Statistics Poland

Statistics Poland

TranStat - smart system for road and maritime transport statistics

using Big Data sources and technology

Dominik Rozkrut , Michał Bis
Statistics Poland

Geneva, 23 June 2023
UNECE CES

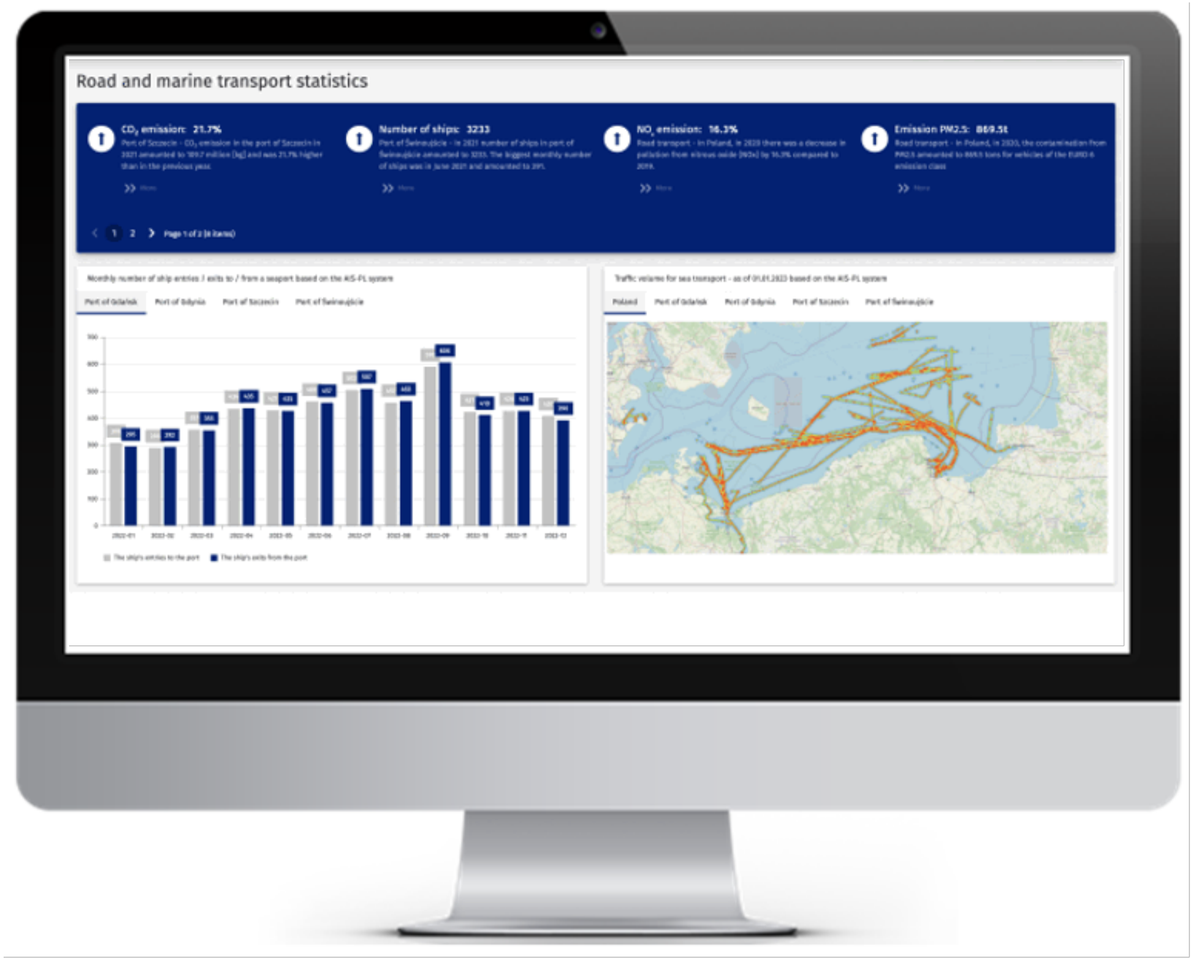
Outline

- TranStat: overview, assumptions, architecture, data processing
- AIS: Automatic Identification System in a nutschell
- Statistics of traffic intensity, emissions, transportation volume in maritime transport
- e-TOLL: electronic toll collection system in a nutschell
- Statistics of traffic in road transport – assumptions, results
- Conclusions

TranStat: Overview

- Implementation of experimental statistics in road and maritime transport
- Decreasing cost by using non-traditional data sources
- More timely and granular statistics
- Consortium:
  - Statistics Poland
    Maritime University of Szczecin
    Cracow University of Technology

https://transtat.stat.gov.pl

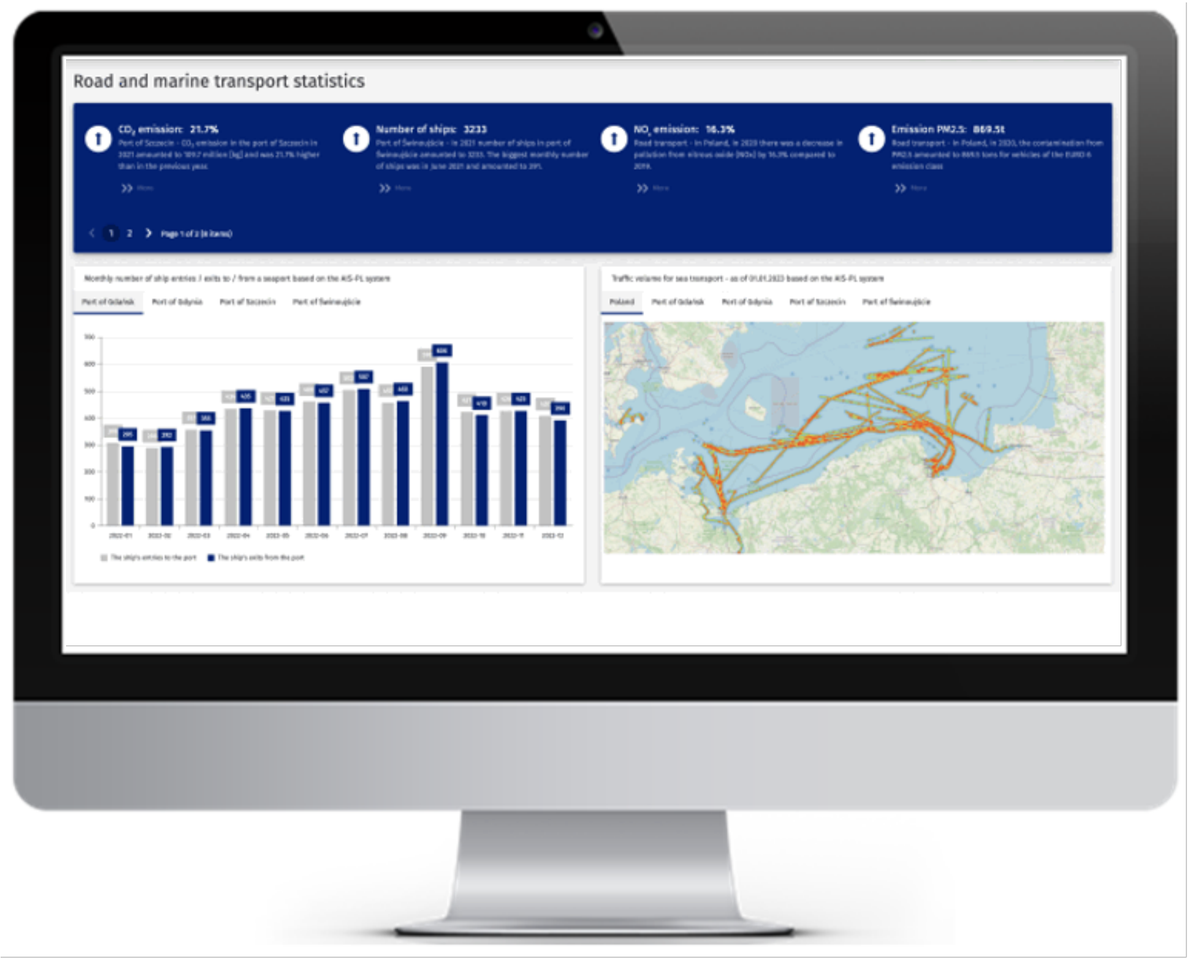


Statistics Poland

TranStat: Overview

- Getting access to sensor data:
  - Automatic Identification System (AIS)
  - e-TOLL electronic toll collection system
- Application of Big Data methods and tools
- Designing a methodology for traffic intensity, transportation volume and emissions estimations

