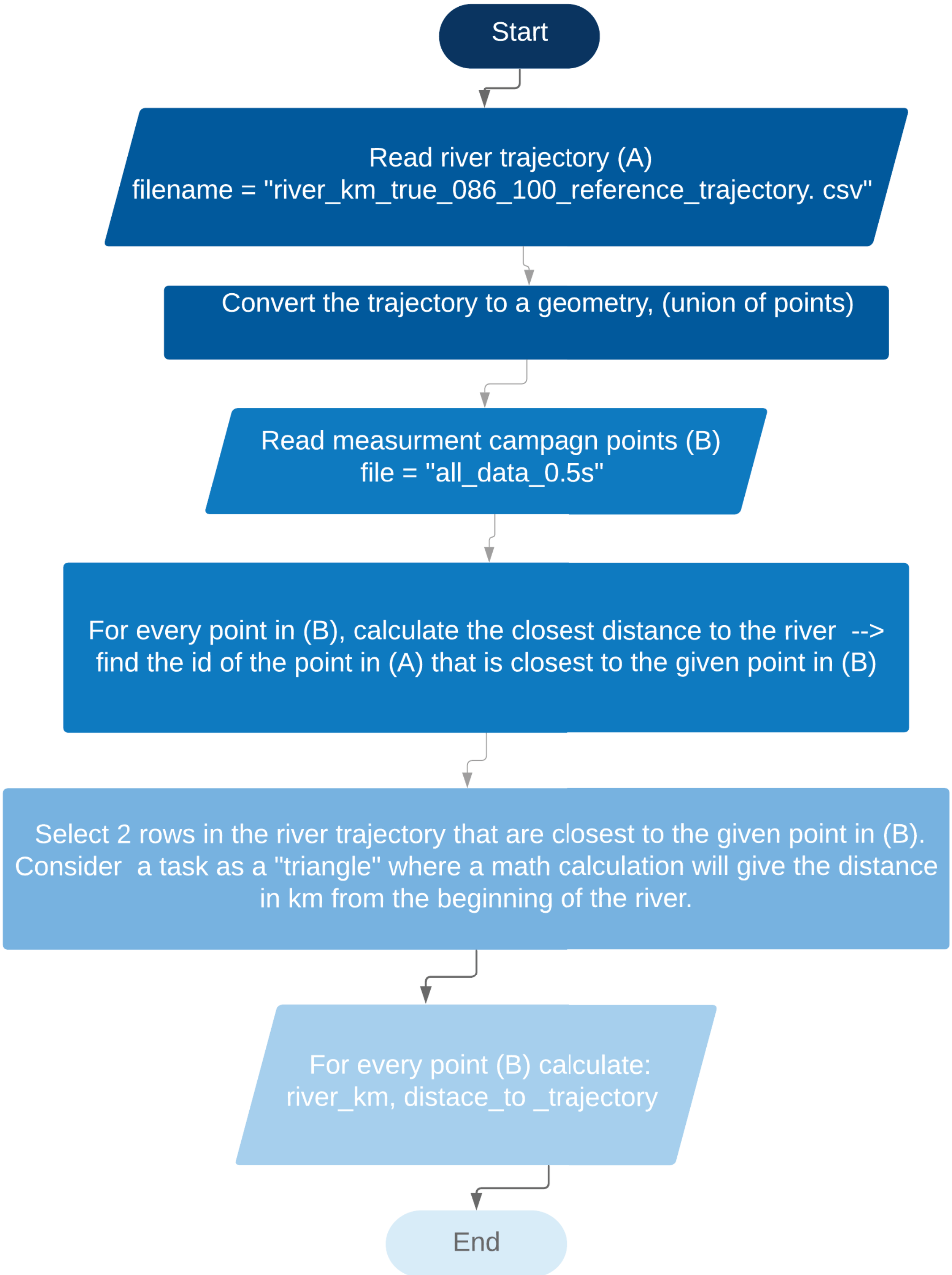


FIGURE 5.10

5.7 The Frequency of the Passage across the Reference Trajectory

The reason for investigating the frequency of the passage across each part of the river was because this parameter has considerable influence on the evaluation of

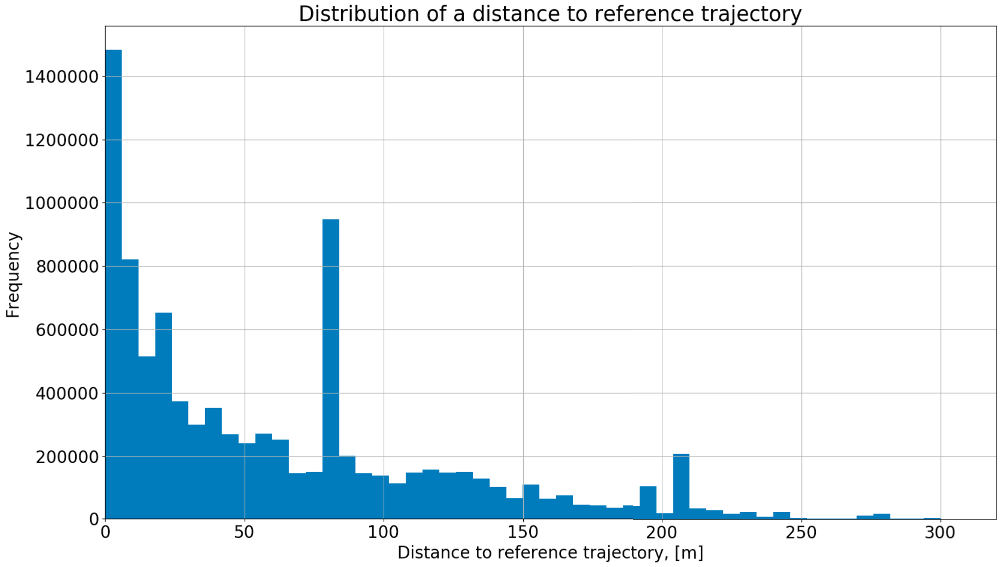


FIGURE 5.11: Distribution of distance to reference trajectory in meters

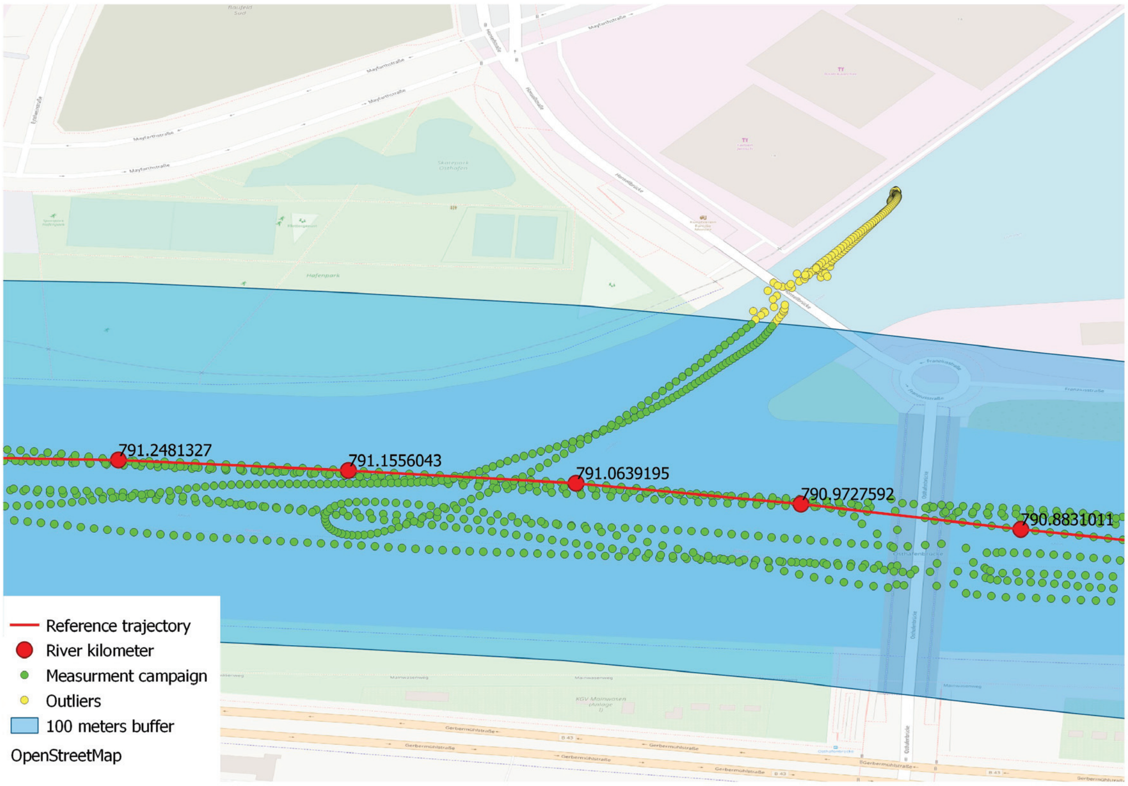


FIGURE 5.12: 150 meters buffer

the internet availability. Thus, the number of passes the vessel took along each evaluated section of the river was determined. Figure 5.13 indicates shows each river kilometer with repect to each time stamp. This illustrates all the tracks and their frequencies during the 1-year measurement campaign.

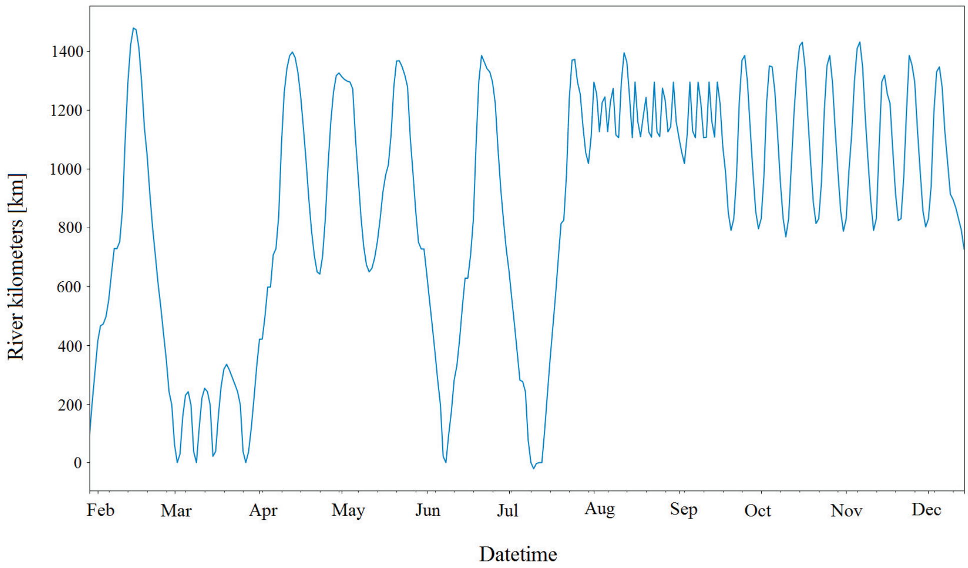


FIGURE 5.13: 1 year measurement campaign El-Nino

From figure 5.13, we can see that the most frequently passed part is the segment from 800 to 1400 kilometer. In this part of the river, the vessel passed 54 times. The 800th kilometer was situated in Frankfurt am Main and the 1400th kilometer was in the city of Antwerpen.

5.8 Algorithm for Splitting Tracks

For the next assessment, it's important to have a tool for dividing data into specific tracks. Since we are working with big data, this tool should be fast and automatic. An algorithm for splitting tracks was developed for this purpose. This algorithm is based on the monotonic function and is applied with the following conditions (see table 5.2).

Condition	**Movement upstream**	**Movement downstream**	**Mooring**
Increase	True	False	False
Decrease	False	True	False
Non - increase	False	True	False
Non - decrease	True	False	False
Monotonic	True	True	False

TABLE 5.2: Conditions for splitting data into separately tracks

The algorithm was applied to the river kilometer parameter behavior. So when this parameter monotonically increasing, we suggest that this is upstream movement and that when it is monotonically decreasing, it is the start of a new track that is downstream (see Fig.5.14). Each change of down or upstream was stored as a new track. When the function is not monotonic we suggest that the vessel is maneuvering or mooring. To analyze the river kilometer behavior, the data were divided into small groups with a variable parameter for the number of taken points. As a standard, were take 300 points for 2 Hz data.

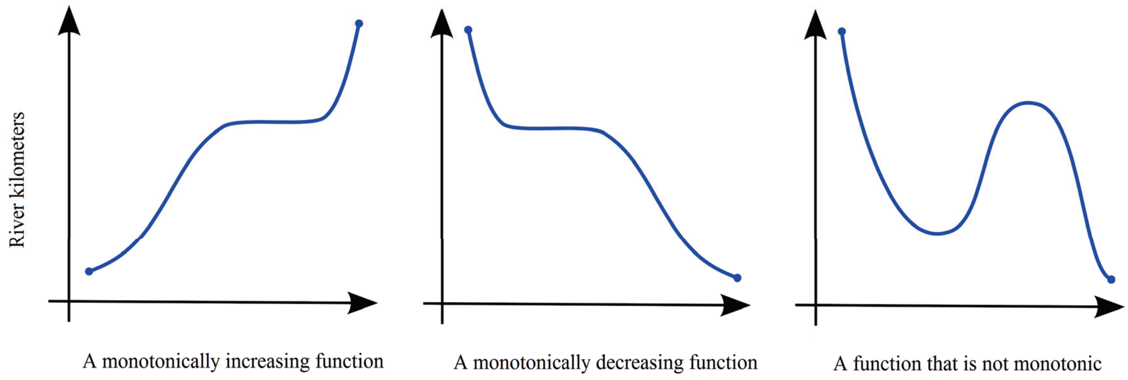


FIGURE 5.14: Behaviour of river kilometer

Chapter 6

Evaluation of Availability of GSM Connection

In this chapter, using the previously presented methods and tools, we will estimate the availability of GSM connections along the entire waterway. The parameters that were used for the evaluation of the connection availability will be described in this part. This chapter also contains the analysis of the internet connection in some specific parts of the river, and different environmental scenarios are analyzed with an emphasis on the influence of waterway constructions such as bridges and waterway locks.

6.1 Parameters for Evaluation the Availability of Communication Channels

The central issue of this work is the detection of areas of the waterway path without a stable internet connection. As described before, the high accuracy positioning provided by RTK measurements depends on the reception of correction data. The indicator which shows the number of epochs/second from the last received correction data set is the Age of Correction [21]. This parameter will be used in the next chapters for evaluating the connection availability. Another parameter that will be checked to evaluate communication is the number of satellites tracked. This parameter can be considered as an indicator of areas with open sky.

6.2 Age of Differential Correction Data

The main parameter that can show the availability of internet connection is the Age of Differential Correction Data. The Age of Correction Data represents how many seconds before the differential correction data are calculated by the DGPS reference base station. With the chance of a constant correction stream, age of differential GPS data will point out link latency. It is computed by subtracting the time when the correction message has been generated from the current receiver time. It is an essential tool to debug connectivity issues.

The Age of Differential Correction Data varies in range from 0.5 to 60 seconds. For precise positioning, the age can be in the range from 0.5 to 1.5 seconds. When the connection was lost, then the value starts growing from 1.5 to 60 seconds and after that the receiver switches to Single Point Positioning(SPP).

When the GSM connection is lost, the influence on the precision can be seen immediately in the altitude measurement.

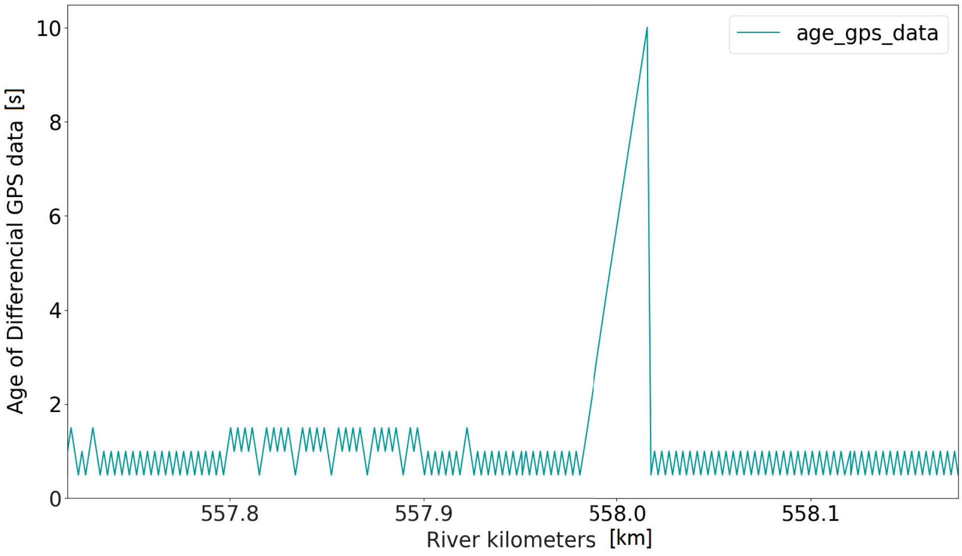


FIGURE 6.1: Loss of internet connection

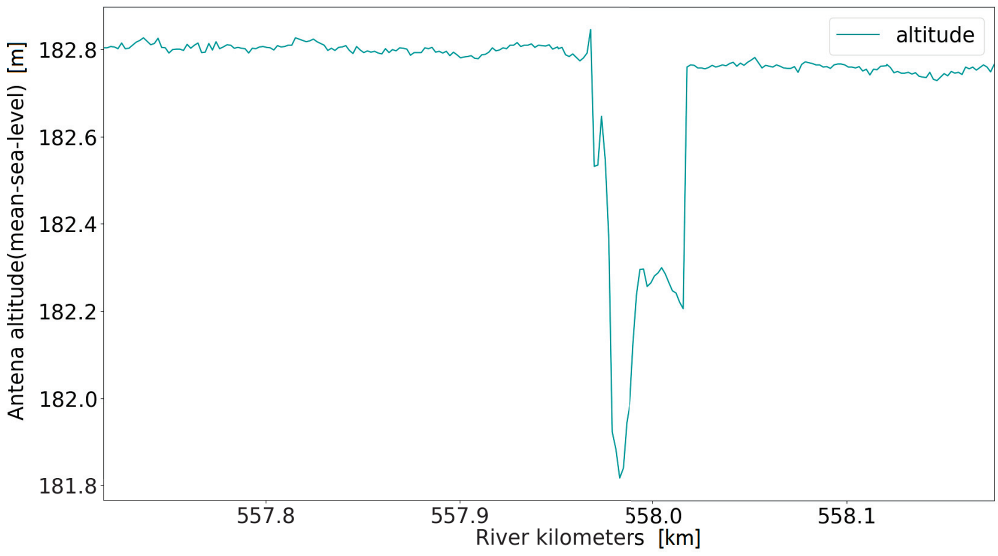


FIGURE 6.2: Disturbance on the accuracy of measuring altitude after increasing Age of Differential GPS data

In figure 6.1, the general behaviour the Age of Differential GPS data in one part of the waterway is represented. We can see that the common variation of the value is in the range from 0.5 to 1.5 seconds. Also, we can see the peak of 10 seconds in the figure 6.1, which represents the disturbance in the broadcasting signal. To see how it influences the altitude measurement,the altitude is plotted for the same area (see Fig. 6.2). It is evident from the graph that when the age of correction data is in the range from 0.5 to 1.5 seconds, there is no disturbance in the altitude measurement, but as soon as this indicator starts increasing we can observe at the same point an immediate increase in the altitude value. This shows how the losing the connection impacts the precision. When the age of correction data increases, the accuracy of the RTK system decreases. These corrections should be in the range of 0.5 to 1.5 seconds to have centimeter positioning accuracy. In the next sections, we will use this range of Age of Differential GPS data to evaluate the ability of the system to provide such accuracy in the waterway tack.

