Marine
Management
Organisation

Spatial trends in shipping activity (AIS derived shipping activity – data standards)

November 2013

This page has been intentionally left blank.

Spatial Trends in Shipping Activity (AIS Derived Shipping Activity - Data Standards)

MMO Project No: 1042

Marine Management Organisation

Project funded by: The Marine Management Organisation

abp mer marine environmental research

Report prepared by: ABP Marine Environmental Research Ltd

Disclaimer

This report should be cited as

MMO (2013). Spatial trends in shipping activity (AIS derived shipping activity - data standards). A report produced for the Marine Management Organisation, pp 47. MMO Project No: 1042. ISBN: 978-1-909452-12-1.

First published November 2013.

Acknowledgement

The authors would like to thank the kind assistance offered by the Maritime and Coastguard Agency in providing data and advising on its network of receiving stations.

Contents

Figures

Tables

List of acronyms

ABPmer	ABP Marine Environmental Research Ltd
AIS	Automatic Identification System
ASCII	American Standard Code
COG	Course over the Ground
DG	Dangerous Goods
DTE	Data Terminal Equipment
ETRS	European Terrestrial Reference System
GIS	Geographic Information System
GLONASS	Globalnaya Navigatsionnay Sputnikovaya Sistema
GPS	Global Positioning System
HS	Harmful Substances
HSC	High Speed Craft
ICO	Information Commissioners Office
ID	Identification
IMO	International Maritime Organization
MCA	Maritime and Coastguard Agency
MEDIN	Marine Environmental Data and Information Network
MP	Marine Pollutants
MMO	Marine Management Organisation
MMSI	Maritime Mobile Service Identity
NM	Nautical Miles
QA	Quality Assurance
QGIS	Quantum Geographic Information System
ROT	Rate of Turn
SART	Search and Rescue Transponder
SEP	Strategic Evidence Plan
SOG	Speed over the Ground
SQL	Structured Query Language
VHF	Very High Frequency
WIG	Wing-In-Ground
XML	Extensible Markup Language

Executive Summary

The Marine Management Organisation (MMO) commissioned ABP Marine Environmental Research Ltd (ABPmer) to develop approaches for the mapping of shipping activity using Automatic Identification System (AIS) data supplied by the Maritime and Coastguard Agency (MCA). This work has been commissioned by the MMO to provide a national AIS derived dataset for the mapping and analysis of spatial shipping trends. The aim of this report is to document the methodology developed for processing the MCA provided UK AIS dataset to create spatial datasets which show shipping density and shipping routes. Through the use of a common methodology, AIS processing and analysis can be conducted by UK devolved administrations to a consistent and commonly agreed approach.

The main text of this ‘AIS Derived Shipping Activity - Data Standards’ report should be read alongside Annex 2 - a ‘Draft data guideline for the presentation of AIS Derived Vessel Spatial Data’. This data guideline has been developed as an additional output of this MMO commissioned project which the MMO will use in a wider consultation with industry and regulators to try and reach agreement on how derived data should be produced and archived. This draft guideline provides a suggested definition for the format of AIS derived vessel spatial data and its representation within mapped products and builds on the exisiting standards for capture and display of AIS raw data.

The processing of AIS data to create a national dataset to describe shipping spatial trends has been achieved through the creation of vessel vector transits (vessel journeys defined by measuring Speed Over the Ground (SOG)). These vessel vector lines have attributes to describe the vessel type, vessel dimensions and transit classification. Through the combination of vessel transit data, a density grid at a resolution of 0.5km has been produced. This allows the creation of a range of spatial data products from the available transit line attributes. Following the code of practice in the Information Commissioners Office publication on anonymisation; vessel transit line identification has been anonymised.

To define a yearly AIS dataset, seven complete days (of 24 hours) from the start of April, July and October 2010; plus January 2011 were provided by the MCA. This reduces seasonality effects by providing four representative weeks from four months across the year.

The derived AIS data products have been provided in a Geodatabase with the intention that this will be made freely accessible (within the bounds of data protection rules) in line with the Government’s principles on open data.

1. Introduction

The Marine Management Organisation (MMO) has commissioned ABP Marine Environmental Research Ltd (ABPmer) to develop approaches for the mapping of shipping activity using Automatic Identification System (AIS) data as recorded by the Maritime and Coastguard Agency (MCA). This work has been commissioned by the MMO to provide AIS derived shipping activity datasets detailing transit routes and shipping density in the South Coast Inshore and Offshore Marine Plan Areas.

A secondary objective of this project has been to outline the methods and procedures employed in the development of these AIS derived shipping activity data products. These are detailed in this data standards report and in a draft data guideline (Annex 2) which attempts to specify an archive standard.

The MMO aim to use this data standards report and its annexes to promote try and facilitate a (UK) agreed standard approach to the production of AIS derived shipping activity data as well as their representation and format for archival.

This report will also be used to inform interested parties of the methodology applied, the processing steps used to generate the end outputs from vessel AIS positional and voyage data.

Comments on the content or this report are encouraged to be sent to the MMO from industry, Government planning organisations, Government departments and sectoral representatives from ports, shipping and marine safety.

This Data Standards Report covers the following topics:

- Data standards for outputting spatial data products
- Data source overview
- AIS processing and procedural guidance
- Confidence assessment methodology
- GIS data interpretation and validation.

The following sections provide the detailed methodology describing AIS processing, GIS data interpretation and validation to produce the product deliverables identified in Section 1.1.

1.1 Spatial data product deliverables

The final product deliverables of this contract comprise of the following:

- Draft data guideline
- Data processing template
- XML Metadata
- GIS data layers (Geodatabase)
- Data Confidence Assessment.

2. Data Standards for Outputting Spatial Data Products

The requirement of the data product is to produce a high-resolution data set with a grid resolution of 0.5km by 0.5km. The project deliverables will provide spatial layers depicting commercial shipping movements from AIS data (records) collected by the MCA using its network of AIS receiving stations. The product deliverables would therefore provide data for a specific area or region, clipped from a national dataset, which has been processed to the standards laid out in this report.

To define a yearly AIS dataset, seven complete days (of 24 hours) from the start of each month of April, July and October 2010; plus January 2011 were provided by the MCA. This reduces seasonality effects by providing four representative weeks from four months across the year. This approach is consistent with previous AIS data deliverables and thereby allows comparison between previously processed AIS datasets (for example, data contained within the MaritimeData.co.uk website, created for the Department of Energy and Climate Change (DECC)) and the dataset from this project.

2.1 Spatial data products

The following spatial data products will be produced by the processing procedure outlined in the document:

- Density Data Grid. The generation of a data grid as described fully in Section 6, including a list of grid attributes. The data grid allows interrogation of shipping routes by overall transit density, vessel type, vessel draught and vessel transit class. An example of all vessel traffic as a Density Data Grid output is presented in Figure 8, and an example of vessel draught as a Vector Data output is presented in Figure 9.
- Vessel transits (vector lines) produced as a compiled layer for the whole dataset. An example output of Transit Route Classification is presented in Figure 10. Vessel transits have been classified into the following types:
    1. UK to UK traffic - transits that originated **and** terminated at a UK port or anchorage
    2. UK to non-UK traffic - transits of vessels that have in a week period (seven days of 24 hours), included at least one transit that originated **or** terminated at a UK port or anchorage
    3. Transitory traffic - transits of vessels that have not in a week period (seven days of 24 hours) originated or terminated a transit at a UK port or anchorage.

It should be noted that a requirement of using MCA collected data is vessel identification attributes need to anonymised. The identity of individual vessels within the finalised GIS data layer deliverables must maintain commercial confidential and should follow the code of practice detailed in the Information Commissioners Office publication on anonymisation (ICO 2012).

3. Data Source Overview

AIS data provides a source of geo-referenced information which can be used to spatially depict vessel movements within receiving range of the transmission. The technology works with transponders which automatically broadcast information at regular intervals via a Very High Frequency (VHF) transmitter. AIS signals are broadly classified as ‘Class A’ or ‘Class B’. AIS-A is carried by international voyaging ships with gross tonnage (GT) of 300 or more tonnes, and all passenger ships regardless of size. AIS-B is carried by smaller vessels and is aimed at the leisure and fishing sector. AIS-A has been identified as the primary data input to this project but is referred to generically as ‘AIS’ data throughout this report. Four seven day periods of AIS data for January, April, July and October form the underlying dataset provided by the Maritime and Coastguard Agency (MCA). The following AIS data message types are provided:

- Message Type 1, 2 and 3 – Positional Report
- Message Type 5 – Voyage Message.

In each case the data is provided in the format:

- Time and Date of message reception
- Maritime Mobile Service Identity (MMSI) vessel identification number
- Message type
- AIS message string.

Tables 1 and 2 provide the data fields from each message type.

Table 1 details the components of the Positional Report. It is from the positional data report that the vessel location is identified and can be presented as spatial ‘point’ data. Through interrogation of the MMSI number, individual vessel locations can be identified and a track plot produced.

Table 1: Position report fields (message type 1, 2 and 3).

Parameter	Number of Bits	Description
Message ID	6	Identifier for this message 1, 2 or 3
Repeat indicator	2	Used by the repeater to indicate how many times a message has been repeated
User Identification	30	MMSI number
Navigational status	4	0 = under way using engine, 1 = at anchor, 2 = not under command, 3 = restricted manoeuvrability, 4 = constrained by her draught, 5 = moored, 6 = aground, 7 = engaged in fishing, 8 = under way sailing, 9 = reserved for future amendment of navigational status for ships carrying dangerous goods (DG), harmful substances (HS), or marine pollutants (MP), or International Maritime Organization (IMO) hazard or pollutant Category C, high speed craft (HSC), 10 = reserved for future

Parameter	Number of Bits	Description
		amendment of navigational status for ships carrying DG, HS or MP, or IMO hazard or pollutant category A, wing in grand (WIG); 11-13 = reserved for future use, 14 = AIS-SART (active), 15 = not defined = default (also used by AIS-SART under test)
Rate of turn $(ROT)_{AIS}$	8	0 to +126 = turning right at up to 708 deg per min or higher 0 to -126 = turning left at up to 708 deg per min or higher Values between 0 and 708 deg per min coded by ROT_{AIS} = 4.733 SQRT(ROT_{sensor}) degrees per min where ROT_{sensor} is the Rate of Turn as input by an external Rate of Turn Indicator (TI). ROT_{AIS} is rounded to the nearest integer value. +127 = turning right at more than 5 deg per 30s -127 = turning left at more than 5 deg per 30s -128 (80 hex) indicates no turn information available (default)
Speed over the Ground (SOG)	10	Speed over ground in 1/10 knot steps (0-102.2 knots) 1 023 = not available, 1 022 = 102.2 knots or higher
Position Accuracy	1	The position accuracy (PA) flag 1 = high (> 10 m) 0 = low (< 10 m) 0 = default
Longitude	28	Longitude in 1/10 000 min (+/-180°, East = positive, West = negative). 181= not available = default
Latitude	27	Latitude in 1/10 000 min (+/-90°, North = positive, South = negative). 91° = not available = default
Course over the Ground (COG)	12	Course over ground in 1/10 = (0-3599). 3600 = not available = default. 3 601-4 095 should not be used
True heading	9	Degrees (0-359) (511 indicates not available = default)
Time stamp	6	UTC second when the report was generated by the electronic position system (EPFS) (0-59, or 60 if time stamp is not available, which should also be the default value, or 61 if positioning system is in manual input mode, or 62 if electronic position fixing system operates in estimated (dead reckoning) mode, or 63 if the positioning system is inoperative)
special manoeuvre indicator	2	0 = not available = default 1 = not engaged in special manoeuvre 2 = engaged in special manoeuvre
Spare	3	Not used. Should be set to zero. Reserved for future use
Note: an example of a typical MCA formatted position report is as follows: 2010-04-01 00:00:00, 304678000,1,14RT4L0026033jnLvuVqIoMh08Rq		
Source: United States Coastguard, 2013		

Table 2 provides the components of the Voyage message. Components of the Voyage Message are decoded and stored for association with the appropriate Positional Data reports (Table 1).