https://transtat.stat.gov.pl



Statistics Poland

TranStat: Foundations

- Implementation of open standards
- Technological neutrality (vendor lock-in)
- Compliance with applicable laws
- Modular construction
- Easy expansion with new system functionalities in the future
- Ensuring an appropriate level of security
- Scalable Big Data solutions: volume, velocity and variety

TranStat: Architecture

- The TranStat IT system has been developed and implemented up to the level of the regular production environment.
- The system consist of

1. Data collection and processing subsystem responsible for the following subprocesses:
2. Data presentation and analysis subsystem
3. Data presentation and analysis subsystem

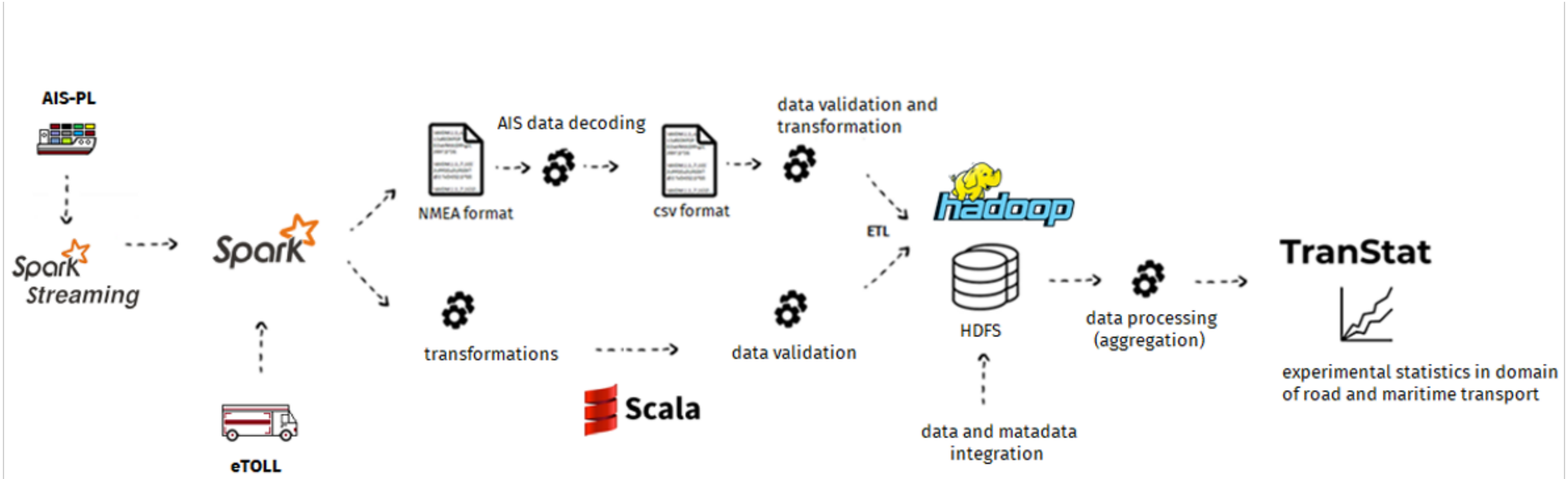
TranStat: Architecture

1. Data collection and processing subsystem
   - decoding AIS data
   - processing stream data from sensors
   - integration, validation, transformation and aggregation of data
2. Internal data presentation and analysis subsystem
   - enabling data exploration and visualization as well as statistical analyzes using the RStudio and Apache Zeppelin

TranStat: Architecture

1. External data presentation and analysis subsystem
   - portal intended for external users
   - data presentation and exploration
   - https://transtat.stat.gov.pl

TranStat: Data flow and processing



Automatic Identification System in a nutschell

- Automatic Identification System (AIS)
  - system of automatic identification used on ships for the electronic exchange of information between nearby ships, AIS base stations and satellites
- According to the requirements defined in Chapter V of the SOLAS Convention developed by the IMO, the AIS system should be installed on
  - all ships of 300 gross tonnage and more used in international shipping
  - all ships of 500 gross tonnage and more not used in international shipping
  - all passenger ships, regardless of size

Automatic Identification System in a nutschell

- The basic applications of the AIS system
  - strengthening of navigation safety (anti-collision system)
  - vessel traffic management support for coastal Vessel Traffic Service (VTS)
- Data source availability: legal basis
  - Regulation of the Minister of Maritime Economy and Inland Navigation of September 26, 2018 on the National System for Monitoring Vessel Movement and Information Transmission

