Institute of Geodesy,
University of Stuttgart

Ecole Nationale des
Sciences Géographiques

Rapport de stage

Cycle des Ingénieurs diplômés de l'ENSG 2ème année

Mapping a cargo-ship network to detect tsunamis

Tasnîme-Jenna Louartani

May - August 2022

☒ Non confidentiel ☐ Confidentiel IGN ☐ Confidentiel Industrie ☐ Jusqu'au ...

ECOLE NATIONALE DES SCIENCES GÉOGRAPHIQUES
6-8 Avenue Blaise Pascal - Cité Descartes - 77420 Champs-sur-Marne
Téléphone 01 64 15 31 00 Télécopie 01 64 15 31 07

Jury

Président de jury :

Laurent Breton

Commanditaire :

Bruce Thomas, James Foster

Encadrement de stage :

Bruce Thomas

Enseignant référent :

Victor Coindet

Rapporteur expert :

Responsable pédagogique du cycle Ingénieur :

Jean-François Hangouët, IGN/ENSG/PEGI

Gestion du stage :

Delphine Genes, Claire Driessens, IGN/ENSG/DFI

Stage 2A du 23 Mai 2022 au 12 Août 2022

Diffusion web : ☒ Internet ☒ Intranet

Situation du document :

Rapport de stage de fin d'études présenté en fin de 2ème année du cycle des Ingénieurs

Nombres de pages : 54 pages dont 55 d'annexes

Système hôte : LaTeX

Modifications :

EDITION	REVISION	DATE	PAGES MODIFIEES
1	0	09/2016	Création

Abstract

Recent tsunami events are studied in depth to improve tsunami early warning systems and tsunami models using GNSS technology. Many devastating tsunamis are generated over, or under the oceans, however our observing capacity is located on land: the oceans being a geodetic desert. To fill this gap, we could use geodetic techniques on board of ships to track changes in sea-surface height and detect even small, around 10 cm amplitude, tsunamis. Using one year of Automatic Idenfitication System (AIS) data of commercial ships crossing the Pacific Ocean, this study aims to realize a statistical coverage showing a density of ships in a world grid. The maps are plotted with PyGMT and identify locations of ships on an hourly average per year, season and month. The maps show a permanent presence of ships next to the coast, especially in East Asia, and a pronounced absence in the southern Pacific Ocean. A seasonal pattern is detected with more ships on sea in winter and autumn. Afterwards, by comparing their locations to the tsunami sources, first insights can be drawn on how effective this geodetic technique could help early warning systems.

Key words: AIS, cargo-ship, PyGMT, GNSS, tsunami warning system

Contents

Glossary and useful acronyms

ENSG École Nationale des Sciences Géographiques, National School of Geographic Sciences

GNSS Global Navigation Satellite Systems

GPS Global Positionning System

IMO International Maritime Organization

AIS Automatic Identification System

DART Deep-ocean Assessment and Reporting of Tsunamis

CPT Color Palette Table

Introduction

0.0.1 The institute

From 23rd May 2022 to 12th August 2022, I realized an internship at the Institute of Geodesy at the University of Stuttgart. The institute gathers several research works mainly based on hydrology but also on geophysics, field in which my training course was based. It is thus with the Prof. Dr. James Foster and my internship supervisor, the Research Assistant Bruce Thomas that I worked during these 12 weeks. During a retreat, I was also able to get interested in other research work and discover different issues in geodesy.

0.0.2 Project Context

Following some of the greatest earthquakes and tsunamis resulting in the loss of many lives (e.g. the Sumatra-Andaman earthquake on the 26th December 2004), researchers and engineers have developed better tsunami warning systems (Banerjee et al., 2005). Indeed, technologies based on Global Navigation Satellite System (GNSS) like Deep-ocean Assessment and Reporting of Tsunamis (DART) buoys which can provide tsunami data in real-time (Bernard and Titov, 2015). It has been proved that real-time GNSS data are effective because they provide robust and rapid estimation of tsunami measurements (Tsushima and Ohta, 2014). GNSS data plays several indispensable roles in the tsunami warning system. Firstly, when it comes to detect affected areas, GPS technology is essential to have an accurate real-time position, which allows the prediction of possible tsunami locations through modeling, in order to provide early warnings (Hébert et al., 2020). Secondly, GNSS have the ability to detect sea level changes with some reliability, speed and robustness in order to set up a tsunami early warning system (Tsushima and Ohta, 2014). Moreover, it has been discovered that tsunami currents can be detected using navigation data derived from the Automatic Identification System (AIS) (Inazu et al., 2020). Since many sources and impacts of tsunamis are located in the Pacific Ocean and that commercial ships represent a large coverage of oceans, AIS data exploitation is useful to provide a cost-effective and easy-to-implement tsunami detection system with real-time, high-throughput estimation detection (Foster et al., 2012).

0.0.3 Goal

My mission was to map a cargo ships network in order to determinate the relevance of monitoring certain ships to improve tsunami warning systems.

0.0.4 Requirements Specification

The goal is to map a network of cargo ships in the Pacific Ocean in order to clearly see where most ships are on an hourly average. In order to have an easy visualization of the network, the Pacific Ocean will be gridded and discriminated by the number of ships (color code), like a heat map. The colors will represent the hourly average of the ships in a block of the grid. The map will be centered and fixed on the Pacific Ocean. The grid block size will be 500km by 500km to start with as it represents the distance between the current DART buoys. Then, the grid size will change to see if it is a relevant parameter or not. The hourly average will be realized over all the year of data, then focused on seasons and months.(Figure 1) Besides, in order to have the visual and the numbers, there will be a table for each criterion that represents the coordinates of a block in the grid with the number of ships it contains.(Figure B.1)

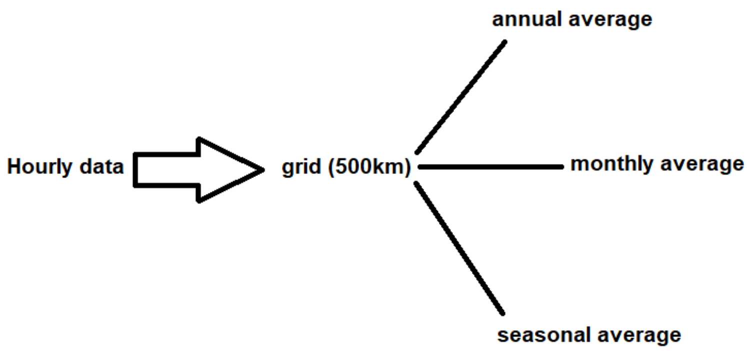


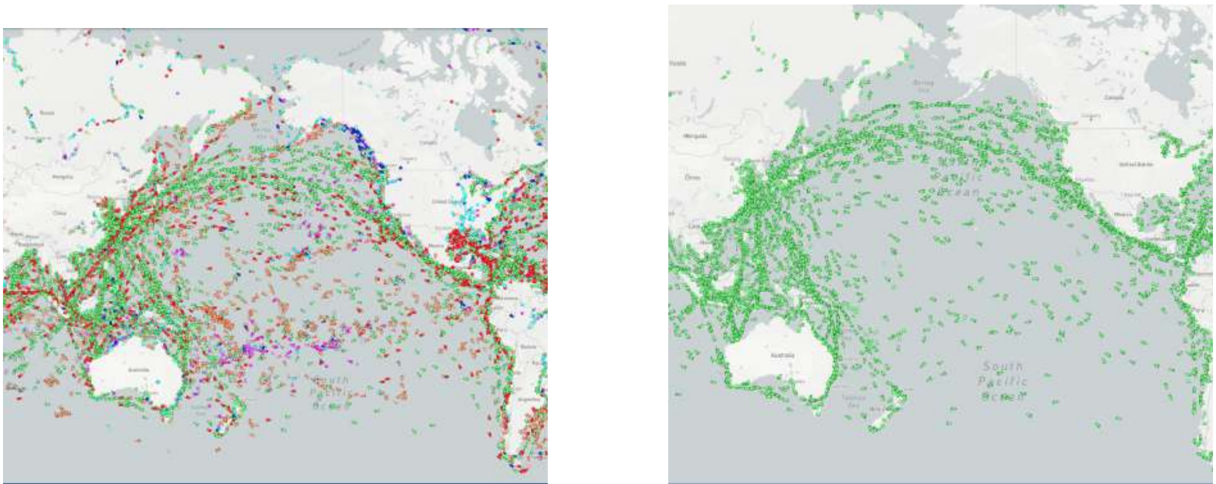
Figure 1: A general overview of how the work is distributed according to the criteria.

Data and IT support

Chapter 1

1.1 AIS Data set

Research based on the largest and deadliest earthquakes in Sumatra, Indonesia (2004) and Tohoku, Japan (2011) has uncovered that AIS data can be used to detect tsunami currents. As the International Maritime Organization (IMO) officially states : *"vessels whose gross tonnages are basically greater than 300 gross tonnages have obligation to transmit the AIS messages using very-high-frequency maritime bands"*. (Inazu et al., 2020) AIS data are the navigation data that delivers specific information used in order to detect tsunami currents. They are available on websites like Marine Traffic (Figure 1.1). These transmitted AIS messages can be received by terrestrial stations and by satellites (Inazu et al., 2020).



(a) All type of ships

(b) Filtered for cargo ships

Figure 1.1: The website Marine Traffic shows a map of ships location in real time

There are two types of information in this data: dynamic and static. (Inazu et al., 2018) The one we pay most attention to in order to map is the horizontal position (dynamic information) which is the longitude and the latitude.
The AIS file provided by Prof. Dr. James Foster goes from the 15th October 2018 to the 15th October 2019. Each file represents the coordinates of cargo ships in the Pacific Ocean for one hour. So for one day, the records go from hour 0 to hour 23. As there are 365 days of AIS, the number of files is 8760 (365*24) files. The files which represented the year 2018 begin with "AIS_2018" and those for the year 2019 begin with "AIS_2019". After the precision of the year, there is a precision of the day and the hour. For instance, for the records of cargo ships during the 15th hour of the 15th October 2018: as the 15th October is the 288th day of the year 2018, the file will be named "AIS_2018_288_15" (Figure 1.2). The files are in *.latlon* format. The *.latlon* format means that a line in a file represents the geographical coordinates of a ship (latitude and longitude). (Figure 1.3)

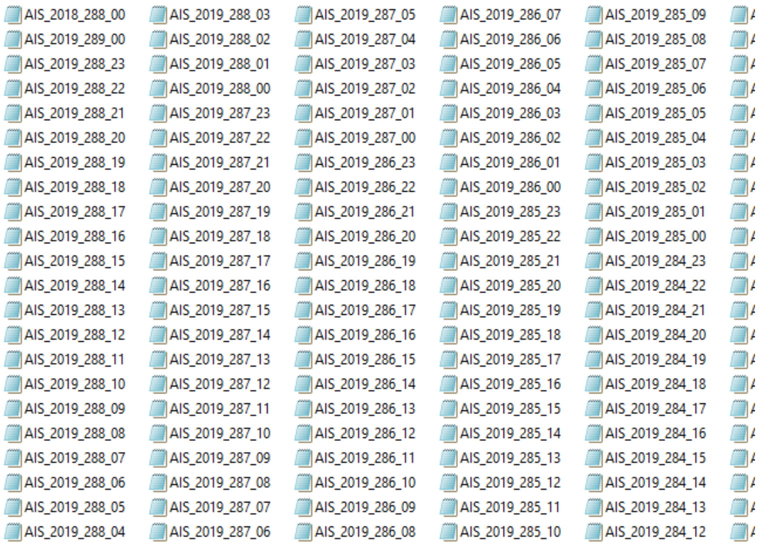


Figure 1.2: Screenshot of the AIS data

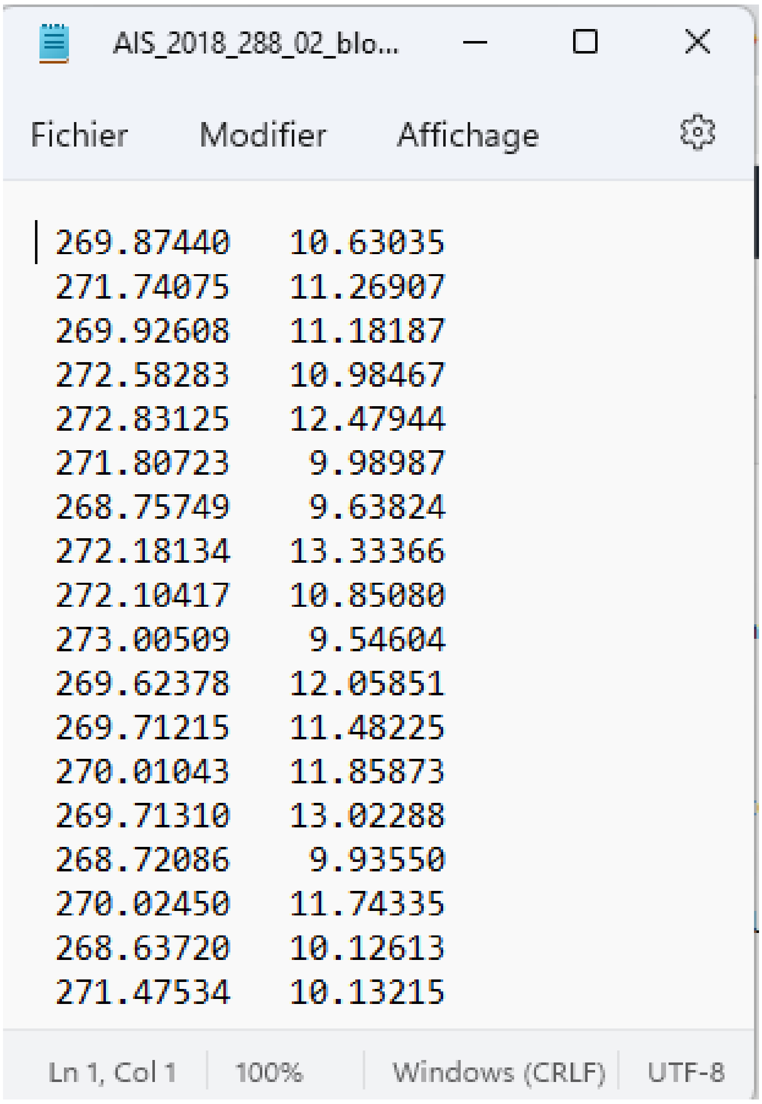


Figure 1.3: Screenshot of one AIS file

1.2 Computer hardware

The solution proposed is to use Python with the library PyGMT. *"PyGMT is a library for processing geospatial and geophysical data and making publication quality maps and figures. It provides a Pythonic interface for the Generic Mapping Tools (GMT)"*. (Uieda et al., 2022) Indeed, Python is a common tool and by using it, the codes can be easily used in order to create others maps with other data. Moreover, PyGMT allows to obtain maps respecting the cartographic semiology by processing a lot of data. Furthermore, it is possible to work on Windows and Linux but I used Windows with an intel core i5 processor for the plotting because it provides a better visual quality when presenting the maps to my supervisors.

Methodological aspect : Setting up the maps

Chapter 2

2.1 Setting up the grid

By doing a grid, the distribution of ships in the Pacific can be quantified and it will be possible to make statistics. First of all, a grid will be created with a set of data (1 day for instance) in order to have a pattern of the grid in the targeted region (here it is : [80,300,-90,90], so the longitudes go to 80 to 300 and the latitudes go to 90 to -90 in order to have a view on the Pacific Ocean). The parameter *spacing* is referring to the size of the blocks (the distance between two nodes): here it is 500 km to begin.

