EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

AIS-Driven Maritime Insights: Improving Italian Ports Traffic Analysis

Istat | Istituto Nazionale di Statistica

6 June 2024

Session 20 – Big Data

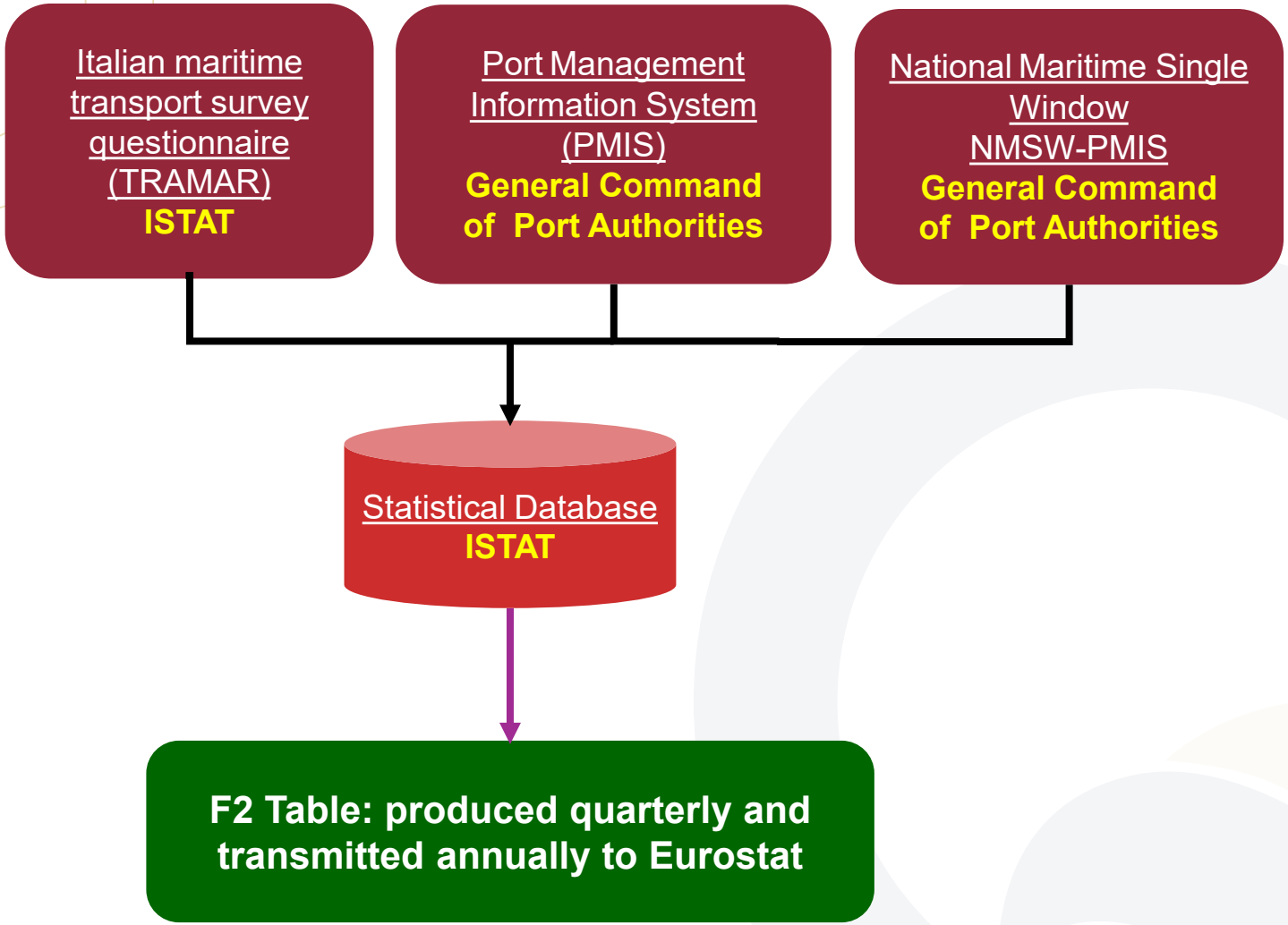
INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Italy - Maritime Transport Survey (TRAMAR)