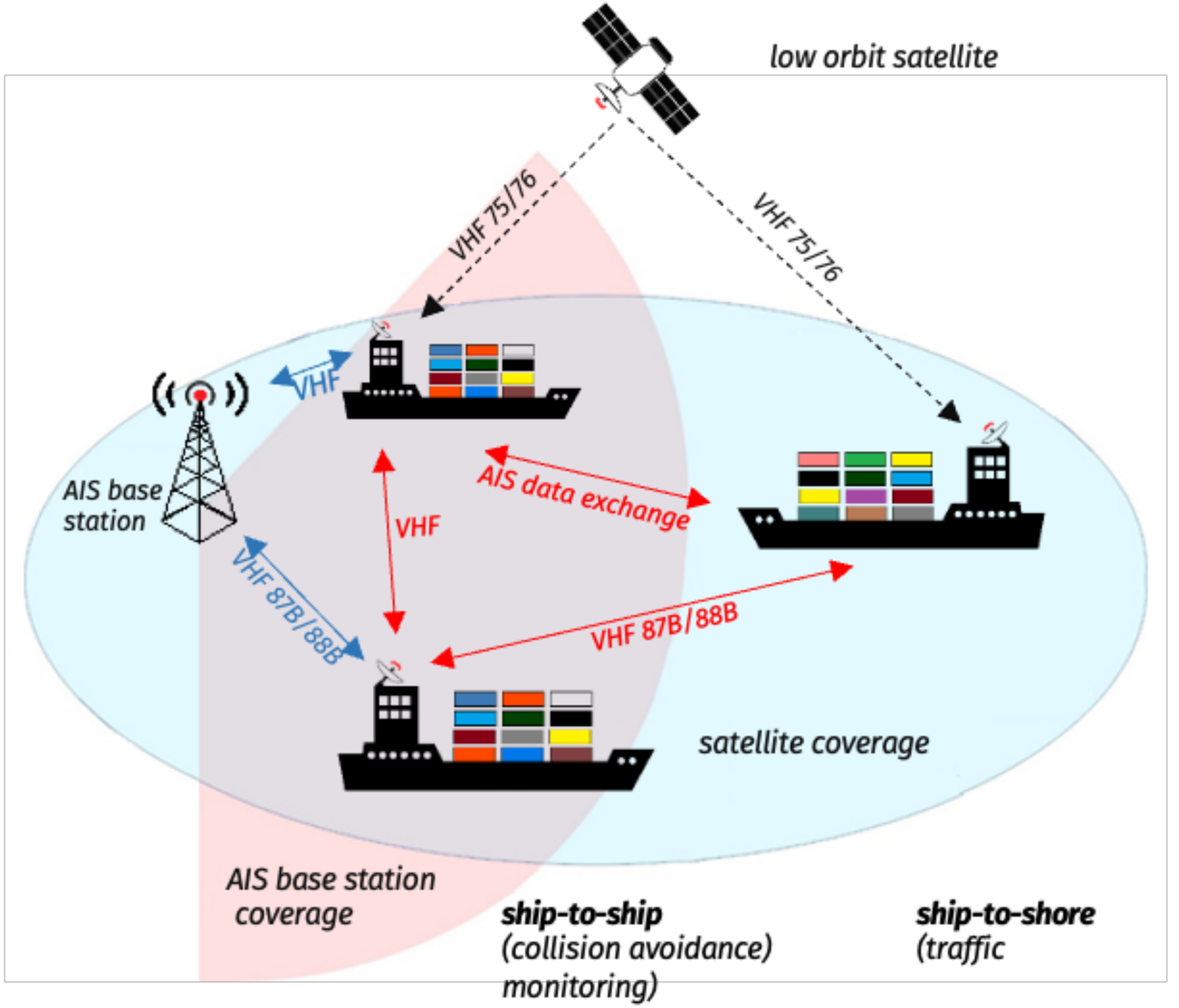


Figure 2. AIS – the principlr of work

Source : own study

Dynamic data

- Information on ship movements
- Automatically transmitted
- Every 2 to 10 seconds depending on vessel's speed
- Every 3 to 6 minutes when anchored

- Maritime Mobile Service Identity number (MMSI)
- AIS navigational status
- Rate of turn
- Speed over ground
- Position coordinates (longitude/latitude)
- Course over ground
- Heading
- Bearing at own position
- UTC second

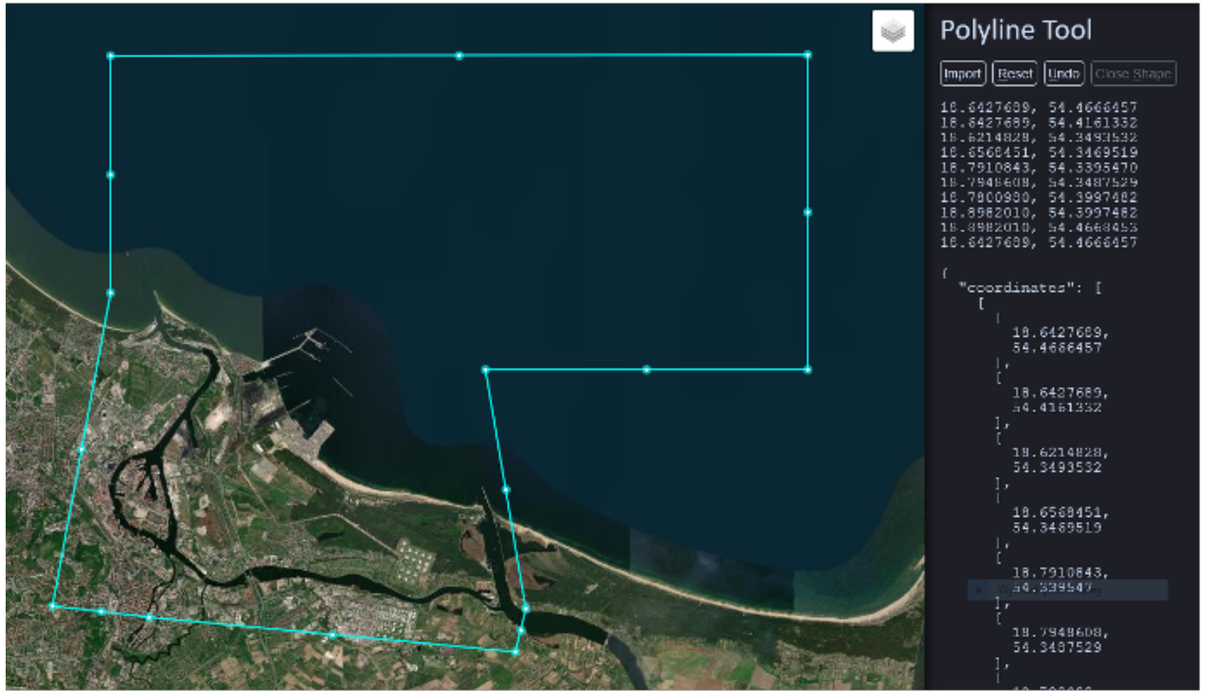
Static data

- Information on ship characteristic
- Manually transmitted
- Every 6 minutes

- International Maritime Organisation number (IMO)
- Call sign, Name, Type
- Dimensions
- Location of the positioning system's antenna on board the vessel
- Type of positioning system
- Draught
- Destination
- ETA (estimated time of arrival)

