Multi-Sensor Fusion for Seamless Navigation in Railway Domain

Gorka De Miguel Aramburu

MULTI-SENSOR FUSION FOR SEAMLESS NAVIGATION IN RAILWAY DOMAIN

DISSERTATION

submitted for the degree of Doctor of Philosophy of the University of Navarra by

Gorka De Miguel Aramburu

under the supervision of

Dr. Iñigo Adín Marcos and

Dr. Jon Goya Odriozola

Donostia-San Sebastián, July 2020

tecnun
Universidad de Navarra

UNIVERSIDAD DE NAVARRA
ESCUELA SUPERIOR DE INGENIEROS
SAN SEBASTIÁN

Multi-Sensor Fusion for Seamless Navigation in Railway Domain

DISSERTATION

submitted for the Degree of Doctor of Philosophy by

GORKA DE MIGUEL
ARAMBURU

under the supervision of

PhD. Iñigo Adín Marcos and

PhD. Jon Goya Odriozola

Donostia–San Sebastián, July 2020

To my family

Acknowledgment

With these lines, I would like to express my sincere gratitude to all institutions and individuals who have made this research work possible.

First of all, I would like to express my sincere gratitude to CEIT (Asociación Centro Tecnológico Ceit) and TECNUN (School of Engineering of the University of Navarra) for the opportunity they have given me to perform this research work.

Likewise, I would like to express my gratitude to the European Commission's Framework Programme H2020 and its initiative Shift2Rail by the funding received. Particularly, my deep gratitude to the EU projects ERSAT-EAV, FR8RAIL and X2Rail-2 for the chance they have given me to be part of their teams. From this point in advance, the acknowledgements are written in my mother tongue in order to express them more sincerely.

Lehenik eta behin, eskerrik beroenak nire ikerketa zuzendari izan diren Iñigo Adin eta Jon Goyarentzat eskainitako laguntza, denbora, pazientzia eta une zailetan emandako adoreagatik. Denborarik ez zegoenean iribarre batez niri denbora egiteagatik, mila esker.

Laborategiko lankideei, zuen laguntzarik gabe ezinezkoa izango zen tesi honi amaiera ematea. Gaur egun bertan lanean jarraitzen dezuenok azken hilabeteetan behar nituen animoak eman dizkidazue, baita teknikoki laguntza behar izan dudanean. Orain laborategian lan egiten ez zaudenei ere mila esker tesia egitera animatzeagatik eta CEIT-en lanean hasi nintzenean emandako harreragatik. Bereziki eskerrak Pauli nire ondoan debugeatzen emandako orduengatik.

Jaizki Mendizabali eskerrak nigan konfidantza izateagatik graduko eta masterreko proiektuak egin ondoren CEIT-en lan egiteko aukera emateagatik. Egunez egun ikasi eta lan-munduan hazten jarraitzeko tresnak emateagatik.

Mila esker Iñigo Gutiérrez eta Noemí Pérez TECNUN-eko irakasleei, beraiekin elkarlanean ikasgai baten parte izateko aukera luzatzeagatik. Irakaskuntzaren mundu ederrean lehen pausuak zuekin ematea plazerra izan da.

Kuadrilako lagunei, eskerrik asko momentu txar zein onetan bertan egoteagatik, ni entretenitzeagatik eta burua husten laguntzeagatik. Mundialak zarete!

Unibertsitateko lagunei mila esker, tesi honen zati txiki bat zuena da. Zuek gabe ez nintzen Ingeniari izango. Mila esker tesiaren inguruan arduratzeagatik.

Azkenik, mila esker familiarteko guztioi. Aita, Ama, Ander, mila esker beti bertan egoteagatik eta ni jasateagatik momentu on eta txarretan. Hutsaren truke beti bertan egoteagatik, behar nuenean laguntzeko eta momenturik txarrenetan ni berriro indarrez betetzeko. Ama, Aita, txikitatik erakutsi didazue lanak bere fruituak ematen dituela eta tesi hau zuena da. Ander, anai zaharrenek txikienei erakusten dietela esan ohi da, nire kasuan alderantziz izan da. Zure izateko moduak, galdetzeko moduak, laguntza emateko moduak, tesi hau aurrera eramateko indarra eta irakaspena eman dit. Bukatzeko, mila esker Natalia lehen unetik tesia egitera animatzeagatik eta nigan konfidantza izateagatik. Egunez egun emandako laguntza, indarra eta animo horiek gabe ezin izango nuke tesi hau bukatu. Bihotz bihotzez mila esker lauoi, zuek gabe nire bizitza zailagoa izango zen.

Tesi hau zuek goztiona da. Eskerrik asko denoi!.

Summary

The European Union (EU) aims at making railway a more attractive transportation method by improving its efficiency and reducing its costs. These achievements could be covered with the migration from ETCS level 2 to ETCS level 3. Many projects related to new positioning systems have been funded by The European Union. Most of these positioning systems are based on GNSS, due to the key role that GNSS will play in the migration to ETCS level 3.

One of the problems on using only GNSS systems is the lack of availability of them. During railway operation, there are areas with potential GNSS outages, such as urban canyons, woods or other possible signal blockers and disturbances. Moreover, it is a fact that GNSS signals are not reachable, nor reliable in tunnels or indoor environments. For GNSS to be able to have a key role in the next years in railway security, the afore mentioned lack of availability has to be solved.

To cope with this issue, a multi-sensor approach with software enhancements is proposed in this dissertation. The objective of this research work deals with fusing different sensors and creating new software strategies to achieve a higher availability with the best possible accuracy. The seamless position will benefit in all the operation modes, from the train station to a harsh environment for satellites, during the train operation.

The scope of the dissertation is to create a multi-sensor positioning system including GNSS, Inertial Measurement Unit (IMU), and Ultra Wide Band (UWB) with other software techniques to obtain a position estimation with a 100% availability for railway systems.

This work shows the different steps from the study of the state of the art, going through the implementation, and ending with the performance analysis of the algorithm developed. This research work has been conducted under different European projects such as ERSAT-EAV, FR8RAIL or X2Rail-2, in which CEIT has participated.

Contents

1 INTRODUCTION AND CONTEXT OF THE RESEARCH WORK 1

1.1 Introduction 2

1.1 MOTIVATIONS 4

1.2 OUTLINE OF THE RESEARCH WORK 5

2 STATE OF THE ART 9

2.1 GNSS Positioning Systems 10

2.1.1 GPS 10

2.1.2 GLONASS 13

2.1.3 GALILEO 14

2.1.4 BeiDou Chinese COMPASS System 16

2.1.5 GNSS Observations 20

2.1.6 Computation methods for GNSS Positioning 30

2.2 IMU 48

2.2.1 Clipping 49

2.3 Ultra Wide Band 54

2.4 Map Based Methods 56

2.4.1 Map-Aiding 56

2.4.2 Map-Matching 58

2.5 IMU and GNSS data fusion 59

2.5.1 Loosely coupled integration 59
2.5.2 Tightly coupled integration 59
2.5.3 Deeply coupled integration 60
2.5.4 IMU and GNSS Fusion Algorithm 61
2.5.5 IMU/GNSS integration 62

2.6 GNSS and UWB data fusion 91

2.7 RAIM 93
2.7.1 Fault Detection and Exclusion 93
2.7.2 Integrity Engine 94

2.8 Commertial positioning systems available for different level of users 98

3 OBJECTIVES 101

3.1 Introduction 102

3.2 Objectives 103

4 DESIGN AND IMPLEMENTATION OF THE ALGORITHM 105

4.1 Preliminary decisions 106

4.2 Positioning algorithm 108
4.2.1 Check Data Availability 109
4.2.2 Read Data 118
4.2.3 GPS Data Treatment 120
4.2.4 IMU Data Treatment 138
4.2.5 UWB Data Treatment 140

4.2.6 Kalman Filter 141
4.2.7 Known Blocked Scenarios (KBS) 144
4.2.8 Data Output 149

5 MEASUREMENT CAMPAIGNS FOR THE SYSTEM VALIDATION AND PERFORMANCE EVALUATION 153

5.1 Scenario 1: EuskoTren Line 154
5.1.1 Measuring Equipment 155
5.1.2 Ground Truth Generation Equipment 158

5.2 Scenario 2: Nottingham Geospatial Institute Roof Trial Site 159
5.2.1 Measuring Equipment 160
5.2.2 Ground Truth generation equipment 165

5.3 Scenario Comparison 166

5.4 Testing methodology 167
5.4.1 Tests performed 167
5.4.2 Test template 167

5.5 Testing plan for reference journeys 168
5.5.1 Test list 168
5.5.2 Test definitions 169

6 INPUT DATA ANALYSIS 171

6.1 Introduction 172

6.2 GPS Data Analysis 173

6.2.1 Available satellites 173
6.2.2 Elevation of the satellites 174
6.2.3 SNR of the satellites 177
6.2.4 SNR and Elevation joint implication 181
6.2.5 Standard deviation of the position 182
6.2.6 HDOP 185
6.2.7 Number of observables 186

6.3 IMU Data Analysis 187
6.3.1 Accelerometer data analysis 187
6.3.2 Gyroscope data analysis 189
6.3.3 Yaw data analysis 192
6.3.4 Bias influence in accelerometer data 193
6.3.5 Number of observables 197

6.4 UWB Data Analysis 198
6.4.1 Residuals 198
6.4.2 Number of Observables 199
6.4.3 Standard deviation of the position 200
6.4.4 HDOP 205

7 FIELD-TEST RESULTS AND PERFORMANCE EVALUATION 207

7.1 Performance of the algorithm 208
7.1.1 TS1_EUS_1100 208
7.1.2 TS1_EUS_1101 210
7.1.3 TS2_NOT_1000 212
7.1.4 TS2_NOT_0100 215
7.1.5 TS2_NOT_0010 217

7.1.6 TS2_NOT_1100 220
7.1.7 TS2_NOT_1101 223
7.1.8 TS2_NOT_1010 225
7.1.9 TS2_NOT_1011 228
7.1.10 TS2_NOT_0110 230
7.1.11 TS2_NOT_0111 233
7.1.12 TS2_NOT_1110 233
7.1.13 TS2_NOT_1111 236

7.2 Performance Discussion 239

8 CONCLUSIONS AND FUTURE GUIDELINES 243

8.1 Conclusions 244

8.2 Future Guidelines 249

A PUBLICATIONS 261

A.1 International Journal Papers 263

A.2 International Conference Papers 313

List of Figures

1 INTRODUCTION AND CONTEXT OF THE RESEARCH WORK 1

2 STATE OF THE ART 9

Figure 2.1 ECEF Coordinate frame 13
Figure 2.2 Satellite to user ranges 31
Figure 2.3 Example SnapShot Positioning fo Real GPS Rail Data (EATS n.d.) 46
Figure 2.4 Example Kalman Filter Positioning for Real GPS Rail Data (EATS n.d.) 46
Figure 2.5 Loosely coupled integration scheme (Noureldin, Karamat, and Georgy 2013) 59
Figure 2.6 Tightly coupled integration scheme (Noureldin, Karamat, and Georgy 2013) 60
Figure 2.7 Deeply coupled integration (Noureldin, Karamat, and Georgy 2013) 61
Figure 2.8 Open-loop architecture (Noureldin, Karamat, and Georgy 2013) 62
Figure 2.9 Closed-loop architecture (Noureldin, Karamat, and Georgy 2013) 63
Figure 2.10 Tightly coupled IMU/GNSS integration (Noureldin, Karamat, and Georgy 2013) 64
Figure 2.11 Overall implementation of a tightly coupled IMU/GNSS integration (Noureldin, Karamat, and Georgy 2013) 64
Figure 2.12 Mechanization of the IMU in the inertial frame (Noureldin, Karamat, and Georgy 2013) 82
Figure 2.13 Mechanization in the e-frame (Noureldin, Karamat, and Georgy 2013) 84
Figure 2.14 Mechanization in ECEF coordinates (Noureldin, Karamat, and Georgy 2013) 90

3 OBJECTIVES 101

4 DESIGN AND IMPLEMENTATION OF THE ALGORITHM 105

Figure 4.1 Overview of the positioning algorithm 108

Figure 4.2 Overview of the data availability check 110
Figure 4.3 Overview of the read data functionality 119
Figure 4.4 GPS data treatment overview flowchart 120
Figure 4.5 Overview of Satellite Position computation flowchart 125
Figure 4.6 Least-squares flowchart 131
Figure 4.7 Flowchart of the Only Leave One Out method 132
Figure 4.8 Flowchart of the SNR and Elevation cutoff 133
Figure 4.9 Flowchart of the Klobuchar model algorithm 137
Figure 4.10 IMU data treatment flowchart 140
Figure 4.11 Kalman Filter Initialization flowchart 142
Figure 4.12 Kalman filter execution flowchart 144
Figure 4.13 Known Blocked Scenarios flowchart 147

5 MEASUREMENT CAMPAIGNS FOR THE SYSTEM VALIDATION AND PERFORMANCE EVALUATION 153

Figure 5.1 General environment in EuskoTren Lane 154
Figure 5.2 Section of Eusko Trenbide Sarea (ETS) network 155
Figure 5.3 EuskoTren Serie 900 Locomotive 155
Figure 5.4 Nottingham Geospatial Institute Roof Trial Site 159
Figure 5.5 Antenna placement in the mock train 159
Figure 5.6 Nottingham Geospatial Institute Trial Site Top View 160
Figure 5.7 Equipment placement in the mock train 163
Figure 5.8 UWB Antenna placement 164
Figure 5.9 UWB anchor placement example 164

6 INPUT DATA ANALYSIS 171

Figure 6.1 Number of Satellites received in each measurement for GPS 173
Figure 6.2 Mean elevation for the satellites with data available 175
Figure 6.3 Number of satellites discarded due to bad elevation 176
Figure 6.4 Number of satellites used for positioning after performing the elevation mask 176
Figure 6.5 Mean elevation for satellites used 177
Figure 6.6 Mean SNR for available satellites 178
Figure 6.7 Number of satellites discarded due to bad SNR 179
Figure 6.8 Number of satellites used after applying SNR mask 179
Figure 6.9 Mean SNR for satellites used 180
Figure 6.10 Number of satellites discarded due to bad SNR and elevation 181
Figure 6.11 Number of satellites used after SNR and elevation joint mask 182
Figure 6.12 Standard deviation in X-axis 183
Figure 6.13 Standard deviation in Y-axis 184
Figure 6.14 Standard deviation in Z-axis 185
Figure 6.15 Horizontal Dilution of precision (HDOP) for the satellite constellation in use 186
Figure 6.16 Acceleration in X-axis 187
Figure 6.17 Acceleration in Y-axis 188
Figure 6.18 Acceleration in Z-axis 189
Figure 6.19 Gyroscope angular acceleration in X-axis 190
Figure 6.20 Gyroscope angular acceleration in Y-axis 191
Figure 6.21 Gyroscope angular acceleration in Z-axis 191
Figure 6.22 Yaw angle 192
Figure 6.23 Acceleration in the X-axis after bias subtraction 193
Figure 6.24 Acceleration in the Y-axis after bias subtraction 194
Figure 6.25 Acceleration in the Z-axis after bias subtraction 195
Figure 6.26 Angular acceleration in the X-axis after bias subtraction 195

Figure 6.27 Angular acceleration in the Y-axis after bias subtraction 196
Figure 6.28 Angular acceleration in the Z-axis after bias subtraction 197
Figure 6.29 Residuals of the positions 198
Figure 6.30 Residuals of acceptable measurements 199
Figure 6.31 Standard deviation in X-axis 200
Figure 6.32 Standard deviation in Y-axis 201
Figure 6.33 Standard deviation in Z-axis 202
Figure 6.34 Standard deviation in X-axis after using the residuals mask 203
Figure 6.35 Standard deviation in Y-axis after using the residuals mask 204
Figure 6.36 Standard deviation in Z-axis after using the residuals mask 204
Figure 6.37 Horizontal Dilution of Precision 205
Figure 6.38 Horizontal Dilution of Precision after residuals mask 206

7 FIELD-TEST RESULTS AND PERFORMANCE EVALUATION 207

Figure 7.1 Performance of the algorithm for test TS1_EUS_1100 208
Figure 7.2 Histogram of the position error for test TS1_EUS_1100 209
Figure 7.3 CDF of the test TS1_EUS_1100 209
Figure 7.4 Performance of the algorithm for test TS1_EUS_1101 210
Figure 7.5 Histogram of the position error for test TS1_EUS_1101 211
Figure 7.6 CDF of the test TS1_EUS_1101 211
Figure 7.7 Performance in KBS zone, Test TS1_EUS_1100 vs TS1_EUS_1101 212
Figure 7.8 Performance of the algorithm for test TS2_NOT_1000 213
Figure 7.9 Histogram of the position error for test TS2_NOT_1000 214
Figure 7.10 CDF of the test TS2_NOT_1000 214
Figure 7.11 Performance of the algorithm for test TS2_NOT_0100 215
Figure 7.12 Histogram of the position error for test TS2_NOT_0100 216
Figure 7.13 CDF of the test TS2_NOT_0100 217

Figure 7.14 Performance of the algorithm for test TS2_NOT_0010 218
Figure 7.15 Histogram of the position error for test TS2_NOT_0010 219
Figure 7.16 CDF of the test TS2_NOT_0010 219
Figure 7.17 Performance of the algorithm for test TS2_NOT_1100 221
Figure 7.18 Histogram of the position error for test TS2_NOT_1100 222
Figure 7.19 CDF of the test TS2_NOT_1100 222
Figure 7.20 Performance of the algorithm for test TS2_NOT_1101 223
Figure 7.21 Histogram of the position error for test TS2_NOT_1101 224
Figure 7.22 CDF of the test TS2_NOT_1101 225
Figure 7.23 Performance of the algorithm for test TS2_NOT_1010 226
Figure 7.24 Histogram of the position error for test TS2_NOT_1010 227
Figure 7.25 CDF of the test TS2_NOT_1010 227
Figure 7.26 Performance of the algorithm for test TS2_NOT_1011 228
Figure 7.27 Histogram of the position error for test TS2_NOT_1011 229
Figure 7.28 CDF of the test TS2_NOT_1011 230
Figure 7.29 Performance of the algorithm for test TS2_NOT_0110 231
Figure 7.30 Histogram of the position error for test TS2_NOT_0110 232
Figure 7.31 CDF of the test TS2_NOT_0110 232
Figure 7.32 Performance of the algorithm for test TS2_NOT_1110 234
Figure 7.33 Histogram of the position error for test TS2_NOT_1110 235
Figure 7.34 CDF of the test TS2_NOT_1110 235
Figure 7.35 Performance of the algorithm for test TS2_NOT_1111 236
Figure 7.36 Histogram of the position error for test TS2_NOT_1111 237
Figure 7.37 CDF of the test TS2_NOT_1111 238

8 CONCLUSIONS AND FUTURE GUIDELINES 243

A PUBLICATIONS 261

List of Tables

1 INTRODUCTION AND CONTEXT OF THE RESEARCH WORK 1

2 STATE OF THE ART 9

Table 2.1 Comparison between GPS, GLONASS, GALILEO and BEIDOU 19
Table 2.2 GNSS observables 20
Table 2.3 PROs and CONs of GNSS Algorithm options (EATS n.d.) 45

3 OBJECTIVES 101

4 DESIGN AND IMPLEMENTATION OF THE ALGORITHM 105

Table 4.1 Format of GPS data in the input files 113
Table 4.2 Format of IMU data in the input files 115
Table 4.3 Format of UWB data in the input files 116
Table 4.4 Format of Ephemerides File data 118
Table 4.5 Format of the Known Blocked Scenarios database 149
Table 4.6 Format of the Data Output file 152

5 MEASUREMENT CAMPAIGNS FOR THE SYSTEM VALIDATION AND PERFORMANCE EVALUATION 153

Table 5.1 Accelerometers Performance 157

Table 5.2 Angular Rate (gyroscope) Performance 158
Table 5.3 Accelerometers Performance 161
Table 5.4 Angular Rate (gyroscope) Performance 162
Table 5.5 Scenario Comparison 166
Table 5.6 Scenario Comparison 166
Table 5.7 Description of the test cases' naming template 167
Table 5.8 List of test cases 168

6 INPUT DATA ANALYSIS 171

7 FIELD-TEST RESULTS AND PERFORMANCE EVALUATION 207

Table 7.1 Algorithm vs RTKLib comparison for test TS2_NOT_1000 215
Table 7.2 Algorithm vs RTKLib comparison for test TS2_NOT_0100 217
Table 7.3 Algorithm vs RTKLib comparison for test TS2_NOT_0010 220
Table 7.4 Algorithm vs RTKLib comparison for test TS2_NOT_1100 223
Table 7.5 Algorithm vs RTKLib comparison for test TS2_NOT_1101 225
Table 7.6 Algorithm vs RTKLib comparison for test TS2_NOT_1010 228
Table 7.7 Algorithm vs RTKLib comparison for test TS2_NOT_1011 230
Table 7.8 Algorithm vs RTKLib comparison for test TS2_NOT_0110 233
Table 7.9 Algorithm vs RTKLib comparison for test TS2_NOT_1110 236
Table 7.10 Algorithm vs RTKLib comparison for test TS2_NOT_1111 238

8 CONCLUSIONS AND FUTURE GUIDELINES 243

Table 8.1 Comparison of different positioning systems available vs presented positioning algorithm 245
Table 8.2 Performance of the algorithm vs RTKLib performance 246

A PUBLICATIONS 261

Glossary

3D	Three-Dimensional
3GPP	3rd Generation Partnership Project
4D	Four-Dimensional
A-GPS	Assisted GPS
AHRS	Attitude and Heading Reference System
AI	Artificial Intelligence
AoA	Angle of Arrival
BDS	BeiDou Navigation Satellite System
BDT	BeiDou Time
BTS	Base Transceiver Station
CDDS	Commercial Data Distribution Service
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CNSS	Compass Navigation Satellite System
CS	Commertial Service
CSNPC	China Satellite Navigation Project Centre
CGCS2000	China Geodetic Coordinate System 2000
DCM	Direction Cosine Matrix
DEM	Digital Elevation Model
DoD	Department of Defense

DOP	Dilution of Precision
ECEF	Earth-Centered, Earth-Fixed
EGNOS	European Geostationary Navigation Overlay Service
EIRENE	European Integrated Railway Radio Enhanced Network
EKF	Extended Kalman Filter
ENU	East North Up Coordinate System
E-OTD	Enhanced Observed Time Difference
ERA	European Railway Agency
e-RAIM	Enhanced Receiver Autonomous Integrity Monitoring
ERRAC	European Rail Research Advisory Council
ERSAT	ERTMS on SATELLITE – Enabling Application & Validation
ERTMS	European Rail Traffic Management System
ETCS	European Train Control System
ETS	Eusko Trenbide Sareak
EU	European Union
FCC	Federal Commission of Communications
FDE	Fault Detection and Exclusion
FDMA	Frequency Division Multiple Access
GDP	Gross Domestic Product
GEO	Geostationary Orbit
GLONASS	Global'naya Navigatsionnaya Sputnikovaya Sistema
GNSS	Global Navigation Satellite System

GPS	Global Positioning System
GSM-R	Global System for Mobile Communications – Railway
GST	Galileo System Time
GT	Ground Truth
GTRF	Galileo Terrestrial Reference System
HAS	High Accuracy Service
HDOP	Horizontal Dilution of Precision
HPL	Horizontal Protection Level
ICE	Intercity Express
IGSO	Inclined Geosynchronous Satellite Orbit
IMU	Inertial Measurement unit
INS	Inertial Navigation Service
IoT	Internet of the Things
IOV	In Orbit Validation
IPP	Ionospheric Pierce Point
IR-UWB	Impulse Radio Ultra Wide Band
ITRF	International Terrestrial Reference Frame
ITRS	International Terrestrial Reference System
KBS	Known Blocked Scenarios
KF	Kalman Filter
LAN	Local Area Network
LCS	Location Services
LLF	Local Level Frame

LOS	Line of Sight
LS	Least-Squares method
MBOC	Multiplexed Binary Offset Carrier
MC-UWB	Multi Carrier Ultra Wide Band
MEO	Medium Earth Orbit
MHSS	Multiple Hypothesis Solution Separation
MORANE	MObile RAdio for Railways Networks in Europe
MS	Mobile Station
NAVSTAR	Navigation Satellite Timing and Ranging
NGI	Nottingham Geospatial Institute
NLOS	Non-Line of Sight
OBU	On-Board Unit
OFDM	Orthogonal Frequency Division Multiplexing
OS	Open Service
OTDOA-IPDL	Observed Time Difference of Arrival - Idle Period in Down Link
PF	Particle Filter
PRN	Pseudo-Random Noise
PRS	Public Regulated Service
PSD	Power Spectral Density
PVT	Position, Velocity and Time
PZ-90	Earth Parameter System 1990
QP	Quadratic Programming

RAIM	Receiver Autonomous Integrity Monitoring
RDSS	Radio Determination Satellite System
RIMS	Ranging and Integrity Monitoring Stations
RMS	Root Mean Square
RMSE	Root Mean Square Error
RRAIM	Relative Receiver Autonomous Integrity Monitoring
RRC	Radio Resource Control
RRLP	Radio Resource LCS Protocol
RSPB	Reference Signal Power Budget
RSS	Received Signal Strength
RTCA	Radio Technical Commission for Aeronautics
RTD	Relative Time Difference
RTK	Real Time Kinematics
RTT	Round Trip Time
SA	Selective Availability
SAR	Search and Rescue Service
SBAS	Satellite Based Augmentation System
SDR	Software Defined Radio
SI	International System of Units
SIL	Safety Integrity Level
SMLS	Serving Mobile Location Centre
SoL	Safety of Life
SPS	Standard Positioning Service

STPS	Smart Train Positioning System
SU	Soviet Union
SV	Space Vehicle
TAI	International Atomic Time
TDOA	Time Difference of Arrival
TGD	Timing Group Delay
THR	Tolerable Hazard Rate
ToA	Time of Arrival
UERE	User Equivalent Range Error
UK	United Kingdom
UKF	Unscented Kalman Filter
UMTS	Universal Mobile Telecommunications System
US	United Sates
USNO	United States Naval Observatory
UTC	Universal Coordinated Time
U-TDOA	Uplink Time Difference of Arrival
UWB	Ultra Wide Band
WCT	Wireless Communications Technology
WGS84	World Geodetic System 1984

Chapter 1

Introduction and context of the research work

This chapter includes the current concerns of the railway sector regarding positioning. It includes the description of the framework within this research work lies, ERSAT-EAV, FR8RAIL, and X2R2 projects. Thereafter the research work motivation is discussed and finally, the outline of the Thesis is presented.

1.1 INTRODUCTION

The European Union (EU) aims at making railway a more attractive transportation method by improving its efficiency and reducing its costs.

In this context, Shift2Rail, the main railway innovation initiative in EU, plans to increase the attractiveness of the railway transportation by doubling the railway capacity, cutting the life-cycle costs of railway transports by as much as 50% and increasing reliability and punctuality by as much as 50% (Shift2Rail n.d.). Two of the Innovation Programs are concerned by the progresses shown in this document: IP2 and IP5 of this initiative dealing with control and communication systems and freight rail respectively. In the former, control, command and communication systems will go beyond merely being a contributor to the control and safe separation of trains, and become a flexible, real-time, intelligent traffic management and decision support system. And in the latter, the main challenge is to acquire a new service-oriented profile for rail freight services based on excellence in on-time delivery at competitive prices, interweaving its operations with other transport modes, addressing end-user needs by incorporating innovative value-added services, among others. In both cases, one of the enablers is the on-board Global Navigation Satellite System (GNSS) based positioning system, the improvement of their performance and the combination with other technologies to reach every business' case requirements. For instance, these new services resulting from IP2 and IP5 roadmap will have an important impact on the migration from European Train Control System, ETCS level 2 to ETCS level 3. This migration will allow a descent in the infrastructure costs up to 25% to regional and freight dedicated lines and efficiency improvements of more than 50% (Ramdas et al. 2010).

As a base, the suitability of GNSS systems for railway applications is being or has been analyzed by several European projects. The framework within this research work lies, includes some of the above mentioned European projects, such as ERSAT-EAV, FR8RAIL and X2R2 projects.

ERTMS on SATELLITE – Enabling Application & Validation (ERSAT-EAV)

ERSAT-EAV (ERSAT-EAV 2017) aimed to reuse the ETCS odometry and the virtual balise concept to eliminate fixed balises. Augmentation networks such as EGNOS (European Geostationary Navigation Overlay Service) are also assessed to verify and validate different GNSS solutions to guarantee the positioning functions in areas in which the GNSS signal is not accurate enough.

Development of Functional Requirements for Sustainable and Attractive European Rail Freight (FR8RAIL)

The FR8RAIL project (FR8RAIL n.d.) is part of the Shift2Rail Research and Innovation Action. The main aim of the project was the development of the functional requirements for a sustainable European rail freight. Its objectives are the reduction the 10% of the cost, the reduction of 20% in time variations during dwelling and the increase of the logistic chain information system to 100%. The objectives of the project were achieved by developing six different areas (Business Analytics, KPIs, Top Level Requirements; Condition Based and Predictive Maintenance; Telematics & Electrification; Running Gear, Core and Extended Market Wagon; Automatic Coupling and High level System Architecture and Integration.) in which positioning was included within the telematics and electrification part.

Enhancing railway signalling systems based on train satellite positioning, on-board safe train integrity, formal methods approach and standard interfaces, enhancing Traffic Management System functions (X2RAIL-2)

X2RAIL-2 (X2RAIL-2 n.d.) aims to improve the performance at a railway system level by introducing new functionalities that should revolutionize the signalling and automation concepts in the future. The key technologies cover GNSS applications in railway and its combination with other advanced technologies for implementing new signalling functionalities.

1.1 MOTIVATIONS

By 2030 mobility and business around it will be 20% of the worldwide Gross Domestic Product (GDP), which more or less will be 26.6 billion dollars (García 2020). Moreover, nowadays, the sector is one of the pillars of the global economy with more or less 15 billion dollars and a sixth of the global GDP. In this context, railways and trains plays a key role, and positioning systems are going to become even more crucial.

The European Rail Research Advisory Council (ERRAC) has defined some goals in different areas to increase the efficiency of the current railway infrastructure improving its capacity and enhancing its competitiveness (ERRAC 2012):

- Railway infrastructure should provide the technical capability to increase its traffic. For that, product innovation in control command and in passenger and freight rolling stock solutions should be done. Significant investment in research and development would lead to innovation.
- The performance of the railway infrastructure should be improved to absorb a bigger share of traffic. New solutions should be developed to reduce the lifetime cost of the infrastructure and target the new interoperability requirements approved in terms of safety and security, reliability, maintainability, and interoperability.
- To improve the information management systems with high-quality services and implementation of intelligent mobility concepts involving customer information for freight and passenger services for enhanced accessibility and availability.
- To develop innovative and advanced rolling stock, signalling and infrastructure solutions with cost-competitive technologies, including retrofitting solutions. To this end, research and innovation will have to improve the performance of products, production processes and reduced life cycle costs (benefiting from economies of scale) to improve the economic attractiveness of the rail transport mode.
- To reduce cost on product certification and validation, fleet operation and maintenance costs; this being one of the most significant areas of expenditure for the

railways. This will potentially release substantial capital for further investment that supports a greater modal shift.

The points above presented could be covered with the migration from ETCS level 2 to ETCS level 3. Many projects related to new positioning systems have been funded by the European Union. And the most of these positioning systems are based on GNSS and its already deployed infrastructure, due to the high potential of playing a key role that GNSS plays in the fulfilment of ETCS level 3 requirements .

One of the problems on using only GNSS systems is the lack of availability. During railway operation, there are areas with potential GNSS outages, such as urban canyons, woods or other possible line of sight signal blockers and disturbances. Moreover, it is a fact that GNSS signals are not reachable in tunnels or indoor environments, and if any data is received there are not reliable at all. For GNSS to be able to have a key role in the next years in railway operation, the before mentioned lack of availability has to be solved.

The scope of the dissertation is to create a multi-sensor positioning system including GNSS, Inertial Measurement Unit (IMU), and Ultra Wide Band (UWB) with other software techniques to obtain a position estimation with a 100% availability and to assess its results for the use into the railway domain.

1.2 OUTLINE OF THE RESEARCH WORK

This subsection includes the outline followed in this dissertation:

Chapter 1: Introduction and context of the research work

This chapter includes the current concerns of the railway sector regarding positioning. It includes the description of the framework within this research work lies, ERSAT-EAV, FR8RAIL, and X2R2 projects. Thereafter the research work motivation is discussed and finally, the outline of the Thesis is presented.

Chapter 2: State of the art

This chapter presents the basic concepts of positioning systems and other technologies and techniques that have been studied during the elaboration of this thesis. The aim is to provide the reader with a basic overview of different positioning systems and their fusion. These definitions would help the reader to understand the contribution of this research work.

Chapter 3: Objectives

Chapter 3 presents the objectives of this work. After analyzing the information introduced in Chapter 1 and Chapter 2, the gaps identified in the literature have been analyzed and possible solutions to them will be stated in pursuit of evolving the positioning strategies and adapting them to the new challenges of the railway transport applications.

Chapter 4: Design and implementation of the algorithm

This chapter describes the design and implementation of the positioning algorithm created. First, the whole architecture of the algorithm is introduced. Then, the implementation and functionality of each of the modules included in the algorithm are detailed.

Chapter 5: Measurement campaigns for system validation and performance evaluation

This chapter describes the organization of the measurement campaigns that have been done in order to evaluate the performance of the proposed system. The chapter describes the different environments in which the system has been tested and the equipment used in each of them. In this way, a good characterization of the areas in which the algorithms are tested is introduced.

Chapter 6: Preliminary data analysis

This chapter includes the data analysis of the input data obtained from the measurement campaigns done and to be used in the algorithms. The objective is to check that the data obtained are coherent and to show how each of the modules introduced in the algorithm affects the resulting positioning function.

Chapter 7: Field-test results and performance evaluation

This chapter shows and analyzes the results obtained by the positioning algorithm in the measurement campaigns explained in Chapter 5. Moreover, it analyzes the performance obtained by the algorithm in comparison with the ground truths of each of the scenarios.

Chapter 8: Conclusions and future guidelines

This chapter summarizes the conclusions achieved in the presented dissertation. Moreover, it includes some open research areas that might be interesting to carry out future work on them.

Chapter 2

State of the Art

This chapter presents the basic concepts of positioning systems and other technologies and techniques that have been studied during the elaboration of this thesis. The aim is to provide the reader with a basic overview of different positioning systems and their fusion. These definitions would help the reader to understand the contribution of this research work.

2.1 GNSS POSITIONING SYSTEMS

2.1.1 GPS

GPS is the simplified acronym of Global Positioning System and is the first system of the Global Navigation Satellite Systems (GNSS) which was developed by the Department of Defense (DoD), United States. The full name of GPS is NAVigation Satellite Timing and Ranging/Global Navigation System (NAVSTAR/GPS). GPS was originally designed as a navigation system for the U.S. military users. However, later on, it was also available for civilians, and now is a dual-use system for both military and civilian users. GPS is a one-way-ranging (passive) system that provides worldwide, 24-hours real-time, continuous, accurate, three-dimensional position, velocity and timing information to the users with appropriate receivers. GPS consists of three segments, namely, the Space Segment, the Control Segment, and the User Segment. The Space Segment deals with the launching of satellites. The Control Segment monitors and manages the operation of the satellites. The User Segment relates to both military and civil receiver equipment development.

The baseline GPS constellation contains 24 Medium Earth Orbit (MEO) satellites available 95% of the time in six Earth-centered orbital planes with a radius of 26,560 km (i.e. about 20,163 km above the Earth). To ensure this commitment, the Air Force has been flying 31 operational GPS satellites. Each plane hosts four satellites. The orbits are nearly circular and are separated by 60o around the equator with a 55º inclination relative to the equatorial plane. The GPS satellites have an orbit period of one-half of a sidereal day or 11 hours, 58 minutes. The first GPS satellite was launched on 22nd February 1978 belonging to the first generation of GPS satellites, indicated as Block I. Block I comprised 12 satellites: 11 were successfully launched from 1978 to 1985, while one failed to be launched. The purpose of Block I was to

build up the ground track network and test the GPS receiver performance and the possibility of the GPS operation (GPS.gov: Space Segment n.d.).

The second generation of GPS satellites is known as Block II/IIA which was developed for the first operational constellation. Block IIA ("A" denotes "advance") is an advanced version of Block II with an enhancement of the navigation message storage. The first Block II satellite was successfully launched on 14th February 1989 and the first Block IIA satellite was launched on 26th November 1990. From 1989 to 1997, a total of 28 Block II/IIA satellites were launched. The full operation constellation of GPS was declared in April 1995.

A GPS satellite transmits navigation signals on two carrier frequencies (or sine waves) called L1 (1575.42 MHz), the primary frequency, and L2 (1227.60 MHz), the secondary frequency. These signals are generated synchronously, and if both signals are received by a user, the ionospheric delay can be calibrated. Most civilian users, however, for example in the railway and road domains, only used one frequency (normally L1) due to the cost of the dual frequencies receivers, but this has been changing in the 2010's decade. The carrier frequencies are modulated by spread spectrum codes with unique PRN sequences (or PRN codes) associated with each Space Vehicle (SV), and navigation message data. All SVs translate into the use of a CDMA (Code Division Multiple Access) technique.

As for the GPS modernization program development (Block IIR-M, Block IIF, and GPS III), a new civil code will be modulated on the L2 frequency (known as L2C). A new military code (M-code) will also be broadcast on both L1 and L2 frequencies. The availability of two civilian codes will allow a stand-alone GPS receiver to calibrate the ionospheric delay. Furthzer, GPS Block IIF satellites provide a new third frequency (L5, 1176.45 MHz), targeted primarily at safety-critical applications. GPS III satellites will also provide this signal as well as a fourth civil signal, L1C. L1C features a Multiplexed Binary Offset Carrier (MBOC) modulation scheme that enables international cooperation while protecting U.S. national security interests.

The design will improve mobile GPS reception in cities and other challenging environments(GPS.gov: New Civil Signals n.d.).

GPS provides two services separately. One is the Standard Positioning Service (SPS) for the civil users. The other one is the Precise Positioning Service for the DoD authorized military and government agency users which can provide better accuracy than the SPS. The SPS is free of charge to all users worldwide. It provided an accuracy of about 100m (95%) in the horizontal plane and 156m (95%) in the vertical plane until May 2000 because selective availability (SA) technology intentionally degraded performance. On 1st May 2000, SA was discontinued by U.S. Currently, the performance of the SPS is about 3.8m (1σ) in the horizontal plane and 6.2m (1σ) in the vertical plane (P. D. Groves 2008).

The World Geodetic System 1984 (WGS84) is the earth model used for GPS applications. In WGS84, the coordinate origin is located at the center of the mass of Earth. It provides an ellipsoidal model of the Earth's shape in which the semi-major axis “a” is 6,378.137 km and the semi-minor axis “b” is 6,356.7523142 km. The position of a GPS receiver can be presented in WGS84 as the parameters of latitude, longitude, and height (which are also called geodetic coordinates), or using the Earth-centered Earth-fixed (ECEF) coordinate system. In ECEF the positive Z-axis points to the North Pole and the X-Y plane is the equatorial plane in which the X-axis is along the prime meridian (Kaplan and Hegarty 2006). The position of a receiver in ECEF coordinates (x,y,z) can be converted from geodetic coordinates (λ,ϕ,h) by the following equation.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \dfrac{a \cdot \cos\lambda}{\sqrt{1+(1-e^2)\cdot \tan^2\phi}} + h \cdot \cos\lambda \cdot \cos\phi \\ \dfrac{a \cdot \sin\lambda}{\sqrt{1+(1-e^2)\cdot \tan^2\phi}} + h \cdot \sin\lambda \cdot \cos\phi \\ \dfrac{a\cdot(1-e^2)\cdot\sin\phi}{\sqrt{1-e^2\cdot\sin^2\phi}} + h\cdot\sin\phi \end{bmatrix} \tag{2.1}$$

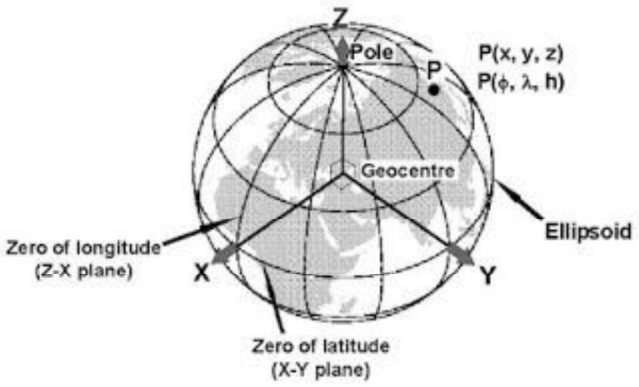


Figure 2.1 ECEF Coordinate frame

GPS uses its time reference frame, known as GPS time. It is expressed by the GPS week number and seconds of the week. It is synchronized to Coordinated Universal Time, U.S. Naval Observatory (UTC USNO) but with an offset due to leap seconds. Currently, the leap seconds offset is 17 seconds. The difference between GPS time and UTC is included in the GPS navigation message.

2.1.2 GLONASS

GLONASS is the Russian counterpart system to GPS. The Russian acronym stands for GLObal'naya NAvigatsionnaya Sputnikovaya Sistema. GLONASS was developed to provide position, velocity and timing determination. It was initially developed by the former Soviet Union and currently is operated by the Russian government. The development of GLONASS started in 1976 and the first GLONASS satellite was launched on 12th October 1982. Originally, GLONASS was designed, again, for Soviet Union military users, but now it also serves as a dual-use system for both civil and military users. The nominal constellation is composed of 24 active MEO satellites (21 active satellites + 3 active spares) in three orbital planes separated by 120 degrees. The satellites operate in circular 19,100-km orbits at an inclination of 64.8 degrees to the Earth's surface. The GLONASS satellites have an orbit period of about 11 hours and 15 minutes. GLONASS provides a continuous

navigation service with global coverage(GLONASS Space Segment - Navipedia n.d.).

GLONASS was declared fully operational in February 1996. However, due to the financial problems of the Russian government, the system fell rapidly in the following five years without sufficient constellation maintenance. By the end of 2001, it operated with only eight satellites. On 20th August 2001, the Russian government decided to rebuild GLONASS. In October 2011, the full 24 satellite constellation was restored.

The time reference systems are different between GLONASS and other GNSS. The GLONASS time system is linked to Coordinated Universal Time, Soviet Union (UTC SU). GLONASS also uses its coordinate frame to express the position of the satellites, namely the Earth Parameter System 1990 (PZ-90). The maximum difference between PZ-90 and WGS84 could be 20m on the Earth surface. These differences need to be considered for the integration of GLONASS with other GNSS.

2.1.3 GALILEO

Galileo is a European GNSS, initially built by the European Union and European Space Agency. It is an independent system from GPS and any other country's satellite-based positioning system. Unlike GPS and GLONASS, the Galileo system is specifically designed for civilian use, providing high accuracy and globally available positioning services. The idea of Galileo began in the early 1990s, and the different concepts for Galileo were unified to one by the agreement of four EU countries (the United Kingdom, German, Italy, and France) at the end of 1999. In the year 2000, the feasibility and definition phases of the Galileo system were finally completed. The Galileo system will consist of 30 MEO satellites, divided within three operational orbital planes at an altitude of 23,616 km above the Earth's surface and

with an inclination of 56 degrees (Galileo General Introduction - Navipedia n.d.). The orbit period for a Galileo satellite is about 14 hours and 22 minutes. Each orbital plane will host nine operational satellites and one active spare satellite. Currently, twenty-two satellites have been launched. Even if the system is nowadays operative, the full operational constellation is expected to be ready by the end of 2020 (Constellation Information | European GNSS Service Centre n.d.).

The Galileo system has planned to provide five major services which are defined as Open Service (OS), High Accuracy Service (HAS), Safety of Life (SOL), Commercial Service (CS), Public Regulated Service (PRS), and the support to Search and Rescue Service (SAR) (ESA - Galileo services n.d.). These services are and will be provided worldwide and independently from other satellite navigation systems by using the signals broadcast by the Galileo satellites when they will be available. The OS is designed for mass-market users, to provide PVT information that can be accessed free of direct charge, suitable for applications such as in-car navigation and location system in mobile phones. The OS will be available for all the users who have Galileo capable receivers. However, if the receivers are integrated with other GNSS, navigation performances are improved especially in severe environments, such as urban canyon areas and forests. The HAS is a service complementing the OS with an additional navigation signal and added services in a different frequency band to the OS. The SoL improves the OS performance giving timely warnings to the user when it fails to meet some integrity margins. The CS gives access to two additional signals to allow a higher data rate to improve accuracy. Moreover, it uses encrypted signals. The PRS is restricted to authorized users that need a high level of service continuity. It is encrypted and designed to be robust against jamming mechanisms and it is intended to be used by security and strategic infrastructures. The SAR is the worldwide search and rescue service provided by GALILEO (Galileo General Introduction - Navipedia n.d.).

The geodetic coordinate reference frame for Galileo is called the Galileo Terrestrial Reference System (GTRF), which is also an independent realization of

the International Terrestrial Reference System (ITRS) (IERS - IERS - The International Terrestrial Reference System (ITRS) n.d.). GTRF only differs from WGS-84 (which is also a realization of the ITRS) by a few centimetres. Therefore, GTRF and WGS-84 are compatible with this accuracy level for most users.

The time reference frame of the Galileo system is also different from the GPS time system. Galileo uses Galileo System Time (GST) based on the international atomic time (TAI) while GPS time is based on UTC USNO. Currently, the Galileo IOV satellites have begun broadcasting the difference between these two time systems.

2.1.4 BeiDou Chinese COMPASS System

Compass Navigation Satellite System (CNSS) is the Chinese-developed GNSS. BeiDou is the Chinese name for this system; therefore it is also known as BeiDou navigation satellite system (BDS). BDS is a multistage program operated by the China Satellite Navigation Project Centre (CSNPC). The first stage, called BeiDou-1, is a navigation system at the test stage. This means that until June 2020 is only going to be available for local operation (Directions 2020: BeiDou in the new era of globalization - GPS World : GPS World n.d.) (Government 2012). Three prototype BeiDou-1 satellites were launched between October 2000 and May 2003. BeiDou-1 was fully operational at the beginning of 2004 and provided services to users over China and surrounding areas. Therefore, the BeiDou-1 system was originally a regional satellite navigation system. Unlike GPS, GLONASS and Galileo, which are passive-systems employing one-way TOA measurements, the BeiDou-1 system provides a radio determination satellite service (RDSS) which requires two-way range measurements to avoid synchronizing the receiver clock. With an estimated user altitude, the RDSS requires only two satellites to locate the two-dimensional user position at the operation center.

In October 2006, the China National Space Administration decided to upgrade and fully implement the BeiDou-1 system to the next stage, known as the BeiDou-2 system. The BeiDou-2 system is the fourth GNSS in the world (Beidou satellite navigation system to cover whole world in 2020 2011). The current design for the Beidou-2 system consists of a constellation of 27 MEO satellites, 3 Inclined Geosynchronous (IGSO) satellites, and five Geostationary (GEO) satellites. The MEO satellites will be equally split to six orbital planes at an altitude of about 21,500 km above the Earth's surface and with an inclination of 55 degrees. The BeiDou-2 system, when fully operational, will provide two services: an open civilian service, and one for military/government users. In December 2012, fourteen BeiDou-2 satellites were in service, providing regional GNSS PVT solutions to China and surrounding areas. A second version of the BDS was released in December 2013, covering two civilian signals namely B1I (centered at 1561.098 MHz) and B2I (centered at 1207.140 MHz). A performance standard for the current regional capability was also released.

The geodetic coordinate reference frame for BDS is the China Geodetic Coordinate System 2000 (CGCS2000) which is based on the CGCS2000 ellipsoid with $a = 6378137.0\,\text{m}$, $f = 1/298.257222101$ being *a* the semi-major axis, and *f* the flattening of the Earth.

BDS also uses its time reference system which is called BeiDou navigation satellite system Time (BDT). The BDT uses the international system of units (SI) seconds as the basic unit, and it is a continuous time scale. The start epoch of BDT is 00:00:00 on 01/01/2006 (UTC). It is synchronized with UTC through UTC NTSC. The difference is controlled within 100 nanoseconds. The leap seconds are broadcasted in the navigation message (China Satellite Navigation Office 2012). Table 2.1 gives a comparison of different GNSS types.

	GPS	GLONASS
Number of satellites	24	24
Orbital planes	6	3
Satellite per orbital plane	4	8
Inclination of orbital (deg)	55	64.8
Altitude (km)	20163	19100
Orbital period	11h58m	11h15m
Repeat ground path	8 sidereal days	1 sidereal day
Signal separation technique	CDMA	FDMA
Satellite coordinate frame	WGS84	PZ-90
Time reference	GPST	GLONASST
Carrier frequency	L1: 1575.42 MHz L2: 1227.60 MHz	L1: 1602.5625-1615.5 MHz L2: 1246.4375-1256.5 MHz
C/A Code rate	1.023 MHz	0.511 MHz
P-code rate	10.23 MHz	5.11 MHz

	GALILEO	BEIDOU
Number of satellites	30	5 GEO, 3 IGSO and 27 MEO
Orbital planes	3	6
Satellite per orbital plane	10	5
Inclination of orbital (deg)	56	55 for IGSO and MEO, 0 for GEO
Altitude (km)	23222	21528 for MEO and 35786 for GEO and IGSO
Orbital period	14h	12 h 53 min 24 s
Repeat ground path	10 sidereal days	-
Signal separation technique	AltBOC	MBOC
Satellite coordinate frame	GTRF	CGCS2000
Time reference	GST	BDT
Carrier frequency	E1: 1575.42 MHz E6: 1278.75 MHz E5:1191.795 MHz E5a: 1176.45 MHz E5b: 1207.14 MHz	B1: 1575.42 MHz B2: 1191.79 MHz B3: 1268.52 MHz
C/A Code rate	1.023MHz	1.023MHz
P-code rate	10.23 MHz	10.23 MHz

Table 2.1 Comparison between GPS, GLONASS, GALILEO and BEIDOU

2.1.5 GNSS Observations

Standard GNSS receivers provide up to four observations per satellite and per frequency at a specified rate, typically 1Hz, as described by the table below (Table 2.2):

Observable	Definition	Units
Pseudorange	Distance from receiver antenna to satellite antenna including receiver and satellite clock offsets, and other biases.	Meters
Carrier phase	The carrier phase measurement is a measure of the range between a satellite and receiver expressed in units of cycles of the carrier frequency.	Wavelengths (cycles)
Doppler	Doppler shift with respect to nominal signal frequency; positive for approaching satellites. Used for calculation of user velocity.	Hz
Signal strength	Measurement of the strength of the received signal, dependent on the degree of thermal, background and intermodulation noise to which the signal has been subjected	dB/Hz

Table 2.2 GNSS observables

2.1.5.1 Pseudoranges

A pseudorange is the distance measured between the satellite and the receiver in a transmission epoch, obtained by multiplying the time of arrival of the signal with

the speed of light. The time of flight is computed aligning the PRN sent by the satellite and the one generated by the receiver. Using P-code, accuracies up to 30cm can be obtained when positioning with pseudoranges (Wells et al. 1999).

Pseudoranges are basic GNSS observables used for navigation, time and satellite movement solution. Pseudoranges are affected by tropospheric and ionospheric delays (Leick, Rapoport, and Tatarnikov 2015).

The signal sent by the satellites travels across the atmosphere that will introduce some delay in its reception. The environment near the receiver will produce reflections that will make the distance from the satellite to the receiver larger, what at the end is a larger travel time for the signal. All these effects should be then taken into account when correlating the received signal with the replica generated in the receiver. The functional model for pseudorange observables is introduced in the following equation:

$$p = \rho + c(dt - dT) + d_{ATMOSPHERE} + \epsilon_p \quad (2.2)$$

Where

p … Pseudorange observable

ρ … Real satellite-receiver distance

dt, dT … Satellite and receiver corrections in the GNSS time scale

$d_{ATMOSPHERE}$ … Effect in the distance due to the atmosphere

c … Light speed in vacuum

ϵ_p … Distance error correction produced by random effects

The coordinates of the receiver appear implicit in the real satellite to receiver distance

$$p = \|\rho_S - \rho_E\| = \sqrt{(x_s - x_E)^2 + (y_s - y_E)^2 + (z_s - z_E)^2} \quad (2.3)$$

With (x_S, y_S, z_S) and $(x_{E,}\, y_E, z_E)$ being the ephemeris of the satellite and the receiver position.

The satellite ephemeris required for the position computation is sent in the navigation message. They can also be received via radio or IP from different companies that perform precise satellite tracking. The satellite coordinates can be easily computed using a mathematical model.

Using the following second-order polynomial,

$$dt = a_0 + a_1(t - t_0) + a_2(t - t_0)^2 \quad (2.4)$$

The existing time error between timestamps is minimized. The coefficients of the polynomial are sent in the navigation message and are updated by GNSS control stations.

The atmosphere affects pseudoranges in two parts of it; ionosphere and troposphere. Ionospheric effects can be reduced by using observables in frequency band L1 and L2 or modeling the effects. On the other hand, tropospheric effects can only be modeled.

The new mathematical model including the atmospheric effects is now given by the following equation for each of the satellites.

$$p^i = \sqrt{(x_S^i(t) - x_E)^2 + (y_S^i(t) - y_E)^2 + (z_S^i(t) - z_E)^2} + c \cdot dT + \epsilon \quad (2.5)$$

Being the term e, the sum of all the random errors explained before that could vary the time of arrival of the signal (Berrocoso, Ramírez, and Pérez-Peña n.d.).

2.1.5.2 Carrier Phase

Carrier phase measurements are obtained comparing the phase of the received signal with the phase generated by the receiver's oscillator. The carrier has a higher

frequency than the emitted codes which leads to better accuracy in the satellite-receiver distance calculus than the one obtained using pseudorange measurements. However, the phase measurement is affected by the uncertainty of the emitted cycles from the satellite until the receiver does the signal acquisition. This concept is known as ambiguity. The carrier phase measurement is composed of a certain number of full cycles, a fractional part of the last cycle and the ambiguity recorded. The recorded total phase multiplied by the wavelength of the signal, once ambiguity has been solved, gives the position of the satellites (Langley 2008) (Teunissen 1998) (Wells et al. 1999).

Phase measurements can also be done during a period instead of doing them in a time instant. In that case, the Doppler frequency could be measured. The carrier frequency received is variable depending on the distance between the satellites and the receiver, and can be used to determine the position of the satellites (Berrocoso, Ramírez, and Pérez-Peña n.d.).

GNSS satellites emit a sinusoidal signal generated in their oscillators. The frequency of these signals is the derivative of the phase, and so the phase of the signal is the integral of the frequency, as represented by the following equation (Teunissen and Kleusberg 1998).

$$f(t) = \frac{d\phi(t)}{dt}, \phi(t) = \int_{t_0}^{t} f(\tau)d\tau + \phi(t_0) \quad (2.6)$$

Where $\phi(t_0)$ is the number of cycles of the initial phase of the signal in time instant t_0. From the oscillator-measured phase, the time can be obtained using the following function

$$\psi(t) = \frac{\phi(t) - \phi(t_0)}{f_0} \quad (2.7)$$

Where f_0 is the nominal frequency of the satellite's oscillator, which is the frequency not affected by the Doppler Effect or other effects generated by the satellite translation movement.

If the oscillator's frequency is constant and equal to its nominal frequency we have

$$\phi(t) - \phi(t_0) = \int_{t_0}^{t} f(\tau)d\tau = f_0(t - t_0) \tag{2.8}$$

Where

$$\psi(t) = t - t_0 \Leftrightarrow\ t = \psi(t) + t_0 \tag{2.9}$$

Being t the exact time in the GNSS scale and $\psi(t)$ the time obtained from the measured phase. However, the oscillator's frequency is not always equal to the nominal frequency, and so there will be a frequency drift Δf comparing to the nominal frequency.

$$f(t) = f_0 + \Delta f(t) \tag{2.10}$$

And therefore,

$$\psi(t) = \frac{1}{f_0}\int_{t_0}^{t} f(\tau)d\tau = \frac{1}{f_0}\int_{t_0}^{t} \big(f_0 + \Delta f(\tau)\big)d\tau = t - t_0 + \int_{t_0}^{t} \frac{\Delta f(\tau)}{f_0} d\tau \tag{2.11}$$

The frequency drift will also induce a time drift modeled in the following way,

$$\Delta\psi(t) = \int_{t_0}^{t} \frac{\Delta f(\tau)}{f_0}\, d\tau \tag{2.12}$$

So we can obtain a relationship between GNSS time and oscillator's time.

$$\psi(t) = t + \Delta\psi(t) - t_0 \tag{2.13}$$

Summing up, the time obtained from the phase measurement, is the addition of the real time scale, the frequency drift, and the initial phase drift. Naming the oscillator's error $d\psi(t)$ we have,

$$\psi(t) = t + d\psi(t) \tag{2.14}$$

Where, as before, $\psi(t)$ is the GNSS time scale, t is the exact time scale and $d\psi(t)$ is the oscillator's error. For simplifying the notation, the oscillators allocated in the satellites are going to be expressed as follows,

$$t^k(t) = t + dt^k(t) \tag{2.15}$$

And the oscillators allocated in the receivers in the following way,

$$t_i(t) = t + dt_i(t) \tag{2.16}$$

As explained before, the total carrier phase is an integer number of cycles $Int(\phi, t_0, t)$, starting the count when the signal t_0 is received until the instant t; a fractional part $Fr(\phi)$, referred to the signal acquisition instant; and an unknown integer number of cycles $N(t_0)$ corresponding to the time elapsed between the signal emission and the reception in the receiver. These number of cycles are called initial ambiguity and are different for each receiver and satellites.

$$\phi_{total} = Fr(\phi) + Int(\phi, t_0, t) + N(t_0) \tag{2.17}$$

The ambiguity value is not computable so $\phi_{measurement}$, the carrier observable is

$$\phi_{measurement} = Fr(\phi) + Int(\phi, t_0, t) \tag{2.18}$$

And so

$$\phi_{total} = \phi_{measurement} + N(t_0) \tag{2.19}$$

The phase of the observed carrier $\phi_{measured} = \phi_i^k$ is the difference between the phase generated by the receiver's oscillator ϕ_i and the phase generated by the satellite$k$,ϕ_k.

$$\phi_{measurement} = \phi_i^k = \phi_i(t) - \phi^k(t - \tau_i^k) \quad (2.20)$$

Being all the phases represented in cycles. Reformulating, we have the following,

$$\begin{aligned} \phi_i(t) &= f_0 t_i(t) + \phi_i(t_0) = f_0\big(t + dt_i(t)\big) + \phi_i(t_0) \\ \phi^k\big(t - \tau_i^k\big) &= f_0\left(t - \tau_i^k + dt^k\big(t - \tau_i^k\big)\right) + \phi^k(t_0) \end{aligned} \quad (2.21)$$

And then the measured carrier phase is,

$$\phi_i^k = f_0\big[\tau_i^k + dt_i(t) - dt^k(t - \tau_i^k)\big] + [\phi_i(t_0) - \phi^k(t_0)] \quad (2.22)$$

Multiplying the measured carrier phase with the wavelength, $\lambda = \frac{c}{f_0}$ the measured carrier phase equation is obtained in length units.

$$\lambda\phi_i^k = c\tau_i^k + c\big[dt_i(t) - dt^k(t - \tau_i^k)\big] + \lambda[\phi_i(t_0) - \phi^k(t_0)] \quad (2.23)$$

In an analog way to the pseudorange positioning, $c\tau_i^k$ and $c[dt_i(t) - dt^k\big(t - \tau_i^k\big)]$ show the traveled distance from the satellite to the receiver and the oscillator errors in the satellites and the receiver respectively. The term $\lambda[\phi_i(t_0) - \phi^k(t_0)]$ is the distance term corresponding to the fractional part of the measured phase in the reception time.

Moreover, the refraction index of the ionosphere is

$$n_\phi = 1 - \alpha\frac{D_e}{f^2} \quad (2.24)$$

Where α is a constant, D_e is the electronic density, and f is the emission frequency (Langley 1998).With the multiplication of the constants α and D_e having a value of 40,3.

Integrating the expression above during the traveled distance,

$$\rho_\phi^{ion} = \int \left(1 - \alpha\frac{D}{f^2}\right) ds = \rho^{ion} - d_{ion} \quad (2.25)$$

And substituting the terms corresponding to the distance in the equation 25 introduced before for distance results in,

$$\lambda\phi_i^k = \rho_i^k\left(t, t-\tau_i^k\right) - d_{ion} + d_{trop} + c\left[dt_i(t) - dt^k\left(t-\tau_i^k\right)\right] + \lambda[\phi_i(t_0) - \phi^k(t_0)] \quad (2.26)$$

And the observation equation for the carrier total phase can now be written in the following way:

$$\phi_i^k(t) = \rho_i^k\left(t, t-\tau_i^k\right) - d_{ion} + d_{trop} + c\left[dt_i(t) - dt^k\left(t-\tau_i^k\right)\right] + \lambda[\phi_i(t_0) - \phi^k(t_0)] + \lambda N_i^k \quad (2.27)$$

Implementing the following considerations,

$$\Phi = \Phi_i^k(t), \qquad \rho = \rho_i^k\left(t, t-\tau_i^k\right), \qquad c(dt-dT) = c\left[dt_i(t) - dt^k(t-\tau_i^k)\right] \quad (2.28)$$

The observation equation can be simplified to the following terms,(Wells et al. 1999)

$$\Phi = \rho + c(dt - dT) + \lambda N - d_{ion} + d_{trop} + \epsilon \quad (2.29)$$

The obtained equation is analog to the one obtained in pseudorange positioning. ρ is the distance between the satellite and the receiver and dt and dT are the errors introduced by the satellite and receiver oscillators. The parameter ϵ represents the random errors accumulated during the whole calculus. Finally, the parameter N is the number of complete cycles counted until the receiver acquires the signal. That is the only difference between the pseudorange and carrier phase positioning.

Note that the ionosphere correction term is a substraction instead of an addition in the pseudorange positioning model. That occurs due to the pseudorange ionosphere refraction model,

$$n_p = 1 + \alpha\frac{D_e}{f^2} \quad (2.30)$$

And so,

$$\rho_p^{ion} = \int \left(1 + \alpha \frac{D_e}{f^2}\right) ds = \rho^{ion} + d_{ion} \tag{2.31}$$

Prior to being processed within a navigation engine, these observables must undergo a certain amount of pre-processing. This usually involves the following steps where a stand-alone solution (no augmentation source) is concerned (not necessarily in this order):

- Application of elevation mask (removal of satellites below it from current epoch calculations) – this requires an approximate user position which can be based on the previous epoch solution when available to calculate satellite elevations.
- Convert Doppler shift observations to range rates using the following formula:

$$\dot{\rho}_i = -c \cdot \frac{D_i}{f} \tag{2.32}$$

where,

$\dot{\rho}_i$ = range-rate for Satellite i;

D_i = Doppler observable for Satellite i;

f = frequency of signal;

c = speed of light.

- Correct pseudoranges for the ionospheric delay effect. This is typically done using one of the following two methods:
  - Single-frequency case: use of a model. The most common is the Klobuchar model, for which parameters are broadcast in the GPS navigation message. This model can be applied to any GNSS.

GLONASS does not transmit such model parameters, however, Galileo broadcasts parameters for its model, NeQuick.

- Dual-frequency case: formation of ionospheric-free observable. When observations on a second frequency are available, they may be used in combination with those of the first frequency to form pseudoranges which are free from ionospheric delay:

$$\rho_{IF,i} = \frac{{f_1}^2}{{f_1}^2 - {f_2}^2} \cdot \rho_{1,i} - \frac{{f_2}^2}{{f_1}^2 - {f_2}^2} \cdot \rho_{2,i} \quad (2.33)$$

where,

$\rho_{IF,i}$ = ionosphere-free pseudorange for satellite i;

$\rho_{1,i}$ = pseudorange on the first frequency for satellite i;

$\rho_{2,i}$ = pseudorange on the second frequency for satellite i;

f_1 = first frequency (MHz);

f_2 = second frequency (MHz).

- Correct ionosphere-corrected pseudoranges for tropospheric effect, for example, using the RTCA model described in (RTCA 2006).
- In the case of multi-constellation solutions, it may be chosen to apply system time corrections to pseudoranges at this point to achieve synchronization.
- If GLONASS is being used within a multi-constellation solution then a correction for the leap seconds present in GLONASS system time may be made here.
- Apply satellite clock offset and drift corrections obtained from navigation data to pseudoranges.

- Calculate User Equivalent Range Error (UERE) budget based on internal models (in the stand-alone case) and based on calculated satellite elevations.

2.1.6 Computation methods for GNSS Positioning

As already presented in the last paragraphs, positioning with GNSS is based on the Time-of-Arrival (TOA) concept which is used to calculate the ranges between the user and satellites. The range is derived from measured time difference by comparing the received PRN codes in the satellite signal and the receiver-generated PRN code. In Figure 2.2, r_i represents the geometric range between the i^{th} satellite and the user. If the satellite time, at which the signal was transmitted, is designed as t_s , and the receiver time, at which the signal arrived at the receiver, is designed as t_u ; then for each range:

$$r = c \cdot (t_u - t_s) = c \cdot \Delta t \tag{2.34}$$

However, the receiver clock and the satellite atomic clock will have a bias error concerning the true system time. The transmitted signal in space will also be delayed by various other error components, such as ionospheric error, tropospheric error, and multipath error. Therefore, the observed range, as opposed to the geometric range, is called a pseudorange measurement, which is denoted as ρ and can be expressed as:

$$\rho_i = r_i + c \cdot (t_u - t_s) + d\rho_i + I_{\rho_i} + T_{\rho_i} + \varepsilon_{\rho_i} \tag{2.35}$$

where,

t_u is the receiver clock offset with respect to the system time

t_s is the satellite clock offset with respect to the system time

$d\rho_i$ is the ephemeris error

I_{ρ_i} is the ionospheric delay

T_{ρ_i} is the tropospheric delay

ε_{ρ_i} are the errors including multipath, hardware bias, and receiver noise.

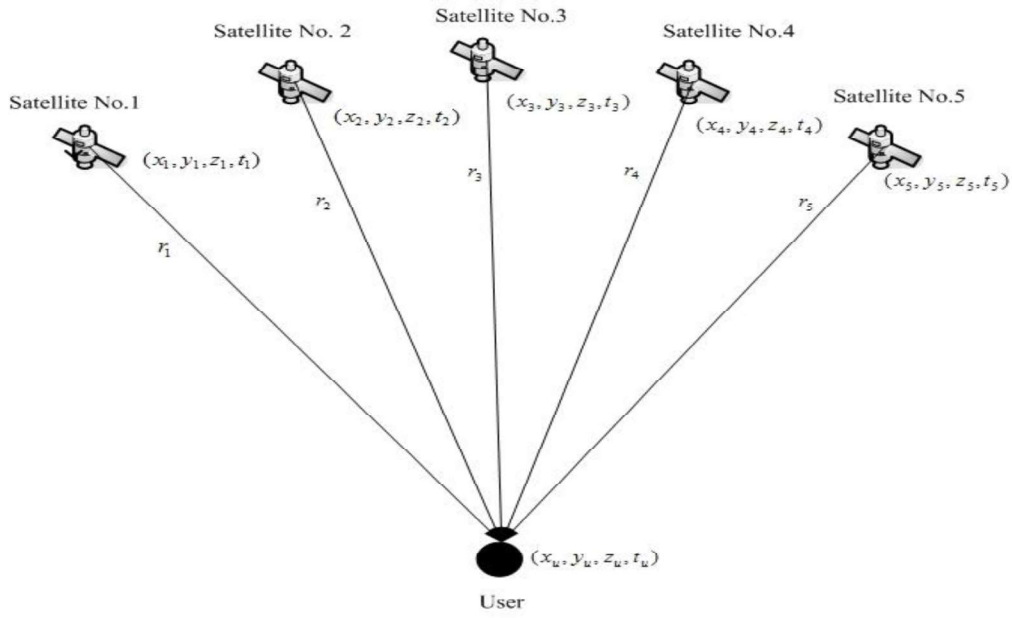


Figure 2.2 Satellite to user ranges

The receiver clock error is the offset of the receiver clock from the reference GNSS time. Depending on the quality of the oscillator used in the GNSS receiver,

the error could range from 200 ns up to a few ms, and the measured ranges could, therefore, vary from a few meters to a few thousand kilometers. It is normally treated as an unknown parameter in the position computations. Hence, equation (2.35) can be expressed as

$$\rho_i = r_i + c \cdot t_u + \varepsilon_i \tag{2.36}$$

where ε_i is the composite of errors produced by, e.g. atmospheric delays, multipath, satellite ephemeris errors.

Two principal methods for the computation of a user position solution using GNSS signals may be considered; namely the snapshot method and the Kalman filter.

2.1.6.1 SnapShot method for position computation

2.1.6.1.1 Position Engine

This section outlines the principles of a snapshot least-squares algorithm. In Figure 2.2, the satellites' positions are expressed as (x_i, y_i, z_i) and the user position is denoted as (x_u, y_u, z_u), for which x, y, z are the values in the Earth-Centered Earth-Fixed (ECEF) coordinate system. Therefore, the geometric range, r_i, can be computed by the following equation:

$$r_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} \tag{2.37}$$

Substituting Equation 2.36 into Equation 2.37, the pseudorange equation can be expressed as

$$\rho_i = \begin{cases} \sqrt{(x_1 - x_u)^2 + (y_1 - y_u)^2 + (z_1 - z_u)^2} + c \cdot t_u + \varepsilon_1 \\ \sqrt{(x_2 - x_u)^2 + (y_2 - y_u)^2 + (z_2 - z_u)^2} + c \cdot t_u + \varepsilon_2 \\ \sqrt{(x_3 - x_u)^2 + (y_3 - y_u)^2 + (z_3 - z_u)^2} + c \cdot t_u + \varepsilon_3 \\ \vdots \\ \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} + c \cdot t_u + \varepsilon_i \end{cases} \tag{2.38}$$

Where i depends on the number of satellites that have been tracked in the view. Therefore, the GNSS navigation systems can compute a three-dimensional (3D) position only when the GNSS antenna can receive at least four different satellite signals to solve four unknowns, comprising three coordinates of the user position (x_u, y_u, z_u) and one receiver clock offset (t_u). Normally, the GNSS solution can be solved through an iterative least-squares (LS) method. This iterative process only uses information for the instant of time in which data has been received. It does not use any information from points in time before or after. This is the reason to call it a SnapShot method for position estimation. To use the LS method, the nonlinear mathematical model described by Equation 2.38 can be denoted by the following expression:

$$f(X) = l \tag{2.39}$$

where $X = (x_u, y_u, z_u, ct_u)^T$ are the parameters and $l = (\rho_1, \rho_2, \cdots, \rho_i)^T$ are the observations. If $X_0 = (x_0, y_0, z_0, ct_0)^T$ is assumed as the approximate estimated

coordinates for the user and the associated estimate predicted receiver clock offset, then:

$$X = X_0 + \Delta X \tag{2.40}$$

where $\Delta X = (\delta x_u, \delta y_u, \delta z_u, c \cdot \delta t_u)^T$. Therefore,

$$f(X) = f(X_0 + \Delta X) \tag{2.41}$$

The right-hand function can be linearized around the approximate parameters X_0 by using Taylor series, giving:

$$A\Delta X = B + v \tag{2.42}$$

Where

$$A = \begin{bmatrix} \frac{\delta f_1}{\delta x_u} & \frac{\delta f_1}{\delta y_u} & \frac{\delta f_1}{\delta z_u} & \frac{\delta f_1}{\delta t_u} \\ \frac{\delta f_2}{\delta x_u} & \frac{\delta f_2}{\delta y_u} & \frac{\delta f_2}{\delta z_u} & \frac{\delta f_2}{\delta t_u} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\delta f_i}{\delta x_u} & \frac{\delta f_i}{\delta y_u} & \frac{\delta f_i}{\delta z_u} & \frac{\delta f_i}{\delta t_u} \end{bmatrix} \tag{2.43}$$

$$\frac{\delta f_i}{\delta x_u} = \frac{x_i - x_u}{\sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z - z_u)^2}} \tag{2.44}$$

$$\frac{\delta f_i}{\delta y_u} = \frac{y_i - y_u}{\sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z - z_u)^2}} \tag{2.45}$$

$$\frac{\delta f_i}{\delta z_u} = \frac{z_i - z_u}{\sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z - z_u)^2}} \tag{2.46}$$

$$\frac{\delta f_i}{\delta t_u} = 1 \tag{2.47}$$

A is referred to as the design matrix, which contains the direction vectors pointing from the approximate user position to the available satellites, and $B = [\rho_1 - \hat{\rho}_1, \rho_2 - \hat{\rho}_2, \cdots, \rho_i - \hat{\rho}_i]^T$. v represents residual pseudorange errors and is assumed as normally distributed with zero mean and variance $Cov(l)$, which is the variance-covariance matrix of the observations. Generally, the $Cov(l)$ is assumed diagonal, which means the observations are uncorrelated. Therefore, the variance-covariance matrix of the pseudoranges can be shown as:

$$Cov(l) = \begin{bmatrix} \sigma_{\rho_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{\rho_2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{\rho_i}^2 \end{bmatrix} \tag{2.48}$$

A small standard error associated with an observation means that a high weight is assigned to it. The weight matrix would be $W = Cov(l)^{-1}$.

The least-squares method aims to minimize the function $v^T W v$. Taking Equation 2.42 to replace v gives:

$$\begin{aligned} v^T W v &= (A\Delta X - B)^T W (A\Delta X - B) \\ &= (\Delta X^T A^T - B^T)(WA\Delta X - WB) \\ &= \Delta X^T A^T WA\Delta X - B^T WA\Delta X \\ &\quad - \Delta X^T A^T WB + B^T WB \end{aligned} \tag{2.49}$$

The minimum of $v^T Wv$ must occur at a value of ΔX that gives a zero for the gradient. Hence, setting the gradient to be zero and seeking a value that will minimize $v^T Wv$ gives:

$$\frac{\delta v^T Wv}{\delta \Delta X} = 2A^T WA\Delta X - 2A^T WB = 0 \tag{2.50}$$

$$\Delta X = (A^T WA)^{-1} A^T WB \tag{2.51}$$

The improved estimate ΔX should be used to iterate until the change in the estimate is sufficiently small. Once the unknowns ΔX are obtained, the user coordinates and the receiver clock offset can be computed by Equation 2.40.

2.1.6.2 Kalman Filter

The Kalman filter technique for GNSS stand-alone positioning provides increased solution availability concerning snapshot positioning. This is achieved through the use of a prediction component based on previous solutions, allowing a solution to be formed when there are insufficient satellites available for an instantaneous snapshot solution.

The Kalman filter method is presented in this section. The Kalman filter method used in (EATS n.d.) has been taken as reference. This includes an innovation fault detection and exclusion (FDE) technique and a comprehensive integrity algorithm, to provide continuously accurate positioning and high integrity performance. The algorithm is described in paragraph Position Engine (2.1.6.2.1) and Integrity Engine (Subsection **Error! Reference source not found.**).

2.1.6.2.1 Position Engine

The GNSS Kalman filter is used to continuously estimate the user position, velocity, receiver clock offset and clock drift by using GNSS observations only. For the position calculation, the 8 navigation states are considered at every epoch as:

$$X = \left[x, y, z, t_u, v_x, v_y, v_z, t_v\right]^T \tag{2.52}$$

Where

x is the Cartesian x coordinate of the vehicle position, in meters

y is the Cartesian y coordinate of the vehicle position, in meters

z is the Cartesian z coordinate of the vehicle position, in meters

t_u is the receiver clock offset, in seconds

v_x is vehicle velocity in the direction of the Cartesian x-axis, in m/s

v_y is vehicle velocity in the direction of the Cartesian y-axis, in m/s

v_z is vehicle velocity in the direction of the Cartesian z-axis, in m/s

t_v is receiver clock drift, in seconds/second.

The vehicle dynamic model used to update state estimates over Δt (between epochs *t-1* and *t*) may be described at a basic level (prior to the consideration of acceleration) using the following equations:

$$\begin{cases} x_t = x_{t-1} + v_{xt-1} \cdot \Delta t \\ y_t = y_{t-1} + v_{yt-1} \cdot \Delta t \\ z_t = z_{t-1} + v_{zt-1} \cdot \Delta t \\ t = t_{t-1} + t_{vt-1} \cdot \Delta t \\ v_{xt} = v_{xt-1} \\ v_{yt} = v_{yt-1} \\ v_{zt} = v_{zt-1} \\ t_{vt} = t_{vt-1} \end{cases} \tag{2.53}$$

These equations can be rewritten in matrix form as:

$$X_t = \Phi_{t-1} X_{t-1} + w_t, \quad w_t \sim N(0, Q_{w,t}) \tag{2.54}$$

Where,

X_t is the state vector at epoch *t*,

Φ_{t-1} is the state transition matrix which defines the state transition from epoch *t-1* to epoch *t*, given as

$$\Phi_{t-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.55}$$

$Q_{w,t}$ is the state noise covariance matrix which is used to model the acceleration:

$$Q_{w,t} = \begin{bmatrix} \frac{\Delta t^4 \cdot \sigma_a^2}{4} & 0 & 0 & 0 & \frac{\Delta t^3 \cdot \sigma_a^2}{2} & 0 & 0 & 0 \\ 0 & \frac{\Delta t^4 \cdot \sigma_a^2}{4} & 0 & 0 & 0 & \frac{\Delta t^3 \cdot \sigma_a^2}{2} & 0 & 0 \\ 0 & 0 & \frac{\Delta t^4 \cdot \sigma_a^2}{4} & 0 & 0 & 0 & \frac{\Delta t^3 \cdot \sigma_a^2}{2} & 0 \\ 0 & 0 & 0 & \frac{\Delta t^4 \cdot \sigma_t^2}{4} & 0 & 0 & 0 & \frac{\Delta t^3 \cdot \sigma_t^2}{2} \\ \frac{\Delta t^3 \cdot \sigma_a^2}{2} & 0 & 0 & 0 & \Delta t^2 \cdot \sigma_a^2 & 0 & 0 & 0 \\ 0 & \frac{\Delta t^3 \cdot \sigma_a^2}{2} & 0 & 0 & 0 & \Delta t^2 \cdot \sigma_a^2 & 0 & 0 \\ 0 & 0 & \frac{\Delta t^3 \cdot \sigma_a^2}{2} & 0 & 0 & 0 & \Delta t^2 \cdot \sigma_a^2 & 0 \\ 0 & 0 & 0 & \frac{\Delta t^3 \cdot \sigma_t^2}{2} & 0 & 0 & 0 & \Delta t^2 \cdot \sigma_a^2 \end{bmatrix} \quad (2.56)$$

Where,

σ_a is the standard deviation of vehicle acceleration.

σ_t is the standard deviation of the receiver clock.

Therefore, the GNSS Kalman Filter navigation engine is conducted using the following steps:

State vector prediction:

$$\hat{X}_t^- = \Phi_{t-1}\hat{X}_{t-1}^- \quad (2.57)$$

Variance-covariance matrix prediction:

$$P_t^- = \Phi_{t-1}P_{t-1}\Phi_{t-1}^T + Q_{w,t} \quad (2.58)$$

Kalman gain matrix calculation:

$$K_t = P_t^- G_t^T (G_t P_t^- G_t^T + R_t)^{-1} \quad (2.59)$$

State estimate update using measurements:

$$\hat{X}_t = \hat{X}_t^- + K_t\left(Z_t - g(\hat{X}_t^-)\right) \quad (2.60)$$

$$P_t = (I - K_t G_t) \cdot P_t^- \cdot (I - K_t G_t)^T + K_t R_t K_t^T \quad (2.61)$$

Where,

Z_t = vector of pseudorange and Doppler measurements at epoch *t*

g = function for computing measurement vector based on state estimates

G_t = measurement matrix at epoch *t*

R_t = measurement noise covariance matrix at epoch *t* (obtained from the measured noises)

Within the GNSS Kalman Filter navigation engine, measurement innovations may be formed for each ranging source used. They are normalized and simply compared to a threshold in order to detect potential faults. This process is performed before the update of the state estimates using the measurements. Should a normalized innovation exceed the threshold then the relevant measurements (pseudorange and Doppler) are excluded from the solution computation process at the epoch in question.

At epoch *t* the measurement innovation vector is formed as:

$$\delta p_t^- = p_t - g_p(\hat{X}_{\Delta,t}^-) \quad (2.62)$$

Where,

p_t = vector of pseudorange measurements at epoch *t*

g_p = function for computing pseudorange measurement vector based on 4D state estimates (3D position and receiver clock offset)

$\hat{X}_{4,t}^-$ = 4D state estimates at epoch *t* (typically those of epoch *t-1* propagated forward using the transition matrix), prior to update using measurements

These innovations may then be normalized using the innovation covariance matrix. For measurement *i,* this is performed as follows at epoch *t*:

$$y_{t,i}^- = \frac{\delta p_{t,i}^-}{\sqrt{C_{t,i,i}^-}} \tag{2.63}$$

Where,

$$C_t^- = G_{p,t} P_t^- G_{p,t}^T + R_{p,t} \tag{2.64}$$

$G_{p,t}$ = pseudorange (and 4 state) measurement matrix at epoch *t*

P_t^- = error covariance matrix at epoch *t*, before update, using measurements

$R_{p,t}$ = pseudorange measurement noise covariance matrix at epoch *t*

In practice this method may be applied within an iteration loop of a Kalman Filter navigation engine, with the innovations formed at each iteration as follows:

$$\delta p_t^- = \Delta p_t - G_{p,t} \Delta \hat{X}_{4,t}^- \tag{2.65}$$

Δp_t = 'observed minus computed' vector: the difference between pseudorange observations at epoch *t* and computed pseudoranges based on the most recent estimate of the receiver position

$\Delta\hat{X}_{4,t}^{-}$ = difference between prediction (that of epoch *t-1* propagated forward to epoch *t*) and most recent iteration loop estimation of 4D state estimates

As the iteration loop progresses, the components of Δp_t should reduce as the state estimates improve. However, any significant measurement errors will remain apparent and should be easily detectable. $G_{p,t}$ and $\Delta\hat{X}_{4,t}^{-}$ will also be updated upon each iteration, reflecting the new position estimates.

2.1.6.3 Multi-Constellation GNSS Positioning

As already introduced before, all GNSS systems use the TOA concept to derive the receiver position, but each of them has their geodetic reference system and satellite clock system. Therefore, the alignment of coordinate systems and the synchronization of time systems are needed when multi-GNSS are combined.

2.1.6.3.1 Terrestrial Reference System Alignment

The WGS84 (GPS), GTRF (Galileo) and CGCS2000 (BeiDou) are all based on ITRF, differing from ITRF by only a few centimeters. Therefore, they are compatible with most users with this accuracy level.

From 20th September 2007, GLONASS implemented the new version of PZ-90 (i.e. PZ-90.02) as the terrestrial reference system. PZ-90.02 is aligned with ITRF2000, but it contains an origin shift of about 0.4m. Therefore, PZ-90.02 can be transformed to ITRF2000 by the following equation:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{ITRF2000} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{PZ-90.02} + \begin{bmatrix} -0.36 \\ 0.08 \\ 0.18 \end{bmatrix} \quad (2.66)$$

2.1.6.3.2 Time System Synchronization

Synchronization is necessary for multi-constellation GNSS solutions. There are two options to achieve this; synchronization by the information from the navigation message and extra state for inter-system clock bias. Both options are shown here below.

2.1.6.3.3 Synchronization by the information from the navigation message

In the GPS navigation data messages, the GPS-UTC time corrections (i.e. 'GPUT' in navigation header) are included. Similarly, for Galileo, the GST-UTC time corrections (i.e. 'GAUT') will be broadcast in the Galileo navigation messages. The case is the same for GLONASS and BeiDou, with GLONASS-UTC (i.e. 'GLUT') and BDT-UTC (i.e. 'COUT') being found in their navigation messages, respectively. By applying the above corrections, the time systems can be synchronized to the same UTC frame. This is a highly recommended option since there is no need to introduce any extra state in the navigation engine for the combined GNSS system. The availability of a solution is maximized since each extra state removes a degree of freedom. Additionally, if GST-GPS, GLONASS-GPS and BDT-GPS time corrections are available in the various navigation messages, the different GNSS measurements can also be synchronized to GPS time.

2.1.6.3.4 Extra State for Inter-system Clock Bias

If the time corrections from the navigation message are not available, the inter-system clock biases have to be treated as an extra state in the navigation engine to estimate the individual clock bias for each constellation (Liu et al. 2017).

2.1.6.4 Trade-Off Analysis for computation methodology

A choice must be made for the high-level approach to the GNSS PVT algorithm, i.e. whether a snapshot or Kalman filter method will be used. The driver here is integrity as this is the most demanding aspect in railway applications for positioning. ERA Subset 088 (UNISIG 2008) defines the safety requirement as Safety Integrity Level (SIL) 4. A SIL indicates the required degree of confidence that a system will meet its specified safety functions with respect to systematic failures. SIL indicates the probability of failure with the mission time. The THR (Tolerable Hazard Rate THR per hour and per function) applicable to SIL4 is defined as $10^{-9} \leq THR \leq 10^{-8}$.

Standard ERTMS positioning (odometry-based) is designed against this specification and for any alternative source of positioning it may reasonably be assumed that the specification also applies. A tolerable hazard rate of 10^{-9}/hour is two orders of magnitude more demanding than that of GNSS-based aircraft approach procedures which highlights the challenge involved in designing Smart Train Positioning.

Table 2.3 presents the pros and cons of the two approaches with consideration of the application to railway positioning.

	SnapShot	Kalman Filter
PROs	• Simple to implement • Simple to model errors • No initialization issues • No issues regarding time to recover following outliers	• High availability positioning • Stable solution in difficult environments • Well suited to train motion profile • Existing fault detection methods can handle multiple simultaneous faults
CONs	• No positioning when insufficient satellites available • Existing fault detection methods for multiple simultaneous faults are computationally demanding	• More challenging to implement • Requires tuning according to application • Harder to model errors • Solution can take time to initialize • Solution can take time to recover from outliers

Table 2.3 PROs and CONs of GNSS Algorithm options (EATS n.d.)

Figure 2.3 and Figure 2.4 below illustrate positioning results for GPS data recorded on an operational passenger train in the UK using a basic snapshot algorithm (including the RAIM algorithm described in Section Integrity Engine) and the Kalman filter created by NSL (prior to any tuning for railway application) respectively (EATS n.d.). Individual 1Hz solutions are overlaid on aerial photography, shown as red and blue points for the snapshot and Kalman filter respectively. It may be seen that the snapshot algorithm is not able to provide solutions for the stretch of track in the middle of the image which corresponds to the train traveling through a tunnel. There are also two outlier solutions to the right of the image caused by only being four satellites available due to the train remerging from the tunnel and requiring some time to require satellites. RAIM integrity is not available on these occasions as it requires a minimum of five satellites. In contrast, the Kalman filter can continuously provide a solution through the tunnel and does not suffer from the outliers, with the solution remaining close to the track throughout.

This is a good example of where a filtered solution is advantageous over a snapshot solution, and such scenarios are not uncommon in the railway environment.

Figure 2.3 Example SnapShot Positioning fo Real GPS Rail Data (EATS n.d.)

Figure 2.4 Example Kalman Filter Positioning for Real GPS Rail Data (EATS n.d.)