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

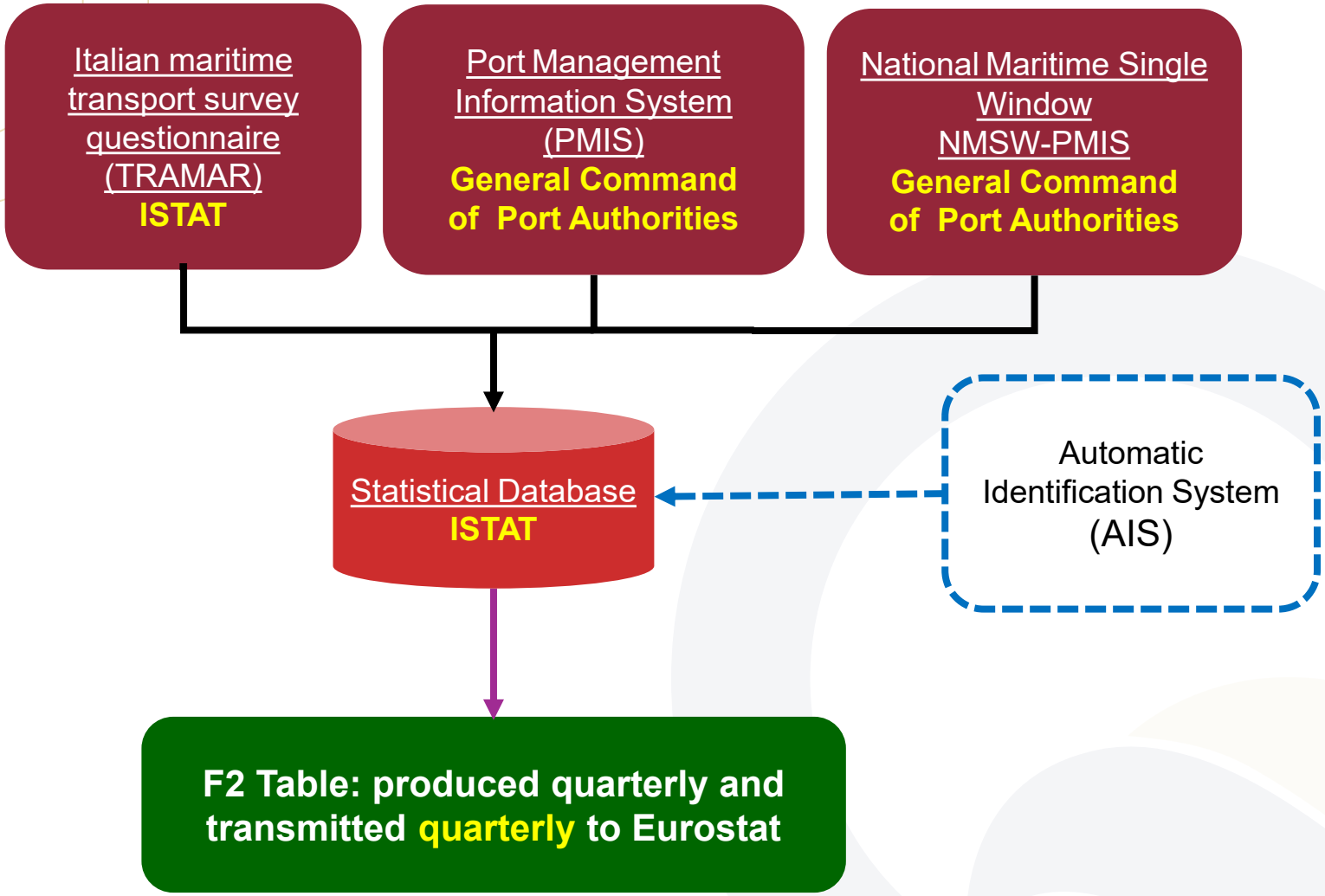
eurostat

The conference is partly financed by the European Union

Istat | Istituto Nazionale di Statistica

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

AIS in the Maritime Transport Survey



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

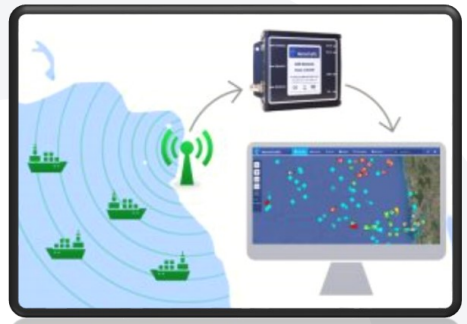
The conference is partly financed by the European Union

