FFI-RAPPORT

16/02061

Density mapping of ship traffic

–

Morten Aronsen
Knut Landmark

Density mapping of ship traffic

Morten Aronsen
Knut Landmark

Norwegian Defence Research Establishment (FFI)

23 November 2016

Keywords

Density map
AIS
GeoPackage
Shape
Vector

FFI-RAPPORT:
2016/02061

Project number:
1415

ISBN:
P: 978-82-464-2826-0
E: 978-82-464-2827-7

Approved by:
Elling Tveit, *Director*
Atle Ommundsen, *Research Manager*

Summary

An increasing number of satellites relay information from the Automatic Identification System (AIS) anti-collision system. This has resulted in a global coverage of ship position data, and several commercial service providers currently offer ship traffic analysis products based on such data. While these services are useful in work to maintain maritime situational awareness, military users may have to avoid them due to concerns about disclosing information about their areas of interest etc. The aim of the present work was to develop software for creating traffic analysis products, for military use, on in-house hardware.

The Norwegian Defence Research Establishment (FFI) maintains a database of information, available to government users, from vessel reporting systems like the AIS, Long Range Identification and Tracking (LRIT), and Vessel Monitoring System (VMS). Data on ship positions and associated metadata, accumulated over time, may be used to create density maps of ship traffic. In this context, a density map is a geographical map where the graphics primitives (pixels or polygons) represent the expected number of ships in an area at a given time, or related quantities. Density values are plotted using color gradients that reveal patterns in ship traffic. In particular, density maps are useful for analysing patterns of life, i.e., habitual behavior.

Providing user-friendly, on-demand density map services presents several challenges due to the large amount of data that must be processed: during a single month, the number of messages received from unclassified data sources approaches one billion. The report describes measures for reducing the amount of data and preparing data for density map products, including aggregation, quantization, and indexing. In particular, a subdivison of the Earth's surface into cells of constant physical size (0.1×0.1 nm^2) is defined. The grid provides an index scheme for position data and a way to quantize ship tracks, thereby reducing the amount of data. For example, a single ship track is represented as the set of grid cells it intersects and the length of time it visits each cell. Moreover, grid cells correspond to graphics primitives (polygons) in density maps, with associated metadata, that can be aggregated in large-scale maps, hence increasing performance and reducing file size.

A range of prototype processing services have been implemented and made available for the Norwegian Defence Joint Head Quarters, the Royal Norwegian Coast Guard, and other users with similar interests. These services create and deliver density map products on demand based on input parameters in network requests from end users. Based on the experience gained from this work, some recommendations for the deployment of density mapping services on military networks are provided. A separate classified document, related to this report, describes military operational use of these types of products.

Sammendrag

Stadig flere satellitter overfører data fra anti-kollisjonssystemet Automatic Identification System (AIS). Dette har gitt et stort tilfang av posisjonsdata fra skip verden over, og flere kommersielle tjenesteleverandører tilbyr nå produkter for skipstrafikkanalyse basert på slike data. Slike tjenester kan være nyttige for å oppnå maritim situasjonsforståelse, men de kan ikke alltid benyttes av militære analytikere grunnet behovet for å skjerme informajon om f.eks. hvilke områder som er av interesse. Forsvarets forskningsinstitutt (FFI) har derfor utviklet programvare for analyse av skipstrafikk for militære formål på interne IT-systemer.

FFI drifter en database med informasjon fra fartøyrapporteringssystemer som AIS, Long Range Identification and Tracking (LRIT), og Vessel Monitoring System (VMS). Posisjonsdata og tilhørende metadata fra fartøyer, samlet opp over tid, kan brukes til å lage kart over skipstrafikktetthet (tetthetsplott). I denne sammenhengen er et tetthetsplott et geografisk kart der grafikken (piksler eller polygoner) viser forventet antall fartøy i et område på et gitt tidspunkt, eller relaterte størrelser. Skipstettheten plottes med fargeskalaer som viser mønstrene og intensiteten i skipstrafikken. Spesielt er slike kart nyttige for å analysere "patterns of life", dvs. normale (typiske) bevegelser, oppførsel og mønstre.

Det er flere tekniske utfordringer knyttet til å lage brukervennlige, fleksible tjenester for plotting av skipstrafikktetthet. Dette skyldes at mengden data som må prosesseres er stor; det registreres nærmere en milliard nye meldinger i databasen hver måned. Denne rapporten beskriver diverse tiltak som er gjort for å organisere og komprimere data for å gjøre tetthetsplotting mindre tid- og beregningskrevende, bl.a. aggregering, kvantisering og indeksering av data. Spesielt innføres en inndeling av jordoverflaten (ellipsoiden) i kvadratiske celler med uniform fysisk størrelse på 0.1×0.1 nm^2. Dette griddet fungerer som en søkeindeks for posisjonsdata. Videre kan fartøybaner diskretiseres (kvantiseres), hvilket innebærer at datamengden reduseres ved at posisjonsnøyaktigheten også reduseres. En fartøybane representeres ved mengden celler den passerer gjennom og tiden (antall sekunder) fartøyet er i hver celle. Hver celle svarer til et polygon i tetthetsplottet, med tilhørende attributter. Polygonene kan slås sammen når kart i stor skala skal produseres. Dette bidrar til å øke ytelsen og redusere filstørrelsen til det ferdige produktet.

FFI har implementert en rekke prosesseringtjenester som er gjort tilgjengelig for Forsvaret Operative Hovedkvarter, Kystvakten og andre brukere med tilsvarende behov. Disse nettverksbaserte tjenestene genererer og leverer trafikktetthetskart etter spesifikasjoner som brukerne setter i sine forespørsler. Basert på erfaringene fra dette arbeidet anbefales det at tilsvarende prosesseringstjenester, for både dynamisk generering av kart og levering av ferdigproduserte standardkart, opprettes på de militære nettverkene. Et eget gradert dokument knyttet til denne rapporten beskriver militær operasjonell bruk av disse produktene.

Contents

1 Introduction

The Norwegian Defence Research Establishment (FFI) maintains a database of information from vessel reporting systems like the Automatic Identification System (AIS), Long Range Identification and Tracking (LRIT), and Vessel Monitoring System (VMS). Data on ship positions and associated metadata, accumulated over time and space, may be used to create density maps of ship traffic. Density maps are graphical representations of spatial or spatio-temporal data. The density value (a numerical quantity) is shown using colors or color gradients, and represents the expected number of vessels in a region at a given time, or related quantities. Such maps are useful tools for visual and quantitative analysis of ship behavior, traffic patterns, and the efficiency of shipping routes. In particular, density maps can be used to determine traffic patterns in sensitive areas, such as military areas of interest or marine protected areas. Density maps based on ship position information, as well as general maps of other attributes, such as vessel speed, are important tools for analysing patterns of life, i.e., habitual behavior.

Ship traffic analysis products are currently offered by several commercial service providers. This is due in part to the present global coverage of ship position data, available from an increasing number of satellites with AIS receivers. However, military users may wish to avoid disclosing their areas of interest or leaving other electronic traces on the internet. Their use of commercial traffic analysis products may be constrained by such concerns. The present work was therefore aimed at developing software for creating density maps, for military use, on in-house hardware. As a result, a range of prototype processing services have been made available for the Norwegian Defence Joint Head Quarters, the Royal Norwegian Coast Guard, and other users with similar interests. These services create and deliver density map products on demand based on input parameters in network requests from end users.

This report describes how the FFI has used sensor data available to government users for creating traffic density maps and other ship data map products. Providing user-friendly, on-demand density map services presents several challenges due to the large amount of data that must be processed: during a single month, the number of messages received from unclassified data sources approaches one billion. The report describes measures for reducing the amount of data and preparing data for density map products, including aggregation, quantization, and indexing. A separate classified document, related to this report, describes military operational use of these types of products.

A brief overview of the different data sources and sensor systems (AIS, LRIT, VMS) is provided in Chapter 2. Density maps can be created for specific classes of ships (or activity), and ships are labeled according to an international ship classification system, the IHS Statcode 5, which is introduced in Chapter 2.4. Sensor data processing and preparations for density mapping are described in Chapter 3. In particular, a subdivison of the Earth's surface into cells of constant physical size (0.1×0.1 nm) is defined. The grid provides an index scheme for position data and a way to quantize ship tracks, thereby reducing the amount of data. For example, a single ship track is represented as the set of grid cells it intersects and the length of time it visits each cell. Moreover, grid cells correspond to graphics primitives (polygons) in density maps, with associated metadata, that can be aggregated in large-scale maps, hence increasing performance and reducing file size. Density mapping is described in Chapter 4, including scale considerations and cell aggregation, image formats, and the definition and interpretation of density as implemented in this project.

Several examples of map product are presented in Chapter 4.4. Based on the experience gained from this work, some recommendations for the deployment of density mapping services on military networks are given in Chapter 5.

2 Sensors and ship information

Density maps created in this project reflect the available sensors in the relevant area. The data used in this project have been limited by area and type in order to produce best results for Norwegian areas of interests. Shore-based systems and data feed from e.g. NATO and other contributors have been limited to include only the northern part of Europe. For other areas it is recommended to include local sensor data in the density maps. The product examples presented in this report show how to generate and use such maps.

The density map products are based on sensor data in combination with ship information and type classification.

2.1 Automatic Identification System (AIS)

AIS is a system primarily aimed at increasing the safety for ships and protection of the environment. According to International Maritime Organization (IMO) rules, ships of more than 300 gross tons are required to carry an AIS transponder broadcasting key vessel data such as identity, speed and heading to other vessels in their vicinity. The AIS periodically transmits information that is received by other ships in the vicinity and by onshore and satellite-based stations[1].

2.1.1 Norwegian land-based AIS-network

The Norwegian Coastal Administration (NCA) has established a land-based network consisting of 44 AIS base stations[2]. The land-based stations cover the area from the Norwegian baseline and out to 40–60 nautical miles from shore. Base stations or AIS receivers are also found on several islands, including Spitsbergen, as well as offshore structures.

2.1.2 Norwegian AIS-satellites

Figure 2.1 *AIS Sat-1.*

[1] http://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.1371-5-201402-I!!PDF-E.pdf

[2] http://www.kystverket.no/Maritime-tjenester/Meldings--og-informasjonstjenester/AIS/

FFI has contributed in developing two orbiting nano satellites (20x20x20cm) equipped with AIS-receivers in order to monitor ship traffic also outside the range of the coastal base stations. Besides monitoring ship traffic in the Norwegian areas of interest, the AIS satellites also help establish a global maritime situational awareness picture. The update rate is less than 90 minutes in northerns waters (depending on the distance between the orbiting AIS satellites). Figure 2.2 shows the orbit of these satellites and the coverage far north[3].

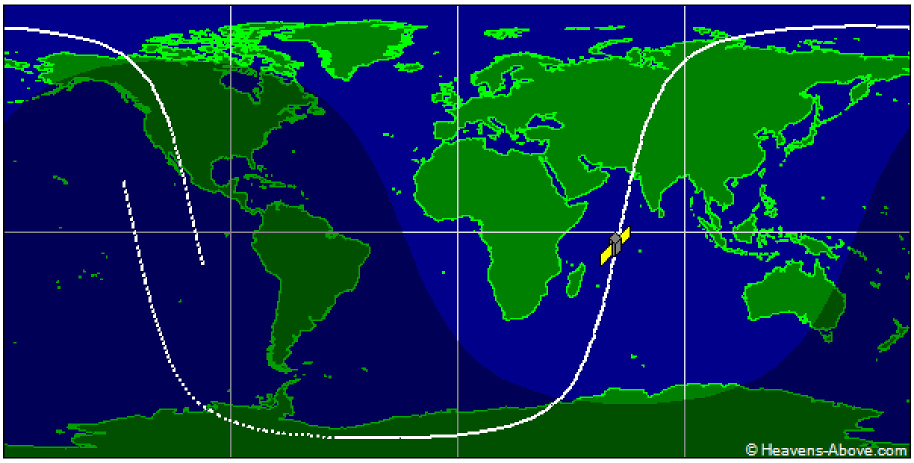


Figure 2.2 *Figure shows AISSat-1 and AISSat-2 orbit.*

2.1.3 International Space Station AIS receiver

FFI jointly with Kongsberg Seatex AS also developed an AIS receiver for the International Space Station (ISS). The coverage of northern waters is limited to approximately 68°north and 68°south. Figure 2.3 shows the orbit of the ISS and its north/south limitation[3].

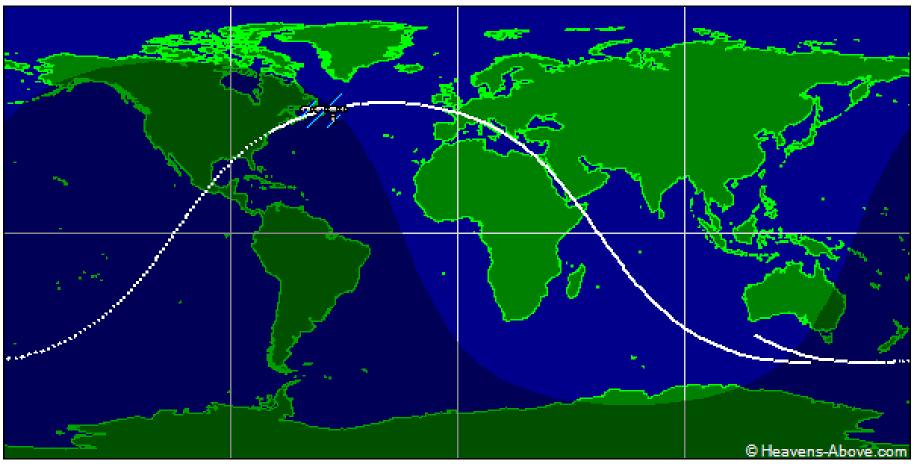


Figure 2.3 *ISS orbit.*

[3] http://www.heavens-above.com/

2.2 Vessel Monitoring System (VMS)

VMS is a system for tracking and monitoring the activities of commercial fishing vessels. The main function is to monitor the movements of fishing vessels (position, course, and speed), but it is also used for communicating catch reports.

