AIS REPORT

TECHNICAL REPORT
TASK C2.1.1 AIS DATA ANALYSIS I

COWI
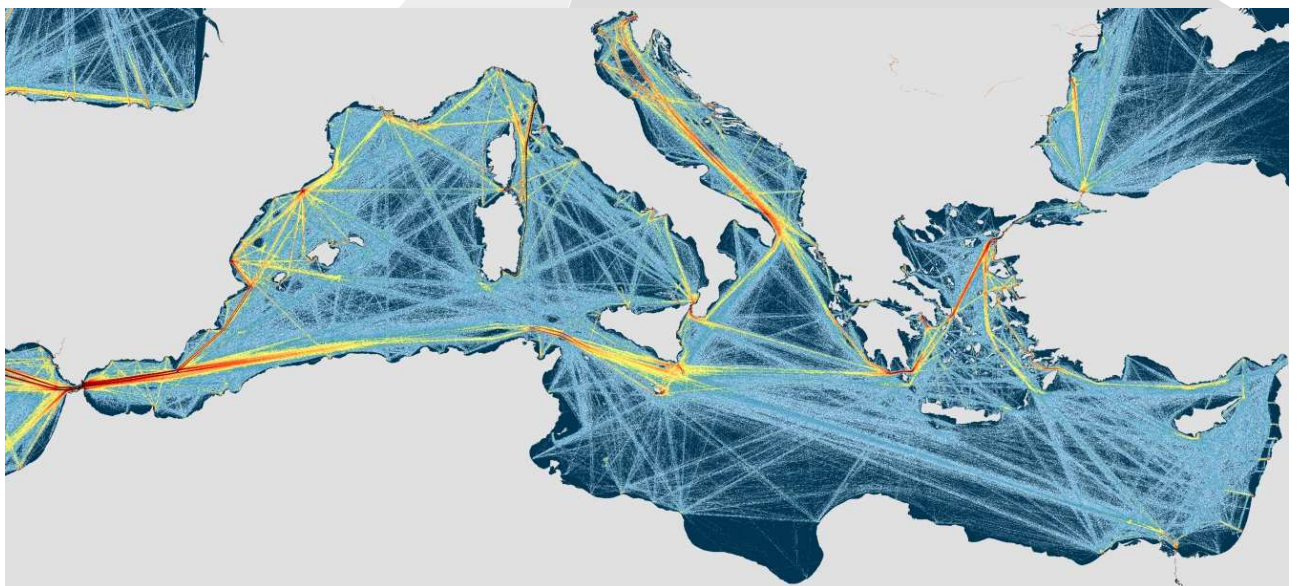
APRIL 2014
MEDESS-4MS

AIS REPORT

CONTENTS

1   About

The MEDESS-4MS (Mediterranean Decision Support System for Marine Safety) project is dedicated to the maritime risks prevention and strengthening of maritime safety related to oil spill pollution in the Mediterranean.

The present technical report describes part one of the ship traffic analysis in the Mediterranean Sea. This includes analysis of the AIS data, deriving a route net and developing the idealised traffic on this basis. In part two of the traffic analysis information about the individual vessels will be analysed in detail and representative vessels will be chosen for the model of the various accident scenarios required later in the risk analysis of oil spills in the project area.

2   Introduction

Modelling the ship traffic in an appropriate way is one of the corner stones of the risk analysis. As in the earlier BRISK project covering the Baltic Sea (BRISK, 2012) and the BE-AWARE project covering the North Sea (BE-AWARE, 2014), it is based on AIS ship traffic data. AIS (Automatic Identification System) consists of position messages broadcast by each single vessel, with information on identity, position, speed over ground, course over ground etc. AIS has been introduced as part of IMO's International Convention for Safety of Life at Sea (SOLAS) and is compulsory for all cargo vessels with a gross tonnage of 300 tons or more as well as all passenger vessels regardless of size. The intention is to increase the safety of vessels operating close to each other. In addition to this primary purpose, it is possible to collect AIS data by means of coast stations, which can be used to establish a comprehensive ship traffic database. The methodology described in this note requires the availability of such a database.

It is in the nature of such a database that it is very extensive and that its raw content cannot be applied directly in any ship accident risk model. This discrepancy is solved by generating a discrete route net covering the whole sea area and associating the individual AIS traces with the nearest net segments. The resulting route-based traffic description provides an unmatched basis for the following ship accident risk analysis.

The present report describes

- the applied/required data sources (Section 3)

- the AIS data analysis including the generation of the discrete route net (Section 4)

- the idealised traffic model (Section 5)

3   Ship traffic data

AIS data

The AIS database operated by the Italian Coast Guard is the primary data source for establishing the traffic model, /MEDESS AIS, 2013/. It holds records of AIS messages of all AIS-equipped vessels in the Mediterranean Sea. Data were provided for a six-month period covering both winter and summer months. This is assumed to provide the adequate basis to describe seasonal differences and provide statistically significant amount of data.

A period lasting from 1 February 2013 to 31 July 2013 is applied as reference period, since these were the six most recent months at the time the MEDESS-4MS project was initiated.

IHS Fairplay data

The World Shipping Encyclopaedia (WSE) issued by IHS Fairplay is a database containing information on a large number of parameters for each vessel. Since every vessel has a unique IMO number, which is both used in WSE and for AIS, it is possible to determine relevant vessel characteristics for the vessels recorded in the AIS data base (type, size, geometry, single or double hull etc.).

The WSE has earlier been known as Lloyd's Register, i.e. prior to its purchase by IHS Fairplay.

4   AIS analysis

4.1   Basics

The AIS messages sent by the vessels consist of position reports (POS) and static reports (STAT), as described in Recommendation ITU-T M. 1371-1 issued by the International Telecommunication Union (ITU).

POS reports

POS reports are sent approx. every two seconds and contain information on vessel position, course, speed etc. In this reports, the ship is identified by its MMSI number.

STAT reports

STAT reports are sent every six minutes and contain information about the ship itself, amongst others MMSI and IMO number, name, call sign, size, actual draught, category of potentially hazardous cargo and position of the AIS transmitter relative to the ship.

It has generally been observed that AIS reports, where vessels are supposed to enter data themselves are not always reliable. Information that needs to be updated by the crew (cargo, actual draught, destination etc.) is therefore not necessarily valid, whereas automatically updated information (position, course, speed) can be expected to be more reliable.

4.2   Ship identification

AIS data have been delivered by Italian Coast Guard in a form of 68 Access database files. We received data as a one dataset which means that the POS reports have been compiled with STAT reports. AIS data are down-sampled to approx. 6 minute long intervals. We received more than 368 million of distinct AIS records. Due to the size of the database no filtering on raw AIS has been performed. We have identified ships by extracting distinct MMSI numbers from AIS dataset (17,799). Afterwards ship types and sizes have been identified either by joining with IHS database by use of IMO number or based on properties extracted directly from AIS, in cases where information from IHS has not been available or ship's