6.3 Availability of Internet Connection along the Waterway Pass

For the first overall estimation, the data from the measurement campaign within a buffer of 150 meters along a reference waterway path were taken. Wherein, we evaluate the connection only in the main waterway path.

The availability of internet connection was determined by the following conditions (table6.1):

Availability	**Age of Differential GPS data (sec)**
Stable connection	0.5 - 1.5 s
Loss of connection	1.5 - 60 s

TABLE 6.1: Internet availability conditions

Under these conditions, the waterway pass was split into 50 meter sections. To determine in which 50 meter case the point is situated, equation 6.1:

$$(km * 1000//50) * 50 \tag{6.1}$$

where, integer division was used.

The algorithm that was used in this work checked the mean value of all passes of the vessel in each 50 meter case. At first, the mean value of all points situated in the first 50 meters were calculated, then with the return pass, the procedure was repeated until we proceeded through all the passes in this specific part. In the end, for each 50 meters case, we know the frequency of passing this part, the mean value of the availability of the internet connection, and can calculate the number of the satellites in this area. This algorithm was applied for the whole waterway path by splitting it into 50-meter cases; as a result, we found the mean connection availability for each case, taking into consideration the frequency of the passing vessel.

The outcome of the availability of connection in the waterway area from Linz to Antwerpen is shown in figure 6.3.

The following are the interpretations of the results. The most frequently passed area has the highest value of availability, and some parts that experience poor connectivity should be considered in more detail to determine the cause of the signal

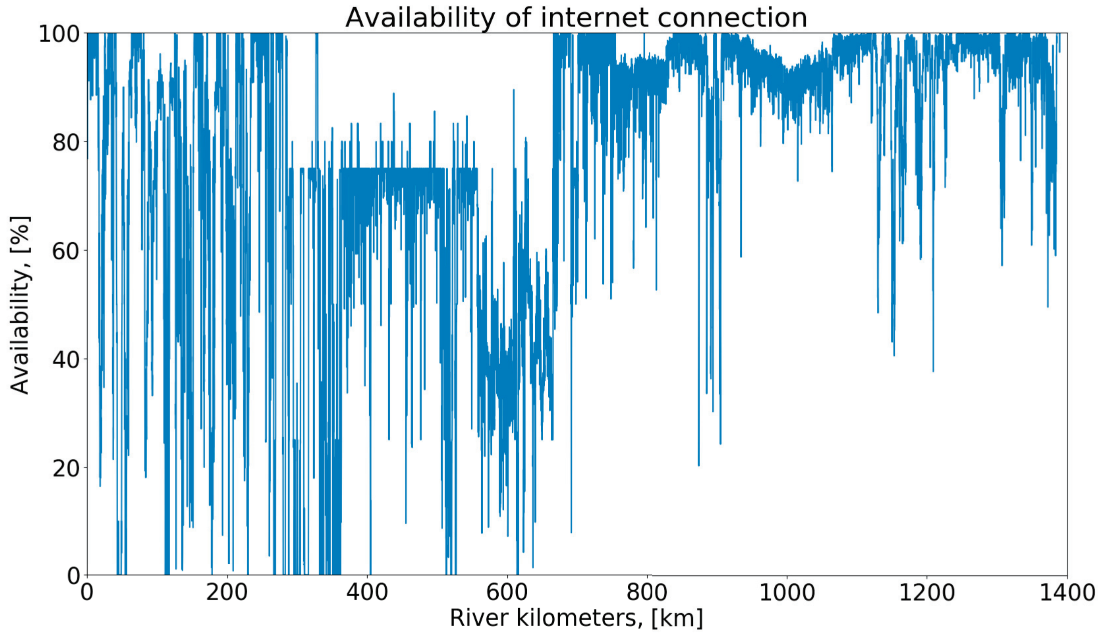


FIGURE 6.3: Availability of internet connection along the waterway pass from Liz to Antwerpen

interruption. And of course, different environmental scenarios can influence on the internet connection: the most common of which will be investigated in section 6.7. It should be noted that the region located from 560 to 600 kilometers is a special test AIS area, and this part will be specifically evaluated in the next chapters.

6.4 Number of Satellites

The number of used satellites can be an indicator of the clarity of the sky and which areas are good locations for communication availability. To validate this assumption, this section we will explore the role of the number of satellites visible to the vessel.

When the communication link is not available and can not provide correction information from the base station, the receiver switched to a stand-alone position.

For the investigation, the mean number of satellites for all points within 50 meters of this location were calculated. When there was a high number of satellites, a stand-alone position was used. Changes in the number of satellites can be compared with the behavior of the age of differential GPS data.

The figure 6.4 shows the opposite curve which we achieve with the age of differential age data. Consequently, this method can also be used for the description of the availability of the communication link.

To see the correlation between these 2 methods, the internet availability achieved from the age of differential data parameter and the mean number of satellites were plotted.

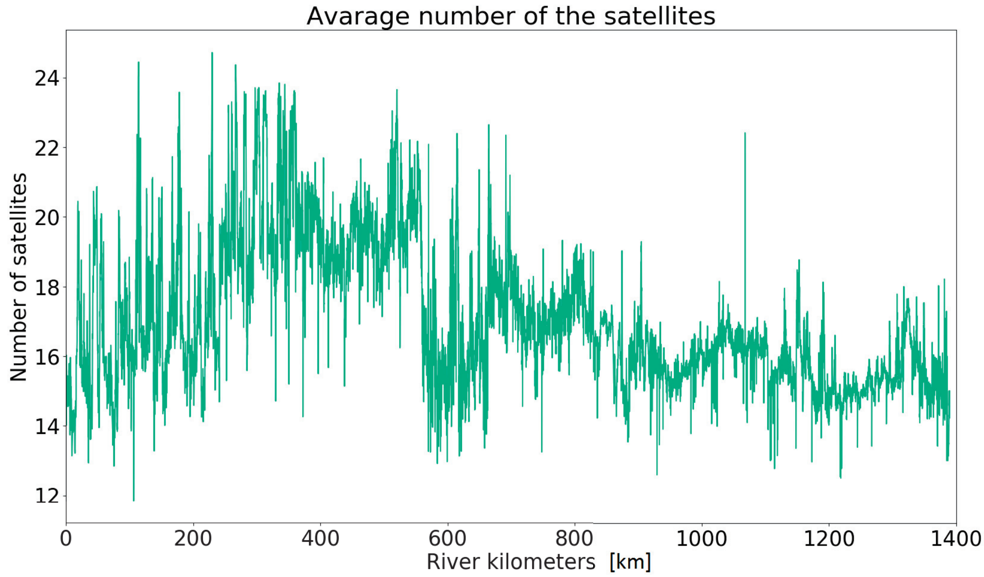


FIGURE 6.4: Number of satellites in view in the waterway path

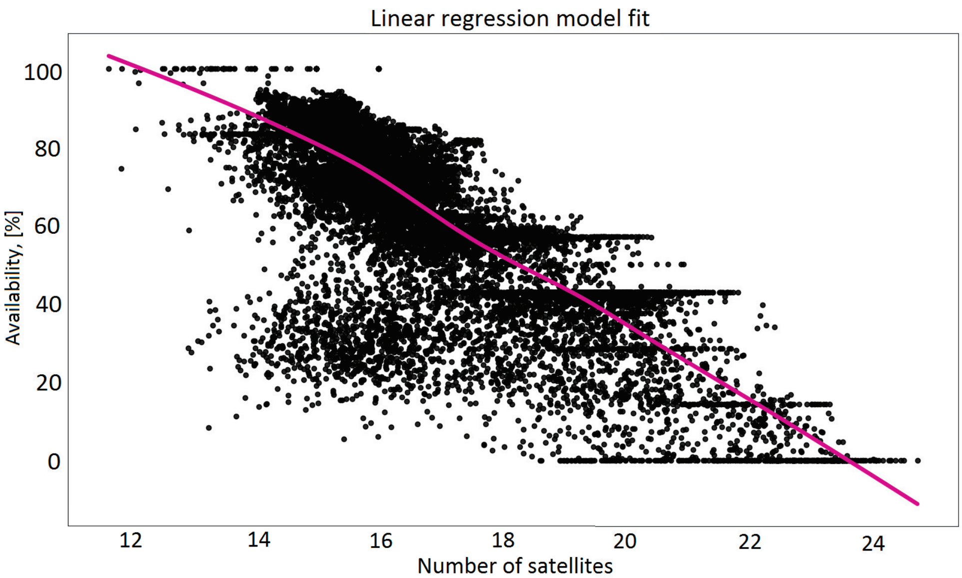


FIGURE 6.5: Liner regression between availability of internet and number of satellites in view

The figure 6.5 shows the dependence on the availability of the Internet and the number of satellites used during the evaluation measurement campaign. However, due to the receiver setup, our assumption about the usage of the number of satellites did not prove to be used in this evaluation. Because we have only the number of used satellites not the information about the satellites in the view. Thus, in the next chapters for the evaluation of an internet link, only the Age of Differential Data will be used.

6.5 Correlation of Internet Availability in the Upstream and Downstream

One assumption that can be checked for the whole pass is that the connection depends on the direction of the voyage: upstream or downstream. For this purpose, two continuous tracks in the upstream and downstream directions were used, and the data were checked separately. To analyze the dependency of the internet connection of different routes, the cross-correlation method was used.

The cross-correlation method of two signals allows for the determination at which time lag the correlation factor is the greatest. The cross-correlation of the normalized discrete-time signals $\overline{x}[n]$ and $\overline{y}[n]$ is expressed as equation 6.2.

$$R_{xy}\left[m\right] = \sum_{n=-\infty}^{\infty} x\left[n\right] y^{\star}\left[n-m\right] \tag{6.2}$$

In our case, $x[n]$ is upstream and $y[n]$ is downstream.

CCross-correlation of any two given signals can be found via graphical techniques. The cross-correlation method was applied to two different datasets from two different observations, the aim is to figure out if those two sets are correlated. We want to cross-correlate them and see if they match in any way. We plotted availability for both datasets' in 100 meters chunks over the top half of the figure to visualize the data. For checking the parts with a loss in connection, missing values after 60 seconds were filled with the value 99. Using matplotlib's xcorr, which in turn uses NumPy's correlate() function, we computed cross correlation and plotted it on the bottom half (see Fig. 6.6).

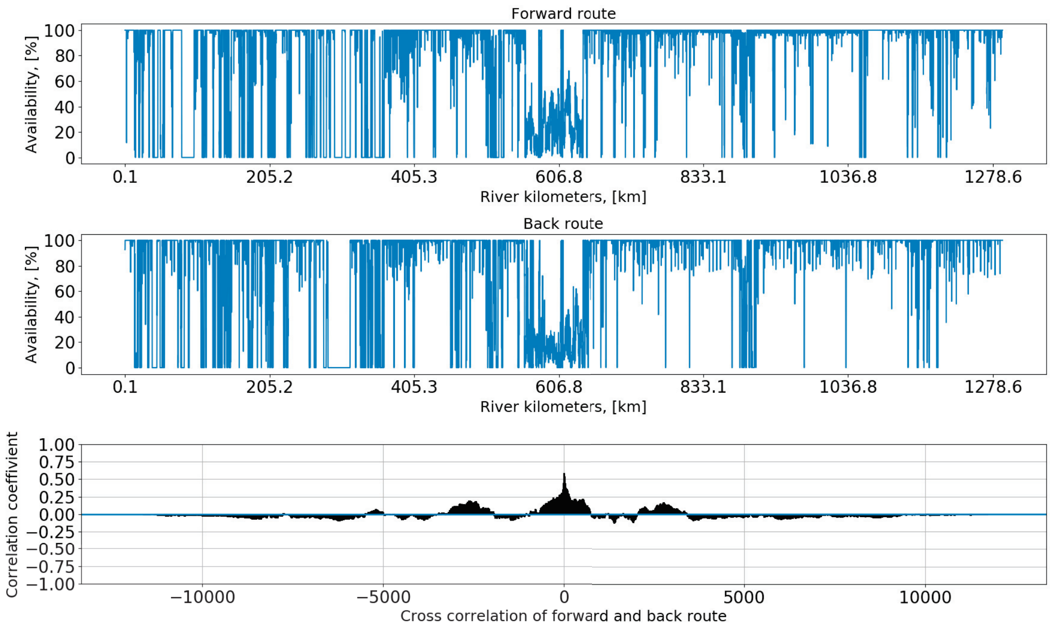


FIGURE 6.6: Correlogram of availability of internet connection forward and back route

We achieved correlation coefficients array that represents the degree of similarity of two datasets. The cross-correlation diagram shows us that these two routes are highly correlated. And at lag 0 the correlation coefficient has its peak.

The average width of the river is around 600 meters, so in this range, the connectivity is the same. Consequently, we can suggest that the parts where the internet link is not stable are constant and don't depend on which direction upstream or downstream the vessel is navigating.

6.6 Evaluation of Access to Communication Channels in the Most Frequently Traveled Parts of the Waterway

From figure 5.13, we can see that the most frequently passed waterway section is from 1100 to 1250 km. For a more precise and accurate evaluation this part of the vehicle track was taken. This part of the river was passed for 54 times.

After applying the algorithm of splitting tracks, which was discussed in section 5.8, we achieved 54 separate tracks over this part of the river. To verify the connectivity behavior in each individual case, the mean internet availability was calculated in the vicinity of 100 meters of each track. In an attempt to identify if the connection depends on how frequently the vessel passes through a given waterway section, a histogram of mean internet availability for each track were created (fig. 6.7).

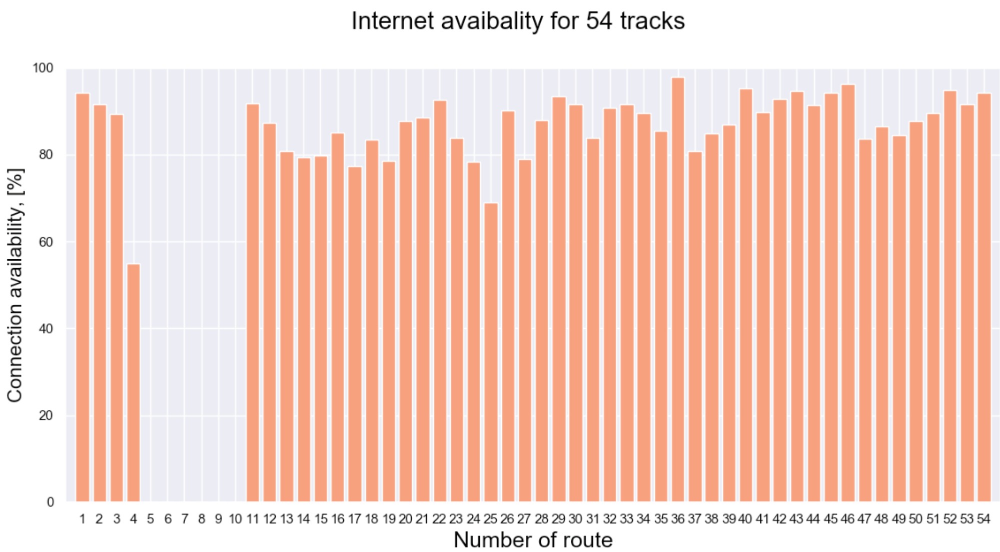


FIGURE 6.7: Distribution of internet availability for 54 tracks

The distribution of internet availability is as follows: 38 tracks have a very high percentage of connection and vary in the range of (100 - 80)%, 9 tracks have some disturbances in connection and were in the range of (80-50)%, and 6 tracks have almost 0% connection. Those with 0% of connection are tracks that are not defined.

6.7 Environmental Scenarios in GNSS for Vehicle Localisation

The analysis of the measurement campaign of the vehicle shows the error source from the user segment. From the user side, different environmental scenarios such as the lack of a signal, shadowing, multipath, and signal propagation effects cause an accuracy degradation in localization.

Scenario	Scenario Description
open area	track laying outside town or city, very good view of tracks
urban track	very close to the buildings
bridges area	loss of visibility
waterway locks area	narrow way with changing in altitude
mountains	on one side or close to one steep mountain causing barrier for signal
harbour area	close to the buildings

TABLE 6.2: Inland waterway track environmental scenarios

The vehicle in a waterway area can be met with different infrastructures and different geographic properties. This can lead to signal interruptions or shadowing effects. Additionally, the vehicle travels through city areas thus having multipath effects.

The accuracy of GNSS receiver locations is highly correlated to where the vehicle is. The environmental scenarios can be categorized and defined as in Table 6.2. The GNSS receiver location performance varies a lot in these environments. To apply reliable and accurate GNSS for vehicle localization, the safety aspects of these environments need to be analyzed and taken into account.

The GNSS receiver will experience different GNSS localization accuracies depending on the environmental scenario. The impact of two waterway locks in close vicinity is shown in figure 6.8. The falling height of both waterway locks is 24.67 meters. This scenario shows what occurs when the receiver is in such challenging waterway area. We can see where the speed was almost 0 and the altitude was reducing in the location of the waterway lock infrastructure. We can see that in the beginning, in both cases where the waterway level is reducing we still have a good age of correction with a variation between 0.5 and 1.5 seconds. When the age grows to 60 s, the receiver switches to the single point positioning, but even with this positioning method, we have a big variation in the number of used satellites. That is caused by the influence of the walls of the waterway lock. At some middle point in the passing waterway lock, the receiver was able to receive corrections and give the position in the RTK or DGPS, but then switch back to SPP.

The above scenario shows the demand for further investigation of influence such waterway infrastructures in the communication link availability. There are many different obstructions in the waterway that can affect the availability of the internet. The prime example that can be investigated is the influence of bridges and waterway locks.