Although in theory, a snapshot approach is more suitable for applications requiring high integrity due to the lack of the potentially unstable prediction component inherent in a Kalman filter, RAIM FDE methods for handling multiple faults are not mature (EATS n.d.). The Kalman filter HPLs are however generally larger and may require further tuning to meet the availability requirement of railway applications.

The Kalman filter also provides a solution with high availability and when properly tuned to the application can provide relatively stable solutions in challenging environments. This should be of particular value in the rail domain since, even though the number of satellites available should be high through the use of four GNSS constellations, there are environments (tunnels, cuttings, covered stations, etc) in which the number of satellites in view will be severely limited. A filter is also well suited to the smooth characteristics of a rail track. Although it is extremely challenging to provide integrity when a small number of satellites are available, there are potential methods to be explored to handle this.

2.2 IMU

An Inertial Measurement Unit (IMU) is a device that measures forces, angular rates and sometimes the orientation of a body. In the case of this thesis, the IMU has three gyroscopes and three accelerometers that are displaced along three mutually orthogonal axes. The main idea of inertial navigation is Newton's First Law: "A body will continue in its state of rest, or of uniform motion in a straight line unless an external force is applied to it". The accelerometers detect the acceleration changes due to the gravity forces and the gyroscopes the changes in the rotational attributes.

Using IMUs has some benefits for positioning; IMUs are autonomous and are not dependent on other devices or visibility of signals. Moreover, it does not need an antenna so it can be located in any place. However, the main problem of the IMUs is their accumulative error. The sources of the errors are mainly two kinds of bias errors and noises:

- Bias errors are constant errors suffered by the sensors in their measurements. Knowing those errors beforehand could be a solution to calibrate the IMU, but there are two types of bias errors, static and dynamic.
    - The static bias is a constant error result of a wrong calibration of the sensors.
    - The dynamic bias is the in-run variation and changes over time. However, the dynamic bias is of about 10% of the static bias, so its influence in the total error is lower. Anyhow, static errors can be compensated as they are always similar, and the dynamic ones are going to be the ones with the biggest influence in the final implementation.
- Noise is the unwanted signal generated from internal electronics that interfere with the measurement of the desired signal. The accelerometer noise will negatively affect the minimum pitch or roll angle resolvable and

introduce velocity and position error caused by the misalignment of the gravity vector. Gyro noise creates orientation angle errors for an INS (Inertial Navigation System) or AHRS (Attitude and Heading Reference System), which affect negatively the projection of the gravity vector and results in velocity and position error (they introduce an error in yaw angles). As almost all of it is random noise (white noise), it will not be possible to predict it in advance. But the positive point is that in general, velocity, position, or pitch or roll error from the accelerometer or gyroscope white noise will be smaller than the other described noise sources for IMU (Raymond Chow 2011).

Dead reckoning is the method used by the Inertial Measurement Units to give a position and velocity estimation on its own. This subsystem can give an accurate trajectory estimation on position, but the information must be accurately known to be reliable, as it suffers from cumulative error. If the information is not accurate enough, the position will neither be accurate, and the next position calculation would be based on an initial wrong position, making that second calculation to have even a higher error.

The position estimation given has an error that will continuously increase due to the accumulation of errors.

2.2.1 Clipping

State estimation is challenging in nonlinear systems based on first principles. The introduction of constraints is a good approach to force the estimates to have physical meaning. This approach is important to delimit the Kalman filter estimations to reasonable limits according to the application in which is going to be used (Kolås, Foss, and Schei 2009).

In this section, constraints in the UKF (Unscented Kalman Filter) are going to be analyzed as it is the most suitable for the railway application. EKF (Extended Kalman Filter) has limitations when using constraints. However, the existing literature

discussing constraint implementation in UKF is limited and discussed clipping techniques are not much. More detailed information about different Kalman filter has been introduced in sections 2.5 and 2.6.

In 1960 Rudolph Kalman presented a recursive state estimation method, the Kalman Filter. A recursive algorithm is used to compute the first and second-order moments (mean and covariance). The system random variables can be consistently estimated by updating the first and second-order moments sequentially by using a linear form estimator (Kolås, Foss, and Schei 2009).

The Kalman filter consists of two parts; forward prediction and correction. The first one computes a predicted estimate of the first (a prior estimation) and second-order moment; and the correction computes the corrected estimates (a posterior estimation). All the existing Kalman Filters follow this structure and need to be initialized with the first and second-order moments before starting the computation.

The UKF (Unscented Kalman Filter) is based on the fact that it is easier to approximate a probability distribution than a nonlinear function (Julier and Uhlmann n.d.). The use of the UKF avoids the Jacobians in the algorithm what could be a computational difficulty. The probability function is approximated by a set of deterministic points that capture the mean and covariance of the distribution, called sigma points. These points are propagated, processing them through the non-linear system model composed by a non-linear function. Some weights are given to these propagated sigma points in order to obtain with them the mean and covariance of the transformed distribution.

In the literature, a symmetric sigma point set is proposed, resulting in 2n+1 sigma points (Julier and Uhlmann n.d.). This approach is precise but demands a big effort in computation as the sigma point generation step is the most computational demanding step in the whole Kalman filter. Calculating the sigma points involves computing the square root of the covariance matrix what may affect the convergence performance.

The estimates in a system should be constrained in order to force them to have a physical meaning according to the field of application. The application of the constraints can be done in one or more of the following processes:

- IMU data recording
- Kalman filter sigma point expansion
- Sigma points binding

2.2.1.1 Constraint methods

A common method for implementing constraints in the Kalman Filter is clipping. The corrected state estimation is set inside some boundaries if it is outside their limits. Other constraint methods such as QP problem solutions can also be considered. However, clipping may produce a lower computational load.

Quadratic Programming (QP) is a special type of mathematical optimization problem that optimizes several variables subject to linear constraints on these variables. It is a type of nonlinear programming.

The general quadratic program can be stated as

$$min_x \mathrm{q(x)} = \frac{1}{2} x^T G x + x^T c \tag{2.67}$$

$$\text{subject to} \quad a_i^T = b_v \quad i \,\epsilon\, \varepsilon \tag{2.68}$$

$$a_i^T \geq b_v \quad i \,\epsilon\, I \tag{2.69}$$

Where:

G is a symmetric nxn matrix,

ε and I are finite sets of indices,

$c, x,$ and $\{a_i\}, i \in \varepsilon \cup I,$ are vectors in $\mathbb{R}^n$.

Quadratic programs can always be solved in a finite amount of computation, but the computation effort required depends on the characteristics of the function being evaluated. If the Hessian matrix G is positive semidefinite, the problem computation requirement is similar to a linear program. The Quadratic Program is called Strictly Convex is this happens. On the other hand, if G is an indefinite matrix, the Quadratic Program is called Nonconvex and the computation load is more challenging due to several stationary points (Nocedal and Wright n.d.).

This procedure results in an improved estimate and the system eventually converges to the true state. Usually, only the estimations are clipped and not the covariance matrix, so the covariance matrix is not precise (Haseltine and Rawlings 2005).

After introducing the clipping strategy, possible limitations for the railway environment are going to be analyzed. An IMU records accelerations, so a limit for reasonable accelerations should be set in order to discard the measurements out of the selected boundaries. For boundary selection, different trains and their maximum accelerations have been examined.

- The first one has been a high-speed train, German ICE. This kind of trains have an average acceleration of 0.3 m/s^2, starting in 0.5 m/s^2 and reducing it 0.1 m/s^2 every 1000m of track. Their maximum deceleration rate is 0.5 m/s^2 (Connor n.d.).
- Regarding to the Intercity Express trains used in the UK, their maximum acceleration is 0.75 m/s^2 at the beginning of the operation and is reduced continuously in a nonlinear way until the maximum speed of the train is reached (Technical 2012).
- Finally, different underground, urban and suburban trains have been analyzed as they are the railway vehicles with the highest acceleration and decelerations. Manchester tram has maximum acceleration and

decelerations of 1.3 m/s^2; Nottingham trams have accelerations of 1.2 m/s^2 and decelerations of 1.4 m/s^2 and the Sheffield Supertram has accelerations of 1.3 m/s^2 and decelerations of 1.5 m/s^2 . This last train is the one with the highest accelerations found in the literature that ensures passenger security and comfort (Powell and Palacín 2015).

Having analyzed different acceleration values in different trains, the maximum accelerations recorded are always less than 1.5 m/s^2 so it seems a good boundary for clipping. Accelerations higher than this recorded by the IMU should be discarded or constrained to the maximum permitted in order to have more accurate results. This approach is valid for the general railway environment, but depending on the train used it should be adapted to obtain better results.

2.3 ULTRA WIDE BAND

The Federal Commission of Communications (FCC) defines Ultra Wide Band (UWB) as any radio technology having at least, a bandwidth of 500MHz. Moreover, technologies with a bandwidth larger than 20% of the central frequency are also considered UBW technologies (FCC 2002). UWB systems employ very low power spectral densities along with very short duration pulses (Duran et al. 2012). The main feature of UWB signals is that they occupy a wider frequency band than conventional signals (ŞahinoĞlu, Gezici, and Güvenç 2008). This is the reason to limit the transmission power below certain values to avoid significant interferences in other systems. The average Power Spectral Density (PSD) must be lower than -41.3 dBm/MHz in the frequency band going from 3.1 GHz to 10.6 GHz in the United States according to the specifications of the Federal Commission of Communications from the beginning. After this regulation appeared, many UWB systems were developed and standardized as the ones introduced in (ECMA 2008) and (Standards Committee of the IEEE Computer Society 2007).

Europe allowed the use of UWB later and under different regulations to the ones ruling in the United States. The bands allowed in Europe are different from the ones in the US and so the spectrum mask also changes. For positioning systems, two bands are relevant; the band going from 3.2 GHz to 4.8 GHz and the band going from 6 GHz to 8.5 GHz. The first band allows maximum Power Spectral Densities of -70 dBm/MHz and the second one allows -41.3 dBm/MHz as maximum just like the US standard (Union n.d.).

UWB signals can be generated in two ways: Single Carrier or Impulse Radio Ultra Wide Band (IR-UWB) and Multi-Band or Multi-Carrier Ultra Wide Band (MC-UWB). The first one uses ultra-short pulses to transmit the information; these pulses are usually shorter than nanoseconds. The second one uses multi-carrier techniques such as Orthogonal Frequency Division Multiplexing (OFDM) based on multiple sinusoids, modulated or not (Zamora Cadenas 2014).

Most of the narrowband systems cannot perceive the UWB signals due to the low energy level in which their signals are transmitted. This makes UWB systems immune to jamming by narrowband systems.

The wide bandwidth of UWB improves the reliability of the information sending process because having different frequency components increases the probabilities of part of the signal avoiding attenuations that do not cover the full spectrum. A wide bandwidth also gives very fine time resolution which allows to distinguish easily direct path signals from reflected signals that have longer flight times and introduce bigger errors when positioning. UWB is a very good tool for indoor positioning and indoor-outdoor transitions or the other way round (Bocquet, Loyez, and Benlarbi-Delaï 2005). This is definitively a good complementary positioning system to GNSS.

UWB systems have been widely used in different works under different conditions and with different results. The positioning system described in (Zwirello et al. 2012) uses UWB based technologies to obtain position estimations with an average accuracy of 2.5cm. In (Zhang et al. 2010) and (Yu et al. 2012) UWB radars have been developed with millimeter accuracies in small measurement areas.

However, not only prototypes or systems for research use are available nowadays. There is a range of different UWB systems available in the market for mass users. Sapphire DART system from Zebra Technologies can work in 2D scenarios within a 200m range and its accuracy is about 30cm according to its datasheet (Zebra n.d.). Ubisense Group sells another UWB system called Dimension4, which claims to have accuracies of about 15cm but operation ranges of 50-100m (Ubisense group n.d.). Finally, Decawave develops another system called EVK1000 that claims to allow long ranges up to 290m and centimeter-level accuracy along with high reliability and low latency. Moreover, it is a low-cost low power system that allows a high tag density enabling lower infrastructure implementation and maintenance costs (Decawave n.d.).

2.4 MAP BASED METHODS

2.4.1 Map-Aiding

Positioning with GNSS in urban areas can have errors of tens of meters due to shadowing, multipath effects or other effects affecting the GNSS signal. Map aiding is a technique that combines traditional GNSS positioning with 3D mapping and GNSS shadow matching. Shadow matching obtains a position comparing the measured signal with a set of positions predicted using 3D mapping. In this case, both pseudorange and signal to noise measurements are used to determine the position (P. Groves et al. 2016).

Over the past years different techniques using 3D mapping have been developed. The simplest form of map aiding is the terrain height aiding (Adjrad and Groves 2017). For most applications the GNSS antenna are found at a certain height above the ground. Using a Digital Elevation Model (DEM) the position can be constrained to a surface. So the position estimation is done into a 2D space and then the height is obtained using the Digital Elevation Model. Removing a dimension from the position solution calculation can also improve the accuracy of the other dimensions. Height aiding can improve not only vertical position components but also horizontal ones in almost a factor of two as the solution estimation is done only for 2 dimensions but using all the available satellites (Adjrad and Groves 2017).

Environmental 3D maps are constructed, usually in urban environments in order to predict which signals are going to be blocked and which ones are going to be visible (LOS and NLOS signals can also be predicted) (Bradbury et al. 2007).

In a first glance taking into account a complete city model when analyzing the availability of the GNSS signals seems to be a great computation effort. However, the computation load can be reduced considering some constraints. For example,

buildings more far away than 300m from the receiver are not relevant or satellites recently used by the user are the first in being analyzed (Wang, Groves, and Ziebart 2012b).

Position is determined comparing the available signals measured with the ones predicted to be available using the 3D map of the environment. This has been experimentally demonstrated in (Ben-Moshe et al. n.d.)(Wang, Groves, and Ziebart 2011)(Suzuki and Kubo 2012)(Wang, Groves, and Ziebart 2012a)(Isaacs et al. 2014)(Wang, Groves, and Ziebart 2015)(Wang n.d.)(Yozevitch and Moshe 2015). Accuracies of around five-ten meters have been achieved in urban areas.

Another way of using 3D mapping is to use the building models to discard the NLOS signals before doing the positioning calculus. Usually, when positioning in urban environments, there is big uncertainty in position estimation. In this areas, the obtained GNSS positions could be used as an area definition for map definition.

Researchers have also used 3D mapping to predict the delay of the NLOS signals received. However, these techniques require a high computation engine unless the area in which it is being used is very small (Suzuki and K. 2013)(Kumar and Petovello 2014)(Hsu, Gu, and Kamijo 2016)(Kumar and Petovello 2016).

In summary, using as many techniques as possible of the ones introduced will be the best solution to obtain an accurate position. Anyhow a trade-off between position accuracy and computing load should be taken into account.

- 3D mapping is a technique useful for single epoch positioning, as for multiple epoch positioning exist many other GNSS augmentation methods that are more accurate and easier to implement (P. D. Groves 2008).

- 3D mapping is intended to be important in multi-epoch positioning such as navigation as it will enable position to be accurately initialized and re-initialized in challenging urban environments (P. Groves et al. 2016).

3D mapping has a great potential to provide accurate positioning in dense urban areas. However, it is not sufficiently reliable and efficient for deployment in commercial products yet. For that it should work in a big range of different environments, provide quality metrics of the data given, has to be able to work in different user equipment with minor changes, be efficient in processing and to have an easy access to 3D maps already constructed (P. D. Groves et al. 2015).

2.4.2 Map-Matching

Map matching enhances the availability for location in railway applications. Applied at a centralized location, it can improve the consistency of the reported locations (Hutchinson 2016). The basic idea of map-matching is to compare the output of a positioning module with the shape of the nearest environment in order to obtain a better position estimation. Two ways of map-matching are used nowadays; Semi-deterministic and Probabilistic.

- Semi-Deterministic map-matching: It assumes that the vehicle is travelling in a road (a railway track when using it for trains). The simplest semi deterministic map-matching is to move the estimated position to the nearest point in the road or railway.
- Probabilistic map-matching: It improves the performance of the semi deterministic ones by incorporating the accuracy of the position estimated to the map information available. Using the error models of the sensors used for positioning, confidence regions are defined for the estimation given. The regions appearing over roads or railways are more likely to be the places in which the receptor is actually located.

2.5 IMU AND GNSS DATA FUSION

Data fusion between IMU and GNSS can be done in many ways. In this section, the most common ones are going to be introduced.

2.5.1 Loosely coupled integration

Each system (GNSS and IMU systems) obtains a position velocity and attitude estimation independently. The information obtained is fused using an optimal estimator in order to obtain a third solution.

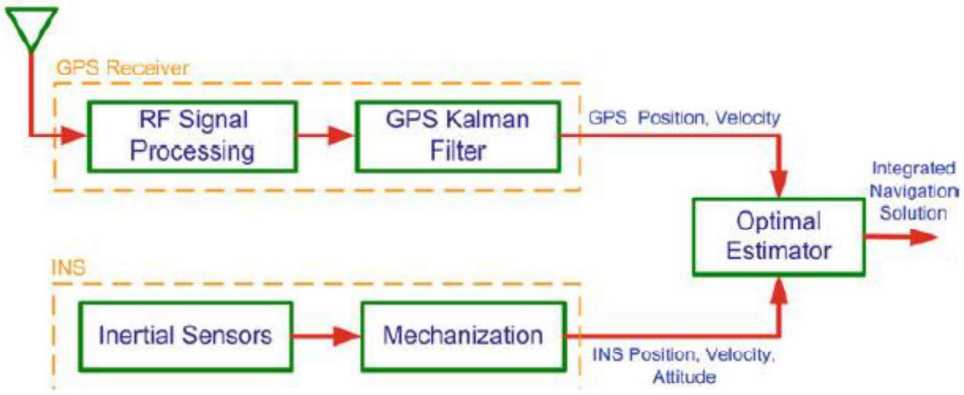


Figure 2.5 Loosely coupled integration scheme (Noureldin, Karamat, and Georgy 2013)

2.5.2 Tightly coupled integration

Pseudo-range and pseudo-range rate from GNSS and accelerations and rotation rates from the IMU are mixed by a single estimator in order to achieve a single position estimation.

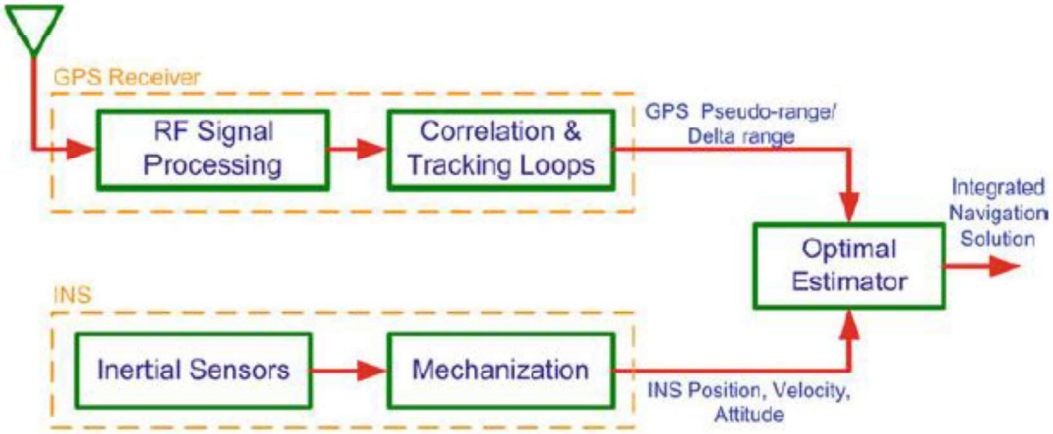


Figure 2.6 Tightly coupled integration scheme (Noureldin, Karamat, and Georgy 2013)

2.5.3 Deeply coupled integration

The main advantage of this method, also known as ultra-tightly coupled integration, is that the dynamics of the host vehicle are compensated in the GNSS tracking loops using the Doppler information. Various configurations for ultra-tightly coupled integration exist. The estimator combines the pseudo-ranges or the In-phase and Quadrature measurements from the GNSS with the navigation parameters obtained from the IMU in order to render the estimated Doppler. This Doppler is used to remove the dynamics from the GNSS signal entering the tracking loops, reducing in this way the carrier tracking loop bandwidth. This integration is more complex and requires access to the GNSS hardware, but it can improve the quality of the raw measurements and the anti-jamming performance comparing with the integrations presented before (Noureldin, Karamat, and Georgy 2013).

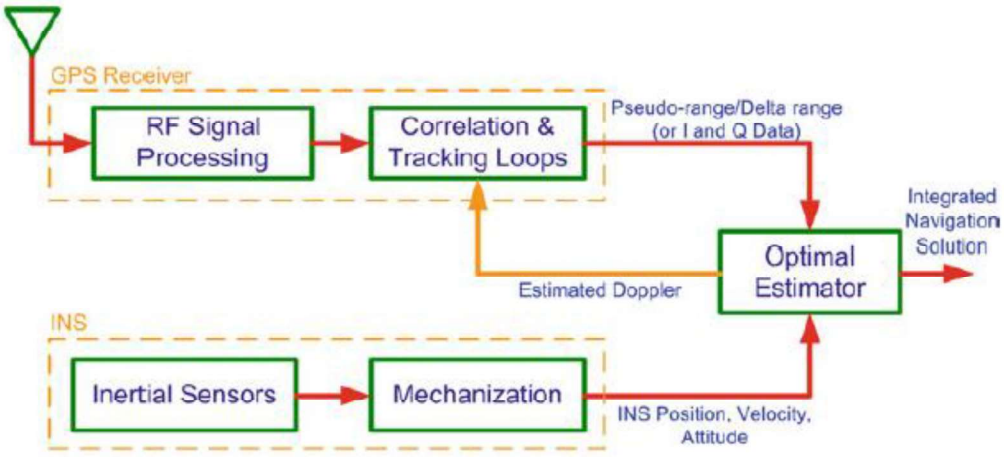


Figure 2.7 Deeply coupled integration (Noureldin, Karamat, and Georgy 2013)

2.5.4 IMU and GNSS Fusion Algorithm

Several algorithms for optimal fusion between IMU and GNSS data exist; the most used ones are the Kalman filter (KF), the particle filter (PF) and the artificial intelligence (AI). Traditionally, the Kalman filter has been used to fuse navigational data (Faragher 2012). Kalman Filter is an optimal recursive algorithm that processes all the available measurements to optimally estimate the current value of the state of interest and obtain the uncertainty of the estimation done.

2.5.5 IMU/GNSS integration

2.5.5.1 Open-loop architecture

In this configuration, position, velocity, and attitude are performed external to the IMU where the estimation errors are subtracted from the IMU solution in each iteration. The errors are not sent back to the IMU. The advantage of this architecture is that the raw IMU solution is capable to support integrity monitoring and continuing service in the event of a problem with the Kalman Filter. However, the errors in the IMU grow larger with time (Section 2.2), making the linearity assumption invalid. So, open-loop architecture is more susceptible to Kalman Filter performance issues.

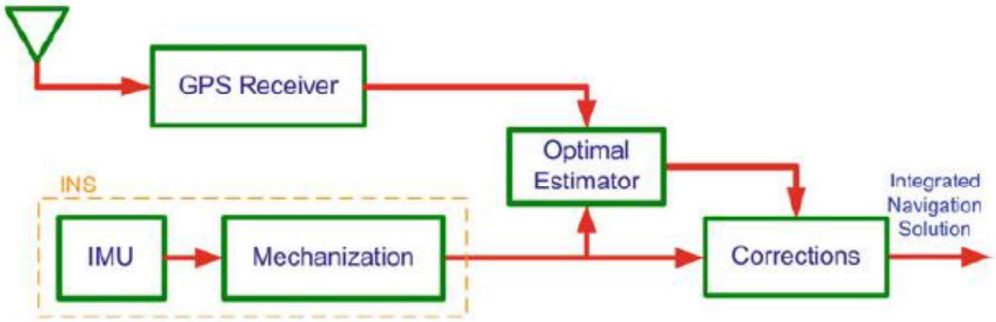


Figure 2.8 Open-loop architecture (Noureldin, Karamat, and Georgy 2013)

2.5.5.2 Closed-loop architecture

In this architecture, based on a Loosely coupled integration, the error estimates from the Kalman filter are fed back in order to correct the IMU eliminating the biases and giving a reference position to the mechanization. The output of the IMU would be the output of the integrated solution. Kalman Filter is reset to zero after the error estimates are fed back. Kalman Filter estimated accelerometer and gyroscope errors

are also fed back to correct the IMU measurements before they are used in mechanization. The errors are applied in every iteration of mechanization and acceleration and gyroscope errors are updated periodically.

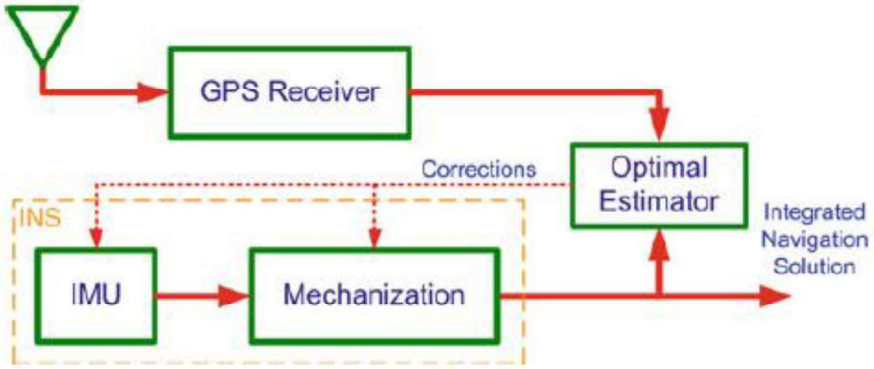


Figure 2.9 Closed-loop architecture (Noureldin, Karamat, and Georgy 2013)

2.5.5.3 Tightly coupled IMU/GNSS Integration

This architecture is based on the tightly coupled integration and has only a common master filter. The difference between the pseudo-range measurements of the GNSS and the corresponding values predicted by the IMU is fed to the Kalman Filter to estimate the errors made by the IMU. The output of the IMU is then corrected with these errors to obtain the integrated navigation solution. Pseudo-range or pseudo-range rate measurements can be used, but usually, both are used as they are complementary to each other. This architecture eliminates the problem of the correlated measurements that appears in loosely coupled approaches due to the cascaded Kalman Filters. This integration can give a GNSS update even if fewer than four satellites are visible. However, it is more difficult to implement as the algorithm implies processing raw GNSS data and there is no standalone GPS solution. Using the same hardware a tightly coupled integration almost always performs better than a loosely coupled one (Petovello n.d.).

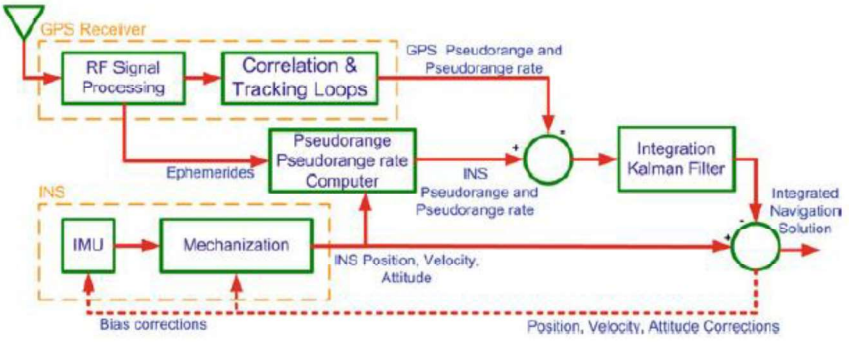


Figure 2.10 Tightly coupled IMU/GNSS integration (Noureldin, Karamat, and Georgy 2013)

The overall implementation of a tightly coupled IMU/GNSS integration is shown in the picture below (Figure 2.11).

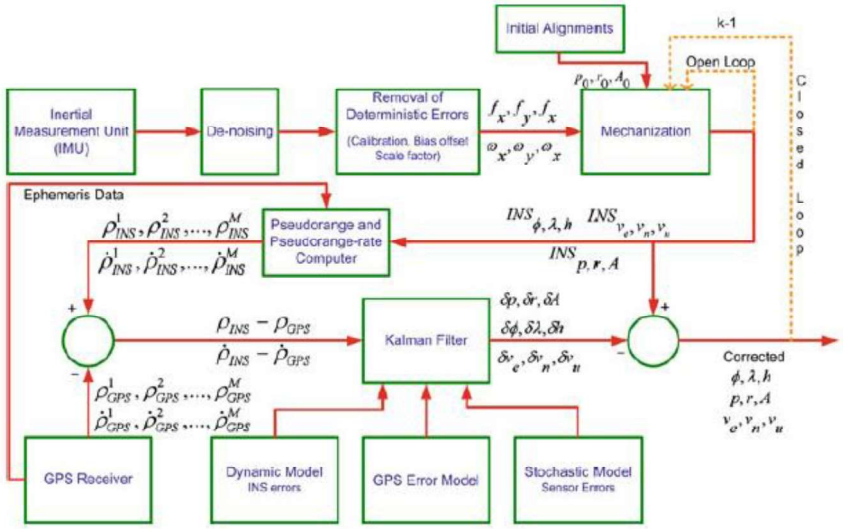


Figure 2.11 Overall implementation of a tightly coupled IMU/GNSS integration (Noureldin, Karamat, and Georgy 2013)

Now, every block shown in Figure 2.11 is going to be explained in order to introduce how to perform a complete integration.

2.5.5.3.1 IMU part

The system model for continuous INS KF is

$$\delta\dot{x} = F_l \delta x_l + G_l w_l \tag{2.70}$$

Where,

x_l the state vector of the IMU;

F_l dynamic coefficient matrix of the IMU model;

G_l noise distribution matrix;

w_l white noise.

$$\delta x^l_{15x1} = [\delta r^l_{3x1}, \delta v^l_{3x1}, \varepsilon^l_{3x1}, \delta\omega_{3x1}, \delta f_{3x1}]^T \tag{2.71}$$

Where,

$\delta r^l = [\delta\varphi, \delta\lambda, \delta h]^T$	is the position error vector
$\delta v^l = \left[\delta v_{e,} \delta v_{n,} \delta v_u\right]^T$	is the Earth-referenced velocity error vector
$\varepsilon^l = [\delta p, \delta r, \delta A]^T$	is the attitude error vector
$\delta\omega = \left[\delta\omega_x, \delta\omega_y, \delta\omega_z\right]^T$	is the gyroscope error vector (consisting of drifts)
$\delta f = \left[\delta f_x, \delta f_y, \delta f_z\right]^T$	is accelerometer error vector (consisting of biases)

w is the unit-variance white Gaussian noise,
the term G is the noise distribution vector (including the variances associated with the state vector)

$$\mathrm{G} = \left[\sigma_{r_{1x3}}, \sigma_{v_{1x3}}, \sigma_{\varepsilon_{1x3}}, \sigma_{\omega_{1x3}}, \sigma_{f_{1x3}}\right] \tag{2.72}$$

The term F is the dynamic coefficient matrix that contains the IMU error models for the position, velocity and attitude and the inertial sensors (Noureldin, Karamat, and Georgy 2013).

$$\mathrm{F} = \begin{bmatrix} 0_{3x3} & F_r & 0_{3x3} & 0_{3x3} & 0_{3x3} \\ 0_{3x3} & 0_{3x3} & F_v & 0_{3x3} & R_b^l \\ 0_{3x3} & F_\varepsilon & 0_{3x3} & R_b^l & 0_{3x3} \\ 0_{3x3} & 0_{3x3} & 0_{3x3} & F_\omega & 0_{3x3} \\ 0_{3x3} & 0_{3x3} & 0_{3x3} & 0_{3x3} & F_f \end{bmatrix} \tag{2.73}$$

2.5.5.3.2 GNSS part

The equation for the Kalman Filter system model for the GNSS is:

$$\delta\dot{x}_G = F_G \delta x_G + G_G w_G \tag{2.74}$$

It can also be written as:

$$\begin{bmatrix} \delta\dot{b}_r \\ \delta\dot{d}_r \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta b_r \\ \delta d_r \end{bmatrix} + \begin{bmatrix} \sigma_b \\ \sigma_d \end{bmatrix} w_G \tag{2.75}$$

Where, δb_r is the bias of the GNSS receiver clock and δd_r is its drift. σ_b is the standard deviation of the white noise for the clock bias and σ_d is the standard deviation of the noise for the clock drift.

Combining IMU and GNSS system models:

$$\delta\dot{x} = F\delta x + Gw$$

$$\begin{bmatrix}\delta\dot{x}_I \\ \delta\dot{x}_G\end{bmatrix} = \begin{bmatrix}F_I & 0 \\ 0 & F_G\end{bmatrix}\begin{bmatrix}\delta x_I \\ \delta x_G\end{bmatrix} + \begin{bmatrix}G_I \\ G_G\end{bmatrix} w \tag{2.76}$$

Inserting the parameters from the IMU and the GNSS we obtain:

$$\begin{bmatrix}\delta\dot{r}^I_{3x1} \\ \delta\dot{v}^I_{3x1} \\ \dot{\varepsilon}^I_{3x1} \\ \delta\dot{\omega}_{3x1} \\ \delta\dot{f}^I_{3x1} \\ \delta\dot{b}_r \\ \delta\dot{d}_r\end{bmatrix} = \begin{bmatrix}I_{3x3} & F_r & 0_{3x3} & 0_{3x3} & 0_{3x3} & 0 & 0 \\ 0_{3x3} & I_{3x3} & F_v & 0_{3x3} & R^I_b & 0 & 0 \\ 0_{3x3} & F_\varepsilon & I_{3x3} & R^I_b & 0_{3x3} & 0 & 0 \\ 0_{3x3} & 0_{3x3} & 0_{3x3} & I_{3x3} + F_\omega & 0_{3x3} & 0 & 0 \\ 0_{3x3} & 0_{3x3} & 0_{3x3} & 0_{3x3} & I_{3x3} + F_f & 0 & 0 \\ 0_{1x3} & 0_{1x3} & 0_{1x3} & 0_{1x3} & 0_{1x3} & 0 & 1 \\ 0_{1x3} & 0_{1x3} & 0_{1x3} & 0_{1x3} & 0_{1x3} & 0 & 0\end{bmatrix}\begin{bmatrix}\delta r^I_{3x1} \\ \delta v^I_{3x1} \\ \varepsilon^I_{3x1} \\ \delta\omega_{3x1} \\ \delta f_{3x1} \\ \delta b_r \\ \delta d_r\end{bmatrix} + \begin{bmatrix}\sigma_{r_{1x3}} \\ \sigma_{v_{1x3}} \\ \sigma_{\varepsilon_{1x3}} \\ \sigma_{\omega_{1x3}} \\ \sigma_{f_{1x3}} \\ \sigma_b \\ \sigma_d\end{bmatrix} w \tag{2.77}$$

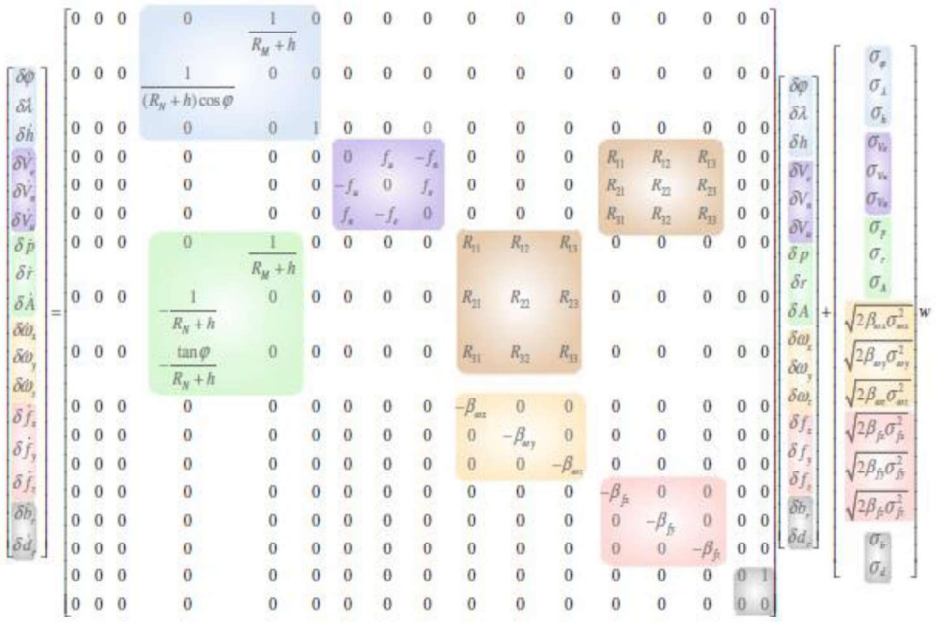
The discrete equation can also be written as:

$$\delta x_k = (I + F\Delta t)\delta x_{k-1} + G\Delta t w_{k-1} \tag{2.78}$$

And the expanded one in discrete form as:

$$\begin{bmatrix} \delta r^I_{3x1} \\ \delta v^I_{3x1} \\ \varepsilon^I_{3x1} \\ \delta \omega_{3x1} \\ \delta f_{3x1} \\ \delta b_r \\ \delta d_r \end{bmatrix}_k = \begin{bmatrix} I_{3x3} & F_r \Delta t & 0_{3x3} & 0_{3x3} & 0_{3x3} & 0 & 0 \\ 0_{3x3} & I_{3x3} & F_v \Delta t & 0_{3x3} & R^I_b \Delta t & 0 & 0 \\ 0_{3x3} & F_\varepsilon \Delta t & I_{3x3} & R^I_b \Delta t & 0_{3x3} & 0 & 0 \\ 0_{3x3} & 0_{3x3} & 0_{3x3} & I_{3x3} + F_\omega \Delta t & 0_{3x3} & 0 & 0 \\ 0_{3x3} & 0_{3x3} & 0_{3x3} & 0_{3x3} & I_{3x3} + F_\omega \Delta t & 0 & 0 \\ 0_{1x3} & 0_{1x3} & 0_{1x3} & 0_{1x3} & 0_{1x3} & 1 & \Delta t \\ 0_{1x3} & 0_{1x3} & 0_{1x3} & 0_{1x3} & 0_{1x3} & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \delta r^I_{3x1} \\ \delta v^I_{3x1} \\ \varepsilon^I_{3x1} \\ \delta \omega_{3x1} \\ \delta f_{3x1} \\ \delta b_r \\ \delta d_r \end{bmatrix}_{k-1} + \begin{bmatrix} \sigma_{r_{1x3}} \\ \sigma_{v_{1x3}} \\ \sigma_{\varepsilon_{1x3}} \\ \sigma_{\omega_{1x3}} \\ \sigma_{f_{1x3}} \\ \sigma_b \\ \sigma_d \end{bmatrix} w_{k-1} \tag{2.79}$$

The completely expanded equation can be seen in the following equation (Noureldin, Karamat, and Georgy 2013):



(2.80)

2.5.5.4 Measurement model

The measurement model in the discrete-time domain is expressed as:

$$\delta z_k = H_k \delta x_k + \eta_k \tag{2.81}$$

Where,

η_k is a vector of measurement noise which is zero-mean with covariance R_k

H_k is the measurement design matrix and describes the linear combinations of state variables that comprise z_k in the absence of noise.

In tightly coupled implementation the available observations are pseudo-ranges and pseudo-range rate measurements. The measurement vector is composed of the difference between these two values predicted by the IMU and the GNSS

$$\delta z = \begin{bmatrix} \delta z_\rho \\ \delta z_{\dot{\rho}} \end{bmatrix} = \begin{bmatrix} \rho_{INS} - \rho_{GPS} \\ \rho_{INS} - \dot{\rho}_{GPS} \end{bmatrix} \tag{2.82}$$

For M satellites:

$$\begin{bmatrix} \delta z_\rho^1 \\ \delta z_\rho^2 \\ \vdots \\ \delta z_\rho^M \\ \delta z_{\dot{\rho}}^1 \\ \delta z_{\dot{\rho}}^2 \\ \vdots \\ \delta z_{\dot{\rho}}^M \end{bmatrix} = \begin{bmatrix} \rho_{INS}^1 - \rho_{GPS}^1 \\ \rho_{INS}^2 - \rho_{GPS}^2 \\ \vdots \\ \rho_{INS}^M - \rho_{GPS}^M \\ \dot{\rho}_{INS}^1 - \dot{\rho}_{GPS}^1 \\ \dot{\rho}_{INS}^2 - \dot{\rho}_{GPS}^2 \\ \vdots \\ \dot{\rho}_{INS}^M - \dot{\rho}_{GPS}^M \end{bmatrix} \tag{2.83}$$

2.5.5.4.1 Pseudo-range measurements:

The pseudo-range measurement vector to implement an error state Kalman Filter has the following form:

$$\delta z_\rho = \rho_{INS} - \rho_{GPS} \tag{2.84}$$

The pseudo-range for the m^{th} satellite to the GNSS receiver can be modeled in the following way:

$$\rho_{GPS}^m = \mathrm{r}^m + c\delta t_r - c\delta t_s + cI^m + cT^m + \tilde{\varepsilon}_\rho^m \tag{2.85}$$

Where

ρ_{GPS}^m is the measured pseudo-range from the m^{th} satellite to the GNSS receiver (meters)

r^m is the actual distance between the receiver antenna at the reception time t_r and the satellite's antenna at the transmit time t^t (meters)

δt_r is the receiver's clock offset (sec)

c is the speed of light (meters/sec)

δt_s is the satellite's clock offset (sec)

I^m is the ionospheric delay (sec)

T^m is the tropospheric delay (sec)

ε_ρ^m is the error due to inexact modelling, receiver noise and multipath

Satellite clock bias and ionospheric errors can be calculated from the satellite's navigation message, and tropospheric can also be estimated. After correcting all errors except receiver errors (noise and clock bias), the corrected pseudo-range is:

$$\rho_{GPS}^m = \mathrm{r}^m + c\delta t_r + \tilde{\varepsilon}_\rho^m \tag{2.86}$$

Where the last term is the total effect of residual errors. The true geometric range from the m^{th} satellite to the receiver is

$$r^m = \sqrt{(x - x^m)^2 + (y - y^m)^2 + (z - z^m)^2} = \|x - x^m\| \tag{2.87}$$

Where

$x = [x, y, z]^T$ is the true receiver position in the e-frame

$x^m = [x^m, y^m, z^m]^T$ is the m^{th} satellite's position in the e-frame

So the equation (2.86) above can be now written as:

$$\rho_{GPS}^m = \sqrt{(x - x^m)^2 + (y - y^m)^2 + (z - z^m)^2} + \delta b_r + \tilde{\varepsilon}_\rho^m \tag{2.88}$$

Where $\delta b_r = c\delta t_r$ is the error in range (meters) due to the receiver's clock bias. The corrected position of the receiver is defined as,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_{INS} - \delta x \\ y_{INS} - \delta y \\ z_{INS} - \delta z \end{bmatrix} \tag{2.89}$$

Where

$x_{INS} = [x_{INS}, y_{INS,} z_{INS}]^T$ is the output of the mechanization

$\delta x = [\delta x, \delta y, \delta z]^T$ is the estimated position error

The pseudo-range equation is not linear, so it must be linearized to use the Kalman Filter by applying a Taylor series aroundx_{INS}. So linearizing the equation (2.88), we obtain the following expression.

$$\begin{aligned}&\rho_{GPS}^m \\ &= \sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2} \\ &+ \frac{(x_{INS} - x^m)(x - x_{INS}) + (y_{INS} - y^m)(y - y_{INS}) + (z_{INS} - z^m)(z - z_{INS})}{\sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2}} \\ &+ \delta b_r + \tilde{\varepsilon}_\rho^m\end{aligned} \tag{2.90}$$

By defining the pseudo-range for the output of the IMU to be,

$$\rho_{INS}^m = \sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2} \tag{2.91}$$

We obtain,

$$\begin{aligned}&\rho_{INS}^m - \rho_{GPS}^m \\ &= \frac{(x_{INS} - x^m)(x - x_{INS}) + (y_{INS} - y^m)(y - y_{INS}) + (z_{INS} - z^m)(z - z_{INS})}{\sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2}} \\ &- \delta b_r + \tilde{\varepsilon}_\rho^m\end{aligned} \tag{2.92}$$

$$1_{INS}^{m} = \begin{bmatrix} 1_{x,INS}^{m} \\ 1_{y,INS}^{m} \\ 1_{z,INS}^{m} \end{bmatrix} = \begin{bmatrix} \dfrac{(x_{INS} - x^{m})}{\sqrt{(x_{INS} - x^{m})^2 + (y_{INS} - y^{m})^2 + (z_{INS} - z^{m})^2}} \\ \dfrac{(y_{INS} - y^{m})}{\sqrt{(x_{INS} - x^{m})^2 + (y_{INS} - y^{m})^2 + (z_{INS} - z^{m})^2}} \\ \dfrac{(z_{INS} - z^{m})}{\sqrt{(x_{INS} - x^{m})^2 + (y_{INS} - y^{m})^2 + (z_{INS} - z^{m})^2}} \end{bmatrix} \tag{2.93}$$

Is the line of sight unit vector from the m^{th} satellite to the receiver's position based upon the output of mechanization. And therefore,

$$\rho_{INS}^{m} - \rho_{GPS}^{m} = 1_{x,INS}^{m}\delta x + 1_{y,INS}^{m}\delta y + 1_{z,INS}^{m}\delta z - \delta b_r + \tilde{\varepsilon}_{\rho}^{m} \tag{2.94}$$

With

$$\begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} = \begin{bmatrix} x_{INS} - x \\ y_{INS} - y \\ z_{INS} - z \end{bmatrix}$$
$$\delta z_{\rho}^{m} = \begin{bmatrix} 1_{x,INS}^{m} & 1_{y,INS}^{m} & 1_{z,INS}^{m} \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} - \delta b_r + \tilde{\varepsilon}_{\rho}^{m} \tag{2.95}$$

When M satellites are visible, the equation (2.95) can be expressed as

$$\delta z_{\rho} = \begin{bmatrix} \rho_{INS}^{1} - \rho_{GPS}^{1} \\ \rho_{INS}^{2} - \rho_{GPS}^{2} \\ \vdots \\ \rho_{INS}^{M} - \rho_{GPS}^{M} \end{bmatrix}$$
$$= \begin{bmatrix} 1_{x,INS}^{1} & 1_{y,INS}^{1} & 1_{z,INS}^{1} \\ 1_{x,INS}^{2} & 1_{y,INS}^{2} & 1_{z,INS}^{2} \\ \vdots & \vdots & \vdots \\ 1_{x,INS}^{M} & 1_{y,INS}^{M} & 1_{z,INS}^{M} \end{bmatrix}_{Mx3} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}_{3x1} - \begin{bmatrix} \delta b_r \\ \delta b_r \\ \vdots \\ \delta b_r \end{bmatrix}_{Mx1} \tag{2.96}$$
$$+ \begin{bmatrix} \tilde{\varepsilon}_{\rho}^{1} \\ \tilde{\varepsilon}_{\rho}^{2} \\ \vdots \\ \tilde{\varepsilon}_{\rho}^{m} \end{bmatrix}_{Mx1}$$

By defining a geometry matrix Y, we get

$$\delta z_\rho = Y_{Mx3}\begin{bmatrix}\delta x\\ \delta y\\ \delta z\end{bmatrix}_{3x1} - \delta b_{r\,Mx1} + \tilde{\varepsilon}_{\rho,Mx1} \tag{2.97}$$

The position in the equation (2.97) above is in ECEF geodetic coordinates, to be able to use it

$$\begin{bmatrix}\delta x\\ \delta y\\ \delta z\end{bmatrix} = \begin{bmatrix} -(R_N+h)\sin\varphi\cos\lambda & -(R_N+h)cos\varphi\sin\lambda & \cos\varphi\cos\lambda\\ -(R_N+h)\sin\varphi\sin\lambda & (R_N+h)\cos\varphi\cos\lambda & \cos\varphi\sin\lambda\\ \{R_N(1-e^2)+h)\}\cos\varphi & 0 & \sin\varphi\end{bmatrix}\begin{bmatrix}\delta\varphi\\ \delta\lambda\\ \delta h\end{bmatrix} \tag{2.98}$$

$$\delta z_\rho = Y_{Mx3}T_{3x3}\begin{bmatrix}\delta x\\ \delta y\\ \delta z\end{bmatrix}_{3x1} - \delta b_{r\,Mx1} + \tilde{\varepsilon}_{\rho,Mx1} \tag{2.99}$$

Where,

$$H^{\rho}_{M\times 3} = G_{M\times 3}T_{3\times 3} \tag{2.100}$$

Pseudo-range rate measurements:

The measurement vector for the pseudo-range rate is the difference between the pseudo-range rate predicted by the INS and the value measured by the GPS.

$$\delta z_{\dot\rho} = \dot\rho_{INS} - \dot\rho_{GPS} \tag{2.101}$$

The Doppler shift produced by satellite and receiver motion is the projection of the relative velocities onto the line of sight, scaled by the transmission frequency and divided by speed of light

$$D^m = \frac{[(v^m - v)\cdot 1^m]L_1}{c} \tag{2.102}$$

Where,

$v^m = \left[v_x^m, v_y^m, v_z^m\right]^T$	is the velocity of the m^{th} satellite in the e-frame
$v = \left[v_x, v_y, v_z\right]^T$	is the true receiver velocity in the e-frame
L_1	is the satellite transmission frequency
c	speed of light
1^m	line of sight unit vector from the m^{th} satellite to the GNSS receiver

Pseudo-range can be computed as,

$$\dot{\rho}^m = -\frac{D^m c}{L_1} \tag{2.103}$$

The pseudo-range can be modeled as

$$\begin{aligned}\dot{\rho}_{GPS}^m &= 1_x^m \cdot (v_x - v_x^m) + 1_y^m \cdot \left(v_y - v_y^m\right) + 1_z^m \cdot (v_z - v_z^m) + c\delta\dot{t}_u + \varepsilon_{\dot{\rho}}^m \\ \dot{\rho}_{GPS}^m &= 1_x^m \cdot (v_x - v_x^m) + 1_y^m \cdot \left(v_y - v_y^m\right) + 1_z^m \cdot (v_z - v_z^m) + \delta d_r + \varepsilon_{\dot{\rho}}^m\end{aligned} \tag{2.104}$$

Where δd_r is the receiver's clock drift in meters/sec and $\varepsilon_{\dot{\rho}}^m$ is the error in observation. The IMU estimated pseudo-range is:

$$\dot{\rho}_{INS}^m = 1_{x,INS}^m \cdot \left(v_{x,INS} - v_x^m\right) + 1_{y,INS}^m \cdot \left(v_{y,INS} - v_y^m\right) + 1_{z,INS}^m \cdot \left(v_{z,INS} - v_z^m\right) \tag{2.105}$$

Where $v_{x,INS}$, $v_{y,INS}$, $v_{z,INS}$ are the receiver's velocities in the e-frame estimated by the IMU. Taking the difference of the IMU-estimated pseudo-range rate and GNSS-measured pseudo-range rate, we get:

$$\dot{\rho}^m_{INS} - \dot{\rho}^m_{GPS} = -1^m_{x,INS} \cdot \left(v_x - v_{x,INS}\right) - 1^m_{y,INS} \cdot \left(v_y - v_{y,INS}\right) - 1^m_{z,INS} \cdot \left(v_z - v_{z,INS}\right) - \delta d_r + \varepsilon^m_{\dot{\rho}} \tag{2.106}$$

$$\dot{\rho}^m_{INS} - \dot{\rho}^m_{GPS} = 1^m_{x,INS} \cdot \left(v_{x,INS} - v_x\right) + 1^m_{y,INS} \cdot \left(v_{y,INS} - v_y\right) + 1^m_{z,INS} \cdot \left(v_{z,INS} - v_z\right) - \delta d_r + \varepsilon^m_{\dot{\rho}} \tag{2.107}$$

$$\delta z^m_{\dot{\rho}} = 1^m_{x,INS} \cdot \delta v_x + 1^m_{y,INS} \cdot \delta v_y + 1^m_{z,INS} \cdot \delta v_z - \delta d_r + \varepsilon^m_{\dot{\rho}} \tag{2.108}$$

Where

$$\delta z^m_{\dot{\rho}} = \dot{\rho}^m_{INS} - \dot{\rho}^m_{GPS} \text{ and } \begin{bmatrix} \delta v_x \\ \delta v_y \\ \delta v_z \end{bmatrix} = \begin{bmatrix} v_{x,INS} - v_x \\ v_{y,INS} - v_y \\ v_{z,INS} - v_z \end{bmatrix}$$

Changing this equation to state-space form,

$$\delta z^m_{\dot{\rho}} = \begin{bmatrix} 1^m_{x,INS} & 1^m_{y,INS} & 1^m_{z,INS} \end{bmatrix} \begin{bmatrix} \delta v_x \\ \delta v_y \\ \delta v_z \end{bmatrix} - \delta d_r + \varepsilon^m_{\dot{\rho}} \tag{2.109}$$

Making it general for M satellites the pseudo-range rate measurement model is

$$\begin{aligned} \delta z_{\dot{\rho}} &= \begin{bmatrix} \dot{\rho}^1_{INS} - \dot{\rho}^1_{GPS} \\ \dot{\rho}^2_{INS} - \dot{\rho}^2_{GPS} \\ \vdots \\ \dot{\rho}^M_{INS} - \dot{\rho}^M_{GPS} \end{bmatrix} \\ &= \begin{bmatrix} 1^1_{x,INS} & 1^1_{y,INS} & 1^1_{z,INS} \\ 1^2_{x,INS} & 1^2_{y,INS} & 1^2_{z,INS} \\ \vdots & \vdots & \vdots \\ 1^M_{x,INS} & 1^M_{y,INS} & 1^M_{z,INS} \end{bmatrix}_{Mx3} \begin{bmatrix} \delta v_x \\ \delta v_y \\ \delta v_z \end{bmatrix} - \begin{bmatrix} \delta d_r \\ \delta d_r \\ \vdots \\ \delta d_r \end{bmatrix}_{Mx1} \\ &+ \begin{bmatrix} \varepsilon^1_{\dot{\rho}} \\ \varepsilon^2_{\dot{\rho}} \\ \vdots \\ \varepsilon^M_{\dot{\rho}} \end{bmatrix}_{Mx1} \end{aligned} \tag{2.110}$$

$$\delta z_{\dot{\rho}} = G_{Mx3} \begin{bmatrix} \delta v_x \\ \delta v_y \\ \delta v_z \end{bmatrix}_{Mx1} - \delta d_{r,Mx1} + \varepsilon^m_{\rho,Mx1} \tag{2.111}$$

The relationship between velocity in l-frame and e-frame is

$$\begin{bmatrix}\delta v_x\\ \delta v_y\\ \delta v_z\end{bmatrix}=\begin{bmatrix}-\sin\lambda & \sin\varphi\cos\lambda & \cos\varphi\cos\lambda\\ \cos\lambda & \sin\varphi\sin\lambda & \cos\varphi\sin\lambda\\ 0 & \cos\varphi & \sin\varphi\end{bmatrix}\begin{bmatrix}\delta v_e\\ \delta v_n\\ \delta v_u\end{bmatrix} \tag{2.112}$$

Where the matrix can be written as R_l^e. And the pseudo-range rate measurement model becomes

$$\delta z_{\dot{\rho}}=G_{Mx3}R_l^e\begin{bmatrix}\delta v_E\\ \delta v_N\\ \delta v_U\end{bmatrix}-\delta d_{r,Mx1}+\varepsilon^m_{\dot{\rho},Mx1} \tag{2.113}$$

2.5.5.5 Overall measurement model

The measurement model for pseudo-range errors and pseudo-range rate errors can be combined to create an overall measurement model as

$$\begin{bmatrix}\rho^1_{INS}-\rho^1_{GPS}\\ \rho^2_{INS}-\rho^2_{GPS}\\ \vdots\\ \rho^M_{INS}-\rho^M_{GPS}\\ \dot{\rho}^1_{INS}-\dot{\rho}^1_{GPS}\\ \dot{\rho}^2_{INS}-\dot{\rho}^2_{GPS}\\ \vdots\\ \dot{\rho}^M_{INS}-\dot{\rho}^M_{GPS}\end{bmatrix}_{2Mx1}=\begin{bmatrix}H^{\rho}_{Mx3} & 0_{Mx3} & 0_{Mx9} & -1_{Mx1} & 0_{Mx1}\\ 0_{Mx3} & H^{\dot{\rho}}_{Mx3} & 0_{Mx9} & 0_{Mx1} & -1_{Mx1}\end{bmatrix}_{2Mx17}\cdot\delta x_{17x1}+\begin{bmatrix}\tilde{\varepsilon}_{\rho,Mx1}\\ \varepsilon_{\dot{\rho},Mx1}\end{bmatrix}_{2Mx1} \tag{2.114}$$

2.5.5.5.1 IMU part

After explaining how to create the Kalman filter and the pseudo-range/pseudo-range rate measurements, now the mechanization of the IMU measurements is going to be introduced.

First of all, the gyroscope measurement model is the following one

$$\widetilde{\omega}_{ib}^{b} = \omega_{ib}^{b} + b_g + S\omega_{ib}^{b} + N\omega_{ib}^{b} + \varepsilon_g \tag{2.115}$$

Where,

$\widetilde{w}_{ib}^{b}$ is the gyroscope measurement vector (deg/h)

w_{ib}^{b} is the true angular rate velocity vector (deg/h)

b_g is the gyroscope instrument bias vector (deg/h)

S is a matrix representing the gyro scale factor

N is a matrix representing non-orthogonality of the gyro triad

ε_g is a vector representing the gyro sensor noise (deg/h)

The matrices N_g and S_g are given as

$$N_g = \begin{bmatrix} 1 & \theta_{g,xy} & \theta_{g,xz} \\ \theta_{g,yx} & 1 & \theta_{g,yz} \\ \theta_{g,zx} & \theta_{g,zy} & 1 \end{bmatrix}$$
$$S_g = \begin{bmatrix} s_{g,x} & 0 & 0 \\ 0 & s_{g,y} & 0 \\ 0 & 0 & s_{g,z} \end{bmatrix} \tag{2.116}$$

Where $\theta_{g,i}$ are the small angles defining the misalignments between the different gyroscope axes i and $s_{g,i}$ are the scale factors for the different gyroscope axes i.

And the accelerometer measurement model is similar to the gyroscope one.

$$\tilde{f}^b = f^b + b_a + S_1 f + S_2 f^2 + N_a f + \delta g + \varepsilon_a \quad (2.117)$$

Where,

$\tilde{f}^b$ is the accelerometer measurement vector (m/s^2)

f^b is the true specific force vector (i.e. observable) (m/s^2)

f is the specific force (m/s^2)

b_a is the accelerometer instrument bias vector (m/s^2)

S_1 is a matrix of the linear scale factor error

S_2 is a matrix of the non-linear scale factor error

N_a is a matrix representing non-orthogonality of the accelerometer triad

δg is the anomalous gravity vector (i.e. deviation from the theoretical gravity value) (m/s^2)

ε_a is a vector representing the accelerometer sensor noise (m/s^2)

The matrices N_a, S_1 and S_2 are

$$N_a = \begin{bmatrix} 1 & \theta_{a,xy} & \theta_{a,xz} \\ \theta_{a,yx} & 1 & \theta_{a,yz} \\ \theta_{a,zx} & \theta_{a,zy} & 1 \end{bmatrix} \quad (2.118)$$

$$S_1 = \begin{bmatrix} s_{1,x} & 0 & 0 \\ 0 & s_{1,y} & 0 \\ 0 & 0 & s_{1,z} \end{bmatrix} \quad (2.119)$$

$$S_2 = \begin{bmatrix} s_{2,x} & 0 & 0 \\ 0 & s_{2,y} & 0 \\ 0 & 0 & s_{2,z} \end{bmatrix} \quad (2.120)$$

Where $\theta_{a,i}$ are the small angles defining the misalignments between the different accelerometer axes i and $S_{1,i}$ and $S_{2,i}$ are the scale factors for the three accelerometers.

For both inertial sensors, the scale factors are considered to be constant but unknown quantities that are uncorrelated between the different sensors. This error can be eliminated using calibration techniques.

2.5.5.5.1.1 IMU mechanization

The output of an accelerometer is the specific force and is given as

$$f^i = a^i - \bar{g}^i \quad (2.121)$$

If $\boldsymbol{a^i} = \ddot{\boldsymbol{r}}^{\boldsymbol{i}}$ it can be rewritten as

$$\ddot{r}^i = f^i + \bar{g}^i \quad (2.122)$$

Where,

$\ddot{r}^i$ is the second derivative of the position vector measured from the origin of the inertial frame to the moving platform

f^i is the specific force

$\bar{g}^i$ is the gravitational vector

The second-order equations can be transformed to a set of first-order equations as follows

$$\dot{r}^i = v^i \tag{2.123}$$
$$\dot{v}^i = f^i + \bar{g}^i$$

The measurements are usually made in the body frame that is assumed to coincide with the sensor triad frame. This measurements can be transformed into the inertial frame using the transformation matrix R_b^i.

$$f^i = R_b^i f^b \tag{2.124}$$

Where R_b^i is a 3x3 rotation matrix that transforms the measurements from the b-frame to the i-frame. The same transformation has to be done with the gravitational vector but that is usually represented in the e-frame or the l-frame.

$$\bar{g}^i = R_e^i \bar{g}^e \tag{2.125}$$

And so we have

$$\dot{v}^i = R_b^i f^b + R_e^i \bar{g}^e \tag{2.126}$$

The rate of change of a transformation matrix is

$$\dot{R}_b^i = R_b^i \Omega_{ib}^b \tag{2.127}$$

With

$$\Omega_{ib}^b = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \tag{2.128}$$

Where w_x, w_y and w_z are the gyroscope measurements in the b-frame. Finally, the mechanization of the i-frame can be summarized as

$$\begin{bmatrix} \dot{r}^i \\ \dot{v}^i \\ \dot{R}_b^i \end{bmatrix} = \begin{bmatrix} v^i \\ R_b^i f^b + R_e^i \bar{g}^e \\ R_b^i \Omega_{ib}^b \end{bmatrix} \tag{2.129}$$

$r^i = [x^i \;\; y^i \;\; z^i]$ is a 3D position vector in the i-frame

$v^i = [v_x^i \;\; v_y^i \;\; v_z^i]$ is a 3D velocity vector in the i-frame

R_b^i is a 3x3 matrix containing the information for the three Euler angles

The following block diagram shows the mechanization of the IMU in the inertial frame.

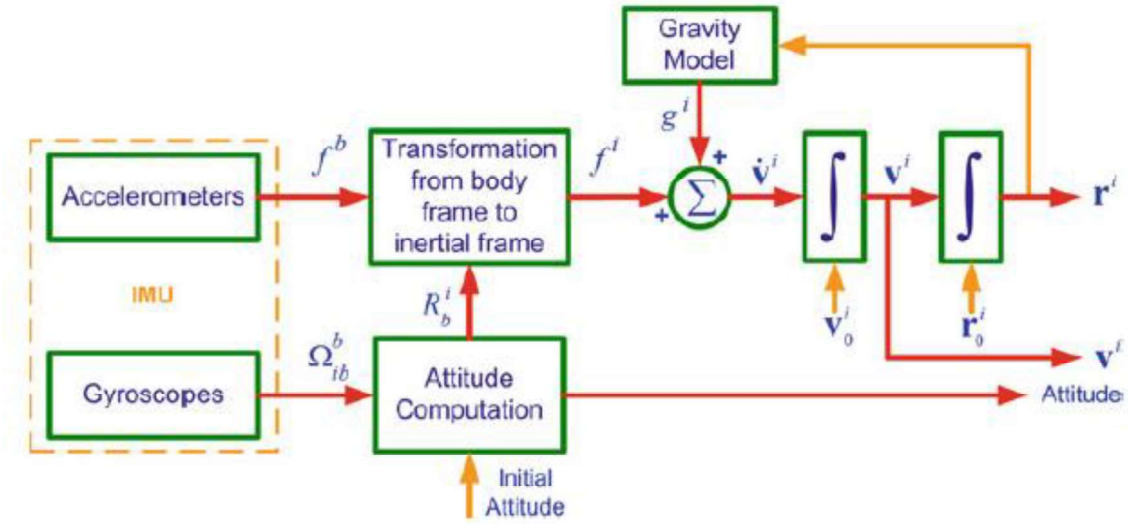


Figure 2.12 Mechanization of the IMU in the inertial frame (Noureldin, Karamat, and Georgy 2013)

2.5.5.5.1.1.1 Mechanization in ECEF frame

A position vector in the e-frame can be transformed to the i-frame by using the rotation matrix R_e^i

$$r^i = R_e^i r^e \tag{2.130}$$

After differentiating twice and rearranging the terms,

$$\ddot{r}^i = R_e^i(\ddot{r}^e + 2\Omega_{ie}^e \dot{r}^e + \dot{\Omega}_{ie}^e r^e + \Omega_{ie}^e \Omega_{ie}^e r^e) \tag{2.131}$$

And substituting in the specific force equation,

$$R_e^i\left(\ddot{r}^e + 2\Omega_{ie}^e \dot{r}^e + \dot{\Omega}_{ie}^e r^e + \Omega_{ie}^e \Omega_{ie}^e r^e\right) = f^i + \bar{g}^i \tag{2.132}$$

Substituting the right hand with their equivalents in the i-frame we get

$$R_e^i\left(\ddot{r}^e + 2\Omega_{ie}^e \dot{r}^e + \dot{\Omega}_{ie}^e r^e + \Omega_{ie}^e \Omega_{ie}^e r^e\right) = R_b^i f^b + R_e^i \bar{g}^e \tag{2.133}$$

The equation can be simplified by letting $R_b^i = R_e^i R_b^e$ and $\dot{\Omega}_{ie}^e r^e = 0$ due to the approximately constant value of the Earth's rotation rate w_{ie}

$$R_e^i\left(\ddot{r}^e + 2\Omega_{ie}^e\dot{r}^e + \dot{\Omega}_{ie}^e r^e + \Omega_{ie}^e\Omega_{ie}^e r^e\right) = R_e^i R_b^e f^b + R_e^i \bar{g}^e$$
$$\ddot{r}^e = R_b^e f^b - 2\Omega_{ie}^e\dot{r}^e + \bar{g}^e - \Omega_{ie}^e\Omega_{ie}^e r^e \quad (2.134)$$

The gravity vector is defined as $g^e = \bar{g}^e - \Omega_{ie}^e\Omega_{ie}^e r^e$, so it can be reduced to

$$\ddot{r}^e = R_b^e f^b - 2\Omega_{ie}^e\dot{r}^e + g^e \quad (2.135)$$

The second-order equation can be separated into two first-order equations

$$\dot{r}^e = v_e$$
$$\dot{v}_e = R_b^e f^b - 2\Omega_{ie}^e v^e + g^e \quad (2.136)$$

The rate of change of the rotation can be given as

$$\dot{R}_b^e = R_b^e\Omega_{eb}^b \quad (2.137)$$

We use the following relationships to write the above expression in terms of angular rate monitored by the gyroscopes.

$$\Omega_{ib}^b = \Omega_{ie}^b + \Omega_{eb}^b$$
$$\Omega_{eb}^b = -\Omega_{ie}^b + \Omega_{ib}^b \quad (2.138)$$
$$\Omega_{eb}^b = \Omega_{ei}^b + \Omega_{ib}^b$$

And substituting the relationships,

$$\dot{R}_b^e = R_b^e(\Omega_{ei}^b + \Omega_{ib}^b) \quad (2.139)$$

The e-frame equations can be summarized as

$$\begin{bmatrix}\dot{r}^e \\ \dot{v}^e \\ \dot{R}_b^e\end{bmatrix} = \begin{bmatrix} v^e \\ R_b^e f^b - 2\Omega_{ie}^e v^e + g^e \\ R_b^e(\Omega_{ei}^b + \Omega_{ib}^b)\end{bmatrix} \quad (2.140)$$

Which represent the mechanization equations in the e-frame where the inputs are the sensed acceleration f^b from the accelerometers and rotation rates $w_{ib}^b$from the gyroscopes. The outputs are the position vector r, the velocity vector v, and the Euler angles, all expressed in the e-frame. The figure below is the block diagram of the mechanization explained before.

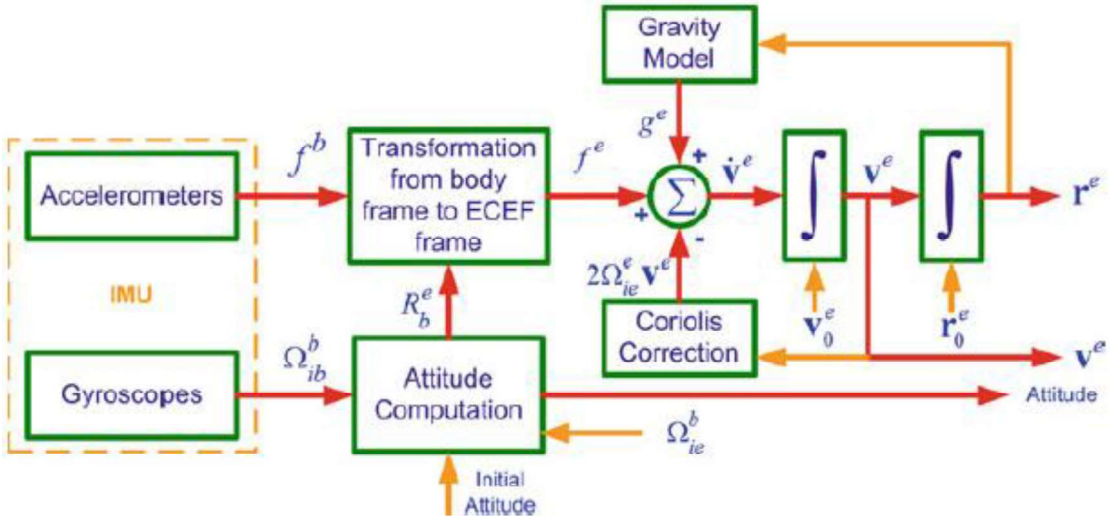


Figure 2.13 Mechanization in the e-frame (Noureldin, Karamat, and Georgy 2013)

2.5.5.5.1.1.2 Mechanization in local-level frame

In many cases to make the mechanization in LLF (Local Level Frame) has advantages:

- It gives a more intuitive navigation solution to the user near the Earth's surface.
- The l-frame axes are aligned with the ENU directions and the attitude angles can be obtained directly at the output of the mechanization equations.
- The computational errors in the navigation parameter on the horizontal plane (E-N) are bounded by the Schuler effect.
- This effect stipulates that the inertial system errors of the horizontal plane components are coupled to produce the Schuler loop and that these errors oscillate at the Schuler frequency 1/5000Hz.

The position vector of a moving plane is expressed in geodetic coordinates in the ECEF frame as

$$r^l = [\varphi \quad \lambda \quad h]^T \tag{2.141}$$

Where φ is the latitude, λ is the longitude and h is the altitude. The rate of change of its position is expressed in terms of the velocity in ENU directions because the platform travels near to the surface of the Earth. The rate of change of the latitude, longitude, and altitude are

$$\begin{aligned} \dot{\varphi} &= \frac{v_n}{R_M + h} \\ \dot{\lambda} &= \frac{v_e}{(R_N + h)\cos\varphi} \\ \dot{h} &= v_u \end{aligned} \tag{2.142}$$

Where,

v_e is the component of the velocity in the east direction

v_n is the component of the velocity in the north direction

v_u is the component of the velocity in the up direction

R_M is the meridian radius of the ellipsoid

R_N is the normal radius of the ellipsoid

Those last equations can be written in matrix notation in the following way

$$\begin{bmatrix} \dot{\varphi} \\ \dot{\lambda} \\ \dot{h} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{R_M + h} & 0 \\ \frac{1}{(R_N + h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_n \\ v_e \\ v_u \end{bmatrix}^l \tag{2.143}$$

$$\dot{r}^l = D^{-1} v^l$$

In which D^{-1} transforms the velocity vector from rectangular coordinates into curvilinear coordinates in the ECEF frame.

2.5.5.5.1.1.3 Velocity Mechanization equations in ECEF frame

The acceleration of the moving body is measured in three mutually orthogonal directions in the b-frame by a three-axis accelerometer. These measurements are the specific force measurements and are given in the b-frame as

$$f^b = \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \tag{2.144}$$

They can be transformed to the local-level frame using the rotation matrix

$$f^l = \begin{bmatrix} f^e \\ f^n \\ f^u \end{bmatrix} = R_b^l f^b = R_b^l \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \tag{2.145}$$

The acceleration components expressed in the LLF cannot directly yield the velocity components of the moving body. That happens due to the following reasons:

- The rotation of the Earth about its spin axis is interpreted in the LLF as the angular velocity w_{ie}^l

$$\omega_{ie}^l = \begin{bmatrix} 0 \\ \omega^e \cos\varphi \\ \omega^e \sin\varphi \end{bmatrix} \tag{2.146}$$

- A change of the LLF with respect to the earth arises from the definition of the local north and vertical directions. The north direction is tangent to the meridian. The vertical direction is normal to the Earth's surface. This effect is interpreted by the angular velocity w_{el}^l

$$\omega_{el}^l = \begin{bmatrix} -\dot{\varphi} \\ \dot{\lambda}\cos\varphi \\ \dot{\lambda}\sin\varphi \end{bmatrix} = \begin{bmatrix} -\dfrac{v_n}{R_M + h} \\ \dfrac{v_e}{R_N + h} \\ \dfrac{v_e \tan\varphi}{R_N + h} \end{bmatrix} \tag{2.147}$$

- The Earth's gravity field is

$$g^l = \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} \tag{2.148}$$

Taking these three factors into consideration, the expression of the time rate of change of the velocity components in the moving body can be derived. The Earth referenced velocity vector can be transformed into the LLF by using the rotation matrix

$$v^l = R_e^l \dot{r}^e \tag{2.149}$$

Where, $v^l = [v_e, v_n, v_u]^T$, and the time derivative is

$$\dot{v}^l = \dot{R}_e^l \dot{r}^e + R_e^l \ddot{r}^e \tag{2.150}$$

Substituting the rate of change of the transformation matrix, $\dot{R}_e^l = R_e^l \Omega_{le}^e$

$$\begin{aligned} \dot{v}^l &= \dot{R}_e^l \Omega_{le}^e \dot{r}^e + R_e^l \ddot{r}^e \\ \dot{v}^l &= \dot{R}_e^l (\Omega_{le}^e \dot{r}^e + \ddot{r}^e) \end{aligned} \tag{2.151}$$

We know that $\Omega_{le}^e = -\Omega_{el}^e$ and $\dot{r}^e = v^e$

$$\dot{v}^l = R_e^l (\ddot{r}^e - \Omega_{el}^e v^e) \tag{2.152}$$

We can transform the position vector r from the ECEF frame into the inertial frame by

$$r^i = R_e^i r^e \tag{2.153}$$

Taking the time derivative and using the following relationship, $\dot{R}_e^i = R_e^i \Omega_{ie}^e$

$$\begin{aligned} \dot{r}^i &= \dot{R}_e^i r^e + R_e^i \dot{r}^e \\ \dot{r}^i &= R_e^i \Omega_{ie}^e r^e + R_e^i \dot{r}^e \\ \dot{r}^i &= R_e^i (\Omega_{ie}^e r^e + \dot{r}^e) \end{aligned} \tag{2.154}$$

Where Ω_{ie}^e is the skew-symmetric matrix corresponding to w_{ie}^e. Taking the second time derivative and rearranging terms, we obtain

$$\ddot{r}^i = R_e^i (\ddot{r}^e + 2\Omega_{ie}^e \dot{r}^e + \dot{\Omega}_{ie}^e r^e + \Omega_{ie}^e \Omega_{ie}^e r^e) \tag{2.155}$$

We know that the Earth's rotation rate is approximately constant, so $\dot{\Omega}_{ie}^e = 0$

$$\ddot{r}^i = R_e^i (\ddot{r}^e + 2\Omega_{ie}^e \dot{r}^e + \Omega_{ie}^e \Omega_{ie}^e r^e) \tag{2.156}$$

Due to Newton's laws, we have the relationship between the second time derivative, the specific force, and the gravitational field vector,

$$\ddot{r}^i = f^i + \bar{g}^i \tag{2.157}$$

And substituting the value of $\ddot{r}^i$

$$\begin{aligned} f^i + g^i &= R_e^i(\ddot{r}^e + 2\Omega_{ie}^e\dot{r}^e + \Omega_{ie}^e\Omega_{ie}^e r^e) \\ R_i^e(f^i + \bar{g}^i) &= (\ddot{r}^e + 2\Omega_{ie}^e\dot{r}^e + \Omega_{ie}^e\Omega_{ie}^e r^e) \\ f^e + \bar{g}^e &= (\ddot{r}^e + 2\Omega_{ie}^e\dot{r}^e + \Omega_{ie}^e\Omega_{ie}^e r^e) \\ \ddot{r}^e &= f^e + \bar{g}^e - \Omega_{ie}^e\Omega_{ie}^e r^e - 2\Omega_{ie}^e\dot{r}^e) \end{aligned} \tag{2.158}$$

The gravitational field vector ($\bar{g}^e$) and the gravity field vector (g^e) are related;

$$g^e = \bar{g}^e - \Omega_{ie}^e\Omega_{ie}^e r^e \tag{2.159}$$

Substituting in the previous equation and using the fact that $\dot{r}^e = v^e$

$$\ddot{r}^e = f^e + g^e - 2\Omega_{ie}^e v^e \tag{2.160}$$

Substituting this expression in the equation for the velocity derivative in the LLF

$$\begin{aligned} \dot{v}^l &= R_e^l(f^e + g^e - 2\Omega_{ie}^e v^e - \Omega_{el}^e v^e) \\ \dot{v}^l &= R_e^l\{f^e + g^e - (2\Omega_{ie}^e + \Omega_{el}^e)v^e\} \\ \dot{v}^l &= R_e^l f^e + R_e^l g^e - R_e^l(2\Omega_{ie}^e + \Omega_{el}^e)v^e \\ \dot{v}^l &= f^l + g^l - R_e^l\big(2R_l^e\Omega_{ie}^l R_e^l + R_l^e\Omega_{el}^l R_e^l\big)v^e \\ \dot{v}^l &= R_b^l f^b + g^l - \big(2R_e^l R_l^e\Omega_{ie}^l + R_e^l R_l^e\Omega_{el}^l\big)R_e^l v^e \\ \dot{v}^l &= R_b^l f^b - \big(2\Omega_{ie}^l + \Omega_{el}^l\big)v^l + g^l \end{aligned} \tag{2.161}$$

With R_b^l the transformation matrix from the b-frame to the l-frame, f^b the specific force measured by the accelerometers in the b-frame and g^l the gravity vector in the l-frame. Ω_{ie}^l and Ω_{el}^l are the skew-symmetric matrices of w_{ie}^l and w_{el}^l.

$$w_{ie}^l = \begin{bmatrix} 0 \\ \omega^e\cos\varphi \\ \omega^e\sin\varphi \end{bmatrix} \rightarrow \Omega_{ie}^l = \begin{bmatrix} 0 & -\omega^e\sin\varphi & \omega^e\cos\varphi \\ \omega^e\sin\varphi & 0 & 0 \\ -\omega^e\cos\varphi & 0 & 0 \end{bmatrix}$$

$$w_{el}^l = \begin{bmatrix} -\dfrac{v_n}{R_M + h} \\ \dfrac{v_e}{R_N + h} \\ \dfrac{v_e\tan\varphi}{R_N + h} \end{bmatrix} \rightarrow \Omega_{el}^l = \begin{bmatrix} 0 & \dfrac{-v_e\tan\varphi}{R_N + h} & \dfrac{v_e}{R_N + h} \\ \dfrac{v_e\tan\varphi}{R_N + h} & 0 & \dfrac{v_n}{R_M + h} \\ \dfrac{-v_e}{R_N + h} & \dfrac{-v_n}{R_M + h} & 0 \end{bmatrix} \tag{2.162}$$

2.5.5.5.1.1.4 Attitude (orientation) mechanization equations

The attitude of the moving body is determined by solving the time derivative equation of the rotation matrix that relates the body frame with the LLF. For local-level mechanization, the following time derivative equation of the transformation matrix should be considered.

$$\dot{R}_b^l = R_b^l \Omega_{lb}^b \tag{2.163}$$

And the angular velocity matrix can be expressed as

$$\begin{aligned} \Omega_{lb}^b &= \Omega_{li}^b + \Omega_{ib}^b \\ \Omega_{lb}^b &= -\Omega_{il}^b + \Omega_{ib}^b \\ \Omega_{lb}^b &= \Omega_{ib}^b - \Omega_{il}^b \end{aligned} \tag{2.164}$$

If we substitute

$$\dot{R}_b^l = R_b^l (\Omega_{ib}^b - \Omega_{il}^b) \tag{2.165}$$

The rotation matrix can be obtained by solving this equation for the attitude angles. Ω_{ib}^b, the rate of rotation with respect to the i-frame, is measured by the gyroscopes. In addition to the angular velocities, the gyroscopes contain the Earth's rotation and the change of orientation of the LLF. So Ω_{il}^b must be subtracted from Ω_{ib}^b to remove these effects. Ω_{il}^b has two parts; the Earth's rotation rate with respect to the i-frame but represented in the body frame (Ω_{ie}^b) and the change of orientation of the LLF with respect to the ECEF frame as expressed in body frame (Ω_{el}^b).

$$\begin{aligned} \Omega_{il}^b &= \Omega_{ie}^b + \Omega_{el}^b \\ \Omega_{ie}^b &= R_l^b \Omega_{ie}^l R_b^l \\ \Omega_{el}^b &= R_l^b \Omega_{el}^l R_b^l \\ \Omega_{il}^b &= R_l^b \Omega_{ie}^l R_b^l + R_l^b \Omega_{el}^l R_b^l \\ \Omega_{il}^b &= R_l^b (\Omega_{ie}^l + \Omega_{el}^l) R_b^l \end{aligned} \tag{2.166}$$

And substituting these relationships in the equation introduced before,

$$\dot{R}_b^l = R_b^l\left[\Omega_{ib}^b - R_l^b(\Omega_{ie}^l + \Omega_{el}^l)R_b^l\right] \tag{2.167}$$

Where Ω_{ib}^b is the skew-symmetric matrix corresponding to the gyroscopic measurement vector.

The previous results can be summarized in the following matrix which shows the mechanization in the local-level frame.

$$\begin{bmatrix} \dot{r}^l \\ \dot{v}^l \\ \dot{R}_b^l \end{bmatrix} = \begin{bmatrix} D^{-1}v^l \\ R_b^l f^b - (2\Omega_{ie}^l + \Omega_{el}^l)v^l + g^l \\ R_b^l(\Omega_{ib}^b + \Omega_{il}^b) \end{bmatrix} \tag{2.168}$$

The position output is in ECEF curvilinear coordinates, the velocity is in l-frame coordinates (ENU) and the attitude angles (yaw, roll, and pitch) are measured with respect to the l-frame.

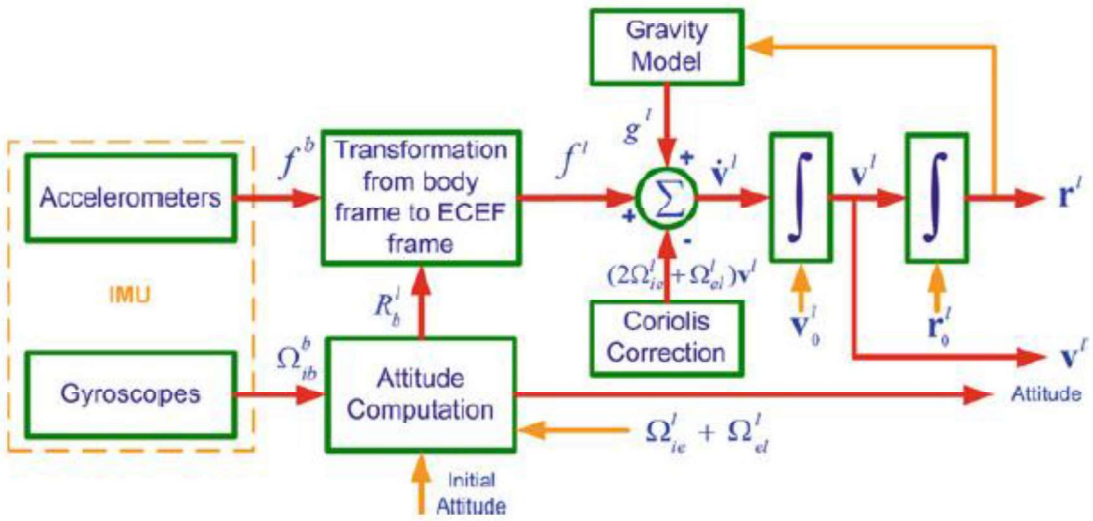


Figure 2.14 Mechanization in ECEF coordinates (Noureldin, Karamat, and Georgy 2013)

2.6 GNSS AND UWB DATA FUSION

Data fusion between GNSS and UWB can be done in multiple ways as explained in Section 2.5 for GNSS and IMU fusion. In the last times, some fusion algorithms have been presented in order to enhance the performance of positioning systems in GNSS compromised environments.

Lack of GNSS signals it's a challenge for implementing positioning systems in urban canyons or tunnels. Nowadays new options have been presented for positioning in GNSS denied areas. (Gao et al. n.d.), introduces a cooperative positioning solution using integrated Ultra Wideband (UWB) and GNSS.

Other works (Macgougan and Klukas 2009) instead of using GNSS stand-alone, use RTK (Real Time Kinematics) GNSS to fuse with UWB. This technique is common in the industry, but limited in application due to signal masking, attenuation and multipath in hostile environments. That is the reason why RTK systems are not useful in urban canyons, forests and congested construction environments. For the case of this thesis, this kind of GNSS is not a good option to do a fusion with UWB as the railway industry has usually operations in this kind of environments. Moreover, RTK systems need more data than the one obtained by a single GNSS receiver. The fusion presented is mostly used in surveying applications.

Another application in which GNSS and UWB data fusion has been presented is the emergency system positioning (Han et al. 2019) (AIOSAT n.d.). For emergency systems such as firefighters, police or security services, indoor/outdoor coordinated positioning systems are needed, and the fusion presented in this section can fulfil those requirements. Those indoor/outdoor transitions could be performed by the presented fusion maintaining a high-accuracy positioning.

After analysing the publications introduced, the use of UWB seems a good solution to maintain high accuracy positioning in environments in which GNSS signal reception is not good or it is denied. However, for the use of UWB some anchors have to be placed in the area in which positioning is performed. This fact makes the

use of UWB not suitable for large areas as a big number of anchors are needed to make the system work. That leads us to the use of UWB fused with GNSS to demanding areas in a railway, or for train stations in which high accuracy is needed due to the high number of trains in a small area.

Moreover, a brief analysis of the state of the art in GNSS and UWB fusion has shown us, that there are not applications with GNSS and UWB for railway available in the literature. That opens a good opportunity in order to analyse if the performance of this kind of fusion in terms of field test could be useful for the railway industry.

2.7 RAIM

Receiver Autonomous Integrity Monitoring (RAIM) is defined as an algorithm that determines the integrity of the GNSS solution. It compares pseudorange measurements to ensure that they are consistent. RAIM algorithms make use of measurement redundancy to check the relative consistency among them and satellite failure detection.