Statistics of traffic intensity in maritime transport

The port of Gdańsk



The port of Gdynia

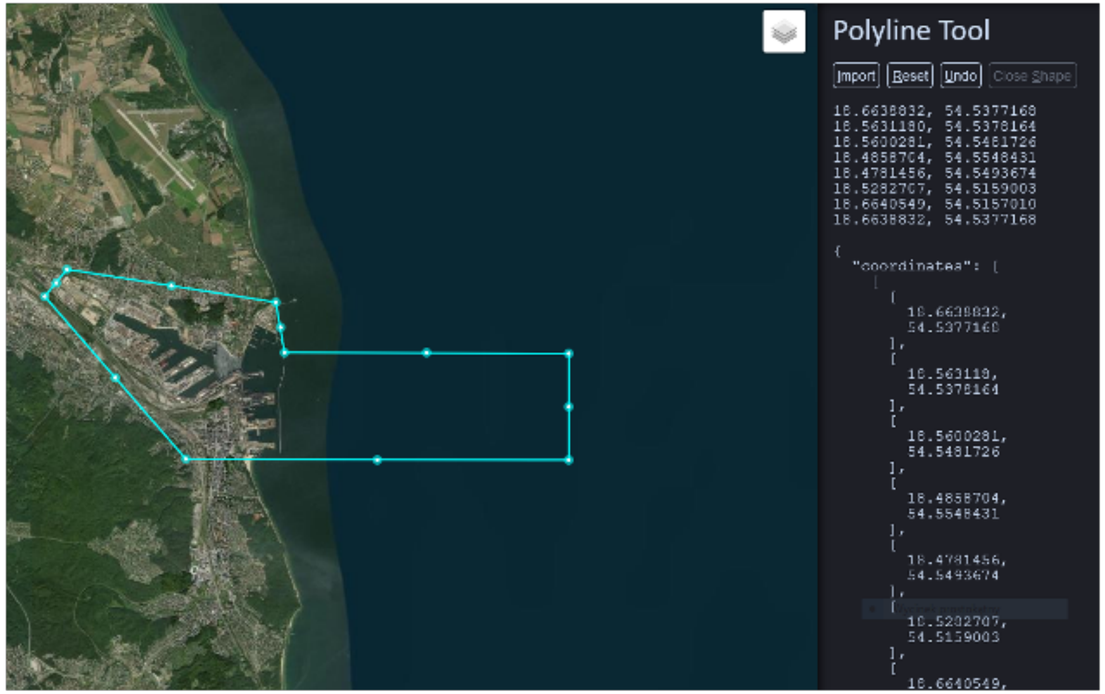
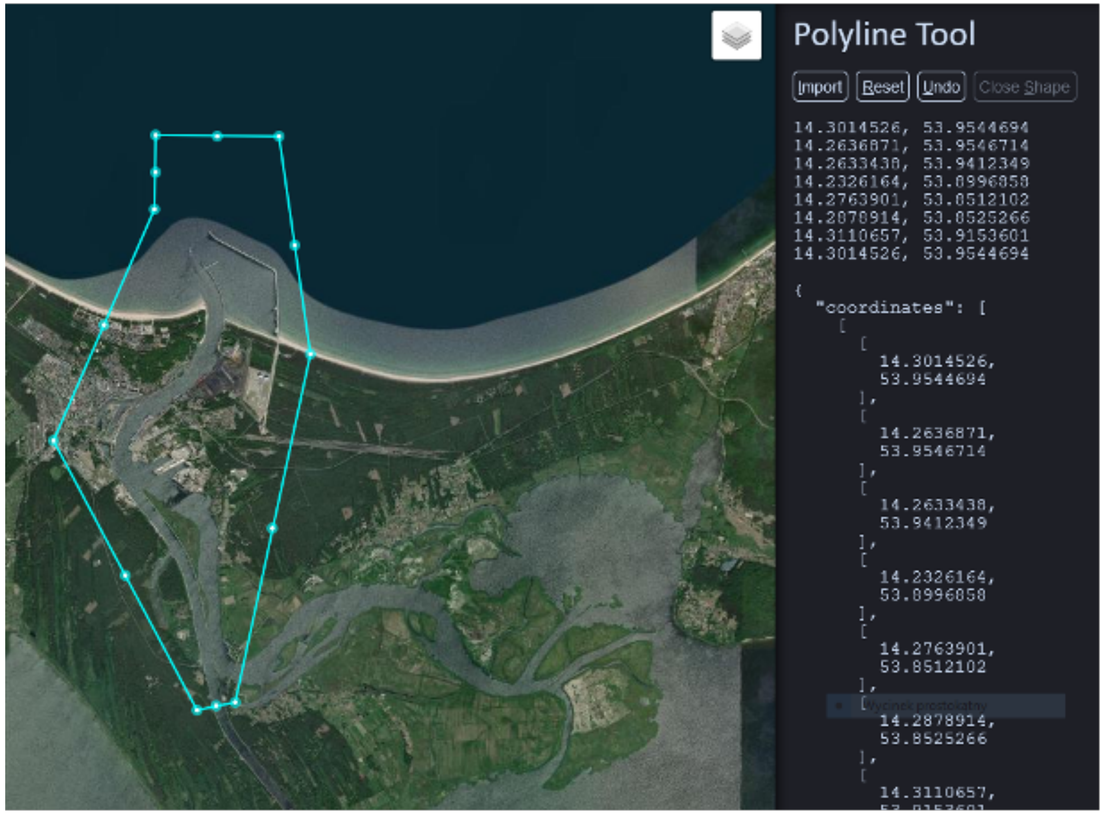


Figure 3. The ports of Gdańsk, Gdynia

Source: own study, generated on the basis of the tool:
https://www.keene.edu/campus/maps/tool/

Statistics of traffic intensity in maritime transport

The port of Świnoujście



The port of Szczecin

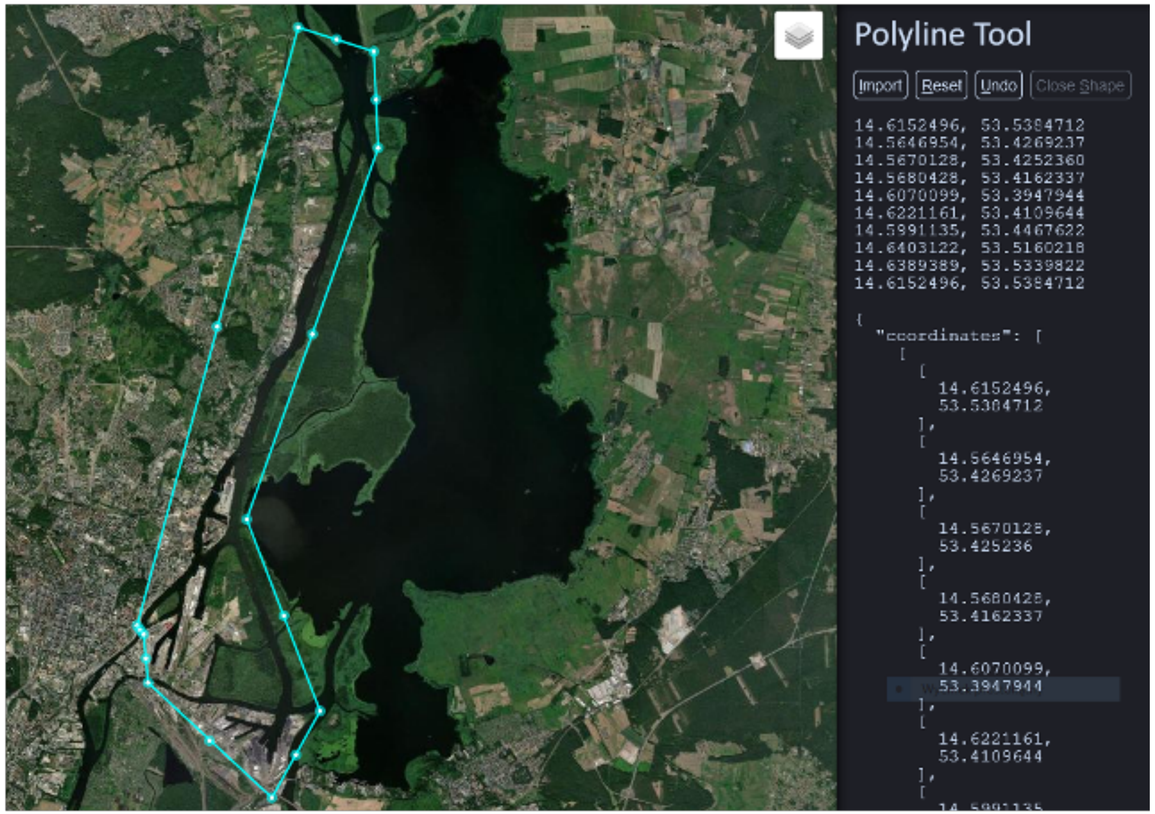


Figure 4. The ports of Świnoujścicie, Szczecin

Source: own study, generated on the basis of the tool:
https://www.keene.edu/campus/maps/tool/