The use of VMS is regulated by different regimes in different parts of the world. In the North-east Atlantic and Barents Sea the North-East Atlantic Fisheries Commission (NEAFC) requires fishing vessels to report their movements and catches through the VMS. The five contracting parties of NEAFC are Denmark (including Faroe Islands and Greenland), Iceland, Norway, Russia and the European Union (EU). The frequency of position/movement reports is minimum once per hour.

FFI receives VMS information through the Norwegian Directorate of Fisheries. The messages are according to the North Atlantic Format (NAF) format[4], but are converted to the AIS-format locally at FFI for further use.

2.3 Long Range Identification and Tracking (LRIT)

LRIT is a global satellite-based system for identification and monitoring of ship traffic. The system was introduced in Norway in 2009 and is a closed system for national governmental agencies[5]. IMO established the system in 2006, which applies to the following ship types engaged on international voyages:

- All passenger ships including high-speed craft,
- cargo ships, including high-speed craft of 300 gross tonnage and above, and
- mobile offshore drilling units.

Fishing vessels and military vessels are exempted from using the LRIT system.

FFI receives LRIT messages from NCA as AIS messages minimum every 6 hours (maximum every 15. minute)[6]. LRIT data from Norwegian ships are received from all over the World, while data from foreign ships are only received when they are bound for a Norwegian port or travelling within 1000 nm of Norwegian waters.

2.4 IHS Fairplay Statcode 5

IHS Fairplay Statcode 5 version 1081 classification coding system is an industry-standard method of coding up and describing ship types[7][8]. The coding system has five levels describing the ship types, and each level has a code containing letters, numbers, or both, and a description of the level type. Further details can be found in sub-chapters 2.4.1-2.4.3.

[4] http://www.naf-format.org/

[5] http://www.kystverket.no/Maritime-tjenester/Meldings--og-informasjonstjenester/Havovervakingssystemet-LRIT/

[6] http://emsa.europa.eu/operations/lrit.html

[7] http://www.ihsfairplay.com/about/imo_standards/s5download.html

[8] https://www.ihs.com/newsletter/maritime-information/innovations-vol1-2011/statcode.html

2.4.1 Statcode level 1

The first level is basic ship description:

- A - Cargo carrying
- B - Work vessel
- W - Non seagoing merchant ships
- X - Non merchant
- Y - Non propelled
- Z - Non-ship structures

2.4.2 Statcode level 2

Level 2 specifies the ship type in more used terms combined with the first level classification:

- A1 - Tankers
- A2 - Bulk carriers
- A3 - Dry cargo/Passenger
- B1 - Fishing
- B2 - Offshore
- B3 - Miscellaneous
- W1 - Inland waterways
- X1 - Non merchant
- Y1 - Non propelled
- Z1 - Non ship structures

2.4.3 Statcode level 3 to 5

The next levels further classify ship types in more detail. In Statcode 5 version 1081, level 3 has 31 unique classifications, level 4 has 136 unique classifications, and level 5 has 320.

Example:
B11A2FS - Stern trawler
B – Working vessel
B1 – Fishing
B11 - Fish catching
B11A - Trawler
B11A2FS - Stern trawler. A vessel for catching fish by trawling with nets handled over the stern.

3 Data handling and gridding

NCA distributes AIS data from base stations and satellites as well as LRIT data to governmental users. FFI has access to these data as well as VMS position reports for research purposes. The amount of raw data is very large; for example, the raw sensor data from June 2015 consist of 776.264.445 messages (each stored as a row in a database table). In addition, interpolated data are introduced to estimate the true ship voyage path when there is a significant time and space gap between position reports. The AIS coverage and update frequency varies depending on whether there is base station coverage or only data from LRIT and AIS satellites.

Making density maps based on this large body of data is computationally intensive and may take much time, while users require up-to-date and relevant products. Density maps delivered as dynamic services, where users may define area and period of interest, ship type and other attributes, must have short response time. To achieve this, data must organized and reduced to manageable size.

Figure 3.1 shows the amount of data rows (messages and grid cells) in the different steps of organizing the data in time and space. This chapter describes the various processing steps.

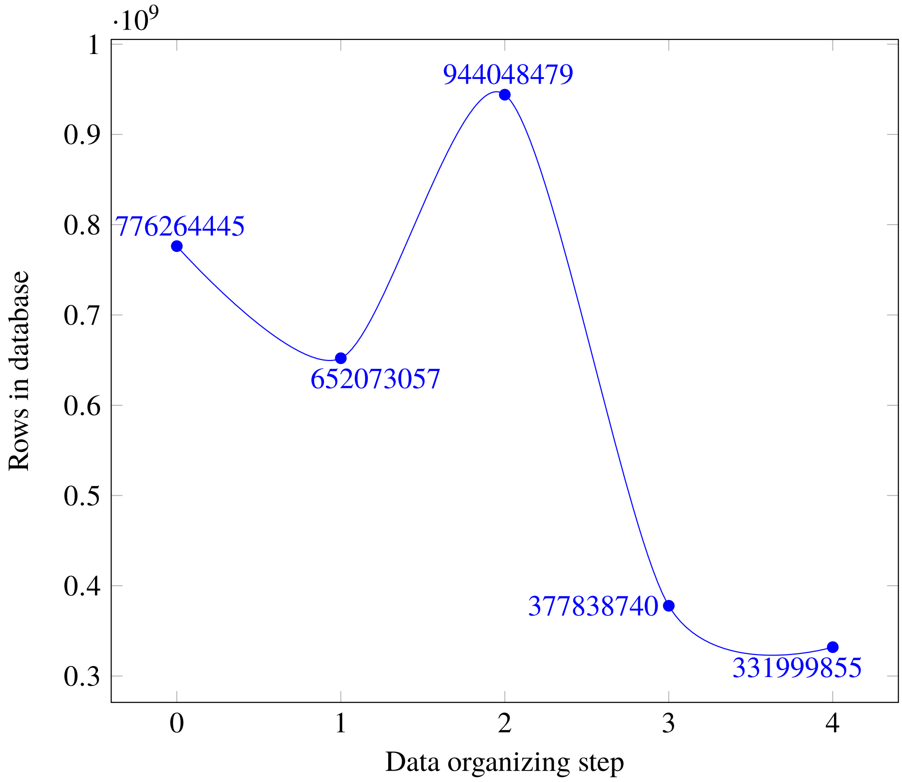


Figure 3.1 Data reduction while organizing in time and space, June 2015

Visualising density patterns in coastal waters, perhaps only a harbour or short stretch of coastline, requires sufficiently high resolution in time and space to discern features like small fjords, inlets and harbour areas. Thus the smallest space resolution was set to a square cell size of one square cable length (1/10 nautical mile or 185.2 meters). Three levels of data processing have been applied in order to enrich and reduce the data:

1. Remove erroneous positions that are reported on land (dry cell).
2. Introduce interpolated cells based on position gaps.
3. (Main gridding process) Extract and aggregate desired attributes (such as time, sensor type, Maritime Mobile Service Identity (MMSI), speed, navigation status, etc.) into hourly data and spatially into cells of one square cable length.
4. Aggregate certain attributes into monthly data for specific use based on end user feedback.

3.1 Dry/wet filter

A reported position may be wrong due to, e.g., signal noise or a faulty GPS receiver. Such errors may cause apparent large jumps in a vessel's position. The same phenomenon may occur when two distinct ships report using identical MMSI numbers, or when a vessel's position is recorded at extended intervals. The latter typically happens when AIS transmissions are received only during the passage of satellites. In all these instances, a vessel's estimated (interpolated) position (Section 3.2) may cross dry land, and such erroneous positions should be removed from the database.

To accomplish this, a global, binary dry/wet raster dataset was created. The raster grid corresponds exactly to the density map grid of 1/10 nm resolution described in Section 3.3. Each pixel (basic cell) has value $p = 1$ if the cell contains some water (ocean, lake, river, canal) and $p = 0$ otherwise. The grid has 216000×108000 pixels and is stored as a 30 MB GeoTiff file with internal compression and one bit per pixel. Checking whether a reported or estimated position is on land is a fast look-up operation. A position is first binned to the density map grid and subsequently rejected if the corresponding pixel in the dry/wet grid is zero.

The dry/wet grid is based on two data sources, the 30 m Global Land Cover dataset by the US Geological Survey (USGS)/University of Maryland [1], and the World Vector Shoreline database (WVS) [2]. The USGS dataset has been derived from Landsat 7 imagery and is the primary source for our dry/wet grid. WVS data are used whereever land cover data is missing. The dry/wet grid was initially computed as a tiled mosaic in a geographical reference system. The mosaic covers all areas containing some land, except the south polar cap, latitude $< -88°$ (due to the coordinate singularity at $-90°$). Each tile is $4° \times 2°$ in extent, and the physical resolution (pixel size) is uniform across the globe (100 m). Each tile was produced by first gridding WVF data, resampling USGS data, and merging the two. Next, the dry/wet grid was initialized with value $p = 0$ for latitude $< -88°$ and $p = 1$ everywhere else. Finally, for each tile in the geodetic mosaic, each pixel was transformed to the dry/wet grid and the corresponding cell value updated.

Figure 3.2 shows the global dry/wet data using the density grid definition. Figure 3.3 shows a single tile from the geographically referenced dry/wet mosaic (100 m resolution). The USGS dataset in particular provides a high level of detail. The dataset contains some noise (errors), but is still useful for removing erroneous ship positions. For June 2015 the data is reduced by approximately 15% before being further processed.

Figure 3.2 Global dry/wet raster dataset using the same grid definition as the density grid.

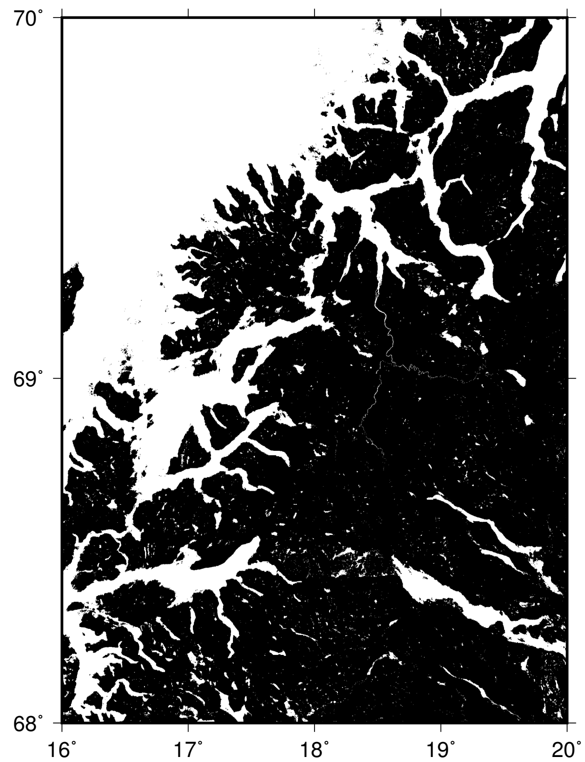


Figure 3.3 Single tile from the geographically referenced dry/wet mosaic with 100 m resolution (from northern Norway).

3.2 Interpolated positions

The distance between two positions reported from a ship may exceed the smallest space resolution (cells of one square cable length). In order to best represent the voyage of a ship, the path between to consecutive positions is calculated and segmented into positions with a distance that is less than 0.1 nm. By introducing these positions density values can be calculated along a ships path rather than only its reported positions, see Figure 3.5–3.9 for more details. To avoid interpolated cells being introduced on land the segmented positions between the two points are all checked against the dry/wet filter. Some segments between consecutive reported positions cross dry cells (land). For June 2015, two million (2.056.932) segments were discarded for this reason to avoid introducing erroneous interpolated positions.

3.3 Main gridding process

The third processing level is a data reduction process. It consists of parsing all raw and interpolated data, one hour at a time, and storing the visit time for every unique MMSI number and unique cell. In other words, within a one-hour interval a vessel track intersects certain grid cells, and the estimated time spent in each cell is stored in the database. Each grid cell has a row and column number that identifies its geographical position, and every grid cell has the same physical size (185.2×185.2 m). Row and column numbers are computed with respect to the origin at 0° N, 0° E. The grid is sparse, i.e., only a small number of cells are visited by ships in any given one-hour interval, so only non-zero (visited) cells are stored in the database.

By virtually gridding the world in cells of identical size this way one may efficiently transform between cell indices and geographical coordinates (latitude,longitude). The grid definition is shown in Figure 3.4. A position within 185.2 meters from the origin in positive direction, east and north, gets column index (X-column) 1 and row index (Y-row) 1. Cells are numbered from 1 (-1 in opposite direction), so the cell index (0,0) is not used. On the equator the number of cells in east-west direction is 216000. In north-south direction the number of cells is 108000. The number of cells in east-west direction depends on latitude due to the ellipsoidal shape of the earth (meridional convergence). The algorithm for computing grid cell indices for a given geographical position (longitude,latitude) is shown in Listing 3.1.

