INSPIRe

WP8 – Assess the feasibility of using opportunistic sources of data to aid DFMC integrity

D 8.1: Crowd-sourced sourced inputs into DFMC integrity, Feasibility report

Prepared for:

Table of Contents

1. Introduction 1

1.1 Context and Objective 1

1.2 Report structure 2

1.3 Abbreviations 3

1.4 Revision History 3

2. Opportunistic data source feasibility assessment 4

2.1 Opportunistic data sources 5

2.1.1 User-level Crowdsourcing 5

2.2.2 Signal-of-opportunity 5

2.2.3 System-level Crowdsourcing 7

2.2.4 Computer Vision Data Sources 8

2.2 Benefits of opportunistic data sources in INSPIRe context 8

2.3 Selecting opportunistic approaches 10

3. System-level crowdsourcing 13

3.1 Data source 15

3.2 Error distributions 15

3.3 Distribution estimation 16

3.4 Goodness-of-fit assessments 18

3.5 Assessments Results 18

3.5.1 Kolmogorov-Smirnov test results 18

3.5.2 Graphical results 20

3.5.3 Availability results 25

3.6 Summary and Conclusion 27

4. User-level Crowdsourcing 28

4.1 Overview 28

4.2 Automatic Identification System 30

4.3 Mathematical model 31

4.4 Imperial College simulation platform 33

4.5 User-level crowd-sourcing integrity performance 36

4.5.1 Evaluation metrices 36
4.5.2 Radar-based crowd-sourcing 37
4.5.3 LIDAR-based crowd-sourcing 39
4.6 Summary 43
5. Crowdsourcing as an e-Navigation Service 47
6. Development and Implementation Plan 52
6.1 System-level crowdsourcing (error characterisation) 52
6.1.1 System-level Adaptive Error Characterisation Framework 52
6.1.2 Mechanism to utilise GEV distribution in the positioning algorithm 52
6.1.3 Logistic Distribution Implementation 53
6.1.4 Carrier Phase Error Characterisation 53
6.1.5 Continuous Improvement 53
6.1.6 Phased Rollout 53
6.1.7 User-level future works 54
6.2 User-level crowd-sources positioning 55
6.2.1 Specify the Lidar specification to support the port phase 55
6.2.2 Define the Lidar failure mode and models 55
6.2.3 Improve the Crowdsourcing Mathematical Model and Simulator Performance 55
6.2.4 Designing a combined set of Lidars based on the '*n*' number of airborne Lidars to measure range precisely. 56
6.2.5 Disseminate the findings to academic, governmental, and industrial stakeholders, including manufacturers 56
7. Exploitation Plan 57
7.1 Overview of the Crowdsourcing Concept 57
7.2 Applications domain 57
7.3 Opportunities for equipment, application, and service providers: 57
8. References 60
9. Appendix 63

1. Introduction

This report presents the findings from an assessment of the feasibility of using opportunistic data sources to enhance DFMC integrity in the maritime sector. This significant research initiative is undertaken as part of Work Package 8 (WP8) in the Integrated Navigation System-of-Systems PNT Integrity for Resilience (INSPIRe) project.

1.1 Context and Objective

WP8 assesses the feasibility of augmenting dual-frequency multi-constellation GNSS integrity monitoring using crowd-sourced integrity data from users. The work focuses on the maritime sector, considering potential expansion into other sectors where integrity is a key performance metric in critical applications. The objectives of WP8 are:

- Assess the feasibility of incorporating crowdsourced user integrity data into the UK-wide dual-frequency, multi-constellation GNSS integrity monitoring system, considering different approaches and taking account of potential drawbacks.
- Produce an outline functional design, including crowd-sourcing algorithms, to incorporate crowdsourcing into the UK-wide dual-frequency, multi-constellation GNSS integrity monitoring system addressed in WP7.
- Develop and run a mock-up model, using a suitable software application, to:
    - Prove the crowd-sourcing concept;
    - Optimise the outline functional design for integrity crowd-sourcing in the UK maritime sector
    - Estimate the potential improvements to the operational performance and coverage of the integrity monitoring system by incorporating crowd-sourcing.
- Develop a conceptual design for an e-Navigation service based on crowd-sourced integrity data organised through the Maritime Connectivity Platform (MCP);
- Produce a development and implementation plan for integrity crowd-sourcing directly for the maritime sector; and explore potential expansion into other sectors and applications with acknowledged requirements for high integrity (such as aviation, rail and connected autonomous vehicles).
- Construct an exploitation plan for the crowd-sourcing concept

1.2 Report structure

This report is structured into seven sections, encompassing the entirety of the work conducted in WP8 (WP8.1 to WP8.7). A visual representation of WP8 activities with the report structure is presented in Figure (1.1). The sections are outlined as follows:

Section 2: Assessing the Feasibility of Opportunistic Data Sources
This section delves into the viability of integrating opportunistic data sources, and evaluating their potential benefits and challenges.

Section 3: System-Level Crowdsourcing
This section utilises crowdsourcing at the system level for error characterisation, the section includes the functional design and associated mathematical models and evaluates the error characterisation using Gaussian, Generalised-t, Generalised Extreme Value (GEV), Logistic, Laplace, and Cauchy distributions.

Section 4: User-Level Crowdsourcing
This section presents the development of the crowdsourcing positioning and integrity layer at the user-level, which is based on computing the position using nearby vessels and range measurements. Performance is evaluated using a simulation platform developed by Imperial College.

Section 5: Crowdsourcing as an e-Navigation Service

This section proposes the implementation of an e-Navigation Service within the Maritime Connectivity Platform (MCP) as a more secure and reliable method for transferring position and integrity information between vessels, compared to the current Automatic Identification System (AIS).

Section 6: Development and Implementation Plan
This section presents a comprehensive plan detailing the steps for developing/improving and implementing crowdsourcing strategies.

Section 7: Exploitation Plan
A strategic plan aimed at maximising the benefits and opportunities arising from the implementation of crowdsourcing.

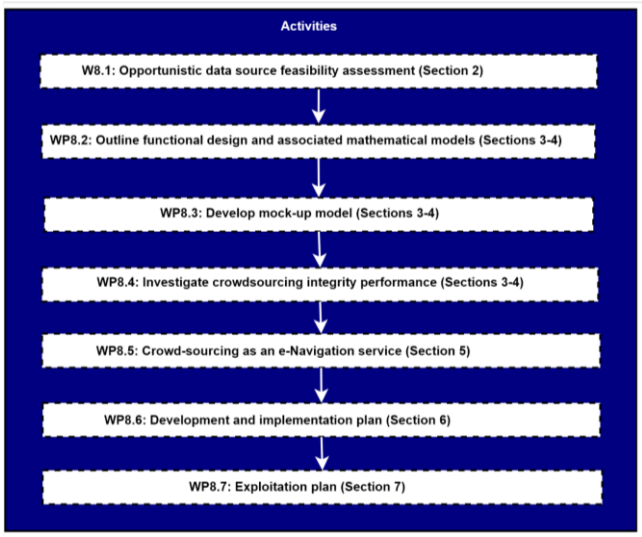


Figure 1.1: WP8 activities with the report structure

1.3 Abbreviations

Abbreviation	Description
AIS	Automatic Identification System
ASF	Additional Secondary Factor
CDF	Cumulative Distribution Function
CORS	Continuously Operating Reference Stations
ECEF	Earth-Centered, Earth-Fixed
eCDF	empirical CDF
eLoran	Enhanced Long Range Navigation
FDE	Failure Detection and Exclusion
GEV	Generalised Extreme Value
GNSS	Global Navigation Satellite System
IPS	Indoor Positioning System
ML	Machine learning
MCP	Maritime Connectivity Platform
MSE	Mean Squared Error
OLS	Ordinary Least Squares
OS	Ordnance Survey
PDF	Probability Density Function
PNT	Positioning, Navigation, and Timing
RMS	Root Mean Square
SPP	Single Point Positioning
SoOP	Signals of Opportunity Positioning
TDoA	Time Difference of Arrival
ToA	Time of Arrival
UWB	Ultra-Wideband

1.4 Revision History

Revision	Author(s)		Date	Section(s)
V0.1	Mamon Alghananim Washington Ochieng	Imperial College	13-11-2023	1,2,3,4,6,7
	Chris Hargreaves	GRAD	13-11-2023	5
V0.2	Mamon Alghananim Washington Ochieng	Imperial College	21-2-2024	1,2,3,4,6,7
	Chris Hargreaves	GRAD	21-2-2024	5

2. Opportunistic data source feasibility assessment

Utilising opportunistic data sources within INSPIRe aims at the first level to support the system in achieving its requirements, especially during the port phase, where GNSS standalone cannot achieve the system requirements, as evidenced in the (INSPIRe-GMVNSL-D4.1-v1.1, 2024; INSPIRe-GMV-D3.1-v1.1, 2023) reports findings. In addition, the opportunistic data sources can enhance system performance, leading to a reduction in the protection level and, subsequently, the required alarm limit, ultimately enhancing maritime operations.

The workflow of the opportunistic data source feasibility assessment, presented in Figure 2.1, has six stages: i) definition of 'opportunistic data sources', ii) identifying relevant opportunistic data sources, iii) describing the data sources, iv) identifying key challenges and limitations, v) highlighting the benefits of opportunistic data sources, and vi) selecting opportunistic approaches.

Starting with the first stage, establishing a clear definition for 'opportunistic data sources' is central to identifying which opportunistic data sources can benefit INSPIRe. The 'opportunistic data sources' can be defined either: as additional PNT (Positioning, Navigation, and Timing) data sources that are not typically utilised as PNT sources within the context of INSPIRe, or as PNT data sources that were not initially designed to serve as PNT sensors (e.g. Bluetooth technology). Both definitions are encompassed in WP8.

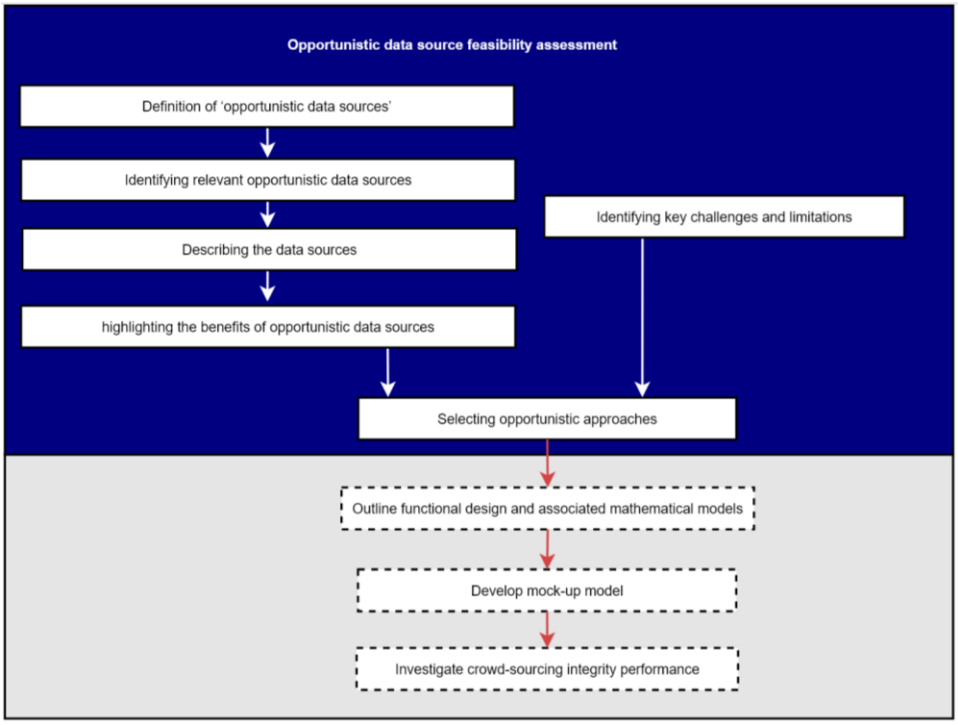


Figure 2.1: The workflow of the opportunistic data source feasibility assessment

The second stage, identifying opportunistic data sources, includes reviewing PNT related literature and products (including applications, services and technologies including those from the research and industry communities). The opportunistic data sources in this report include: Signal-of-opportunity (Indoor position system, eLoran, and AIS shore station), user-level crowdsourcing

(vessels and smartphones), and system-level crowdsourcing (CORS networks). The data sources are described in detail within stage three in Section 2.1.

The fourth and fifth stages, identifying key challenges and limitations in INSPIRe and outlining the benefits of opportunistic data sources enable the selection of approaches used in INSPIRe as discussed in detail in Section 2.3.

2.1 Opportunistic data sources

2.1.1 User-level Crowdsourcing

User-level crowdsourced positioning aims to enhance the accuracy and reliability of positioning services using nearby PNT devices (e.g., GNSS), or to provide positioning information for the users when GNSS is not available. In the maritime sector, the two main crowdsourced data sources are smartphones and nearby vessels/ships. The former can provide position information with a low level of accuracy. In the integrity monitoring domain, Angrisano and Gaglione (2022) evaluated Smartphone GNSS performance in an urban scenario with RAIM application, showing that position RMS error achieved by the smartphone in complex environments after Failure Detection and Exclusion (FDE) varied between 16m – 105m. This study demonstrates that smartphone data cannot provide the required level of performance for INSPIRe, by comparing these results with the performance requirements. In addition, transferring positioning information from nearby smartphones to the vessels is a complex process.

Utilising data from nearby vessels/ships is more reliable than smartphones, as the positioning accuracy of the GNSS devices in vessels/ships is significantly better. Unlike smartphones, the vessels' information is transmitted via Automatic Identification System (AIS) messages, which include positioning and integrity status information. In addition to positioning information, ranging sensors (e.g., Radar) are used to compute the distances between the vessels. Utilising both positioning data and range measurements, the vessels' positions can be computed.

However, to the best of our current knowledge, there are no studies that have evaluated the protection level that can be computed using information from nearby vessels/ships'. Therefore, user-level positioning exploiting crowdsourced data and the corresponding integrity layer are developed and tested in this report in Section 4, presenting in detail the AIS messages and the mathematical models for positioning and integrity monitoring.

2.2.2 Signal-of-opportunity

- **Shore-based Automatic Identification System positioning**

Shore-based Automatic Identification System (AIS) positioning is based on utilising AIS message information from shore stations and measuring Very High Frequency (VHF) radio signals. This technique is based on gauging the transmission delay of the VHF AIS signal emanating from the AIS shore stations. The Additional Secondary Factor (ASF) correction is applied to mitigate any errors in signal propagation. Ultimately, the vessel's location is deduced using a position algorithm that relies on the signal's travel time.

In this context, Hu et al., (2015) developed Shore-based AIS positioning based on the Time Difference of Arrival (TDOA) model, derived form Time of Arrival (ToA). They assessed this approach by utilising ASF correction to calculate the positions of vessels using data from three AIS-based stations. The results show that the RMS error in latitude and longitude were 4.224 and 2.534 meters, respectively. They estimated the position error for dynamic positioning to be 9.851 meters (95%). However, the integrity aspect of this approach is still to evaluated, including computation of protection level and FDE.

- **Enhanced Long Range Navigation**

Enhanced Long Range Navigation (eLoran) is a terrestrial navigation system developed as an alternative to GNSS. eLoran uses low-frequency radio transmitters from multiple locations to deliver navigation services for vessels. Similar to Shore-based AIS positioning, ASF correction can be used to improve eLoran positioning accuracy. Some previous studies have explored using eLoran as an alternative to GNSS (e.g., Johnson et al., 2007; Son et al., 2020); however, eLoran integrity monitoring is still to be explored, with its development facing challenges due to the limited number of eLoran stations.

As it is currently defined, eLoran no longer exists in Europe, where the system was switched off in 2015 (The Maritime Executive, 2016). This deactivation reflects a shift towards satellite-based navigation solutions and a move away from traditional, ground-based systems. Today, only a single Loran station remains operational in Europe, providing a limited scope of services—it offers a timing and data service utilised within the UK. However, there is a resurgence of global interest, including in the UK, in this technology. The UK minister for science, research, and innovation recently announced recently the government's policy framework for greater PNT resilience (Freeman,2023). The Framework includes the development of a proposal for a resilient, terrestrial, and sovereign eLoran system to provide a backup positioning and navigation.

Indoor Positioning System

Indoor Positioning System (IPS) is Signal-of-Opportunity based and can be utilised to support maritime navigation in the port phase. Several IPS technologies have been developed in the last two decades to support GNSS in indoor/complex environments such as Wifi, Bluetooth, RFID, and Ultra-WideBand (UWB). When evaluating IPS technologies, it is crucial to consider the entire end-to-end positioning process. This assessment encompasses factors like infrastructure availability, cost implications, and system requirements including the accuracy, which is influenced by elements such as measurement types, positioning methodology, measurement quantity, network design, node density, and environmental conditions.

Generally speaking, WiFi-based positioning (Li et al., 2016; Talvitie et al;2015) is capable of achieving a moderate accuracy level, ranging from 2 to 10 meters. RFID-based IPS can attain meter-level accuracy in their active mode and accuracy ranging from centimetres to decimetres in passive mode (Chon et al.,2004; Bouet and Dossantos, 2008). However, this technology is hampered by its limited range and substantial implementation costs. Ultra-WideBand (UWB) technology is another option, offering accuracy between centimetres and decimeters, albeit accompanied by significant infrastructure expenses (Krishnan et al.,2007; Alarifi et al, 2016). Bluetooth-based IPS emerges as a cost-effective alternative, providing accuracy from decimeters

to meters (Eyng et al.,2020; Kluge et al.,2020), and incurring considerably lower costs in comparison to UWB and RFID-based IPS. Table (2.1) provides a summary of key IPS technologies, outlining their accuracy levels and key limitations.

In the maritime sector, the deployment of these technologies is feasible and reliable only at the porting level, within a limited range. In addition, the integrity monitoring for these technologies is still to be explored in detail, requiring further research.

Technology	*Accuracy*	*Key limitations*
UWB	cm – dm level (10 -50 cm)	- Infrastructure does not exist. - Metallic and liquid materials cause UWB signal interference - Not very secure
Wifi	meter level (2- 10 m)	- Low accuracy - Not very secure
Bluetooth	meter level (1-3m)	- Low accuracy
RFID / passive	cm to dm level (20 cm – 100 cm)	- Very small coverage small - The RF signal influenced by the antenna. - Cannot be integrated easily with other sensors.

Table 2.1: key IPS technologies accuracy levels and key limitations

2.2.3 System-level Crowdsourcing

The System-level Crowdsources (e.g. CORS network) can be utilised as an opportunistic data source to support system-level integrity monitoring. This includes enhancing error characterisation, which is linked to positioning reliability and performance. In more detail, error characterisation takes place across three phases at the user level in positioning/integrity algorithms, and at two phases within system-level integrity monitoring.

At the system level, error characterisation is defined through Probability Density Functions (PDFs), widely assumed by a Gaussian distribution. Subsequently, these are overbounded and transmitted to users. In addition, error characterisation is used in the Failure Detection and Exclusion (FDE) test statistic process at the system level.

At the user level, error characterisation takes place in three phases (i) measurement error characterisation, (ii) position error characterisation, and (iii) residuals characterisation. The effects of these phases are apparent in the outputs of positioning algorithms, affecting estimated position, position uncertainty, protection level, and integrity alarms. Like at the system level, measurement error at the user level is also characterised by PDFs, generally following a Gaussian distribution. Through uncertainty propagation, position uncertainty is then calculated using the PDFs of measurement error.

Furthermore, the outputs from the positioning algorithm are used as inputs for the integrity layer. This layer includes three main processes: FDE, protection level computation, and integrity alarm

generation. FDE commonly employs residuals-based test statistics, dependent on the characterisation of residuals. Similarly, the computation of protection levels is influenced by the characterisation of errors and residuals, since it is a function of measurements and position error characterisation, and the FDE process.

GNSS errors, nevertheless, are not always well-characterised by the Gaussian distribution especially at the tail (Panagiotakopoulos et al., 2014). Therefore, the Gaussian distribution could have a significant impact on all the elements (FDE, positioning uncertainty, and PL). Thus, several efforts have been made to replace Gaussian with alternatives to improve the performance. Of these efforts, Panagiotakopoulos et al., (2014) evaluated the capability of the GEV distribution to characterise pseudorange error in static mode. The results showed that replacing the Gaussian distribution with the GEV distribution offers several advantages: (i) a more accurate distribution of residual errors and better representation of extreme errors at the tail, (ii) a less conservative safety threshold than the Gaussian based safety threshold derived for a given missed detection probability, and (iii) a better positioning quality. However, the work did not comprehensively and numerically assess distribution fitting and its impact on the system availability.

2.2.4 Computer Vision Data Sources

Computer vision sensors have been widely utilised for localization and positioning through Simultaneous Localization and Mapping (SLAM). SLAM operates by mapping the surrounding environment and determining the user's location within it in real-time. SLAM can be categorised based on the sensors used into vision-based SLAM (Davison,2003; Schlegel et al., 2018), Lidar-based SLAM(Kohlbrecher et al., 2011; Hess et al., 2016), Radar-based SLAM (Hong et al.,2020; Schuster et al.,2016), and hybrid SLAM (Bibby and Reid, 2010; Gallagher et al.,2021). The precision of SLAM is influenced by the number of features captured and the range/depth accuracy in real-time, making Lidar-based SLAM one of the most accurate forms.

In maritime navigation, implementing SLAM has limitations in terms of system availability, as it relies on feature matching, difficult in environments lacking distinct features. In feature-poor environments like those in the ocean and coastal phases, there are not enough distinctive features to match, which may cause the SLAM system to fail. However, there are some features within the port environment. Nevertheless, the development of integrity monitoring in SLAM is still limited requiringfurther research.

2.2 Benefits of opportunistic data sources in INSPIRe context

The opportunistic data sources in Section 2.1 can support maritime navigation at both the user and system levels. Crowdsourced data (from smartphones and nearby vessels) can support the user level across all operational phases (Ocean, Coastal, HEA and Restricted, and Port). Utilising the crowdsourced positioning approach can provide higher performance in the port phase, where the number of nearby vessels is greater than in the other phases.

In addition to crowdsourced data, signals-of-opportunity can support the user level. Some signals-of-opportunity can be used only at the port with imitated range, such as IPS (Wi-Fi, Bluetooth,

and UWB), while eLoran and AIS base stations can be utilised at all operational levels (depending on eLoran availability).

Since computer vision-based positioning (e.g., SLAM) requires features to be extracted and matched with pre-stored ones, this approach can only be applied at the port with limited ranges, due to the difficulty in finding features in other phases. Table (2.2) presents the opportunistic data sources, detailing their contribution to maritime navigation in terms of support at levels (user and/or system) and across different operational phases.