2.7.1 Fault Detection and Exclusion

- Fault Detection: This process checks the reliability of the measurements, using statistical hypothesis or least-squares position estimation (RAIM - Navipedia n.d.).
- Fault Exclusion: This process is only activated if a fault is detected in the process explained before. Its function is to detect the measurements responsible for the fault detected before. In this way, it would be discarded from the navigation solution calculus (RAIM - Navipedia n.d.).

Fault detection and fault exclusion processes should be called iteratively until no faultier measurements are detected.

The enhanced version of RAIM (eRAIM) is also known as Fault Detection and Exclusion (FDE). It uses at least six satellites to detect and exclude a possible faulty satellite without stopping the navigation function. The eRAIM algorithms are derived from the least-squares estimators of the state parameters in a Gauss-Markov KF (WANG 1999) and are used to assess GNSS/INS (see Subsection 2.5.5 for more information about the integration) RAIM performance for a tightly coupled simulation scenario (Hewitson and Wang n.d.). For more information about FDE methods and their strengths and weaknesses, see Appendix A.1.

2.7.2 Integrity Engine

Receiver Autonomous Integrity Monitoring algorithms have been developed for use with GPS positioning and now are extended to all GNSS. The development has principally targeted the field of aviation due to the associated integrity requirements for aviation applications. In this transport area, multipath and NLOS effects are small and RAIM techniques focus on providing a barrier to faults arising at a GNSS system level. RAIM algorithms are designed with the assumption that only one fault may occur at any one time. Multiple-fault algorithms are in development with a view to multi-constellation use, but they tend to be computationally demanding.

Application of RAIM to the rail environment must take into account that the probability of multiple simultaneous faults is orders of magnitude higher than in the aviation case due to the dominance of multipath and NLOS. As no proven snapshot RAIM method exists for dealing with this threat, the implementation considered for trading off position engine options is that of the well-known weighted least-squares method (Hillenbrand and Hofestädt n.d.).

Horizontal Protection Level (HPL) is based on the assumption that Fault Detection and Exclusion process excluded any faulty satellite in the processing. It combines horizontal components of the P matrix into a standard deviation term and multiples it by a safety factor based on the integrity requirement of the application.

Different RAIM algorithms have been investigated in the last few years:

- Least-Square-Residuals (LS) RAIM: It compares the size of the least-square residuals of redundant pseudorange measurements. If a measurement is faulty the residual becomes large.
- Solution Separation RAIM: If more than the minimum 4 satellite signals are received at a time, the position determination can be split into satellite subsets. That supposes that at least one healthy subset is available each time (RAIM - Navipedia n.d.).

- Multiple Hypothesis Solution Separation (MHSS) RAIM: It is a Solution Separation RAIM extension that admits multiple simultaneous faulty measurements (RAIM - Navipedia n.d.).
- Relative RAIM: RRAIM uses carrier phase measurements to propagate older position solutions in time. RAIM is performed in the carrier trajectory to ensure integrity and then protection levels are calculated with the new parameters (RAIM - Navipedia n.d.).
- Absolute RAIM: It is a RAIM extension based on MHSS RAIM. It focuses on the receiver autonomy instead of in the ground segment (RAIM - Navipedia n.d.).
- Isotropy Based Protection Level: It computes the protection level based on the error isotropy considering multiple fault conditions (RAIM - Navipedia n.d.).

2.7.2.1 Snapshot Integrity Engine

Numerous Receiver Autonomous Integrity Monitoring (RAIM) algorithms have been developed in academia and industry for use with GNSS snapshot positioning. Development has principally targeted the field of aviation due to the associated integrity requirements for aviation applications. Since aviation involves a clear-sky antenna view, multipath and NLOS effects are small and RAIM techniques focus on providing a barrier to faults arising at a GNSS system level. As these kinds of multiple faults are very rare, traditional techniques are designed with the assumption that only one fault may occur at any one time. Although snapshot RAIM methods for handling multiple-faults are in development with a view to multi-constellation use, they tend to be computationally demanding.

Application of RAIM to the rail environment must take into account that the probability of multiple simultaneous faults is orders of magnitude higher than in the

aviation case due to the dominance of multipath and NLOS. As no proven snapshot RAIM method exists for dealing with this threat, the implementation considered for trading off position engine options is that of the well-known weighted least-squares method (Hillenbrand and Hofestädt n.d.).

2.7.2.2 Kalman Filter Integrity Engine

Following the actions of the FDE process, a new estimate of the state vector $\hat{X}_t^+$ is computed using the remaining (not excluded) measurements. Correspondingly the state covariance matrix is updated to form P_t^+. In order to form a Horizontal Protection Level (HPL_{GNSSKF}),$HPL_{exclude}$ and HPL_{remain} need to be generated.

$HPL_{exclude}$ is based on the assumption that the FDE process excluded any faulty satellites in the processing. Therefore, it combines the horizontal components of the P_t^+ matrix into a single standard deviation term and multiplies by a 'safety factor' based on the integrity risk requirement of the application. It is calculated as:

$$HPL_{exclude} = \mathrm{K}\sigma_H \tag{2.169}$$

Where,

K is defined as $K(p) = Q^{-1}(1-p)$, where *Q* is the cumulative distribution function of a Rayleigh-distributed random variable with unit standard deviation and *p* is the probability of missed detection requirement of the application

$\sigma_H = \sqrt{\sigma_N^2 + \sigma_E^2}$ where σ_N and σ_E are the standard deviations of the updated state solution Northing and Easting position components as contained in P_t^+.

However, the FDE approach used can only remove the faulty measurements which exceed the threshold. Therefore, there may still be large undetectable biases remaining on the measurements. Additionally, in order to guarantee the maximum

availability of integrity, FDE does not exclude any faulty satellites if at any point there are less than five satellites available. For instance, suppose there are five or fewer satellites in view for 2 consecutive seconds; if there is at least one satellite with a large NLOS effect, and this is detected, the FDE would not exclude any satellites. Therefore, HPL_{remain} is used to help account for non-Gaussian behavior in the underlying range errors. HPL_{remain} is mainly dependent on the maximum undetectable bias, the measurement residuals and the safety factor *K*. It is computed as:

$$HPL_{remain} = \mathrm{K} \cdot Bias_H \cdot b_{residual} \tag{2.170}$$

Where $Bias_H$, the horizontal maximum, is undetectable bias and $b_{residual}$ is the measurement residuals.

Thus, we can derive:

$$HPL_{GNSSKF} = \max(HPL_{exclude}, HPL_{remain}) \tag{2.171}$$

2.8 COMMERTIAL POSITIONING SYSTEMS AVAILABLE FOR DIFFERENT LEVEL OF USERS

The positioning systems existing in the literature, can be grouped in three macrosegments (GSA 2018).

- Mass market: presenting high-volume receivers for consumer devices. Automotive (not safety critical), consumer drones, smartphones, and specialised IoT devices from Health to robotics are all covered.
- Transport safety and liability-critical solutions: presenting receivers built in accordance with standards to deliver such solutions. Automotive, aviation, professional drones, maritime, search and rescue and, new to this issue of the TR, space-borne GNSS applications are all covered.
- High precision and timing solutions: presenting receivers designed to deliver the highest accuracy (position or time) possible. Agriculture, GIS, Surveying and Timing and Synchronisation applications are all covered.

In the first macrosegment, different manufacturers lead the market. Within them, Broadcom in America, or U-blox and STMicroelectronics in Europe fight for that leadership. In the smartphone market, Qualcomm, Broadcom and MediaTek are the most important manufacturers. As an example, Broadcom introduced the BCM47755, a dual-frequency receiver for mobile phones, tablets and wearables with good impact in the market and brands like Xiaomi using it for their devices (World 2017). U-blox and STMicroelectronics in the other hand lead the market in automotive, transportation and Internet of the Things (IoT). U-blox sells different receivers divided in P and T models inside the different updates. U-blox m6T and m6P, m8T and m8P and the new z9P are some of the most sold receivers of the brand. Inside the same generation the hardware is quite similar and the firmware is the one that changes from one model to the other (Explorer 2018).

The second macrosegment repeats some of the most important manufacturers. In aviation Garmin, Thales or Rockwell are well positioned in the market. In maritime environment the receivers are sometimes integrated on the on-board equipment and on automotive, STMicroelectronics and U-blox are in front of the market. In railway not many market projects are developed yet and so regarding the common points that could appear with automotive industry, the research line could be similar.

The third macrosegment is the most complex and expensive. High-end professional receivers appear here and most of them are out of the budget for many projects implementation. However, some of the manufacturers for the low-end receivers also create high-end receivers such as U-blox for precise timing. New companies as Hexagon with their Leica and Novatel brand appear here and also Septentrio is in the top of the market in Europe.

The kind of receiver that is going to be used in the this dissertation will be discussed after presenting the objectives.

Apart from the receivers/systems explained before, some receiver-algorithm combinations are introduced afterwards due to their interest in different ways such as open-source systems, there are only a few available for the mass user to buy. In that way, a brief analysis of other existing systems has been done.

U-blox, besides GNSS receivers, also provides an application called U-center (U-blox n.d.), that is available for free in their website. This application performs positioning for data received from devices created by the firma. It gives plenty of data about the GNSS parameters and the position obtained along with performance indicators. However, the code is not open source and it is not possible for the user to know the algorithm performed inside.

Moreover, one of the most well known is RTKLIB (RTKLIB n.d.). It shows plenty of data about the GNSS signal received along with the obtained position and some performance indicators about it. The code is open and different algorithms can be used to obtain a position estimation, such as least-squares or Kalman filter.

Moreover, the post processing of the data can be performed to obtain better positioning results. The useful point of this open-source application is that the user is totally aware of the algorithms used for positioning and the data treatment done.

The last step related to positioning systems is to analyse a commertial positioning device created specifically for the railway industry. Savvy telematics created a stand-alone telematics device with a running life up to 15 years called SAVVY CargoTrac (SAVVY n.d.). The device is optimized for the railway industry and has a location system using GPS and GLONASS signals. However, its accuracy and positioning algorithm is not open for the user and its performance is not shown by the company.

Some of the presented applications have been tested during the current work and a comparison between the created algorithm within the system and the applications presented has been done in Chapter 7.

Chapter 3

Objectives

Chapter 3 presents the objectives of this work. After analyzing the information introduced in Chapter 1 and Chapter 2, the gaps identified in the literature have been analyzed and possible solutions to them will be stated in pursuit of evolving the positioning strategies and adapting them to the new challenges of the railway transport applications.

3.1 INTRODUCTION

In the literature, different positioning systems can be found to improve transport systems by means of reducing the operation cost and improving the safety of them. These systems involve GNSS, IMU or UWB among others. Some of them have been fused to perform better accuracies along with a bigger availability of the whole system. However, a fusion of the three technologies presented (GNSS, IMU, and UWB) for railway nor other transport systems has not been found in the literature.

One of the main problems when facing positioning for transport applications is the lack of availability in the position estimation availability that appears in certain areas. To cope with this issue, a multi-sensor approach with software enhancements is proposed. The objective of this research work deals with fusing different sensors and creating novel software strategies to achieve a higher availability with the best possible accuracy. The seamless position will benefit in all the operation modes, from the train station to a harsh environment for satellites, during the train operation

3.2 OBJECTIVES

The main objective of this research work is to study the current status of the existing positioning systems for railway systems and to develop and assess a novel solution combining different positioning technologies to obtain a low-cost system improving the features presented by the currently used systems in the transport industry. The system should be able to position the vehicle during the whole journey. Moreover, the accuracy of the system is determined by the operation mode and the transport system in which it is going to be used. Finally, the system is wanted to be low-cost for the state of the art systems in order to be affordable for different transport systems besides the railway industry.

The global objective of this research work has been split in the following partial objectives:

- **Introduction and analysis of the existing positioning systems.**
  A deep analysis of the positioning technologies and algorithms existing to understand the singularities of each of them and face a suitable strategy when doing their fusion. Moreover, different data fusion techniques are also analyzed.

- **Design of a novel positioning system and algorithm.**
  Having understood the different systems, a novel positioning system should be created using the ones that best fit the needs. A proper algorithm has to be developed and assessed to compute positions and fuse the data obtained by the different positioning systems.

- **Improvement of the availability of the system.**
  One of the problems that face positioning systems in transportation nowadays, is the lack of availability in some areas. This can happen when going through tall buildings, canyons (urban or not) and tunnels. A solution to this problem should be given by the presented system and algorithms.

- **Improvement of the accuracy in harsh scenarios for GNSS systems.**
  In addition to the improvement of the availability of the system, progress in terms of accuracy should be done in some harsh scenarios. For that, a novel software technique would be used, based on map aiding.
- **Fusion of GNSS, IMU and UWB technologies for transport systems.**
  A solution for the fusion of the three systems stated should be given to improve the accuracy and availability of the systems existing in the industry from the train station to the normal operation.
- **Test and validation of the proposed system and algorithms.**
  To validate and check the system and algorithm created, some tests need to be performed. These tests are done in different scenarios; harsh environments in which the real potential of the system will be shown.
- **Implementation of a recursive algorithm in terms of accuracy improvement.**
  A recursive algorithm has been implemented to improve the accuracy of the presented system as read in the literature.
- **Implemetation of a Fault Detection and Exclusion (FDE) method in terms of accuracy improvement.**
  Following the same idea, a Fault Detection and Exclusion method (OLOO) has been implemented to improve the accuracy as presented in appendix A.1.
- **Analysis of the performance obtained by the system created.**
  Finally, a deep analysis of the results obtained with the created system should be done. The performance in terms of accuracy and availability should be studied in the different scenarios introduced.

Chapter 4

Design and implementation of the algorithm

This chapter describes the design and implementation of the positioning algorithm created. First, the whole architecture of the algorithm is introduced. Then, the implementation and functionality of each of the modules included in the algorithm are detailed.

4.1 PRELIMINARY DECISIONS

First of all, the decisions made before starting the creation of the algorithm are introduced. The three different positioning subsystems presented in Chapter 2 need to be fused to build the whole positioning algorithm.

The first technology chosen is GNSS and its data are obtained by a low-cost receiver. It is a basic technology used in this algorithm due to its mass deployment, the possibility of low-cost, and wide-range. In this aspect, this thesis wants to present an output usable for all kind of GNSS receivers from the simplest to more complex ones and make it affordable for many applications and industries. Taking into account that, GPS is the GNSS system used in this thesis. GPS has the advantage that is fully deployed and it is the positioning system most used nowadays. The algorithm is wanted to be used with as many different hardware as possible and some of the low-cost hardwares are only compatible with GPS even if nowadays some include Galileo and / or GLONASS constellations. Also thinking in the possibility of more compatible hardware, only L1 frequency has been used. In order to be independent of the receiver software or the processing made by the receiver, pseudoranges and ephemerides data have been used. In any case, the upgrade to multi-frequency and/or multi-constellation alternatives will only benefit the results present in this dissertation in terms of availability and precision obtained.

When using the IMU, also the same options have been taken into account for the algorithm presented in this thesis work: simplicity, and compatibility with different hardware. This is the reason to use only accelerometers and gyroscopes and not using magnetometers. The big metal masses existing in the railway environment imply difficult calibrations in each use of the magnetometers. Most of the low-end IMUs have six degrees of freedom and it is not that usual to find some with nine degrees of freedom. The measurements used are the raw linear and angular accelerations.

Finally, the UWB system used can give the ranges from the tag to the anchors and those have been used to calculate the UWB position estimate. The UWB system chosen is in the same line as the other systems. It is a commertial low-cost kit suitable to evaluate the technology.

The algorithm has been programmed to be used in two different ways, directly connected to the hardware receiving data in real-time, or using data previously recorded in the reference scenarios. In any case, when the raw data are recorded , then many different tests could be performed using that data.

4.2 POSITIONING ALGORITHM

This section introduces the complete algorithm with, first, the main positioning functions, and afterwards the details for those.

Figure 4.1 shows the flow of the algorithm in a summarized way. Six different modules can be seen there; check of data availability, read of the data, data treatment, application of the Kalman filter, use of the KBS, and data output.

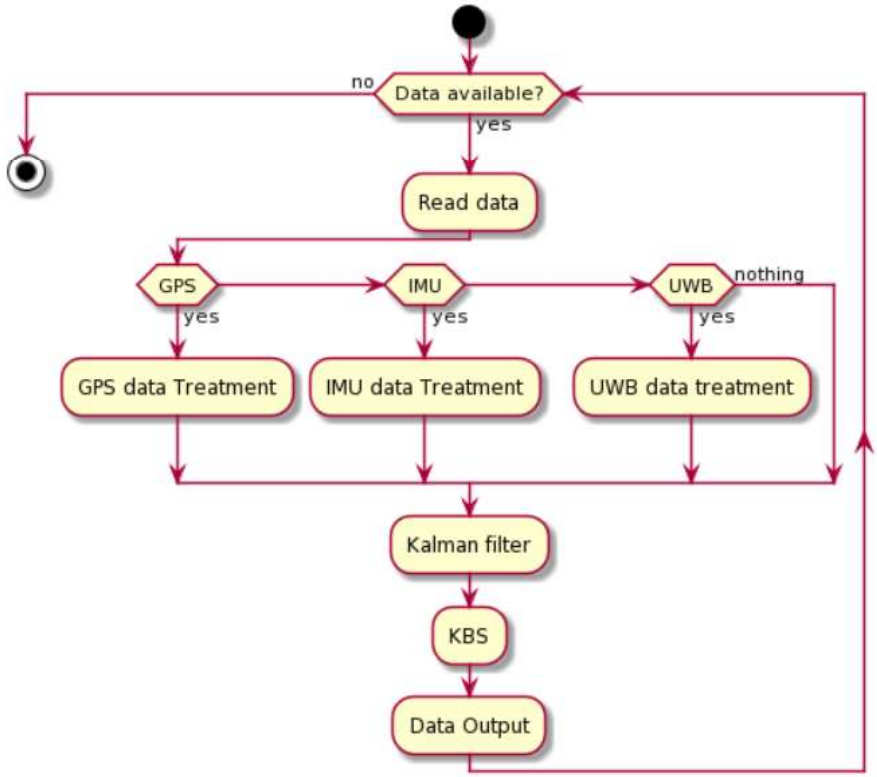


Figure 4.1 Overview of the positioning algorithm

The algorithm first checks if there is any data available. If there is no data available, the behaviour is different depending on the mode of use of the algorithm.

If there is data available, data is read. The data is read from a file with an Input File format created by CEIT that has been used for measuring data from different systems with the same timestamp. A parser to obtain the data available in the input file has been programmed.

Once data is read, a different treatment is given to every line depending on the data available: GPS, IMU or UWB.

Afterwards, a Kalman filter is performed to the input data in order to provide a position estimation. A Kalman filter has been chosen to be performed by the positioning algorithm to fuse the different sensors following the information presented in sections 2.5 and 2.6. Once a position estimation is available, the Known Blocked Scenarios (KBS) part of the algorithm is used. A deep explanation of it is also available in this chapter.

Finally, the data output process is completed. For that, the position estimation obtained through the steps explained before is written in the Output file. The output file created by the algorithm is nowadays a file with a proprietary format created by CEIT to cover the necessities of its performance computation algorithms.

4.2.1 Check Data Availability

The first block of the algorithm is the one that checks if there is some data available or not. As explained before, in this work, data from prerecorded files are used. So the first step is to check if the necessary files are available.

When the files needed are received, their format is checked. If they fulfil the format required, the first part of data availability checking is finished.

The second part of the data availability checking implies the verification of new data available in the data input file. The input files are read line by line, data by data until the end of the file. When the end of the file is reached, no more data will be available and the positioning algorithm will finish the execution.

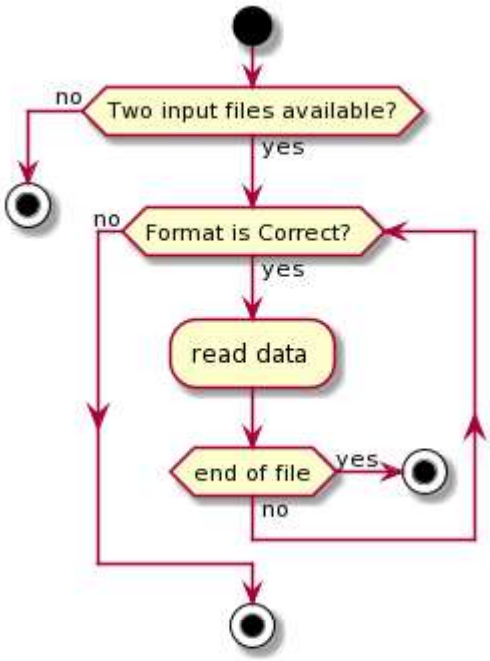


Figure 4.2 Overview of the data availability check

The format used in the input files is different depending on the technology used. The input files for ephemerides have also a different format. However, all the files have been created respecting the .csv (comma-separated value) format.

4.2.1.1 GPS

This section introduces the format of the GPS data included in the input files. The data of all the columns introduced are not used in the final positioning algorithm.

Column number	Name	Description
1	Time	Time of the measurement. Time is given in the following format: hh:mm:ss.fff/ GPS_week/GPS_second_of_week
2	Identifier	A number to identify the technology used. 0 is the identifier for GPS
3	Δ of Time	Time elapsed since the last data measurement in seconds.
4	Time (s)	Time in seconds transcurred since the beginning of the measurement.
5	X(m)	Position calculated by the GPS receiver in ECEF (meters) for the X-axis.
6	Y(m)	Position calculated by the GPS receiver in ECEF (meters) for the Y-axis.
7	Z(m)	Position calculated by the GPS receiver in ECEF (meters) for the Z-axis.
8	Velocity(m/s)	Velocity calculated by the GPS receiver in the driving direction in m/s.
9	Yaw(rad)	Yaw calculated by the GPS receiver in radians
10	Vx(m/s)	Velocity calculated by the GPS receiver in ECEF (meters) for the X-axis.
11	Vy(m/s)	Velocity calculated by the GPS receiver in ECEF (meters) for the Y-axis.

12	Vz(m/s)	Velocity calculated by the GPS receiver in ECEF (meters) for the Z-axis.
13	GPS Time	Time given by GPS receiver in the following format: hhmmss
14	GPS Date	Date given by the GPS receiver in the following format: ddmmyy
15	StdDevX(m)	Standard Deviation of the position in the X-axis given by the GPS receiver in meters.
16	StdDevY(m)	Standard Deviation of the position in the Y-axis given by the GPS receiver in meters.
17	StdDevZ(m)	Standard Deviation of the position in the Z-axis given by the GPS receiver in meters.
18	NumSatUsed	Number of satellites from which data has been received.
19	HPL	7-sigma Horizontal Protection Level calculated from the Standard Deviations given by the GPS receiver.
20	SNR Average(dB)	Average of the SNR value given by all the satellites with data available
21-24-27-…-21+(NumSatUsed*3)	Number of Satellite	Number of the satellite from which data is going to be given in the next columns
22-25-28-…-22+(NumSatUsed*3)	Pseudorange	Pseudorange of the satellite in meters

23-26-29-…-23+(NumSatUsed*3)	SNR	SNR of the satellite in dB

Table 4.1 Format of GPS data in the input files

4.2.1.2 IMU

Column number	Name	Description
1	Time	Time of the measurement. Time is given in the following format: hh:mm:ss.fff
2	Identifier	A number to identify the technology used. 1 is the identifier for IMU
3	Δ of Time	Time elapsed since the last data measurement in seconds.
4	Time (s)	Time in seconds transcurred since the beginning of the measurement.
5	X(m)	Position calculated by the IMU in ECEF (meters) for the X-axis with GPS as reference.
6	Y(m)	Position calculated by the IMU in ECEF (meters) for the Y-axis with GPS as reference.

7	Z(m)	Position calculated by the IMU in ECEF (meters) for the Z-axis with GPS as reference.
8	Velocity(m/s)	Velocity calculated by the IMU in the driving direction in m/s.
9	Yaw(rad)	Yaw angle calculated by the IMU in radians
10	Pitch(rad)	Pitch angle calculated by the IMU in radians
11	Roll(rad)	Roll angle calculated by the IMU in radians
12	AccX(G)	Linear acceleration in the X-axis in Gs
13	AccY(G)	Linear acceleration in the Y-axis in Gs
14	AccZ(G)	Linear acceleration in the Z-axis in Gs
15	GyroX(degree/s)	Angular acceleration in the X-axis in degree/second
16	GyroY(degree/s)	Angular acceleration in the Y-axis in degree/second
17	GyroZ(degree/s)	Angular acceleration in the Z-axis in degree/second
18	MagX(mG)	Magnetic field in the X-axis in milliGauss
19	MagY(mG)	Magnetic field in the Y-axis in milliGauss

20	MagZ(mG)	Magnetic field in the Z-axis in milliGauss
21	Temp(C)	Temperature of the sensor

Table 4.2 Format of IMU data in the input files

4.2.1.3 UWB

Column number	Name	Description
1	Time	Time of the measurement. Time is given in the following format: hh:mm:ss.fff
2	Identifier	A number to identify the technology used. 2 is the identifier for UWB
3	Δ of Time	Time elapsed since the last data measurement in seconds.
4	Time (s)	Time in seconds transcurred since the beginning of the measurement.
5	X(m)	Position calculated by the UWB in ECEF (meters) for the X-axis.
6	Y(m)	Position calculated by the IMU in ECEF (meters) for the Y-axis.
7	Z(m)	Position calculated by the IMU in ECEF (meters) for the Z-axis.

8	StdX(m)	Standard Deviation of the position in the X-axis calculated from UWB measurements.
9	StdY(m)	Standard Deviation of the position in the Y-axis calculated from UWB measurements.
10	StdZ(m)	Standard Deviation of the position in the Z-axis calculated from UWB measurements.
11	Residual	Residual of the position calculated from UWB measurements.

Table 4.3 Format of UWB data in the input files

4.2.1.4 Ephemerides File

Column number	Name	Description
1	Time	Time of validation for ephemerides. Time is given in the following format: hh:mm:ss.mmm
2	Time(s)	Time in seconds transcurred since the first ephemerides received.
3	Svprn	Satellite number
4	Af2	Space vehicle clock drift rate correction coefficient
5	M0	Mean anomaly at the reference time

6	Roota	Square root of the semi-major axis
7	Deltan	Mean motion difference from the computed value
8	Ecc	Eccentricity
9	Omega	Argument of perigee
10	Cuc	Amplitude of the cosine harmonic correction term to the argument of latitude
11	Cus	Amplitude of the sine harmonic correction term to the argument of latitude
12	Crc	Amplitude of the cosine harmonic correction term to the orbit radius
13	Crs	Amplitude of the sine harmonic correction term to the orbit radius
14	I0	Inclination angle at the reference time
15	IDOT	Rate of the inclination angle
16	Cic	Amplitude of the cosine harmonic correction term to the angle of inclination
17	Cis	Amplitude of the sine harmonic correction term to the angle of inclination
18	Omega0	Longitude of ascending node of orbit plane at the weekly epoch

19	Omegadot	Rate of right ascension
20	Toe	Reference time ephemeris
21	Af0	SV Clock Bias Correction Coefficient
22	Af1	SV Clock Drift Correction Coefficient
23	Toc	Time of clock
24	IODE	Issue of data (Ephemeris)
25	Code_on_L2	Code on L2 frequency
26	Weekno	Week number
27	L2_flag	Flag for L2 frequency
28	Svaccur	Space vehicle accuracy
29	Svhealth	Space vehicle health
30	Tgd	Group delay differential
31	Fit_init	Fit interval flag

Table 4.4 Format of Ephemerides File data

4.2.2 Read Data

After checking the data received, is time to read and store it in different structures. First, the ephemeris file is read and the data for that timestamp is stored in a structure. Then, and only during the first iteration, the data from the KBS file is read and stored. As it is only a database and does not change during the time, reading it once is enough. More information about it is introduced in the KBS section (Subsection 4.2.7)

Then, and every execution until the end of the file, data is collected from the input file, one line every single execution. The second column of the line obtained is

read in order to know the type of technology to which belongs the data. Then, taking into account the type of data it is stored in different structures to use the data in the subsequent steps.

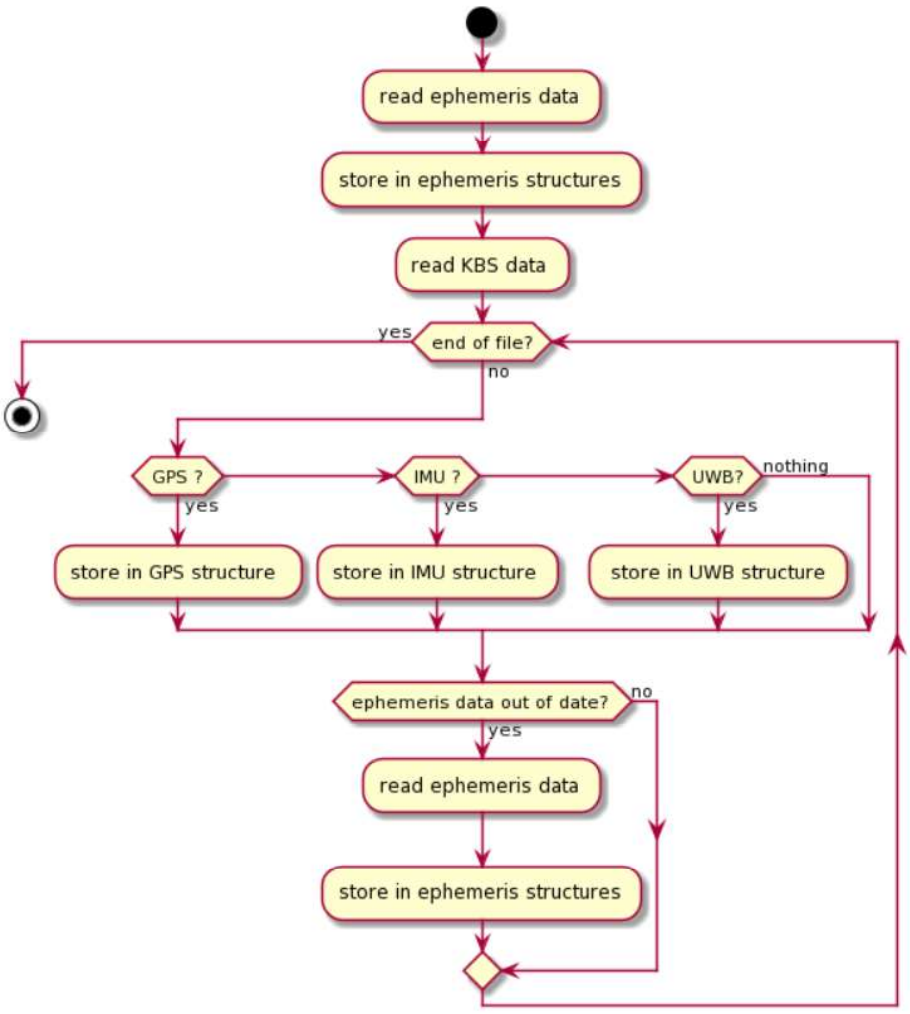


Figure 4.3 Overview of the read data functionality

Once the data available in each timestamp is stored in a structure. A different treatment is given to the data depending on its measuring technology.

4.2.3 GPS Data Treatment

One of the systems used for positioning is GPS. To obtain a GPS position estimation, many steps have to be made. The steps are explained in this section. The objective of this part of the code is to obtain a GPS preliminary position estimation (before performing the Kalman filter) and the necessary statistical to perform the Kalman filter positioning estimation.

A complete diagram of the main steps of the algorithm is presented in Figure 4.4. Then the most important parts are explained one by one.

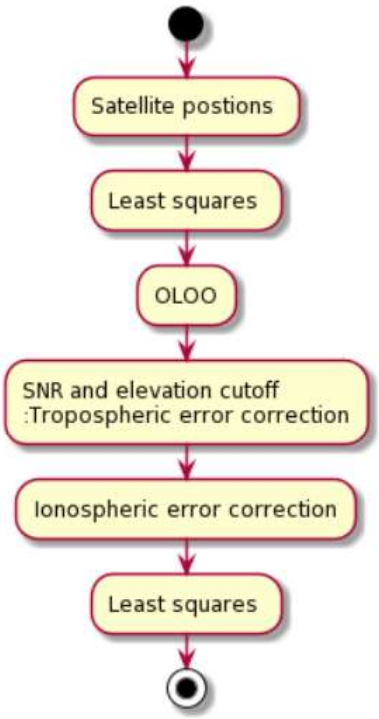


Figure 4.4 GPS data treatment overview flowchart

4.2.3.1 Satellite positions

To obtain a GPS position estimation is necessary to calculate the positions of the satellites to then estimate the position of the rover. First of all, if the signal transmission time is computed. For that, the travel time obtained from the pseudorange is subtracted to the reception time. To that transmission time obtained, different time corrections are made such as the relativistic error correction or the satellite clock error correction along with the Timing Group Delay (TGD). The relativistic error correction is an error that takes into account the eccentric anomaly of the orbit. The satellite clock error correction includes the satellite clock drifts and offsets received in the ephemeris, and the TGD includes delays caused by antennas, cables and/or filters in the satellite, and also is included in the ephemeris data.

Then the following calculus is performed in order to obtain the position of the satellites (Managers n.d.).

Computed mean motion:

$$n_0 = \sqrt{\frac{GM}{A^3}} \tag{4.1}$$

Where *GM* is the WGS-84 value for the product of gravitational constant *G* and the mass of the Earth *M* and *A* is the semi-major axis of the orbit.

Corrected mean motion:

$$n = n_0 + \Delta n \tag{4.2}$$

Where *Δn* is the mean motion difference from the computed value.

Time elapsed since the reference epoch:

$$t_k = t - t_{0e} \tag{4.3}$$

Where t_{0e} is the reference time for the ephemeris parameters

Mean anomaly:

$$M_k = M_0 - nt_k \tag{4.4}$$

Where M_0 is the mean anomaly at the reference time.

Then, Kepler's equation of eccentric anomaly is solved by iteration:

$$E_k = M_k + e \sin E_k \tag{4.5}$$

Where e is the eccentricity.

True anomaly:

$$\cos v_k = \frac{\cos E_k - e}{1 - e \cos E_k} \tag{4.6}$$

The argument of latitude:

$$\Phi_k = v_k + \omega \tag{4.7}$$

Where ω is the argument of the perigee.

The argument of latitude correction:

$$\delta u_k = C_{uc} \cos 2\Phi_k + C_{us} \sin 2\Phi_k \tag{4.8}$$

Where C_{uc} is the amplitude of the cosine harmonic correction term to the argument of latitude and C_{us} is the amplitude of the sine harmonic correction term to the argument of latitude.

Radius correction:

$$\delta r_k = C_{rc} \cos 2\Phi_k + C_{rs} \sin 2\Phi_k \quad (4.9)$$

Where C_{rc} is the amplitude of the cosine harmonic correction term to the orbit radius and C_{rs} is the amplitude of the sine harmonic correction term to the orbit radius.

Inclination correction:

$$\delta i_k = C_{ic} \cos 2\Phi_k + C_{is} \sin 2\Phi_k \quad (4.10)$$

Where C_{ic} is the amplitude of the cosine harmonic correction term to the angle of inclination, and C_{is} is the amplitude of the sine harmonic correction term to the angle of inclination.

Corrected argument of latitude:

$$u_k = \Phi_k + \delta u_k \quad (4.11)$$

Corrected radius:

$$r_k = A(1 - e \cos E_k) + \delta r_k \quad (4.12)$$

Corrected inclination:

$$i_k = i_0 + it_k + \delta i_k \tag{4.13}$$

Where i_0 is the inclination angle at reference time and i is the rate of change of inclination.

Position in the orbital plane:

$$X'_k = r_k \cos u_k \tag{4.14}$$

$$Y'_k = r_k \sin u_k \tag{4.15}$$

Corrected longitude of ascending node:

$$\Omega_k = \Omega_0 + (\dot{\Omega} - \omega_e)t_k - \omega_e t_{0e} \tag{4.16}$$

Where Ω_0 is the longitude of the ascending node at reference time, $\dot{\Omega}$ is the rate of change of right ascension and ω_e is the WGS-84 value of the Earth's rotation rate.

Earth-fixed geocentric satellite coordinate:

$$X_k = X'_k \cos \Omega_k - Y'_k \sin \Omega_k \cos i_k \tag{4.17}$$

$$Y_k = X'_k \sin \Omega_k + Y'_k \cos \Omega_k \cos i_k \tag{4.18}$$

$$Z_k = Y'_k \sin i_k \tag{4.19}$$

Once the satellite positions have been obtained, the earth rotation correction is applied in order to have the definitive satellite positions with which is going to calculate the position estimation. The whole flowchart can be seen in Figure 4.5.

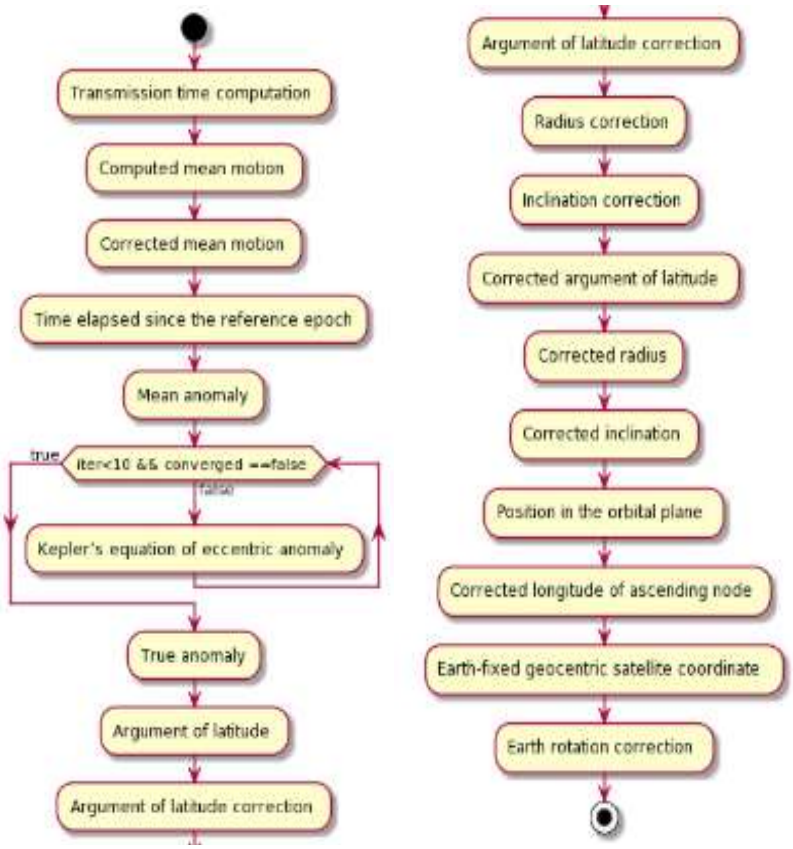


Figure 4.5 Overview of Satellite Position computation flowchart

4.2.3.2 Least-Squares

After obtaining the satellite positions, a preliminary GPS position estimation is done using a least-squares approximation. That part of the algorithm is explained in this section.

The relation between the user's position and the satellites information is represented by the following equation:

$$\rho_i = \|s_i + u\| + ct_u \tag{4.20}$$

where ρ_i is the pseudorange for the satellite i, the satellite position is represented by s_i, the user's position by u, t_u represents the advance of the receiver clock concerning the system time and *c* the speed of the light.

Equation 4.20 can be expanded and expressed by the following equation:

$$\rho_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} + ct_u \tag{4.21}$$

where x_i, y_i and z_i denotes the position of the satellite and x_u, y_u and z_u denotes the user's position.

As can be seen, the obtained equation is not linear. Since the WLS method is used to solve linear equations, the equation must be linearized. For that, an approximate position of the user $(\hat{x}_u, \hat{y}_u, \hat{z}_u)$ is used and a receiver clock offset $\hat{t}_u$ is considered to obtain an approximated pseudorange.

$$\hat{\rho}_i = \sqrt{(x_i - \hat{x}_u)^2 + (y_i - \hat{y}_u)^2 + (z_i - \hat{z}_u)^2} + c\hat{t}_u \tag{4.22}$$

Based on this approach, all the presented unknowns in equation 4.21 can be denoted as:

$$\begin{aligned} x_u &= \hat{x}_u + \Delta x_u \\ y_u &= \hat{y}_u + \Delta y_u \\ z_u &= \hat{z}_u + \Delta z_u \\ t_u &= \hat{t}_u + \Delta t_u \end{aligned} \tag{4.23}$$

Where $\Delta x_u, \Delta y_u, \Delta z_u \; and \; \Delta t_u$ are the difference between the approximate position and time to the PVT solution to be obtained.

Using Taylor series and truncating after the first-order partial derivative to eliminate nonlinear terms the following equation is obtained:

$$\rho_i = \hat{\rho}_i - \frac{x_i - \hat{x}_u}{\hat{r}_i}\Delta x_u - \frac{y_i - \hat{y}_u}{\hat{r}_i}\Delta y_u - \frac{z_i - \hat{z}_u}{\hat{r}_i}\Delta z_u + c\Delta t_u \quad (4.24)$$

where $\hat{r}_i$ is the following:

$$\hat{r}_i = \sqrt{(x_i - \hat{x}_u)^2 + (y_i - \hat{y}_u)^2 + (z_i - \hat{z}_u)^2} \quad (4.25)$$

The system that is now linearized based on the $\Delta x_u, \Delta y_u, \Delta z_u$ and Δt_u can be re-arranged expressed as the following system of linear equations:

$$\hat{\rho}_i - \rho_i = \frac{x_i - \hat{x}_u}{\hat{r}_i}\Delta x_u + \frac{y_i - \hat{y}_u}{\hat{r}_i}\Delta y_u + \frac{z_i - \hat{z}_u}{\hat{r}_i}\Delta z_u - c\Delta t_u$$

$$Ax = bA$$

$$= \begin{bmatrix} \frac{x_1 - \hat{x}_u}{\hat{r}_1} & \frac{y_1 - \hat{y}_u}{\hat{r}_1} & \frac{z_1 - \hat{z}_u}{\hat{r}_1} & 1 \\ \frac{x_2 - \hat{x}_u}{\hat{r}_2} & \frac{y_2 - \hat{y}_u}{\hat{r}_2} & \frac{z_2 - \hat{z}_u}{\hat{r}_2} & 1 \\ \vdots & \vdots & \vdots & 1 \\ \frac{x_n - \hat{x}_u}{\hat{r}_n} & \frac{z - \hat{z}_u}{\hat{r}_n} & \frac{z_n - \hat{z}_u}{\hat{r}_n} & 1 \end{bmatrix} \quad x = \begin{bmatrix} \Delta x_u \\ \Delta y_u \\ \Delta z_u \\ -c\Delta t_u \end{bmatrix} \quad b = \begin{bmatrix} \hat{\rho}_1 - \rho_1 \\ \hat{\rho}_2 - \rho_2 \\ \vdots \\ \hat{\rho}_n - \rho_n \end{bmatrix} \quad (4.26)$$

Then, this linear equation system is solved in order to minimize the mean square error (MSE).

The least-squares method aims to produce an approximation for an overdetermined linear system of equations. This means having more equations than unknowns, in which the elements are perturbed by errors, where it does not exist a unique solution that fulfil all the equations. Least-square method solve algebraic problems defined as the following:

$$a_{11}x_1 + \cdots + a_{1n}x_n = b_1; \ldots; a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \quad (4.27)$$

$$Ax = b$$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}_{nxn} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}_{nx1} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}_{nx1}$$

As mentioned, the system is perturbed by errors, thus the equation 4.27 can be denoted either way by:

$$Ax - b = \varepsilon \text{ or } \varepsilon = b - Ax \tag{4.28}$$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}_{nxn} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}_{nx1} - \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}_{nx1} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{nx1}$$

where ε, states the residuals of the error.

What is intended by the least-square method is to minimize the means square error (MSE).

$$MSE(\varepsilon) = \sum_{i=1}^{n} (Ax_i - b_i)^2 \tag{4.29}$$

$$\varepsilon^T \varepsilon = (Ax - b)^T (Ax - b)$$

$$\varepsilon^T \varepsilon = (x^T A^T - b^T)(Ax - b)$$

$$\varepsilon^T \varepsilon = (x^T A^T - b^T)(Ax - b)$$

$$\varepsilon^T \varepsilon = b^T b - b^T Ax + x^T A^T Ax - x^T A^T b$$

where *T* denotes the transpose of the matrix and *n* the number of equations.

Note that b and x are column vectors of dimension *mx1*. This means that $b^T Ax = x^T A^T b$ are equal and symmetric *(1x1).* Thus, equation 4.29 can be denoted as follows:

$$\varepsilon^T \varepsilon = b^T b + x^T A^T Ax - 2x^T A^T b \tag{4.30}$$

As mentioned, the least-square method aims to minimize the sum of the squared residuals. This means to get the derivate of the sum of squares residual based on the unknowns x, in this case $\frac{\partial[\varepsilon^T \varepsilon]}{\partial x}$*=0,* as follows:

$$\frac{\partial[\varepsilon^T\varepsilon]}{\partial x} = 2A^TAx - 2A^Tb$$

$$\frac{\partial[\varepsilon^T\varepsilon]}{\partial x} = A^TAx - A^Tb \quad (4.31)$$

$$A^TAx = A^Tb$$

Just by pre-multiplying by the inverse of $(A^TA)^{-1}$, the equations will be solved as denoted as follows:

$$x = (A^TA)^{-1}A^Tb \quad (4.32)$$

In this method, all the residuals have the same weight when obtaining the MSE of the residuals.

There are circumstances where the cost function can be modified by other parameters providing more importance to certain conditions (this means higher weights for the residuals). Let's use equation 4.29, including different weights (w_i) to each of the residuals.

$$MSE(\varepsilon) = \sum_{i=1}^{n} w_i(Ax_i - b_i)^2 \quad (4.33)$$

This can be solved using linear algebra by expressing the different w_i, as a diagonal matrix and zeros on the rest of the matrix. The same way the aim is to express the errors as a sum of square errors to minimize this cost function.

$$\varepsilon^TW\varepsilon = (Ax - b)^TW(Ax - b) \quad (4.34)$$

$$\varepsilon^TW\varepsilon = (x^TA^T - b^T)W(Ax - b)$$

$$\varepsilon^TW\varepsilon = (x^TA^T - b^T)W(Ax - b)$$

$$\varepsilon^TW\varepsilon = b^TWb - b^TWAx + x^TA^TWAx - x^TA^TWb$$

As mentioned before, the aim is to minimize the sum of the squared residuals. This means to get the derivate of the sum of squares residual based on the unknowns x, in this case $\frac{\partial[\varepsilon^T W\varepsilon]}{\partial x}$*=0,* as follows:

$$\frac{\partial[\varepsilon^T W\varepsilon]}{\partial x} = 2A^T WAx - 2A^T Wb \tag{4.35}$$

$$\frac{\partial[\varepsilon^T W\varepsilon]}{\partial x} = A^T WAx - A^T Wb$$

$$A^T WAx = A^T Wb$$

Just by pre-multiplying by the inverse of $(A^T WA)^{-1}$, the equations will be solved as denoted as follows:

$$x = (A^T WA)^{-1} A^T Wb \tag{4.36}$$

As it can be seen from equation 4.36, the LS square method is a WLS particularized when the weight matrix W is the identity matrix. In the case of the presented work, the LS method has been weighted using the SNR and elevation of the satellites.

The algorithm can be also expressed as an activity diagram (see Figure 4.6):

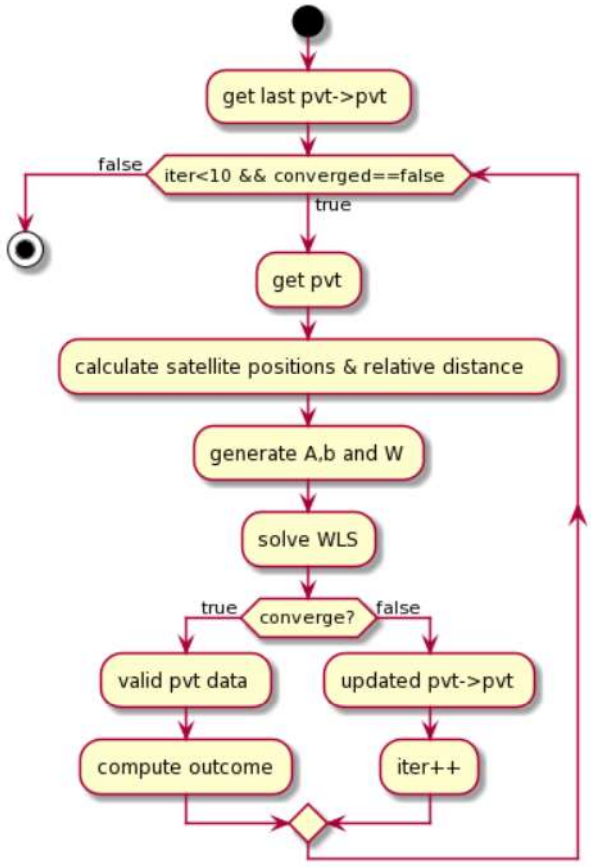


Figure 4.6 Least-squares flowchart

4.2.3.3 OLOO

After the Least-squares solution, when using GPS, also a fault detection and exclusion method has been introduced, the Only Leave One Out (OLOO) method.

It is the only method that assumes one single faulty satellite and cannot be extended to multiple failure cases. During the execution of the test, the maximum value of the test statistic is found and compared with a user-defined threshold value. In case the statistic exceeds the threshold, a rejection of this faulty observation is

done. In any case (exclusion or not), the solution is considered to be reliable, as a maximum of one fault is considered. This method has been chosen as it provides 100% availability; it can be performed every time a GPS measurement has been received.

The flowchart of the part of the algorithm including this method can be found in Figure 4.7.

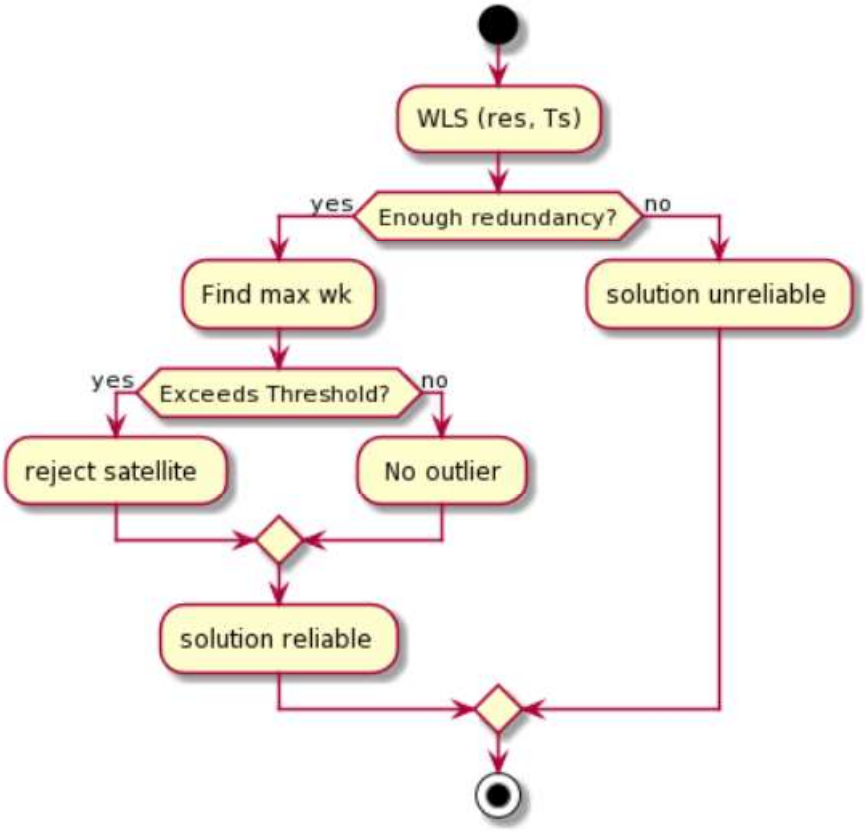


Figure 4.7 Flowchart of the Only Leave One Out method

4.2.3.4 SNR and elevation cutoff

After performing the OLOO method, satellites that are not complying with the SNR and elevation masks imposed are also discarded. For that, the topocentric coordinates of the satellites are calculated in order to obtain their elevation. Then the satellites that do not have an elevation higher than 10 degrees and an SNR higher than 25 dB are discarded. The discrimination of the satellites can be done using only one of the masks, SNR or elevation, or using both. This is chosen depending on the configuration of the algorithm in the same way that the WLS algorithm can be weighted with the elevation, the SNR or with both at the same time. The flowchart of this part of the algorithm can be seen in Figure 4.8.

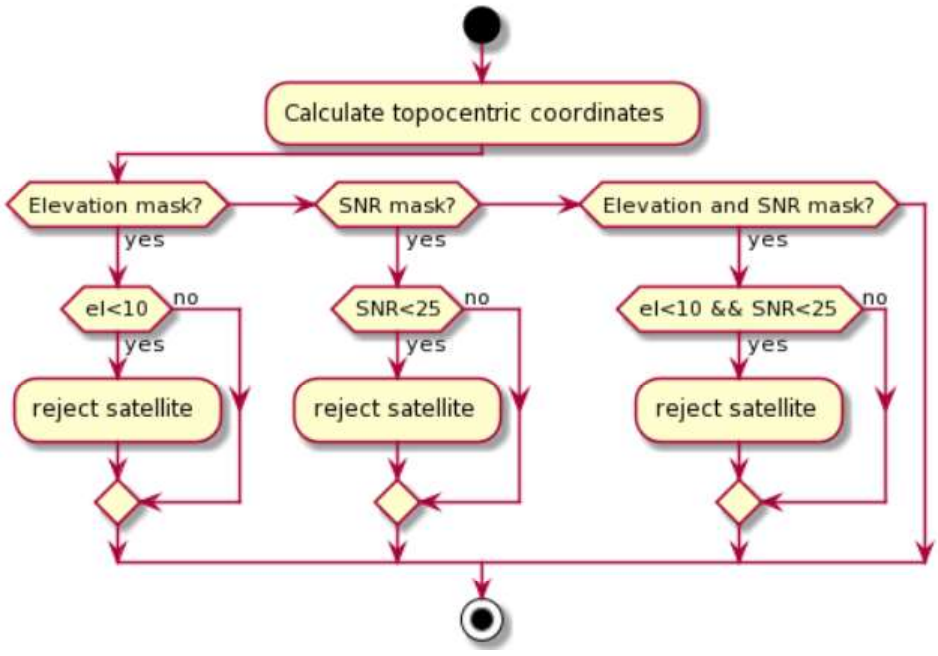


Figure 4.8 Flowchart of the SNR and Elevation cutoff

4.2.3.5 Tropospheric error correction

Afterwards, the pseudorange correction is done to mitigate the errors created by the tropospheric refraction. For that, the Saastamoinen algorithm is used. This algorithm is an algorithm reproducing an a priori tropospheric model, that requires some input parameters. Those inputs can be real meteorological observations, or as it happens, in this case, the parameters have been derived from a standard model of the atmosphere.

The Saastamoinen model (Saastamoinen 2013) assumes that the dry atmosphere is in hydrostatic equilibrium. And so its Zenith Hydrostatic delay in meters can be written as

$$\Delta\rho^2_{hydro} = (0.0022768 \pm 5 * 10^{-7}) \frac{P_0}{g(\phi, H)} \tag{4.37}$$

Where P_0 is the total ground pressure, and the function $g(\phi, H) = (1 - 0.00266\cos(2\varphi) - 2.8 * 10^{-7} * H)$ with φ the latitude and H the altitude.

This computed correction is directly made to the pseudoranges by subtracting its value in meters.

4.2.3.6 Ionospheric error correction

Then, another pseudorange correction is done to mitigate the ionospheric delay. For that, the Klobuchar model is used. GPS satellites end the parameters of the ionospheric model created by Klobuchard for single-frequency users. This model is estimated to reduce the 50% RMS ionospheric range error worldwide (Navipedia-Klobuchar n.d.). The correction is calculated by the following equations.

Calculate the earth-centered angle:

$$\psi = \frac{0.0137}{E + 0.11} - 0.022 \tag{4.38}$$

Where E is the elevation in semicircles.

Compute the latitude of the Ionospheric Pierce Point (IPP):

$$\phi_I = \phi_u + \psi \cos A \tag{4.39}$$

Where A is the azimuth angle and ϕ_u the geodetic latitude

Compute the longitude of the IPP:

$$\lambda_I = \lambda_u + \frac{\psi \sin A}{\cos \phi_I} \tag{4.40}$$

Where λ_u is the longitude

Find the geomagnetic latitude of the IPP:

$$\phi_m = \phi_I + 0.064 \cos(\lambda_I - 1.617) \tag{4.41}$$

Find the local time at the IPP:

$$t = 43200\lambda_I + t_{GPS} \tag{4.42}$$

Where t_{GPS} is the GPS time

Compute the amplitude of ionospheric delay:

$$A_I = \sum_{n=0}^{3} \alpha_n \phi_m^n \tag{4.43}$$

Where α_n are the coefficients broadcasted in the GPS message

Compute the period of ionospheric delay:

$$P_I = \sum_{n=0}^{3} \beta_n \phi_m^n \tag{4.44}$$

Where β_n are the coefficients broadcasted in the GPS message

Compute the phase of ionospheric delay:

$$X_I = \frac{2\pi(t - 50400)}{P_I} \tag{4.45}$$

Compute the slant factor (the line of sight distance along a *slant direction* between two points which are not at the same level relative to a specific datum):

$$F = 1 + 16(0.53 - E)^3 \tag{4.46}$$

Compute the ionospheric time delay:

$$I_{L1_{GPS}} = \begin{cases} \left[5 * 10^{-9} + \sum_{n=0}^{3} \alpha_n \phi_m^n * \left(1 - \frac{X_I^2}{2} + \frac{X_I^4}{24}\right)\right] * F \ ; \ |X_I| \leq 1.57 \\ 5 * 10^{-9} * F \ ; \ |X_I| \geq 1.57 \end{cases} \tag{4.47}$$

All this part of the algorithm can be seen in the flowchart depicted in Figure 4.9.

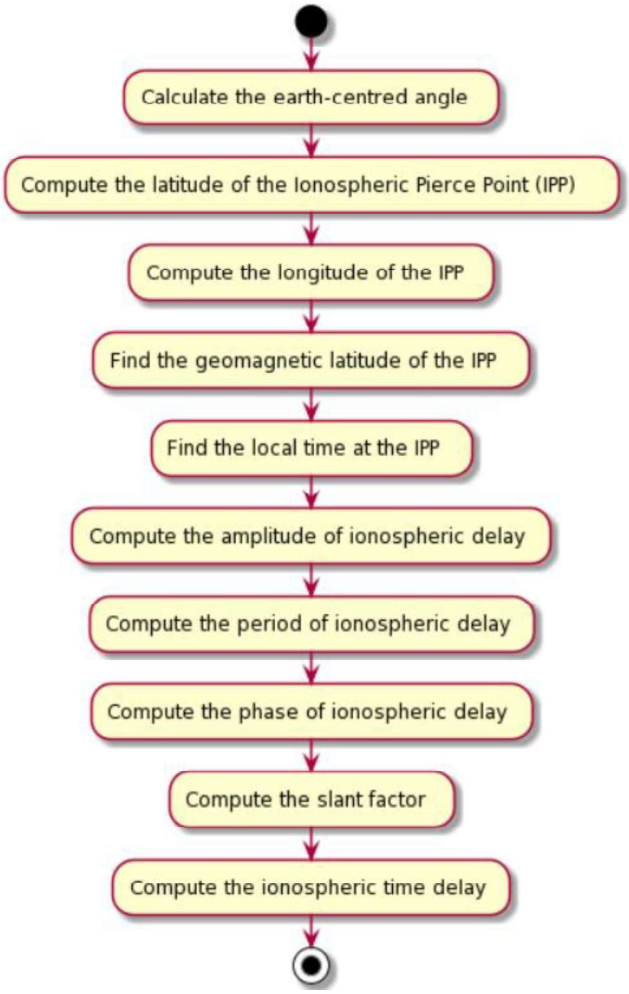


Figure 4.9 Flowchart of the Klobuchar model algorithm

To finish the subsection 4.2.3, the last part of the GPS data treatment, before performing the Kalman filter is introduced. After performing all the corrections explained, and taking as starting point the position computed in paragraph 4.2.3.2, the Least-squares algorithm is performed again with the satellites remaining after the masks applied. This position is the starting point for the Kalman filter of the fusion.

4.2.4 IMU Data Treatment

The second system used for positioning is the IMU. To obtain an IMU position estimation, mechanization is needed. The steps are explained in this section. The objective of this part of the code is to obtain an IMU preliminary position estimation (before performing the Kalman filter of the fusion) and the necessary statistical to perform the Kalman filter positioning estimation. In order to have a position in a global frame, a GPS position calculated before is needed as a reference for the relative movement.

First of all, it is checked if there is movement or not in order to calculate the bias error of the accelerometers and gyroscopes in the three-axis. If the data measured is below some threshold, that acceleration (linear or angular) is used to calculate the bias and then becomes equal to zero as will be explained practically in chapter 6. This step is done until the accelerations overcome the threshold or a set of thousand measurements is used to calculate the bias error.

When the movement starts, the bias error is applied to the accelerations obtained and the accelerations are changed from *'G'*s to m/s^2. Then, the yaw angle is calculated using the data obtained from the gyroscope in the Z-axis. The same is done to obtain the pitch and roll angles with the y-axis and x-axis respectively.

Once the angles have been calculated, the measured accelerations are changed from the body frame to the ENU coordinate system. With these accelerations, an integration scheme is performed in the ENU system.

The velocity in each of the axis is calculated in the following way (Woodman 2007):

$$v(t+\delta t) = v(t) + \delta t * (a(t+\delta t) - G) \tag{4.48}$$

Where $v(t+\delta t)$ is the velocity in the time instant $t+\delta t$, $v(t)$ is the velocity in the time instant t, δt the time elapsed since the last measurement, $a(t+\delta t)$ is the

acceleration in the instant $t + \delta t$, and G is the gravitational component in the axis in which the velocity is being calculated.

And the position in each of the axis is calculated in the following way (Woodman 2007):

$$s(t + \delta t) = s(t) + \delta t * v(t + \delta t) \quad (4.49)$$

Where $s(t + \delta t)$ is the position in the time instant $t + \delta t$, $s(t)$ is the position in the time instant t, δt the time elapsed since the last measurement, $v(t + \delta t)$ is the velocity in the instant $t + \delta t$ calculated before.

Then the position estimation obtained in ENU is changed to the ECEF coordinate system using the reference position. That reference position is the last position estimation made by the algorithm. Afterwards, the standard deviation of the position estimation obtained is calculated for each of the axes in ECEF coordinates from the data obtained from the datasheet of the IMU used for the measurements. Reached this point, all the necessary data to perform the Kalman filter has been calculated from the IMU side. Figure 4.10 shows a brief overview of the IMU data treatment presented in this section.

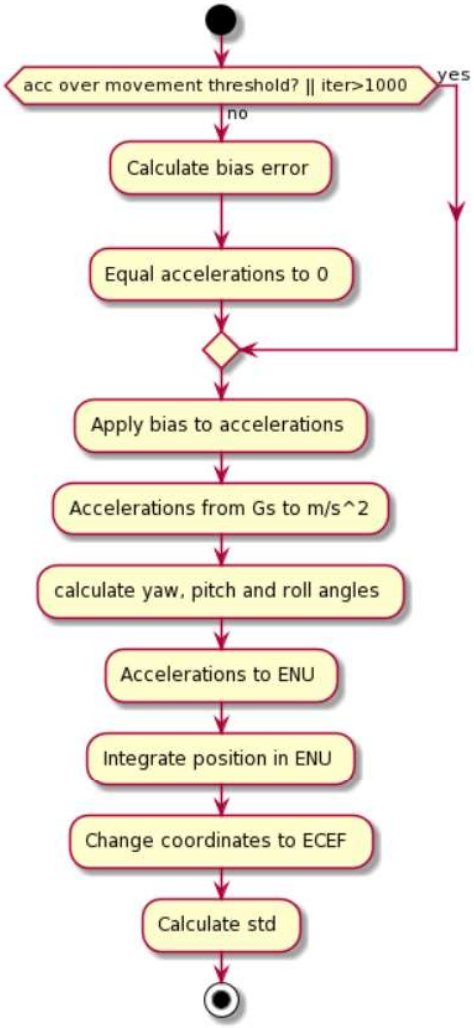


Figure 4.10 IMU data treatment flowchart

4.2.5 UWB Data Treatment

The last data used for positioning is the position and the errors estimation from the UWB system. To obtain these estimations, the ranges from the tag to the anchors are needed. Performing a least-squares algorithm as the one introduced in section

4.2.3.2 a position estimation and its residuals are obtained. In order to omit unnecessary repetitions, the least-squares algorithm is not repeated in this section.

4.2.6 Kalman Filter

After obtaining all the data needed to perform the Kalman filter from the three positioning systems used, the Kalman filter is introduced in this section. The Kalman filter presented is performed in two steps. First, the variables for the Kalman filter are initialized and the Kalman filter itself is executed.

4.2.6.1 Kalman Filter initialization

In the filter initialization, first, the dynamic model of the Kalman filter is initialized, also the same is done for the position and velocity equations of the system dynamics. Then the Kalman filter initial state is computed. The algorithm continues calculating the initial state covariance matrix and the dilution of precision. Afterwards, the model error for the covariance matrix is calculated and propagated in the global coordinates. The flowchart of this Kalman filter initialization can be seen in Figure 4.11.

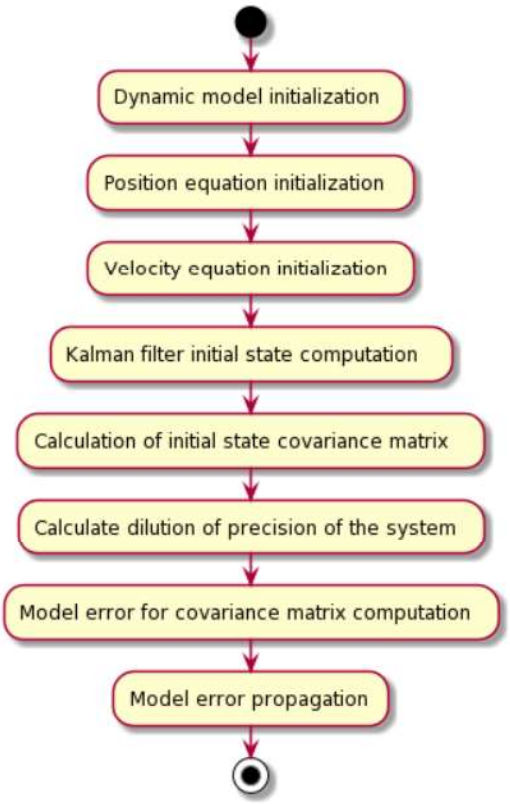


Figure 4.11 Kalman Filter Initialization flowchart

4.2.6.2 Kalman Filter execution

The next step is to perform the Kalman filter itself. The Kalman filter and its different integrations have been explained in Chapter 2. In this dissertation, semi-tightly integration has been done.

Innovation covariance is:

$$K = T * C_{ee} * T' + C_{vv} \tag{4.50}$$

Where T is the system dynamics matrix, C_{ee} is the state covariance matrix and C_{vv} is the covariance of the model error.

The gain matrix is:

$$G = K * H' * (H * K * H' + C_{nn})^{-1} \quad (4.51)$$

Where H is the observation model and C_{nn} is the receiver position error covariance matrix.

The Kalman state estimation is:

$$\widehat{X_{tt}} = (I - G * H) * X_{t1t} + G * y_0 \quad (4.52)$$

Where I is the identity matrix, X_{t1t} is the Kalman prediction made from the state before by means of propagation and y_0 is the measurement done by the technology in use.

The propagation for the next timestamp is calculated as follows:

$$X_{t1t} = T * \widehat{X_{tt}} \quad (4.53)$$

And the state covariance matrix update is:

$$C_{ee} = (I - G * H) * K \quad (4.54)$$

This Kalman filter estimation can be seen in the flowchart depicted in Figure 4.12.

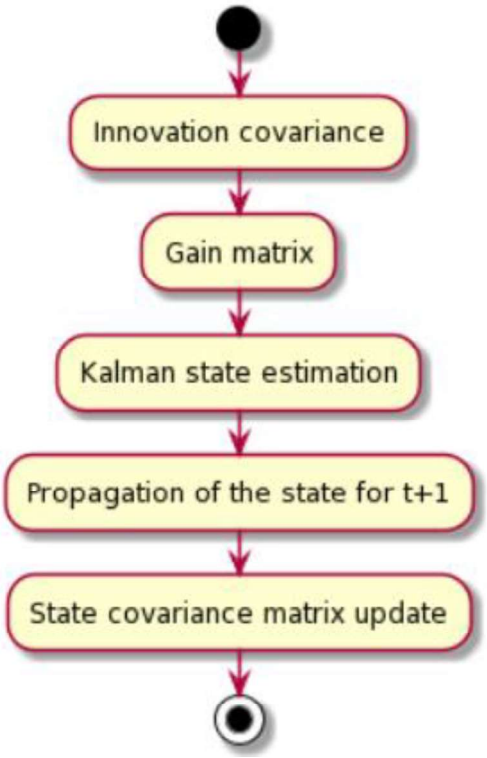


Figure 4.12 Kalman filter execution flowchart

4.2.7 Known Blocked Scenarios (KBS)

Known Blocked Scenarios (KBS) is a software enhancement that uses the knowledge of the zone in which the positioning system is working to forecast areas in which GNSS signals are not available or could give misleading information. With this map-based algorithm, a better choice of the available positioning technologies leads to an enhanced data management and by hence to a system with a better position estimation (de Miguel et al. 2019).

The KBS is a novel map aided positioning enhancement proposed by this dissertation. It helps the positioning algorithm in terms of anticipating blocking scenarios and minimizing the use of GPS misleading information.

Having a reliable heading and position of the train is important before entering a GPS blocked area because this is the absolute reference that is useful for the relative movement brought from the IMU. The orientation of the train is computed before entering blocked scenarios in order to perform an accurate positioning operation. With a non-redundant system, the heading angle is obtained from the last "a priori" reliable position of the positioning algorithm. Thus, detecting erroneous information improves the performance of the solution in GPS blocked areas.

Regarding the use of the KBS, two main steps are defined. The first step is the determination of known block scenarios for the determination of the KBS threshold and the second step is the KBS operation mode to enhance the position estimate in which real-time positioning operation is performed.

4.2.7.1 Determination of the GNSS blocked scenarios and KBS threshold determination

The KBS aims to improve the positioning and it is done under map located known blocked scenarios. Thus, for the KBS to be available, before starting its operation, it is necessary to first detect and create a database with these spots and sections with known blocked scenarios.

In this case, for the determination of a degraded GNSS area, the received signal to noise ratio of the GNSS signal strengths is analyzed. Having a lower SNR increases the probability of having an error in the received information (De Miguel et al. 2017).

Every scenario produces an attenuation based on the surrounding environment. So in the first phase, an analysis of this environment needs to be done. With that purpose, maps are consulted and the trackside is inspected to identify the possible GNSS blockers such as urban canyons, tunnels, woods, or other artificial blockers and constructions. A database with all the areas where possible GNSS outages are expected is constructed and proper position and heading values are introduced

based on reliable track database information. A good example of that kind of databases is b5m in Gipuzkoa where maps with high precision can be found (B5m 2020).

To reduce the computation time for the enhancement of the position estimate, the KBS is only triggered under certain conditions. Thus, based on the analysis of the SNR along with these known block scenarios campaign detection, the sensibility and the KBS threshold is calculated (Arrizabalaga et al. 2016). This threshold is highly dependent on the antenna positioning and the trackside environment. If the antenna has a full view of the sky, the threshold will be higher as the difference between blocked and non-blocked scenarios will be more elevated due to the good health of the received signal in non-blocked scenarios. This is why this dissertation proposes to perform trips at different times of the day, with different satellite constellations in view and different weather to have a reliable threshold configuration.

In a railway with very differentiated parts, in which the train has a line of sight with the satellites every moment apart from the blocked scenarios, two runs will be enough to build an accurate database as the satellite architecture and the surrounding environment will not have a big influence (Miguel et al. 2019). In the case in which the surroundings of the blocked scenarios are GNSS challenging areas, more runs are needed in order to differentiate between blocked and no-blocked areas as the SNR decreases are lower and the threshold must be calculated more carefully. Runs at different times of the day are recommended to evaluate the performance with different satellite architectures.

The sensibility of the KBS is triggered based on the SNR decreases. This means that an SNR fall down compared with the previous epoch determines that GNSS signals are not trustful anymore unless the SNR shortfall is recovered. This allows IMU to have a trustful position once GNSS signals are not available, which improves the position estimation results. The KBS system allows detecting the entrance of the tunnels or other blocking structures where the complementary

positioning sensors take the key role regarding the position estimation. Summing up, the KBS threshold is set-up based on the SNR decreases.

A threshold with a 10% probability of being triggered is chosen using a Cumulative Distribution Function (CDF).

4.2.7.2 KBS operation

Once the KBS threshold is properly selected, the positioning algorithm can enhance the positioning in these blocked scenarios. Figure 4.13 shows the KBS algorithm method which relies on having a more accurate starting point which leads to a more accurate position estimate.

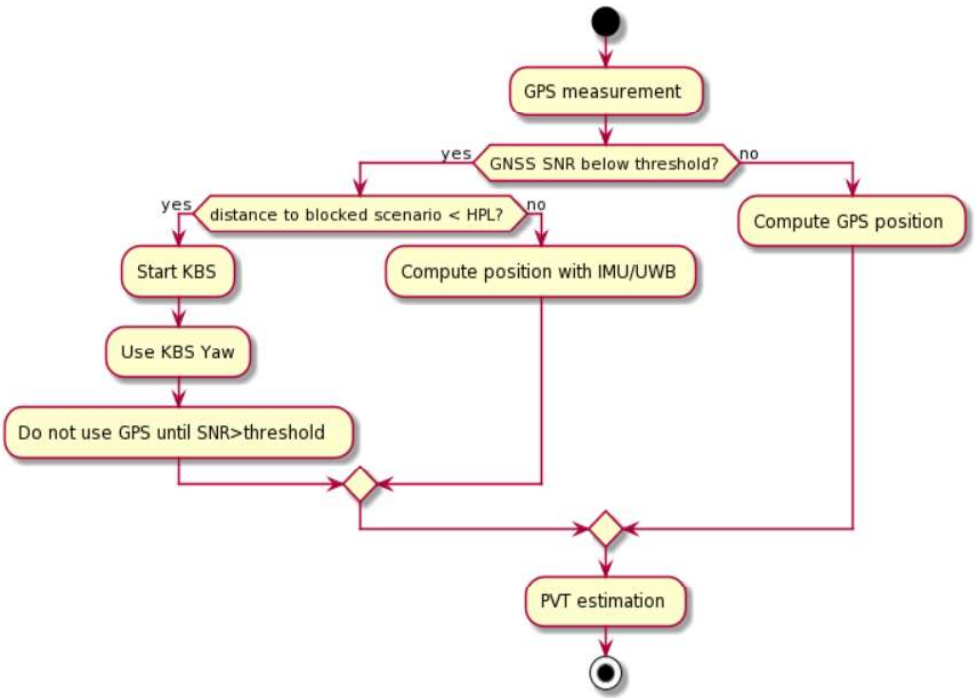


Figure 4.13 Known Blocked Scenarios flowchart

The KBS determines not the valid GNSS, but the degraded GNSS information based on the strong SNR decreases between epochs. The KBS takes advantage of the detection of strong blocking areas, which are mainly tunnels. Tunnels are the main blocking areas, due to their strong effect and their limitation to a defined area. At this point, apart from detecting and discarding the GNSS information a "database" is set up to reduce the errors of the next time epochs where GNSS is not available by resetting the yaw information of the algorithm.

The positioning system proceeds, in the same manner, every time that it receives GPS information. First of all, it compares if the selected triggering option is reached. If not, the GNSS is considered reliable and the algorithm computes the position estimate. If the KBS is triggered, the KBS takes a lookup into the KBS database, and based on the information of the previous reliable estimate and the distance to the different known blocked scenarios entries, it uses the stored information to correct it. The distance to use an entry from the database is, as maximum, the HPL. The HPL is the statistical error bound obtained during the real-time operation of the positioning algorithm.

This methodology is repeated until the KBS detects a valid GPS.

The format of the database mentioned can be seen in Table 4.5.

Column number	Name	Description
1	ID	Identification number for the Known Blocked Scenario
2	LAT	Latitude of the Known Blocked Scenario
3	LON	Longitude of the Known Blocked Scenario
4	ALT	Altitude of the Known Blocked Scenario

5	X_ECEF	Position of the Known Blocked Scenario in ECEF (meters) for the X-axis.
6	Y_ECEF	Position of the Known Blocked Scenario in ECEF (meters) for the Y-axis.
7	Z_ECEF	Position of the Known Blocked Scenario in ECEF (meters) for the Z-axis.
8	Flag	Flag that determines if the point is the entry or exit of a Known Blocked Scenario. 1 is used for entry and 0 for the exit.
9	Description	Description with extra information about the Known Blocked Scenario

Table 4.5 Format of the Known Blocked Scenarios database

4.2.8 Data Output

Finally, the position estimation obtained and other statistics related to it are written in an output file. That output file has a proprietary format created by CEIT and used for many projects. The format and the content of each of the columns are explained in Table 4.6.

Column number	Name	Description
1	UTC-Date	Date of the measurement. Date is given in the following format: yyyy/mm/dd
2	UTC-Time	Time of the measurement. Time is given in the following format: hh:mm:ss.fff/weekno/towinseconds
3	Algorithm ID	Shows the technology with which the position estimation has been done
4	SolStatus	0 for solution OK, 1 for the solution by system dynamics
5	EGNOSStatus	1 for EGNOS data used, 0 when not used
6	ECEFx	Position calculated by the positioning algorithm in ECEF (meters) for the X-axis.
7	ECEFy	Position calculated by the positioning algorithm in ECEF (meters) for the Y-axis.
8	ECEFz	Position calculated by the positioning algorithm in ECEF (meters) for the Z-axis.
9	ECEFvx	Velocity calculated by the positioning algorithm in ECEF (meters) for the X-axis.

10	ECEFvy	Velocity calculated by the positioning algorithm in ECEF (meters) for the Y-axis.
11	ECEFvz	Velocity calculated by the positioning algorithm in ECEF (meters) for the Z-axis.
12	HPL	Horizontal protection level calculated from the standard deviations obtained by the positioning algorithm using a 7-sigma approximation
13	NoSatUsed	Number of satellites used for the position computation
14	NoSatExclude	Number of satellites excluded from the position computation
15	HDOP	Horizontal Dilution of Precision
16	PDC	Reserved column
17	PDC	Reserved column
18	PDC	Reserved column
19	PDC	Reserved column
20	STD_x	Standard deviation calculated by the positioning algorithm in the X-axis
21	STD_y	Standard deviation calculated by the positioning algorithm in the Y-axis
22	STD_z	Standard deviation calculated by the positioning algorithm in the Z-axis

Table 4.6 Format of the Data Output file

Chapter 5

Measurement Campaigns for the System Validation and Performance Evaluation

This chapter describes the organization of the measurement campaigns that have been done in order to evaluate the performance of the proposed system. The chapter describes the different environments in which the system has been tested and the equipment used in each of them. In this way, a good characterization of the areas in which the algorithms are tested is introduced.

5.1 SCENARIO 1: EUSKOTREN LINE

The scenario presented in this section is a demanding environment near Donostia / San Sebastián. The chosen railway is placed between Hendaia and Donostia, in a regional train that puts together France and Spain operated by Eusko Trenbide Sarea (ETS). This railway has a difficult orography since GNSS blockers appear in more than 50% of the track analyzed. Most of the journey is done inside tunnels or canyons; both urban or natural (see Figure 5.1).

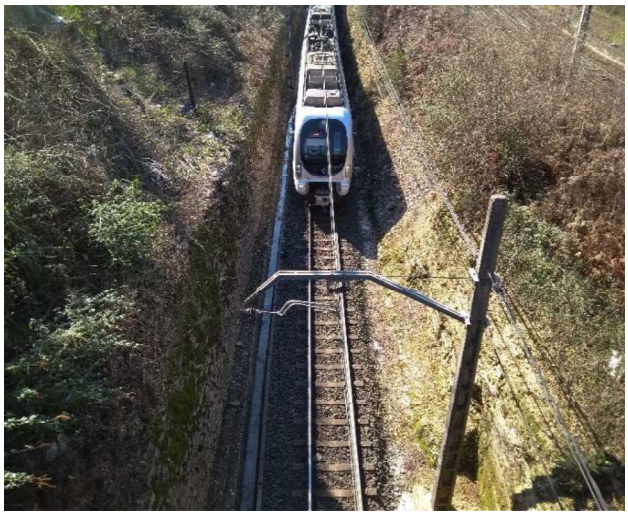


Figure 5.1 General environment in EuskoTren Lane

The trial track is mostly set in an urban environment inside the city of Donostia and the villages near it. However, there are also suburban and rural areas in more than 20km of track. The rural areas are mountainous areas in which the signal reception is not direct in all the cases. The environment in which the algorithm was tested can be considered a GNSS demanding zone in which long periods of lack of signal were recorded.

The tests took place in service for approximately 30 days, in which the full operation of the trains during their usual journeys were recorded. In this way, different satellite constellations in view were tested, and the collected data are meaningful.

The route going from Fanderia station to Errenteria station was chosen for the tests as it is a representative part of the journey because the track is surrounded by buildings and the train travels through a tunnel (see Figure 5.2 (green area)). The degradation of the signal before the entrance to the tunnel affects the performance of the positioning system. The surroundings and the fact that the tunnel is preceded by a curve also also has influence.

Figure 5.2 Section of Eusko Trenbide Sarea (ETS) network

The vehicle used was an electric locomotive Serie 900 from Euskotren. (See Figure 5.3).

Figure 5.3 EuskoTren Serie 900 Locomotive

5.1.1 Measuring Equipment

The hardware subsystems used in this test scenario are low-cost sensors since a good cost-effectiveness for an attractive business case was wanted. The GNSS receiver could be a low-medium end subsystem, which is not usual on the railway applications analyzed. In case a high-end GNSS system is mounted on board, a

better performance would be expected by the algorithm. Physically, the equipment used in the measurements done in the first scenario, includes a multisensor equipment for location estimation that consists of the subsystems presented in the following paragraphs.

5.1.1.1 GPS Receiver

U-Blox EVK-6T-0-001 is the GPS receiver used during the measurements. The two main objectives of the receiver are the following: to record the raw data and to perform a standalone positioning operation. This low-cost device allows GPS and QZSS signals, it has several interfaces such as UART USB or SPI and it is classified as professional-grade in the use in spite of its low-cost. It has raw pseudorange data output and an onboard RTC crystal to have faster warm and hot starts. The receiver claims to have accuracies of 2.5m in position what seems to be fair for this kind of single constellation single frequency low-cost system (U-blox 2015).

The GPS antenna used for this measurement was an u-Blox ANN-MS-0-005 antenna. It is an active antenna with a low noise figure and high gain coverage. Its integration is easy and a background in antenna mounting is not necessary to make it work properly (U-blox 2017).

5.1.1.2 Inertial Measurement Unit (IMU)

The Inertial Measurement Unit (IMU) used was the AIMS Navigation 0817111411 (AIMS n.d.). Its objective is to provide linear accelerations and angular velocities. The IMU employed in this case has 6 degrees of freedom, the manufacturer claims to have a high performance for autonomous and remote operate vehicle systems and it has multiple output interfaces such as CAN 2.0B, RS-232, RS-422 or RS-485. The accelerometer and gyro data provided is fully

temperature compensated. It has a high output rate that reaches the 200Hz. The different performances can be seen in Table 5.1 and Table 5.2.

Accelerometers Performance	
Range	+-2 g
Bias Error	0.6 mg
In-Run Bias Stability	8 μg
Scale Factor Error @ 1 g	0.25 mg
Non-linearity	0.5% of FS
Noise	0.3 mg RMS
Bandwidth	15 Hz
Misalignment	2 mrad

Table 5.1 Accelerometers Performance

Angular Rate (gyroscope) Performance	
Range	120 °/s
Bias Error	0.06 °/s
In-Run Bias Stability	5 °/h
Scale Factor Error	0.06%
Non-linearity	0.25% of FS
Noise	0.3 °/s RMS

Bandwidth	15 Hz
Misalignment	2 mrad

Table 5.2 Angular Rate (gyroscope) Performance

5.1.1.3 Microprocessor

The sensors used to obtain the data for positioning are connected to a microprocessor. In this case, a Raspberry Pi 3B model has been used. Within its specifications the most important ones are the following ones; 1.4 GHz 64-bit ARM 8 Cortex-A53 microprocessor, 1 GB SDRAM, wireless LAN and Bluetooth connectivity, 4 USB ports and 40 GPIO pins among others (Raspberry n.d.).

5.1.2 Ground Truth Generation Equipment

The ground truth for scenario 2 was generated by EuskoTren and Eusko Trenbide Sareak (ETS). It has been generated using the onboard odometer and the track maps in property of EuskoTren and Eusko Trenbide Sareak..

5.2 SCENARIO 2: NOTTINGHAM GEOSPATIAL INSTITUTE ROOF TRIAL SITE

The second scenario presented in this chapter is located in the Nottingham Geospatial Institute. To be more specific in the roof of the Nottingham Geospatial Institute building. Here there is a trial site in which controlled measurements can be done to check systems and/or validate. The environment is known, and so the possible problems can be foretold and known in advance.

Figure 5.4 Nottingham Geospatial Institute Roof Trial Site

The measurements in this environment were performed during August 2018. They were performed using an electrical train and a railway simulating the shape of an 8 (see Figure 5.4), but without crossing the rails. The train placed in the roof is a small train that models the behavior of a real one in terms of movement and dynamics. It travels through a track that models the vibrations that the train would suffer in a real one.

Figure 5.5 Antenna placement in the mock train

Figure 5.6 Nottingham Geospatial Institute Trial Site Top View

5.2.1 Measuring Equipment

The hardware subsystems used are low-cost sensors since a good cost-effectiveness for an attractive business case was wanted. However, in this case, for the obtaining of the ground truth high-end systems were used. Physically, the equipment used in the measurements done in the third scenario, includes a multisensor equipment for location estimation that consists of the following:

5.2.1.1 GPS Receiver

The GPS receiver is the same one that was used in the first scenarios. More information about it can be found in paragraph 5.1.1.1.

5.2.1.2 Inertial Measurement Unit (IMU)

The Inertial Measurement Unit (IMU) used was the Advanced Navigation Orientus (Navigation n.d.). Its objective was to provide linear accelerations and angular velocities. The IMU employed in this case has 9 degrees of freedom, the manufacturer claims to have a high versatibility. The accelerometer and gyro data provided is fully temperature compensated. It has a high output rate for this work that reaches the 1000Hz. The different performances can be seen in Table 5.3 and Table 5.4.

Accelerometers Performance	
Range	+-2 g
Initial Bias Error	<5 mg
In-Run Bias Stability	20 µg
Scale Factor Stability	<0.06%
Non-linearity	< 0.05%
Noise	100 µg/$\sqrt{HZ}$
Bandwidth	400 Hz
Misalignment	<0.05º

Table 5.3 Accelerometers Performance

Angular Rate (gyroscope) Performance	
Range	250 º/s
Initial Bias Error	<0.2º/s
In-Run Bias Stability	3 º/h
Scale Factor Error	0.05%
Non-linearity	0.05%
Noise	0.004 º/s $\sqrt{HZ}$
Bandwidth	400 Hz
Misalignment	<0.05º

Table 5.4 Angular Rate (gyroscope) Performance

5.2.1.3 Microprocessor

The microprocessor used for the measurements performed in this environment was not changed from the environments before. More information about its features can be found in paragraph 5.1.1.3.