The function pygmt.xyz2grd creates the grid with a set of data with a specified size of the blocks. With the parameter *registration*, there is the choice between **pixels** and **gridline**. Indeed, the function has the capacity to count the number of ships in a block. But the parameters make different nodes in the grid : with the *registration*=**pixels** parameter, the nodes created are the corner of the blocks and with **gridline**, the nodes are in the middle of the blocks.(Figure 2.2) Moreover, the count of the ships is related to the area around the node. (Figure 2.1) That is why the *registration*=**gridline** parameter will be used to count the ships and the *registration*=**pixels** parameter will be used to have the coordinates of the blocks of the grid.(Figure 2.4)

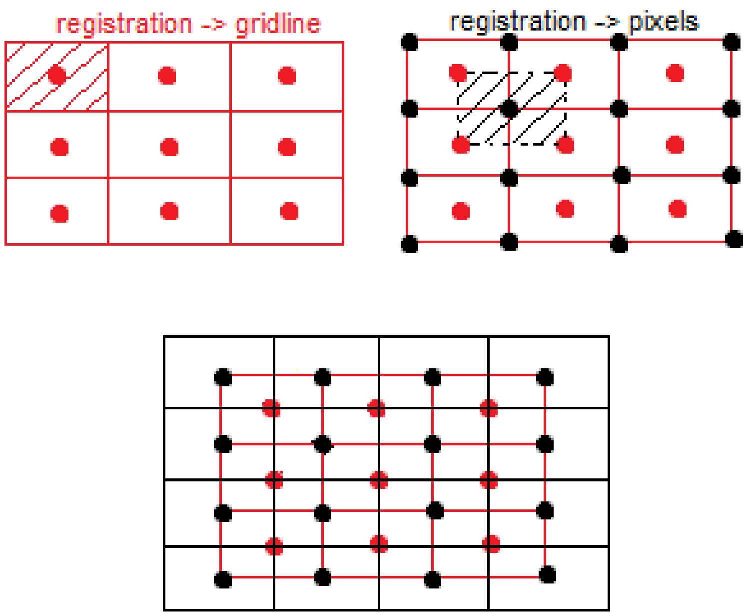


Figure 2.1: The way the ships are counted and the coordinates returned depend on the parameter *registration*

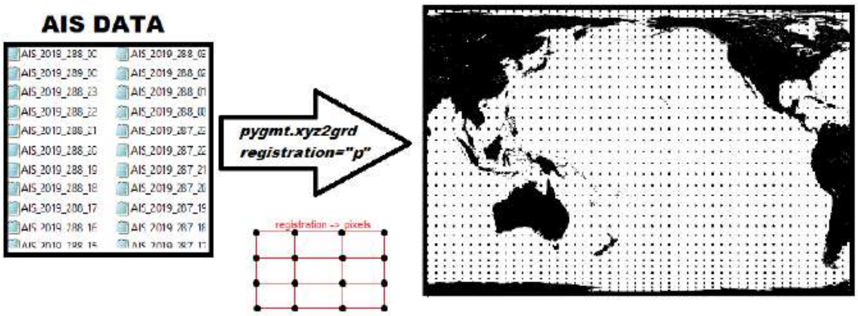
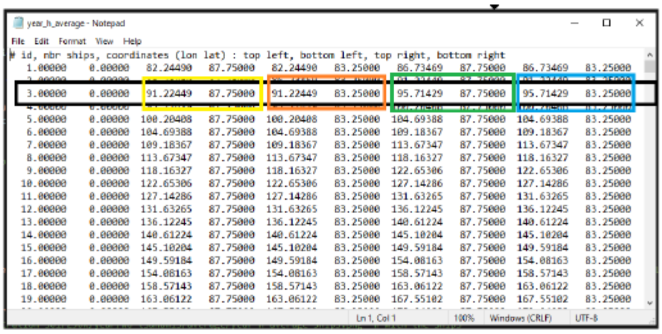


Figure 2.2: The nodes of the grid created with the function pygmt.xyz2grd and the parameter **pixels** ("p")


year_h_average - Notepad

File Edit Format View Help

id, nbr ships, coordinates (lon lat) : top left, bottom left, top right, bottom right

1.00000	0.00000	82.24490	87.75000	82.24490	83.25000	86.73469	87.75000	86.73469	83.25000
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
3.00000	0.00000	91.22449	87.75000	91.22449	83.25000	95.71429	87.75000	95.71429	83.25000
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5.00000	0.00000	100.20408	87.75000	100.20408	83.25000	104.69388	87.75000	104.69388	83.25000
6.00000	0.00000	104.69388	87.75000	104.69388	83.25000	109.18367	87.75000	109.18367	83.25000
7.00000	0.00000	109.18367	87.75000	109.18367	83.25000	113.67347	87.75000	113.67347	83.25000
8.00000	0.00000	113.67347	87.75000	113.67347	83.25000	118.16327	87.75000	118.16327	83.25000
9.00000	0.00000	118.16327	87.75000	118.16327	83.25000	122.65306	87.75000	122.65306	83.25000
10.00000	0.00000	122.65306	87.75000	122.65306	83.25000	127.14286	87.75000	127.14286	83.25000
11.00000	0.00000	127.14286	87.75000	127.14286	83.25000	131.63265	87.75000	131.63265	83.25000
12.00000	0.00000	131.63265	87.75000	131.63265	83.25000	136.12245	87.75000	136.12245	83.25000
13.00000	0.00000	136.12245	87.75000	136.12245	83.25000	140.61224	87.75000	140.61224	83.25000
14.00000	0.00000	140.61224	87.75000	140.61224	83.25000	145.10204	87.75000	145.10204	83.25000
15.00000	0.00000	145.10204	87.75000	145.10204	83.25000	149.59184	87.75000	149.59184	83.25000
16.00000	0.00000	149.59184	87.75000	149.59184	83.25000	154.08163	87.75000	154.08163	83.25000
17.00000	0.00000	154.08163	87.75000	154.08163	83.25000	158.57143	87.75000	158.57143	83.25000
18.00000	0.00000	158.57143	87.75000	158.57143	83.25000	163.06122	87.75000	163.06122	83.25000
19.00000	0.00000	163.06122	87.75000	163.06122	83.25000	167.55102	87.75000	167.55102	83.25000

Ln 1, Col 1 100% Windows (CRLF) UTF-8

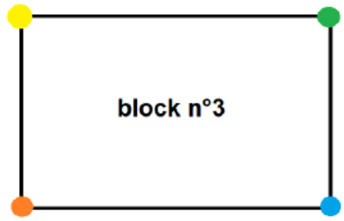


Figure 2.3: The first step is to have the pattern of the grid with a table in which there is the IDs and the coordinates of each block in the grid

With the combination of the two registration parameters, it is possible to have the coordinates of the grid and the number of the ships for each blocks by associating which node of the corners correspond to the node in the middle of the block. Hence, the parameter *registration*=**pixels** will be used only for the pattern of the grid.(Figure 2.3) To get the coordinates of the grid, the function pygmt.grd2xyz returns a table in which 2 columns are the longitude and the latitude of each node of the grid.

Second of all, when the pattern is set, the hourly average of the number of ships can be calculated according to the criteria. Then, the function pygmt.xyz2grd is used to count the number of ships in the blocks with the parameter *registration* = **gridline** this time so that it counts the ships in the block. Then, the function pygmt.grd2xyz returns a table with the coordinates of the nodes in the middle of the blocks and the number of ships associated (Figure 2.4). However with the parameter *registration* = **gridline**, some nodes (which are the middle of the blocks) are in the corner of the grid and so it is impossible to have the coordinates for the top or bottom left corners or right corners. In order to fill correctly the table with the coordinates of the blocks in the grid, some points are not in the grid and have to be deleted. (Figure 2.5) So, now, the table represents the number of ships in each block of the grid and these numbers are written in the matrix with the id of the blocks.(Figure 2.4) At the same time, the table in which there are the coordinates and the IDs of the blocks is filled with the number of ships in the right blocks. (Figure 2.4)

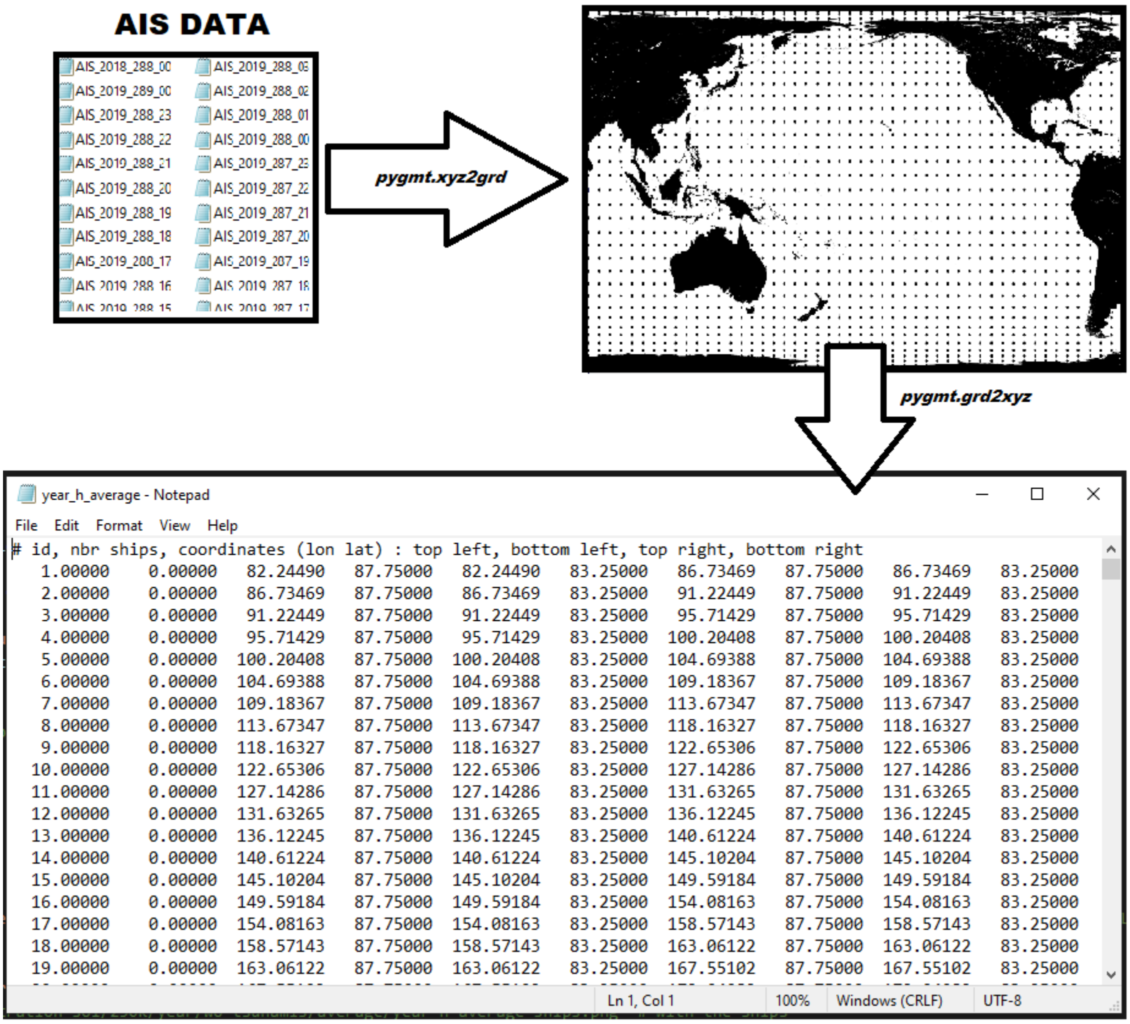


Figure 2.4: The coordinates of the nodes of the grid obtained with the function pygmt.grd2xyz

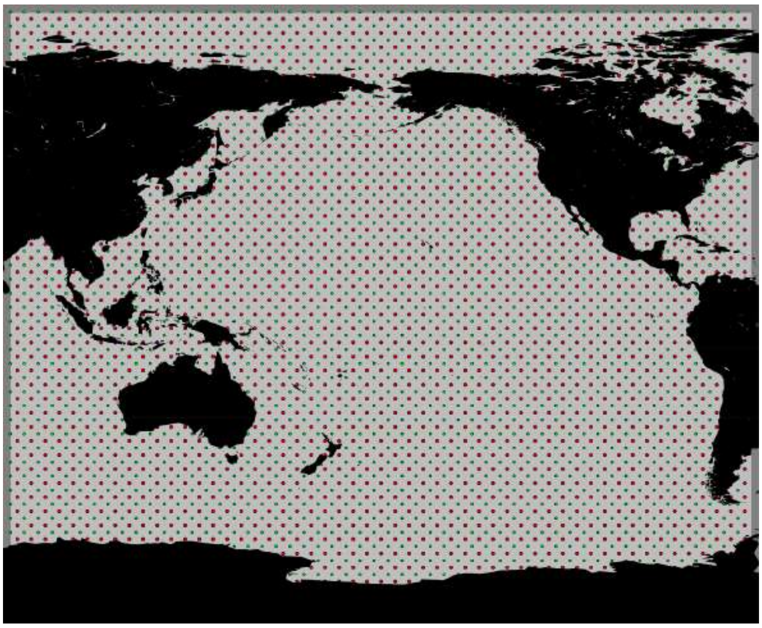


Figure 2.5: The grid obtained after doing the correction (in red: **gridline**, in green: **pixels**)

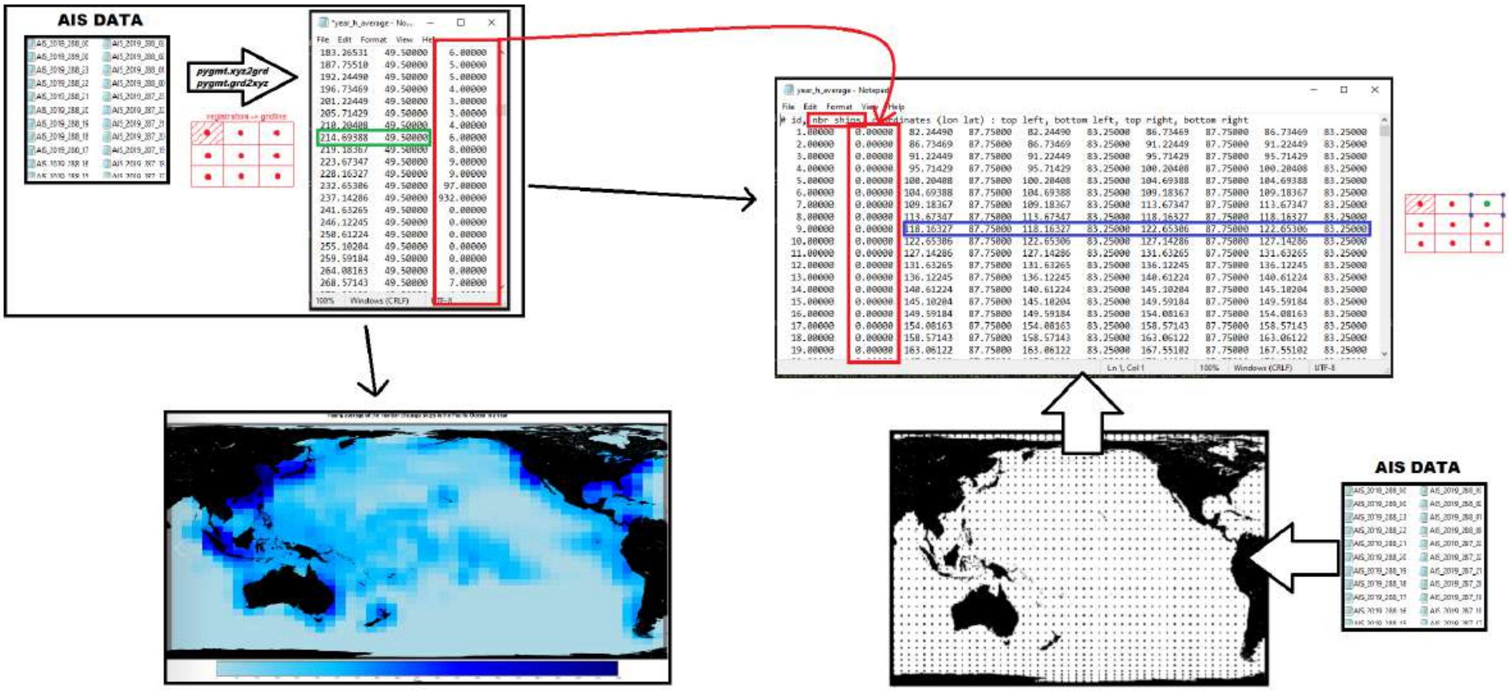


Figure 2.6: Summary diagram of the method used to obtain the maps and the numerical results of the calculation of the average number of ships

2.2 Data processing

The main goal is not only to make maps but also compare and make analysis on the ships' locations in order to improve tsunami early warning systems. Therefore, the more ways there are to compare, the more there are different analyses and so different findings. As the AIS data covers a whole year, the Prof. Dr. James Foster, my supervisor Bruce Thomas, and I discussed about the different possibilities to count the number of ships in the Pacific Ocean. The first criterion is to take into account all the ships in the year and make an hourly average. Thereby, the 8760 AIS data files are read. As in a AIS data file, one line correspond to the coordinates of a ship, so the number of lines correspond to the number of ships. With the function pygmt.xyz2grd, the number of ships, while knowing their coordinates, is counted and distributed according to the blocks created at the same time. Then, per blocks, the total of ships is divided by 8760 to obtain an hourly average.