Opportunistic approach		Level		Operation phase			
		User	System	Ocean	Coastal	HEA and restricted	Port
User-level crowdsourcing	- Vessels (via AIS) - smartphone	✓		✓	✓	✓	✓
Signal of opportunity	IPS	✓					✓
	eLoran	✓		✓	✓	✓	✓
	AIS base stations	✓			✓		✓
System-level crowdsourcing	CORS netwok	✓	✓	✓	✓	✓	✓
Computer vision	SLAM	✓					✓

Table 2.2: opportunistic data sources, detailing their contribution to maritime navigation in terms of support at levels (user and/or system) and across different operational stages

In more detail, the opportunistic data sources can be beneficial for the user and system level in enhancing maritime navigation and operational efficiency. At the user level, the key benefits of the opportunistic data sources are that:

- By leveraging these data sources, it might be feasible to meet future requirements, particularly those pertinent to the port phase. To the best of our knowledge, the Dual Frequency Multi-Constellation (DFMC) GNSS and/or dead reckoning cannot achieve the required level of future performance in the port phase (1m Horizontal accuracy (95%), 2.5 m Horizontal alarm limit, availability >99.8%, continuity >99.97 per 15 minutes, and an integrity risk of 10^(-5)). This underscores the potential need for employing GNSS alternatives besides GNSS to meet these performance criteria. During the port stage, opportunistic data sources should provide a positioning accuracy of 1 meter with an integrity risk of 10^{-5} when the system operates independently. When integrating with GNSS, the candidate technologies should be evaluated using either real data or simulations.
- The opportunistic data sources can contribute to reducing alarm limits, thereby enhancing maritime navigation operations. Put differently, in the case of achieving performance requirements in all operational phases, improving system performance can still lead to a reduction in the alarm limit in the future. This results in a decrease in the minimum required distance between vessels, leading to improved operations, especially in high-density regions (e.g., port phase).
- The opportunistic data sources can be one of the robust layers of defence against various threats such as jamming and spoofing.

In addition to the advantages at the user level, opportunistic data sources offer numerous benefits for system-level integrity monitoring. The following characterise some of the key advantages:

- Improving error characterisation and overbounding. By increasing the amount of data utilised at the system level, the system can more accurately characterise measurement errors. The impact of Gaussian assumption at the system level and user level is significant, as discussed in Section 2.2.3. Thus, there is a need to utilise Gaussian alternatives to achieve a more precise characterisation of measurement errors, ultimately enhancing system performance. This improvement can reflect on user-level performance and support the system in achieving the required performance level, especially at the porting level. In addition, as noted earlier, this enhancement can reduce the alarm limit in the future, leading to improved operations.
- Enhancing the FDE process. The additional number of opportunistic data increases the redundancy which leads to a more robust FDE process. This will be reflected in all user-level outputs including positioning accuracy and protection level, then it is indirectly liked also with availability and continuity.

2.3 Selecting opportunistic approaches

The selection of opportunistic approaches is based on the current limitations of INSPIRe, which may not achieve future performance in the port phase. By reviewing the advantages and limitations of opportunistic approaches, summarised in detail in Table 2.3, two approaches have been selected in terms of priority for INSPIRe: user-level crowdsourcing utilising information from nearby vessels, and system-level crowdsourcing based on utilising the CORS network to improve error characterisation. The former can support system performance in all operation phases, and the latter can improve the FDE process and enhance error characterisation and overbounding, which in turn can enhance system performance at the user level across all phases.

Meanwhile, IPS can support the system within a limited range in the port phase and requires infrastructure. eLoran was switched off in Europe and AIS shore station-based systems integrity monitoring is not well explored, and the level of accuracy of these two technologies cannot meet the requirements in the port phase. Computer vision faces challenges in finding features to match in the ocean and coastal phases, with limited features in the port phase, and Integrity monitoring for SLAM is still not well explored.

Table 2.3: advantages and limitations of opportunistic approaches

Opportunistic approach		Limitations	Advantages
User-level crowdsourcing	Vessels	• As of the current state of knowledge, no integrity monitoring algorithm has been developed yet. • The system's availability is a function of the number of nearby vessels. • To ensure computational safety, a protection level needs to be incorporated into the AIS messages. • Failure modes and modes are not well explored for the range sensors (e.g., Radar or LiDAR), as well as the overbounding of range measurements.	• Based on IMO (2015), AIS information should include the ship's position with an accuracy indication and integrity status. • The range sensors exist in the majority of the vessels. • The performance of this approach is a function of the vessel's density, thus, this approach can provide higher performance in the port phase
	Smartphone	• Provides a positioning solution with a low level of accuracy. • Data transformation between smartphones and vessels is a complex process. • There is a significant impact from signal blockage due to the user's body and the surrounding complex environment. • The protection level is not computed on the smartphone using the best of our technology, which is necessary to determine the vessel's protection level. • Advanced error modelling is not applied to smartphones • the system availability is a function of nearby smartphone numbers	• There is a vast number of smartphone users.
Signal of opportunity	IPS (e.g WiFi, Bluetooth and UWB)	• Required infrastructure • Can be utilised within a limited range at the port level only • There is a weak understanding of indoor positioning integrity monitoring, failure modes, and models to date	• Can provide a high level of accuracy. Of the IPS, UWB can provide a cm-dm accuracy level, and AOA Bluetooth can provide 1-3 meters of accuracy

		• Implementation of IPS in real-time applications remains complex • The majority of IPS systems are highly impacted by metal, especially UWB	• The cost of some IPS infrastructure such as Bluetooth is not expensive
	eLoran	• The system was switched off in Europe in 2015, and only one station works in UK, which is not sufficient for positioning. • Integrity monitoring for eLoran is not well explored, including developing an integrity layer at the user level and understanding the failure mode and model. • Required system-level integrity monitoring to ensure signal quality. •	• Can cover a wider range compared with IPS. • Requires expansive infrastructure.
	AIS base stations	• Integrity monitoring for AIS shore station-based integrity is not well-explored • Required system-level integrity monitoring to ensure the signal's quality	• The AIS shore stations exist in the UK and Europe
system-level crowdsourcing	CORS network	• A limited number of CORS networks in the UK provide 1Hz data, required for integrity monitoring. • Increasing the number of utilised CORS networks at the system level will increase computational costs.	• Can improve the FDE process at the system level. • Can enhance error characterisation and overbounding at the system level, which can improve system performance at the user level in all phases
Computer vision	SLAM	• There are not enough features to match in the ocean and coastal phases, and limited features in the port phase. • Complex Integrity monitoring • computationally expensive.	• Can provide a high level of accuracy

3. System-level crowdsourcing

As mentioned in Section 2.2.3, error characterisation is employed across three phases at the user level in positioning/integrity algorithms, and at two phases within system-level integrity monitoring. At the system level, errors are characterised by PDFs, which are transmitted to users after overbounding the distributions. In addition, error characterisation is used in the FDE test statistic process at the system level. At the user level, error characterisation is utilised in measurement error characterisation, position error characterisation, and residual characterisation. This characterisation impacts all outputs of positioning algorithms, including estimated position, position uncertainty, and protection level. The development of GNSS integrity widely assumes that GNSS error follows a Gaussian distribution; however, this is not always the case, especially at the tail (Panagiotakopoulos et al., 2014).

This section evaluates error characterisation at the system level using Gaussian, Generalised-t, GEV, Logistic, Laplace, and Cauchy distributions. These distributions were selected based on the findings of Alghananim and Ochieng (2023), following evaluation of 31 distributions for their capability to model the distribution of GNSS measurement error.

The distributions in this report have been evaluated in three main aspects: fitting (overall and tail), impact on system availability, and bounding, taking computational complexity into account, as shown in Figure 3.1 The assessments utilised to evaluate these dimensions include the Kolmogorov-Smirnov (KS) test (Kolmogorov,1933), graphical assessment, and availability assessment. The KS test is used to assess the distribution's overall fitting, while the graphical assessment used is to evaluate the fitting (tail and core), and overbounding. The availability assessment is used to evaluate the impact on system availability.

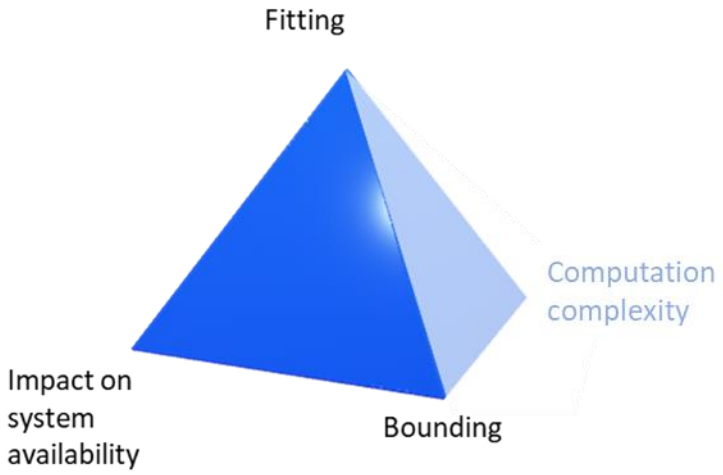


Figure 3.1: error characterisation evaluation aspects

These aspects significantly influence system performance. Overall fitting is correlated with system reliability, and tail fitting is crucial for mapping extreme events which is significant for integrity. The ultimate goal of error characterisation is to strike a balance between these qualities, ensuring safety while maintaining system availability and reliability. Table 3.1 shows the error characterisation aspects, impact on the system, and assessments.

Aspect		Impact on the system	Assessment		
			Graphical assessment	Kolmogorov-Smirnov (KS)	Availability assessment
Fitting	Overall fitting	System reliability	✓	✓	
	Tail fitting	Mapping extreme events	✓		
Bounding		Safety/integrity	✓		
impact on system availability		availability			✓

Table 3.1: error characterisation aspects, impact on the system, and assessments

The system-level error characterisation encompasses five stages: data collection, distribution selection, distribution estimation, goodness of fit assessments, and evaluation, as shown in Figure (3.2). In the first stage, data have been collected from the OS Net CORS network, discussed in Section 3.1. In the second stage, six distributions have been tested: Gaussian, Generalised-t, GEV, Logistic, Laplace, and Cauchy. The Probability Distribution Function (PDF), Cumulative Distribution Function (CDF), and their parameters for these distributions are presented in Section 3.2. Distribution estimation using maximum likelihood is discussed in Section 3.3. The fitting assessment results are summarised in Section 3.4, and Section 3.5 provides a discussion that includes evaluating the tested distributions.

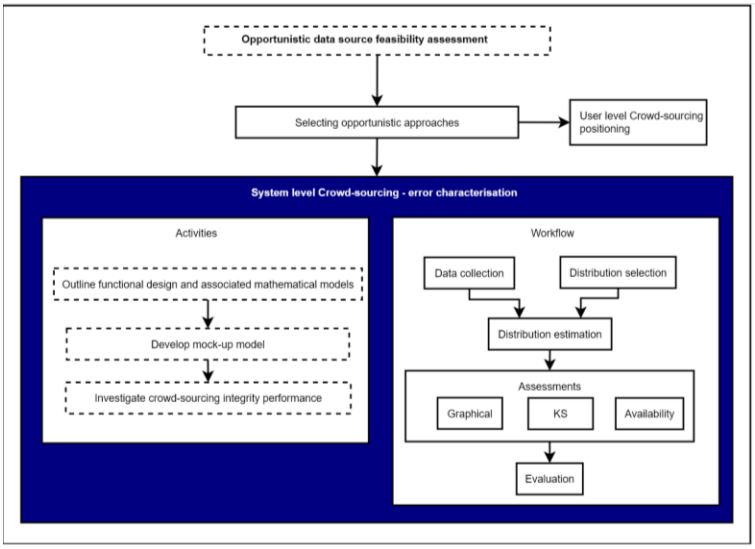


Figure 3.2: system-level error characterisation stages and high level WP8.1-WP8.4 activates

3.1 Data source

The data have been collected and proceeded from 20 OS stations around the UK. The data include 3 hours of raw data (RINEX) files with 30 seconds epochs. Table (3.2) presents data datasets used in this section.

Table 3.2:datasets used in this study

ID	StationID_year_dayoftheyear_startingtime_endingtime
1	AMER_ 2023_220_18_21
2	ANLX_ 2023_255_21_00
3	ATTL_ 2023_215_00_03
4	BUCI_ 2023_252_09_12
5	CAMO_ 2023_228_14_17
6	CARL_ 2023_218_10_13
7	CLAW_ 2023_230_12_15
8	FAUG_ 2023_232_07_10
9	GLAS _2023_242_04_07
10	HOLY _2023_210_06_09
11	LEED_ 2023_240_06_09
12	LEEK_ 2023_212_17_20
13	MANR _2023_238_08_11
14	NCAS _2023_250_0_3
15	SABS _2023_245_12_15
16	SHOE _2023_225_06_09
17	SOTN _2023_235_00_03
18	SWAN _2023_247_15_18
19	SWAS _2023_248_18_21
20	THUS _2023_222_15_18

3.2 Error distributions

As noted earlier, we have employed six distributions in this report: Gaussian, Generalised-t, GEV, Logistic, Laplace, and Cauchy. The following are the key characteristics of these distributions:

- The Gaussian distribution, characterised by its symmetry and two parameters (mean for location and standard deviation for scale), is used across various fields, due to its simplicity in both computation and uncertainty propagation.
- The generalised-t distribution, or t location-scale distribution, is a heavy-tailed distribution defined by three parameters: location, scale, and degrees of freedom. This distribution is particularly advantageous for modelling small sample sizes.
- GEV distribution, which belongs to the scale-location distribution family and originates from extreme value theory, is an asymmetric three-parameter distribution that maps extreme events and their deviation from the sample median.

- The Logistic distribution, commonly used in growth models and logistic regression, is a two-parameter function defined by its location and scale parameters.
- The Laplace distribution, or double-exponential distribution, is a symmetric distribution with widespread applications in engineering, finance, inventory management, quality control, astronomy, biology, and environmental sciences (Kotz et al. 2001).
- Cauchy distribution, also known as the Lorentz distribution, is a symmetric distribution with very heavy tails, making it an excellent tool for outlier detection.

Table (3.3) presents the PDF, CDF, and parameters for each of the six distributions.

3.3 Distribution estimation

In this report, we utilise maximum likelihood - a robust and systematic estimation method - to determine distributions. ML determines the parameters of the PDF, denoted as

$$L_n(\theta|x_1 \dots, x_n) = \prod_{i=1}^{n} f(\theta|x_i) \tag{3.1}$$

where: $(x_1, \dots., x_n)$ is the data sample, $f(x|\theta)$ is the PDF with parameters θ, and n is the sample size.

The essence of ML estimation lies in maximising the log-likelihood function, which is a transformation of the likelihood function for simplification. The log-likelihood function is expressed as:

$$\ell(\theta; x_i) = In\big(L(\theta, x_i)\big) \tag{3.2}$$

To estimate the parameter, the derivative of the log-likelihood function concerning θ is set to zero, as the objective is its maximisation:

$$\frac{\partial \ell}{\partial \theta_i} = 0 \tag{3.3}$$

Then an iterative numerical procedure (e.g., Newton-Raphson) can be employed to solve the equations returned from the partial derivative of the log-likelihood function.

Table 3.3: PDF, CDF, and parameters for distributions used in this study

Distribution	Symmetric/ asymmetric	Parameters	PDF	CDF
Gaussian	symmetric	- location (μ) - Standard deviation (σ)	$\frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}\ dt$
Generalised-t	symmetric	- Location (μ_t) - Scale (σ_t) - Degree of freedom (v).	$\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sigma\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)}\left[\frac{v+\left(\frac{x_i-\mu}{\sigma_t}\right)^2}{v}\right]^{-\frac{v+1}{2}}$	$\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sigma\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)}\int_{-\infty}^{x}\left[\frac{v+\left(\frac{t-\mu_t}{\sigma_t}\right)^2}{v}\right]^{-\frac{v+1}{2}}$ a
Generalised Extreme Value (GEV)	asymmetric	- Shape (ξ), - location (μ), - Scale (σ)	$F(x)=e^{-Q(x)}$ $Q(x)=\begin{cases}\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi} & \xi\neq 0\\ e^{\left(-\frac{x-\mu}{\sigma}\right)} & \xi=0\end{cases}$	$f(x)=\frac{1}{\sigma}\,Q(x)^{\xi+1}\,e^{-Q(x)}$
Logistic	symmetric	- Location (μ) - Scale (s).	$\frac{e^{\left(-\frac{x-\mu}{s}\right)}}{s\left(1+e^{\left(-\frac{x-\mu}{s}\right)}\right)^2}$	$\frac{1}{1+e^{\left(-\frac{x-\mu}{s}\right)}}$
Laplace	symmetric	- Location (μ), - Scale (b)	$f(x)=\frac{1}{2b}\ e^{-\frac{\lvert x-\mu\rvert}{b}}$	$\frac{1}{2}\left[1+sign(x-\mu)*\left(1-e^{-\frac{\lvert x-\mu\rvert}{b}}\right)\right]$
Cauchy	symmetric	- Location (x_0) - Scale (γ)	$\frac{1}{\pi}\ \frac{\gamma}{(x-x0)^2+\gamma^2}$	$(1/2)+\frac{1}{\pi}\,atan(\frac{(x-x_0)}{\gamma})$

3.4 Goodness-of-fit assessments

As highlighted earlier, the goodness-of-fit assessments concentrate on four key areas: overall fit, core fitting, the impact on system availability, and overbounding. The Goodness-of-fit assessment methods are systematically divided into two categories: Graphical assessment and hypothesis tests. The graphical method evaluates the distribution using the PDF, CDF, and empirical CDF (eCDF). This assessment aims to test three aspects: overall fit, core fit, and overbounding.

The KS test is the hypothesis testing approach used in this report to assess the overall fitting. The test examines if a data sample originates from a population following a specified distribution. It is based on computing the largest difference between the empirical distribution and the CDF of the proposed distribution. The formula for the KS test is as follows:

$$D = \max_{n}(D_1, D_2) \tag{3.4}$$

$$D_1 = \max_{n}(F(x_I) - eCDF(x_I)) \tag{3.5}$$

$$D_2 = \max_{n}\left(eCDF(x_I) - F(x_I)\right) \tag{3.6}$$

where: $F(x)$ is the CDF hypothesized distribution, eCDF is the empirical cumulative distribution function.

The hypothesis that the error sample follows a hypothesized error distribution is given by:

- Null-Hypothesis H_0: the error sample $(x_1, \dots, x_n)$ follows a hypothesized error distribution

$$\chi^2 > \chi^2_{1-\alpha,k-s} \tag{3.7}$$

- Alternate Hypothesis H_1: the error sample $(x_1, \dots, x_n)$ does not follow a hypothesized error distribution

$$\chi^2 \geq \chi^2_{1-\alpha,k-c} \tag{3.8}$$

In addition to these techniques, this report also includes an availability assessment as a heuristic evaluation, to measure the impact of the distribution on the system availability by calculating the Confidence Interval (CI) for various confidence levels (CL), discussed in more detail in Section (3.5.3).

3.5 Assessments Results

3.5.1 Kolmogorov-Smirnov test results

The KS test results show that the pseudorange error does not follow any of the six distributions in all 20 datasets and their combination, at 5% significance level. However, the test value, computed from equations (3.4), has been used to evaluate the distribution's overall fitting. Table (3.4) summarises the KS test statistic values for the six distributions within the 20 datasets. Table (3.5) and Figure (3.3) displays the number of times with which each distribution ranked as the best, second best, and third best in terms of dataset fitting based on the KS test results

The KS test values show a mixed result in terms of fitting the measurement error. Generally, the results suggest that among the tested distributions, the Logistic distribution provides the best overall fitting. Based in KS test values, the Logistic distribution shows the best performance between the tested distribution in fitting 12 datasets, and second best for 2 databases, and the third best for 5 databases. Put differently, Logistic was one of the best three distributions in 19 datasets (out of 20). In addition, the Logistic was the best in the combined dataset.

The Cauchy and Laplace distribution ranked as the second-best and third-best distribution, following the Logistic distribution in terms of overall fitting. Specifically, the Laplace distribution showed the second-best distribution in fitting the combination dataset and the Cauchy distribution ranked as the third-best, while the Cauchy distribution showed the best fitting for 4 datasets, second-best for 3 datasets, and third-best for one dataset, while the Laplace distribution provides the best fit for three datasets. Notably, neither the Gaussian nor the generalised t distribution ranked as the best distribution in fitting any of the datasets.

Table 3.4: KS test statistic values for the six distributions within the 20 datasets

Dataset ID	Gaussian	Generalised-t	GEV	Logistic	Cauchy	Laplace
1	0.185	0.228	0.354	**0.181**	0.201	0.219
2	0.156	0.244	0.170	**0.156**	0.171	0.174
3	0.178	0.310	0.187	**0.172**	0.201	0.211
4	0.179	0.361	0.196	**0.165**	0.196	0.214
5	0.231	0.300	0.210	0.205	0.199	**0.193**
6	0.269	0.298	0.421	0.172	**0.170**	0.175
7	0.226	0.226	0.227	0.239	0.226	**0.223**
8	0.182	0.366	0.181	**0.169**	0.201	0.208
9	0.349	0.354	0.214	0.196	**0.173**	0.194
10	0.163	0.167	**0.155**	0.169	0.191	0.239
11	0.176	0.365	0.184	**0.174**	0.218	0.251
12	0.221	0.317	0.210	**0.206**	0.232	0.234
13	0.216	0.275	0.310	**0.178**	0.205	0.212
14	0.188	0.266	0.168	**0.166**	0.196	0.213
15	0.511	0.208	0.574	0.501	**0.176**	0.976
16	0.180	0.402	0.159	**0.159**	0.213	0.255
17	0.191	0.324	0.187	0.183	**0.179**	0.203
18	0.180	0.184	0.197	0.182	0.183	**0.178**
19	0.219	0.207	0.389	**0.184**	0.206	0.225
20	0.184	0.298	0.183	**0.183**	0.185	0.199
Combined	0.219	0.282	0.424	**0.174**	0.180	0.177

Table 3.5: number of times with which each distribution ranked as the best, second best, and third best in terms of dataset fitting based on the KS test results

	Gaussian	Generalised-t	GEV	LOGISTIC	CAUCHY	LAPLACE
Best	0	0	1	12	4	3
second	7	2	5	2	3	1
third	6	2	4	5	1	2

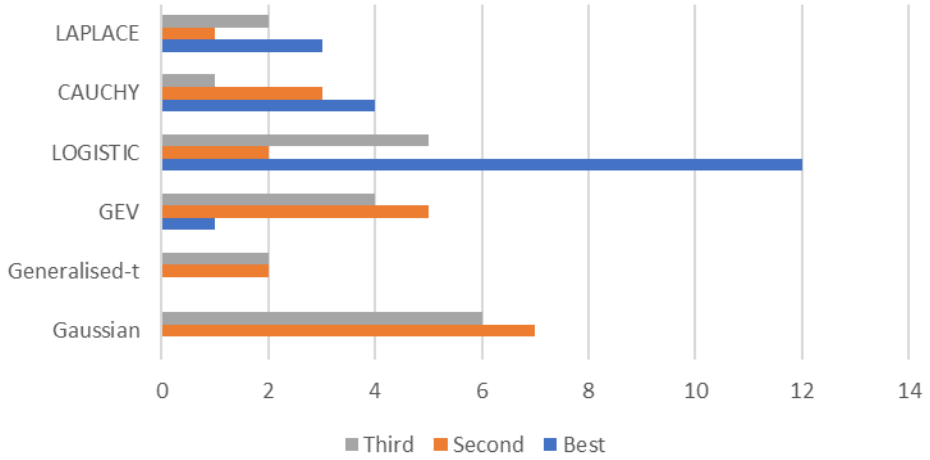


Figure 3.3: number of times with which each distribution ranked as the best, second best, and third best in terms of dataset fitting based on the KS test results

The mixed results overall suggest that no single distribution can accurately fit measurement errors in all cases. This highlights the potential for implementing adaptive characterisation based on an understanding of quality indicators and other factors in real-time operations. Generally, the best distribution in terms of fitting datasets using KS test values is the Logistic distribution, followed by Cauchy, and Laplace.