5.2.1.4 Ultra Wide Band (UWB)

The Ultra Wide Band system used during the measurements in the second environment is a DecaWave TREK 1000 system. It is a Two-Way-Ranging Real-Time Location System, easy and quick to set up. It is the ideal system to build a proof of concept in hours (Decawave n.d.). As 4 boards are included and the navigation mode was used, 3 boards acted as anchors and the last one as a tag. Connected to the tag, the necessary information of the ranges from the tag to the anchors was stored. It has centimeter-level accuracy, it is reliable against multipath and

interferences, and has low latency, usually faster than GNSS systems. Moreover, it has a long-range while having a low energy consumption and it is a low-cost system.

5.2.1.5 Measuring Equipment installation

Two GNSS antennas were installed on the train (see Figure 5.5) to obtain GNSS measurements and a Ground truth using a high-end GNSS receiver and an RTK rover-base system. Both antennas were installed maintaining the same reference position.

The equipment was placed inside the train (see Figure 5.7). Both IMU, GNSS receivers and the processing unit were placed there.

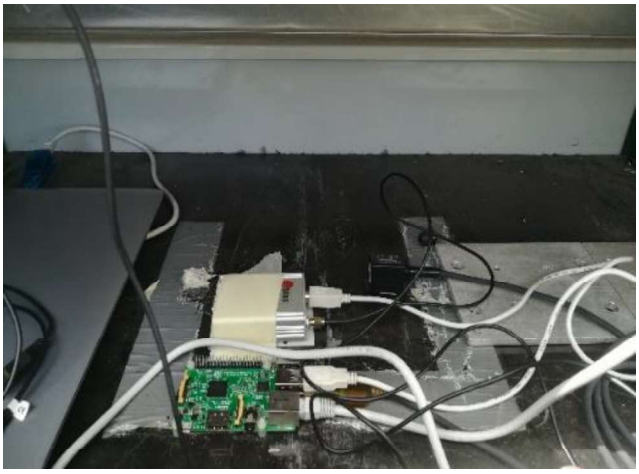


Figure 5.7 Equipment placement in the mock train

For Ultra Wide Band (UWB), the tag for the rover was placed in the roof of the train (see Figure 5.8). Moreover, the UWB anchors were also placed in the limits of the roof, to cover all the areas in which the train can travel.

Once the anchors were placed (see Figure 5.9), their exact points were measured using a Leica Nova MS60 MultiStation. This kind of instrument has a high precision that ensures the correct measurement of the points in which the anchors were placed, and so no extra errors are introduced in the UWB measurements derived from the placement of the anchors. In this case, the position of the anchors was measured with errors below the centimeters.

Figure 5.8 UWB Antenna placement

The environment is not a harsh environment for GNSS or UWB. However, it has a couple of areas in which the signals could suffer multipath effects due to big metal masses (see Figure 5.6). The south limit of the railway is near to the ventilation ducts. These ducts are over 2m in height and the train is not bigger than 1m tall. This fact can block the signals of some of the satellites and create multipath in others. A smaller mass of metal appears in the south wall in which some of the machines used for the air conditioning appear. They are not as tall as the ones on the other side, and they do not cover the full wall

Figure 5.9 UWB anchor placement example

from side to side but could generate multipath in some signals when the train travels near them. They are placed in the middle of the eight when both parts of the railway appear one near the other.

5.2.2 Ground Truth generation equipment

The ground truth for environment 2 was generated by the Nottingham Geospatial Institute (NGI). The ground truth was generated using a Leica Viva GS10 GNSS receiver placed in the same mock train that the measuring system, along with an RTK base station place in the trial site. The antenna provided by the Leica Viva measuring system was used and placed in the same place in which the antenna of the measuring system was. RTK measurements with corrections were done to obtain the Ground truth.

When doing UWB measurements, it is of high importance if you want to do absolute positioning, to know precisely the position of the anchors placed. For that purpose, a Leica Nova MS60 has been used (Leica n.d.). This is an instrument, mainly used by topographers that can give exact positions of places or objects with accuracies of millimeters. With this instrument along with a known position in the trial site, the exact positions of the anchors have been calculated.

5.3 SCENARIO COMPARISON

Finally, a table comparing the introduced scenarios, the hardware used and the complexity of the scenarios in terms of accuracy in positioning systems is presented.

	Name	Complexity	Ground Truth
Scenario 1	ETS Trial Site	High	Odometer
Scenario 2	NGI Roof Trial Site	Medium	RTK measurements with high end GNSS

Table 5.5 Scenario Comparison

	GNSS	IMU	UWB	Microprocessor
Scenario 1	U-blox NEO 6T	AIMS Navigation	Not included	Raspberry Pi 3 B+
Scenario 2	U-blox NEO 6T	Advanced Navigation Orientus	Decawave EVK-1000	Raspberry Pi 3 B+

Table 5.6 Scenario Comparison

5.4 TESTING METHODOLOGY

This section describes the methodology employed during the tests. The methodology includes the identification of tests to be carried on and the definition of a strategy to name the list of tests to be performed.

5.4.1 Tests performed

This section describes the different tests that will be performed by the created algorithm inside the system, following the next testing procedure:

1. The algorithm runs each test with the different systems selected for the test case.
2. The obtained result is compared with the ground truth.
3. Differences are analyzed in order to face conclusions and validate the algorithm.

5.4.2 Test template

A common template for naming the different tests to be performed is defined. The description of all the test cases is based on this template.

TS_SCENARIO_ZZZZ	Test code. TS represents the Test scenario used from the ones introduced in this chapter. SCENARIO represents the short name of the scenario (EUS for EuskoTren and NOT for Nottingham) and ZZZZ represents the technologies used in a binary way in the following order; GNSS, IMU, UWB, KBS (1 if it is used and 0 if not).

Table 5.7 Description of the test cases' naming template

5.5 TESTING PLAN FOR REFERENCE JOURNEYS

This section describes the testing plan by introducing the test template employed to describe the tests, the test list, and the definition of the tests.

5.5.1 Test list

The test cases that have been identified in order to test the performance of the positioning algorithm created appear in Table 5.8.

Test Code	Scenario	GNSS	IMU	UWB	KBS
TS1_EUS_1100	EuskoTren	✓	✓		
TS1_EUS_1101	EuskoTren	✓	✓		✓
TS2_NOT_1000	Nottingham	✓			
TS2_NOT_0100	Nottingham		✓		
TS2_NOT_0010	Nottingham			✓	
TS2_NOT_1100	Nottingham	✓	✓		
TS2_NOT_1101	Nottingham	✓	✓		✓
TS2_NOT_1010	Nottingham	✓		✓	
TS2_NOT_1011	Nottingham	✓		✓	✓
TS2_NOT_0110	Nottingham		✓	✓	
TS2_NOT_0111	Nottingham		✓	✓	✓
TS2_NOT_1110	Nottingham	✓	✓	✓	
TS2_NOT_1111	Nottingham	✓	✓	✓	✓

Table 5.8 List of test cases

5.5.2 Test definitions

In this subsection, the definition of the test methodology is going to be detailed. Only one dataset is shown for each test in order not to lengthen the chapter. However, more datasets have been tested and the results led to the same conclusions. For this reason, the dataset shown presents the performance of the algorithm in a representative way.

These tests are oriented to evaluate the performance of the algorithm in the scenarios presented before. For this, an embedded system, GNSS Pseudorange, IMU and UWB measurement IF files are needed along with the GNSS Ephemeris measurement file. To perform the test, the following steps have to be taken:

1) Format files to the latest version using EditIMUandGPSFormat.m
2) Perform UWB LS positioning (In the case LS measurements are available).
3) Switch on Raspberry Pi
4) Send files to Raspberry.
5) Give permissions to the uploaded files.
6) Execute test.
7) Download obtained OF files from Raspberry Pi.
8) Compare obtained OF file with the Ground Truth.

Chapter 6

Input Data Analysis

This chapter includes the data analysis of the input data obtained from the measurement campaigns done and to be used in the algorithms. The objective is to check that the data obtained are coherent and to show how each of the modules introduced in the algorithm affects the resulting positioning function.

6.1 INTRODUCTION

This chapter shows the data analysis done to validate the data measured in the different scenarios. To fulfil that verification, data obtained from GNSS, IMU and UWB sensors are analysed.

In Chapter 5, two different scenarios have been introduced. The methodology for data acquisition has been the same in the different scenarios and most of the subsystems have also been the same. The only change has been a change in the IMU sensor that does not affect as it behaves similarly as it will be explained in section 6.3. Taking into account these facts presented, the decision of presenting only one of the analysed scenarios has been taken. Because, the presentation of the analysis of all the data recorded in the different scenarios could lead this chapter to unnecessary repetition and tedious length as there has not been detected any relevant change due to the different runs.

The scenario chosen for the raw data analysis has been the scenario 2 located on the roof of the Nottingham Geospatial Institute. This scenario is the chosen one due to its controlled environment. It is a scenario in which the ground truth can be measured and the possible sources of multipath or GNSS misleading signals can be spottable. Moreover, the movement of the rover is performed also in a controlled way and has no interferences of possible train traffic as in other scenarios performed in real railways. Another important fact to take into account in the selection of the scenario is that Scenario 2 is the only one that includes GPS, IMU and UWB positioning systems that are the ones possible to use in the system.

The aim of this chapter is to verify the three different data sources that are used as inputs for the positioning algorithm. Moreover, this analysis shows that the data received is coherent. To fulfil this aim, different data such as the number of available satellites, the behaviour of the accelerations measured or residuals of the position estimation obtained by UWB are analysed.

6.2 GPS DATA ANALYSIS

First of all, data recorded with the GPS is going to be analysed. The GPS is the more obvious and affordable source when it comes to positioning and the better accuracy and availability is desirable from it. Different statistics of the data received are going to be analysed to assess the GPS data received and afterwards perform the positioning algorithm using them.

6.2.1 Available satellites

One of the statistics that are important to analyse the GPS suitability and quality is the number of satellites available. A satellite is available when data from that satellite has been received. It is well known that to obtain a position estimation in 3D, at least data from four satellites is necessary, however, a higher number of satellites is desirable to improve the algorithms. More satellites mean more information, and this is helpful to discard faulty ones and use more reliable data.

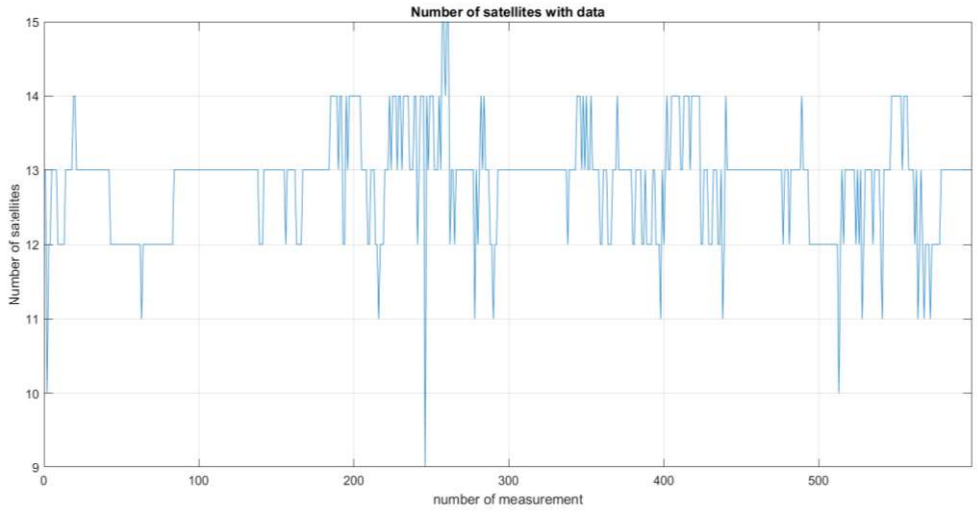


Figure 6.1 Number of Satellites received in each measurement for GPS

Figure 6.1 shows the number of satellites with data available received in each of the measurements computed in the test scenario considered for this chapter. There are no measurements with less than four satellites available. On the contrary, all the measurements have data from nine to fifteen satellites which shows that the scenario has good sky visibility. In conclusion, with this data, there should be no problems to perform the positioning algorithm and obtain a positioning estimation with GPS data in means of the number of available satellites.

6.2.2 Elevation of the satellites

Good availability of satellites does not ensure a good position estimation. It is important to see the elevation of those satellites in order to discard satellites. A satellite with a low elevation is going to introduce misleading data to the positioning algorithm, as there are higher possibilities of receiving its data in non-line-of-sight. An NLOS reception could imply multipath that affects the accuracy of the position estimation. That is the reason to analyse the availability of the satellites by means of their elevation. Reflections in the signal also impact the positioning accuracy and so a good elevation reduces the possibility of reflections in the way.

As a first step, the mean elevation of the satellites available in each of the measurements has been calculated (Figure 6.2). It shows that the values of elevation go from 22 to 31 degrees and its mean value is 27.31 degrees. This is acceptable as the typical elevation masks used for GPS satellites stand between 10 and 15 degrees (Lee n.d.).

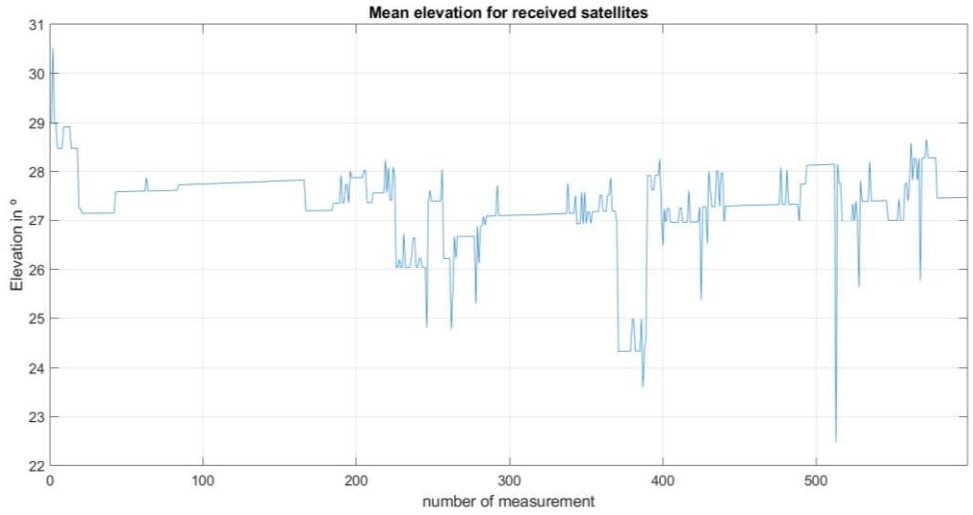


Figure 6.2 Mean elevation for the satellites with data available

However, what seems to be a good elevation, when analysing the data more in deep, can turn into some satellites with a bad elevation that can provide misleading information to the positioning algorithm. As GPS is the basic global positioning system used in the positioning algorithm, a conservative elevation mask of 10 degrees has been used. Therefore, the number of satellites not fulfilling the mask for each measurement can be seen in Figure 6.3. It shows a maximum of 6 and a minimum of 0 satellites with bad elevation can be found. A mean of 4.19 satellites is discarded every measurement.

To see the impact of applying an elevation mask, Figure 6.4, shows the satellites able to be used in the positioning algorithm after performing the elevation mask. It shows that now there are measurements with just 5 available satellites only one more than the minimum necessary to perform the positioning estimation with the GPS. On the other hand, we can see measurements with a maximum of 10 satellites and that usually the data in the scenario analysed has between 7 and 9 satellites. To be more precise, a mean of 8.66 satellites is used in each measurement. This

number of available satellites is enough to perform the positioning estimation but shows a reality with less satellites worse than in subsection 6.2.1.

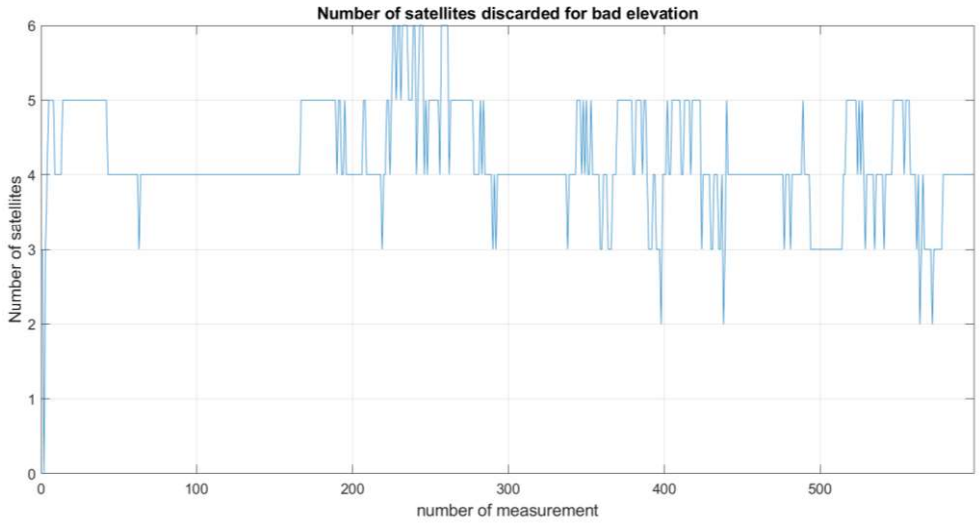


Figure 6.3 Number of satellites discarded due to bad elevation

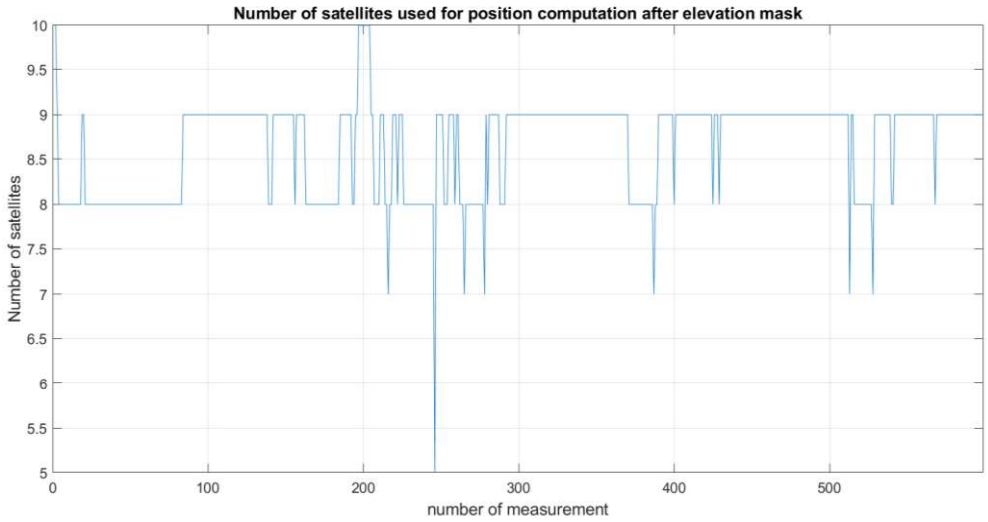


Figure 6.4 Number of satellites used for positioning after performing the elevation mask

Finally, the mean elevation of the satellites used has been analysed in Figure 6.5. In comparison with Figure 6.2, a more constant elevation is depicted. Moreover, the elevation is higher due to the elimination of the satellites with bad elevation, and its mean value has increased to 36.74 degrees.

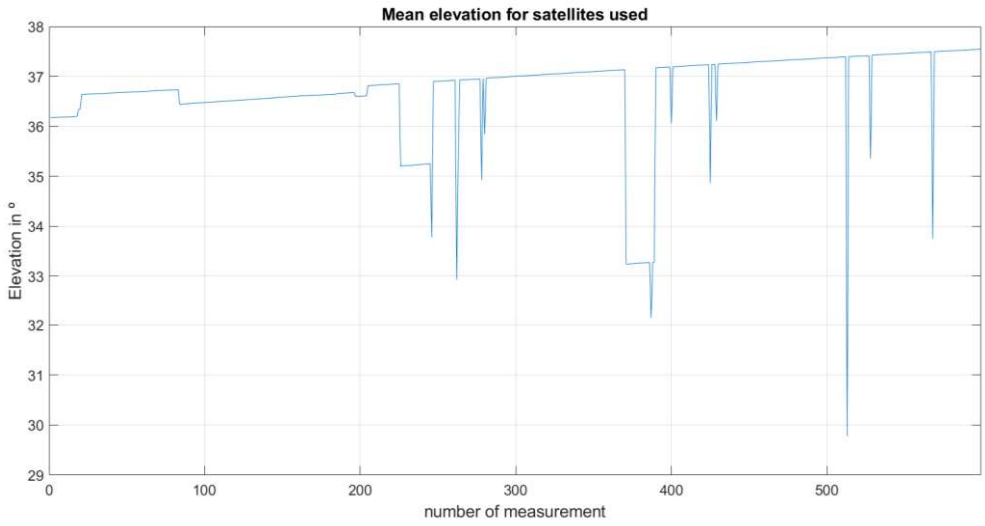


Figure 6.5 Mean elevation for satellites used

6.2.3 SNR of the satellites

The elevation is not the only parameter that should be analysed when deciding if the data received from a satellite is acceptable or not to perform the positioning estimation. In this dissertation, the SNR is analysed to see if the data received is good enough. Even if a satellite has a good elevation, which assures a good line of sight, other phenomenons could appear that would attenuate the power of the signal received providing misleading information to the positioning algorithm.

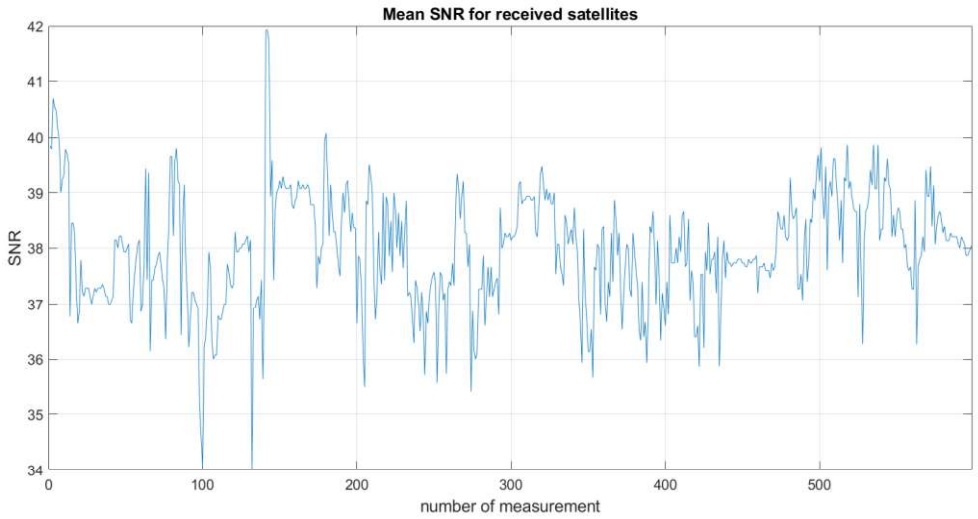


Figure 6.6 Mean SNR for available satellites

Figure 6.6 shows the mean SNR of all the satellites available. It shows an SNR fluctuating from 34dB to 42dB with a mean of 37.97dB. This value is higher than the minimum values used in many GPS positioning systems with SNR masks of 15-20dB. However, this value can be increased by the user in order to compute only the satellites that have better stability on view.

In this thesis, an SNR mask of 25dB has been used and this has filtered some of the satellites previously in view. Figure 6.7 shows the number of satellites discarded for insufficient SNR that moves from 0 to 3 for the set of measurements considered in that test scenario. However, in most of the measurements, 1 or fewer satellites are discarded. Moreover, and changing the view of the analysis, between 12 and 13 satellites are used as the mean value in each of the measurements as shown in Figure 6.8. This is a higher number than the satellites available after using the elevation mask. From this analysis, we can conclude that the elevation mask is usually more restrictive than the SNR mask.

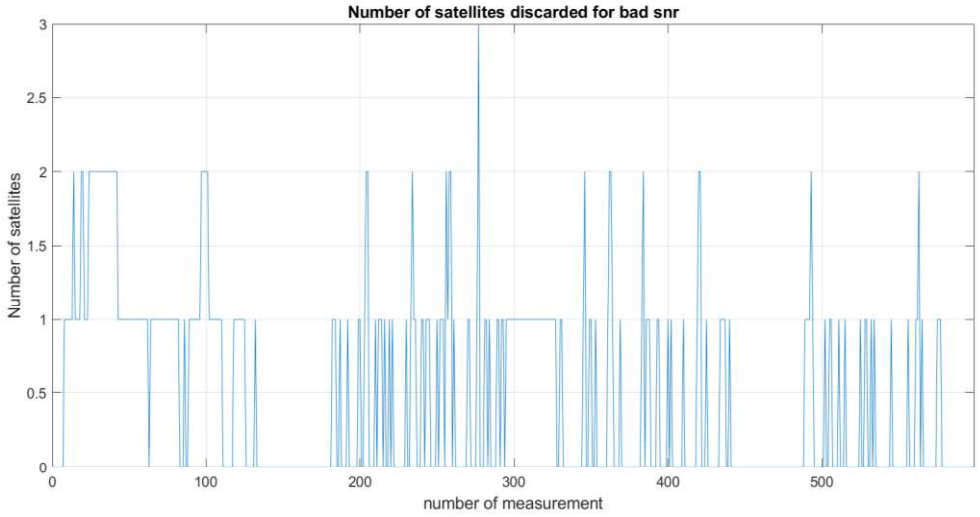


Figure 6.7 Number of satellites discarded due to bad SNR

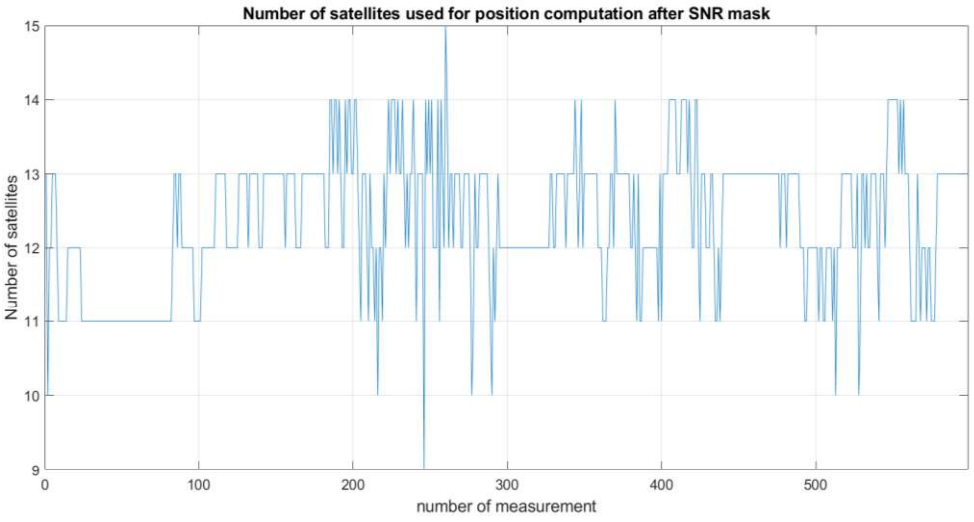


Figure 6.8 Number of satellites used after applying SNR mask

Finally, to finish with the SNR analysis, the mean of the SNR of the satellites used in each measurement after applying the SNR mask is depicted in

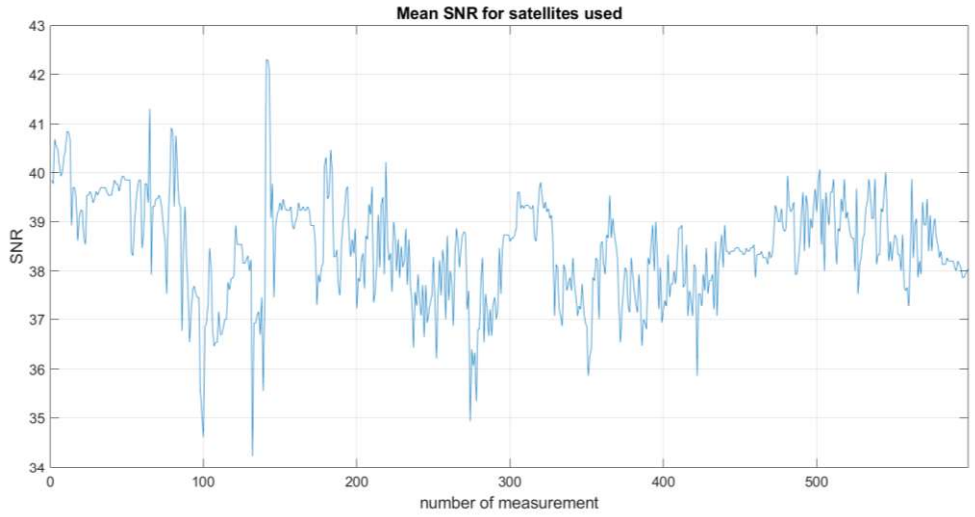


Figure ***6.9***. The mean SNR has increased to 38.44dB which is higher than the value obtained in the case of using all the available satellites.

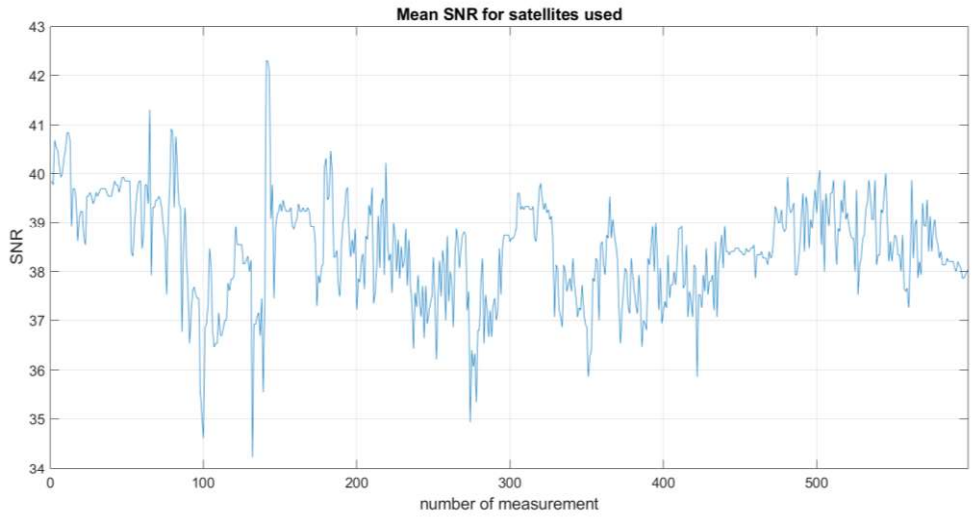


Figure 6.9 Mean SNR for satellites used

6.2.4 SNR and Elevation joint implication

After analysing the SNR and elevation separately, a joint evaluation of both is necessary. It could happen that a satellite with good elevation could have a bad SNR or vice-versa. In order to take into account those possibilities, a joint mask using SNR and elevation thresholds at the same time is going to be used, with the same values than the one introduced in the previous paragraphs. The conclusions obtained are that the most restrictive value, in this case, is the elevation mask. Figure 6.10 shows the satellites discarded by this new mask. It is similar to the one shown by the elevation mask alone, and the mean number of discarded satellites is also similar: 4.24 compared to the previous 4.19. Following the analysis with the satellites in use after the filtering, the behaviour is the same and the mean value of the used satellites is also similar 8.62, compared to 8.66. The maximum and minimum of satellites used is also very similar and continues being acceptable to obtain a position estimate in all the measurements.

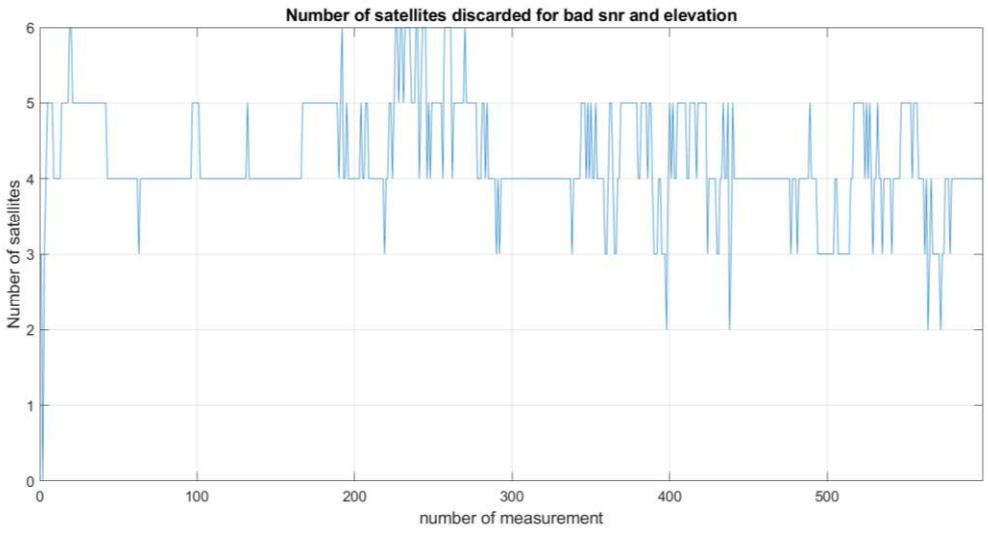


Figure 6.10 Number of satellites discarded due to bad SNR and elevation

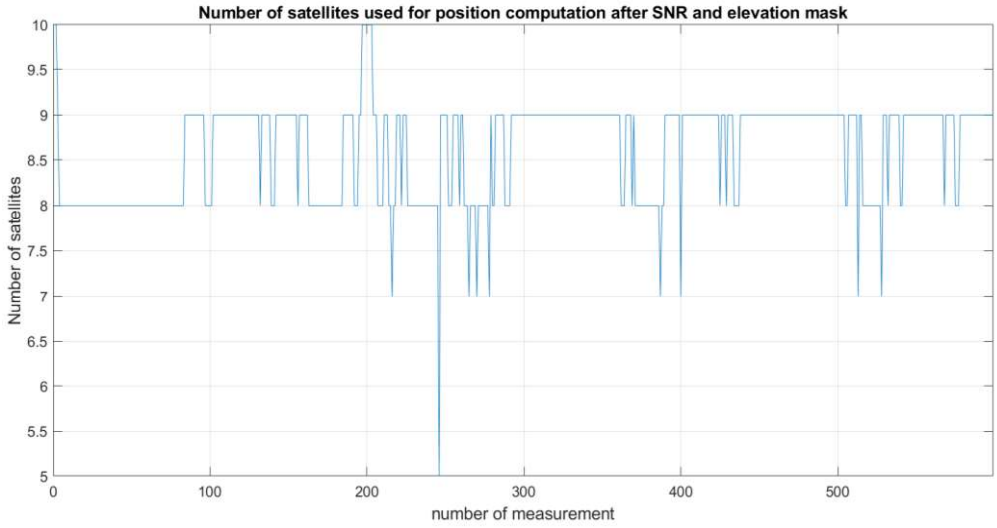


Figure 6.11 Number of satellites used after SNR and elevation joint mask

6.2.5 Standard deviation of the position

The standard deviation of the positions received that are going to be used for the fusion with the other positioning data are analysed. A small standard deviation will show us better confidence in the position estimation given in that axis.

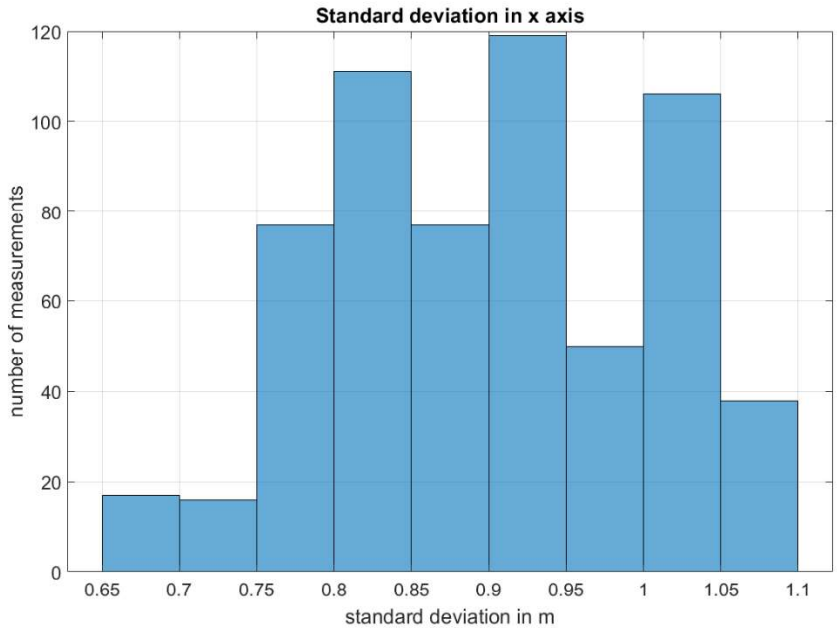


Figure 6.12 Standard deviation in X-axis

Figure 6.12 shows the standard deviation of the position estimation in the X-axis. The standard deviation is under 1.1m during the whole test. Moreover, the mean of the standard deviation is 0.89m.

Similar behaviour happens in the Y-axis. Figure 6.13 shows the standard deviation for the Y-axis. The standard deviations are below 0.55m. The mean standard deviation for this axis is 0.45m.

Analyzing the behaviour of the Z-axis (see Figure 6.14) the standard deviations are below 1.6m, with a mean standard deviation of 1.2m.This means that the position estimation is also going to be good.

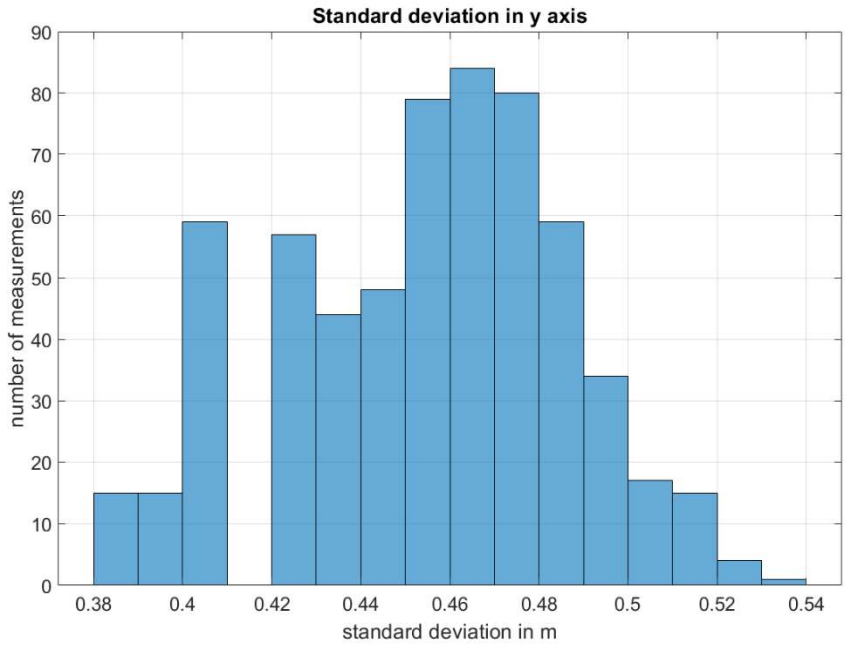


Figure 6.13 Standard deviation in Y-axis

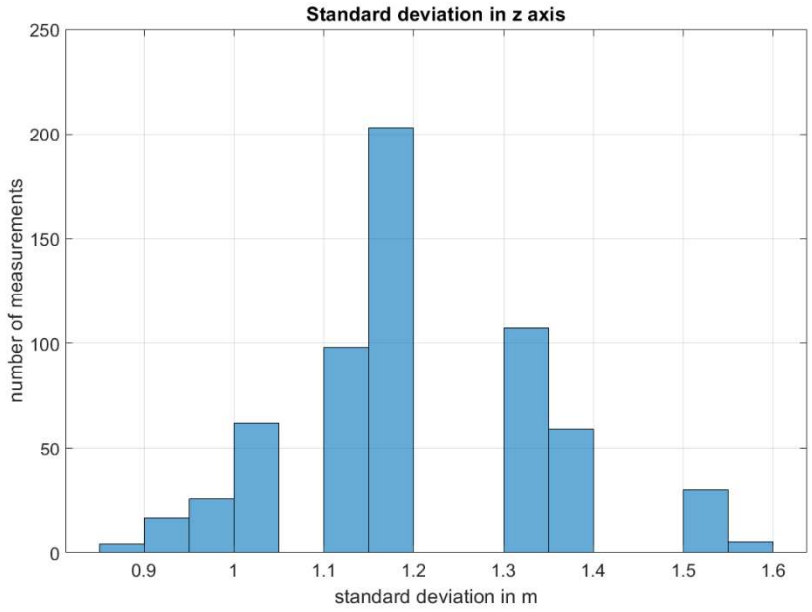


Figure 6.14 Standard deviation in Z-axis

6.2.6 HDOP

The Horizontal Dilution of Precision, HDOP, provides a parameter related to the satellite geometry. Better satellite geometry reduces the uncertainty area and increases the mathematical stability of the algorithm.

The HDOP has different ratings depending on its value and is marked a good when its value is below 5, and ideal when it is below 1 (GIS 2011). Figure 6.15 shows the HDOP of the satellite constellation used during the test. The HDOP for most of the measurements is between 1 and 2, so it is very good. Even in some points, an ideal HDOP can be found.

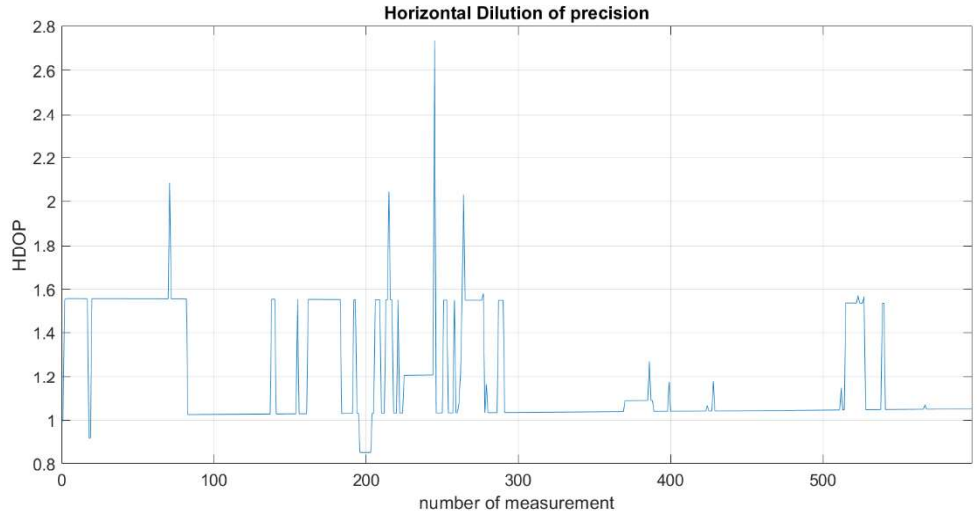


Figure 6.15 Horizontal Dilution of precision (HDOP) for the satellite constellation in use

6.2.7 Number of observables

Finally, the number of observables for the analysed test is going to be examined. In subsection 6.2.1 and subsequent, has been shown that the minimum needed satellites to obtain a position estimation is available in all the measurements. The number of measurements done in this test is 599 and thus there are 599 observables available for the positioning algorithm, one per measurement per second.

6.3 IMU DATA ANALYSIS

The IMU is going to be fused with the rest of the data and it will enhance the availability of the system. Taking into account the fact introduced, different statistics of the data received are going to be analysed to validate the IMU data received and afterwards perform the positioning algorithm with it.

6.3.1 Accelerometer data analysis

First of all the accelerometer data is analysed. Figure 6.16, shows the acceleration measured in the X-axis. The behaviour is coherent with this kind of systems, and there can be detected three-movement stages with stationary stages among them and also at the beginning and the end of the test of this scenario. The most significant conclusion that we can obtain from this data is the existence of a bias error, which is stopped motion has a value of 0.01m/s^2 in X axis, 0.04m/s^2 in Y axis and 0.01m/s^2 in Z axis.

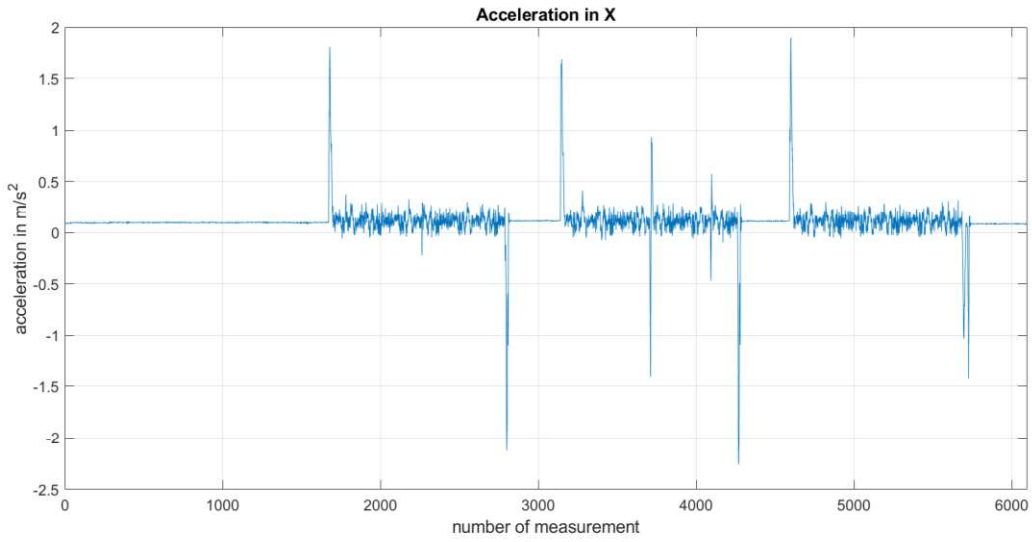


Figure 6.16 Acceleration in X-axis

Figure 6.17, depicts the acceleration in the Y-axis. As in the X-axis, the only anomaly spotted in the behaviour of the acceleration in this axis is the existence of a bias error. From this graph, new data from the test can be obtained. The movement of the rover has a 8-shape (double circular) and a three-time repetition of the journey is done in every movement stage of the test.

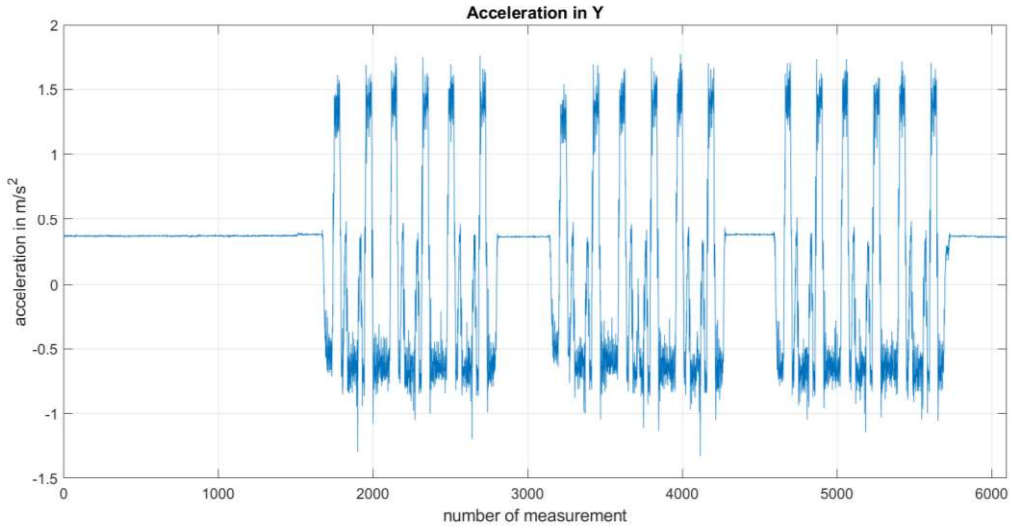


Figure 6.17 Acceleration in Y-axis

To finish with the analysis of the accelerations, the Z-axis is shown in Figure 6.18. The Z-axis is the axis that less influence has in the 2D positioning. However, it measures the vibration absorbed by the rover. An important fact is that the acceleration always contains the gravitational component in this axis and that is the reason to have mean accelerations of -9.7 m/s^2. As in the other axis, a bias error can be found as the gravitational component is of about 9.8 m/s^2. The reason to have a negative acceleration is the orientation of the axis with the positive Z component facing up the body of the IMU.

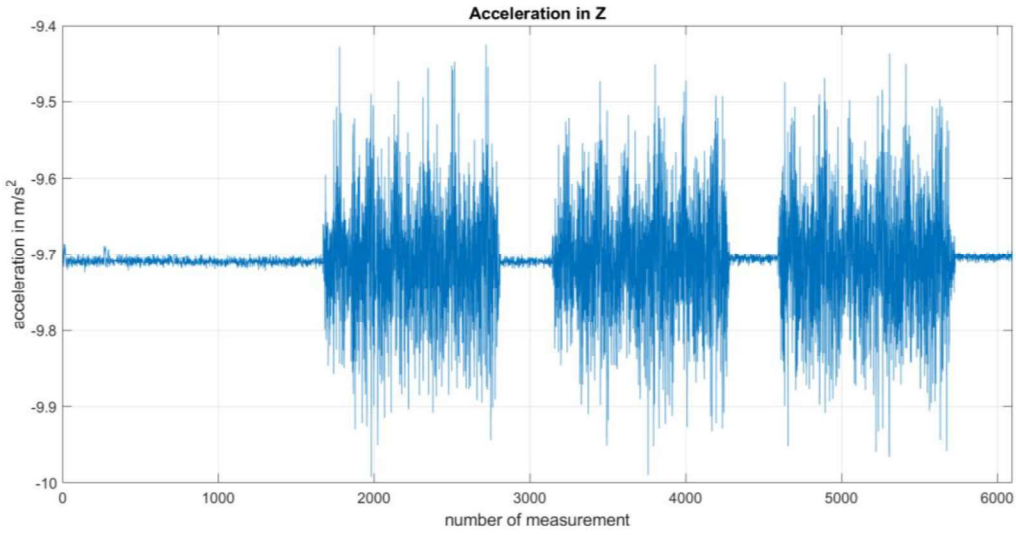


Figure 6.18 Acceleration in Z-axis

After the analysis of the accelerations recorded, the conclusion is that the data recorded are coherent with the test performed and that they can be used as inputs for the positioning computing. Measured bias has values of $0.01 m/s^2$ in X axis, $0.04 m/s^2$ in Y axis and $0.01 m/s^2$ in Z axis, which complies the values showed in the datasheet, <5mG.

6.3.2 Gyroscope data analysis

The gyroscope data analysis is meaningful as this parameter is responsible for calculating the angles of movement of the rover. This dissertation is oriented to the position computation and its accuracy, and so, the angular accelerations in X and Y axis are not as important as the Z-axis that is the one that is going to calculate the heading of the rover.

Figure 6.19, shows the angular accelerations recorded in the X-axis. In this case, the angular accelerations measured are smaller than the angular acceleration of the Z-axis (Figure 6.21), and so is the bias.

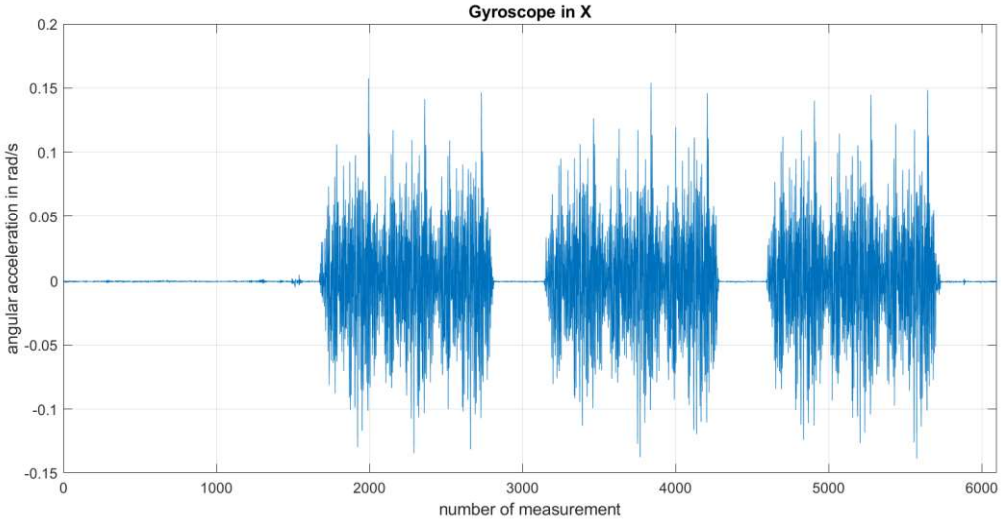


Figure 6.19 Gyroscope angular acceleration in X-axis

Figure 6.20 shows the angular accelerations recorded in the Y-axis. As happened in the X-axis, the angular accelerations measured are small and so the bias error is. The reason for these effects is the before mentioned negligible effect of the angular accelerations in the X and Y axis for a rover installed in a vehicle moving on rail tracks in a flat plane without cambers or course changes.

Finally, Figure 6.21 shows the angular accelerations recorded in the Z-axis, the ones that determine the heading of the rover. In this case, as the recorded data is not negligible, a considerable bias error appears. However, coherent behaviour according to the test performed is depicted as in the other axis. The bias calculated are -0.02º/s in X axis, 0.02 º/s in Y axis and 0.72 º/s in Z axis. Which does not fulfil with the value given in the datasheet in the Z axis. The datasheets shows bias <0.2º/s for all the three components.

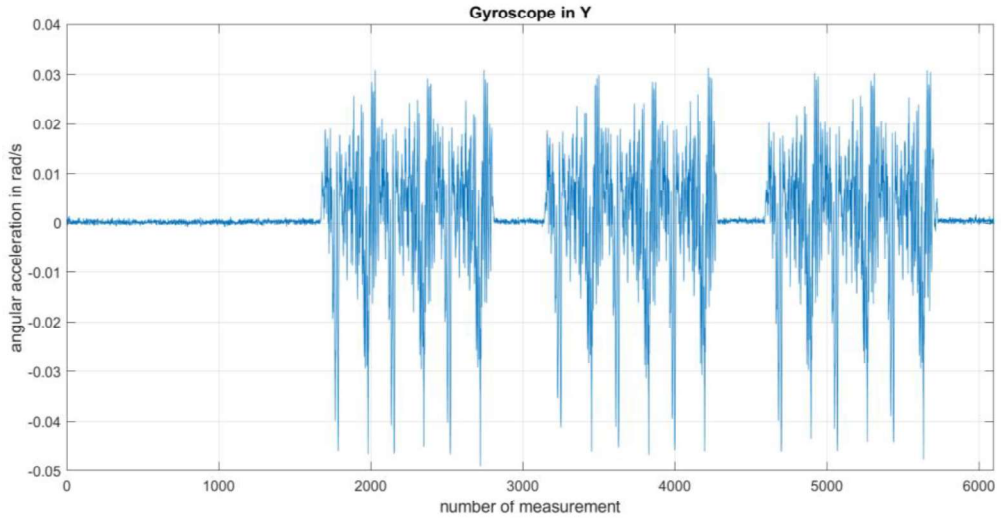


Figure 6.20 Gyroscope angular acceleration in Y-axis

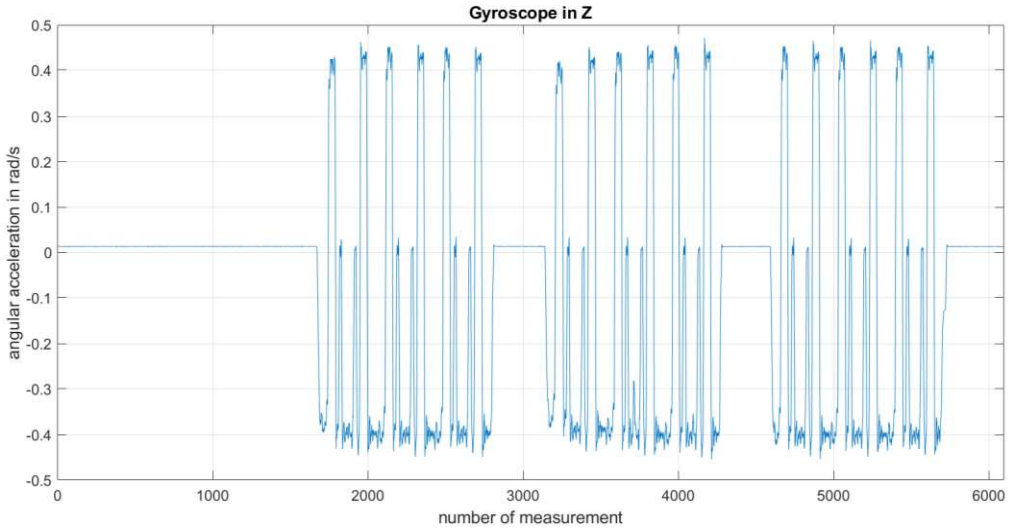


Figure 6.21 Gyroscope angular acceleration in Z-axis

6.3.3 Yaw data analysis

Another significant indicator to assess that the data are coherent with the test scenario is the so-called Yaw angle. The Yaw angle is calculated from the angular accelerations obtained in the Z-axis. The yaw angle shows the angle that the rover has in comparison with the north. That is the reason to also call it “heading”. Figure 6.22, shows the behaviour of the Yaw angle during the whole test. The figure shows perfectly the 3 movement-stationary stages. Doing a deeper analysis an anti-clockwise movement can be detected, and the Yaw changes can be compatible with a journey similar to a flat 8 that is presented in the scenario in which the test has been performed.

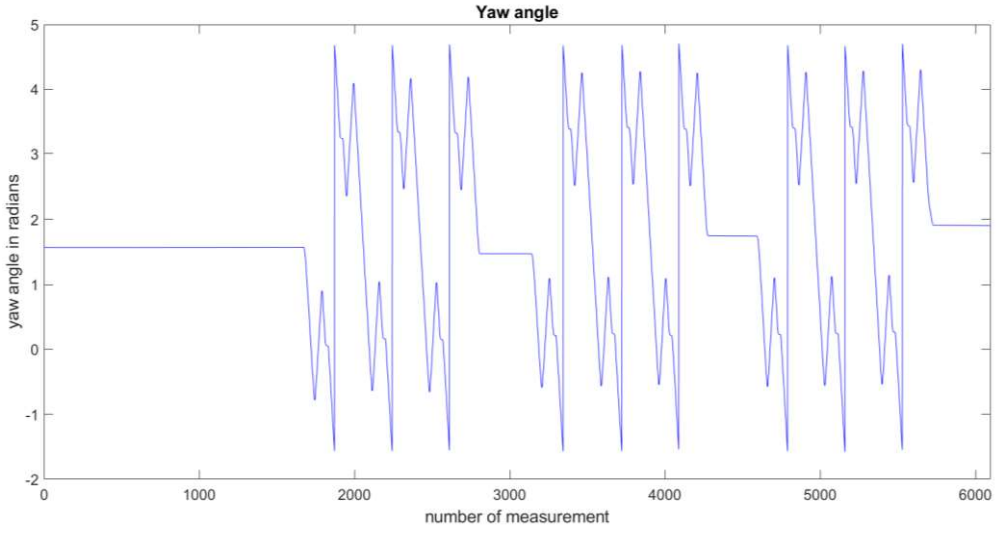


Figure 6.22 Yaw angle

6.3.4 Bias influence in accelerometer data

In subsections 6.3.1 and 6.3.2, a bias error has been detected in the measurements obtained from the test. The bias error is a usual IMU error as presented in section 482.2. However, even if it is usual to have bias errors, it also easy to correct that error, as it is constant during the time. To correct the data obtained and eliminate the bias errors, its value has been calculated. For that, the first 1000 measurements have been used where the rover remains stoped at the same point.

The bias calculated for the acceleration in the X-axis is 0.096 m/s^2. Figure 6.23 shows the new acceleration after subtracting the obtained bias to the measured accelerations. Now, the acceleration in X-axis when the rover is in a stationary stage is around 0 m/s^2. This behaviour models the reality in a better way. It can be concluded that the bias subtraction makes the acceleration more realistic.

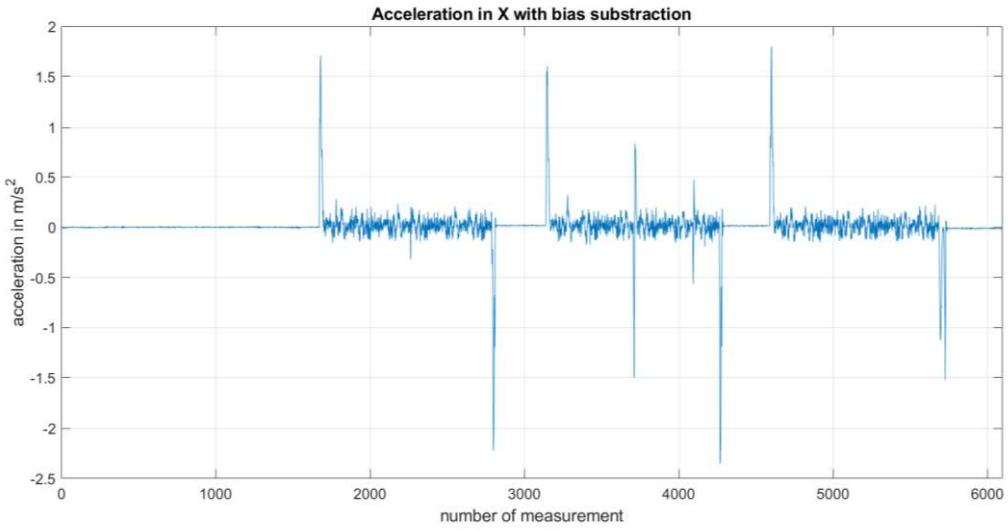


Figure 6.23 Acceleration in the X-axis after bias subtraction

The same calculus has been performed for the Y-axis. In this case, a bias error of 0.37 m/s^2 appears in the measurements. In the same way as for the X-axis, when subtracting the bias error to the measurements, the accelerations get centered in 0 m/s^2.

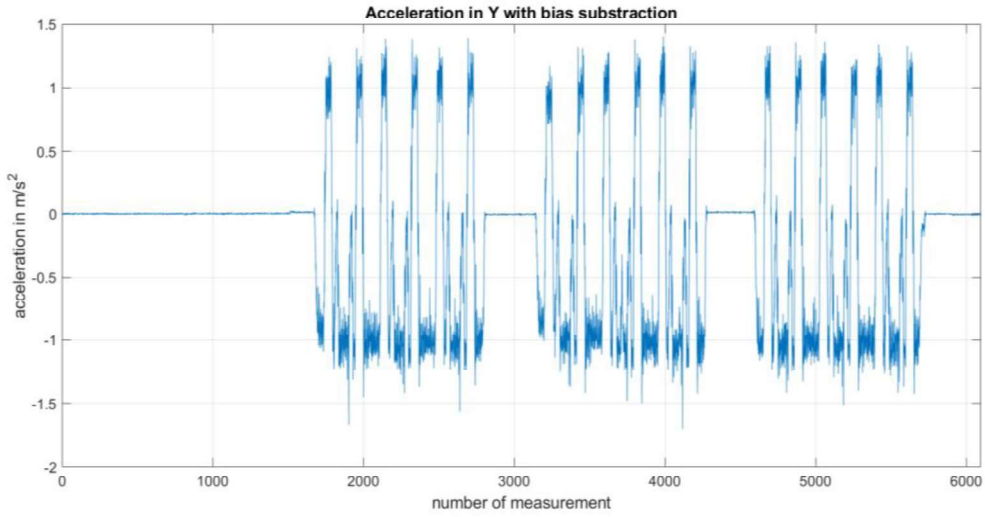


Figure 6.24 Acceleration in the Y-axis after bias subtraction

In the Z-axis the bias error is 0.09 m/s^2. However, subtracting the bias error the acceleration does not get centered in 0 m/s^2. Apart from subtracting the bias error, the gravitational component has also been subtracted in order to center the acceleration in 0 m/s^2 (see Figure 6.25).

The same analysis done with the linear accelerations has been done with the angular accelerations. Figure 6.26, shows the angular accelerations in X-axis after subtracting the bias error of 0.0002 rad/s. As explained before, X-axis has not significant angular accelerations and so the bias error is also smaller than the one on the Z-axis.

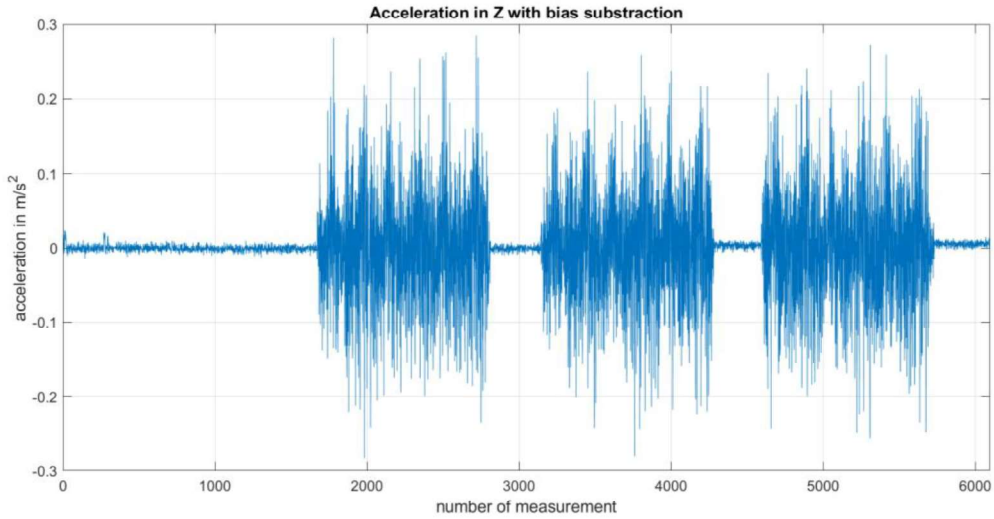


Figure 6.25 Acceleration in the Z-axis after bias subtraction

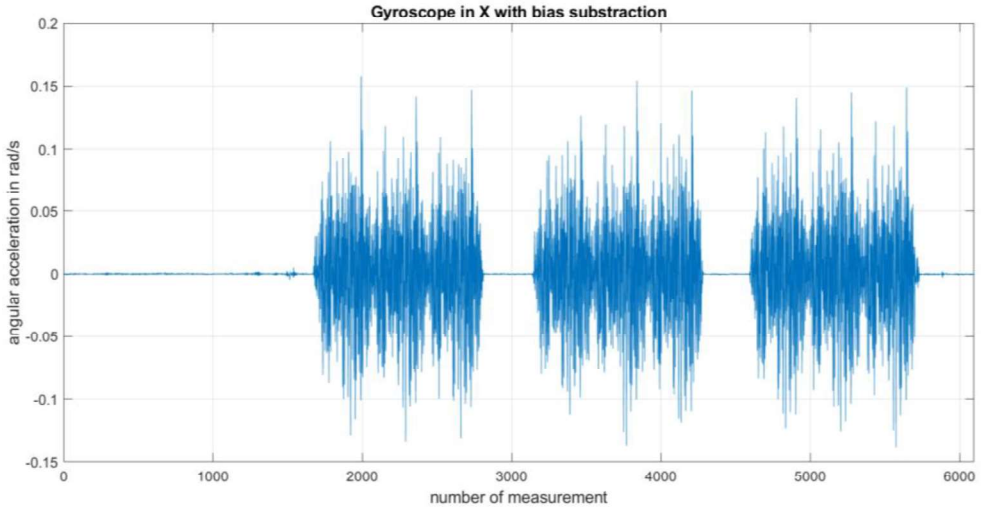


Figure 6.26 Angular acceleration in the X-axis after bias subtraction

The behavior seen in the X-axis is repeated in the Y-axis. Figure 6.27, shows the angular accelerations in Y-axis after subtracting the bias error of 0.0002 rad/s. As explained before, Y-axis has not significant angular accelerations and so the bias error is also much smaller than the one presented for Z-axis.

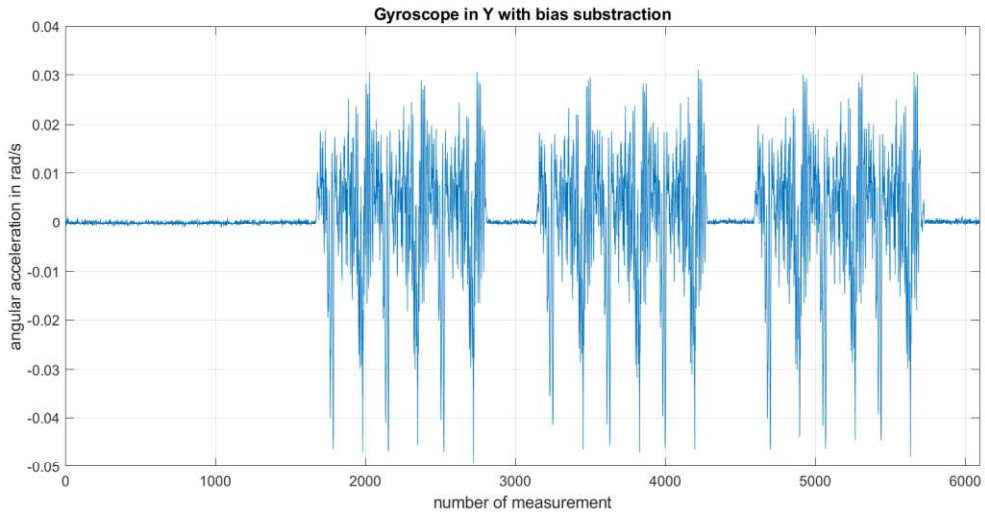


Figure 6.27 Angular acceleration in the Y-axis after bias subtraction

Finally, the Z-axis is analyzed, the one used for the Yaw computation. Figure 6.28, shows the angular accelerations in Z-axis after subtracting the bias error of 0.012 rad/s. In this case, the angular accelerations measured are significant and thus is the bias error calculated. This leads to the correction of the bias error to be more evident in the Z-axis than in the other two axis analyzed.

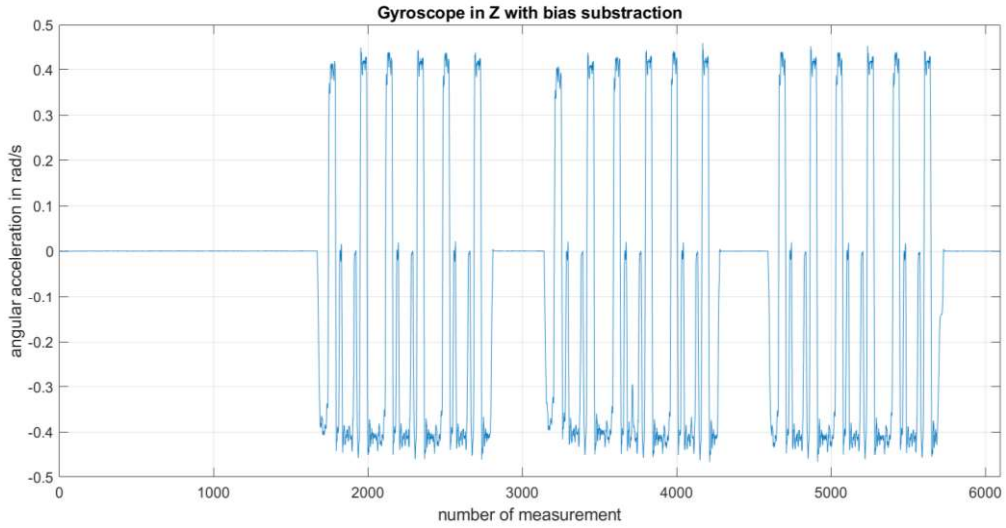


Figure 6.28 Angular acceleration in the Z-axis after bias subtraction

6.3.5 Number of observables

Finally, the number of observables for the analysed test in this scenario is going to be examined. In the case of the IMU, no acceleration measurements are exceeding the maximum accelerations accepted in the railway industry for passenger commodities and introduced in paragraph 2.2.1.1. The number of measurements done in this test is 6098 which is more or less 10 times the number of measurements for the GPS subsystem. This shows the usual operation of the IMU as it is programmed to work with a 10Hz frequency while the GPS is programmed to work with a 1Hz frequency.

6.4 UWB DATA ANALYSIS

UWB is proposed in this thesis as a system used for conflictive spots in which the GPS accuracy and availability is not enough, or directly there is no GPS data. A good example would be for example a train station. Inside train stations, there is usually no proper GPS signal available, and so the use of UWB could give good accuracy maintaining the availability provided by IMU.

UWB positioning system used provides us with a position estimation and some residuals with which some figures of merit have been calculated to fulfil this preliminary analysis and validate the UWB data recorded.

6.4.1 Residuals

First of all, the residuals given by the positioning system are analysed. This statistical gives the measure of how good is the position estimation given by the UWB system. This fact makes the residuals an important factor to take into account.

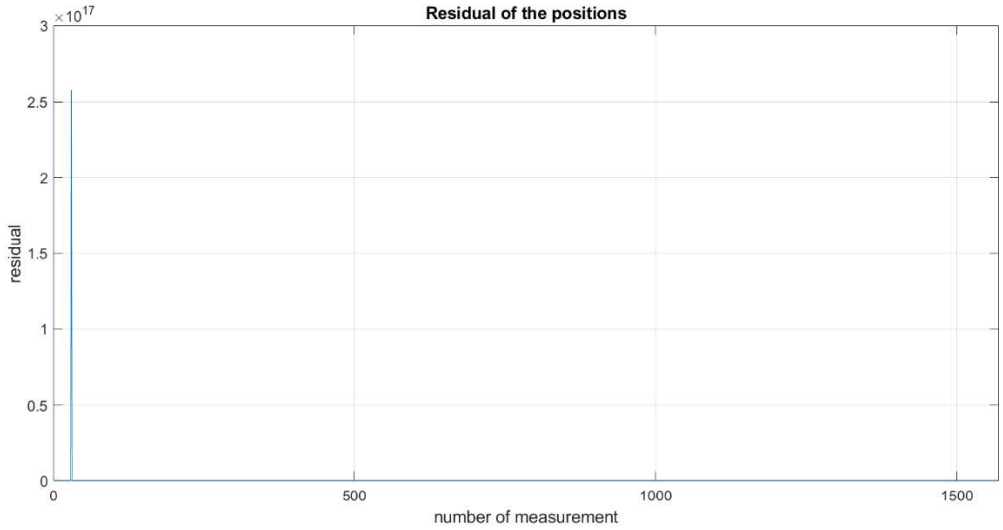


Figure 6.29 Residuals of the positions

Figure 6.29 shows the residuals of the UWB system over the whole test. A big peak appears at the beginning of the test showing unacceptable residuals. The position estimation associated with the residuals in that part of the test could not be reliable. The decision to use a mask of residuals to filter bad measurements has been made, as explained and exposed here below.

As UWB is wanted to be a very precise system used only in conflictive spots or spots without GPS coverage, only estimations with very low residuals, less than 0.0001 are accepted as good estimations. Figure 6.30 shows the new residuals after applying the mask proposed. No outliers appear in this new figure. The position estimations given by these measurements should be of good confidence.

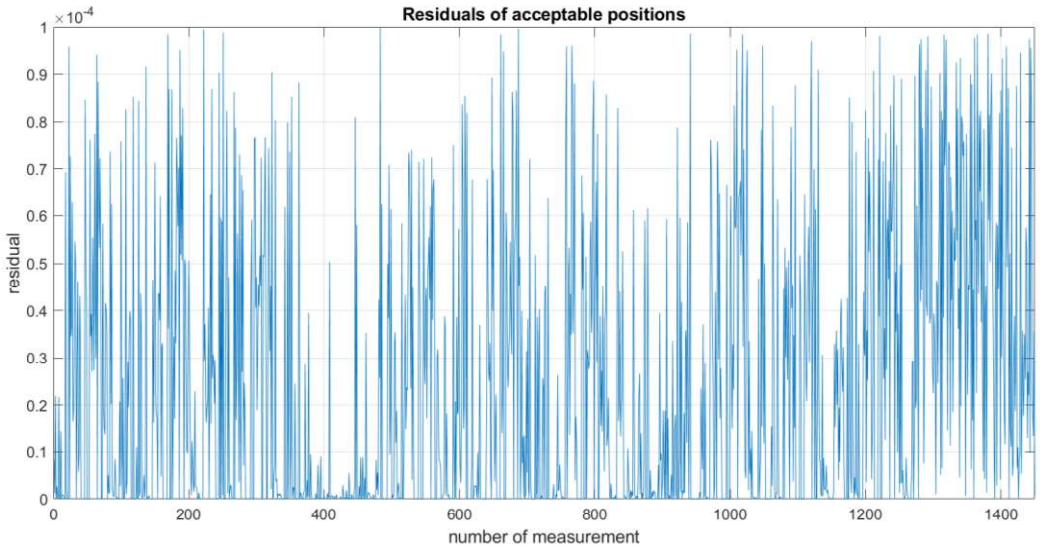


Figure 6.30 Residuals of acceptable measurements

6.4.2 Number of Observables

The number of total UWB measurements for this test is 1570. However not all the measurements are good enough to use their positions. After applying the residuals mask, the number of available measurements is reduced to 1450. It can

be deduced that 120 measurements are not reliable enough for this application. However, a good rate of 2 measurements per second is obtained.

6.4.3 Standard deviation of the position

The standard deviation of the positions received that are going to be used for the fusion with the other positioning data are analysed. A small standard deviation will show us better confidence in the position estimation given in that axis.

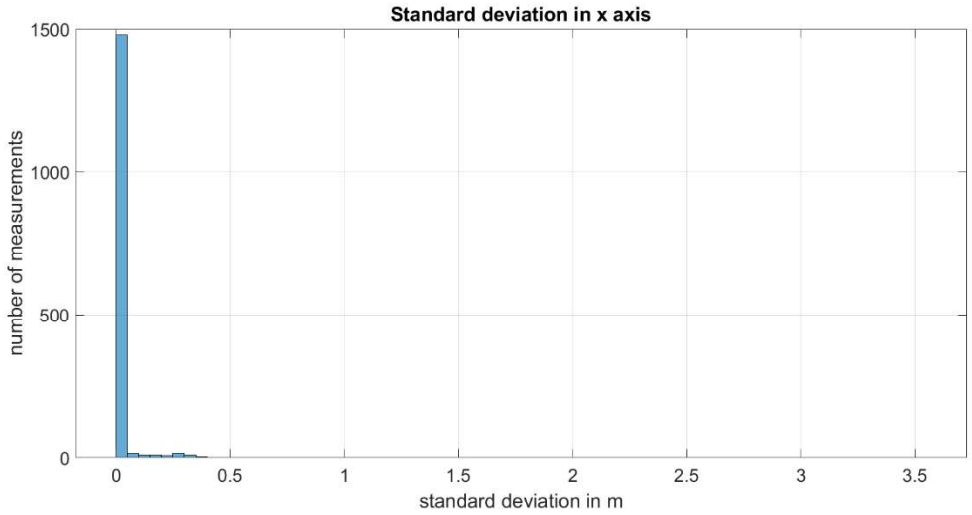


Figure 6.31 Standard deviation in X-axis

Figure 6.31 shows the standard deviation of the position estimation in the X-axis. The standard deviation is under 0.5m during the whole test, after the filtering of the first data. Moreover, the mean of the standard deviation is 0.02m. However, some high standard deviations can be found at the beginning of the test. Those high standard deviations could lead the system to give some bad position estimations.

Similar behaviour happens in the Y-axis. Figure 6.32 shows the standard deviation for the Y-axis. There are some outliers at the beginning of the test and then the standard deviations are below 0.2m. The mean standard deviation for this axis

is 0.0039m what shows very good confidence in the position estimation given for the Y-axis.

Analyzing the behaviour of the Z-axis (see Figure 6.33), something very similar to the other two axes is found. Some outliers at the beginning of the test and standard deviations below 0.5m in the rest, with a low mean standard deviation of 0.018m.This means that the position estimation is also going to be good.

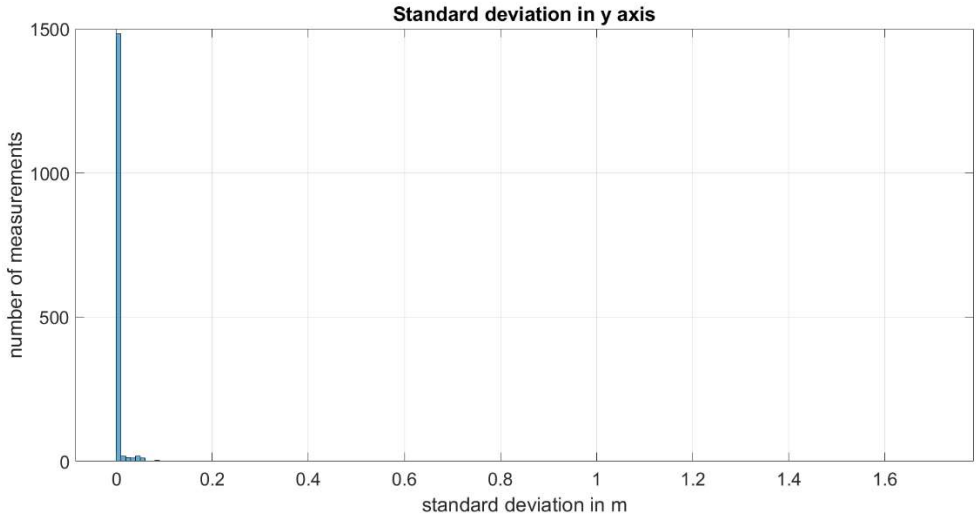


Figure 6.32 Standard deviation in Y-axis

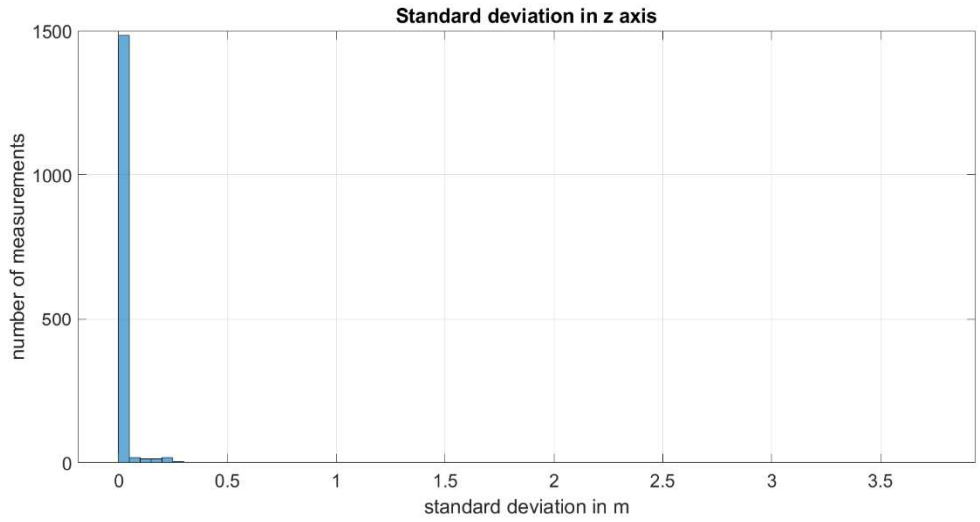


Figure 6.33 Standard deviation in Z-axis

When analyzing the standard deviation of the position in the three-axis, some outliers have been spotted. As explained in subsection 6.4.1, not all the measurements have reliable information to position. So using the residuals mask introduced, only the reliable measurements have been taken into account. After doing it, new standard deviation curves have been graphed.

Figure 6.34 shows the standard deviation in the X-axis after applying the residuals mask. The outliers have been deleted, and all the standard deviations are below 0.01m. The mean standard deviation has also been reduced to 0.0026m. Now all the chosen measurements are reliable and the position estimations will contribute the positioning algorithm.

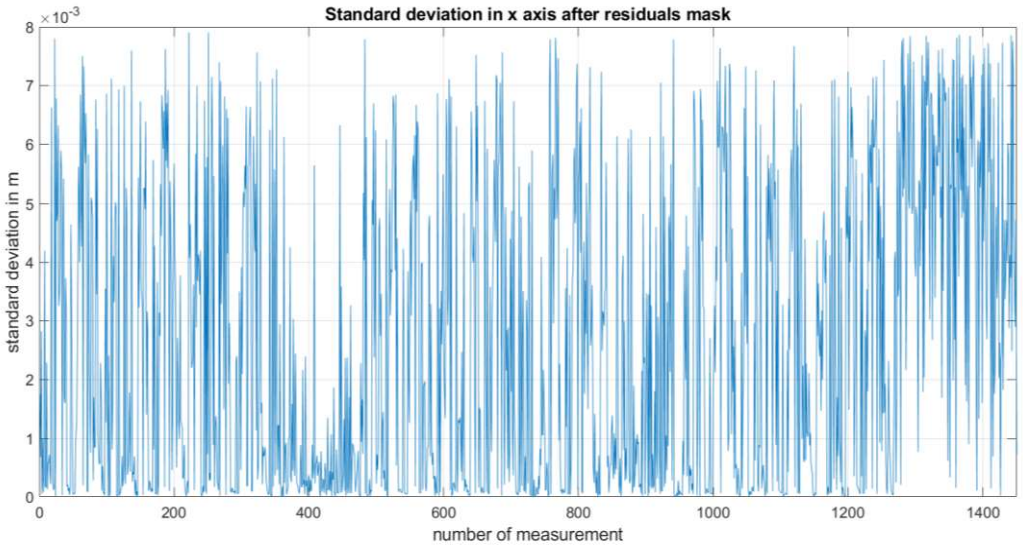


Figure 6.34 Standard deviation in X-axis after using the residuals mask

Similar behaviour happens in the Y-axis. Figure 6.35 shows the standard deviation for the Y-axis after applying the residuals mask. The outliers at the beginning have been deleted and the standard deviations are below 0.01m. The mean standard deviation for this axis is 0.0009m what shows very good confidence in the position estimation given for the Y-axis.

Analyzing the behaviour of the Z-axis (see Figure 6.36), something very similar to the other two axes is found. The outliers have disappeared, and the standard deviations are always below 0.01m with a low mean standard deviation of 0.0019m. This means that the position estimation is also going to be good.