Listing 3.1 *Source code for computing grid cell indices for a given geographical position.*

```
public void ClosestGridPoint(double longitude, double latitude, out int xcol, out int ↩
    yrow)
  {
    xcol = Convert.ToInt32(Math.Floor(Math.Cos(DegreeToRadian(latitude))*(Math.Abs(↩
        longitude)*600)));
    xcol = longitude < 0 ? -xcol-1 : xcol+1;
    yrow = Convert.ToInt32(Math.Floor(Math.Abs(latitude)*600));
    yrow = latitude < 0 ? -yrow-1 : yrow+1;
  }
```

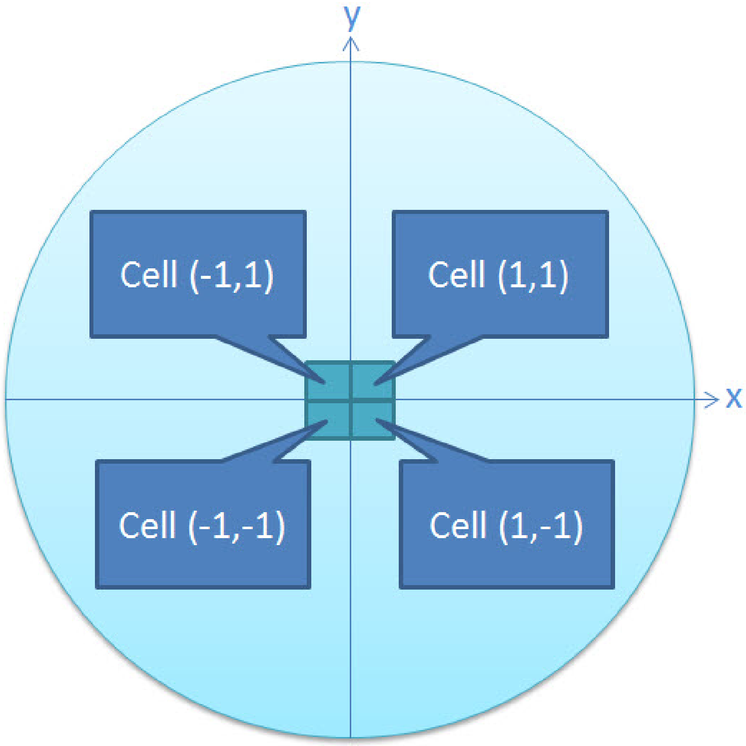


Figure 3.4 *Grid cell explanation. Each cell has the size 1 square cable length.*

For every one-hour interval, the following data are stored for each cell and unique MMSI number, using day-based tables in a PostgreSQL spatial database (PostGIS extension):

- x-column value
- y-row value
- date
- hour
- MMSI
- seconds in cell
- navigation status
- sensor type(s)
- speed over ground
- grid point type

The grid point type indicates whether the cell is based on an actual reported position (true) or interpolated calculated position (false). Figures 3.5-3.9 illustrate the process going from raw to gridded data.

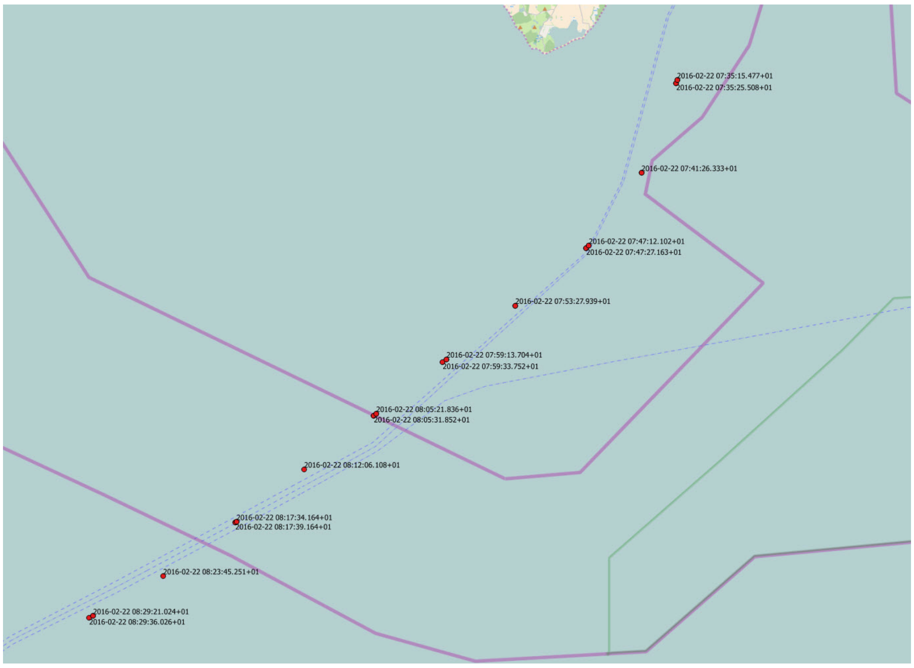


Figure 3.5 *Raw data from the cruise ferry Color Fantasy*

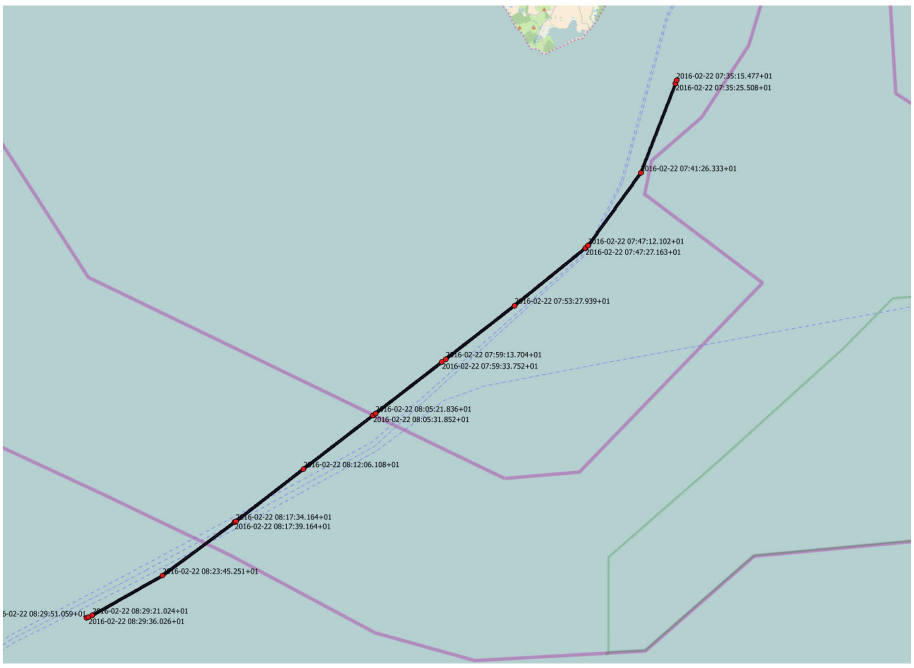


Figure 3.6 *Intermediate positions between reported positions are made by dividing the linestring path between two positions into segments. The segment length is less than 90 m.*

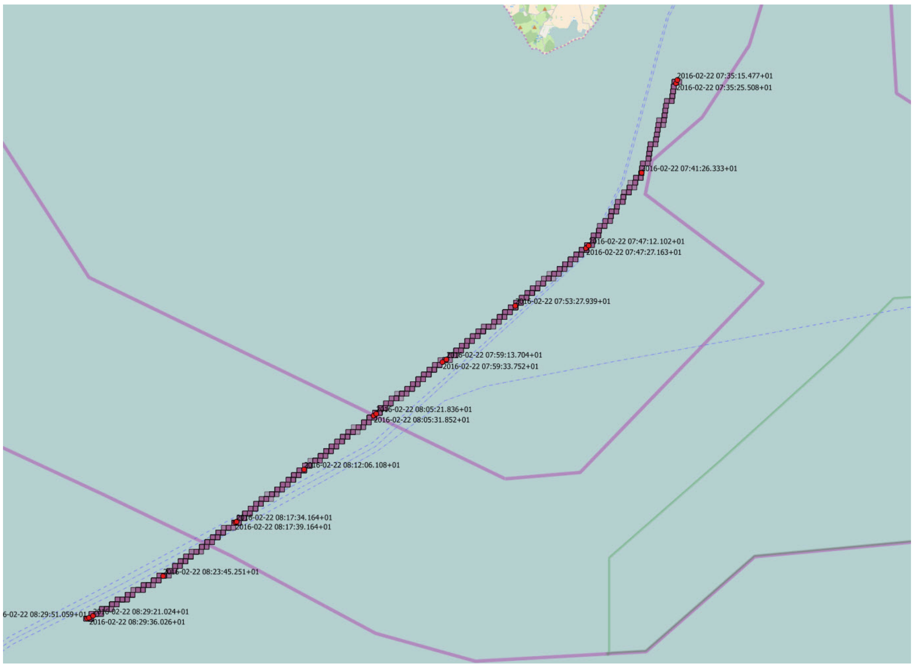


Figure 3.7 *Each position is tied (moved) to the closest grid cell center. The cell square is drawn to show its extent (coverage).*

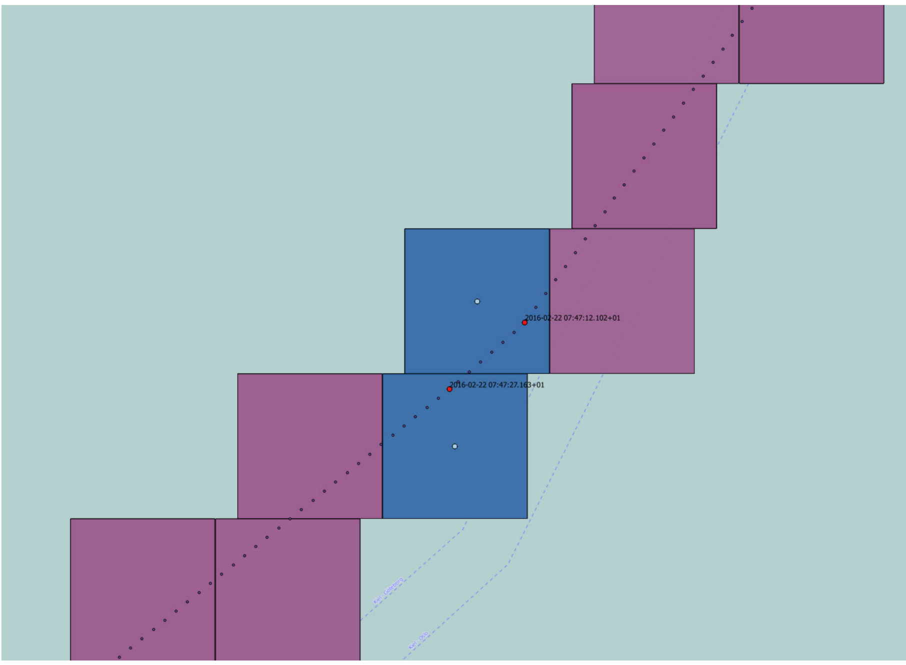


Figure 3.8 *Close-up of Color Fantasys reported and interpolated positions, with grid cells shown as containing either reported positions (blue color) or only interpolated positions (purple color).*

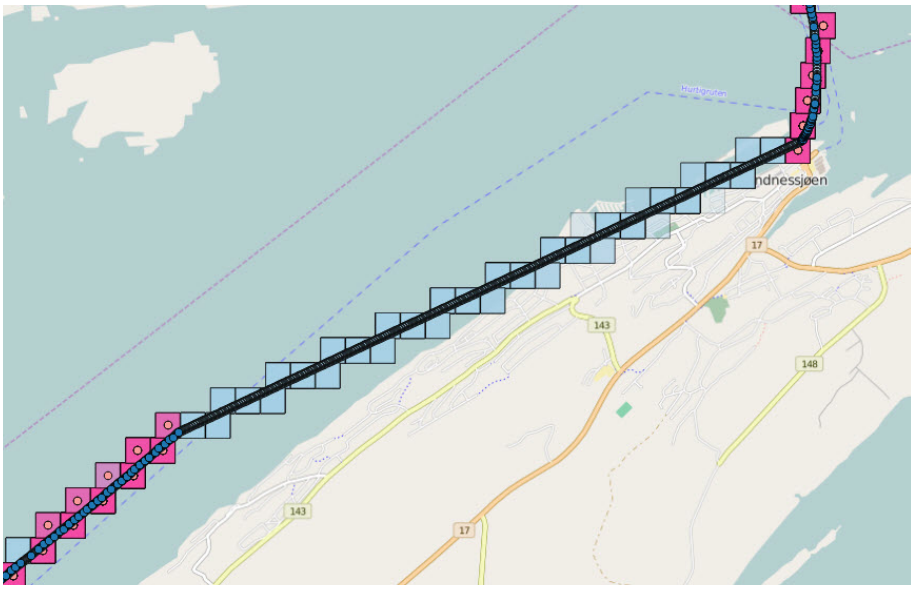


Figure 3.9 *False interpolated path due to a large gap between reported positions (light blue squares).*

Interpolation between reported positions can sometimes create false paths such as the one shown in Figure 3.9, where a ship appears to sail over land. A lot of logic has been added to the gridding process that eliminates most of these problems. For example, interpolated positions are only computed when the elapsed time between two position reports is less than eight hours. This limit was chosen based on the reporting interval of he LRIT feed (six hours), with an additional two hours to account for delays.