3.5.2 Graphical results

As safety is the first concern in mission-critical applications, overall fitting is not sufficient to evaluate the distributions. Consequently, this section incorporates graphical assessments to evaluate distribution bounding and tail fitting. The overall fitting will be re-evaluated and compared with the results of the KS test, since the KS test has limitations in precisely assessing overall fitting, as it is based on the maximum difference between the CDF and eCDF.

Figure (3.4-3.7) shows a key result of the graphical assessment for 5 datasets. The appendix presents the results of the rest of the database (database 6-20). Figure (3.8) presents the graphical assessment of the combined dataset. The results show that logistic, Cauchy, and Laplace showed the best overall fitting, confirming KS test results. The Generalised t distribution showed an extremely good fitting for the distribution core.

The GEV distribution showed the best performance in fitting for the tails and mapping the extreme events, and provides the best bound in 19 datasets out of 20, followed by Gaussian and Logistic distributions. This suggests using these distributions for mission-critical applications, considering the need to apply overbounding.

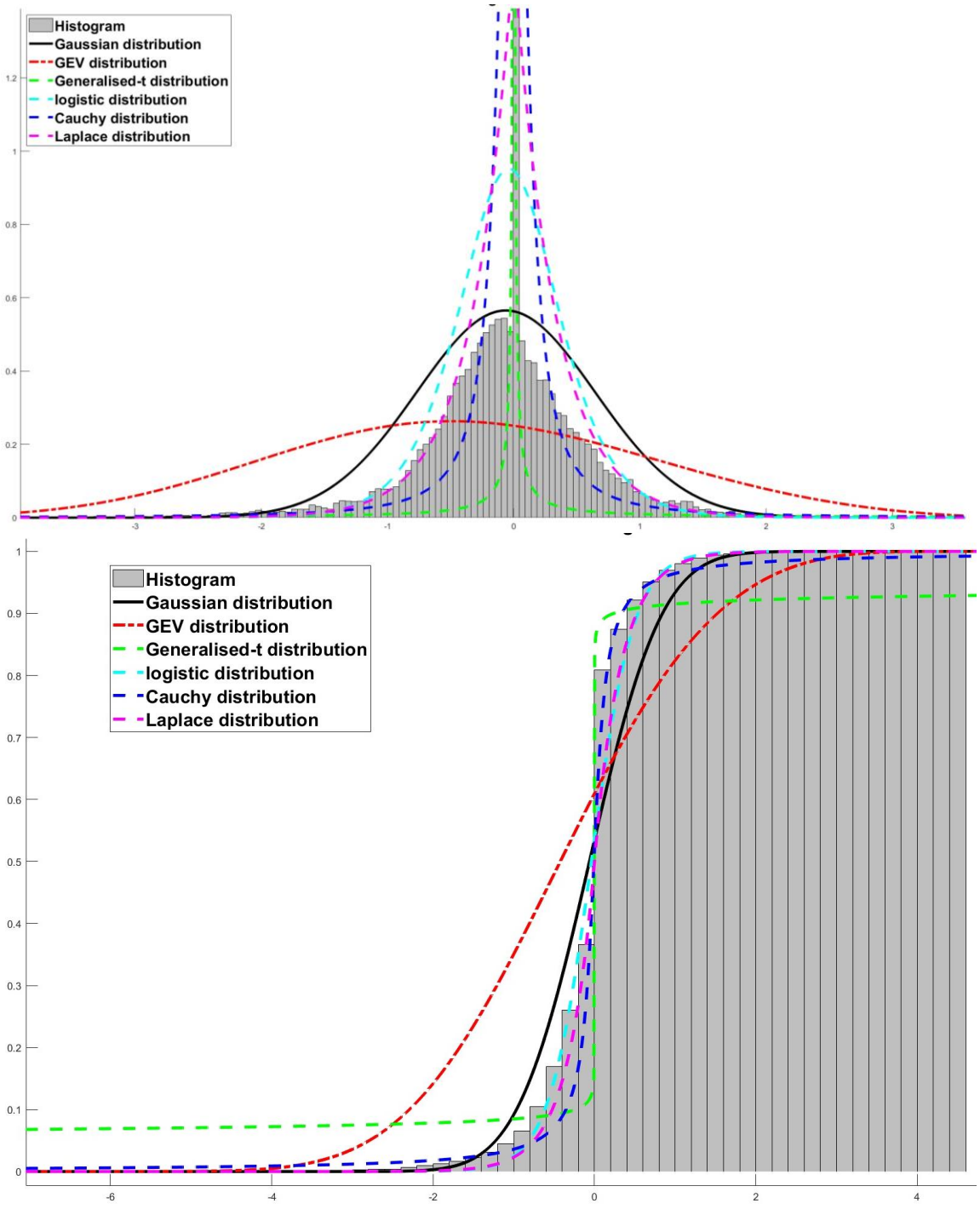


Figure 3.4: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 1

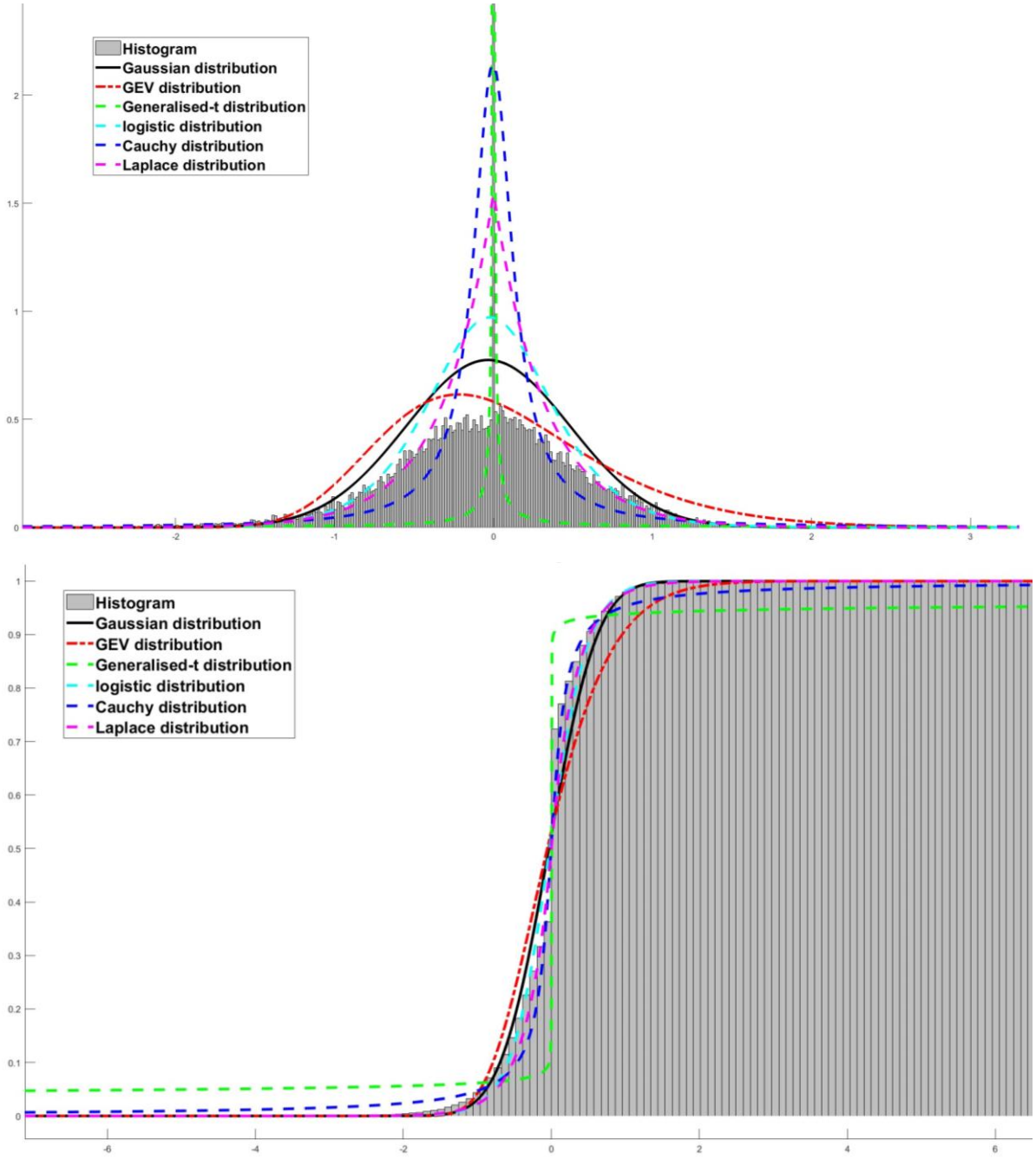


Figure 3.5: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 2

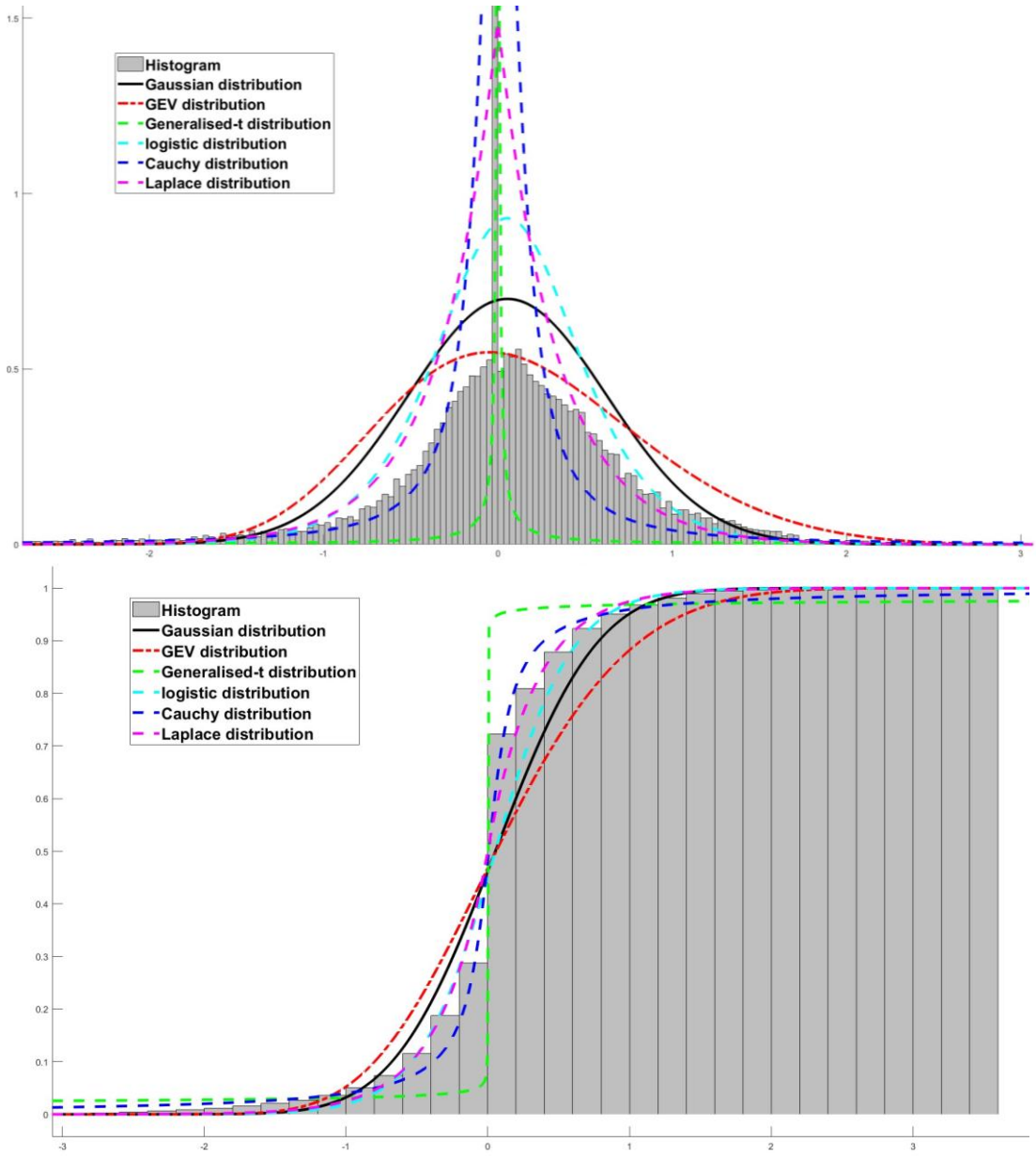


Figure 3.6: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 3

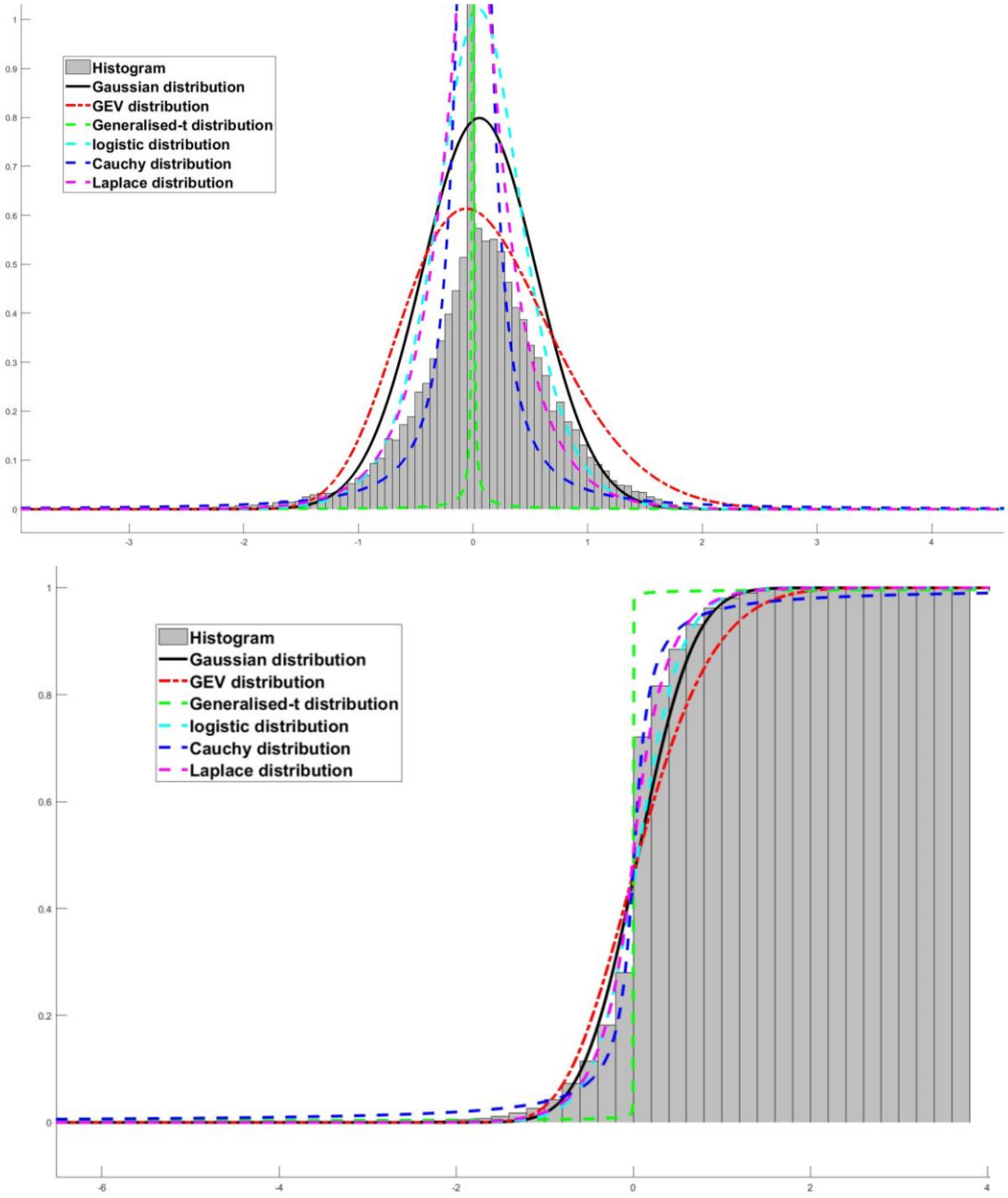


Figure 3.7: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 4

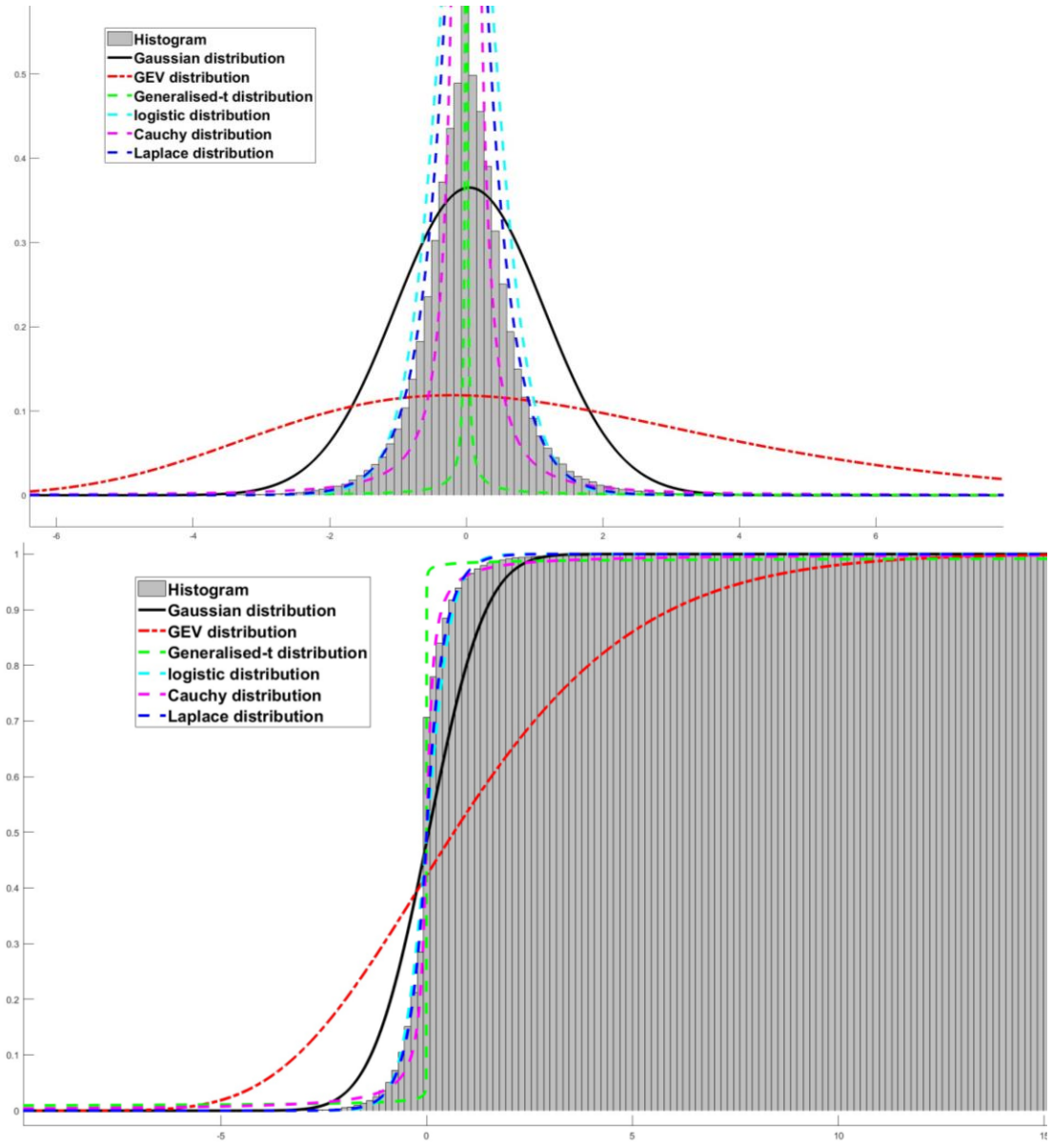


Figure 3.8: pseudorange error characterisation using the six distributions in PDF and CDF domain, combined datasets

3.5.3 Availability results

As noted earlier, the availability test aims to assess the impact of the distributions on system availability. The availability test involves computing the confidence interval at a confidence level that aligns with the integrity risk requirements. In this section, the test is applied for integrity risks of $10^{-4}, 10^{-5}, 10^{-6}, and\ 10^{-7}$. Therefore, it entails computing the confidence interval for $1-10^{-4}, 1-10^{-5}, 1-10^{-6}, and\ 1-10^{-7}$. However, while this assessment may not establish a precise link between availability and distribution, it can reliably exclude certain distributions that significantly reduce system availability.

Table (3.6) presents the results of the availability assessment for the 20 datasets and the combined dataset. Across all tested integrity risk levels and datasets, the findings indicate that the implementation of the Generalised t and Cauchy distributions can reduce system availability to the point where the system almost fails to operate in all cases. This outcome is a result of the heavy-tail phenomena associated with these two distributions. Consequently, Generalised t and Cauchy are not considered reliable for characterising GNSS measurement error.

Focusing on the integrity risk of 10^{-5} to 10^{-4} level (as required in INSPIRe 10^{-5} per 3 hours), the results for the remaining distributions (Gaussian, Logistic, Laplace, and GEV) show that Gaussian and Logistic distributions have the lowest confidence intervals among the four, followed by Laplace, while GEV has the highest. This outcome aligns with the findings of the graphical assessment, as GEV provides better bounds and mappings of extreme events. Since overbounding is the first concern in mission-critical applications and the protection level is computed form an overbounding distribution, these results do not indicate a direct link between these four distributions and system availability. To accurately project the distribution to the protection level, a comparison of the confidence levels should be made for overbounded-Gaussian, overbounded-Logistic, overbounded-Laplace, and overbounded-GEV.

In conclusion, the results indicate that Generalised t and Cauchy are not reliable for characterising GNSS measurement errors. The other distributions should be re-evaluated in the future, taking into account overbounding. This is discussed in Section 5.1.