Table 2: Ship static and voyage fields (message type 5).

Parameter	Number of Bits	Description
Message ID	6	Identifier for this Message 5; always 5
Repeat indicator	2	Used by the repeater to indicate how many times a message has been repeated.
User ID	30	MMSI number
AIS version indicator	2	0 = station compliant with Recommendation ITU-R M.1371-1 1 = station compliant with Recommendation ITU-R M.1371-3 2-3 = station compliant with future editions
IMO number	30	1-999999999; 0 = not available = default
Call sign	42	7 bit ASCII characters, @@@@@@@ = not available = default
Name	120	Maximum 20 characters 6 bit ASCII
Type of ship and cargo type	8	0 = not available or no ship = default 1-99 = as defined in 'MEDIN definition list' 100-199 = reserved, for regional use 200-255 = reserved, for future use
Overall dimension	30	Reference point for reported position Also indicates the dimension of ship (m)
Type of electronic fixing device	4	0 = undefined (default) 1 = GPS 2 = GLONASS 3 = combined GPS/GLONASS 4 = Loran-C 5 = Chayka 6 = Integrated Navigation System 7 = Surveyed 8 = Galileo 9-15 = not used
Expected Time of Arrival (ETA)	20	Estimated time of arrival; MMDDHHMM UTC Bits 19-16: month; 1-12; 0 = not available = default Bits 15-11: day; 1-31; 0 = not available = default Bits 10-6: hour; 0-23; 24 = not available = default Bits 5-0: minute; 0-59; 60 = not available = default
Maximum present static draught	8	In 1/10 m, 255 = draught 25.5 m or greater, 0 = not available = default; in accordance with IMO Resolution A.851
Destination	120	Maximum 20 characters using 6-bit ASCII; @@@@@@@@@@@@@@@@@@@@ = not available
Data terminal equipment (DTE)	1	Data terminal equipment (DTE) ready (0 = available, 1 = not available = default)
Spare	1	Spare. Not used. Should be set to zero. Reserved for future use
Note: an example of a typical MCA formatted voyage report is as follows: 2010-04-01 00:01:25, 538003268,5,58157A02?T19KM17J21<uE@Q0u9B22222222221IJPV?M4p;NQS3miDm3iU888888888880		
		Source: United States Coastguard, 2013

Ship and Cargo attributes included in the voyage message, require further decoding as shown in Table 3.

Table 3 provides the ship types contained within the AIS Voyage Message. The AIS interface on the vessel is user controlled, and therefore set by an operator onboard the vessel. Table 3 provides the available selection. It should be noted that an operator onboard a vessel can change its transmitted ship type depending on the operation or cargo it is involved with. It is not uncommon for ship type to change either as the vessel/cargo changes or resulting from user input error onboard the vessel.

Table 3: Vessel type.

Digit	Type
1X	Reserved for future use
2X	Wing In Ground (WIG)
3X	**Vessel [Sub-type list shown below]**
4X	High Speed Craft (HSC)
5X	**Special craft [Sub-type list shown below]**
6X	Passenger ships
7X	Cargo ships
8X	Tanker(s)
9X	Other types of ship
3X	**Vessel [Sub-types]**
30	Vessel fishing
31	Vessel towing
32	Vessel towing and length of the tow exceeds 200m or breadth exceeds 25m
33	Vessel engaged in dredging or underwater operations
34	Vessel engaged in diving operations
35	Vessel engaged in military operations
36	Sailing vessel
37	Pleasure craft
5X	**Special craft [Sub-types]**
50	Pilot vessel
51	Search and rescue vessels
52	Tugs
53	Port tenders
54	Vessels with anti-pollution facilities or equipment
55	Law enforcement vessels
56	Spare for assignments to local vessels
57	Spare for assignments to local vessels
58	Medical transports (as defined in the 1949 Geneva Conventions and Additional Protocols)
59	Ships according to RR Resolution No.18 (Mob-83)

As a refinement of Ship Type 5, ‘Special Craft’ descriptors are available. This allows identification of vessels within a narrow descriptive banding, providing useful attributes for AIS classification. These descriptions are shown in Table 3.

3.1 Vessel grouping

To facilitate appropriate display of AIS data by vessel type, AIS data has been reduced into sub-sets of 'type' using the information contained within the received signal. This provides for consistent project outputs and reporting; the vessel groups chosen are as shown in Table 4.

Table 4: Project selected vessel grouping.

Group	Type	Description
0	Unknown	
1	Non-Port service craft	Search and rescue vessels Towing & Towing where length of the tow exceeds 200m or breadth exceeds 25m Medical transports (as defined in the 1949 Geneva Conventions and Additional Protocols) Ships according to Resolution No. 18 (Mob-83) Other special craft (56 & 57 reserved for local use)
2	Port service craft	Pilot vessels Tugs Port tenders and vessels with anti-pollution facilities or equipment
3	Vessels engaged in dredging or underwater operations	Vessels engaged in dredging or underwater operations Vessels engaged in diving operations (see note)
4	High Speed Craft	
5	Military or law enforcement vessels	
6	Passenger vessels	
7	Cargo vessels	
8	Tankers	
9	Excluded vessels	Fishing vessels Pleasure craft Sailing vessels
Note: 'Vessels engaged in diving operations' are predominantly commercial diving operations, for example, related to underwater maintenance.		

The project scope required the display of commercial vessels, and therefore vessel types not considered within this category were removed. Each vessel type removed from the AIS dataset was excluded for the stated reason shown in Table 5.

Table 5: Excluded vessels groups.

Vessel group	Reason for exclusion
Fishing	Fishing vessels are not required to carry AIS beacons as a result AIS data will not provide an accurate representation of fishing activity. All fishing vessels >15m are required to carry Vessel Monitoring System (VMS) beacons and in addition fishing activity is recorded by MMO and IFCA surveillance operations. Fishing activity data is the focus of additional MMO work.
Pleasure craft	Pleasure craft are not required to carry AIS beacons and as a result AIS data will not provide an accurate representation of pleasure craft activity. A review of recreational activity data (MMO1043) is being undertaken concurrent with this this project which is targeting the collection of data on activities including recreational boating.
Sailing	Sailing Vessels are not required to carry AIS beacons and as a result AIS data will not provide an accurate representation of sailing activity. The MMO currently source sailing routes data (modelled) from the RYA. A review of recreational activity data (MMO1043) is being undertaken concurrent with this this project and additional data may be identified to strengthen this dataset.

4. AIS Processing and Procedural Guidance

The following section outlines the methodology used during this project for processing AIS data. It has been broken into sub-sections and provides an overview of the processing techniques and data outputs generated.

4.1 AIS processing

The AIS data processing consists of the following stages:

1. Identification of positional data inaccuracy
2. Identification of vessel transits
3. Quality assurance of vessel transit lines
4. Simplification of transit information
5. Identification of appropriate vessel voyage statistics
6. Identification and reporting of vessel AIS processing statistics.

At the completion of the sixth stage of data processing, the resultant records are held within a SQL database and passed into a GIS environment to produce the required spatial data outputs.

4.2 Identification of positional data inaccuracy

The accuracy of the AIS information is reliant on the accuracy of the vessel's positioning system and the correct encoding, transmission, reception and decoding of that position. Occasional erroneous position reports are observed within the dataset. A small percentage of position reports are not credible within a UK received AIS dataset (for example, they are in the Southern Hemisphere) or lie outside the area of interest. This type of invalid data is removed from the dataset by filtering for all position reports that lie outside of the following range:

- -10 Degrees Longitude to +10 Degrees Longitude
- +47 Degree Latitude to +65 Degree Latitude.

It was initially expected that position reports located inland away from coastal areas would also be erroneous. However, investigation of the data indicated the majority of inland AIS posisition reports which appeared to be errors were in fact associated with inland waterways or in some cases, aircraft broadcasting an AIS signal. Aircraft reports were filtered out using a maximum plausible vessel speed restriction. Terrestrial position reports were filtered out using a coastal boundary mask in GIS.

Filtering the data in this manner typically removes approximately 1.5% of the position reports.

4.3 Identification of vessel transits

By plotting sequential vessel positional data using MMSI ID numbers, vessel transit routes can be identified. This provides vector lines relating to each vessel transit.

The advantage of identifying vessel transits rather than simply identifying the number of position reports or the number of vessels in a particular area is that this provides greater understanding of vessel movement and a history of vessel movements. For example, a vessel moving slowly through an area may produce a high density of position reports for a single transit where a high speed craft may produce the same density of position reports but this may represent many repeated transits back and forth through the area.

The beginning and end of each vessel transit is identified using the same method outlined by Calder and Schwehr (US Hydro 2009). When the reported SOG exceeds 0.5 knots the start of the transit is initiated. The transit normally continues until the reported SOG falls below 0.2 knots for more than five minutes. The transit can also be terminated for a number of other reasons. The full list of termination reasons is as shown below with their appropriate code, to be recorded with each transit record:

0. End of the period of the input dataset
1. Position outside the area of interest
2. Contact lost
3. Intermittent contact (position reports greater than ten minutes apart. Start of an **Inferred transit**)
4. Intermittent contact finished, normal contact restored (End of an **Inferred transit**)
5. Normal transit termination. SOG falls below 0.2 knots for more than five minutes
6. QA checks on consecutive positions result in repeated invalid points (See Section 4.4.5).

Where time separation of positional reports exceeds ten minutes (normally due to intermittent reception of a reduced strength signal) then the transit between points becomes less certain and the true path of the vessel can only be inferred. These **inferred** transits have a more stringent quality assurance check applied for this period of lost reception whereby two values, the vessel’s reported speed (average for this transit segment) and calculated speed and distance between consecutive points, are compared. If the difference between these two values exceeds 10% then the position report is discarded.

Once a transit has been identified a further series of checks are applied and if necessary the transit is discarded. These checks are designed to simplify and tidy the dataset to remove short transits (vessels moving berths, ranging or swinging at anchor) or transit fragments. Transits will be excluded if:

- Minimum transit length is less than 250m
- There are fewer than three position reports per transit
- There are fewer than five position reports per vessel.