Despite adding data by estimating the intermediate cells on a ship's track, the gridding process reduces the number of database rows by almost 50 % in the June 2015 example (Figure 3.1).

3.4 Anomaly considerations

In order to avoid false paths between ships reporting identical MMSI numbers, the estimated speed and reported speed between two positions are taken into account. If the distance between two consecutive positions (reported with identical MMSI numbers) cannot be traversed in the given time interval, the two positions are treated as pertaining to distinct ships. Consequently interpolation is not carried out between two such postions. Some MMSI numbers that occur frequently (by default), such as 0, often result in such position jumps, and each received position will be associated with the previous position that is closest in time and distance. This way different ships reporting with identical MMSI numbers are included in the density maps, and the intermediate positions are more correctly represented. There are still anomalies that are hard to avoid, such as GPS errors that yield false tracks or no track (constant position).

Such errors will affect the end product. The more data included in the density maps (i.e., using data accumulated over longer periods), the less significant and conspicuous such errors will be, but in some cases they are still visible and should be recognized as anomalies.

3.5 Aggregation of specific attributes

In order to reduce the amount of database table rows even more, selected attributes are aggregated into monthly tables that collect unique MMSI numbers in a cell, their total time in each cell, as well as maximum registered speed over ground for each vessel and type of sensor:

- x-column value
- y-row value
- MMSI
- total seconds in cell
- sensor type(s)
- maximum speed over ground in cell
- grid point type

Based on these data tables, an end user can request a variety of density map products by selecting data based on selected attributes. For example, since the MMSI number is preserved and can be related to ship type, density maps for specific ship types can be created. After this second level of data aggregation, the number of table rows for June 2015 is 35.970.952, which is only about 43% the size of the raw data. With this second level of processing, users cannot plot data for periods shorter than one month, which is a limitation compared to the daily tables with hourly data.

4 Density mapping

4.1 Image format considerations

There are two main types of graphics file formats, namely raster and vector formats. Raster images are more common and are widely used on the web. Vector graphics are common for images that will be applied to a physical product. Both raster and vector formats have advantages and disadvantages described below. Eventually, based on advantages and disadvantages, a vector format was chosen for creating density maps.

4.1.1 Raster presentation

A raster graphic is an image that is a rectangular array of many tiny colored squares, known as pixels. The most common type of raster graphic is the digital photograph. Figure 4.1 shows a raster image with a coarse representation of ship density based on a short period of data from the harbour of Bergen city. Some advantages of raster formats are:

- Images are rich in detail and color nuances.
- Images can be precisely edited by modifying individual pixels.

Some disadvantages of raster formats are:

- Images become blurry when enlarged.
- Files may be large, depending on resolution.

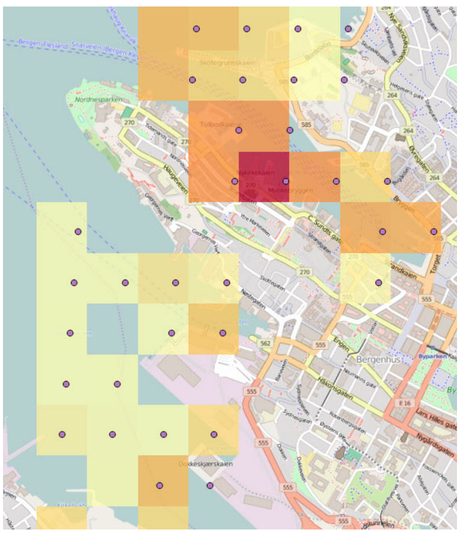


Figure 4.1 A raster representation of density map cells, with a superposed vector data layer showing the raster cell origins.

4.1.2 Vector presentation

A vector graphic is encoded as a collection of shapes (objects), e.g. points, lines and curves, effectively a set of scale and device-independent drawing instructions. While a computer screen is similar to a raster (a rectangular collection of pixels), a vector graphics renderer uses mathematical formulas to map vector shapes to pixels dependent on scale (zoom) and view. As a result, vector graphics are scalable and do not become blurry or grainy when the zoom level changes. Fonts, logos and e.g. engineering designs are common types of vector graphics.

A square polygon is encoded as four corner points and attributes such as color or line width. On the other hand, a raster graphic of a 1 in $\times$ 1 in square with a typical resolution of 300 pixels per inch (PPI) contains $300 \times 300 = 90000$ pixels. Thus vector representations often preferable when the data are sparse, as is the case with the density grid. Some advantages of vector formats are:

- Images are infinitely scalable (scale/zoom does not affect image quality).
- Files are smaller when there are few objects.
- Metadata can be attached to individual objects.
- Shapes are editable. All original data is usually kept when stored to file.

Some disadvantages of vector images are:

- Vector representation is not well-suited for complex images with many color nuances. Good results can be achieved by combining raster and vector data.
- Styling options are limited. This is often done by raster effects.

For presentation of density maps the vector format was chosen, due to the sparsity of (visited) grid cells, the representation of basic data as polygons, and the need for attaching metadata to individual polygons. Another factor is scalability, which makes it possible to reuse a product of a given resolution at different zoom levels without degrading the product.

4.2 Vector data file format

Since the amount of data may be large, a spatial database-oriented file format was preferred. Spatial database-oriented files are easily queried and reused in different areas (within the extent of the file content), and easily merged in order to accumulate data over a larger time period. The two file formats were selected and are presented in the following sections, GeoPackage and ESRI Shapefile. GeoPackage is the main format for this project. The Shapefile format is the more established format and widely supported by GIS software, while GeoPackage is a fairly new standard.

4.2.1 GeoPackage

GeoPackage is an open, non-proprietary, platform-independent and standards-based data format for geographic information system based on the SQLite database format. The GeoPackage specification is managed by the Open Geospatial Consortium[9] with the backing of the US military[10] and published in 2014[11].

Spatialite is a spatial extension to the SQLite database format providing functions and other geodatabase functionality (similar to the PostGIS extension on the PostgreSQL database)[12]. GeoPackage is as such an extended Spatialite database (*.gpkg) containing data, metadata, and spatial index tables with specific definitions, integrity assertions, format limitations, and content constraints. The GeoPackage format can used for both vector data and imagery/raster data at various scales. The format is also extensible through the use of custom extensions. Due to the lightweight design it is easily deployable on hand-held devices such as mobile phones and tablets.

4.2.2 ESRI Shapefile

The shapefile geospatial vector data format is developed and regulated by the Environmental Systems Research Institute (ESRI), a Geographic Information System (GIS) software company. A shapefile describes vector features such as points, lines, and polygons and associated attributes. The shapefile is actually a format consisting of a collection of files. There are three mandatory files with the filename extensions *.shp*, *.shx* and *.dbf*. The shapes are stored in the *.shp* file. The *.shx* file contains a spatial index so that data can be searched more quickly. The *.dbf* file contains the attribute/metadata information. There are several non-mandatory files, including the *.prj* file that contains coordinate system and projection information; the *.prj* file is also used in the density map products.

4.3 Creating the vector files

Chapter 3 described the organization of the sensor data. The data are stored in databases and retrieved via database queries when density maps are created. To allow end users select time period, area of interest, and other limiting or describing factors when generating density maps, several software programs were developed for

- creating GeoPackage/shapefile;
- merging GeoPackages/shapefiles;
- extracting GeoPackage/shapefile layers (i.e., creating a new GeoPackage/shapefile based on desired attributes, e.g. a specific ship type);
- masking GeoPackage/shapefile (clipping a GeoPackage/shapefile with a land mask and recalculating density along the coastal lines).

Each of these programs are published as Web Processing Services (WPSs)[13]. Small or moderately large areas, including typical coastal areas of interest such as the Oslofjord, are easily mapped

[9] http://www.GeoPackage.org/spec/
[10] http://www.c4isrnet.com/story/military-tech/geoint/2015/08/13/edge-pushing-geo-data-disadvantaged-troops/31547355/
[11] https://en.wikipedia.org/wiki/GeoPackage
[12] https://www.gaia-gis.it/fossil/libspatialite/index
[13] www.opengeospatial.org/standards/wps

within minutes. Larger areas can also be made on demand if data are somehow limited, e.g., to a range of MMSI numbers. Products based on very large amounts of data (large areas, heavy traffic) must be produced in advance to ensure swift delivery.

4.3.1 Density calculation and interpretation

As described in Chapter 3, space and time are discretized with spatial resolution (basic cell size) $\Delta x \times \Delta y = (1/10) \times (1/10)$ nm^2 and time resolution $\Delta T = 1$ h. Hence any vessel position (x, y, t), where t denotes calendar time, is associated with a unique space-time cell $c = C(x, y, t)$ of size $\Delta x \times \Delta y \times \Delta T$. If a ship track intersects a space-time cell, c, the ship MMSI number, time τ spent in the cell, maximum SOG and other attributes are recorded in the database.

The traffic density is computed from the time fraction $\tau/\Delta T$: Suppose $N(c)$ ship tracks intersect cell c with visiting times $\tau_1, \tau_2, \ldots, \tau_{N(c)}$, respectively. The space and time-dependent density is defined as

$$\rho(x, y, t) = \frac{1}{\Delta x \Delta y \Delta T} \sum_{k=1}^{N(c)} \tau_k, \tag{4.1}$$

where $c = C(x, y, t)$. Clearly ρ is constant within each cell (see ClosestGridPoint in Section 3.3). The aggregated time fraction

$$\rho_T(x, y, t) = \frac{1}{\Delta T} \sum_{k=1}^{N(c)} \tau_k = \rho(x, y, t)\Delta x \Delta y \tag{4.2}$$

can be interpreted as follows: If someone took a photograph of the area of cell c at time t, then the expected number of ships in the image is $\rho_T(x, y, t)$. The inverse value $1/\rho_T$ is the number of photographs that must be taken around time t to observe on average a single ship (any activity at all) in the cell. For example, $\rho_T(x, y, t) = 2$ implies that we expect to observe two ships in the cell containing (x, y) at the instant of time t. If $\rho_T(x, y, t) = 0.1$, we expect that a single ship would be revealed in a series of $1/\rho_T = 10$ snapshots during the interval ΔT.

Various density maps can be obtained by averaging ρ over space and time. For arbitrarily fine resolution, averaging is accomplished by integrating ρ with respect to space and/or time. Due to the finite discretization, the integrals are replaced by sums. The density at position (x, y), averaged over the interval $[T_1, T_2]$, is

$$\rho(x, y; T_1, T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \rho(x, y, t) \, \mathrm{d}t. \tag{4.3}$$

Assuming $[T_1, T_2]$ is partitioned into M segments of length ΔT, so that $t_n = T_1 + n\Delta T$, $t_M = T_2$, Eqn. (4.3) becomes

$$\rho(x, y; T_1, T_2) = \frac{1}{M} \sum_{n=0}^{M-1} \rho(x, y, t_n). \tag{4.4}$$

The interpretation of $\rho(x, y; T_1, T_2)\Delta x \Delta y$ is the same as for Eqn. (4.2) above. In fact, Eqns. (4.4) and (4.1) are equivalent if we choose time resolution $\Delta T = T_2 - T_1$. Similarly, the density averaged over some geographical area Ω at time t is

$$\rho_\Omega(t) = \frac{1}{\mathrm{area}(\Omega)} \int_\Omega \rho(x, y, t) \, \mathrm{d}A. \tag{4.5}$$

Assuming Ω constitutes K basic cells $c(x_1, y_1), \ldots, c(x_K, y_K)$ at time t, Eqn. (4.5) becomes

$$\rho_\Omega(t) = \frac{\Delta x \Delta y}{\mathrm{area}(\Omega)} \sum_{k=1}^{K} \rho(x_k, y_k, t) = \frac{1}{\mathrm{area}(\Omega)} \sum_{k=1}^{K} \rho_T(x_k, y_k, t). \tag{4.6}$$

Hence $\rho_\Omega(t)\mathrm{area}(\Omega) = \int_\Omega \rho(x, y, t)\, \mathrm{d}A$ is the expected number of ships seen if we observe Ω at instant of time t. Finally, the space-time-averaged density,

$$\rho_\Omega(T_1, T_2) = \frac{1}{(T_2 - T_1)\mathrm{area}(\Omega)} \int_{T_1}^{T_2} \int_\Omega \rho(x, y, t)\, \mathrm{d}t\, \mathrm{d}A, \tag{4.7}$$

is such that $\rho_\Omega(T_1, T_2)\mathrm{area}(\Omega)$ is the expected number of ships observed in a single snapshot of Ω during the interval $T_2 - T_1$.

For an arbitrary geographical area Ω, the above discussion can be summarized as follows:

- The density $\rho_\Omega(t)$ defined by (4.1) and (4.5) is such that $\rho_\Omega(t)\mathrm{area}(\Omega)$ is the expected number of ships observed in Ω at instant of time t.
- A series of $N_{\mathrm{obs}} = N_{\mathrm{ship}}/\rho_\Omega(t)\mathrm{area}(\Omega)$ observations ("photographs") is needed on average for N_{ship} detections in Ω (these detections are not necessarily of distinct ships).
- The time-averaged density $\rho_\Omega(T_1, T_2)$ has the same interpretation as $\rho_\Omega(t)$ for observations in the interval $T_1 \leq t \leq T_2$.

For density maps, Ω will correspond to a graphics primitive, i.e. a pixel, point, or filled polygon. There are two natural alternatives for such primitive regions: Ω_1 is a rectangular aggregate of $n_x \times n_y$ basic cells, and Ω_2 is the subset of $n_x \times n_y$ basic cells that contain ship tracks (i.e., empty cells are discarded). Since $\Omega_2 \subseteq \Omega_1$, the first alternative implies more smoothing.