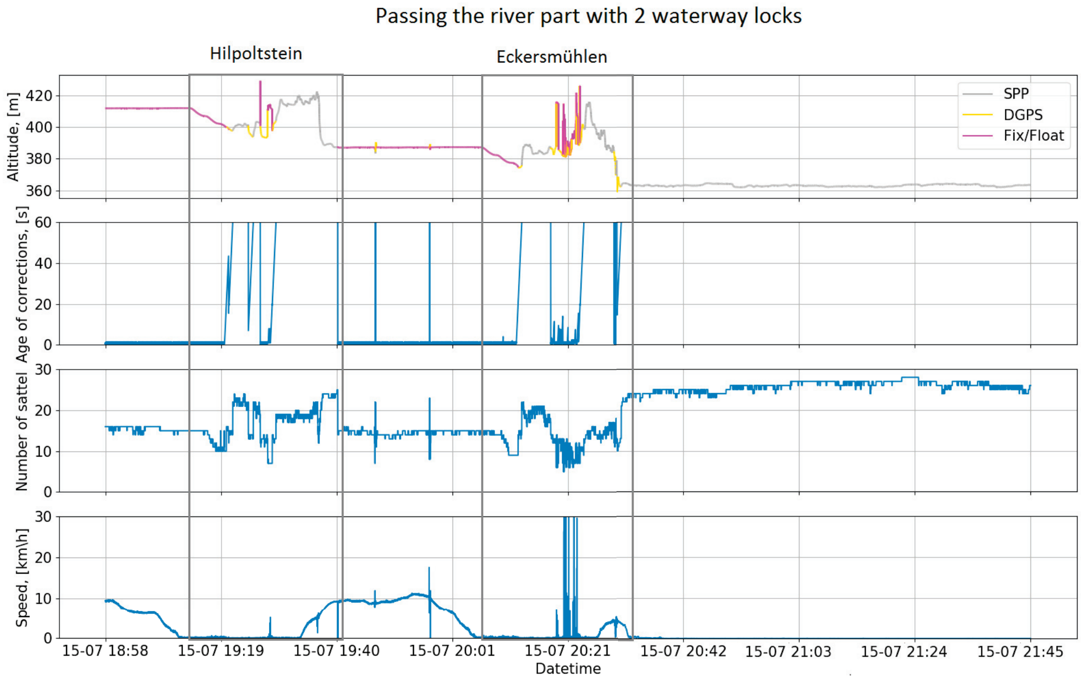


FIGURE 6.8: The waterway locks scenario

6.8 Availability of the Internet in the Waterway Infrastructure Areas

Inland waterways of Europe and Germany are intersected by bridges in numerous places. Most of the older bridges, in particular, are sustained by piers in the fairways or at their edges. The bridges' influence on the ship's behavior and the nautical skills needed for passing them are often undervalued.

Bridge piers are man-made constructions in the river cross-section, disturbing its discharge pattern. In this area, the water level is not stable and moreover, due to the bridge construction, vortices can appear in the water surface. Another situation that requires high precision is the interactions with passing ships, especially in sophisticated bridge constructions.

Another waterway construction that can be evaluated is waterway locks. These are waterway navigation constructions in canals and canalized rivers that aid ships to overcome the height difference (lift height) between the upstream and downstream levels of two stretches of water. German locks that are constructed for large ships have a clear width of 12.5 m. The available widths of river locks are in some cases (Rhine and Danube: 24m) even larger. In Germany, the available length of a lock chamber, i.e. the clearance between both gates minus a safety distance on both sides is commonly in the range between 100m and 300m [35].

Due to the restrictive width of the waterway lock structure, it requires precise navigation and timely response to any influences that could make the ship deviate from its course.

The influence of the described waterway infrastructures was investigated, and with the help of the GIS platform, the area from Linz to Antwerp were digitized and transformed to the river kilometers system; 275 bridges and 59 waterway locks were found. The figure 6.9 shows the average internet availability and the positions of waterway locks and bridges. It is obvious from figure 6.9 that in the area where most of the waterway infrastructure is located, there is low internet access.

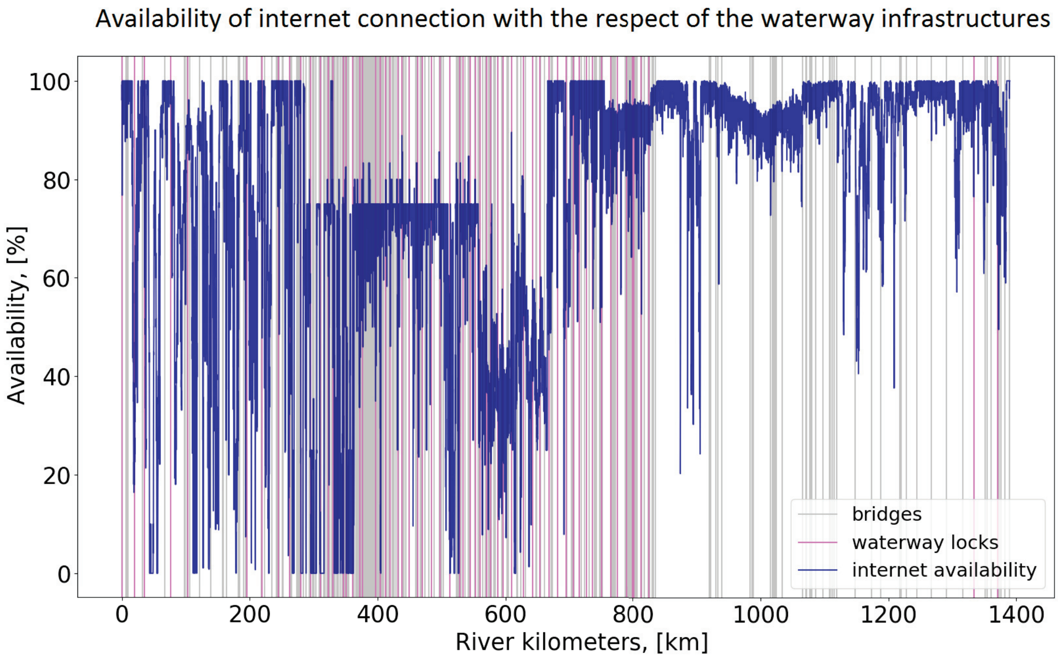


FIGURE 6.9: Availability of internet and the positions of waterway locks and bridges in the waterway path

For investigating this more in detail, every 100-meter chunk of the waterway pass was checked to determine the existence of a bridge or waterway lock. After this, a special column was created that classifies these areas as: bridge areas, waterway, locks areas, and open sky areas.

The area parameter enables us to check if this waterway infrastructure has an effect on the availability of connection. The histogram 6.10 and 6.11 shows an increase in the frequency of low availability of internet in both types of waterway infrastructure, yet the internet availability is pretty high in areas without waterway infrastructure (see Fig.6.12). The mean percentage of internet connection in the bridge are is 48.6%, in the waterway lock 41,8% and in the open sky area 58.2%.

But this still does not give us considerable certainty as to the loss of connection being dependent on a specific part waterway infrastructure. To enable us to suggest that the loss of connection occurs due to infrastructure, the mean value of internet connection. Of the areas within 100 meters before and after waterway infrastructure was calculated. This investigation may show that either the infrastructure of the waterway causes a loss of Internet connection or that the whole area has a poor Internet connection.

The delta value that represents this difference of connection was taken from the next equation:

$$\Delta = availability(v) - availability(i) \tag{6.3}$$

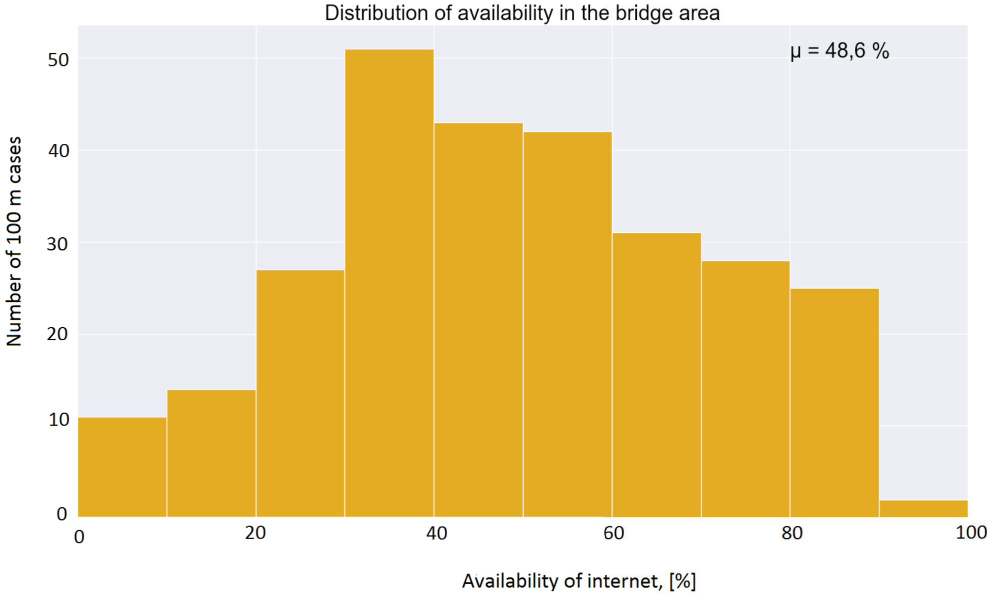


FIGURE 6.10: Availability of internet in the bridges area

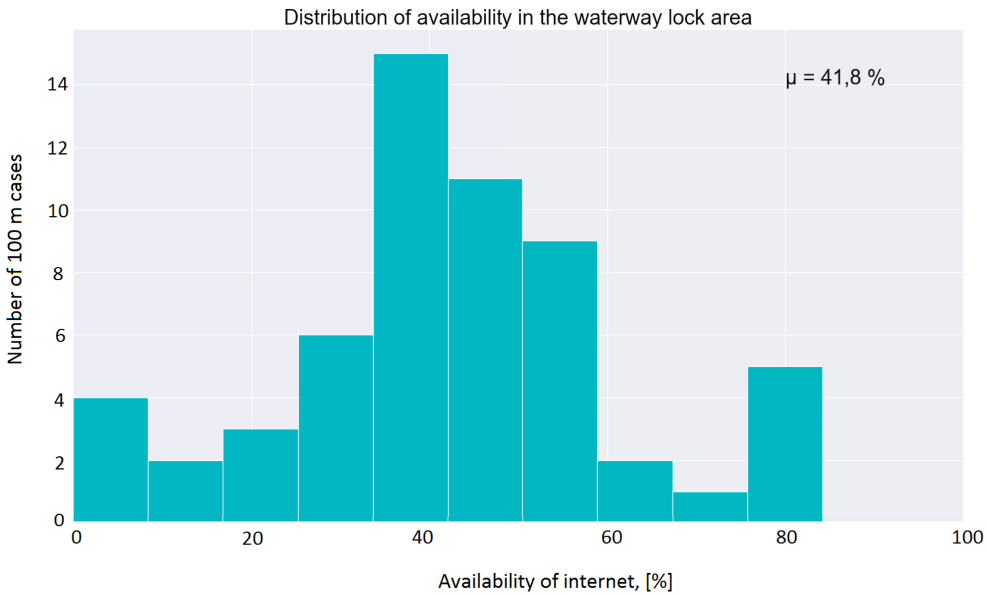


FIGURE 6.11: Availability of internet in the waterway locks area

where,
availability(v) - availability in the vicinity area (mean value of 100 meters before and after infrastructure area),
availability(i) - 100 meter around the infrastructure area.

The figure 6.13 shows that bridges area has less connectivity than the areas that is before the bridge and after the bridge. The mean internet availability in the bridge area is 48.6%, when in the vicinity of the bridge the internet availability is

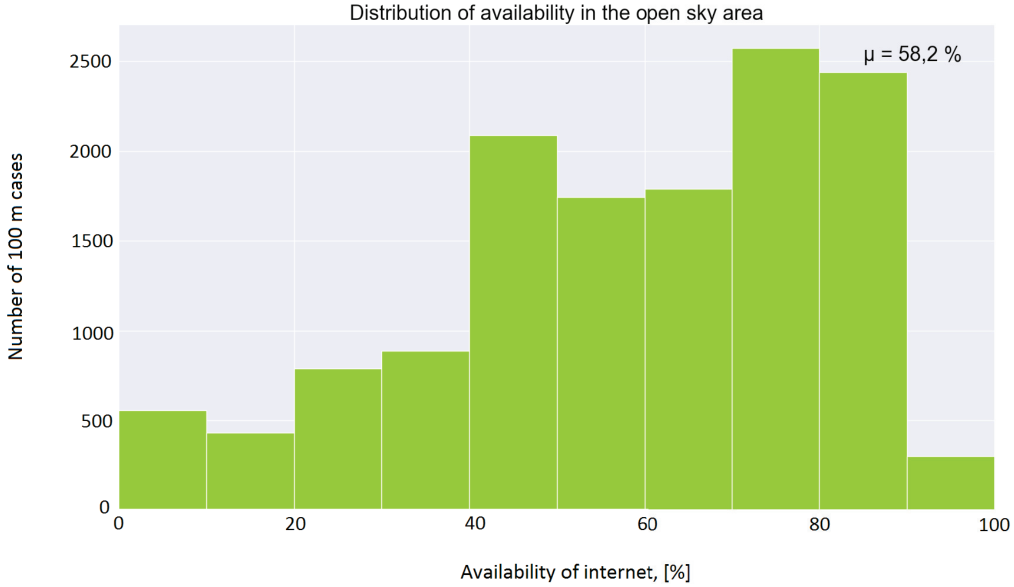


FIGURE 6.12: Availability of internet in the open sky area

51.4%.

For waterway locks, this difference is not as significant, but we can suggest that it has some impact on the connectivity (fig. 6.14). The mean internet availability in the waterway lock area is 41.8%, when in the vicinity of the waterway lock the internet availability is 42.6%. The waterway locks are huge waterway infrastructure, and with such investigation, it is hard to see the direct impact because we need to know the exact position where the waterway lock infrastructure starts and ends.

The next that can be checked is in which range the loss of connection occurs. Additionally, construction specifications have to be investigated in more detail.

For analyzing the influence bridges have, the part with a high average availability of the internet in the section between 1000 and 1300 km was taken. In this area, there are 32 bridges. After GIS analyzing we received the next statistic (table. 6.3):

Parameters	**High influence**	**Slightly influence**	**No influence**
Number of bridges	9	11	14
Range of loss connection	10-40 %	< 10 %	0 %
Width of the bridges	35 - 75 m	20 - 35 m	< 25m

TABLE 6.3: Bridges influence on loss connection

In approximately 56% of cases from the chosen area, the bridges disturb the connectivity. Also, bridges can create shadow effects that are out of the signal flow. As an assumption, we suggest that different constructions and materials for bridges can have different influences on the stability of the connection. The pedestrian bridges have almost no effect on the signal.

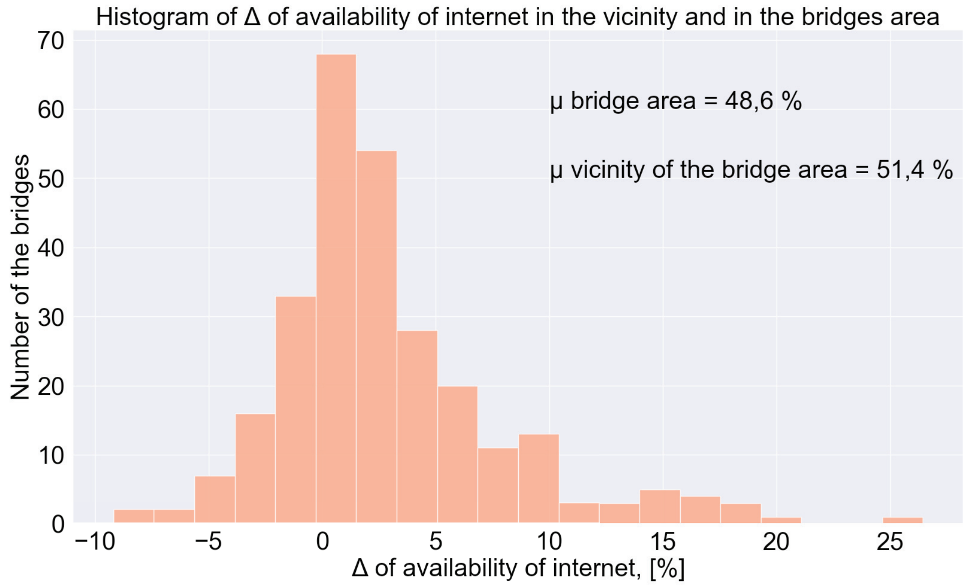


FIGURE 6.13: Mean Δ availability of internet in the waterway locks area

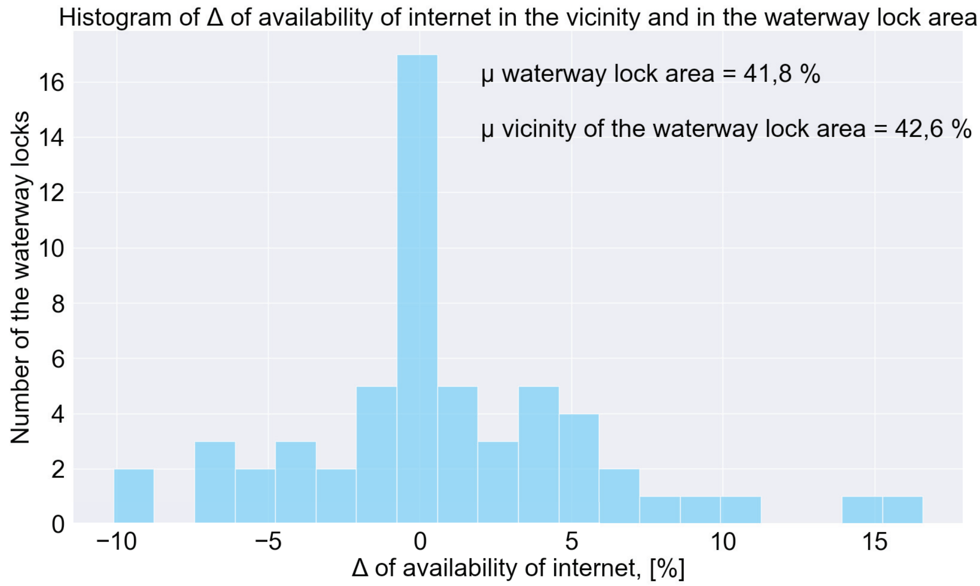


FIGURE 6.14: Mean Δ availability of internet in the waterway locks area

The example of the different bridge constructions is presented in figure 6.15. For example, the Cologne Rodenkirchen bridge(a) is 52,3 meters wide and has 40% less connectivity than the area in the vicinity of the bridge connection. Yet, the Rheinkniebrücke has almost no influence on the connection, and the width of this bridge is 28,9 meters.

The figure 6.16 shows the Old bridge in the Würzburg with the waterway lock. In this case, the loss of connection is 15%. In such areas, it is hard to determine

(a) Cologne Rodenkirchen Bridge (b) Düsseldorf Rheinkniebrücke

FIGURE 6.15: Different bridges constructions

the influence of which infrastructure reduces internet availability. Such areas are generally highly challenging for inland navigation.

FIGURE 6.16: Old bridge in the Würzburg with the waterway lock

For the investigation of the connection availability on the waterway locks, an area with 41 waterway locks was taken, and we receive the next statistic:

Parameters	High influence	Slightly influence	No influence
Number of waterway locks	9	17	15
Range of loss connection	35 - 50 %	8 - 32 %	0 - 6 %
Falling height	3.30 - 24.67 m	2.75 - 10.34 m	2.74 - 4.01 m

TABLE 6.4: Waterway locks influence on loss connection

In 60% of the cases, the waterway locks have an influence on the loss of connection. The falling height varies from 2.74 to 24.67 meters. As we can see, there is no straight correlation between the falling height and the range of a loss of connection. But, it is obvious that the highest percentage of influence has the waterway locks with the highest falling height. As an example, one of the biggest falling height waterway locks is represented in figure 6.17. This waterway lock has 24.67 meters of falling height.