Table 3.6: availability assessment results in meters for the 20 datasets and the combined one

		Integrity risk			
		10^{-4}	10^{-5}	10^{-6}	10^{-7}
Gaussian	Average	6	7	8	8
	min	3	3	3	4
	max	28	32	35	38
GEV	Average	9	10	12	13
	min	3	3	3	3
	max	34	38	43	57
Logistic	Average	6	7	8	10
	min	3	4	5	5
	max	17	21	25	29
Laplace	Average	7	8	9	8
	min	3	4	4	4
	max	19	23	25	26
Generalised t	Average	7.44543×10^{39}	$6.24816E\times 10^{51}$	5.24645×10^{63}	4.40575×10^{75}
	min	50	154	473	1459
	max	$1.48784E\times 10^{41}$	1.24946×10^{53}	1.04926×10^{65}	8.81147×10^{76}
Cauchy	Average	2170	44677	53180	131692
	min	380	1235	2522	3410
	max	4065	262563	385935	1915867

3.6 Summary and Conclusion

Error characterisation plays a crucial role in positioning and integrity monitoring algorithms, impacting the quality of all outputs, including protection levels and uncertainties. GNSS pseudorange error at the system level has widely been assumed to follow a Gaussian distribution. Hoewever, this is not always the case as the results presented in this section demonstrate.

Five distributions, including Generalised-t, GEV, Logistic, Laplace, and Cauchy, have been tested against the Gaussian distribution. These distributions were evaluated in three main aspects: fitting (overall and tail), their impact on system availability, and overbounding. The primary objective of error characterisation is to strike a balance between ensuring safety while maintaining system availability and reliability. To achieve this, three types of tests were applied in this section: KS tests, graphical assessment, and availability assessment.

The availability assessments revealed that the Generalised-t and Cauchy distributions are not reliable choices for error characterisation. Among the remaining distributions (Gaussian, GEV, Logistic, and Laplace), the results indicated that the GEV distribution offers the best performance in bounding and mapping extreme events. When considering fitting results, Logistic and Laplace distributions generally outperformed GEV and Gaussian distributions in most datasets.

Considering computational complexity of real-time applications, the GEV distribution is more complex than the other distributions. Therefore, we suggest utilising Logistic, Laplace, and Gaussian distributions based on current computational capabilities. Additionally, we recommend future research to simplify the computational complexity associated with the GEV distribution, discussed in more detail in Section 5.1.

Highlighting the mixed results in terms of fitting, this suggests the potential development of an adaptive error characterisation approach in the future. This can be achieved by using quality indicators to dynamically select one of the four distributions (Gaussian, GEV, Logistic, and Laplace) in real-time. Finally, in cases where an adaptive error characterisation method has not been developed, we recommend using the Logistic distribution based on the findings of this report, with consideration of the need for overbouding.

4. User-level Crowdsourcing

4.1 Overview

As mentioned in Section 2.2.1, user-level crowd-sourced positioning is based on using nearby vessel information to enhance the positioning and integrity algorithms in the maritime sector. This is especially relevant because the current GNSS cannot provide the required level of performance, particularly at the port. Furthermore, the proposed user crowd-sourced positioning concept in this report can also be expanded to include other sectors where integrity is a critical performance metric.

The fundamental concept of crowd-sourced positioning is based on leveraging nearby vessel positioning and integrity information, which can be transmitted via the Automatic Identification System (AIS). In addition, it also encompasses range measurements that can be computed using Radar and/or Lidar technology. According to the International Maritime Organization (IMO, 2015), AIS information should include the ship's position with an accuracy indication and integrity status, as discussed in Section 4.2.

In simpler terms, this layer contributes to system safety by establishing connections between all nearby vessel positions, moving beyond the traditional approach of solely linking vessel positions with satellites, as shown in Figure (4.1). Furthermore, user-level crowd-sourced positioning can enhance the system's situational awareness due to its foundation in relative positioning modes, and this layer can be extended to include tracking and predicting the locations of nearby vessels. In addition, the adaptive relative positioning utilised in this layer provides a basis for further developments of anti-spoofing and anti-jamming techniques, ensuring a more secure and reliable system.

This approach is introduced as an additional layer to the two-integrity monitoring layers in INSPIRe. In real-time operation, the system level integrity sends to the user the required integrity information to ensure safe operation. Then at the user-level, the GNSS devices receive this information to compute the position and protection level. The proposed layer then is based on transferring positioning information between the vessels to compute the protection level within a third layer. Figure (4.2) presents the functional architecture of these three layers.

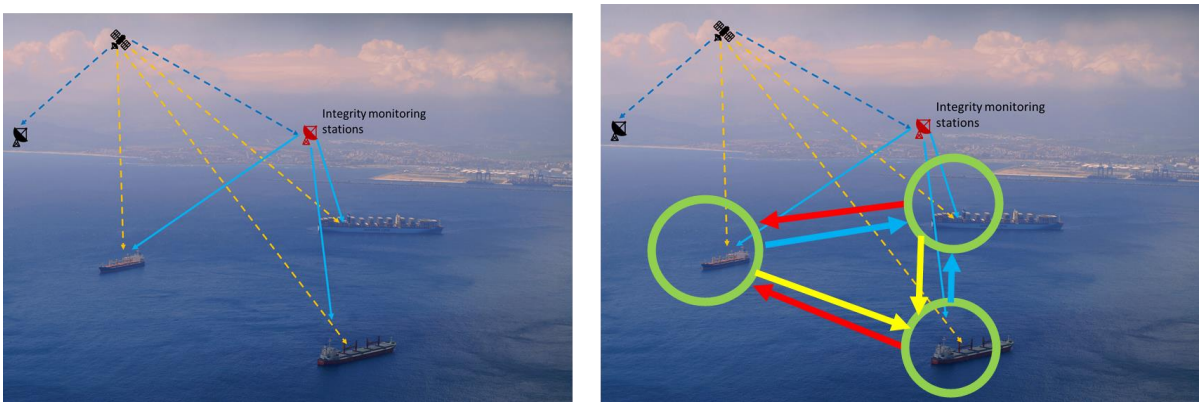


Figure 4.1: The left side illustrates the traditional approach of linking vessel positions to satellites, while the right side demonstrates the contribution of crowd-sourced positioning in establishing connections between the positions of all nearby vessels in addition to satellites

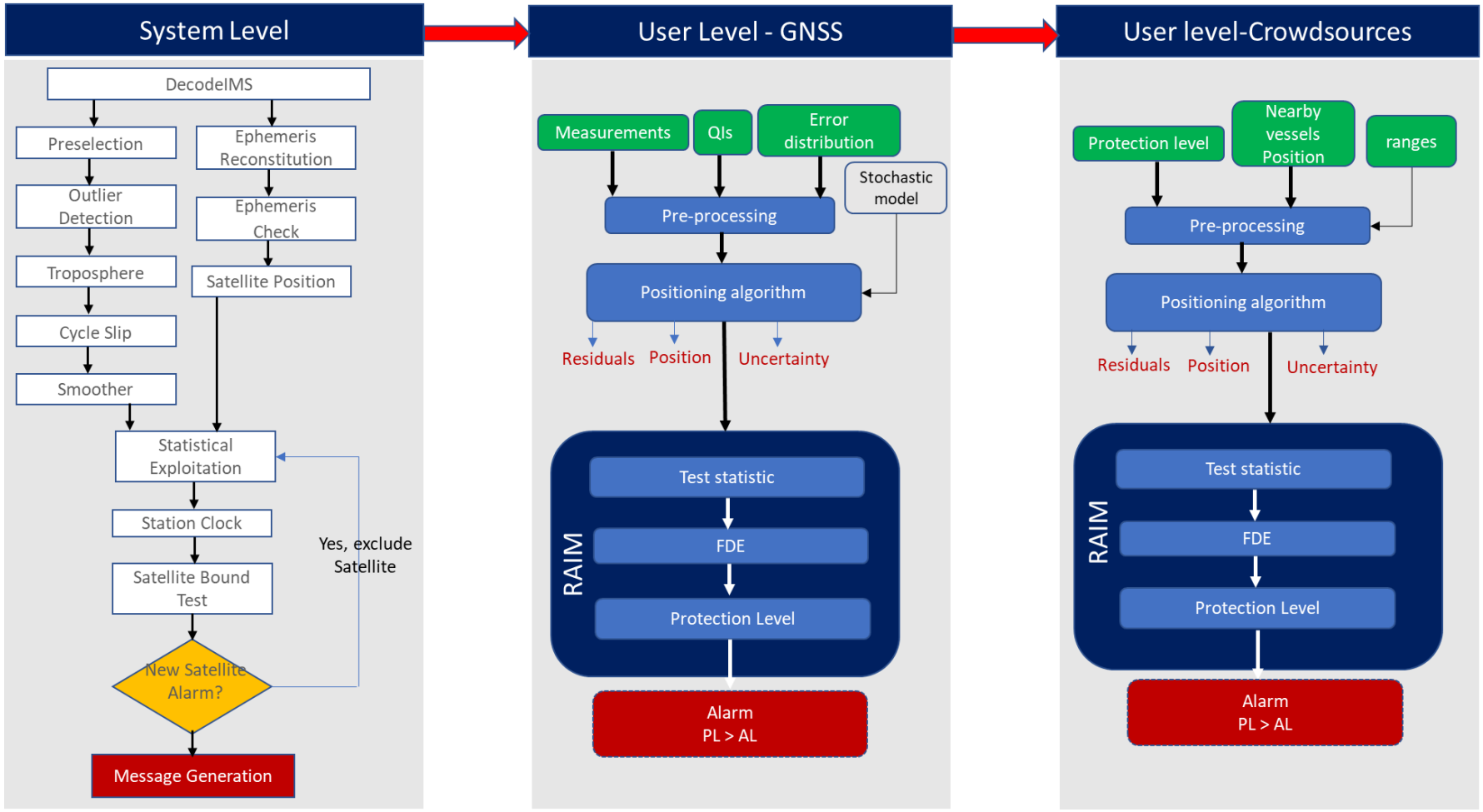


Figure 4.2: functional architecture of these three layers, the system level architecture developed within WP7 by [(Niemann, 2023)]

The developed mathematical model, discussed in Section 4.3, has been tested via a simulator developed by Imperial College for this purpose. The simulator is designed to handle various configuration parameters (e.g. number of nearby vessels, geometry, range sensors) to demonstrate the developed approach in various operation conditions, this will be discussed in detail in Section 4.4. The investigation of user-level crowdsourcing integrity performance based on LiDAR and Radar range sensors is summarised in Section 4.5.

4.2 Automatic Identification System

The IMO, as per its 2015 guidelines (IMO,2015), specifies the mandatory and optional elements of the AIS, as follows:

- Mandatory Information
  - Ship's Position: The exact location of the ship, accompanied by an indication of the accuracy of this data and the integrity status of the positioning system.
  - Time in UTC: The Coordinated Universal Time at which the other data parameters are valid.
  - Course Over Ground: The direction in which the ship is moving over the bottom.
  - Speed Over Ground: The ship's speed is measured relative to the ground.
  - Heading: The direction towards which the ship is pointing.
  - Navigational Status: This includes various statuses like Not Under Command, at anchor, etc., and requires manual input.
  - Rate of Turn: This is provided where available, indicating how fast the ship is turning.
  - IMO Number (where available): A unique reference for ships and for registered ship owners and management companies.
  - Call Sign & Name: The ship’s radio call sign and its name.
  - Length and Beam: The dimensions of the ship.
  - Type of Ship: The specific type or classification of the ship.
  - Location of Position-Fixing Antenna: This should be specified in relation to the ship’s structure (aft of bow and port or starboard of centerline).
  - Ship’s Draught: The vertical distance between the waterline and the bottom of the hull.
  - Hazardous Cargo (type): Information on any hazardous cargo on board.
  - Destination and ETA: The ship’s destination and estimated time of arrival, are provided at the master’s discretion.
- Optional Information
  - Angle of Heel: The degree to which the ship is tilted to one side, provided where available.
  - Pitch and Roll: The up or down and side-to-side tilting of the ship, respectively, provided where available.
  - Route Plan: Information about the planned route of the ship.

4.3 Mathematical model

The functional model utilised in the crowd-sourced positioning layer is based on trilateration, a positioning fixing solution that leverages the coordinates of nearby vessels and measured ranges. The mathematical representation of the function models is given as follows:

$$r = \sqrt{(x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2} + \delta.c \quad (4.1)$$

where: r is the measured range to the nearby vessel, (x, y, z) is the coordinates of the unknown vessel, (x_n, y_n, z_n) are the coordinates of the nearby vessel, c is the speed of light, and δ is a free parameter representing the clock error. The minimum number of required nearby vessels is four.

The linearised form of 4.1 is given by:

$$f(x + \Delta x) \approx f(x) + J\Delta x \quad (4.2)$$

In this context, the weighted nonlinear Ordinary Least Squares (OLS) solution is given by:

$$\hat{x} = x + \Delta x \quad (4.3)$$

$$\Delta x = (J^t W J)^{-1} J^t W b \quad (4.4)$$

$$b = f(x_0) - BWb \quad (4.5)$$

$$W = Q_\iota^{-1} \quad (4.6)$$

where:$\hat{x}$ is the updated estimate of the parameters after applying the correction Δx, J is the Jacobian matrix, W is the weight matrix, B is the observation parameters, and Q_ι is the covariance matrix of the observations.

The covariance matrix of the observations (Q_ι) is a diagonal matrix comprised of the variances of the range measurements. These variances can be estimated through a combination of the sensor range variance ($\sigma_{i^2_{range\ sensor}}$) and the range error caused by the positional errors of nearby vessels, as shown in Figure (4.3). This can be expressed as:

$$\sigma_{i^2_{range}} = \sigma_{i^2_{range\ sensor}} + \sigma_{i^2_{range\ GNSS}} \quad (4.7)$$

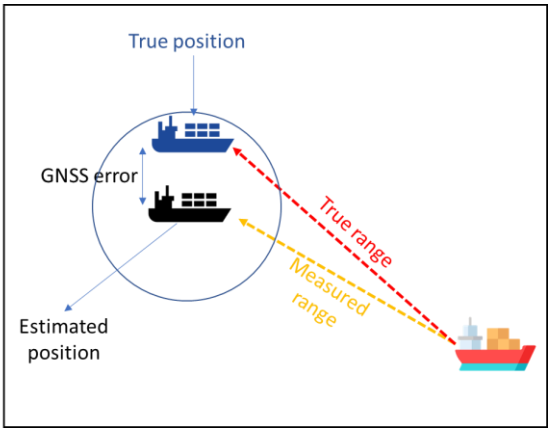


Figure 4.3: the range error component

The range error caused form nearby vessel GNSS positioning error can be estimated/propagated as:

$$\sigma_{i}^{2}{}_{gps} = \left(\frac{\partial r}{\partial x}\,\sigma_{x_i}\right)^2 + \left(\frac{\partial r}{\partial y}\,\sigma_{y_i}\right)^2 + \left(\frac{\partial r}{\partial z}\,\sigma_{z_i}\right)^2 \tag{4.8}$$

Following the position solution in equations (4.3), the Mean Squared Error (MSE) and covariance matrix of the estimated parameters($Q_{\hat{x}}$) are given by:

$$MSE = \frac{B' * (W - W * J * (J' * W * J)^{-1} * J' * W) * B}{m - n} \tag{4.9}$$

$$Q_{\hat{x}} = (J' * W * J)^{-1} * MSE \tag{4.10}$$

To transfer the covariance matrix of the estimated parameters $Q_{\hat{x}-WGS84}$ from Earth-Centered Earth-fixed (ECEF) coordinates to WGS84, rotation matrix (R) can be used as follows:

$$Q_{\hat{x}-WGS84} = R \times Q_{\hat{x}_{ij\ for\ i,j=1,2,3}} \times R' \tag{4.11}$$

where: $Q_{\hat{x}_{ij\ for\ i,j=1,2,3}}$ represents the subset of the first three columns and rows of $Q_{\hat{x}}$, encompassing the 3D positioning errors.

For FDE, employing the test statistics utilised for the all-in-view solution in GNSS (Blanch et al.,2012), the chi-square test statistic (χ^2)for the all-in-view set and the test threshold (T_{χ^2}) can be computed using the probability of false alarm (P_{fa}) as follows:

$$\chi^2 = y^t(B' * (W - W * J * (J' * W * J)^{-1} * J' * W) * B)y \tag{4.12}$$

$$F\left(T_{\chi^2}, degree\ of\ freedom\right) = 1 - P_{fa} \tag{4.13}$$

Utilising Equations (4.11) and (4.12), a fault is detected when the test statistic (χ^2) exceeds the threshold (T_{χ^2}). Subsequent to the FDE stage, the Horizontal P rotection Level (HPL) can be computed as follows:

$$HPL = \sqrt{PL_x^2 + PL_y^2} \tag{4.14}$$

$$PL_x = k_x * \sigma_x + p_1 * P_{fa} \tag{4.15}$$

$$PL_y = k_y * \sigma_y + p_2 * P_{fa} \tag{4.16}$$

$$proj._{bnominal} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = (J^t W J)^{-1} J^t W b_{nomial} \tag{4.17}$$

Where: $proj._{bnominal}$ is the projection nominal bias (b_{nomial}) in the position domain, $k_x\ and\ k_y$ are factors that reflect the probability of missed detection and required confidant interval, the latter estimated based on required integrity risk. In the equations (4.14,4.15), the σ_x and σ_y should be

an overbounded standard devision, which should be computed from an overbounded range standard division to esure the safety. Put differently, to compute the protection level, the input range parameters should be overbounded in equation (4.7,4.8).

4.4 Imperial College simulation platform

As previously mentioned, a simulation platform has been developed to evaluate the developed approach. This simulator incorporates a variety of configuration parameters, enabling the simulation of diverse scenarios under different operational conditions. The interface of the simulator is shown in Figure (4.4). The input configuration parameters and sensor accuracy are:

- **Configuration parameters**
  - Number of Nearby Vessels
  - Minimum and Maximum Distance: The range of distances between the vessel (of unknown position) and the nearby vessels.
  - Minimum Distance Between Nearby Vessels
  - Simulation Reference Locations (Window): Used to centre the simulation around this specific point.
  - Elevation Range: Refers to the elevation difference between vessels, crucial for enhancing the simulation's reliability.
  - Geometry: Includes three options; strong, weak, and random. This is vital for understanding the impact of geometry on computations and for accommodating various operational scenarios.
  - Number of Scenarios: This parameter offers the flexibility to generate any number of scenarios and save the results in CSV files, enabling the generation of a million of scenarios based on the selected configurations.
  - Output Filename: The name of the CSV file containing the results.
- **Sensor's accuracy:**
  - Nearby Vessels GNSS Positioning accuracy: These parameters are utilised to estimate the positions of nearby vessels and to model the range errors resulting from positional errors in the functional model, this includes two parameters:
    - Nearby vessels' GNSS positioning horizontal accuracy
    - Nearby vessels' GNSS positioning vertical accuracy
  - Range Sensors Accuracy: Used to simulate the ranges based on sensor accuracy, include:
    - Sensor type: consists of two options (Radar IMO standard, customise).
    - Sensor accuracy, in case the customise option is selected, the sensor accuracy can be inserted manually

The outputs of the simulation are summarised as follows:

- KML files present the simulated vessel's true position and estimated position, opened in Google earth. Figure (4.5) presents an example of the KML file output in google earth
- A web map presents the simulated vessels, as shown in Figure (4.6)

- Tables present a key output of the simulation, as shown in Figure(4.7). These tables include
    - nearby vessels estimated position (in WGS84)
    - nearby vessels' true position (in WGS84)
    - nearby vessels GNSS horizontal error
    - nearby vessels GNSS vertical error
    - estimated range
    - true range
    - range error
- Estimated position of the unknown vessels in WGS84 and ECEF coordinate system
- Horizontal and vertical estimated standard division from the least squares.
- Horizontal and vertical true error
- Horizontal protection level
- CVS files include the all results