IMO number was unknown. 17,108 MMSI numbers belonged to ships will well defined ship type and 15,190 of them were larger than 300 GT. 1,102 MMSI number could not be identified via IHS database. This identification of the ships is necessary since the traffic model takes into account only ships larger than 300 GT sailing within analysis area. Daily variation of number of AIS reports received from vessels larger than 300 GT and most probably larger than 300 GT is presented in
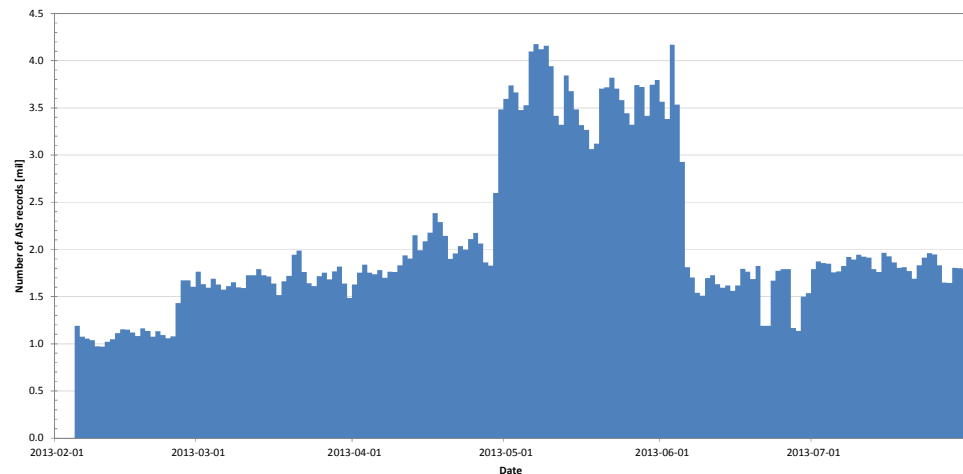

Figure 4-1   Variation in number of daily AIS reports in period 1 February 2013-31 July 2013.

4.3   Traffic intensity

As a basis for the further analysis, it is necessary to determine the resulting traffic density for the Mediterranean Sea. This density should – apart from confirming a correct data processing – be suitable as decision basis for the generation of a route net and the following data analysis (Section 4.4).

The density is determined by following the trace of a specific vessel – longitude & latitude – and registering its path across a predefined quadratic grid. This approach is implemented by simply rounding the trace coordinates to the nearest multiples of the cell length (Δlong and Δlatt) in the grid net (see Figure 4-2).

Values of Δlong and Δlatt have been chosen to secure the grid cell size of approx. 500 m x 500 m. Close to 8 million cells have been defined to cover Mediterranean Sea. It should be noted that only cells with at least one AIS record have been created.

Each AIS record can be assigned to a single cell in the grid and counted afterwards. Count of AIS records per cell have been performed, but the counts per cell cannot be directly translated into the number of passages populating the route (considering constant sampling rate and vessels with different velocities then the distance between sequential AIS records is different, therefore number of "footprints" per length is different for those vessels).
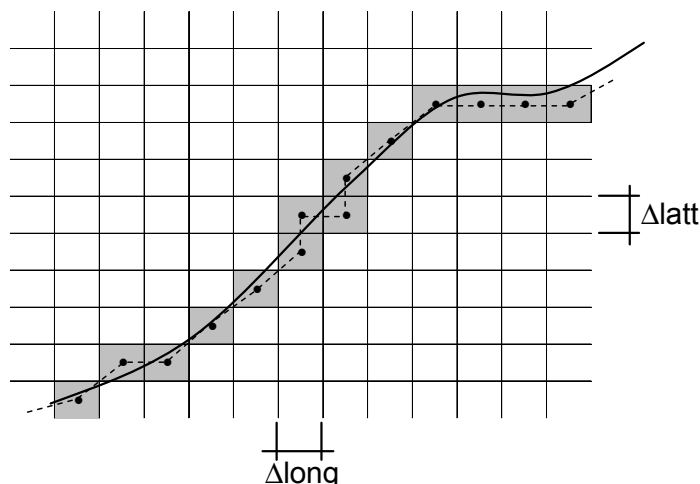
Figure 4-2   Digitalisation of a vessel track in order to determine the traffic density

A simple count of the recorded AIS vessel passages yields a traffic intensity plot based on all received data presented in Figure 4-3. The colour scale is not linear and therefore not only most trafficked routes are visible.

One can easily notice that the traffic has a tendency to concentrate along routes. It is especially well distinguishable in narrow navigation channels such as sounds (ex. Strait of Elafonisos – south-western exit from the Aegean Sea) or in areas with existing traffic separations schemes (ex. Straits of Gibraltar TSS, North of Cap Bon TSS or North Adriatic Sea TSS). However, the tendency of following clearly distinguishable routes is general, since vessels always follow the most direct possible route between two destinations and since the number of relevant destinations is limited. One can also notice areas where the traffic intensity seems to diminish which can be caused by large distance to the shore and hence some of the transmitted radio AIS reports are missing or not correctly recorded (ex. Levantine Sea or Ionian Sea). The diminished traffic intensity can be also caused by lack or reduced number of AIS terrestrial base stations (ex. South of Libyan Sea). Although main routes can easily be identified in Figure 4-3, creation of the whole route net on this basis would be very difficult. Different types of ships such as merchant, passenger/ferry, cruise and offshore vessels have different sailing patterns and the contributions from different ship types cannot easily be distinguished. Therefore separate maps for different tankers (Figure 4-4), general cargo (Figure 4-5), container (Figure 4-6), passenger (Figure 4-7), cruise (Figure 4-8) and offshore (Figure 4-9) vessels have been prepared. Separation between those different types results in much clearer traffic patterns. Also less populated routes are easily distinguishable now. One can also notice major differences between traffic patterns of ex. tanker, cruise and offshore vessels.
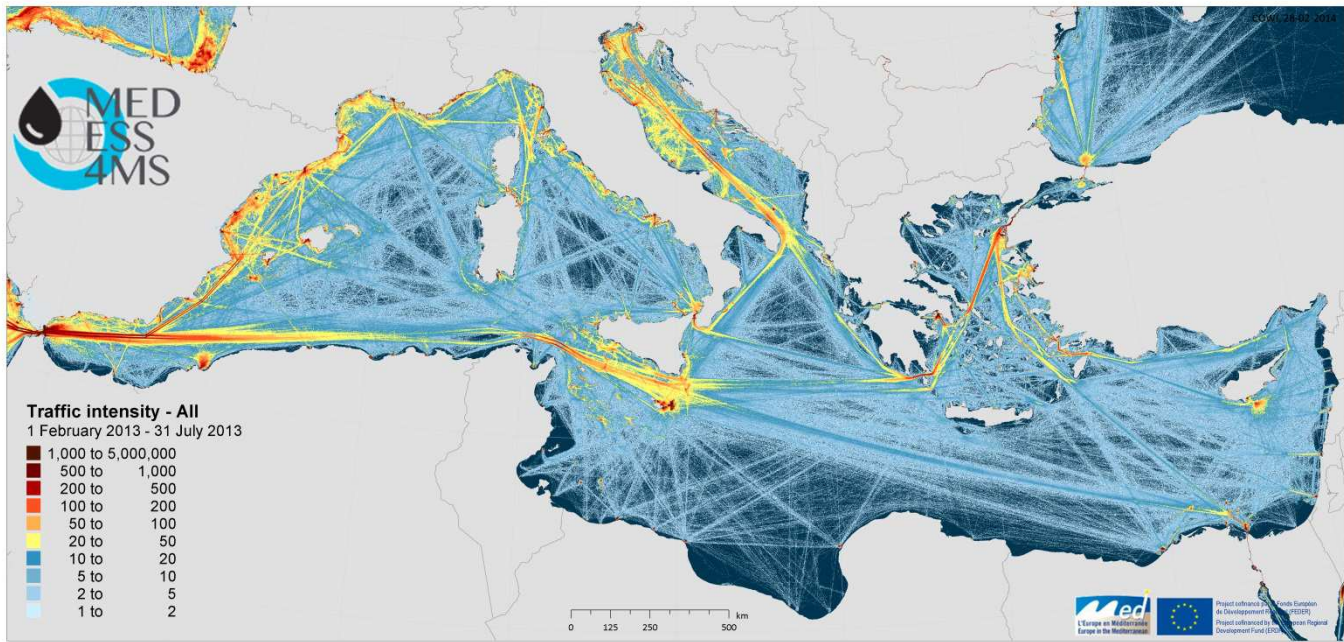
Figure 4-3   Map of traffic intensity based on counted number of all AIS records per cell /MEDESS AIS, 2013/.
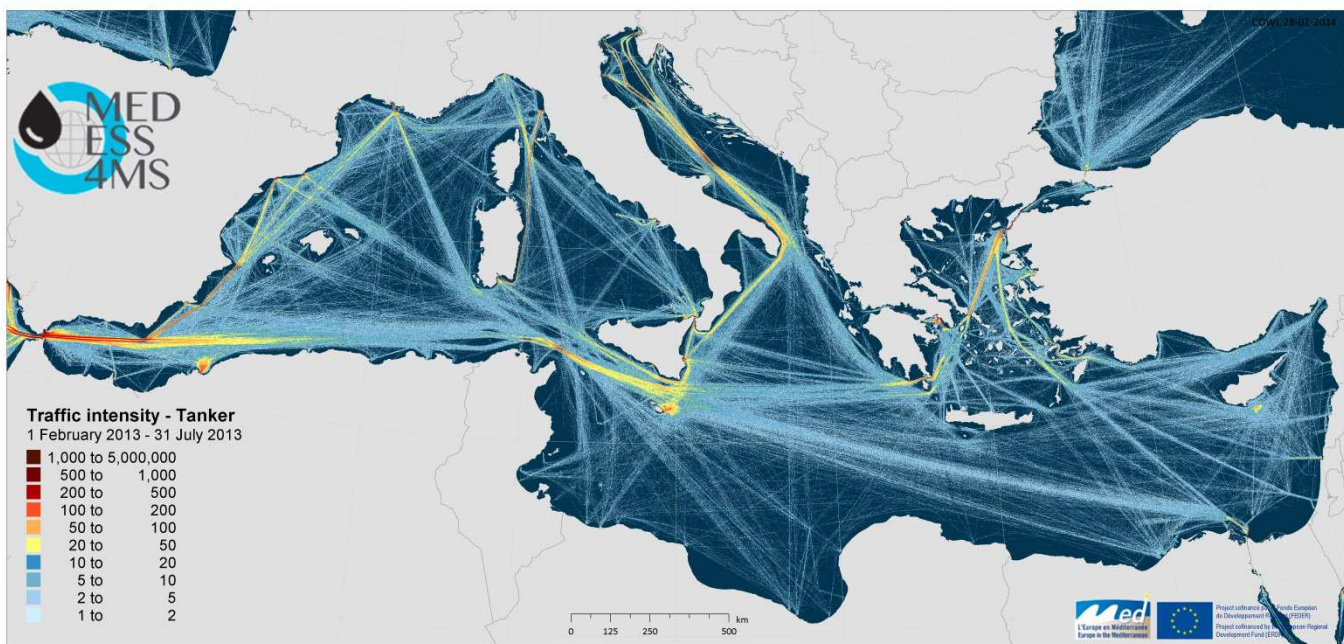

