DM

Low-cost Internet of Things and Snapshot Geolocation Pipeline in Marine Sensing
MASTER DISSERTATION

Pedro José Gouveia Pinto
MASTER IN INFORMATICS ENGINEERING

UNIVERSIDADE da MADEIRA
A Nossa Universidade
www.uma.pt

February | 2022

Low-cost Internet of Things and Snapshot Geolocation Pipeline in Marine Sensing

MASTER DISSERTATION

Pedro José Gouveia Pinto
MASTER IN INFORMATICS ENGINEERING

ORIENTATION
Marko Radeta

CO-ORIENTATION
Filipe Magno Gouveia Quintal

Low-cost Internet of Things and Snapshot Geolocation Pipeline in Marine Sensing

Pedro Pinto

supervisionado por

Prof. Dr. Marko Radeta and Prof. Dr. Filipe Quintal

3 de maio de 2022

Abstract

Biologging and biotelemetry are commonly used as methods to assess marine biodiversity population. However, current state-of-the-art devices (commonly referred to as tags) remain at the greater cost of production while geolocation and georeferencing methods use proprietary satellite constellations, remain expensive and are prone to greater battery usage. This dissertation enhances such state-of-the-art devices, providing affordable tags for multipurpose usage. Dissertation contribution is two-fold. In first, it describes the design of low-cost telecommunication system comprised from tag emitters and land receivers, evaluated during the sea-vessel field trips in pelagic area of Madeira island. In second, it also describes the software pipeline for deducing the position of tags, leveraging the raw signal from obtained GPS receivers.

Keywords: Ubiquitous Computing · Internet of Things (IoT) · Marine Sensing · Long-Range protocol (LoRa) · Fastloc GPS.

Resumo

Bio-logging e biotelemetria são métodos de grande importância como métodos de avaliação da biodiversidade marítima. No entanto, os dispositivos atuais normalmente referidos por tags permanecem com um elevado custo de produção e, são suscetíveis a elevado consumo de energia. Esta dissertação procura melhorar o bio-logging e a biotelemetria para a estimativa da biodiversidade marítima, com três contribuições distintas: (i) realizar análise em detalhe de sistemas de última geração de bio-logging e de biotelemetria, (ii) desenvolver um sistema inovador usando Internet of things (IoT) e Long Range (LoRa), e (iii) melhorar o sistema fastloc com computação no CPU da tag, para estimar a posição de mamíferos marítimos na superfície do mar. O príncipal objetivo é reduzir o custo de tais sistemas de detecção, explorando o IoT, LoRa e fastloc na cricação de bio-loggers e sistemas de biotelemetria.

Keywords: Ubiquitous Computing · Internet of Things (IoT) · Marine Sensing · Long-Range protocol (LoRa) · Fastloc GPS.

Acknowledgements

---

[1] http://wave-labs.org

Table of Contents

List of Figures

List of Tables

1 Introduction

Ways to obtain the location of important locations has always been a concern throughout humanity progress, visible in many events in our history. From Indians when they used smoke signals, seafaring navigators using the angles of the stars and sun progressing, up until our modern days using satellites along with position providing services. Currently, it is being employed in multiple areas ranging from automated vehicles (hi-tech mechanised field) to nature and animal study (natural field). It quickly became part of the core of our society, not only as an assistant tool to reach a specific location but also in improving several fields responsible for the quality of life (i.e. tectonic plate observation, fishing boats control, package delivery location). This dissertation explores the usage of biologging and biotelemetry in marine science, towards remotely monitoring significant information by means of electronic tags attached to marine species [5]. Current state-of-the-art technology is supported by satellite tracking, providing information about large-scale movements of marine species. Although Fastloc Global Positioning Systems (GPS)[2] have solved most of the limitations inherited from such solutions [12], it remains at the high cost - e.g. Wildlife Computing SPLASH10-F-333 transmitting tag is of the cost of appx. EUR 2000 including additional ARGOS monthly subscription fee (appx. EUR 100), seen in Figure 1. The aforementioned tag is a Low Impact Minimally Percutaneous Electronic Transmitter (LIMPET), being a satellite transmitting tag widely used for cetacean tracking. LIMPET tags are designed to be deployed using Dan-Inject CO2 rifles, which remain obtrusive and unethical. Being small in size and weight (56x50x27, 69g), they can be deployed high on the dorsal fin enabling frequent transmissions to the Argos satellites. Such tags can obtain geolocation snapshots, using ARGOS satellite constellation, depth, temperature, with appx. 112 days of autonomy. Further, Fastloc technology remains close-sourced, therefore lessening research efforts. Commercial off-the-shelf (COTS) tags, as the one mentioned previously, are used to track marine mammals, providing opportunities for ecologists to monitor behavioral, physiological, and environmental information. However, they remain at a great cost, are obtrusive to the animals while the technology which allows the fast detection of the animals on sea-surface remains proprietary.
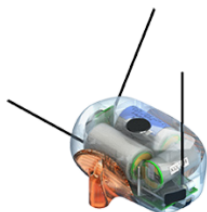
Fig. 1: Commercial Off-the-shelf (COTS) tag - SPLASH10-F-333.

---

[2]e.g. `http://www.wildtracker.com/Homex.html`

The dissertation contributes by leveraging the Internet of Things (IoT) in reducing the cost of production of COTS tags, including the pipeline to estimate the position of the tag from raw satellite data. Below, there will be a clarification about the terminology and georeferencing methods, including the procedures of how the positions are usually obtained, current satellite providers, as well as important terms which will be used for navigation of terrestrial objects using satellite constellations.

1.1 Used Geolocation Vocabulary

In below, will be provided a brief description of used terminology regarding the geolocation estimation and trilateration, introducing the reader to the topic. Paragraphs outline existing satellite positioning providers, GNSS, ephemeris, two-line element (TLE), space-track organization and almanac, which will be further used in the dissertation.

1.1.1 Satellite Positioning Providers

Satellite positioning providers are typically Satellite networks (from now referred to as constellations) that allow attaining the receiver's positioning worldwide, from the 31 Global Positioning System (GPS) satellites[3]. GPS broadcasts in two distinct bands, namely L1 at 1575.42MHz and L2 at 1227.60MHz, however, civilians used to be restricted to direct L1 measurements. This introduced limitations in the accuracy of such measurements because receivers were unable to correct for delays to the signal being caused by the ionosphere [18]. Measurements on both bands enable accuracy improvements [1], however, it was restricted for US military and certain authorized agencies, meaning that aforementioned services could influence the system's public signal at will, which is commonly referred to as *Selective Availability (SA)*[4]. Currently, SA has been deactivated, but there still are some constraints for civilian and unauthorized users, emerging the question of how to obtain accurate position measurements where there are not enough satellites. Global Navigation Satellite Systems (GNSS) were developed towards these problems.

1.1.2 GNSS

GNSS stands for a global network of satellite-based position systems, including ranging signals (e.g. calculating the distance to the satellites) and navigation messages. The navigation messages comprise of *ephemeris* data used to calculate the position of each satellite in orbit and *almanac* information regarding the time and status of the entire constellation. As previously mentioned, the GPS was developed by the US military, however additional countries provided similar efforts,

---

[3]https://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/gnss/gps/howitworks

[4]https://www.gps.gov/systems/gps/modernization/sa/

resulting in more open satellite-based position systems. Highlighting the Global Navigation Satellite System (GLONASS) from Russia, Galileo from the European Union, BeiDou from China, and the Quasi-Zenith Satellite System (QZSS) from Japan. Table 4 contains information regarding the most commonly known GNSS Providers. Additionally, constellations are supported by space-based augmentation systems (SBAS) and ground-based augmentation systems (GBAS) which assist existing infrastructure with geostationary (GEO) or geosynchronous satellites [10]. This dissertation will explore the usage of two GNSS receivers, with the former being for GPS only and the latter being for multiple constellations.

Table 1: Global satellite position providers (GNSS)

System	Constellation	Nº Satellites	Frequency Range
Galileo	Walker 24/3/1	24	E1-I, E1-Q, E5a, E5b, E6-I, and E6-Q
GPS	Expandable 24	31	L1, L2, L3, L4, and L5
GLONASS	GLONASS- M/K1	24	L1OF, L1SF, L2SF, and L3OC
BeiDou	BeiDou- 1/2/3	35	E1, E2, E5B, and E6
QZSS	QZSS- 1/2/3/4	4	L1, L1C/A, L1C, L2C, L5C, and E6

1.1.3   Ephemeris and Almanac

The term Ephemeris is often used in astronomy and astronavigation, and it refers to a table that consists of the trajectory of naturally occurring astronomical objects and artificial satellites. The use of the Ephemeris dates back to the Babylonian civilization[5] where there were calculations on the motion of Jupiter being conceived, resulting in what can be considered a primitive Ephemeris. Nowadays, modern Ephemerides are created by combining the progress made in computing technology and the theory behind the motion of celestial bodies[6]. The Ephemeris that will be used consists primarily of the trajectory of the artificial satellites. These Ephemerides can typically be obtained either directly by decoding the broadcast Ephemeris or through alternative means via a communication network, e.g. Assisted GNSS (AGNSS)[7]. While the Ephemeris is required by the GPS receiver to decode its position, it requires one from each one of at least four satellites it has established a connection (minimum required to obtain position). Also, it further requires the Almanac corresponding to the specific constellation it is communicating with to obtain status of the entire satellites and their general position. This is obtained faster than the Ephemeris due to its smaller size, since it only contains coarse information about the satellite constellation, which is

---

[5]https://www.discovermagazine.com/the-sciences/ancient-babylonian-astronomers-were-way-ahead-of-their-time

[6]https://www.britannica.com/science/ephemeris

[7]Further explanation on GPS / A-GPS Technology, https://www.rohde-schwarz.com/us/technologies/satellite-navigation/gps-a-gps/gps-a-gps-technology-/gps___a_gps_technology_55418.html

accurate enough for a receiver to generate a list of visible satellites and to calculate its positioning faster.

1.1.4 Two-Line Element set (TLE)

A two-line element set (TLE) is a data format for Earth-orbiting elements, providing them in a specific moment in time, commonly referred as an Epoch time. These are used with Simplified Perturbations Models (SGP, SGP4, SDP4, SGP8, and SDP8) allowing to calculated the element's position and velocity vectors, in a given time (in the epoch time system). A TLE message is decoded following the dictionary on Table 2 and it is structured in two (2) lines and sixty-nine (69) columns (lines 1 and 2), for clearer understanding of the corresponding satellite and a easier visualization of the data when exposed on a list, it is often preceded by a line of twenty-four (24) identifying the satellite (line 0). Such messages are used as external data in this study, allowing the calculation of the position of the satellites. An example of TLE provider is seen with Space-track [8]. It is a civil service deployed by the Unite States Space Surveillance Network that allows the request of TLE elements of all the tracked space objects through API request. These objects are all identified and go from active and inactive satellites to space debris. The TLE used in this study were all obtained from this API allowing the use of a reliable location for the satellites in the algorithms.

Simplified Perturbations Models are a set of five mathematical models that allow the calculation of a space object, by predicting the effect of perturbations caused by the Earth's shape, drag, radiation, and gravitation effects from other celestial bodies. This set of models is often referred to collectively as SGP4 due to the frequency of use of that model particularly with two-line element sets produced by NORAD and NASA. SGP4 will be the model used in the algorithms presented in this thesis.

1.1.5 One Socket Protocol (OSP)

One socket Protocol is the designation for the mode the device receives the messages from the satellite. Some devices are capable of using several modes but never simultaneously. OSP was developed for $SiRFstarIV^{TM}$ a gps architecture developed by Qualcomm. All the devices used in this study were used in $SiRFstarIV^{TM}$ OSP mode[9].

---

[8] https://www.space-track.org/

[9] https://mt-system.ru/sites/default/files/docs/documents/sim18%20module%20osp%20manual%20(cs-129291-dc-8)%5B1%5D.pdf

Table 2: TLE format and decoding.

Line	Columns	Content
0	01–24	Satellite name
1	01	Line number
	03–07	Satellite catalog number
	08	Classification (U: unclassified, C: classified, S: secret)
	10–11	International Designator (last two digits of launch year)
	12–14	International Designator (launch number of the year)
	15–17	International Designator (piece of the launch)
	19–20	Epoch year (last two digits of year)
	21–32	Epoch (day of the year and fractional portion of the day)
	34–43	First derivative of mean motion; the ballistic coefficient
	45–52	Second derivative of mean motion (decimal point assumed)
	54–61	B*, the drag term, or radiation pressure coefficient (decimal point assumed)
	63–63	Ephemeris type (always zero; only used in undistributed TLE data)
	65–68	Element set number, incremented when a new TLE is generated for this object
	69	Checksum (modulo 10)
2	01	Line number
	03–07	Satellite Catalog number
	09–16	Inclination (degrees)
	18–25	Right ascension of the ascending node (degrees)
	27–33	Eccentricity (decimal point assumed)
	35–42	Argument of perigee (degrees)
	44–51	Mean anomaly (degrees)
	53–63	Mean motion (revolutions per day)
	64–68	Revolution number at epoch (revolutions)
	69	Checksum (modulo 10)

1.2 Ranging

Location estimation will be conducted using errors in estimation, pseudoranges and trilateration. Each one is briefly addressed below.

1.2.1 Errors in Estimation

Ionospheric Delay. The ionosphere is a zone in the Earth atmosphere that extends from a height around 60 kilometers up to 2000 kilometers high. Its components make it so the propagation speed varies depending on the electron density, which is higher during the day due to the X and UV rays radiation. In this study, such delay will be solved through the use of the raw data received from the satellite.

Clock Bias. Satellite clock errors are prone to cause additional errors in GNSS measurements. Such errors are common to all receivers observing the same satellite and can be removed through further differentiation between the receivers. When applying the satellite clock correction in the navigation message, such leads to the satellite clock errors. Similarly to the ionospheric delay, such error will be also solved through the usage of raw satellite data.