Istat | Istituto Nazionale di Statistica

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

What is AIS (Automatic Identification System)

- An **automatic tracking system on ships**, used in the **maritime world for safety** and management purposes.
- A **Big Data source** containing information regarding the main features and the identity of a **ship**, as well as its **location** and **speed** during navigation, **landing** and **anchorage**



Data Source

Provider: United Nation Global Platform (UNGP)

Content:
- **Live and global** archive data since **1st December 2018**
- **Observations** of all kind of ships registered about **every 10 minutes**

Environment: platform based on SPARK that, through parallel processing, allows to perform efficient queries and elaborations.

Static data					Dinamic data					Travel related data	
IMO	MMSI	CALLSIGN	VESSEL NAME	VESSEL TYPE	TIME	COORDINATES	NAVIG. STATUS	SPEED	COURSE	DESTINATION	DRAFT
8401561	20110115	ZAD4L	FROJDI II	Cargo	04/06/2023 19:45	41.1323 16.8530	MOORED	0	258	Ravenna	null

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

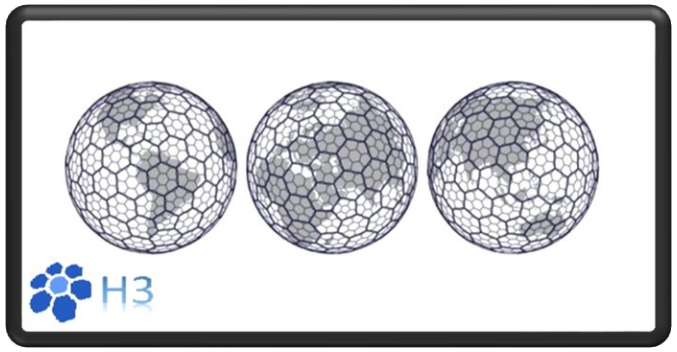
EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

H3: Hexagonal hierarchical geospatial indexing system

UN-AIS dataset provides also ships' position at **multiple resolutions**, through the **H3 index**.

H3 (Hexagonal hierarchical geospatial indexing system - https://h3geo.org) is a geospatial indexing system, developed by **Uber Technologies**, that approximates the **GPS coordinates** through a **hexagonal tessellation** of the earth's surface.

The H3 index **identifies** the hexagon containing the **ship's coordinates** and the **hexagon size** depends on the adopted **resolution**.



From:
latitude/longitude

41.89880/12.49513

To:
hexagon

ID:89402W2030Z032120

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly
financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Methodology

Consecutive signals from the same ship draw a **route** in the sea.

A route is identified by an arrival to and a departure from a port (**port call**).

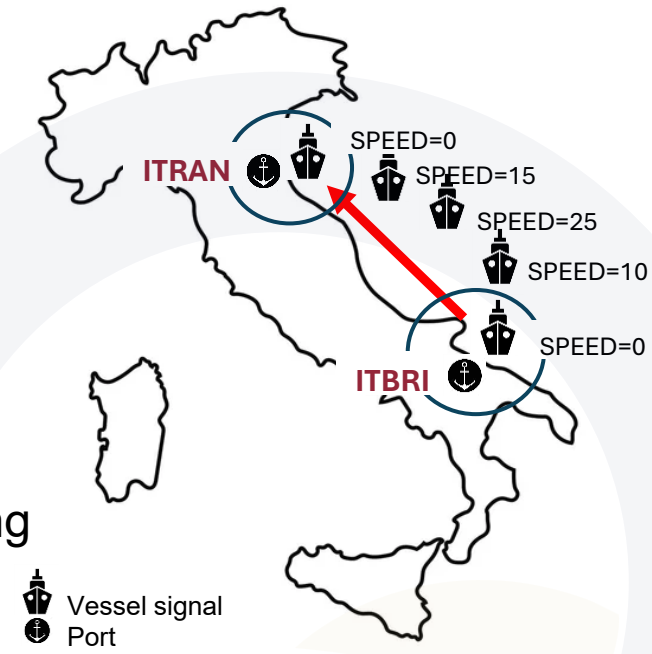
A **port call** is given by a ship being in a port area at speed near to 0.