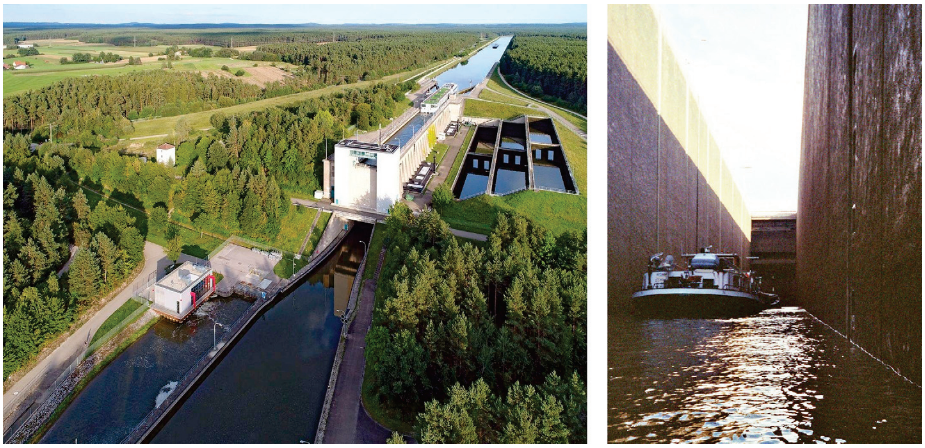


FIGURE 6.17: Leerstetten waterway lock

6.9 Poor Internet Connection Areas

For the purpose of analyzing some specific parts with bad connectivity, 100 meter groups were taken and the availability of the signal in each specific voyage was checked. The lowest internet supply is in the 200 - 500 km waterway path. First, it should be checked if the connection availability is always poor or if it can have a good connection.

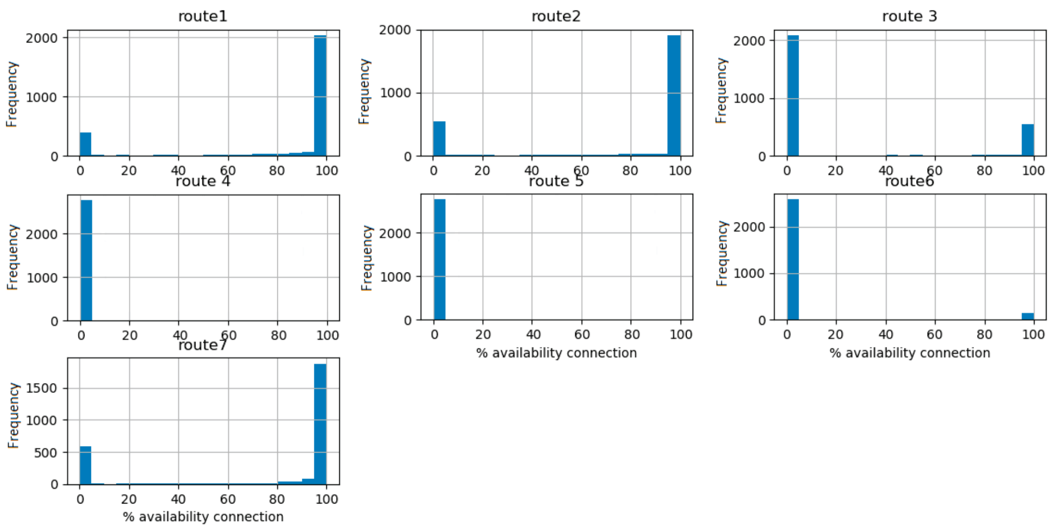


FIGURE 6.18: Frequency of availability of connection in each route

As we can see from figure 6.18, only 3 routes have a significant number of 100% connectivity. The other 4 routes have almost 0% connection. The specific parts of the river should be checked for defining the cause of the loss and lack of the signal. For analyzing these areas, box plots were built that display variations in samples of a statistical population. So, we can see where the internet connection is stable low and where it varies randomly from 0 to 100. One of the parts is the boxplot 6.19

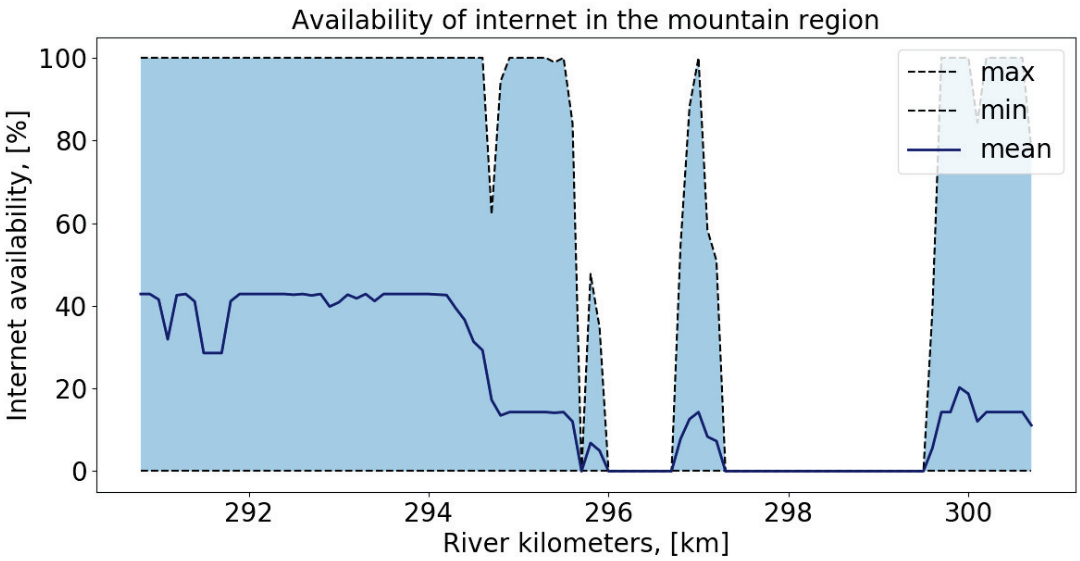


FIGURE 6.19: Availability of the internet with poor connection

The figure 6.19 shows that at the first part, the connection ranges from 0 to 100% and is not stable. So the mean value is about 40%. But in the second part, the connection is almost 0%. We can see that there are some small part around 100%, but most of the value is 0% connection.

Next, we should check the location of this area. From figure 6.20, we can see that in the area where we have almost 0% connection, the river makes a large bend loop to the north around the mountain. And this loop is exactly the area where we lost the connection from 294.8 km to 299.7 km. The mountain that is situated is named Speckelsberg and is 531m tall. Consequently, this area is poor for GSM signalling, and we can see that this area is constantly without connection.

In figure 6.21 the city area with bridges and waterway lock and rural area is represented, and it clearly shows the effect of different environments on the internet availability.

6.10 Map of Availability of GSM Connection

The analysis of bridges and waterway locks shows that there are a number of waterway infrastructures that have no influence on the availability of connection. To have a more clear understanding of where the reliable and poor parts of connection in the waterway path are located, a digital and interactive map that visualizes connectivity and the location of GSM stations in the main waterway path built.

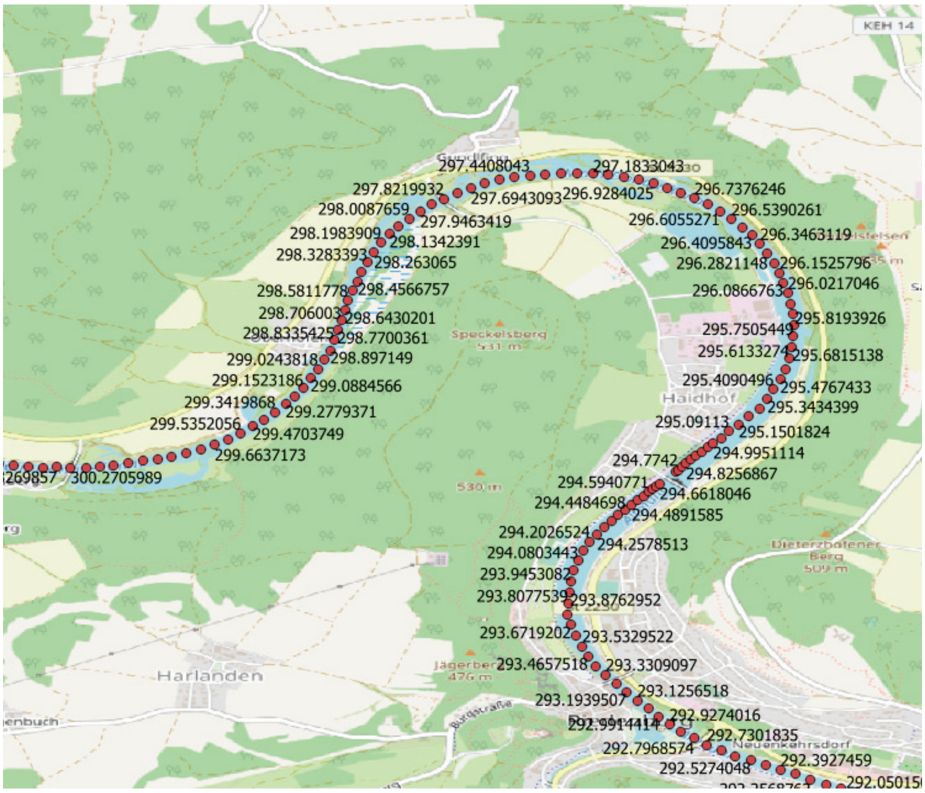


FIGURE 6.20: QGIS: OpenStreetMap with river kilometers layer

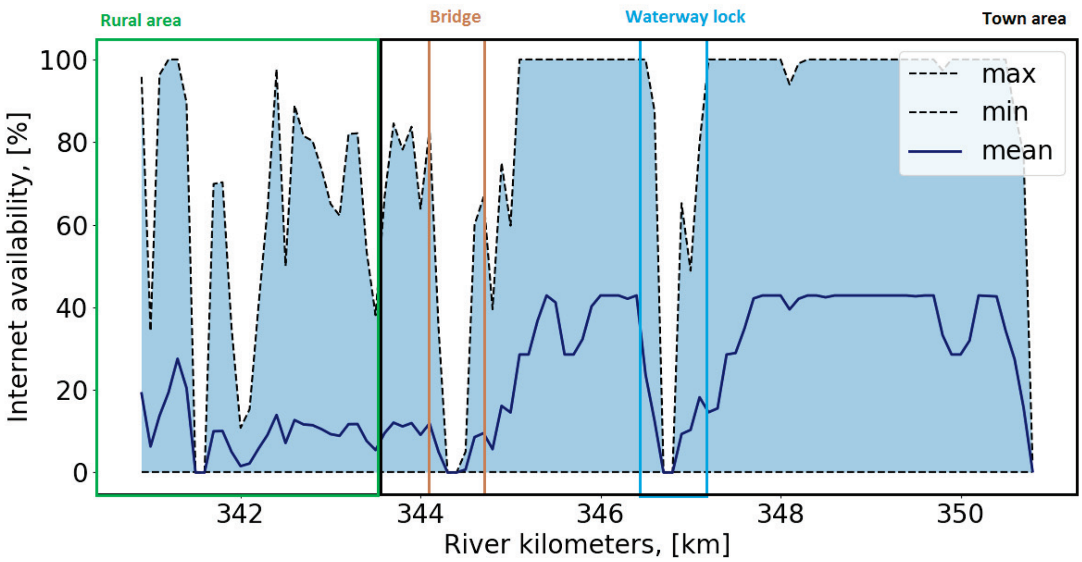


FIGURE 6.21: Availability of internet with poor connection in rural and town area

Folium python library was chosen. This library allows for the manipulation of data in Python, and then visualize it in on a Leaflet map. Using maps instead of other forms of charting has a lot of advantages as highlighting trends, uncovering patterns, and revealing realities not visible before when it comes to spatial data. It clarifies aspects of the data, rather than just simplifying the data itself.

Python geospatial development was done by data manipulation and analysis using geopandas, shapely, pandas, and other Python libraries. The purpose of each library is shown in table 6.5

Library	**Functionality**
Pandas	help to structure and manipulated data in a DataFrame format
geopandas(gpd)	allow To work with spatial data in a DataFrame
geopandas(GeoDataFrame)	enable creation a GeoDataFrame from a DataFrame
shapely.geometry(LineString)	creation line geometries that can be used in a GeoDataFrame
folium	to generate a Leaflet-based map with the analysed data

TABLE 6.5: Python libraries and there functionality

A line segment requires two pairs of of XY coordinates. First, for each row we set a special geometry type from a shapely point. Then, we create a gdf (GeoDataFrame) where the XY coordinates for a given row can be found in the geometry column. Then with the help of LineString() constructor, we start creating lines. With help of the function make_lines(), we can take two sets of XY coordinates from a GeoDataFrame and add them to a shapely LineString() constructor. As a result, we have a DataFrame. Then we loop through a GeoDataFrame, passing consecutive XY coordinate pairs to make_lines(). Now we have a new geometry column containing shapely LineStrings. The achieved DataFrame then needs to be converted into a GeoDataFrame.(see fig.6.22)

id	availability	geometry
0	95.2062	LINESTRING (14.34603282 48.28115323, 14.34589752 48.28159378)
1	93.0731	LINESTRING (14.34589752 48.28159378, 14.34579723 48.28203845)
2	85.7076	LINESTRING (14.34579723 48.28203845, 14.34568596 48.28248195)
3	83.5983	LINESTRING (14.34568596 48.28248195, 14.3455701 48.28292496)

FIGURE 6.22: Geometry line segments

And then all the LineStrings in a Leaflet map are plotted.

On the resulting map, each line segment is symbolized by the availability of connection value using folium's color mapping functionality. It also has a scaled color bar on the top, which rescales from green (100%) to red(0%) the availability value.(see fig.6.23)

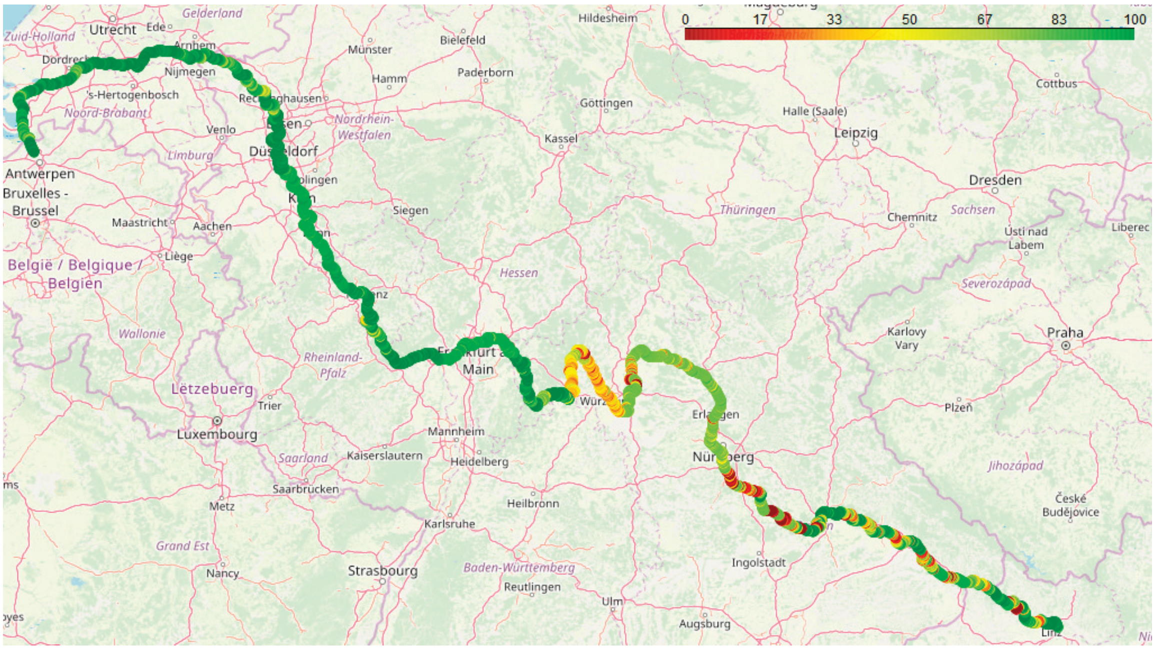


FIGURE 6.23: Availability of connection in the Linz - Antwerpen waterway path in Leaflet map

Chapter 7

Evaluation of the AIS Test Area

As was outlined in the previous chapters, the measurement campaign evaluated in this research was within the LAESSI project. The on-board assistance functions designed to provide reliable assistance to the skipper require accurate, integrity-tested GNSS RTK correction data, as well as relevant waterway information (WSI). Based on these specifications, the requirements for land-based technologies and communication could be derived. One part of the LAESSI project focused on establishing a communication channel for transmitting the high data rate GNSS data and waterway information over a modified AIS shore network. The AIS / VDES communication infrastructure was established on the Main River in an approximately 100 kilometer section. In this chapter, we will validate and evaluate the capacity of the AIS/VDES communication channel. The analysis in this chapter provides information about how the system works on regular trips over this area. The capacity of the channels for transmitting corrections will be evaluated here, and the message error rate of received RTCM and AIS messages will be simulated.

7.1 RTK Corrections Broadcasting via AIS

Due to the requirements of the reliable and stable communication link, in the LAESSI project, five test transmission infrastructure AIS/VDES were established. This area is the test-bed area for checking the availability of communication links. Data streams are in the RTCM 3.1 format, and the RTCM messages contain observations for the GPS and GLONASS satellite systems. The data rate for transmitting messages is 1 Hz and is provided via the AIS/VDES communication channel. Under ideal signal reception conditions, the achievable position accuracy for the user is 1-3 cm. For project-specific purposes, the bandwidth was optimized, non-relevant message types were filtered, and one-to-one assignable station identifiers were assigned in the message headers. The RTCM 10403.x format defines a "transport layer", which should ensure the correct and reliable transmission of data. For this purpose, a frame structure is used, in which the actual message is embedded. This structure consists of a fixed preamble (8 bits), 6 reserved / undefined bits, the message length (10 bits), the data message itself (0-1023 bytes), and a checksum (24 bits). Each individual RTCM message (in particular GPS or GLONASS RTK observations = RTCM messages 1004 and 1012) uses this structure for data transmission. For the integrity marking of the messages, the reserved bits are used in this project.

The successfully received RTCM-3 message is decoded by using five binary messages, which can only give us the final sentence containing the necessary information if the message is completely received.

The message type #8 (Binary Message) is used to send the integrity checked GNSS RTK correction data as well as the waterway information via AIS. The AIS MT #8 is variable in its length, depending on the amount of data to be transferred.

The individual messages should occupy a maximum of 5 slots. Below is a description of the AIS MT format # 8 [32]. The figure 7.1 shows the format used for AIS messages.