Figure 4-4   Traffic density map for tankers in the Mediterranean Sea based on the recorded traffic /MEDESS AIS, 2013/
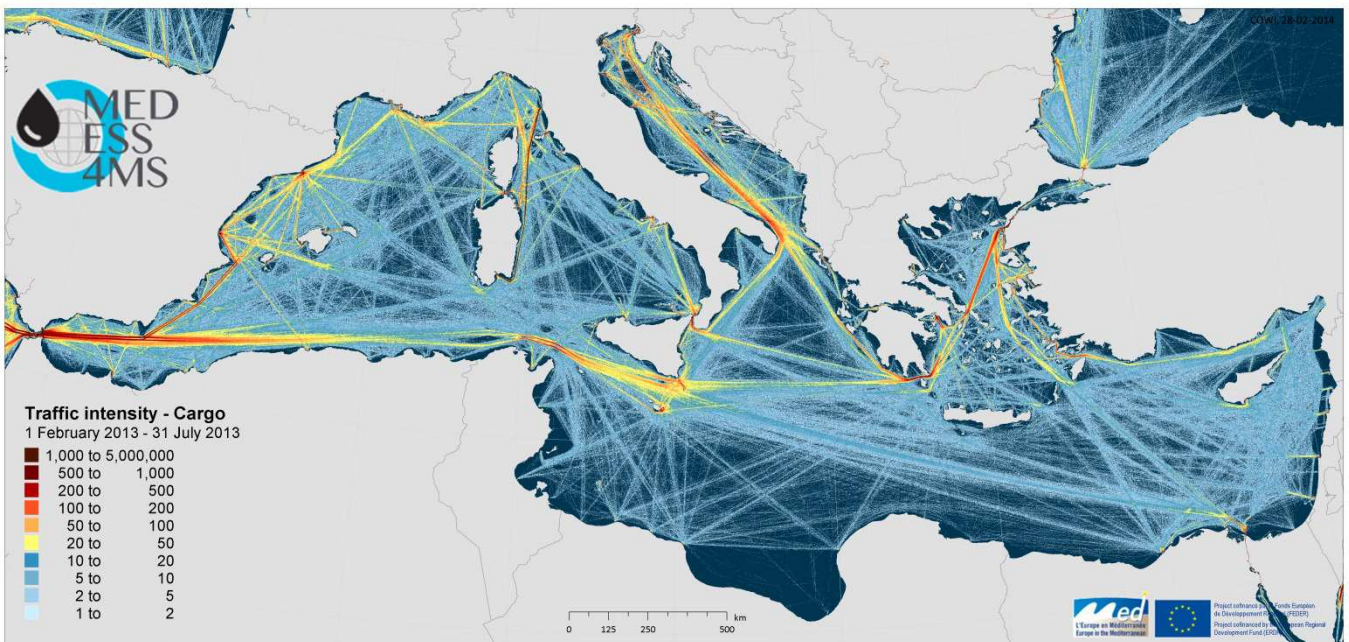
Figure 4-5   *Traffic density map for all general cargo ships in the Mediterranean Sea based on the recorded traffic /MEDESS AIS, 2013/*
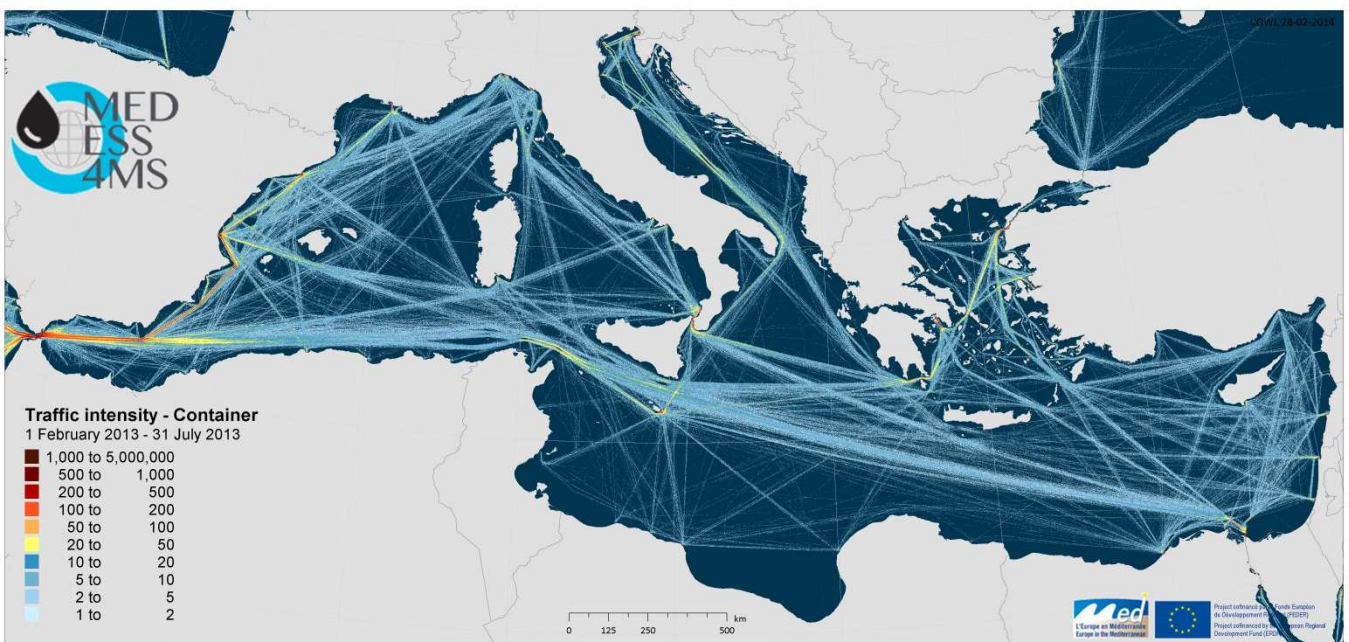