Tropospheric Delay. The troposhpere is located in between the Earth surface and around 60 kilometers high. This zone influences the signal with a multitude of variables like temperature, pressure and humidity. Due to the volatile nature of this delay (easily affected by the weather conditions on the areas the signal travels through) it is extremely difficult to obtain a reliable enough value to implement in this study, therefore the solution implemented was based around using reference points which is a reliable alternative however requiring more than one device.

1.2.2 Pseudoranges

Pseudorange is the distance obtained by the GNSS receivers (i.e. biotelemetry tags) from the gps signal, this distance requires compensation for several errors (clock bias, Ionospheric Delay, Tropospheric Delay...). It can be an underestimation or an overestimation of the real range, as depicted in fig. 2 pseudorange 1 and 2, respectively. To determine their position, GNSS receiver typically determine the range of at least four satellites as well as their positions at time of transmitting. Such ranges can be calculated for any point in time, and are obtained multiplying the speed of light by the time the signal has taken from the satellite to the receiver. Pseudoranges are prone to errors that will corrected when used for trilateration, briefly addressed below and depicted further in Section 4.

Fig. 2: Illustration of the pseudorange possibilities.

1.2.3 Trilateration

Trilateration is the technique used by global positioning systems to calculate the object's position using the distance from the receiver to the satellites. To provide a functional global positioning system, the GNSS constellations are designed and deployed to fulfill the minimum requirement of three satellites being simultaneously visible at every position. However, while three satellites are the least amount used for calculation, an accurate measurement is ensured using at least four satellites. Three satellites are enough for two-dimensional calculations, however in reality we require three-dimensional measurements because GPS satellites broadcast signals as a sphere, which could possibly originate multiple points after calculating the intersection each with different heights, this is then filtered with the radius of the fourth satellite. If only three satellites are available, Earth's position and radius is used to simulate the fourth one and therefore obtain a location on a three-dimensional space. Following image 3 demonstrates a two-dimensional example for trilateration, measuring the intersection of distances with three different satellites (to not be confused with triangulation which measures angles). Triangulation is the process of two or more receiving antennas extracting directions from the received signals. The accuracy of measurements increases as more data is available, which requires more antennas that end up increasing the cost of the system. This implies that it is necessary to find a balance between cost and performance when choosing how to use such a method. Trilateration will be used during the creation of a snapshot pipeline in the forthcoming dissertation text.

1.3 Used Coordinate Systems

Several coordinate systems are used throughout the dissertation: Geographic Coordinate System (GCS), Earth-Centered Inertial (ECI), Earth-centered, Earth-fixed (ECEF), where the reader is briefly introduced in all of such.

(a) GPS satellites (b) One satellite (c) Two satellites (d) Three satellites
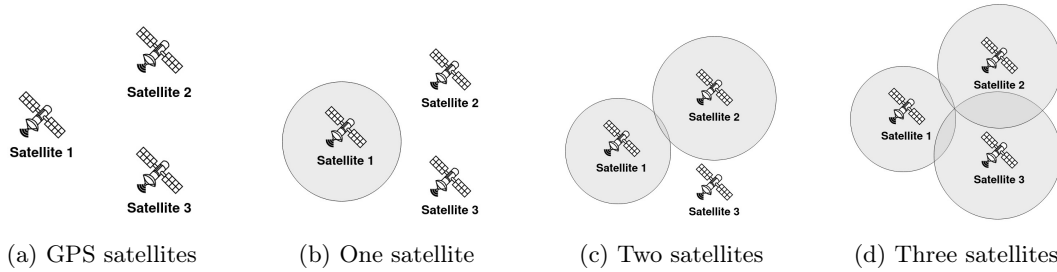
Fig. 3: Steps to obtain an accurate position using three satellites in a two-dimensional space where intersection indicates the correct location. Images (left to right): (a) Initial satellites with known positions; (b) Message from satellite one, where the position can be anything inside the radius of the circle; (c) Intersection between satellites one and two suggests two points as the possible measured location; (d) Intersection between all satellites pinpoints a precise location.

1.3.1   Geographic coordinate system (GCS)

GCS is a ellipsoidal coordinate system that measures positions directly on the Earth, through latitude (used for North and South), longitude (used for East and West) and altitude. Both latitude and longitude are calculated based on angles while altitude is measured by the distance from the surface of the Earth. Latitude is the angle between the equatorial plane and the line that passes through the designated point and the center of the Earth, while longitude is the angle of the meridian that passes through the specified location that is East or West of a reference meridian, usually the prime meridian (0° longitude) is used.

1.3.2   Earth-centered, Earth-fixed coordinate system (ECEF)

The Earth-centered, Earth-fixed coordinate system is a cartesian spatial reference system for representing positions in the proximity of the Earth. Global Positioning System uses an ECEF that is designated as World Geodetic System (WGS 84), for its parameters it uses the center of mass of the Earth as its origin. The Z axis is the line betwen the North and South Poles. The X axis is in the plane of the equator, passing through the origin and extends from 180° longitude (negative) to the prime meridian (positive). While the Y axis also is in the plane of the equator and passes through the origin it instead extends from 90°W longitude (negative) to 90°E longitude (positive). Visible in Fig. 4.

1.3.3   Earth-Centered Inertial coordinate system (ECI)

Earth Centered inertial is the designation for the coordinate frames that use the center of mass of the Earth as its origin where its axis are fixed according to the stars. Unlike ECEF, this system axis do not rotate with the earth, exemplified in Fig. 5. An ECI often used together with TLE, as well as the one being used in this study, is the true equator, mean equinox (TEME) coordinate system.
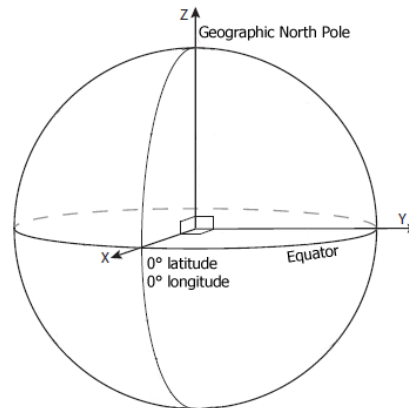
Fig. 4: Illustration of the ECEF coordinate system.
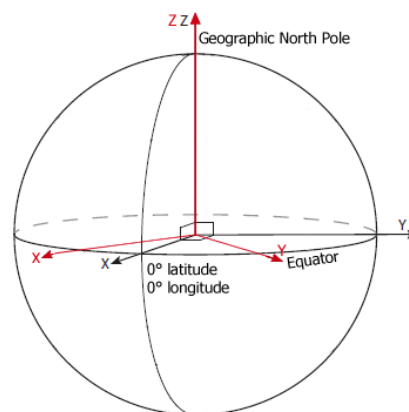
Fig. 5: Illustration of the ECI coordinate system - black axis is represents the earth axis and the red axis represents a posible position for the ECI axis.

1.4 Objectives and Research Questions

The core objectives of this dissertation are to prepare IoT devices for LoRa telemetry and enable the georeferencing from the calculation on onboard CPU, allowing the IoT device to obtain the latitude and longitude as snapshot receivers. Conversely, the main hypothesis of the research is that it is possible to advance the state-of-the-art in biotelemetry, applying snapshot receivers, IoT and LoRa, allowing the computation of location estimation to be performed on local CPU. Two major research questions are provided:

– **[RQ1] - How to provide more affordable marine sensors?**
  Through diverse Internet of Things (IoT) modules, the dissertation creates land receivers and off-shore sensors, providing the experiments in tracking one marine vessel.

– **[RQ2] - How to create an accessible snapshot pipeline for GNSS receivers?**
  Through diverse algorithms, the dissertation implements the pipeline and studies the performance in georeferencing the tags using raw satellite coordinates, obtained in a controlled environment.

1.5 Dissertation Structure

In Section 1, the dissertation outlined the problem of the high cost of COTS devices applied in marine sensing and provided some of the key terminologies which are used throughout the reported study. In forthcoming, current efforts and pitfalls of IoT devices in marine sensing are presented (Section 2). Next, IoT device for emitters and receivers (Section 3) and snapshot pipeline (Section 4) is presented. Conversely, carried out in-the-wild study of IoT device is described (Section 5), and experimental setup for using the snapshot geolocation pipeline is depicted (Section 6). Finally, results from both the IoT device and the pipeline are discussed where the dissertation states the obtained contributions, addressing aforementioned research questions (Section 7).

2 Related Work

The literature review describes the current biotelemetry (i.e. transmitting data from marine species) and biologging (i.e. storing data on marine species) systems, and the effort of IoT and LoRa in approaching such endeavor. Furthermore, an outline of the snapshot receiver is presented.

In general, there are three types of distinct systems used in biotelemetry (radio-telemetry): VHF radio tracking, satellite tracking, and GPS [7]. Table 3 summarizes aforementioned technologies and compares them with biologging tags, which are currently employed in monitoring the wildlife. It is possible to conclude that each system presents a correlation between the level of spatial detail and the cost. This factor plus the restrictions that come into play when considering the weight and lifespan of the device, end up reducing the possibilities available for the experts when studying a specific range of species.

Table 3: Comparison of different technologies used in biologging and biotelemetry.

| | Biologging | Biotelemetry | | |
	Archival Tags	VHF	Argos PTT	GPS
Power Source	Batteries	Batteries / Solar Cell	Batteries / Solar Cell	Batteries/ Solar Cell
Weight (g)	0.3 to 3	0.2 to 100	2 to 50 / 45 to 105	17 to 50
Lifespan	5 days to 2 years	Few days to 4 years	2 to 3 years / 40 days to 3 years	up to 3 years
Range (km)	N/A	5 to 25	Global	Global
Location Method	GLS	Triangulation / Homing	Doppler Effect	GPS
Positioning Accuracy	Low	Medium	High	Very High
Price (USD)	12 to 600	180 to 300	2900 to 4450 / 2550 to 2950	2000 to 8000

Radio-tracking uses a radio signal to provide information about the marine taxa to marine biologists. This tracking system represents any kind of communication system which allows transmitting the data using sender (emitter) and receiver. The emitter systems consist of: (i) a radio transmitter, (ii) a power source, (iii) and a propagating antenna. On another hand, a receiver system includes a power source, a receiving antenna, and a signal receiver with a reception indicator. These two parts are commonly used in-situ validation, whereby adjusting the transmitters

to different frequencies it is possible to achieve a proper link identification between the emitter and receiver. It is also important to reflect on which sender-receiver system is best suited for the experiment to occur [4]. Thomas et al. [20] follows three criteria that can lead an expert to the best choice available allowing to achieve adequate communication: (i) specification of the data that is required for the project; (ii) understanding the constraints imposed by the species and locations of the study; (iii) deduction of the cost of the various tracking methods available. Based on such constraints, it is possible to evaluate which of the technologies to use to obtain an acceptable basis for the proposed solution in this study.

In regards to the range specification, signals located in a band between 30 to 300 MHz VHF transmitters emit radio-frequency which is received by an antenna with a ground-to-ground range of 5-10 km and an air-to-ground range of 15-25 km [14]. VHF enables animal tracking using two main methods, homing and triangulation. As explained in [13], the former technique involves following a signal towards its origin. This assumes that experts are deployed, and using the receiver to track the signal. As experts are closing in on the transmitter, which correlates to an increase of signal strength, the receiver gain is gradually reduced to get a better orientation towards the signal. This is a process that requires undergoing constant iterations until the subject of the experiment is located. Triangulation is the process of determining the location of a point by measuring only angles relative to known points at either end of a fixed baseline. The position point is then fixed as the third point of a triangle with the usage of the previously obtained angles and the side defined by the baseline edges.

Conversely, VHF reports a margin of error between 200-600 m when estimating the location through triangulation and homing, and has proven to be effective in the studies of species with low movement with few acquisition costs and transmitters prices ranging from US$ 180 to US$ 300 [19]. In terms of battery autonomy, these systems rely on Lithium batteries [14] which have a longer lifetime or use a combination of solar energy technology plus rechargeable batteries ensuring a 24 hours signal output, at least until the other components start to malfunction. The power source is a critical planning point to take into consideration, as the options that have a longer lifetime can sometimes be heavier and therefore, will result in a device's total weight which ends up being an ethical concern among scholars. This weight should be properly assessed while being cautious of the 3-5% rule related to the transmitter weight [3] (the transmitter should be at most 5% of the animal weight). VHF transmitters' weight can be as little as 0.2 g [15] and have a lifespan of 18-22 days or as much as 100 g with a corresponding lifespan of up to 4 years[10].

---

[10]https://atstrack.com/tracking-products/transmitters/product-transmitters.aspx?serie=A1500

VHF is a very effective way of tracking animals that do not have migratory tendencies, in this case, if the goal is real-time monitoring is better to use Satellite tracking [21], although this implies a considerable increase in the system final cost.

The satellite-based system Argos was implemented in 1978 [2] and has global coverage through three subsystems: the Platform Transmitter Terminals (PTT), the space segment, and the ground segment. PTT are attached to animals to transmit radio signals in the ultra-high frequency band which is visible to the Argos satellites. These satellites cover 100% of the earth's surface with a visibility diameter of 5000 km [11]. When a transmitter enters the range of a satellite, it has approximately a window of 10 to 12 minutes to send the frequency data and the required timestamps (Doppler effect)[12]. The data is then downlinked to the Argos processing centers, where the location is obtained. This information is assessed through a least-squares method analysis and assigned to one of the several location classes that represent the range in which the position was estimated. David Nicholls mentions [16] that Argos satellites can locate any transmitter signal between 1 to 14 times per day with a high position accuracy with a margin of error of 1 km. However, there is a different deviation error in latitudinal and longitudinal axes depending on which location class[13] is used. This is claimed to be 250 m for location class 3, between 250 m and 500 m for location class 2, between 500 m and 1500 m for location class 1, and more than 1500 m for location class 0[14].These different precisions come with different services provided by Argos.