Parameter	Number of bits	Description		
Message ID	6	Identifier for Message 8; always 8		
Repeat indicator	2	Used by the repeater to indicate how many times a message has been repeated. See § 4.6.1, Annex 2; 0-3; default = 0; 3 = do not repeat any more		
Source ID	30	MMSI number of source station		
Spare	2	Not used. Should be set to zero. Reserved for future use		
Binary data	Maximum 968	Application identifier	16 bits	Should be as described in § 2.1, Annex 5
		Application data	Maximum 952 bits	Application specific data
Maximum number of bits	Maximum 1 008	Occupies up to 3 slots, or up to 5 slots when able to use FATDMA reservations. For Class B "SO" mobile AIS stations the length of the message should not exceed 3 slots For Class B "CS" mobile AIS stations should not transmit		

FIGURE 7.1: Format description AIS MT # 8:Binary broadcast message [32]

This AIS message provides the framework for any application data and is already being used to send so-called "Application Specific Messages (ASM)".

The GNSS RTK correction data, which is in the RTCM format, is encoded as a complete frame of data, which, because of the amount of data (up to 8232 bits), has made it necessary to split it into multiple ASMs. Variable occupancy of the AIS channels also allows for the more efficient utilization of the available bandwidth. A corresponding software module (client) operating on the in-flight side has also been developed and forms the interface between the communication channel and the driver assistance functions. The GNSS RTK correction data is again provided as a data stream in the RTCM 3 format as the input for the on-board PNT calculation. By contrast, the waterway information is stored as a machine-readable file in an XML format so that it can be used for the display / warning in the driver assistance system[2].

AIS transceivers apply two different frequencies, VHF maritime channels 87B (161.975 MHz) and 88B (162.025 MHz). For this test area, the following were taken as backup frequencies 2024(161.800) and 2025(161.850). These bandwidths were specially chosen only for communication with our equipment. The shifted bandwidth test does not influence the overall AIS communication system with other vessels. The figure 7.2 shows the VDE frequency used for broadcasting AIS messages in the test bed area.

Since 2002 (Class A mandatory) data communication based on AIS has existed. This system has some disadvantages:

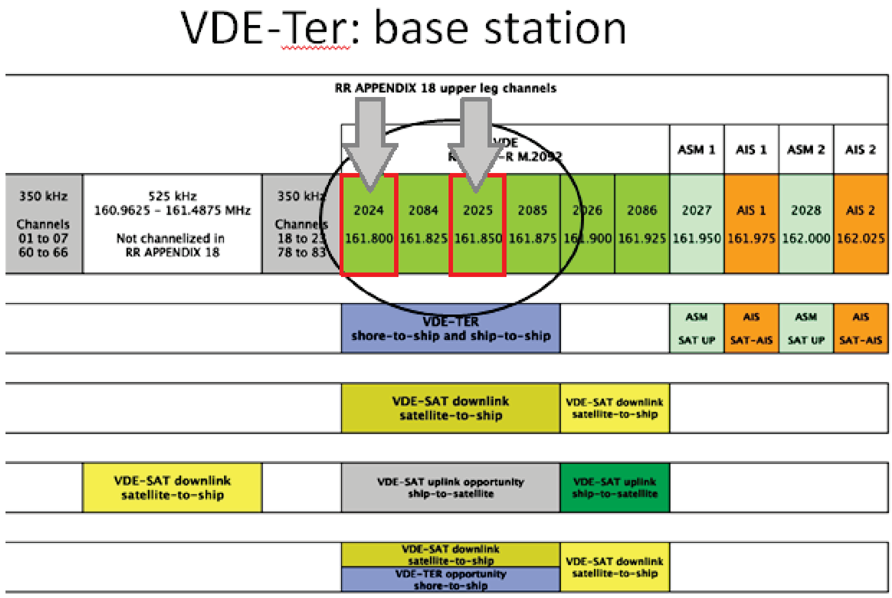


FIGURE 7.2: VDES range

- No error correction
- 25 kHz narrowband
- GMSK modulation (= GSM (2G))

7.2 Test-Bed Area

In one part of the Main river, back up antennas have been used to transmit AIS messages on VDES frequencies. The five selected AIS base stations are located on an approximately 100 km section of the Main between Würzburg and Lengfurt. In addition, a sixth VRS solution was defined in the middle of the measurement area, which provided the correction data via mobile radio (Internet) and was used as a backup system.

The system was designed as follows, the corrections were transmitted over AIS stations. When the Age of Differential GPS data broadcasted over AIS was more than 20 seconds the system switch to the backup GSM connection.

The map, in figure 7.3, shows an overview of the test area.

The data is transmitting in a range of about 7 Kbit/s using an update rate of 1-2 seconds.

7.3 Evaluation of the Test Area Channels Capacity

During the time of the measurement campaign, the vessel passed the test-bed area seven times. In order to identify the coverage of the tested frequency in this area, we checked the connection to the AIS station. The percentage of connection to each station and the coverage of river kilometers are represented in the table 7.1.

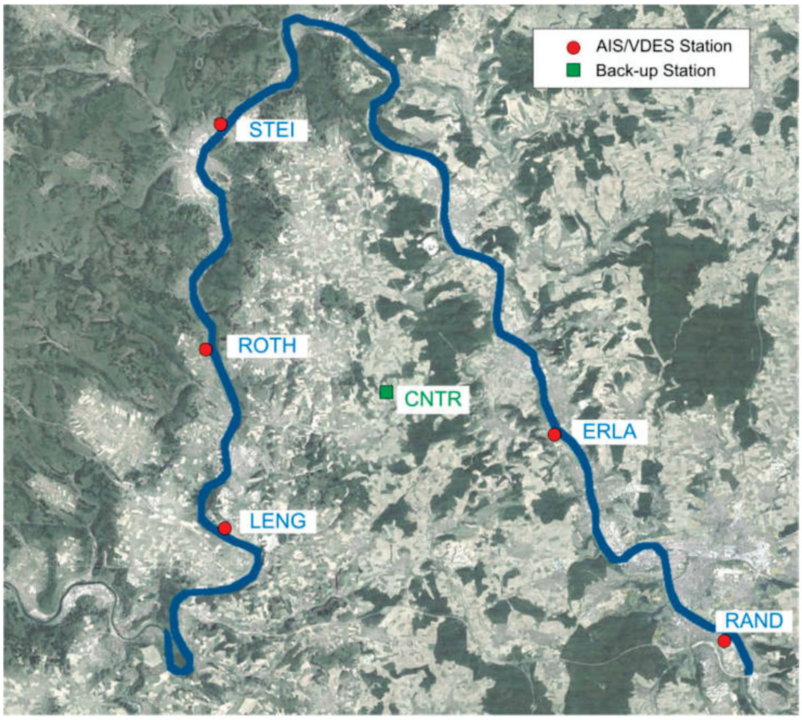


FIGURE 7.3: AIS stations in test area

AIS station	Frequency of connection to AIS stations	Coverage
LENG	4 %	15 km
ROTH	9,2%	15 km
STEI	22,3%	25km
ERLA	17,2%	35km
RAND	6,7%	20 km

TABLE 7.1: Connection to the AIS stations

Throughout this time, the vessel received corrections in almost 60% of cases through AIS stations. Moreover, the signal coverage of each station of river kilometers variates as well. The STEI and ERLA stations have the highest percentage of the frequency of connection to AIS stations and also the biggest coverage river kilometers.

To distinguish the connection in each individual voyage were plotted river kilometers against the date time.

As we can see from figure 7.4, there were 2 voyages where the AIS stations were not receiving correction data. The coverage of broadcasting corrections for each AIS station is stable in each route. The AIS stations did not fully cover the area, and we can see that between the end of one station and the beginning of another station a GSM backup connection is always used. Unfortunately, there are also some data gaps in the measurement campaign that may have arisen for various technical reasons. The lack of data in this part cannot be assessed.

Next, we should check the behavior of the Age of Differential GPS data in each voyage. The age of differential GPS data varies from 0.5 to 60 seconds. However,

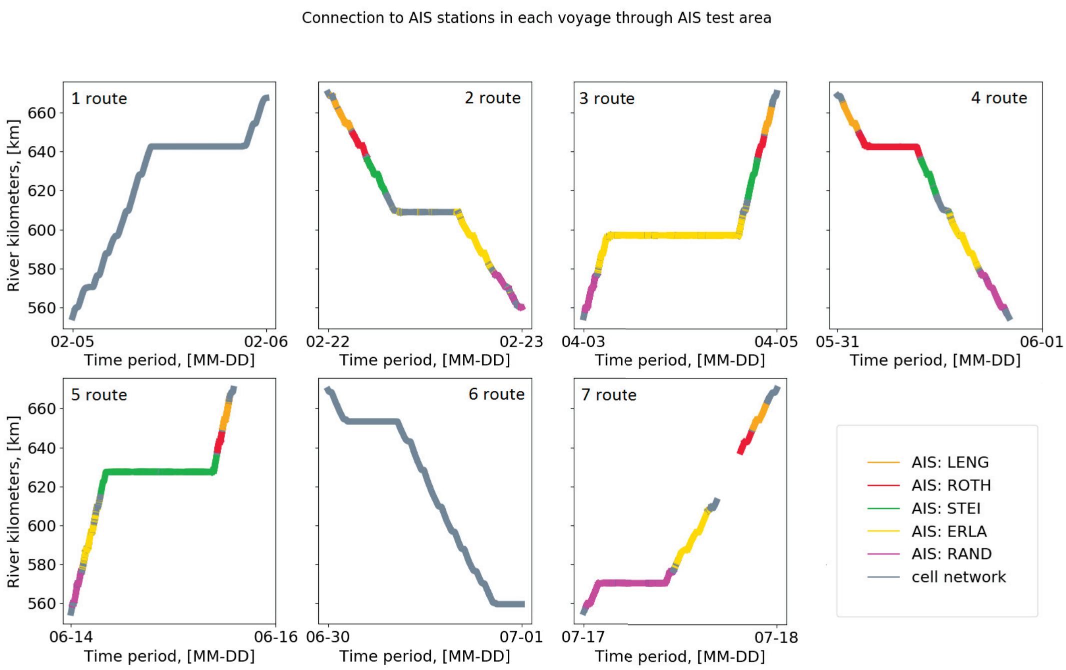


FIGURE 7.4: Connection to the AIS stations in each voyage

when the connection is lost after 60 s, the receiver didn't assign any value to the NMEA message about the time of receiving the correction data. For the purpose of plotting the behavior of the overall Age of Differential GPS data for each individual voyage, missing data were assigned the value of 99 to represent the loss of connection.

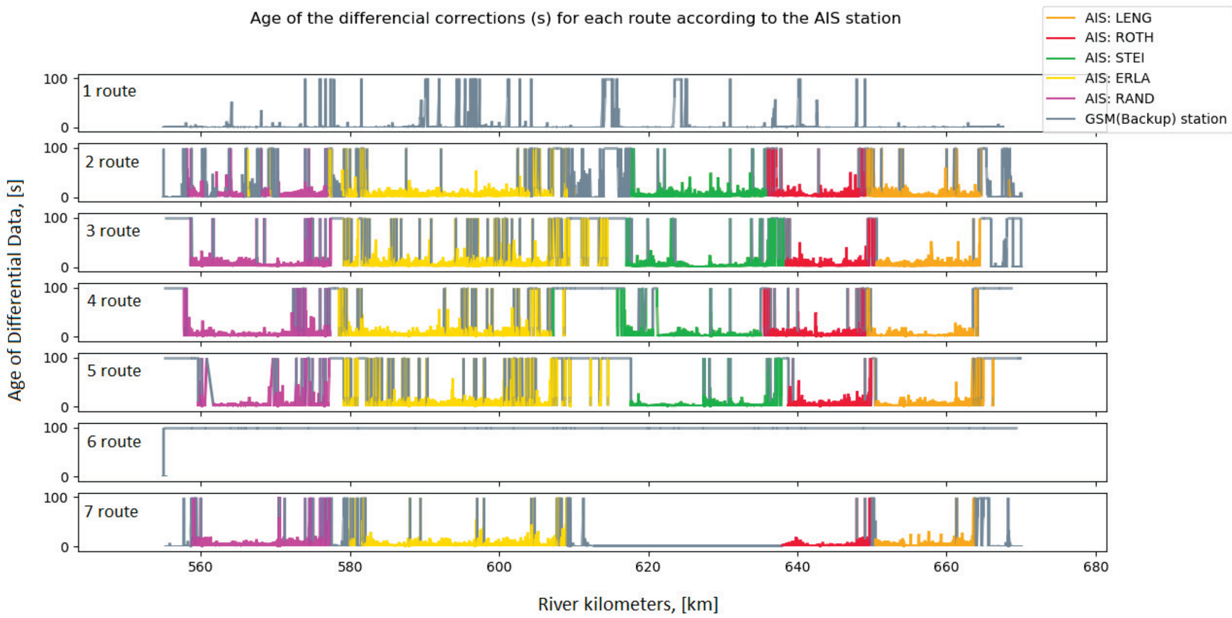


FIGURE 7.5: Connection to the AIS stations

The figure 7.5 describes the age of differential GPS data for each individual voyage.

In figure 7.5, the spikes represent a loss of connection in routes at the same river kilometer value. The first route that uses only a GSM connection shows the best connection, and those that used AIS show the same behavior in all routes. In the sixth route, we have almost no connection, but we cannot rule out that there might problems with the receiver on that voyage as the reason for this outage is not yet entirely understood.

7.4 Estimation of each Station's Capability

We consider that the strength of the signal depends on the vessel's position according to the station. It is obvious that each station has its own coverage area, and as we saw in the previous chapter, each station has its own signal range.

For investigating this assumption, for each point the difference from the current position on the waterway to the used AIS station was calculated. For each station, we plot the mean Age of Differential GPS Data according to the kilometer distance to AIS station. Additionally, the standard deviation for each group of 100 meters was calculated.

The figure 7.6 shows that even in the vicinity of the station, the mean Age of Differential GPS Data is approximately 3 seconds. From -2.5 to -5.0 there are significant jumps in the data. That can be interpreted that with such mean value like 5 seconds we can not have precise positioning. In this place we lost the accuracy of positioning.

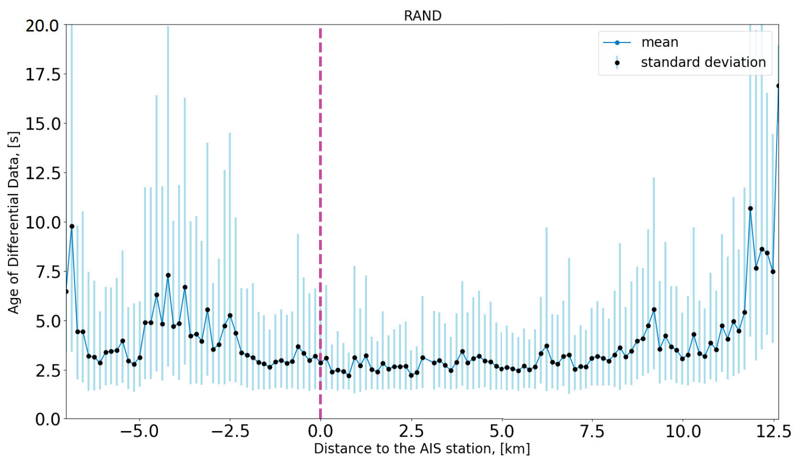


FIGURE 7.6: Connection to the AIS station RAND

The figure 7.7 shows the distribution for ERLA station. Here we can see more jumps in the data even in the close vicinity to the station. It is curious that in the -20 kilometer we have quite good value around 2.5 mean seconds. And from -20 to -30 we still can connect to the ERLA station, but the connection is quite not stable. And the peaks are very deviated from each other.

The ROTH station is presented in figure 7.8, and in the vicinity of the station, around 3 kilometers, the signal is quite stable. But then we can see that further away from the station, the signal losing the power..

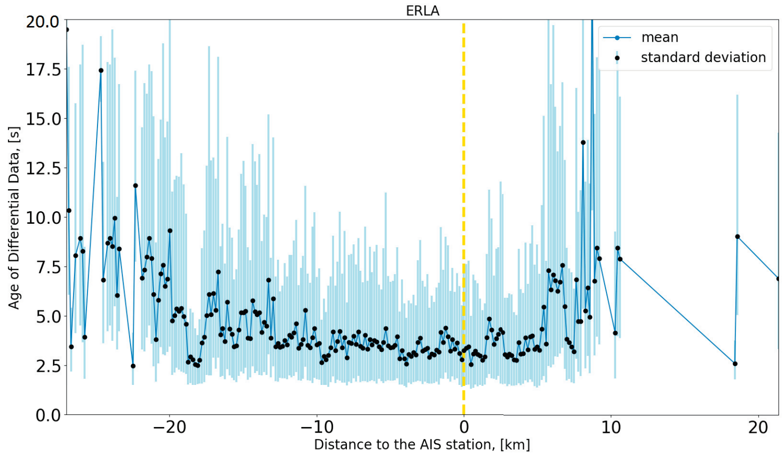


FIGURE 7.7: Connection to the AIS station ERLA

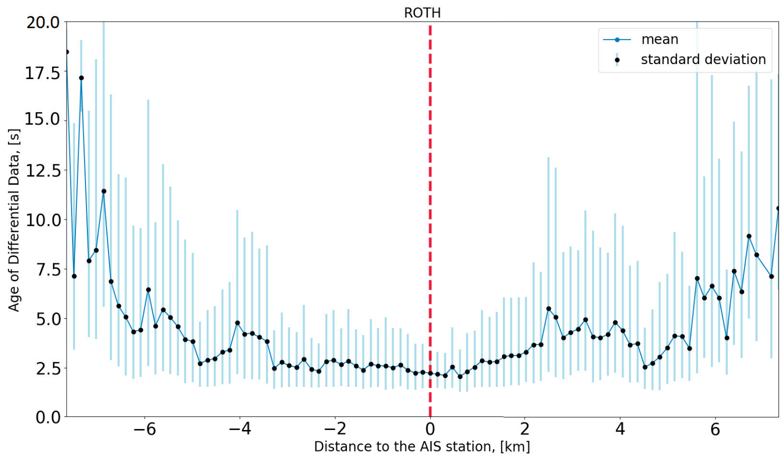


FIGURE 7.8: Connection to the AIS station ROTH

The STEI station, even at the exact position of the station, has jumps in the signal (fig. 7.9). At 5 kilometers away from the station, the mean Age of Differential GPS Data starts growing from 3 to almost 10 seconds.

The distribution of Age of Differential GPS Data in the RAND station is shown in figure 7.10. In the forward direction from the station, there is quite a stable connection, but in another direction we can see that it has reduced. And from -2.5 kilometer the deviation of the data and the mean value is growing.

These results show that our assumption that the connection depends on the distance to the station is proven correct, and that the average Age of Differential corrections is higher in areas more distant from the AIS stations.