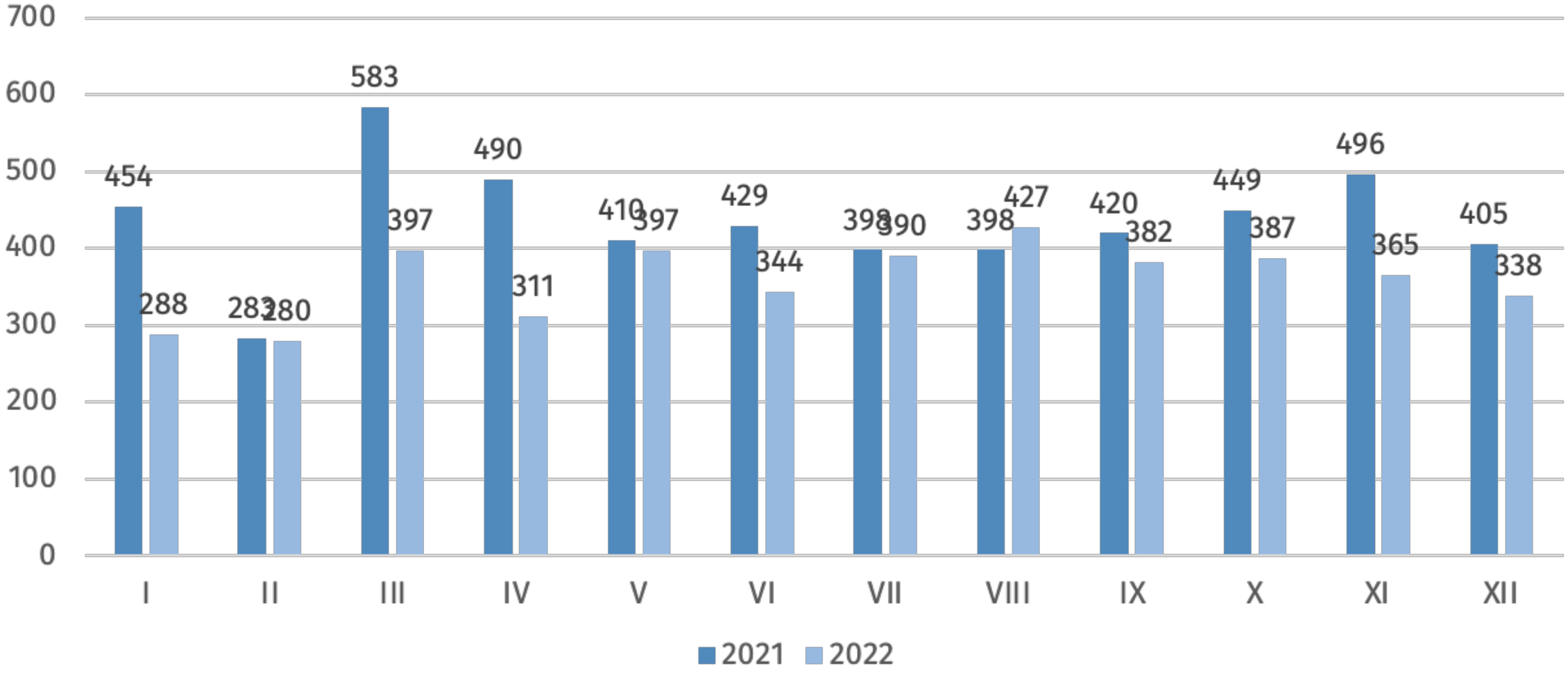
Statistics of traffic intensity in maritime transport

- Traffic intensity is understood as the intensity of the stream, defined as the number of transport units
  - passing through the line delimiting a given area in a certain period of time
- Implementation in TranStat application
  - Location: ports of Gdańsk, Gdynia, Szczecin, Świnoujście.
  - Data source: Automatic Identification System (AIS)

Statistics of traffic intensity in maritime transport

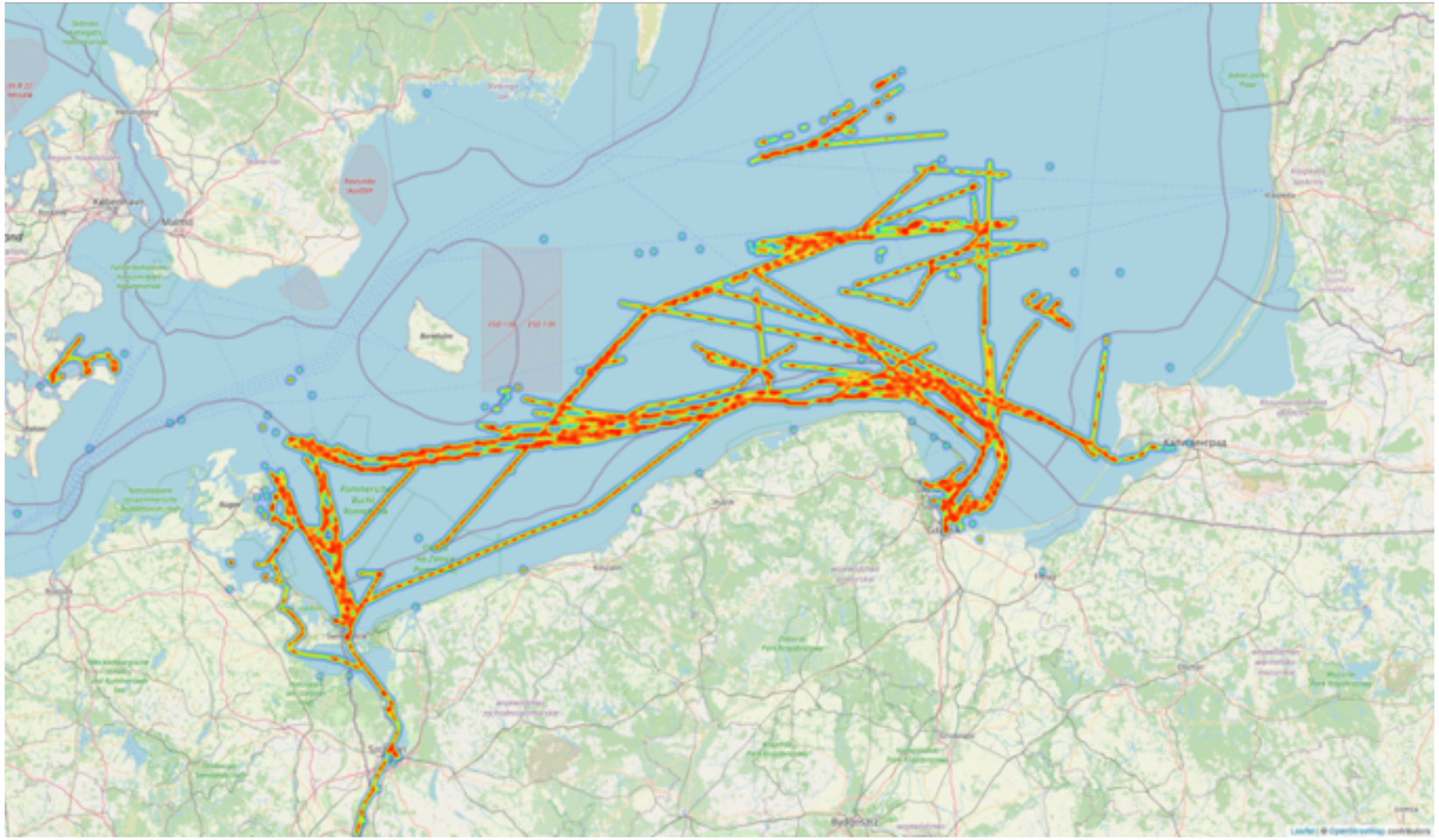
- Algorithms developed for the traffic intensity, lead to, among others
  - variables
    - number of ships at the port
    - number of calls by ships
- breakdowns
  - time: day, month, quarter, year
  - location: ports of Gdańsk, Gdynia, Szczecin, Świnoujście
  - means of transportation: by type of ships, by country of flag