The choice of these various parameters will determine the number of position reports and transits which are successfully pre-processed and progress to the GIS processing stage. By varying these parameters, the number of inferred transits will be increased or decreased within the dataset.

4.4 Quality assurance of vessel transit lines

The accuracy of the resultant vessel transit line is affected by any inherent error within the dataset; the following variables provide potential for erroneous transit lines:

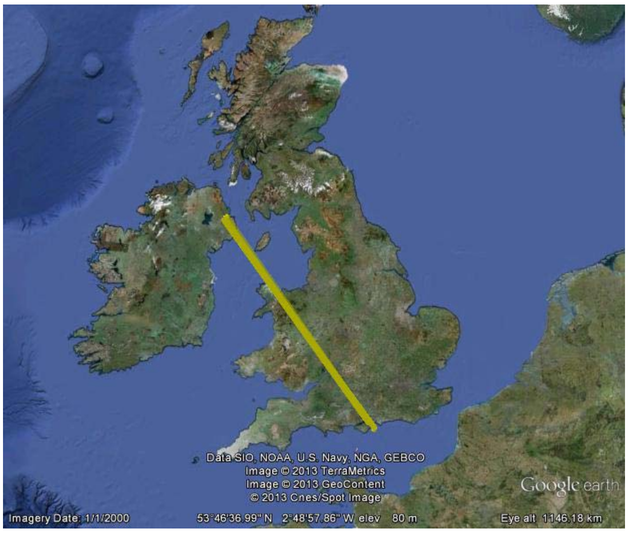
- Invalid and multiple MMSI numbers
- Incorrect signal reception time stamp at the receiving station, resulting in incorrect sequence of transit position points
- Multiple reception
- Loss of transit position (due to landmass obstruction, signal fade or transmission ceases temporarily).

Examples of these data problems are given below.

4.4.1 Invalid and multiple MMSI numbers

Experience of processing large AIS datasets has shown that the data was found to include limited examples of vessels transmitting invalid MMSI vessel ID numbers. These numbers should be unique to each vessel or AIS station. The majority of these issues are removed from the dataset by matching with the vessel type identification (where vessel types are ‘unknown’) and hence filtered from the data. Figure 1 shows an example of two vessels transmitting the same MMSI number within the same temporal period, with the appearance that the vessel transits between these sequential vessel positions.

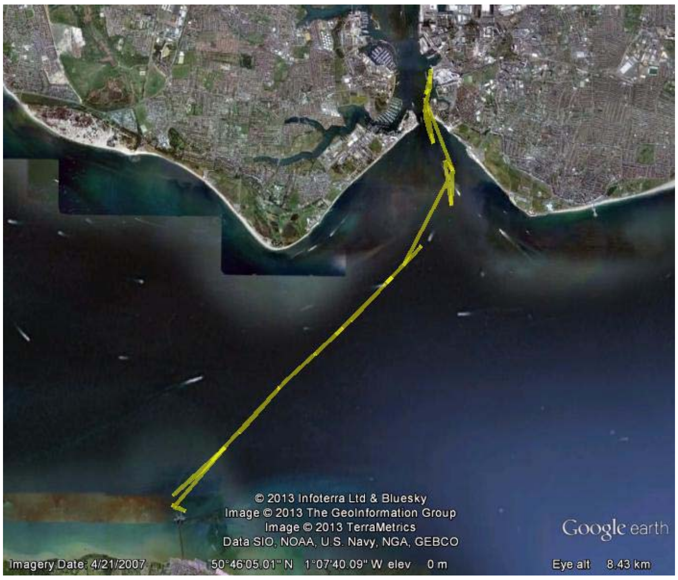
Figure 1: Example of a vessel transit resulting from duplicate MMSI numbers.



4.4.2 Incorrect signal reception time stamp

Figure 2 shows an example of incorrect sequencing of vessel transit position reports and also the duplication of a position report. These problems typically occur in high traffic areas where signal reception and processing causes delays in recording the data and hence an incorrect time stamp is applied.

Figure 2: Example of a vessel transit position sequencing errors.



4.4.3 Multiple reception

Multiple reception can occur as a result of data corruption or multiple reception where two receiving stations date/time stamp are not correctly synchronised. Data corruption can occur where high traffic density results in delayed processing of the vessel data. Figure 3 provides a view of MCA AIS receiving station network location with a 20 and 40 Nautical Mile radius added for reference.

The AIS data used for this analysis was provided by the MCA who have a wide ranging shore based reception and monitoring network. Their network of reception stations are shown in Figure 3. The actual reception range will be dependent on the strength of the vessel transmitter, the height of the transmitting and receiving antenna plus the prevailing weather conditions. It is considered likely that a 40nm reception radius will be achieved by the MCA AIS network subject to the factors previously mentioned.

4.4.4 Loss of transit position

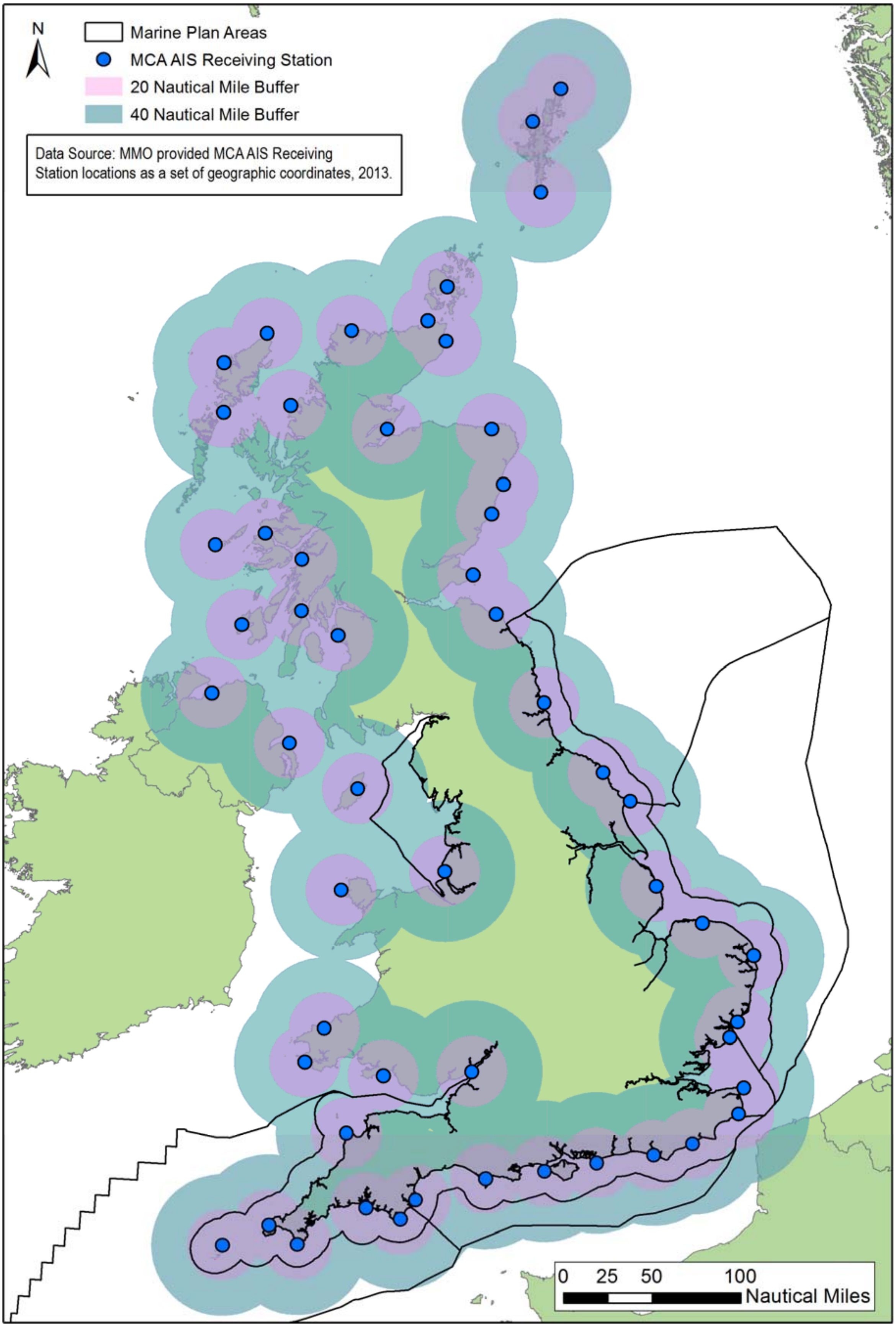
Vessel transit errors can occur due to a loss of signal reception for a range of reasons including: landmass obstruction, signal fade or transmission ceases temporarily or the vessel moves out of range of the receiving station.

4.4.5 AIS data processing quality assurance

The positional accuracy issues identified in Section 4.4.1 to 4.4.4 are addressed by a sequence of data validation and Quality Assurance (QA) steps which are applied to each transit segment. These QA processes are as follows:

- Position reports outside of the area of interest (as defined above) are removed.
- Invalid and Multiple MMSI Numbers are filtered by identification of unknown vessel types and deletion of duplicate position reports.
- Vessel reported speed (average for this transit segment) is compared to the calculated speed and distance between consecutive points of a **normal** transit. If the difference between the two values exceeds 50%, then the position report is discarded.
- Note: this check is separate from that described above for **inferred** transits, where time separation of positional reports exceeds ten minutes. In this situation the test is more stringent and if the difference between these two values exceeds 10% then the position report is discarded.
- Vessel reported COG is compared to the calculated transit bearing between consecutive points. If the acceptable tolerance (±30 Deg) between these two values is exceeded then the position report is discarded.
- Vessel reported SOG and calculated SOG between consecutive points is compared to maximum possible vessel speed. If either of the speed values exceeds the maximum, then the position report is discarded. The maximum speeds used for this purpose is according to the vessel type group as follows:
    - 35 knots – Non-port service craft: dredging or underwater operations (including diving); passenger; cargo; tankers
    - 50 knots - port service craft: high speed craft; military or law enforcement vessels.

Figure 3: MCA AIS message reception network.



4.5 Simplification of transit information

When a valid transit has been identified it initially contains all the valid position reports between the start and end of the transit. In normal circumstances this transit represents a vessel movement from a stationary position (berth or mooring) to a final stationary position.

