



Port Detection



Justas Tamulis

Vilnius University

30/05/2025

Task Overview

- **Goal** : detect seaports in 15 M-row AIS log by clustering low-speed vessel positions.
- **Parallelism** : eight ProcessPoolExecutor workers load & filter 300 k-row chunks concurrently.
- **Pipeline** : noise filtering → per-vessel batch building → spatial DBSCAN ($\varepsilon=1$ km) → prune clusters with < 10 ships.
- Output : port statistics table and port map.

Filtering & Port Detection

Filtering :

- Group by MMSI, sort by time ; ignore Base Stations.
- Compute distance & Δt ; keep pairs with speed ≤ 5 kmh.
- Remove “jumps” (0 s but > 10 m displacement).
- Aggregate consecutive low-speed points into *batches*.
- Merge temporally separated batches that share the same location.

Clustering :

- Step 1 : DBSCAN on each vessel's batches to collapse repeated visits.
- Step 2 : global DBSCAN with $\varepsilon = 1$ km and `min_samples=10`.
- Discard clusters with < 10 unique MMSI (treated as noise/fishing grounds).

Port map

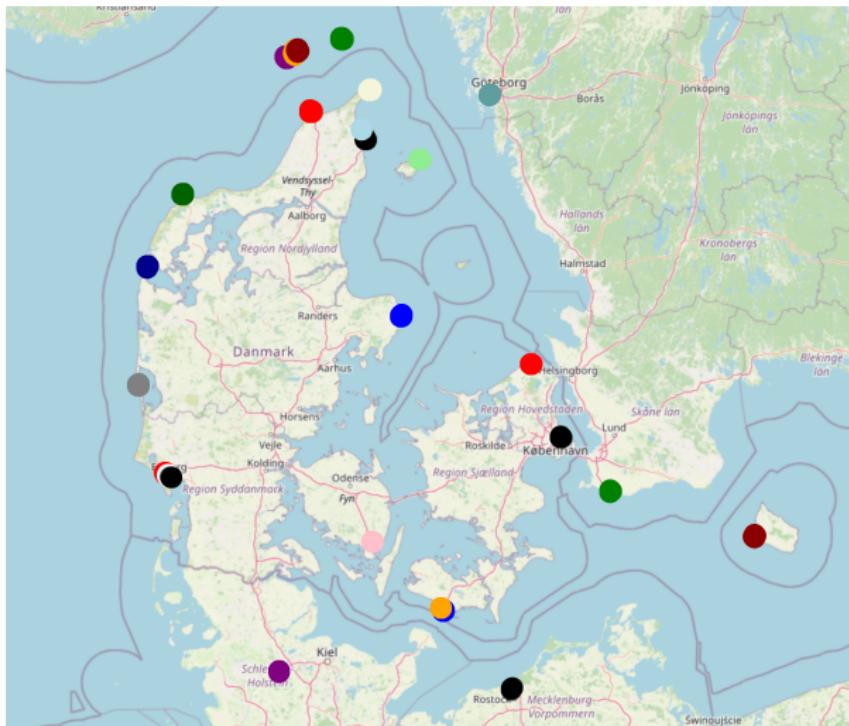


Figure – 26 ports detected in total. Most correspond to real harbours; a few clusters cover fishing grounds and one lies on an inland river.

Port statistics

Port ID	Area (km ²)	Unique Vessels	Total Batches	Data Points
23	4.48	16	25	1 176
18	2.40	51	55	254 426
0	2.34	27	27	139 413
4	0.09	14	14	13 668
14	0.05	18	18	108 488
17	0.04	15	15	96 632

Table – Three largest and three smallest ports by area.

- The number of batches represents batch count after step 1 of clustering, we can see that some of the vessels are still recorded multiple times in the same port. There are only 4 of these cases in biggest real port (id 18).
- The total data points represents how many signals there were in total from that port.
- However, these numbers represent only the portion that was not filtered out, therefore the real number of vessels in the port is likely much bigger.

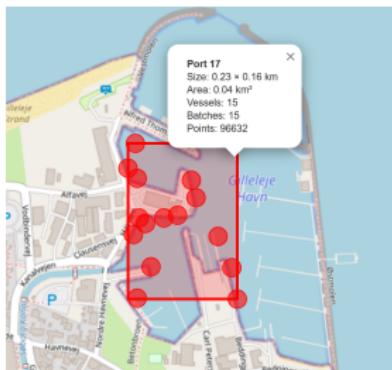
Specific ports



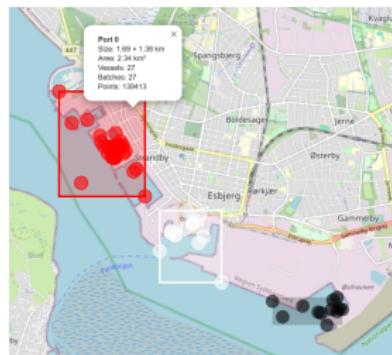
Port 23 — largest area,
fishing grounds



Port 18 — largest real port
with most vessels



Port 17 — smallest



Port 0 — multiple nearby ports