GPS tracking devices can provide precise location data about a marine species, by accessing transmissions from the total 24 satellites, that integrate the satellite constellation. When at least four satellites are in the range of the transmitter, GPS can provide a location with a margin of error below 30 m using trilateration. The data containing the location is transferred by associating the GPS with a datalink method such as Argos, Global System for mobile, VHF, or by simply storing the information on board for later study.

When used in the ecology field GPS telemetry systems have major cost implications compared to the other systems. A single tracking device can range from around 2000 to 8000 US$ depending on the features required for the deployment [9]. The expensive costs per unit of GPS have directly affected the sample sizes of marine taxa used in studies. GPS devices have issues with the lifespan of the device since GPS is a power-intensive location method. To minimize this problem many devices have programmable duty-cycles that collect a few locations per day and proceed to sleep, reducing

---

[11]http://www.argos-system.org

[12]when the distance between a satellite and a transmitter shorten, the frequency of which the transmitted signal is measured by the onboard receiver is higher than the normal transmission frequency, and lower when the distance becomes longer

[13]From the beginning of Argos service, locations have been classified according to the following criteria: (i) type of location (Argos or GPS); (ii) estimated error; and (iii) number of messages received during the pass.

[14]www.argos-system.org/manual/

the power consumption, therefore augmenting its lifespan. These devices can be supported by either batteries or solar cells, however, that depends on the environment in which it was deployed, in which solar cells may turn out to be non-viable options [20].

While GNSS is capable to obtain coordinates from typically 3-4 satellites, there are several methods how such coordinates are obtained: using cold-start, hot-start, or Fastloc. While the former two are adequate for terrestrial navigation, they lack in estimating the GPS coordinates for succinct moments (e.g. when a marine animal is above the surface for less than a second). For this reason, Fastloc receivers (snapshot) are used to obtain raw data from satellites, for later calculating the position internally. For instance, one recent study (Eichelberger et al. 2019) described the Fastloc receiver setup, which can obtain georeferencing from all known constellations [8]. While such technology typically remains proprietary in for-profit biotelemetry companies, there are no studies on how snapshot receivers perform on IoT devices used for marine sensing.

Indeed, long-range (LoRa) protocol has the opportunity to be used for marine sensing, reaching 84 km when tracking marine objects at the sea surface [17]. Such implies the need for a traditional emitter-receiver system which will be explored throughout this dissertation. In the remainder, this dissertation will provide the IoT sensor, describing the pipeline for snapshot receivers, leveraging LoRa.

Existing Snapshot GNSS Receivers. Previously described satellite systems' receivers generally take several seconds to lock with required satellites and therefore estimate a position. Although it may not be identified as a problem in most cases when tracking marine taxa that are only briefly visible when surface, it's not possible to communicate with the satellites in such a time frame. Fastloc-GPS is a major breakthrough for such marine species, overcoming the aforementioned problem, through rapid (milliseconds) acquisition of GPS data. Receivers developed for Fastloc-GPS manage to grab a snapshot of the radio signals for post-process and post-compress them onboard the tag, thus reducing time-costs in communication. Such methodology means that processing and compression continue after the animal has dived and have no impact on communication with GPS systems. Furthermore, Fastloc-GPS accuracy is substantially better than conventional Argos tracking or light-based geolocation providing new insights into small-scale movement patterns of marine taxa [6]. This dissertation explores the creation of an own snapshot pipeline, analyzing the raw GNSS receiver values.

3 Low-cost IoT Sensor Design

As to uphold the commitment of developing a low cost device the LoPy4 board was used (encompassing ESP32 CPU and LoRa transceiver chip). The system consisted of the emitters (hereinafter tags) and receivers (hereinafter gateways), both of which included storage units (SD cards), allowing its data to be stored locally for later retrieval in case the gateways could not connect to the internet due to some difficulty, as well as to verify if the tags correctly sent all the data they obtained. Also, waterproofing boxes were used, securing the devices from heat and humidity. In further, tags were assisted further with the Global Positioning System (GPS) sensor, Inertial Measurement Unit (IMU), custom-made LoRa antenna, ADC regulator, and buoy. Further rationale for using LoPy4 in our experiments was motivated by the fact that it has been already used to estimate the location of sea-vessels [17]. Below, we describe each component individually with more details.

3.1 Microcontroller Unit: LoPy4

This board was used to reduce the cost of the current state-of-the-art marine bio-logging devices. This board was developed by Pycom, containing Espressif ESP32 chipset, providing enough computing power required for the data collection. It also includes a Semtech LoRa transceiver $SX1276$ for using LoRa protocol, and a wifi module that allows the usage of external APIs to send the data into the Wave Labs database[15]. Base code for this MCU was further expanded from online sources[16]. In further, as the designed system required a constant writing of data and immediate response to the sensor it required multitasking which was achieved using protothreading which was further adopted from the online sources[17] , this method was chosen due to protothreads being less memory demanding than threads. The cost of the item is appx. EUR 40.

3.2 Flash Board Unit: Custom Board

While LoPy4 was originally designed for the usage of MicroPython programming language, several prior empirical tests suggested that MicroPython is not the most efficient for long-term deployments as it affects battery autonomy. For that reason, a custom board expansion board was created, allowing to upload of the C-written code directly onto the LoPy4 MCU, using Arduino IDE. In addition, the a FTDI module was added, allowing a hardware interface to flash the code between the device and the computer. The cost of the device is appx. 10 EUR encompassing the SD card.

---

[15] http://wave-labs.org
[16] https://github.com/espressif/arduino-esp32
[17] http://dunkels.com/adam/pt/index.html

(a) Front view
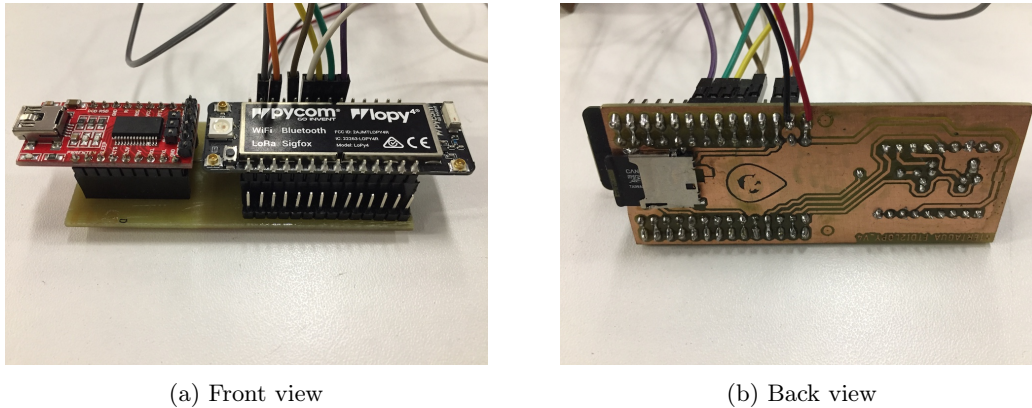
(b) Back view

Fig. 6: Expansion board (from left to right): (a) front view of expansion board containing the FTDI module and LoPy4; (b) back view of expansion board containing the SD card module.

(a) Lantronix GPS

(b) Wurth GNSS
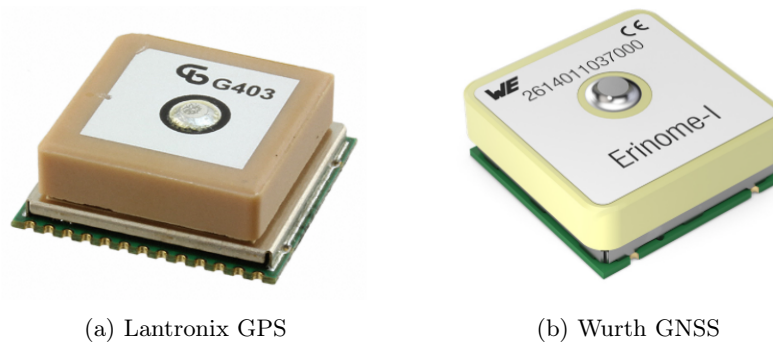
Fig. 7: Used GPS modules in experiments.

3.3 Geolocation Unit: GPS receiver

As for the GNSS receivers, two separate units were benchmarked. The former contained the Lantronix module A2235-H (in OSP mode). This module is a low-cost component, capable of obtaining the satellite pings only from the GPS constellations (not using the other constellations such as Beidu, Galileo, etc). This module, therefore, comes with more latency needed to obtain the signal. Code for the GPS reading was expanded from the online available code[18]. The cost of the device was appx. EUR 15. The latter GNSS receiver was Erinome-I by Würth which allowed other constellations, GPS, GLONASS, Galileo, BeiDou. The price of the component was appx. EUR 20. For both GNSS receivers, expansion boards were made, allowing the interface with the LoPy4. Examples of such sensors are seen in Figure 7. From both GNSS receivers the raw values were collected with satellite position coordinates, pseudo ranges, time of sending, and time of retrieval.

---

[18]https://github.com/SlashDevin/NeoGPS

3.4 Inertial Measurement Unit: IMU

The inertial measurement unit used on this project is the $9DoF$ sensor stick developed by Sparkfun, the sensor used is the $LSM9DS1$ with 9 different integrated sensors being: 3-axis gyroscope, 3-axis magnetometer, and 3-axis accelerometer. The collected sample rate of each was in $10Hz$. The calibration procedure for each sensor was done by collecting data samples using FIFO, 32 for accelerometer and gyroscope and 128 for magnetometer, averaging them and scaling them to the respective units gs for accelerometer, deg/s for gyroscope, and $Gauss$ for the magnetometer. This resulted in biases that are subtracted to make the measurements more accurate and remove the errors that exist when there are variations in the initial placement. Base code for the IMU was used and further expanded from the online sources[19]. The cost of the single item was appx. EUR 20.
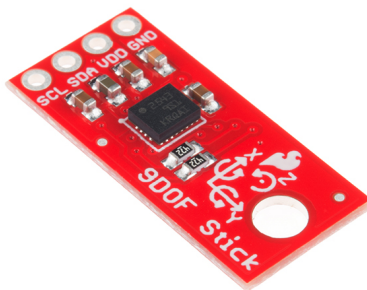
Fig. 8: Sparkfun 9DoF sensor stick

3.5 Memory Unit: SD Card

SD card used was the standard $32Gb$ by Kingston, formatted in Ex-FAT. It was envisioned that each tag and gateway may be retrieved, and therefore that it will be possible to assess the collected data. Moreover, IMU was stored on the SD card due to the required duty cycle which can not be transmitted through the LoRa. Base code for the SD card was expanded further from the online code[20]. The cost of the SD card was appx. EUR 10.

3.6 Telecommunication Unit: LoRa and Flexible Antenna

Although the dissertation addresses the experimental setup with tracking sea vessels, it was further constrained to be prepared with deployments for marine species, where 1/4 wavelength antenna

---

[19]https://github.com/sparkfun/SparkFun_LSM9DS1_Arduino_Library
[20]https://github.com/greiman/SdFat

size was used. A custom-based antenna was created from fishing wire made from stainless steel and nylon, being externally accessible and allowed to bend 180 degrees reducing the impact of waves. Base code for communicating with the LoRa chip was expanded from the online code[21]. The cost of a roll of fishing wire (for tag antenna) is appx. EUR 5, while regular LoRa antenna for gateways is appx EUR 20.

(a) RSSI

(b) SNR
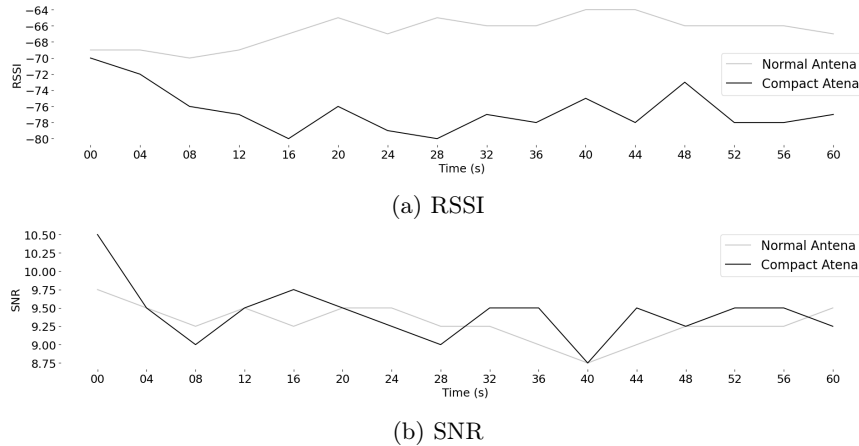
Fig. 9: Difference between 1/4 wavelength helical antenna (compact) and 1/4 wavelength lora antenna (normal) for RSSI and SNR.

Although changes in RSSI and SNR are present (Figures 9a and 9b), the results are still inside the acceptable range for deployment. From these, we can conclude that both RSSI and SNR are coherent for both normal and compact antennas, therefore the smallest form factor is considered acceptable for further study.

3.7   Capacitive Sensor: Water Detection

In addition to other modules, the electrical capacitance was tested with an additional open circuit module, which conductivity increases with the presence of salt. The rationale for using such a sensor was constrained for a future use case, detecting the sensor to be outside of the water, to start the data transmission. Several tests were carried out and presented below.