The second criterion is to make an hourly average month by month. The method is the same as for the year but the AIS data files are read according to the months. As the files cover the year from the 15th October 2018 to the 15th October 2019, the files are separated according to the first day of the month and the end of the month, while taking into account the number of days in the month in 2018 or in 2019 (28 days for February, 30 days for November, April, June and 31 days for the others). Moreover for the month of October, the last 15 days of October 2018 are combined with the first 15 days of October 2019 so that it makes only one month. After counting the number of ships month by month, the total is divided by the number of files read for each month to make an hourly average. So for a 31-day-month, the number of AIS data files read is 744 (31*24) so the total of ships is divided by 744.

The last criterion is to make 4 maps based on the seasons. The method is quite similar as the one used to have the monthly hourly average of the number of ships. However, the AIS data file are read by taking into account the first day and the last day of the season. As for the month of October, days from the year 2018 and the year 2019 are combined to make the fall season: the first days of Autumn go from the 23rd September 2019 to the 15th October 2019 and the last days of Autumn go from the 15th October 2018 to the 20 December 2018. (Table 2.1) The average is also make the same way as for the months.

Dates of the seasons 2018/2019		
Season name	Beginning	End
Winter	21 December 2018 *AIS_2018_355_00*	19 March 2019 AIS_2019_078_23
Spring	20 March 2019 AIS_2018_079_00	20 June 2019 AIS_2019_171_23
Summer	21 June 2019 AIS_2018_172_00	22 September 2019 AIS_2019_265_23
Fall	23 September 2019 AIS_2018_266_00 And 15 October 2018 AIS_2018_288_00	15 October 2019 AIS_2019_288_23 20 December 2018 AIS_2019_354_23

Table 2.1: The dates of the 4 seasons 2018/2019 with their names in AIS data files

2.3 Setting up the display

First of all, before obtaining the method explained before, an initial map is drawn as a test in order to have a first visualization of the distribution of ships in the Pacific Ocean. (Figure 2.7)

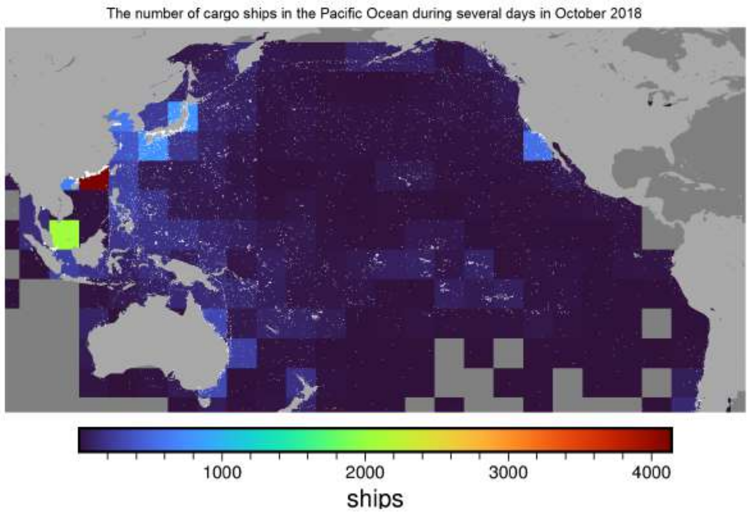


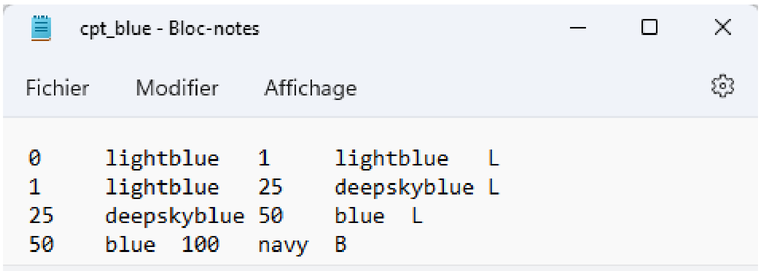
Figure 2.7: The first maps were tests to have a first visualization of the distribution of the ships in the Pacific Oceans

Having multiple colors to show where most of the ships are can be confusing, there are some places where the ships are the most present (near the coasts) and some others where there are no ships at all (South of the Pacific Ocean) and some parameters need to be change (the projection and the scale).

As shown in the initial map (Figure 2.7), the distribution of the ships is not uniform so the main solution to plot a map where the distribution is quite represented is to set a logarithmic scale.

Moreover, in order to obtain colored maps, a Color Palette Table file (CPT) is needed to range the colors according to the number of ships. As the main goal is to have a clear visualization of where the ships are in the Pacific Ocean, the logic of the colors is to have the darker/hotter colors in the blocks where there is a lot of ships and so the lighter/colder colors in the blocks where there is less ships. This logic is similar to heatmaps. However, having multiple colors means having too many information and it can be confusing: these maps don't represent the temperature in the Ocean but the number of ships. Thereby, only blue is chosen: the color blue. Having a gradient of the same color is better to avoid information overload. The color lightblue represents the blocks where less ships are, and the color navy represents the blocks where most of the ships are. The CPT file is useful to contain the colors chosen for the maps and the scale. Indeed, the scale can be adapted, and according to the analysis, the scale can be modified. Here, the scale is set from 0 to 100 because of the gap between the highest value and the lower value that is too huge and cause problem to truncate the colorbar. The CPT file is like a text file so the parameters can be easily written in it. (Figure 2.8)

However, there are blocks in the Pacific Ocean that are coloured in grey.(Figure 2.7) This means that in some places there is no data. It is possible to configure PyGMT, when plotting the maps, the blocks where there is no data to be white (Figure 2.9) or grey (Figure 2.10) or when the average is equal to 0 to be also in white (Figure 2.11). Another option would be to put in red all the blocks that contain more than 100 boats to highlight these areas and color the blocks in grey where there is no data (Figure 2.12). Finally, we decide that, to have a clear visualization, the background color (when there is no data) is lightblue and the foreground color (when there is more than 100 ships) is navy so that we have a homogeneous gradient of blue.



cpt_blue - Bloc-notes

Fichier Modifier Affichage

```
0	lightblue	1	lightblue	L
1	lightblue	25	deepskyblue L
25	deepskyblue 50	blue	L
50	blue	100	navy	B
```

Figure 2.8: The colors and the scale are written this way in the CPT file

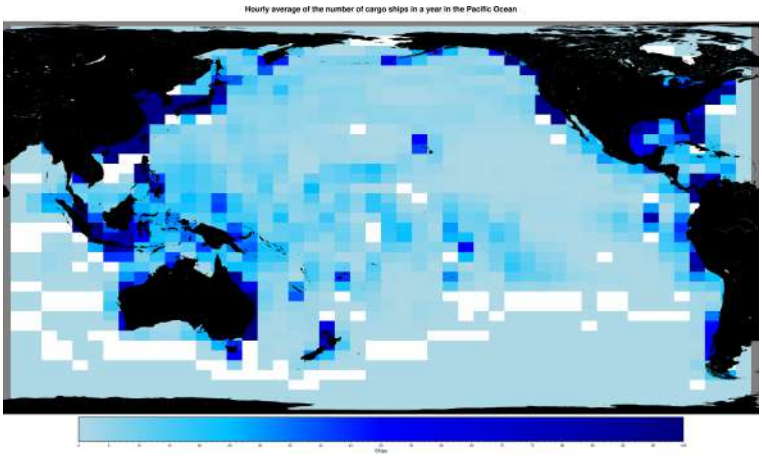


Figure 2.9: The blocks of the grid are white when there is no data.

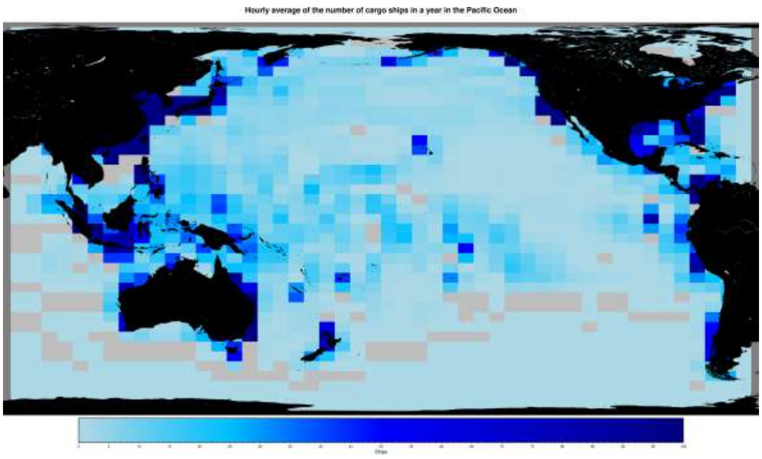


Figure 2.10: The blocks of the grid are grey when there is no data.

All these parameters for the colors and for the scale go to the functions pygmt.makecpt and pygmt.Figure.colorbar. Furthermore, maps are meant to be the most realistic representation to support research, so it is important to have a good projection because otherwise the size of the blocks in the grid is wrong. The projection used is the Cylindrical equal-area. The projection by default is the Plate Carrée but it is not representative of the real distance between each blocks (500 km for this example). Indeed, in order to have the right scale and to represent the most the real localisation of the ships in the Pacific Ocean, it is better to use the cylindrical equal area projection. Now that all the parameters are set, the maps can be plotted.

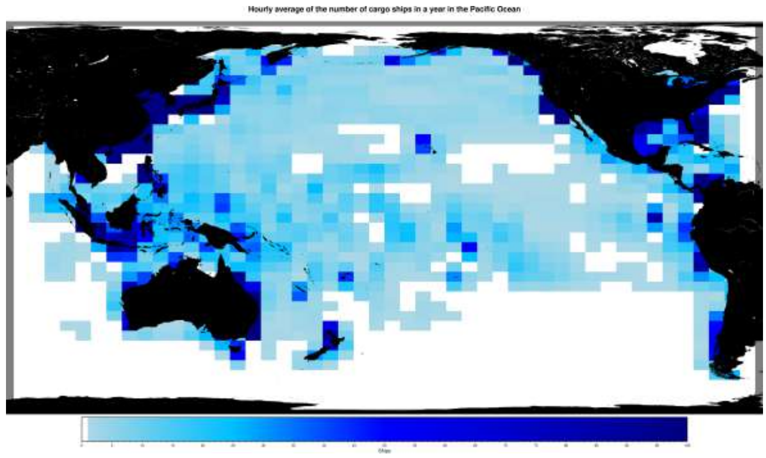


Figure 2.11: The blocks of the grid are white when there is no data and when there is no ships in average.

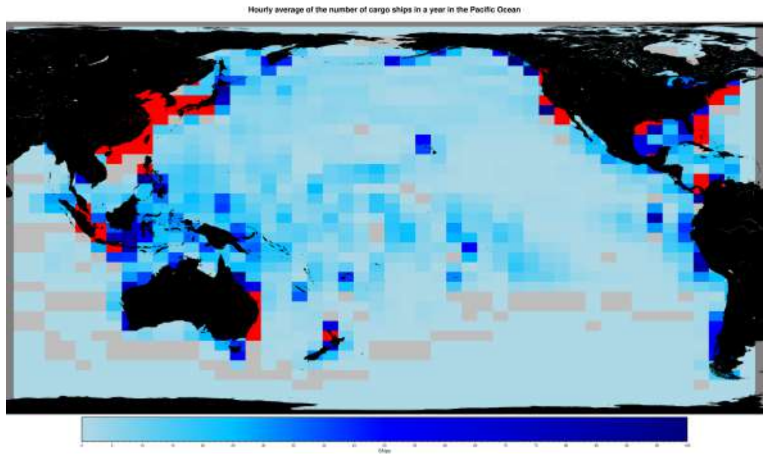


Figure 2.12: The blocks of the grid are grey when there is no data and red when there is more than 100 ships in average.

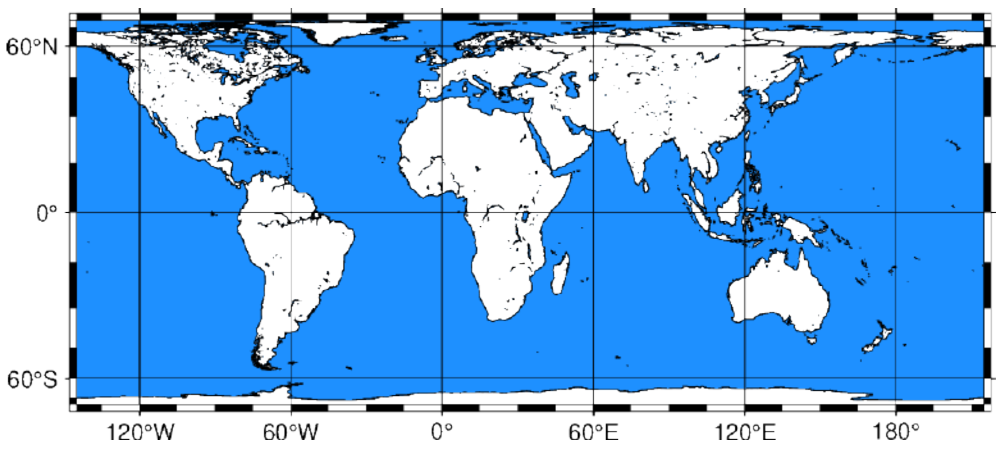


Figure 2.13: The Cylindrical Equal-Area projection used for the mapping

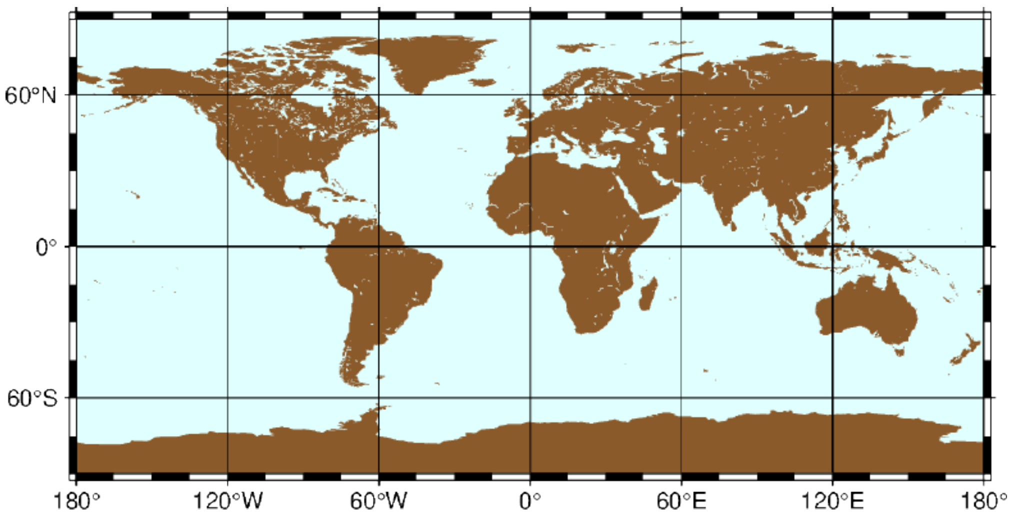


Figure 2.14: The Cylindrical Equidistant (Plate Carrée) projection used for the mapping

Results and analysis

Chapter 3

3.1 General overview

(donner exemple dans le rapport de 250k par exemple !) carte min et max analyse ?? zones vides, mettre en avant endroits ou il y ales sources de tsunamis et la ou il n y a pas la possibilite de suivre les mesures gps car pas de bateaux