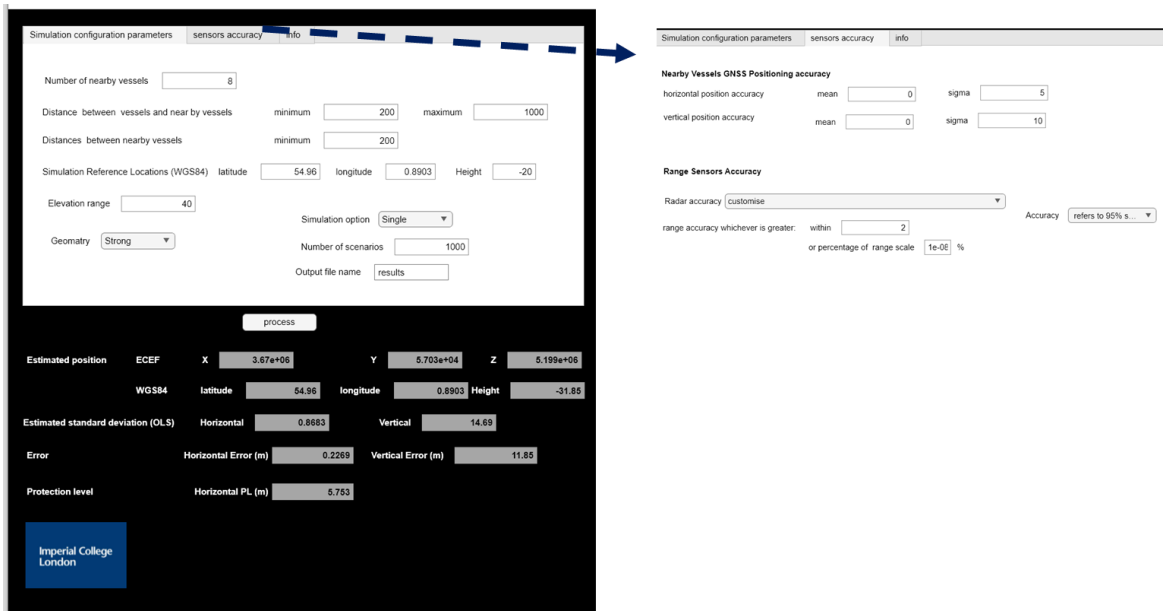


Figure 4.4: Imperial College simulation platform

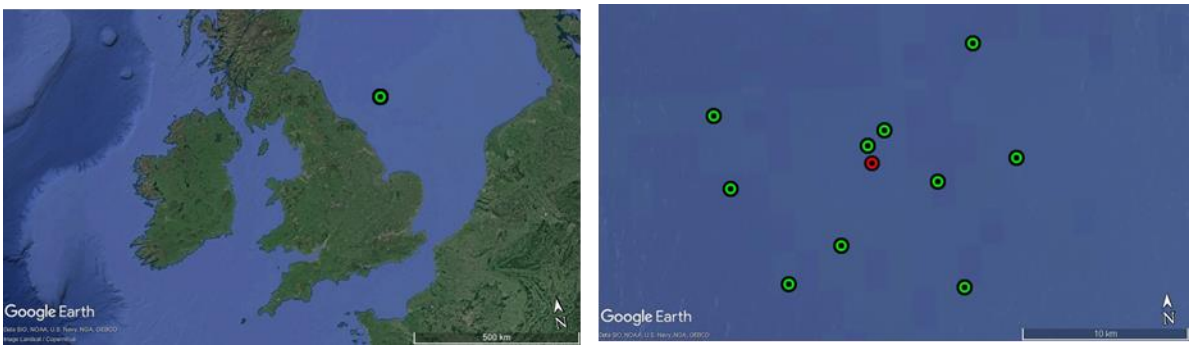


Figure 4.5: Sample of KML output file from the simulator

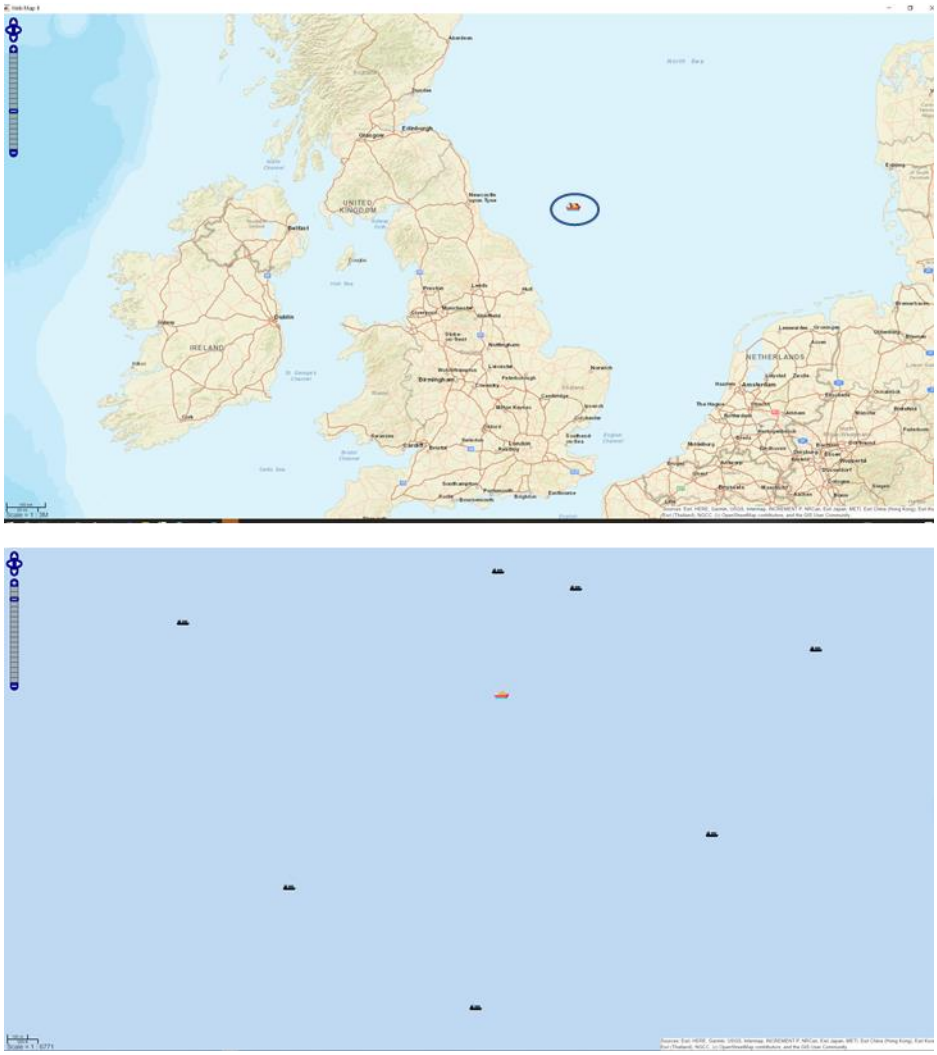
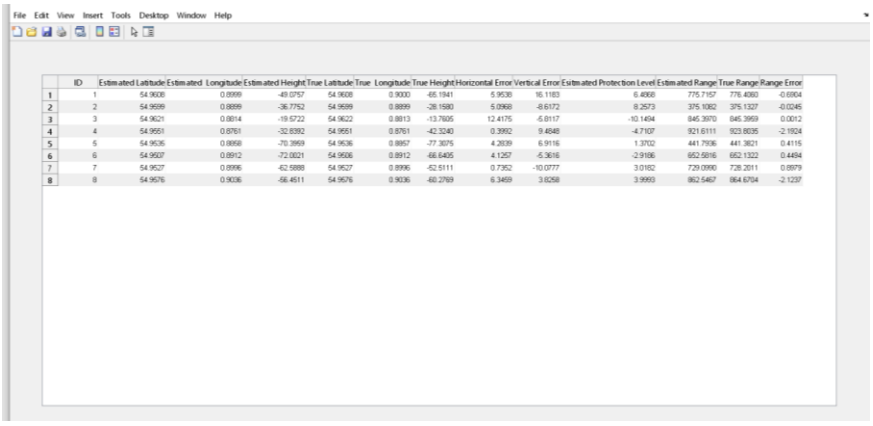


Figure 4.6: Sample of Web map generated from the simulator



	ID	Estimated Latitude	Estimated Longitude	Estimated Height	True Latitude	True Longitude	True Height	Horizontal Error	Vertical Error	Estimated Protection Level	Estimated Range	True Range	Range Error
1	1	54.9608	0.8999	-49.0757	54.9608	0.9000	-66.1941	5.9538	16.1183	6.4868	775.7157	776.4060	-0.6904
2	2	54.9599	0.8899	-36.7752	54.9599	0.8899	-28.1580	5.0968	-8.6172	8.2573	375.1082	375.1327	-0.0245
3	3	54.9621	0.8814	-19.5722	54.9622	0.8813	-13.7605	12.4175	-5.8117	-10.1494	845.3970	845.3969	0.0012
4	4	54.9551	0.8763	-32.8382	54.9551	0.8761	-42.3240	0.3992	9.4948	-4.7107	921.6111	923.8035	-2.1924
5	5	54.9536	0.8858	-70.3959	54.9536	0.8857	-77.3075	4.2839	6.9116	1.3702	441.7936	441.3821	0.4115
6	6	54.9507	0.8912	-72.0021	54.9506	0.8912	-66.6405	4.1257	-5.3616	-2.9106	652.5816	652.1322	0.4494
7	7	54.9527	0.8996	-62.5888	54.9527	0.8996	-52.5111	0.7352	-10.0777	3.0182	729.0990	728.2011	0.8979
8	8	54.9576	0.9036	-56.4511	54.9576	0.9036	-60.2769	6.3459	3.8258	3.9993	862.5467	864.6704	-2.1237

Figure 4.7: Output table presents a key output of the simulation

4.5 User-level crowd-sourcing integrity performance

This section conducts an in-depth analysis of the developed approach, taking into account a variety of configuration parameters and range accuracies. The results derived from the Radar-based positioning solution are discussed in Section 4.5.1, while the Lidar-based solution is discussed in Section 4.5.2.

4.5.1 Evaluation metrices

The evaluation matrices are used in this section; protection level, position estimated horizontal standard division, and availability – in case n number of nearby vessels achieved. The protection level has been computed using equations (4.14-4.16), assuming the $k_x$and k_y factors equal by 6.625, and nominal bias by 10% of the estimated range error. Since the k factors reflect the probability of missed detection and required confidant interval, it has been assumed by multiplication of 4.4167 (factor represents the confidant interval at integrity risk 10^{-5}) and 1.5 (factor reflects the probability of missed detection). The former factor is content based on the integrity risk requirements; the latter is an assumption that can cover the worst-case scenario and requires future investigation to precisely estimate this factor. In addition, the P_{fa} refers to the continuity budget allocated to the chi-square test, which is assumed to be 10^{-8} , as in the development of GNSS ARAIM (Blanch et al.,2012).

The availability ($\text{availability}_{\text{n}}$) in this context does not equate to system availability; instead, it refers to the crowd-sourcing availability given that the input number of vessels is met. The overall crowd-sourcing availability ($\text{availability}_{\text{crowd sourcing}}$) can be calculated using the probabilities of having a specific number of vessels $P(n_{vessels})$ and the availability results obtained from n number of nearby vessels, and can be given by:

$$\text{availability}_{\text{crowd sourcing}} = \sum_{n=5}^{k-1} P(n_{vessels}) \times \text{availability}_{\text{n}} + P(n_{vessels} \geq k) \times \text{availability}_{\text{k}} \quad (4.18)$$

where: k is the maximum number of vessels that can be utilised within the computation, which can be identified based on the maximum required computational power/process.

When the developed approach in used besides the GNSS, without applying the integration in the measurements domain, the total system availability then can be given by:

$$\text{availability}_{\text{system}} = \text{P}(\text{crowd sourcing ava.}) + \text{P}(\text{GNSS ava.}) - P(\text{crowd sourcing ava} \cap \text{GNSS ava}) \quad (4.19)$$

where: the $\text{P}(\text{crowd sourcing ava.})$ is the total crowd-sourcing availability, and $\text{P}(\text{GNSS ava.})$ is the GNSS availability.

The above availability models can be applied in various alarm limits/operation phases. In this section, the availability ($\text{availability}_{\text{n}}$) is evaluated for three alarm limits 2.5 m, 5 m, and 25 m. Given that the pseudorange GNSS stand-alone approach in INSPIRe cannot meet the required alarm limits during the port phase, and considering that the requirements in the maritime sector should be achievable, this section assesses a 5-meter alarm limit in addition to the 2.5 meters alarm limit (required for port phase) and a 25 alarm limit (required for coastal, ocean, and HEA and restricted water phases).

4.5.2 Radar-based crowd-sourcing

This section evaluates the developed approach using two classes of radar range accuracies: the IMO standard, as specified in RESOLUTION MSC.192(79), and the FURUNO X-S-band radar models FR-2115-B, 2125-B, 2155-B, 2135S-B. The latter represents radars that offer higher accuracy than the standard IMO specification. Table (4.1) presents the radar accuracies used in evaluating the proposed approach.

Table 4.1: radar accuracies used in evaluating the proposed approach.

Radar	Range accuracy
IMO standard, as specified in RESOLUTION MSC.192(79)	within 30 m or 1% of the range scale, whichever is greater;
FURUNO (X-S-band radar) Models:FR-2115-B/2125-B/2155-B*/2135S-B	within 15 m or 1% of the range scale, whichever is greater;

The investigation was conducted across three scenarios with three cases of nearby vessel numbers (5, 6, and 7), simulating 100 'strong geometry' scenarios for each, in a total of 1200 scenarios. The simulated distances between the vessel and nearby vessels ranged from 100 to 1000 meters, selected to be within ranges that radar can provide the highest level of accuracy, as shown in Figure (4.8).

Table (4.2) summaries the average positioning standard deviation, protection level, and availability at the three-alarm limits for the tested scenarios. Figures (4.9,4.10) present the improvement in protection level and position accuracy with increasing the number of nearby vessels in each tested scenario. The findings indicate that the radar-based crowdsourcing approach ends up with a low level of accuracy and does not significantly enhance maritime navigation. The highest availability, achieved using the FURUNO radar with 10 nearby vessels at a 25-meter alarm limit, is quite low (lower than 25%). Furthermore, the average protection level across all cases was substantially higher than the required ones.

Table 4.2: summary of the average positioning standard deviation, protection level, and availability at the three-alarm limits

Lidar range accuracy (95%)	Number of vessels	Integrity risk	Position standard division (m) – average	PL (m) – average	Availability – in case n number of nearby vessels achieved		
					Alarm limit 25 (m)	Alarm limit 5 (m)	Alarm limit 2.5 (m)
IMO	5	10^{-5}	14.40	95.37	<15%	<5%	<4%
	6	10^{-5}	13.65	90.48	<15%	<5%	<4%
	10	10^{-5}	11.42	75.69	<15%	<5%	<4%
FURUNO	5	10^{-5}	8.33	52.61	<25%	<5%	<4%
	6	10^{-5}	7.52	49.85	<25%	<5%	<4%
	10	10^{-5}	5.68	37.63	<25%	<5%	<4%

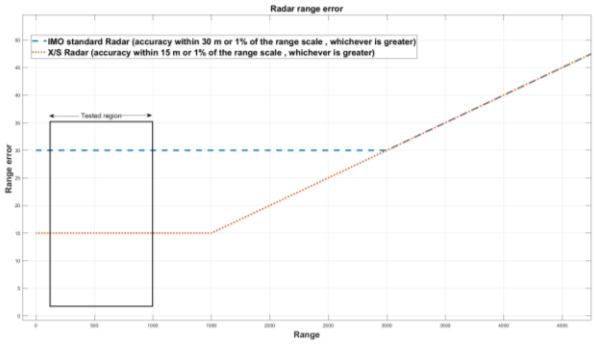


Figure 4.8: Radar Accuracy for the IMO standard radar and accurate X/S radar

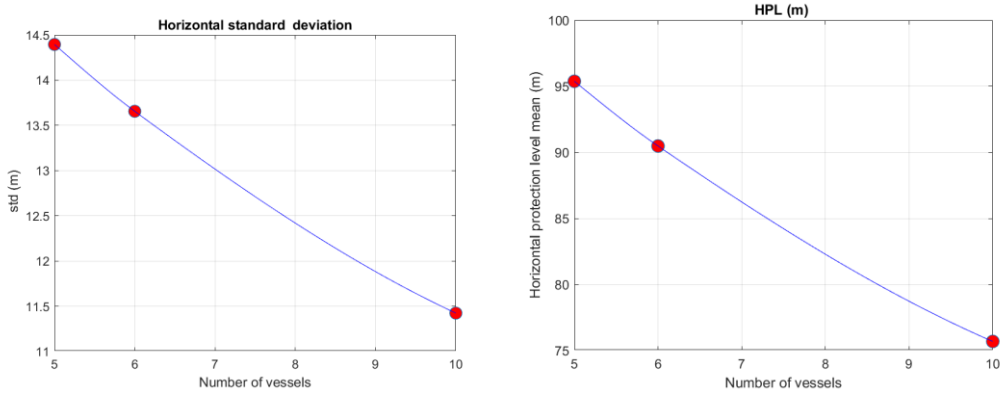


Figure 4.9: average horizontal standard deviation and protection level using IMO standard Radar (accuracy within 30 m or 1% of the range scale, whichever is greater)

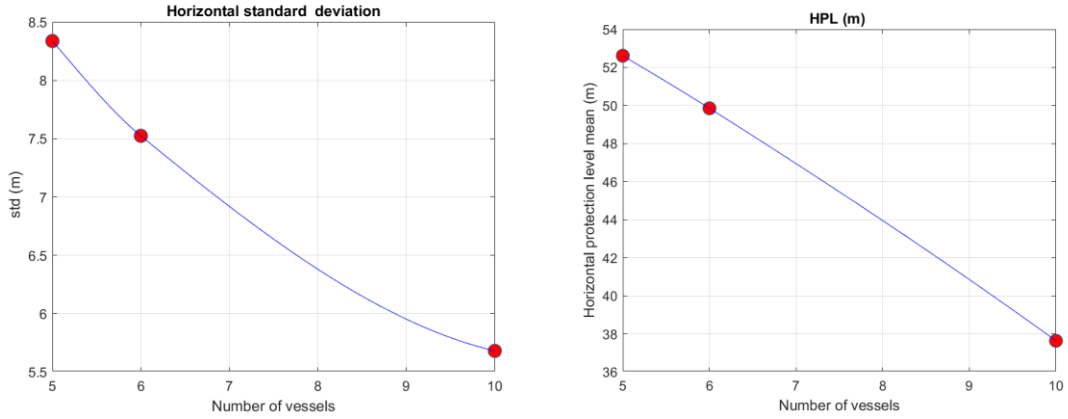


Figure 4.10: average horizontal standard deviation and protection level using X/S Radar (accuracy within 30 m or 1% of the range scale, whichever is greater), distance to nearby vessels ranges 100-1000

4.5.3 LIDAR-based crowd-sourcing

This section tests the developed approach using Lidar across various accuracy levels (0.5, 1, 1.5, and 2m at 95% CL). The initial investigation involves all four accuracy levels across three cases of nearby vessel numbers (5, 6, and 7), with 100 strong geometry mode scenarios simulated for each case, in a total of 1200 cases. The simulated distance between the vessel and nearby vessels ranged from 100 to 1000m.

Table (4.3) provides a summary of the average positioning standard deviation, protection level, and availability at the alarm limits (25, 5, and 2.5), based on the four lidar range accuracy levels (0.5, 1, 1.5, and 2m at 95%CL). Figures (4.8-4.11) present the improvement in protection level and position accuracy with an increase in the number of nearby vessels in each tested scenario.

Table 4.3: summary of the average positioning standard deviation, protection level, and availability at the three-alarm limits

Lidar range accuracy (95%)	Number of vessels	Position standard division mean (m)	PL mean (m)	Availability – in case n number of nearby vessels achieved		
				Alarm limit 25 (m)	Alarm limit 5 (m)	Alarm limit 2.5 (m)
2	5	1.36	9.03	95%	36%	<20%
	6	1.23	8.17	96%	45%	<20%
	10	0.82	5.46	99%	57%	<20%
1.5	5	1.19	7.89	97%	52%	23%
	6	0.93	6.13	98%	48%	12%
	10	0.65	4.28	99%	88%	24%
1	5	0.80	5.30	97%	70%	50%
	6	0.60	3.99	100%	76%	37%
	10	0.38	2.49	100%	99%	56%
0.5	5	0.50	3.33	99%	86%	63%
	6	0.47	3.14	99%	93%	75%
	10	0.31	2.05	100%	99%	95%

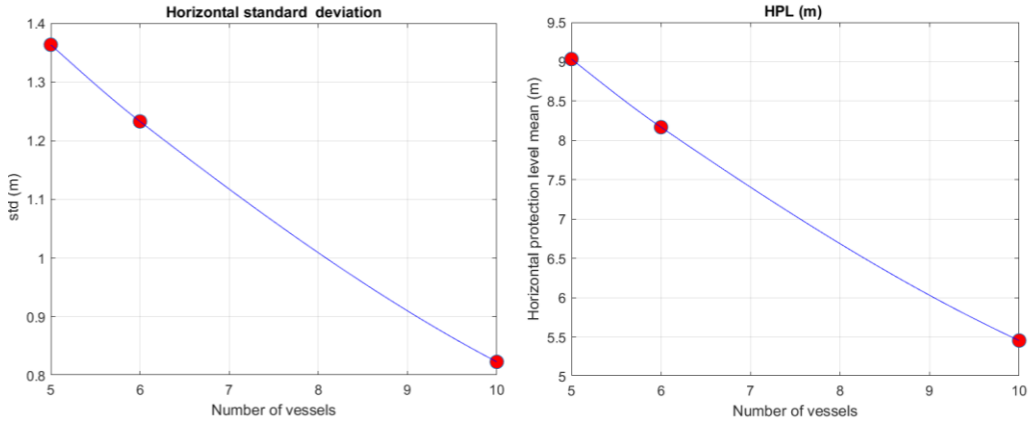


Figure 4.11: Average Horizontal Standard Deviation and horizontal protection level using Lidar with 2m Accuracy (95%)

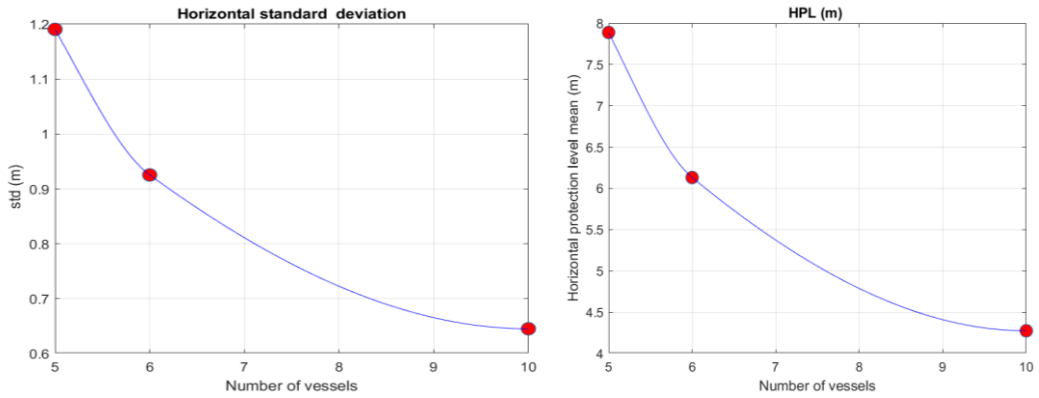


Figure 4.12: Average Horizontal Standard Deviation and horizontal protection level using Lidar with 1.5 m Accuracy (95%)

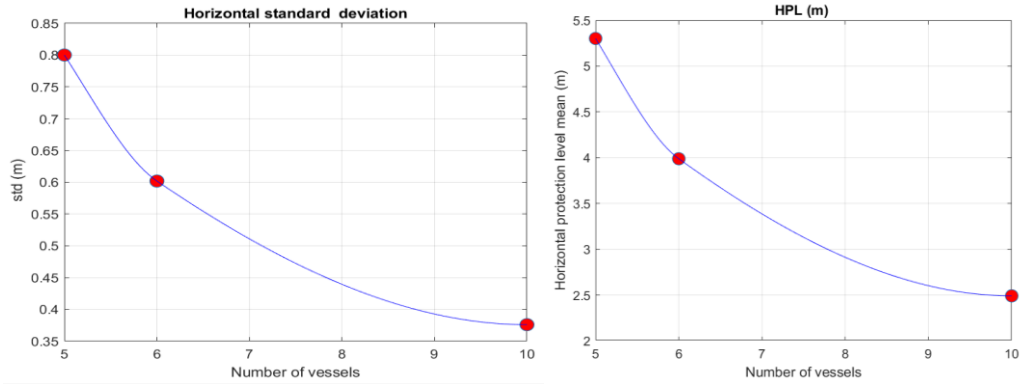


Figure 4.13: Average horizontal standard deviation and horizontal protection level using lidar with 1m accuracy (95%)

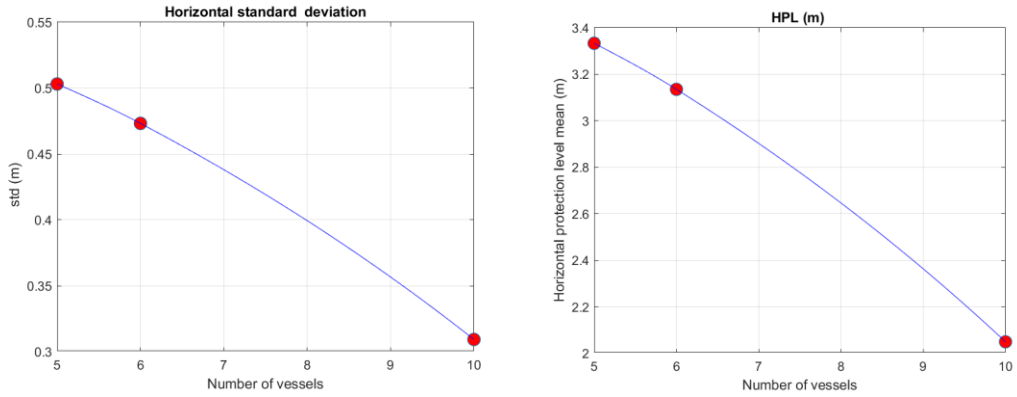


Figure 4.14: Average Horizontal Standard Deviation and protection using Lidar with 0.5m Accuracy (95%)

The results show that the average protection level and standard deviation vary based on the number of vessels and range accuracy, ranging from 2m to 9m for protection level, and 0.3m to 1.4m for standard deviation. The results demonstrate that the developed approach delivers a high-performance level at a 25-meter alarm limit and appears promising for 2.5 and 5-meter alarm limits when the range is measured with better than 1-meter accuracy (95%).