Capacitive sensor results (Figures 10a and 10b) indicate that there is an equilibrium in both methodologies (as it is possible to distinct when the object as surfaced in each), suggesting that it is possible to identify moments where there is a transition between surface and underwater settings. Further, such results were obtained using a low-cost circuit (< 1 EUR), providing state-of-the-art benefits towards low-cost marine sensing. Its application reduces power consumption costs, due to

[21]https://github.com/sandeepmistry/arduino-LoRa

(a) Polished

(b) Base
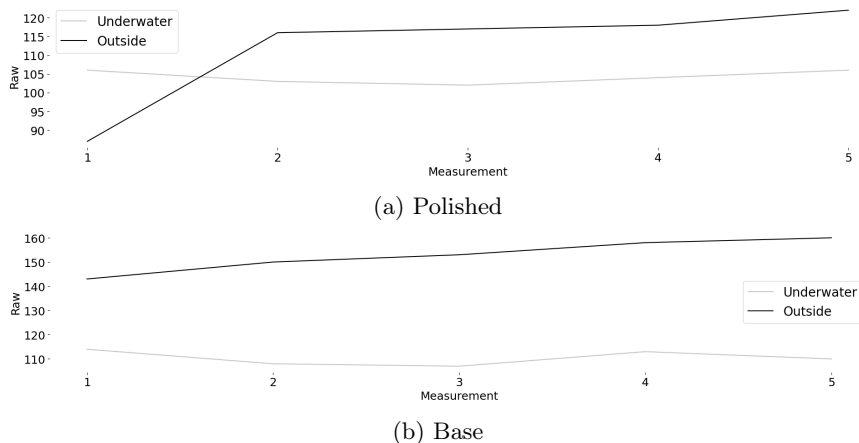
Fig. 10: Comparison between performance between capacitive sensors when they are factory new (base) and protected with an water protection polishing layer (polished).

establishing GNSS communication with satellites only when such communication is available on the surface.

3.8 Power Autonomy Unit: Battery

Since gateways were prone to heat exposure and increase in temperature, a Li-ion battery with 3.7 V 2500 mAh was used across all devices which are planned to be installed at the land-based locations. Conversely, all devices which were subject to sea deployment were powered by a Li-Po battery with 3.7 V 2600 mAh. The cost of an individual battery is appx. EUR 10. Both batteries were rechargeable, while the gateway battery was recharged with an additional 3.5W 6V 583mA Monocrystalline Mini Solar Panel Photovoltaic Panel, with appx. EUR 10 cost. In figure 11, it is possible to see that power consumption remains with all components and is stable during experiments' lifetime, without achieving extreme values which may invalidate experiments' results.

3.9 Encasing Unit: Waterproof Box

All aforementioned components were securely stored inside of the $Ip67$ containers (Figure 12a), capable of withstanding the severe weather condition. Such assures the longevity of the used components to be deployed for long-term periods at both sea-surface and land receivers. The cost of the item is appx. EUR 10.

3.10 Glider Unit: Buoy

Once all components were gathered, a waterproof box with all modules was placed onto the bodyboard (Figure 12b), securely connecting it with additional screws and bolts. Usage of this
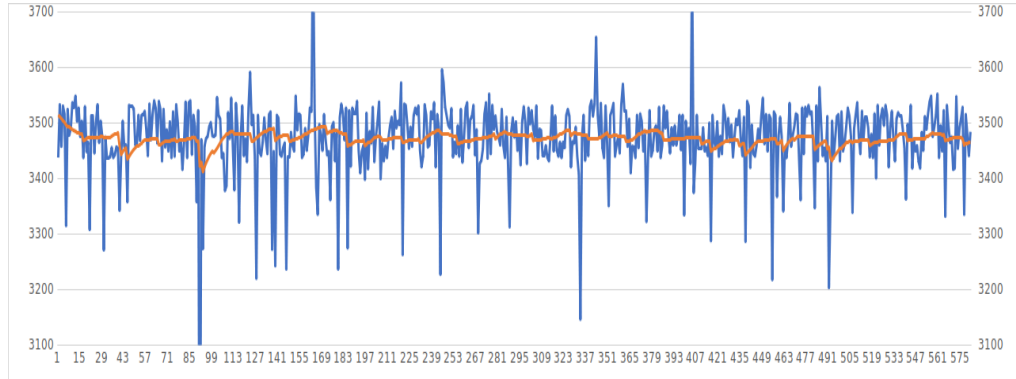
Fig. 11: Battery Status. X - minutes Y - milliamps

board served to provide the draggable sensor, mimicking the marine mammal at the sea surface, to be deployed from the sea vessels. The cost of the item is appx. EUR 10.

3.11 Obtained Tags and Gateways

All aforementioned components were stacked together into two separate IoT devices, being *tags* (emitters) and *gateways* (receivers), where one full setup of single tag and gateway final price costs appx. EUR 170. Both of them were placed into waterproof Ip67 boxes. Two tags were made for obtaining the signal from both the sea surface as well as from the sea-vessel mast. The rationale for creating both tags was to simulate the realistic position of marine species (e.g. turtle) at the sea surface, in this case, the bodyboard. The hypothesis was that the signal from the sea surface will be subject to loss of data points due to the presence of waves, and due to the loss of the line-of-sight when the sea vessel is between the bodyboard and the land receiver. The time of building each tag and gateway is appx. 5 days.

In further, two tags were made with the former being mounted on a purchased bodyboard, while being dragged from the sea vessel, simulating the marine animal (Figure 12c). The latter was mounted onto a sea-vessel mast (Figure 12d) ,taking in consideration the fact that the mast being metallic, allowing the sensing during the typical whale-watching activities. On land, 4 receivers were deployed along the coast of Madeira island, facing the south sea (Figure 13). Both the tags and receivers were storing the data onto the SD card, from which two receivers were connected to the Wireless Local Area Network (WLAN) through an access point created on the mobile phone, allowing them to work as gateways enabling data upload to the Wave Labs server[22].

Provided cost of a single tag and gateway are affordable (appx. EUR 170) and can be used in the operational environment. In the remainder, the dissertation will perform in-situ validation,

---

[22]http://wave-labs.org
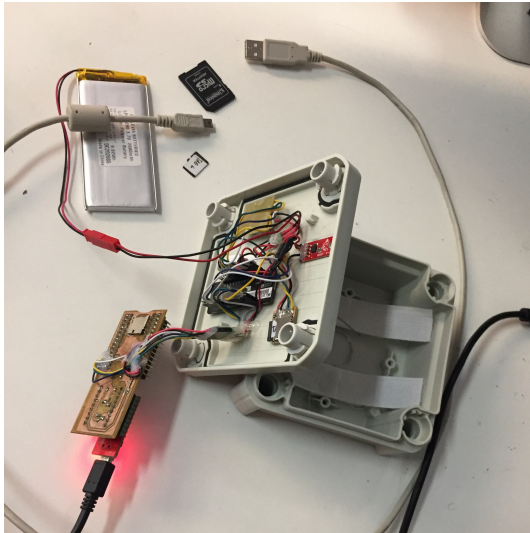
(a) Tag components
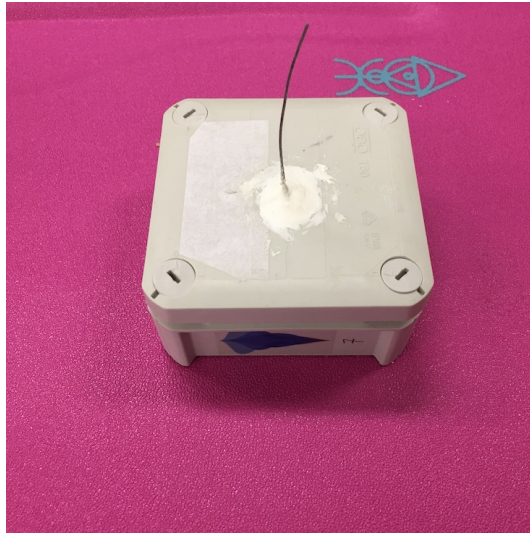
(b) Tag external
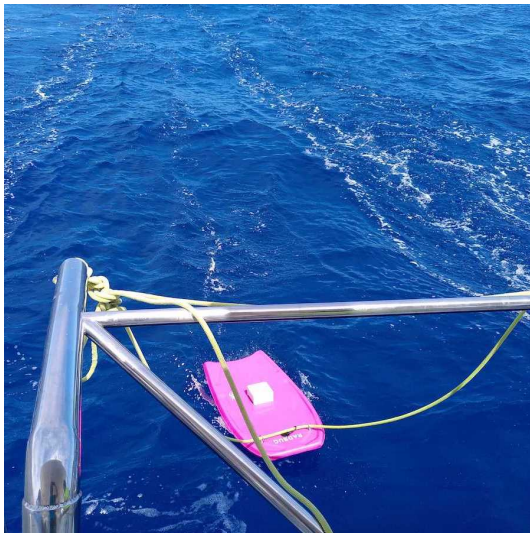
(c) Tag 6 (bodyboard)

(d) Tag 7 (mast)

Fig. 12: Used tags during the experiment. Images (left to right): (a) Ip67 waterproof case with integrated LoPy4, LiPo battery, SD card, GPS and IMU modules; (b) Custom made 1/4 wavelength antenna; (c) In-the-wild deployment on bodyboard; (d) In-the-wild deployment on seavessel mast.

(a) GW 8

(b) GW 9

(c) GW 10

(d) GW 11

Fig. 13: Installation of solar land receivers during the tests in Madeira island: (from left to right): (a) Quinta do Lorde, (b) Fajã dos Padres, (c) Palheiro Golf, and (d) Meliá Hotel.

while first describing the pipeline and set of procedures for processing raw GPS signal, including the used trilateration and optimization algorithms.

4 Snapshot Geolocalization Pipeline

Geolocalization devices when applied to the marine fauna are faced with many restrictions (i.e. size, battery lifetime, processing data). For the low budget, an IoT device proposed by this dissertation meeting such constraint, most of the data processing was handled on the server and not on the device itself , as the time frame when marine fauna surfaces is sometimes a matter of seconds (where the device needs to capture data and send any prepared LoRa messages) and the computation required to calculate the position is intensive, this allows it to be more responsive to it's sensors. This resulted in a faster time of data capture and a faster preparation of the LoRa message, ensuing more reliable data emission. In below, it will be explained in detail the steps required for the algorithm implemented in this study as well as an in-depth explanation of which steps are required for each method.

Rationale. Typically, each GNSS receiver computes the latitude and longitude on the device, which is prone to significant battery consumption. This results in significant time to lock the coordinate depending on which scenario of operation it stands. These scenarios can be commonly specified as cold or factory, warm or normal and hot or standby. Cold refers to when the device has inaccurate estimates of its position, velocity, the time or which satellites are visible, forcing it to search for all satellites and receiving the almanac when it acquires a signal, taking up to 12.5 minutes to obtain a complete almanac. If the device has estimates of the current time withing 20 seconds it is currently in a warm scenario, taking up to 30 seconds to obtain a correct location. When all the estimates are correct the device is in a hot scenario and it only takes up to 5 seconds to calculate the position. Taking in consideration that GPS signals do not work under water it is plausible to say the devices used to study marine fauna will be in a cold scenario, which is not enough to detect marine species which appear abruptly at the sea surface, or in a hot scenario that is extremely taxing on the battery reducing its lifetime. Initial rationale for the creation of snapshot pipeline was to allow the computation of the position to not be performed by the tag, but rather, on a server side, thus improving the battery lifetime. GPS raw signal payload can be compressed and stored into a LoRa message, allowing the gateways to retrieve the raw data from tags and send them to the local server, which performs the estimation of both the latitude and longitude, and ultimately depicts the data on an existing dashboard[23].

Message IDs. The method of location in this study makes use of the raw data from observable satellites which emit signals (i.e.) to the receiver tags. All raw GNSS data are transmitted from satellites in a compressed message, where each message has a different ID. Such serves to identify the group and type of content that such message contains. Throughout upcoming experiments, different messages will be used in several methods with the ultimate goal to depict the pros and

---

[23]http://wave-labs.org

cons of each method. Table 4 summarizes the message IDs which are collected from the typical raw GNSS signal. In forthcoming, several methods are showcased, where each of them calculates the position of the satellites in a different manner. Some will use exclusively raw data provided by the messages sent by the satellites while other will make use of external accessible data.

Table 4: Message ID structure.

Message ID	Description	Usage
7	Clock Status Data (Response to Poll)	Provides time measurements such as clock bias and drift. Used for error correction.
28	Navigation Library Measurement Data	Provides the Pseudorange and carrier phase used in trilateration, and GPS software time that will be used for error correction.
30	Navigation Library SV State Data	Provides the computed satellite position. Also contains velocity, ionospheric delay and clock bias for error correction.

4.1   Step 1: Parsing TLE

The precise satellite positions required are calculated through the use of a SGP4 perturbation model oriented library. Such outputs the position in a True Equinox, Mean Equinox (TEME) coordinate frame, which is a Earth Centered inertial frame (ECIF). This library requires a TLE containing the information about the queried satellite and a Julian date of the time of study.

4.2   Step 2: Pseudorange Retrieval

Message ID 28. The pseudorange used in this experiments is obtained through the use of message $ID28$ from the raw data sent from the satellite (Table 5), it goes by the name of *Navigation Library Measurement Data* in devices that can use SiRFstarIV™ which both devices used in these experiments are set to be used. It contains the pseudorange, the id of the satellite that sent the data, including the GPS software time. When used in the algorithm of the methods that will be explained in this thesis, the satellite ID allows a direct association of a satellite and its corresponding pseudorange, resulting in a more trustworthy data. Important to note is that such pseudorange does not contain ionospheric, tropospheric nor clock bias corrections.