The following equation shows an example for a given area Ω over the time period of January where the scale factor is 5 (see Section 4.3.3). Ω represents the 5 x 5 basic cell area and the density is represented as constant over the time period January 1st (T_1)- January 31st (T_2). M is total amount of hours in January (744).

$$\rho_\Omega(T_1, T_2) = \frac{1}{(T_2 - T_1)\mathrm{area}(\Omega)} \sum_{k=1}^{M} \sum_{i=1}^{5} \sum_{j=1}^{5} \left(\sum_{s=1}^{N(k,i,j)} \tau_{k,i,j}^{s} \right) \tag{4.8}$$

Here $\tau_{k,i,j}^{s}$ is visit time of ship track s in basic cell (i, j) in area Ω at time step k. $N(k, i, j)$ is the number of ship tracks intersecting the space-time cell (k, i, j).

4.3.2 Preprocessed areas of interest

Density maps of a large areas provide overview and enable end users to identify and order smaller areas of interest. Overview maps are pre-computed for easy reuse, transfer, and delivery through WPS. Current products are Europe GeoPackages containing density maps for ship types at level 2-5 of IHS Fairplay Statcode 5, see Chapter 2.4. Ships not identified with a level 2–5 ship type classification are collected as "unknown". If the software has a complete IHF Fairplay database connecting MMSI numbers and ship types, the unknown class will include only ships with invalid MMSI numbers, i.e. MMSI numbers yet to be reported or known outside own country distributing authority.

4.3.2.1 Implementation of Statcode 5 in the vector file output

The calculated density of each ship type is stored in the output file. Shapefiles are limited to 10 character attribute name length, while GeoPackage has no such limitation. Since a ship type classification varies in length and sometimes exceeds the shapefile limitation, the ship type code is used for shapefiles, while the complete ship type name is used for GeoPackages.

4.3.3 Density map cell size—scale factor

Ideally the basic cell size of 1 square cable length should be maintained in all products. A single product might then provide both world overview and fine details from harbor areas of interest, and hence be highly scalable. However, such a product would be too large for a GIS tool. Each polygon representing a square cell contains position information and associated attributes, with corresponding data storage demands. The more polygons, the larger the GeoPackage/shapefile will be. To reduce the file size, density values are computed for larger cells consisting of $S \times S$ basic cells, where S is referred to as the scale factor. Figure 4.2 illustrates the scale factors used in this work. For small areas the scale factor is $S = 1$ [$(0.1\ \text{nm})^2$, in yellow]. For medium large areas, like Europe, the scale factor is $S = 5$ [$(0.5\ \text{nm})^2$, in red]. For very large areas the scale factor is $S = 10$ ($1\ \text{nm}^2$, in green). Cells with scale factor $S = 5$ are five times larger in the longitudinal and latitudinal directions, and cells with scale factor $S = 10$ are 10 times larger in each direction.

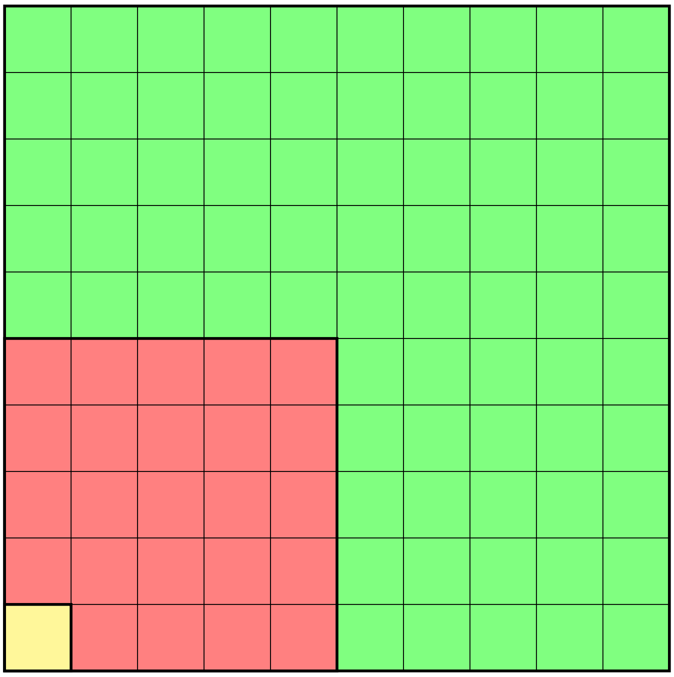


Figure 4.2 *Illustration of scale factor $S = 1$ (yellow), $S = 5$ (red), and $S = 10$ (green), corresponding to cell width 1/10 nm, 1/2 nm, and 1 nm, respectively.*

There are at least two different ways to aggregate cells (see also Section 4.3.1):

1. Form the union of all *non-empty* basic cells ($S = 1$) within a larger cell ($S = 5$ or $S = 10$) (Figure 4.3, left),
2. Create a new square parent polygon with $S = 5$ or $S = 10$ (Figure 4.3, right).

In either case the density is recomputed as the average density of all aggregated cells. Since the square parent polygon may contain empty cells, the aggregate density is typically lower than for union polygons; this is an effect of smoothing. Since the union effectively preserves the non-empty geometry of the $S = 1$ level, this method saves less space. In addition, computing the union is computationally intensive for large sets of geometries.

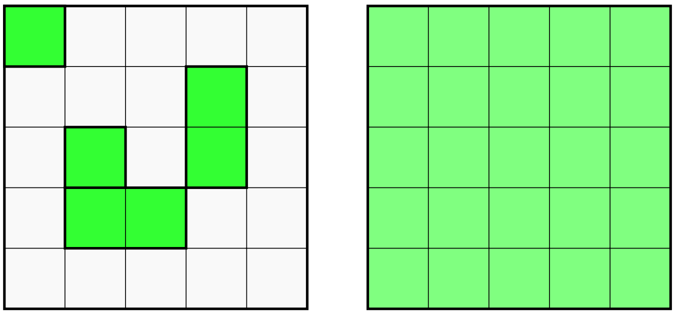


Figure 4.3 *At scale factor $S = 5$ the multipolygon in the left figure contains six cells with scale factor $= 1$. The right figure shows a square parent polygon that contains 5×5 cells with scale factor $S = 1$, including empty cells (no reported or interpolated positions). The density value of the multipolygon area will be higher than that of the parent polygon.*

The parent polygon that includes empty cells, with density computed as the average for the complete area, generally has a much simpler geometry, i.e., a square polygon defined by four corner positions. One problem with this method is that in coastal areas some empty cell may cover land and hence should not contribute to the aggregate density. This can be amended by clipping the polygon shape with a land mask (a polygon or raster representing land areas). Since the density is calculated based on the total area of the new polygon, each clipped polygon needs to be recalculated based on the new area covering the sea, where ship activity is actually possible.

The fact that part of the polygon may cover land will be an issue on every scale level. Narrow sounds, less than 185 meters across, needs to be masked and recalculated to be entirely correct. The difference in density may be slight when considering the density map as a whole, but should be taken into account nevertheless.

Figure 4.4 shows a narrow sound and the polygons covering the surrounding land areas.

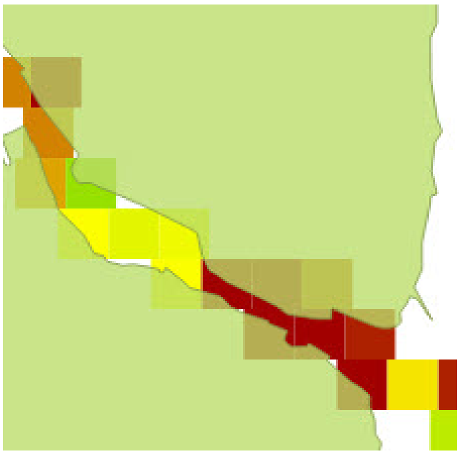


Figure 4.4 *Density polygons partly covering land in a narrow sound. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

Figure 4.5 shows the same sound clipped with a land mask. The density has not been recomputed in this example. Every cell whose area is reduced will get a larger density value, and hence a shift in the color spectrum.

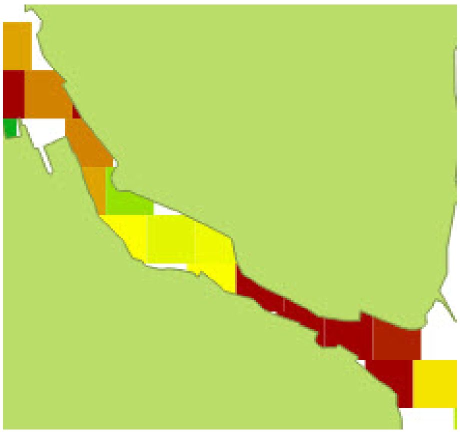


Figure 4.5 *Density polygons clipped with a land mask, now only covering sea area. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

4.3.3.1 File sizes for different scale factors and methods

Compared to creating the parent polygon that includes empty cells, taking the union of non-empty cells yields the better visual impression, even when zooming in on a small area. This section presents density maps of the entire World followed by close-ups of the south-eastern UK. The maps were made using different methods and scale factors. The maps have different color scales and hence are not directly comparable. Figure 4.6 shows a world density map with scale factor $S = 30$ and union of non-empty cells.

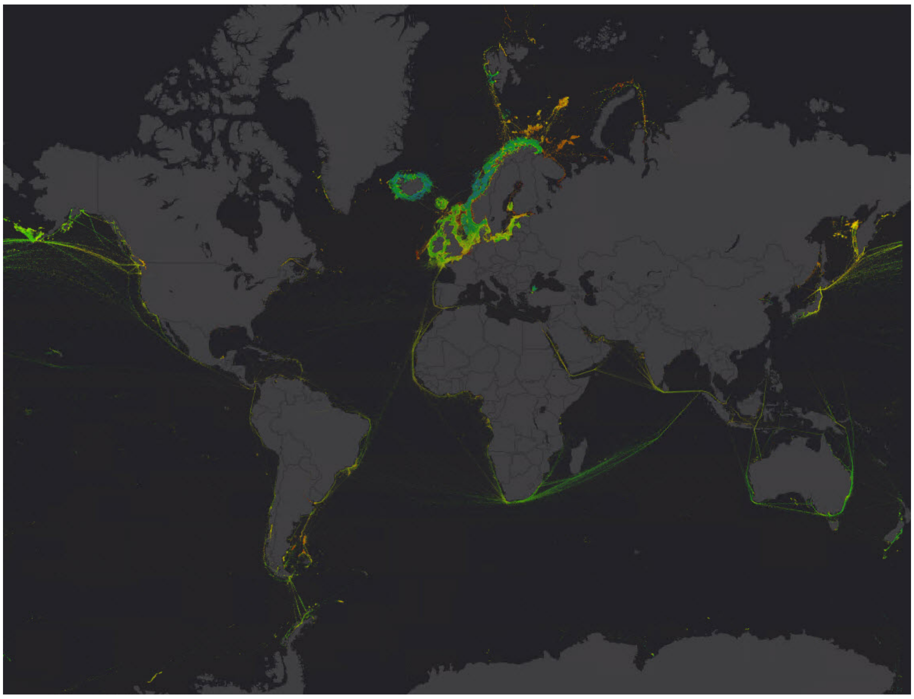


Figure 4.6 *World density map, based on one month of data, using the union contributing (non-empty) cells and scale factor $S = 30$. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

Figure 4.7 shows the same density map zoomed in on south-eastern UK. The map is still readable and contains many fine details even at this scale.

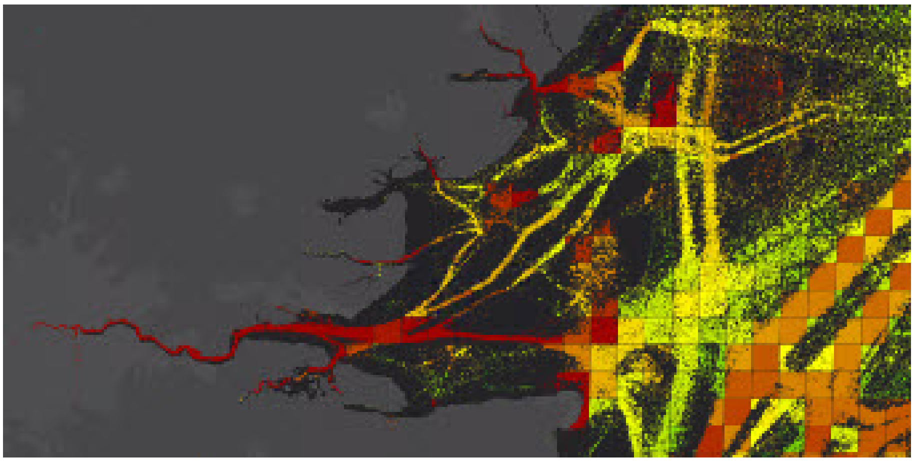


Figure 4.7 *World density map zoomed in on the south-eastern UK. The map is based on one month of data, using the union contributing (non-empty) cells and scale factor $S = 30$. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

Figure 4.8 is a World map using parent polygons and the same scale factor $S = 30$, and it is evident that this method works well at this scale. However, when zooming in on the south-eastern UK the result is not satisfactory. Consequently a different scale factor or aggregation method must be used to enable end users to find areas of interest at this scale starting from overview maps.