Figure 4-6   *Traffic density plot for container vessels traffic in the Mediterranean Sea based on the recorded traffic /MEDESS AIS, 2013/*
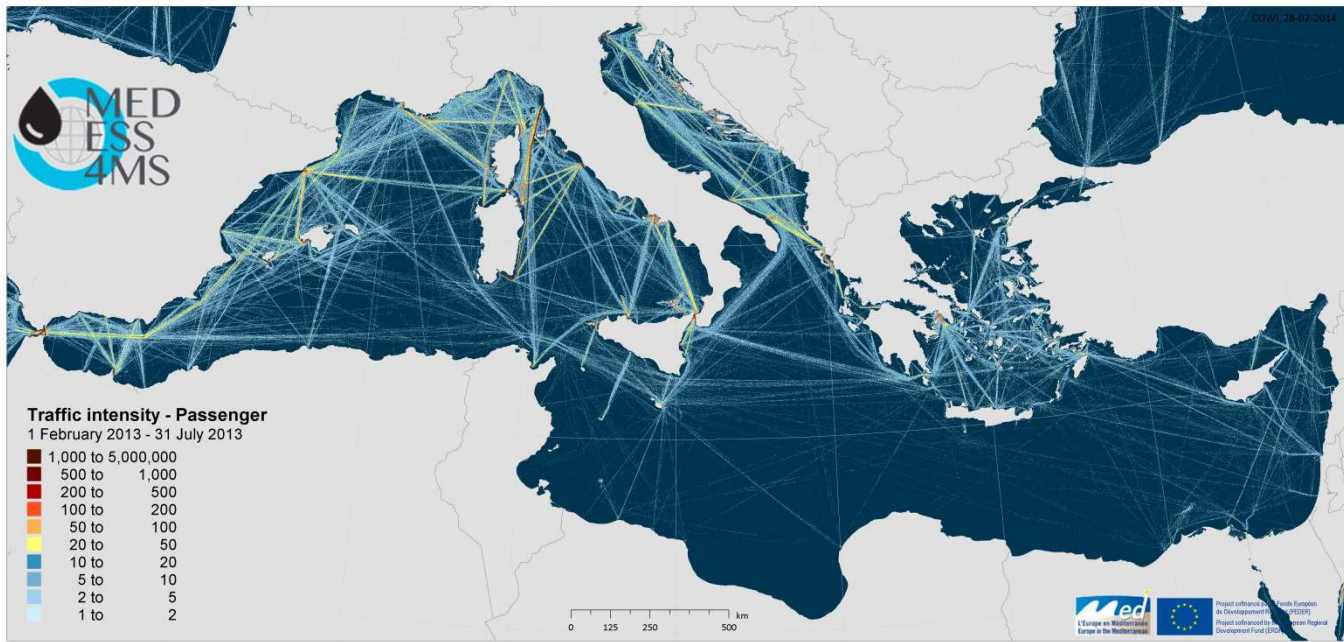
Figure 4-7   Traffic density plot for passengger vessels traffic in the Mediterranean Sea based on the recorded traffic /MEDESS AIS, 2013/
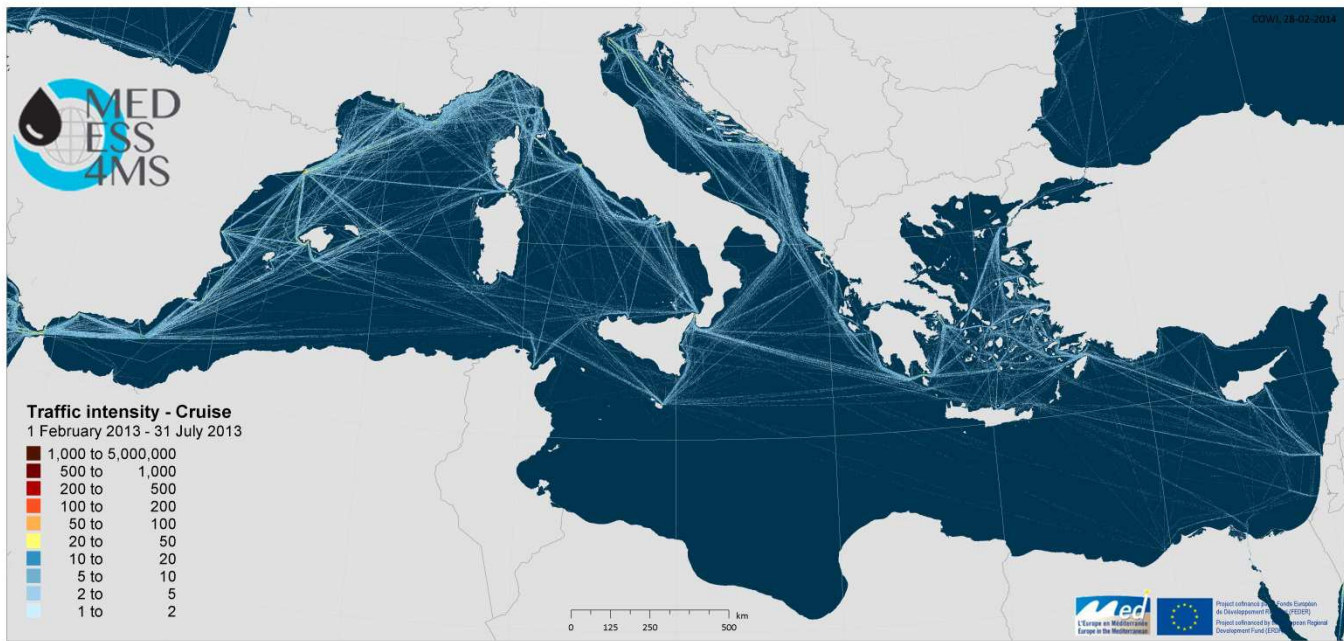