Table 5: Navigation Library Measurement Data - Message ID 28.
Where U - Unsigned int, Dbl and D - Double and Sgl - Single

Name	Bytes	Unit
Message ID	1 U	
Channel	1 U	
Time Tag	4 U	ms
Satellite ID	1 U	
GPS Software Time	8 Dbl	sec
Pseudorange	8 Dbl	m
Carrier Frequency	4 Sgl	m/s
Carrier Phase	8 Dbl	m
Time in Track	2 U	ms
Sync Flags	1 D	
C/N0 1	1 U	dB-HZ
C/N0 2	1 U	dB-HZ
C/N0 3	1 U	dB-HZ
C/N0 4	1 U	dB-HZ
C/N0 5	1 U	dB-HZ
C/N0 6	1 U	dB-HZ
C/N0 7	1 U	dB-HZ
C/N0 8	1 U	dB-HZ
C/N0 9	1 U	dB-HZ
C/N0 10	1 U	dB-HZ
Delta Range Interval	2 U	ms
Mean Delta Range Time	2 U	ms
Extrapolation Time	2 Sgl	ms
Phase Error Count	1 U	
Low Power Count	1 U	

Message IDs 7 and 30. Some of the methods will require more than just the aforementioned message IDs. For such, it will make use of message $ID7$ and message $ID30$, designated by *Clock Status Data (Response to Poll)* and *Navigation Library SV State Data* respectively. Message $ID7$ reports the actual time of measurement in GPS time and computed clock bias and drift (Table 6), that will be used for pseudorange correction. Message $ID30$ contains the computed satellite position and velocity at the GPS time it is reported (Table 7). This will allow the development of a method that relies only on the data retrieved from the satellite for the location calculation, replacing the use of TLE.

Table 6: Clock Status Data (Response to Poll) - Message ID 7.
Where U - Unsigned int

Name	Bytes	Unit
Message ID	1 U	
Extended GPS Week	2 U	
GPS TOW	4 U	s
SVs	1 U	
Clock Drift	4 U	Hz
Clock Bias	4 U	ns
Estimated GPS Time	4 U	ms

Table 7: Navigation Library SV State Data - Message ID 30.
Where U - Unsigned int, Dbl and D - Double and Sgl - Single

Name	Bytes	Unit
Message ID	1 U	
Satellite ID	1 U	
GPS Time	8 Dbl	sec
Position X	8 Dbl	m
Position Y	8 Dbl	m
Position Z	8 Dbl	m
Velocity X	8 Dbl	m/s
Velocity Y	8 Dbl	m/s
Velocity Z	8 Dbl	m/s
Clock Bias	8 Dbl	sec
Clock Drift	84 Sgl	s/s
Ephemeris Flag	1 D	ms
Reserved	4 Sgl	
Reserved	4 Sgl	
Ionospheric Delay	4 Sgl	m

4.3 Step 3: Pseudorange Correction

Once the pseudoranges are retrieved, they are prone to further corrections. In below, several methods are depicted, using: (i) raw pseudorange for reference, (ii) external clock bias correction, (iii) reference method, (iv) satellite data, and (v) satellite data with reference.

4.3.1 No Correction - Raw pseudorange with Lantronix GPS

This method makes use of the Lantronix GPS (Figure 7a) in a one socket protocol (OSP), and records the messages received from the satellites. In these it is possible to obtain the raw pseudorange captured by the device without any correction. Such method will be used as a starting point and also as a baseline to compare which method proves to be the better.

4.3.2 External Clock Bias Correction Method

As the pseudorange measured by the device is related to the time the signal takes from the emitting satellite to the receiving device (i.e. tags), it is expected that even the smallest delay would affect the measurement, resulting in greater shift of latitude and longitude. The Lantronix device used in this thesis does not have the precision of the atomic clocks in each satellite, therefore it has a small time deviation even after achieving sync (synchronizing its internal clock by the satellite clock). To correct this deviation and correct the pseudorange it was incorporated into the algorithm a sum making use of an average clock bias for each satellite originating from external files obtained from European Space Agency (ESA) [24] in which it depicts the clock bias for each satellite for the entirety of the day, in time intervals of fifteen minutes. Such is given with the expression:

$$pseudorange_c = c * (t * 10^6) + pseudorange_o, \tag{1}$$

where c - constant speed of light in meters per second, t - clock bias from the external file in nanoseconds, $pseudorange_o$ - obtained pseudorange in meters and $pseudorange_c$ - corrected pseudorange in meters.

4.3.3 Reference Method

The pseudorange obtained by a GPS device is the result of the sum of the real range and several errors (clock bias, ionospheric delay, tropospheric delay). As a way to remove such errors, two devices will be used in this method: (i) the former being the reference which is at a known location (latitude and longitude), and (ii) the latter, which acts as the tag with an unknown location. Since the reference position is indeed known, it is further used to calculate the real range

---

[24] http://navigation-office.esa.int/Products.html

from the device to each satellite (equation 2). Then, from the calculated range is subtracted the pseudorange (calculated through the usage of TLE) which results in the value that represents all the errors for that measurement as seen in next expressions:

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

where $range$ - true range from the device to the satellite, X_s, Y_s, Z_s - satellite coordinates obtained through TLE, and X_r, Y_r, Z_r - reference device known coordinates.

$$error_r = range - pseudorange_r, \tag{3}$$

where $error_r$ - calculated error, $range$ - previously calculated range, and $pseudorange_r$ - pseudorange obtained by the reference device. This value is then added to the pseudorange of the tag resulting in a more accurate pseudorange.

$$pseudorange_c = pseudorange_o + error_r. \tag{4}$$

where $pseudorange_c$ - corrected pseudorange in meters, $pseudorange_o$ - pseudorange obtained by the tag, and $error_r$ - previously calculated error. Illustration of the reference method is shown in the Figure 14.
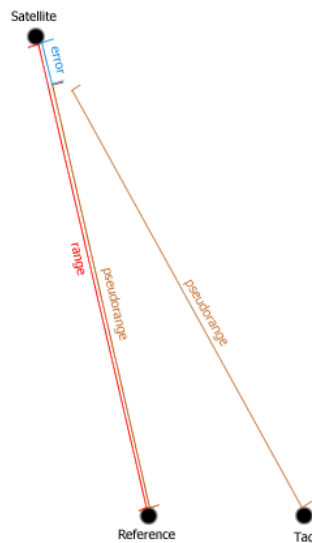
Fig. 14: Illustration of the reference method.

4.3.4   Satellite Data Method

This method was developed with the goal of removing external dependencies required to calculate satellite position and compensate for clock bias error. To achieve this, the data used for the pseudorange correction all originated from the satellite messages (i.e. message IDs 7, 28 and 30) without using TLE calculation for the position and external clock bias. To achieve this it is required to correct the satellite position received as well as correcting the pseudorange with the data reported by the satellite. To obtain a correct pseudorange to the raw pseudorange, next expression is used:

$$pseudorange_c = pseudorange_o - c*(t_{bias}/10^9) + (i/1000), \tag{5}$$

where c - constant speed of light in meters, t_{bias} - clock bias from message $ID7$ in nanoseconds, $pseudorange_o$ - obtained pseudorange from message $ID28$ in meters, i - ionospheric delay from Message $ID30$ in meters, and $pseudorange_c$ - corrected pseudorange in meters.

To obtain a correct the position of the satellites that will be used for the linearization, several steps are required to obtain them. These position are calculated in a 3 coordinate axis (X, Y and Z). For these equations it will be used data received from all three messages previously explained and will make use of the satellite time and it's clock bias.

$$PosX_c = PosX_o + \left(((t_{gps} - t_{bias}/10^9) - (t_{gps} - (pseudorange_o/c)) * vX_o\right) \tag{6}$$

$$PosY_c = PosY_o + \left(((t_{gps} - t_{bias}/10^9) - (t_{gps} - (pseudorange_o/c)) * vY_o\right) \tag{7}$$

$$PosZ_c = PosZ_o + \left(((t_{gps} - t_{bias}/10^9) - (t_{gps} - (pseudorange_o/c)) * vZ_o\right) \tag{8}$$

where c - constant speed of light in meters, t_{gps} - Gps Software time from message ID 28, t_{bias} - clock bias from message ID 7 in nanoseconds, $pseudorange_o$ - obtained pseudorange from message ID 28 in meters, vX_o, vY_o, vZ_o - satellite velocity in each axis obtain from Message $ID30$, $PosX_o, PosY_o, PosZ_o$ - satellite position coordinates from Message $ID30$ in meters, and $PosX_c, PosY_c, PosZ_c$ - corrected coordinates in meters.

4.3.5   Satellite Data with Reference Method

Although the previous method corrects the pseudorange using most of the components, it still lacks to reach the realistic range, as it does not incorporate the tropospheric delay. Tropospheric delay brings further complications as it oscillates abruptly making it challenging to obtain a realistic value. To overcome such difficulty an altered version the aforementioned reference method (topic

4.3.3) was used as an auxiliary method, the subtle change were made on the equation 2. These were made due to the complete removal of external dependencies , by having calculated the satellites position from the data received instead of obtaining it through the usage of TLE, resulting in equation 9:

$$range = \sqrt{(PosX_c - X_r)^2 + (PosY_c - Y_r)^2 + (PosZ_c - Z_r)^2}, \tag{9}$$

where $range$ - true range from the device to the satellite, $PosX_c, PosY_c, PosZ_c$ - satellite coordinates obtained through equations 6, 7 and 8, and X_r, Y_r, Z_r - reference device known coordinates.

4.4 Step 4: Trilateration Procedure

After obtaining the position of the satellite and the corresponding pseudorange it is possible to proceed to the calculation of the position of devices on surface (i.e. tags). For this purpose, the intersection of the spheres is calculated, originated on the satellite positions with radii being the corresponding pseudorange of each satellite (will be further presented in Figure 15).

4.4.1 Shifting Algorithm

As to cover for the possibility of the order of the pseudoranges being different from the order of the real ranges , as a error in the order would create a different location point and not the real one as the satellites would not have their real distance from the device taken in consideration, a shifting algorithm was implemented. This would change the order of the satellites if the resulting point was outside of the acceptable height limit ($20000km - 26000km$), to achieve this using the furthest satellite as reference it would then be calculated the missing distance between its corresponding pseudorange and the maximum limit ($26000km$) then the resulting value is added to the remaining satellites, and the pseudorange of the satellite used as reference would be set to the minimum limit ($20000km$). This cycle was set to iterated to a maximum of five times since the number of satellites being used was four, this would ensure it would end in the same starting order although with different heights resulting from the order change the algorythm implements. In case of not finding a valid point even after the five cycles the point is discarded and labeled as invalid the total of these was used to verify the precision of the method. A different and more complex shifting algorithm that ran through all possible combinations of the satellites distances was also tested, but ended up being discarded as it required a lot more resources, but the main issue was the creation of position points that while being incorrect were inside the acceptable height limit. This way they were being accept as valid having a impact on the precision. For all these reasons the simpler and more reliable version of the algorithm was applied.

4.4.2 Limited Data Linearization (4 Satellites)

In next, following the procedure by Kaplan [11], for obtained 4 satellite pseudoranges and 4 satellite positions, we calculate the position of the tag at the sea surface from the next set of nonlinear equations:

$$\rho_1 = \sqrt{(x_1 - x_{tag})^2 + (y_1 - y_{tag})^2 + (z_1 - z_{tag})^2} + ct_{tag}, \tag{10}$$

$$\rho_2 = \sqrt{(x_2 - x_{tag})^2 + (y_2 - y_{tag})^2 + (z_2 - z_{tag})^2} + ct_{tag}, \tag{11}$$

$$\rho_3 = \sqrt{(x_3 - x_{tag})^2 + (y_3 - y_{tag})^2 + (z_3 - z_{tag})^2} + ct_{tag}, \tag{12}$$

$$\rho_4 = \sqrt{(x_4 - x_{tag})^2 + (y_4 - y_{tag})^2 + (z_4 - z_{tag})^2} + ct_{tag}, \tag{13}$$

where c - constant speed of light, t_{tag} - time offset, $\rho_1, \rho_2, \rho_3, \rho_4$ - obtained pseudoranges, x, y, z - satellite positions and $x_{tag}, y_{tag}, z_{tag}$ - positions of the tag at the sea surface. By expanding aforementioned equations into Taylor series about the approximate position, we can obtain the position offset (Δx_{tag}, Δy_{tag} and Δz_{tag}) as linear functions of the known coordinates and pseudorange measurements:

$$\Delta \rho_1 = a_{x1} \Delta x_{tag} + a_{y1} \Delta y_{tag} + a_{z1} \Delta z_{tag} - c \Delta t_{tag}, \tag{14}$$

$$\Delta \rho_2 = a_{x2} \Delta x_{tag} + a_{y2} \Delta y_{tag} + a_{z2} \Delta z_{tag} - c \Delta t_{tag}, \tag{15}$$

$$\Delta \rho_3 = a_{x3} \Delta x_{tag} + a_{y3} \Delta y_{tag} + a_{z3} \Delta z_{tag} - c \Delta t_{tag}, \tag{16}$$

$$\Delta \rho_4 = a_{x4} \Delta x_{tag} + a_{y4} \Delta y_{tag} + a_{z4} \Delta z_{tag} - c \Delta t_{tag}, \tag{17}$$

where a_{x1}, a_{x2}, a_{x3} and a_{x4} denote the direction cosines of the unit vector pointing from the approximate user position to the satellite. Once the unknowns (Δx_{tag}, Δy_{tag}, Δz_{tag}, and Δt_{tag}) are computed, obtaining tag's coordinates ($x_{tag}, y_{tag}, z_{tag}$) and the receiver clock offset (t_{tag}).