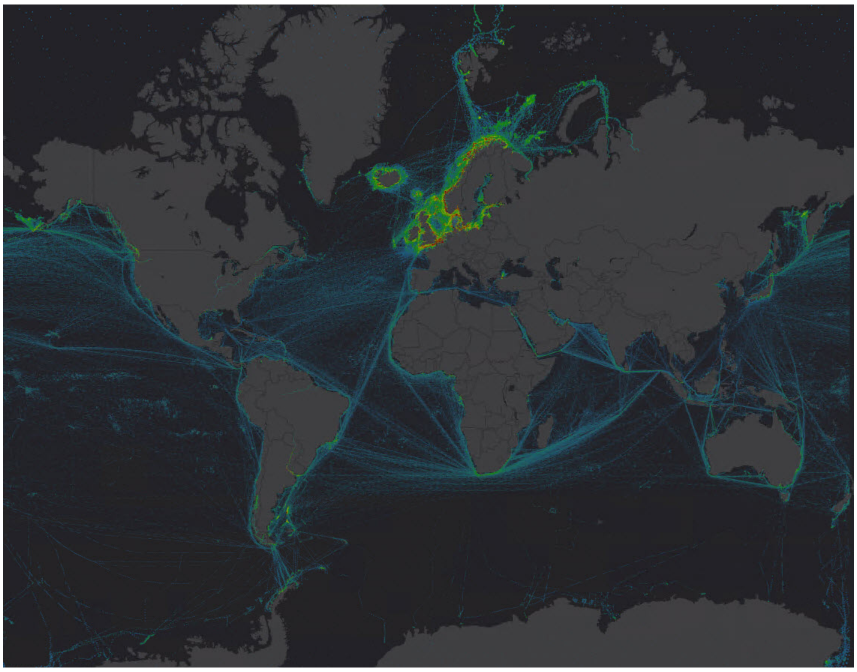


Figure 4.8 *World density map, based on one month of data, using the parent polygon method with scale factor $S = 30$. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

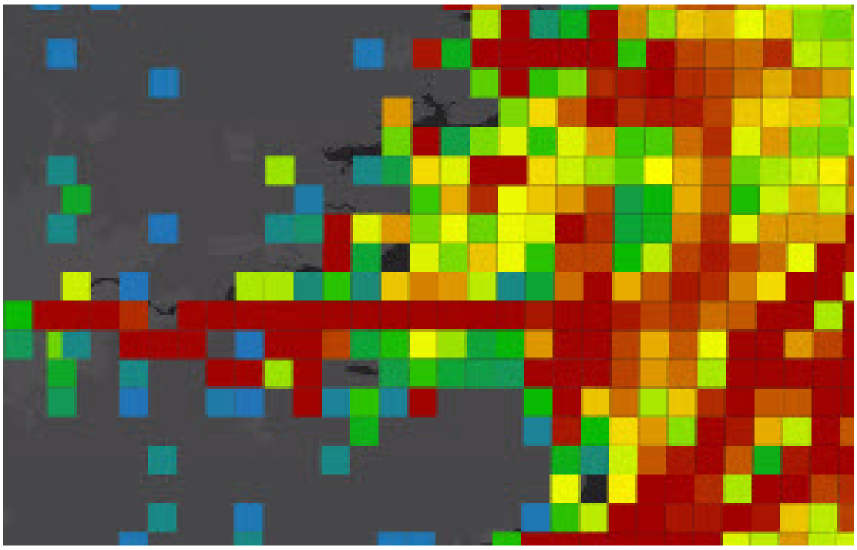


Figure 4.9 *World density map zoomed in on the south-eastern UK. The map is based on one month of data, using the parent polygon aggregation method and scale factor $S = 30$. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

Figures 4.10–4.11 show the same World density map using the parent polygon method with scale factor $S = 10$. The size of each polygon is 1 nm × 1 nm. On a global scale (Figure 4.10) the result is visually comparable to that obtained with $S = 30$, while on a regional scale (Figure 4.11) the result is comparable to that obtained with the union method and $S = 30$. Figure 4.11 provides sufficient detail for end users to find areas of interest, and scale factor $S = 10$ is hence more appropriate than $S = 30$ when using the parent polygon method. Using the union method with a large scale factor results in very large files and is generally not feasible.

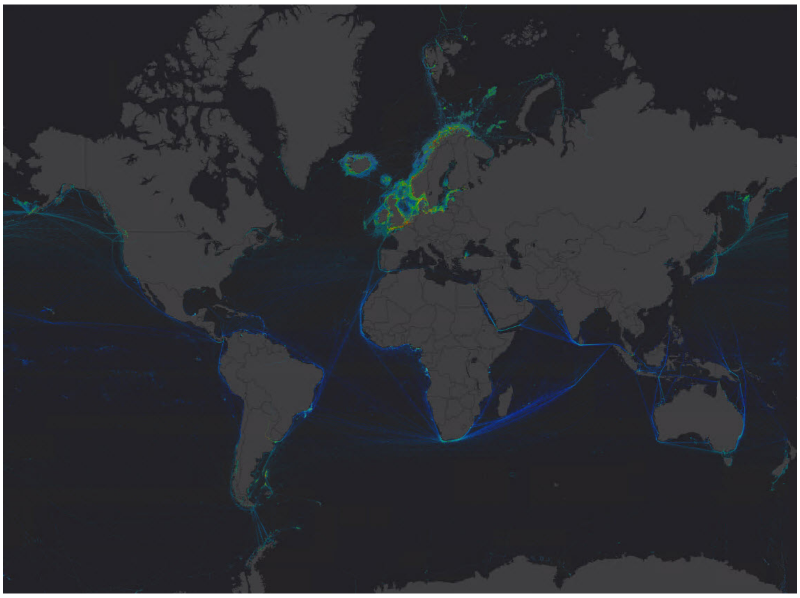


Figure 4.10 *A world density map (one month) using union new parent polygons method on scale factor 10. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

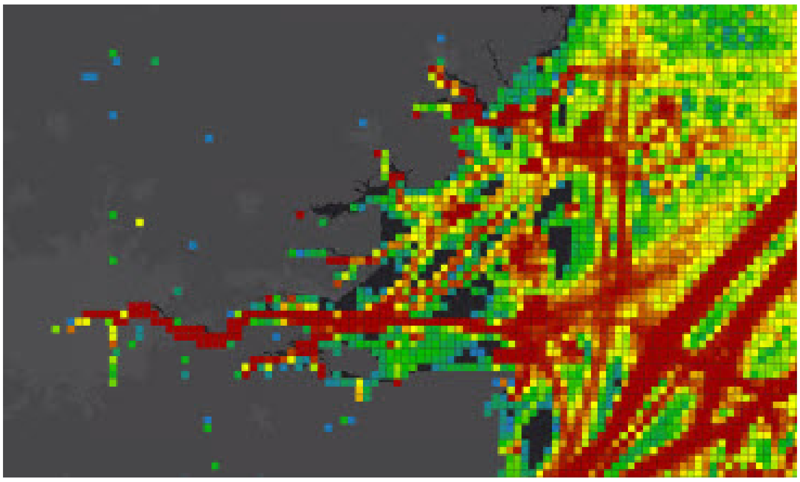


Figure 4.11 *A world density map (one month) zoomed in on the south east UK using union new parent polygons method on scale factor 10. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

Scale factor (S)	Aggregation method	Uncompressed size	Compressed size
30	Union	2400 MB	407 MB
30	Parent polygon	568 MB	91 MB
10	Parent polygon	1280 MB	208 MB

Table 4.1 *File size for World density maps in shapefile format, with different scale factors and aggregation methods. The maps were based on data from February 2015.*

Arguably the union method yields better results, but the data compression rate is correspondingly low. The file size examples in Table 4.1, for shapefiles made from February 2015 data, illustrate this clearly. The files include attributes for 11 different ship type classifications as well as a total density value and cell position information. The compression rates are similar for the GeoPackage format. Files produced with scale factor $S = 10$ and parent polygons are about 50 % smaller than files produced with scale factor $S = 30$ and the union method. By comparison, the density map for the south-eastern UK area alone (Figure 4.12) is about 60 MB uncompressed and 9 MB compressed when using scale factor $S = 1$ (i.e., cell width 0.1 nm and no aggregation).

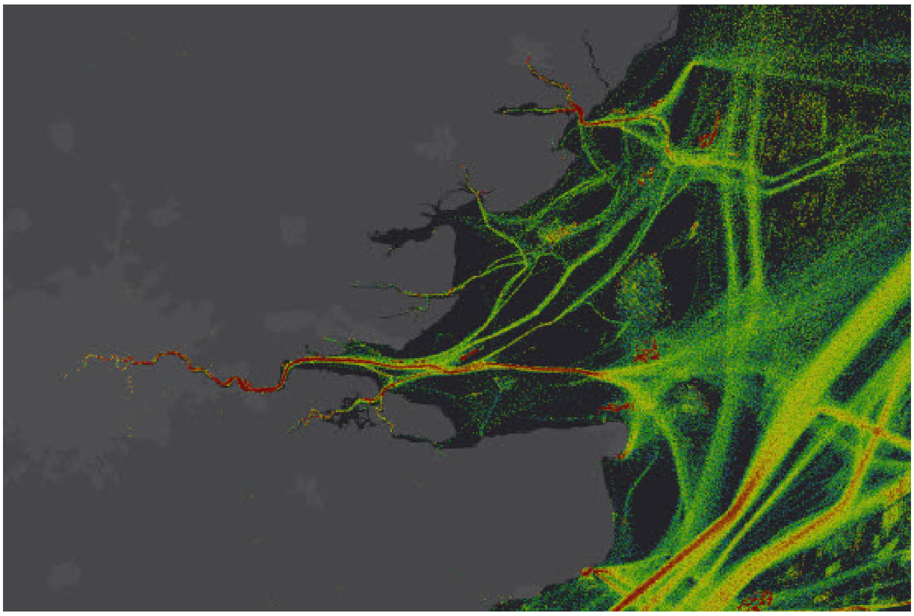


Figure 4.12 *South-eastern UK density map with scale factor $S = 1$. The color gradient ranges from blue (low density) via green and yellow to red (high density).*

4.4 Product examples

4.4.1 Preprocessed density maps

Producing density maps can be time-consuming, but maps can be produced in advance and re-used in the services exposed to end users. Single-attribute “layers”, such as one ship type, can be extracted and delivered. Other products can be made by extracting sub-regions from existing density maps, this would also reduce the file size. The scale factor would stay the same and has to be taken into consideration by the end user. Using a World density map for a small area might not provide sufficient detail, as in Figure 4.9.

4.4.1.1 Example use case

The map in Figure 4.13 is taken from the pre-processed Europe 2015 density map and shows ship types with offshore classification. The parent ship type classification provides a quick overview of an activity or traffic pattern specific for that type.

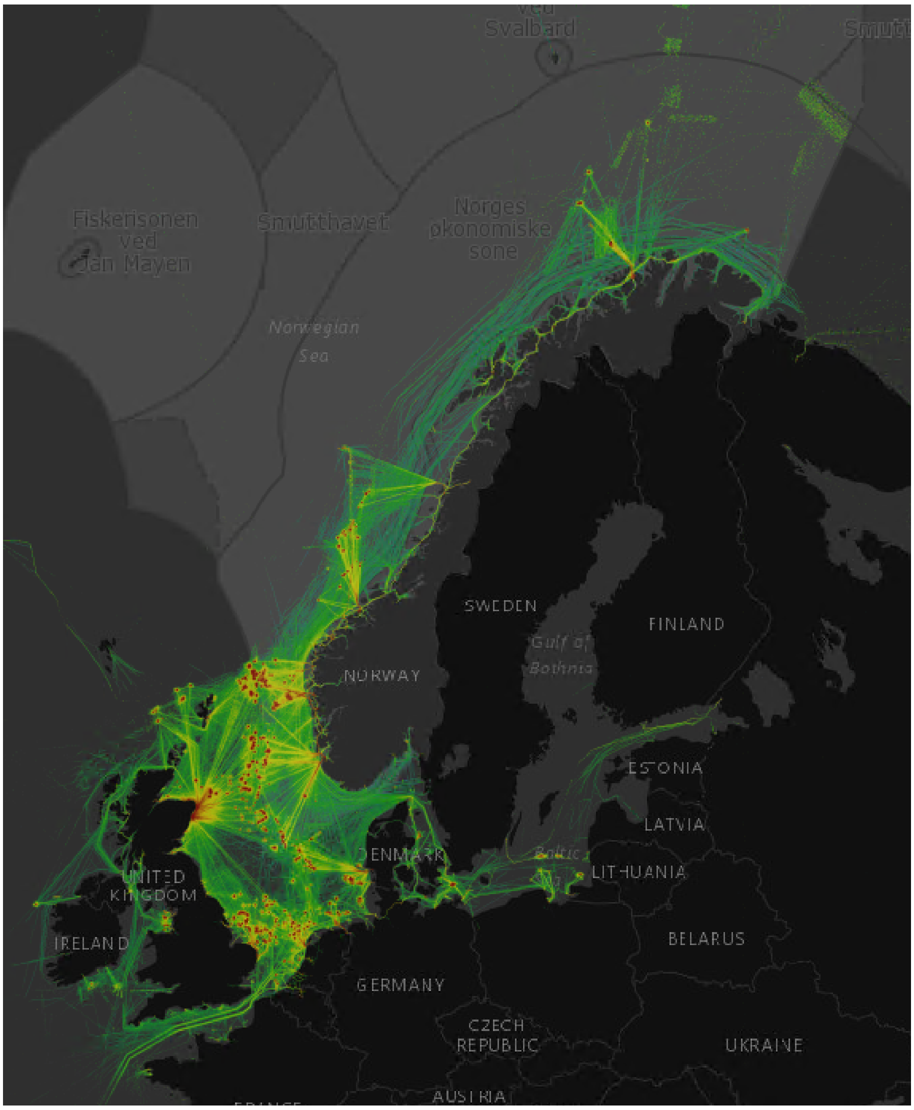


Figure 4.13 *Density map based on 2015 data for ships with offshore type classification (Statcode 5 type B2, Offshore). The scale factor is $S = 5$.*

A user may find a particular area more interesting and worth further analysis. The user may zoom in on this area using a GIS tool, as shown in Figure 4.14.

