

Carsten Bullemer, C-Sigma EMSA 2017 Lisbon



We have the best (ais) data in town ! how can we benchmark AIS data ?

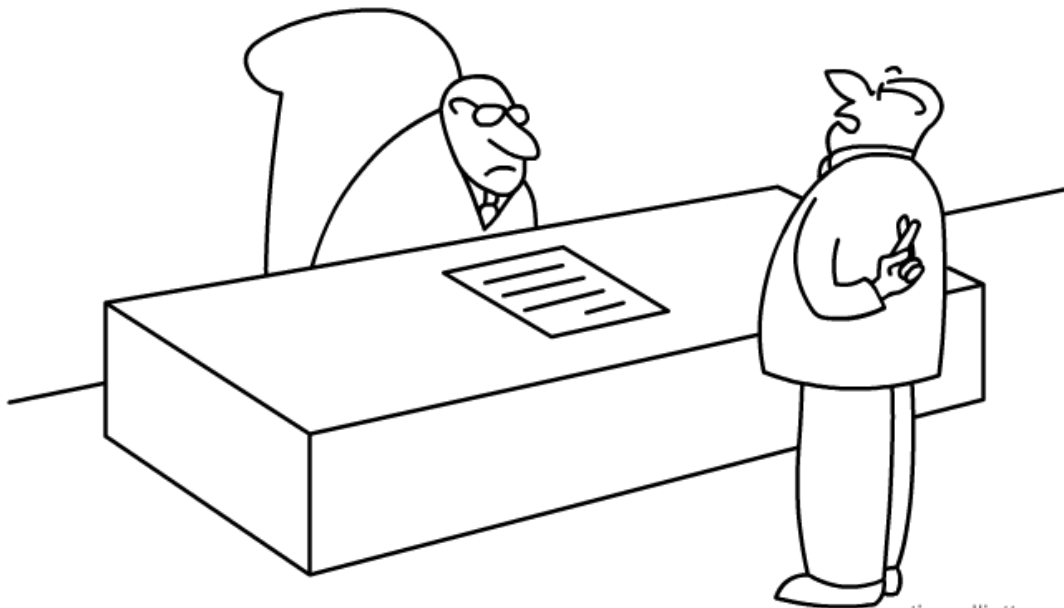
We have the best (ais) data in town ! how can we benchmark AIS ?



timoelliott.com

“I think you’ll find that mine is bigger...”

We have the best (ais) data in town ! how can we benchmark AIS ?



timoelliott.com

"Yes sir, you can absolutely trust those numbers"



Data quality refers to the condition of a set of values of qualitative or quantitative variables. There are many definitions of data quality but data is generally considered high quality if it is "fit for [its] intended uses in operations, decision making and planning".^[1]

Alternatively, **data is deemed of high quality if it correctly represents the real-world construct to which it refers.**

People's views on data quality can often be in disagreement, even when discussing the same set of data used for the same purpose. Data cleansing may be required in order to ensure data quality.^[2]



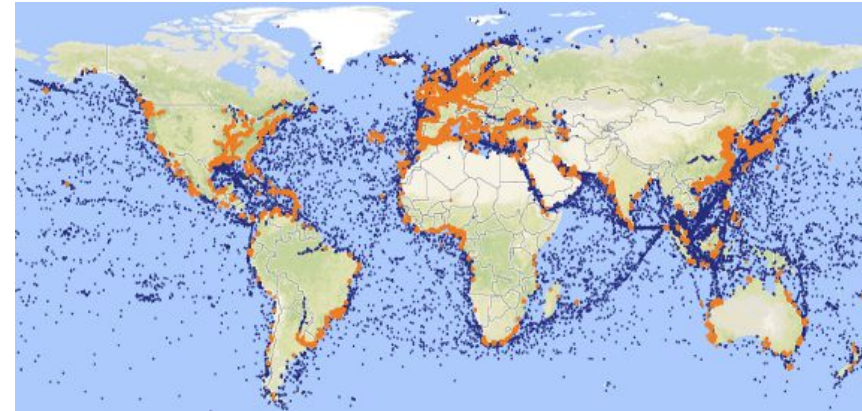
We have the best (ais) data in town ! how can we benchmark AIS ?