Number of calls of ships to the port of Szczecin by months in 2021, 2022



Statistics Poland

Traffic intensity of vessels for Poland as of January 1, 2023



Statistics Poland

Traffic intensity of vessels for port of Świnoujście as of January 1, 2023

Statistics Poland

Statistics of emissions in marine transport: Foundations

- To estimate the amount of pollution emitted by means of sea transport
  - transport ships with gross tonnage of GT 100 and more were analyzed
  - for ports of Gdańsk, Gdynia, Szczecin, Świnoujście
- Data sources
  - Automatic Identification System
  - Information on ship parameters
    - a solution based on machine learning algorithms has been designed, on top of AIS data

Statistics of emissions in marine transport: Models

- reference model (detailed)
  - requires the preparation of a matrix of characteristic technical parameters dedicated to the ship, enabling the determination of the value of individual emissions
- specific model
  - using machine learning on a representative set of data from the REFERENCE model. The input parameters are the basic parameters of AIS messages and the output parameters are the emission values

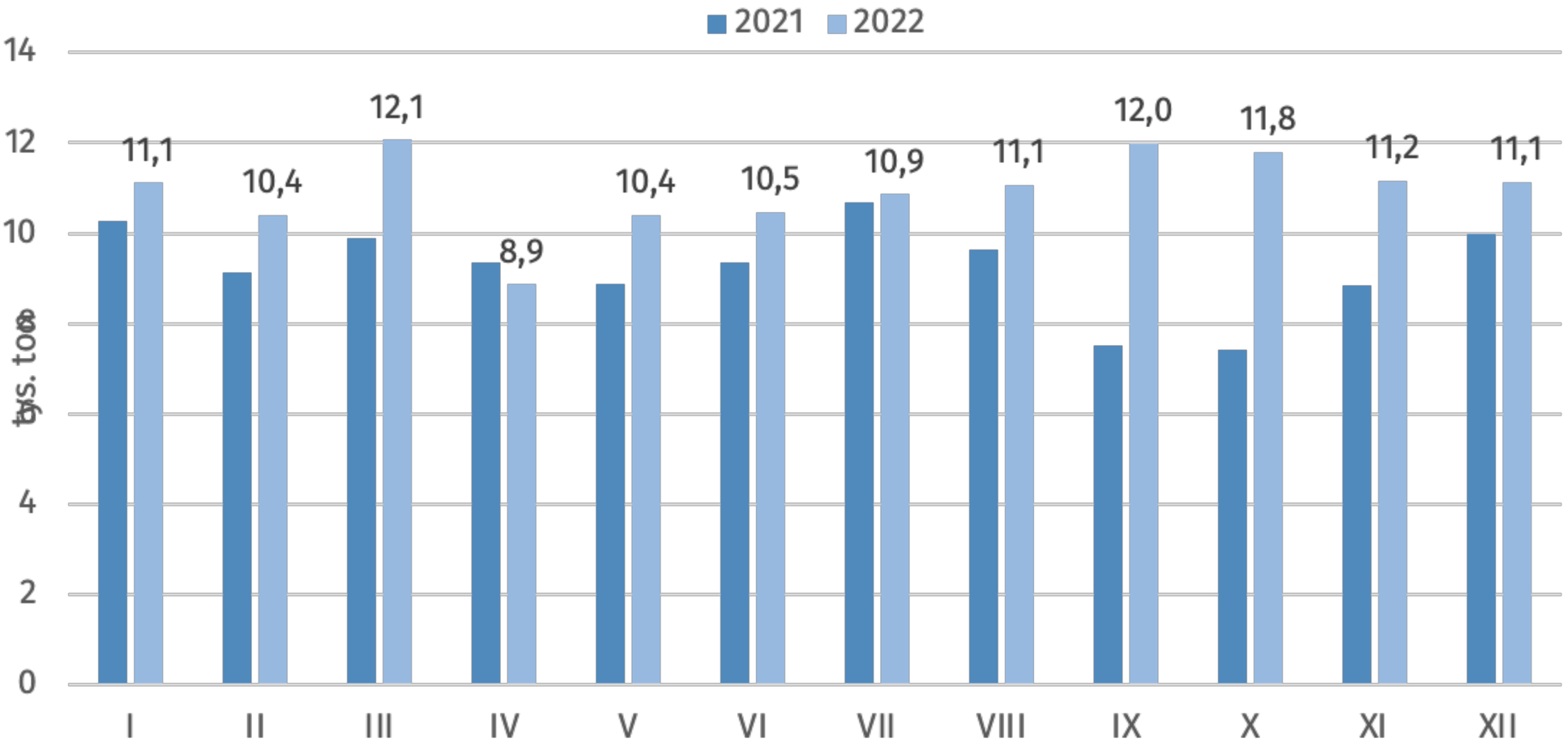
Statistics of emissions in marine transport: Models

- generic model
  - used when the specific model obtains limit values or the input data are outside the accepted range, e.g. ship length over 300 m

Statistics of emissions in marine transport: Results

- Variables
  - NOx emissions (nitrates, nitrites)
  - SOx emissions (sulphates, sulphites)
  - CO2 emission (carbon dioxide)
  - PM (particulate matter) emissions
- Breakdowns
  - time: day, month, quarter, year
  - spatial - ports located on the Polish coast
  - means of sea transport by types

Monthly CO2 emissions in the port of Szczecin in 2021 and 2022



Statistics Poland

Statistics of transportation volume in maritime transport

- Tonne-kilometre (tkm) is the unit of measure representing the transport of one tonne of cargo in a ship over one kilometre
- Passenger-kilometre (pkm) is the unit of measure representing the transport of one passenger in a ship over one kilometre
- We need to know about : the amount of cargo (loaded/unloaded) and the ship’s route
- Location: ports of Gdańsk, Gdynia, Szczecin, Świnoujście
- Data source: Automatic Identification System (AIS), Maritime transport data set based on Directive 2009/42/EC of the European Parliament and of the Council of 6 May 2009 on statistical returns in respect of carriage of goods and passenger by sea

Statistics of transportation volume in maritime transport

- The transportation volume estimation model implements the presentation of possible ship routes in the form of a directed (weighted) graph, where the vertices of the graph are navigation points or quays, and the edges are straight sections between them.
- Each edge contains the coordinates of the start and end points, and the weight is the distance between individual nodes, calculated by the Haversine formula.
- The graph consists of 9 859 vertices covering the entire globe.
- There are 10 731 connections between the vertices.
- Ports are vertices that have been described with UNLOCODE.
- There are 3 564 ports included in the graph.
- The sum of the weights of the edges of the graph is 1 088 864 km.