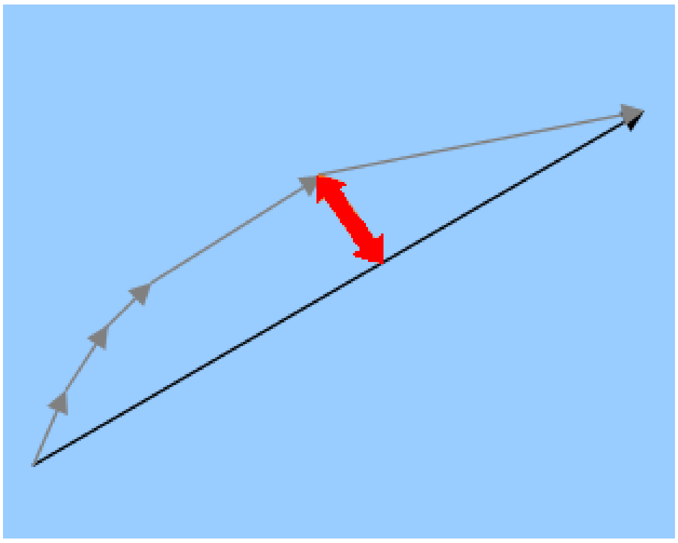
Dependent on the necessary resolution of the output required, this transit can be simplified. For example, if the transit consists of a consistent vessel course (a straight line) between the start and end of the transit then the intermediate position reports are redundant. Removal of these redundant position reports allows for savings in data storage and subsequent processing.

The transit is simplified using the following measures:

- Distance between consecutive position reports exceeds 50m
- Allowable vessel track error tolerance of 50m.

Figure 4 shows a diagram to illustrate track error. Any intermediate position reports can be deleted from the transit until the resulting transit (shown in black) is more than the acceptable tolerance from the reported points defining the original vessel track. This resulting track error is shown in red in Figure 4.

Figure 4: Vessel track error diagram.



4.6 Identification of appropriate vessel voyage statistics

Analysis of a typical data set has shown that vessels change their reported vessel type during the duration of the dataset. Changes of transmitted vessel type may be to reflect changes in operational status (for example where a vessel stops engaging in diving operations) or may simply be incorrect or missing type information transmitted in error. Table 6 shows the number of vessels that changed their vessel type within a typical sample dataset. The records show for the data sample that 97% of vessels report one type and are consistent over the dataset. The remaining 3% change their type, up to four times, during the test sample dataset.

Table 6: Multiple vessel type

Number of reported vessel (group) types per vessel	Number of vessels	Percentage of sample (7047)
1	6851	97.20%
2	184	2.60%
3	8	0.10%
4	4	0.06%

Vessel type identification and vessel dimensions associated with each vessel transit are based upon that type and size which are reported for the longest period of time (most commonly recorded) i.e. if the vessel type and size reported changes over time, then the unique combination of type and size which has the earliest to latest time duration will be used.

Vessel draught is most likely to be voyage specific and time sensitive. For this reason the vessel draught associated with each transit is selected from the time varying data based on the mid transit time. If data is not available for this particular period, for whatever reason, then the smallest draught recorded in the dataset for that vessel is used as the default value. The choice of the smallest draught as the default value is justified for two reasons:

1. Where draught values are provided they are generally accurate and potentially time sensitive. However it is not uncommon for the vessel draught to be transmitted as 0m and therefore the change of draught value indicates that the indicated draught is unreliable. If the vessel's draught is reported as 0m, this has been recorded with the transit meta-data.
2. Where the vessel draught changes over time and is potentially unreliable it was considered preferable to select the smallest draught to lower the significance of the vessel transit. i.e. the vessel is identified and grouped with smaller shallower draught vessels. If the maximum draught was selected as the default, this could subsequently be mis-interpreted as a vessel heavily loaded with cargo.

If this default value is used for the vessel draught this is indicated in the transit attribute table, VoyageDataSource field.

4.7 Reporting of vessel AIS processing statistics

The processing statistics for each vessel are also collated and recorded for data validation and quality checks. These statistics include:

- Total number of position reports
- Total number of valid position reports
- Number of invalid points failing the QA checks
- Total number of points saves
- Number of points deleted due to acceptable track error tolerance

- Number of points deleted due to acceptable point separation tolerance
- Number of points deleted for each of the QA checks
- Total number of transits identified for this vessel
- Total number of transits deleted as being too short
- Total number of transits deleted as they contained too few points.

After processing of the AIS position information, identification of transits and their subsequent simplification, a normal ratio of original data positions to saved data positions, is 30:1. The majority of these deleted points are associated with stationary vessels.

5. Confidence Assessment Methodology

The AIS system was developed primarily to aid safe navigation of vessels. Much of the information contained within the transmission provides quick and easy identification of the vessel, its navigational status and the relative risk posed by its cargo. Voyage data is largely user entered, and therefore has inherent limitations due to human data entering error or misrepresented information.

The confidence assessment of a particular AIS dataset should be undertaken and consider the following principles:

- Methodology Confidence
- Timeliness Confidence
- Spatial Confidence
- Completeness Confidence
- Confidence in Quality Standards.

This type of assessment might reasonably conclude a generally high level of confidence in the dataset. However specific details of any particular vessel or transit will be reliant on the underlying information which can be broadly classified as follows:

- Best quality - automatically generated data (for example: position, COG)
- Variable quality - one time user entered data (for example: ship name, MMSI)
- Lower quality - frequently entered user data (for example: destination, draught).

A measure of accuracy and therefore confidence is inherent within the processed data, where time separation of positional reports exceeds ten minutes the true path of the vessel can only be inferred. Inferred transits are identified in the dataset as having a position density of zero. Transits with a low or zero position density value (position reports per km) have a correspondingly lower confidence level.

6. GIS Data Interpretation and Validation

The spatial data Products have been developed in the form of GIS data layers to provide a means of identifying spatial trends in shipping activity for the study area. A GIS based spatial analysis was conducted using the decoded and processed AIS data provided as a SQL database.

The basis of the analysis relies on the creation of transit lines through the connection of sequential vessel positions. This allows the construction of a density grid which stores the transit density. Each density grid cell is visualised into a colour scale thereby providing a spatial distribution.

6.1 Data processing

The transits were created by plotting each AIS vessel position, and then connecting the positions together to form a line, using a common attribute and sequential field. The common attribute, transit ID, determined which points were grouped together to form a line, while the sequence field, DateTime, determined the order in which the points were connected.

The AIS vessel positions were plotted using latitude and longitude values, provided as geographic coordinates in decimal degrees, using the World Geodetic System (WGS 84). It was decided that all further data would be constructed using the Universal Transverse Mercator 30N (UTM30N) coordinate reference system. UTM30N is particularly well suited for GIS data reliant on accurate distance and measuring capabilities, and is a popular choice when coverage pertains to that of mainland UK. It successfully provided the means necessary to build a precise and accurate density grids.

6.1.1 Construction of transits

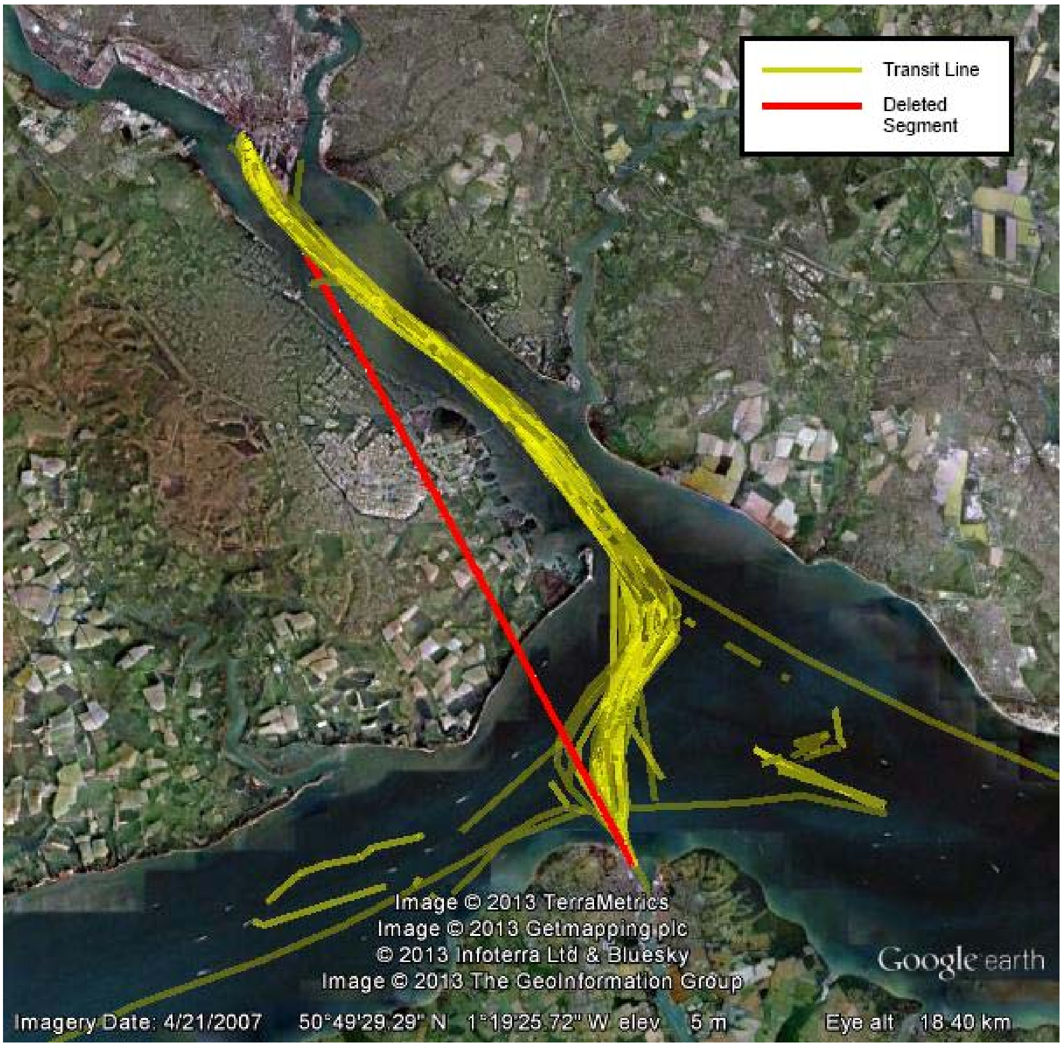
The method used to construct the transits was tested to ensure that transits were created in the correct structure and sequence. All geospatial data was prepared using ArcGIS, but as a measure of quality assurance, a sample of transits around the Isle of Wight was constructed using two other GIS applications. These two applications proved that the algorithm used by ArcGIS is correct, and suitable to construct the transits. The method used to construct the transit lines was the ArcGIS 10.1 points to line tool. This tool took the vessel points and created a polyline feature class using two attributes as constraints. The Transit ID was used to group associated transits together, and the DateTime attribute was used to determine the order in which to connect the points together.

Once the transits were created, a data cleaning exercise was undertaken to ensure high confidence in the data, described below.

6.1.2 Cleaning of transits

In certain circumstances, transits can contain segments that include or cross inland areas. This may be due to incorrect vessel position data or more commonly missing or poorly sequenced data. The indicated transit then deviates from the actual path taken by the vessel. These erroneous transit segments need to be removed from the dataset. An example of a segment that was removed from a transit can be seen in Figure 5.