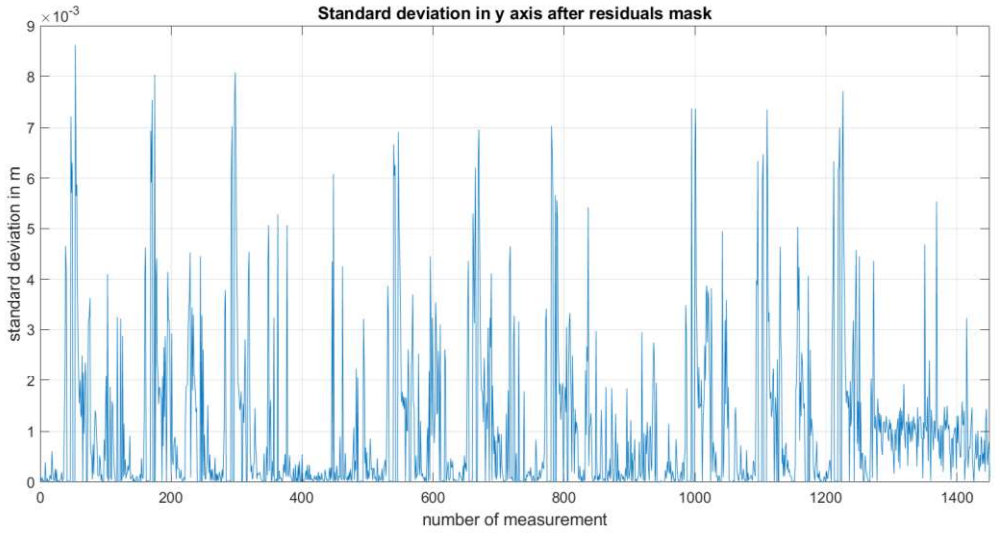


Figure 6.35 Standard deviation in Y-axis after using the residuals mask

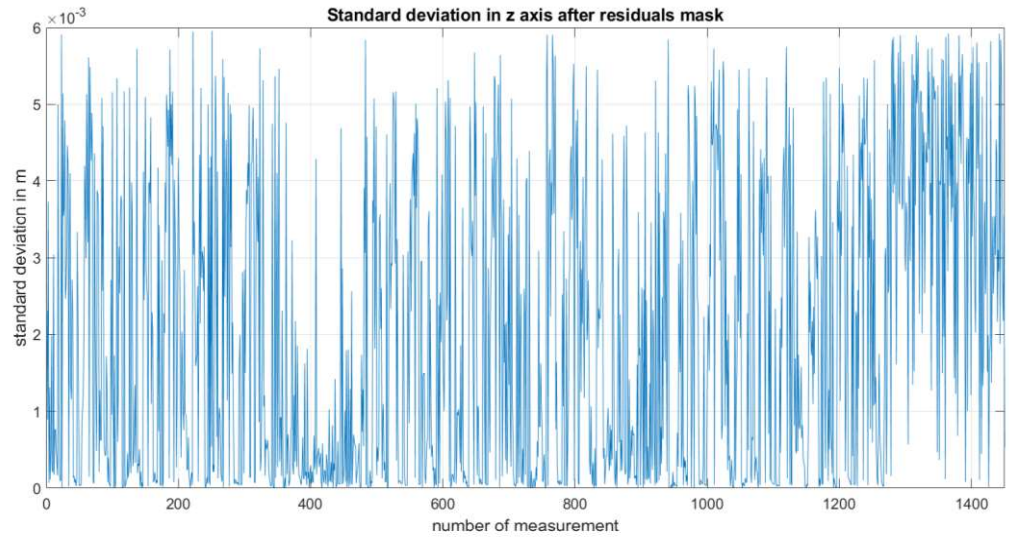


Figure 6.36 Standard deviation in Z-axis after using the residuals mask

6.4.4 HDOP

The Horizontal Dilution of Precision, HDOP, specifies the navigation satellite geometry on a positional measurement precision. In other words, it says how good is the satellite geometry. In the UWB case, there are no satellites, but there are anchors that will act in the same way as satellites when positioning the rover.

The HDOP has different ratings depending on its value and is marked a good when its value is below 5, and ideal when it is below 1 (GIS 2011). Figure 6.37 shows the HDOP of the anchors placed for the test. The HDOP for all the measurements is below 1, so it is ideal, except at the beginning of the test where some outliers are found. In that area, the HDOP is rated only as good.

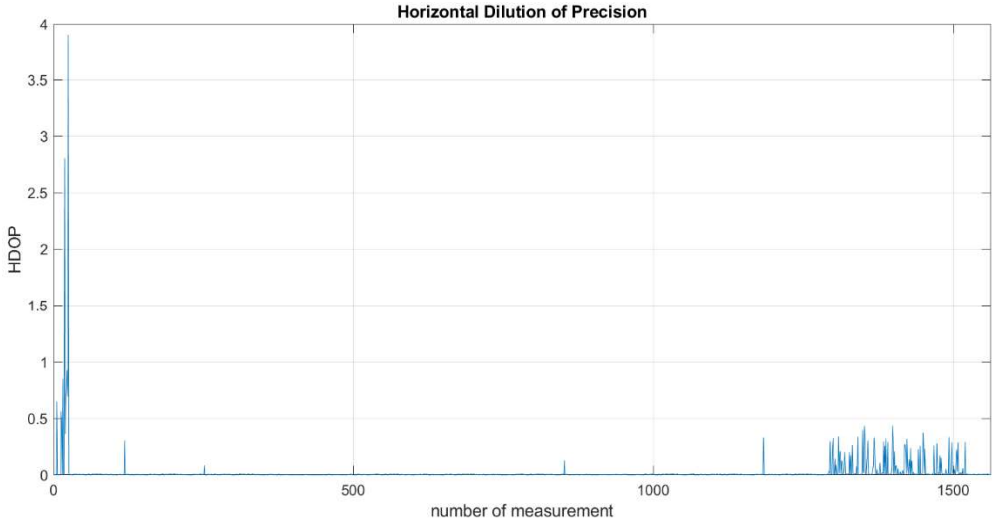


Figure 6.37 Horizontal Dilution of Precision

As a very high accuracy is wanted in the UWB system, measurements with the highest possible HDOP are wanted. With that purpose, the residuals mask introduced in subsection 6.4.1 is also used here. The result is depicted in Figure 6.38. After applying the mask, and taking into account only the measurements with

good residuals, all the computed HDOP for the measurements are below 1 and thus they are rated as ideal.

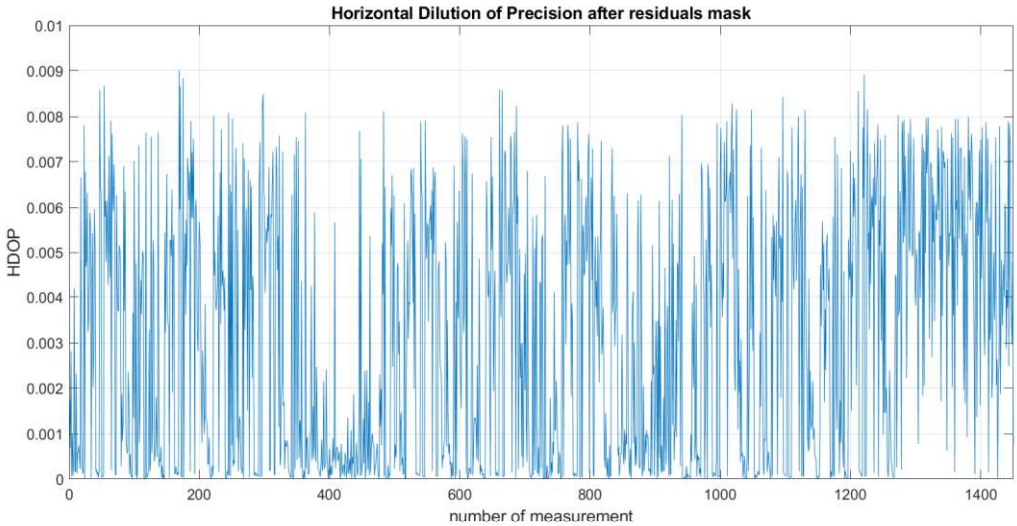


Figure 6.38 Horizontal Dilution of Precision after residuals mask

Chapter 7

Field-test results and performance evaluation

This chapter shows and analyzes the results obtained by the positioning algorithm in the measurement campaigns explained in Chapter 5. Moreover, it analyzes the performance obtained by the algorithm in comparison with the ground truths of each of the scenarios.

7.1 PERFORMANCE OF THE ALGORITHM

This section shows the performance of the algorithm in the different tests defined in subsection 5.5.1.

7.1.1 TS1_EUS_1100

The results of this test show the performance of the algorithm when using a GPS and IMU fusion. Figure 7.1 shows the performance of the presented algorithm showing the error committed for each measurement.

This scenario has a mean error of 9.88 meters and a maximum error peak that reaches 50 meters. During this peak, the train travels through a tunnel.

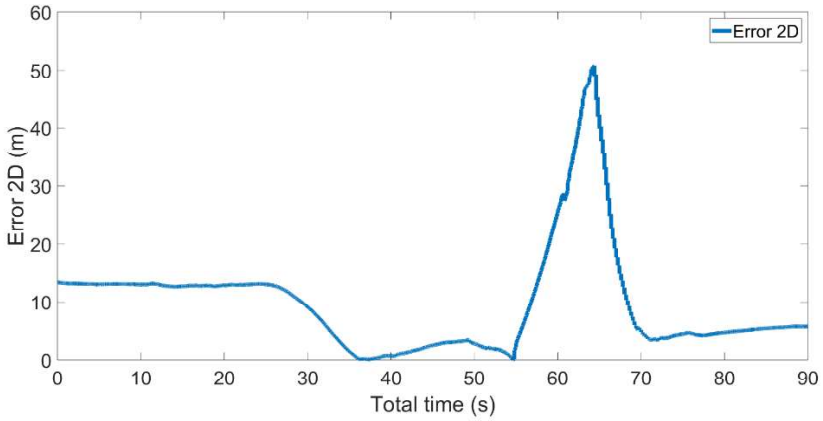


Figure 7.1 Performance of the algorithm for test TS1_EUS_1100

Figure 7.2 analyses the frequency of the different errors during this journey. The error is mainly spread at low error distances below twelve meters; however, there is a long tail with errors that reach the aforementioned value of 50 meters.

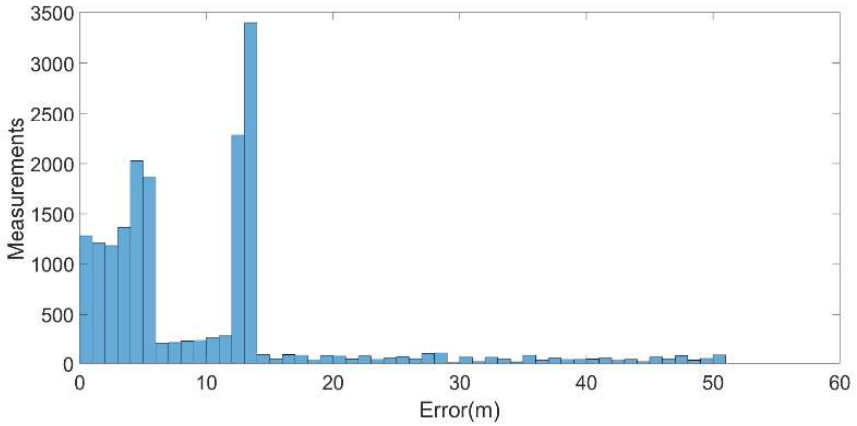


Figure 7.2 Histogram of the position error for test TS1_EUS_1100

Analyzing the cumulative distribution function (CDF) it can be stated that the CDF(95%) is around 31 meters (see Figure 7.3).

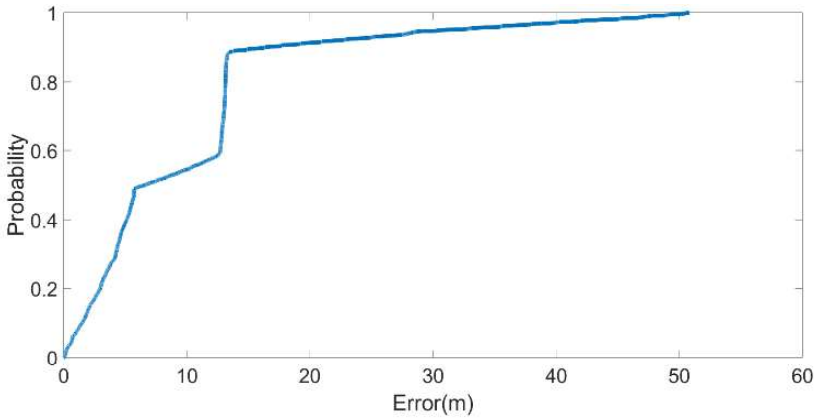


Figure 7.3 CDF of the test TS1_EUS_1100

The error obtained in the tunnel environment decreases the performance significantly and thus improving this section will have a great impact on the overall picture.

7.1.2 TS1_EUS_1101

The results of this test show the performance of the algorithm when using a GPS and IMU fusion along with KBS. Figure 7.4 shows the performance of the presented algorithm showing the error committed for each measurement.

In this case, the accuracy is significantly better than in the test before in which KBS was not used, even if a certain peak can be seen with around 12 meters of maximum error. This improvement makes the mean error to be reduced down to 5.54 meters.

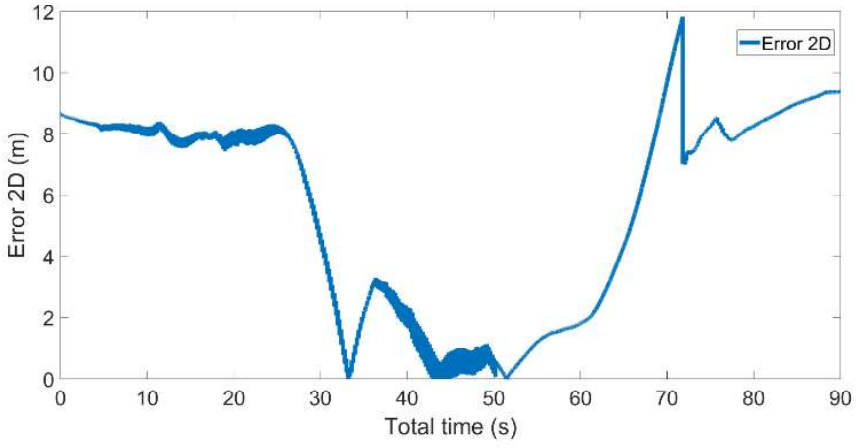


Figure 7.4 Performance of the algorithm for test TS1_EUS_1101

In the same way, analyzing the frequencies of the errors shown by the histogram in Figure 7.5, the maximum error is set in 12 meters which are lower than the preceding case where the maximum error is set in 50 meters.

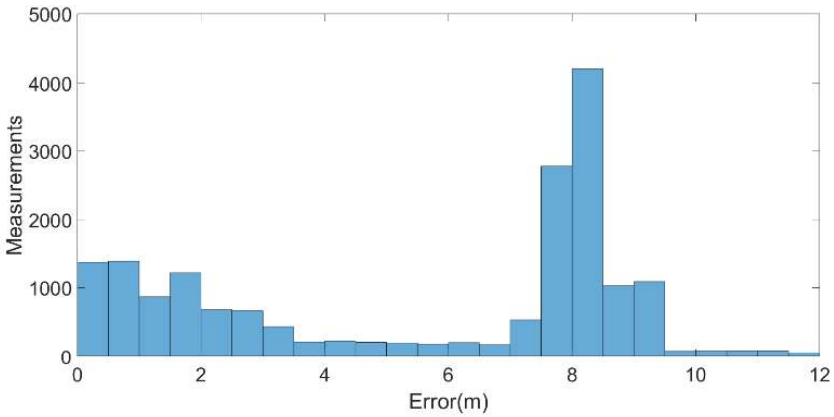


Figure 7.5 Histogram of the position error for test TS1_EUS_1101

The CDF obtained from the error information bound the error at 95% in 9.27 meters (see Figure 7.6).

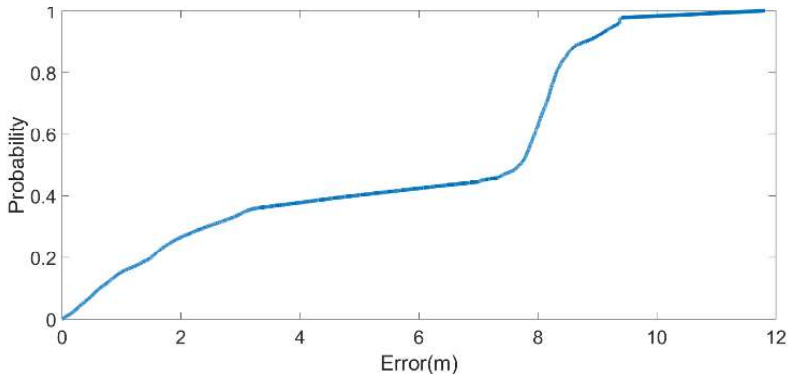


Figure 7.6 CDF of the test TS1_EUS_1101

In tests TS1_EUS_1100and TS1_EUS_1101, the whole journey has been analyzed. However, the KBS is effective when used in singular places in which GPS signal is not available, such as tunnels.

However, to analyze the whole journey could blur the focus from the important part of the KBS enhancement. That is the reason to show the graphs in the area in which the KBS enters in operation. Figure 7.7 shows the KBS operation zone that is analyzed.

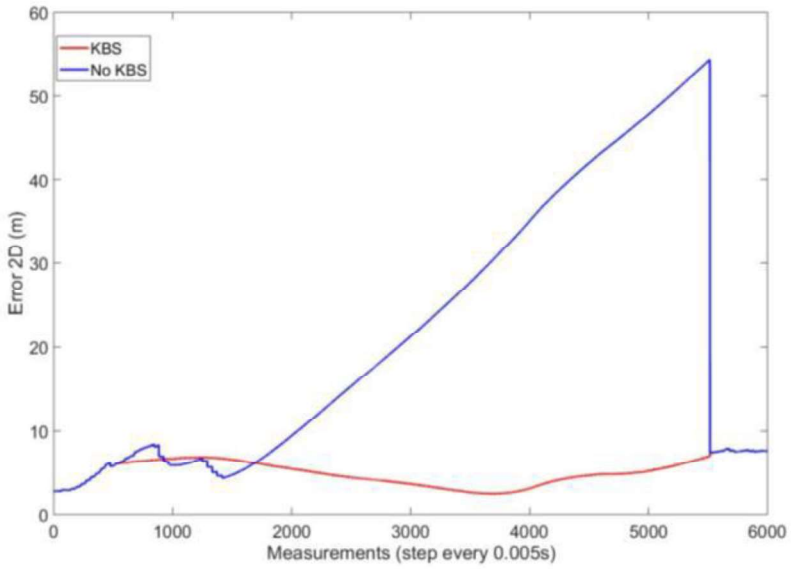


Figure 7.7 Performance in KBS zone, Test TS1_EUS_1100 vs TS1_EUS_1101

The accuracy results of the algorithm without the KBS show an error peak that reached 54.29 meters. The mean error of this section is 21.15 meters.

On the other hand, when using KBS, accuracy is significantly better. The maximum error peak is lower than 7.96 meters and the main improvement is that the mean error is reduced to 5.07 meters

7.1.3 TS2_NOT_1000

The results of this test show the performance of the GPS stand-alone. Figure 7.8 shows the performance of the presented algorithm using GPS (blue) against the

Ground Truth (GT) (red). The GPS gives a position estimation that draws a path similar to the one given by the GT.

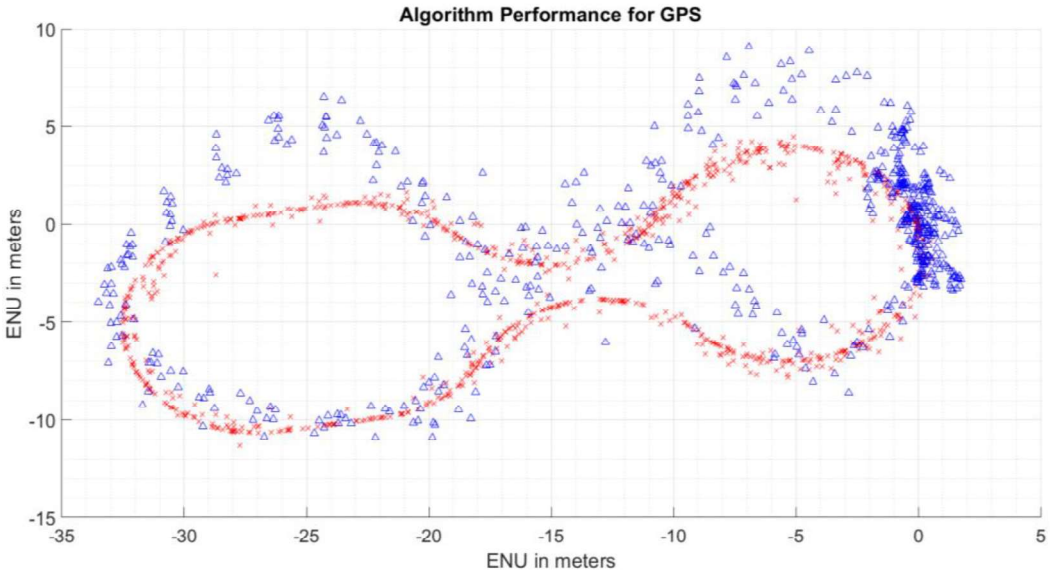


Figure 7.8 Performance of the algorithm for test TS2_NOT_1000

The mean of the error is 0.95 meters. However, there are errors up to 5.05 meters and the standard deviation of the error is 1.04 meters. Moreover, the number of measurements is small for some areas or applications in which one measurement per second could not be enough.

When analyzing the error histogram (see Figure 7.9), the biggest part of the measurements have errors lower than a meter due to the good positioning accuracy of the algorithm when it uses GPS in static environments. Those static parts, at the beginning and at the end of the test, have accuracies below a meter which shift the overal results towards this low error. However, there are position estimations with accuracies under a meter also during the movement stages. However, the multipath errors in the middle part of the scenario due to the reflection of the signals in the different ventilation ducts lead some estimations to higher positioning errors.

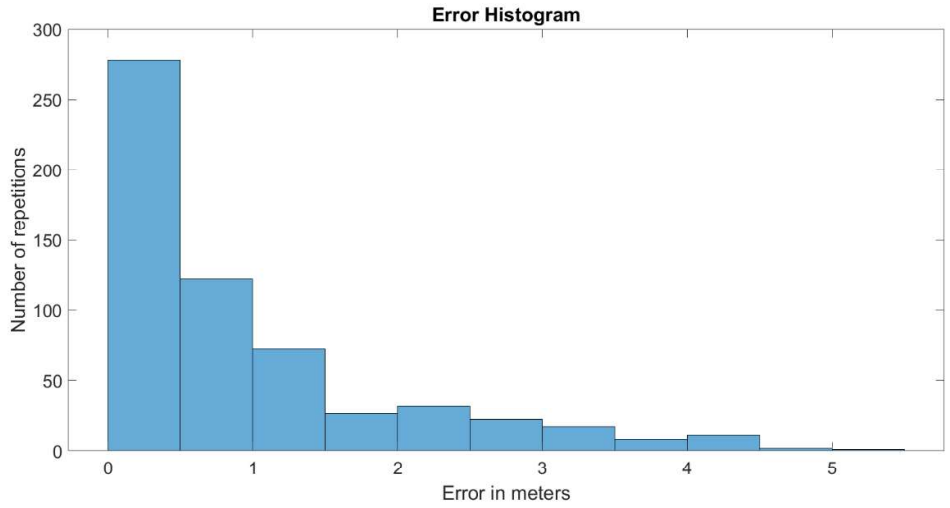


Figure 7.9 Histogram of the position error for test TS2_NOT_1000

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.10), shows that more than 95% of the position estimations have errors under four meters, and more than 90% of the position estimation errors are under 2.5 meters.

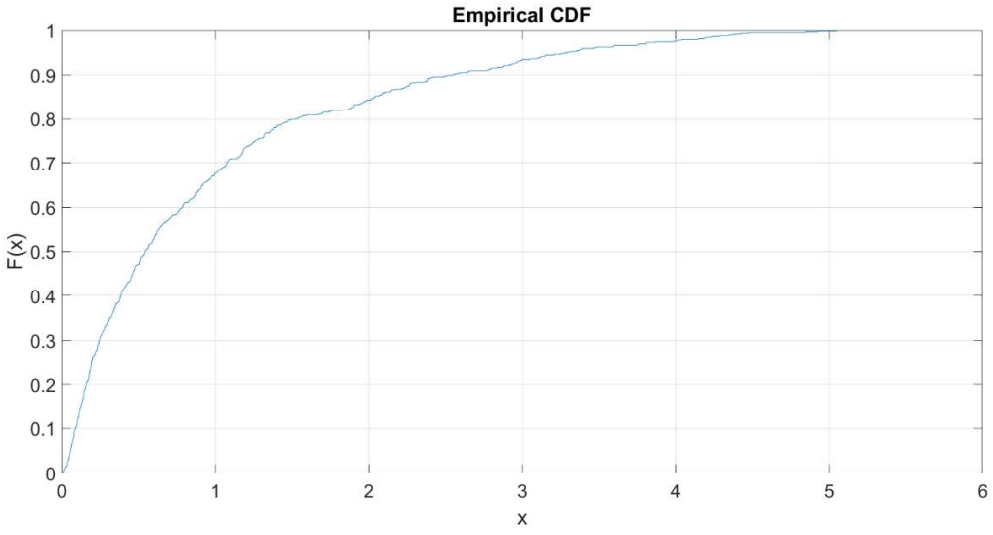


Figure 7.10 CDF of the test TS2_NOT_1000

Table 7.1 shows a brief comparison of the results obtained by this algorithm and an open-source algorithm called RTKLib (presented in Chapter 2) in the present test.

System	Mean Error	Mean std	Measurements/s
GPS (G. De Miguel)	0.9505m	1.0419m	1

Table 7.1 Algorithm vs RTKLib comparison for test TS2_NOT_1000

7.1.4 TS2_NOT_0100

The results of this test show the performance of the IMU stand-alone. Figure 7.11 shows the performance of the presented algorithm using IMU (black) against the Ground Truth (GT) (red). The algorithm gives a position estimation that starts drawing a path similar to the one given by the GT but displaced. Then, when time advances, the position estimation with IMU derives due to the cumulative errors introduced by this kind of sensors. In the end, the position estimation is out of the figure presented.

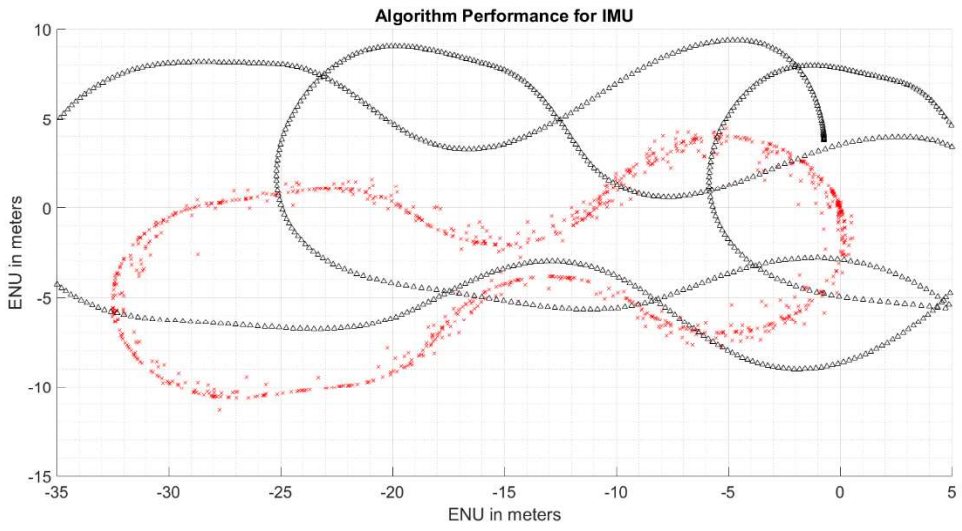


Figure 7.11 Performance of the algorithm for test TS2_NOT_0100

The mean of the error is 107.27 meters. However, there are errors up to 469.88 meters and the standard deviation of the error is 132.56 meters. The use of an IMU gives to the system a bigger number of position estimations. Fusion with other absolute positioning systems could take advantage of the mentioned

When analyzing the error histogram (see Figure 7.12), the error grows exponentially. It is seen that IMU stand alone is not a valid positioning option.

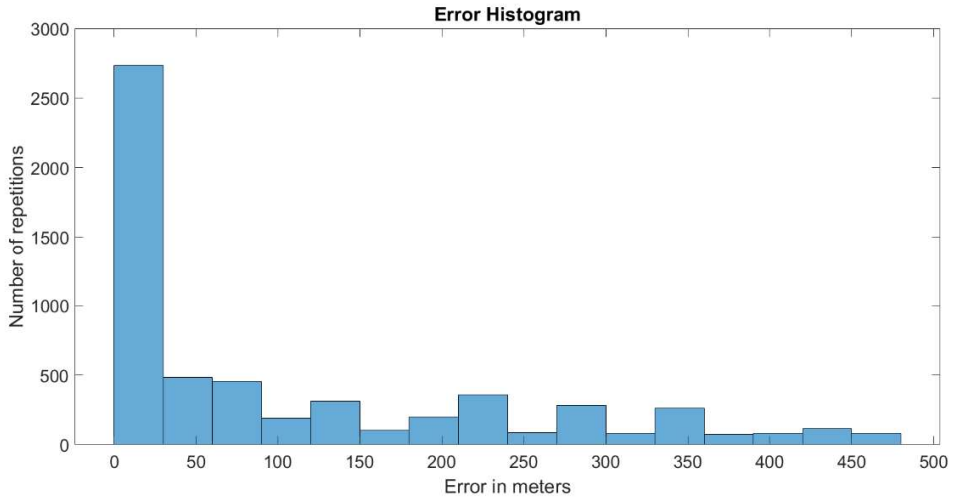


Figure 7.12 Histogram of the position error for test TS2_NOT_0100

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.13), shows that 95% of the position estimations have errors lower than 400 meters corroborating that IMU can not be used alone to position in big time-lapses.

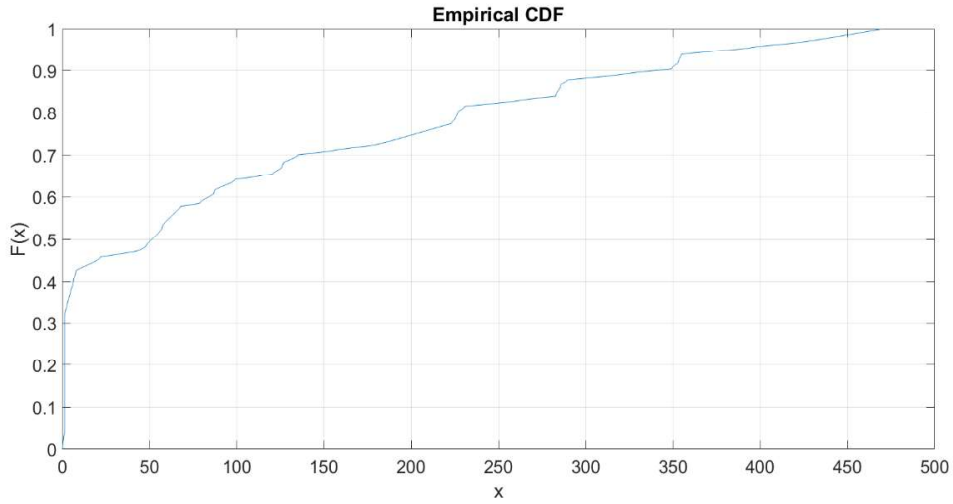


Figure 7.13 CDF of the test TS2_NOT_0100

Table 7.2 shows a brief comparison of the results obtained by this algorithm and an open-source algorithm called RTKLib in the present test.

System	Mean Error	Mean std	Measurements/s
IMU (G. De Miguel)	107.2744m	132.5698m	9.99

Table 7.2 Algorithm vs RTKLib comparison for test TS2_NOT_0100

7.1.5 TS2_NOT_0010

The results of this test show the performance of the UWB stand-alone. Figure 7.14 shows the performance of the presented algorithm using UWB (green) against the Ground Truth (GT) (red). The UWB gives a position estimation that draws a path which nearly matches the one given by the GT.

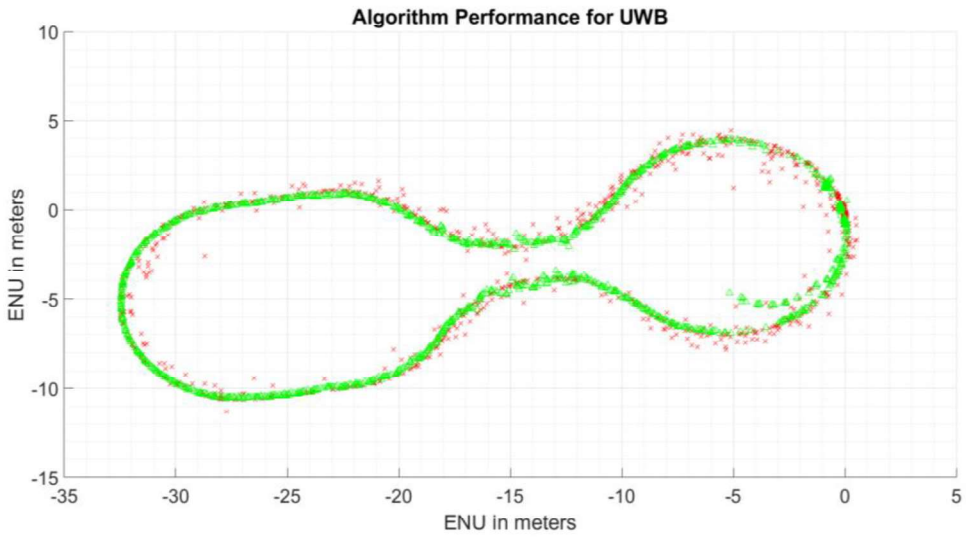


Figure 7.14 Performance of the algorithm for test TS2_NOT_0010

The mean of the error is 0.12 meters. However, there are errors up to 0.77 meters and the standard deviation of the error is 0.07 meters. The number of measurements is also acceptable.

When analyzing the error histogram (see Figure 7.15), most of the measurements have errors lower than half a meter due to the good positioning accuracy of the algorithm when using UWB. The accuracy of the measurements is stable during the whole test both in static and moving stages.

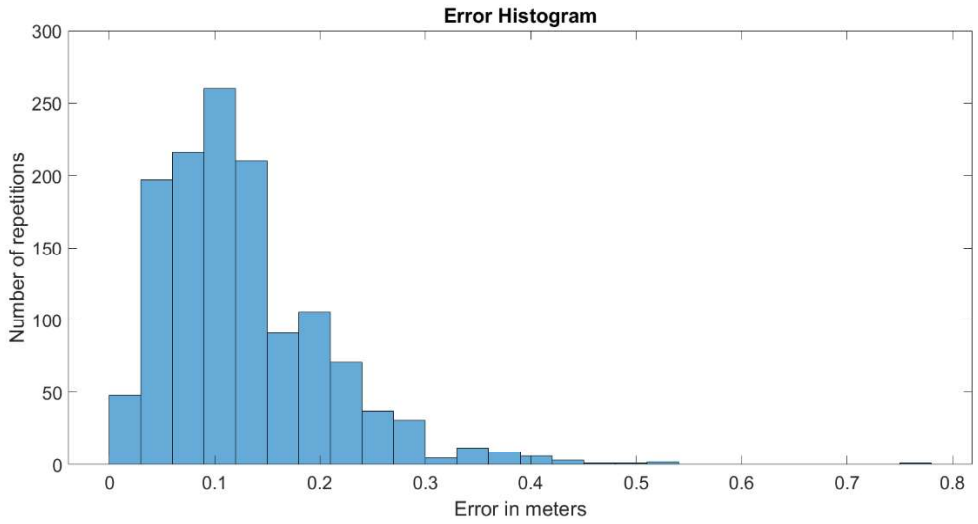


Figure 7.15 Histogram of the position error for test TS2_NOT_0010

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.16), shows that more than 95% of the position estimations have errors under thirty centimeters.

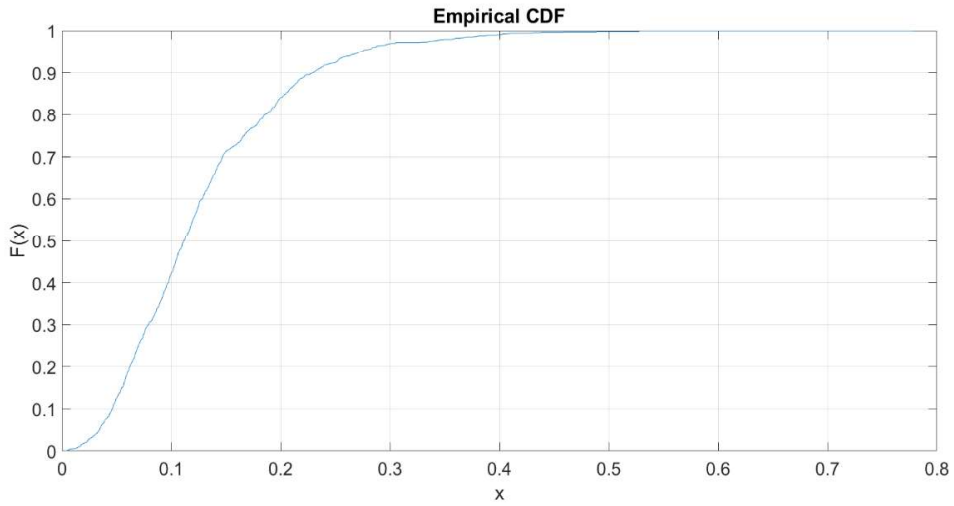


Figure 7.16 CDF of the test TS2_NOT_0010

Table 7.3 shows a brief summary of the performance achieved by the algorithm.

System	Mean Error	Mean std	Measurements/s
UWB (G. De Miguel)	0.1278m	0.0797m	2.21

Table 7.3 Algorithm vs RTKLib comparison for test TS2_NOT_0010

7.1.6 TS2_NOT_1100

The results of this test show the performance of GPS and IMU fusion. Figure 7.17 shows the performance of the presented algorithm using GPS (blue) and IMU (black) fusion against the Ground Truth (GT) (red). The algorithm gives a position estimation that draws a path in which the GT can be recognized. However, the obtained accuracy is not the best.

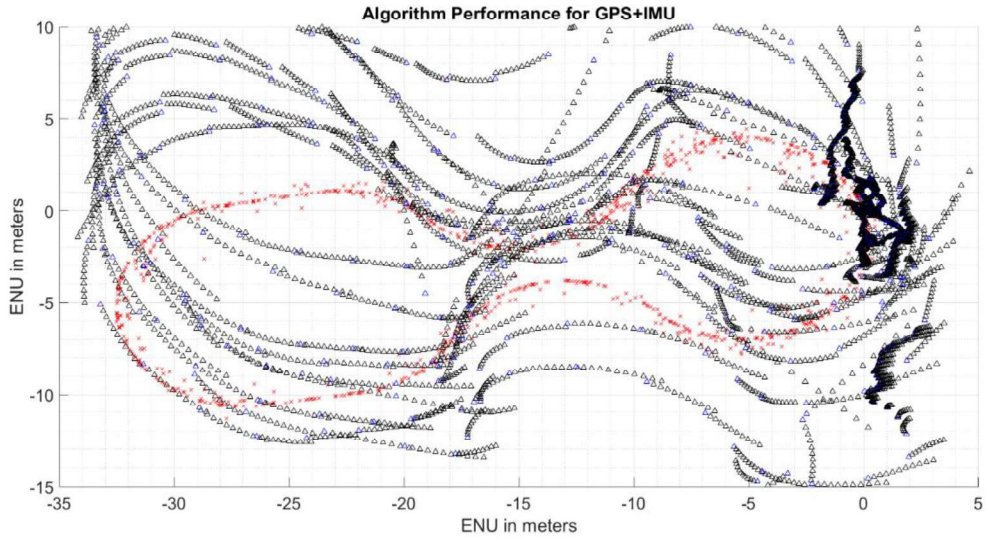


Figure 7.17 Performance of the algorithm for test TS2_NOT_1100

The mean of the error is 2.33 meters. However, there are errors up to 18.50 meters and the standard deviation of the error is 3.01 meters. The number of measurements is big due to the fusion of two systems.

When analyzing the error histogram (see Figure 7.18), most of the measurements have errors lower than four meters. However, some position estimations are bigger due to long IMU stand-alone times or misleading GPS data. A misleading yaw angle calculated by the GPS can lead the next estimations done by the IMU to be of bad quality.

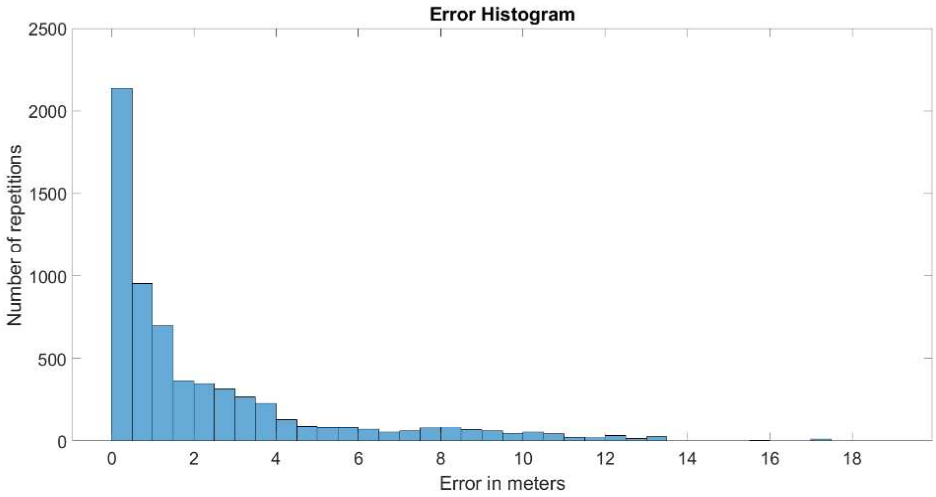


Figure 7.18 Histogram of the position error for test TS2_NOT_1100

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.19), shows that more than 95% of the position estimations have errors under nine meters and more than 80% of the measurements have errors lower than four meters.

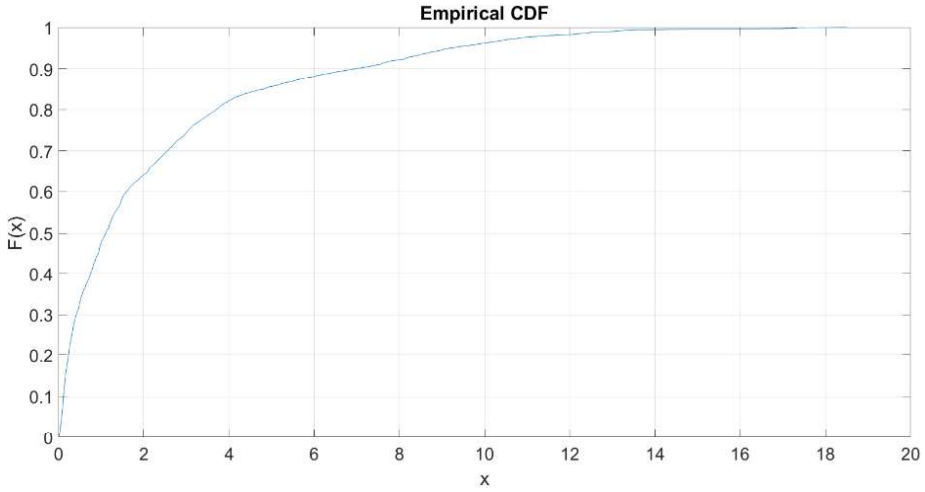


Figure 7.19 CDF of the test TS2_NOT_1100

Table 7.4 shows a brief summary of the performance achieved by the algorithm.

System	Mean Error	Mean std	Measurements/s
GPS+IMU (G. De Miguel)	2.3375m	3.0107m	10.99

Table 7.4 Algorithm vs RTKLib comparison for test TS2_NOT_1100

7.1.7 TS2_NOT_1101

The results of this test show the performance of the GPS and IMU fusion along with KBS. Figure 7.20 shows the performance of the presented algorithm using GPS(blue) and IMU (black) fusion against the Ground Truth (GT) (red). The algorithm gives a position estimation that draws a path in which the GT can be recognized. However, the accuracy obtained is not optimal.

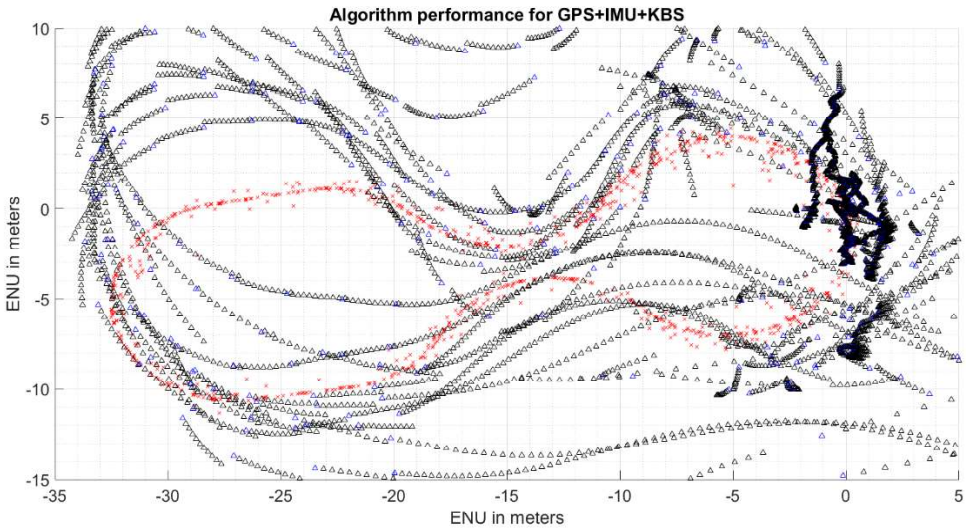


Figure 7.20 Performance of the algorithm for test TS2_NOT_1101

The mean of the error is 2.30 meters. However, there are errors up to 16.90 meters and the standard deviation of the error is 2.88 meters. The number of measurements is big due to the fusion of two systems.

When analyzing the error histogram (see Figure 7.21), most of the measurements have errors lower than four meters. However, some position estimations are bigger due to long IMU stand-alone times.

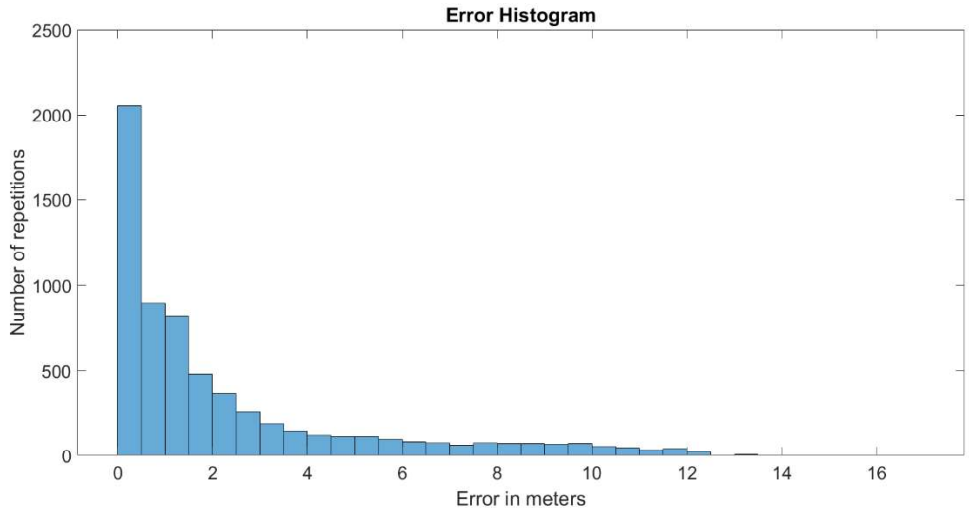


Figure 7.21 Histogram of the position error for test TS2_NOT_1101

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.22), shows that more than 95% of the position estimations have errors under nine meters and more than 80% of the measurements have errors lower than four meters.

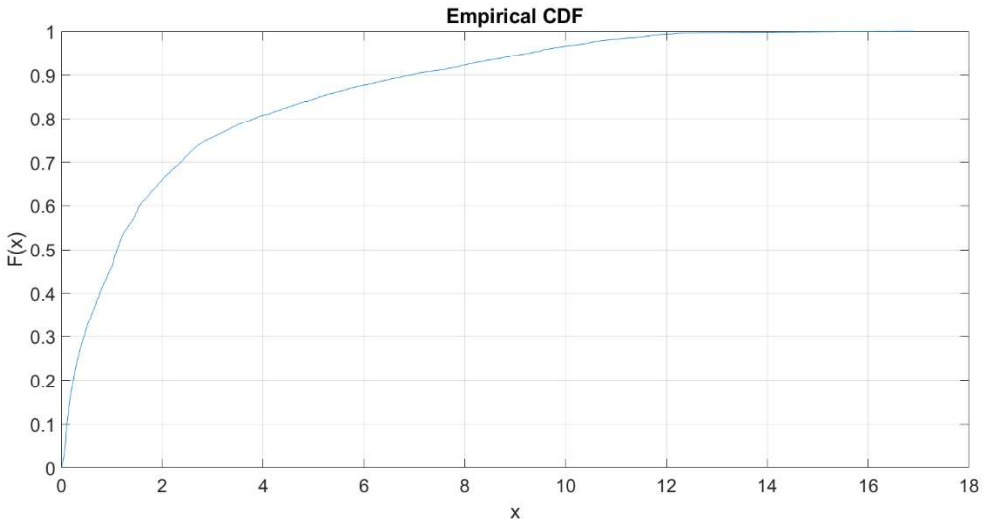


Figure 7.22 CDF of the test TS2_NOT_1101

Table 7.5 shows a brief summary of the performance achieved by the algorithm.

System	Mean Error	Mean std	Measurements/s
GPS+IMU+KBS (G. De Miguel)	2.2991m	2.8813m	10.89

Table 7.5 Algorithm vs RTKLib comparison for test TS2_NOT_1101

7.1.8 TS2_NOT_1010

The results of this test show the performance of GPS and UWB fusion. Figure 7.23 shows the performance of the presented algorithm using GPS (blue) and UWB (green) fusion against the Ground Truth (GT) (red). The algorithm gives a position estimation that draws a path in which the GT and the position estimation are nearly the same in most of the time of the test.

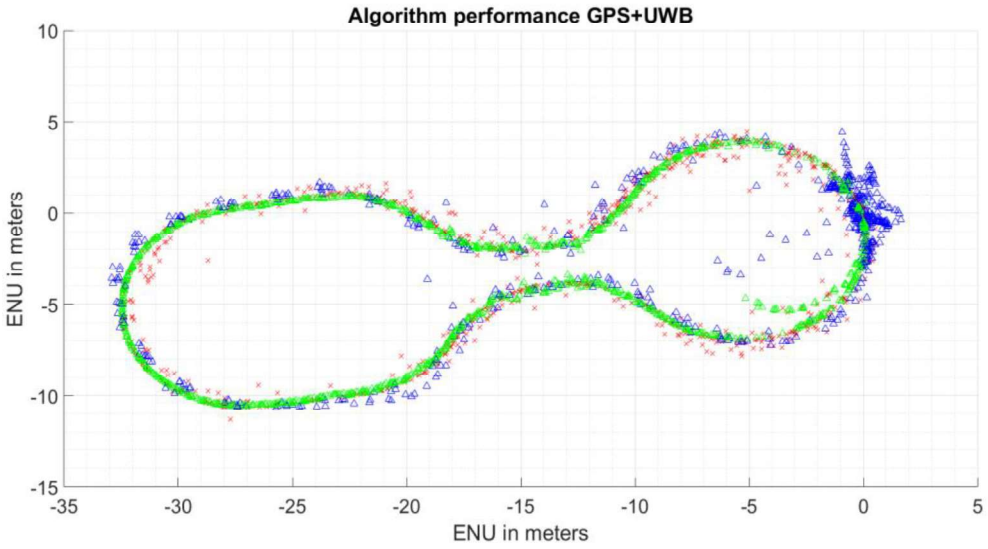


Figure 7.23 Performance of the algorithm for test TS2_NOT_1010

The mean of the error is 0.18 meters. However, there are errors up to 3.16 meters and the standard deviation of the error is 0.23 meters. The number of measurements is average due to the fusion of two systems.

When analyzing the error histogram (see Figure 7.24), most of the measurements have errors lower than half a meter. The accuracy of the measurements is stable during the whole test both in static and moving stages.

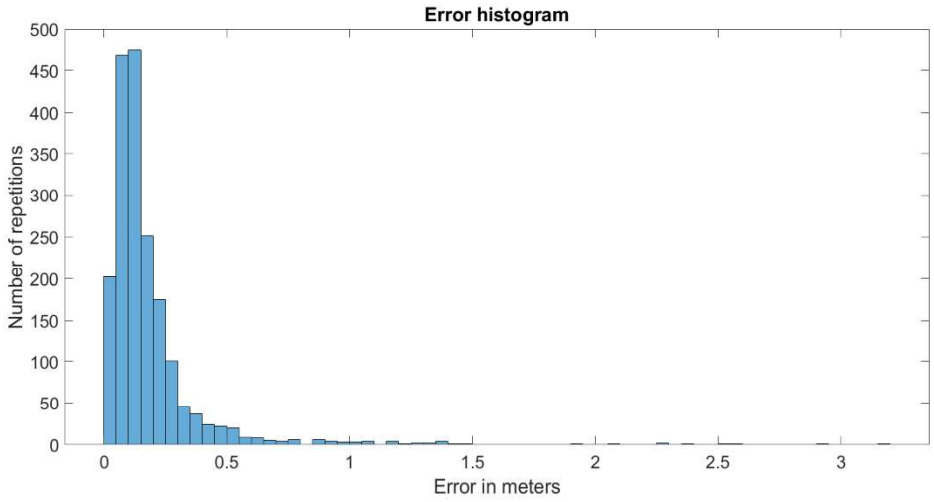


Figure 7.24 Histogram of the position error for test TS2_NOT_1010

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.25), shows that more than 95% of the position estimations have errors under fifty centimeters.

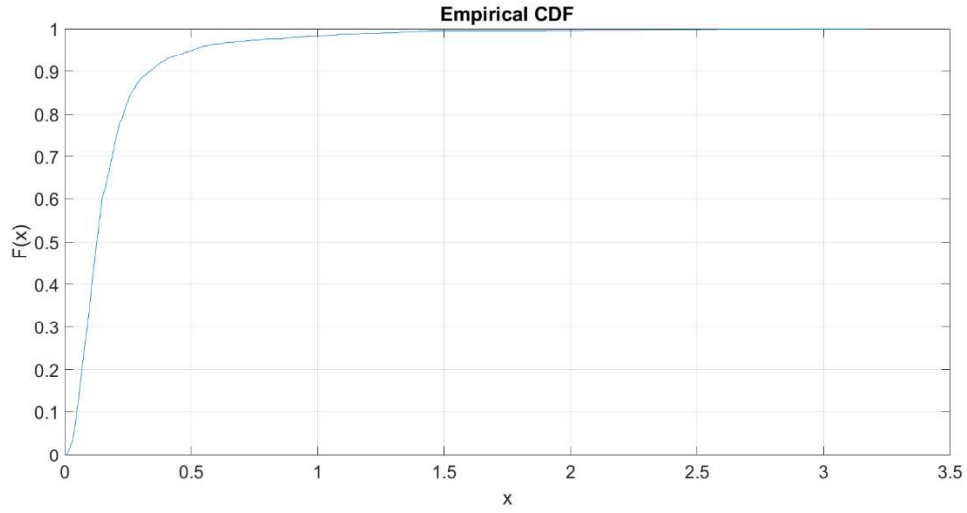


Figure 7.25 CDF of the test TS2_NOT_1010

Table 7.6 shows a brief summary of the performance achieved by the algorithm.

System	Mean Error	Mean std	Measurements/s
GPS+UWB (G. De Miguel)	0.1862m	0.2376m	3.21

Table 7.6 Algorithm vs RTKLib comparison for test TS2_NOT_1010

7.1.9 TS2_NOT_1011

The results of this test show the performance of the GPS and UWB fusion along with KBS. Figure 7.26 shows the performance of the presented algorithm using GPS (blue) and UWB (green) fusion against the Ground Truth (GT) (red). The algorithm gives a position estimation that draws a path in which the GT and the position estimation are nearly the same in most of the time of the test.

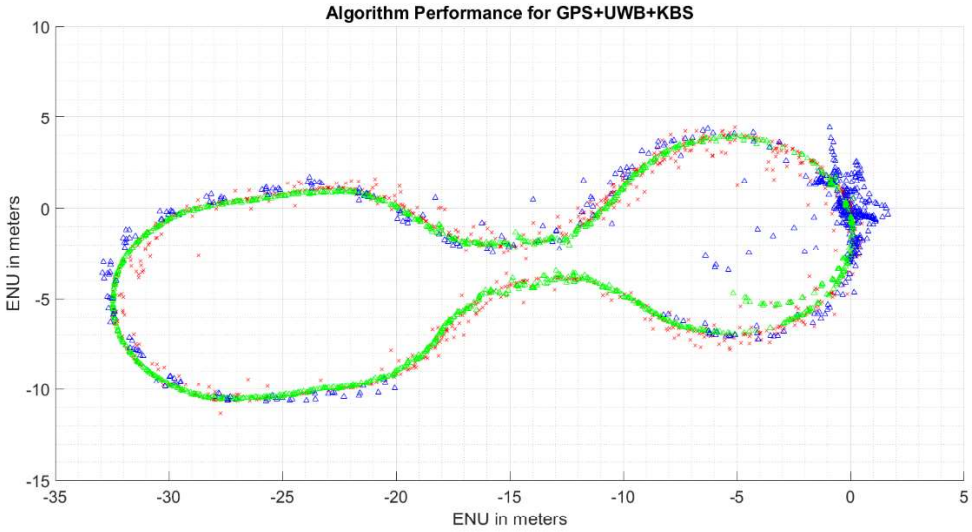


Figure 7.26 Performance of the algorithm for test TS2_NOT_1011

The mean of the error is 0.18 meters. However, there are errors up to 3.16 meters and the standard deviation of the error is 0.23 meters. The number of measurements is average due to the fusion of two systems.

When analyzing the error histogram (see Figure 7.27), most of the measurements have errors lower than half a meter. The accuracy of the measurements is stable during the whole test both in static and moving stages.

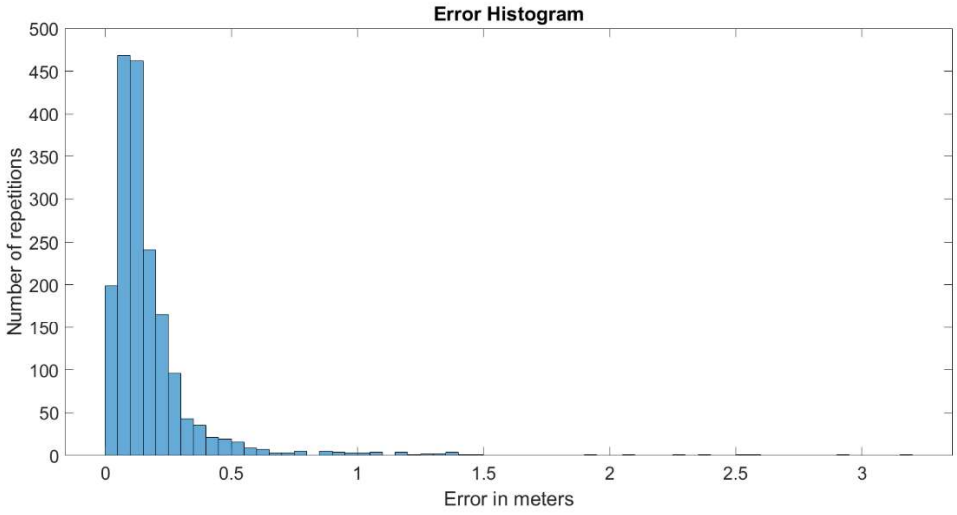


Figure 7.27 Histogram of the position error for test TS2_NOT_1011

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.28), shows that more than 95% of the position estimations have errors under fifty centimeters.

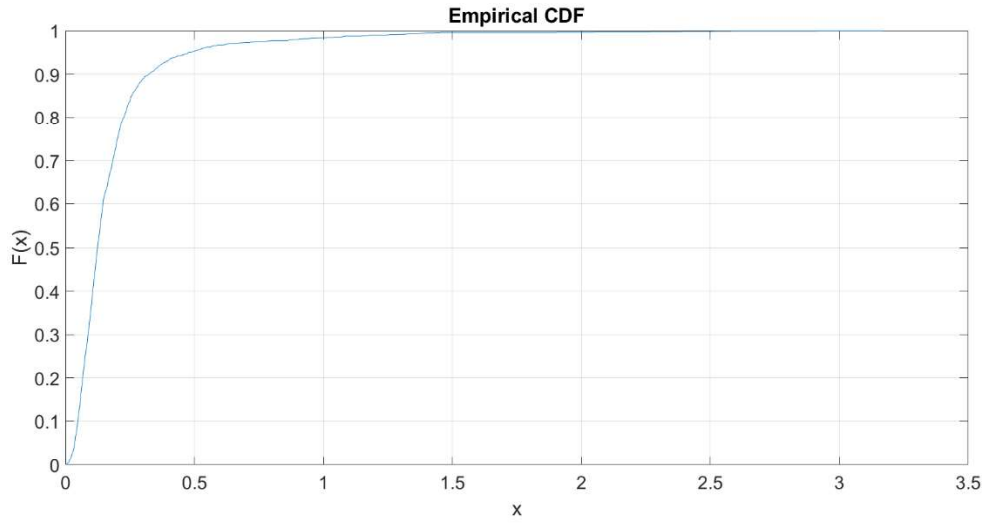


Figure 7.28 CDF of the test TS2_NOT_1011

Table 7.7 shows a brief summary of the performance achieved by the algorithm.

System	Mean Error	Mean std	Measurements/s
GPS+UWB+KBS (G. De Miguel)	0.1815m	0.2330m	3.1

Table 7.7 Algorithm vs RTKLib comparison for test TS2_NOT_1011

7.1.10 TS2_NOT_0110

The results of this test show the performance of the IMU and UWB fusion. Figure 7.29 shows the performance of the presented algorithm using IMU (black) and UWB (green) fusion against the Ground Truth (GT) (red). The algorithm gives a position estimation that draws a path in which the GT and the position estimation are nearly the same in most of the time of the test.

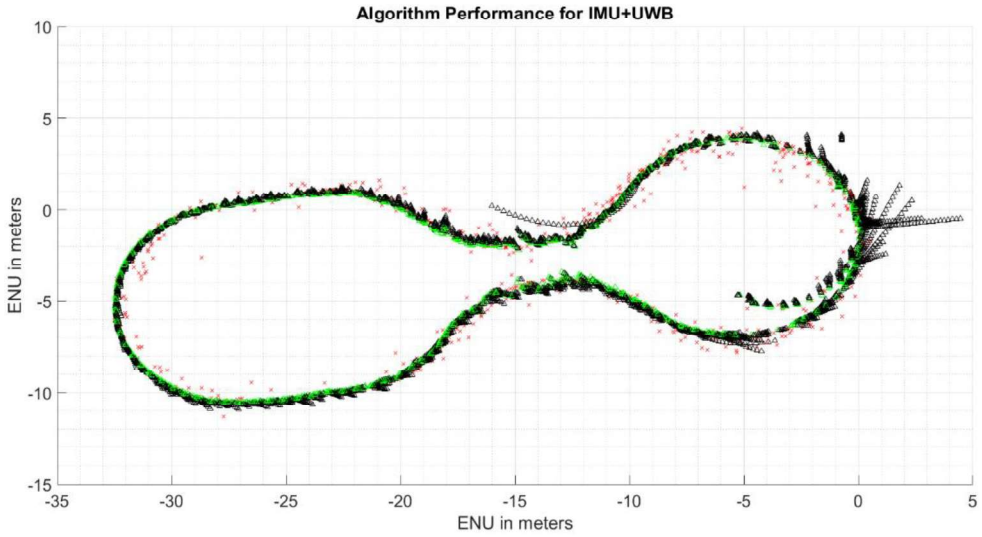


Figure 7.29 Performance of the algorithm for test TS2_NOT_0110

The mean of the error is 0.46 meters. However, there are errors up to 3.96 meters and the standard deviation of the error is 0.59 meters. The number of measurements is high due to the fusion of two systems with a high frequency of data output.

When analyzing the error histogram (see Figure 7.30), most of the measurements have errors lower than half a meter. However, there is a peak around 1.5 meters due to the initial and final drift of the IMU position while UWB is not in use.

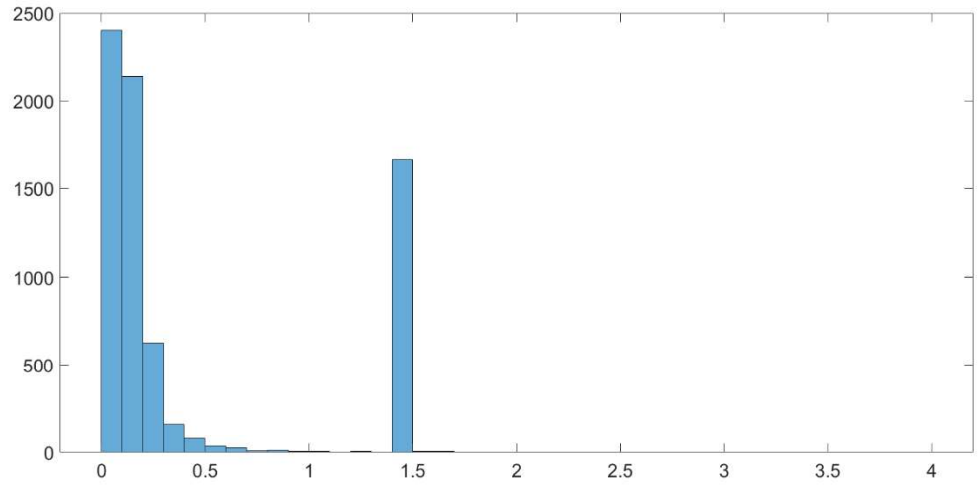


Figure 7.30 Histogram of the position error for test TS2_NOT_0110

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.31), shows that the 95% of the position estimations have errors unde a meter and a half, and that more than 75% of the position estimations have errors under thirty centimeters.

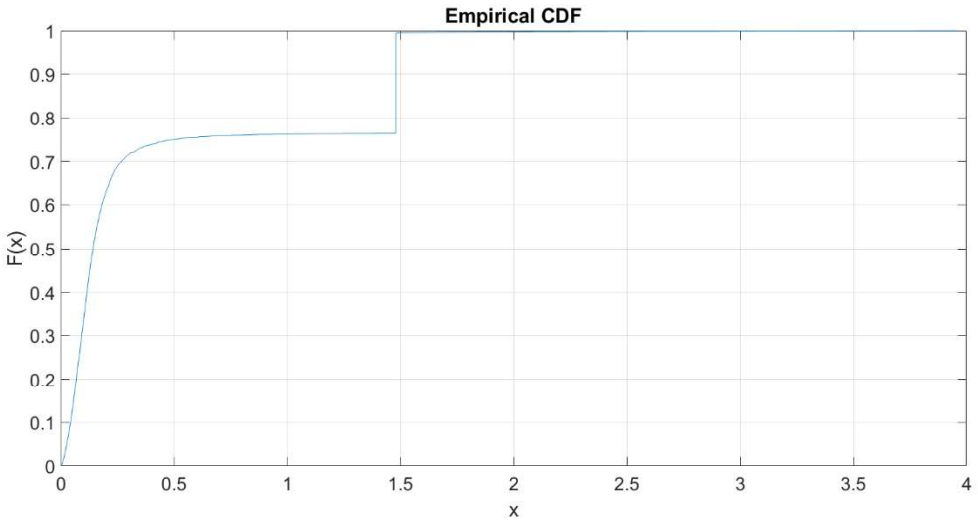


Figure 7.31 CDF of the test TS2_NOT_0110

Table 7.8 shows a brief summary of the performance achieved by the algorithm.

System	Mean Error	Mean std	Measurements/s
IMU+UWB (G. De Miguel)	0.4569m	0.5896m	12.2

Table 7.8 Algorithm vs RTKLib comparison for test TS2_NOT_0110

7.1.11 TS2_NOT_0111

Due to the nature of the KBS, explained in chapter 4, it has no sense to use it without GPS measurements. This test even if it was one of the permutations in the test, is not done.

7.1.12 TS2_NOT_1110

The results of this test show the performance of GPS, IMU, and UWB fusion. Figure 7.32 shows the performance of the presented algorithm using GPS(blue), IMU (black), and UWB (green) fusion against the Ground Truth (GT) (red). The algorithm gives a position estimation that draws a path in which the GT and the position estimation are nearly the same in most of the time of the test.

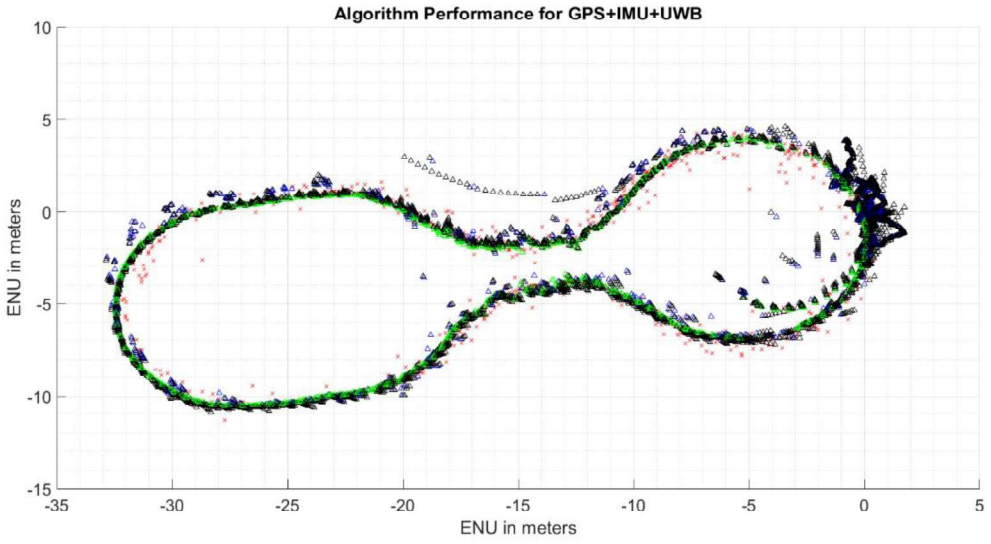


Figure 7.32 Performance of the algorithm for test TS2_NOT_1110

The mean of the error is 0.235 meters. However, there are errors up to 2.53 meters and the standard deviation of the error is 0.29 meters. The number of measurements is high due to the fusion of three systems, two of them with a high frequency of data output.

When analyzing the error histogram (see Figure 7.33), most of the measurements have errors lower than half a meter.

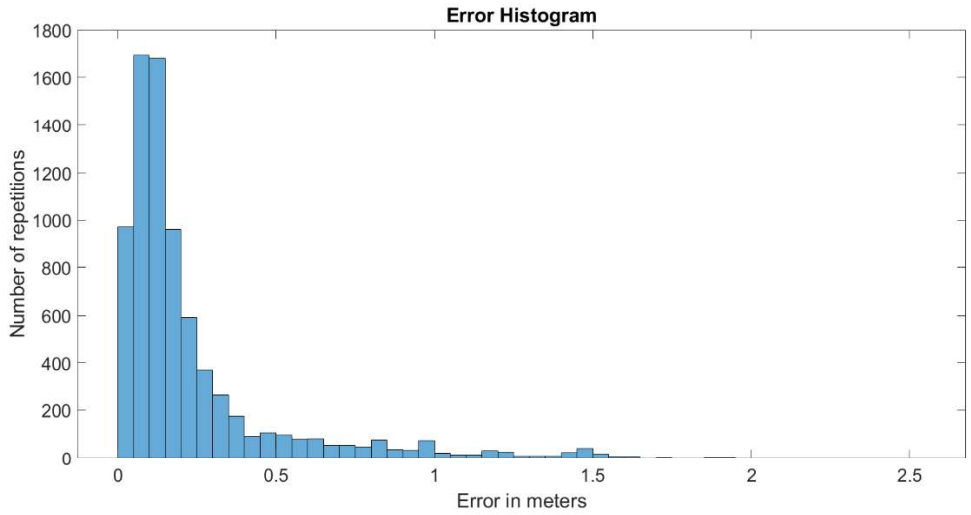


Figure 7.33 Histogram of the position error for test TS2_NOT_1110

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.34), shows that more than 75% of the position estimations have errors under twenty-five centimeters. Moreover, more than 95% of the position estimation errors are below one meter.

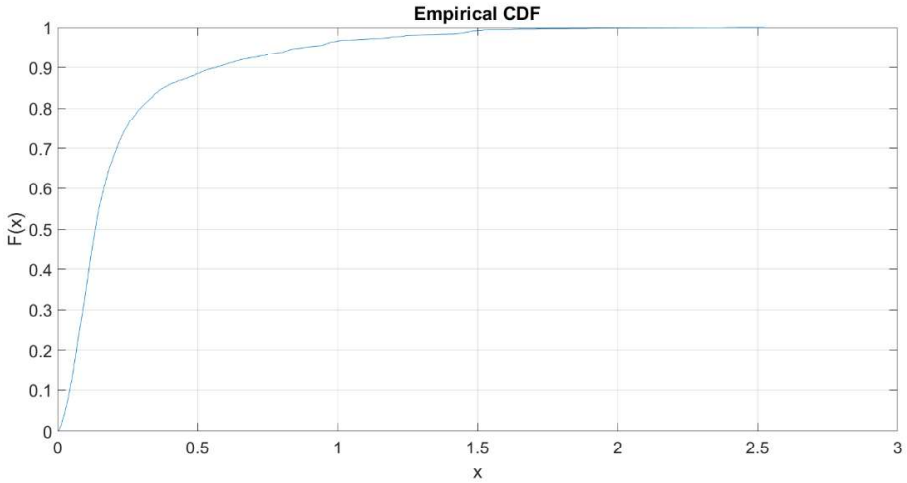


Figure 7.34 CDF of the test TS2_NOT_1110

Table 7.9 shows a brief summary of the performance achieved by the algorithm.

System	Mean Error	Mean std	Measurements/s
GPS+IMU+UWB (G. De Miguel)	0.2349m	0.2931m	13.2

Table 7.9 Algorithm vs RTKLib comparison for test TS2_NOT_1110

7.1.13 TS2_NOT_1111

The results of this test show the performance of GPS, IMU, and UWB fusion along with KBS. Figure 7.35 shows the performance of the presented algorithm using GPS(blue), IMU (black), and UWB (green) fusion against the Ground Truth (GT) (red). The algorithm gives a position estimation that draws a path in which the GT and the position estimation are nearly the same in most of the time of the test.

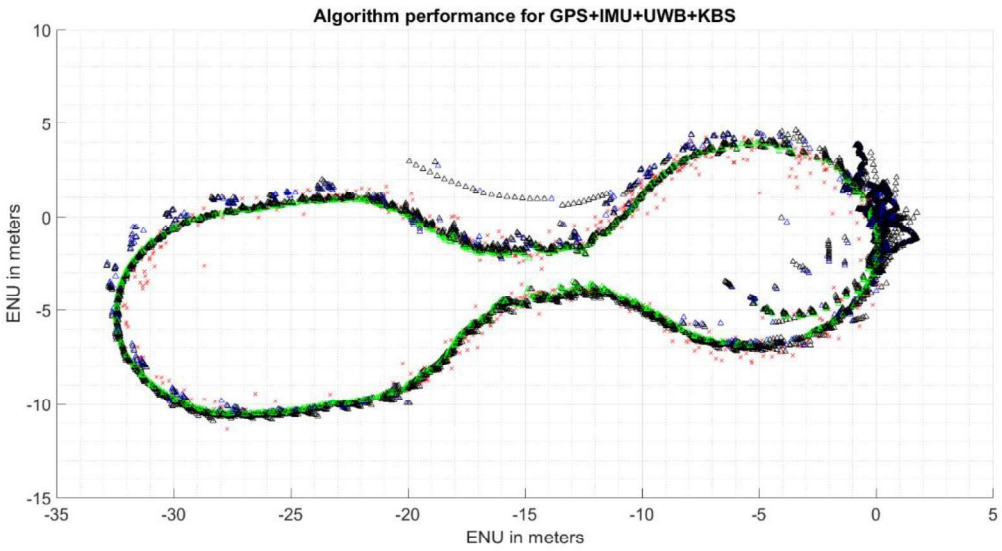


Figure 7.35 Performance of the algorithm for test TS2_NOT_1111

The mean of the error is 0.233 meters. However, there are errors up to 2.52 meters and the standard deviation of the error is 0.29 meters. The number of measurements is high due to the fusion of three systems, two of them with a high frequency of data output.

When analyzing the error histogram (see Figure 7.36), most of the measurements have errors lower than half a meter.

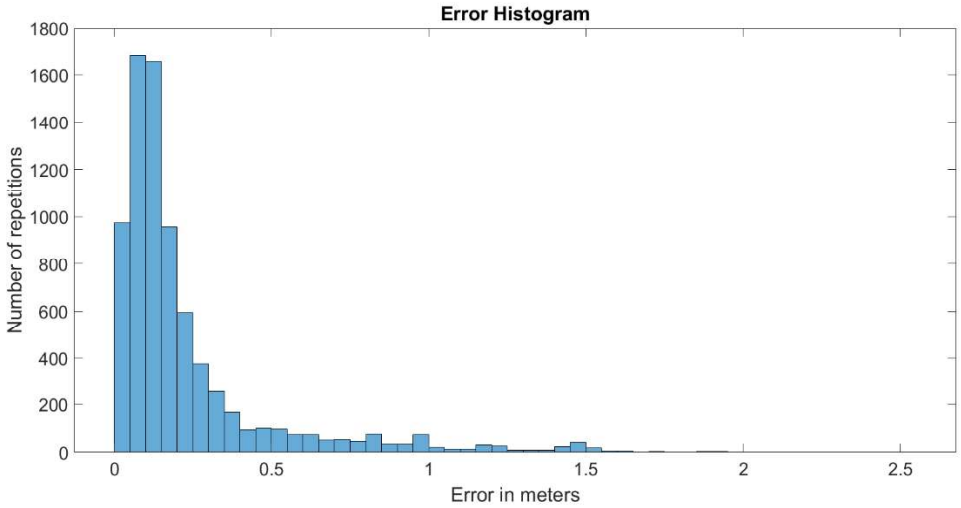


Figure 7.36 Histogram of the position error for test TS2_NOT_1111

Finally, the Cumulative Distribution Function (CDF) (see Figure 7.37), shows that more than 75% of the position estimations have errors under twenty-five centimeters. Moreover, more than 95% of the position estimation errors are below one meter.

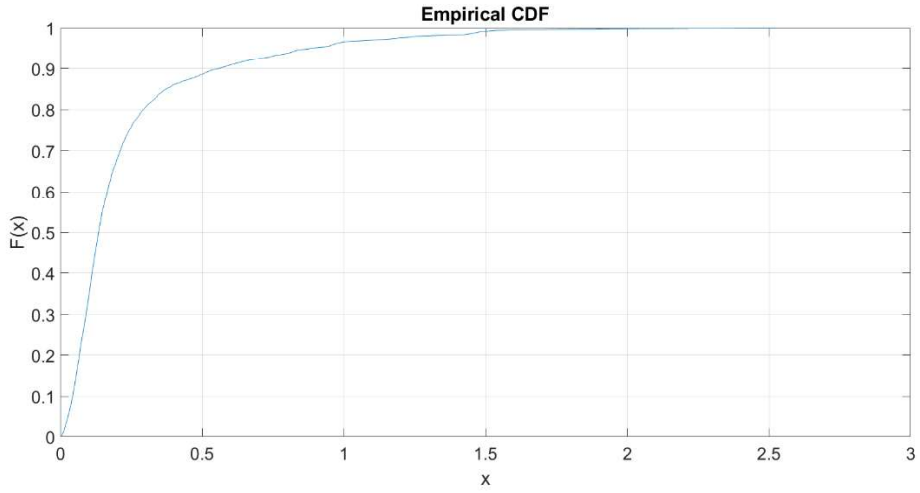


Figure 7.37 CDF of the test TS2_NOT_1111

Table 7.10 shows a brief summary of the performance achieved by the algorithm.

System	Mean Error	Mean std	measurements/s
GPS+IMU+UWB+KBS (G. De Miguel)	0.2336m	0.2910m	13.1

Table 7.10 Algorithm vs RTKLib comparison for test TS2_NOT_1111

7.2 PERFORMANCE DISCUSSION

After presenting the results of the tests and the performance of the algorithm in them, a brief discussion is done to analyze the strengths and weaknesses of the proposed algorithm.

The algorithm uses three positioning subsystems and a software technique to enhance the performance of their fusion in certain situations. Each subsystem has pros and cons, also each of the different fusions have them. All the possible permutations have been done in order to analyze them:

- GPS:
  - Pros: It is a global system; that makes it very useful as it has global coverage. It is cheap to implement and has reasonably good accuracy.
  - Cons: It has problems when positioning in areas in which the receiver has not an open sky view. Moreover, it is not available to use indoors or in tunnels. Output frequency could be not enough for some applications.
- IMU:
  - Pros: It has high rate output frequencies. It can perform positioning estimations indoor.
  - Cons: It needs a reference position to perform the positioning estimation in global frames. Its accuracy in stand-alone is poor and needs to reset regularly in order not to accumulate big errors.
- UWB:
  - Pros: It has accuracy with under meter errors with remarkably low error margins, with good output frequencies. Indoor positioning is possible maintaining the accuracies obtained outdoors.

  - Cons: It is not a global system. Anchors have to be installed in order to obtain a position. Each anchor has a limited range in which it can position the rover (max of about 100m) what increments the cost of Hardware installation.
- GPS+IMU:
  - Pros: This kind of fusion copes with the problem of having a low output frequency and also solves the problem of positioning where there is no open sky view.
  - Cons: The accuracy is not good enough for some environments of the railway operation such as the start of mission or track selectivity.
- GPS+UWB:
  - Pros: This fusion provides good accuracies in indoor and outdoor environments while global coverage is provided.
  - Cons: The cost of the system as many anchors have to be placed to obtain good accuracies with good continuity during the journey.
- IMU+UWB:
  - Pros: This fusión has a good accuracy along with a high output frequency and the possibility of indoor and outdoor positioning.
  - Cons: The positioning capability is available for small areas as anchors have to be placed. The big number of anchors needed can lead the resulting system to be expensive for some applications.
- GPS+IMU+UWB:
  - Pros: This fusion collects the pros of the three systems and so has a good positioning accuracy with a high output frequency and global availability.

  - Cons: Computational needs are higher than the other fusions as it uses more systems and more data are processed. More systems also imply a higher economic cost.
- KBS:
  - Pros: To use the KBS is not necessary to have any extra hardware, it is a software technique. Moreover, it can be used in any environment.
  - Cons: It is necessary to create a database before the operation with KBS.

In order to enhance the performance of the fusion alternatives introduced, KBS was developed. This software positioning enhancement uses map-aiding to obtain better performances when there are transitions into GNSS blocked spots. It has been created to be used in a railway environment taking into account its particularities.

In the tests performed in Nottingham, the impact of the KBS is not remarkable because the environments in which the train travels are GNSS friendly and the KBS is not triggered. In the tests performed in Nottingham, the GPS outages introduced are small and the use of UWB reduces the impact of the KBS. However, in this scenario, an enhancement in the order of centimetres can be seen in the results presented. Moreover, the second scenario is a mock railway in which the turns are sharp. In real railways, the turns cannot be so sharp for security issues. Sharp turns lead the IMU to have misleading positioning estimations and not to use the KBS most efficiently. Moreover, the accelerations accepted in different railway operations are limited to ensure security and comfort (See paragraph 2.2.1.1).

However, in scenario 1 (EuskoTren), the impact of the KBS is clear. A real railway is presented, there, with a tunnel. This kind of environment is the ideal environment to obtain the best performance for the KBS. The obtained results show that the performance of the KBS improves significantly the accuracy of the GNSS

and IMU fusion algorithm by detecting misleading information of the GNSS. The maximum error reached is five times lower and the mean error is improved by 50%.

After the analysis presented, this research work proposes a system using GPS, IMU, and UWB fusion along with KBS for the railway domain. GPS and IMU with KBS would be suitable to position in most of the railway, including some tunnels as shown in the tests performed for scenario 1.

For the start of mission and indoor positioning in stations, the use of UWB gives sufficient accuracy to be track selective, as shown in the tests done for scenario 2 in which all three systems are used. This strategy can also be useful for critical points in the railway in which high accuracies are needed such as track shifts or railway yards.

In long tunnels in which the IMU working stand-alone could accumulate errors during a long time, the installation of some UWB anchors could improve significantly the performance of the system as it could restart the IMU position reference and the accumulated error.

Chapter 8

Conclusions and future guidelines

This chapter summarizes the conclusions achieved in the presented dissertation. Moreover, it includes some open research areas that might be interesting to carry out future work on them.

8.1 CONCLUSIONS

Along with this dissertation, different positioning systems available in the literature have been analyzed. The analysis of the state of the art has pointed out the need for a positioning system capable to be track selective at least in some specific areas of the railway operation. Moreover, the impossibility in terms of accuracy or cost of the available systems in the literature to cope with these requirements has shown the necessity of a new approach.

Different fusion techniques have been studied to analyze if they could reach the requirements presented, but there was no record of accessible references for the railway domain. The open-source software with better performance found has been RTKLib (RTKLIB n.d.), which uses many enhancement techniques for GNSS positioning.

This thesis work has proposed and implemented a positioning algorithm with three positioning technique fusion (GPS, IMU, and UWB), along with a map aided software positioning enhancement technique called KBS. It can be stated that it makes progress beyond the state of the art for medium-low end positioning algorithms as demonstrated in Table 8.1. Additionally, the proposed algorithm has been proven to:

- Increase the number of positioning systems fused and has introduced a novel software technique to enhance the positioning under known blocked scenarios (KBS).
- Increase the accuracy in comparison with an open-source positioning software widely used as RTKLib (Table 8.2). Moreover, when using the full system along with KBS it has been proved to be track selective as the total error made when positioning is below a meter. Table 8.2 shows the improvement in the accuracy of the systems against RTKLIB for different railway operations introduced afterwards.

- The use of KBS reduces the number of faulty measurements as it has a FDE like behaviour, that discards GNSS measurements with misleading information.
- Be an affordable system to be used in many applications as the positioning hardware used is a medium-low-end.

	U-Center (U-blox)	RTKLib	CargoTrac (SAVVY)	G. De Miguel
Systems:				
GPS	Yes	Yes	Yes	Yes
IMU	No	No	No	Yes
UWB	No	No	No	Yes
KBS	No	No	No	Yes
GPS multi-frequency multi-constellation	Yes	Yes	Multi-constellation	No
Accuracy	1-meter	2-3meter	N/A	Under a meter
Availability indoor	No	No	No	Yes
Open-Source	No	Yes	No	No
Positioning engine	Not available	LS/Kalman	Not available	LS/Kalman
Use sector	General	General	Railway	Railway

Table 8.1 Comparison of different positioning systems available vs presented positioning algorithm

System	Mean Error	Mean std	Measurements/s
RTKLib	4.84m	3.77m	1
U-Center	4.54m	1.90m	1
GPS+IMU (G. De Miguel)	2.33m	3.01m	10.99
GPS+IMU+KBS (G. De Miguel)	2.29m	2.88m	10.89
IMU+UWB (G. De Miguel)	0.45m	0.58m	12.2
GPS+IMU+UWB+KBS (G. De Miguel)	0.23m	0.29m	13.1

Table 8.2 Performance of the algorithm vs RTKLib performance

This research thesis has analyzed different combinations for fusion and their performance in Chapter 7. Taking into account the performance results obtained, a combination of different fusions for different railway operations in order to be cost-effective and maintain its necessities has been proposed:

- GPS and IMU fusion is proposed for common operation as it is suitable to provide a reliable position estimation in most of the railway, including some short tunnels.
- For the start of mission and indoor positioning in stations, the use of UWB is proposed. It gives sufficient accuracy to be track selective. The cost of installing UWB anchors in stations is affordable within the retro-fitting or evolution of a signalling system, and its fusion with IMU is appropriate for this kind of scenario. Moreover, the system could take advantage of the KBS in the transition from outdoor to indoor where GPS is not reliable enough for this requirements. This strategy can also

be useful for critical points in the railway in which high accuracies are needed such as track shifts or railway yards.