Statistics of transportation volume in maritime transport

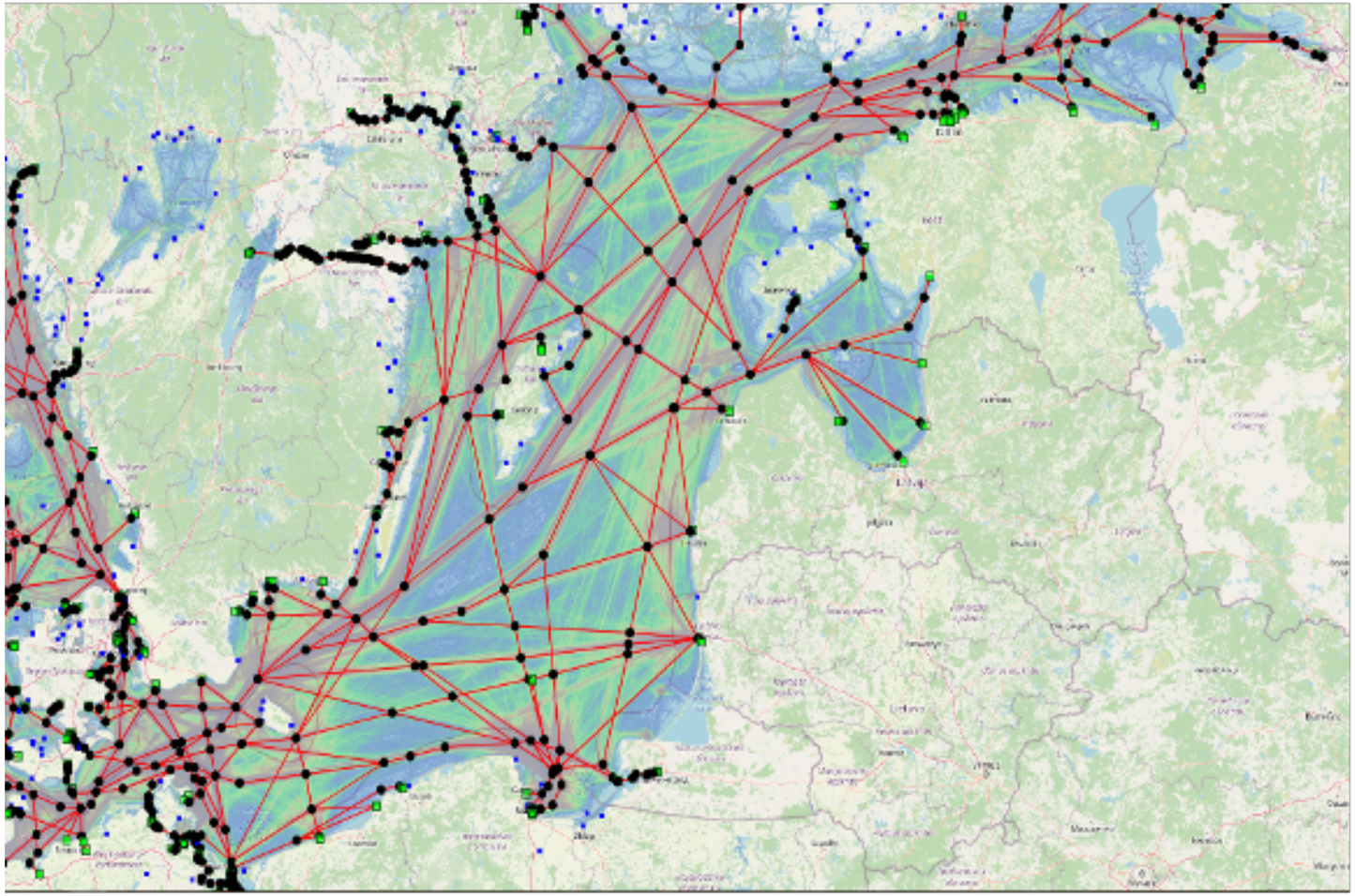


Figure 7. *Graph visualization for the Baltic Sea*

Source: Maritime University of Szczecin

Statistics of transportation volume in maritime transport

- Implementation of port distance estimation based on directed graph:
  - 1. determining the weights of the edges of a graph
    the Haversine formula.
  - 2. finding the shortest path in a graph- the Dijkstra's algorithmof transportation volume in maritime transport

A - start vertex , C - end vertex

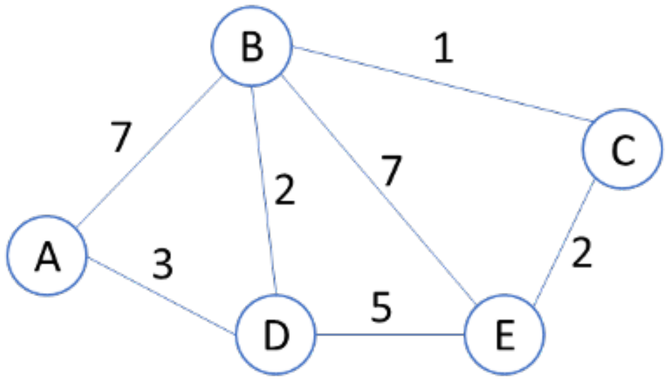


Figure 8. Graph with the representation of weights

Calculated shortest path from A to C is 6
and goes through vertices A, D, B and C

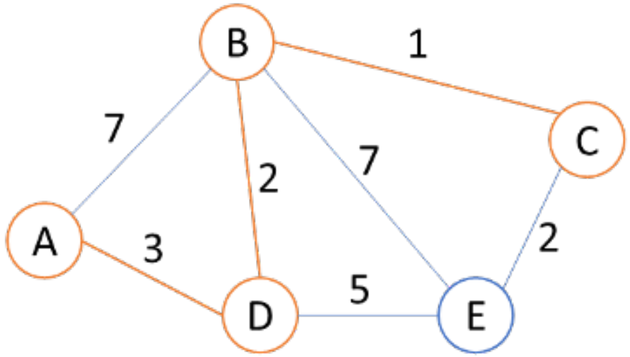
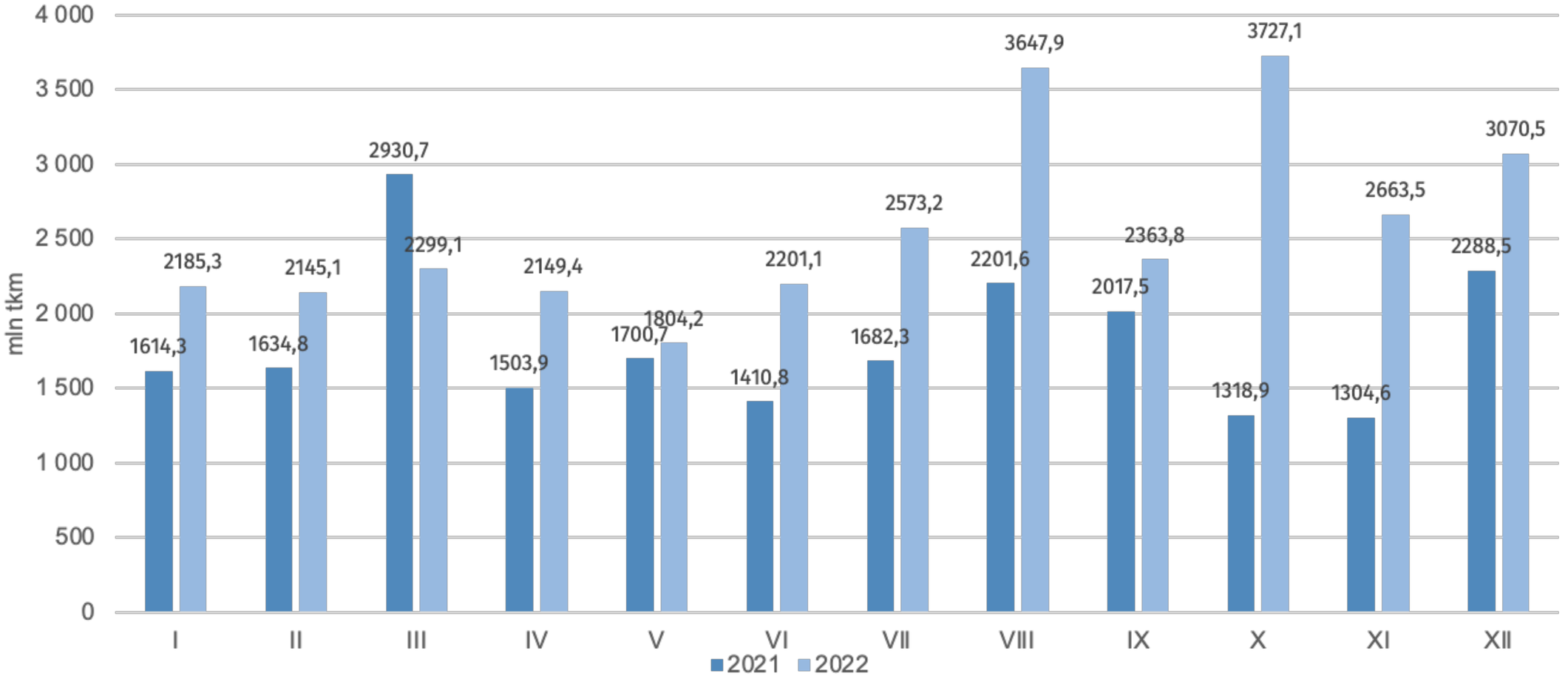