At first, the goal is to map the cargo ships with a grid where the blocks have a size of 500km because the DART buoys are set with this distance in the Pacific Ocean. The goal is also to map the annual, the monthly and the seasonal hourly average. The first map plotted is the hourly average of the number of cargo ships in a year.(Figure 3.1)

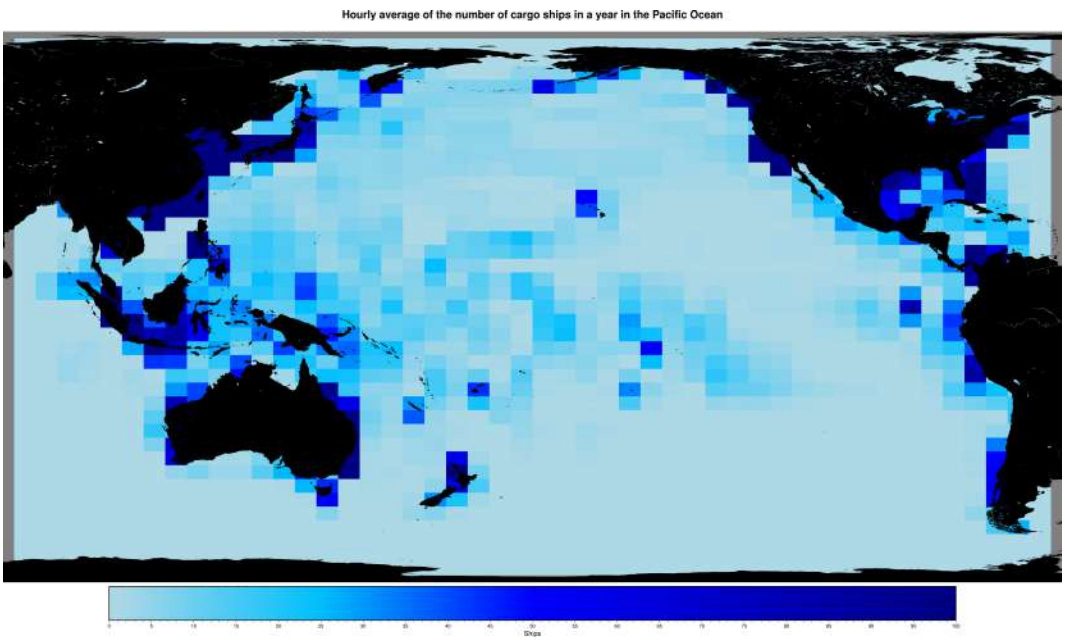


Figure 3.1: The hourly average of the number of cargo ships in the Pacific Ocean in a year

The locations of the ships are also plotted in order to compare and verify if the result drawn is visually correct. On the second map, the ships are plotted as little points in green.(Figure 3.2) Thanks to this map, it is easy to see where the ships are and if it is coherent with the colors of the blocks. Near the coasts in Indonesia, Australia, the USA, China, and Japan, the blocks are in navy color and there is a lot of green points. It means that there is a lot of ships. For instance, near China and Japan, the hourly average is around 10,000 ships in a year (Figure 3.3).

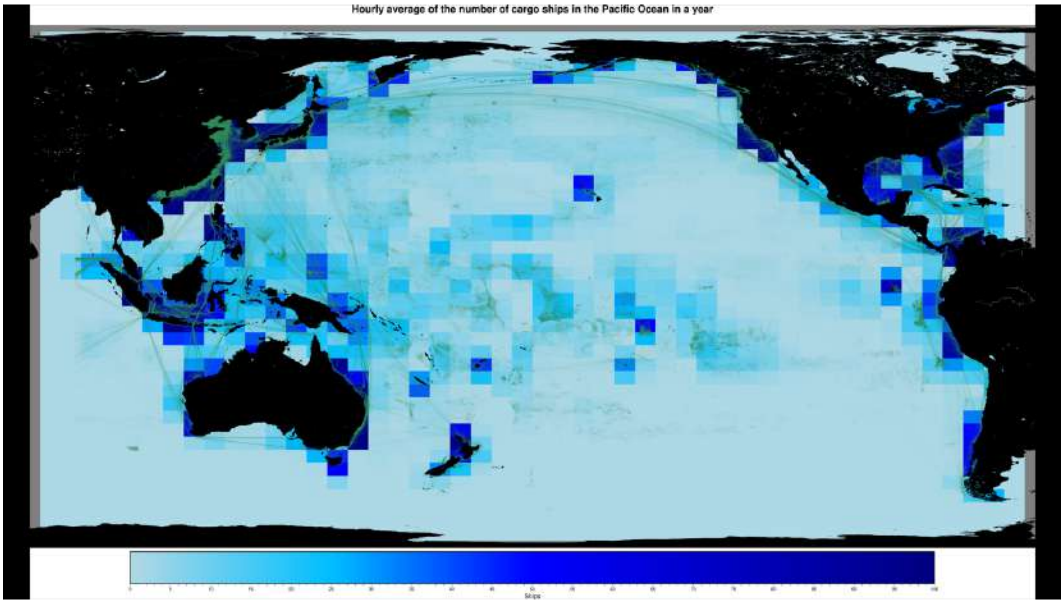


Figure 3.2: The hourly average of the number of cargo ships in the Pacific Ocean in a year with the ships plotted in green

With the website MarineTraffic, it is also shown the distribution of cargo ships in the coast of China.(Figure 3.4) The reason why the hourly average in a year at this location is so important is because there are harbors and a lot of boats stay at the same place during several days. This reason applies on all coasts, such as Indonesia or Japan. It is also viewable on Google Maps with a satellite image.(Figure 3.5) Indeed, East Asia has a large network of commercial ships in the ports but also along the rivers.

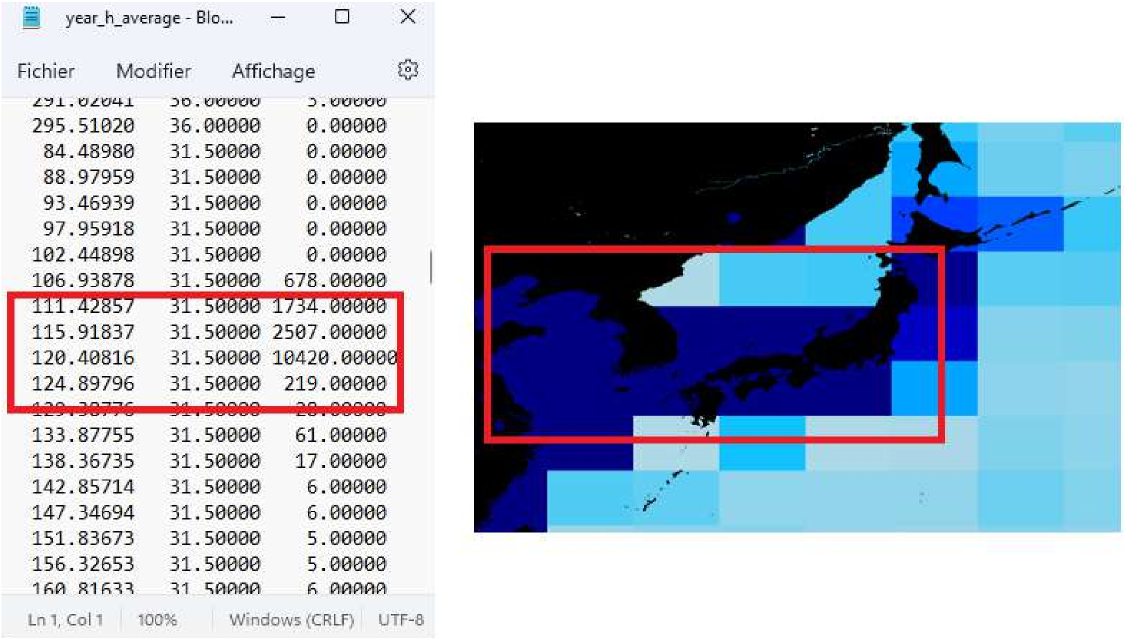


Figure 3.3: The table associated to the map show the hourly average of the number of ships

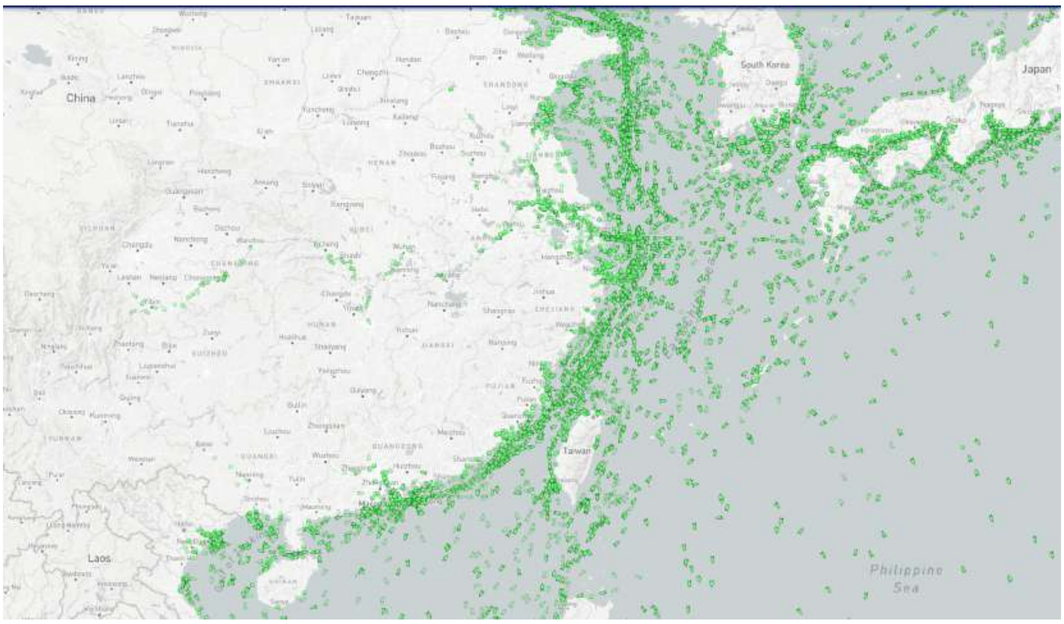


Figure 3.4: The coast of China and the cargo ships in green on the website MarineTraffic

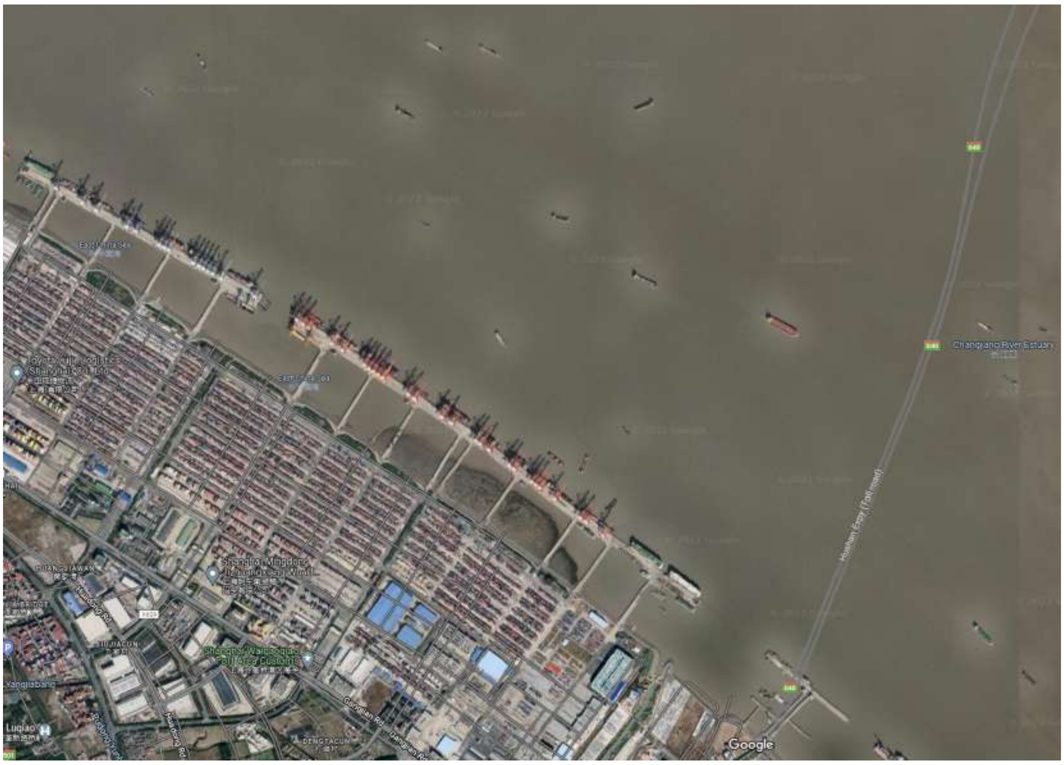


Figure 3.5: A satellite image from the port of Shanghai in China on Google Maps

In order to better represent the distribution of ships in each block, there is a histogram that shows the number of blocks that has a certain number of ships in it (Figure 3.6).

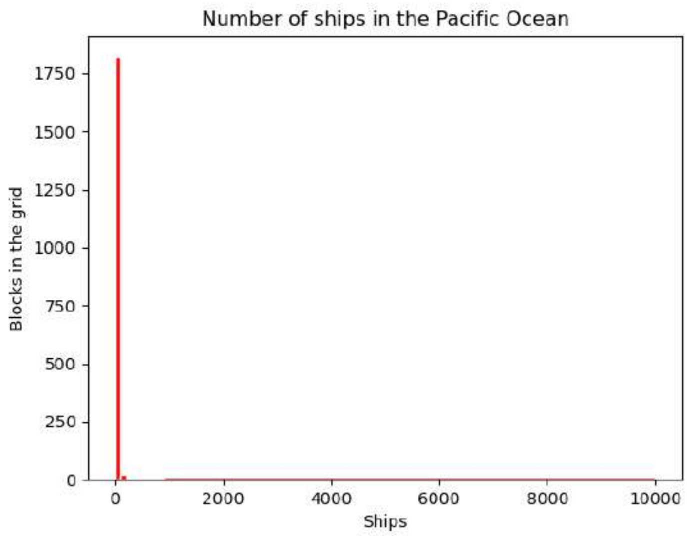


Figure 3.6: Histogram that plots the number of blocks of the grid that have 0 to 10,000 ships in it.

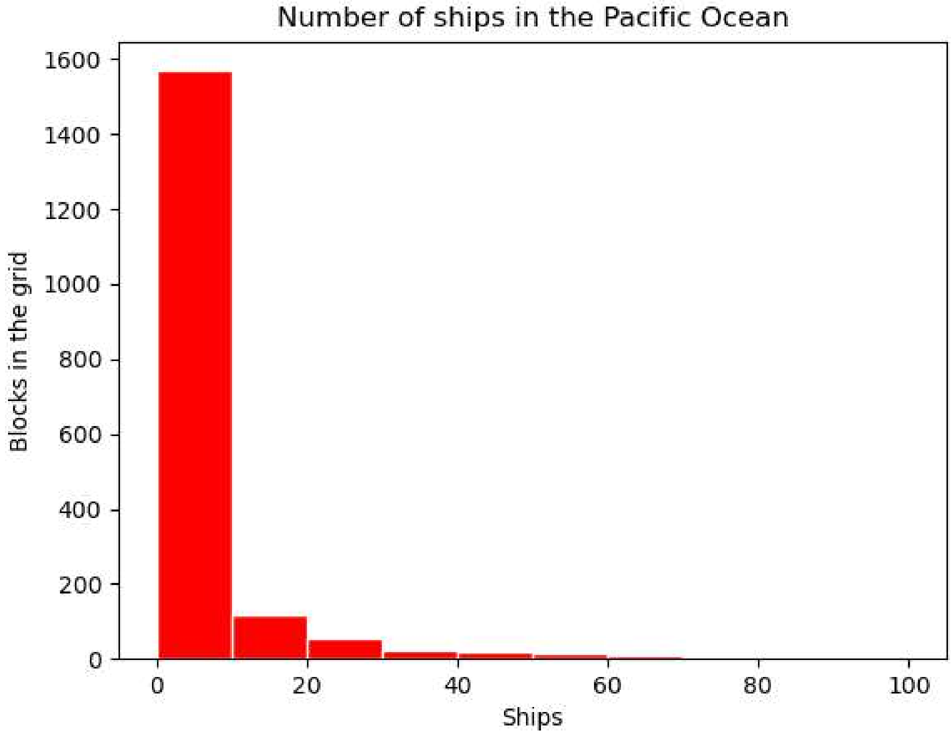


Figure 3.7: Histogram that plots the number of blocks of the grid that have 0 to 100 ships in it.