4.4.3 Multiple Satellite Linearization (Beyond 4 Satellites)

Obtaining position can be also obtained using beyond existing 4 satellites. Previous equations may be placed into the form of matrix by:

$$
\Delta\rho = \begin{bmatrix} \Delta\rho_1 \\ \\ \Delta\rho_2 \\ \\ \Delta\rho_3 \\ \\ \Delta\rho_4 \end{bmatrix} \qquad H = \begin{bmatrix} a_{x1}\ a_{y1}\ a_{z1}\ 1 \\ \\ a_{x2}\ a_{y2}\ a_{z2}\ 1 \\ \\ a_{x3}\ a_{y3}\ a_{z3}\ 1 \\ \\ a_{x4}\ a_{y4}\ a_{z4}\ 1 \end{bmatrix} \qquad \Delta x = \begin{bmatrix} \Delta x_{tag} \\ \\ \Delta y_{tag} \\ \\ \Delta z_{tag} \\ \\ -c\Delta t_{tag} \end{bmatrix} , \tag{18}
$$

which has a solution:

$$
\Delta x = H^{-1}\Delta\rho. \tag{19}
$$

Indeed, true user-to-satellite measurements are prone to errors such as measurement noise, deviation of the satellite path from the reported ephemeris and multipath. Least square method (LSM) is further applied to improve the estimates of unknowns. In the case of multiple satellites, LSM is applied where the H matrix in expression 18 becomes:

$$
H = \begin{bmatrix} a_{x1}\ a_{y1}\ a_{z1}\ 1 \\ \\ a_{x2}\ a_{y2}\ a_{z2}\ 1 \\ \\ \vdots \\ \\ a_{xn}\ a_{yn}\ a_{zn}\ 1 \end{bmatrix} ,
$$

where n is the number of satellites. LSM is further obtained by mutiplying both sides on the left of expression 19 by the matrix transpose of H, yielding:

$$
\Delta x = (H^T H)^{-1} H^T \Delta\rho, \tag{20}
$$

which is the LSM formulation for Δx as a function of $\Delta\rho$.

4.4.4   Conversion from ECEF to LLH coordinates

Once the tag positions are obtained, further step is to convert the Earth-centered Earth-fixed (ECEF) coordinates into latitude-longitude-heigh (LLH) coordinates. Two radii were used: (i) $a = 6378137$ - radius at poles in meters and (ii) $b = 6356752.314245$ - being the radius at equator in meters. Longitude (λ) is obtained by the expression:

$$\lambda = atan2(Y, X), \tag{21}$$

while the height is obtained from the a circular relationship involving N, which is a function of latitude:

$$h = \frac{p}{cos\phi} - N, \tag{22}$$

where $p = \sqrt{X^2 + Y^2}$, and latitude (ϕ) by:

$$\phi = arctan\big((Z/p)/1 - e^2N/(N + h)\big), \tag{23}$$

where $e^2 = 1 - \frac{b^2}{a^2}$ is the square of the first numerical eccentricity of the ellipsoid. An example of the obtained coordinates may be seen in image below (Figure 15), which depicts the position of the tag on a sphere (black dots against blue circle), starting with the initial coordinates from the earth core , the position of the satellites used in the process (blue, yellow, red and green dots on the left) and the resulting time error of the process throughout the calculation (right graph).
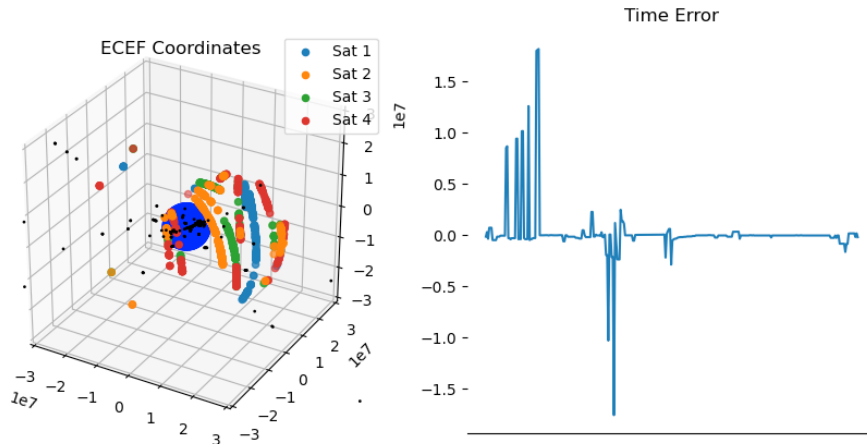
Fig. 15: Performed trilateration and time error.

4.5 Method Procedures

In below, the steps required for each method will be depicted in table 8.

Table 8: Method Procedures.

From top to bottom: (a) TLE – Use of TLE extenal data (topic 4.1), (b) MessageIDs – messages used to retrieve Pseudorange (topic 4.2), (c) Shift – shift algorithm (topic 4.4.1), (d) L. Linearization – Limited Data Linearization (4 Satellites) (topic 4.4.2), (e) M. Linearization – Multiple Satellite Linearization (Beyond 4 Satellites) (topic 4.4.3) and (f) ECEF conv – Conversion from ECEF to LLH coordinates (topic 4.4.4).

#	1	2	3	4	5
Method	Raw Pseudorange	External Clock Bias Correction	Reference	Satellite Data	Satellite Data with Reference
(a) TLE	-	X	X	-	-
(b) MessageIDs	28	28	28	7,28,30	7,28,30
(c) Shift	-	X	X	-	-
(d) L. Linearization	X	X	X	X	-
(e) M. Linearization	-	-	-	-	X
(f) ECEF conv.	X	X	X	X	X

5 IoT Device - Experimental Setup and Validation

In forthcoming, experimental procedures and data inquiry is depicted for both the IoT sensor (in-the-wild study) and snapshot geolocation pipeline (controlled environment).

5.1 Experiments with IoT Sensor

Three experiments were performed on an IoT device, being: (i) custom antenna attenuation, (ii) capacitive tests, and (iii) in-the-wild validation.

5.1.1 Experiment 1: LoRa Antenna Attenuation

Flexible and smaller antenna to be used for aquatic species observation, and will use one commercial antenna, and a re-factored antenna where the external coating and the amplifier were taken of (Figure 16). The objective being to test either the custom made antenna is ready for deployment or not, by comparing the RSSI and SNR obtained from both antennas. As of preparation for this experiment two tables are placed at the distance of 5 meters in the confined area as to ensure a controlled environment of five square meters. On top of one of them a microcontroller is used with LoRa chips and the commercial antenna, this one will work as a receiver and will be the same throughout the experiment. On top of the other table is placed a microcontroller is used with LoRa chips as this will be the sender the antenna used will be first be the commercial antenna, for a baseline, followed by the custom antenna. Each testing phase will be made making the sender device send payload messages once in a 4 second duty cycle, the data was gathered for 10 min for each phase.

(a) 1/4 Wavelength   (b) Helical (exterior)   (c) Helical (interior)   (d) Fishing Antenna
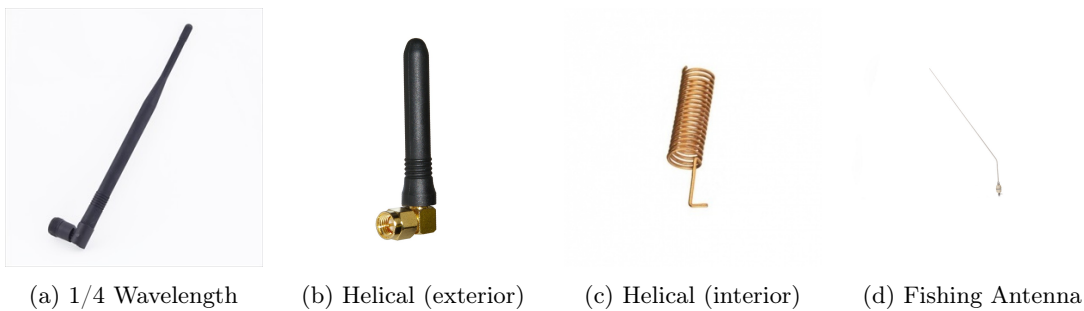
Fig. 16: Used LoRa 868 Mhz antennas (from left to right): (a) regular 1/4 wavelength (appx. 21cm); (b) helical antenna (exterior); (c) helical antenna (interior); (d) custom made fishing antenna from stainless steel and nylon.

Another solution that fit the requirements of size for an antenna suitable for the device being conceived was a commercial 1/4 wave helical antenna. Still using the commercial antenna as a baseline this experiment will compare the previously acquired data with the newly data obtained from using the antenna in its complete state and without its cover. Since this experiment complements the prior antenna experiment, the environment used is the same. Two tables are placed at the distance of 5 meters in the confined area

as to ensure a controlled environment of five square meters. On top of one of them a microcontroller is used with LoRa chips and the commercial antenna, this one will work as a receiver and will be the same throughout the experiment. On top of the other table is placed a microcontroller is used with LoRa chips as this will be the sender the antenna used will be first be the helical antenna while covered, followed by the data collection while without the cover. Each testing phase will be made making the sender device send payload messages once in 4 second duty cycle, the data was gathered for 10 min for each phase.

Table 9 contains the average values in both SNR (signal-to-noise ratio) and RSSI (received signal strength indication) obtained after both phases of Experiment 1. The displayed absolute difference between the respective antenna and the baseline antenna.

Table 9: Results from Experiment 1. LoRa Antenna Attenuation

Antenna	Avg. SNR	Avg RSSI	Highest ABS. Difference (SNR)	Highest ABS. Difference (RSSI)
Commercial	9.29	-55.64	—	—
Re-Factored	9.31	-66.39	1.75	19
1/4 Helical (without cover)	9.29	-55.64	3.25	12
1/4 Helical (with cover)	9.29	-55.64	1.75	10

From the results we can observe that the commercial antenna provides a reasonably stronger signal than the other options while having a similar signal to noise ratio.

5.1.2 Experiment 2: Capacitive Tests

A solution found to extend the battery lifetime was to have a way to control the device and control the device's uptime, this experiment is to see if the chosen solution of having a capacitive sensor control the wake up routine is reliable. For it to be reliable it is require to have a noticeable different in readings while underwater and when it surfaces.

Since salt water was an expected condition for the deployment, the test was carried out using salt water obtained from Machico Sand Beach. The objective for this test was to verify if the capacitive sensor had a significant conductive difference between the two expected situations, inside the water for long periods of time (marine animal dive period) and outside for small periods (marine animal breathing period). The capacitive sensors used were identical in manufacturing process with the only difference being that one had a layer of red nail polish, being this the theoretical solution to increment the conductive difference. Both sensors were placed in the salt water ,without touching any surface or material that would affect its conductivity, only using its own wires bent around the top of the recipient to support the weight, for

a full submersion for a period of one hour. After this period was due the sensors were connected to a microcontroller which was set to read the conductive value measured by the sensor, values were obtained while underwater for two minutes as the value was stable and the conditions were the same, and when the sensor was removed from the salt water for another two minutes simulating the surfacing of marine fauna that can range from a few seconds to a few minutes. As to not affect the values the sensors and their cables were not handled by hand while values were being read.

5.1.3   Experiment 3: In-the-wild Validation

In-situ validation was performed in the pelagic zone of Funchal area on sailing boat sea vessel. The former was in vicinity to the Funchal harbour, while the latter one was in the pelagic area in front of the Funchal. Obtained tests are depicted in Figure 17, where it is possible to seethe comparison between the obtained data points from the SD card against the data points obtained directly LoRa transmission, where the LoRa payload messages reached the gateways on coast. As suspected, obtained data points using the LoRa telemetry were subject to data loss, due to a loss of line-of-sight (LOS).

5.2   Experiments with Snapshot Geoolocation Pipeline

Similar to the IoT experiments, total three experiments were also carried out with the snapshot geolocation pipeline: (i) Lantronix outdoor worst case scenario, (ii) Lantronix outdoor best case scenario, and (iii) Wurth outdoor worst case scenario.

5.2.1   Lantronix Outdoor (worst case scenario)

In order to prepare for scenarios where the device would obtain a small amount of satellites, a chair was set in the front garden of ARDITI where one device would be placed during the entirety of the experiment running in One Socket Protocol (OSP) mode. This place is located in a valley and next to a tall building which limits the amount of satellite signals visible, and also causes a more delayed reception in some of the satellite data. This data was used with the methods explained in the topics 4.3.1 and 4.3.2.

5.2.2   Lantronix Outdoor (best case scenario)

To ensure a clear sky view with no terrain interference two devices were placed on the rooftop of UMa, again running in One Socket Protocol (OSP) mode. This environment would emulate the case where the tag would have a clear view of the satellites in the sea without any animal or human interference (boats), and would fulfill the clear sky and known location requirements for the reference device. This data was used with the method explained in the topic 4.3.3 as it allowed to acquire more data for the reference.