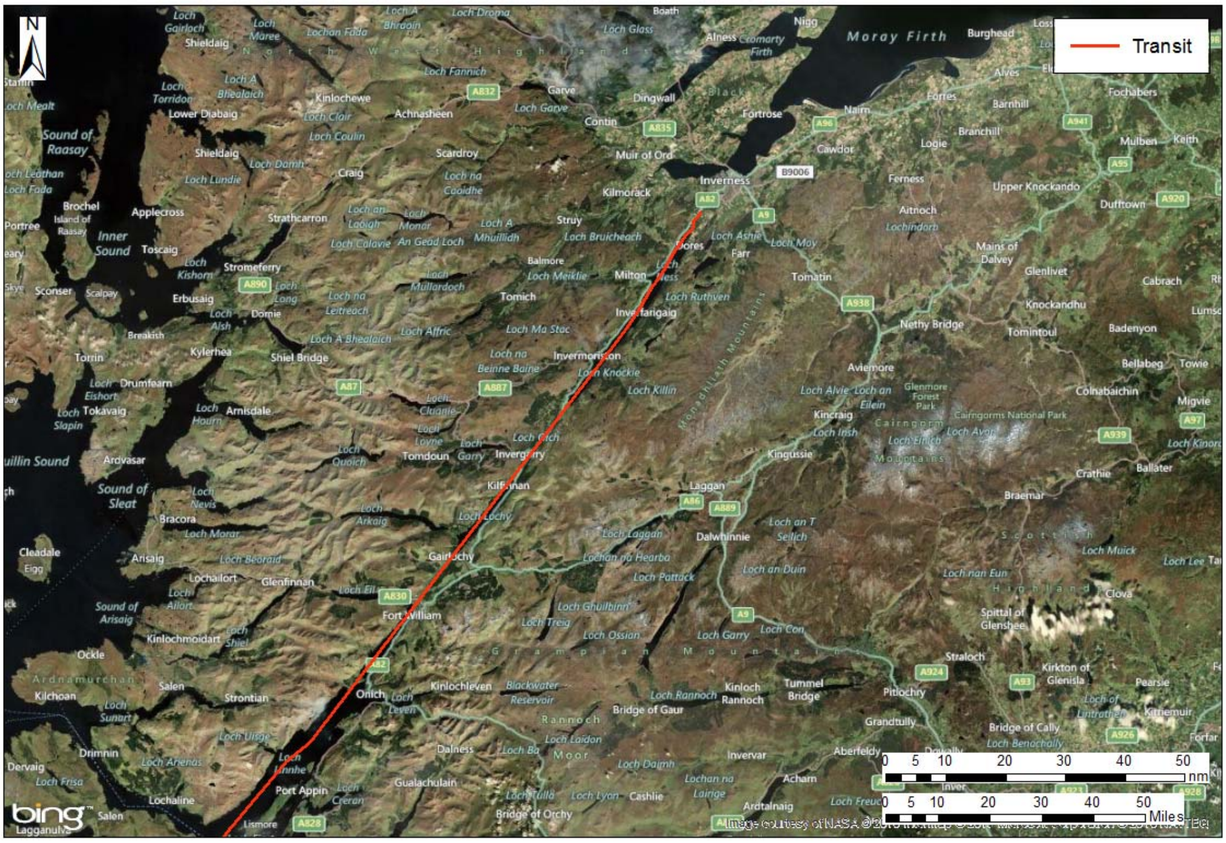
Figure 5: Removed transit segments (in red).



An ‘automated’ approach was used to remove this type of erroneous sections from transits. This is achieved by dividing transits into segments, these were overlaid with a high resolution UK map (polygon). This layer contained a detailed coastline, as well as canals, waterways and channels that run throughout the UK.

All transit segments that intersected with the UK landmass were removed, except those that passed through existing water bodies. From a full UK AIS dataset, the following example (Figure 6) shows a segment from the ‘Caledonian Canal’ with AIS data which has been retained.

Figure 6: Transit segment passing through the Caledonian Canal.



On successful removal of all flawed segments, the remaining segments pertaining to each transit line were merged into a single transit feature.

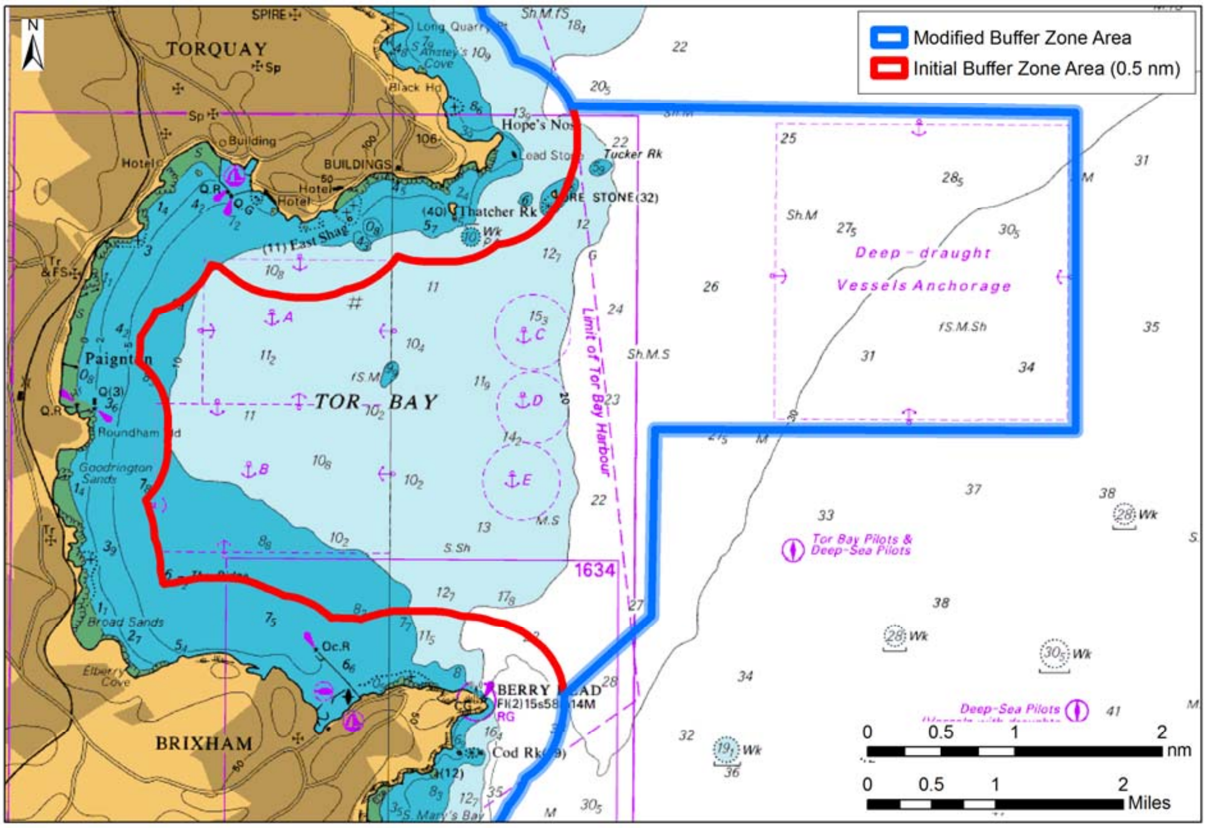
6.1.3 Classification of transits

The transit classification process is designed to identify shipping that are involved in UK trade and those that are only transiting through UK waters. The classification involves the following steps:

- Vessel transits are converted to nodes (one node for start of transit, one node for end of transit).
- Zoned areas were created to define the transit classifications. A UK zone, based on a 0.5 nautical mile buffer of the United Kingdom was created and is used to define an area that includes all UK port destinations and anchorages. This area definition is edited to include anchorage sites that fall outside of the buffer 0.5 nautical mile zone, but are still associated with UK seas. An example of this can be found in Torbay and is illustrated in Figure 7.
- The transit nodes are intersected with the UK port and anchorage zone, resulting in two new attributes per node, starting zone and ending zone. These attributes work together to determine whether the transit commences or terminates within UK waters. This information is used to create a transit history and hence determine and store the transit classification, as described in Section 2.1.

To relate the classification information from the transit nodes to the transit lines, a table join is required. This join is based on Transit ID, and results in the copying of transit classification between the two features.

Figure 7: Modified buffer zone area to include anchorage sites.



6.1.4 Determination of transit density

Transit Density is determined by means of a spatial join. A spatial join is a method for combining information between GIS layers based on their spatial location. This allows for the number of transits intersecting a grid cell to be stored as an attribute. The transit lines do not need to necessarily originate or terminate in the cell, but merely come into contact with it. The method was validated to ensure accuracy of operation before applying it to the dataset.

6.2 GIS data products

The data product specification requires a high-resolution data set with a grid resolution of 0.5km by 0.5km. For the specified data grid, two core data layers are required as follows:

- A density grid provided as a polygon feature class (vector format)
- Vessel transits, provided as a polyline feature class (vector format).

Both of the above datasets contain the attributes necessary for both simple and complex visual data display within the ArcGIS environment. These features are stored within an ESRI File Geodatabase, compatible with version 9.3.1 of ESRI software.

6.2.1 Density grid

The primary purpose of the density grid is to map the density of vessel transits for the period of time represented within the data. This is achieved through a spatial join which intersects the transit lines with the grid, and counts the number of times a transit line passes through each individual cell. On completion of the join, each grid cell can be displayed as a different colour using the transit density attribute as the parameter. When viewed at its full spatial extent, the data can give a clear indication as to which areas experience the most or least amount of traffic.

In addition to its primary use of storing transit density for all vessel transits as a collective, further analysis is undertaken on the transits resulting in a more comprehensive, detailed and ultimately more useful data product.

Additional information is provided by further classifying the density of vessel transits by their vessel type. This provides nine more attributes to the grid, each storing the density of that specific vessel type. For example, a particular grid cell A, might indicate a total vessel transit density of 300, composed of 250 cargo vessels and 50 tankers.

Further transits density attributes are calculated using the vessel draught and the vessel transit classification (VTC).

All attributes pertaining to the grid can be seen in Table 7.

Table 7: Summary of density grid.

Attribute	Description
Cell ID	Unique key used to identify each grid cell.
Total Density (Number of transits/week)	The total number of transits that intersect the cell.
Ship Type 1 Density (Number of transits/week)	The number of transits that intersected the cell where the vessel was of type “Port service craft”.
Ship Type 2 Density (Number of transits/week)	The number of transits that intersected the cell where the vessel was of type “Dredging or underwater operations (including diving)”.
Ship Type 3 Density (Number of transits/week)	The number of transits that intersected the cell where the vessel was of type “HSC”.
Ship Type 4 Density (Number of transits/week)	The number of transits that intersected the cell where the vessel was of type “Military or Law enforcement vessels”.
Ship Type 5 Density (Number of transits/week)	The number of transits that intersected the cell where the vessel was of type “Passenger”.
Ship Type 6 Density (Number of transits/week)	The number of transits that intersected the cell where the vessel was of type “Cargo”.
Ship Type 7 Density (Number of transits/week)	The number of transits that intersected the cell where the vessel was of type “Tankers”.
Small Draught Density	The number of transits that intersected the cell where the

Attribute	Description
(Number of transits/week)	draught of the vessel was between 0 and 7.99 metres.
Medium Draught Density (Number of transits/week)	The number of transits that intersected the cell where the draught of the vessel was between 8 and 11.99 metres.
Large Vessel Density (Number of transits/week)	The number of transits that intersected the cell where the draught of the vessel was between 12 and 25.5 metres.
VTC UK UK Density (Number of transits/week)	The number of transits that intersected the cell where the transit start and finishes with the UK 0.5nm buffer zone
VTC UK Density (Number of transits/week)	The number of transits that intersected the cell where the vessel has at least one transit within the data period that started or finished with the UK 0.5nm buffer zone
VTC Non UK Density (Number of transits/week)	The number of transits that intersected the cell where the vessel has no transits within the data period that started or finished with the UK 0.5nm buffer zone.