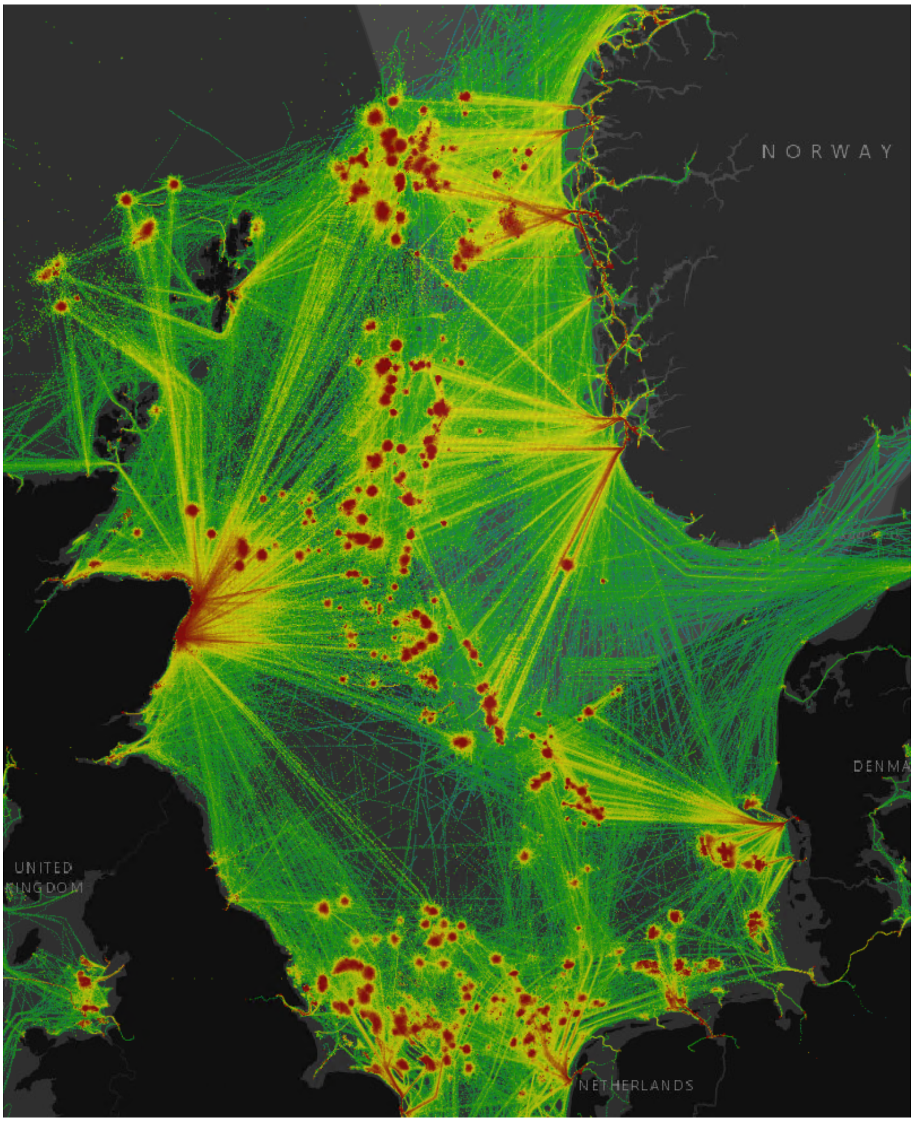


Figure 4.14 *Density map based on 2015 data for ships with offshore type classification (Statcode 5 type B2, Offshore), zoomed in on the North Sea. The scale factor is $S = 5$.*

After zooming in on this area and selecting the northern-most offshore installations for further analysis, a new density map can be created on demand, see Chapter 4.4.2. Using ship type classification B2 (Offshore) as limiting attribute and presenting ship types at level 5 in the Statcode 5 classification, the end user has the possibility to investigate in more detail the offshore activity of interest. Figure 4.15 shows a density map for January 2015 based on data from platform supply ship classified as B21A2OS in Statcode 5. This particular ship type has the largest contribution in the overview map in Figure 4.13.

Figure 4.15 Density map showing January of 2015 with only platform supply offshore ship type, classification B21A2OS, Northern Norwegian Sea. Scale factor 1.

4.4.2 Density maps created on demand

Applying limiting factors, or filters, may reduce the file size and processing time. Maps may be generated with a variety of filter combinations and are not easily pre-produced. Instead such density maps are created on demand via WPS. Figure 4.16 shows a density map created using limiting factors. Specifically, the data have been restricted to ships classified as Liquefied Gas tanker ships.

Figure 4.16 Density map showing Liquefied Gas tankers in January 2016. The scale factor is $S = 10$.

Figure 4.17 shows a selection of polygon attributes corresponding to different ship types, and their associated density values. Data can be filtered based on these attribute in order to visualise traffic patterns for specific ship types.

Feature	Value
Europe_2016_01_lvl3	
fid	188951
(Derived)	
(Actions)	
fid	188951
X	330
Y	7057
MAXSOG	15.5
DENSITY	0.000104345878136201
BULK DRY	4.31600955794504e-06
BULK DRY/LIQUID	
CHEMICAL	6.75029868578256e-06
CONTAINER	
DREDGING	
FISH CATCHING	4.36081242532855e-06
GENERAL CARGO	3.05406212664277e-05
INLAND WATERWAYS DRY CARGO/PASSENGER	
INLAND WATERWAYS OTHER NON SEAGOING	
INLAND WATERWAYS TANKER	
LIQUEFIED GAS	6.8847072879331e-06
NON MERCHANT SHIPS	
NON PROPELLED	
NON SHIP STRUCTURES	
OFFSHORE SUPPLY	3.9426523297491e-06
OIL	4.92831541218638e-06
OTHER ACTIVITIES	
OTHER BULK DRY	
OTHER DRY CARGO	
OTHER FISHING	
OTHER LIQUIDS	
OTHER OFFSHORE	5.10752688172043e-06
PASSENGER	
PASSENGER/GENERAL CARGO	
PASSENGER/RO-RO CARGO	2.89725209080048e-06
REFRIGERATED CARGO	3.01672640382318e-06
RESEARCH	
RO-RO CARGO	6.45161290322581e-06
SELF DISCHARGING BULK DRY	5.97371565113501e-06
TOWING/PUSHING	1.56063321385902e-05
UNKNOWN	3.56929510155317e-06

Figure 4.17 *Attributes in a density map made with ship classification types at level 3 in the Statcode 5. The X and Y attribute refer to the column (X) and row (Y) cell position value with scale factor* $S = 5$*. The density attribute refers to the overall density, while specific ship type densities are provided if the type is present (known). Ships with unknown ship type classification contributes to the UNKNOWN density value. MAXSOG refers to the maximum registered speed in the selected cell during the aggregation period.*

4.4.2.1 Fisheries

Figures 4.18–4.21 show fishing activity in a part of the Norwegian Sea/Barents Sea. The maps are based on data from the first quarter of 2016. The following limiting attributes, or filters, were applied:

- Select ship types at level 5 in Statcode 5 limited by level 3 codes B11 (Fish Catching) and B12 (Other Fishing). Ships in this category include:
  Factory Stern Trawler
  Stern Trawler
  Trawler
  Fishing Vessel
  Fish Factory Ship
  Fish Carrier
  Live Fish Carrier (Well Boat)
  Fish Farm Support Vessel
  Fishery Patrol Vessel
  Fishery Research Vessel
  Fishery Support Vessel
  Seal Catcher
  Whale Catcher
  Kelp Dredger
  Pearl Shells Carrier
- Include only cells with speed less than 5 knots.
- Limit data by geographical area of interest.

Figure 4.18 shows an overall density map including all ship types. Note that ships without of unknown type were also included; such ships include small fishing boats that report on AIS but are not classified by IHS Fairplay.

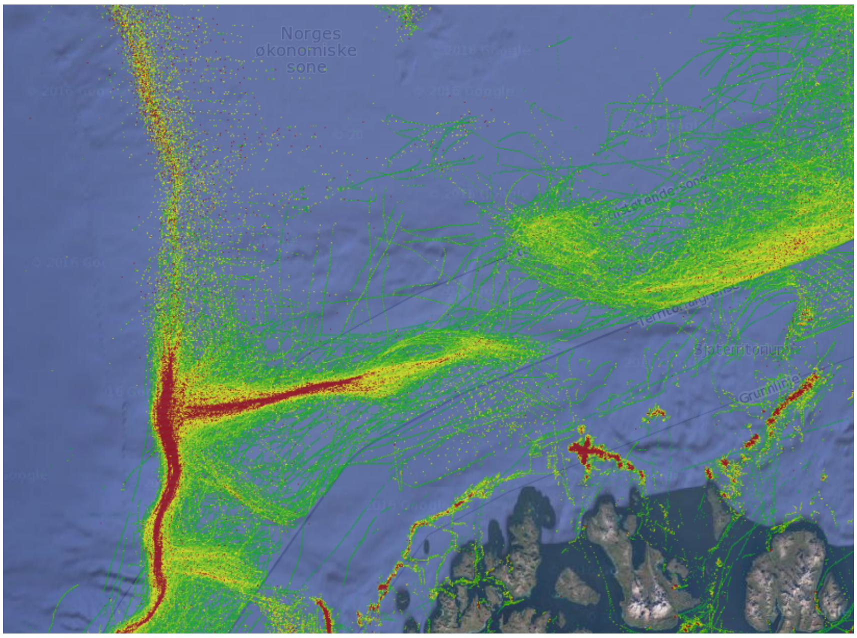


Figure 4.18 *Density map including all ship types limited by ship type level 3 codes B11 (Fish catching) and B12 (Other Fishing). Based on data from the first quarter of 2016.*

Figure 4.19 shows regular trawler activity and clearly reflects fishing regulations. These vessels are not allowed to fish in the territorial waters.

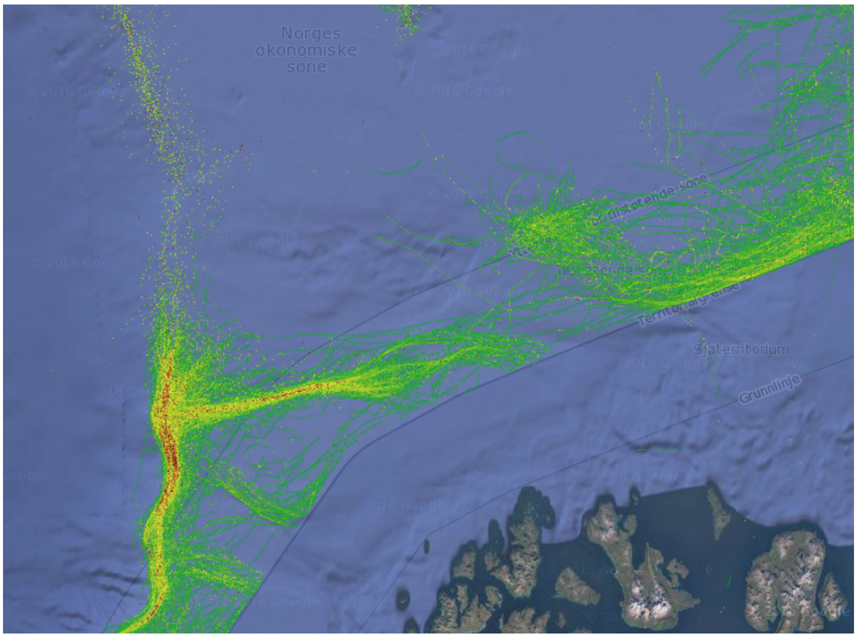


Figure 4.19 *Density map including all trawlers (not including factory or stern trawlers). Based on data from the first quarter of 2016.*

Other regulations apply to regular fishing vessels, as can be seen in Figure 4.20.

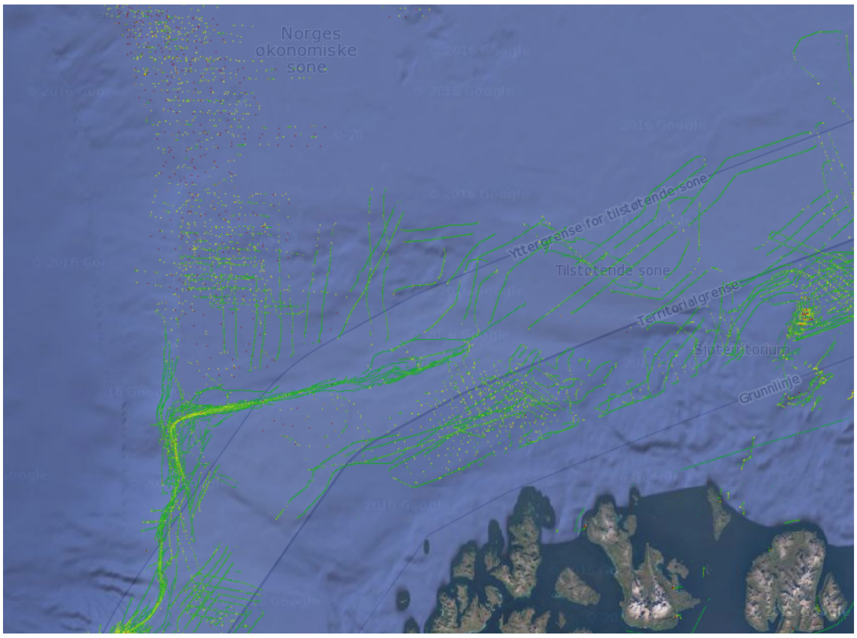


Figure 4.20 *Density map including all boats classified as Fishing Vessel (B11B2FV). Based on data from the first quarter of 2016.*

The last example (Figure 4.21) shows ships of unknown ship type, believed to include, among others, small unclassified fishing boats. According to fishery experts, this map shows mainly small fishing boat activity within territorial waters.