Figure 4-8   Traffic density plot for cruise vessels traffic in the Mediterranean Sea based on the recorded traffic /MEDESS AIS, 2013/
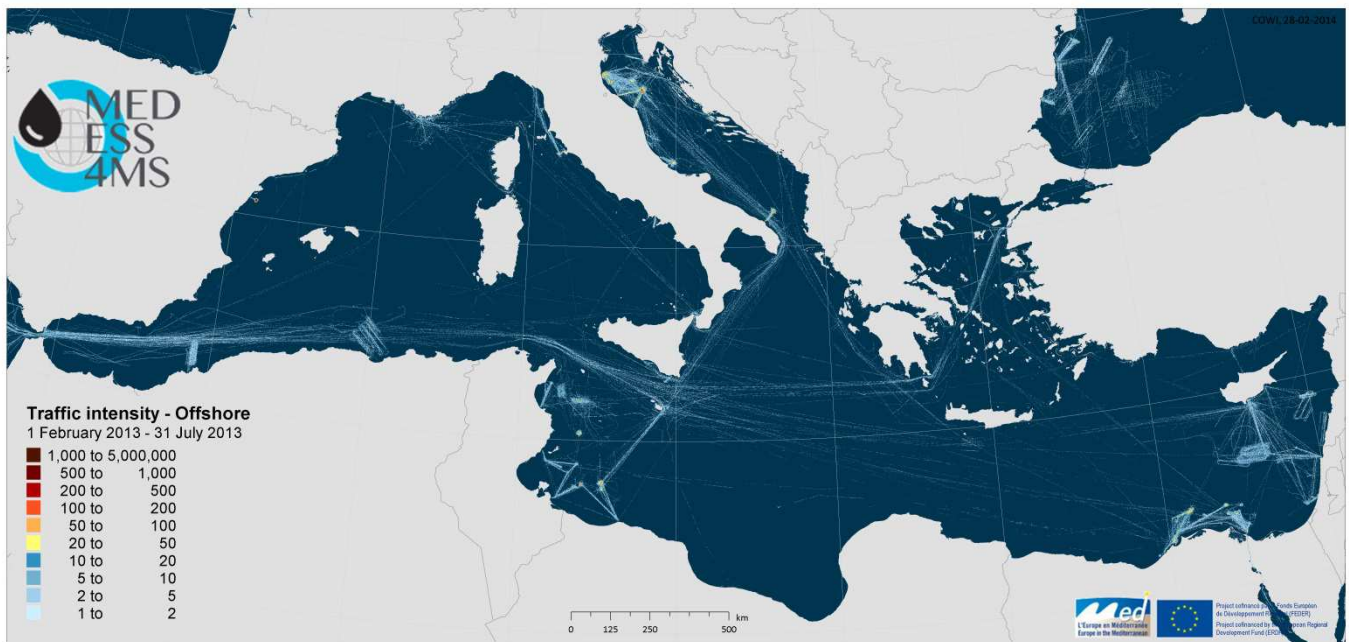
Figure 4-9   Traffic density plot for offshore traffic in the Mediterranean Sea based on the recorded traffic /MEDESS AIS, 2013/

It should be emphasized that intensity maps are only serving as a background enabling creation of idealised route net that can represent the vessel traffic.

4.4   Route generation and analysis

The tendency of the traffic is to concentrate along routes and this indicates that a populated idealised route net could be a representative approximation of the sea traffic. On some routes traffic can be spread loosely to both sides of the route axis, but this does not cause any conceptual concerns.

Route generation and analysis means:

- definition a geographic route net, which can represent the vessel movements in the Mediterranean Sea with good precision

- mathematical analysis of the route net, i.e. to determine the shortest possible paths through the net between two locations

- mapping the AIS trace, i.e. to associate each AIS point with a route net segment.

- determination of various relevant statistics for each route segment, e.g. the distribution of the vessels' deviation from the route segment axis.

Definition of the route net
This work is done manually by creating a route net on a series of background maps consisting of an intensity plots and the sea charts. The route net has to

accommodate all traffic patterns of all vessel types. This work is performed in a
GIS programme (MapInfo). Once the route net has been defined, its geometry is
exported to Excel (combined with Visual Basic for Applications) for further
analysis and in order to check its consistency (all route ends meeting in one node
shall have the same coordinates).

Figure 4-10   The route net created for the Mediterranean Sea

Figure 4-10 shows the route net developed for the Mediterranean Sea.

The route net consists of two types of elements:

- nodes (defined by their longitude and latitude)
- route segments connecting the nodes

The developed route net for the project area consists of 5301 route segments and
2582 nodes.

Analysis of the route net

The route net defines different possible ways through the project area and the
concept of "the shortest way" between two nodes in the route net is a useful
support function for associating the AIS points to route segments.

The shortest way between two nodes is determined by means of a simple iterative
algorithm based on Markov network logic. The results are deposited in two
separate matrixes which are created for the particular route net. The two matrixes
are:

- *NN(i,j)* matrix describes that the shortest way from node *i* to node *j* starts by going from node *i* to *NN(i,j)*

- *MLi,j)* matrix contains the length of the shortest way from node *i* to node *j*.

Representation of the vessel's passage through the route net consists of a list of used route segments. Therefore representation can be stored as a table with the sequence of route segments:

TrackNo	IMO	Time	RouteNo
…	…	…	…
100	9322255	2013-02-11 10:10:05	-4134
100	9322255	2013-02-11 11:10:39	-4119
100	9322255	2013-02-11 11:34:57	4088
100	9322255	2013-02-11 12:35:51	4024
100	9322255	2013-02-11 14:38:33	-4000
100	9322255	2013-02-11 15:20:39	-3942
…	…	…	…

Adding a sign in front of the route segments is a simple way of marking the passage direction.

Systematic mapping of the AIS traces

With the above-described basis it is possible to map the individual AIS traces systematically. As a first step, it needs to be defined, when a track – i.e. a sequence of AIS points – can be concluded to represent a coherent journey. This definition needs to take the possibility of data transmission interruptions into account (see Figure 4-11). It would simplify the mapping procedure significantly to neglect missing sequences. However, this would result in a systematic underestimation of the traffic in certain area, if e.g. one local coast station has been out of order during a certain period of time. Furthermore, information about the total journey and its origin and destination would get lost.
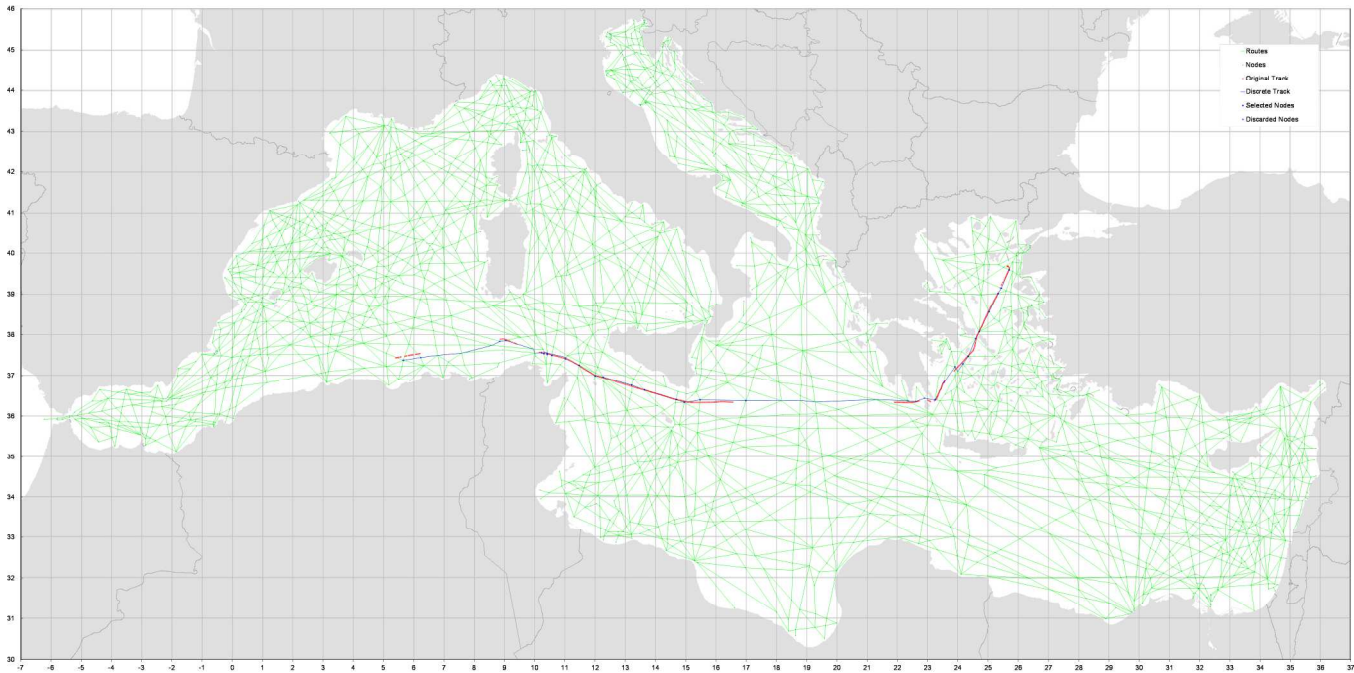
Figure 4-11   Example of AIS points of an identified trace with pronounced transmission interruptions