The primary output to be derived from the GIS data products is the representation of these transit densities. An example of how a density attribute layer might be presented is shown in Figure 8. Another useful and highly informative output is the identification of transits by vessel draught. An example of a Vessel Draught as Vector Data output is presented in Figure 9. Figure 10 shows Vector Data output of vessel transits illustrating the different vessel transit classification. The transit line figures (Figures 9 and 10) show a proportion of broken transit lines. The AIS processing methodology joins together positional data into a vector line, if any position data is missing or fails QA thresholds for reliability, the information is not presented and hence a broken transit line is evident.

Figures 8, 9 and 10 have been generated from an example AIS test dataset used in the development of the Data Standards report. These figures, whilst based on real AIS data, should not be used for planning purposes and are illustrative only.

6.2.2 Transit lines

Transits contain attributes that store information which can be used to visually determine spatial trends in shipping activities. These trends become apparent when symbolising the transits with different colours using one of the attributes provided. Each transit within the dataset has the following attributes as shown in Table 8. The data will be subject to an anonymisation process to remove reference to specific vessel identification (MMSI and/or IMO numbers, ship name and call sign).

Table 8: Summary of transit dataset.

Attribute	Description
Transit ID	Unique key used to identify transits.
Vessel Type	AIS signal derived ship type.
Vessel Length (metres)	AIS signal derived, entered by operator onboard vessel, normally entered and left as a static value.
Vessel Width (metres)	AIS signal derived, entered by operator onboard vessel, normally entered and left as a static value.
Vessel Draught (metres)	AIS signal derived, entered by operator onboard vessel, could be changed every transit or during transit.
Start Time (date time)	The time the transit started.
End Time (date time)	The time the transit ended.
Vessel Transit Classification (VTC)	1) Transits that originate and terminate at a UK port (UK 0.5nm buffer zone where this includes anchorage within UK waters). 2) Transits where the vessel has at least one transit within the data period that originated or terminated at a UK port (UK 0.5nm buffer zone where this includes anchorage within UK waters), but excludes Transits of VTC 1 3) Transits where the vessel **has not** had any transits within the data period that originated or terminated at a UK port (UK 0.5nm buffer zone where this includes anchorage within UK waters)
Position Density (positions / km)	The number of vessel position reports per km that are represented by the transit. A value of zero indicates an inferred transit
Transit End Reason	This code indicates the reason the transit was terminated. 0. End of the period of the input dataset 1. Position outside the area of interest 2. Contact lost 3. Intermittent contact (position reports greater than ten minutes apart) 4. Intermittent contact finished, normal contact restored 5. Normal transit termination. SOG falls below 0.2 knots for more than 5mins 6. QA checks on consecutive positions result in repeated invalid points
Voyage Data Source	This code indicates if the data for the vessel information (draught, etc.) is derived from data associated with the mid point (in time) of the transit (code 1) or from a period of time before or after the transit duration (code 2).

To provide a layer of vessels classified by draught, the thresholds shown in Table 9 have been used for each vessel type to allow a large/medium/small classification.

Table 9: Vessel draught.

Vessel type	Large (m)	Medium (m)	Small (m)
Tanker	>16	8-16	<8
Cargo Ship (Container/Bulk)	>16	8-16	<8
Passenger (Cruise Ship)	>8	5-8	<5
High Speed Craft (HSC)	>4	2-4	<2
Non-Port service craft	>7	2-7	<2
Military or Law enforcement vessels	>7	4-7	<4
Port Service Craft (Pilot/Tug)	>6	2-6	<2
Vessels engaged in dredging or underwater operations	>7	4-7	<4

7. Conclusion

The focus of this report is to describe the data standards to which AIS data are processed for the creation of mapped AIS derived products. The data standards are applicable to both AIS-A and AIS-B data sources. The report provides a commentary on the identification of positional inaccuracies, vessel transit errors and the quality assurance steps required to ensure vessel track (transit) confidence. To manage data size, the report discusses simplification and thinning of transit information. Finally, the process for assigning voyage information to vessel transit lines are discussed, which includes vessel type, size and transit classification.

The procedures and guidance shown here have been reviewed initially by a select group of key shipping stakeholders as part of this project. Annex 2: 'Draft data guideline for the presentation of AIS Derived Vessel Spatial Data' has been developed to facilitate wider discussion.

8. References

Department of Energy and Climate Change (DECC), 2013. Accessed on 18/07/2013. http://www.maritimedata.co.uk/anonymous/DataOve.aspx

ICO (2012). Anonymisation: Managing data protection risk. Code of practice. Information Commissioner's Office, November 2012.

MEDIN (2013). MEDIN data guideline for the presentation of AIS Derived Vessel Spatial Data. Version 1.0, March 2013.

US Hydro (2009). Calder, B., Schwehr, K., Traffic Analysis for the Calibration of Risk Assessment Methods, US Hydro 2009, May 11-14, Norfolk, VA.

Annex A: Maps

Figure 8: An example of all vessel traffic as a vessel density grid.

Figure 9: An example output of vessel draught (using cargo) as transit lines.

Figure 10: An example output of vessel transit class as transit lines.

Annex B: Draft data guideline

Title	Draft data guideline for the presentation of AIS Derived Vessel Spatial Data
Author(s)	Monty Smedley, Jonathan Baggott, Gordon Osborn
Document Owners	
Reviewed By	
Date Reviewed	
Version	1.0
Date Approved and Published on MEDIN Website	
Summary	This guideline defines the format of Automatic Identification System (AIS) derived vessel spatial data and its representation within mapped products.
Keywords	AIS, Vessel, Shipping

1. Scope and data format for submission to DAC

This guideline covers the format of Automatic Identification System (AIS) derived vessel data and the spatial representation of these datasets within mapping products. This information will be used to better understand spatial and temporal trends in shipping activity, with potential for use in economic, social and environmental studies. Collection and decoding of AIS data are not included in the scope of this guidance.

To submit data to a MEDIN Data Archive Centre (DAC) or transfer to other organisations the raw data is to be provided in a file type format which can be read by a Geographic Information System (GIS).

1.1 Background to Data Guidelines

The Marine Environmental Data and Information Network (MEDIN) is working towards creating a framework of consistent standards covering the major types of data collected and used to describe the marine environment around the UK. The principle benefits of this suite of standards are:

- Allows contracting organisation to easily specify a format that data should be returned in that can be readily used and includes all relevant attributes;
- Provides a consistent format for contractors to work to (rather than a different format for each contract);

- Data can be readily exported to Data Archiving Centres and other users; and
- Instils good practice amongst users.

Each standard defines the data and information that must be stored with a particular data type to ensure it can be readily used and reused. As this type of information is specific for different data types, guidelines are developed for each type. This document describes the format used for AIS derived spatial data and its representation within mapped products.

MEDIN provides standards for data and products. This product guideline describes the necessary information which should be provided when a product has been derived from raw data. The information in the tables below should be included in a MEDIN discovery metadata record.

2. Definition List

A number of terms used in the guideline are specific to shipping and have been drawn from commonly used nautical terms. Whilst these terms can differ between commercial, recreational and naval use, where possible the most commonly accepted variant has been presented. Table 1 provides definitions for specific terms need to understand the context of this draft data guideline, this is not an exhaustive list and will be added to over time as the guideline is revised.

Table 1 Definition List

Term	Meaning
Call sign	A unique designation for a transmitting station normally associated with a broadcasting and radio communications.
Draught	The minimum depth of water a vessel requires to float (without any allowance for under keel clearance) and is a measure from the bottom of the keel to the waterline. This parameter is entered into the AIS user interface on board the vessel by an operator, and can be changed to match the vessel draught. The closely related term 'trim' is defined as the difference between the forward and aft draughts.
IMO Number	The International Maritime Organization (IMO) number is a unique identifier for ships and for registered ship management companies. The IMO vessel identification number scheme was introduced in 1987; the number remains unchanged upon transfer of the ship to other flag(s) and is inserted in the ship's certificates.
MMSI Number	A Maritime Mobile Service Identity (MMSI) is a series of nine digits which are sent in digital form over a radio frequency channel to uniquely identify ship stations, ship earth stations, coast stations, coast earth stations and group calls.
Ship Type	AIS ship type definition is based on the combination of a two digit type code (XX) and this includes: **Type** 1X - Reserved for future use 2X - Wing In Ground (WIG) 3X - Vessel [Sub-type list shown below]

Term	Meaning
	4X - High Speed Craft (HSC) 5X - Special Craft [list shown below]] 6X - Passenger ships 7X - Cargo ships 8X - Tanker(s) 9X - Other types of ship **3X - Vessel [Sub-type]** 30 - Vessel Fishing 31 - Vessel Towing 32 - Vessel Towing and length of the tow exceeds 200m or breadth exceeds 25m 33 - Vessel Engaged in dredging or underwater operations 34 - Vessel Engaged in diving operations 35 - Vessel Engaged in military operations 36 - Sailing Vessel 37 - Pleasure craft Vessel **5X - Special craft [Sub-type]** 50 - Pilot vessel 51 - Search and rescue vessels 52 - Tugs 53 - Port tenders 54 - Vessels with anti-pollution facilities or equipment 55 - Law enforcement vessels 56 - Spare for assignments to local vessels 57 - Spare for assignments to local vessels 58 - Medical transports (as defined in the 1949 Geneva Conventions and Additional Protocols) 59 - Ships according to RR Resolution No.18 (Mob-83)
Transit	Transit starts when the Speed Over the Group (SOG) is greater than 0.5 knots and continues until the vessel SOG falls below 0.2 knots for more than 5mins (or reception of the AIS signal has been lost or compromised).
Underway	Underway means that a vessel is not at anchor, or made fast to the shore, or aground.
Vessel	Vessel includes every description of water craft, including non-displacement craft, wing-in-ground-effect (WIG) vehicle, and seaplanes, used or capable of being used as a means of transportation on water.
Wing-In-Ground (WIG)	Wing-In-Ground (WIG) craft means a multimodal craft which, in its main operational mode, flies in close proximity to the surface by utilizing surface-effect action.