Main goal:

Build routes from AIS data

Expected output:

Identify and quantify ships arriving to and departing from Italian ports



IMO	VESSEL TYPE	DEPARTURE PORT	ARRIVAL PORT	DEPARTURE DATE	ARRIVAL DATE
8401561	Cargo	ITBRI (Bari)	ITRAN (Ravenna)	04/09/2023	05/09/2023
9483712	Passenger	ITGOA (Genova)	ILOLB (Olbia)	05/09/2023	06/09/2023

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

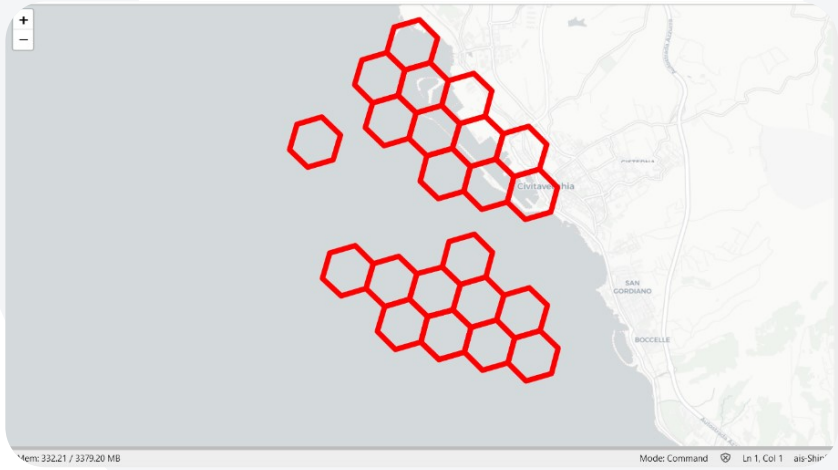
Methodology: routes from AIS – Problem 1

Identify the area of each **port**, where a port is identified by a point (latitude, longitude).

Solution

Build the port area as a groups of **hexagons of the H3 system** containing the position of ships that in the AIS data result to be **stationary near the point identifying a port**.

We have moved from a pair of coordinates (latitude/longitude) to a redefined geographic area that allows to improve the accuracy of the port calls identification.



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Methodology: routes from AIS – Problem 2

The **large amount of data** makes processing difficult, e.g. the number of database records for a **quarter exceeds 2 billions** records.

Solution

Filtering AIS data to collect observations related to:

- ships of some **specific types** (e.g. no fishing vessels and yachts).
- ships **passing close to Italian ports**.
- ships moving with a **speed close to 0 knots**, as we are interested in **ships stationary in the port areas**.

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly
financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Critical Issues: AIS vs Eurostat

Different coverage between Eurostat and AIS, both for vessels and ports.

	Eurostat	AIS
Vessel	Vessel with **GT>100**	AIS is mandatory for vessel with **GT>300**
Port	**Statistical port**: one port may include a group of arrival points	AIS ports are considered one per **point of arrival**

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Critical Issues: lack of data 1/2

This refers to a time period (even hours) during which no AIS observations are available for a vessel.

What happens if the ship disappears in a port area during this temporary “missing”?

In the example port calls in the port of Olbia are lost.