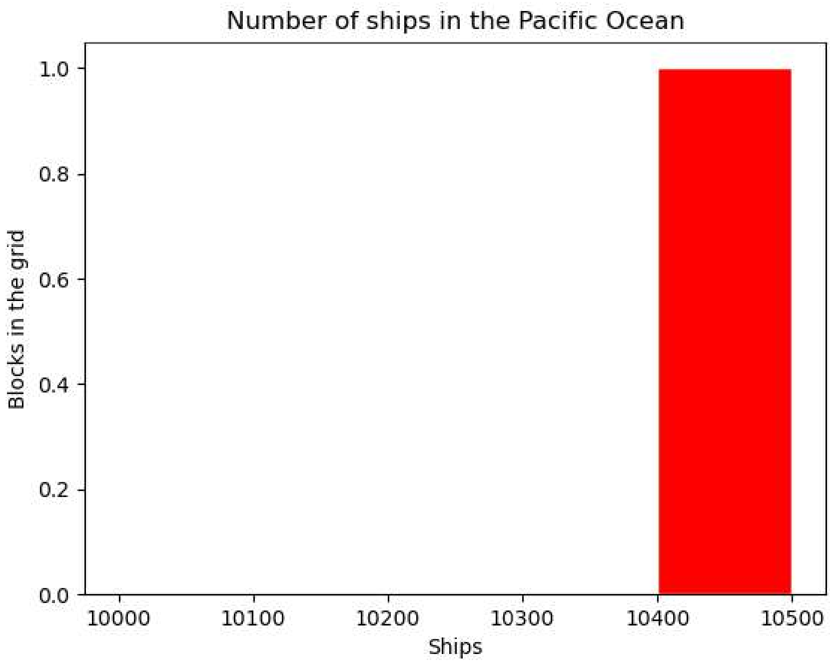
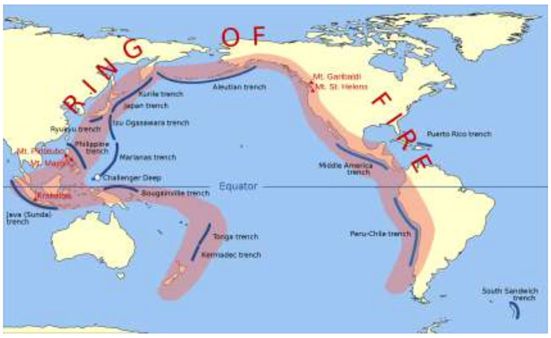


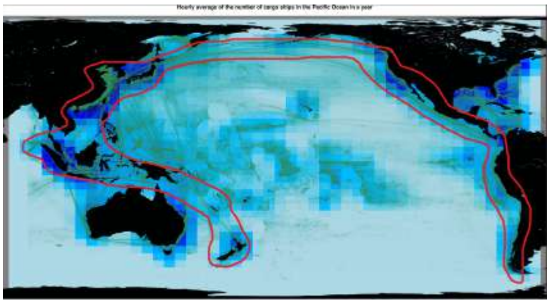
Figure 3.8: Histogram that plots the number of blocks of the grid that have 10,000 to 10,500 ships in it.

First of all, there is a gap: the majority of the blocks contains very few ships. There are at least 1,750 blocks that have less than 100 ships in it and very few that have between 100 and 10,000 ships in it.(Figure 3.6) With a zoom in a little interval of 0 to 100 ships, around 1,500 blocks have a total of ships inferior or equal to 10 ships (Figure 3.7). Also, there is only one block that has between 10,400 and 10,500 ships (Figure 3.8). For a grid with blocks of 500km per 500km, there are 1,872 blocks but at least 90% of the blocks have less than 20 ships in average in a year.

Furthermore, as it can be seen in Figure 3.2, the path that the boats take runs along the Pacific Ring of Fire (Figure 3.9). The Ring of Fire is a volcanic belt where all seismic activity is present and we can see that a lot of ships pass along this line. In general, the network of cargo ships is mainly in the north of the ocean (between north America and east Asia).



(a) The Pacific Ring of Fire.

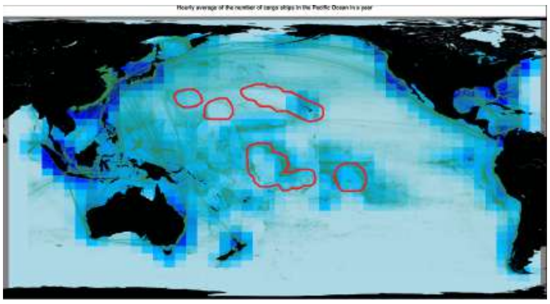


(b) The Pacific Ring of Fire is encircled in red in the final map.

Figure 3.9: The path followed by most of the cargo ships is similar to the Pacific Ring of Fire.

By the way, it is possible to see that there are circular lines around some Pacific islands. Indeed, these are certainly protected areas that the ships bypass (Figure 3.10).

(a) The protected areas in the Pacific Ocean, MPAtlas 2022.



(b) The areas circled in red refer to certain protected areas in the Pacific.

Figure 3.10: There is a correlation between the trajectory of cargo ships and protected sea areas in red circles.

3.2 Seasonal trend

At first sight, with the map of the hourly average number of cargo ships in the Pacific Ocean over the course of a year (Figure 3.1), it is possible to guess the pattern that cargo ships create and where most of the ships are located in a year. This helps researchers a lot because they will know how many ships are in certain locations in a year and decide whether or not it is relevant to track ships in order to detect tsunamis. However, the overview of the hourly average over a whole year does not ensure that at certain locations the number of ships is always the same. Assuming that there are 400 ships on average on the east coast of Japan in a year, there may be 10 ships on average in July and 600 ships on average in November. Therefore, the second and third criterion for mapping cargo ships is to do it by month and by season in order to have another overview of the hourly average. The scale is the same, so that it is easier to compare the number of ships between months or between seasons.

The first results obtained with the maps in December and in July show us that there are darker blue blocks that appear and disappear in the middle of the Ocean. It means that the number of ships between these two months varies a lot, but also in Winter and in Summer.

On one hand, as it is possible to see, the number of cargo ships varies throughout the year. In October, November, December, January and February, there are new horizontal lines of darker blue blocks in the north of the Ocean.(Figure 3.11) These lines are also present in the east of the Ocean in Autumn and Winter.(Figure 3.12) On the other hand, these lines become lighter in the sense that the blocks become light blue again from April (Figure 3.11) and Spring (Figure 3.12). Indeed, according to the gradient of blue, the number of ships increases from Fall, hence the dark blue from October, and decreases from Spring, hence the light blue from April. In other words, there is a seasonal trend of the variation of the number of ships in the Pacific Ocean. There are more commercial ships sailing in autumn/winter than in spring/summer. This trend might be explained with the commercial exchanges due to the end of year festivities such as Christmas or New Year for instance.

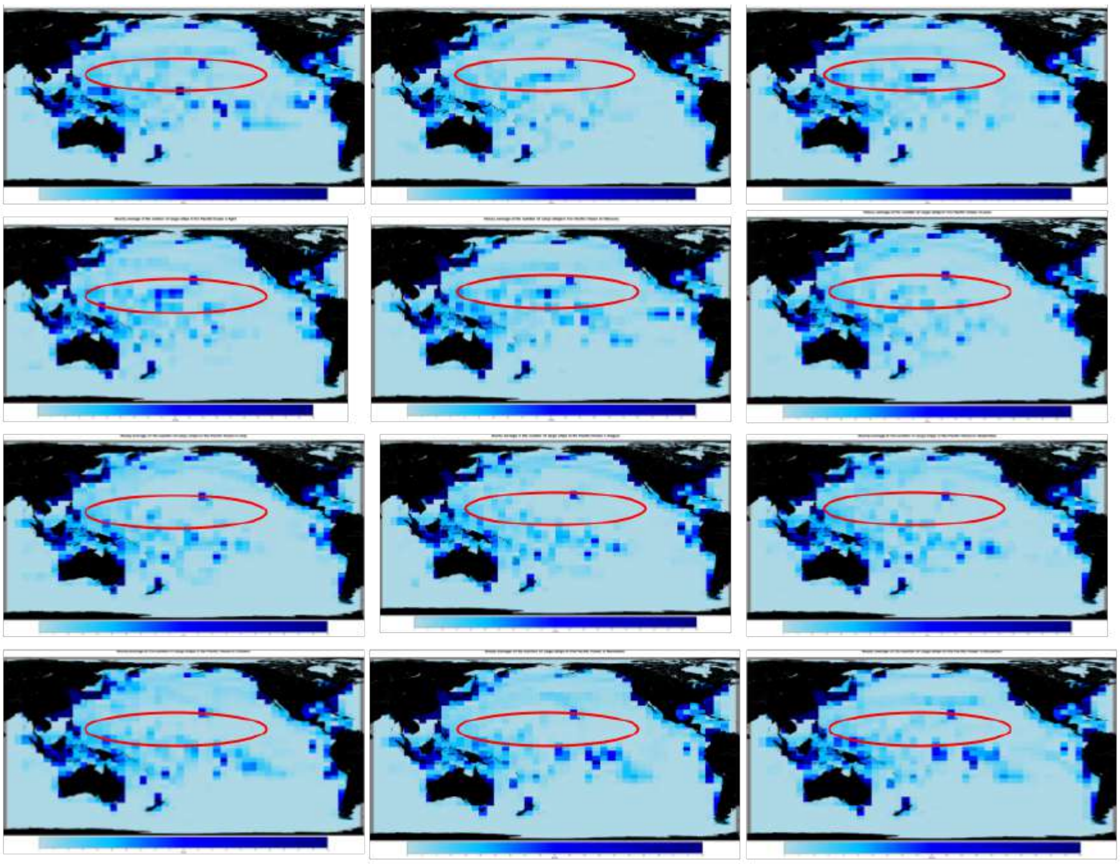


Figure 3.11: The hourly average of the number of cargo ships increases from November and decreases from April

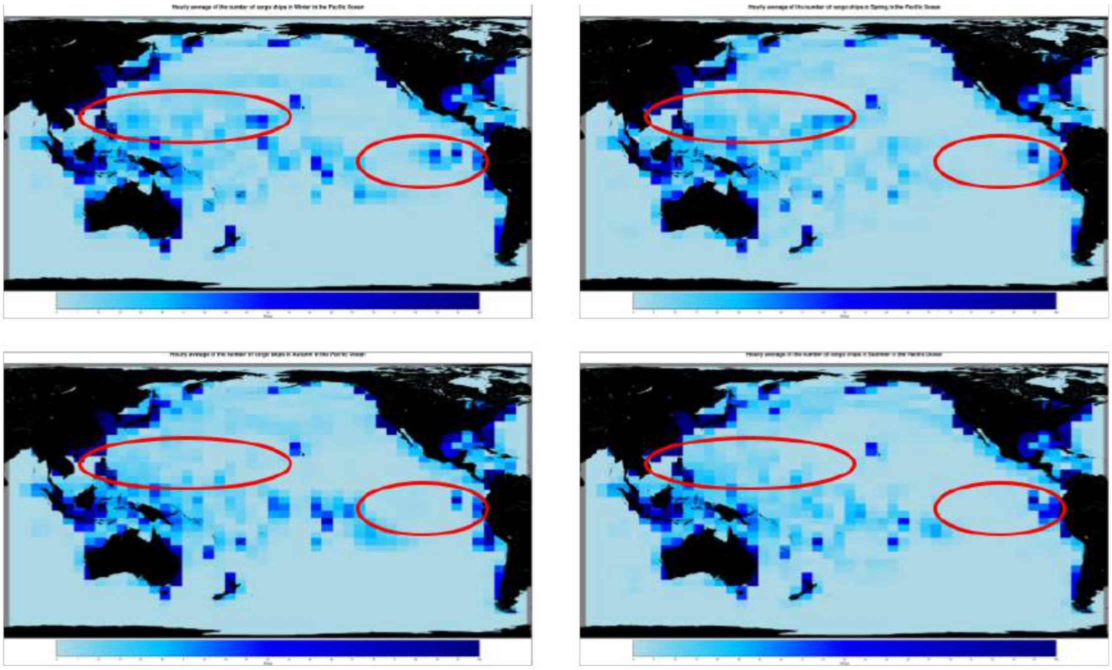


Figure 3.12: The hourly average of the number of cargo ships increases from Autumn and decreases from Spring

3.3 Comparison with the sources of tsunamis

3.3.1 Distribution of tsunamis in the Pacific Ocean

The objective of mapping the cargo ships network in the Pacific Ocean is to improve the tsunami warning system. Knowing the location of most ships is not enough, it must be complemented by fine analysis of comparison with tsunami sources. Thereby, we must overlay the maps obtained before with the locations of the sources of tsunamis.

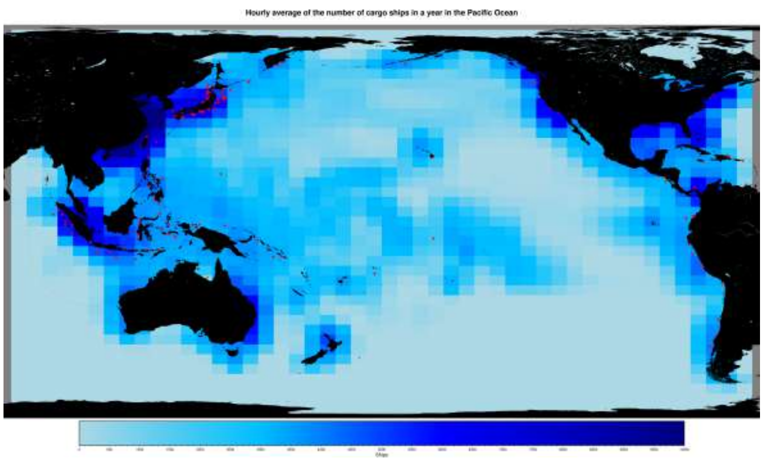


Figure 3.13: The localisation of the sources of tsunamis

The distribution of tsunamis in the Pacific Ocean clearly shows the correlation between the subduction zones in the Pacific Ring of Fire (Figure 3.6) and the locations of tsunami sources. At first sight, we can say that areas where there are many cargo ships are also tsunamigenic areas, which is helpful because most of the cargo ships are in the risk areas.(Figure 3.13) However, most of the tsunamis are located in blocks that contain between 0 and 100 ships.(Figure 3.14)

They have estimated 11% of the commercial fleet that participate in the program in order to give

GPS data and the need of 5 ships to detect and confirm a tsunami. (Foster et al., 2012) So we can approximate by saying we need at least 50 ships per block to detect a tsunami.