2.1 Using this Product Guideline

This guideline is split into sections which refer to the process states of AIS data transfer to describe mapped products. These are identified as:

- **AIS Data** – a summary of AIS data format and types;
- **Spatial Accuracy** – provides a qualitative description of expected spatial accuracy of AIS positional data;
- **AIS Processing** – a summary of AIS processing steps to arrive at derived data for spatial representation; and

- **AIS Derived Data** – provides the details table layout of AIS derived products within a database structure for use in a GIS environment.

The AIS Data, Spatial Accuracy and AIS Processing are common to all derived AIS products. The processing steps can be varied to achieved the required level of confidence in the accuracy of the AIS data, ultimately, the only way to verify spatial records is through corroboration with a secondary data source (typically, a radar survey to positively identify targets/AIS records). Comparisons can be carried out between AIS datasets and third party databases to compare record fields, and provide missing fields with additional values.

3. AIS Data

AIS is a maritime navigation safety communications system standardized by the International Telecommunication Union (ITU) and adopted by the International Maritime Organization (IMO) to provide vessel information.

AIS data provides an accessible source of spatial information which can be used to depict vessel movements within receiving range of the transmission. The technology works with transponders which automatically broadcast information at regular intervals via a Very High Frequency (VHF) transmitter. Vessel transmitted AIS signals are broadly classified as 'Class A' or 'Class B'. AIS-A is carried by international voyaging ships with gross tonnage (GT) of 300 or more tonnes, and all passenger ships regardless of size. AIS-B is carried by smaller vessels and is aimed at the leisure and fishing sector. This MEDIN standard does not deal with the decoding of AIS data, the focus is the display of derived products from processed data. The following section provides information on the message content, relevant to post-processing and display.

There are currently 27 AIS message types; these are identified in Table 2.

Table 2 AIS Message Types

Message ID	Name	Description
1	Position report	Scheduled position report; (Class A shipborne mobile equipment)
2	Position report	Assigned scheduled position report; (Class A shipborne mobile equipment)
3	Position report	Special position report, response to interrogation; (Class A shipborne mobile equipment)
4	Base station report	Position, UTC, date and current slot number of base station
5	Static and voyage related data	Scheduled static and voyage related vessel data report; (Class A shipborne mobile equipment)
6	Binary addressed message	Binary data for addressed communication

Message ID	Name	Description
7	Binary acknowledgement	Acknowledgement of received addressed binary data
8	Binary broadcast message	Binary data for broadcast communication
9	Standard SAR aircraft position report	Position report for airborne stations involved in SAR operations, only
10	UTC/date inquiry	Request UTC and date
11	UTC/date response	Current UTC and date if available
12	Addressed safety related message	Safety related data for addressed communication
13	Safety related acknowledgement	Acknowledgement of received addressed safety related message
14	Safety related broadcast message	Safety related data for broadcast communication
15	Interrogation	Request for a specific message type (can result in multiple responses from one or several stations)(4)
16	Assignment mode command	Assignment of a specific report behaviour by competent authority using a Base station
17	DGNSS broadcast binary message	DGNSS corrections provided by a base station
18	Standard Class B equipment position report	Standard position report for Class B shipborne mobile equipment to be used instead of Messages 1, 2, 3(8)
19	Extended Class B equipment position report	Extended position report for class B shipborne mobile equipment; contains additional static information
20	Data link management message	Reserve slots for Base station(s)
21	Aids-to-navigation report	Position and status report for aids-to-navigation
22	Channel management(6)	Management of channels and transceiver modes by a Base station
23	Group assignment command	Assignment of a specific report behaviour by competent authority using a Base station to a specific group of mobiles
24	Static data report	Additional data assigned to an MMSI. Part A: Name. Part B: Static Data
25	Single slot binary message	Short unscheduled binary data transmission (Broadcast or addressed)
26	Multiple slot binary message with Communications State	Scheduled binary data transmission (Broadcast or addressed)
27	Position report for long range applications	Scheduled position report (Class A shipborne mobile equipment outside base station coverage)
28-63	Undefined; Reserved for future use	N/A
Source: United States Coastguard, 2013		

Class A Messages are comprised of the following message types:

- Message Type 1, 2 and 3 – Positional Report
- Message Type 5 – Voyage Message

Class B Messages are comprised of the following message types:

- Message Type 14 – Safety Related Message
- Message Type 18 – Standard Class B CS Position Report
- Message Type 19 – Extended Class B Equipment Position Report
- Message Type 24 – Class B CS Static Data Report

The AIS coded message string needs to be associated with the time and date of message reception.

Tables 3 and 4 provide the decoded data fields from each message type. It is through the position reports that the spatial aspect of AIS is established.

Table 3 Position Report Fields (Message Type 1, 2 and 3)

Parameter	Number of Bits	Description
Message ID	6	Identifier for this message 1, 2 or 3
Repeat indicator	2	Used by the repeater to indicate how many times a message has been repeated.
User Identification	30	MMSI number
Navigational status	4	0 = under way using engine, 1 = at anchor, 2 = not under command, 3 = restricted manoeuvrability, 4 = constrained by her draught, 5 = moored, 6 = aground, 7 = engaged in fishing, 8 = under way sailing, 9 = reserved for future amendment of navigational status for ships carrying dangerous goods (DG), harmful substances (HS), or marine pollutants (MP), or International Maritime Organization (IMO) hazard or pollutant category C, high speed craft (HSC), 10 = reserved for future amendment of navigational status for ships carrying DG, HS or MP, or IMO hazard or pollutant category A, wing in grand (WIG); 11-13 = reserved for future use, 14 = AIS-SART (active), 15 = not defined = default (also used by AIS-SART under test)
Rate of turn $(ROT)_{AIS}$	8	0 to +126 = turning right at up to 708 deg per min or higher 0 to -126 = turning left at up to 708 deg per min or higher Values between 0 and 708 deg per min coded by ROT_{AIS} = 4.733 SQRT(ROT_{sensor}) degrees per min where ROT_{sensor} is the Rate of Turn as input by an external Rate of Turn Indicator (TI). ROT_{AIS} is rounded to the nearest integer value. +127 = turning right at more than 5 deg per 30s -127 = turning left at more than 5 deg per 30s -128 (80 hex) indicates no turn information available (default).
Speed over the Ground (SOG)	10	Speed over ground in 1/10 knot steps (0-102.2 knots) 1 023 = not available, 1 022 = 102.2 knots or higher
Position Accuracy	1	The position accuracy (PA) flag 1 = high (> 10 m)

Parameter	Number of Bits	Description
		0 = low (< 10 m) 0 = default
Longitude	28	Longitude in 1/10 000 min (+/-180°, East = positive, West = negative). 181= not available = default
Latitude	27	Latitude in 1/10 000 min (+/-90°, North = positive, South = negative). 91° = not available = default
COG	12	Course over ground in 1/10 = (0-3599). 3600 = not available = default. 3 601-4 095 should not be used.
True heading	9	Degrees (0-359) (511 indicates not available = default)
Time stamp	6	UTC second when the report was generated by the electronic position system (EPFS) (0-59, or 60 if time stamp is not available, which should also be the default value, or 61 if positioning system is in manual input mode, or 62 if electronic position fixing system operates in estimated (dead reckoning) mode, or 63 if the positioning system is inoperative)
special manoeuvre indicator	2	0 = not available = default 1 = not engaged in special manoeuvre 2 = engaged in special manoeuvre
Spare	3	Not used. Should be set to zero. Reserved for future use
Source: United States Coastguard, 2013		

Table 4 Ship Static and Voyage Fields (Message Type 5)

Parameter	Number of Bits	Description
Message ID	6	Identifier for this Message 5; always 5
Repeat indicator	2	Used by the repeater to indicate how many times a message has been repeated.
User ID	30	MMSI number
AIS version indicator	2	0 = station compliant with Recommendation ITU-R M.1371-1 1 = station compliant with Recommendation ITU-R M.1371-3 2-3 = station compliant with future editions
IMO number	30	1-999999999; 0 = not available = default
Call sign	42	7 bit ASCII characters, @@@@@@@ = not available = default
Name	120	Maximum 20 characters 6 bit ASCII
Type of ship and cargo type	8	0 = not available or no ship = default 1-99 = as defined in 'MEDIN definition list' 100-199 = reserved, for regional use 200-255 = reserved, for future use
Overall dimension	30	Reference point for reported position. Also indicates the dimension of ship (m)
Type of electronic fixing device	4	0 = undefined (default) 1 = GPS 2 = GLONASS 3 = combined GPS/GLONASS 4 = Loran-C 5 = Chayka 6 = Integrated Navigation System 7 = Surveyed 8 = Galileo, 9-15 = not used
Expected Time of Arrival (ETA)	20	Estimated time of arrival; MMDDHHMM UTC Bits 19-16: month; 1-12; 0 = not available = default Bits 15-11: day; 1-31; 0 = not available = default Bits 10-6: hour; 0-23; 24 = not available = default Bits 5-0: minute; 0-59; 60 = not available = default
Maximum present static draught	8	In 1/10 m, 255 = draught 25.5 m or greater, 0 = not available = default; in accordance with IMO Resolution A.851
Destination	120	Maximum 20 characters using 6-bit ASCII; @@@@@@@@@@@@@@@@@@@@ = not available
Data terminal equipment (DTE)	1	Data terminal equipment (DTE) ready (0 = available, 1 = not available = default)
Spare	1	Spare. Not used. Should be set to zero. Reserved for future use.
Source: United States Coastguard, 2013		

AIS-B units autonomously transmit Messages 18 and 24, these message types are defined in Table 5 for position fields, and Table 6 for static data fields (which is split into Part 1 and Part B). AIS-B positional and voyage data is used to provide AIS-B derived data products which can then be spatially represented.