The results above stem from an initial investigation involving 100 scenarios for each configuration, intended to identify a suitable level of range accuracy for use in the developed model. Given that the results of the Lidar with less than 1 m accuracy (95%) are seen as promising for a port phase.

The performance of the developed approach depends on several factors, including but not limited to range accuracy, nearby vessels' position accuracy, their distribution (geometry), and distances to nearby vessels. The latter is significant and can significantly impact the developed approach performance in certain cases. This factor is generally overlooked in GNSS positioning because the distances between receivers and satellites are relatively within similar ranges. In more detail, in previous crowdsourcing tests involving vessels within a 100–1000 meters range, thus, the range from some nearby vessels might be ten times greater than others, potentially impacting system performance. Such scenarios, with ranges ten times greater than others, do not occur in GNSS. Therefore, the distances between the vessel and nearby ones should be carefully considered in positioning. Put differently, this factor is one dimension that should be considered alongside classical factors like geometry and range measurements.

To evaluate the impact of distances to nearby vessels, an additional test was performed using distances to nearby vessels between 100-300 meters, and the results were compared with the previous experiments (based on distances of 100-1000 meters). The tests utilised LiDAR with 0.5-meter accuracy (95%) in strong geometry mode. The findings revealed that when the distance range to nearby vessels was reduced to closer ranges (100-300 meters), there was a remarkable improvement in system performance across all metrics, as presented in Table (4.4). Figure (4.15) presents the 100-300 m range results.

Table 4.4: Outputs for average std, protection level, and availability for distance ranges 100-300 and 100-1000 m.

Lidar range accuracy (95%)	Distance range to nearby vessels (m)	Number of vessels	Position standard division mean (m)	PL mean (m)	Availability – in case n number of nearby vessels achieved		
					Alarm limit 25 (m)	Alarm limit 5 (m)	Alarm limit 2.5 (m)
0.5	100-1000	5	0.50	3.33	99%	86%	63%
		6	0.47	3.14	99%	93%	75%
		10	0.31	2.05	100%	99%	95%
0.5	100-300	5	0.42	2.84	99%	95%	69%
		6	0.28	1.87	100%	98%	75%
		10	0.23	1.58	100%	99.5%	96%

Further investigation was conducted to test the geometry impact using LiDAR with a range accuracy of 0.5 meters (95%), using distances from nearby vessels ranging from 100-300 meters. The results indicate that geometry significantly impacts computational performance, as shown in Table (4.5). Figures (4.15, 4.16) present the average horizontal standard deviation and protection level for the weak and strong geometry solutions.

Table 4.5: Outputs from crowdsourcing solutions under weak and strong geometry conditions.

Lidar range accuracy (95%)	Geometry	Distance range to nearby vessels (m)	Number of vessels	Position standard division mean (m)	PL mean (m)	Availability – in case n number of nearby vessels achieved		
						Alarm limit 25 (m)	Alarm limit 5 (m)	Alarm limit 2.5 (m)
0.5	Strong	100-300	5	0.43	2.84	99%	95%	69%
			6	0.28	1.87	100%	98%	75%
			10	0.23	1.58	100%	99.5%	96%
0.5	Weak		5	1.99	13.19	97%	57%	31%
			6	1.39	9.22	93%	65%	33%
			10	0.75	4.97	96%	93%	57%

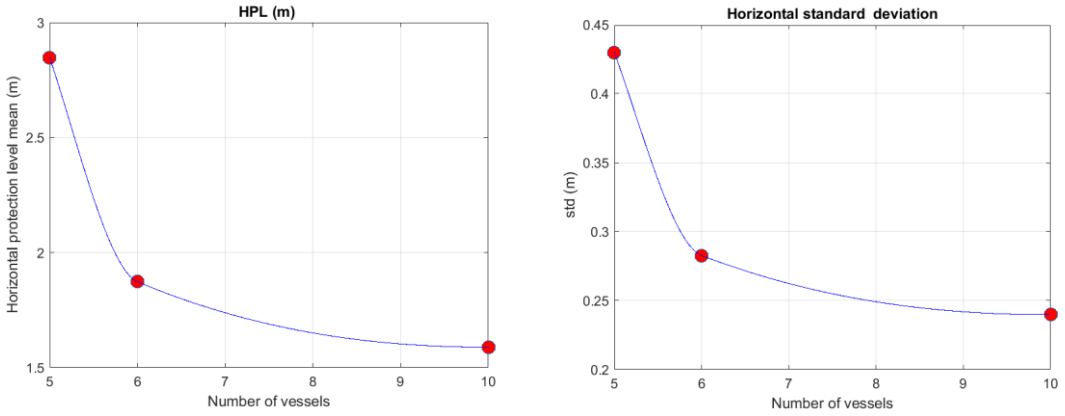


Figure 4.15: Average Horizontal Standard Deviation and Protection Level using LiDAR with 0.5m Accuracy (95%) for distances of 100-300 meters to nearby vessels in strong geometry mode.

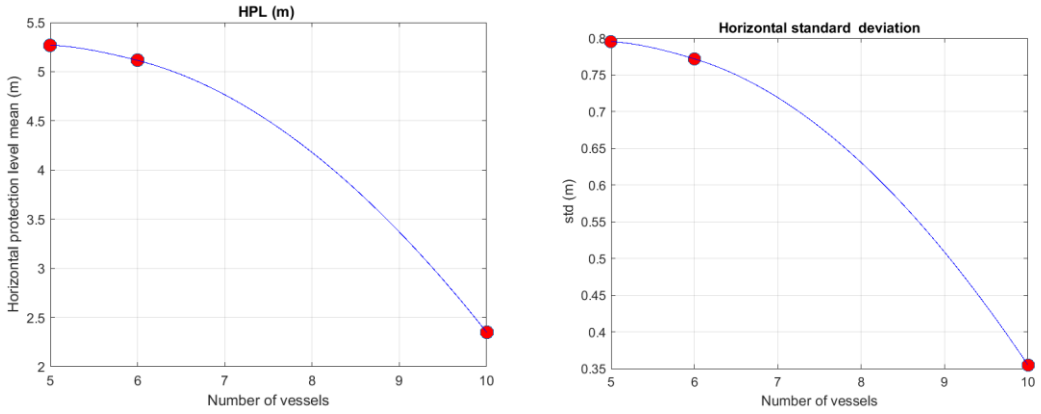


Figure 4.16 Average Horizontal Standard Deviation and Protection Level using LiDAR with 0.5m Accuracy (95%) for distances of 100-300 meters to nearby vessels in weak geometry mode.

The aforementioned results demonstrate that the geometry and distances to nearby vessels impact the positioning performance. This suggests using a measurement selection framework based on these factors to maintain high performance, especially when there is a high number of

nearby vessels and a solution separation approach is not employed to address multiple simultaneous failures.

In addition to the above mentioned tests, in-depth testing was conducted on LiDAR systems with 1-meter and 0.5-meter accuracy levels (95% CL). This testing involved three scenarios with varying numbers of nearby vessels: 5, 6, and 10. For each scenario, 10,000 simulations were generated, totaling 60,000 simulations. The tests were conducted with nearby vessel ranges of 200 to 1000 meters, under strong geometry conditions. The outcomes of these tests are summarised in Table (4.6). These results, confirming the initial investigation reutls, demonstrate the potential of the developed approach in enhancing maritime navigation.

Table 4.6: Summary of the Average Positioning Standard Deviation, Protection Level, and Availability at Three Alarm Limits. This table presents data from 10,000 scenarios for each specified number of nearby vessels, amounting to a total of 60,000 scenarios.

Lidar range accuracy (95%)	Number of vessels	Position standard division mean (m)	PL mean (m)	Availability – in case n number of nearby vessels achieved		
				Alarm limit 25 (m)	Alarm limit 5 (m)	Alarm limit 2.5 (m)
1	5	0.6496	4.3044	0.9833	0.6985	0.4099
	6	0.5909	3.9152	0.9890	0.7749	0.4100
	10	0.4108	2.7222	0.99	0.9654	0.5306
0.5	5	0.3816	2.5286	0.9885	0.8852	0.6778
	6	0.3486	2.3100	0.9896	0.9342	0.7524
	10	0.3400	2.2525	0.9904	0.9368	0.7608

4.6 Summary

This section introduces a novel crowdsourcing positioning framework with an integrated integrity monitoring layer. The approach underwent testing with a simulator developed by Imperial College, designed to handle various range accuracy/sensors (e.g. Radar and Lidar) and operational conditions such as geometry, and distance range to nearby vessels. Key outputs of the simulator include estimated position accuracy and horizontal protection level. In addition, the simulator was developed to manage a large number of scenarios.

The approach was subjected to 30 experiments (in total 62,400 scenarios), each experiment incorporating different sensors with their accuracy levels and various configuration parameters that reflect operational conditions, including geometry and the distance range to nearby vessels. Three evaluation metrics were employed to assess the results: protection level, positioning accuracy, and availability at three alarm limits (2.5 m, 5 m, 25 m). As noted in Section 4.5.1, the availability in this context does not equate to system availability; instead, it refers to the crowd-sourcing availability given that the input number of vessels is met.

Table (4.6) summarises the results of the 24 experiments, including average standard deviation, protection level, and availability at the specified alarm limits. The experiment's aims and their findings can be summarised as follows:

- **Experiments 1-6:** These cases were set up for an initial assessment of Radar-based crowdsourcing positioning using the radar as IMO standard (as specified in RESOLUTION MSC.192(79)) and FURUNO Radar (FR-2115-B, 2125-B, 2155-B, 2135S-B) that offer higher accuracy than the standard IMO specification. The results of these experiments show that the Radar-based crowdsourcing positioning provides a low level of performance, compared with the required level within the maritime phases.
- **Experiments 7-18:** These cases investigated LiDAR-based crowdsourcing positioning across various LiDAR accuracy levels (0.5, 1, 1.5, and 2 meters at 95% confidence level). The investigation covered all four accuracy levels across three scenarios of nearby vessel numbers (5, 6, and 7), with 100 strong geometry mode scenarios simulated for each scenario. The results suggest that this approach can deliver high performance at a 25-meter alarm limit and shows promise for 2.5 and 5-meter alarm limits when the range is measured with better than 1-meter accuracy (95%).
- **Experiments 19-21**: These cases assessed the influence of the distance to nearby vessels on the developed approach's performance. The findings from these cases underscore the significant impact of this factor on positioning accuracy.
- **Experiments 22-24:** These cases assessed the impact of the geometry factor. The outcomes of these experiments demonstrate that geometry factors play a significant role in positioning performance. Therefore, the geometry factor should be considered for measurement selection when solution separation is not applied. These aspects will be discussed in more detail in the implementation plan and future work in Section (5.2).
- **Experiments 25-30:** These cases investigated in-depth LiDAR-based crowdsourcing positioning across two LiDAR accuracy levels (0.5 and 1 meter at 95% confidence level). The investigation covered two accuracy levels across three scenarios of nearby vessel numbers (5, 6, and 7), with 10,000 strong geometry mode scenarios simulated for each scenario. The results confirm the initial investigation results, suggesting that this approach can deliver high performance at a 25-meter alarm limit and shows promise for 2.5 and 5-meter alarm limits.

ID	Target range sensor	Range accuracy (m) (95%)	Aim	Number of nearby vessels	Distances range between the vessel and nearby ones (m)	Geometry		Number of Seniors	Std	HPL	Availability		
						Strong	Weak				Alarm limit 25 (m)	Alarm limit 5 (m)	Alarm limit 2.5 (m)
1	Radar (IMO standard)	within 30 m or 1% of the range scale, whichever is greater;	Investigation Radar-based	5	100-1000	✓		100	14	95	<15%	<5%	<4%
2				6	100-1000	✓		100	13	90	<15%	<5%	<4%
3				10	100-1000	✓		100	11	75	<15%	<5%	<4%
4	Radar (FURUNO)	within 15 m or 1% of the range scale , whichever is greater;	investigation Radar-based	5	100-1000	✓		100	8	52	<25%	<5%	<4%
5				6	100-1000	✓		100	7	49	<25%	<5%	<4%
6				10	100-1000	✓		100	5	37	<25%	<5%	<4%
7	Lidar	2	Investigation Lidar-based	5	100-1000	✓		100	1.36	9.03	95%	36%	<20%
8				6	100-1000	✓		100	1.23	8.17	96%	45%	<20%
9				10	100-1000	✓		100	0.82	5.46	99%	57%	<20%
10		1.5	Investigation Lidar-based	5	100-1000	✓		100	1.19	7.89	97%	52%	23%
11				6	100-1000	✓		100	0.93	6.13	98%	48%	12%
12				10	100-1000	✓		100	0.65	4.28	99%	88%	24%
13		1	Investigation Lidar-based	5	100-1000	✓		100	0.80	5.30	97%	70%	50%
14				6	100-1000	✓		100	0.60	3.99	100%	76%	37%
15				10	100-1000	✓		100	0.38	2.49	100%	99%	56%
16		0.5	Investigation Lidar-based	5	100-1000	✓		100	0.50	3.33	99%	86%	63%
17				6	100-1000	✓		100	0.47	3.14	99%	93%	75%
18				10	100-1000	✓		100	0.31	2.05	100%	99%	95%

ID	Target range sensor	Range accuracy (m) (95%)	Aim	Number of nearby vessels	Distances range between the vessel and nearby ones (m)	Geometry		Number of Seniors	Std	HPL	Availability		
						Strong	Weak				Alarm limit 25 (m)	Alarm limit 5 (m)	Alarm limit 2.5 (m)
19	Lidar	0.5	Evaluating 'distances to nearby vessels' factor impact	5	100-300	✓		100	0.42	2.85	99%	95%	69%
20				6	100-300	✓		100	0.28	1.87	100%	98%	75%
21				10	100-300	✓		100	0.24	1.58	100%	99.5%	96%
22			Evaluating Geometry factor impact	5	100-300		✓	100	1.99	13.19	97%	57%	31%
23				6	100-300		✓	100	1.39	9.22	93%	65%	33%
24				10	100-300		✓	100	0.75	4.97	96%	93%	57%
25		1	In depth investigation	5	200-1000	✓		**10,000**	0.64	4.30	98.3%	69.8%	40%
26				6	200-1000	✓		**10,000**	0.59	3.91	98.9%	77%	41%
27				10	200-1000	✓		**10,000**	0.41	2.72	99%	96%	53%
28		0.5	In depth investigation	5	200-1000	✓		**10,000**	0.38	2.52	98.8%	88%	67%
29				6	200-1000	✓		**10,000**	0.34	2.31	98.9%	93%	75%
30				10	200-1000	✓		**10,000**	0.34	2.25	99%	94%	76%

5. Crowdsourcing as an e-Navigation Service

One of the significant drawbacks of the proposed crowd-sourcing architecture is the reliance on the AIS system to transfer position and integrity data between vessels. AIS is based on a published standard, the signal is open entirely un-encrypted, and broadcasting of spoofed or misleading AIS messages is actually quite trivial. Corruption of AIS messages by noise or interference in the channel is also possible, as only simple checksums are applied and very little message authentication is provided.

Furthermore, reliance on the AIS reported position of other vessels is not recommended for the purposes of collision avoidance, and no provision is made in the COLREGS [IMO Convention on the International Regulations for Preventing Collisions at Sea, 1972] for the use of AIS information at sea. Mis-use of AIS information as a sole means either for situational awareness, or for enacting collision avoidance may be considered a form of *VHF assisted collision* [IMO AIS Guidelines A.1106(29)]. It is recommended [MCA Marine Guidance Notice MGN324 (M+F) Amendment 1] that mariners may use AIS in conjunction with other information to assist with situational awareness and collision avoidance, but reliance on it as a sole means should be avoided.

It is proposed that a more secure, and reliable method of transferring position and integrity information between vessels would be to implement this data transfer as an *e-Navigation Service* within the Maritime Connectivity Platform (MCP).

The MCP is a conceptual, carrier agnostic, data communications and dissemination platform for sending electronic data between ships and shore-based infrastructure. This data is intended to underpin a number of automated e-Navigation services, with individual users electing to subscribe to receive data based on their navigation needs or geographical location.

The MCP itself consists of three components:

1. Maritime Identity Registry (MIR) serves the security of the MCP by containing a registry of unique identities for all maritime vessels and users and also acts as the certificate authority for the public-private-key cryptography, which underpins all communications over the MCP.

2. Maritime Service Registry (MSR), performs the same task of maintaining an identity registry for service providers, and enables the maritime user to look up and subscribe to various e-Navigation services.

3. Maritime Messaging Service (MMS) provides the architectural interface between shore-based service providers, and the MCP user vessels.

All three of these aspects is important for the safe dissemination of integrity information between users, so we shall spend some time to describe each.

Figure 5.1: Basic Architecture of the MCP

The Maritime Identity Registry (MIR) records the unique identity of every entity (ship, device, organisation or service provider) that will access the MCP. It also provides the public key infrastructure (PKI) for a system of public-private-key cryptography to ensure data messages are secure, authenticated, and end-to-end encrypted.

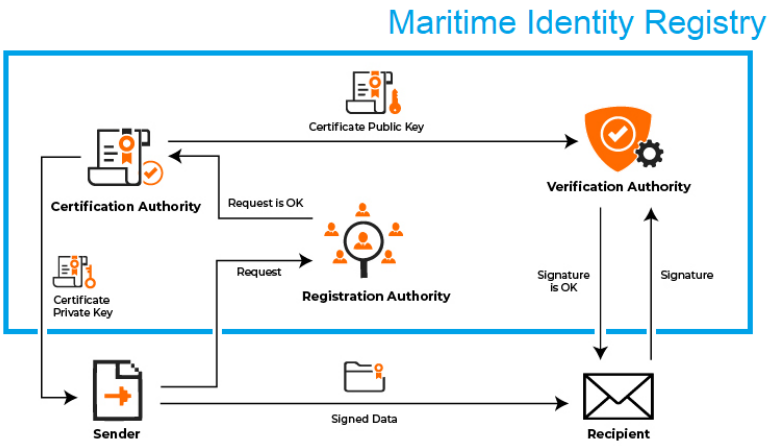


Figure 5.2: Basic public key infrastructure, showing the MIR boundary

The functional operation of the MIR is out of the scope of this project, and it will not be described in full. It interfaces to the user by providing private key certification, and also the public key infrastructure to verify the authenticity of the broadcast data. This ensures two things:

1. The positioning and integrity information is provided in encrypted messages along a secure channel, ensuring no corruption of the data can happen during transmission. Should the broadcast message become corrupted, the decryption service can detect this with a very high level of reliability.

2. The identify of the sender can be uniquely verified, making unauthorised transmission of false or misleading information to the user virtually impossible to achieve by a 3rd party.

The Maritime Service Registry (MSR) provides descriptions of the various e-Navigation services that are available to the mariner in both human and machine-readable formats. It is effectively a phone book; app store; or portfolio of available services open to the users of the MCP to discover and to which they may subscribe. It also includes machine-readable service specifications that detail how the e-Navigation service functions, and how the required data is broadcast and managed [IALA guideline 1128]. Three main parts of the service description are as follows:

1. Specification: A technology-agnostic description of a service on a logical level.
2. Technical Design: A description of the technology-bound, actual realization of a service on a technical level.
3. Instance: Information about the actual URI and other relevant data about a specific running service instance.

The e-Navigation Service user will access the MSR to obtain this information in order to establish an instance of the Crowd Sourced Integrity service it describes. This access will also be via the identity registry, and end-to-end encrypted for security.

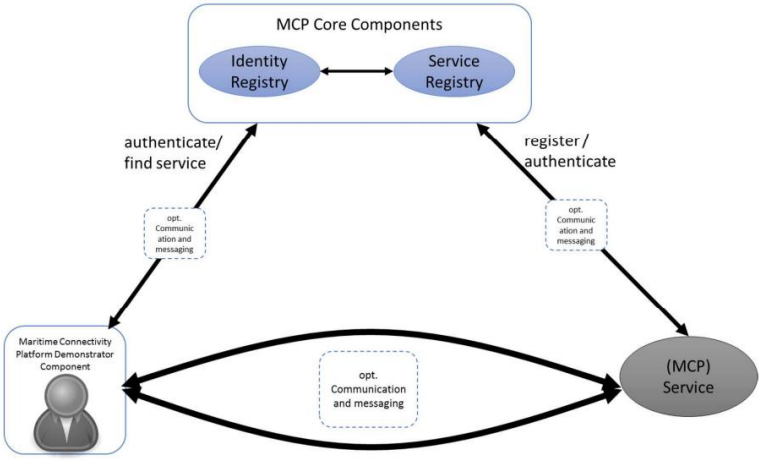


Figure 5.3: Basic functional design of an MCP user accessing an e-navigation service.

This report will not go into further details of the functional aspects of the MSR, since this is strictly out of scope of the INSPIRe project. More details of the operation of the three components of the MCP can be found in the Efficiensea2 website: (https://efficiensea2.org).

The MCP, and associated e-Navigation services are built upon a number of physical data communications links, any and all of which may be employed interchangeably to facilitate data transfer to and from the ship. The system is designed to be carrier agnostic, and whichever communication route is available to the vessel at any one time will be employed to transfer data from ship-to-ship and from ship-to-shore.

- VHF Data Exchange System (VDES), makes use of a number of marine VHF frequency channels to broadcast data from ship to ship and ship to shore. It builds upon the Automatic Identification System (AIS), which uses VHF to communicate a ship's identity, position and routing information among other data. VDES itself has two forms:
    - Terrestrial: VDE-TER uses two blocks of four adjacent VHF channels for data exchange when within range of VHF communications.
    - Satellite: VDE-SAT uses wider blocks of the VHF band for communications when outside VHF range of a base-station, and for higher bandwidth data backhaul.
- Mobile telecoms. When within range of shore-based conventional LTE (3G, 4G, 5G) data communications transmitters, the vessel will be able to exchange data securely over the internet via IP.
- Satellite telecoms. When outside of coverage of conventional mobile telecoms, the vessel will be able to make use of satellite-based communications. This route induces additional call-and-response (ping) delays and may be unsuitable for some applications.