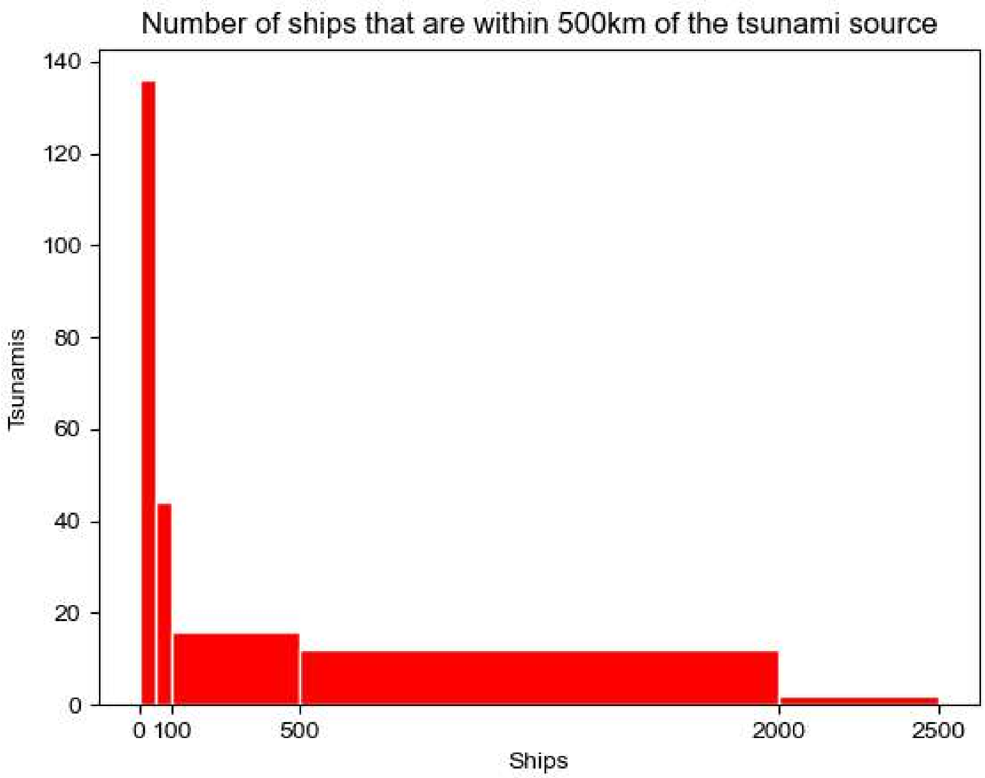
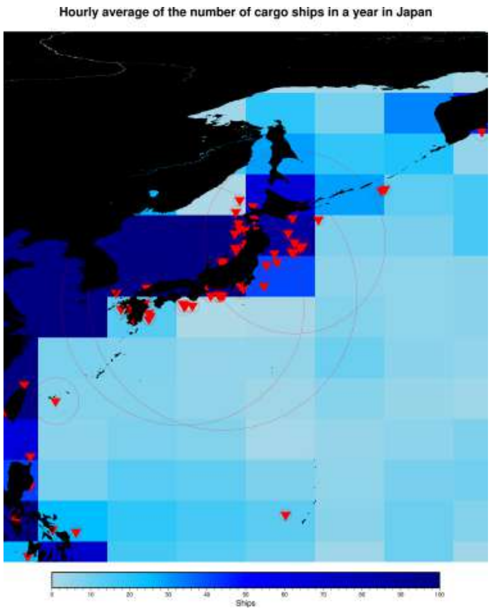


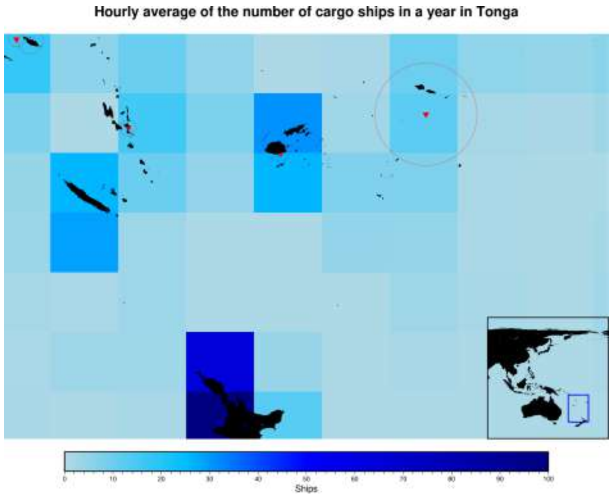
Figure 3.14: Histogram that shows the number of ships that are close to 500km from a source of tsunami

3.3.2 Local zoom

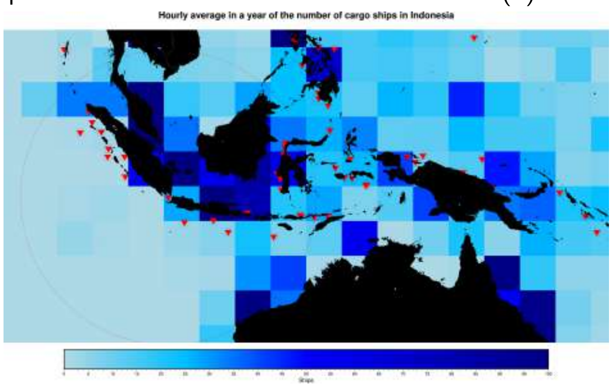
Moreover, the most at-risk areas are zoomed in to see more specifically where the sources of tsunamis are located: in Japan, in Indonesia and in Tonga where a recent tsunami occurred in 2011 and even in 2022.(Figure 3.15)



(a) Zoom in Japan



(b) Zoom in Tonga



(c) Zoom in Indonesia

Figure 3.15: Zoom in the three most tsunamigenic areas in the Pacific Ocean

Before zooming in on these areas, with the global view of tsunami sources (Figure 3.13), we can guess that most tsunamis are located where most cargo ships are. However, by zooming in, we can see that this is not always the case. In Japan, this theory holds, but in Indonesia, the tsunamis are located in the lighter blue blocks (Figure 3.15), which means that the number of ships is very small compared to what we thought.

Conclusion and Discussion

3.4 Conclusion

To conclude, the maps representing the hourly averages of the number of cargo ships in the Pacific Ocean were made, and on different criteria: according to the year, according to the month and according to the season. It is even possible to calculate the standard deviation and the minimum and maximum of the number of cargo ships. But also, there is the possibility to change the main criterion of the size of blocks in the grid, the basis of all calculations of the distribution of the number of boats in the ocean. Thanks to these maps, first analyses were made and it was possible to get some statistics. However, it is still necessary to push the analysis and to make more maps according to the criteria in order to reach a finalization of the tsunami detection system. In a year and with the hourly average criterion, per block, there is an average of 24 ships. But this is not a representative average because there are some blocks where the number of ships is zero. In addition, in order to detect and confirm a tsunami, at least 50 boats are required in a block and indeed, there are 71 blocks that validate this request, mainly located on the coast. (Figure 3.16)

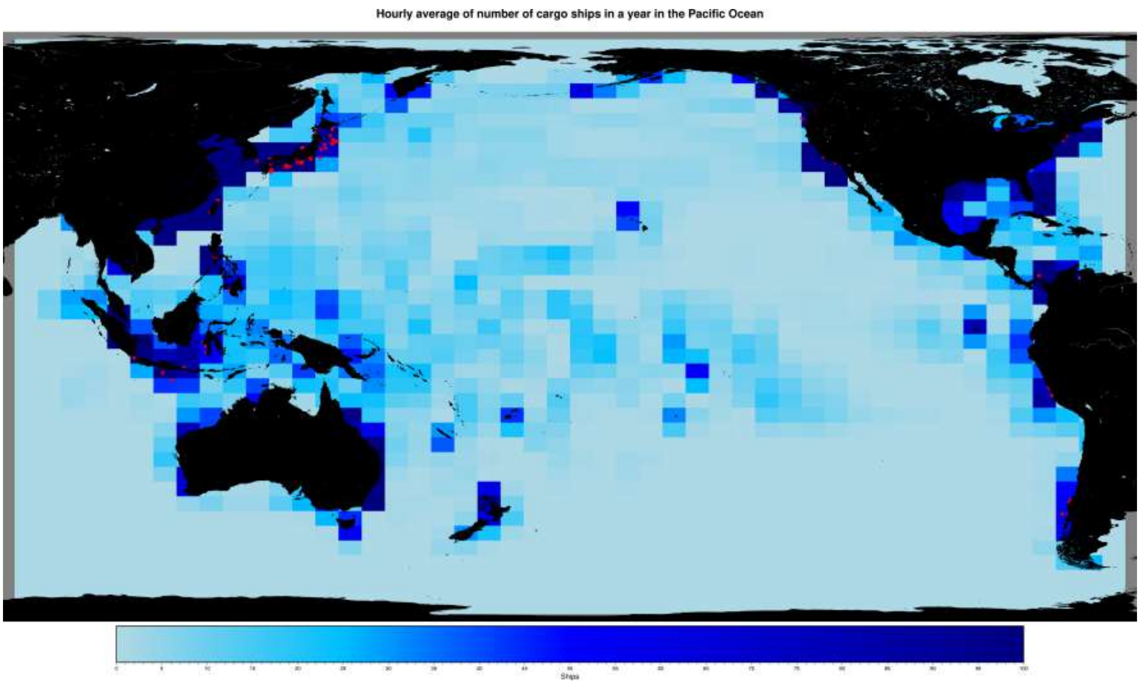


Figure 3.16: The location of tsunami sources in blocks that have at least 50 ships

3.5 Discussion

By comparing the journeys that the cargo ships would probably make and the marked out routes that the ships take on the maps, we can see some incoherence. The source of comparison is the website *MarineTraffic* and it is possible to filter the vessels by seeing only the commercial ones. For example, in the Fiji Islands, the website shows that there are very few commercial boats (Figure 3.17) while on the map where the boats are drawn in green, we can still see that there are a number of them (Figure 3.2). If the filter of the boats on the site allows to select the other boats (fishing and passenger), it is possible to see many more boats having the path described before. (Figure 3.18)

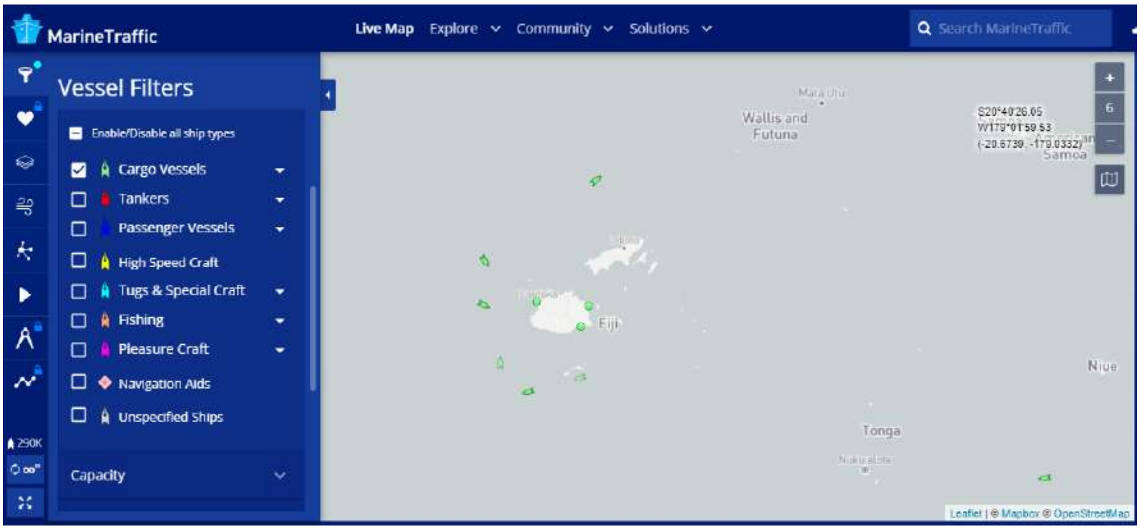


Figure 3.17: The cargo ships near the Fidji Islands, MarineTraffic

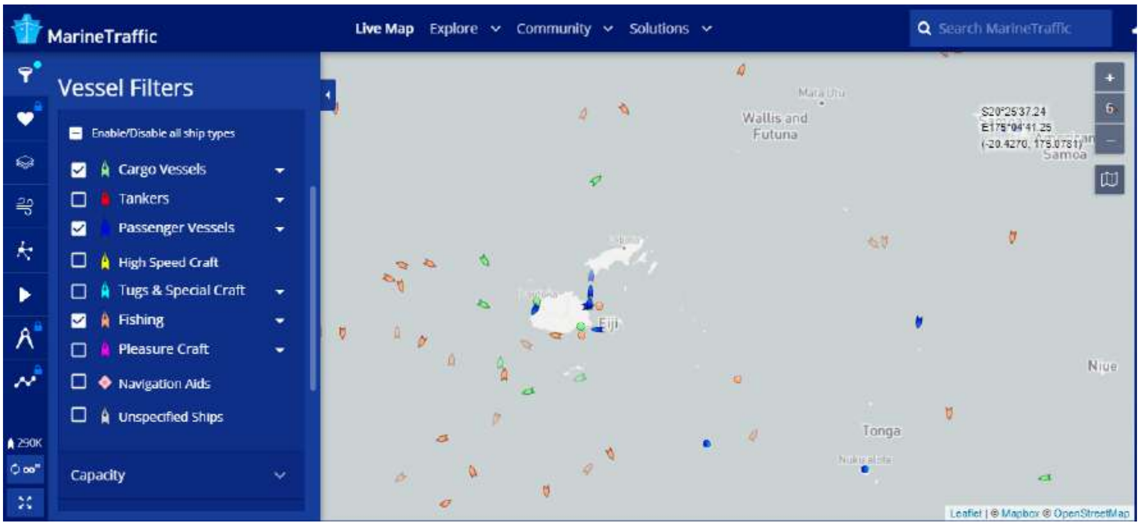


Figure 3.18: The cargo, passenger and fishing ships near the Fidji Islands, MarineTraffic

This is true for other places in the Pacific. So we came up with a theory based on a data problem: either there is a lack of data, or the AIS data does not correspond only to commercial vessels...

Bibliography

Banerjee, P., Pollitz, F. F., & Burgmann, R. (2005). The size and duration of the sumatra-andaman earthquake from far-field static offsets. *Science*, *308*(5729), 1769–1772.

Kubo, M., Cho, I.-S., Sakakibara, S., Kobayashi, E., & Koshimura, S. (2005). The influence of tsunamis on moored ships and ports. *Journal of Navigation and Port Research*, *29*(4), 319–325.

Blewitt, G., Kreemer, C., Hammond, W. C., Plag, H.-P., Stein, S., & Okal, E. (2006). Rapid determination of earthquake magnitude using gps for tsunami warning systems. *Geophysical Research Letters*, *33*(11).

Daud, M. E., Sagiya, T., Kimata, F., & Kato, T. (2008). Long-baseline quasi-real time kinematic gps data analysis for early tsunami warning. *Earth, planets and space*, *60*(12), 1191–1195.

Blewitt, G., Hammond, W. C., Kreemer, C., Plag, H.-P., Stein, S., & Okal, E. (2009). Gps for real-time earthquake source determination and tsunami warning systems. *Journal of Geodesy*, *83*(3), 335–343.

Foster, J. H., Brooks, B. A., Wang, D., Carter, G. S., & Merrifield, M. A. (2012). Improving tsunami warning using commercial ships. *Geophysical Research Letters*, *39*(9).

Löfgren, J. S., Haas, R., & Scherneck, H.-G. (2014). Sea level time series and ocean tide analysis from multipath signals at five gps sites in different parts of the world. *Journal of Geodynamics*, *80*, 66–80.

Tsushima, H., & Ohta, Y. (2014). Review on near-field tsunami forecasting from offshore tsunami data and onshore gnss data for tsunami early warning. *Journal of Disaster Research*, *9*(3), 339–357.

Bernard, E., & Titov, V. (2015). Evolution of tsunami warning systems and products. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*(2053), 20140371.

Inazu, D., Waseda, T., Hibiya, T., & Ohta, Y. (2016). Assessment of gnss-based height data of multiple ships for measuring and forecasting great tsunamis. *Geoscience Letters*, *3*(1), 1–13.

Rochette, J., Binet, T., & Diazabakana, A. (2016). Etude d'optimisation d'une aire marine gérée à.

Inazu, D., Ikeya, T., Waseda, T., Hibiya, T., & Shigihara, Y. (2018). Measuring offshore tsunami currents using ship navigation records. *Progress in Earth and Planetary Science*, *5*(1), 1–11.

Yamazaki, Y., Cheung, K. F., & Lay, T. (2018). A self-consistent fault slip model for the 2011 tohoku earthquake and tsunami. *Journal of Geophysical Research: Solid Earth*, *123*(2), 1435–1458.

Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556–5564.

Hébert, H., Occhipinti, G., Schindelé, F., Gailler, A., Pinel-Puysségur, B., Gupta, H., Rolland, L., Lognonné, P., Lavigne, F., Meilianda, E., et al. (2020). Contributions of space missions to better tsunami science: observations, models and warnings. *Surveys in Geophysics*, *41*(6), 1535–1581.

Inazu, D., Ikeya, T., Iseki, T., & Waseda, T. (2020). Extracting clearer tsunami currents from shipborne automatic identification system data using ship yaw and equation of ship response. *Earth, Planets and Space*, *72*(1), 1–13.