Table 5 Standard Class B Equipment Position Fields (Message Type 18)

Parameter	Number of Bits	Description
Message ID	6	Identifier for Message 18; always 18
Repeat indicator	2	Shall be 0 for "CS" transmissions
User ID	30	MMSI number
Spare	8	Not used. Should be set to zero. Reserved for future use
SOG	10	Speed over ground in 1/10 knot steps (0-102.2 knots) 1 023 = not available, 1 022 = 102.2 knots or higher
Position accuracy	1	1 = high (> 10 m) 0 = low (< 10 m) 0 = default
Longitude	28	Longitude in 1/10 000 min (180, East = positive, West = negative; 181= not available = default
Latitude	27	Latitude in 1/10 000 min (90, North = positive, South = negative; 91= not available = default
COG	12	Course over ground in 1/10= (0-3 599). 3 600 = not available = default; 3 601-4 095 should not be used
True heading	9	Degrees (0-359) (511 indicates not available = default)
Time stamp	6	UTC second when the report was generated by the EPFS (0-59 or 60 if time stamp is not available, which should also be the default value or 61 if positioning system is in manual input mode or 62 if electronic position fixing system operates in estimated (dead reckoning) mode or 63 if the positioning system is inoperative) 61, 62, 63 are not used by "CS" AIS
Spare	2	Not used. Should be set to zero. Reserved for future use
Class B unit flag	1	0 = Class B SOTDMA unit 1 = Class B "CS" unit
Class B display flag	1	0 = No display available; not capable of displaying Message 12 and 14 1 = Equipped with integrated display displaying Message 12 and 14
Class B DSC flag	1	0 = Not equipped with DSC function 1 = Equipped with DSC function (dedicated or time-shared)
Class B band flag	1	0 = Capable of operating over the upper 525 kHz band of the marine band 1 = Capable of operating over the whole marine band (irrelevant if "Class B Message 22 flag" is 0)
Class B Message 22 flag	1	0 = No frequency management via Message 22 , operating on AIS1, AIS2 only 1 = Frequency management via Message 22
Mode flag	1	0 = Station operating in autonomous and continuous mode = default 1 = Station operating in assigned mode
Receiver autonomous integrity monitoring (RAIM-flag)	1	RAIM (Receiver autonomous integrity monitoring) flag of electronic position fixing device; 0 = RAIM not in use = default; 1 = RAIM in use
Communication state selector flag	1	0 = SOTDMA communication state follows 1 = ITDMA communication state follows (always "1" for Class-B "CS")

Parameter	Number of Bits	Description
Communication state	19	SOTDMA communication state. Because Class B "CS" does not use any Communication State information, this field shall be filled with the following value: 1100000000000000110.
Source: United States Coastguard, 2013		

Table 6 Static Data Fields (Message Type 24)

Parameter	Number of Bits	Description
Part A		
Message ID	6	Identifier for Message 24; always 24
Repeat indicator	2	Used by the repeater to indicate how many times a message has been repeated. 0 = default; 3 = do not repeat any more
User ID	30	MMSI number
Part number	2	Identifier for the message part number; always 0 for Part A
Name	120	Name of the MMSI-registered vessel. Maximum 20 characters 6-bit ASCII, @@@@@@@@@@@@@@@@@@@@ = not available = default
Part B		
Message ID	6	Identifier for Message 24; always 24
Repeat indicator	2	Used by the repeater to indicate how many times a message has been repeated. 0 = default; 3 = do not repeat any more
User ID	30	MMSI number
Part number	2	Identifier for the message part number; always 1 for Part B
Type of ship and cargo type	8	0 = not available or no ship = default 1-99 = as defined in § 3.3.2 100-199 = reserved, for regional use 200-255 = reserved, for future use
Vendor ID	42	Unique identification of the Unit by a number as defined by the manufacturer (option; "@@@@@@@" = not available = default)
Call sign	42	Call sign of the MMSI-registered vessel. 7 X 6 bit ASCII characters, "@@@@@@@" = not available = default
Dimension of ship/reference for position. Or, for unregistered daughter vessels, use the MMSI of the mother ship.	30	Dimensions of ship in meters and reference point for reported position or for an unregistered daughter vessel, use the MMSI of the associated mother ship in this data field
Message ID	6	Identifier for Message 24; always 24
Source: United States Coastguard, 2013		

3.1 Spatial Accuracy

The underlying positional accuracy of the dataset is ultimately dependant on the accuracy of navigation equipment on each vessel. AIS-A will have a higher accuracy confidence level due to its installation, connection to ship board positioning technology and testing regime. The International Convention for the Safety of Life at Sea (SOLAS) Chapter V Regulation (19.2) requires ships to carry a receiver for a global navigation satellite system or a terrestrial radio-navigation system, or other means, suitable for use at all times throughout the intended voyage to establish and update the ship's position by automatic means. AIS transponders are connected to the vessel's positioning system, providing automated position fixing. SOLAS Chapter V Regulation (18.9) requires an annual test of shipborne AIS-A systems, which provides an annual check on spatial confidence associated with AIS transmission.

AIS-B spatial accuracy is dependant on the vessels position system accuracy. Irrespective of AIS-A or AIS-B data reception, quality assurance steps should be built into any AIS data interpretation routine to ensure that derived data is checked for errors of positional accuracy, so far as this is possible.

3.2 AIS Processing

AIS data, once decoded is processed to provide positional data records. Individual positional records can be further processed to provide vessel transit tracks. The following provides a summary of AIS decoding and processing steps prior to presenting the derived AIS data.

1. Decoding of AIS signal (positional reports / voyage information);
2. Joining ship static and voyage information with positional reports;
3. Identification of vessel transits;
4. Quality assurance of vessel transit lines;
5. Identification and reporting of vessel AIS processing statistics.

At the completion of the last stage of data processing, the resultant records (depending on record count) are stored within a database and passed into a GIS environment as AIS derived data to produce the required Spatial Data outputs.

3.3 AIS Derived Data

AIS data can be provided in the following formats:

Vessel Positions (point data);
Transit Lines (polyline data); and

Density Grid (polygon data).

The tables below outline the data fields, a description and where available a term list and/or format given at the end of each field which should be used to store the data. Each field is either mandatory, conditional or optional as indicated by M, C, or O respectively. Conditional means that the field must be completed if a value is known. [MEDIN note: the following text may need amending/deleting - at the time of writing a spreadsheet has not been created] In the absence of an existing spreadsheet or database to hold the below information, it is recommended that the template available to download from the MEDIN website is used. Instructions are provided in the template.

3.4 Vessel Positions

Vessel Positions are used to construct Transit Lines. They contain a uniquely identifiable attribute named Transit ID, a Date Time attribute that contains the date and time the position was recorded, and Latitude and Longitude coordinates provided in the World Geodetic System 1984 (WGS84) geographic coordinate system. At least two Vessel Positions are required to successfully construct a single Transit Line..

Attribute	M, C, O	Description	Recommended Term List or Format
Transit identifier	M	A non-unique identifier which specifies which transit, multiple vessel positions relate to.	Number; (e.g. 1)
Date and time	M	The date and time of derivation.	Date; dd-mm-yyyy; (e.g. 01-10-2013 13:33:00)
Latitude of vessel position (WGS84 decimal degrees)	M	The latitude of the vessel position at the time of derivation. Units are positive north. Six decimal places are recommended.	Decimal degrees; minimum of four and maximum of seven decimal places. (50.798252)
Longitude of vessel position (WGS84 decimal degrees)	M	The latitude of the vessel position at the time of derivation. Units are positive east. Six decimal places are recommended.	Decimal degrees; minimum of four and maximum of seven decimal places. (0.937410)

3.5 Transit Lines

Transit Lines visually represent the routes taken by vessels. A single vessel can have multiple transits. Transit Lines are used to determine the values stored in the Density Grid.

Attribute	M, C, O	Description	Recommended Term List or Format
Transit identifier	M	A unique identifier for the transit.	Number; (e.g. 1)
Vessel Type	C	AIS signal derived ship type.	Number; (e.g. 8) – “Tanker”
Vessel Length (metres)	C	AIS signal derived. Derived from two values entered by operator onboard vessel, normally entered and left as a static value.	Number; (e.g. 177) Units = metres
Vessel Width (metres)	C	AIS signal derived. Derived from two values	Number; (e.g. 31)

Attribute	M, C, O	Description	Recommended Term List or Format
		entered by operator onboard vessel, normally entered and left as a static value.	Units = metres
Vessel Draught (metres)	C	AIS signal derived, entered by operator onboard vessel, could be changed every transit or during transit.	Number; (e.g. 7.4) Units = metres
Start Time	M	The time the transit started.	Date; dd-mm-yyyy; (e.g. 01-10-2013 10:33:00)
End Time	M	The time the transit ended.	Date; dd-mm-yyyy; (e.g. 01-10-2013 12:31:00)
Vessel Transit Classification (VTC)	O	Classification that identifies the complete route or combination of routes taken by a vessel in transit(s), from beginning to end. The classification is determined by intersecting the transits with a 0.5 nautical mile buffer of a high resolution UK Coastline layer. The coastline is measured at MHWS, and should be edited to include all anchorage sites.	Attribute domains should be used for enforcing data integrity. Domain type - Coded domain (specifies a valid set of options for the VTC). 1 - UK to UK traffic - transits that originated **and** terminated at a UK port or anchorage; 2 - UK to non-UK traffic - transits of vessels that have with in a week period (7 days of 24 hours), included at least one transit that originated **or** terminated at a UK port or anchorage; and 3 - Transitory traffic - transits of vessels that have not in a week period (7 days of 24 hours) originated or terminated a transit at a UK port or anchorage.
Position Density (positions / km)	M	The number of vessel position reports per km that are represented by the transit. A value of zero indicates an inferred transit (where the time interval between position reports exceeds 10 minutes).	Number; (e.g. 7)
Transit End Reason	C	Indicates the reason the transit was terminated.	Attribute domains should be used for enforcing data integrity. Domain type - Coded domain (specifies a valid set of options for the Transit End Reason). 0 - End of the period of the input dataset; 1 - Position outside the area of interest;

Attribute	M, C, O	Description	Recommended Term List or Format
			2 - Contact lost; 3 - Intermittent contact (position reports greater than 10 minutes apart); 4 - Intermittent contact finished, normal contact restored; 5 - Normal transit termination. SOG falls below 0.2 knots for more than 5mins; 6 - QA checks on consecutive positions result in repeated invalid points.
Voyage Data Source	M	This code indicates if the data for the vessel information (draught, etc) is derived from data associated with the mid point (in time) of the transit (code 1) or from a period of time before or after the transit duration (code 2).	Attribute domains should be used for enforcing data integrity. Domain type - Coded domain (specifies a valid set of options for the Voyage Data Source). 1 - Derived from mid point of transit; 2 - Derived before or after the duration of the transit.

3.6 AIS Density Grid

The primary purpose of the density grid is to map the density of vessel transits for the period of time represented within the data. This is achieved through a spatial join which intersects the transit lines with the grid, and counts the number of times a transit line passes through each individual cell. On completion of the join, each grid cell can be displayed as a different colour using the transit density attribute as the parameter. In addition to the density grid's primary use of storing transit density, further analysis can be undertaken on the transits resulting in a more comprehensive, detailed and more useful data product. The density grid can also be displayed using other attributes, such as type, dimension or Vessel Transit Classification to provide a more explicit density grid view.

Attribute	M, C, O	Description	Recommended Term List or Format
Cell ID	M	Unique key used to identify each grid cell.	Number; (e.g. 1)
Density	M	A value that holds the number of transits that intersect the grid cell. The depth of analysis undertaken and resultant attributes created depend on the requirements of the project. The attributes contained in the Transit Line allow for the following density variations: ▪ Ship Type Density: Number of transits that intersect the grid cell, over a specified period of time; for example where the vessel type is a 'Cargo'. ▪ Dimension (draught, length, beam) Density: Number of transits that intersect the grid cell, over a specified period of time, for example, where the vessel draught is small. ▪ VTC Density: Number of transits that intersect the grid cell, over a specified period of time; for example where the transit had a Vessel Transit Classification of 1.	Number; (e.g. 200)

References

United States Coastguard, 2013. 'AIS Messages'. Accessed: 11 March 2013. http://www.navcen.uscg.gov/?pageName=AISMessages. Last Updated 17 March 2011