Therefore, the mapping procedure is refined in order to handle interrupted traces and to interpolate the missing sections. When an individual trace is identified, the following conditions are applied:

- The time difference between two successive AIS points must not exceed 72 hours

- An approximate vessel speed v_{appr} is calculated as the distance between two points divided by the time difference between the two messages. The two points are considered as part of the same trace if

  - $v_{appr} > 0$ knots (the ship does not stand still)

  - v_{appr} is finite (i.e. not very large, which would indicate an unrealistic jump and therefore an error)

With these conditions, the most significant errors are filtered away and the trace is interrupted, if the vessel stops. The latter is chosen in order to obtain two separate traces in case a vessel is lying still in a port or at anchor.
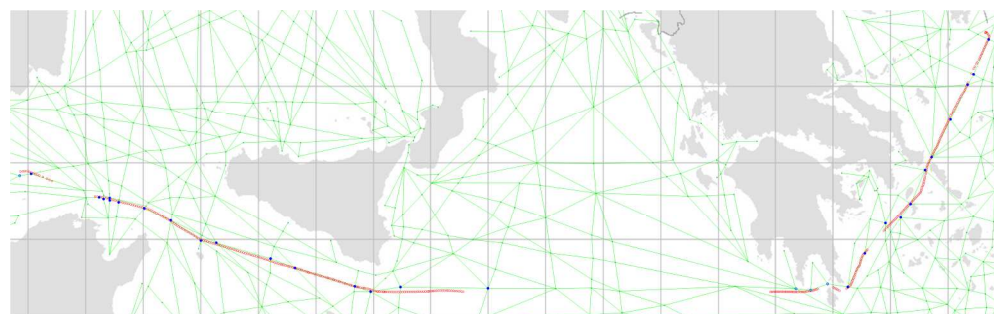

Figure 4-12   Determination of which nodes in the route net are close to the AIS trace

When a sequence of AIS points has been recognised as a continuous track (as shown in Figure 4-11), an algorithm determines, which nodes are passed at the closest distance (see Figure 4-12).

To limit the number of analyses, it has been necessary to simplify and optimize this determination of the closest node to the track. It is done by determination of the closest node for a number of points evenly distributed in the area of the analysis and storing results in the table. For this purpose the centres of the cells used for creation of traffic density mapping are used. In the analysis of AIS points the cell the AIS point belongs to is first determined and via above mentioned reference table the closest route net nodes is then found. This discrete grid used in this approximation method has sufficiently good resolution (approx. 500m in both directions) to assure that no major error happen while determining the node closest to the AIS trace.

Once the sequence of nodes in the route net has been determined, another algorithm removes unrealistic outcomes caused by the mathematical logics in the first algorithm (see Figure 4-13). Another typical misinterpretation are vessels that seem to sail into a "dead end", i.e. by following a route segment first in one direction and then into the opposite direction before continuing. This error is equally removed.
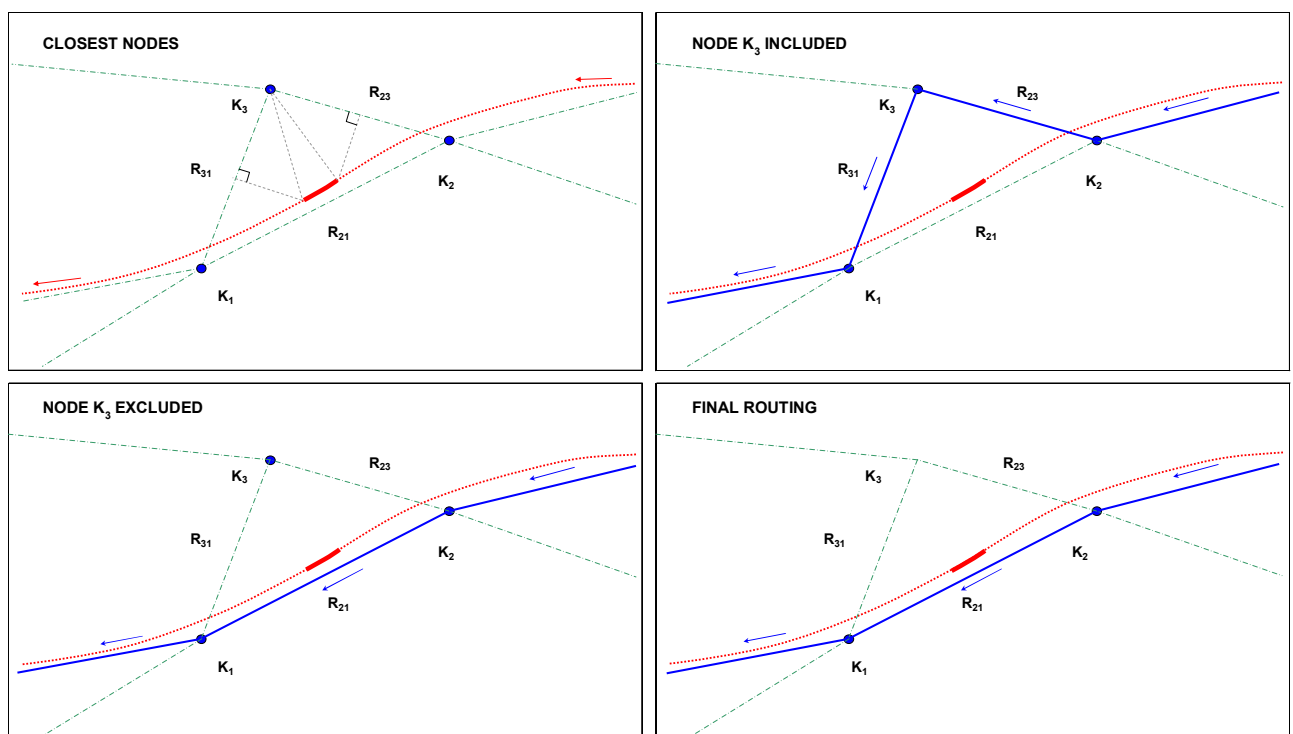

Figure 4-13   An example showing how the closest node (K_3) can mislead the mapping algorithm

During the route mapping procedure it is determined, which AIS points can be associated with which route segment passages. This information is subsequently used for determining the mean value and spreading of the average geometrical distance between the points and the ideal line in the route net. These statistics are

required for the calculation of the collision frequency of vessels sailing along the same route segment.