Real route

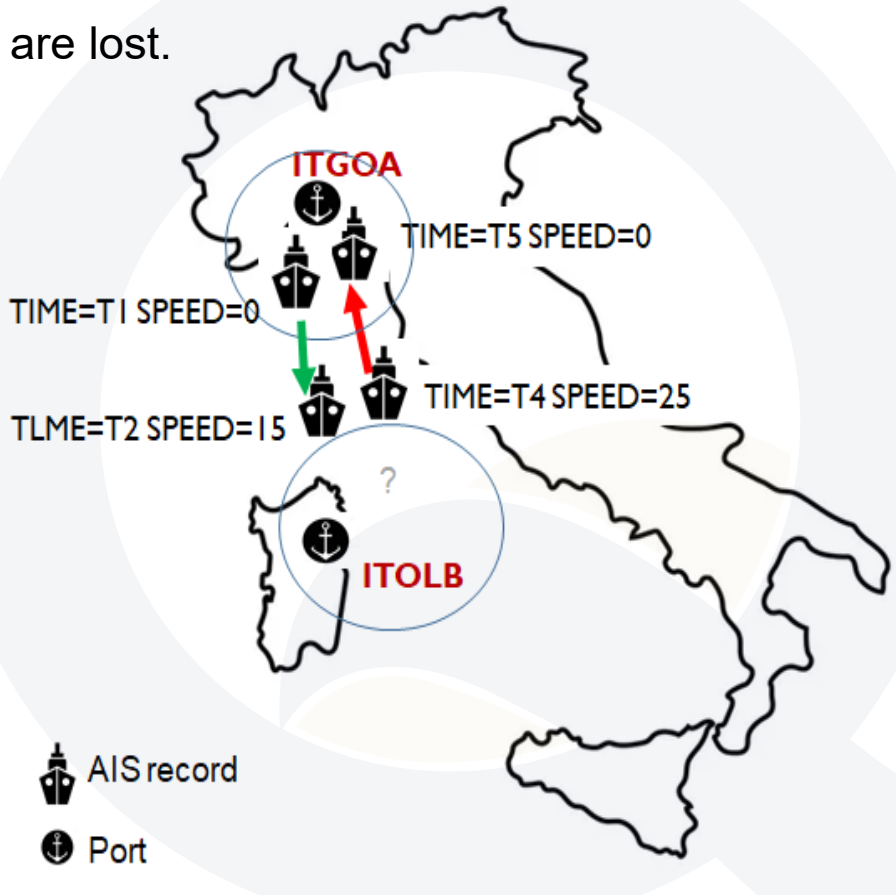
ITGOA (Genoa) – ITOLB (Olbia)

– ITGOA (Genoa)

Routes rebuilt by the algorithm

ITGOA (Genoa) – ITGOA (Genoa)

=> no navigation



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

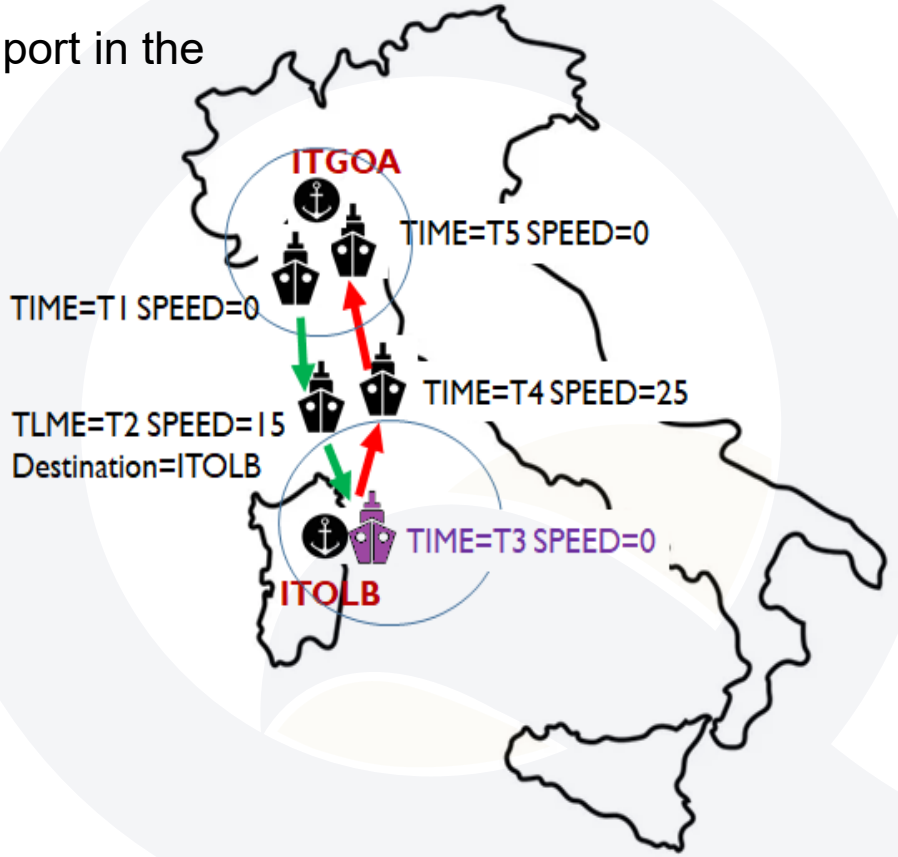
EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Critical Issues: lack of data 2/2

Port (e.g. of arrival) imputation:

- If the attribute "**destination**" in the last **AIS** record before the missing one clearly indicates a port, this is designated as the port of arrival.