- In long tunnels in which the IMU working stand-alone could accumulate errors during a long time, the installation of some UWB anchors is proposed to improve significantly the performance of the system as UWB anchors can be used to restart the IMU position reference and the accumulated error.

Having a look back into Chapter 3, in which objectives were introduced, the presented thesis fulfils the objectives presented:

- **Introduction and analysis of the existing positioning systems.**
  During chapter 2 the positioning systems existing and the possible fusion between them have been analyzed.

- **Design of a novel positioning system and algorithm.**
  A novel system and algorithm has been implemented and explained in Chapter 4. Moreover, a novel techniques and the fusion of three different positioning systems has been done.

- **Improvement of the availability of the system.**
  As three systems have been used to position, the availability of the resulting system has been increased in comparison with the availability of GPS stand-alone positioning. In Table 8.2 the availability for each system for the same time length is shown.

- **Fusion of GNSS, IMU, and UWB technologies for transport systems.**
  The three positioning systems introduced have been successfully fused.

- **Improvement of the accuracy in harsh scenarios for GNSS systems.**
  With the introduction of the KBS and a proper fusion of the three technologies, the accuracy has been significantly improved as shown in Chapter 7.

- **Test and validation of the proposed system and algorithms.**
  The system has been tested in Chapter 7 and its results have been analyzed for diverse reasonable scenario for the railway operation and conditions.

- **Analysis of the difference between using snapshot and recursive algorithms in terms of accuracy.**
  Taking into account the literature presented in Chapter 2, a recursive algorithm has appeared to be better in terms of positioning accuracy.

- **Analysis of the accuracy improvement when using a Fault Detection and Exclusion (FDE) method.**
  Taking into account the literature presented in Chapter 2, the use of an FDE method helps to obtain a better accuracy as faulty satellites are discarded.

- **Analysis of the performance obtained by the system created.**
  An analysis of the full system with the fusion of GPS, IMU, and UWB along with the KBS strategy has been presented in Chapter 7.

8.2 FUTURE GUIDELINES

The main objective of this research work was to present and assess a multi-sensor positioning algorithm for the railway industry.

During this work, several topics have been found that deserve further research:

- **Extension of the compatible Input files**
  The presented algorithm uses a proprietary format for the input files used. The adequation of the code to use standardized input files such as RINEX files would lead the algorithm to wider use possibilities. Moreover, the introduction of parsers for more hardware to use in real-time would make the algorithm more flexible.

- **Multi-frequency and Multi-constellation positioning**
  The algorithm introduced in the present investigation work uses GNSS technology in a single-frequency (L1) and single-constellation (GPS) mode. The use of a multi-frequency multi-constellation approach would lead the algorithm to better positioning accuracies for positioning estimations involving GNSS. Moreover, the use of multi-constellation GNSS positioning would improve the availability of the GNSS subsystem.

- **Implementation of a tighter Kalman filter**
  The analysis of the state of the art has pointed out that the use of a tighter Kalman filter would lead the system to better accuracies. However, a higher computational effort would be required.

- **Use of machine learning for the selection of the technology used**
  The proposed algorithm selects the weight given to each technology depending on the availability of them and the standard deviation of the

measurements. Moreover, the KBS software strategy is also taken into account. The use of machine learning could help to explore other weighting strategies that might provide better performance of the algorithm.

- **Implementation of antispoofing techniques**
  With the generalization of the use of GNSS positioning techniques, spoofing strategies have appeared to be used by system crackers. Even if the presented algorithm has fused of more than one positioning system that could cope with a jamming attack, a spoofing attack could not be that easily overcome. That is the reason to recommend further work in antispoofing techniques to make the presented algorithm secure enough to be used in the railway environment.

- **Introduction of Integrity evaluation techniques**
  During this dissertation, an analysis of the accuracy performance of the positioning estimation algorithm has been presented. However, the quality of a positioning system is not only measured in terms of accuracy but also in terms of integrity. A brief integrity analysis has been done by providing a horizontal protection level using 7-sigma factor for the system, with hopeful results. However, the analysis needed to validate the integrity of the system requires further work.

References

Adjrad, Mounir, and Paul D. Groves. 2017. "Enhancing Least Squares GNSS Positioning with 3D Mapping without Accurate Prior Knowledge." *Navigation* 64(1): 75–91. http://doi.wiley.com/10.1002/navi.178 (April 12, 2020).

AIOSAT. "Home - AIOSAT." http://www.aiosat.eu/ (April 17, 2020).

Arrizabalaga, Saioa, Gorka De Miguel, Jon Goya, and Leticia Zamora-cadenas. 2016. "Data Acquisition for Complementary Positioning System in GNSS-Denied Areas." *11th World Congress on Railway Research (WCRR 2016)*: 1–6.

B5m. 2020. "Web 1:5000." https://b5m.gipuzkoa.eus/web5000/es/ (June 24, 2020).

"Beidou Satellite Navigation System to Cover Whole World in 2020." 2011. Eng.chinamil.com.cn.

Ben-Moshe, Boaz, Elazar Elkin, Harel Levi, and Ayal Weissman. *Improving Accuracy of GNSS Devices in Urban Canyons* *.

Berrocoso, M, ME Ramírez, and A Pérez-Peña. "El Sistema de Posicionamiento Global."

Bocquet, Michael, Christophe Loyez, and Aziz Benlarbi-Delaï. 2005. "Millimeter Wave Up-Converted UWB Based Positioning System." In *ACM International Conference Proceeding Series*, New York, New York, USA: ACM Press, 293–96. http://portal.acm.org/citation.cfm?doid=1107548.1107619 (April 12, 2020).

Bradbury, J. et al. 2007. "Code Multipath Modelling in the Urban Environment Using Large Virtual Reality City Models: Determining the Local Environment." *Journal of Navigation* 60(1): 95–105.

China Satellite Navigation Office. 2012. *BeiDou Navigation Satellite System Signal In Space Interface Control Document*.

Connor, Piers. “Rules for High Speed Line Capacity or, HOw to Get a Realistic Capacity Figure for a High Speed Line.” (3): 1–8.

“Constellation Information | European GNSS Service Centre.” https://www.gsc-europa.eu/system-service-status/constellation-information (April 13, 2020).

Decawave. “EVK1000 Evaluation Kit - Decawave.” https://www.decawave.com/product/evk1000-evaluation-kit/ (April 12, 2020).

“Directions 2020: BeiDou in the New Era of Globalization - GPS World : GPS World.” https://www.gpsworld.com/directions-2020-beidou-in-the-new-era-of-globalization/ (April 13, 2020).

Duran, Mauricio A.Caceres et al. 2012. “Terrestrial Network-Based Positioning and Navigation.” In *Satellite and Terrestrial Radio Positioning Techniques*, Elsevier Ltd, 75–153.

EATS. “Final Report Summary - EATS (ETCS Advanced Testing and Smart Train Positioning System) | Report Summary | EATS | FP7 | CORDIS | European Commission.” https://cordis.europa.eu/project/id/314219/reporting/es (April 10, 2020).

ECMA. 2008. *High Rate Ultra Wideband PHY and MAC Standard*. www.ecma-international.org (April 12, 2020).

ERRAC. 2012. “Rail Route 2050: The Sustainable Backbone of the Single European Transport Area | ERRAC.” *Rail Route 2050*. https://errac.org/publications/rail-route-2050-the-sustainable-backbone-of-the-single-european-transport-area/ (April 5, 2020).

“ERSAT-EAV.” 2017. http://www.ersat-eav.eu/ (April 7, 2020).

“ESA - Galileo Services.” https://www.esa.int/Applications/Navigation/Galileo/Galileo_services (April 13, 2020).

Explorer, RTKLib. 2018. “Comparison of the U-Blox M8T to the u-Blox M8P –

Rtklibexplorer." https://rtklibexplorer.wordpress.com/2018/08/14/comparison-of-the-u-blox-m8t-to-the-u-blox-m8p/ (May 20, 2020).

Faragher, Ramsey. 2012. "Understanding the Basis of the Kalman Filter via a Simple and Intuitive Derivation [Lecture Notes]." *IEEE Signal Processing Magazine* 29(5): 128–32.

FCC. 2002. "Federal Communications Commission of the United States. Revision of Part 15 of the Commission's Rules Regarding Ultra WideBand Transmission Systems." : 98–153.

FR8RAIL. "Development of Functional Requirements for Sustainable and Attractive European Rail Freight | FR8RAIL Project | H2020 | CORDIS | European Commission." https://cordis.europa.eu/project/id/730617 (April 7, 2020).

"Galileo General Introduction - Navipedia." https://gssc.esa.int/navipedia/index.php/Galileo_General_Introduction (April 13, 2020).

Gao, Yang et al. *UWB/GNSS-Based Cooperative Positioning Method for V2X Applications*.

GIS. 2011. "Dilution of Precision (GPS) - GIS Wiki | The GIS Encyclopedia." http://wiki.gis.com/wiki/index.php/Dilution_of_precision_(GPS) (May 22, 2020).

"GLONASS Space Segment - Navipedia." https://gssc.esa.int/navipedia/index.php/GLONASS_Space_Segment (April 13, 2020).

Government, Chinese. 2012. "Beidou Roadmap." http://www.beidou.gov.cn/2012/12/14/2012121481ba700d7ca84dfc9ab2ab9ff33d2772.html.

"GPS.Gov: New Civil Signals." https://www.gps.gov/systems/gps/modernization/civilsignals/ (April 13, 2020).

"GPS.Gov: Space Segment." https://www.gps.gov/systems/gps/space/ (April 13,

2020).

Groves, Paul D. 2008. GNSS Technology and Applications Series *Principles of GNSS Inertial and Multi-Sensor Integrated Navigation Systems - GNSS Technology and Applications*. Artech House.

Groves, Paul D, Lei Wang, Mounir Adjrad, and Claire Ellul. 2015. *GNSS Shadow Matching: The Challenges Ahead*.

Groves, PD, M Adjrad, H Gao, and CD Ellul. 2016. "Intelligent GNSS Positioning Using 3D Mapping and Context Detection for Better Accuracy in Dense Urban Environments." *In: (Proceedings) INC 16: International Navigation Conference, 8-10 November 2016, Glasgow, UK. Royal Institute of Navigation (2016)* .

GSA. 2018. "USER TECHNOLOGY REPORT A U T O M AT I O N." https://www.gsa.europa.eu/contact-us, (May 20, 2020).

Han, Houzeng et al. 2019. "An Emergency Seamless Positioning Technique Based on Ad Hoc UWB Networking Using Robust EKF." *Sensors (Switzerland)* 19(14).

Haseltine, Eric L, and James B Rawlings. 2005. "Critical Evaluation of Extended Kalman Filtering and Moving-Horizon Estimation." *Ind. Eng. Chem. Res* (8): 2451–60. http://pubs.acs.org (April 12, 2020).

Hewitson, Steve, and Jiling Wang. *Extended Receiver Autonomous Integrity Monitoring (ERAIM) for GNSS/INS Integration*.

Hillenbrand, Wolfgang, and Holm Hofestädt. *GSM-R Traffic Model for Radio-Based Train Operation*.

Hsu, Li Ta, Yanlei Gu, and Shunsuke Kamijo. 2016. "3D Building Model-Based Pedestrian Positioning Method Using GPS/GLONASS/QZSS and Its Reliability Calculation." *GPS Solutions* 20(3): 413–28.

Hutchinson, Michael. 2016. "Demonstration of Advanced GNSS-Based Location on

the UK ERTMS Test Train." In *INC Glasgow 2016*,.

"IERS - IERS - The International Terrestrial Reference System (ITRS)." https://www.iers.org/IERS/EN/Science/ITRS/ITRS.html (April 13, 2020).

Isaacs, Jason T. et al. 2014. "Bayesian Localization and Mapping Using GNSS SNR Measurements." In *Record - IEEE PLANS, Position Location and Navigation Symposium*, Institute of Electrical and Electronics Engineers Inc., 445–51.

Julier, Simon J, and Jeffrey K Uhlmann. *A New Extension of the Kalman Filter to Nonlinear Systems*.

Kaplan, Elliott D., and C. (Christopher J.) Hegarty. 2006. *Understanding GPS : Principles and Applications*. Artech House.

Kolås, S., B. A. Foss, and T. S. Schei. 2009. "Constrained Nonlinear State Estimation Based on the UKF Approach." *Computers and Chemical Engineering* 33(8): 1386–1401.

Kumar, Rakesh, and Mark G. Petovello. 2016. "Sensitivity Analysis of 3D Building Modelassisted Snapshot Positioning." In *29th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2016*, Institute of Navigation, 1285–95.

Kumar, Rakesh, and Mark G Petovello. 2014. *A Novel GNSS Positioning Technique for Improved Accuracy in Urban Canyon Scenarios Using 3D City Model*.

Langley, Richard B. 1998. "GPS Receivers and the Observables." In *GPS for Geodesy*, Springer Berlin Heidelberg, 151–85.

Langley, Richard B. 2008. "Propagation of the GPS Signals." In *GPS for Geodesy*, Springer Berlin Heidelberg, 103–40.

Lee, Kevin. "What Is an Elevation Mask for GPS?" *itstillworks*. https://itstillworks.com/elevation-mask-gps-18007.html (June 25, 2020).

Leica. "Leica Nova MS60: La Primera MultiEstación de Autoaprendizaje Del Mundo

| Leica Geosystems.” https://leica-geosystems.com/es-es/products/total-stations/multistation/leica-nova-ms60 (April 18, 2020).

Leick, Alfred, Lev Rapoport, and Dmitry Tatarnikov. 2015. GPS Satellite Surveying: Fourth Edition *GPS Satellite Surveying: Fourth Edition*. Hoboken, NJ, USA: wiley. http://doi.wiley.com/10.1002/9781119018612 (April 10, 2020).

Liu, Hui et al. 2017. “Accounting for Inter-System Bias in DGNSS Positioning with GPS/GLONASS/BDS/Galileo.” *Journal of Navigation* 70(4): 686–98.

Macgougan, Glenn D, and Richard Klukas. 2009. *Method and Apparatus for High Precision GNSS/UWB Surveying*.

Managers, Applied GPS for Engineers and Project. *Calculation of Satellite Position from Ephemeris Data*.

de Miguel, Gorka et al. 2019. “Map-Aided Software Enhancement for Autonomous GNSS Complementary Positioning System for Railway.” *IEEE Transactions on Vehicular Technology* 68(12): 11611–20.

Miguel, Gorka de et al. 2019. “Map-Aided Software Enhancement for Autonomous GNSS Complementary Positioning System for Railway.” *IEEE Transactions on Vehicular Technology*. https://doi.org/10.1109/TVT.2019.2940621.

De Miguel, Gorka et al. 2017. “GNSS Complementary Positioning System Performance in Railway Domain.” In *Proceedings of 2017 15th International Conference on ITS Telecommunications, ITST 2017*, Institute of Electrical and Electronics Engineers Inc.

Navigation, Advanced. *ADVANCED NAVIGATION-ORIENTUS*.

Navipedia-Klobuchar. “Klobuchar Ionospheric Model - Navipedia.” https://gssc.esa.int/navipedia/index.php/Klobuchar_Ionospheric_Model (May 2, 2020).

Nocedal, Jorge, and Stephen J Wright. *Numerical Optimization Second Edition*.

Noureldin, Aboelmagd, Tashfeen B. Karamat, and Jacques Georgy. 2013. Fundamentals of Inertial Navigation, Satellite-Based Positioning and their Integration *Fundamentals of Inertial Navigation, Satellite-Based Positioning and Their Integration*. Springer Berlin Heidelberg.

Petovello, Mark G. *Real-Time Integration of a Tactical-Grade IMU and GPS for High-Accuracy Positioning and Navigation*.

Powell, J. P., and R. Palacín. 2015. "Passenger Stability Within Moving Railway Vehicles: Limits on Maximum Longitudinal Acceleration." *Urban Rail Transit* 1(2): 95–103.

"RAIM - Navipedia." https://gssc.esa.int/navipedia//index.php/RAIM (April 12, 2020).

Ramdas, V et al. 2010. *Transport Research Laboratory CLIENT PROJECT REPORT CPR798 ERTMS Level 3 Risks and Benefits to UK Railways Final Report*.

Raymond Chow. 2011. *Evaluating Inertial Measurement Units*. www.tmworld.com (April 12, 2020).

RTCA. 2006. *MINIMUM OPERATIONAL PERFORMANCE STANDARDS FOR GLOBAL POSITIONING SYSTEM/WIDE AREA AUGMENTATION SYSTEM AIRBORNE EQUIPMENT*. www.rtca.org (April 10, 2020).

RTKLIB. "RTKLIB: An Open Source Program Package for GNSS Positioning." http://www.rtklib.com/ (April 17, 2020).

Saastamoinen, J. 2013. "Atmospheric Correction for the Troposphere and Stratosphere in Radio Ranging Satellites." In *GMS*, , 247–51.

ŞahinoĞlu, Zafer, Sinan Gezici, and Ismail Güvenç. 2008. 9780521873093 Ultra-Wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols *Ultra-Wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols*. Cambridge University Press.

SAVVY. "SAVVY® CargoTrac Telematic Device | Operating Life of up to >15 Years."

https://www.savvy-telematics.com/en/savvy-cargo-trac.html (April 17, 2020).

Shift2Rail. “Mission and Objectives - Shift2Rail.” https://shift2rail.org/about-shift2rail/mission-and-objectives/ (April 10, 2020).

Standards Committee of the IEEE Computer Society, Man. 2007. *IEEE Standard for Information Technology-Telecommunications and Information Exchange between Systems-Local and Metropolitan Area Networks-Specific Requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs) Amendment 1: Add Alternate PHYs IEEE Computer Society*.

Suzuki, Taro, and Nobuaki K. 2013. “Correcting GNSS Multipath Errors Using a 3D Surface Model and Particle Filter.” : 1583–95.

Suzuki, Taro, and Nobuaki Kubo. 2012. “GNSS Positioning with Multipath Simulation Using 3D Surface Model in Urban Canyon.” : 438–47.

Technical, IEP. 2012. *Intercity Express Programme. Train Technical Specification*.

Teunissen, Peter J. G. 1998. “GPS Carrier Phase Ambiguity Fixing Concepts.” In *GPS for Geodesy*, Springer Berlin Heidelberg, 319–88.

Teunissen, Peter J. G., and Alfred Kleusberg. 1998. “GPS Observation Equations and Positioning Concepts.” In *GPS for Geodesy*, Springer Berlin Heidelberg, 187–229.

U-blox. “U-Center | u-Blox.” https://www.u-blox.com/en/product/u-center (April 17, 2020).

Ubisense group. *Ubisense Dimension 4 Datasheet*.

Union, European. *COMMISSION IMPLEMENTING DECISION (EU) 2019/ 785 - of 14 May 2019 - on the Harmonisation of Radio Spectrum for Equipment Using Ultra-Wideband Technology in the Union and Repealing Decision 2007/ 131/ EC - (Notified under Document C(2019) 3461)*.

UNISIG. 2008. *ETCS Application Levels 1 & 2 - Safety Analysis Part 3 - THR Apportionment-SUBSET 088 Part 3*.

WANG, JINLING. 1999. "Stochastic Modeling for Real-Time Kinematic GPS/GLONASS Positioning." *Navigation* 46(4): 297–305. http://doi.wiley.com/10.1002/j.2161-4296.1999.tb02416.x (April 15, 2020).

Wang, Lei. *Investigation of Shadow Matching for GNSS Positioning in Urban Canyons*.

Wang, Lei, Paul D. Groves, and Marek Ziebart. 2011. "GNSS Shadow Matching Using a 3D Model of London in Urban Canyons."

Wang, Lei, Paul D. Groves, and Marek K. Ziebart. 2015. "Smartphone Shadow Matching for Better Cross-Street GNSS Positioning in Urban Environments." *Journal of Navigation* 68(3): 411–33.

Wang, Lei, Paul D Groves, and K Ziebart. 2012a. *GNSS Shadow Matching: Improving Urban Positioning Accuracy Using a 3D City Model with Optimized Visibility Prediction Scoring*.

Wang, Lei, Paul D Groves, and Marek K Ziebart. 2012b. "Multi-Constellation GNSS Performance Evaluation for Urban Canyons Using Large Virtual Reality City Models."

Wells, David et al. 1999. *GUIDE TO GPS POSITIONING*.

Woodman, Oliver J. 2007. "Number 696 An Introduction to Inertial Navigation An Introduction to Inertial Navigation." http://www.cl.cam.ac.uk/http://www.cl.cam.ac.uk/techreports/ (May 2, 2020).

World, GPS. 2017. "Broadcom Launches Dual-Frequency GNSS Receiver : GPS World." https://www.gpsworld.com/broadcom-launches-dual-frequency-gnss-receiver-for-mass-market/ (May 20, 2020).

"X2RAIL-2." https://projects.shift2rail.org/s2r_ip2_n.aspx?p=X2RAIL-2 (April 7, 2020).

Yozevitch, Roi, and Boaz ben Moshe. 2015. “A Robust Shadow Matching Algorithm for *GNSS* Positioning.” *Navigation* 62(2): 95–109. http://doi.wiley.com/10.1002/navi.85 (April 12, 2020).

Yu, Yinan et al. 2012. “A Compact UWB Indoor and Through-Wall Radar with Precise Ranging and Tracking.” *International Journal of Antennas and Propagation* 2012: 11.

Zamora Cadenas, Leticia. 2014. “Radiofrequency-Based Indoor Location Systems for Ambient Assisted Living Applications.” : 1.

Zebra. “Tecnología Ultra Wideband (UWB) | Zebra.” https://www.zebra.com/es/es/products/location-technologies/ultra-wideband.html (April 12, 2020).

Zhang, Cemin et al. 2010. “Real-Time Noncoherent UWB Positioning Radar with Millimeter Range Accuracy: Theory and Experiment.” *IEEE Transactions on Microwave Theory and Techniques* 58(1): 9–20.

Zwirello, Lukasz, Tom Schipper, Marlene Harter, and Thomas Zwick. 2012. “UWB Localization System for Indoor Applications: Concept, Realization and Analysis.” *Journal of Electrical and Computer Engineering* 2012: 11.

Appendix A

Publications

Papers in journals and conferences resulting from the author's research time are included in chronological order in this appendix.

A.1 INTERNATIONAL JOURNAL PAPERS

- Jon Goya, Gorka De Miguel, Saioa Arrizabalaga, Leticia Zamora-Cadenas, Iñigo Adín and Jaizki Mendizabal. Methodology and Key Performance Indicators (KPIs) for Railway On-Board Positioning Systems. *IEEE Transactions on Intelligent Transportation Systems* (Q1). March 2018.
- Gorka De Miguel, Jon Goya, Nerea Fernández, Saioa Arrizabalaga, Jaizki Mendizabal and Iñigo Adín. Map Aided Software Enhancement for Autonomous GNSS Complementary Positioning System for Railway. *IEEE Transactions on Vehicular Technology* (Q1). October 2019.
- Paul Zabalegui, Gorka De Miguel, Alejandro Pérez, Jaizki Mendizabal, Jon Goya and Iñigo Adín. A review of the evolution of the integrity methods applied in GNSS. *IEEE Access* (Q1). March 2020
- Nerea Fernández-Berrueta, Jon Goya, Javier Añorga, Saioa Arrizabalaga, Gorka De Miguel and Jaizki Mendizabal. An Overview of Current IP Network Emulators for the Validation of Railways Wireless Communications. *IEEE Access* (Q1). June 2020

Methodology and Key Performance Indicators (KPIs) for Railway On-Board Positioning Systems.

Jon Goya, Gorka De Miguel, Saioa Arrizabalaga, Leticia Zamora-Cadenas, Iñigo Adín and Jaizki Mendizabal.

IEEE Transactions on Intelligent Transportation Systems (Q1). March 2018.

IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 19, NO. 12, DECEMBER 2018 4035

Methodology and Key Performance Indicators (KPIs) for Railway On-Board Positioning Systems

Jon Goya, Gorka De Miguel, Saioa Arrizabalaga, Leticia Zamora-Cadenas, Iñigo Adin, and Jaizki Mendizabal

***Abstract*— The European Union (EU) is bolstering the railway sector with the aim of making it a direct competitor of the aviation sector. For that to occur, railway efficiency has to be improved by means of increasing capacity and reducing operational expenditure. Tracks are currently used below their maximum capacity. Given this fact and the EU's goals for the railway sector, research on solutions for on-board positioning system based on global navigation satellite systems (GNSS) have arisen in recent years. By taking advantage of GNSS, safety critical positioning systems will be able to use the infrastructure more efficiently. However, GNSS based positioning systems still cannot fulfill current normative validation processes, mainly, due to the fact that GNSS based positioning performance evaluation is not compatible with the key performance indicators (KPIs) used to assess railway systems performance: reliability, availability, maintainability, and safety. This paper proposes a methodology and unified key performance indicators (KPIs). Additionally, it shows real examples to address this issue. It aims to fill the gap between the current railway standardization process and any on-board positioning system.**

***Index Terms*— Key performance indicators, RAMS, GNSS, EN50126, on-board position.**

I. Introduction

THE European Union (EU) in order to provide a more attractive rail transportation method for end-users, has established several goals: the improvement of the efficiency of transport, the reduction of infrastructure costs, the decrease of operation time, the increase of the safety requirements and the standardization of quality controls.

One of these goals is to use the existing railway infrastructure more efficiently with the aim of increasing current capacity. This improvement will reduce operation delays and other disturbances that affect the transport quality. Focusing on the augmentation of track capacity by means of a more efficient use, the European Rail Research Advisory Council (ERRAC) has defined certain milestones [1].

One of the solutions relies on improving the on-board positioning systems based on Global Navigation Satellite Systems (GNSS) to meet the requirements defined by the European Train Control System (ETCS). GNSS based on-board position is believed to be one of the technologies that will make the migration from ETCS L2 to ETCS L3 possible. Some benefits of ETCS L3 migration are: a 25% decrease in the infrastructure costs and an 50% increase railway efficiency [2], [3]. More than 30 GNSS-related research projects have been funded by the EU, such as GRAIL-2, EATS, ERSAT-EAV and NGTC [4]–[7].

However, the implementation of GNSS based on-board train positioning has not yet been successful. The main issue is the validation of the requirements specifications when applying the standardization norm EN50126 [8].

This paper proposes a methodology and key performance indicators (KPIs) that are compliant with the EN50126 [8] standardization process for railway on-board positioning systems for safety critical applications.

The rest of this article is organized as follows. Section II reviews the literature on the methodology and KPIs definition. Section III proposes a methodology for the standardization process and defines KPIs for on-board positioning systems. Section IV provides examples of the methodology under different railway scenarios. Finally, in Section V and Section VI the proposal discussion and conclusions are presented.

II. Railway Standardization Process Related Work: Methodology and KPI

In this section the methodology and KPIs definition in the literature are presented. For each of these, three different sources are analyzed: the railway standardization norm, scientific research in the literature and road sector. Maritime and aviation sectors have been discarded from the study because they do not share as many similarities with railway like road intelligent transportation systems (ITS) applications do.

A. CENELEC EN50126 [8]

EN50126 norm defines a systematic verification and validation (V&V) process on which Reliability, Availability, Maintainability and Safety (RAMS) measures are controlled during the life-cycle of a system for railway.

The main issue in the validation process is to determine the performance results defined as: Reliability, Availability, Maintainability and Safety (RAMS) requirements for the GNSS-based positioning systems. Different performance specification parameters are defined in EN50126 [8] and [9] for

Manuscript received November 11, 2016; revised April 20, 2017, July 26, 2017, and November 21, 2017; accepted February 11, 2018. Date of publication March 7, 2018; date of current version November 27, 2018. This work was supported by the European Community's Framework Program H2020/FR8RAIL project under Grant 730617. The Associate Editor for this paper was D. Chen. *(Corresponding author: Jon Goya.)*

The authors are with the Ceit, 20018 San Sebastián, Spain, and also with the Universidad de Navarra, 20018 San Sebastián, Spain (e-mail: jgoya@ceit.es; gdemiguel@ceit.es; sarrizabalaga@ceit.es; lzamora@ceit.es; iadin@ceit.es; jmendizabal@ceit.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2018.2806187

GNSS QoS performance evaluation. At this point the process is stuck.

A new systematic performance assessment that meets all terms defined by EN50126 [8] is mandatory.

B. Scientific Research Works

Several scientific researches have dealt with the task of applying the current standardization process to GNSS based on-board positioning systems.

There is already a road-map for the development of a GNSS solution that theoretically satisfies the EN50126 [8] norm and follows the verification and validation process [10].

Even if the conceptual application of the standardization process is clear, the "System Requirement" and the "System validation" stages have not been validated yet, due to the nature of the GNSS based positioning systems. These systems depend on the operating environment. For example, additional undesired effects that derived from the railway infrastructures should be considered, such as the situations where signal are blocked (tunnel) or where the GNSS systems are degraded. Thus, characterizing the system requirements is a complex task due to the fact that RAMS terms differ from the ones defined for GNSS, which are inherited from aviation sector [9]. Thus, a link between both RAMS and GNSS is needed.

Several research studies have focused their efforts on trying to map RAMS onto GNSS [11], [12]. For this purpose, the positioning system is defined as a model state machine; so the terms can be evaluated numerically providing a relation between RAMS and GNSS performance measures [10], [12], [13]. However, some of these studies also pointed out that the mapped terms do not always fit together exactly and remarked differences between RAMS and GNSS terms [11], [14]. For example availability, where in GNSS the focus is on providing a position estimate and in RAMS additional integrity requirement considerations are applied [11].

C. Standardization of Performance of GNSS for ITS [15]

Road transport sector is also dealing with the task of standardizing GNSS based positioning systems [15]. Currently, only the ISO 26262 [16] safety norm is applied.

The validation is generated based on operational scenarios, error models and metrics. By defining operational scenarios, the simulation tests are reduced to the cases where the positioning system is stressed by means of having a harsh environment or a minimum performance requirement.

This methodology only focuses on the simulation and performance characterization of the positioning system. However, the proposed metrics should be valid for the simulation and field-test performance evaluation, allowing straightforward comparison between both. Currently, the "classical" approach to assess the performance is adopted where "Accuracy," "Availability" and "Integrity" terms are selected.

Accuracy is measured using the cumulative distribution function (cdf) of the position error. Availability is defined as the percentage of the time the system is capable of providing a position estimate, velocity and time (PVT) and finally, for the integrity analysis the so-called "Standford Plot" is used, where the position error and the confidence interval of each epoch are compared.

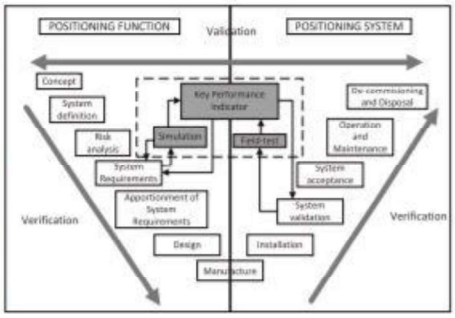


Fig. 1. "V" Representation proposal.

Additionally, the definition of "Performance Classes" is proposed [15]. This would define a certain criterion (such as having the cdf of the position error) to classify or determine the positioning system suitability for an application.

III. Methodology and Key Performance Indicators

The methodology that this article proposes for evaluating the performance of on-board positioning systems for safety critical application is the inclusion of an intermediate stage inside the V model defined by the EN50126 [8]. This proposal aims to provide a methodology that allows the validation of the "System Requirements" and "System Validation". In this process it is important to create a common framework where the requirements are measured in the same terms (see Fig. 1).

The proposed methodology is divided in two sections, the "Positioning Function" and the "Positioning System". This division is carried out due to the fact that in the simulation stage the system is modeled and the "Positioning System" is the real system used on field-test.

Benefits of the intermediate stage:

- Generation of common framework that enables the validation of the requirements.
- Independence of the Verification and Validation stage.
- System requirements definition using a simulator.
- Capability of testing different operational scenarios.
- Capability of simulating a large number of iterations to obtain relevant requirements.
- Generation of expected figures and results.
- Early-stage development error detection.
- Capability of using field-test data during the simulation.
- Reduction of field-test trials during the validation.
- Reduction of the time and cost of the validation process.

Additionally, the characterization of on-board positioning system, by means of the KPIs, must be carried out independently of technologies used. Thus, a black-box approach is applied as a generic evaluation procedure. Only the outcomes of the black-box are evaluated, without any knowledge of

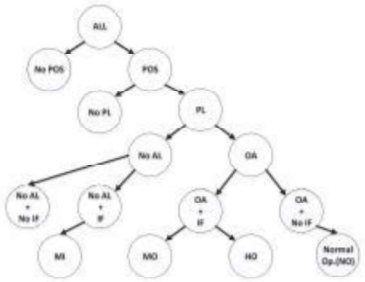


Fig. 2. Tree diagram for the states definition.

its internal workings. This means that external information sources' effects, such as signal blocking, multi-path, shadowing, etc. are not directly considered.

Train positioning system calculates the traveled distance and a confidence interval that bounds the train's position error [17]. In addition to this, it is known that the error estimate shall be equal to, or lower, than ± (5m + 5 % traveled distance) [18]. This operational requirement bounds the maximum error that can be handled by establishing an alert limit (AL).

In the literature, positioning outcomes are classified as Up state, Degraded state or Faulty state for subsequent use by the KPIs to characterize the positioning system [10], [12], [13].

However, up to now, the parameters that can be used during the performance analysis are:

- Position estimate (POS)
- Position error (PE)
- Protection Level (PL)
- Operational requirement (Alert Limit (AL))

Fig. 2 depicts the possible number of outcomes based on the detected parameters, which can be greater than the three states proposed in the literature [10], [12], [13].

The operational requirement (AL) is not adopted yet for ETCS level 3. However, GNSS advisory group for railway has proposed operational requirement values for several scenarios [19]. Even though AL is not yet defined, it can be useful to determine a possible performance limit of a particular technology during the KPIs' analysis.

The states shown in Fig. 2 are defined as follows:

- (No POS) The positioning system is not able to provide a position estimate.
- (No PL) The positioning system is not able to provide a protection level.
- (No AL) The protection level is greater than the AL, thus the generated position estimate does not fulfill the operational requirement.
- (Operational Available (OA)) The protection level is lower than the alert limit.
- (Integrity Fault (IF)) The position error is greater than the protection level.
- (Misleading Information (MI)) There is an integrity fault and the protection level is greater than the alert limit. It is only considered to be wrong information.
- (Misleading Operation (MO)) There is an integrity fault and the position error is lower than the alert limit, which points outs the maximum tolerable position error for the current operation.
- (Hazardous Operation (HO)) There is an integrity fault and the position error is greater than the alert limit. Thus, this should be considered as high risk epoch.
- (Normal Operation (NO)) The protection level is lower than the alert limit and greater than the position error.

The on-board positioning system's number of states is increased up to 7 different states. States are considered to be the branch end-points in Fig. 2. Having a detailed classification is useful for precisely defining all the possibilities, especially the ones related to the integrity of the position estimate.

The next step is to define the KPIs based on the generated outcome classification. The selected KPIs are: Accuracy, Protection Level, Availability, Reliability, Integrity and Continuity.

A. Accuracy

The term Accuracy is widely used in the literature. However, in real-time operation, accuracy is a term that cannot be measured as there is no true reference to be compared with. Accuracy can be only calculated when the true reference position is available. This is only achievable when results are post-processed against the true reference. Thus, the accuracy is defined as follows:

Accuracy (PE): Degree of compliance between the true reference and the estimate for any given instant of time.

In the railway sector, the horizontal position error is the one that best describes the requirements for accuracy. As the train movement is linked to a physical track placed on the ground, the Vertical Error can be neglected. Horizontal Position Error (HPE) is defined as the difference between the train horizontal position estimate and the train real horizontal position at any given instant of time. Consequently, Horizontal Position Error (HPE) is defined as follows:

$$HPE = \|\overline{PR}_{x,y}\| = \sqrt{(x_{t,p} - x_{t,r})^2 + (y_{t,p} - y_{t,r})^2} \quad (1)$$

where P and R are points with x, y coordinates to define the position estimate P and the reference or true position R. The subindex t represents the time instant where the accuracy is evaluated. Subindex p is used to define the position estimate and subindex r to define the reference or true position.

The accuracy term is used to represent the error between the position estimate and true reference point of the trajectory. Cumulative distribution function is usually used in order to obtain a position estimate error bound for a given percentage of the time.

B. Protection Level

The aim of the protection level is to provide a bounding value that ensures that the real position is limited by it. This term is more interesting for the real time on-board positioning systems due to the fact that this term is provided during the operation for each epoch, in contrast to the accuracy, which cannot be assessed in real time. The protection level is defined as the following:

Protection Level (PL): The value that limits the uncertainty of the real position. The protection level calculation is based on the safety requirements that the navigation system is targeting.The aim is to ensure that the Safety Integrity Level (SIL) requirement is fulfilled.

The protection level function (or integrity function) is outside the scope of this research work, and it is a research area in its own. The simplest approach is defined as [20]:

$$PL = K \cdot d \tag{2}$$

where d is an estimate of the uncertainty in the error (related with the second order statistic of the error) and K is a constant factor that is linked to a statistical distribution to ensure the PL is bounding the error [20].

In rail sector, PL is defined as confidence interval. This parameter limits the position uncertainties that appear during the evaluation of the traveled distance carried out because of the balise and odometer [17].

C. Availability

It can be assumed that the on-board positioning systems that are being studied are intended to be used in safety-critical applications. This fact provides a clear requirement that having PL provision is a must. Therefore, availability is defined as the following:

Availability: The ability of the positioning system to provide a position estimate with the corresponding valid confidence metric (PL) calculated over a time interval (T).

$$A(T) = 100 \cdot \frac{1}{T} \cdot \sum_{t=0}^{T} a(t) \tag{3}$$

where, $a(t)$ represents the instant availability of the positioning system (for a given time instant (t) within the time interval T). The confidence metric PL is considered valid when its value is greater than zero. A position estimate is defined as unavailable when it has not been possible to determine any protection level (represented by $PL = -1$) or the expected position estimate has not been provided. This is because positioning systems used for railway applications are assumed to provide fixed rate position estimates to train position in real-time. Thus, the instant availability $a(t)$ measures if the positioning system provides a position estimate and a PL in each of the expected time epochs. These are the minimum requirements for using epoch information for safety-critical applications.

At this point, the defined *"Availability"* determines if an epoch fulfills the minimum requirements to use it in a safety-critical context. Instant availability is defined as the following:

$$a(t) = \begin{cases} 1 \text{ if } \exists \text{Pos. Estimation } \& \ \exists PL \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

In order to see whether it is usable during normal operations in railway environment, the integrity and operational availability terms have been defined.

D. Operational Availability

Related to the availability, the operational availability includes an extra requirement to an available position estimate for the particular on-going application or operation mode. The performance limitation of a solution is defined by the operation or the application. Aviation sector has defined different sets of requirements for several operations (en-route, approach, arrival and departure, etc.) [9], [21]. This means that the position estimate and position uncertainty are assessed depending on the current aircraft operation mode. The same way, road sector is proposing different performance classes [15].

In this case, the information about the operational availability terms is not fixed as the value in each operation type is different. This could be understood as a parametrized analysis of the availability using operational extra requirements. This is useful for detecting the possible improvement area of the positioning system under study using the same operational requirement, compared with the availability term.

This means that the operational availability of a positioning system is the percentage of the time that the systems fulfill certain AL requirement. The AL requirements definition is linked to the final application of the positioning system. For an epoch to be what is called op. available, it also needs to fulfill the integrity requirements. Op. availability is defined as:

Op. Availability: is the ability of the positioning system to provide a train position solution with a corresponding confidence metric (PL) under the following conditions ($0 < \text{PL} < \text{AL}$) and ($\text{PL} > \text{PE}$) over a given time interval (T).

$$OA(T) = 100 \cdot \frac{1}{T} \cdot \sum_{t=0}^{T} oa(t) \tag{5}$$

where,

$$oa(t) = \begin{cases} 1 \text{ if } 0 < PL < AL \ \& \ PL > PE \\ 0 \text{ otherwise} \end{cases} \tag{6}$$

The AL sets the maximum error that the application could handle without having any safety risk. During the real operation, the PL provides the error upper bound. This parameter analysis determines the current performance of the position estimate and integrity related concerns.

E. Integrity

Safety is the most important requirement in railway. The integrity term is the most studied parameter due to the fact that it has a direct effect in the safety of the system; the clearest example is that EN50126, where the integrity is related to the *"safety integrity"* [8]. However, the detected integrity faults need to be further studied to detect the real impact that can have in the overall system safety analysis.

Integrity fault studies the trust that can be placed in the systems, by means of having a PL greater than the PE. Therefore, Integrity is defined as the following:

Integrity: the minimum ensured probability where the confidence metric (e.g. PL.) is greater that the position error.

$$I(T) = 100 \cdot (1 - \frac{1}{T} \cdot \sum_{t=0}^{T} i(t)) \tag{7}$$

where $i(t)$, instantaneous integrity, in case that the AL is defined, can be defined as the set of points or sum of the subsets that do not fulfill integrity, where depending on the type could be considered dangerous. Non-integrity epochs are defined with PL < PE. If the AL is defined, the epochs that would violate the integrity are classified depending on the risk that could cause.

$$i(t) = \begin{cases} 1 \text{ if } PL < PE \\ 0 \text{ otherwise} \end{cases} \tag{8}$$

$$i(t) = i_{mo}(t) + i_{mi}(t) + i_{ho}(t) \tag{9}$$

$$i_{mo}(t) = \begin{cases} 1 \text{ if } PL < PE \ \& \ PE < AL \\ 0 \text{ otherwise} \end{cases} \tag{10}$$

$$i_{mi}(t) = \begin{cases} 1 \text{ if } PL < PE \ \& \ PL > AL \\ 0 \text{ otherwise} \end{cases} \tag{11}$$

$$i_{ho}(t) = \begin{cases} 1 \text{ if } PL < AL < PE \\ 0 \text{ otherwise} \end{cases} \tag{12}$$

where $i_{mo}(t)$ is defined as misleading operations, $i_{mi}(t)$ as misleading information and $i_{ho}(t)$ as hazardous operations. The integrity check is required when location solution is being developed for safety-critical system.

In real time, the integrity of each epoch cannot be assessed. The safety requirement is defined during the design and implementation of the positioning function, however it can only be obtained during the post-process analysis. Integrity is not achieved every time the protection level is lower than the position error. This evaluation notes the number of times the positioning system makes an error that could have undesired effects and could compromise the safety.

F. Continuity

The requirement of evaluating the possibility of having a continuous operationally available estimate in the next time interval of Δt seconds has been taken from the aviation sector. In railway sector, there is still not any kind of operation mode definition. However, functions that rely on the positioning could need this indicator, such as the virtual coupling function that tries to join two trains on-the-go.

Continuity: The continuity is the probability that the specified system performance will be maintained for the duration of a phase of operation, presuming that the system was available at the beginning of that phase [9].

The term Failure is reserved for the RAMS analysis where the aim is to assess the Reliability of the systems. The first change that will be made for the consistency of the EN50126 [8] inside the KPIs that are being proposed, is to differentiate failure events and fault events.

Just as a reminder, fault events are considered epochs that match with one of the following conditions:

- Unavailable epoch.
- Operational unavailable epoch.

The Continuity term assesses the operational available epochs' disruption in a defined Continuity Time Interval (CTI). Currently, there are two approaches to assess the continuity:

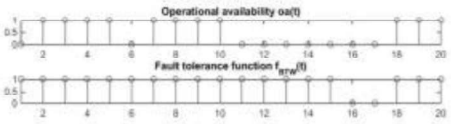


Fig. 3. Fault event example with BTW of 5 epochs.

- *Mean Continuity*: The mean continuity is the probability of having a fault during a CTI using the Mean Time Between faults (MTBf).

$$MC(T) = 1 - \frac{CTI}{MTBf_{BTW}(T)} \tag{13}$$

where, the MTBf ($MTBf_{BTW}$) may be calculated over a given time interval (T).

$$MTBf_{BTW}(T) = \frac{\text{Total time}}{\text{Number of faults}} \tag{14}$$

Total time may be defined also as the sum of the duration of multiple simulation iterations.

A fault event is defined as any period lasting for more than Burst Time Window (BTW) seconds where the obtained positions are unavailable or op. unavailable (also called fault points). This means that the fault event may have different duration, always higher than BTW seconds. Number of faults is equal to the addition of all fault events that have occurred across all operation duration in case there is more than one iteration.

- *Continuity*: is defined as the number of epochs without any faults within the defined CTI during the entire operation duration (T):

$$C_{CTI}(T) = 100 \cdot \frac{1}{T - CTI} \cdot \sum_{t=CTI}^{T} c(t) \tag{15}$$

where, the continuity at instant t is defined as:

$$c(t) = \prod_{i=t-CTI}^{t} f_{BTW}(t) \tag{16}$$

To analyze the fault events and their duration, $f_{BTW}(t)$ fault tolerance function is defined for a given BTW:

$$f_{BTW}(t) = \begin{cases} 1 \text{ if } t > BTW \ \& \ \sum_{i=t-BTW}^{t} oa(t) \geq 1 \\ 1 \text{ if } t \leq BTW \\ 0 \text{ otherwise} \end{cases} \tag{17}$$

If $f_{BTW}(t) = 0$, there is a fault point; if $f_{BTW}(t) = 1$, there is not any fault point. The fault event is defined as one or more contiguous fault points. For example, in case that there are 5 contiguous fault points, there is a single fault event with a duration of 5 epochs (see Fig. 3).

The defined continuity is more precise that the mean continuity, however both can be used.

In the next section the implemented methodology and visualization are presented by means of practical examples.

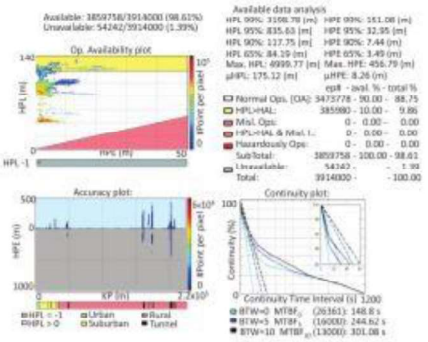


Fig. 4. EATS [5] results graphical representation.

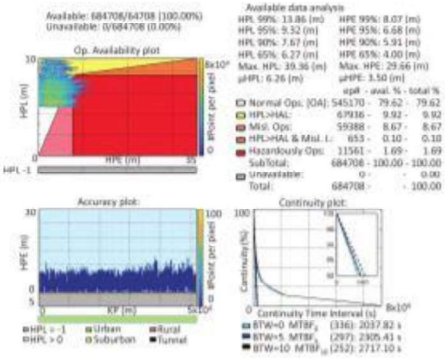


Fig. 5. ERSAT-EAV [6] results graphical representation.

IV. Proof of Concept

The proposed KPI methodology has been included in the Railway Advanced Navigation System Simulator (RANSS) [22]. Using this simulation platform practical examples have been carried out to demonstrate the applicability of the methodology in real cases and different railway scenarios.

- EATS [5]: EATS project advanced ETCS on-board equipment laboratory testing and the use of satellite positioning technologies together with other technologies in ERTMS.
- ERSAT-EAV [6]: ERSAT-EAV verified the suitability of EGNSS in making possible cost-efficient and economically sustainable signaling solutions for railway safety-critical applications, fostering the competitiveness and the innovation of the European rail industries.

Fig. 4 depicts EATS [5] project results using the systematic methodology under a full simulated scenario.

Fig. 5 depicts ERSAT-EAV [6] project results using the systematic methodology under a field-test scenario.

The following list terms are depicted in the graphical representation and numerically in both Fig. 4 and Fig. 5 in order to have a better characterization of the system performance:

- Accuracy: in the bottom-left where the distance vs accuracy is depicted, it also provides the error distribution.
- Protection Level: the representation of this term is covered in the top-left area, where accuracy vs. protection level is depicted.
- Availability and Op. Availability: the representation of this term is covered in the top-left area, where accuracy vs protection level is depicted, additionally the alert limits have to be defined to obtain the op. availability.
- Integrity: the representation of this term is covered in the top-left area, where the capabilities of the protection level calculation are compared to the obtained accuracy.
- Continuity: the representation of this term is covered in the bottom-right area, where the capabilities of the system of having continuous op. available position estimate is evaluated along the operation.

According to the results obtained during the simulations the use of signals opportunity increased the availability, op. availability and continuity of the systems but they were not accurate enough for strict accuracy requirement on tunnel environments. Additionally, the importance of improving the protection level calculation was detected, as the performance really should be evaluated using this parameter as it plays a key role for the safety decision-making.

V. Proposal Discussion

It is important to remark that the proposed intermediate stage does not interfere with the standardization process as it allows producing more realistic "System Requirements" based on simulations and the possibility to compare them directly in the "System Validation".

The relevant decision-making parameter used during the operation is changed from the position error which is studied in the literature, to the confidence interval as it is the unique quality measurement available during the train's operation. This means, that from the point of view of integrity, the calculation of a confidence interval will really determine the capabilities of the system. Having an accurate system with an erroneous or high protection level calculation is worthless for safety-critical applications.

The KPIs have been proposed for an absolute positioning system, but at present railway uses the traveled distance in order to provide a position. The main issue for the applicability of the methodology is how absolute positions are translated to traveled distance. Fluctuation of the absolute position can increase the traveled distance. The rest of the methodology can be used by defining the position error as the difference between the reference traveled distance and the estimated traveled distance.

Despite the discussion of the simulation platform is out of the scope of the methodology and KPIs, it is important to remark the most important capability of RANSS related with this topic. RANSS platform is capable of fully working under a fully simulated scenario or with field-test data,

providing the KPIs results and allowing the comparison of both. Moreover, the simulation platform compliance with the reality is compulsory in order to obtain significant results, mainly when simulation is the main testing procedure. Modeling GNSS effects related with the different operational environment, such as multi-path, atmospheric delays, etc., has crucial importance in order to succeed with the proposed methodology.

The end-user is responsible for interpreting the results of the simulation platform and their implications, as in the field-test additional errors could appear, such as system failures, not expected undesired signal effects, etc. [22]. The addition of these failures is not an easy task to tackle as every component of the system would need to be characterized in detail. This means that the Reliability and the Maintainability are not considered during the simulation.

This means that the system requirements should be considered as the upper-bound of the achievable performance of the positioning system. If the positioning function does not fulfill certain requirement the positioning system would not fulfill it either. This is due to the fact that models are simplifications of the complex real situations.

Another consideration is the Safety and Integrity. The proposed KPI are focused only on measuring the integrity failures of the positioning system protection level estimation. It can be assumed that the protection level estimation can be studied as an "element" defined in the RAMS process and a SIL requirement can be defined. RAIM techniques that are inherited from the use of the GNSS based systems will enhance the integrity performance. However, it should be mentioned that the Safety analysis is a more complex task, where the frequency of a risk and its implication must be measured. Additionally, the RAMS proposed safety analysis procedure should be based on a generally accepted principle, such as "As Low As Reasonably Practicable" (ALARP), "globally at least as good" (Globalement Au Moins Aussi Bon or GAMAB) or Minimum Endogenous Mortality" (MEM). This means that the integrity assessment of the KPIs is not the RAMS Safety evaluation process but it is reasonable enough to assess the protection level estimation element using a SIL requirement. Thanks to the capabilities of RANSS the number of epochs can be high enough to be able to check if the safety requirements are fulfilled.

It is important to analyze the source and purpose of the KPIs and compare the proposed ones with current RAMS performance measures.

Table I indicates the purpose of the KPIs and the sources that were used as guidelines when defining the KPI. It does not mean that the KPIs were directly defined in these sources.

Fig. 6 depicts the relation between RAMS (on the left) and the proposed KPIs (on the right).

As it is presented, the definition of the proposed KPIs are related with the RAMS performance measures and the intermediate approach is complementary to the current standardization process. Additionally, the proposed methodology could also be valid for ITS systems, which could have benefits and create a unique standardization process for safety-critical positioning systems for both road and rail sector.

TABLE I
KPI SUMMARY

KPI	Source	Purpose
Accuracy	ISO 5725-1 [23]	Evaluate the compliance of the provided position compared to a ground-truth (post-processing).
Protection Level	FRP2014 [9]	Provide a confidence interval to the position estimate.
Availability	EN50126 [8] FRP2014 [9]	Evaluate if the system is available to work even if the whole set of information is not provided.
Op. Availability	EN50126 [8] FRP2014 [9]	Once the operational requirements are defined, in which percentage of the time the system is available.
Integrity	FRP2014 [9]	Comparison between the accuracy and the protection level to detect the proper performance of the integrity function that is in charge of managing the train positioning safety-critical function (post-processing).
Continuity	FRP2014 [9]	Determine the capability of the positioning system to provide a position estimate during the following X seconds.

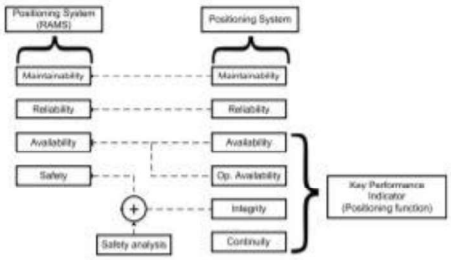


Fig. 6. RAMS and KPI relation.

VI. CONCLUSION

This paper presents a methodology and key performance indicators proposal for on-board positioning that can be integrated in the currently used EN50126 standardization process [8]. The methodology adds an intermediate stage that will allow the system requirements to be defined independently from the positioning technology used. Moreover, these key performance indicators are able to measure field-test information in the same terms, allowing a direct requirements' validation.

Apart from the proposed methodology and KPI, performance measures results are presented for the data analysis providing the end-user the necessary information to draw practical conclusions to make decisions about the positioning system. Different examples have demonstrated the versatility of the proposed methodology.

The alignment with the presented road transport standardization draft is important from the point of view that both share the same objectives. Thus, the use of the same performance evaluation methodology is important as it also allows each

sector to apply its own safety requirements. This approach creates a more competitive market for positioning systems which is beneficial to all related industries.

ACKNOWLEDGMENT

The authors want to thank Michael Hutchinson and Yuheng Zheng from Nottingham Scientific Limited.

REFERENCES

[1] G. Travaini, D. Schut, and A. McNaughton, "Rail route 2050: The sustainable backbone of the single European transport area," ERRAC, Paris, France, Tech. Rep., 2013.

[2] V. Ramdas, T. Bradbury, S. Denniss, D. Chapman, R. Bloomfield, and D. Fisher, "ERTMS level 3 risks and benefits to UK railways," Transp. Res. Lab., Wokingham, U.K., Final Rep. CPR798, 2010.

[3] L. Aliadiere and R. Heinisch, "Global perspectives for ERTMS ETCS and GSM-R," in *Proc. ERTMS Annu. Conf.*, Berne, Switzerland, 2007.

[4] L. Marradi *et al.*, "GNSS for enhanced odometry: The GRAIL-2 results," in *Proc. 6th ESA Workshop Satellite Navigat. Technol. Eur. Workshop GNSS Signals Signal Process., (NAVITEC)*, Dec. 2012, pp. 1–7.

[5] S. Arrizabalaga *et al.*, "Development of an advanced testing system and smart train positioning system for ETCS applications," in *Proc. 6th Transp. Res. Arena*, Paris, France, 2014, pp. 1894–1903.

[6] I. Adin, J. Mendizabal, G. de Miguel, J. Goya, L. Zamora, and S. Arrizabalaga, "Complementary positioning system in GNSS-denied areas," in *Proc. 6th Transp. Res. Arena (TRA)*, Warsaw, Poland, 2016, pp. 4562–4571.

[7] P. Gurnik, "Next generation train control (NGTC): More effective railways through the convergence of main-line and urban train control systems," *Transp. Res. Procedia*, vol. 14, pp. 1855–1864, Jan. 2016.

[8] *Railway Applications. The Specification and Demonstration of Reliability, Availability, Maintainability and Safety (RAMS). Generic RAMS Process*, document BS EN 50126-1:2017, CENELEC, 2008.

[9] A. B. Carter, A. T. Foxx, and J. C. Johnson, "Federal navigation plan 2014," Dept. Defense, Dept. Homeland Secur., Dept. Transp., Springfield, MI, USA, Tech. Rep. DOT-VNTSC-OST-R-15-01, 2014.

[10] J. Beugin and J. Marais, "Simulation-based evaluation of dependability and safety properties of satellite technologies for railway localization," *Transp. Res. C, Emerg. Technol.*, vol. 22, pp. 42–57, Jun. 2012.

[11] A. Filip, J. Beugin, J. Marais, and H. Mocek, "A relation among GNSS quality measures and railway RAMS attributes," in *Proc. CERGAL*, 2008, pp. 2–3.

[12] D. Lu and E. Schnieder, "Performance evaluation of GNSS for train localization," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 1054–1059, Feb. 2015.

[13] D. Lu, F. Grasso Toro, and E. Schnieder, "RAMS evaluation of GNSS for railway localisation," in *Proc. IEEE Int. Conf. Intell. Rail Transp. (ICIRT)*, Aug./Sep. 2013, pp. 209–214.

[14] J. Beugin, A. Filip, J. Marais, and M. Berbineau, "Galileo for railway operations: Question about the positioning performances analogy with the RAMS requirements allocated to safety applications," *Eur. Transp. Res. Rev.*, vol. 2, no. 2, pp. 93–102, 2010.

[15] F. Peyret, "Standardization of performances of GNSS-based positioning terminals for ITS applications at CEN/CENELEC/TC5," in *Proc. Intell. Transp. Syst. World Congr.*, 2013, pp. 1–10.

[16] "Road vehicles—Functional safety," Standard ISO 26262-10:2012, 2012.

[17] UNISIG, "ERTMS/ETCS documentation, system requirements specification, SUBSET-026," UNISIG, Brussels, Belgium, Tech. Rep. Baseline 3-v3.4.0, 2014.

[18] UNISIG, "ERTMS/ETCS documentation, performance requirements for interoperability, SUBSET-041," UNISIG, Brussels, Belgium, Tech. Rep. Baseline 3-v3.1.0, 2012.

[19] J.-M. Wiss *et al.*, "Requirements of rail applications," in *Proc. GNSS Rail User Forum*, May 2000.

[20] F. Kneissl, G. Hein, and C. Stöber, "Combined Integrity of GPS and Galileo," in *Proc. Inside GNSS*, 2010, pp. 52–63. [Online]. Available: http://www.insidegnss.com/auto/IGM_wp-janfeb10.pdf

[21] International Civil Aviation Organization, *Required Navigation Performance Authorization Required (RNP AR) Procedure Design Manual*, 1st ed. Montreal, QC, Canada: ICAO, 2009.

[22] J. Goya, L. Zamora-Cadenas, S. Arrizabalaga, A. Bruzález, J. Meléndez, and J. Mendizabal, "Advanced train location simulator (ATLAS) for developing, testing and validating on-board railway location systems," *Eur. Transport Res. Rev.*, vol. 7, p. 24, Sep. 2015. [Online]. Available: http://link.springer.com/10.1007/s12544-015-0173-5

[23] *Accuracy (Trueness and Precision) of Measurement Methods and Results. Part 1: General Principles and Definitions*, document. 5725-1, 1994.

Jon Goya received the M.Sc. degree in telecommunications engineering in 2011 and the Ph.D. degree from the Universidad de Navarra in 2016. He is currently a Lecturer with the TECNUN, Universidad de Navarra, and also a Researcher with the CEIT-IK4. His professional research activity lies in the simulation of on-board positioning system and performance analysis for railway. He has participated in FP7 projects coordinated by CEIT and is now actively participating in H2020 European funded and Shift2Rail projects in Railway signaling and Positioning topics: ERSAT-EAV and FR8RAIL.

Gorka de Miguel received the M.Sc. degree in telecommunications engineering from the School of Engineering of San Sebastián (TECNUN), Spain, in 2015. He joined the CEIT-IK4 Research Centre, San Sebastián, in 2015, where he is currently a Researcher Assistant with the Electronics and Communications Department. His research activity lies in the field of positioning and software development.

Saioa Arrizabalaga received the M.Sc. degree in telecommunication engineering from the Faculty of Engineering in Bilbao, University of the Basque Country, in 2003, and the Ph.D. degree in engineering from the TECNUN in 2009. She is currently a Lecturer with the TECNUN, Universidad de Navarra, and also a Researcher with the CEIT-IK4. She is involved in the participation of European research projects, direction of doctoral theses, scientific, and technical publications in national and international journals and conferences. She has authored the book *Multi-Service QoS Architecture for a Multi-Dwelling Gateway*.

Leticia Zamora-Cadenas received the B.Sc. and M.Sc. degrees in telecommunications engineering from the University of Valladolid, in 2005 and 2008, respectively, and the Ph.D. degree in electrical and electronics engineering from the Universidad de Navarra in 2014. Her research interests include localization systems, positioning algorithms and Bayesian filtering, and data fusion and signal processing for telecommunication systems.

Iñigo Adin received the M.Sc. degree in electronics engineering in 2003 and the Ph.D. degree from the Universidad de Navarra in 2007. His research interests include safety-critical designs, with special interest in electromagnetic compatibility and transport interoperability. He has authored or co-authored one patent, two technical book, an invited chapter, and 19 articles in journals and international conferences. He was the Coordinator of the FP7 European Project TREND: Test of Rolling stock ElectromagNetic compatibility for cross-Domain interoperability contract number 285259.

Jaizki Mendizabal received the M.Sc. and Ph.D. degrees in electrical engineering from the TECNUN, Universidad de Navarra, San Sebastián, Spain, in 2000 and 2006, respectively. He joined the Fraunhofer Institut fr Integrierte Schaltungen, Erlangen, Germany, and SANYO Electric Ltd, Gifu, Japan, as a RF-IC Designer. Nowadays, he is currently with the CEIT-IK4, San Sebastián, Spain, where his research interests include communications and signaling systems for the railway industry. He is also lecturing Communications Electronics and Communications via Radio with the TECNUN, University of Navarra.

Map-Aided Software Enhancement for Autonomous GNSS Complementary Positioning System for Railway.

Gorka De Miguel, Jon Goya, Nerea Fernández, Saioa Arrizabalaga, Jaizki Mendizabal and Iñigo Adín.

IEEE Transactions on Vehicular Technology (Q1). October 2019.

IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 68, NO. 12, DECEMBER 2019 11611

Map-Aided Software Enhancement for Autonomous GNSS Complementary Positioning System for Railway

Gorka de Miguel, Jon Goya, Nerea Fernández, Saioa Arrizabalaga, Jaizki Mendizabal, and Iñigo Adin

***Abstract*—Independently on the business case addressed, one of the main drawbacks of the railway use cases that need continuous Global Navigation Satellite Systems data is the lack of availability for the 100% of the time of the journey. Additionally, the integrity assessment of the position estimation given is also mandatory for safety critical applications. Thus, tunnels and multipath effects are one of the most challenging situations for the continuous positioning systems. In this context, an autonomous on-board Complementary Positioning System has been proposed to overcome the limitation of Global Navigation Satellite System based positioning systems. This paper proposes a positioning enhancement solution by means of fusing data from the satellite navigation system and inertial measurement units. That hybrid solution provides higher availability and accuracy to the positioning specially on known blocked scenarios, such as tunnels, or urban canyons, by means of a novel environment aware map aided software technique named Known Blocked Scenarios algorithm ... This paper describes the Complementary Positioning System and the field test carried out in a challenging environment to validate the enhancement proposed by the authors, which demonstrate the benefits that this system has in known harsh environments for railways.**

***Index Terms*—Availability, CPS, GNSS, IMU, KBS, Multi-Sensor, On-Board Positioning, Positioning, Railway.**

I. Introduction

THE EUROPEAN Union (EU) aims at making railway a more attractive transportation method by improving its efficiency and reducing its costs. The European Rail Research Advisory Council (ERRAC) has defined some goals in order to increase the efficiency of the current railway infrastructure by means of track capacity augmentation [1].

In this context, SHIFT2RAIL, the main railway innovation initiative in EU plans to increase the attractiveness of the railway transportation by means of doubling the railway capacity, cutting the life-cycle costs of railway transports by as much as 50% and increasing reliability and punctuality by as much as 50% [2]. This could be applied to, at least, two of the Innovation Programs: IP2 and IP5 of this initiative dealing with control and communication systems and freight rail respectively. In the former, control, command and communication systems will go beyond merely being a contributor to the control and safe separation of trains, and become a flexible, real-time, intelligent traffic management and decision support system. And in the latter, the main challenge is to acquire a new service-oriented profile for rail freight services based on excellence in on-time delivery at competitive prices, interweaving its operations with other transport modes, addressing end-user needs by incorporating innovative value-added services, among others. In both cases, one of the enablers is the on-board Global Navigation Satellite System (GNSS) based positioning system and the improvement of their performance in order to reach every business' case requirements. For instance, these new services resulting from IP2 and IP5 roadmap will have an important impact in the migration from ETCS level 2 to ETCS level 3. This migration will allow a descent in the infrastructure costs up to 25% to regional and freight dedicated lines and efficiency improvements of more than 50% [3].

As a base, the suitability of GNSS systems for railway applications is being or has been analyzed by several European projects such as:

- GRAIL2 [4]–[7] aimed at developing and validating an ETCS odometry system prototype based on GNSS technology.
- SATLOC [8]–[10] aimed at developing and validating the use of GNSS in low traffic lines signaling and train control. The usefulness of the presented system on a real line was shown. However, the use of complementary positioning techniques was presented as essential. A future use of route maps and virtual balises was proposed.
- GaLoROI [9]–[14] project obtained the position estimation using GNSS (including Galileo technology), Eddy Current Sensors (ECS) and track-matching techniques. The tests performed in order to analyze the Reliability, Availability, Maintainability and Safety (RAMS) showed that GNSS performance was not good enough to meet railway requirements in harsh environment, showing the need of additional sensor fusion in order to increase the availability of the system.
- 3InSat [15]–[17] developed a Localization Determination System (LDS). The LDS is a fusion of GNSS systems with

Manuscript received February 1, 2019; revised May 30, 2019; accepted August 27, 2019. Date of publication October 14, 2019; date of current version December 17, 2019. This work was supported by SIA project as part of the European H2020 framework of projects, funded by European Commission (EC) and managed by European GNSS Agency (GSA), under contract number 776402. The review of this article was coordinated by Guest Editors of the Special Section on Smart Rail Mobility. *(Corresponding author: Gorka de Miguel.)*

The authors are with the CEIT-IK4 Research Centre, 20018 Donostia-San Sebastián, Spain, and also with the TECNUN, Universidad de Navarra, 20018 Donostia-San Sebastián, Spain (e-mail: gdemiguel@ceit.es; jgoya@ceit.es; nfernandez@ceit.es; sarrizabalaga@ceit.es; jmendizabal@ceit.es; iadin@ceit.es).

Digital Object Identifier 10.1109/TVT.2019.2940621

augmentation techniques and integrity detection solutions. The project aimed to reach the requirements of the European Rail Traffic Management System (ERTMS)/ETCS level 2 trackside balise. This project gave as a result a GNSS solution and a theoretical modeling concerning the Safety Integrity Level (SIL). The results wanted to integrate the satellite based localization systems in ERTMS-ETCS environment.

- ERSAT-EAV [18] aimed to reuse the ETCS odometry and the virtual balise concept in order to eliminate fixed balises. Augmentation networks such as EGNOS (European Geostationary Navigation Overlay Service) are also used to verify and validate different GNSS solutions in order to guarantee the positioning functions in areas in which GNSS signal is not accurate enough.
- RHINOS [19], [20] aimed to increase the use of EGNSS (European Global Navigation Satellite System) to support safety critical train localization functions for train control. RHINOS pillar was the GNSS infrastructure used for aviation with additional layer to meet the railway requirements.
- STARS [21], [22] developed an approach to characterize the GNSS performance in the railway environment. The project also quantified the economic benefits of using GNSS in ERTMS.
- X2RAIL-2 [23] aims to improve the performance at a railway system level by introducing new functionalities that should revolutionize the signaling and automation concepts in the future. The key technologies cover GNSS application in railway and advanced technologies for implementing new signaling functionalities.
- ERSAT-GCC [24], [25] aims to speed up the certification process of EGNSS according to ERTMS rules. The main goal is to define and certify a standard process, methodology and the related toolset for classifying track areas as suitable for locating Virtual Balises (VB), and to launch an operational line with integrated satellite technology by 2020.

One of the main problems of the on-board satellite positioning systems is the reduced coverage in urban environments or in difficult orography such as tunnels or canyons. In order to solve these limitations of the GNSS only positioning systems, the positioning enhancement proposed in this paper takes into consideration the previously mentioned papers and is based on a fused GNSS and Inertial Measurement Units (IMUs) solution which helps increasing the availability of the position data.

The use of on-board satellite positioning systems in the projects listed is mainly focused on the objective of having a positioning solution suitable for Virtual Balise (VB) [26].The positioning solutions used to provide the VB usually generate a continuous Position, Velocity and Time (PVT) estimate. However, the VB solution requires only areas where the VB can be triggered and therefore only a spot positioning system is required. For that reason, the continuous positioning is not mandatory but usually considered in order to reach a reliable solution as the PVT cannot only rely on GNSS. However, there are a number of other applications and services that require continuous positioning such as the odometry service. Taking this into account, the map aided software solution presented in this article pretends to broaden the scope and to tackle the issues of the GNSS based positioning system for a continuous positioning solution that could be used not only for the VB detection.

The progress beyond the state of the art of this work is the addition of a specific environment aware map aided software technique to the GNSS and IMU positioning. This software technique, called KBS algorithm, improves the accuracy in the entrance of Known Blocked Scenarios. All has been implemented in a Complementary Positioning System (CPS).

The paper is structured as follows:

- Section II introduces the Complementary Positioning System, the technologies used and its subsystems.
- Section III describes the Complementary Positioning System including both hardware and software.
- Section IV defines the validation tests and describes the environments in which they have been performed.
- Section V shows the results obtained after the tests.
- Section VI exposes the conclusions arisen from the work presented.

II. CPS and Technologies Used

This section introduces the Complementary Positioning System (CPS) solution and the technologies used to obtain the enhanced position estimation.

CPS is an autonomous positioning system as it can perform a full available positioning operation in a standalone mode without the need of other systems already integrated in the trains. The hardware subsystems used in the CPS are low cost sensors due to the fact that it was designed to have a good cost-effectiveness for an attractive business case. Even the GNSS receiver could be a low-medium end subsystem, which is not usual on the railway applications analyzed in the introduction's projects and papers. Moreover, a wider use is enabled, as trains without GNSS onboard positioning systems could use the CPS and the ones that already had GNSS installed would only need an expansion. In case a high end GNSS system mounted on board leads to a better CPS performance to obtain the best possible positioning solution, in lines or applications in which a high performance is required.

The CPS overcomes the GNSS obscured areas thanks to the introduction of the inertial sensors, software positioning enhancement techniques and map aided environment awareness. This is achieved by using the technologies listed below:

- Global Navigation Satellite Systems (GNSS)
- Inertial Measurement Units (IMU)
- Known Blocked Scenarios (KBS)

A. GNSS

GNSS are systems that use satellites to compute a position with global coverage thanks to a GNSS receiver. They can be used for providing timing, position, navigation or tracking information. It is the main technology used for navigation and positioning services and it is widely spread in the daily life. In the near future there will be four GNSS fully operative: GPS (USA),

GLONASS (Russia), GALILEO (EU) and BeiDou/COMPASS (China).

Standard GNSS receivers provide up to four observations per satellite and per frequency at a specified rate that depends on the application:

- Pseudorange: Measurement of the distance from receiver antenna to satellite antenna including receiver and satellite clock offsets, and other biases expressed in meters [27].
- Carrier Phase: Measurement of the range between a satellite and receiver expressed in units of cycles of the carrier frequency [27].
- Doppler: It is the Doppler shift with respect to the nominal signal frequency; it is positive for approaching satellites. Usually used for the calculation of the user velocity expressed in Hz [27].
- Signal Strength: Measurement of the strength of the received signal, dependent on the degree of thermal, background and intermodulation noise to which the signal has been subjected expressed in dBm [27].

B. IMU

An IMU is a device that has three gyroscopes and three accelerometers that are displaced along three mutually orthogonal axes. The main idea of inertial navigation is the Newton's First Law: "A body will continue in its state of rest, or of uniform motion in a straight line, unless an external force is applied to it". The accelerometers detect the acceleration changes due to the gravity forces and the gyroscopes the changes in the rotational attributes. An IMU is autonomous and is not dependent on other devices or signals visibility. Moreover, it has no need of an antenna so it can be mounted on-board in a stable place. However, the main drawback of the IMU is its accumulative error. The sources of the errors are mainly two:

- Bias errors: constant errors suffered by the sensors in their measurements. Knowing those errors beforehand could be a solution in order to calibrate the IMU. The static bias can be easily corrected as it is constant and a result of a wrong calibration of the sensors. On the other hand, the dynamic bias is the in-run variation and changes over the time. Even if the drift error is about the 10% of the static bias error, it is the one with the highest influence in the final implementation, as it cannot be easily corrected [28].
- Noise: another unwanted signal generated from internal electronics that interferes with measurement of the desired signal. As almost all of it is random noise (white noise), it will not be possible to predict it in advance. Nonetheless, the positive point is that in general, velocity, position, or pitch or roll error from the accelerometer or gyroscope white noise will be smaller than the other described correctable noise sources for IMU [29]. All in all, the whole sum of noises will have a lower effect than the bias errors in the IMU.

C. Known Blocked Scenarios

Known Blocked Scenarios (KBS) is a software enhancement that uses the knowledge of the zone in which the complementary positioning system is working to forecast areas in which GNSS signals are not available or could give misleading information. With this map based program, a better choice of the available positioning technologies leads to an enhanced data management and by hence to a system with a better position estimation (further details in III.C).

Fig. 1. Positioning and line characterization system hardware for the prototype implementation.

III. CPS Architecture

This section describes CPS architecture including both hardware and software.

A. CPS Hardware

Fig. 1 shows the hardware architecture of the CPS. Physically, it includes a multisensor equipment for location estimation that consists of the following:

- GNSS receiver (UBLOX EVK-6T-0-001): Responsible for providing an absolute position to the system. This position is used as a reference position while the GNSS system is available and its standard deviation is under a certain level. The receiver used in the CPS for the measurements, showed in the upcoming Sections IV and V, is a low-mid range GNSS receiver. The main two objectives of the receiver are the following: to record the raw data for the later assessment of the KBS threshold (Section C.1) and to perform a standalone positioning operation if there is no access to the on-board GNSS positioning system (Section C.2).
- Inertial Measurement Unit (IMU) (AIMS Navigation 0817111411): To perform a relative position estimation using the linear accelerations and angular velocities computed along with the gravity when the CPS crosses a GNSS blocked area. It is also used when a higher frequency of positioning is wanted making the availability of the system higher. The IMU employed is a low cost system.

Fig. 2. Antennas placed in ALN 668 3114 roof (Courtesy of ASTS).

- Microprocessor: The sensors used to obtain the data for positioning are connected to a microprocessor. It is the device in charge of receiving all the data from the sensors and fusing them to have a precise position estimation. The software of the positioning techniques such as the GNSS, IMU and the KBS are executed in it. The databases needed for the KBS are stored in it as well.
- AC/DC energy converter: The used sensors and microprocessor need DC power to perform their operation. Usually the power supplies available in trains are 220V AC power supplies. The introduction of this converter deals with this issue and helps with the discontinuities appearing in the power supply of the trains as it can maintain the DC power supply required for some seconds.
- GSM/UMTS Modem: CPS has also the ability to send and receive information in real time through a communication modem connected to a GSM/UMTS antenna. The data management is done through this modem, and the CPS can be remotely managed.

Excluding the IMU that needs to be placed in a stable place of the vehicle, the rest of the modules are placed into a 25 cm × 16 cm × 20.5 cm metallic container in order to make the system portable. CPS has its own GNSS receiver, and in this way it can perform stand-alone positioning.

The system uses two antennas to obtain all the needed information to work stand-alone (See Fig. 2). A dual GSM/UMTS antenna and a GNSS antenna, both compatible with the bands that are used. In the case in which the CPS is working as a complementary system to the existing on-board GNSS, the on-board GNSS antenna could be used in order to reuse the available equipment. Nevertheless, having a redundant GNSS system could help in terms of integrity due to the existence of two systems working in a similar way what gives backup to the system.

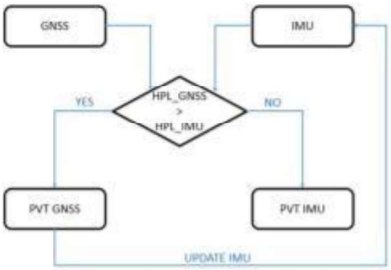


Fig. 3. GNSS+IMU integration.

B. GNSS and IMU Based Positioning Algorithm

The GNSS and IMU based positioning algorithm is the base of the CPS when providing a 100% available position. The KBS is built over this first GNSS and IMU based positioning algorithm, as an enhancement to the data that are shared by it.

This GNSS and IMU module uses the data received from a GNSS receiver in order to compute the position of the train in which it is installed. Nonetheless, taking into account that GNSS positioning has performance problems under harsh environments such as canyons or tunnels, the information received from an IMU is included in the positioning algorithm. The IMU performs a dead reckoning algorithm from a known starting point given by GNSS. This approach ensures a 100% availability in the positioning solution. However, the computed position is not always accurate enough for the system to work fulfilling the requirements of a given use case. For example, when getting into a tunnel, the last GNSS position could contain misleading information due to the degraded GNSS information obtained in the proximity of the blocking scenario. This initial misleading situation leads the IMU to give a wrong direction and position when performing the dead reckoning algorithm. IMUs have accumulative errors and having an erroneous heading direction incurs in an even bigger erroneous position estimate. That is the reason to introduce a Known Blocked Scenarios (KBS) algorithm to overcome these difficulties. Fig. 3 shows a diagram in which the situation above is explained. When doing the GNSS and IMU integration, the HPL of both technologies are compared. If the HPL of the GNSS is lower than the HPL of the IMU, the PVT obtained from the GNSS is the one given by the system. And the IMU is updated with it. In the other hand, if the HPL of the IMU is lower than the HPL of the GNSS, the PVT is obtained from the dead-reckoning of the IMU.

C. KBS

The KBS is the novel map aided software positioning enhancement proposed by this paper. It helps the GNSS and IMU positioning algorithm in terms of anticipating blocking scenarios and minimizing the use of GNSS misleading information.

Having a reliable heading and position of the train is important before entering a GNSS blocked area. The orientation of the train

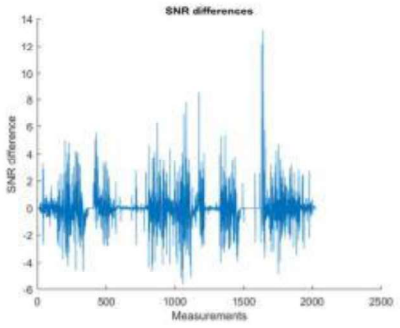


Fig. 4. Measured SNR differences in a complete scenario with different environments for GNSS.

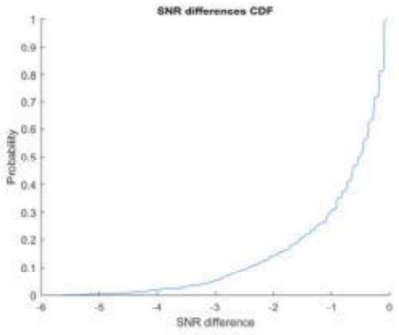


Fig. 5. CDF of SNR decreases between epochs for Threshold detection.

is computed before entering blocked scenarios in order to perform an accurate positioning operation. With a non-redundant system, the heading angle is obtained from the last "a priori" reliable position of the GNSS receiver. Thus, detecting erroneous information improves the performance of the solution in GNSS blocked areas.

Regarding the use of the KBS, two main steps are defined. The first step is the determination of known block scenarios for the determination of the KBS threshold and the second step is the KBS operation mode to enhance the position estimate in which real time positioning operation is performed.

1) Determination of the GNSS Blocked Scenarios and KBS Threshold Determination: The aim of the KBS is to improve the positioning and it is done under map located known blocked scenarios. Thus, for the KBS to be available, before starting its operation, it is necessary to first detect and create a database with these spots and sections with known blocked scenarios.

In this case, for the determination of a degraded GNSS area, the received signal to noise ratio (SNR) of the GNSS signal strengths is analyzed. Having a lower SNR increases the probability of having error in the received information [30].

Fig. 4 shows a real case for the known blocked scenarios. It shows the SNR difference between a GNSS measurement and the previous one.

Every scenario produces an attenuation based on the surrounding environment. So in the first phase, an analysis of this environment needs to be done. With that purpose, maps are consulted and the trackside is inspected with the aim of identifying the possible GNSS blockers such as urban canyons, tunnels, woods or other artificial blockers and constructions. A database with all the areas where possible GNSS outages are expected is constructed and proper position and heading values are introduced based on reliable track database information.

With the aim of reducing the computation time for the enhancement of the position estimate, the KBS is only triggered under certain conditions. Thus, based on the analysis of the SNR along these known block scenarios campaigns detection (see Fig. 4), the sensibility and the KBS threshold is calculated [31]. This threshold is highly dependent on the antenna positioning and the trackside environment. If the antenna has a full view of the sky, the threshold will be higher as the difference between blocked and non-blocked scenarios will be more elevated due to the good health of the received signal in non-blocked scenarios. This is why the authors propose to perform trips in different times of the day, with different satellite constellations in view and different weather to have a reliable threshold configuration.

In a railway with very differentiated parts, in which the train has line of sight with the satellites every moment apart from the blocked scenarios, 2 runs will be enough to build an accurate database as the satellite architecture and the surrounding environment will not have a big influence. In the case in which the surroundings of the blocked scenarios are GNSS challenging areas, more runs are needed in order to differ between blocked and no-blocked areas as the SNR decreases are lower, and the threshold must be calculated more carefully. Runs in different times of the day are recommended in order to evaluate the performance with different satellite architectures.

The sensibility of the KBS is triggered based on the SNR decreases. This means that a SNR fall down compared with the previous epoch determines that GNSS signals are not trustful anymore unless the SNR shortfall is recovered. This allows complementary positioning sensors to have a trustful position once GNSS signals are not available, which improves the position estimation results. The KBS system allows detecting the entrance of the tunnels or other blocking structures where the complementary positioning sensors take the key role regarding to the position estimation. Summing up, the KBS threshold is set-up based on the SNR decreases.

For the particular case of the test campaign shown in Fig. 4, the selected threshold to trigger the software enhancement tool was -2.5 dB between two consecutive SNR measurements. To select that threshold, a Cumulative Density Function (CDF) is used. The threshold is selected to be the one which marks the 10% of probability (Fig. 5).

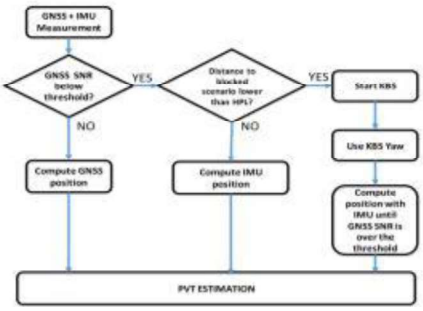


Fig. 6. KBS operation.

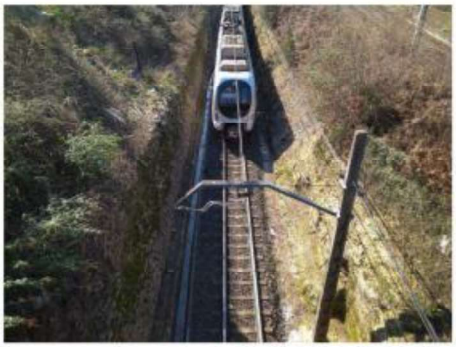


Fig. 7. General environment in EuskoTren Lane.

Fig. 8. Section of Eusko Trenbide Sarea (ETS) network.

2) KBS Operation: Once the KBS threshold is properly selected, the positioning algorithm is able to enhance the positioning in these blocked scenarios. Fig. 6 shows the KBS algorithm methods which relies on having a more accurate starting point which leads to a more accurate position estimate.

The KBS determines not the valid GNSS, but the degraded GNSS information based on the strong SNR decreases between epochs. The KBS takes advantage of the detection of strong blocking areas, which are mainly tunnels. Tunnels are the main blocking areas, due to their strong effect and their limitation to a defined area. At this point, apart from detecting and discarding the GNSS information a "database" is set up to reduce the errors of the next time epochs where GNSS is not available by resetting the yaw information of the algorithm.

The positioning system proceeds in the same manner every time that it receives a GNSS information. First of all, it compares if the selected triggering option is reached. If not, the GNSS is considered reliable and the algorithm computes the position estimate. If the KBS is triggered, the KBS takes a look up into the KBS database and based in the information of the previous reliable estimate and the distance to the different known blocked scenarios entries, it uses the stored information to correct it. The distance to use an entry from the database is, as maximum, the HPL (Horizontal Protection Level). The HPL is the statistical error bound obtained during the real time operation of the positioning algorithm.

This methodology is repeated until the KBS detects a valid GNSS.

IV. Test Definition

This section shows the tests proposed and the measurements done in order to validate the CPS algorithms. A demanding environment near Donostia / San Sebastian (Spain), was selected. In this section the track and train selection and ground truth generation are introduced.

A. Track and Train Selection

The chosen railway is placed between Hendaia and Donostia, in a regional train that puts together France and Spain operated by Eusko Trenbide Sarea (ETS). This railway has a difficult orography since GNSS blockers appear in more than 50% of the track analysed. Most of the journey is done inside tunnels or canyons; both urban or natural (see Fig. 7).

The trial track is mostly set in an urban environment inside the city of Donostia and the villages near it. However, there are also suburban and rural areas in the more than 20km of track. The rural areas are mountainous areas in which the signal reception is not direct in all the cases. The environment in which the algorithm was tested can be considered a GNSS demanding zone in which long periods of lack of signal were recorded.

The tests took place in service during approximately 30 days, in which the full operation of the trains during their usual journeys were recorded. In this way, different satellite constellations in view were tested, and the collected data was more realistic.

The route going from Fanderia station to Errenteria station (Fig. 8) has been chosen for the tests as it is a representative part of the journey because the track is surrounded by buildings and the train travels through a tunnel (see Fig. 8 (green area)). The degradation of the signal prior to the entrance to the tunnel affects the performance of the positioning system. The surroundings

Fig. 9. EuskoTren Serie 900 Locomotive.