Use Real World Constructs for validating AIS Data



Fleet

Geography



We have the best (ais) data in town ! how can we benchmark AIS ?



use World Fleet for ais benchmarking



KPI : How long does it take to cover 50%, 75%, 85%, 90%, 95% of the fleet segments

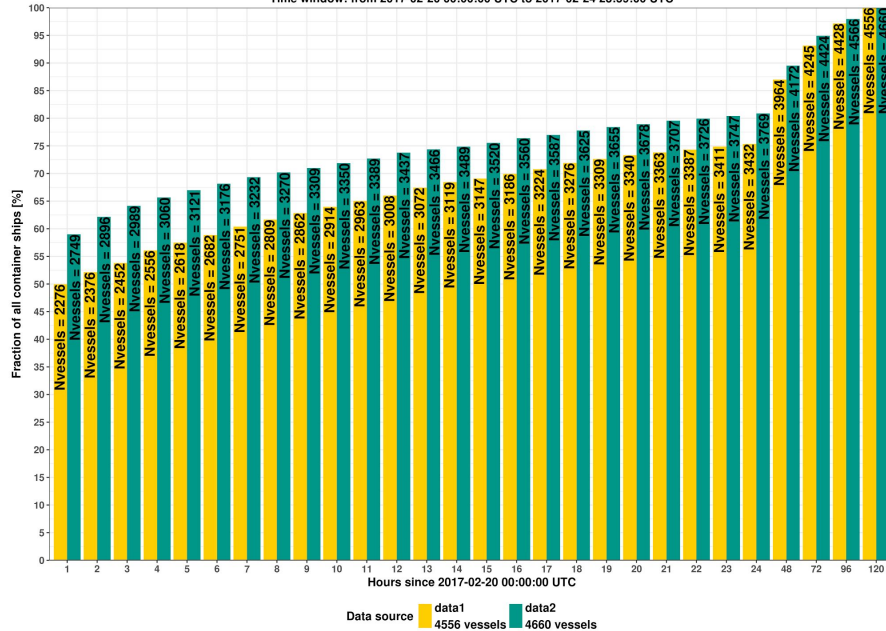
We have the best (ais) data in town ! how can we benchmark AIS ?



Container Fleet

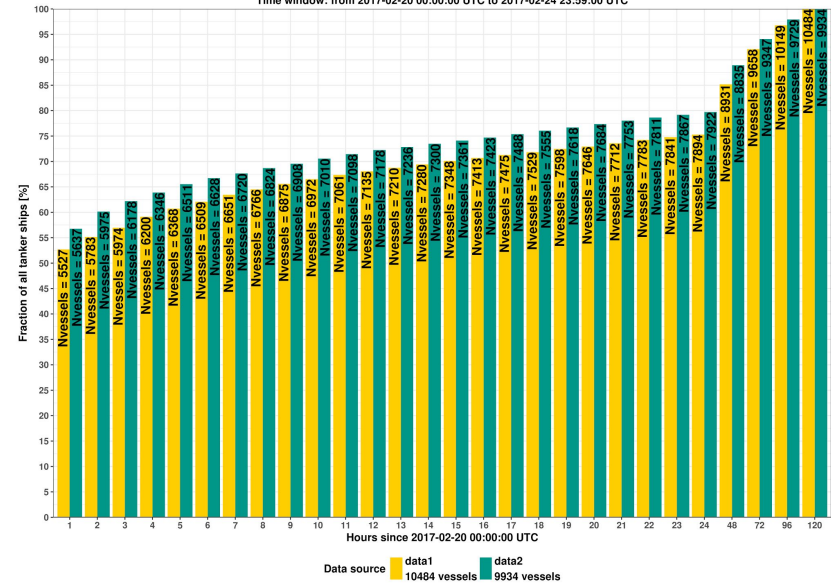
AIS geocoverage for container ships

All unique container ships with correct imo analyzed
Time window: from 2017-02-20 00:00:00 UTC to 2017-02-24 23:59:00 UTC



AIS geocoverage for tanker ships

All unique tanker ships with correct imo analyzed
Time window: from 2017-02-20 00:00:00 UTC to 2017-02-24 23:59:00 UTC