Key results for track mapping usage of AIS reports

Total number of AIS reports	368,031,865
Number of AIS reports from ships chosen for AIS mapping	226,620,439
AIS reports with identified track	74,049,371
Identified route passages	388,025

5   The resulting traffic model

The resulting traffic model is essentially described as a database table containing all identified route passages (events, where a vessel passes a route segment) combined with information about passage direction and vessel characteristics from the World Shipping Encyclopaedia (WSE). Using this detailed model has the following advantages:

- traffic surveys can be performed very flexibly based on the detailed ship characteristics from the WSE

- the actual journeys of the respective vessels are contained in the description, since sequences of route passages are tied together by a common track number and the date information

- the passage of the vessels through the respective nodes in the route net – i.e. on which route segment does a vessel arrive at a node and on which route segment does it continue – are contained in the description and can be used in the ship collision model

The database provides traffic data for the calculation of accident and spill frequencies, which are directly dependent upon the traffic, its volume and composition.

In order to display the content of the traffic model, different tables can be extracted – the aggregated transport activity (sailed nautical miles) and the distribution of the traffic on specific routes to different ship types and sizes.

Classification of ships

The information on the identified vessels that can be found in the World Shipping Encyclopaedia is more detailed than what is meaningful in the context of the risk analysis. This broad classification is reduced to 24 different types as shown in Table 5-1. Type 25 "unknown" is not used in the final traffic model, but is used in order to classify the remaining group that cannot be identified during the model establishment.

The ship groups introduced in Table 5-1 are used for preparing statistics and results.

Table 5-1   Ship types used in the model (left) and general groups of types used for preparing statistics and results (right)

Type ID	Type description
1	Work vessel
2	Car transport
3	Bulk
4	Bulk/Oil
5	Container
6	Fishing vessel
7	Ferry
8	Ferry/Ro-Ro
9	Cruise ship
10	Reefer
11	Nuclear fuel
12	Offshore
13	Ro-Ro
14	Tug
15	General cargo
16	Navy
17	Tanker, food
18	Tanker, gas
19	Tanker, chemical/products
20	Tanker, chemical
21	Tanker, product
22	Tanker, crude oil
23	Tanker, others
24	Others
25	Unknown

Vessel group	Type description
Tankers	Bulk/oil
	Tanker, food
	Tanker, gas
	Tanker, chemical/prod.
	Tanker, chemical
	Tanker, product
	Tanker, crude oil
	Tanker, others
Bulk carriers	Bulk
General cargo	General cargo
Packed cargo	Car transport
	Container
	Reefer
	Nuclear fuel
	Offshore
	Ro-Ro
Ferry and passanger traffic	Ferry
	Ferry/Ro-Ro
	Cruise ship
Others	Work vessel
	Fishing vessel
	Tug
	Navy
	Others
	Unknown

Within a GIS system, traffic can be illustrated graphically for individual traffic segments. Mapped traffic of all vessels is presented in Figure 5-1. Total traffic intensity on route segments for oil and chemical tanker, container and general cargo vessels, cruise and passenger ships as well as offshore vessels is presented in

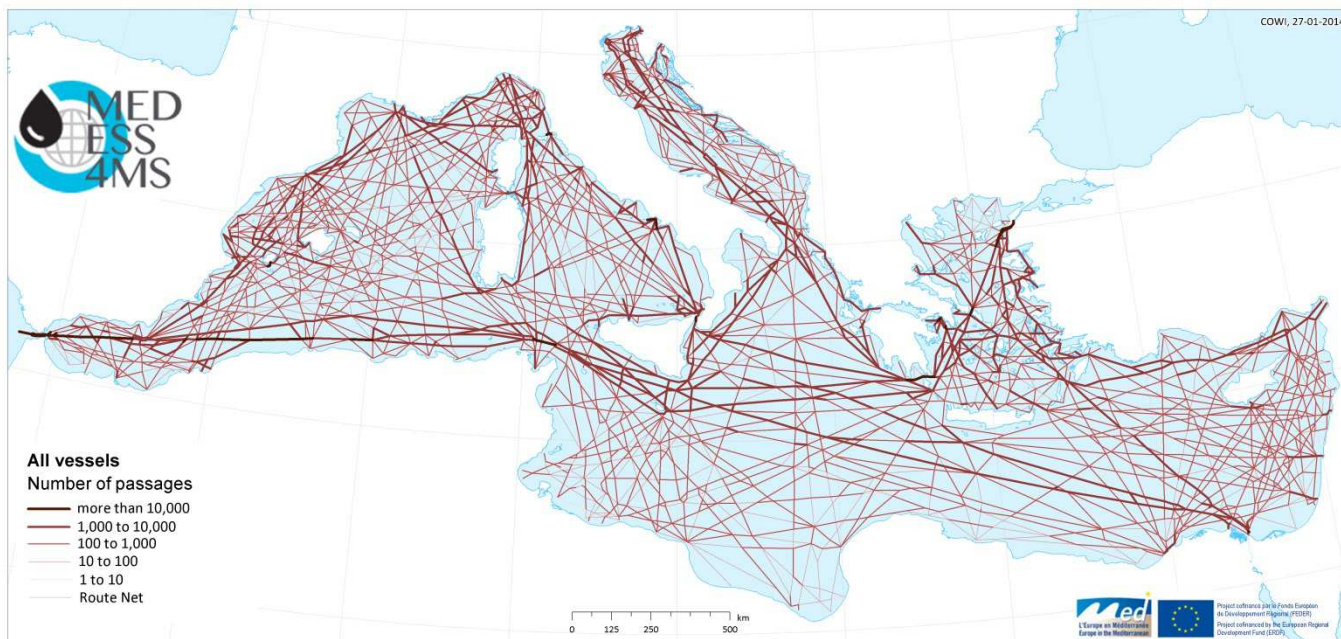Figure 5-2, Figure 5-3, Figure 5-4 and Figure 5-5 respectively.

Figure 5-1   Map of the total traffic intensity on route segments of all vessels larger than 300 GT in period 1 February 2013 – 31 July 2013.
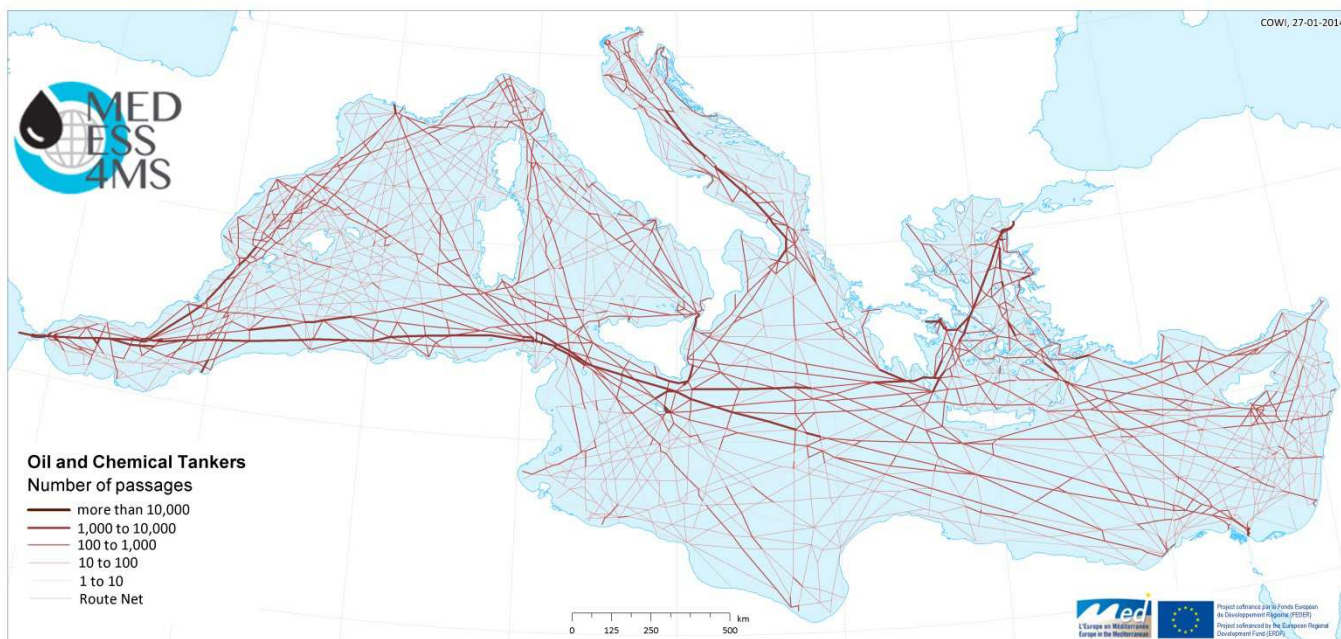