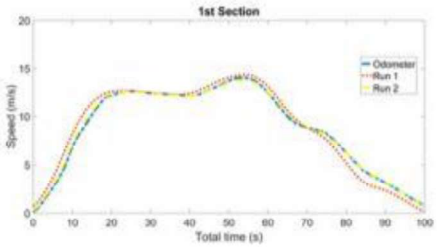


Fig. 10. Speed profile comparison.

and the fact that the tunnel is preceded by a curve also influences the result without the software enhancement proposed by this paper.

The vehicle used was an electric locomotive Serie 900 from Euskotren. (See Fig. 9).

B. Ground Truth Generation

For this measurement campaign, the ground truth is obtained using independent sensors installed on-board. In this case, the Euskotren on-board odometer [32] has been used.

First, the information source for this ground truth was independent to the information used by the CPS in order to ensure a non-dependent behaviour of the error.

The odometer data were interpolated to obtain a smoother speed profile. Additionally, in order to check the proper behaviour of this profile it has been compared and validated with other GNSS based speed profiles from other journeys carried out during this period. (see Fig. 10) and used to determine the starting point of the journey.

The obtained accuracy after interpolating the information and smoothing the behaviour depicts a representative and consistent behaviour compared with other runs. Afterwards, this speed profile is projected into positions in combination with a track database composed by the geolocated information (x,y,z and

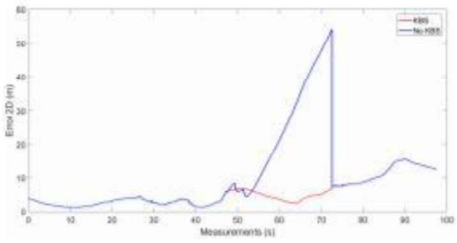


Fig. 11. Error 2D full journey.

travelled distance) to obtain the set of timestamped trustful positions which are considered as the ground truth, in order to compare them with the PVT estimates obtained using the CPS [33]. The section under study in this work is 100 seconds long so the drift error of the odometer is affecting the accuracy of the ground truth less than the order of magnitude of the resulting data from the CPS solution shown in the results section. For the measurement campaign shown here, no slope and no slippery effects are reported in the section under study at a 10m/s average speed, so a speed error margin of 0.333% could be applied for the 100 seconds from one stop to the other. That represents a maximum error of $+/-$ 3.33m of error with the odometer algorithm, which is lower than the order of magnitude of the error computed by CPS.

V. Results

This section presents the obtained results. The results here shown have been obtained from a 1 month 20h/day run on a CAF 900 series train, operated by Euskotren over Eusko Trenbide Sareak infrastructure.

First of all, the results obtained without the use of the KBS will be presented to later on be compared with the results when the KBS enhancement technique is applied. During this performance evaluation, accuracy and availability results are going to be discussed as they are the main drivers for the railway applications, as presented in the introduction. The accuracy error is assessed by the Euclidean distance in the horizontal plane between the ground truth position and the position estimate.

It is important to mention that the analyzed sector is part of a longer journey and the initial errors are dependent on the previous position estimations. However, the tunnel is reached between the 50^{th} and 70^{th} second. A change of behavior in the error can be observed during the mentioned seconds in Fig. 11.

A. Measurement Campaign Analysis

To be able to make a good comparison between both systems, it is important to use the same observables generated during the test runs when using KBS and when not. Fig. 11 shows the 2D errors computed for both KBS and non KBS operation modes during the whole journey.

However, to analyse the whole journey could quit the focus from the important part of the paper. That is the reason to show

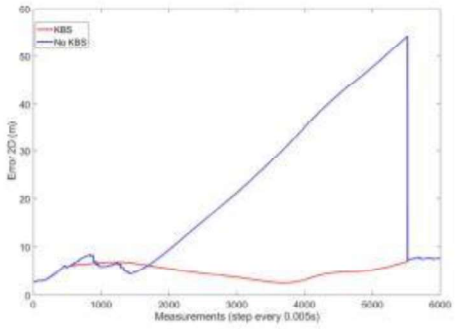


Fig. 12. Error 2D in KBS operation zone.

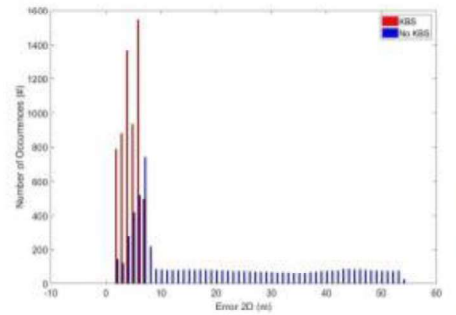


Fig. 13. Histogram of the error 2D.

the graphs in the area in which the KBS enters in operation. The focus will show the real improvement of the KBS. Fig. 12 shows the KBS operation zone that is analysed.

The accuracy results of the CPS systems without the KBS, showed an error peak that reached 54.29 meters. It is important to remark that even if the tunnel is almost straight using misleading information can lead to this undesired effects (see Fig. 12). The mean error of this section is 21.15 meters.

In the other hand, when using KBS, the accuracy is significantly better. The maximum error peak is lower than 7.96 meters and the main improvement is that the mean error is reduced to 5.07 meters (See Fig. 12).

Fig. 13 analyses the frequency of the different error during this journey. The error is mainly spread at low error distances; however, when KBS is not being used, there is a long tail with error that reach the aforementioned value of more than 50 meters.

When using KBS, the error is also mainly spread at low error distances; but the maximum error reaches a value of 8 meters, which is smaller than the maximum error obtained without using the KBS.

Having the error as a distribution function, the CDF presents the probability that the error will take a value lower than x. And

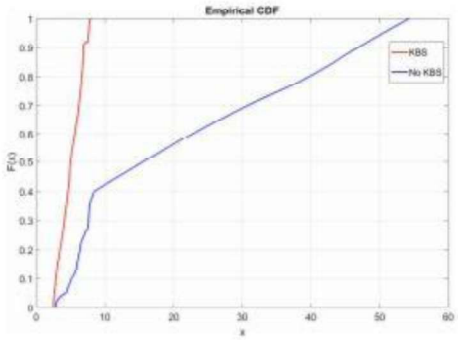


Fig. 14. CDF of the empirical error 2D.

here it is based on the empirical data obtained. It can be stated that the CDF (95%) is around 50 meters (see Fig. 14). This means that the 95 percent of the time the error is lower or equal to 50 meters.

The CDF obtained from the error information when the KBS is used, bounds the error at 95% in 8 meters (See Fig. 14).

As the graphs show, when the KBS is not used, the error obtained in the tunnel environment decreases the performance significantly and thus improving this matter will have a great impact on the overall picture.

The performance of the KBS improves significantly the performance of the GNSS and IMU dead reckoning algorithm by detecting misleading information of the GNSS.

VI. Conclusion

This paper presents a progress for the positioning systems in the railway sector. One of the key aspects that is important to remark is the need of complementary information sources, on top of GNSS, for the continuous position to provide fully availability at the expense of the accuracy error performance. For most of the railway applications, the availability of the systems is as important as the accuracy. It is necessary to have access to the position of the train during the whole journey. So as to overcome these limitations, the Complementary Positioning System (CPS) has been introduced in this paper with a map aid enhancement software called KBS algorithm.

The CPS gives a positioning alternative to overcome the use of degrade GNSS position due to the introduction of a Known Blocking Scenario software module combined with GNSS and IMU algorithm.

The system presented in this work provides a step forward in railway positioning with a low cost solution approach that allows the discrimination of all the GNSS degraded situations enhancing the availability and overall performance in harsh environments.

The use of the CPS system and its software enhancement modules improves the existing on-board systems and can lead

the way to eliminate the trackside equipment which can be useful for applications such as regional train lines or freight tracking in which the requirements are more relaxed than for ETCS level 3.

During the measurement campaign in EuskoTren, the suitability of the KBS to detect and mitigate GNSS shadowed areas in demanding environments is proved with the data of the improvement shown in section V. The maximum error reached is more than six times lower and the mean error is four times lower when the KBS is used. Taking into account the introduced results, the KBS gives the possibility to a positioning system to go through larger tunnels without having to change the existing alert limits. Moreover, the KBS opens the possibility to have better operation limits in the existing positioning systems.

Acknowledgment

The authors would like to express special thanks to their colleagues from EuskoTrenbide Sarea (ETS) and EuskoTren who allowed both measuring campaigns in the project's framework and the use of their infrastructure.

References

[1] ERRAC, "Rail route 2050: The sustainable backbone of the single European transport area," Eur. Rail Res. Advisory Council, Paris, France, Tech. Rep., 2013. [Online]. Available: http://www.errac.org/wp-content/uploads/2013/11/D9-SRRA-RAILROUTE2050.pdf

[2] "Shift2Rail mission and objectives." [Online]. Available: https://shift2rail.org/about-shift2rail/mission-and-objectives/. Accessed on: Oct. 28, 2019.

[3] V. Amdas, T. Bradbury, S. Denniss, D. Chapman, R. Bloomfield, and D. Fisher, "ERTMS level 3 risks and benefits to UK railways (final report)," Transp. Res. Lab., Eur. Railway Traffic Manage. Syst., France, Paris, Tech. Rep. PPCA00004, 2010.

[4] L. Marradi *et al.*, "GNSS for enhanced odometry: The GRAIL-2 results," in *Proc. NAVITEC 2012 Eur. Workshop GNSS Signals Signal Process.*, pp. 1–7.

[5] E. González, C. Prados, V. Antón, and B. Kennes, "GRAIL-2: Enhanced odometry based on GNSS," *Procedia - Social Behavioral Sci. Transp. Res. Arena 2012*, vol. 48, pp. 880–887, 2012.

[6] "GSA. GRAIL-2—Summary." 2012. [Online]. Available: http://grail2.ineco.es/Grail2/html/main.jsp. Accessed on: Oct. 28, 2019.

[7] A. Urech, *GNSS in Rail Safety Applications*. Madrid, Spain: I Rail Technological Forum For Internationalization, 2011

[8] G. Barbu and J. Marais, *The SATLOC Project*, Paris, France: *Transport Research Arena*, 2014.

[9] "GSA. GRAIL-2—Summary." [Online]. Available: https://www.gsa.europa.eu/satellite-based-operation-and-management-local-low-traffic-lines. Accessed on: Oct. 28, 2019.

[10] J. Marais, C. Meurie, A. Flancquart, S. Lithgow, and G. Barbu, "SATLOC. Demonstration of an innovative concept specification and results of simulation tools applied on photographed geo-referenced particular route environment," 2014. [Online]. Available: http://hal.univ-smb.fr/IFSTTAR/hal-01061356. Accessed on: Jan. 22, 2019.

[11] T. P. K. Nguyen, J. Beugin, and J. Marais, "Dependability evaluation of a GNSS and ECS based localisation unit for railway vehicles," in *Proc. 13th Int. Conf. ITS Telecommun.*, pp. 474–479.

[12] "GSC. GaLoROI—Summary," 2014. [Online]. Available: https://www.gsc-europa.eu/markets/rd/fp7-3rd-call-projects/galoroi-galileo-localization-for-railway-operation-innovation. Accessed on: Jan. 22, 2019.

[13] D. Lu and E. Schnieder, "Performance evaluation of GNSS for train localization," *IEEE Trans. Intell. Transp. Syst.* vol. 16, no. 2, pp. 154–1059, Apr. 2015.

[14] "GSA. GaLoROI—Summary." [Online]. Available: https://www.gsa.europa.eu/galileo-localisation-railway-operation-innovation. Accessed on: Oct. 28, 2019.

[15] A. Neri, A. Filip, F. Rispoli, and A. M. Vegni, "An analytical evaluation for hazardous failure rate in a satellite-based train positioning system with reference to the ERTMS train control systems," in *Proc. 25th Int. Tech. Meeting Satell. Division Inst. Navigation*, pp. 2770–2784.

[16] F. Rispoli, M. Castorina, A. Neri, A. Filip, and G. Di Mambro, F. Senesi, "Recent progress in application of GNSS and advanced communications for railway signaling," in *Proc. 23rd Int. Conf. Radioelektronika*, 2013, pp. 13–22.

[17] N. Kassabian, L. L. Presti, and F. Rispoli, "Augmented GNSS differential corrections minimum meansquare error estimation sensitivity to spatial correlationmodeling errors," *Open Access Sensors*, vol. 14, pp. 10258–10272, Jun. 2014.

[18] "ERSAT EAV grant agreement No 640747," [Online]. Available: http://www.ersat-eav.eu/. Accessed on: Jan. 22, 2019.

[19] "RHINOS project summary," [Online]. Available: https://cordis.europa.eu/project/rcn/199612_es.html/. Accessed on: Jan. 22, 2019.

[20] P. Salvatori, C. Stallo, S. Pullen, S. Lo, and P. Enge, "Use of SBAS corrections with local-area monitoring for railway guidance and control applications." in *Proc. IEEE/ION Position, Location Navigation Symp.*, Apr. 2018, pp. 23–26.

[21] STARS Project. [Online]. Available: http://www.stars-rail.eu/project/. Accessed on: Jan. 22, 2019.

[22] B. Stamm and P. Gurnik, "STARS project—Satellite technology for advanced railway signalling," [Online]. Available: http://www.stars-rail.eu/wp-content/uploads/2017/05/STARS-project-Satellite-technology-for-advanced-railway-signalling.pdf. Accessed on: Jan. 22, 2019.

[23] "X2RAIL-2 – Objectives," 2018. [Online]. Available: http://projects.shift2rail.org/s2r_ip2_n.aspx?p=X2RAIL-2. Accessed on: Jan. 22, 2019.

[24] "ERTMS on SATELLITE galileo game changer," Periodic Rep. for period 1. [Online]. Available: https://cordis.europa.eu/project/rcn/212900_en.html. Accessed on: Jan. 22, 2019.

[25] J. Marais, R. Hmida, A. Flancquart, S. Sabina, and M. Ciaffi, "Video-based classification of railway track areas for GNSS-based virtual balise solutions in the ERSAT GGC project," in *Proc. Int. Tech. Meeting Inst. Navigation*, Reston, VA, USA, Jan. 2018, pp. 196–205.

[26] J. Goya, L. Zamora-Cadenas, S. Arrizabalaga, A. Brazález, J. Meléndez, and J. Mendizabal, "Advanced train location simulator (ATLAS) for developing, testing and validating on-board railway location systems," *Eur. Transp. Res. Rev.*, vol. 7, 2015, Art. no. 24.

[27] D.E. Wells *et al.*, *Guide to GPS Positioning*. Fredericton, NB, Canada: Canadian GPS Associates, 1986

[28] A. Noureldin, T. B. Karamat, and J. Georgy, *Fundamentals of Inertial Navigation, Satellite-Based Positioning and Their Integration*. Berlin, Germany: Springer-Verlag, 2013.

[29] R. Chow, *Evaluating Inertial Measurement Units*. Long Beach, CA, USA: Epson Electron. Amer., Nov. 2011.

[30] G. De Miguel, J. Goya, J. Uranga, I. Adin, and J. Mendizabal, "GNSS complementary positioning system performance in railway domain," in *Proc. 15th Int. Conf. ITS Telecommun.*, May 2017, pp. 1–7.

[31] S. Arrizabalaga, I. Adin, G. De Miguel, J. Goya, L. Zamora-Cadenas, and J. Mendizabal, "Data acquisition for complementary positioning system in GNSS denied areas," in *Proc. World Congr. Railway Res.*, 2016, Milan, Italy, pp. 4562–4571.

[32] European Railway Agency, "Performance requirements for interoperability," European Railway Agency, Valenciennes, France, Subset-041, Jun. 2016.

[33] J. Goya, G. D. Miguel, S. Arrizahalaga, L. Zamora-Cadenas, I. Adin and J. Mendizabal, "Methodology and key performance indicators (KPIs) for railway on-board positioning systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 19. no. 12. pp. 4035–4042, Dec. 2018.

Gorka de Miguel received the M.Sc. degree in telecommunications engineering from TECNUN (School of Engineering of San Sebastián), University of Navarra, Donostia-San Sebastián, Spain, in 2015. In 2015, he joined the CEIT-IK4 Research Centre, Donostia-San Sebastián, Spain, where he is currently a Research Assistant and a Ph.D. Student within the Transport and Sustainable Mobility Group. He is also an Assistant Lecturer in Electronic Fabrication Systems with TECNUN. His research interests include the field of positioning and software development. He is now actively participating in H2020 European funded projects in Railway signaling and positioning topics.

Jon Goya received the M.Sc. degree in telecommunications engineering in 2011 and the Ph.D. degree in 2016 from the University of Navarra, Donostia-San Sebastián, Spain. He is currently a Lecturer with TECNUN, University of Navarra, and a Researcher with CEIT-IK4 Research Centre, Donostia-San Sebastián, Spain. His professional research interests include the simulation of on-board positioning system and performance analysis for railway. He has participated in FP7 projects coordinated by CEIT (EATS) and is now actively participating in H2020 European funded and Shift2Rail projects in railway signaling and positioning topics: ERSAT-EAV and FR8RAIL.

Nerea Fernández received the M.Sc. degree in telecommunications engineering from the Faculty of Engineering in Bilbao, University of the Basque Country, Bilbao, Spain, in 2016. In 2017, she joined the CEIT-IK4 Research Centre, Donostia-San Sebastián, Spain, where she is currently a Research Assistant within the Transport and Sustainable Mobility Group. Her research interests include the field of railway signaling specially in zero on-site testing and adaptable communication systems.

Saioa Arrizabalaga received the M.Sc. degree in telecommunication engineering from the Faculty of Engineering in Bilbao, University of the Basque Country, Bilbao, Spain, in 2003, and the Ph.D. degree in engineering from TECNUN, University of Navarra, Donostia-San Sebastián, Spain, in 2009. She is currently a Lecturer with TECNUN and a Researcher with CEIT-IK4 Research Centre, Donostia-San Sebastián, Spain. She is involved in the participation of European research projects, direction of doctoral theses, scientific and technical publications in national and international journals and conferences. She is the Author of the book *Multi-Service QoS Architecture for a Multi-Dwelling Gateway*.

Jaizki Mendizabal received the M.Sc. and Ph.D. degrees in electrical engineering from TECNUN, University of Navarra, Donostia-San Sebastián, Spain, in 2000 and 2006, respectively. He was with Fraunhofer Institut für Integrierte Schaltungen, Erlangen, Germany, and SANYO Electric Ltd, Gifu, Japan, as an RF-IC Designer. He is currently with CEIT-IK4 Research Centre, Donostia-San Sebastián, Spain. He is also a Lecturer of Communications Electronics and Communications via Radio with TECNUN. His research interests include communications and signaling systems for the railway industry.

Iñigo Adin received the M.Sc. degree in electronics engineering in 2003 and the Ph.D. degree in 2007 from the University of Navarra, Donostia-San Sebastián, Spain. He is the Author or Co-Author of one patent, two technical books, an invited chapter, and 40 articles in journals and international conferences. His research interests include safety-critical designs, with special interest in positioning, communications, electromagnetic compatibility, and transport interoperability. He was the Coordinator of the FP7 European Project TREND and now coordinates the H2020 AIOSAT for GSA.

A Review of the Evolution of the Integrity Methods Applied in GNSS

Paul Zabalegui, Gorka De Miguel, Alejandro Pérez, Jaizki Mendizabal, Jon Goya and Iñigo Adín.

IEEE Access (Q1). March 2020.

IEEE Access

Received January 21, 2020, accepted February 13, 2020, date of publication March 2, 2020, date of current version March 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2977455

A Review of the Evolution of the Integrity Methods Applied in GNSS

PAUL ZABALEGUI, GORKA DE MIGUEL, ALEJANDRO PÉREZ, JAIZKI MENDIZABAL, (Member, IEEE), JON GOYA, AND IÑIGO ADIN, (Member, IEEE)

CEIT, 20018 Donostia-San Sebastián, Spain
Electrical and Electronic Engineering Department, TECNUN School of Engineering, Universidad de Navarra, 20018 Donostia-San Sebastián, Spain

Corresponding author: Paul Zabalegui (pzabalegui@ceit.es)

This work was supported in part by the European projects: FR8Rail-II, which received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 826206; and by autonomous indoor and outdoor safety tracking system (AIOSAT), which received funding from the European GNSS Agency under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 776425.

ABSTRACT The use of GNSS technologies has been spreading over time up to a point in which a huge diversity of applications require their use. Due to this demand, GNSS has turned into a more reliable technology, as multiple aspects of it have evolved. Integrity has become a vital aspect of being considered when using GNSS. The following document gathers and shows different aspects of integrity in terms of GNSS. The paper mainly focuses on the description of different receiver autonomous integrity monitoring methods. For this purpose, basic concepts and possible GNSS error sources (and their corresponding solutions) are introduced. Afterward, an explanation and a classification of the integrity monitoring techniques is given, where the fault detection and exclusion methods and different protection level computation formulas are analyzed.

INDEX TERMS Fault detection and exclusion, GNSS, integrity monitoring, protection level, receiver autonomous integrity monitoring.

I. INTRODUCTION

Nowadays, the demanding requirements of industry and society urge positioning systems to be safer, more reliable, and even faster. In this scenario, GNSS technologies are considerably spreading since they emerged. The amount of applications and systems that rely on GNSS PVT solutions is substantially growing [1], [2], and users benefit from the evolutions on this topic.

For the aim of fulfilling the high demanding requirements for positioning applications, multiple technologies and techniques have been developed to increase the confidence of GNSS in terms of accuracy, integrity, and continuity. This paper focuses on the review and explanation of the different fault detection and exclusion (FDE) methods found in the literature to improve the integrity of the estimated GNSS PVT solution. These methods are based on the detection and exclusion, when possible, of faulty satellites for the computation of the PVT solution. According to [3], continuity is "a system's ability to function without interruption." As it is dependent on the positioning system instead of the processing of the data, it will not be discussed in the document. Even if accuracy improving techniques are not discussed in this paper, it is important to mention that the fault detection and exclusion methods can improve the accuracy of the solution. Note that most of the time, position, velocity, and time will be considered as the GNSS solution; nevertheless, the discussed methods will mainly refer to integrity in terms of position.

GNSS integrity methods were developed to improve aircraft navigation. Nowadays, these research branches have spread to multiple ground transportation systems. It is due to this that many references about the topic are focused on ground transportation, as the harsh environments and the movement of these systems lead to more dynamic and changing, thus, challenging scenarios.

According to the last issue of the federal radio navigation plan [4], "Integrity is the measure of the **trust** that can be placed in the correctness of the information supplied by a navigation system. Integrity includes the ability of the system to provide timely warnings to users when the system should not be used for navigation".

The associate editor coordinating the review of this manuscript and approving it for publication was Hayder Al-Hraishawi.

Integrity is connected and defined by a set of parameters that are dependent on the target application. The following list contains the ones that are going to be used in this paper [5], [6]:

Alert Limit (AL) is the error tolerance a system has, which cannot be exceeded without issuing a warning.

Time to Alert (TTA) is the maximum allowable time elapsed from the moment the integrity threshold is crossed until the alert is issued.

Integrity Risk (IR) is the probability that the position error exceeds the Alert Limit.

Position Error (PE) is the difference between the measured position and the real position, also known as ground truth.

Protection Level (PL) is the position error that the algorithm guarantees that it will not be exceeded without being detected.

False alert (FA) is the event in which an alert is issued without surpassing the alert limit.

Missed detection (MD) is the event in which there is a fault that is not detected by the algorithm.

Positioning failure (PF) is the event in which the positioning error exceeds the defined Protection Level.

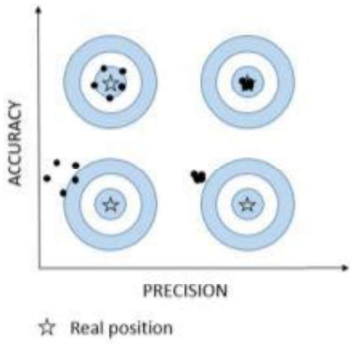


FIGURE 1. Precision vs. Accuracy.

The positioning error is a term that appears in some of the previous definitions. This error is understood as the "**correctness**" of the position. Two terms are commonly used to refer to the "correctness": **accuracy and precision**, which they do not mean the same. As shown in FIGURE 1, the accuracy, on the one hand, is the degree of proximity of the computed solutions to the real position. The precision, on the other hand, is the proximity between the computed solutions between each other. In statistical terms, assuming an undetermined probability distribution characterized by its mean and variance, the accuracy can be understood as the displacement of the distribution's mean from the real position, whereas the precision would be related to the width of the distribution and its variance.

When talking about positioning systems, a combination of precision and accuracy is usually pursued. Nevertheless, systems are not ideal; thus, positioning errors are always present. These errors, as seen in FIGURE 1, are usually bounded by the accuracy rather than by the precision, as low accuracies often lead to bigger distances from the ground truth than what low precision does. Consequently, positioning systems tend to define and set an error tolerance threshold (AL) that should not be surpassed by the error in order to maintain the integrity of the system.

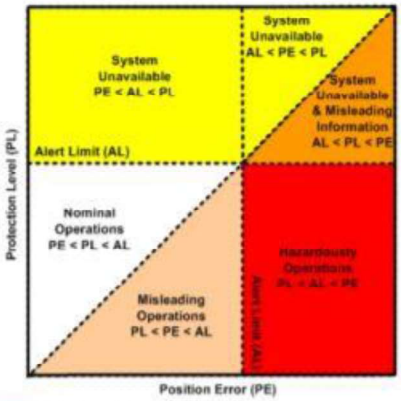


FIGURE 2. The Stanford diagram [7].

The Stanford diagram shown in FIGURE 2 specifies the integrity and availability criteria that describe the performance of a system, as a function of a user-defined alert-limit. It describes the system's state according to the relation between the tolerable error, the estimated error, and the real error.

The expected nominal operation mode, according to the diagram, implies having a positioning error smaller than the calculated protection level and a protection level which is smaller than the alert limit.

II. INTEGRITY FOR GNSS POSITIONING

Incorrect integrity monitoring could lead to devastating consequences for safety-critical applications scenarios. This means that integrity is dependent on the quality of the input information/signals and the detection and mitigation techniques implemented in the positioning algorithm for the multiple error sources that receivers need to handle. The following section mentions the different error sources that could degrade the input signals together with multiple error mitigation techniques that are used. Afterward, different kinds of integrity systems are discussed in section B, with further details on the autonomous method, as it is commonly found in transportation systems and intelligent transport system (ITS)

applications. Finally, due to the relevance of the mentioned method, a classification of these techniques is carried out in section C.

A. CHARACTERIZATION AND REDUCTION OF THE MEASUREMENT ERROR SOURCES

As it is known and referred to in the literature, GNSS signals suffer from multiple error sources before reaching the receiver's antenna. Some of the most common error sources found are random noise, shadowing, multipath, NLOS (Non-Line-Of-Sight), interferences and attenuation due to signal in space and the receiver's surroundings (e.g., skylines, canyons, tunnels, etc.). These phenomena are closely linked to GNSS position performance and integrity evaluation.

Consequently, a wide variety of techniques have been developed in order to reduce or mitigate the aforementioned error sources:

1) SIGNAL WEIGHTING METHODS

These methods are commonly used as criteria of signal reliability, giving proportional weights to each measurement according to certain criteria (e.g., received signal to noise ratio) intending to relay more in better quality signals. This weighting is usually done by modifying/introducing a covariance matrix according to the mentioned criteria. This covariance matrix reflects the errors that affect each satellite's signal and, as it is assumed that the errors that affect each satellite do not affect the rest of them, it tends to be diagonal

$$\begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_m \end{pmatrix},$$

where σ_m is the variance of pseudorange measurement errors, and m is the number of observations [8].

These values can be computed according to models that consider different physical parameters to weight the variance. The author shows in [9] a method that employs the inherent information in the carrier-to-noise-power-density ratio (C/N0) in order to estimate the random errors. This model represents the variance of an undifferenced phase observation as

$$\sigma_i^2 = V_i + C_i 10^{-\frac{C_i/N_0}{10}}, \tag{1}$$

where $V_i[m^2]$ and $C_i\left[m^2\,Hz\right]$ are parameters proper to the model that must be estimated and may vary according to the scenario. Further C/N0 based research is found in literature in [8]–[11].

Together with C/N0, satellite elevation can also be used to estimate the quality of the received signal, assuming that the higher the elevation, the more trustworthy it is. This is represented as [13]

$$\sigma_i^2 = \frac{1}{\mathrm{SIN}(\theta)^2}, \tag{2}$$

where θ is the satellite elevation angle.

The paper in [9] presents a novel hybridization weighting model that takes into account satellite elevation, the signal C/N0 and a LOS/NLOS indicator obtained from the Urban Trench Model (UTM)

$$\sigma^2 = k \cdot \frac{10^{-0.1 \cdot \frac{C}{N0}}}{\mathrm{SIN}(\theta)^2}, \tag{3}$$

where k is the LOS/NLOS indicator; $k = 1$ for LOS signals and $k = 0.5$ for NLOS signals.

The author performs in [13] a comparison between the presented weighting methods.

2) MULTIPATH AND NLOS RELATED TECHNIQUES

Multipath and NLOS mitigation and exploitation techniques are especially used in urban areas, where signal LOS is usually blocked, and signal beam rebounds are commonly found (multipath).

Four different strategies are employed to deal with these phenomena: **ignorer** and **avoidance** of the mentioned beams [14], **mitigation** of these through HW design (correlator modifications [15], correlator banks [14], [16]–[18], diverse polarization antennae [19]–[21], spatial diversity [22]–[24], etc.) or using quality parameters [25] and the **identification** of the NLOS beams for the later elimination or exploitation.

Elimination oriented identification techniques are more commonly used in cases where an LOS component is correctly received, and NLOS components can interfere with it, whereas the exploitation ones are more useful when no LOS beam is received and, thus, NLOS beams are the only information input [26].

The mentioned methods can be sub-classified into three main groups according to the required input: the ones that require a physical signal (usually implemented in HW), the ones that require the parameters obtained from processing the signals such as pseudorange, C/N0, Doppler, etc. (usually implemented in SW) and those which combine both GNSS data and external aid methods such as INS, map aiding, ray tracing, etc. (usually implemented in SW).

Despite the efforts to mitigate the effect of the error sources, the possibility of faulty measurements remains. These faults could imply misleading positions; consequently, error detecting and solving methods have been developed to ensure the integrity of the system.

B. INTEGRITY SYSTEMS

Integrity was originally developed for the aeronautical domain, where a single failure could cause severe fatalities. As mentioned in the introduction, integrity provides the user with timely warnings regarding the reliability of the navigation system through its navigation message. The employed health reporting of the GNSS system may not be well-timed to be considered appropriate for a real-time application that requires a quick failure reaction; thus, two different approaches have been developed to provide integrity: an autonomous one based on self-consistency check of redundant measurements, and a ground-station network-based

one [8], [27]. The diverse needs have resulted in the development of two different external-aid-based approaches.

Satellite-Based Augmentation Systems (SBAS), on the one hand, rely on the integrity-related information provided by geostationary satellites, as they improve the accuracy, reliability, and integrity of the GPS signal [28]. The geostationary satellites also provide ranging capabilities, so that they can be used as GNSS satellites to increase the availability [27].

FIGURE 3. SBAS architecture [74].

These systems (FIGURE 3) are based on regional networks of strategically distributed reference stations that measure GPS signal-in-space (SIS) for the purpose of computing the error estimates at the master stations [29].

Ground-Based Augmentation System (GBAS), on the other hand, provide integrity-related information data based on a locally located ground elements. In contrast to SBAS's regional distribution, GBAS employs a small number of reference stations that perform measurements and later corrections at just one location, which is why these systems can be usually found at airports [5], [27].

While providing integrity assurance is the aim of GBAS, it also increases the accuracy and precision below 1 m.

Receiver Autonomous Integrity Monitoring (RAIM) is a technology developed to estimate the integrity of, originally, the GPS. The main aim is to provide the receiver with the ability to perform a self-contained fault detection by comparing each GPS measurement to the other available measurements [27]. It is, thus, based on the consistency check of redundant range measurements [30]. There are two possible pseudorange error scenarios, the fault-free scenario and the faulty one, being the first one affected only by the nominal errors that are modeled as zero-mean independent Gaussian distributions, whereas, in the second ones, a bias is added to some of the range measurements [30].

According to [31], the inputs to RAIM algorithms are the standard deviation of the measurement noise, the measurement geometry, and a threshold that defines the probabilities for a false alert and a missed detection. The main outputs of RAIM algorithms are the protection levels (PL) and the ability to provide an alert.

Although there are many RAIM schemes (see section C), most RAIM algorithms include an error measurement and its corresponding bound in the form of a protection level (PL).

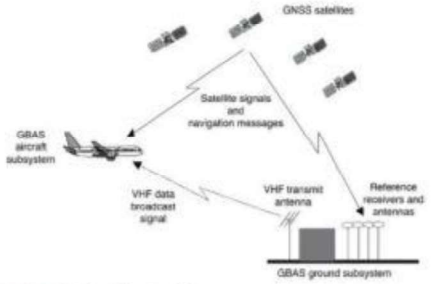


FIGURE 4. GBAS architecture [75].

These PL can be computed in different ways and can even be calculated without applying RAIM, as it will be discussed in Subsection 2.

As RAIM can only bound the positioning error with an *a priori* estimation, the previously mentioned FDE methods have been developed as an extension of RAIM to reduce the positioning error.

The following section will first explain the classical RAIM method, followed by different ways to compute PL. Together with this, the fault detection and exclusion methods will be explained.

1) CLASSICAL RAIM

The principle of the classical RAIM method is shown in FIGURE 5. Each of the diagonal slopes is related to each observable satellite, where the bigger the gradient of the slope, the bigger the position error caused by the ranging error from the said satellite. It is assumed that, in the worst-case scenario, a failure could remain after the FDE test. Consequently, the PL is expected to be such that overbounds the error caused by a faulty measurement at the biggest sloped satellite [32].

In the case in which there was one only faulty measurement and the rest were ideal, the protection level should be located at PR_0, at the point in which the slope crosses the threshold used in the FDE test (T_D) [32].

Nevertheless, as the rest of the measurements are not ideal, they introduce an error that can be modeled as zero-mean Gaussian variables (gray ellipses) around a new mean value. This value is called p_{bias}, and it is a deterministic value that depends on the number of visible satellites [33].

Taking into account the mentioned matters, the PL should be set so that the probability of misdetection (P_{md}) is as small as possible or, in other words, such that reduces the Integrity Risk.

2) COMPUTATION OF THE PROTECTION LEVEL (PL)

The PL is an upper bound estimation of the positioning error. This estimation is used to raise an alert flag whenever a predefined threshold is surpassed.

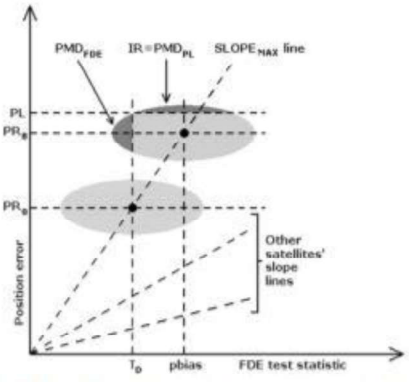


FIGURE 5. Classic RAIM scheme and the corresponding protection level computation [32].

Even though different definitions of PL can be found for the different applications, all of them are used for integrity purposes [20], [21]. The protection level is usually found broken down into its two components: the Vertical Protection Level (VPL) and the Horizontal Protection Level (HPL).

The PL is computed weighting different parameters according to the application or the scenario. Depending on the characteristics of the error source, the weight of this in the solution error may be different; for example, the PL can be decomposed into a noise component usually calculated according to the propagation error and a biased error part that can be computed in different ways according to the chosen model [12]:

$$PL = K(P_{md}) \cdot d_{major} + \max_i(SLOPE_i) \cdot p_{bias}. \quad (4)$$

The former term of the sum, which corresponds to the measurement noise, is computed according to the error propagation with external SBAS data aid [34]. K is an inflation parameter related to integrity risk, dependent on the probability of misdetection (P_{md}). d_{major} denotes the error uncertainty described by the semi-major axis of the error ellipse and it can be computed with the 2D elements of the variance matrix as

$$d_{major} = \sqrt{\sigma_e^2 + \sigma_n^2} \quad (5)$$

The $SLOPE_i$ parameter is a characteristic line, particular to each satellite, result of plotting the Horizontal Positioning Error (HPE) against a test statistic, as can be seen in FIGURE 5 [33].

Brown's method [36] or any of its variations are usually used to compute the $SLOPE_i$ parameter, which guarantees integrity by only accepting geometries that provide adequate redundancy to determine if there is an error on any channel of the receiver [37].

Variations of this method are found in the literature. Nevertheless, plenty of them show a similar structure, with the same slope-based method as a base [38].

Three different methods for the computation of the PL are discussed in [31]. These methods are the Brenner's method [34], which equals the PL to the largest error in the horizontal plane but does not identify the faulty satellite; Brown's method (previously mentioned) and Lee's method [39], which takes into account the fact that the missed detection probability depends on the bias error magnitude, and the maximum occurs before the bias error reaches the value that determines HPL in these methods [31]. Together with this, it also proposes a new method that contains the deterministic part defined by Brown's method's SLOPE and the stochastic part defined by noise. This method differs from (2) in the first term (the noise component) in the use of the integral of a statistic noise distribution instead of external SBAS data.

The computation of the PL can also be found to be based on the real-time processing of the SBAS broadcast data, as augmentation systems as EGNOS provide correction information for all pseudoranges. For this purpose, the RTCA standard differentiates two modes of operation; the non-precision approach and the precision approach. The author in [28] computes the PL as following:

$$PL = K_H \cdot d_{major} = 6.18 \cdot d_{major} \quad (6)$$

where the K_H is computed by the Rayleigh cumulative distribution function, assuming a non-precision approach as [40]

$$\begin{aligned} K_H &= RayleighCDF^{-1}(1 - P_{md}) \\ &= RayleighCDF^{-1}(1 - 5 \cdot 10^{-9}). \end{aligned} \quad (7)$$

a: ISOTROPY BASED PROTECTION LEVEL

This patented method is based on the hypothesis of the entropy on the measurement residuals, what is to say, that this vector can point in any direction equiprobably [30], [41]. This method, unlike others, does not assume any gaussianity on the measurement distributions [42]. According to this method, the PL is computed as follows [30], [41], [43]:

$$PL = k \cdot \|r\| \cdot XDOP. \quad (8)$$

The first parameter, k, is the isotropy confidence ratio (ICR), a parameter that ensures a bound of the error estimation by a confidence level (1-α) and the size of the residual vector, being α the integrity risk. r is the residual vector of the least square estimation. The later parameter, XDOP, is the employed type of dilution of precision (HDOP if HPL is computed and VDOP in the case of VPL).

Reference [42] shows a variant of this method as

$$PL = k \cdot \|r\| \cdot \sqrt{\lambda_{\max}^{(H^T H)^{-1}}}, \quad (9)$$

where the XDOP parameter is substituted by the largest eigenvalue of $H^T H$ (H is the observation matrix used in the linear observation equation).

The PLs mentioned previously were all based on GNSS measurements. Nevertheless, other sensor observations can

be added to the equation in order to obtain an estimation closer to reality. An (ITS) application focused PL is shown in [44]. This method adds IMU and vehicle odometer measurements to the RTK positioning and the corresponding integrity monitoring.

b: KALMAN INTEGRATED PROTECTION LEVEL

This patented method is an extension of the IBPL that allows both the use of GNSS-standalone and GNSS/INS navigation [45]. It estimates the error using a zero-mean multivariate Student distribution instead of a Gaussian one to improve the robustness against outliers [42]. It takes into account the temporal correlation of measurements, which, according to its author, is a requirement in order to compute a correct bound to the Kalman solution errors [46]. The results of this method, when applied to the case of a GNSS low-cost multi-constellation receiver, are also shown in this publication.

3) FAULT DETECTION AND EXCLUSION (FDE) METHODS

FDE methods are extensions of RAIM used not only to detect faulty satellites but also to exclude them from the navigation solutions so that the system's operation is not interrupted due to an incorrect PVT solution. The FDE scheme is divided into a global and a local test, where the global test (GT) is used to check if there is any fault (for which a minimum of five satellites is required) and the local test (LT) is used to identify it (for what six satellites are needed). A common approach to performing the global test is to use a test statistic, based on the Normalized Sum of Squared Error (NSSE), and check if this variable, multiplied by a variance factor and by the degrees of freedom (n-p), follows a centrally chi-squared distribution or not [47]. This test statistic is computed as

$$t = \tilde{r}^T \Sigma^{-1} \tilde{r}, \tag{10}$$

where r represents the range residual of the measurement and Σ represents the covariance matrix of the measurement errors.

In a failure-free situation, in other words, in case the global test passes the predefined threshold, the test statistic will follow a central chi-squared distribution (H_0) (FIGURE 6). This would mean that the solution would be computed without faulty satellites, obtaining, as a consequence, a reliable position estimation. On the other hand, if the global test is failed, which means that a failure has been detected, the test statistic will be non-centrally chi-squared distributed (H_a), with a non-centrality parameter called λ [47]. FIGURE 6 shows a statistical view of the GT and the behavior of the test statistic in both the faulty (H_a) and non-faulty (H_0) scenarios. α and β are the probabilities of detecting correctly and incorrectly the failure in the system, respectively. n is the number of satellites that are observed, and p is the number of parameters to be estimated.

There are multiple FDE schemes in which GT and LT are combined in different ways [12], [30]. In case of detecting a failure during the GT, the next FDE techniques can be carried out [12]:

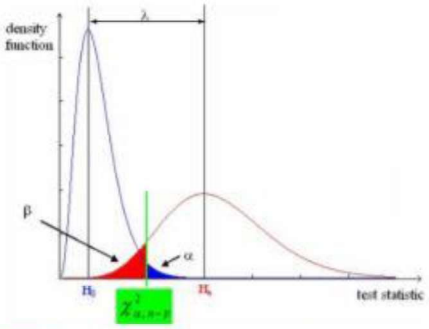


FIGURE 6. A central and a non-central chi-square distribution used for failure detection during the global test [47].

–**Subset testing (ST)** [13], [47], [48]: the GT is recomputed using subsets of more than four satellites, and the one with the most satellites and the smallest test statistic from the ones that pass the GT is selected.

–**Sequential local test (SLT)** [13], [47]: in each iteration of the LT, the one that further exceeds a given threshold with the biggest LT statistic will be eliminated.

–**Forward-backward test (FB)** [13], [48]: The forward loop performs the sequential local test while the backward one reintroduces the eliminated satellites to find the optimal combination from all the possibilities.

–**Danish method** [13], [48]: an iteratively reweighting is performed on the pre-estimated measurement variance based on the ration between an LT statistic and the local threshold. A point that should be emphasized is that there is no exclusion in this method and, if the algorithm cannot converge after iterating, the solution is considered unreliable.

Although all the mentioned FDE methods are based on the same GT – LT principle, each introduced variation has a different result. The following table summarizes the results shown in [13] and classifies these methods from 1 to 4 as a function of seven different aspects, being 1 the best and 4 the worst.

The fact that the results shown in TABLE 1 have been taken from a single measurement implies that some of the values can change between scenarios, as a single sample is not enough to make proper statistical statements. Nevertheless, it is still useful to have an early idea about the behavior of each of the methods and performance of each of them according to the application.

The mentioned methods may not be robust enough, as it is assumed that one outlier cannot be spread into the rest of the residuals. The author shows in [49] a possible solution to the mentioned issue by the Leave One Block Out approach. When possible, the observations are decomposed into uncorrelated blocks, so that the effect of a faulty observation is isolated.

TABLE 1. Comparison of the FDE methods.

		ST	SLT	F-B	Danish
Availability		1	4	2	3
Computation time		4	3	1	2
HPE	Min.	1	2	2	4
	Max.	1	4	2	3
	Mean	1	4	2	3
	Median	2	1	3	4
	Std. Dev.	1	4	2	3

Real-time applications, for example, could require lower computation times, as the detection of faults (and the consequent improvement in solution) and the recalculation of the solution should be done fast enough not to exceed the reaction time.

Life involving applications, on the other hand, could require high availability of the method, as a faulty measurement could imply an error that could lead to fatal consequences. Note that, according to [3], availability is "the percentage of time a signal fulfills the above accuracy, integrity and continuity criteria."

C. CLASSIFICATION OF NOVEL RAIM TECHNIQUES

Originally, the RAIM method was developed for airplane integrity monitoring. In this original scenario, GPS was just employed in an environment where the only signal failure could happen due to the malfunctioning of a satellite. Nowadays, multiple constellations are used, which implies a considerable amount of satellite combination. Moreover, as GNSS positioning and RAIM methods are commonly used on ground applications, where multiple error sources can be found, the RAIM methods have been improved to deal with the mentioned issues.

These techniques, which primary emphasis is the failure detection and exclusion, can be classified in multiple ways according to different parameters. When talking about the required amount of observations for the computation, one can distinguish two categories [27]:

–**A recursive scheme.** Typically a Kalman or a particle filter, which uses the history of the measurement data. Kalman filters are linear-quadratic estimators; thus, they are best for estimating linear systems with Gaussian noise [50]. They have much lower computational requirements than particle filters but are less flexible. Particle filters, on the other hand, can handle almost any kind of model, by discretizing the problem into individual "particles." This process may have as drawbacks high computational requirements.

–**A snapshot scheme.** A scheme in which the estimation of the position and the time of the receiver are based on the current measurements and satellite data. The main advantage of this scheme is that it allows an instantaneous position fix as it does not depend on more data than the corresponding to that same epoch [9]. Regarding its performance, nevertheless, it is usually outperformed by the recursive scheme. This scheme is usually used to calculate the initial position in the recursive scheme, based, often, on the Least Squares Estimation (LSE). E.g: LS background [48], Parkinson's LS residual method, Sturza's parity method [51], Brenner's parity method [52], Maximum residual method [48], solution separation method [53]–[55] etc.

The recursive scheme gives a more accurate position estimation than the snapshot one; therefore, the first is commonly used to detect rapidly growing measurement errors by monitoring the residuals. These residuals, also called innovations, are the differences between the current measurements and the predicted ones based on the history of the measurements. The recursive scheme, nevertheless, using innovations as test statistics, fails to catch slowly growing measurement errors called soft failures or ramp type because of attempting to adjust the solution to match faulty past measurements. Soft failures can be detected by the snapshot scheme. Two schemes should be used in parallel to achieve better fault detection.

No Classical RAIM implementation could fulfill the demanding vertical navigation requirements that were needed in aviation. Therefore, the second generation of RAIM techniques was developed that could ensure integrity in both lateral and vertical navigation. An example of this are the ARAIM and RRAIM methods that were presented as a possible solution [56]. According to GEAS [27], ARAIM with MHSS was decided as the principal architecture. RRAIM, on the other hand, was decided to be used whenever ARAIM was not available. In the following section, these two methods will be discussed, among other novel RAIM methods.

1) ADVANCED RAIM (ARAIM)

With the deployment of multiple new constellations, new RAIM methods have emerged. The Advanced Receiver Autonomous Integrity Monitoring (ARAIM) method expands the RAIM method to more constellations than GPS. This combination of constellations contributes to a better performance of horizontal guidance than RAIM based just on GPS [49], [57], [58].

The ARAIM has been developed from the solution separation, as this method turns out to be easier to modify for providing the improved competencies wanted for ARAIM [27], [59]. The solution separation method is based on the estimation of the position by means of computing it with all the available satellite measurements on the one hand and, on the other hand, computing the solution with all satellites except one. This solution separation method relies on three threshold test for each fault mode, one for each coordinate [55], [60]. ARAIM algorithms assume the possibility of having multiple-signal faults, not as in classical RAIM. The algorithm shown in [27] considers that the probability of multiple-signal fault threat is small, as it assumes that these are mitigated by other methods such as core-constellation design, ground monitoring, or separate airborne evaluation of broadcast data. In [38], on the other hand, the author assumes that the possibility of suffering multiple failures is

not negligible and, as a consequence, it discusses how should be the algorithm configured to detect them.

This approach takes advantage of the multi-frequency signals to compute an ionosphere-free combination to obtain higher accuracies. Moreover, it can use carrier-smoothed code measurements for both fault detection and positioning [61].

One of the main differences between ARAIM and RAIM is the capacity of ARAIM to adjust its operation to different requirements. Integrity parameters, fault probabilities, and even the probability of missed detection are fixed in RAIM, whereas in ARAIM, these parameters vary according to the requirements [62]. The authors describe in [62] the minimum operational performance standards.

The airborne algorithms for ARAIM are already mature enough for their use [63]. The full architecture for its deployment, on the other hand, is still being developed and standardized to support a correct operation [57]. Consequently, developing this architecture will be the current and future work in this area for the use of this technology.

2) RELATIVE RAIM (RRAIM)

The relative RAIM (RRAIM) is a technique that was developed to handle the GPS data latency problem [56], [64].

It is based on the propagation of older pseudoranges forward in time by using precise carrier phase measurements [61]. For this purpose, integrity measurements are performed periodically, when available, and during these intervals (called coasting time), RRAIM takes the difference between the accumulated carrier phase and the original value to estimate the new positions [64]. The accumulated uncertainty during this interval can be split into three main sources; the change of noise and multipath levels, the change in the tropospheric error, and the satellite clock drift [56].

Two main variants of the RRAIM algorithms are found in literature: a Range Domain RRAIM that is based on a Chi-square (χ^2) RAIM method (introduced in [56]), and a Position Domain RRAIM based on a solution separation RAIM method (introduced in [64]).

3) EXTENDED RAIM (ERAIM)

Due to the fact that satellite navigation depends on radiofrequency signals, degradation of these could lead to a faulty PVT solution or even to no solution scenario. As a consequence, multiple transportation systems in the ITS world have started to integrate multiple sensors (such as INS or odometers) to complete and improve the performance of GNSS only systems.

Thus, the purpose of adapting RAIM to GNSS/INS integrated systems gave rise to the creation of ERAIM. This method is based on the least-squares theory, which is used to find the best estimators of the state parameters in a Kalman Filter [55], [56]. Once having characterized the filter, and based on the new measurement, integrity monitoring is performed, including outlier detection and identification, reliability, and separability [55], [56].

The term reliability is used to quantify the minimum detectable bias that stipulates, with a high confidence level, the lower bound for detectable outliers [66].

Separability, on the other hand, is the capacity of separating measurements, in the case of a faulty one, so that good measurements are not incorrectly understood as a fault [65].

References [65] and [66] show empirical results of the method; the first one does, nevertheless, a comparison for ERAIM vs. RAIM and its corresponding analysis.

A different meaning is given to ERAIM in [27], where ERAIM means RRAIM-Extended ARAIM. The main goal of this method is to use RRAIM's coasting during ARAIM's unavailability intervals. Together with improving availability, this method leads to tighter detection thresholds as the carrier-phase is added via RRAIM. The tighter detection results in smaller protection levels than using just ARAIM.

4) CARRIER BASED RAIM (CRAIM)

Even if most RAIM methods are based on code measurements, other architectures exist, such as the Carrier-Phase based RAIM (CRAIM) [46], [47]. This method's main characteristic is the use of the carrier-phase, which is much more precise than code measurements, as it is more robust against noise. Nevertheless, it is not always available as an absolute measurement without external aid, due to its ambiguities, especially difficult to compute in harsh environments.

The CRAIM method proposed in [67], which is a carrier-phase based algorithm for high-accuracy positioning based on Kalman filtering, only allows failure detection and does not provide failure identification. This is why [68] introduces a new variation of this method that permits both failure detection and failure identification using extended w-test detectors. These detectors are based on the use of two test-statistics for the EKF, which are respectively presented for the code (c) and carrier phase (p) double-difference (DD). These statistics can be deduced as follows:

$$w_{ij} = \left| \frac{e_j^T R_i^{-1} r_{ik}}{e_j^T R_i^{-1} W_{ik} R_i^{-1} e_j} \right|, \quad i = c, p \tag{11}$$

where the unit vector represents the use of a certain measurement or, in other words, if $e_i = 1$, the i_{th} measurement is used. R_i represents the measurement noise covariance and W_{ik} is the reduction of W_k, a parameter that represents the variance-covariance of the innovation r_{ik}. W_k is a weighting matrix that considers the covariance and measurement noise of the measurement residual.

In a faulty-free case, the test statistic follows a standard Gaussian distribution. In a faulty case, on the other hand, it follows a non-central Gaussian distribution [69]. These statistics can be used to define a threshold using the P_{FA} in order to detect a fault in both the carrier phase and the code.

This innovation, which makes this method efficient for multi-failure scenarios, improves integrity and reliability.

Both papers [67], [68] provide an extensive analysis of their algorithms and the behavior of the employed

TABLE 2. RAIM technique classification and characteristics.

RAIM type	References	Measurement	Algorithm	FDE type / Tolerated faults	External input	Navigation	Constellations	Frequencies
Classical RAIM	[32][33][39] [45][46]	Code	LS / WLS methods	FDE / Single fault	No	LNAV	GPS	1
Advanced RAIM	[47][48][49] [38][40][42]	Code	Solution separation (SS) or Multiple hypothesis Solution Separation (MHSS)	Real time FDE / Multiple faults	Integrity data from Integrity Support Messages	LPV-200	Multiple	Multiple
Relative RAIM	[50][47] [48][49]	Carrier	Chi-square method or SS method	FDE / Multiple faults	External monitors	LPV-200	GPS	-
Extended RAIM	[18][51][52]	Code	Least square initialized EKF	FDE / Multiple faults	Multiple sensors	-	Multiple	-
Carrier based RAIM	[46][47] [53][54]	Carrier	Ambiguity Resolution algorithms (LAMBDA) Extended Kalman Filter (EKF)	FD (no exclusion) / Multiple faults	No	LNAV	Multiple	Multiple
Time RAIM	[55][56][57]	Code and doppler	WLS and variations	Forward backward FDE / -	No	-	Multiple	-
Vision-Aided RAIM	[58]	Code	Real Time Kinematics (RTK)	Fault detection / single fault assumption in [58] but multiple faults could be detected	Vision system provided landmarks and RTK required corrections	LPV-200	GPS	-

test statistics. Furthermore, the author shows in [67], an example result of the obtained protection level during some measurements.

5) TIME RAIM (TRAIM)

The RAIM concept is usually related to positioning; however, it can also be applied to time matters. The addition of a new level of system reliability to timing applications is proposed in [70]. The author presents an algorithm that detects and removes the satellites that exceed a previously defined time residual threshold. It is also mentioned that the algorithm allows the end-user to predict the overall system time accuracy and that the algorithm is able to predict the time error performance of the receiver both in faulty and fault-free scenarios.

A potential T-RAIM approach for a multi-constellation scenario is discussed in [71]. According to the author, the benefits of multi-constellation T-RAIM are shown in hard environments.

Experimental proof of the performance of T-RAIM is given in [72].

6) VISION-AIDED RAIM (VA-RAIM)

It is mentioned in [73] that the performance of existing RAIM methods could not be acceptable during the landing phase of a flight as a consequence of the lack of observations. This is why it proposes a new RAIM method called Vision-Aided RAIM (VA-RAIM), which employs computer vision systems to match landmarks with photographs in order to obtain additional measurements. This method improves availability, as it introduces the landmarks as pseudo-satellites so that the vision system can model the landmark-receiver distance in an analogous way as done in the GPS. These vision measurements are used to expand the GPS measurement equations in order to improve integrity. This is a concept that is easily transferable to other transport means.

This method assumes that the test statistic defined by the NSSE follows a chi-squared distribution. As shown in

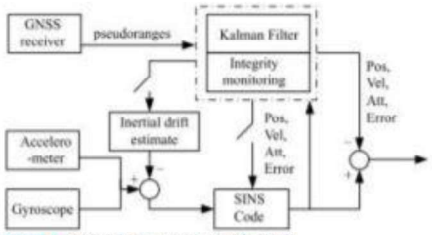


FIGURE 7. GNSS/INS integrated system block [65].

Classical RAIM, this test-statistic is compared to a user-defined threshold T_{NSSE} that is dependent on P_{FA}; so that whenever it exceeds the threshold, a fault will be detected.

The author states that simulation results show that the proposed method outperforms classical RAIM both in terms of fault detection rate and in terms of availability.

III. CONCLUSION

Positioning accuracy is a vital requirement in multiple subjects inside society. As a consequence, so it is the failure detection and the computation of an accurate upper bound of the positioning error.

GNSS techniques have been developed up to a point in which the classical RAIM scheme may not be able to fulfill its duty. As a consequence, research is being carried out on new RAIM schemes that cover the aspects that classical RAIM does not consider, such as multiple faults, multiple frequencies, or augmentation systems.

This paper shows an up to date introduction to GNSS positioning integrity and the respective concept definitions. Together with these explanations, multiple error sources that could cause an incorrect position have been discussed and classified.

Moreover, a statistical approach to the main fault detection and exclusion techniques has been shown in Section B. Besides this summary, a classification (TABLE 2) and an

explanation of the principal among these methods found in the literature are done. Finally, as it is a relevant measurement for navigation (together with the GNSS PVT solution), different error upper bounds or protection levels have been shown.

To sum up, this survey gathers and discusses the main receiver autonomous integrity monitoring methods found in the literature. This initial study can be understood as a state of the art chapter on the matter and may be useful for creating a knowledge base about this topic. Future work can be done both in the development of each of the presented techniques and in the empirical comparison of these. This future research should especially focus on the computation of the protection level, as current models use to overbound the error in a loosely way, not being able to properly adapt to the error. In an ideal case, the PL's curve would be tangent to the error at every moment, tightly bounding the error.

ACRONYMS

AL	Alert Limit
C/N0	Carrier to Noise Ratio
DOP	Dilution of Precision
FA	False Alarm
FB	Forward Backward
FDE	Fault Detection and Exclusion
GBAS	Ground Based Augmentation System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GT	Global Test
HPE	Horizontal Positioning Error
HPL	Horizontal Protection Level
HW	Hardware
IBPL	Isotropy Based Protection Level
ICR	Isotropy Confidence Radio
IM	Integrity Monitoring
IMU	Inertia Measuring Unit
INS	Inertial Navigation System
IR	Integrity Risk
ITS	Intelligent Transport System
KIPL	Kalman Integrated Protection Level
LOS	Line of Sight
LT	Local Test
MD	Misdetection
NLOS	Non Line of Sight
NSSE	Normalized Sum of Squared Errors
PE	Positioning Error
PF	Positioning Failure
PL	Protection Level
PVT	Position, Velocity and Time
RAIM	Receiver Autonomous Integrity Monitoring
RTCA	Radio Technical Commission for Aeronautics
RTK	Real Time Kinematics
SBAS	Satellite Based Augmentation System
SIS	Signal in Space
SLT	Sequential Local Testing
ST	Subset Testing
SW	Software
TTA	Time to Alert
UTM	Urban Trench Model
VPL	Vertical Protection Level

REFERENCES

[1] M. Kaloop, E. Elbeltagi, J. Hu, and A. Elrefai, "Recent advances of structures monitoring and evaluation using GPS-time series monitoring systems: A review," *ISPRS Int. J. Geo-Inf.*, vol. 6, no. 12, p. 382, Nov. 2017.
[2] J. Liu, B. Cai, D. Lu, and J. Wang, "Integrity of GNSS-based train positioning: From GNSS to sensor integration," in *Proc. Eur. Navigat. Conf. (ENC)*, May 2017, pp. 48–56.
[3] *What is GNSS? | European Global Navigation Satellite Systems Agency*. Accessed: Oct. 23, 2019. [Online]. Available: https://www.gsa.europa.eu/european-gnss/what-gnss
[4] *Federal Radionavigation Plan 2017*, Dept. Defense, Dept. Homeland Secur., Dept. Transp., Arlington, VA, USA, 2017.
[5] *2008 Federal Radionavigation Plan*, Dept. Defense, Dept. Homeland Secur., Dept. Transp., Arlington, VA, USA, 2008, pp. 1–184.
[6] *International Standards and Recommended Practices, Annex 10 to the Convention on International Civil Aviation, Volume 1, Radio Navigation Aids*, I-ICAO, 2006.
[7] *NAVIPEDIA*. Accessed: Apr. 25, 2019. [Online]. Available: https://gssc.esa.int/navipedia/index.php/Main_Page
[8] J. J. Spilker, Jr., P. Axelrad, B. W. Parkinson, and P. Enge, Eds., *Global Positioning System: Theory and Applications*, vol. I. Washington, DC, USA: American Institute of Aeronautics and Astronautics, 1996.
[9] A. Wieser and F. K. Brunner, "An extended weight model for GPS phase observations," Earth Planets Space, vol. 52, pp. 777–782, Jul. 2000.
[10] H. Hartinger and F. K. Brunner, "Variances of GPS phase observations: The SIGMA-ε model," *GPS Solutions*, vol. 2, no. 4, pp. 35–43, Apr. 1999.
[11] F. K. Brunner, H. Hartinger, and L. Troyer, "GPS signal diffraction modelling: The stochastic SIGMA-Δ model," *J. Geodesy*, vol. 73, no. 5, pp. 259–267, Jun. 1999.
[12] N. Zhu, D. Bétaille, J. Marais, and M. Berbineau, "GNSS integrity enhancement for urban transport applications by error characterization and fault detection and exclusion (FDE)," Géolo-calisation et Navigation dans l'Espace et le Temps, Journées Scientifiques 2018 de l'URSI, Paris, France, Mar. 2018, p. 11.
[13] N. Zhu, J. Marais, D. Bétaille, and M. Berbineau, "Evaluation and comparison of GNSS navigation algorithms including FDE for urban transport applications," in *Proc. Int. Tech. Meeting Inst. Navigat.*, Mar. 2017, p. 19.
[14] A. Gallay, "Pour une ethnoarchéologie théorique: Mérites et limites de l'analogie ethnographique," *Collect. Des Hespérides*, vol. 2, no. 3, p. 388, 2011.
[15] S. Kim, S. Yoon, and S. Y. Kim, "A novel multipath mitigated side-peak cancellation scheme for BOC(kn, n) in GNSS," in *Proc. 9th Int. Conf. Adv. Commun. Technol.*, vol. 2, Feb. 2007, pp. 1258–1262.
[16] D. Betaille, F. Peyret, M. Ortiz, S. Miquel, and L. Fontenay, "A new modeling based on urban trenches to improve GNSS positioning quality of service in cities," *IEEE Intell. Transp. Syst. Mag.*, vol. 5, no. 3, pp. 59–70, 2013.
[17] M. Z. Bhuiyan, X. Hu, E. S. Lohan, and M. Renfors, "Multipath mitigation performance of multi-correlator based code tracking algorithms in closed and open loop model," in *Proc. Eur. Wireless Conf.*, May 2009, pp. 84–89.
[18] D. Egea, G. Seco-Granados, and J. A. Lopez-Salcedo, "Single- and multi-correlator sequential tests for signal integrity in multi-antenna GNSS receivers," in *Proc. Int. Conf. Localization GNSS (ICL-GNSS)*, Jun. 2014, pp. 1–6.
[19] M. Ibraheam, B. Bieske, K. Blau, E. Schafer, A. Jager, S. I. Butt, R. Stephan, and M. A. Hein, "Feasibility of dual-polarized antenna arrays for GNSS receivers at low elevations," in *Proc. 11th Eur. Conf. Antennas Propag. (EUCAP)*, vol. 2017, no. 978, Mar. 2017, pp. 857–861.
[20] P. D. Groves, "Novel multipath mitigation methods using a dual-polarization antenna," in *Proc. 23rd Int. Tech. Meet. Satell. Div. Inst. Navigat.*, pp. 140–151, Sep. 2010.
[21] Z. Jiang and P. D. Groves, "NLOS GPS signal detection using a dual-polarisation antenna," *GPS Solutions*, vol. 18, no. 1, pp. 15–26, 2014.
[22] N. Vagle, A. Broumandan, A. Jafarnia-Jahromi, and G. Lachapelle, "Performance analysis of GNSS multipath mitigation using antenna arrays," *J. Global Positioning Syst.*, vol. 14, no. 1, p. 4, Dec. 2016.

[23] S. Daneshmand, A. Broumandan, N. Sokhandan, and G. Lachapelle, "GNSS multipath mitigation with a moving antenna array," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 1, pp. 693–698, Jan. 2013.

[24] J. K. Ray, M. E. Cannon, and P. Fenton, "GPS code and carrier multipath mitigation using a multiantenna system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 1, pp. 183–195, 2001.

[25] C. Rost and L. Wanninger, "Carrier phase multipath corrections based on GNSS signal quality measurements to improve CORS observations," in *Proc. IEEE/ION Position, Location Navigat. Symp.*, no. 1, May 2010, pp. 1162–1167.

[26] C. Enneking and F. Antreich, "Exploiting WSSUS multipath for GNSS ranging," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 7663–7676, Sep. 2017.

[27] "Phase II of the GNSS evolutionary architecture study phase II of the GNSS evolutionary," U.S. Government, G. P. Office, Tech. Rep., Feb. 2010.

[28] J. Santa, B. Ubeda, R. Toledo, and A. G. Skarmeta, "Monitoring the position integrity in road transport localization based services," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2006, pp. 2801–2805.

[29] X. Li, Y. Qiao, and T. Wang, "Analysis on the integrity simulation of dual-frequency WAAS," in *Proc. IEEE 2nd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Mar. 2017, pp. 2078–2082.

[30] N. Zhu, J. Marais, D. Betaille, and M. Berbineau, "GNSS position integrity in urban environments: A review of literature," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 9, pp. 2762–2778, Sep. 2018.

[31] V. Perepetchai, "Global positioning system receiver autonomous integrity monitoring," School Comput. Sci., McGill Univ., Nat. Library Canada, Montreal, QC, Canada, 2000.

[32] *GMV | GNSS Based Air Navigation Systems*. Accessed: May 10, 2019. [Online]. Available: https://www.gmv.com/en/Sectors/aeronautics/GNSSnavigation/

[33] E. D. Kaplan, "Understanding GPS: Principles and applications," in *Global Positioning System I*, C. J. Hegarty, Ed., 2nd ed., 2006.

[34] *Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment*, document RTCA DO-229, 2008.

[35] *Aeronautical Telecommunications Annex 10 to the Convention on International Civil Aviation International Civil Aviation Organization International Standards and Recommended Practices and Procedures for Air Navigation Services Volume II*, Communication Procedures Including Those With PANS Status, Montreal, QC, Canada, 2001.

[36] R. G. Brown, *GPS RAIM: Calculation of Thresholds and Protection Radius Using Chi-Square Methods; a Geometric Approach*. Washington, DC, USA: Radio Technical Commission for Aeronautics, 1994.

[37] T. Walter and P. Enge, "Weighted RAIM for precision approach," in *Proc. 8th Int. Tech. Meeting Satell. Div. Inst. Navig. (ION GPS)*, vol. 8, 1995, pp. 1995–2004.

[38] J. Blanch, T. Walter, and P. Enge, "RAIM with optimal integrity and continuity allocations under multiple failures," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 3, pp. 1235–1247, Jul. 2010.

[39] Y. Lee, "New techniques relating fault detection and exclusion performance to GPS primary means integrity requirements," in *Proc. 8th Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GPS)*, Sep. 1995, pp. 1929–1939.

[40] H. Yu and J. Wang, "A new method to compute horizontal protection level based on vertical projection," in *Proc. IEEE Adv. Inf. Manage., Communicates, Electron. Autom. Control Conf. (IMCEC)*, Oct. 2016, pp. 901–905.

[41] M. A. Sáenz, "Method for autonomous determination of protection levels for GNSS positioning based on navigation residuals and an isotropic confidence ratio," GMV Aerosp., Defende SA, Madrid, Spain, Tech. Rep., 2010.

[42] A. Welte, S. P. Xu, and P. Bonnifait, "Protection levels for high integrity localization for autonomous driving," M.S. thesis, ENSTA Bretagne, Univ. Angers, Angers, France, 2017.

[43] E. Shytermeja, A. Garcia-Pena, and O. Julien, "Proposed architecture for integrity monitoring of a GNSS/MEMS system with a fisheye camera in urban environment," in *Proc. Int. Conf. Localization GNSS (ICL-GNSS)*, Jun. 2014, pp. 1–6.

[44] A. El-Mowafy and N. Kubo, "Integrity monitoring for positioning of intelligent transport systems using integrated RTK-GNSS, IMU and vehicle odometer," *IET Intell. Transp. Syst.*, vol. 12, no. 8, pp. 901–908, Oct. 2018.

[45] P. F. N. Madrid, "Method for computing an error bound of a Kalman Filter based GNSS position solution," GMV Aerosp., Defende SA, Madrid, Spain, Tech. Rep., 2016.

[46] P. F. N. Madrid, M. D. L. Samper, and M. M. R. Merino, "New approach for integrity bounds computation applied to advanced precise positioning applications," in *Proc. 28th Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GNSS)*, 2015, pp. 2821–2834.

[47] H. Kuusniemi and G. Lachapelle, "GNSS signal reliability testing in urban and indoor environments," in *Proc. ION NTM*, Aug. 2004, pp. 1–15.

[48] H. Kuusniemi, "User-level reliability and quality monitoring in satellite-based personal navigation," Tech. Rep., Jun. 2005, p. 220.
H. Kuusniemi, "User-level reliability and quality monitoring in satellite-based personal navigation," Inst. Digit. Comput. Syst. Tampere Univ. Technol., Tampere, Finland, Tech. Rep., Jun. 2005.

[49] L. Biagi and S. Caldera, "An efficient leave one block out approach to identify outliers," *J. Appl. Geodesy*, vol. 7, no. 1, pp. 11–19, Jan. 2013.

[50] F. Gustafsson, "Particle filter theory and practice with positioning applications," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 25, no. 7, pp. 53–82, Jul. 2010.

[51] M. A. Sturza, "Navigation system integrity monitoring using redundant measurements," *Navigation*, vol. 35, no. 4, pp. 483–501, 1988.

[52] M. Brenner, "Integrated GPS/inertial fault detection availability," *Navigation*, vol. 43, no. 2, pp. 111–130, 1996.

[53] M. Joerger and B. Pervan, "Fault detection and exclusion using solution separation and chi-squared ARAIM," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 2, pp. 726–742, Apr. 2016.

[54] M. Joerger and B. Pervan, "Solution separation and chi-squared ARAIM for fault detection and exclusion," in *Proc. IEEE/ION Position, Location Navigat. Symp. (PLANS)*, May 2014, pp. 294–307.

[55] J. Blanch, T. Walter, Y. Lee, B. Pervan, M. Rippl, and A. Spletter, "Advanced RAIM user algorithm description: Integrity support message processing, fault detection, exclusion, and protection level calculation," in *Proc. 25th Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GNSS)*, Sep. 2012, pp. 2828–2849.

[56] *GNSS Evolutionary Architecture Study (GEAS), Phase I*, FAA GNSS Group, USA, 2008.

[57] T. Walter, "An overview of advanced receiver autonomous integrity monitoring (ARAIM)," in *Proc. ION Pacific PNT Meeting*, May 2019, pp. 896–914.

[58] Y. Zhai, X. Zhan, J. Chang, and B. Pervan, "ARAIM with more than two constellations," in *Proc. ION Pacific PNT Meeting*, May 2019, pp. 925–941.

[59] J. Blanch, T. Walter, and P. Enge, "A formula for solution separation without subset solutions for advanced RAIM," in *Proc. IEEE/ION Position, Location Navigat. Symp. (PLANS)*, Apr. 2018, pp. 316–326.

[60] V. Kropp and G. Berz, "Optimized MHSS ARAIM user algorithms: Assumptions, protection level calculation and availability analysis," in *Proc. IEEE/ION Position, Location Navigat. Symp. (PLANS)*, May 2014, pp. 308–323.

[61] L. Gratton, M. Joerger, and B. Pervan, "Carrier phase relative RAIM algorithms and protection level derivation," *J. Navigat.*, vol. 63, no. 2, pp. 215–231, Feb. 2010.

[62] J. Blanch, T. Walter, G. Berz, J. Burns, B. Clark, M. Joerger, M. Mabilleau, I. Martini, C. Milner, B. Pervan, and Y. Lee, "Development of advanced RAIM minimum operational performance standards," in *Proc. 32nd Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GNSS)*, Oct. 2019, pp. 1381–1391.

[63] T. Walter, J. Blanch, and P. Enge, "A framework for analyzing architectures that support ARAIM," in *Proc. 25th Int. Tech. Meet. Satell. Div. Inst. Navig. (ION GNSS)*, vol. 4, 2012, pp. 2850–2857.

[64] Y. C. Lee, "A position domain relative RAIM method," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 85–97, Jan. 2011.

[65] H. Liu, W. Ye, and H. Wang, "Integrity monitoring using ERAIM for GNSS/inertial system," *Aircr. Eng. Aerosp. Technol.*, vol. 84, no. 5, pp. 287–292, Aug. 2012.

[66] S. Hewitson and J. Wang, "eRAIM . . . for GNSS/INS integration," *J. Surv. Eng.*, pp. 13–22, Feb. 2010.

[67] S. Feng, W. Ochieng, T. Moore, C. Hill, and C. Hide, "Carrier phase-based integrity monitoring for high-accuracy positioning," *GPS Solutions*, vol. 13, no. 1, pp. 13–22, 2009.

[68] H. Liu, X. Meng, Z. Chen, S. Stephenson, and P. Peltola, "A closed-loop EKF and multi-failure diagnosis approach for cooperative GNSS positioning," *GPS Solutions*, vol. 20, no. 4, pp. 795–805, Sep. 2015.

[69] H. Yun, Y. Yun, and C. Kee, "Carrier phase-based RAIM algorithm using a Gaussian sum filter," *J. Navigat.*, vol. 64, no. 1, pp. 75–90, 2011.

[70] G. J. Geier, T. M. King, H. L. Kennedy, R. D. Thomas, and B. R. Mcnamara, "Prediction of the time accuracy and integrity of Gps timing," in *Proc. IEEE Int. Freq. Control Symp. (49th Annu. Symp.)*, May/Jun. 1995, pp. 266–274.
[71] C. Gioia and D. Borio, "Multi-constellation T-RAIM: An experimental evaluation," in *Proc. 30th Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GNSS)*, Nov. 2017, pp. 4248–4256.
[72] P. Vyskocil and J. Sebesta, "Relative timing characteristics of GPS timing modules for time synchronization application," in *Proc. Int. Workshop Satell. Space Commun.*, Sep. 2009, pp. 230–234.
[73] L. Fu, J. Zhang, R. Li, X. Cao, and J. Wang, "Vision-aided RAIM: A new method for GPS integrity monitoring in approach and landing phase," *Sensors*, vol. 15, no. 9, pp. 22854–22873, Sep. 2015.
[74] Novatel. *Novatel Correct with SBAS and DGPS*. Accessed: Jun. 9, 2019. [Online]. Available: https://www.novatel.com/products/novatel-correct-with-sbas-and-dgps/
[75] *GPS World*. Accessed: Jun. 10, 2019. [Online]. Available: https://www.gpsworld.com/innovationground-based-augmentation/

PAUL ZABALEGUI received the M.Sc. degree in telecommunications engineering from the School of Engineering of San Sebastián (TECNUN), Spain, in 2019. He is currently pursuing the Ph.D. degree with the Transport and Sustainable Mobility Group. He joined the CEIT-IK4 Research Centre, San Sebastián, in 2019. He is currently a Researcher Assistant with the Transport and Sustainable Mobility Group, Universidad de Navarra. He is now actively participating in H2020 European funded projects in railway signaling and positioning topics. His research activity includes the field of positioning and software development.

GORKA DE MIGUEL received the M.Sc. degree in telecommunications engineering from the School of Engineering of San Sebastián (TECNUN), Spain, in 2015. He is currently pursuing the Ph.D. degree with the Transport and Sustainable Mobility Group. He joined the CEIT-IK4 Research Centre, San Sebastián, in 2015. He is currently a Researcher Assistant with the Transport and Sustainable Mobility Group. He is also a Lecturer Assistant in electronic fabrication systems with TECNUN, University of Navarra. He is now actively participating in H2020 European funded projects in Railway signaling and positioning topics. His research activity includes the field of positioning and software development.

ALEJANDRO PÉREZ received the degree in audiovisual systems of telecommunications engineering from the Universidad Europea de Madrid (UEM), Spain, in 2014. He joined the CEIT-IK4 Research Centre, San Sebastián, in 2018. He is currently a Transfer Engineer with the Electronic Systems and Communications Group, Universidad de Navarra. He is also working in European Project H2020 AIOSAT for GSA. His research activities include the field of positioning, electronic, communications, and software development.

JAIZKI MENDIZABAL (Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from TECNUN, in 2000 and 2006, respectively, and the Ph.D. degree RF design for GNSS systems. He joined Fraunhofer IIS-A, Germany, from 2000 to 2002, and SANYO Electric Ltd., Japan, from 2005 to 2006 as an RFIC Designer. He is currently a Researcher with CEIT and also a Lecturer with TECNUN, University of Navarra. He has been engaged in the SIL4 design of a safety-critical receiver for an ERTMS BTM for high-speed trains (8M€project with eight partners), he was the Project Coordinator of the FP7 EATS project dealing with the introduction of GNSS to onboard ETCS. He is currently working in several Shift2Rail projects apart from being member of the board. He has participated in more than 20 research projects, where six of them dealt with GNSS for railways. He has directed five Ph.D. theses and more than 20 master's and graduate theses. He has authored or coauthored two patents and 35 scientific and technical publications and has authored one book and two book chapters. His current research interests at CEIT include GNSS, wireless communication, and safety-critical systems for the railway and maritime industry.

JON GOYA received the M.Sc. degree in telecommunications engineering and the Ph.D. degree from the Universidad de Navarra, in 2011 and 2016, respectively. He is currently a Lecturer with TECNUN, Universidad de Navarra, and also a Researcher with CEIT-IK4. His professional research activities include the simulation of onboard positioning systems and performance analysis for railway. He has participated in FP7 projects coordinated by CEIT and is now actively participating in H2020 European funded, such as AIOSAT, and Shift2Rail projects in railway signaling and positioning topics: X2RAIL-2 and FR8RAIL.

IÑIGO ADIN (Member, IEEE) received the M.Sc. degree in electronics engineering and the Ph.D. degree from the University of Navarra, in 2003 and 2007, respectively. He is currently a Lecturer with TECNUN, University of Navarra, and also a Researcher at CEIT-IK4. He has authored or coauthored one patent, two technical books, an invited chapter, and 40 articles in journals and international conferences. He was the Coordinator of the FP7 European Project TREND and now coordinates the H2020 AIOSAT for GSA. His research interests include safety-critical designs, with a special interest in positioning, communications, electromagnetic compatibility, and transport interoperability.

An Overview of Current IP Network Emulators for the Validation of Railways Wireless Communications

Nerea Fernández-Berrueta, Jon Goya, Javier Añorga, Saioa Arrizabalaga, Gorka De Miguel and Jaizki Mendizabal

IEEE Access (Q1). June2020

IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Dot Number

An Overview of Current IP Network Emulators for the Validation of Railways Wireless Communications

Nerea Fernández-Berrueta[1,2], Jon Goya[1,2], Javier Añorga[1,2], Saioa Arrizabalaga[1,2], Gorka de Miguel[1,2], Jaizki Mendizabal[1,2]

[1]Ceit -Basque Research and Technology Alliance (BRTA), Manuel Lardizabal 15, 20018 Donostia / San Sebastián, Spain

[2]Universidad de Navarra, Tecnun, Manuel Lardizabal 13, 20018 Donostia / San Sebastián, Spain

Corresponding author: N. Fernández-Berrueta (e-mail: nfernandez@ceit.es).

ABSTRACT Communication technologies are in continuous evolution and as well, the different applications making use of them. In order to succeed with the roll-out of the communication-based applications, it is required that the communications technologies are intensively tested and validated before deployment. Current strategies for testing and validation cover field tests and laboratory tests. Railways is also taking advantage of the communication technologies evolution, and therefore, there is a need for having testing and validation strategies adapted to the railway environment, especially for safety-critical applications. Field tests and laboratory tests also apply in Railways. In the frame of laboratory tests, this paper includes an overview of different network emulators existing currently in the market. Furthermore, an analysis of the gaps of the network emulators with regards to the needs of the railways environment is also included. The goal of this paper is to show that network emulators are a flexible cost-effective solution for communication technologies testing purposes. Additionally, this paper also shows that there is a need to adapt current emulators to the railway environment in order to test and validate the future railway applications based on communication technologies.

INDEX TERMS IP Communication, Laboratory testing, Network Emulator, Railways, Wireless Communications

I. INTRODUCTION

Nowadays, communication technologies are in continuous evolution, offering a higher performance that allows getting further services. In order to adopt these new services, the communication technology in use has to be tested to ensure the required performance. However, due to the time constraints and cost of infrastructure deployment, the real performance of new communication technologies are not simple to be tested.

Railways is also aiming to move forwards in this field of the communication technologies to make railways more efficient, safer and profitable [1], even if the main stakeholders interested in the deployment of these new communication technologies are mobile operators. Future railway services are pushing restrictive requirements such as large bandwidth (high data rate transmission capabilities) [2] and low latency [3]. To research these type of requirements (among others) in the railway communication field, some European projects were built, such as Shift2Rail [4] and FRMCS [5]. Shift2Rail project covers the future adaptable communications for railways, but also several different topics such as train integrity, interconnections between the different components inside the train car apart from the communication technologies. FRMCS project is also analyzing new mobile communications capabilities for railways.

With regard to validation, particularly, Shift2Rail includes in its innovation program, a research activity related to the development of a new laboratory test framework, which comprises simulation tools and testing procedures for carrying out open tests with clear operational rules and simple certification of test results. It aims to minimize on-site testing (with the objective of zero on-site testing) by performing full laboratory test processes, even when systems comprise subcomponents of different suppliers. This effort is carried out with the aim also to reduce costs and time

compared to on-site testing as well as increasing the number and type of tests that can be done. Focusing on the unfeasible in terms of cost and time. In order to overcome this issue, the number of tools and strategies developed for testing the communication in the laboratory is increasing. These different strategies for testing in the laboratory are simulation, Hardware (HW) in the loop, and emulation (explained in section III).

It is important to find a laboratory validation tool in order to search for limitations and improvements of a new technology, approaching the closest to the on-site testing in a controlled laboratory environment. Nevertheless, the performance of a communication network does not depend exclusively on the technology but also on the different impairments that the environment causes to, particularly, in wireless networks, for example, the multipath effect [6].

Especially railways do not present a clean environment. In fact, railways present the same problems as other areas (e.g., multipath effect) but additionally, some effects arise just only in this area, e.g., the effect in the communications of transient electromagnetic (EM) disturbances produced by the sliding contact between catenary and pantograph [7][8].

In this paper, different railway applications and services that will be enabled with the new communication technologies are explained (section II) and, the different test strategies that could be applied to test them are listed (section III). Moreover, current network emulators are presented (section IV) to, later on, compare their main features and characteristics (section V). Finally, conclusions are drawn (section VI).

II. RAILWAY APPLICATIONS AND SERVICES

Railway applications and services based on communication technologies are mainly focused on the safety-critical operation or passenger services related to the comfort. Due to these different goals, the applications making use of communication technologies have different set of requirements. The different railway applications could be classified into two types: critical related to safety and non-critical related to non-safety and passengers comfort.

On the one hand, the ones caring about the safety-critical operations are more restrictive than the ones related to the passengers comfort dealing with high availability, low delay, and low packet loss [9]. This type of applications includes signaling systems such as ERTMS (European Rail Traffic Management System) or CBTC (Communications-Based Train Control). Any of these systems could send messages containing, e.g., information for an emergency brake, being a critical message. Safety-critical applications use communications to interact between the train and the control center at trackside. Taking into account that this kind of operation could put passenger's life at risk, communication technologies have to be intensively tested to characterize the performance in different scenarios. These scenarios could be related to different communication technologies, channel communication technologies, testing all candidate technologies for an application makes the on-site testing access modes, cyber-security attacks, or different environment and impairments conditions, which lead to a huge number of combinations.

On the other hand, non-critical applications as tele-maintenance or passenger connectivity do not request such critical restrictions; if the data arrives with a higher delay time than the critical ones (e.g., 4s) it does not have a critical impact.

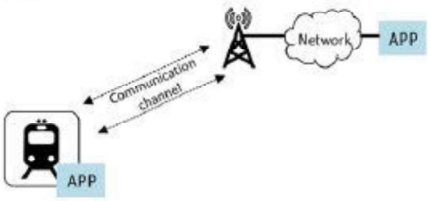


FIGURE 1. Communication scheme: train to ground and ground to train

In the ideal case, the communication between train and ground is continuously working and never disrupted. However, in the real world, communication can be affected by the environment the train is operating, namely a forest, urban canyon, or conflict areas in terms of perturbations. All these areas have a negative effect on the communication channel, reducing the performance of the communication due to multipath effects or interferences in the signal transmitted having as result a degradation or even loss of the received information.

The high-level communication architecture between the train and the ground with any application (APP) is shown in Fig. 1. A communication channel is set-up between the train and the base-station transceiver to enable the bi-directional information flow. The communication channel that allows sending data between the train and the ground can be implemented by different technologies. For example, in mainline tracks, the current ERTMS, consists of ETCS (European Train Control System), focused on the signaling protocols), and its communication technology based on GSM-R [9]. Although GSM-R is the current communication channel for ERTMS, it is becoming obsolete due to its limitations [10] [11]:

1) GSM-R cannot provide advanced services and is not able to be adapted to new requirements, such as critical video application, which needs more data rate than the GSM-R one. The maximum transmission rate of GSM-R per connection is 9.6 kbps, which is sufficient only for applications with low data rate demands such as ETCS. As well, message delay is in the range of 400 ms, which is too high to support any real-time application and emergency communication [10].

2) Interferences: the interference between GSM-R and GSM public network increases because both railways and public operators aim at having coverage along the rail tracks [10]. Theoretically, such interference can be avoided if public operators do not use frequency bands adjacent to those of GSM-R for the areas close to rail tracks; however, this is not well implemented in practice. As well, other EM interferences disturb GSM-R, such as the sliding contact between the catenary and the pantograph [12].
3) Capacity limitation (spectrum): GSM-R has no sufficient resources for the next-generation railway system, where each train will need to establish a continuous data connection with a Radio Block Center (RBC), and each RBC connection needs to constantly occupy one-time slot [10] [11].

Because of these limitations, new technologies for railways are being researched in order to achieve the capabilities required by new services. Due to the current global communication architectures, these new communications technologies for railways are moving towards Internet Protocol (IP) based solutions, such as LTE-R. This technology is being specially researched in railways due to the performance and level of maturity of LTE, such as [13] stated. Moreover, the first LTE-R network of China is scheduled for 2020 [10]. Furthermore, some European projects are researching which solutions along different possible technologies will be the optimal ones such as the Shift2Rail project.

III. TESTING STRATEGIES AND EMULATORS

This section introduces the current communication technologies testing strategies and network emulator characteristics.

A. TESTING STRATEGIES

When a new technology or application is developed, it has to be validated before putting it into service. In fact, the testing strategy for the system under test is crucial to have a correct validation. This can be done by different strategies, so it is important to clearly understand the difference between them in order to know the different capabilities that each of them offers. These strategies are implemented in different environments starting from the laboratory by means of simulation, HW in the loop (HIL) and emulation, and finally, when the laboratory tests have been successfully passed, on-site testing. Fig. 2 shows a complete proposal of validation taken into account the different strategies named above.

FIGURE 2. Complete testing strategy.