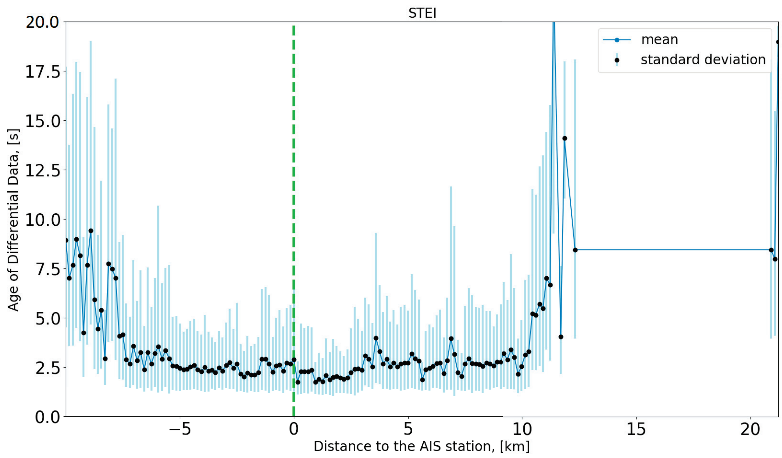


FIGURE 7.9: Connection to the AIS station STEI

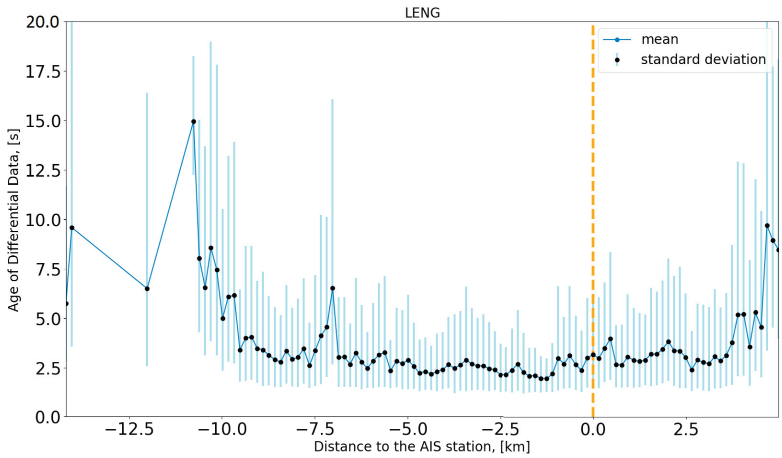


FIGURE 7.10: Connection to the AIS station LENG

7.5 Evaluation of the Transmission of AIS messages

According to previous results, we can assume that in the range from 1 to 1.5 seconds, we successfully received messages from the AIS station. When this parameter grows, we lose the communication channel. This factor can be used to estimate the percentage of successfully received RTCM-3 messages. Figure 7.11 shows the general approach that was used to get the percentage of messages received. This approach uses the average age of the received corrections. Table 7.2 shows an example of the average transmission time percentages of RTCM-3 sentences.

According to the calculated average for each time period, we got the following dependency, which is shown in the figure 7.12. This curve represents the analytic behavior of the percentage of successfully received messages according to the age of the received RTCM-3 message. We see that after about 3 seconds the probability of receiving decoded RTCM-3 sentences increases very quickly to 0 %.

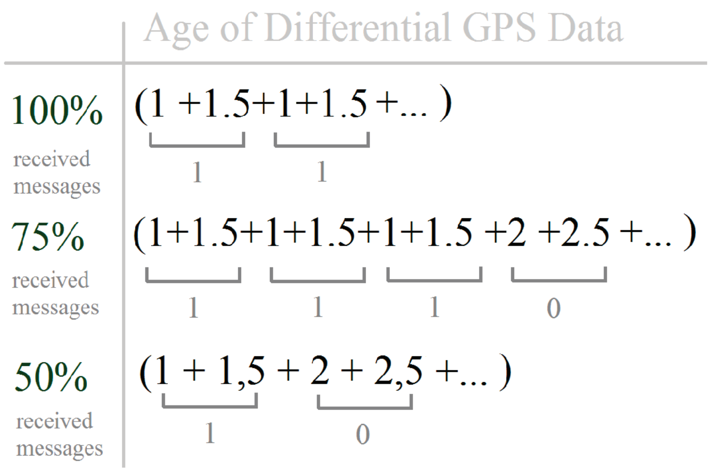


FIGURE 7.11: Percentage of successfully received messages according to the age of received corrections

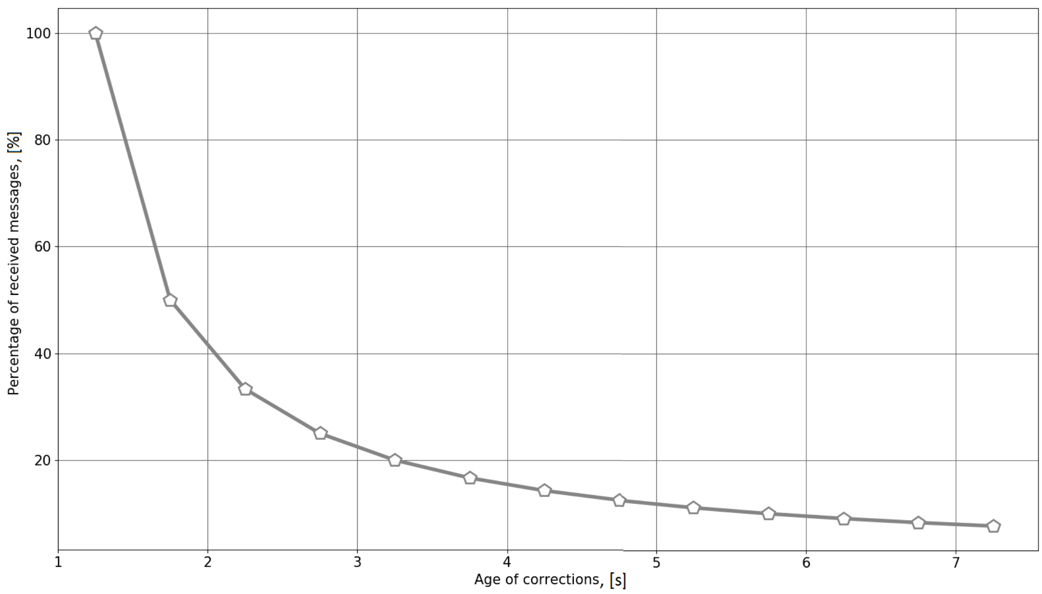


FIGURE 7.12: Percentage of successfully received messages according to the age of received corrections

	Mean Age of Differential Data	Successfully received,%	Range of messages
1	1.2 s	100%	all are received
2	1.75 s	50%	2 received out of 4
3	2.25 s	25%	1 received out of 4

TABLE 7.2: Transmission RTCM-3 sentences

We want to check what is the capability of the communication channel to broadcast the AIS messages and then achieve RTCM-3 sentences. According to this a function that describes the behavior of the distribution of the dependency of the successfully received messages percentage according to the age of received corrections was discovered. We suggest that according to the observation data, the next equation gives us the percentage of successfully received messages according to the mean Age of Differential Data in the appropriate river range:

$$y = \frac{1}{2 * x - 1.5} * 100\% \tag{7.1}$$

To obtain the next equation, an expression was first derived for the growing mean time delay. And then for the appropriate time case, the probability of the percentage of successfully received RTCM-3 sentences was calculated. Derivation of the equation is as follows:

$$x = 0.5 * z + 0.75 \tag{7.2}$$

where, z is the number of the case between two-time epochs for receiving RTCM-3 message.

$$x = 0.5 * \frac{100}{y} + 0.75 \tag{7.3}$$

$$2x = \frac{100}{y} + 1.5 \tag{7.4}$$

$$y = \frac{100}{2 * x - 1.5} \tag{7.5}$$

7.6 Simulation of the Message Error Rate

As far as we know, due to the description of the setup system 5, broadcasted messages can be decoded to one complete RTCM-3 sentence. When we have some loss in the connection to the broadcasted station and lost at least one or more AIS messages, the receiver can not demodulate any more the RTCM-3 sentences. In this case, it is important to investigate the dependence of the error in receiving a complete RTCM-3 sentence due to AIS messages. The background for this simulation is the process of broadcasting AIS messages and decoding RTCM-3 sentences. In this simulation, we used the binomial distribution.

The simulation depends on the number of random samples and the repetitions of the simulation. The experimental probability gets closer and closer to the theoretical probability as we conduct more and more experiments. This is often referred to as The Law of Large Numbers. If we only have a few experiments, it's very possible that our experimental probability could be different from our theoretical probability or even very different. But as we have many more experiments, thousands or more of experiments, the probability that the experimental and the theoretical probabilities are very different, goes down dramatically.

$$P \mathrel{\hat{=}} \text{Probability of AIS loss messages} \tag{7.6}$$

$$1 - P \hat{=} \text{Probability of AIS correctly received messages} \tag{7.7}$$

Each trial can be clearly classified as either a success or failure. Each trial clearly has one of two discrete outcomes. The equation 7.6 represents loss of AIS messages and the equation 7.7 represents the probability of correctly received AIS messages

This approach didn't depend on the amount of broadcasted AIS messages, because with different amounts of transmitted AIS messages, but with the same percentage of loss, the decoded RTCM-3 sentence probability can highly vary. With fewer AIS messages in a worse case, the probability of decoding an RTCM-3 is extremely low. The worst case is when every 5th message is lost, then no RTCM-3 sentence can be decoded. The best case is when we lose the last consecutive messages. For example if we have 1000 AIS messages then we can decode 200 RTCM-3 sentences. But let's suppose that we have lost 20 % of these messages then in the best case we still can achieve 160 sentences when we have lost last consecutive messages. But if this error spread on every 5th message then we will receive 0 sentences. So the probability is spread between 80% and 0%.

For the simulation of the dependency of the message error rate, the binomial distribution were used.

The Binomial Distribution is represented by equation 7.8,

$$P(E) = \binom{n}{k} p^k (1-p)^{n-k} \tag{7.8}$$

where,
n - number of trials
k - number of success/fail (depend on a definition)

When we apply the equation 7.8 to our case we will achieve the next equation:

$$E(x) = \#RTCM3 * (1-p)^5 \tag{7.9}$$

After implementing equation 7.8, the error curve that is shown in figure 7.13 was discovered.

A model that is shown in the figure represents how the error in broadcasting the AIS messages impacts the quantity of successfully decoded RTCM-3 sentences. We also can see how is change the dependency of the chosen amount of required slots for decode complete RTCM-3 message.

The behavior of the curve will depend on how many AIS messages are required for one RTCM-3 sentence. The figures 7.14 and 7.15 shows different dependencies according to the required number of slots for completing one RTCM-3 sentence. This method represents a useful tool for investigating the capability of this communication channel approach.

Next, we can investigate the dependency of the AIS message rate according to the bit error rate. For this, the equation 7.8 with a power of 1008 was used. We achieve the next dependency represented in figure 7.16.

As we can see, the 0.01% bit error rate already causes an AIS message error rate of 10%. And the 0.1% bit error rate leads to almost 100 % of the AIS message error.

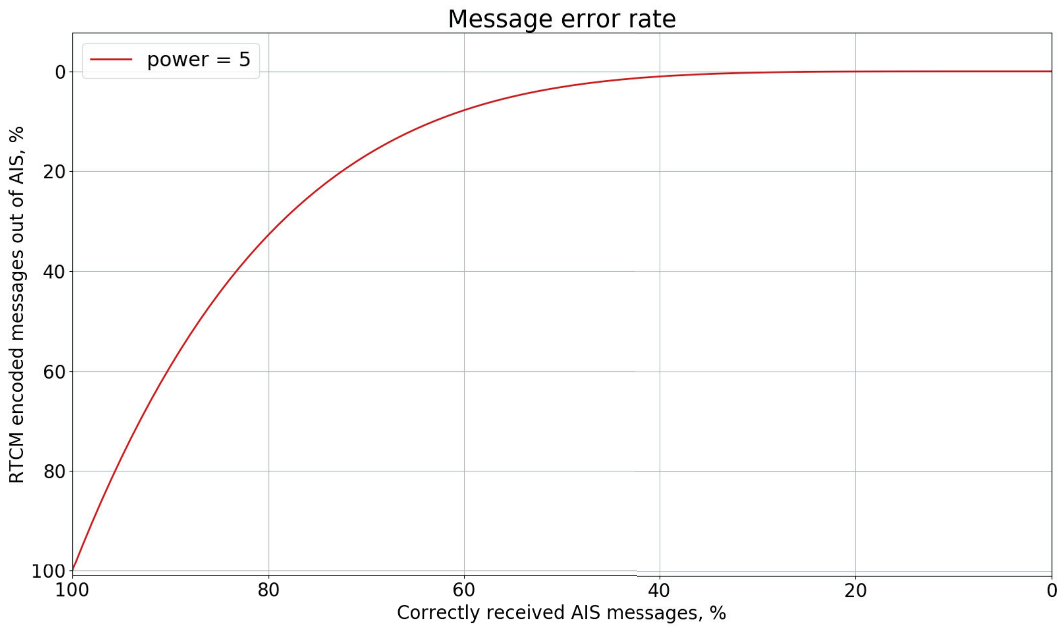


FIGURE 7.13: Message error rate with required 5 slots for one RTCM-3 sentence

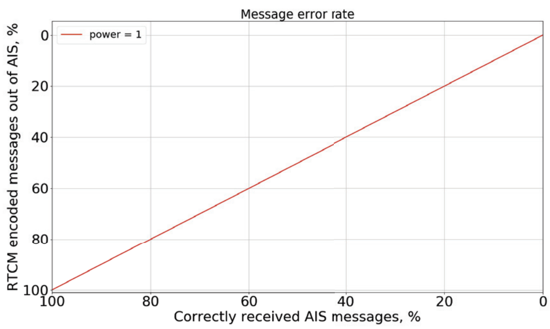


FIGURE 7.14: Message error rate with required 1 slot for one RTCM-3 sentence

FIGURE 7.15: Message error rate with required 12 slots for one RTCM-3 sentence

The bit error rate is the primary error that causes all next errors in decoding AIS and RTCM-3 messages.

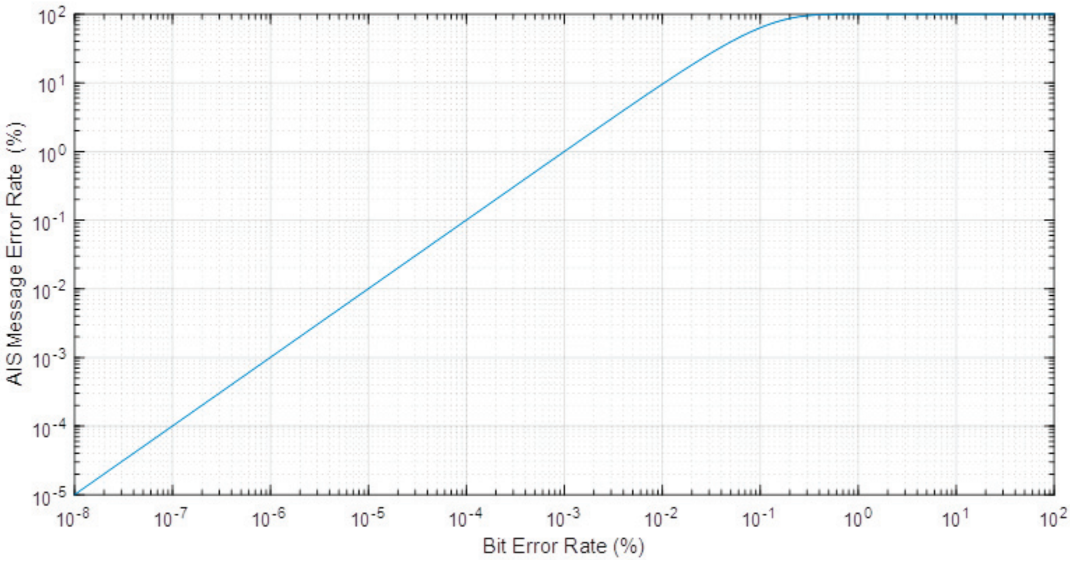


FIGURE 7.16: Message error rate in AIS message according to achieved bits

7.7 Analysis of each AIS station Communication Capability

To analyze the communication capabilities for each station, the mean and standard deviation of the age of the differential corrections were used. Also, the range of received AIS and RTCM-3 messages was calculated for each station.

The figure 7.17 shows the communication capability for the AIS ERLA station.

We can see that the mean age of differential data is around 5 seconds in the vicinity of the AIS station. This value increases with distance from the station. The capability to receive AIS messages is around 70%, and around 20% of the received AIS messages are available to be decoded RTCM messages. The capability of the communication link is not sufficient to be implemented, and the strength of the link should be improved.

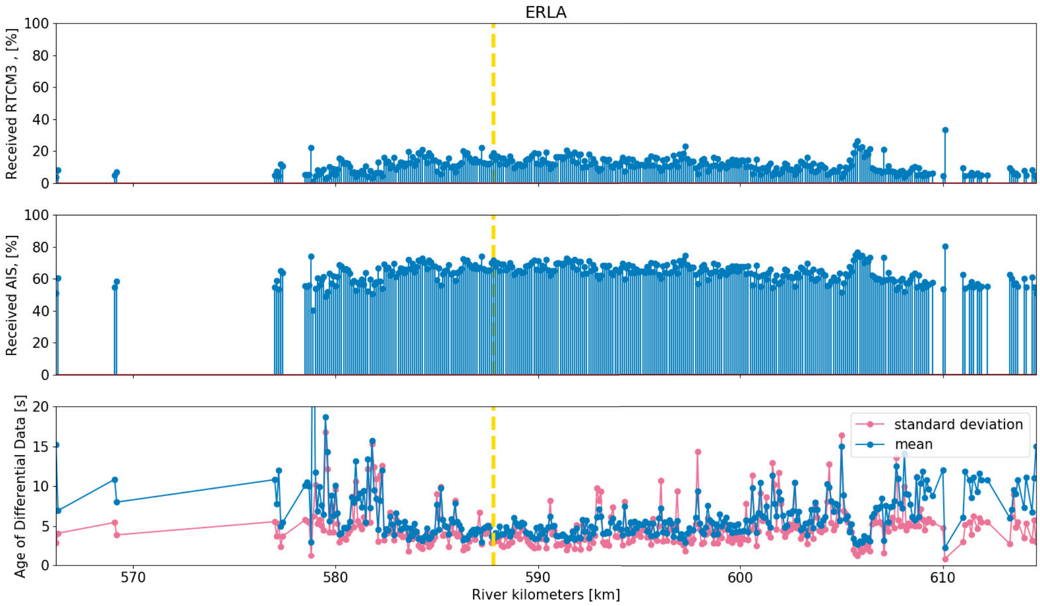


FIGURE 7.17: Capability of AIS station ERLA (station location indicated with dashed line)