Figure 5-2   Map of the total traffic intensity on route segments of oil and chemical tankers larger than 300 GT in period 1 February 2013 – 31 July 2013.
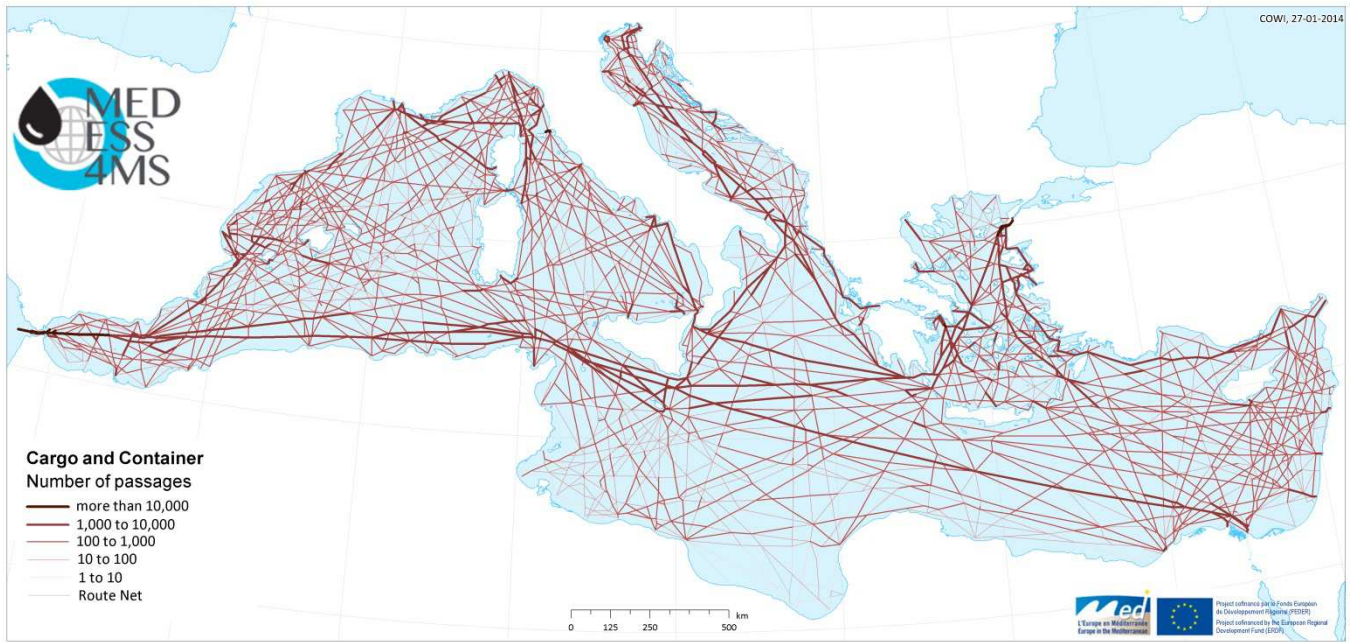
Figure 5-3   Map of the total traffic intensity on route segments of container and general cargo vessels larger than 300 GT in period 1 February 2013 – 31 July 2013.

Figure 5-4   Map of the total traffic intensity on route segments of cruise and passanger vessels larger than 300 GT in period 1 February 2013 – 31 July 2013.
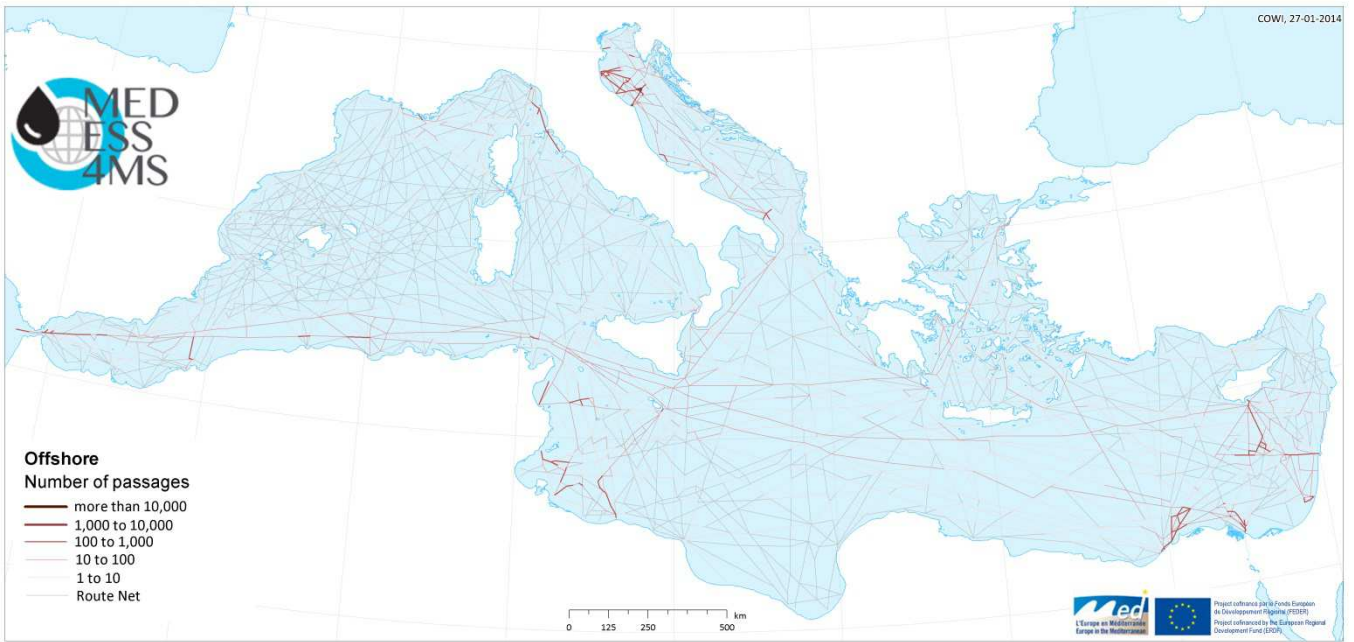
*Figure 5-5   Map of the total traffic intensity on route segments of offshore vessels larger
than 300 GT in period 1 February 2013 – 31 July 2013.*

6   References

/COWI, 2007/   *Risikoanalyse: Olie- og kemikalieforurening i danske farvande (In Danish) [Risk analysis: Oil and chemical pollution in Danish waters]*, prepared for Danish Ministry of Defence by COWI, COWI report 63743-1-01, October 2007

/BRISK, 2012/   *Project on sub-regional risk of spill of oil and hazardous substances in the Baltic Sea (BRISK): Risk Method note.* COWI for Admiral Danish Fleet HQ, report no. 70618-3.1.1, rev. 3.0, April 2012

/BE-AWARE, 2014/   (BE-AWARE):*Technical Sub Report 1 – Ship Traffic.* COWI for Bonn Agreement Secretariat, February 2014

/MEDESS AIS, 2013/   AIS database from the project area covering the period 1 February 2013 to 31 July 2013. Received from LTJG Antonio VOLLERO, Italian Coast Guard - Headquarters