5.2.3   Wurth Outdoor (worst case scenario)

As the Wurth device allowed the possibility of using several satellite constellations its reception was expected to be better than the previously used Lantronix. To test this it was set in the same conditions as

(a) TAG6 (SD card)

(b) TAG7 (SD card)
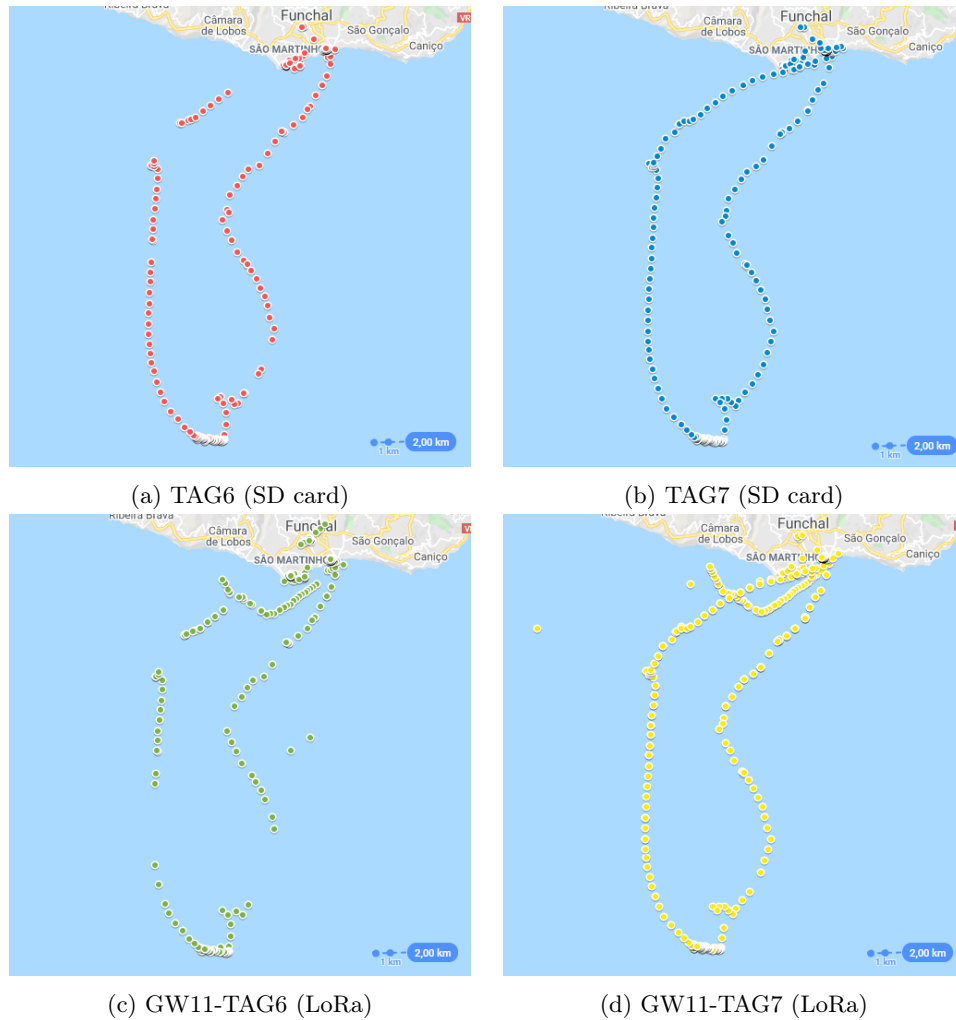
(c) GW11-TAG6 (LoRa)

(d) GW11-TAG7 (LoRa)

Fig. 17: Completed field study test (from left to right): (a) SD card readings of bodyboard tag obtained GPS locations; (b) SD card readings of pole tag obtained GPS locations; (c) Land gateway readings of received LoRa messages from bodyboard tag; (d) Land gateway readings of received LoRa messages from mast tag.
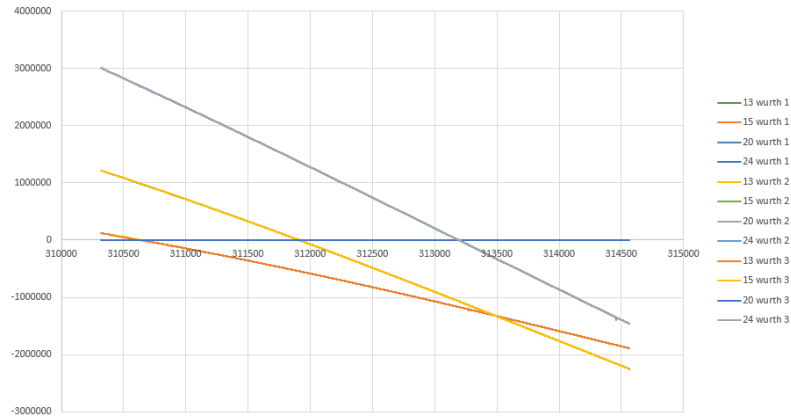
Fig. 18: Satellite pseudorange comparison.

the Lantronix worst case scenario test, where a chair was set in the front garden of ARDITI where three devices would be placed during the entirety of the experiment running in One Socket Protocol (OSP) mode. As this data collection would be used for methods that exclude external data (topics 4.3.4 and 4.3.5), as to ensure that all devices where obtaining similar data a small data collection was acquired and, from it picked four satellites (ids 13, 15, 20 and 24) along with their measured pseudoranges. For each device it was created a comparison using satellite id 20 pseudorange as a baseline and then created a graph with all their values resulting in fig. 18. It is visible that all the satellites with the same ids are depicted with the same difference from our baseline (satellite 20).

6 Snapshot Geolocation Pipeline - Experimental Setup and Validation

In next, apparatus for geolocation is described with benchmarks of different methods for snapshor geolocation including: (i) Raw Pseudorange, (ii) Shifting Algorithm implementation, (iii) External Clock Bias Correction, (iv) Reference, (v) Satellite Data, (vi) Satellite Data with Reference, and (vii) Multiple Satellite Linearization. Each one of them is benchmarked, with maximum obtained accuracy (accuracy in this case is how far the most distant point was from the real location) and lowest possible error in terms of distance for two levels: 100m and 1000m. Summary of obtained results is depicted in Table 10. Also, for each method, latitude and longitude coordinates are depicted.

Table 10: Obtained accuracy and error benchmarks from performed geolocation methods. From top to bottom: (a) Sensor – used receivers (L-Lantronix GPS, W-Wurth GNSS), (b) TotPoints – amount of total sampled points (collected), (b) TotVPoints – amount of total valid mapped points calculated within the acceptable height, (c) Points100m – amount of total points within 100m, (d) Points1km – amount of total points within 1km, (e) Duration(s) – time to obtain sampled points (seconds), (f) FigureNo – reference figure depicting the obtained points on a map when applying the method, (g) Efficiency – points successfully calculated with the linearization (valid points) / total points * 100, and (h) precision – how close the furthest one is to the ground truth. Best method is using multiple satellite linearization.

#	1	2	3	4	5	6	7
Method	Raw Pseudorange	Shifting Algorithm	External Clock Bias Correction	Reference	Satellite Data	Satellite Data with Reference	Multiple Satellite Linearization
(a) Sensor	L	L	L	L	W	W	W
(b) TotPoints	602	602	602	573	7000	7011	**7011**
(c) TotVPoints	130	332	287	185	126	1241	**4929**
(d) Points100m	0	0	0	34	2	470	**2910**
(e) Points1km	30	66	70	181	79	1241	**4929**
(f) Duration(s)	15	15	15	15	10	10	**10**
(g) FigureNo.	19	20,21	22	23	24	25	**26**
(h) **Efficiency**	21,59	55,15	47,67	32,29	1,80	17,70	**70,30**
(i) **Precision**	17788km	16361km	17233km	1120m	15100m	469m	**394m**

6.1 Raw Pseudorange Method

With method using solely Lantronix GPS, Figure 19 depicts the obtained latitude and longitude using SPG4 model without corrections of pseudoranges and Lantronix GPS receiver. In this method the lin-
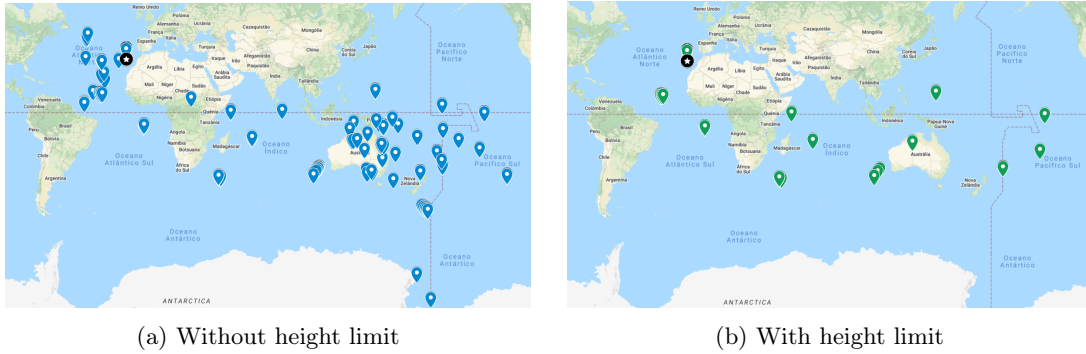
(a) Without height limit (b) With height limit

Fig. 19: Obtained points using Simple Linearization Method (Lantronix GPS).

earization used is the one depicted in the pipeline (see Section 4.4) and it was iterated ten times. It is observable in the image 19a that there are two clusters, one being an antipod of the other, and one of the cluster is close to the correct position (indicated with the black icon). This suggests that there is a need in making the correction to remove the antipods. Figure 19b depicts same experimental conditions as the previous test, where the calculation was iterated twenty times. Such method reduced the amount of total points decreasing the efficiency, however providing less antipod points resulting in an increase of precision although still not acceptable to be considered for deployment.

6.2 Shifting Algorithm Incorporation

As to the implemented shifting algorithm, main goal was to change the order of the satellites if the resulting point was outside of the acceptable height limit (e.g. within $1km$), as we can observe in Figure 20a this resulted in a great improvement in the selection of valid point with minimal points in antipod locations. To verify if there was any possible manner to influence the result with latitude and longitude limitations both were applied resulting in Figure 20b. While the antipods were removed completely as expected, there was no change in the calculated points on the island as we can perceive in Figures 21a and 21b. Given the substantial improvement this method provided, it became part of the linearization process for all the methods , although only applying the height limit as it the only limit that is plausible in a world wide location system scenario, as applying latitude and longitude limits imply that the calculated position is vaguely known or expected.

6.3 External Clock Bias Correction Method

In this method, to the obtained pseudorange from the satellite was added the distance equivalent to the error compensation for clock bias, this measurement is depicted in the external files obtained from NASA in which it depicts the clock bias for each satellite for the entirety of the day, in time intervals of fifteen minutes. While this makes way for some errors the further from the measurement time it is, it helps in reducing the antipods in a considerable amount as we can observe in Figure 22a. Then it was applied a height limit in the same manner it was applied in the SGP4 method, which resulted in a great increase in
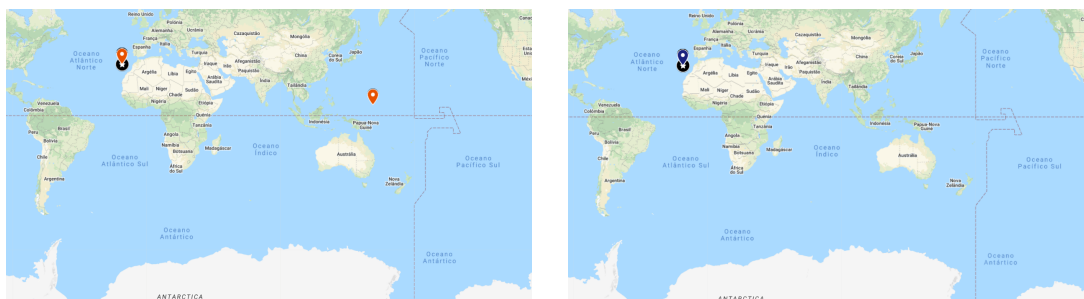
(a) with height limit

(b) with height limit and window limit

Fig. 20: Obtained points using Shifting Algorithm Method.
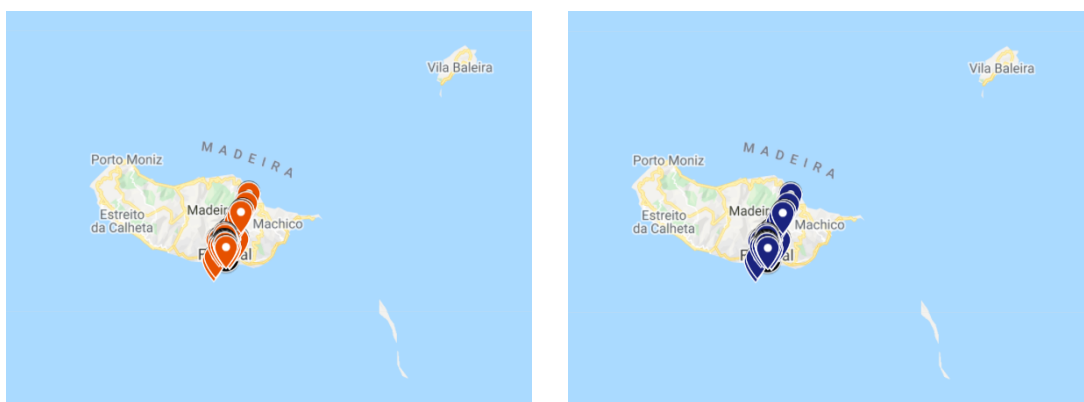
(a) with height limit

(b) with height limit and window limit

Fig. 21: Obtained points using Shifting Algorithm Method, zoomed onto Madeira island.

precision while keeping most of the efficiency. This is visible in Figure 22b as it shows less antipods and the removal of the points in the Atlantic ocean.

(a) no limitations

(b) with height limitation
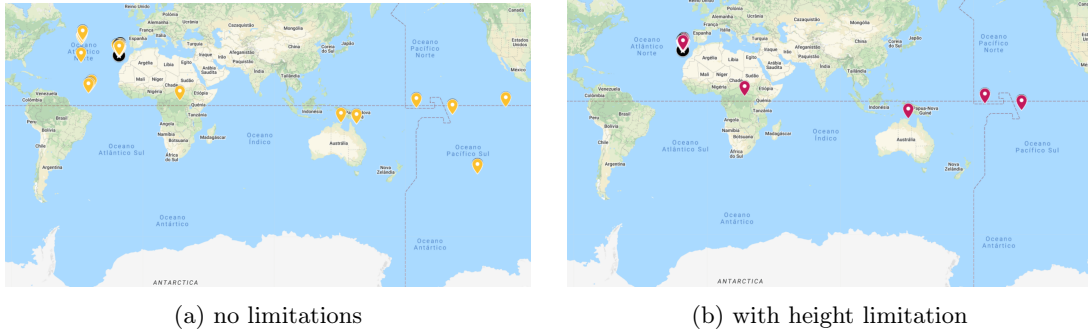
Fig. 22: Obtained points using External Clock Bias Correction Method (Lantronix GPS).

6.4   Reference Method

When performing such method an increase in efficiency was expected due to making use of a known position and its real calculated range. The results depicted in Figure 23d confirms this as most of the points were successfully located with an increase in precision, making this experiment the biggest leap towards an acceptable and reliable solution. The data used as reference and the raw data from the tag are also illustrated in Figures 23d and 23b respectfully, as well as both point collections in Figure 23c.