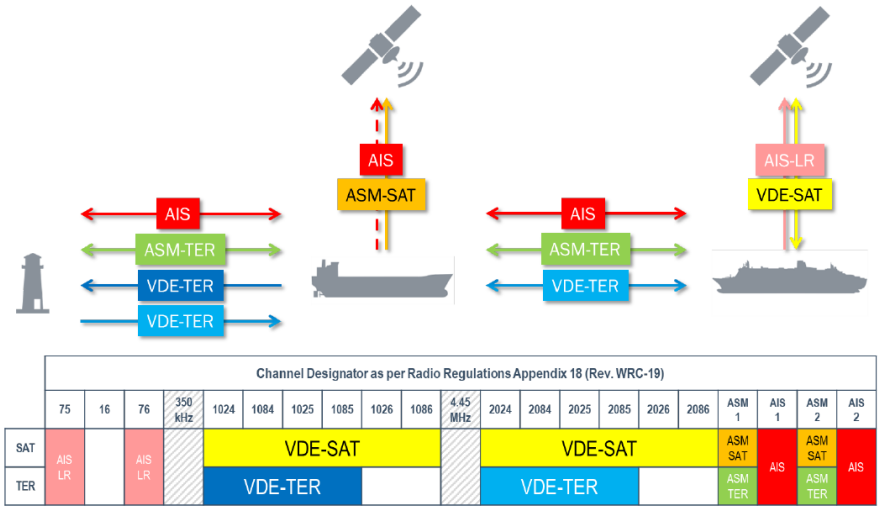


Figure 5.4: The VHF Data Exchange System (VDES)

Modified hardware is required on-board the ship to interface to the e-Navigation service, acquire the integrity information, and pass it to the navigation integrity processor. The existing communications networks between the various users receiver will automatically route and deliver the data.

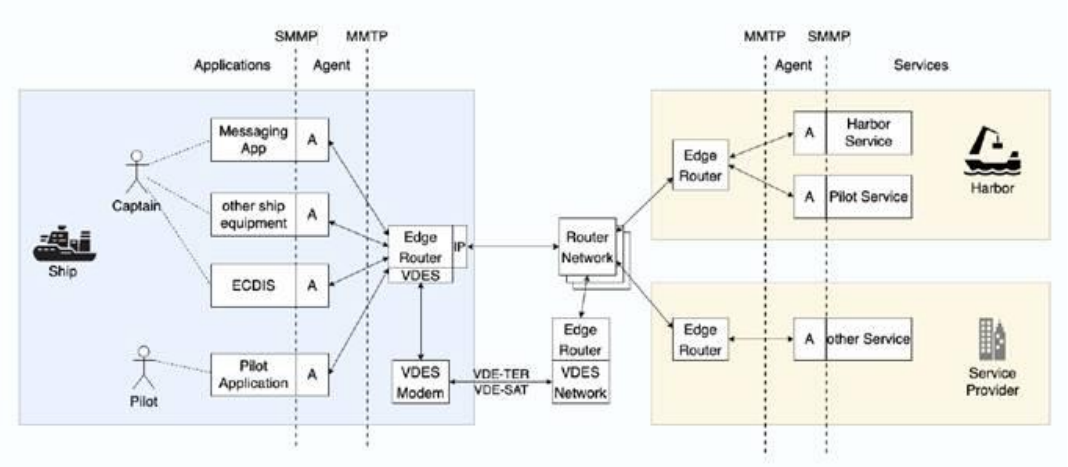


Figure 5.5: e-Navigation Maritime Messaging Service (MMS) Architecture, showing both VDES and conventional mobile telecoms links between a ship and shore-based service providers.

There is some concern over the use of either VDE-SAT or conventional satellite communications (Iridium, or satellite telecoms such as One Web). The issue is that the transportation network may not be able to guarantee a delivery time for the data. This is particularly a problem for VDE-SAT, which operates store-and-forward packet routing and may result in several seconds delay between the integrity data being sent and being received by the user vessel.

Communications via a dedicated VDES link would be the most reliable way to disseminate the data, but there may be concerns over the available bandwidth, especially if new position updates are to be disseminated from multiple nearby vessels at 1Hz update rates. Data compression and an appropriate message structure may have to be designed to facilitate an acceptable rate of data transfer.

6. Development and Implementation Plan

6.1 System-level crowdsourcing (error characterisation)

The findings of this report have underscored the need for error characterisation within the integrity monitoring framework, as they reveal that measurement errors do not follow a Gaussian distribution, requiring a paradigm shift towards an adaptive error characterisation framework. This will have an impact on a broader PNT landscape and applications, especially mission-critical ones. While the error characterisation has been applied at the system level in this report, future work should also evaluate user-level error characterisation.

This section presents the future development and implementation plan to address these findings and guide future work towards enhancing the accuracy, integrity, availability, and continuity of mission-critical navigation including the maritime sector.

6.1.1 System-level Adaptive Error Characterisation Framework

Objective

To develop a robust adaptive error characterisation framework capable of selecting the error distribution based on real-time quality indicators, this should account for the distribution overbounding.

Strategy

Future research to implement Machine Learning (ML) for the adoptive error distribution selection based on quality indicators in real-time operation. This requires a representative dataset for training purposes, and evaluating various ML methods including deep learning ones. In addition, this required selecting the quality indicators that are related to the error distributions. This should be done by evaluating the performance of overbounded Gaussian distributions against overbounded alternatives.

6.1.2 Mechanism to utilise GEV distribution in the positioning algorithm

Objective

To devise a mechanism that simplifies the utilisation of the GEV distribution, which has shown the best results in terms of mapping extreme events. This can improve the system performance including accuracy, integrity, availability, and continuity.

Strategy

Future research to develop fast propagation models for GEV distribution - ML can be employed in this process.

6.1.3 Logistic Distribution Implementation

Objective

To apply the Logistic distribution for error characterisation, incorporating overbounding to maintain safety, if an adaptive error characterisation framework (see Section 6.1.1) is unavailable. This approach can be used based on the findings of this report (see Section 3.6).

Implementation strategy

Integrate the Logistic distribution into the existing system, applying overbounding techniques to guarantee the system's safety. This involves transmitting the Logistic scale parameter through integrity messages, as an alternative to the Gaussian distribution scale parameter. At the user level, modifications are required in the stochastic model's variances to base them on the Logistic distribution. Consequently, the position solution will be presented using the Logistic distribution, this must be taken into account at protection level models.

6.1.4 Carrier Phase Error Characterisation

Objective

To investigate the error characterisation for carrier phase measurements error, can support high accuracy carrier-phase integrity monitoring in the future.

Strategy

Conduct an error characterisation framework for carrier phase measurement errors based on the assessments used in this report, taking into account the aspects discussed in Section 3. This should evaluate the Gaussian assumption against alternatives and quantify the impact on system performance.

6.1.5 Continuous Improvement

Objective

In case the ML-based adaptive error characterisation framework is developed (see Section 6.1.1), the model should be improved with the increasing data volume over time.

Strategy

Implement regular updates to the framework, incorporating expanded datasets, and establishing a feedback loop with end-users and stakeholders to continuously gather insights and suggestions for improvement.

6.1.6 Phased Rollout

Objective

To ensure a smooth transition and integration of the new adaptive error characterisation framework into the existing system.

Strategy

Conduct pilot testing of the new framework in controlled environments to ensure stability and performance.

6.1.7 User-level future works

The future work outlined in the aforementioned development and implementation strategies concentrates on system-level error characterisation using crowd-sourced datasets. Although this report evaluates error distribution at the system level, the findings indicate potential directions for future work at the user level. Given that error distribution at the user level is more complex and distorted than at the system level, and considering the findings in this report, it can be seen that user-level error distributions do not follow Gaussian distributions. Based on this, the following are key areas for future work within the user-level framework:

- **Developing a User-Level Adaptive Error Characterisation Framework at the Measurement's Domain**

This framework aims to select the **measurement error distribution** based on quality indicators and advanced selection models (e.g. ML/deep learning) that reflect the measurement quality and operational environments (including complex environments). The latter might be not significantly considered in the system-level adaptive error characterisation framework, making the user-level framework more complex. In addition, the system-level framework does not take into account user-level errors, which is crucial for ensuring safe operations.

- **Developing a User-Level Adaptive Error Characterisation Framework at the Position Domain**

This framework should develop a model to predict position error directly using quality indicators and deep learning, this can potentially enhance the system performance and speed compared to the measurement domain framework. The development of this framework is more complex than the aforementioned ones as the inputs for the deep learning model will include multiple quality indicators (each measurement having its quality indicators).

- **Developing an ML-Based Protection Level Prediction Model**

Future work should also focus on developing and testing ML-based protection level prediction models, utilising quality indicators. The aim of this approach is to address the multidimensional, nonlinear, and complex protection level computation. This should be done by ensuring the bounding of the estimated ML models and considering the integrity risk requirement.

6.2 User-level crowd-sources positioning

The findings of this report (refer to Section 4.5.2) indicate that Lidar-based crowdsourcing is a promising method to enhance maritime navigation and overcome the current limitations of standalone GNSS in achieving the required performance level during the port phase. However, most long-range Lidars are designed for airborne applications, as there are very few applications that require precise long-range measurements with 360-degree coverage. Therefore, the developed approach opens a new dimension for the industry to manufacture 360-degree long-range Lidars. Currently, airborne Lidars can be utilised by combining a number of Lidars to achieve 360-degree coverage. The following points summarise some of the key future work and implementation plans

6.2.1 Specify the Lidar specification to support the port phase

Objective

To define the Lidar specification including accuracy to support the port phase.

Strategy

This should be done via a sensitivity analysis using a suitable platform, like the Imperial simulator, to define precisely the required level of accuracy.

6.2.2 Define the Lidar failure mode and models

Objective

To delineate the failure modes and models of Lidar systems, as these aspects remain underexplored deeply to date.

Strategy

Conduct research to identify and understand Lidar failure modes and models, assess their impact on measurement quality, and evaluate the probability of such failures.

6.2.3 Improve the Crowdsourcing Mathematical Model and Simulator Performance

Objective

To enhance the crowdsourcing mathematical model and simulator performance, thereby improving system accuracy and availability.

Strategy

- Incorporate defined LiDAR failure modes and models: Update the simulator to include the specific failure modes and models associated with LiDAR technology. This will allow for more accurate simulation of real-world scenarios.
- Refine protection level computation assumptions: This involves refining the k-factor and nominal bias assumptions to increase the protection level reliability.
- Develop a Measurements Selection Model: Create a model to select measurements based on quality indicators, including the geometry of the vessel distribution and the distance to nearby vessels.

- Develop a Solution Separation Integrity Model: Design an integrity model that utilises solution separation to address multiple simultaneous failures.

6.2.4 Designing a combined set of Lidars based on the '*n*' number of airborne Lidars to measure range precisely.

Objective

Designing a combined set of Lidars based on the '*n*' number of airborne Lidars to measure range precisely.

Strategy

Conduct research to integrate a set of Lidars to ensure simultaneous 360-degree coverage and accurate range measurement collection.

6.2.5 Future works shall consider the vessels speed in the user-level positioning and integrity monitoring algorithm

6.2.6 The protection level information shall be transferred between vessels in the future to ensure the integrity and the level of confidence in the positioning information of nearby vessels. In addition, taking into account the overbounding of the range measurement is essential to ensure safe navigation.

6.2.7 Disseminate the findings to academic, governmental, and industrial stakeholders, including manufacturers.

To publish the report outcomes among academic, governmental, and industrial entities, including manufacturers, to showcase novel applications for long-range 360-degree Lidar, potentially catalysing new manufacturing directions

7. Exploitation Plan

7.1 Overview of the Crowdsourcing Concept

Crowdsourcing can be categorised into user-level and system-level types, which can be defined as follows:

- User-level crowdsourcing relies on leveraging nearby GNSS devices to support the user-level navigation system. This can involve using the positioning information from nearby smartphones or, as implemented in this report, information from nearby vessels.
- System-level crowdsourcing involves the use of any available PNT sources to support system-level integrity monitoring, such as employing a CORS network.

7.2 Applications domain

Crowdsourcing offers the potential to support a wide range of mission-critical applications. The system-level crowdsourcing can support all mission-critical applications including aviation, maritime, autonomous vehicles, robotics, precision agriculture, and autonomous drone operations. This is because error characterisation and overbounding are essential elements of any positioning algorithm, particularly those critical to mission ones. Table (6.1) summarises the applications domain mentioned in the Government Office of Science study with its integrity requirements.

For applications meeting the requirements (e.g., aviation), system-level crowdsourcing will reduce the alarm limit requirements in future, which can enhance operations. For those not yet meeting requirements, such as autonomous vehicles, robotics, and precision agriculture, system-level crowdsourcing will be a key element in achieving necessary performance levels in the future.

User-level crowdsourcing positioning has the potential to support a wide range of PNT applications in the future such as autonomous vehicles, robotics, and Visually Impaired Persons (VIP), especially as smart cities evolve to integrate communication among vehicles, pedestrians, and infrastructure. The advantage of this approach is rooted in the limitation that standalone GNSS systems cannot achieve the requirements for high-accuracy applications, especially within challenging urban environments.

7.3 Opportunities for equipment, application, and service providers:

System-level crowdsourcing offers opportunities for SBAS providers, such as ESA, to utilise CORS networks to improve error characterisation and deliver more reliable and safer distributions to users. In the UK, the focus on system-level crowdsourcing underscores the need for a robust CORS network with enhanced data update rates and precision. This provides Ordnance Survey and its commercial partners (Hexagon, Trimble, Topcon, AXIO-NET, SoilEssentials, Premium Positioning) the opportunity to improve the OS Net CORS network's performance to achieve the required level of performance for system-level crowdsourcing, including offering enhanced precision and update rates to users.

In addition, system-level crowdsourcing offers opportunities for equipment providers to enhance positioning performance by utilising Gaussian alternatives as novel features. If Gaussian

alternatives are provided to users, algorithms must be adapted to handle these alternatives, which can improve positioning performance. Of the equipments providers can benefits from this feature include: Ublox, Hexagon, Trimble, Stonex, Teseo, NovAtel, Furuno, Bosch Sensortec, STMicroelectronics, Quectel, Telit, SkyTraq, Swift Navigation, Septentrio, and JAVAD GNSS.

Sector	Application	Required Accuracy	Required Integrity risk
Telecoms	LTE-A & 5G	500ns	$< 10^{-9}$
	LTE-TDD	1.5µs	$< 10^{-9}$
	Billing & alarms	1ms - 500ms	$< 10^{-9}$
Finance	High-frequency trading	0.1ms	$< 10^{-9}$
	Electronic trading	1ms	$< 10^{-9}$
	ATM transactions	1ms - 1s	$< 10^{-9}$
	Voice trading	1s	$< 10^{-9}$
Energy	Infrastructure monitoring	0.1m	10^{-6}
	Oil rig dynamic positioning	0.1m	10^{-7}
	Oil & gas exploration	5m	$< 10^{-5}$
	Oil & gas supply	50m	10^{-5}
	Protection systems Wide area monitoring	1µs	$< 10^{-7}$
	Infrastructure monitoring & control	>1µs	$< 10^{-7}$
Emergency Services	Lane-level vehicle navigation	0.5m	10^{-7}
	Identifying the position of incidents	2m -5m	10^{-7}
	Dynamic routing	5m	10^{-5}
Food and Farming	Automatic farm machines	0.5m	$10^{-5} - 10^{-4}$
	Pest control & prevention	0.5m - 1.0m	$10^{-5} - 10^{-4}$
	Yield mapping & plot mapping	1.0m	10^{-4}
	Tracking of goods & vehicles	50m	10^{-4}
Aviation	CaT II / III PA/surface movement	1.0m	10^{-9}
	CAT I precision approach (PA)	4.0m	$10^{-9} - 10^{-7}$
	Approach operations with vertical guidance	16.0m	$10^{-9} - 10^{-7}$
	Non-precision approach	220m	10^{-7}
	En route, terminal	740m	10^{-7}
	En route, navigation	3700m	10^{-7}
Road	Lane-level vehicle navigation	0.5m	10^{-7}
	Dynamic routing	5m	10^{-4}
	Fleet management	>50m	10^{-3}
Rail	Infrastructure Customer applications monitoring	>1m	10^{-6} - 10^{-5}
	High-density command & control	1m - 5m	10^{-9}
	Low-density command & control	5m- 50m	10^{-7} -10^{-6}
	Customer applications	5m - 100m	10^{-6} - 10^{-5}
	Asset management	50m - 100m	10^{-7} -10^{-6}

Maritime	Automatic docking	0.1m	10^{-7}
	Port operations	0.1m - 1.0m	10^{-6}
	Cargo handling	0.1m	$10^{-6} - 10^{-4}$
	Port approach	5.0m	10^{-6} - 10^{-4}
	Icebreakers	5.0m -10m	-
	Coastal navigation	10m	10^{-6} - 10^{-4}
	Search & rescue	10m - 50m	10^{-7} - 10^{-6}
	Ocean navigation	50m	10^{-4}
Surveying and Mass-market applications	Surveying	0.1m	10^{-6}
	Sports	5.0m - 50m	10^{-4}
	Social networking	50m	10^{-4}
Autonomous road vehicles and drones	Autonomous road vehicles	0.1m - 1.0m	10^{-7}
	Drones	5.0m - 10m	10^{-7} - 10^{-6}
New space	Spacecraft docking	0.1m - 0.5m	10^{-7}
	Satellite orbit determination	0.5m	10^{-6}
	Space navigation	50m	10^{-5}

Table 6.1: GNSS applications accuracy and integrity requirements, approximated from study graphs (Whitty and Walport,2018).

In user-level crowdsourcing, the approach developed in this report recommends utilising Lidar to measure distances in order to achieve the required performance level during the porting stage. As previously mentioned, most long-range Lidars are designed for airborne applications, with very few requiring precise long-range measurements with 360-degree coverage. Consequently, the approach introduced here paves the way for the industry to innovate and produce 360-degree long-range Lidars. This can offer opportunities for Lidar manufacturers includes: Velodyne Lidar, Luminar Technologies, Leica Geosystems, SICK AG, Teledyne Optech, RIEGL Laser Measurement Systems, Quanergy Systems, Trimble, FARO Technologies, and Hesai Technology

User-level crowdsourcing has significant potential to extend into various industries, with autonomous driving being a particularly promising area. The following are key research and development providers in the autonomous driving sector:

- Waymo
- Cruise Automation
- Tesla
- Aptiv
- Uber Advanced Technologies Group
- NVIDIA Corporation
- Mobileye
- Baidu Apollo
- Argo AI
- Zoox
- Aurora - Founded by former leads from Google, Tesla, and Uber,
- Nuro

8. References

ALARIFI, A., AL-SALMAN, A., ALSALEH, M., ALNAFESSAH, A., AL-HADHRAMI, S., AL-AMMAR, M. A. & AL-KHALIFA, H. S. 2016. Ultra-wideband indoor positioning technologies: Analysis and recent advances. Sensors, 16, 707.

ANGRISANO, A. & GAGLIONE, S. 2022. Smartphone GNSS performance in an urban scenario with RAIM application. Sensors, 22, 786.

Bibby, C. and Reid, I., 2010, May. A hybrid SLAM representation for dynamic marine environments. In 2010 IEEE International Conference on Robotics and Automation (pp. 257-264). IEEE.

Blanch, J., Walter, T., Enge, P., Lee, Y., Pervan, B., Rippl, M. and Spletter, A., 2012, September. Advanced RAIM user algorithm description: Integrity support message processing, fault detection, exclusion, and protection level calculation. In Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012) (pp. 2828-2849).

BOUET, M. & DOS SANTOS, A. L. RFID tags: Positioning principles and localization techniques. 2008 1st IFIP Wireless Days, 2008. IEEE, 1-5.

CHON, H. D., JUN, S., JUNG, H. & AN, S. W. 2004. Using RFID for accurate positioning. Positioning, 1.

Davison, 2003, October. Real-time simultaneous localisation and mapping with a single camera. In Proceedings Ninth IEEE International Conference on Computer Vision (pp. 1403-1410). IEEE.

EYNG, A. C., RAYEL, O. K., OROSKI, E. & REBELATTO, J. L. Kalman filtering-aided hybrid indoor positioning system with fingerprinting and multilateration. 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), 2020. IEEE, 1-5.

Freeman, G., 2023. Government Policy Framework for Greater Position, Navigation and Timing (PNT) Resilience. [online] 18 October. Available at: https://questions-statements.parliament.uk/written-statements/detail/2023-10-18/hcws1073 [Accessed Date].

Gallagher, L., Kumar, V.R., Yogamani, S. and McDonald, J.B., 2021, August. A hybrid sparse-dense monocular slam system for autonomous driving. In 2021 European Conference on Mobile Robots (ECMR) (pp. 1-8). IEEE.

Hess, W., Kohler, D., Rapp, H. and Andor, D., 2016, May. Real-time loop closure in 2D LIDAR SLAM. In 2016 IEEE international conference on robotics and automation (ICRA) (pp. 1271-1278). IEEE.

Hong, Z., Petillot, Y. and Wang, S., 2020, October. Radarslam: Radar based large-scale slam in all weathers. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 5164-5170). IEEE.

IMO, 2015, Resolution, A., 2015. 1106 (29)“Revised guidelines for the operational use of shipborne automatic identification systems (AIS)”. IMO, London.

INSPIRe-GMV-D3.1-v1.1, 2023. D3.1 – Technical Report of Developments and Test of DFMC M(G)RAIM. [pdf] Prepared by GMV INSPIRe Team. Version 1.1, October 2023. Approved by M Pattinson, Authorized by T Richardson. Submitted to ESA

INSPIRe-GMVNSL-D4.1-v1.1, 2024. D4.1 – Technical Report of Developments and Test of DFMC GNSS and DR VAIM. [pdf] Prepared by GMV INSPIRe Team. Version 1.1, January 2024. Approved by M Pattinson, Authorized by T Richardson. Submitted to ESA.

Johnson, G.W., Swaszek, P.F., Hartnett, R.J., Shalaev, R. and Wiggins, M., 2007, May. An Evaluation of eLoran as a Backup to GPS. In 2007 IEEE Conference on Technologies for Homeland Security (pp. 95-100). IEEE.

KLUGE, T., GROBA, C. & SPRINGER, T. Trilateration, fingerprinting, and centroid: taking indoor positioning with bluetooth LE to the wild. 2020 IEEE 21st International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM), 2020. IEEE, 264-272.

Kohlbrecher, S., Von Stryk, O., Meyer, J. and Klingauf, U., 2011, November. A flexible and scalable SLAM system with full 3D motion estimation. In 2011 IEEE international symposium on safety, security, and rescue robotics (pp. 155-160). IEEE.

Kolmogorov, 1933. Sulla determinazione empirica di una legge didistribuzione. Giorn Dell'inst Ital Degli Att, 4, pp.89-91.

Kotz, S., Kozubowski, T. and Podgórski, K., 2001. The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance (No. 183). Springer Science & Business Media.

KRISHNAN, S., SHARMA, P., GUOPING, Z. & WOON, O. H. A UWB based localization system for indoor robot navigation. 2007 IEEE International Conference on Ultra-Wideband, 2007. IEEE, 77-82.

LI, D., YAN, Y., ZHANG, B., LI, C. & XU, P. 2016. Measurement-based AP deployment mechanism for fingerprint-based indoor location systems. KSII Transactions on Internet and Information Systems (TIIS), 10, 1611-1629.

Niemann, P. 2023. WP7 - GNSS DFMC Integrity Monitoring D-CGI-001-001. Taylor Airey Limited. [Unpublished].

Schlegel, D., Colosi, M. and Grisetti, G., 2018, May. Proslam: Graph SLAM from a programmer's Perspective. In 2018 IEEE international conference on robotics and automation (ICRA) (pp. 3833-3840). IEEE.