Data source data1 10484 vessels data2 9934 vessels

Information**Messages****Frequency**

Static or Voyage related

5,24

Every 6 min or, when data has been amended, on request

Dynamic

See below

Safety

As required

Long Range

27

30 mins

**Ships dynamic conditions****Not Changing Course****Changing Course**

At anchor or moored and moving less than 3 knots

3 min

3 min

At anchor or moored and moving faster than 3 knots

10 secs

10 secs

0 to 14 knots

10 secs

3 1/3 secs

14 to 23 knots

6 secs

2 secs

Over 23 knots

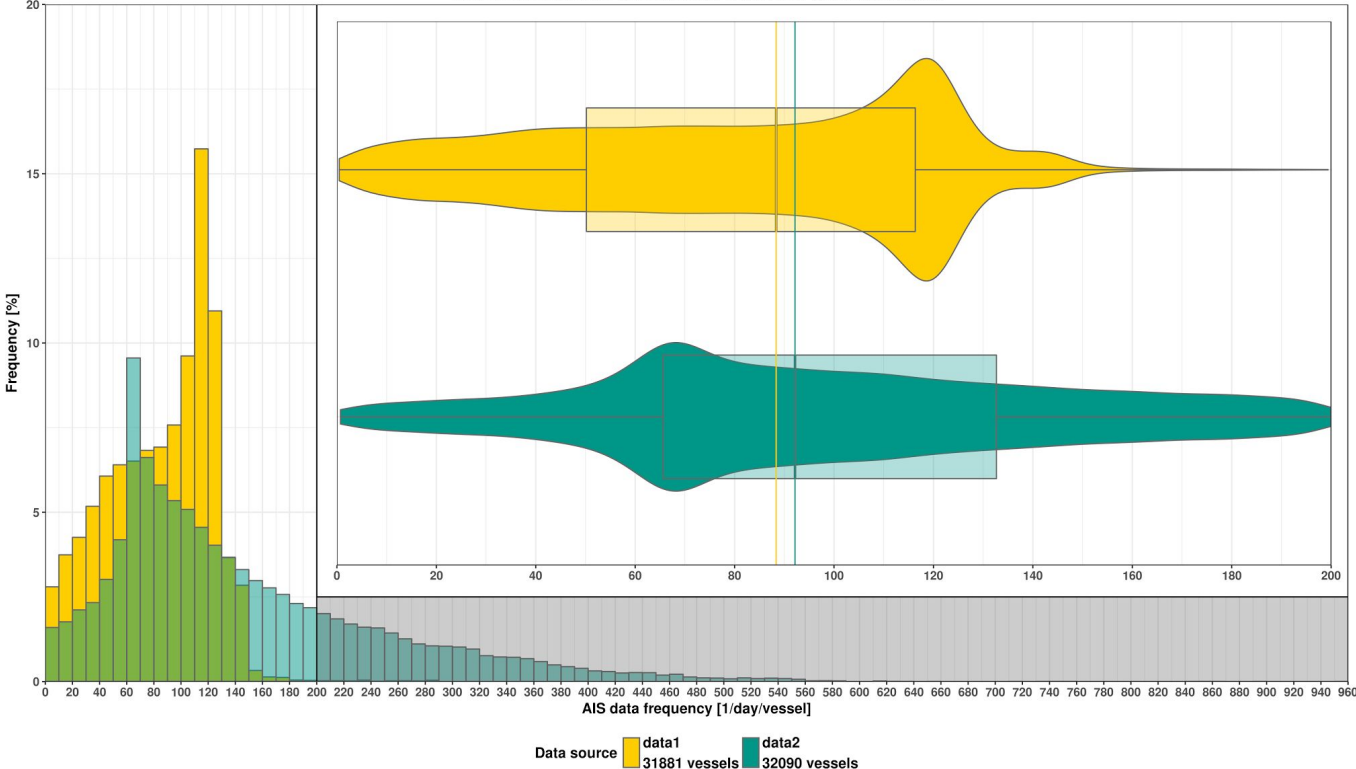
2 secs

2 secs

We have the best (ais) data in town ! how can we benchmark AIS ?



T-AIS signals sending frequency distribution
Common vessels with correct imo (32214 in total) analyzed
Time window: from 2017-02-20 00:00:00 UTC to 2017-02-24 23:59:00 UTC



We have the best (ais) data in town ! how can we benchmark AIS ?