- The port of arrival will be the most frequent port in the vessel's history, given the port of departure



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Results 1/2

In **2022**, the number of arrivals obtained from **AIS** data was approximately **69%** of the number of arrivals counted in **Table F2** produced for **Eurostat**, i.e. 295,819 with AIS vs 504,411 from data source based on Istat, for 56 ports.

UNLOcode	Port Name	F2 Arrivals	AIS Arrivals	Difference
ITBRI	Bari	2,424	2,325	-4.1%
ITCAG	Cagliari	1,690	1,691	+0.1%
ITCVV	Civitavecchia	2,815	2,558	-9.1%
ITGAE	Gaeta	226	365	+61.5%
ITGEA	Gela	103	3	-97.1%
ITGIT	Gioia Tauro	2,077	2,044	-1.6%
ITLIV	Livorno	6,035	6,122	+1.4%
ITMDA	La Maddalena	16,442	944	-94.3%
ITNAP	Napoli	34,367	32,225	-6.2%
ITPFX	Porto Foxi	769	1,095	+42.4%
ITPRJ	Capri	14,730	13,725	-6.8%
ITRAN	Ravenna	3,583	3,420	-4.5%

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

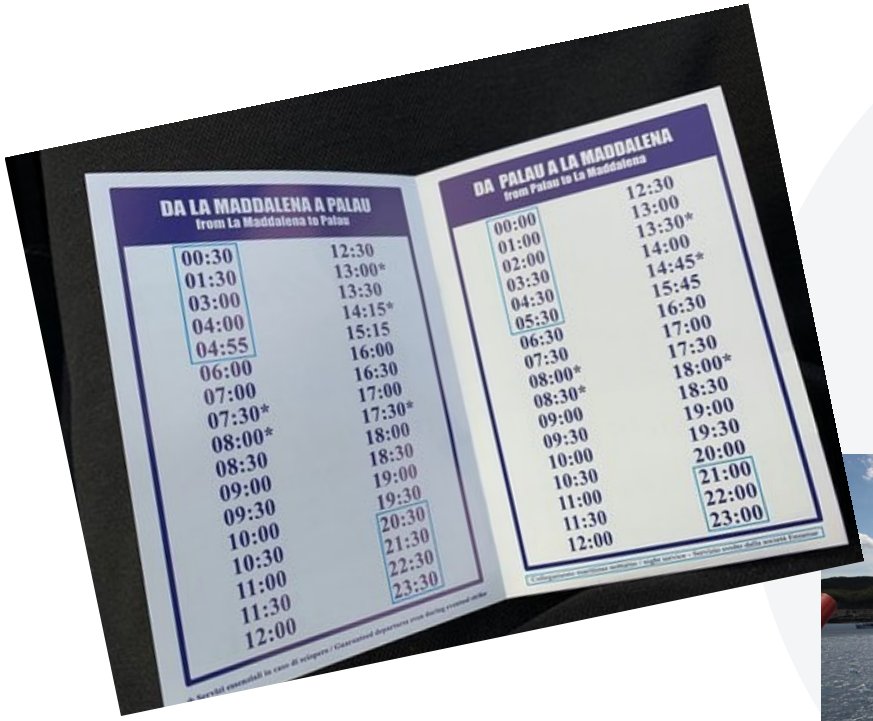
eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Results 2/2

If we don't take into account ports involved in short and frequent trips, the arrivals calculated by AIS even exceed those in F2, i.e. 117,685 arrivals via AIS vs 114,262 in F2.



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Future steps

- Use Deep Learning models to improve arrival ports imputation.
- Improve the quality of results on long tracks.
- Define a more sensitive algorithm to manage short and frequent tracks.
- Use alternative sources to account for ships operating scheduled voyages, e.g. ferries.