Knowing the differences between all these strategies is key to understand what each can offer. It should be pointed out that in the literature [14]-[17], there is no agreement with regards to the simulation and emulation definitions. Therefore, the following definitions of simulation, HIL, and emulation are used in this paper:

1) Firstly, simulation is defined as a software-based technique that is based in an analytical model, which represents the key characteristics, behaviors, and functions of the real system describing them using mathematical tools for a system virtualization in a controlled environment. Simulation usually does not require specialized hardware equipment, which can significantly reduce the final cost of this experimental technique. The disadvantage of this approach is the fact that the real environment needs to be generalized to describe the real system.
2) Then, the hardware in the loop (HIL) is a technique where real hardware is present in the simulation loop, being the testing and the evaluation of the system carried out in real-time [17]. HIL is most often used in the development and testing of embedded systems when those systems cannot be tested easily, thoroughly, and repeatable in their operational environments [18]. This testing strategy is more focused on the subsystem level than in a whole system under test. The system under test is thought as the final assembled product.
3) The emulation is an experimental technique, which replicates the same inputs and outputs as the real system with the same performance, e.g., being able to test in real-time. The main goal is to test a whole system.
4) Finally, on-site testing is defined as testing the whole system in the real world with a real environment. The on-site testing in railways is difficult and expensive. Some reasons are the following ones: the non-full availability of the infrastructure and the dynamic change of the environment.

Therefore, it can be said that the emulation and the simulation are the options to check the performance of a whole system under test. However, HW in the loop is not the most suitable strategy for testing a whole system under test but for testing a specific equipment meaning subsystem testing.

Nevertheless, these laboratory testing strategies have some advantages and disadvantages versus on-site testing, as it is shown in Table I.

TABLE I
ADVANTAGES AND DISADVANTAGES OF LABORATORY VS ON-SITE TESTING

	Laboratory	On-site
Advantages	Costs, Save time & Repeatability	Real world
Disadvantages	Difficult to identify all the possible conditions	High costs and time Real conditions (unpredictable and not repeatable)

The main advantages of laboratory strategies are the reduction of costs and time and the option of repeating the test the times that the user desires. However, for an effective testing an important disadvantage is the need to overcome the challenge of having a realistic environment prior to deployment in order to reproduce coherent network conditions. For this purpose, the emulation can be considered because it is close to the real world, and it replicates the real-time performance of the emulated system. However, it is hard to obtain in the emulation the same results as the on-site testing due to the fact that the real environment conditions produce unpredictable impairments. In fact, one of the main challenges is to accurately identify which are the best and worst use cases conditions in which networks and devices will have to work in.

Comparing emulation and simulation, it can be easily seen that the emulator has to be in real-time, because it employed the final equipment, while simulator can be faster, slower or in real-time because it employs simulation models. This implies that the emulator is expected to have the same behavior (or close to) of a real system, same inputs, and outputs as the real system, allowing a closer real performance evaluation.

With regards to on-site testing, due to the costs, availability, and not repeatability among others, it is clear that the fact of shifting tests from on-site testing to the laboratory is necessary. In this field of moving on-site testing to the laboratory there are a number of European projects such as EATS [20] included in FP7 EU research funding program; and VITE [21], X2RAIL-1 [22] and X2RAIL-3 included in Shift2Rail initiative [4]. The main objective is to reduce on-site tests for signaling systems (in these cases, focused on ETCS railway application), leading to reduce overall testing costs under the zero on-site testing concept. In fact, the key objective of zero on-site testing is to perform functional and non-functional tests (component test, integration test, and system test) in laboratory, instead of testing on-site, in order to save time and costs without compromising safety [23].

Therefore, bringing to the laboratory (emulation or simulation) at least a representative number of use cases from the real world is required. Nevertheless, it should be taken into account that some of the use cases could only be tested on real site testing; they are not possible to be shifted to laboratory such as brake testing.

The complete and most desirable testing process would be to pass through every laboratory testing strategy allowing to characterize the infrastructure performance before taking the deployment decision, and once the deployment is completed on-site test could be carried out. However, a trade-off is necessary due to the cost restrictions of covering all the testing strategies proposed. Therefore, considering that the emulator is the testing strategy closest to the real world, it is considered the best solution to shift the on-site testing to the laboratory.

B. NETWORK EMULATORS CHARACTERISTICS

As it is stated in section II, different elements along the track cause different effects on the communication channel. Some of these effects that disturb the communication channel at RF level are multipath or EM effects. These disturbances caused in real world have to be taken into account when testing. Communication Network Emulation tools are capable of inserting different impairments to the channel with the aim of reproducing real-world condition effects, approaching the best to reality.

As mentioned in section II, the current trend is to move forward to IP based connection schemes. When employing an IP emulator, these RF disturbances are present but indirectly; although the disturbances affect the channel directly at RF level, they do it indirectly at IP level. Hence, the RF interferences are present and possible to be translated into IP network impairments if the research of how they affect to IP level is developed. A list of IP level emulators is listed in section IV.

With regards to the IP level the network emulators, the following IP impairments are defined:

1) Packet delay: is the time required for each bit of the packet to traverse the network or a segment of the network, independent of the packet size [24].
2) Bandwidth: in terms of data network, bandwidth quantifies the data rate at which a network link or a network path can transfer; the amount of data a link or network path can deliver per unit of time [25].
3) Packet corruption: deletion, insertion, modification, re-ordering in a packet in terms of bits.
4) Packet loss: the failure of a packet to traverse the network to its destination [24][26]; it is measured as a percentage of lost packets with respect to sent packets. It occurs when one or more packets transmitted over an IP network fail to arrive at their destination. Packet loss is typically caused by what is generally referred to as network congestion but also because of distance or poor line quality. Excessive packet loss is perceived as disconnections (broken or missing communication).

Thanks to the possibility of modifying these impairments, it is possible to change the conditions of the environment, so it enables to test both the train to ground signaling application and the communication technology that is used to transmit the different messages. Therefore, in order to select the most suitable network emulator, it is needed to have a look into the characteristics of the different current applications that are being developed for the sector (explained in section II), in this case, railway area.

Some criteria can be applied in order to choose the best suitable network emulator. In fact, the three features shown in Fig. 3 are the main criteria from the functionality point of view. These functional features are explained below:

1) Performance: the degree to which a system or component accomplishes its designated functions within

given constraints, such as speed, accuracy, or memory usage [27].
2) Precision: refers to the closeness of two or more measurements to each other.
3) Repeatability: the user can carry out the same test under the same conditions the times that the user wants.

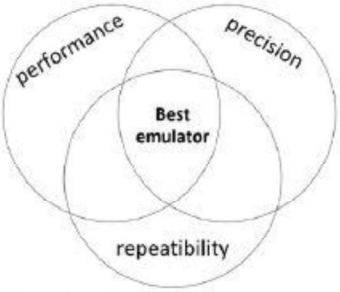


FIGURE 3. Criteria for the best emulator.

If an emulator has a high performance, high precision, and high repeatability, it is for sure an emulator closer to the real world. Nevertheless, these features are not the only ones when choosing a suitable emulator; other important non-functional criteria are cost and time constraints (not represented in Fig. 3).

Different network emulators, which are currently in the market, are explained and compared in the following sections, section IV and section V, respectively.

IV. CURRENT IP NETWORK EMULATORS TESTING

In this section, a number of IP-based network emulators available currently in the market are listed and briefly described.

First, the difference between HW and SW emulators is explained according to Fig. 4, where the definition of terms of high and low depends on the explained concept. For example, low physical ports means that the HW emulator has a fixed number of ports, and it is not possible to increase them.

Usually, HW based emulators are standalone devices, implemented on dedicated HW platforms, equipped with an operating system and software suitably optimized and customized. Thanks to these characteristics, these solutions usually grant excellent performance and accuracy certified by the manufacturer. On the other hand, they usually are very expensive [28].

However, despite the apparent limits of SW network emulators, they are widely used in many contexts and especially for research applications. The main reason is their being open source and, consequently, they can be easily modified by the researcher for their goal [28]; the SW emulator could be adapted to the user's requirements. As well, due to the high portability of the SW emulator, as they are computer programs, it can be installed in any device: they just need a host hardware platform to be executed.

In the rest of the terms, the concept of high carries positive meaning; therefore, being low negative meaning.

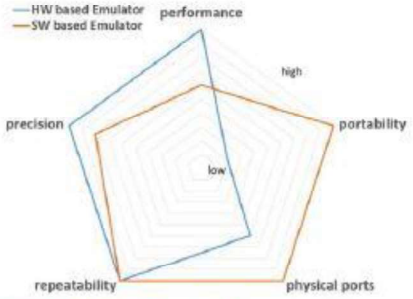


FIGURE 4. Comparison of HW-based vs. SW-based emulators.

The features of both types of emulators are not only those explained in section III (performance, repeatability, and precision) but also portability and limitation of ports:
1) Portability: the ease with which a system or component can be transferred from one hardware or software environment to another [29].
2) Limitation of ports: HW based emulators are limited physically with a fixed number of ports. However, the SW ones are device-independent, so they can be deployed in any device with the possibility to add more ports.
3) Performance and precision: HW emulators, respecting SW ones, are dedicated devices to the emulation task, and its operating system may be optimized for this purpose offering higher execution speed and higher accuracy of the effects being introduced [30]. However, SW based ones have to share the processor with other tasks.

Another classification of the emulators refers to proprietary or commercial network emulators and open source network emulators. Both terms have to described as they are two relevant points in the emulators. The proprietary or commercial network emulators are the ones which its software is owned by the individual or the organization that developed it. In contrast, the open-source emulators refer to the ones that have been developed and tested through open collaboration meaning anyone with the required academic knowledge can access the source code, modify it, and distribute his own version of the updated code.

Firstly, the proprietary emulators are listed as follows:
1) Spirent Network Emulator [31]: it provides industry-leading flexibility in building and modelling these complex real-life systems enabling the user to emulate

networks and the real-world conditions under which applications and platforms need to perform.

2) PacketStorm IP Network Emulators [32], [33]: it reproduces the unfavorable conditions of IP Networks and WANs in a controllable and repeatable lab setting. The emulator recreates the dynamic behavior of the Internet such that any network model can be reproduced, including those models that change with traffic, time, or the behavior of another traffic flow.
3) IXIA [34]: it allows users to accurately emulate the real network conditions that occur over live production LAN/WAN networks. By emulating realistic and worst-case network conditions in the lab, users can validate and test performance of new hardware, protocols, and applications to prevent failures in production networks.
4) NetDisturb [35]: it allows disturbing flows over IP networks, helping to study the behavior of applications, devices, or services in a disturbed network environment.
5) Linktropy 8510 [36], [37]: it emulates terrestrial, wireless, satellite, internet, and other wide area networks to test applications under a spectrum of real-world conditions.

Then, the open-source emulators are the following ones:

6) WANem [38], [39]: it is a tool, which brings the Internet into the user's development/test/lab environment. It emulates internet under conditions within the user's control so that the user can check application performance and availability.
7) Dummynet [40]: it is a live network emulation tool, initially designed for testing networking protocols, and since then used for a variety of applications, including bandwidth management. It enforces queue and bandwidth limitations, delays, packet losses, and multipath effects.
8) NIST Net [41]: it is a general-purpose tool for emulating performance dynamics in IP networks.
9) NetEm [42], [43]: it provides network emulation functionality for testing protocols by emulating the properties of wide-area networks.

V. COMPARISON BETWEEN THE DIFFERENT EMULATORS

Each of the emulators listed in section IV has different aspects; thus, a comparison is needed in order to see which one is the best option according to the user requirements. In this section, a comparison is shown focusing on the main IP impairments (explained in section III) that almost all the network emulators can modify. However, modifying these impairments does not just affect the emulated communication channel, but also the performance of the different applications that go through it, e.g., video or voice, can be worsened.

In this section, a comparison by IP impairments of many emulators is shown; nevertheless, throughout the literature more specific comparison between emulators involving a specific aspect exists (usually just two emulators at the same time and just one IP impairment) [44][45].

In Table II and III, it is shown the different characteristics of each network emulator (from section IV) according to the information provided by its corresponding vendor.

TABLE II
COMPARISON BETWEEN THE DIFFERENT NETWORK EMULATORS I

Network Emulator		Spirent	PacketStorm6rg	PacketStorm HurricaneV	IXIA	NetDisturb	Linktropy 8510
Type of emulator	Commercial	Yes	Yes	Yes	Yes	Yes	Yes
	Open Source	No	No	No	No	No	No
GUI friendly		Yes	Yes	Yes	-[1]	Yes	Yes
HW/SW based		HW	HW	HW	HW	SW	HW
Configurable network impairments	Delay	From 0 ms up to 4/10 s depending on 10GbE/1GbE	From 0 up to 1000 ms resolution of 100 ns	From 0 up to 10s resolution of us	Max 2s/20s/30s depending on 1GE/10GE/100 MBE	From 1ms up to 10s	From 0 ms up to 10 sec. in increments of 0.1ms
	Loss (%)	From 0.1% up to 100% in increments of 0.1%	From 0 up to 100% in increments of 0.01	From 0 up to 100% in increments of 0.01	Min 2.232*10e-8 %	Yes	From 0 up to 100% in increments of 0.0001%
	Duplication	Yes	Yes	-[1]	Yes	Yes	Yes
	Packet corruption	Yes	Yes	Yes	Yes	Yes	Yes
	Reordering	Yes	Yes	Yes	Yes	Yes	Yes
	Limited BW	From 500bps up to 1Gbps	Up to 7.475 Gbps[2]	Up to 4.848 Gbps[2]	-[1]	Up to 1Gbps[3]	From 300 bps up to (depending on

1 Not provided by the vendor
2 Being packet size 64 bytes
3 Bandwidth data number depends on the Network Interface Card of the device where the SW emulator is installed: Ethernet/FastEthernet/GigaEthernet
4 Not longer actively maintained

							the license key of the product) in increments of 1 bps
	Disconnection	Yes	Yes	Yes	Yes	Yes	Yes

TABLE III.
COMPARISON BETWEEN THE DIFFERENT NETWORK EMULATORS II

Network Emulator		**WANem**	**Dummynet**	**NIST Net[4]**	**NetEM**
Type of emulator	**Commercial**	No	No	No	No
	Open Source	Yes	Yes	Yes	Yes
GUI friendly		Yes (basic & adbance mode)	No	-[1]	No
HW/SW based		SW	SW	-[1]	SW
Configurable network impairments	**Delay**	Yes	Yes	Yes	Yes
	Loss (%)	Yes min 1/2[32]	Yes	Yes	Min 0.0000000232%
	Duplication	Yes	Yes	Yes	Yes
	Packet corruption	Yes	-[1]	-[1]	Yes
	Reordering	Yes	-[1]	Yes	Yes
	Limited BW	Yes[3]: min 120 kbps	Yes[3]	Yes[3]	Yes[3]
	Disconnection	Yes	-[1]	-[1]	-[1]

In Tables II and III, there are some relevant differences with regards to the specific parameters between the analyzed emulators:

1) The most obvious one is that some emulators are commercial and the other ones are open source. This usually is one important feature of the emulator depending on the user's budget and support. In Table II and III there are six proprietary emulators and four open-source emulators.
2) The network emulator with higher accuracy, in terms of the configurable network impairments, would allow more precision in the results. Therefore, having high repeatability due to the same inputs would carry the same outputs. For example, in this comparison, PacketStorm6xg is the more accurate emulator in terms of delay due to the accurate resolution that is capable of providing to the user.
3) Another difference, which requires a more in-depth analysis, is the selection of HW or SW based emulators (See Fig. 4). The HW ones, e.g., Spirent emulator, have a given hardware device with some physical characteristics as up to 16 ports; if the user needs more ports, it is not possible to add more. However, the SW emulator can be adapted to any desired device, e.g. NetEm emulator.
4) The quality of the available documentation of the emulator could also be a drawback, i.e., some emulators as NetEm whose source code is constantly modified, and there is no well-documented description of its functionalities.
5) Another important topic is whether the emulator is a generic emulator that could be applied for a specific sector, or it is specific for a given sector. In the case of the emulators compared, all of them are generic.

The rest of the features, in general, differ ones from others and depend on the user; it could be more or less interesting as the ease of the network emulator usage. For example, the applications environment of the network emulator. The emulators are used as well to test the performance of different types of applications. Depending on the application that the user wants to test, the choice of the emulator would be different. Another example, real-time audio information is delay-sensitive, and excessive packet loss decline in voice quality [46]. Therefore the IP impairments, which the user has to focus on, are delay (and jitter) and excessive packet loss, which affect the quality and produce distortion. Another example of applications is the real-time video, which normally has more data to transmit. The IP impairments that degrade the different types of videos are the same as in audio applications but including another aspect: the bit data rate normally is higher than in audio, therefore, the bandwidth [47]. However, other types of applications such as exchange and retrieved information (e.g., FTP, mails) are focused on loss and bit error rate [48][49].

These examples would need a different choice of network emulators. Consequently, the real-time audio and video applications would be suitable to be tested with a network emulator which delay is accurate such as PacketStorm HurricaneV or PacketStorm6xg (which has high BW to test video applications). However, for data transfer applications, a network emulator having the option of packet corruption and more accuracy in loss would be more suitable such as Linktropy 8510 or NetEm. Throughout the literature, it can be found that many experiences have been carried out with different applications and the network emulators exposed in section V as [50]-[55] state in order to test the performance of the target applications.

The network emulators allow the application to test them in different network environments allowing the research of

the network performance impact of the application service. However, the common characteristic is the fact that all of them can repeat tests.

From the railway point of view, it can be stated that none of the emulators analyzed includes railway aspects, i.e., the impairments that the user could introduce in the emulator are in a static position not in movement as a train does. Therefore, due to the different environments present on the track, dynamic configurable impairments are needed as input in the emulator in order to represent the effects of these areas, such as tunnels or urban areas, matching with the train position. In this manner, the effects in the on-site testing could be shifted to the laboratory replaying the real scenario and allowing to test different applications. Therefore, it has to be taken into account that this need is missing and has to be developed.

VI. CONCLUSIONS

The future railways' applications, being more restrictive the safety-critical ones, request demanding requirements for the communication technologies, such as large bandwidth, low delay, and high availability. Therefore, in order to test the functions of the different applications based on communications, the performance of the communication technology has to be validated prior to putting it into service. Currently, extensive on-site tests are considered for the validation, however, due to the high cost, time demand, and lack of repeatability, on-site tests could be considered unfeasible for the whole validation process. In order to overcome this drawback, this paper proposes a testing strategy where, firstly, laboratory tests are performed, and then on-site tests getting as result that the network emulator is the best option at laboratory level to approaches the closest to the real world and allowing to reduce the on-site tests.

As it is shown in this paper, many network emulators are currently available in the market, being able to modify different parameters of the communication channel. Choosing the suitable network emulator depends on the user requirements, e.g., what application has to be performed and/or the budget. By means of the analysis and comparison of the different characteristics of the emulator shown in this paper it can be stated that there is currently not a universal emulator consequently, the user should be able to choose the most suitable network emulator according to his/her requirements.

With regard to the railways case, it has been shown that no one of them is specific for the railway environment. The need of a railway network emulator is given by the necessity of testing different environments that the train could be passing through (tunnels, urban areas, suburban areas…), each one is, indirectly (due to e.g., multipath effect) modifying the characteristics of the communication channel. Because of this, a network emulator being able to take into account the location of the train will improve the test and validation process. In this manner, it is possible to include the negative effects of the perturbations listed beforehand related to the corresponding railway environment, e.g., the tunnel in a specific position reducing the coverage. Nowadays, the need of a specific emulator for railways still exists.

REFERENCES

[1] Juan Moreno et al., "A survey on Future Railway Radio Communications services: Challenges and Opportunities", IEEE Communications Magazine Oct. 2015.
[2] Bo Ai et al., "Future Railways Services-Oriented Mobile Communications Network", IEEE Communications Magazine, October 2015
[3] Juyeop Kim et al., "Automatic Train Control over LTE: Design and Performance Evaluation", IEEE Communications Magazine Oct. 2015
[4] Shift2Rail initiative. Available in https://shift2rail.org/
[5] FRMCS project. Available in https://uic.org/frmcs
[6] Nerea Fernandez et al., "Survey of Environmental Effects in Railway communications", 13th International Workshop Nets4Cars/Nets4Trains/Nets4Aircrafts 2018
[7] V. Deniau, H. Fridhi, M. Heddebaut, J. Rioult, I. ADIN, and J. Rodriguez, "Analysis and Modelling of the EM Interferences Produced above a Train associated to the Contact between the Catenary and the Pantograph," EMC Eur. 2013, p. 6p, 2013.
[8] T. Hammi, N. Ben Slimen, V. Deniau, J. Rioult, and S. Dudoyer, "Comparison between GSM-R coverage level and EM noise level in railway environment," 2009 9th Int. Conf. Intell. Transp. Syst. Telecommun. ITST 2009, pp. 123–128, 2009.
[9] S. F. Ruesche, J. Steuer, and K. Jobmann, "A Packet-Switched Approach to a Train Control System," IEEE Veh. Technol. Mag., no. September, pp. 37–46, 2008.
[10] Ruisi He et al., "High-Speed Railway Communications", IEEE vehicular technology magazine, September 2016
[11] C. Maton, "Study on migration of railway radio communication system from GSM-R to other solutions", ERA 2015 04 1 RS
[12] S. Dudoyer et al., "Study of the susceptibility of the GSM-R communications face to the electromagnetic interferences of the rail environment," IEEE Trans. Electromagn. Compat., vol. 54, no. 3, pp. 667–676, 2012.
[13] J. Yan, D. Li, Y. Xu, and J. Chen, "Performance Evaluation of LTE-R System with Mobile Relay for Differentiated Services" 2019 28th Wirel. Opt. Commun. Conf., no. Wocc, pp. 1–5, 2019.
[14] Muhammad Imran et. al, "A Survey of Simulators, Emulators and Testbeds for Wireless Sensor Networks", 2010 International Symposium on Information Technology
[15] Shie-Yuan Wang et al., "EstiNet OpenFlow Network Simulator and Emulator", IEEE Communications Magazine September 2013
[16] Ryosuke Koshimura et al., "Development of Web–QoE evaluation system for wireless LAN with combination of simulator and network emulator", IEEE 2nd Global Conference on Consumer Electronics (GCCE) 2013
[17] P. Cika, "Network Emulator of Transmission Parameters of Data Networks," *2018 10th Int. Congr. Ultra Mod. Telecommun. Control Syst. Work.*, no. 1, pp. 1–6, 2018.
[18] M. Bacic, "On hardware-in-the-loop simulation", 44th IEEE Conference on Decision and Control, and the European Control Conference 2005
[19] Jim A. Ledin, "Hardware-in-the-Loop Simulation, Embedded Systems Programming", Embedded Systems Programming Feb. 1999
[20] EATS project. Available in http://www.eats-eu.org/
[21] VITE project. Available in https://shift2rail.org/project/vite/
[22] "Deliverable D3.1 User & System Requirements (Telecommunications)", X2RAIL-1: Start-up activities for Advanced Signaling and Automation Systems
[23] X2RAIL-1, "Deliverable D6.1 Current test condition and Benchmarking report"

[24] ITU-T, "ITU-T Rec. G.1050: Network model for evaluating multimedia transmission performance over Internet Protocol," 2016.
[25] R. P. C. D. M. M. kc Claffy, "Bandwidth Estimation: Metrics, Measurement Techniques, and Tools," IEEE Netw., pp. 27–35, 2003.
[26] J. Gozdecki, A. Jajszczyk, and R. Stankiewicz, "Quality of service terminology in IP networks," IEEE Commun. Mag., vol. 41, no. 3, pp. 153–159, 2003.
[27] ISO/IEC/ IEEE 24765, "International Standard Systems and software engineering — Vocabulary," 2013.
[28] D. Capriglione, G. Cerro, L. Ferrigno, and G. Miele, "Experimental Analysis of Software Network Emulators in Packet Delay Emulation," pp. 0–5, 2017.
[29] ITU-T, "Terms and definitions"
[30] Razvan Beuran, "Introduction to Network Emulation. 2012".
[31] Spirent, "Spirent Network Emulator datasheet," pp. 1–8.
[32] PacketStorm, "PacketStorm6xg Network emulator datasheet"
[33] PacketStorm, "PacketStorm Hurricane V Network emulator datasheet"
[34] IXIA, "Network Emulator II TM - Impairment Emulation Problem : knowing how networks" pp. 1–8.
[35] P. D. S. Ltd., "NetDisturb Impairment Emulator / Simulator Software for IP Networks (IPv4 & IPv6)," Configurations, vol. 1, no. 650, pp. 1–19.
[36] A. Technologies, "Linktropy 8510 Hardware Guide."
[37] A. Technologies, "Linktropy WAN Emulator User's Guide," no. 3610.
[38] TATA, "WANem 2.0 Wide Area Network Emulator Performance Engineering Research Centre," 2008.
[39] H. K. Kalita and M. K. Nambiar, "Designing WANem : A Wide Area Network Emulator tool," pp. 4–7, 2011.
[40] M. Carbone and L. Rizzo, "Dummynet revisited," ACM SIGCOMM Comput. Commun. Rev., vol. 40, no. 2, p. 12, 2010.
[41] M. Carson and D. Santay, "NIST Net: a Linux-based network emulation tool," ACM SIGCOMM Comput. Commun. Rev., vol. 33, no. 3, pp. 111–126, 2003.
[42] S. Hemminger, "Network Emulation with NetEm," 6th Aust. Natl. Linux Conf., no. April, pp. 1–9, 2005.
[43] A. Jurgelionis, J. Laulajainen, M. Hirvonen, and A. I. Wang, "An Empirical Study of NetEm Network Emulation Functionalities," *2011 Proc. 20th Int. Conf. Comput. Commun. Networks*, pp. 1–6, 2011.
[44] J. Shaikh, T. N. Minhas, P. Arlos, and M. Fiedler, "Evaluation of Delay Performance of Traffic Shapers," *2010 2nd Int. Work. Secur. Commun. Networks*, pp. 1–8, 2010.
[45] T. N. Minhas, M. Fiedler, J. Shaikh, and P. Arlos, "Evaluation of Throughput Performance of Traffic Shapers," *2011 7th Int. Wirel. Commun. Mob. Comput. Conf.*, pp. 1596–1600, 2011.
[46] W. Zhang, Y. Chang, Y. Liu, and Y. Tian, "Perceived QoS Assessment for Voip Networks," *2013 15th IEEE Int. Conf. Commun. Technol.*, pp. 707–711.
[47] J. Okyere-benya, M. Aldiabat, V. Menkovski, G. Exarchakos, and A. Liotta, "Video Quality Degradation on IPTV Networks," *2012 Int. Conf. Comput. Netw. Commun.*, pp. 702–707, 2012.
[48] Sakti and D. R. Prajitno, "Experiment of Networked Control System (NCS) Using Network Emulator," *2016 3rd Int. Conf. Inf. Technol. Comput. Electr. Eng.*, pp. 100–105, 2016.
[49] S. Rimac-drlje, "Applications Classification and QoS Requirements," pp. 517–522, 2002.
[50] C. Feng, L. Huang, X. Tan, and A. T. Jitter, "Network Packet Level based Video Quality Assessment Metric," *2015 First Int. Conf. Comput. Intell. Theory, Syst. Appl.*, vol. 1, pp. 149–152, 2015.
[51] S. Hao, X. Huang, and Y. Ma, "Application QoS for Video : Benchmarking , Control Strategies and Business Systems," *2009 5th Int. Conf. Wirel. Commun. Netw. Mob. Comput.*, pp. 1–4, 2009.
[52] H. Larijani, "Voice Quality in VoIP Networks Based on Random Neural Networks," *2010 Ninth Int. Conf. Networks*, pp. 89–92, 2010.
[53] R. F. Sari, P. Libratu, and P. Wirya, "Performance Analysis of Session Initiation Protocol on Emulation Network using NIST NET," pp. 506–510, 2007.
[54] M. Marchese, "Study and performance evaluation of TCP modifications and tuning over satellite links.," pp. 129–133, 2000.
[55] K. Matsuo and K. Tajima, "A Study of Network Emulation for IP Network-based Control System of a Small Motor," *2014 17th Int. Conf. Electr. Mach. Syst.*, pp. 2216–2220, 2014.

NEREA FERNÁNDEZ-BERRUETA received the M.Sc. degree in Telecommunications Engineering from Faculty of Engineering in Bilbao (University of the Basque Country) in 2016. She joined the CEIT-IK4 Research Centre in San Sebastián in 2017, and she is currently a PhD student within the Transport and Sustainable Mobility group. Her research activity lies in the field of railway signaling specially in zero on-site testing and adaptable communication systems.

JON GOYA received the M.Sc. degree in telecommunications engineering in 2011 and the Ph.D. degree from the Universidad de Navarra in 2016. He is currently a Lecturer with the TECNUN, Universidad de Navarra, and also a Researcher with the CEIT-IK4. His professional research activity lies in the simulation of on-board positioning system and performance analysis for railway. He has participated in FP7 projects coordinated by CEIT and is now actively participating in H2020 European funded, such as AIOSAT, and Shift2Rail projects in Railway signaling and Positioning topics: X2RAIL-2 and FR8RAIL.

JAVIER AÑORGA received his MSc degree in Telecommunications Engineering from Tecnun , University of Navarra, Spain, in 2011 and obtained his PhD degree from the University of Navarra in 2015. He is in possession of the CEH (Certified Ethical Hacker) certification with license number ECC9021470888. He joined the Ceit-IK4 Research Centre in San Sebastián in 2011, in the Electronics and Communications Department. His professional research activity is in the field of communication protocols, QoS, QoE, Residential Gateways, Embedded Systems, Information Technology, and Cybersecurity. He has published several articles in international journals and conferences. Currently, he is also lecturer at the Engineering School of the University of Navarra.

SAIOA ARRIZABALAGA received her MSc in Telecommunication Engineering from the Faculty of Engineering in Bilbao (University of the Basque Country), Spain, in 2003, and the PhD. in Engineering from Tecnun, University of Navarra, Spain, in 2009. She joined the CEIT Research Centre in San Sebastián in 2003, and she is currently a Researcher and Project Manager with the Electronic and Communication Department of CEIT. On the other hand, she also worked as a teaching collaborator from 2005 to 2009 and from 2009 as a Lecturer at Tecnun (University of Navarra), also in San Sebastián. She has published several papers in international journals and also written a book entitled "Multi-Service QoS Architecture for a Multi-Dwelling Gateway" (Saarbrücken, Germany, VDM Verlag Dr. Müller, 2010). Her professional research activity lies in the field of communication protocols, QoS, Cloud Computing, Embedded Systems and Information Technology.

GORKA DE MIGUEL received the M.Sc. degree in Telecommunications Engineering from TECNUN (School of Engineering of San Sebastián), Spain, in 2015. He joined the CEIT-IK4 Research Centre in San Sebastián in 2015, and he is currently a Researcher Assistant and Ph.D student within the Transport and Sustainable Mobility group. His research activity lies in the field of positioning and software development. He is now actively participating in H2020 European funded and Shift2Rail projects in Railway signaling and Positioning topics.

JAIZKI MENDIZABAL received his M.Sc. and Ph.D. degrees in Electrical Engineering from TECNUN (University of Navarra, San Sebastian, Spain) in 2000 and 2006 respectively. He joined Fraunhofer Institut, Germany) and SANYO Electric Ltd, Japan as RF-IC designer. Nowadays, he is at CEIT-IK4, in San Sebastián (Spain) where his research interests include communications and electronic systems. He is lecturing "Communications Electronics" and "Communications via Radio" at TECNUN (University of Navarra).

A.2 INTERNATIONAL CONFERENCE PAPERS

- S. Azpiazu, G. de Miguel, J. Goya, G. Solas, S. Arrizabalaga, J. Mendizabal, "GPS and EGNOS positioning and integrity determination algorithm for on-board railway applications", *DCIS 2013: Design of Circuits and Integrated Systems*, San Sebastian (SPAIN) 25-27 november 2013
- J. Goya, L. Zamora, S. Arrizabalaga, G. de Miguel, I. Adin, J. Mendizabal, "BTS infrastructure impact analysis and design of an advanced BTS filtering for UKF-based positioning algorithm", *9th International Conference on Circuits, Systems, Communications and Computers, (CSCC 2015),* Zakynthos Island, Greece, July 16-20, 2015.
- I. Adin, J. Mendizabal, G. de Miguel, J. Goya, L. Zamora, S. Arrizabalaga. "Complementary Positioning System in GNSS-denied Areas". *Transportation Research Procedia*. Volume 14, 2016, Pages 4562–4571. *Transport Research Arena TRA2016*.
- R. González, Peter Lubrani, ESSP SaS and Gorka de Miguel, Iñigo Adín, Jaizki Mendizabal CEIT "EGNOS positioning in rail domain (ERSAT EAV)", *Royal institute of Navigation, INC 2016* Glasgow
- Mendizabal J., Solas G., Valdivia L.J., de Miguel G., Uranga J., Adin I. (2016) ETCS's Eurobalise-BTM and Euroloop-LTM Airgap Noise and Interferences Review. In: Mendizabal J. et al. (eds) *Communication Technologies for Vehicles. Nets4Cars/Nets4Trains/Nets4Aircraft 2016. Lecture Notes in Computer Science*, vol 9669. Springer, Cham
- De Miguel, G., Goya, J., Uranga, J., Alvarado, U., Adin, I., & Mendizabal, J. (2017). GNSS Complementary Positioning System performance in railway domain. *In Proceedings of 2017 15th International Conference on ITS Telecommunications, ITST 2017*. https://doi.org/10.1109/ITST.2017.7972204

- Mendizabal, J., Goya, J., De Miguel, G., Valdivia, L., Arrizabalaga, S., & Adin, I. (2017). Virtual testing of the on-board ETCS with GNSS based Train integrity determination. *In Proceedings of 2017 15th International Conference on ITS Telecommunications, ITST 2017*. https://doi.org/10.1109/ITST.2017.7972203
- Mendizabal, J., de Miguel, G., Uranga, J., Sedano, B., Goya, J., & Adin, I. (2017). "On-board electromagnetic interference field-test and evaluation of a non-electrified railway regional line". *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 10222 LNCS). https://doi.org/10.1007/978-3-319-56880-5_9
- F. Parrilla, M. Alonso, D. Batista, A. Alberdi, J. Goya, G. De Miguel and J.Mendizabal. "Technologies Evaluation for Freight Train's Wireless Backbone". *Communication Technologies for Vehicles. 13th International Workshop, Nets4Cars/Nets4Trains/Nets4Aircraft 2018*, Madrid, Spain, May 17-18, 2018, Proceedings. ISBN: 978-3-319-90370-5
- Goya, J., Mendizabal, J., Adin, I., De Miguel, G., Roth, Michael H., Ademeit, A.M., Groos, J.C., "On-board positioning strategies based on GNSS low-cost receivers for rail freight transport". *In: Transport Research Arena 2018. Transport Research Arena (TRA) 2018*, 16.-19. April 2018, Wien, Österreich.

GNSS Complementary Positioning System performance in railway domain

De Miguel, G., Goya, J., Uranga, J., Alvarado, U., Adin, I., & Mendizabal, J.

In Proceedings of 2017 15th International Conference on ITS Telecommunications, ITST 2017.

https://doi.org/10.1109/ITST.2017.7972204

GNSS Complementary Positioning System Performance in Railway Domain

De Miguel, G.
Transport and Energy Monitoring
CEIT and Tecnun (University of Navarra)
Donostia-San Sebastián, Spain
gdemiguel@ceit.es

Goya, J.
Transport and Energy Monitoring
CEIT and Tecnun (University of Navarra)
Donostia-San Sebastián, Spain
jgoya@ceit.es

Uranga, J.
Transport and Energy Monitoring
CEIT and Tecnun (University of Navarra)
Donostia-San Sebastián, Spain
juranga@ceit.es

Alvarado, U.
Transport and Energy Monitoring
CEIT and Tecnun (University of Navarra)
Donostia-San Sebastián, Spain
ualvarado@ceit.es

Mendizabal, J.
Transport and Energy Monitoring
CEIT and Tecnun (University of Navarra)
Donostia-San Sebastián, Spain
jmendizabal@ceit.es

Adin, I.
Transport and Energy Monitoring
CEIT and Tecnun (University of Navarra)
Donostia-San Sebastián, Spain
iadin@ceit.es

Abstract — **An on-board Complementary Positioning System (CPS) has been proposed to overcome the limitation of GNSS positioning systems. This paper deals with the test phase of CPS to guarantee the localization functions in areas in which GNSS signals are not available. In order to analyze the suitability of CPS its performance in a railway operational line has been evaluated. The CPS is based mainly on GNSS, Inertial Measurement Units (IMU) and Wireless Communications Technologies (WCT) as input sources for localisation. The field tests have been carried out in an operational regional line and availability has been increased as it was the purpose.**

Keywords—Railway; positioning; GNSS; Multi-sensor; IMU; GSM; UMTS; Kalman filter; Cold start of mission;

I. INTRODUCTION

European Union (EU) has established many goals in railway industry, such as the improvement of the transport efficiency or the reduction of costs in order to make it a more attractive transportation method. Regarding the efficiency, the European Rail Research Advisory Council (ERRAC) [1] has defined some milestones aiming an augmentation in track's capacity by means of a more efficient use of them.

A proposal is to improve the on-board GNSS based positioning systems to fulfill the European Train Control System (ETCS) requirements. On-board GNSS positioning systems are expected to be the key factor for ETCS level 2 to ETCS level 3 migration. ETCS level 3 allows the infrastructure costs to decrease up to 25% the level 2 costs [2]. Furthermore, with this migration the railway efficiency could be increased in more than a 50%.

The suitability of GNSS systems for railway safety applications is being analyzed by several European projects such as:

ERSAT-EAV [3] Main challenges in ERSAT are to reuse the ETCS odometry by using the virtual balise concept to eliminate fixed balises; use the public EGNOS augmentation network along with local augmentation networks, and to verify and validate alternative GNSS solutions to guarantee the positioning functions in areas in which GNSS signal is not available.

GRAIL2 [4][5][6][7] aimed to develop and validate a GNSS-based ETCS odometry systems prototype. The objectives of SATLOC [8][9] were to develop and validate the use of GNSS in low traffic lines signaling and train control. Satloc showed the feasibility of the system presented on a real line. The need of complementary positioning techniques were shown as essential and a future use of route maps and virtual balises was proposed.

GaLoROI [10][11][12] project based the position solution on a GNSS receiver, including Galileo technology, combined Eddy Current Sensor (ECS) and track-matching techniques. The RAMS tests carried out during the project determined that the performances of GNSS could not meet the railway requirements in harsh environment, showing the need of additional sensor fusion in order to increase the availability of the system.

3InSat [13][14] developed a Localization Determination System (LDS). It combines GNSS systems with augmentation techniques and integrity detection solutions. The objective of the project was to fulfill the requirements of the European Rail Traffic Management System (ERTMS)/ETCS Level 2 trackside balise. A GNSS solution and a theoretical modeling concerning the Safety Integrity Level were developed in order to integrate the satellite based localization systems in ERTMS-ETCS environment.

The main drawback detected in the use of on-board satellite positioning systems is the inability of providing a position in every point of a railway [15]. They have coverage problems in urban environments or in difficult orography such as tunnels or canyons.

In order to solve this limitations, a number of other sensors such as IMUs or WCT have been proposed to build the Complementary Positioning System. The accuracy and reliability of this kind of sensors is not as good as the one provided by GNSS systems but on the other hand they offer an increased availability.

The paper is structured as follows:

- Section II describes the complementary positioning system proposed.
- Section III characterizes the environment and the railway in which the tests were done.
- Section IV shows the results obtained.
- Section V describes the conclusions drew from the work presented.

II. COMPLEMENTARY POSITIONING SYSTEM DESCRIPTION

In this section the complementary positioning system proposed to overcome the limitation of GNSS is introduced and its hardware described.

In Figure 1 the positioning system developed by Ceit is shown. It includes a multisensor approach for location estimation that consis of the following:

- GNSS receiver: responsible for providing a position to the system to use it as reference position while the GNSS system is available.
- Inertial Measurement Unit (IMU): to perform a position estimation using the accelerations computed when the CPS crosses a GNSS blocked area.
- GSM/UMTS receiver (WCT receiver): , used to give a position when GNSS and IMU position estimation is not reliable, but also to perform a cold start of mission when satellite based positioning is not available; and a microprocessor in charge of the performance of the algorithms developed for the position estimation computation using the inputs presented before.

Excluding the IMU the rest of the modules are placed into a 25cm x 16cm x 20.5cm metallic container in order to make the system portable. As it has its own GNSS receiver (a mid-range receiver, usually less accurate than other on-board GNSS receivers), CPS can perform a stand-alone operation and fully position the train without external aid.

Two different architectures have been proposed; one for the characterization of the railway in which the CPS is used and the CPS itself for the positioning operation in the line.

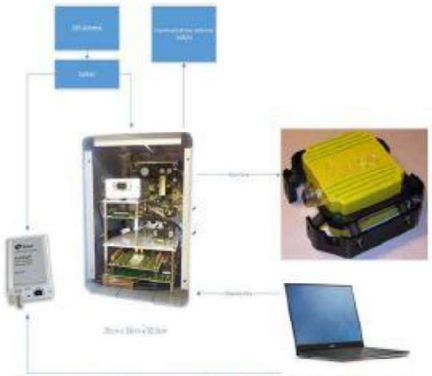


Figure 1 Positioning and line characterization system hardware

The main difference between both architectures is the need of different elements for the different functions to be carried out by the system. A PC and an external GNSS receiver are introduced in the characterization architecture in order to monitor the GNSS status during the first journeys and identify zones in which GNSS signal issues, such as tunnels or natural

blockers occur. These zones are then marked as Known Blocked Scenarios (KBS) and an algorithm to deal with them has been developed.

In both cases the CPS subsystem sends and receives communication data through the communication modem connected to the antenna, and obtains the GNSS data from the GNSS antenna, as it has to work stand alone for that test campaign.

Moreover, some of the equipment mounted in the train has also been used, and it is also described. In particular, the antennas used for GNSS and GSM data acquisition were placed on the train roof, see Figure 2.

Figure 2 Antennas placed in ALN 668 3114 roof (Courtesy of ASTS)

III. TEST CAMPAIGN

The Ceit test campaign was performed during three days in the Sardinia Trail railway site, from 24.10.2016 until 26.10.2016.[16]

The test plan followed during the campaing can be seen in Table 1. Further information about each of the test types is given in Section III.

Day	Journey	Test Description
24	1	KBS calibration
24	2	KBS calibration validation
25	1	CPS subsystem average operation
25	2	CPS subsystem average operation
25	3	KBS operation analysis
26	1	Cold start of mission analysis
26	2	KBS operation analysis
26	3	CPS subsystem average operation

Table 1 Test plan performed

General details of the test campaign are:

- Time frame: 24.10.2016. – 26.10.2016
  - Cagliari-San Gavino trips:
    - 8:05, 13:05, 16:05
  - San Gavino-Cagliari trips:
    - 9:10, 14:10, 17:10
- Dynamic vehicle: Diesel locomotive ALN 668-3114 from Trenitalia. (See Figure 3)

Figure 3 Diesel locomotive ALN 668-3114 from Trenitalia

- Railway line: Cagliari – San Gavino, nearly 50km. (See Figure 4)

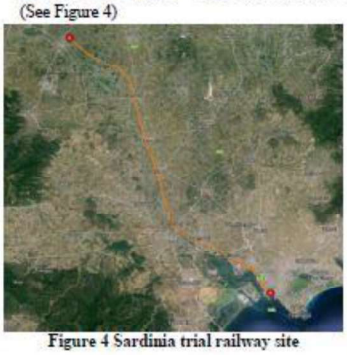


Figure 4 Sardinia trial railway site

The pilot line where the GNSS Measuring Campaign was performed is placed in an open-view-sky area. It is settled in a rural environment as shown in Figure 5. Even though, some industrial facilities can be found during the journey, these buildings are not as high as to obstacle GNSS signals. Then, the test railway track is a GNSS friendly track as there are no tunnels or canyons in which a full loss of GNSS signal is expected. However, some bridges that affect to GNSS signals can be found when the track goes through villages. Anyhow, there are challenging scenarios during the track. In Figure 6, a train station with the buildings very close to the track is shown. In addition, a shed roof and a bridge crossing the track are shown. This kind of environment is quite challenging as GNSS signals could be totally or partially blocked due to the facilities along the trackside. Furthermore, in the case that the signals are received, most of them are going to be non-line of sight or multipath signals due to the multiple objects in the signal trajectory in which it could reflect [17].

Figure 5 General landscape in the trackside

Figure 6 Possible GNSS blockers

During the measurement campaign, additionally the ground truth was obtained with some equipment on board (antennas). In the tests presented in this paper, a GPS only antenna was used by the CPS as the multiconstellation-multiband antennas were used for the ground truth creation. This fact has some relevance in the results obtained as the accuracy achieved using a multiconstellation-multiband antenna is higher. The test campaign carried out was divided in two parts:

- Determination of the GNSS blocked scenarios
- CPS subsystem operation analysis including average operation, KBS operation and cold start of mission.

A. Determination of the GNSS blocked scenarios

In this first test phase, the configuration and thresholds required for the correct operation of the CPS are determined. For this purpose, the real time configuration architecture is used. It is connected to a GNSS antenna (in this test it is just a GPS only antenna) to receive the GNSS signal for on-board positioning and a communications antenna (GSM/UMTS in this test) to receive the observables from the base stations for on-board positioning.[18]

The CPS and the auxiliary GNSS receiver have to be able to determine the minimum threshold to trigger the KBS functionality. This threshold is highly dependent on the antenna positioning. If the antenna has a full view of the sky, the threshold will be higher as the difference between blocked and non-blocked scenarios will be higher due to the good health of the received signal in non-blocked scenarios.

This test was performed during the first journey in a round trip from Cagliari to San-Gavino, what makes a total of nearly 100km. The reason to perform the tests in a round trip is that the environmental obstacles could interfere in a different way depending on the antenna orientation.

B. CPS subsystem average operation

In this test phase the CPS has already been configured and can be used to position the train stand-alone. For this purpose, the real time operation architecture is used. In this case it is also necessary to connect it to a GNSS antenna (in this test is a GPS only antenna) and a communications antenna (GSM/UMTS).

The CPS determines the positioning and continuity output data as a result of its inputs. Three different round trips were performed in order to test the performance of the average CPS subsystem operation under different environmental conditions and satellite geometries. Making a total of more or less 300km of test.

IV. RESULTS

In this section the obtained results are presented. Two fucntions of the sysetma are analysed:

- KBS operation
- Cols start of mission

A. KBS operation analysis

KBS operation was analyzed once the average operation was reliable (see Section IV). For this purpose, the real time architecture is used. The suitability of the KBS to detect and mitigate GNSS shadowed areas was proved.

As in the other test types more than one round trip was tested under this type of operation to analyze its performance under different environments. At the end the operation was tested in about 200km of railway.

Regarding to the positioning system one of the first steps is the calibration of the KBS positioning technique to detect the SNR threshold applicable for the field tests to fix the KBS sensibility for which KBS is triggered. As KBS is based on mean SNR of the satellites the mean SNR and the SNR difference with the previous epoch.

The sensibility of the KBS should be configured to be triggered in case the SNR difference falls down considerably compared with the previous epoch and thus assuming that the GNSS signals are not trustful anymore unless the SNR shortfall is recovered. This allows that complementary positioning sensors, such as IMU or WCT, have a trustful reference beginning for its use improving the position estimation results once GNSS signals are not available. The KBS system allows detecting the entrance of the tunnels or other blocking structures where the complementary positioning sensors take the key role regarding to the position estimation. For the particular case of the test campaigns the selected threshold between two consecutive SNR epoch was set to -4 dB (Figure 7 depicted in green).

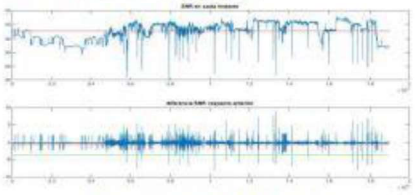


Figure 7 KBS Threshold obtention

The rest of the section describes the performance results of the CPS system during the test campaign. The most representative cases are depicted and discussed in order to avoid the repetition of cases that do not shed light on the performance of the CPS.

The first case under study is the average behavior of the CPS system along the test campaigns allowing the further comparison with the other representative situations. Figure 8 depicts the average CPS performance where the accuracy error is bounded to 10 meters. During the test campaign the CPS system worked only with the GPS signal, however the use of GNSS can improve the CPS system accuracy performance.

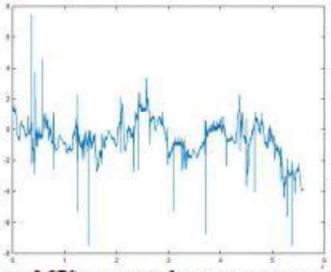


Figure 8 CPS average performance accuracy error

Figure 9 show the histogram of the obtained accuracy errors. As it can be observed the error follows a Gaussian distribution. The absolute mean error 0.95 meters and standard deviation of 0.84meters.

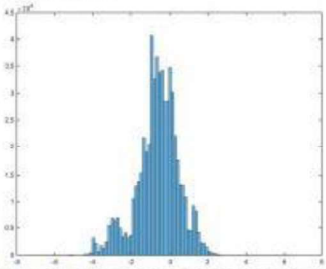


Figure 9 CPS average performance histogram

Thanks to the complementary sensors employed, the system availability and continuity reaches 100%.

Once the average CPS performance is described other significant test campaigns are presented.

Next test campaign provides the behavior of an undetected degraded GPS signal. The aim of the KBS is to detect the degraded GPS signals to avoid the use of them in the position estimation. Once the SNR signal level drops down the error of the position estimation increases and thus the KBS should detect this situation. Figure 10 shows a peak of 35 meters of accuracy error. This error must be avoided and a lower KBS conditions should be applied. However, as the SNR is also dependent on other factors and not only in the position. It is important to have in mind that some peaks could steal appear.

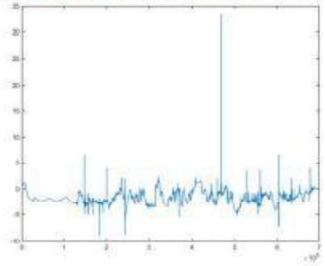


Figure 10 CPS performance during a not detected degraded GPS signal

Regarding to the error accuracy histogram, it is clear that the error is still very low and the peak is a punctual situation for which the CPS systems is recovered very fast (see Figure 11).

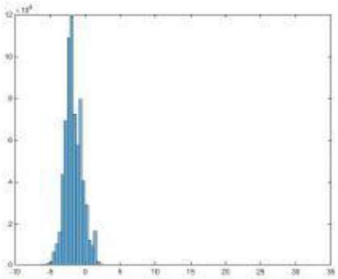


Figure 11 CPS performance histogram during a not detected degraded GPS signal

B. Cold start of mission analysis

Cold start of mission was analyzed using the real time architecture but without GNSS signal availability. For that purpose, at the beginning of a journey, the GNSS signal was discarded. As there was not a reference position given by the GNSS, the IMU was not able to estimate the railway position. WCT was the system used to obtain an absolute position.

This test was only performed once as the satellite geometry would not influence the WCT performance and the beginning of the journey is always performed in the same station.

One of the most relevant test campaign for the CPS is the cold start mission where the CPS is no able to obtain any GPS signals. This cause will force the system to use the WCT position estimation to obtain an absolute position estimation. Figure 12 describes this effect during the beginning of the journey. As it can be observed the error is high compared with the average CPS behavior.

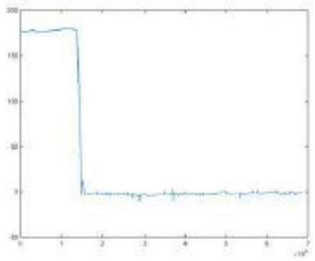


Figure 12 CPS performance for the cold-start of mission

Figure 13 depicts the same behavior translated into the histogram diagram. Compared to the second described test campaign the amount of epoch with accuracy error of 175 meters lasts more than one epoch being clearly observed as a second lobe around 175 accuracy error.

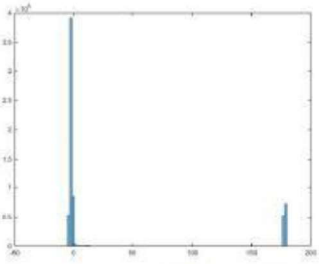


Figure 13 CPS histogram for cold-start of mission

The described three test campaign scenarios are the most significant situation in which CPS works. The average CPS performance of the system is promising as there are not any augmentation techniques or FDE techniques used. However, determining the correctness of the position estimation and cold start mission situations are still topics that need to further researched.

V. CONCLUSIONS

This paper presents the progress for the introduction of GNSS based positioning systems in railway. One of the key aspects that is important to remark is the need of complementary information sources for the continuous position and speed monitoring to provide to the system being available all the time at the expense of the accuracy error performance.

Field-test campaigns allows the definition system sensibility threshold for the GNSS blocking areas. Punctual GNSS signal degradation are situations that positioning systems already are dealing and solving. However, the main challenge is to overpass the GNSS blocking situation. Because of the dependability of GNSS signals of the environment these undesired effects should be handled and discarded in order to provide a high integrity system. Regarding to this, CPS provides a low cost solution approach that allows the discrimination of some of this situation. The inclusion of additional fault exclusion techniques will enhance the proposed solution even if the availability of the system can be affected. In order to address this issue, the proposed system uses additional information sources that increase the availability of the on-board positioning system. IMU and WCT based data fusion strategy implemented for the CPS provides a solution to reach the 100% availability, which is one of the goals set to the system to satisfy the requirements of safety critical applications.

Moreover, for the application case of the safety cold start of mission procedures the accuracy error should be reduced considerably. Further research will help to enhance the performance that already is improved regarding to the position estimation based on mobile network infrastructure data. The inclusion of positioning techniques that will take advantage of

the static situation of the train during the cold-start of mission must be addressed in order to improve the first fixed position estimation when GNSS is not available.

ACKNOWLEDGMENT

This research was supported by the ERSAT-EAV project as part of the European H2020 framework of projects funded by the European Commission (EC) and managed by the European GNSS Agency (GSA).
The author wants to express especial thanks to our colleagues from ASTS, RFI, Trenitalia, Sogei, Radiolabs and ESSP who allowed both measuring campaigns in the project's framework and the use of their infrastructure.

REFERENCES

[1] ERRAC, "RAIL ROUTE 2050: THE SUSTAINABLE BACKBONE OF THE SINGLE EUROPEAN TRANSPORT AREA," Tech. Rep., 2013. [Online]. Available: http://www.errac.org/wp-content/uploads/2013/11/D9-SRRARAILROUTE2050.pdf

[2] V. amdas, T. Bradbury, S. Denniss, D. Chapman, R. Bloomfield, and D. Fisher, "ERTMS LEVEL 3 RISKS AND BENEFITS TO UK RAILWAYS (FINAL REPORT)," Transport Research Laboratory, Tech. Rep., 2010. [Online]. Available: http://www.trl.co.uk/media/773946/cpr798 -ertms level 3 risks and benefits to uk railways.pdf

[3] ERSAT EAV Grant Agreement Nº 640747.

[4] GNSS FOR ENHANCED ODOMETRY: THE GRAIL-2 RESULTS, Livio Marradi, Andrea Galimberti, Lucio Foglia, Alberto Zin, Cristina Pecchioni, Manuel Doronzo, Emilio Javier Gonzalez Garcia-Consuegra, Mouna Lekchiri. NAVITEC 2012 and European Workshop on GNSS Signals and Signal Processing.

[5] GRAIL-2: ENHANCED ODOMETRY BASED ON GNSS, Emilio Gonzalez, Celso Prados, Virginia Anton, Boris Kennes. Procedia - Social and Behavioral Sciences. Transport Research Arena 2012.

[6] GSA. GRAIL-2 – Summary. 2012. Available online: www.gsa.eeuropa.eu/ gnss-based-enhanced-odometry-rail

[7] GNSS IN RAIL SAFETY APPLICATIONS, Alvaro Urech, I Rail Technological Forum For Internationalization, Madrid, 2011

[8] THE SATLOC PROJECT, George Barbu, Juliette Marais. Transport Research Arena 2014, Paris.

[9] GSC. SATLOC – Summary. Available online: https://www.gsc-europa.eu/markets/rd/fp7-3rd-call-projects/satloc-satellite-based-operation-and-management-of-local-low

[10] DEPENDABILITY EVALUATION OF A GNSS AND ECS BASED LOCALISATION UNIT FOR RAILWAY VEHICLES, T.P.K.Nguyen, J. Beugin, J. Marais. 13th International Conference on ITS Telecommunications (ITST)

[11] GSC. GaLoROI – Summary, 2014. Available online: https://www.gsc-europa.eu/markets/rd/fp7-3rd-call-projects/galoroi-galileo-localization-for-railway-operation-innovation

[12] GALILEO LOCALISATION FOR RAILWAY OPERATION INNOVATION, Uwe Becker, 2012.

[13] AN ANALYTICAL EVALUATION FOR HAZARDOUS FAILURE RATE IN A SATELLITE-BASED TRAIN POSITIONING SYSTEM WITH REFERENCE TO THE ERTMS TRAIN CONTROL SYSTEMS, A. Neri, A. Filip, F. Rispoli, A.M. Vegni. Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012)

[14] RECENT PROGRESS IN APPLICATION OF GNSS AND ADVANCED COMMUNICATIONS FOR RAILWAY SIGNALING. F. Rispoli, M. Castorina, A. Neri, A. Filip, G. Di Mambro, F. Senesi. 23rd International Conference Radioelektronika 2013

[15] ADVANCED TRAIN LOCATION SIMULATOR (ATLAS) FOR DEVELOPING, TESTING AND VALIDATING ON-BOARD RAILWAY LOCATION SYSTEMS, Jon Goya, Leticia Zamora-Cadenas Saioa Arrizabalaga Alfonso Brazalez Juan Melendez Jaizki Mendizabal. Eur. Transp. Res. Rev. DOI 10.1007/s12544-015-0173-5

[16] "EGNOS POSITIONING IN RAIL DOMAIN (ERSAT EAV)", R. Gonzalez, Peter Lubrani, ESSP SaS and Gorka de Miguel, Iñigo Adin, Jaizki Mendizabal CEIT, Royal institute of Navigation, INC 2016 Glasgow.

[17] ON-BOARD ELECTROMAGNETIC INTERFERENCE FIELD-TEST AND EVALUATION OF A NON-ELECTRIFIED RAILWAY REGIONAL LINE. Jaizki Mendizabal, Gorka De Miguel, Julen Uranga, Jon Goya and Iñigo Adin. Nets4workshop 2017, Toulouse, France.

[18] DATA ACQUISITION FOR COMPLEMENTARY POSITIONING SYSTEM IN GNSSDENIED AREAS S. Arrizabalaga, I. Adin, G. De Miguel, J. Goya, L. Zamora-Cadenas, J. Mendizabal. World Congress on Railway Research (WCRR) 2016, Milan, Italy.

Complementary Positioning System in GNSS-denied Areas

I. Adin, J. Mendizabal, G. de Miguel, J. Goya, L. Zamora and S. Arrizabalaga

Transportation Research Procedia. Volume 14, 2016, Pages 4562–4571. Transport Research Arena TRA2016

Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 14 (2016) 4562 – 4571

Transportation Research Procedia

www.elsevier.com/locate/procedia

6th Transport Research Arena April 18-21, 2016

Complementary positioning system in GNSS-denied areas

I. Adin[a,*], J. Mendizabal[a], G. de Miguel[a], J. Goya[a], L. Zamora[a], S. Arrizabalaga[a]

[a]Ceit and Tecnun, Paseo Mikeletegi 48, 20009 San Sebastian, Spain

Abstract

The suitability of EGNSS (including EGNOS and Galileo early services) for safety railway applications has been analysed by several European projects such as Grail, Grail-2, Satloc, 3InSat, and it is currently being analysed by ERSAT EAV. ERSAT EAV in particular is addressing the following three challenges beyond the state-of-the-art: a) reuse of ETCS odometry by adding the virtual balise concept to eliminate the fixed balise along the line; b) adoption of the public EGNOS Augmentation network "upgraded" with Local augmentation networks to fulfil the railways' requirements; c) verification and validation of alternative GNSS solutions to guarantee localization functions in areas where the GNSS signal is not available and/or subject to interference. This paper focuses on the third of these challenges, addressing the design and the implementation of a complementary positioning system (CPS) for GNSS-denied areas. This system is based mainly on the Public Land Mobile Network (PLMN) information but also incorporates additional source information (such as predefined GNSS blocked zone information and GNSS positioning information for calibration purposes, when available) in order to enhance positioning performance.

This paper deals with the different steps taken to create the CPS:

1. Analysis of the alternatives for complementary positioning systems when GNSS is not available
2. Simulation and development of algorithms for the CPS in the lab
3. Creation of a prototype for the CPS based on COTS components

More specifically, after analysing the state-of-the-art technologies, the algorithms for the selected solution will be implemented in the laboratory. In order to analyse the theoretical performance of the algorithms, the environment and observables for the algorithms will also be modelled in the simulation platform. The performance of the CPS will be analysed in terms of accuracy, reliability, continuity and integrity in a way similar to the GNSS systems. Finally, the algorithms will be implemented in a prototype based

* Corresponding author. Tel.: 0034943212800.
E-mail address: iadin@ceit.es

Peer-review under responsibility of Road and Bridge Research Institute (IBDiM)
doi:10.1016/j.trpro.2016.05.379

on COTS components so that the CPS will not be a theoretical solution but also a real solution that could be useful for the on-board positioning function in trains when GNSS solutions are not available.

Keywords: railway positioning; GNSS-denied areas; data fusion; IMU; radio communication observables; fingerprinting

Nomenclature

AoA	Angle of Arrival
BTS	Base Transceiver Station
CDMA	Code Division Multiple Access
CPS	Complementary Positioning System
DCM	Data Correlation Method
EIRENE	European Integrated Railway Radio Enhanced Network
GNSS	Global Navigation Satellite System
GSM-R	Global System for Mobile Communications
IMU	Inertial Measurement Unit
MORANE	MObile RAdio for Railways Networks in Europe
RSS	Received Signal Strength
LDS	Location Determination System
QM	Quality Measurement
TDoA	Time Difference of Arrival
ToA	Time of Arrival
ToF	Time of Flight
UMTS	Universal Mobile Telecommunication System
WCT	Wireless Communication Technologies

1. Introduction

One of the main challenges of ERTMS Level 3 is defining a suitable positioning system given such restrictive safety and availability requirements. The inclusion of GNSS technologies is one of the unavoidable steps forward that the railway sector has to face. However, the particularities of train itineraries force the GNSS receiver to complement other devices. In that sense, this paper shows the progress made in that field in the first period of the European project ERSAT-EAV (ERSAT-EAV 2015).

There are two main alternatives for location in non-GNSS based systems. On the one hand, wired sensors such as odometers and inertial measurement units (IMU) provide continuous positioning after directly processing their data. With such systems, the quality of the device and sensors affects the dependability/reliability of the estimation error. On the other hand, some wireless devices are able to estimate position using a more complex operation. Recently, the EATS Project (ETCS Advanced Testing and Smart Train Positioning System) (S. Arrizabalaga et. al. 2014) has, for the first time, introduced the GSM-R and UMTS (WCT) technologies in the on-board location system.

These technologies are mainly based on the analysis of the specific physical characteristics of radio signals, which are called observables, such as Received Signal Strength (RSS), Time of Flight (ToF) and Angle of Arrival (AoA). The wireless location methods presented in section 3 use the observables introduced in section 2, either alone or in combination, in order to provide a position (Bensky, A., 2008).

In any case, in each circumstance, the set of available measurements affects the position estimation technique or signal processing strategy used to solve the user's positioning problem. Three main groups of algorithms can be distinguished: geometric approaches, dead reckoning, and fingerprinting (Mauricio A. Caceres Duran et. al. 2012). The first and second technique types estimate the user's position using the information (e.g. time, angle, and odometer data) extracted directly from the signal received wirelessly or through wired input. The third group first requires a measurement campaign be carried out in order to build a database of location fingerprints, and then the location of the user is estimated by matching the received measurement (usually power measurements) with the closest one stored

in the database. The algorithms and techniques are shown in section 3, the complimentary positioning system (CPS) in presented in section 4, and the plans for implementation, validation and future deployment are described in the sections 5–7.

2. Technologies and information needed for non-GNSS-based location systems

This section describes the wireless communication technologies (GSM-R and UMTS) and devices such as the IMU and the odometer that are considered as sources of information that can be used for positioning.

2.1. Wireless communication technologies: GSM-R and UMTS

GSM-R (Global System for Mobile Communications-Railway) is used for communication between train and railway regulation control centres as part of the European Rail Traffic Management System (ERTMS). GSM-R is based on GSM technology and on EIRENE (European Integrated Railway Radio Enhanced Network) and MORANE (MObile RAdio for Railways Networks in Europe) specifications. The task of railway operators in EIRENE is to define the GSM-R system requirements and the functional requirements thus guaranteeing the interoperability between the railway networks.

Special requirements for GSM-R networks derive from the requirements of applications using GSM-R, such as 95% coverage for 95% of the time in a designated coverage area with a level of above -90dBm. Moreover, coverage must exists inside tunnels (3GPP TS 44.031 2012). The high availability of GSM-R in railway environments makes it interesting for location purposes, although given that it is based on GSM technology, the accuracy obtained with this technology is much worse than the accuracy obtained by the GNSS system. However, it can be useful when higher accuracy technologies are not available.

In analysing the viability of the GSM-R system for positioning, we looked at previous analyses of GSM technology in order to select the observables that potentially could be obtained from the GSM-R transceiver in the train or that could be sent to the transceiver by the surrounding infrastructure. The information regarding potential observables will be integrated into the position calculation algorithm presented in the next section."

In contrast, location accuracy from third generation cellular devices (UMTS) is significantly better than that achieved in GSM and CDMA. The inherent accuracy of UMTS is greater because of the increased signal bandwidth and shorter bit period. These features significantly improve the ability to distinguish the line-of-sight signal among multipath returns in reception. The Universal Mobile Telecommunication System (UMTS) supports the following location methods (3GPP TS 25.305 2012):

- Cell-ID
- Uplink TDOA (U-TDOA)
- Observed TDOA-idle period downlink (OTDOA-IPDL)
- Network-assisted (A-GPS)

Cell-ID is the most basic positioning technology for cellular networks (also for GSM-R). Each base station knows which cell the handset is in, and the handset also knows what cell it is in. The accuracy of this method depends on the dimensions of the cell, where smaller cells yield higher accuracy. Table 1 shows the sizes of some types of cells. It can be easily observed that the position obtained using this method will be very variable depending on the location.

Table 1. Cell types according to their sizes (3GPP TS 25.305 2012).

Cell Type	Cell dimension (km)
Large macrocell	3–30
Small macrocell	1–3
Microcell	0.1–1
Picocell	0.01–0.1
Nanocell	0.01–0.001

The Cell-ID method is more accurate in urban environments where there are more UMTS users and more cells are required to provide proper service. The need to have a higher number of UMTS cells is behind the need to make smaller cells, which substantially enhances positioning accuracy. On the other hand, in rural environments macrocells are common and thus this method is nearly unusable. Apart from that, a user terminal is not always connected to the nearest antenna. In some conditions the terminal can decide to connect to a neighbouring antenna, worsening the accuracy in positioning. Other enhancement methods can be defined for Cell-ID methodology, but this paper does not address any of them in depth. The base assumed for the work on ERSAT-CPS is Cell-ID RTT.

On the other hand, both U-TDOA and OTDOA-IPDL methods are based on Time Difference of Arrival methods. The first one, Uplink Time Difference of Arrival (U-TDOA), involves using the uplink signal and doing the time computation in the network. The second one, Observed Time Difference of Arrival (OTDOA-IPDL), uses the downlink signal measurements, and the computation that obtains the position can be done in the handset. Further information about these two methods can be found in the deliverables from the EATS project (S. Arrizabalaga et. al. 2014), but they are not representative for the needs of the complimentary positioning system presented here. However, they are compared to the Cell-ID method in Table 2. TDOA is found by finding the maximum cross correlation of a received signal with a replica of the known transmitted signal. In a handset-based system, the handset estimates the relative time delays of the input sequences from three or more base stations. Usually there is no synchronization between the handset and base stations clocks, so to estimate a position with TDOA the handset must receive the transmission times from each base station according to a common clock and their positions. The handset could also transmit its observed time differences to a special terminal that has all the information necessary to calculate the handset's position.

Table 2. Positioning methods characteristics (3GPP TS 25.305 2012).

Positioning Methods	Reliability	Latency	Applicability	Accuracy vs. Positioning Scenario		
				Rural	Suburban	Urban
Cell-Id – RTT	High	1–5 s	High	250 m–35 km	250–2500 m	50–550 m
OTDOA-IPL	Medium	<10 s	Medium	50–150 m	50–250 m	50–300 m
U-TDOA	Medium	<10 s	Low	50–120 m	40–50 m	40–50 m

2.2. Inertial measurement unit

An inertial measurement unit (IMU) is a device that has three gyroscopes and three accelerometers that are displaced along three mutually orthogonal axes. The main idea of inertial navigation is Newton's First Law: "A body will continue in its state of rest, or of uniform motion in a straight line, unless an external force is applied to it". The accelerometers detect the acceleration changes due to the forces of gravity and the gyroscopes detect the changes in the rotational attributes. The combination of these parameters can easily be used in algorithms to estimate a position, as the next section will explain.

Using IMUs has some benefits; they are autonomous and do not dependent on other devices or signal visibility. They do not need an antenna, so they can be placed anywhere as it is not necessary to have a clean sight for signal reception.

On the other hand, the main problem with IMUs is accumulative error. There are two main sources of error, namely bias error and noise:

- Bias errors are constant errors suffered by the sensors in their measurements. Bias can be static or dynamic. Static bias is a constant error resulting from incorrectly calibrated sensors, whereas dynamic bias is the in-run variation and changes over time. Dynamic bias is about 10% of static bias, so its influence in the total error is lower.
- Noise is the unwanted signal generated from internal electronics, which interferes with measurement of the desired signal. In general, velocity, position, and pitch-or-roll error from the accelerometer or gyroscope white noise will be smaller than the bias and constant errors described before.

2.3. Odometer

Odometry is the process by which the position of wheeled vehicles is estimated. This term is usually used in order to speak about the distance travelled and the position relative to a known point at which the trip started. To have good odometry data, it is necessary to have a precise data acquisition system, well-calibrated equipment and good data processing.

The odometer is the device that is responsible for giving the velocity of the train at any moment. This element is already present in all the trains equipped with the modern signalling systems. Specifically, inside an ERTMS system, it meets the following security requirements:

- The odometer requires a precision of +/-(5m+5%*s), s being the distance (in meters) from the last reference point to the position of the train.
- The precision for the on-board velocity should be better than +/-2Km/h for velocities lower than 30Km/h and it can increase linearly until the +/-12Km/h when the train reaches 500Km/h.
- Finally, it can support an acceleration error of +/-4 〖m/s〗^2.

3. Analysis of non-GNSS based location algorithms

This section presents three different algorithms, each of which independently calculates a user's position by using the data that is required by each approach. Their input data are directly extracted from the technologies' observables and other information systems, such as those presented in the previous section.

3.1. Geometric approach to wireless location algorithms

Consider that we want to calculate the (x_i, y_i, z_i) position of the user by using the observations presented in the previous section. Assuming perfect observations, i.e. that the calculation of these estimates is made without any error, the position of the user can be found by using simple geometric deterministic considerations:

- Using TOA or RTT observations, the user's position will be the intersection of the spheres centred in each BTS and with radio signals the distance between each BTS and the user. In a 3D space, a minimum of four BTSs are needed to obtain four ranging estimates. Thus, in a 3D space, the real distance $d_{(i,j)}$ between user i and BTS j is equal to:

$$d^2_{(i,j)} = (x_i - x_{BTS_j})^2 + (y_i - y_{BTS_j})^2 + (z_i - z_{BTS_j})^2 \quad (1)$$

where (x_i, y_i, z_i) are the user coordinates and $(x_{BTS_j}, y_{BTS_j}, z_{BTS_j})$ are the BTS coordinates.

- Using TDOA observations, the user's position will be the intersection of the hyperboloids describing the time difference measurements between four or more BTSs in a 3D space. The time differences of arrival are a set of points with constant range-difference from each BTS j to the BTS reference $m = 1$. In a 3D space, the real $d_{(i,j,m)}$ is equal to:

$$\begin{aligned} d_{(i,j,m)} &= d_{(i,j)} - d_{(i,m)} \\ &= \sqrt{(x_i - x_{BTS_j})^2 + (y_i - y_{BTS_j})^2 + (z_i - z_{BTS_j})^2} \\ &- \sqrt{(x_i - x_{BTS_m})^2 + (y_i - y_{BTS_m})^2 + (z_i - z_{BTS_m})^2} \end{aligned} \quad (2)$$

Unfortunately, in the real world the observations present errors and the corresponding spheres, hyperboloids or lines do not intersect in a unique position. To solve the equations and provide the position estimation, different techniques are used, such as Least Square Error (LSE) (Guowei Shen et. al. 2008), Taylor Series (Guowei Shen et. al. 2008), Recursive Least Square (RLS) (Yuechun Chu et. al.2005) or more advanced techniques such as particle filters (H. Wang et. al. 2011) and Kalman filters (S. M. Kay 1993), (G. Welch et. al. 2006).

3.2. Dead reckoning

Inertial navigation systems are integrated systems able to detect position and velocity of a unit in movement, due to the information obtained from odometers, gyroscopes and accelerometers. They calculate the position via dead reckoning using the data received from the motion sensors and the rotation sensors. To do that, IMUs and odometers are used.

Dead reckoning is a method that calculates the current position of an object based on previous measurements. It is able to give the best available trajectory estimation on position, but the information must be accurate to be reliable, as it suffers from accumulative error. If the information is not accurate enough, the position will not be accurate, and the calculation of the next position would be based on an incorrect initial position, making that second calculation have an even higher error.

Dead reckoning could be sufficiently accurate for a short period of time, but the longer it runs, the error in position estimation will continuously increase due to the accumulation of errors.

3.3. Fingerprinting

Fingerprinting is a Data Correlation Method (DCM). The idea behind database correlation is to store the parameters obtained from a signal seen by a mobile station from the whole coverage area of the location. The algorithm uses a correlation approach to obtain the location (B.D.S.Lakmali et. al. 2007). This method involves three steps:

- Obtaining fingerprints: The stored signal information is called a "fingerprint". Depending on the cellular system, fingerprints include signal strength, signal time delay, etc. Measurements can be taken by the network or by the mobile station. As previously mentioned, the received signal strength (RSS) at the mobile station is the selected fingerprint variable for the technologies mentioned in previous sections. The fingerprint database consists, in this case, of the GPS coordinates of a location and the received signal strength of the base station it is connected to and other neighbouring base stations. A sliding window method can be used to reduce the distance between two fingerprints and resolution can be increased by overlapping the measurements (B.D.S.Lakmali et. al. 2007).
- Database preparation: The database can be prepared in two ways. The first one is by taking measurements in the location where the algorithm will be used afterwards. The second one is by using predictive databases, which use propagation models in different scenarios to predict how signals are going to behave in that environment.
- Location estimation: The RSS of each of the cells and the input fingerprint should be given to the DCM algorithm. The objective is to find the stored fingerprints that best match the input fingerprint. This can be done by measuring the signal distance between each database fingerprint and the input fingerprint. Another way to do this is by defining a valid RSS range for each cell in each fingerprint, and then determining whether or not an input RSS value is inside this valid RSS range. After identifying the nearest neighbours by using one of these two approaches, the location is estimated using the weighted k-nearest neighbours (WkNN) method, as in (3):

$$p = \sum p_i w_i \tag{3}$$

where p=Estimated location, p_i=Location of the i^{th} nearest neighbour and w_i= Weight of the i^{th} nearest neighbour

4. CPS architecture in the ERSAT-EAV system

CPS is the ERSAT-EAV's (ERSAT-EAV 2015) on-board LDS (Location Determination System) block, which provides a position whenever GNSS is not available. ERSAT-CPS will continuously provide a PVT estimation and a QM (Quality Measurement). The PVT will be employed to provide Position, Velocity and Time Data; on the other hand, a quality measurement will be a data of the quality of the position provided. An overview of the architecture is briefly explained in that section.

4.1. CPS architecture

The CPS approach differs depending on the nature of the GNSS availability. It relies on different strategies depending on whether GNSS is blocked or not. For non-blocked scenarios, the CPS will be synchronized, calibrated and aided by GNSS via the Universal PVT strategy. When GNSS is blocked the CPS will be used alone.

CPS will require PVT and QM from the ERSAT-EAV core (SIL4 EGNSS-based Enhanced Localisation for Railway). This function will be used to calibrate the CPS when GNSS is available, as the position accuracy offered by GNSS is higher than the one given by CPS. The data received will be employed by the CPS as a reference position to estimate the position based on data or information from other sensors. QM will be used in order to determine the quality of the position data. SIL4 EGNSS-based Enhanced Localisation for Railway will be also used for the CPS time synchronization. Time synchronization is the key factor in order to provide the data with the same time reference.

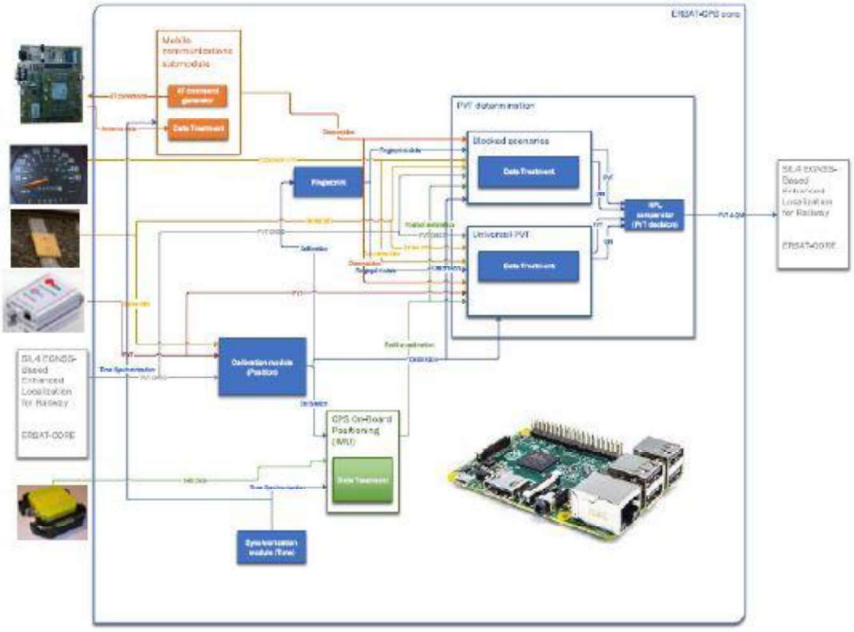


Fig. 1. ERSAT CPS block diagram.

Fig. 1 shows the block diagram that is going to be included in the ERSAT-CPS core, as well as the implementation foreseen for the processor and the peripherals based on COTS.

The CPS will complete the following functions:

- CPS will power on together with the complete ERSAT-EAV system.
- Communications between the core and the CPS will be take place through a wired interface. The PVT and QM will be transmitted in a bidirectional way from the CPS to the core and vice versa.
- CPS will synchronize its own time reference to the ERSAT-EAV core time reference.
- CPS will calibrate its reference position to the one provided by the ERSAT-EAV core, whenever the provided position quality is higher than the one estimated by the CPS.
- CPS will estimate the train PVT and QM by means of different sources, which were introduced in the previous sections. It will have an IMU and the corresponding processing to obtain the position using the dead reckoning strategy. By means of the observables obtained from WCT-TLC, it will estimate absolute position. It is foreseen that observables such as GSM-RXLEV or GSM-TA will be employed as part of a fingerprinting positioning strategy.

The SIL4 EGNSS-Based Enhanced Localization for Railway is not a part of the CPS; it is the complete system, but it gives data to the CPS in order to calibrate and time-synchronize the CPS. On the other hand, the CPS will return the position estimation obtained from all the systems that compose it to the SIL4 EGNSS-Based Enhanced Localization for Railway. It will provide Position Time and Velocity to the CPS and will be developed by other partners in the project.

5. Simulation strategy for the CPS algorithm

This section describes the ATLAS platform employed for the system simulation and lists the system performance indicators employed.

5.1. ATLAS platform

For the development and analysis of new location system algorithms, the ATLAS platform (J. Goya et. al. 2015) has been developed as part of the EATS project (S. Arrizabalaga et. al. 2014).

The Advanced Train Location Simulator (ATLAS) contains four independent and exchangeable modules to cover the entire process for location algorithm testing: the route simulator module, the input generator module, the position estimator module and the performance analysis module:

- The route simulator module obtains the real trajectory of the train; it has support for providing real trajectories for many places in the train (e.g: several receivers placed on the roof of the train in the different cars).
- The Input Generator module generates the observables/measurements that the particular location system to be developed/tested needs. If there are several receivers, the inputs for each of the receivers are generated. Different error patterns can be used and hundreds of iterations can be performed in order to provide statistically significant results at the end.
- The position estimator module is directly related to the location system: it offers a position estimation for the input parameters provided by the previous module. ATLAS makes it possible to test several alternatives simultaneously in order to compare their performance.
- The performance analysis module offers numerical and graphical results for the statistical analysis of the location algorithm results compared with the true/real trajectory that was generated in the first module.

The information exchange between the modules is carried out by using a fixed structure of folders and file names that are handled automatically by the platform. This modularity permits the modules to be easily exchangeable as long as they fulfil the predefined file format. In addition, ATLAS is useful not only for wholly simulated scenarios; it also carries out a performance analysis of location algorithms based on information

collected in field testing, as is the case for the ERSAT-CPS, which is introduced here. Field testing information can also be combined and/or compared with simulated results, as it offers many possible combinations.

5.2. CPS algorithm performance evaluation

As already highlighted in the description of the ATLAS platform, this tool has a performance analysis module that provides not only basic statistical analysis of individual shots but also grouped statistical analysis for several iterations. The basic statistical parameters are the following:

- The position distance error, defined as the Euclidean distance between the real position calculated and the estimated one at each time instant along the trajectory.
- RMSE (Root Mean Square Error) is a measure of the square root of the deviation of the estimated position about the true position, so it combines both the variance and the bias.
- CDF (Cumulative Distribution Function) is the probability of having a distance error in positioning that is lower than a certain value.

When several runs have been carried out within the same configuration, the grouped analysis can be run in order to analyse overall results at a glance. Utilities for comparing different configurations for one location system simultaneously or for comparing several location systems are also integrated into the platform. Additionally, the ATLAS platform performance analysis has also been upgraded to include more performance parameters that are especially common in GNSS algorithms: accuracy, reliability, integrity and continuity values are calculated, and Stanford Plots are included in the analysis.

In those conditions, the CPS algorithm will be evaluated in ATLAS by including real input data from the measurement campaigns, which should reflect a quite reliable behaviour for its final performance in the prototype deployment.

6. Future prototype deployment

Once the prototype introduced in the previous section is implemented and its software is validated in ATLAS, it will be deployed in a real environment. More specifically, the ERSAT project has a work package dedicated to the validation of the systems through a measurement campaign on board a train and another test period using the DLR train simulator.

During deployment, the performance of the CPS will be analysed and the differences between the laboratory test, its validation and the pilot test will be understood. The plan includes the following procedure:

- The real trajectory will be collected by using high-end GNSS receivers and/or map-matching techniques so it can be used as the common "truth" reference for both laboratory and pilot tests.
- The measurements used by the location system will be collected in order to provide the possibility of fine-tuning the location system in the laboratory by using the ATLAS platform.
- The laboratory-based (simulated) location algorithm results and field-test-based results will be compared by using the ATLAS performance analysis module, providing useful information for future enhancements for both the simulation platform and the deployed location system.

7. Conclusions

After analysing the state-of-the-art technologies and the information that a specific equipped train is able provide, the Complementary Positioning System was designed and its implementation with COTS components was proposed. Three algorithms – geometric approaches to wireless location, dead reckoning, and fingerprinting – need to be used in accordance with the available information at every moment.

Prior to deployment and validation in the field, the theoretical performance of the algorithms, the environment, and the observables for the algorithms could be modelled in the simulation platform called ATLAS. This would evaluate

the performance of the CPS in terms of accuracy, reliability and continuity, indicators similar to the GNSS system which CPS has to complement.

Acknowledgements

This work was supported by the European Community's Framework Program FP7/2007-2013 in the frame of the EATS project under grant agreement 31419, and also by the European Community's Framework H2020 in the frame of the ERSAT-EAV project under grant agreement 640747.

References

3GPP TS 25.305 Stage 2 functional specification of User Equipment (UE) positioning in UTRAN Universal Mobile Telecommunications System (UMTS) – Cell-ID location technique, limits and benefits: an experimental study E. revisani, A.Vitaletti, version 11.0.0 Release 11, 2012-09.

3GPP TS 44.031: Digital cellular telecommunications system (Phase 2+): Location Services (LCS): Mobile Station (MS) – Serving Mobile: Radio Resource LCS Protocol (RRPL) Location Centre (SMLC) 3GPP V11.0.0 Release 11, 2012-10.

Arrizabalaga, S., Mendizabal, J., Pinte, S., Sanchez, J., Gonzalez, J., Bauer, J., Themistokleous, M. and Lowe, D. "Development of an Advanced Testing System and Smart Train Positioning System for ETCS applications," TRA2014, 2014.

Bensky, A., 2008. Wireless positioning technologies and applications. Boston, Mass. Artech House.

Caceres Duran, M.A., D'Amico, A.A., Dardari, D., Rydstr¨om, M., Sottile, F., Ström, E.G. and Taponecco L. In Davide Dardari, Marco Luise, and Emanuela Falletti, editors, Satellite and Terrestrial Radio Positioning Techniques, Chapter 3 – terrestrial network-based positioning and navigation, pages 75–153. Academic Press, Oxford, 1 edition 2012.

ERSAT-EAV H2020 Galileo Call 2014, funded by European Commission with contract number 640747 2015-2017.

Goya, J., Zamora-Cadenas, L., Arrizabalaga, S., Braz´alez, A., Meléndez, J. and Mendizabal, J. "Advanced Train Location Simulator (ATLAS) for developing, testing and validating on-board railway location systems," European Transport Research Review, vol. 7, no. 3, p. 24, 2015. [Online]. Available: http://link.springer.com/10.1007/s12544-015-0173-5.

Guowei Shen, Zetik, R. and Thoma, R.S. "Performance comparison of TOA and TDOA based location estimation algorithms in LOS environment," in Positioning, Navigation and Communication, 2008. WPNC 2008. 5th Workshop On, 2008, pp. 71–78.

Kay, S.M. Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice Hall, 1993.

Lakmali, B.D.S., Wijesinghe, W.H.M.P., De Silva, K.U.M., Liyanagama, K.G., Dias, S.A.D. Design, Implementation & Testing of Positioning Techniques in Mobile Networks, Information and Automation for Sustainability, 2007. ICIAFS 2007. Third International Conference on 4–6 Dec. 2007.

Welch, G. and Bishop, G. "An introduction to the kalman filter." 2006.

Wang, H. and Hou, H. "Experimental analysis of beamforming in high-speed railway communication," in Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium On, Sept. 2011, pp. 745–749.

Yuechun Chu and Ganz, A. "A UWB-based 3D location system for indoor environments," in Broadband Networks, 2005. BroadNets 2005. 2nd International Conference On, 2005, pp. 1147–1155 Vol. 2.

tecnun
Universidad
de Navarra

Pº Manuel Lardizabal, 13.
20018 Donostia-San Sebastián, Spain
Tel.943 219 877
Fax 943 311 442
www.tecnun.es