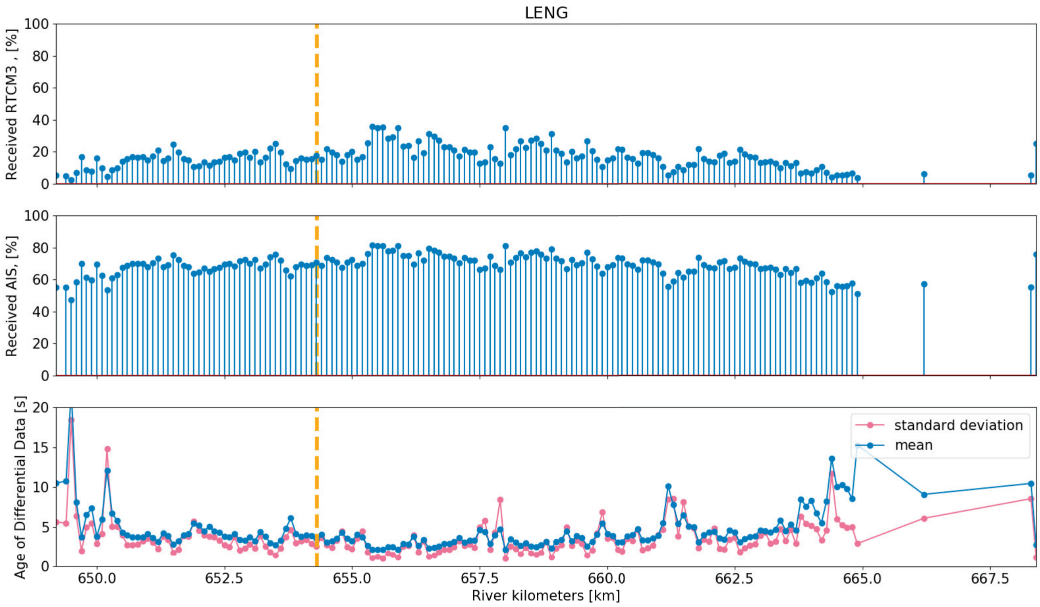


FIGURE 7.18: Capability of AIS station LENG (station location indicated with dashed line)

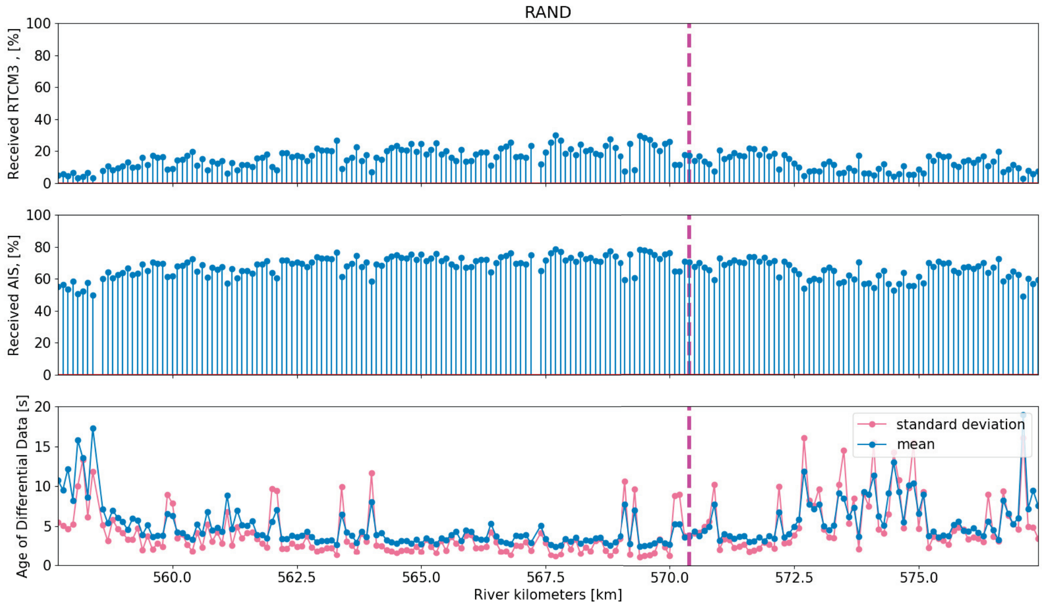


FIGURE 7.19: Capability of AIS station RAND (station location indicated with dashed line)

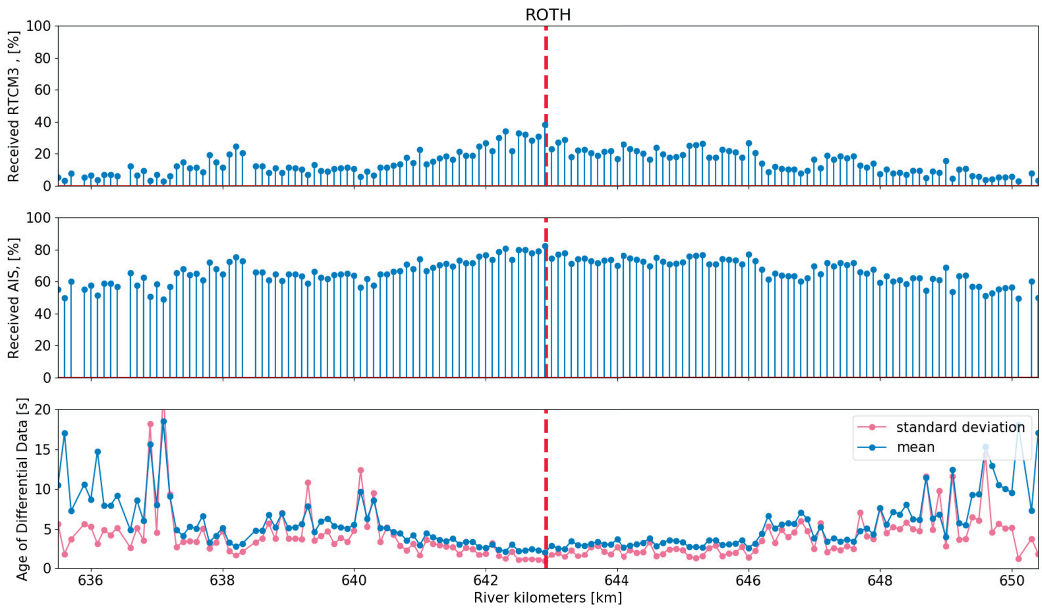


FIGURE 7.20: Capability of AIS station ROTH (station location indicated with dashed line)

FIGURE 7.21: Capability of AIS station STEI (station location indicated with dashed line)

7.8 Bit Rate Error in each AIS Station

The primary cause of errors in successfully achieving RTK corrections is due to the bit transmission.

Below, the figures 7.22 - 7.26 represent the distribution of the bit error rate for each station. We can see that even 0.2% of bit error cause around 80 % of AIS messages to be received and then out of that only 20% of RTCM3 messages are decoded.

We can see that all stations have the same range of the bit error rate (0.1 - 0.3)%.

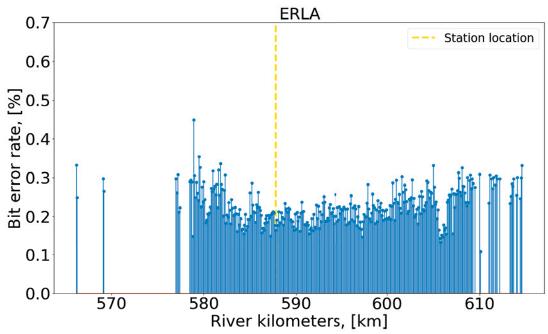


FIGURE 7.22: Percentage of bit error, ERLA station

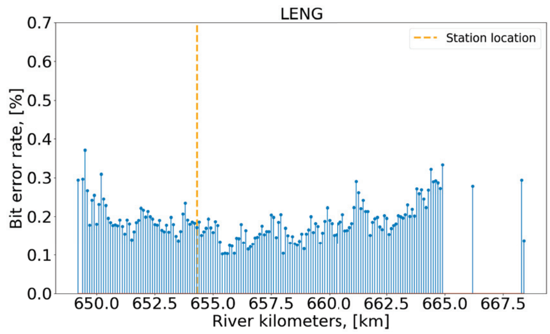


FIGURE 7.23: Percentage of bit error, LENG station

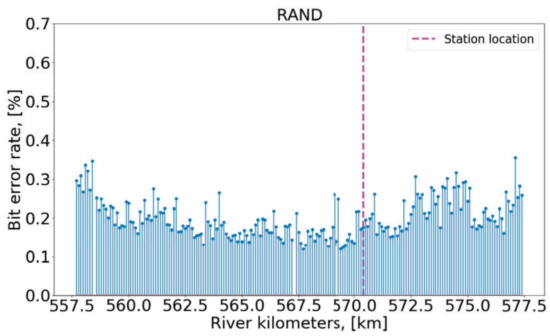


FIGURE 7.24: Percentage of bit error, RAND station

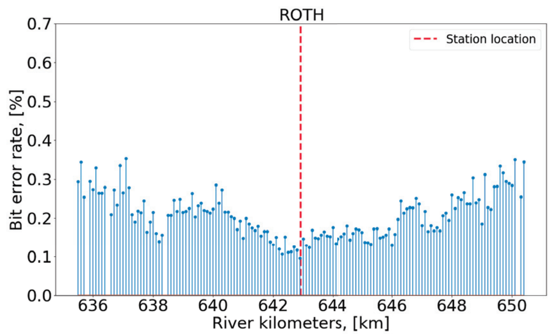


FIGURE 7.25: Percentage of bit error, ROTH station

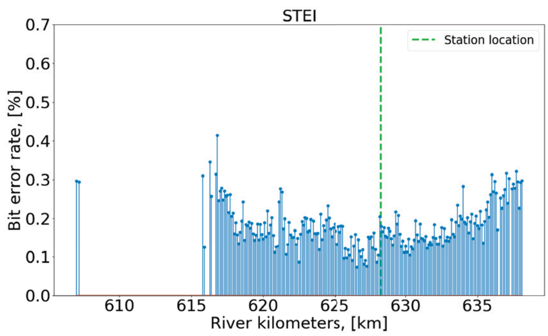


FIGURE 7.26: Percentage of bit error, STEI station

Chapter 8

Summary and Conclusions

In this investigation, precise RTK positioning in the main Europe waterway Rhine–Main–Danube Canal was evaluated. This research emphasizes the importance of improving communication channels due to upcoming automatization and autonomous navigation. In this work, two communication channels, GSM and AIS, were evaluated. For this investigation, a one-year measurement campaign within LAESSI project was used. The measurement campaign used a Trimble BX982 GNSS receiver, Switch in parallel with a mobile phone router, PC, AIS / VDES Tranceiver, and two GNSS antennas. These were established in the wheelhouse of the vessel. The navigation messages during the measurement campaign were stored in a daily SQLite3 database. We have devised a methodology that allows us to evaluate and manipulate big data. First, data was extracted from the each daily database and the GGA(Global Positioning System Fix Data), HDT(Heading true), and VTG(Navigation data) NMEA messages were decoded and stored in a CSV(Comma-separated values) format. Then the one-year measurement campaign was visualized with the QGIS Python Plugin - TimeManager and FFmpeg multimedia framework. For the evaluation, one reference trajectory, which was one continuous route from Linz to Antwerpen, was utilized, the total length of which was 1392,68 km. The river kilometers and distances were recalculated according to the reference trajectory. To accomplish this, an algorithm was developed using the two nearest points with known river kilometers on the reference trajectory to one of the points on the measurement campaigns. Then, considering the three points as a triangle, the measured river kilometer and distance were mapped to the reference trajectory. After that, all data outside a buffer of 150 meters around the reference trajectory were removed to ignore data from harbors and small channels and only consider data in the main waterway. Additionally, as it was discovered the receiver was occasionally switched off(), this data were eliminated from the estimation of the overall internet availability. Finally, an algorithm for splitting voyages was developed based on the behavior of river kilometers and the monotonic function.

The evaluation of communication channels was based on the Age of Differential Corrections. The Age of Differential GPS Data is computed by subtracting the time when the correction message has been generated from the current receiver time. The update rate of corrections is in range from 1 HZ to 2 Hz and corresponds the data range between 0.5 s to 1.5 s. In case of failure of Internet connection the correction data age will be out of this range. To calculate internet availability, the overall waterway pass was divided into 100 meters chunks and the mean value of the internet availability based on the Age of Differential Corrections was determined, taking into consideration the frequency of the vessel passing this 100-meter chunk. As a result, our study provides a map of the availability of GSM connection in the main Europe inland waterway. Additionally, different communication specifications were investigated. First, it was checked whether the connection depends on whether the vessel is travelling downstream or upstream. To determine this, two-directional voyages were taken, and the availability of the connection was cross-correlated. This investigation showed that the connection doesn't depend on the direction of the vehicle route, and it is quite stable in the vicinity of the width of the river, which can be up to 600 meters.

The assessment of the availability of the connection shows us there is a dependency on the connection according to different environmental scenarios. In this research, we investigate the influence of waterway infrastructure such as bridges and waterway locks. In our investigated waterway pass, 275 bridges and 59 waterway locks were accounted for. The intersection of this infrastructure with the internet connection shows in most cases it has a significant negative influence on the connection. This was determined by checking the mean values of availability of internet connection in the 100 meter chunks according to the environmental scenario in this area. The mean value of internet availability in the bridge area is 48.66%, in the waterway infrastructure - 41.6%, in the open sky area - 58.2%. The influence of the bridges on the connection shows it varies according to the type of bridge structure and the width of the bridge. The bridges with a width of 35 -75 meters have a range of lost connection 10 - 40%, and those with 20-35 meters width have less than 10% loss of connection. Bridges with less than 25 meters have almost no influence on the connection. The influence on the connection of the waterway locks varies according to the falling height. This investigation has some overlapping ranges due to the huge dimensions of the infrastructure and because sometimes there is also a bridge near the waterway lock, which has a special effect on the connection. The following statistics were obtained: the waterway locks with a falling height from 3 to 4 meters have a negative influence on the connection of only 0-6%. A falling height of 3 - 10 meters can decrease by 8 - 32% the connection, and when the falling height is 4 - 25 meters, the influence on the connection is 35-50%. Also, the poorly connected areas were examined in greater detail by investigating the effects mountains and hills are likely to have. Finally, the connection is better near the towns than in rural areas.

In another part of this master thesis, a specific test bed area was investigated where the corrections where transmitted over five AIS base stations in a 100 km section between Würzburg and Lengfurt. In these station, the availability of the backup VDES frequencies, 2024(161.800 MHz) and 2025(161.850 MHz), were tested to transmit RTK corrections. For the transmission of correction messages, the RTCM-3 format, which was decoded out of 5 AIS messages type#8 (Binary Message), was used. The total length of the AIS message was 1008 bits. This is the first study of test AIS areas with such a setup established within the LAESSI project. The vessel passed this special area 7 times and has the following frequencies of connection to this AIS stations: LENG - 4%, ROTH - 9.2%, STEI - 22.3%, ERLA - 17.2%, RAND - 6.7%. The vessel received 60% of the corrections through AIS stations. In the other cases, due to the specification of the applied algorithm for the connection, GSM was used as a backup after more than 20 seconds of Age of Differential GPS Data. The evaluation of each station shows different coverage of transmitted corrections over AIS, and it varies from 15 to 35 kilometers. The estimation of each stations capability of transmission corrections shows, that even in the vicinity of the station, the Age of Differential GPS data is approximately 2.5 seconds, which can not satisfy the requirements for precise positioning. By increasing the distance from the station more than 5 kilometers, the Age of Differential GPS data increases significantly.

According to the Age of Differential GPS data, we were able to recalculate the received RTCM-3 messages. According to technical specifications that the measurement rate is higher than the data transmitting rate, we assume that from 1 to 1.5 seconds we successfully receive one RTCM-3 message. Out of the mean Age of Differential GPS data, we recalculate the percentage of received RTCM-3 messages. Then, out of the simulated message error rate, we recalculated the percentage of received AIS messages and the bit error rate. Analysis of each AIS station's capability shows the following performance: the bit error rate for all stations varies in the range of 0.1 - 0.3%, which causes that we can successfully receive only 70% of AIS messages. In the end, the availability to decode RTCM messages out of the received AIS messages varies around 20%. The transmission of AIS messages over AIS stations is sufficiently executed, but this setup does not satisfactorily allow for the decoding of RTCM-3 messages from the received AIS messages.

An AIS test area with this connection capability cannot satisfy the requirements for RTK transmission correction. This study shows communication channels need to be improved. Also, in the existing test setup, there is no error correction, thus the redundancy check can significantly improve the system.

The GSM communication channel, in contrast to AIS, was not made for inland navigation. The map shows the investigated coverage of GSM connection is not good in all areas. There are a lot of blind zones without any connection. This is crucial to knowledge in case of navigation under bridges and in waterway locks, where the high accuracy is especially necessary for autonomous navigation and driver assistance functions.

Chapter 9

Future work

The results gained in this thesis indicate the GSM and AIS communication links capability should be improved. Due to the specifications of new projects to be implemented for inland navigation, the communication channels should suit the different application requirements. The investigated communication channels capability can be used in the upcoming project SciPPPer [24]. The aim of the joint project SciPPPer is the development of a driver assistance system for the automation of the lock movement of inland waterway vessels.

The technological basis for the new driver assistance function will require, first of all, a reliable communication link. The necessary correction data will be made available via the innovative VHF Data Exchange Systems (VDES). The VDE data channel is part of the new VHF Data Exchange System (VDES), which is currently undergoing standardization. The advantages of this system will be its protection against errors bits, capability to work with more frequencies, and availability worldwide.

The future development of the communication system requires the execution of additional measurement campaigns especially in the water areas where autonomous vessels will potentially cruise. In addition to the Age of Corrections a direct measurement of the communication channel characteristic like signal strength would be beneficial.

Appendix A

Python libraries

Essential python libraries that where used for the applications:

- **Datetime**: implements a number of functions to deal with dates, times and time intervals. Whenever is needed to manipulate dates or time, is required to import datetime function.
- **Folium**: it enables easy visualize data that has been manipulated in Python on an interactive leaflet map. It allows both the binding of data to a map for choropleth visualizations as well as passing rich vector/raster/HTML visualizations as markers on the map. The library has plenty of built-in tilesets from OpenStreetMap, Mapbox, and Stamen, and encourages custom tilesets with Mapbox or Cloudmade API keys. Folium supports both Image, Video, GeoJSON, and TopoJSON overlays.
- **Geopandas**:GeoPandas enables adding a spatial geometry data type to Pandas and allows performing spatial operations on these types, using shapely. With this library, it is possible to generate intermediate GIS files and plots with GeoPandas and then shift over to QGIS.
- **Math**: produces access to some common maths functions and constants in Python, which can be used during code for more complex mathematical computations.
- **Matplotlib**: is one of the most popular Python libraries for building plots and other 2D data visualizations. The plots are interactive; it has zoom in option on a section of the plot and a pan around the plot using the toolbar in the plot window.
- **Numpy**: abbreviation for Numerical Python is the foundational package for scientific computing in Python.
- **Pandas**: provide rich data structures and functions created to make working with structured data fast, easy, and expressive. The primary object in pandas that was used in this work is the DataFrame. It is a two- dimensional tabular, column-oriented data structure with both row and column labels. Pandas allow flexible data manipulation capabilities of spreadsheets and relational databases (such as SQL). It presents sophisticated indexing functionality to make it simple to reshape, slice and dice, perform aggregations, and select subsets of data. Pandas is a primary tool that was used in this investigation for data analysis.

- **Pynmea2**: is a python library for the NMEA 0183 protocol. This library allows parsing individual NMEA sentences using the parse(data, check=False) function, which takes a string containing an NMEA 0183 sentence and returns an NMEASentence object.