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly
financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Thanks

Angela Pappagallo, Istat, angela.pappagallo@istat.it
Norina Salamone , Istat, norina.salamone@istat.it
Luca Valentino , Istat, luca.valentino@istat.it
Francesco Sisti , Istat, francesco.sisti@istat.it
Mauro Bruno , Istat, mauro.bruno@istat.it

Istat | Istituto Nazionale di Statistica

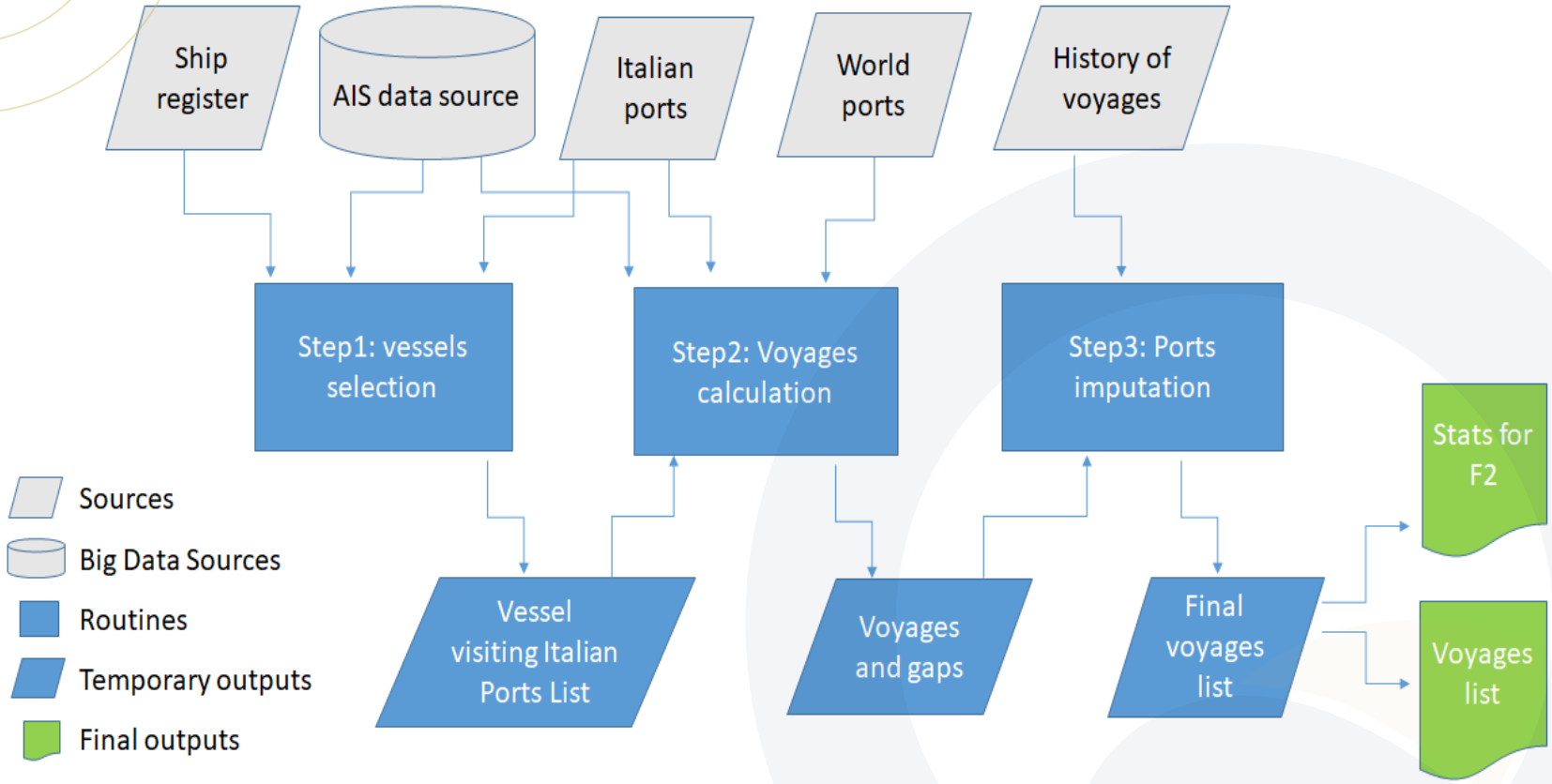
Instituto Nacional de Estatística
Statistics Portugal

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Methodology: Pipeline to process AIS data



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union