Position Data

54% (after 15 minutes)

43 hours to get 100%

112% Position Updates

Static Data

34% (after 15 minutes)

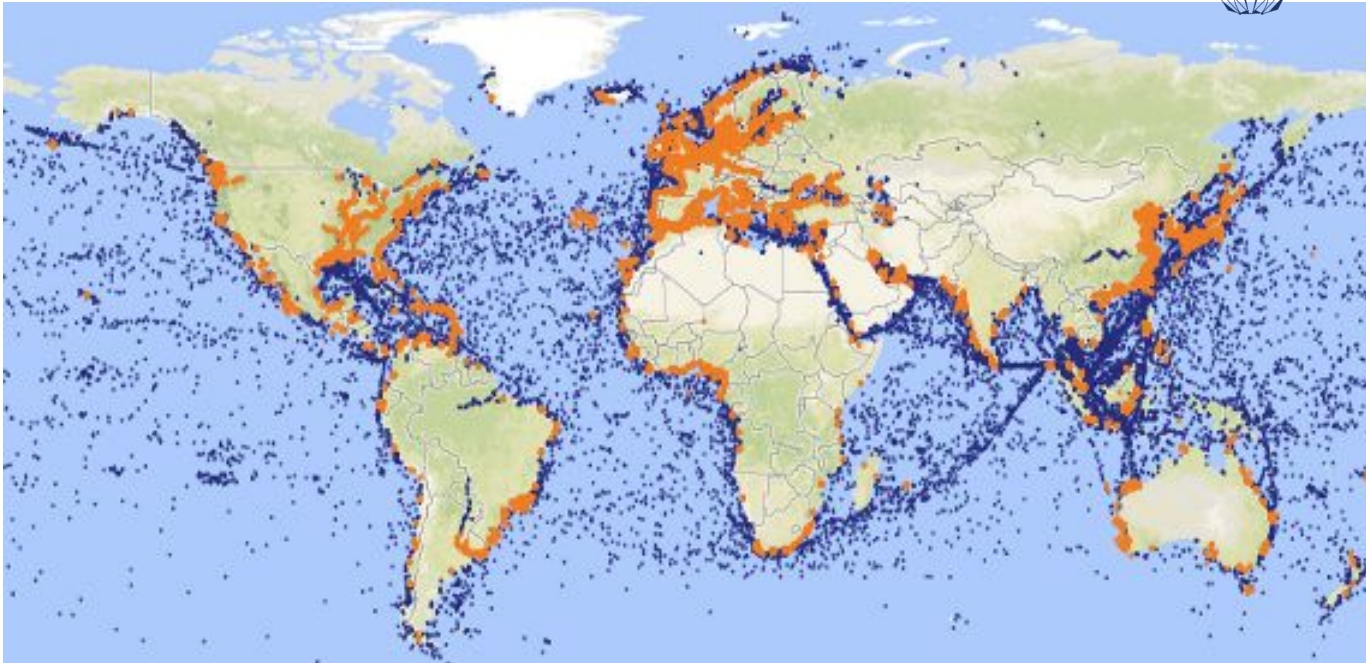
93% after 43 hours

89% Static Updates

Average over a certain
time period



use Geography for AIS Benchmarking



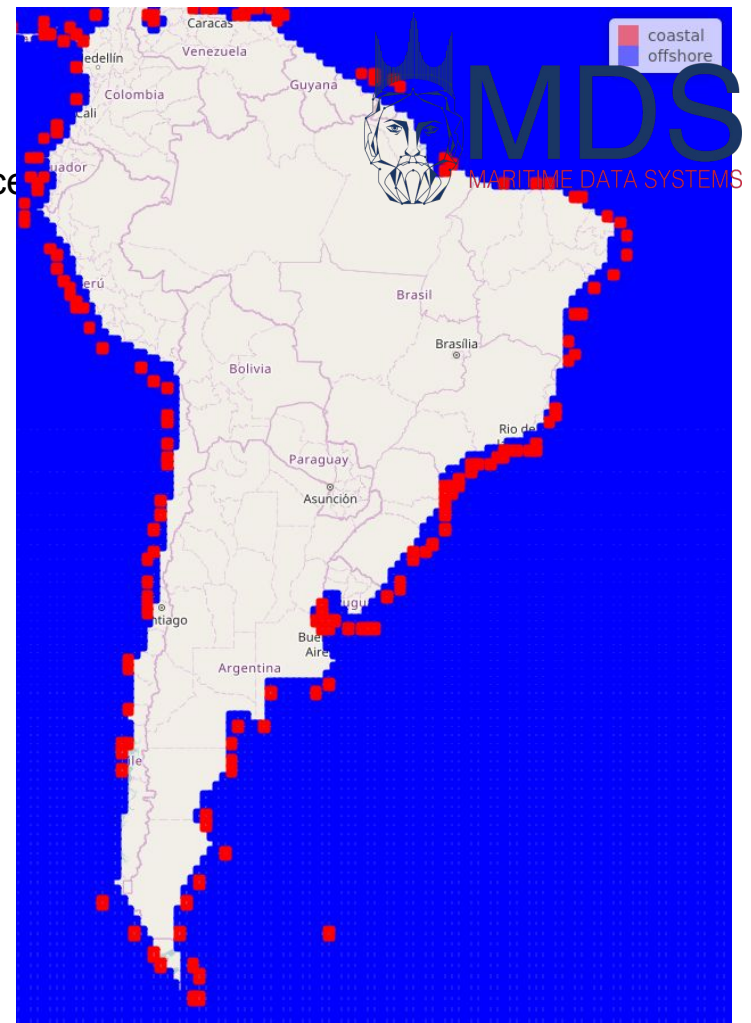
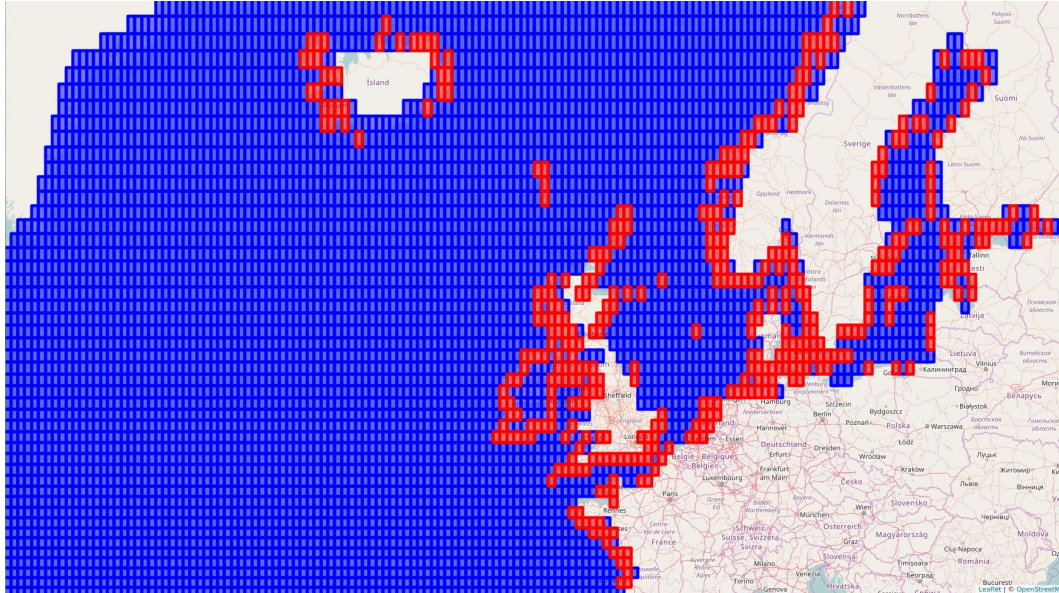
KPI: How many coastal / non coastal gridareas are covered after 5min, 1 hour, ... 24 hours

We have the best (ais) data in town ! how can we benchmark AIS ?

Image of 0.5' coastal boxes with Ports and offshore areas

Assumption: in a perfect world after 6 minutes all vessels are received

after 12 minutes static messages of all ships are received



Port (Coastal) Box

3.300



120 ais reports in Average in
15 minutes
+-23% variance

Offshore (Coastal) Box



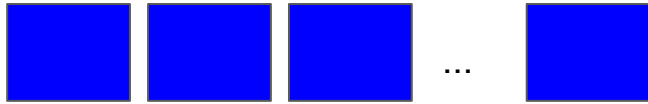
8 ais reports in Average
+-46% variance

We have the best (ais) data in town ! how can we benchmark AIS ?

AIS Position Reports



34% after 15 minutes



14% after 15 minutes



54% after 24 hours



24% after 24 hours



We have the best (ais) data in town ! how can we benchmark AIS ?

Position Data

54% (after 15 minutes)

43 hours to get 100%

112% Position Updates



Coastal Coverage

34% after 15 minutes

Static Data

34% (after 15 minutes)

93% after 43 hours

89% Static Updates



offshore Coverage

14% after 15 minutes