Uieda, L., Tian, D., Leong, W. J., Jones, M., Schlitzer, W., Grund, M., Toney, L., Yao, J., Magen, Y., Materna, K., Newton, T., Anant, A., Ziebarth, M., Quinn, J., & Wessel, P. (2022). *PyGMT: A Python interface for the Generic Mapping Tools* (Version v0.7.0). Zenodo. https://doi.org/10.5281/zenodo.6702566

List of Figures

List of Tables

Annexes

Appendix A

Maps with a grid that has blocks of 500km by 500km

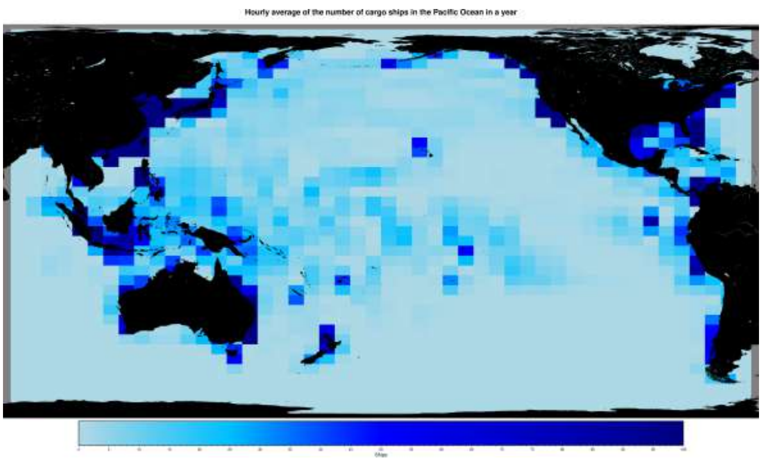


Figure A.1: Map made by taking as a criterion the year and making the hourly average.

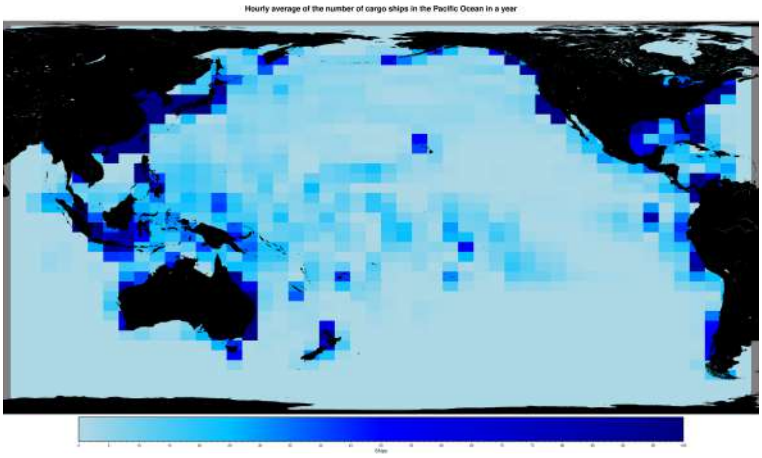


Figure A.2: Map made by taking as a criterion the year and making the hourly average.

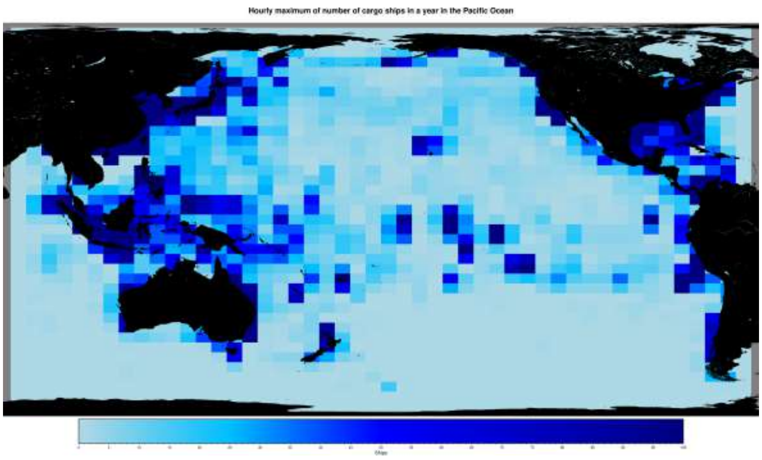


Figure A.3: Map made by taking as a criterion the year and making the hourly maximum.

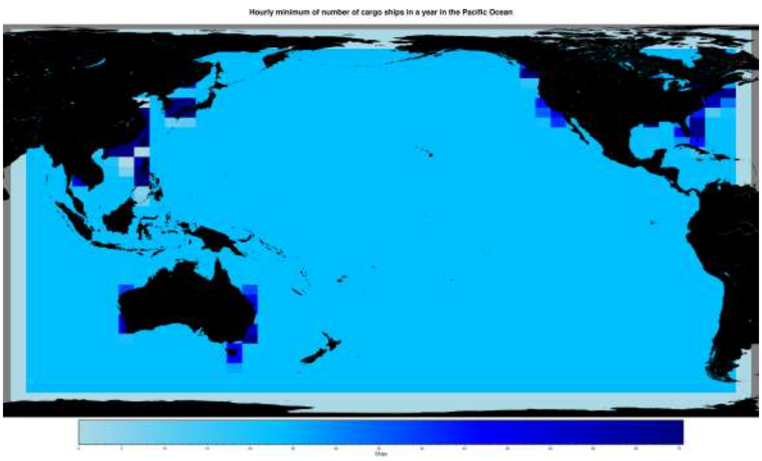


Figure A.4: Map made by taking as a criterion the year and making the hourly minimum.

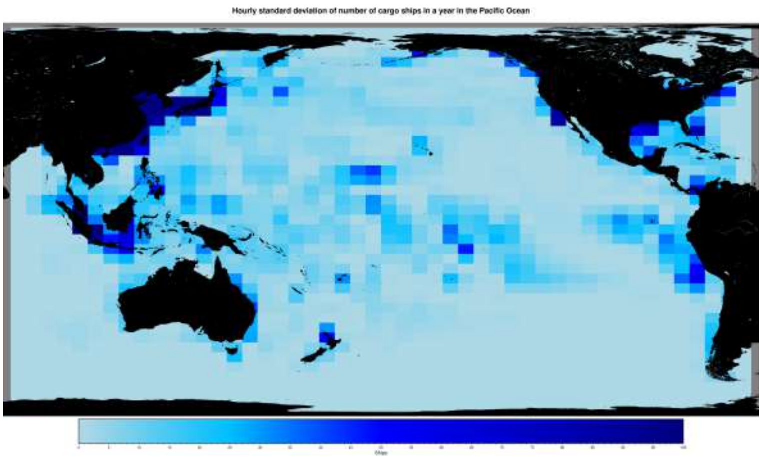


Figure A.5: Map made by taking as a criterion the year and making the hourly standard deviation.

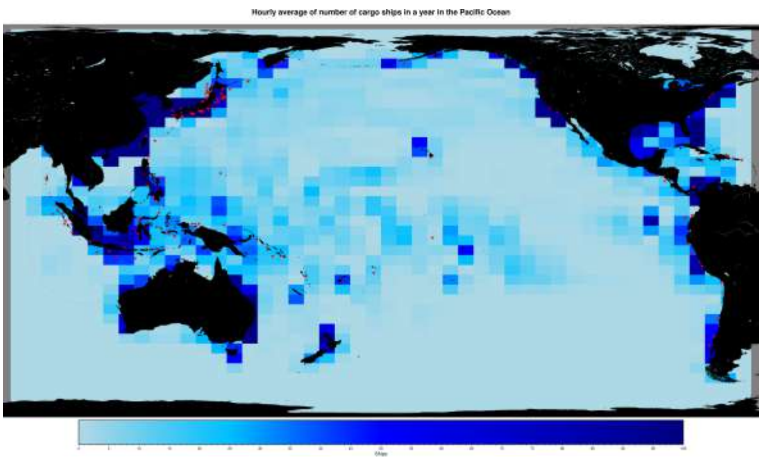


Figure A.6: The tsunamis are plotted in red triangle and the size of the circle depends on the number of deaths.

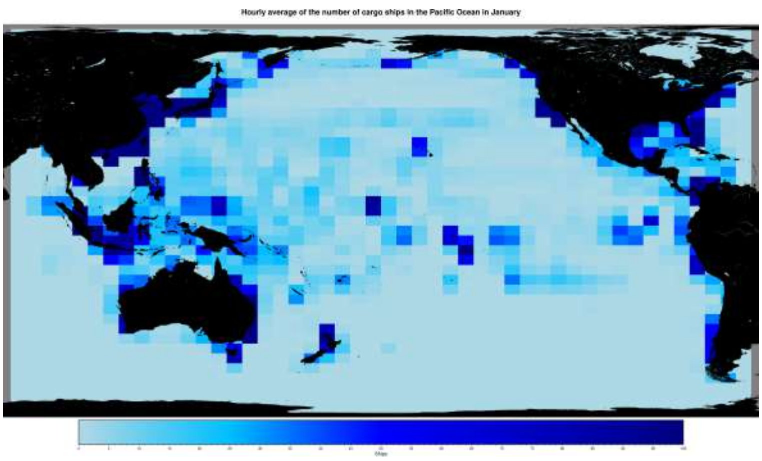


Figure A.7: Map made by taking as a criterion the month of January and making the hourly average.

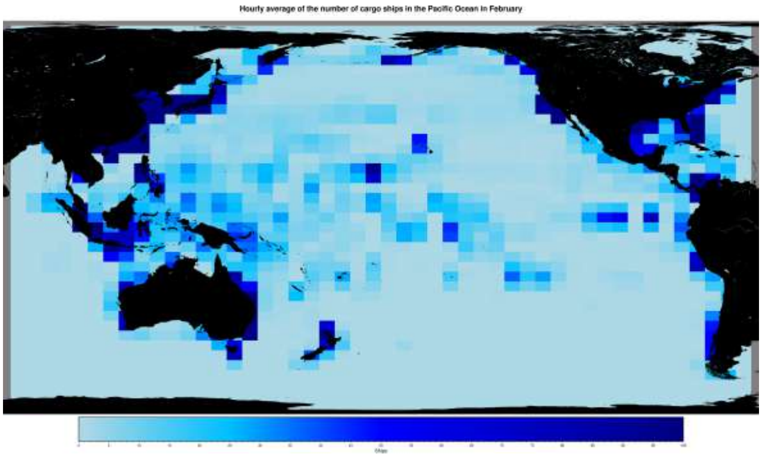


Figure A.8: Map made by taking as a criterion the month of February and making the hourly average.

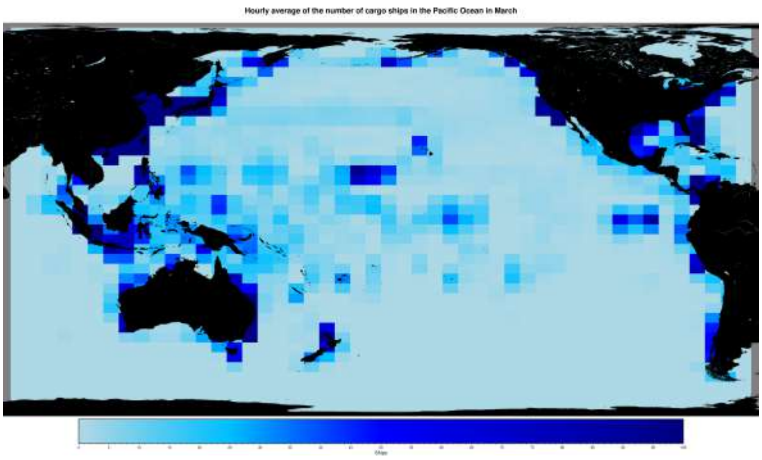


Figure A.9: Map made by taking as a criterion the month of March and making the hourly average.

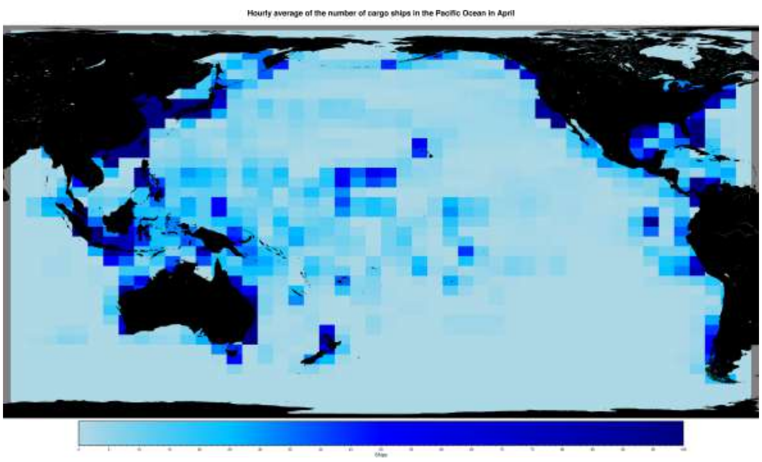


Figure A.10: Map made by taking as a criterion the month of April and making the hourly average.

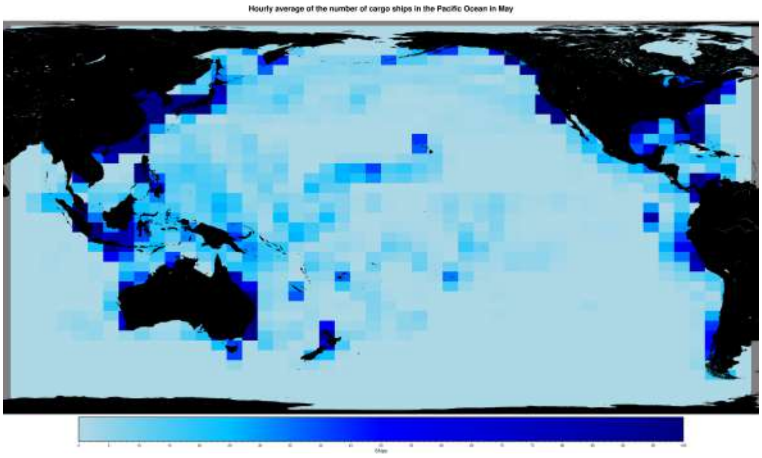


Figure A.11: Map made by taking as a criterion the month of May and making the hourly average.

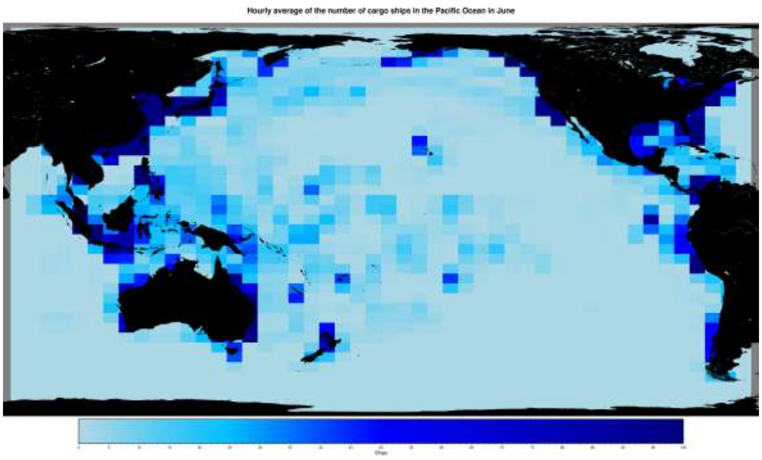


Figure A.12: Map made by taking as a criterion the month of June and making the hourly average.