- **Scipy**: is a scientific library for Python. This library includes modules for statistics, optimization, integration, linear algebra, Fourier transforms, signal and image processing, and more. SciPy depends on NumPy, which provides convenient and fast N-dimensional array manipulation.

- **Shapely**: allows to create shapely geometry objects (e.g. Point, Polygon, Multipolygon) and manipulate with them, e.g. buffer, calculate the area or an intersection, etc.

- **Sqlite3**: is a C library that implements a lightweight disk-based database that doesn't demand a separate server process and enables reaching the database using a nonstandard variant of the SQL query language. For using the module, first needed to produce a Connection object that represents the database.

- **Utm**: is a bidirectional UTM-WGS84 converter for python.
  The syntax to convert from WGS84 to UTM:
  `utm.from_latlon(LATITUDE, LONGITUDE)`.
  The return has the form `(EASTING, NORTHING, ZONE NUMBER, ZONE LETTER)`.

  The syntax to convert from UTM to WGS84:
  `utm.to_latlon(EASTING, NORTHING, ZONE NUMBER, ZONE LETTER)`. The return has the form `(LATITUDE, LONGITUDE)`.

  The speed of utm converting library: 0.4 - 0.5 sec.

Appendix B

Overall system setup

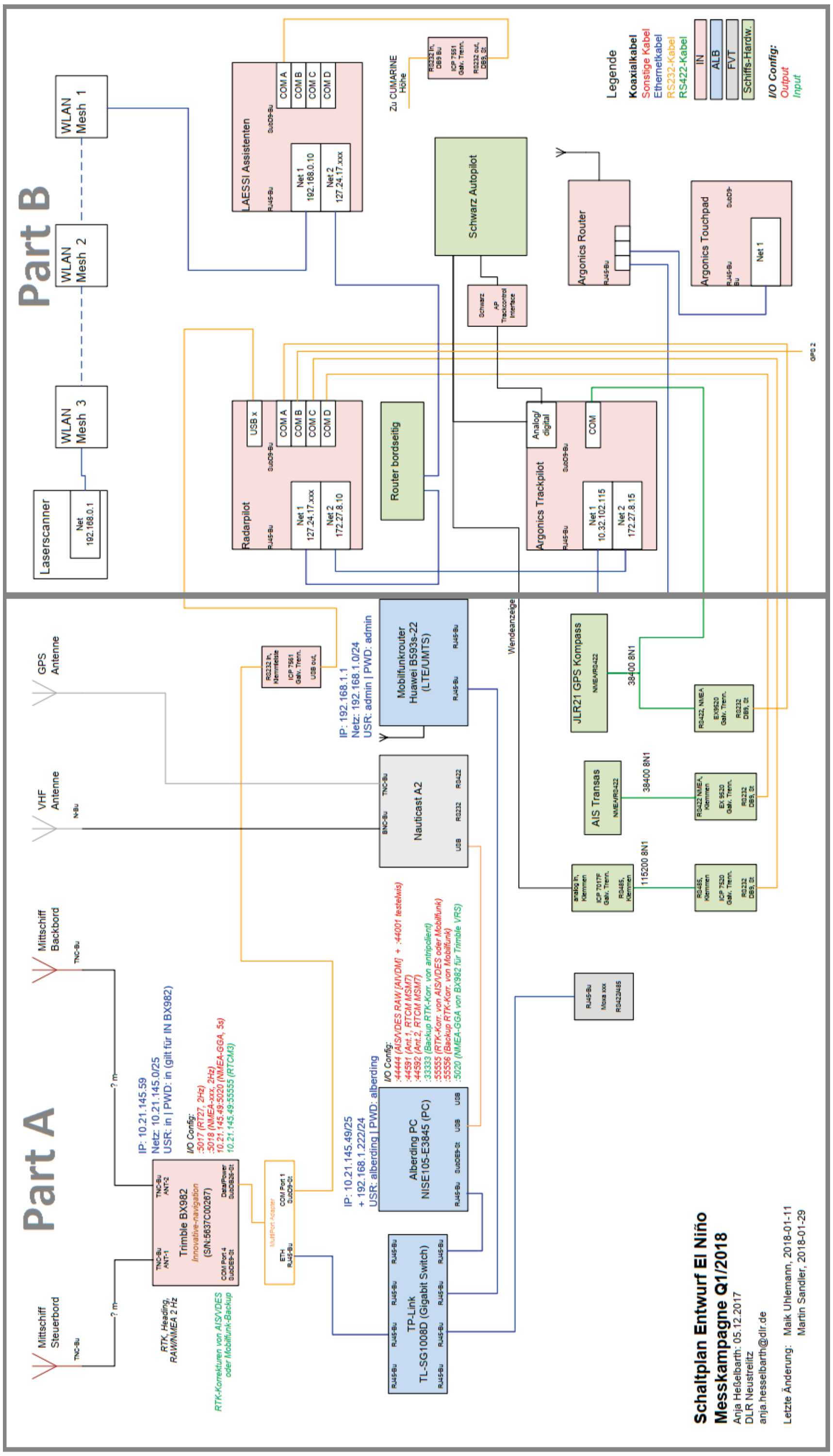


FIGURE B.1: Overall system setup

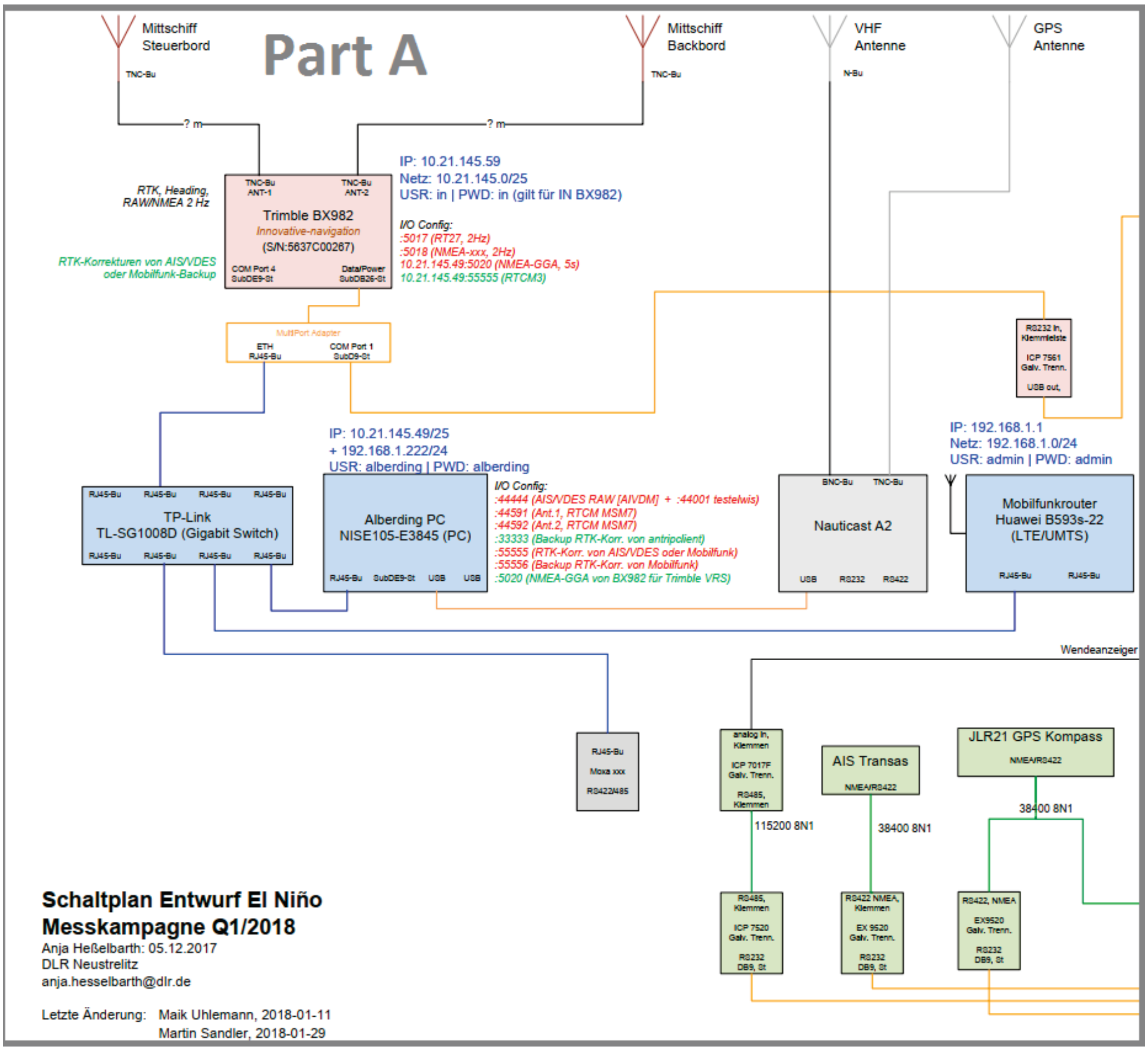


FIGURE B.2: Overall system setup part A

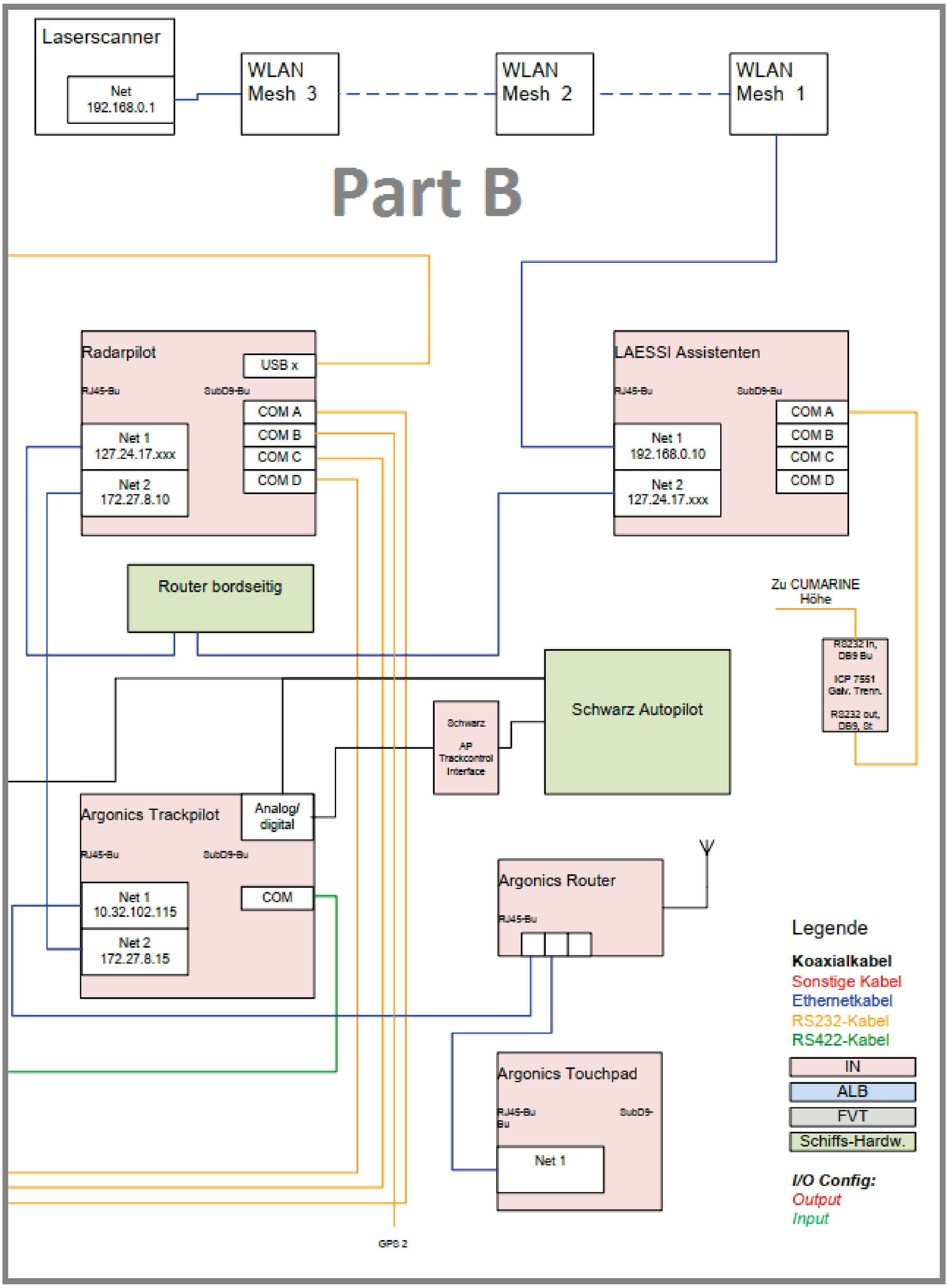


FIGURE B.3: Overall system setup part B

Bibliography

[1] Alberding GmbH 2019. *AutonomSOW*. URL: https://www.autonomsow.de/ (visited on 11/02/2019).

[2] Jürgen Alberding. "Verbundprojekt „Leit- und Assistenzsysteme zur Erhöhung der Sicherheit der Schifffahrt auf Inlandwasserstraßen" (LAESSI) Teilvorhaben: Landseitige Dienste und Datenübertragung Schlussbericht". In: (), p. 28.

[3] John D Bossler. *Manual of geospatial science and technology*. Taylor & Francis, 2002.

[4] U.S. Coast Guard Navigation Center. *AIS MESSAGES*. URL: https://www.navcen.uscg.gov/?pageName=AISMessages (visited on 10/23/2019).

[5] DLR Institute of Communications and Navigation. *LAESSI Project Overview*. URL: https://www.dlr.de/kn/en/desktopdefault.aspx/tabid-4309/3222_read-45318 (visited on 09/25/2019).

[6] DLR. *Nautical Systems Department Institute of Communication and Navigation, DLR*. URL: http://www.dlr.de/kn/en/desktopdefault.aspx/tabid-2204/ (visited on 01/27/2016).

[7] DLR. "NMEA 0183 Standard for Interfacing Marine Electronic Devices". In: (November 1, 2008), p. 254.

[8] Ahmed El-Rabbany. *Introduction to GPS: the global positioning system*. Artech house, 2002.

[9] Netherlands Maritime Technology Foundation. *NOVIMAR vessel train*. URL: https://novimar.eu/ (visited on 11/02/2019).

[10] Paul D Groves. *Principles of GNSS, inertial, and multisensor integrated navigation systems*. Artech house, 2013.

[11] Abbas Harati-Mokhtari et al. "Automatic Identification System (AIS): data reliability and human error implications". In: *The Journal of Navigation* 60.3 (2007), pp. 373–389.

[12] Bernhard Hofmann-Wellenhof, Herbert Lichtenegger, and Elmar Wasle. *GNSS global navigation satellite systems: GPS, GLONASS, Galileo, and more*. Springer Science & Business Media, 2007.

[13] IMO. "Resolution A.1106(29) adopted on 2nd December 2015. Revised guidelines for the onboard operational use of shipboard automatic identification systems (AIS)". In: (2015).

[14] Elliott Kaplan and Christopher Hegarty. *Understanding GPS: principles and applications*. Artech house, 2005.

[15] KONGSBERG. *Autonomous shipping*. URL: https://www.kongsberg.com/maritime/support/themes/autonomous-shipping/ (visited on 11/11/2019).

[16] SOLAS Konvention. *International Convention for the Safety of Life at Sea*. 2004.

[17] Khaled Mohamed Abdel Mageed. "Accuracy Evaluation between GPS Virtual Reference Station (VRS) and GPS Real Time Kinamatic (RTK) Techniques". In: *World Appl. Sci. J* 24 (2013), pp. 1154–1162.

[18] Massterly. *What are the benefits of adopting autonomy technology for the maritime industry?* URL: https://shipinsight.com/articles/what-are-the-benefits-of-adopting-autonomy-technology-for-the-maritime-industry (visited on 11/11/2019).

[19] Michael Mayer, Axel Nothnagel, and Bernhard Heck. "Bezugssysteme für GNSS". In: ().

[20] United Dtates Coast Fuard NAVIGATION CENYER. *How AIS works*. 1999. URL: https://www.navcen.uscg.gov/?pageName=AISworks (visited on 10/25/2019).

[21] Bradford W Parkinson et al. *Global positioning system: Theory and applications, Volume II*. American Institute of Aeronautics and Astronautics, 1996.

[22] Lokukaluge P Perera. "Autonomous Ship Navigation under Deep Learning and the challenges in COLREGs". In: *ASME 2018 37th International Conference on Ocean, Offshore and Arctic Engineering*. American Society of Mechanical Engineers. 2018, V11BT12A005–V11BT12A005.

[23] *Project MUNIN*. URL: http://www.unmanned-ship.org/munin/ (visited on 10/23/2019).

[24] *Project SciPPPer*. URL: http://www.scippper.de/index.php/2-uncategorised/5-zielsetzung-scippper (visited on 10/18/2019).

[25] Jung Putz. "Inland Waterway Transport". In: (September 20, 2016), p. 18.

[26] CCNR PRESS RELEASE. *FIRST INTERNATIONAL DEFINITION OF LEVELS OF AUTOMATION IN INLAND NAVIGATION*. 18 December 2018. URL: https://www.ccr-zkr.org/files/documents/cpresse/cp20181219en.pdf (visited on 10/27/2019).

[27] GIS Resources. *How VRS Work?* URL: http://www.gisresources.com/how-vrs-work/ (visited on 11/02/2019).

[28] Central Commission for the Navigation of the Rhine (CCNR). *ANNUAL REPORT 2018 INLAND NAVIGATION IN EUROPE*. Tech. rep. Sept. 2018, pp. 1–156. URL: https://www.inland-navigation-market.org.

[29] RTCM RTCM-SC104. *Recommended Standards for Differential GNSS Service (v. 2.3)*. 2001.

[30] J Sanz Subirana, JM Juan Zornoza, and M Hernández-Pajares. "GNSS Data Processing, Volume I: Fundamentals and Algorithms". In: *ESA Communications, ESTEC, Noordwijk, Netherlands* (2013), pp. 145–161.

[31] SEAFAR. *SEAFAR*. URL: https://www.seafar.eu/ (visited on 11/11/2019).

[32] M Series. *Technical characteristics for an automatic identification system using time-division multiple access in the VHF maritime mobile band*. 2010.

[33] Peter Teunissen and Oliver Montenbruck. *Springer handbook of global navigation satellite systems*. Springer, 2017.

[34] *User Guide: Trimble BX982GNSS Receiver Enclosure*. December 2015.

[35] VBW. *Driving Dynamics of Inland Vessels, Vessel Behaviour on European Inland Waterways and Waterway Infrastructure with Special Respect to German Waterways*. Bundesanstalt für Wasserbau Federal Waterways Engineering and Research Institute, 2016.

[36] Wikipedia contributors. *Automatic identification system — Wikipedia, The Free Encyclopedia*. [Online; accessed 27-October-2019]. 2019. URL: https://en.wikipedia.org/w/index.php?title=Automatic_identification_system&oldid=928050508.

[37] Wikipedia contributors. *River mile — Wikipedia, The Free Encyclopedia*. [Online; accessed 28-October-2019]. 2019. URL: https://en.wikipedia.org/w/index.php?title=River_mile&oldid=917459405.