Schuster, F., Keller, C.G., Rapp, M., Haueis, M. and Curio, C., 2016, November. Landmark based radar SLAM using graph optimization. In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) (pp. 2559-2564). IEEE.

Son, P.W., Park, S.G., Han, Y. and Seo, K., 2020. eLoran: Resilient positioning, navigation, and timing infrastructure in maritime areas. IEEE Access, 8, pp.193708-193716.

TALVITIE, J., RENFORS, M. & LOHAN, E. S. 2015. Novel indoor positioning mechanism via spectral compression. IEEE Communications Letters, 20, 352-355.

The Maritime Executive. (2016). Europe Gives Up on eLoran. [online] Available at: https://maritime-executive.com/article/europe-gives-up-on-eloran [Accessed 2 Nov. 2023].

Whitty, C. and Walport, M., 2018. Satellite-derived Time and Position: A Study of Critical Dependencies. Government Office for Science: London, UK.

9. Appendix

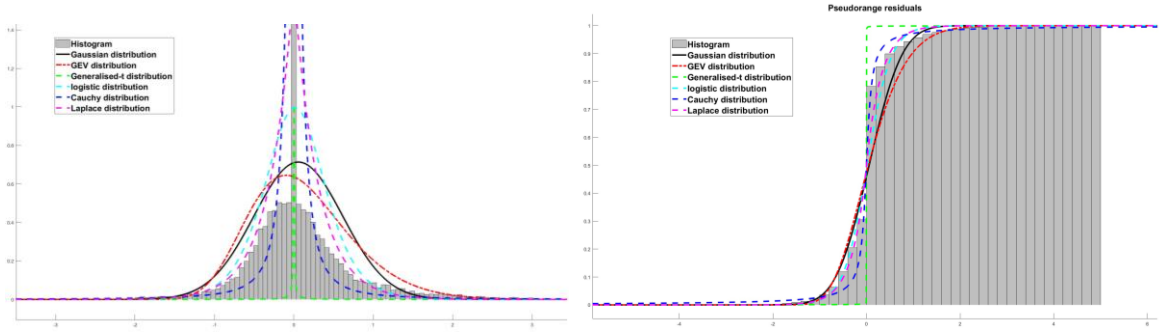


Figure 9.1: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 5

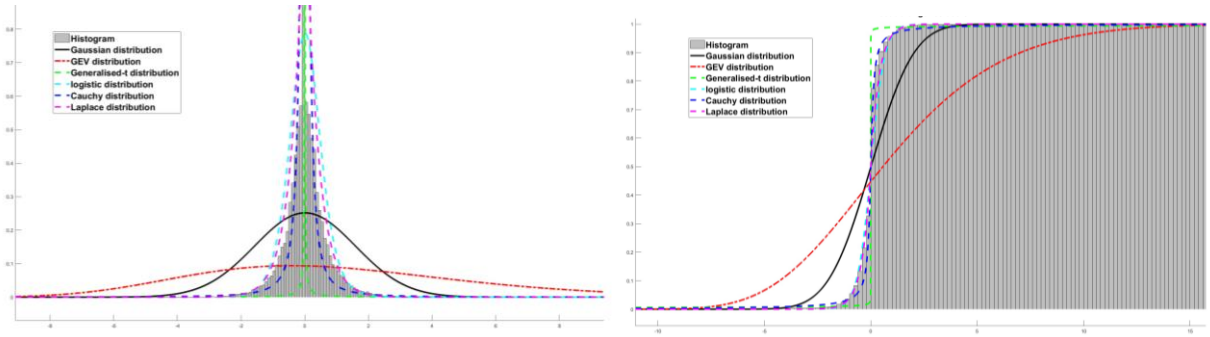


Figure 9.2: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 6

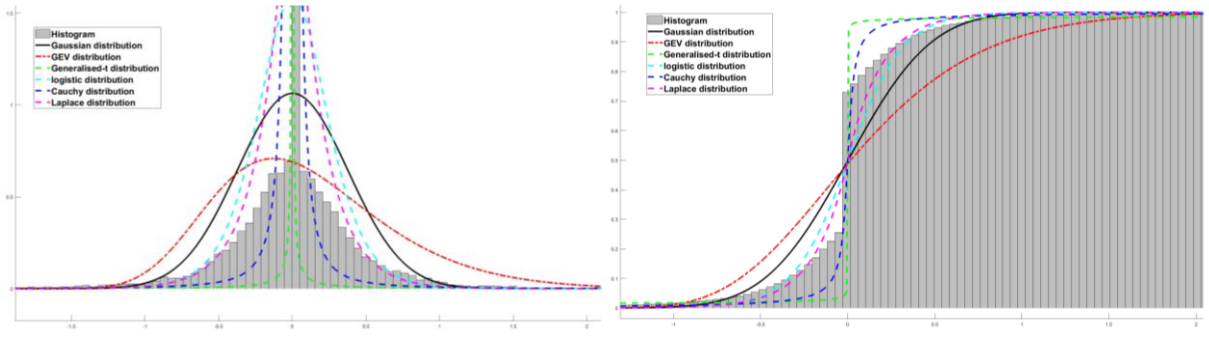


Figure 9.3: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 7

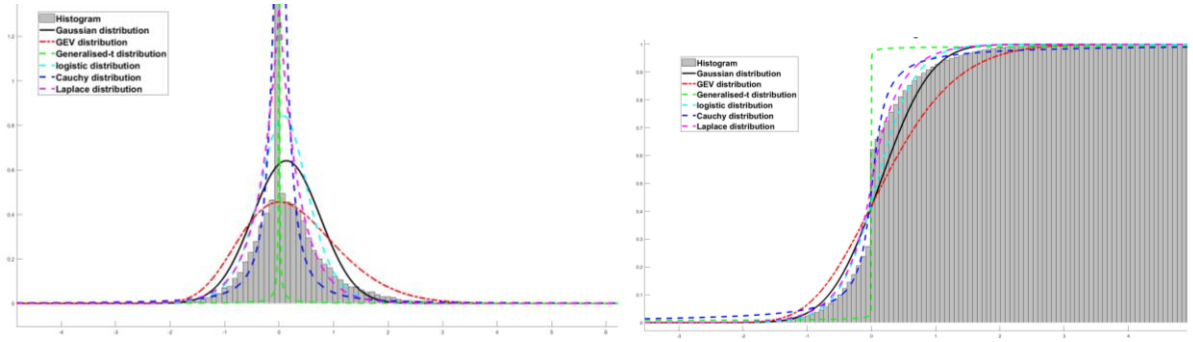


Figure 9.4: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 8

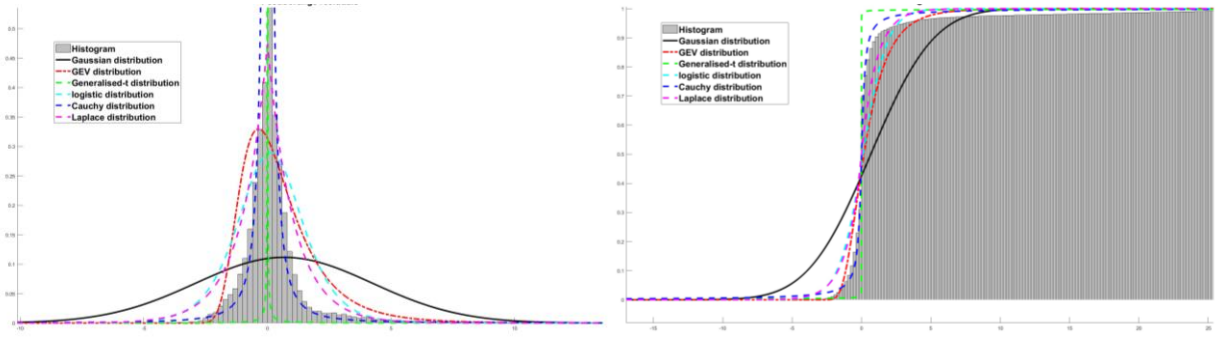


Figure 9.5: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 9

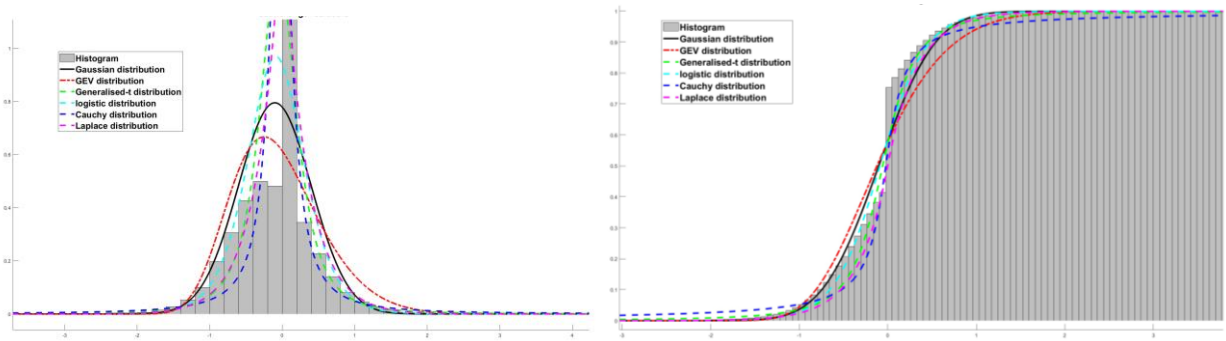


Figure 9.6: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 10

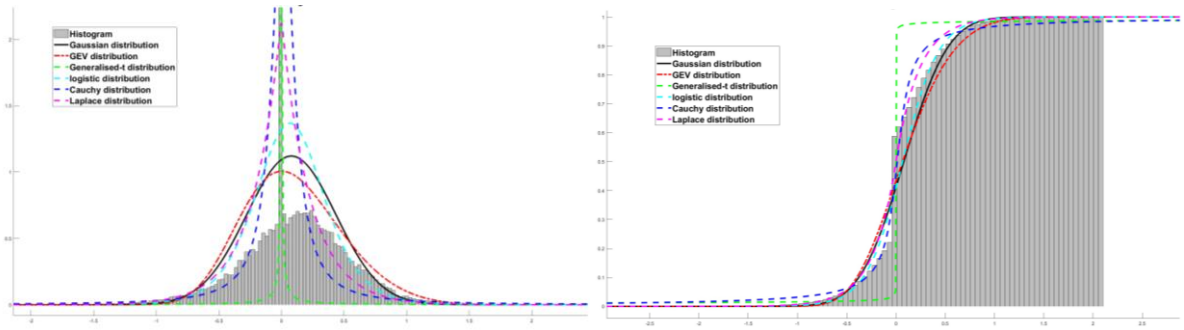


Figure 9.7: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 11

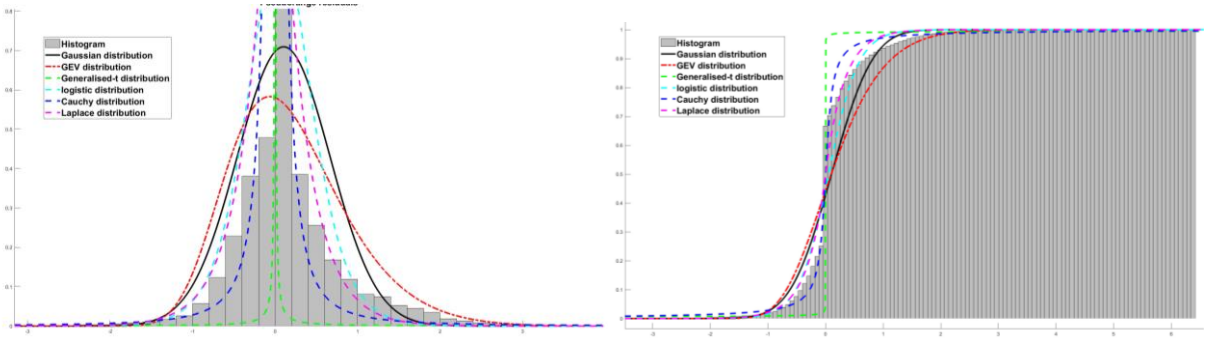


Figure 9.8: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 12

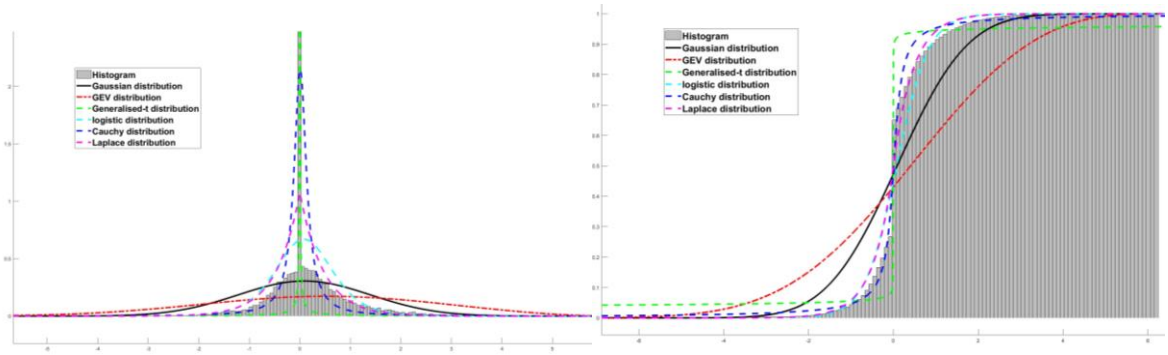


Figure 9.9: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 13

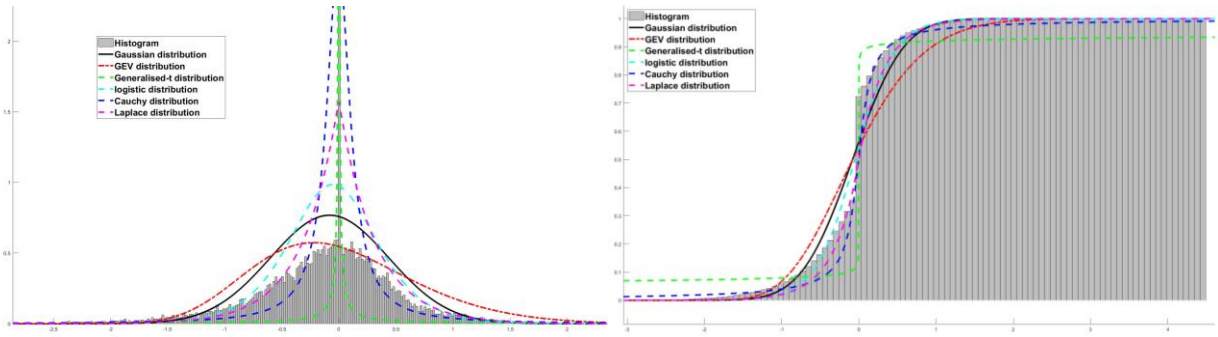


Figure 9.10: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 14

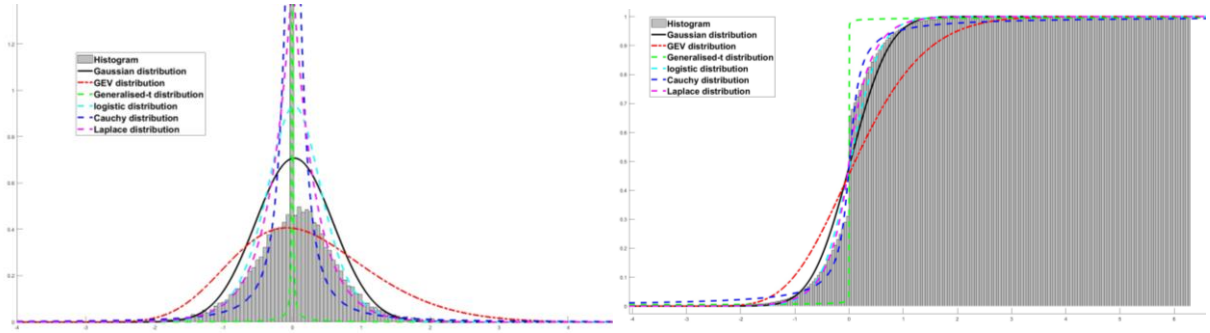


Figure 9.11: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 15

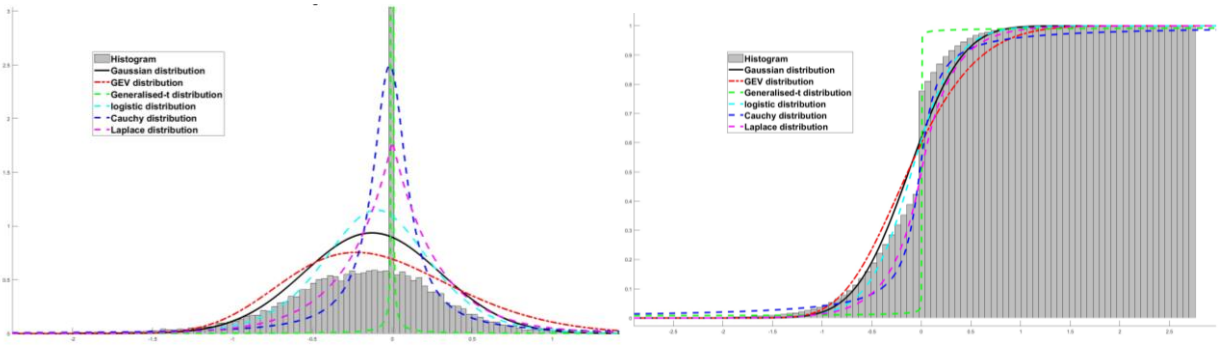


Figure 9.12: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 16

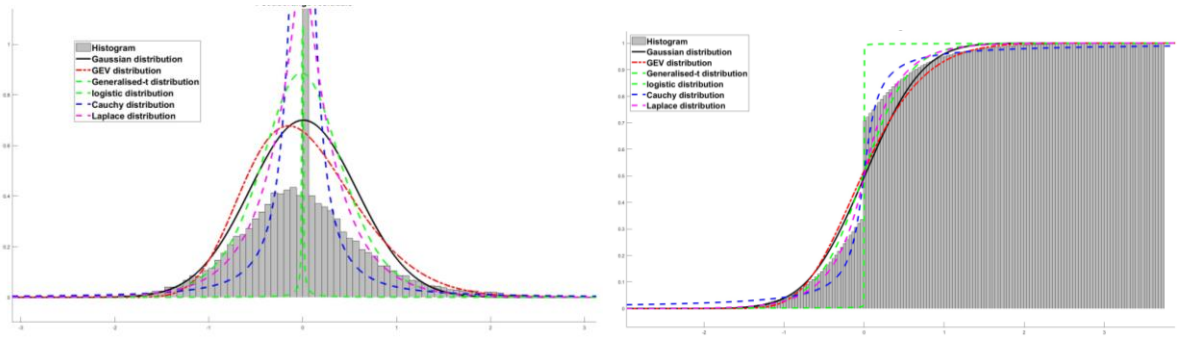


Figure 9.13: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 17

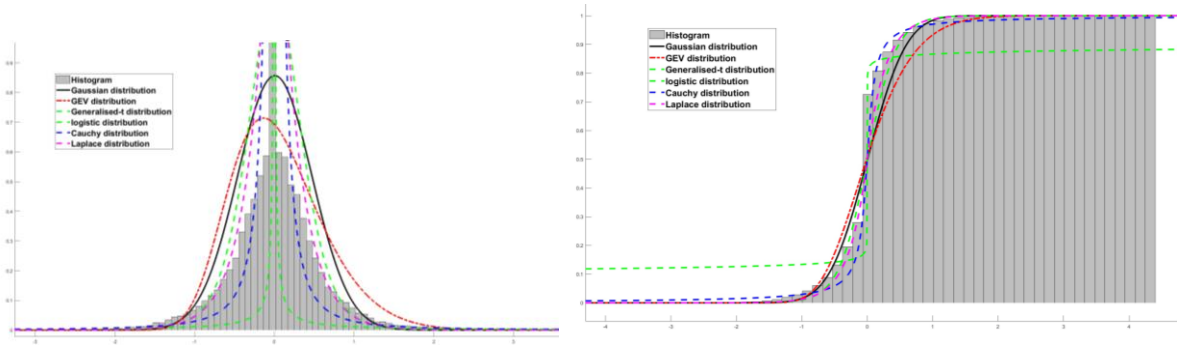


Figure 9.14: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 18

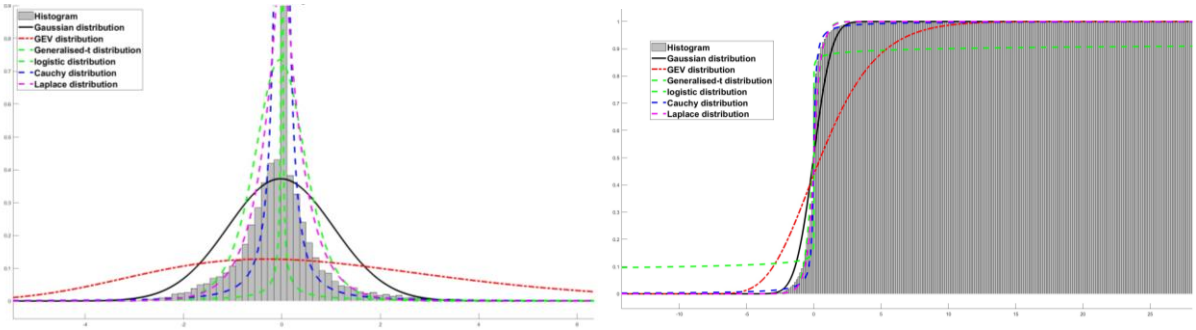


Figure 9.15: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 19

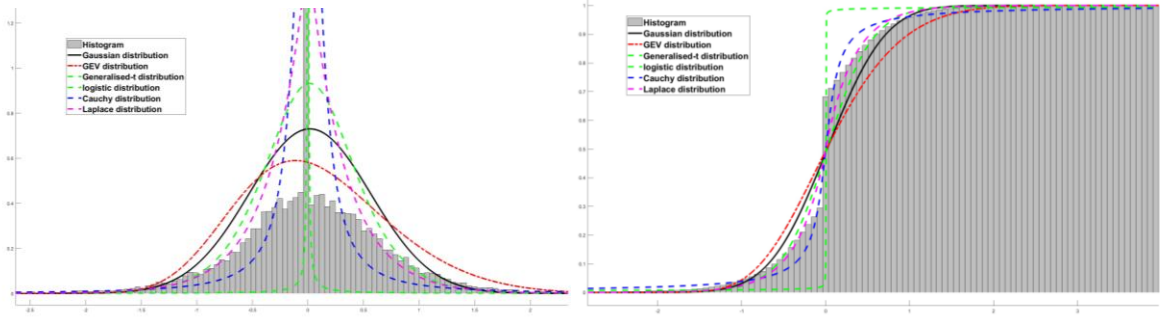


Figure 9.16: pseudorange error characterisation using the six distributions in PDF and CDF domain, dataset 20