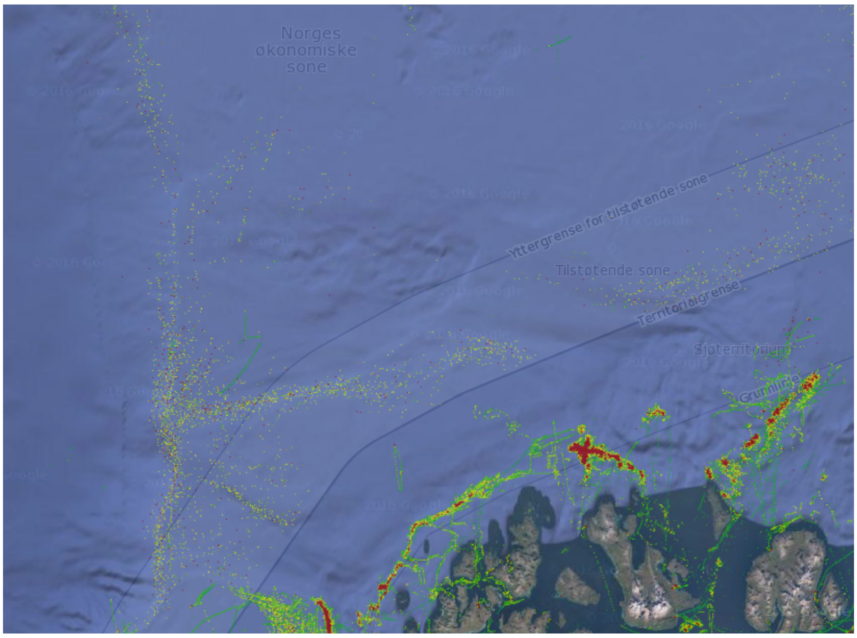


Figure 4.21 *Density map including all fishing vessels not classified by IHS Fairplay. Based on data from the first quarter of 2016.*

4.4.2.2 Other types of maps

The vessel report database can be used in numerous applications besides density maps. Some examples are shown below.

Sensor coverage

Selecting a single sensor type, and rendering all polygons in the resulting density map with the same color, yields a sensor coverage map for a given period of time. This is possible provided all data from the specific sensor, in the relevant period and area, have been stored in the database. Figures 4.22–4.24 show the coverage of 1) all sensors (Figure 4.22), 2) only AISSat-1/2 (Figure 4.23), and 3) the AIS receiver on board the ISS (Figure 4.24). These maps comprise data from January 2016 and were produced by selecting only grid cells from reported positions and excluding intermediate (interpolated) cells. Figure 4.23 clearly shows the good coverage of AISSat-1 and 2 at high latitudes. Figure 4.24 shows the limited coverage of the ISS AIS receiver in the north and south.

Figure 4.22 *Coverage map for all unclassified sensors. Based on data from January 2016.*

Figure 4.23 *Coverage map for AISSat-1/2. Based on data from January 2016.*

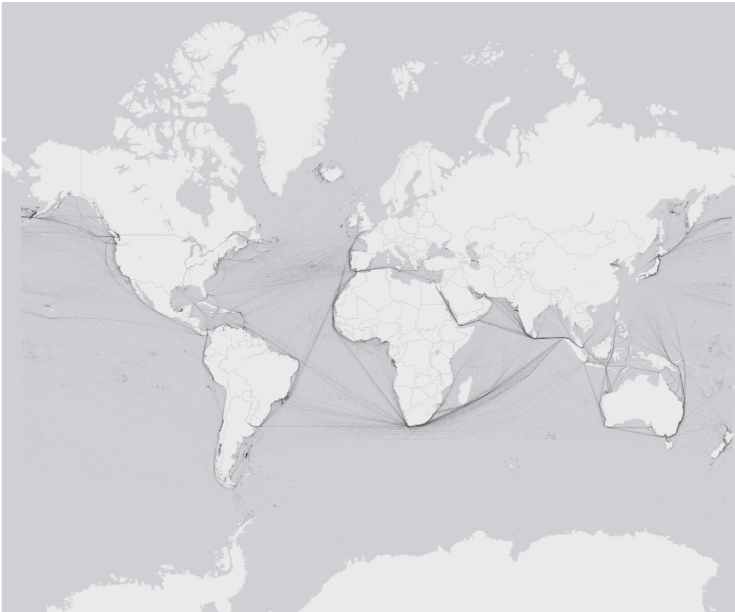


Figure 4.24 *Coverage map for AIS receiver on board the ISS. Based on data from January 2016.*

Maximum registered speed

The pre-produced density maps include a layer showing the maximum registered speed in every cell with the given scale factor. Figure 4.25 shows part of the Europe 2015 (one year of data) map of maximum registered speed.

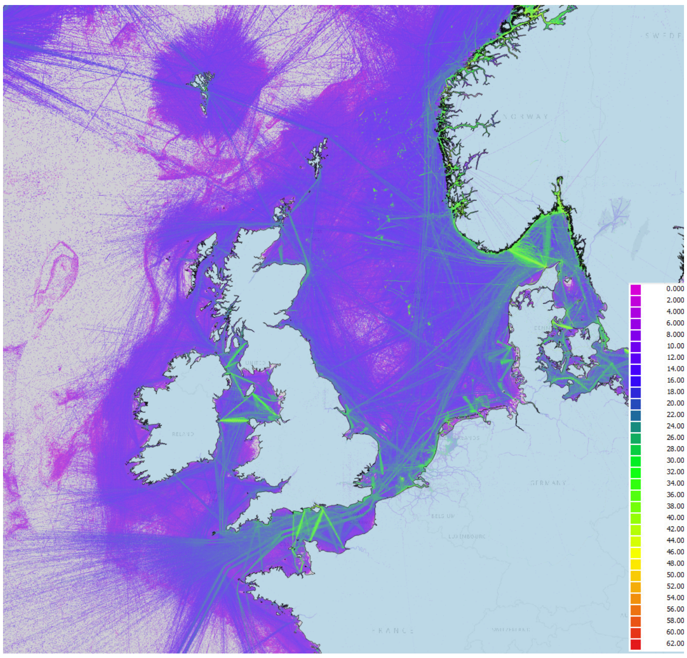


Figure 4.25 *Maximum registered speed in Europe 2015 in knots [nautical miles per hour].*

5 Conclusion and future work

Providing user-friendly, on-demand density map services presents several challenges. First of all creating density maps of ship traffic is a compute-intensive process. A large amount of data must be processed, stored and managed with care to enhance performance; resulting in a much lower and tolerable total operational cost compared to raw data crunching. Secondly, an application that supports on-demand remote computations with high performance and low overhead must be developed and deployed.

Much effort in this work has therefore been put into reducing the amount of data while maintaining sufficient resolution and content to satisfy military application requirements. One purpose of this project has been to consider the deployment of density map services also on closed military networks with local sensor databases. This is made possible with the software and methodology developed in this project combined with available or soon-to-be-introduced service-oriented technology on the existing military platforms and networks. Such service-oriented architecture, supporting the web processing technology is the recommended solution for handling density map requests from end users. To meet operational requirements, we suggest a combination of services supporting both the delivery of standardized, pre-produced maps and the dynamic on-the-fly creation of user-specified maps.

To detect anomalies in the maritime picture, the "normal" situation must be established and understood. Density maps based on larger data sets are less sensitive to outliers, i.e., anomalies. Hence future work should include analyzing several years' worth of data and creating density maps that help establish a robust picture of the normal situation. Analysis of ship track data, density maps in particular, may uncover, e.g., expected areas of operation, routes of ship movement, and expected usage of ports. Density map products have several military applications that are described in a separate classified document.

References

[1] M. Hansen *et al.*, “High-resolution global maps of 21st-century forest cover change,” *Science*, vol. 342, no. 6160, pp. 850–853, 2013. [Online]. Available: http://www.sciencemag.org/content/342/6160/850.abstract

[2] P. Wessel and W. H. F. Smith, “A global self-consistent, hierarchical, high-resolution shoreline database,” *J. Geophys. Res.*, vol. 101, no. B4, pp. 8741–8743, 1996.

Abbreviations

AIS Automatic Identification System
ESRI Environmental Systems Research Institute
EU European Union
FFI Norwegian Defence Research Establishment
FFI Forsvarets forskningsinstitutt
GIS Geographic Information System
IMO International Maritime Organization
ISS International Space Station
LRIT Long Range Identification and Tracking
MMSI Maritime Mobile Service Identity
NAF North Atlantic Format
NCA Norwegian Coastal Administration
NEAFC North-East Atlantic Fisheries Commission
PPI pixels per inch
VMS Vessel Monitoring System
WPS Web Processing Service

About FFI

The Norwegian Defence Research Establishment (FFI) was founded 11th of April 1946. It is organised as an administrative agency subordinate to the Ministry of Defence.

FFI's MISSION

FFI is the prime institution responsible for defence related research in Norway. Its principal mission is to carry out research and development to meet the requirements of the Armed Forces. FFI has the role of chief adviser to the political and military leadership. In particular, the institute shall focus on aspects of the development in science and technology that can influence our security policy or defence planning.

FFI's VISION

FFI turns knowledge and ideas into an efficient defence.

FFI's CHARACTERISTICS

Creative, daring, broad-minded and responsible.

Om FFI

Forsvarets forskningsinstitutt ble etablert 11. april 1946. Instituttet er organisert som et forvaltningsorgan med særskilte fullmakter underlagt Forsvarsdepartementet.

FFIs FORMÅL

Forsvarets forskningsinstitutt er Forsvarets sentrale forskningsinstitusjon og har som formål å drive forskning og utvikling for Forsvarets behov. Videre er FFI rådgiver overfor Forsvarets strategiske ledelse. Spesielt skal instituttet følge opp trekk ved vitenskapelig og militærteknisk utvikling som kan påvirke forutsetningene for sikkerhetspolitikken eller forsvarsplanleggingen.

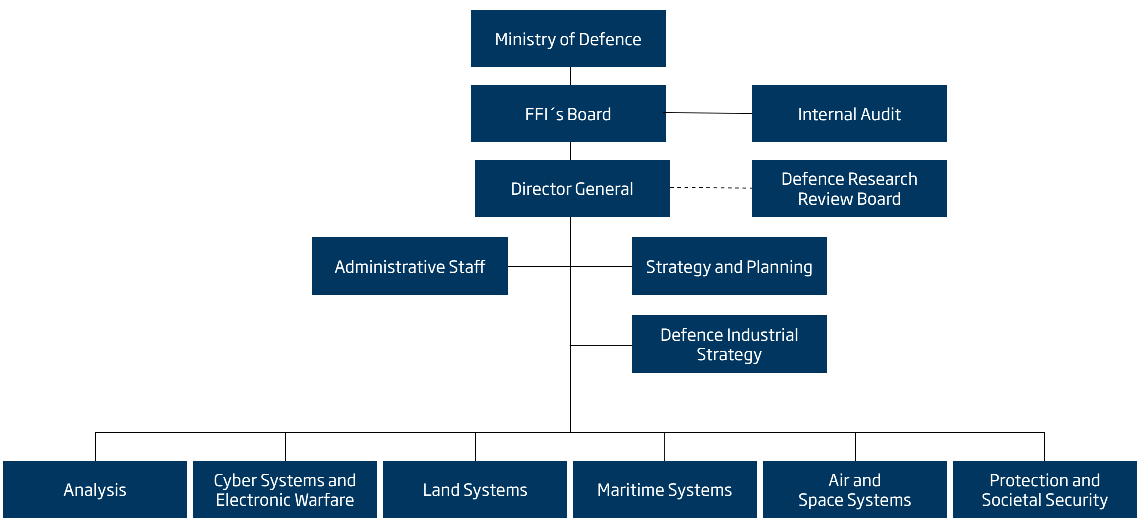
FFIs VISJON

FFI gjør kunnskap og ideer til et effektivt forsvar.

FFIs VERDIER

Skapende, drivende, vidsynt og ansvarlig.

FFI's organisation



Forsvarets forskningsinstitutt
Postboks 25
2027 Kjeller

Besøksadresse:
Instituttveien 20
2007 Kjeller

Telefon: 63 80 70 00
Telefaks: 63 80 71 15
Epost: ffi@ffi.no

Norwegian Defence Research Establishment (FFI)
P.O. Box 25
NO-2027 Kjeller

Office address:
Instituttveien 20
N-2007 Kjeller

Telephone: +47 63 80 70 00
Telefax: +47 63 80 71 15
Email: ffi@ffi.no

FFI Forsvarets forskningsinstitutt
Norwegian Defence Research Establishment