(a) reference points

(b) tag points

(c) points comparison

(d) resulting points

Fig. 23: Obtained points using Reference Method.

6.5   Satellite Data Method

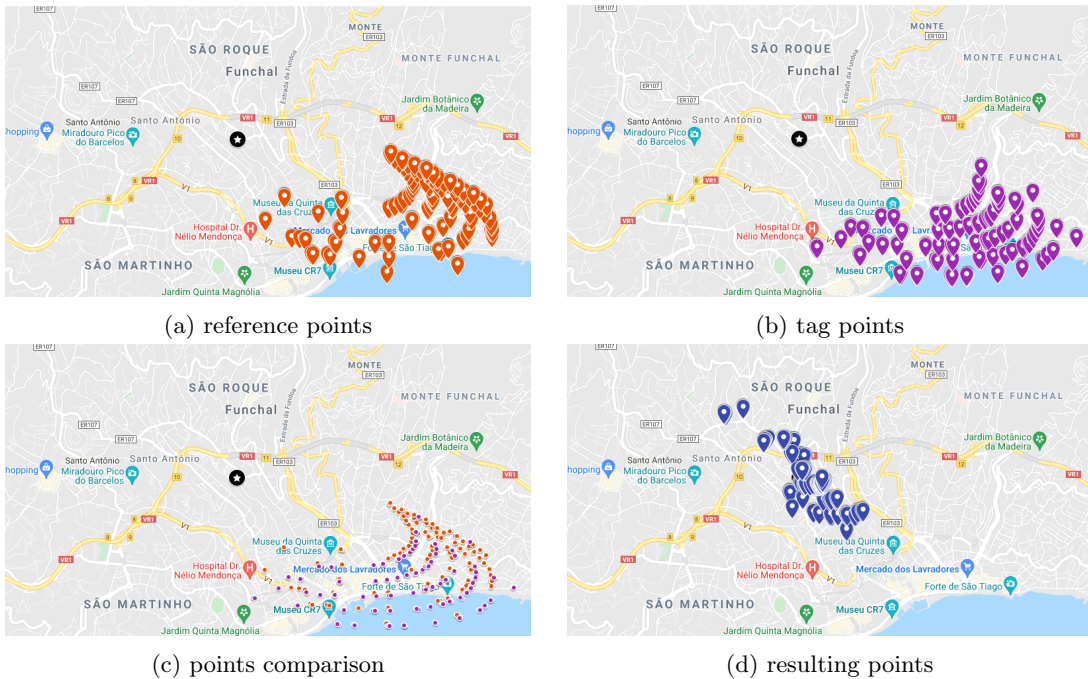In Figure 24 we can observe that although the efficiency greatly decreased there is a cluster near the ground truth. Taking into consideration the results from the previous methods we can infer that there is an issue with this pseudorange correction since the decrease in efficiency is too steep (Table 10). A possible assumption would be the importance of tropospheric delay compensation in this method, given that it is the only error not taken in consideration while correcting the pseudorange in this experiment. However, obtained results indicate that it is possible to achieve an acceptable position using only the satellite data, since the precision is slightly better than the external clock bias correction method.
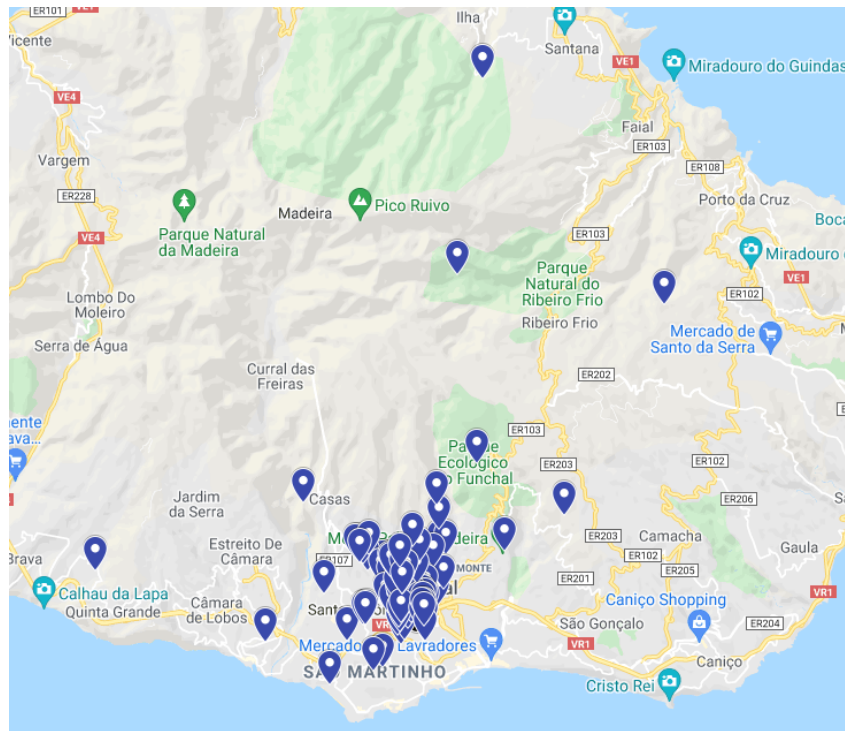
Fig. 24: Obtained points using Satellite Data Method.

6.6   Satellite Data with Reference Method

In Figure 25 it is perceivable that the reference method is responsible for the increase in precision. Such is expected since previously used reference method used to correct ionospheric delay, clock bias and tropospheric delay. In here, it is only responsible for the tropospheric delay and with the reduction on the variables that it affects there is a decrease in errors acquired through calculation. This experiment also suggest that the Wurth GNSS receiver device used throughout the study allows for a more precise result although at the cost of decreasing the efficiency (Table 10).

Fig. 25: Obtained points using Satellite Data With Reference Method.

Fig. 26: Obtained points using Multiple satellites incorporation.

6.7 Multiple Satellites Linearization Incorporation

Last but not least, in Figure 26 it is possible to see more condensed points around the ground truth area. Proposed method used the combination of multiple satellites (> 4), which clearly indicates the advantage of Wurth GNSS of Lantronix GPS. Table 10 indicates the highest possible precision (394 metes) and efficiency (70% of points were at the proper height and at the proper distance). Such results indicate that usage of multiple satellites lead to an increase of overall pipeline accuracy and that it is possible to use limited amount of IoT resources and time to obtain the raw data from the satellites, process them and allocate the object at the sea surface within an affordable perimeter.

7   Discussion and Conclusions

This dissertation contributes by providing the IoT device and snapshot geolocation pipeline, allowing more affordable marine sensing. IoT device was validated in in-the-wild setting, having two tags and four land receivers, when tracking one marine vessel. Snapshot geolocation pipeline was validated in controlled environment, having compared two satellite receivers – Lantronix GPS and Wurth GNSS. Multiple algorithms were benchmarked, allowing the trilateration of obtained raw satellite data.

Finding Analysis. Obtained results indicate that it is possible to perform tracking of marine objects at the sea surface using the proposed affordable system based on IoT devices and the snapshot geolocation pipeline. Throughout the dissertation, IoT device challenged further the existing COTS devices, typically used in marine monitoring. Also, usage of raw GNSS data and proposed geolocation pipeline is feasible to be used when tracking surface marine objects, achieving a precision radius of 394 meters and an efficiency of 70.30 percent which is an acceptable results to be deployed. IoT system is described with the actual cost of EUR < 200, which is the order of magnitude compared to the COTS devices presented in Section 1. Snapshot position mechanism is an alternative approach to existing expensive ARGOS systems, which can leverage existing constellations without introducing additional fees or charges. This validates the completion of the core objectives, being accomplished the preparation of the IoT devices for LoRa telemetry and georeferencing from the calculation on onboard CPU allowing the creation of a low cost snapshot receiver.

Research Question Contributions. Reported dissertation effort thus answers to [RQ1] – How to provide more affordable marine sensors – by gathering diverse Internet of Things (IoT) modules and creating land receivers and off-shore sensors for tracking marine vessel. The decrease in performance from the less expensive sensors requires a more carefully planed system architecture, as their processing unit is incapable of handling all the features a hi-end commercial off-the-shelf tag presents without server side assistance. Moreover, it provides initial answers to the [RQ2] – How to create an accessible snapshot pipeline for GNSS receivers – by benchmarking several existing methods used for trilateration. All the methods developed in this thesis present a gradual evolution towards a reliable snapshot pipeline, with each procedure presenting less errors than the previous. This is achievable due to being possible to identify key components when comparing benchmarks.

Study Limitations and Future Works. Clearly, proposed system has several constraints. Although it is claimed that the raw GPS coordinates processing may provide the greater power autonomy, such tests were not carried out and thus were not validated during this dissertation. Moreover, error estimation which depends on tropospheric delay were not addressed in this dissertation, due to the difficulty of obtaining a reliable data source (due to access to the weather conditions). Furthermore, proposed IoT devices remain constrained by the depth. Current COTS devices allow greater submersion of the device (reaching operational <1km of depth) while reported device was tested solely at the sea surface. The proposed system is thus more appropriate for tracking vessels than to be used for the movement ecology with small emerging time intervals (i.e. tracking loggerhead sea turtle trajectories). For overcoming the greater

depths, the device will be submerged into epoxy resin, allowing further analysis and assessments using the proposed IoT device tags. Finally, from empirical tests it was noted that the GNSS sensor may obtain points in less than 3 seconds, however as the average time throughout the experiments was 10 seconds it is required to do more precise data gathering that should further validate such claims. As to take this research a step further it would be necessary to verify is the snapshot solution provides a justifiable battery lifetime increase, and if a snapshot pipeline taken a step further so that works around every error incorporated in the pseudorange, mainly tropospheric delay, is a viable contender for commercial off-the-self tags.

References

[1] Christophe Adrados, Irène Girard, Jean-Paul Gendner, and Georges Janeau. Global positioning system (gps) location accuracy improvement due to selective availability removal. *Comptes Rendus Biologies*, 325(2):165–170, 2002.

[2] Introduction Argos and The Argos. Give Vision. 1970.

[3] Douglas G. Barron, Jeffrey D. Brawn, and Patrick J. Weatherhead. Meta-analysis of transmitter effects on avian behaviour and ecology. *Methods in Ecology and Evolution*, 1(2):180–187, 2010.

[4] Steven J. Cooke, S.G. Hich, M.C. Lucas, and M. Lutcavage. Chapter 18 - Biotelemetry and Biologging. *Fisheries Techniques*, (January):819–860, 2012.

[5] Steven J Cooke, SG Hinch, Martyn C LuCaS, and M Lutcavage. Biotelemetry and biologging. *Fisheries techniques, 3rd edition. American Fisheries Society, Bethesda, Maryland*, pages 819–860, 2012.

[6] Antoine M. Dujon, R. Todd Lindstrom, and Graeme C. Hays. The accuracy of fastloc-gps locations and implications for animal tracking. *Methods in Ecology and Evolution*, 5(11):1162–1169, 2014.

[7] Vladimir Dyo, Stephen A Ellwood, David W Macdonald, Andrew C Markham, Cecilia Mascolo, Bence Pásztor, Niki Trigoni, and Ricklef Wohlers. Poster Abstract : Wildlife and Environmental Monitoring using RFID and WSN Technology. *Distributed Computing*, pages 7–8, 2009.

[8] Manuel Eichelberger, Ferdinand von Hagen, and Roger Wattenhofer. Multi-year gps tracking using a coin cell. In *Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications*, pages 141–146, 2019.

[9] Mark Hebblewhite and Daniel T. Haydon. Distinguishing technology from biology: A critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2303–2312, 2010.

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

[11] Elliott D Kaplan and Christopher Hegarty. *Understanding GPS/GNSS: principles and applications.* Artech house, 2017.

[12] Carey E Kuhn, Devin S Johnson, Rolf R Ream, and Thomas S Gelatt. Advances in the tracking of marine species: using gps locations to evaluate satellite track data and a continuous-time movement model. *Marine Ecology Progress Series*, 393:97–109, 2009.

[13] L. David Mech. *Handbook of Animal Radio-Tracking.* University of Minnesota Press, ned - new edition edition, 1983.

[14] L David Mech and Shannon M Barber. a Critique of Wildlife Radio-Tracking and Its Use in National Parks a Report To the U.S. National Park Service. (January 2002):1–83, 2002.

[15] B. Naef-Daenzer. Miniaturization (0.2 g) and evaluation of attachment techniques of telemetry transmitters. *Journal of Experimental Biology*, 208(21):4063–4068, 2005.

[16] D G Nicholls. Satellite tracking of large seabirds - a practical guide. *Conver*, (August):22, 1994.

[17] Marko Radeta, Miguel Ribeiro, Dinarte Vasconcelos, Hildegardo Noronha, and Nuno Jardim Nunes. Loraquatica: Studying range and location estimation using lora and iot in aquatic sensing. In *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 1–6. IEEE, 2020.

[18] C Rizos, MB Higgins, and S Hewitson. New gnss developments and their impact on survey service providers and surveyors. In *The National biennial Conference of the Spatial Sciences Institute, Melbourne*. Citeseer, 2005.

[19] Gregory M. Skupien, Kimberly M. Andrews, and Terry M. Norton. Benefits and biases of VHF and GPS telemetry: A case study of American alligator spatial ecology. *Wildlife Society Bulletin*, 40(4):772–780, 2016.

[20] Bindi Thomas, John D. Holland, and Edward O. Minot. Wildlife tracking technology options and cost considerations. *Wildlife Research*, 38(8):653–663, 2011.

[21] By Kath Walkerl, Graeme Elliottl, David Nichollsz, Durn Murray, and Peter Dilksl. ALBATROSS (Diomedea exulans) FROM THE AUCKLAND ISLANDS : PRELIMINARY RESULTS. (Gales), 1993.