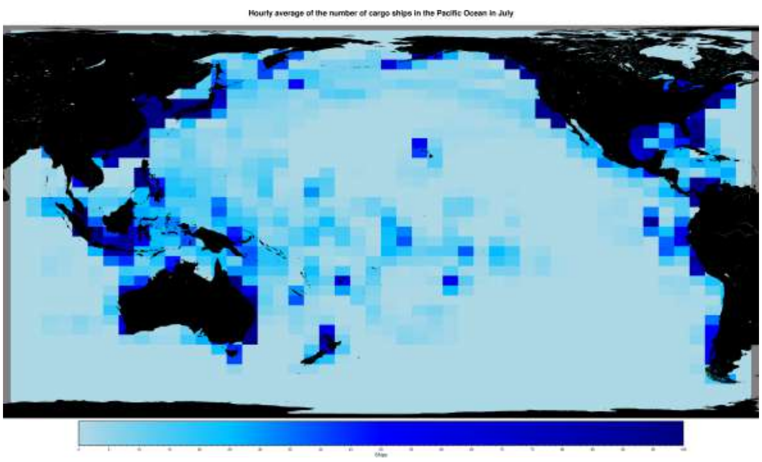


Figure A.13: Map made by taking as a criterion the month of July and making the hourly average.

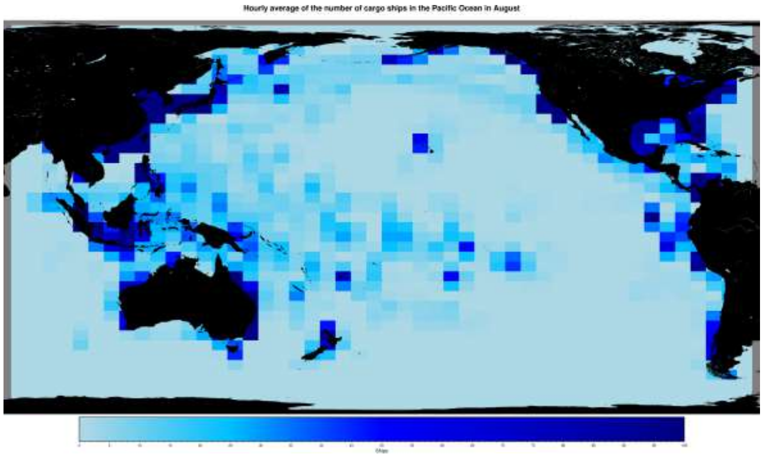


Figure A.14: Map made by taking as a criterion the month of August and making the hourly average.

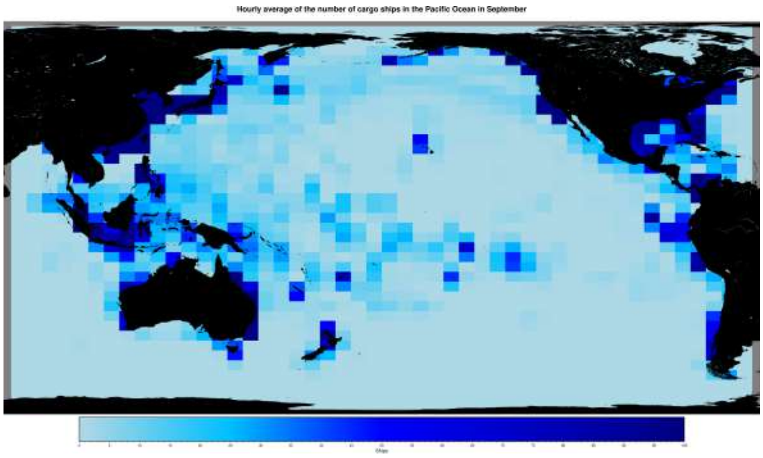


Figure A.15: Map made by taking as a criterion the month of September and making the hourly average.

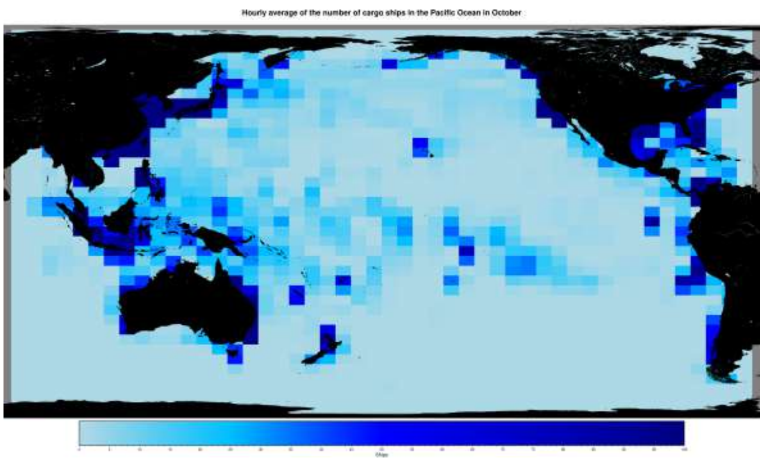


Figure A.16: Map made by taking as a criterion the month of October and making the hourly average.

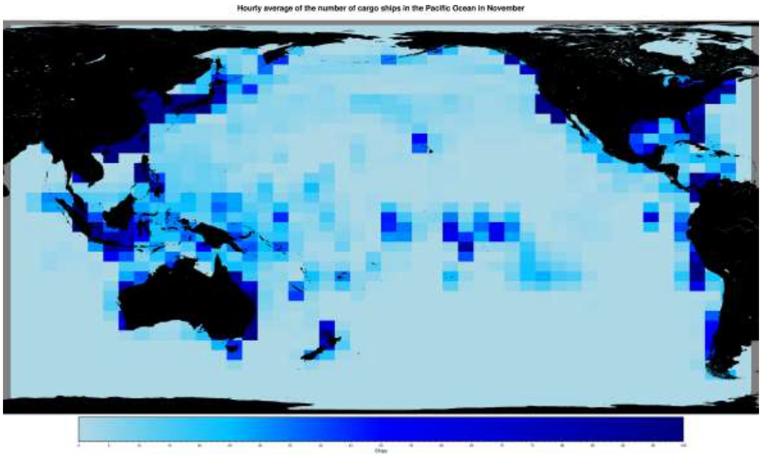


Figure A.17: Map made by taking as a criterion the month of November and making the hourly average.

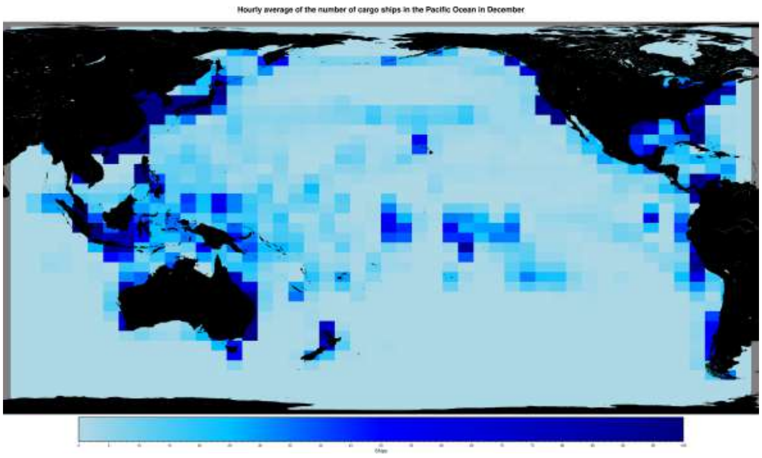


Figure A.18: Map made by taking as a criterion the month of December and making the hourly average.

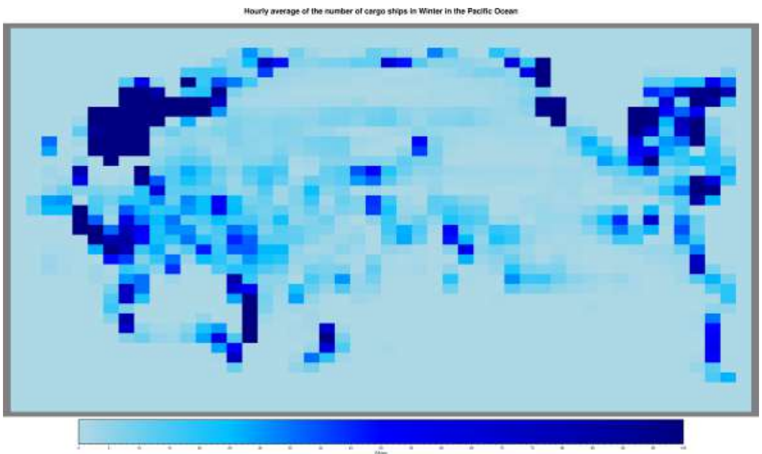


Figure A.19: Map made by taking as a criterion the season of Winter and making the hourly average.

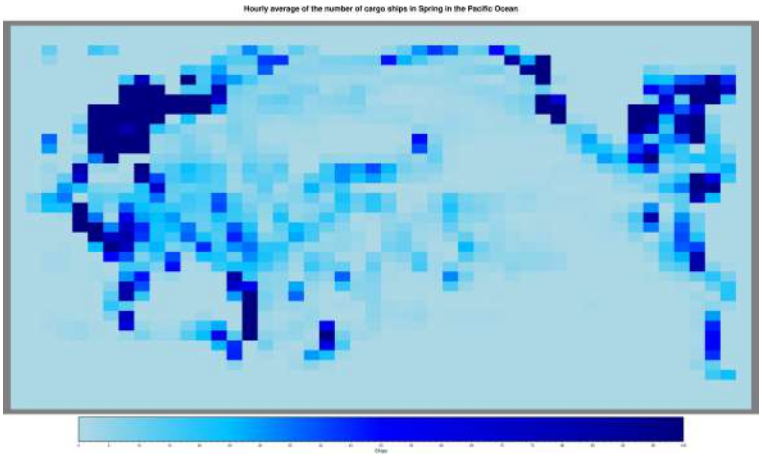


Figure A.20: Map made by taking as a criterion the season of Spring and making the hourly average.

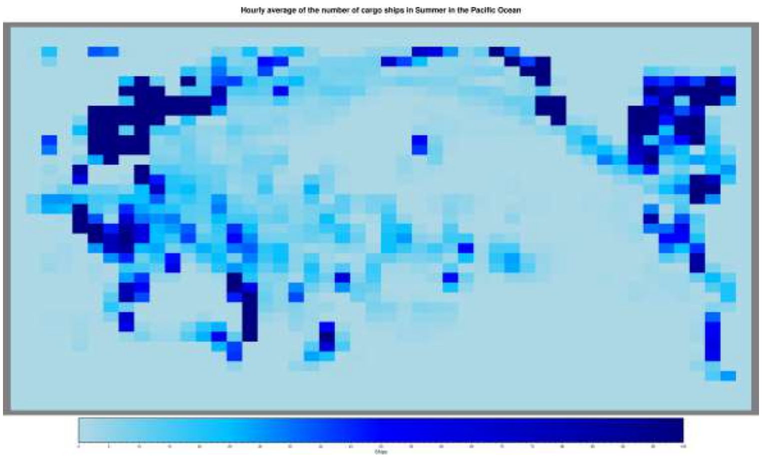


Figure A.21: Map made by taking as a criterion the season of Summer and making the hourly average.

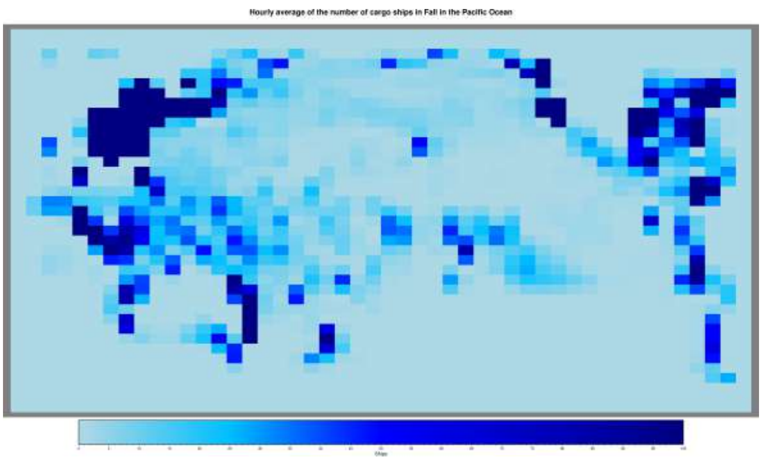


Figure A.22: Map made by taking as a criterion the season of Fall and making the hourly average.

Appendix B
Planning

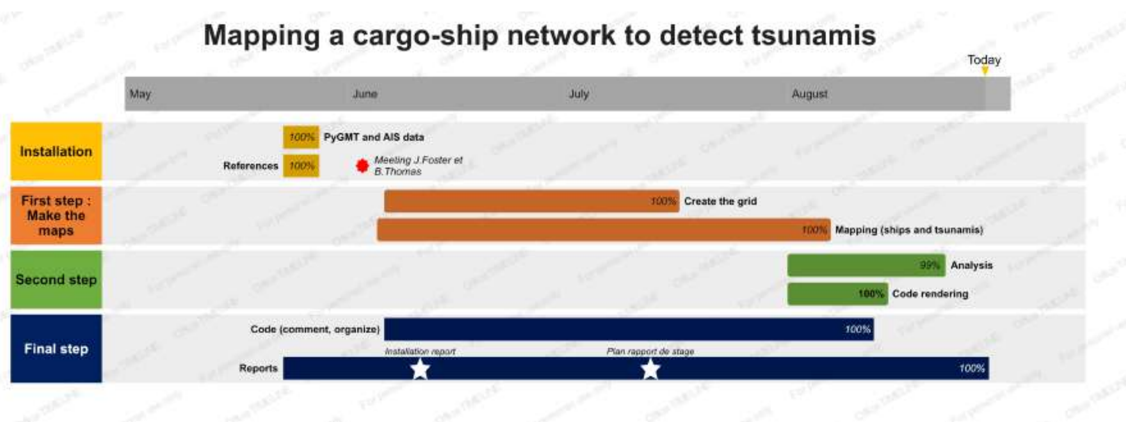


Figure B.1: The distribution of work

Appendix C
Map with a grid that has blocks of 250km by 250km

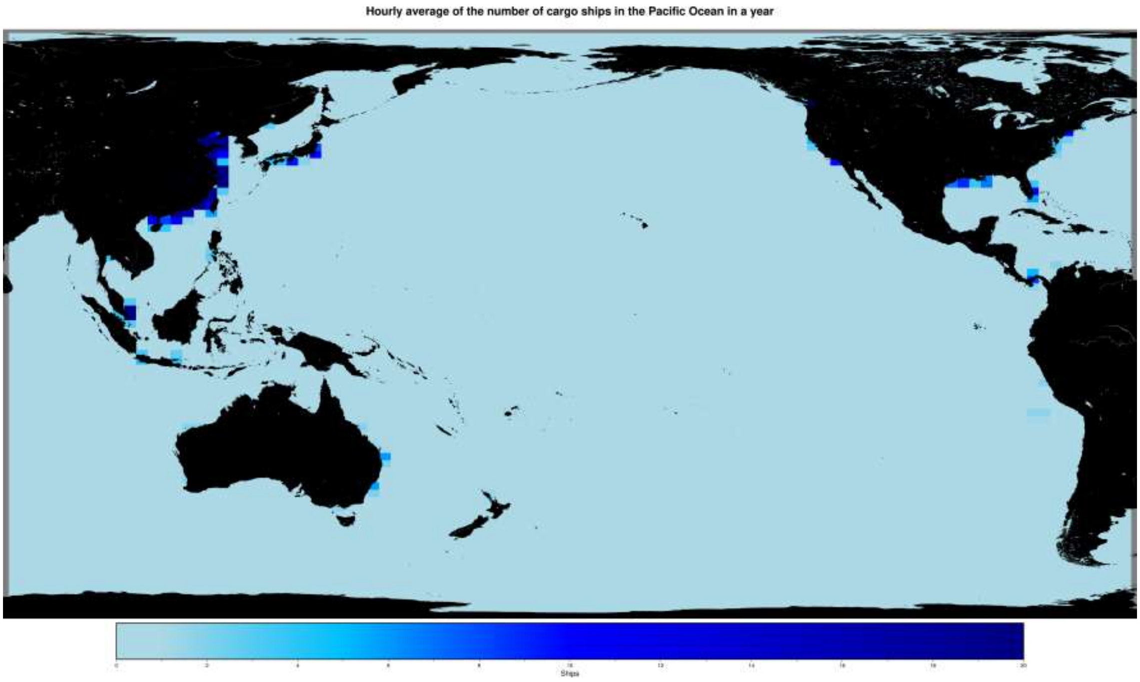


Figure C.1: Map with a grid that has blocks of 250km by 250km