Figure 9. A graph with a representation of the shortest path

Statistics of transportation volume in maritime transport

- Variables
  - transportation volume for cargo and passengers,
  - avarage transport distance for 1 tonne of cargo in kilometers
  - avarage transport distance for 1 passenger in kilometers
- Breakdowns
  - time: day, month, quarter, year
  - location: ports of Gdańsk, Gdynia, Szczecin, Świnoujście
  - means of maritime transportation - by type, by flag, by gross tonnage
  - type of cargo – cargo group, commodity group

Transportation volume in relations to the port of Szczecin



Statistics Poland

e-TOLL – electronic toll collection system

e-TOLL is an advanced solution developed, implemented, maintained and monitored by the Head of the National Revenue Administration

It is based on the Global Navigation Satellite System for user position location with the use of virtual gates.

In the TranStat system, we only test heavy vehicles with a maximum permissible weight of more than 3,5 tons, including buses.

The length of the paid sections is currently approximately 3,677 km

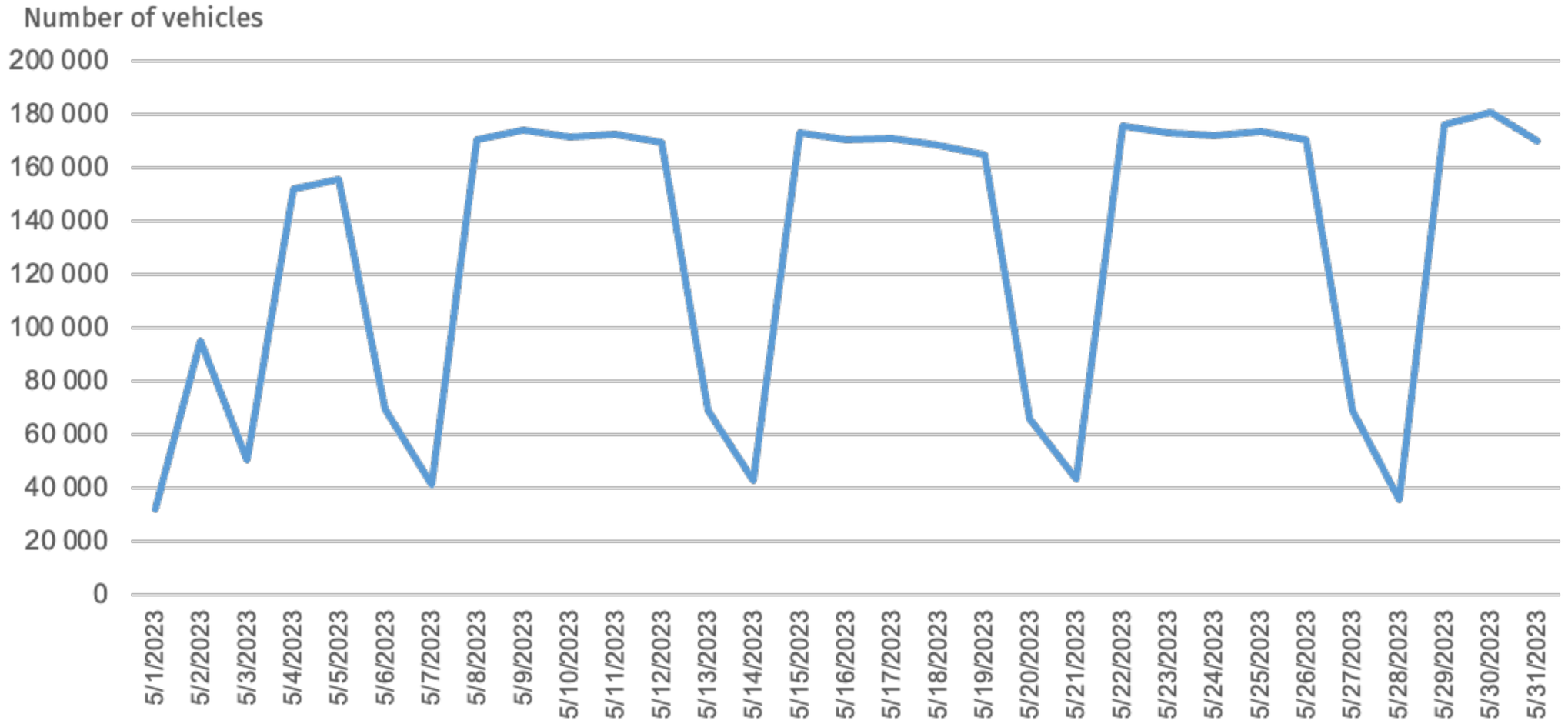
Statistics of traffic in road transport: Assumptions

- In total, there are 951 virtual gates on motorways, expressways and national roads covered by the e-TOLL system.
- In order to create statistics on traffic volume, it was assumed that a vehicle made a trip under the e-TOLL system if it was registered in at least 2 transactions from the analyzed dataset.
- As a result of the developed algorithms for the traffic the following variables and breakdowns are obtained, among others:
- variables:
  - number of transactions - the number of toll transactions for vehicles subjected to toll, registered on the toll section;
  - number of vehicles - unique number of vehicle occurrences at a toll collection point or section

Statistics of traffic in road transport: Assumptions

- Breakdowns:
  - time: day, week, month
  - spatial: road number
  - categories of vehicles according to payload groups (GVW)
    - coaches, capacity group 30, with more than 9 seats (including the driver),
    - heavy duty vehicles:
      - load group 41 – heavy duty vehicles with a GVW above 3.5 tons and below 12 tons,
      - load group 42 - heavy duty vehicles with a GVW above 3.5 tonnes and below 12 tonnes with the physical ability to tow a trailer,
      - load group 50 – heavy duty vehicles with a GVW over 12 tons.
  - categories of vehicles according to the Euro emission class (0 - 6) - European emission standard specifying the standards of permissible emissions in new vehicles sold in the EU and the European Economic Area.

Daily traffic volume on the road network covered by the e-TOLL system by the number of vehicles > 3.5 tonnes - May 2023



Conclusions

- The implementation of the TranStat project in the field of maritime statistics has enriched the current statistical production carried out by Statistics Poland through
  - access to streaming Big Data source related to maritime transport (AIS)
  - implementation of the necessary Big Data technology for sensory data enabling an automatic process of data flow, validation and processing
  - development of traffic intensity, transportation volume and emissions models in maritime transport with the use of sensory data
  - development of algorithms enabling generation of new statistics and obtaining new knowledge in the field of maritime transport statistics by using the correlation of multiple data sources
  - reduction of research costs thanks to the use of modern technology in the collection and processing of non-statistical sources (AIS)

Statistics Poland

Thank you !

Dominik A. Rozkrut
President, Statistics Poland

Statistics Poland