IOS-ISI 2024
MEXICO CONFERENCE

Using new and innovative data sources for statistics on container transport

- Statistics Netherlands -

INEGI
International Statistical Institute
IAOS
IMPROVING OFFICIAL STATISTICS

Statistics Netherlands

Port of Rotterdam

	Port	Volume 2021 (Million TEU)
1	Shanghai, China	47.03
2	Singapore	37.49
3	Ningbo-Zhoushan, China	31.07
4	Shenzhen, China	28.77
5	Guangzhou Harbor, China	24.18
6	Busan, South Korea	22.71
7	Qingdao, China	23.71
8	Hong Kong, S.A.R, China	17.8
9	Tianjin, China	20.27
10	Rotterdam, The Netherlands	15.3

15 million containers continue their journey through Netherlands

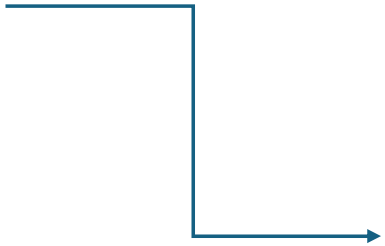
- Road transport
- Inland Waterways
- Rail transport

Port of Rotterdam

Policy making and decisión making

- **Infrastructure policy** (by the Dutch Government)

- **Monitoring modal shift containers** (European Green Deal - sustainable transportation)

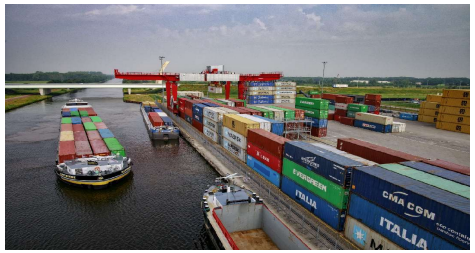


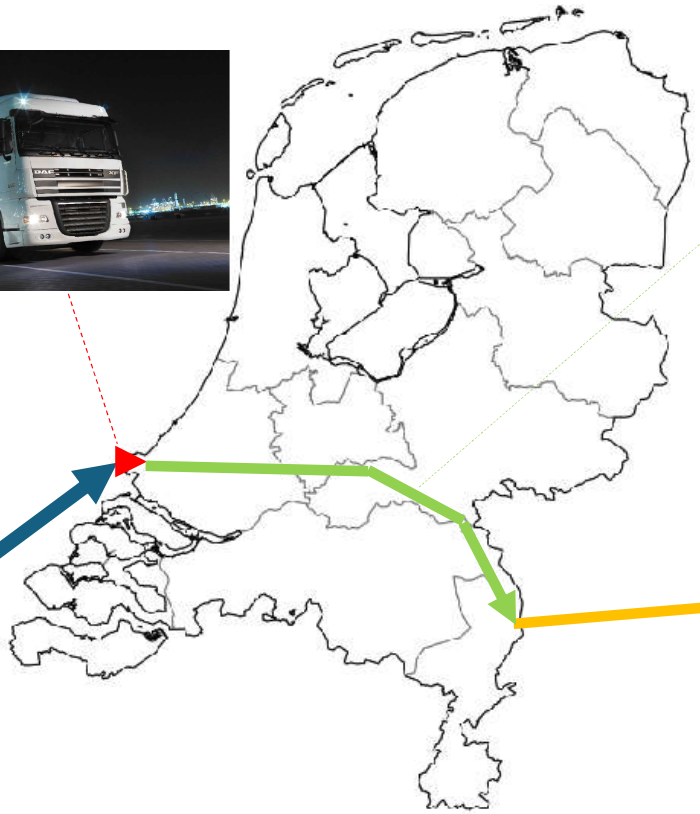
container Project → Container chains

'If things go well for the port of Rotterdam, the European economy will do well.'

URSULA VON DER LEYEN, President of the European Commission

The need: insights into multimodal container chains





History container project

Before 2018

Starting point

Import Control System data
Export Control system data
- Custom data -

IVS Next
(AIS and Lock data)

ProRail (rail infra)
Rail operators data

Road transport
sample survey

IOS-ISI 2024 MEXICO CONFERENCE

History container project

2018

Eurostat grant constructing container chains:
Combine registrations and collected data sources at micro level

- *Issues with coverage data*
- *Detail of the data in different modals*

Conclusion: Too many gaps

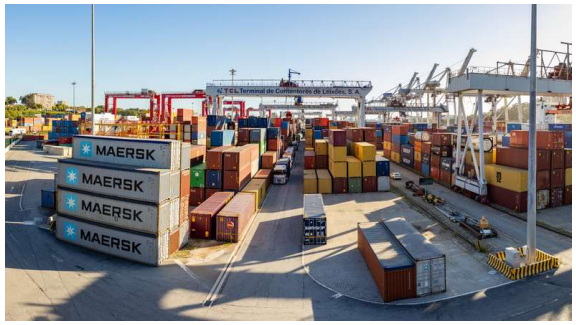
eurostat

History container project

2020

Pilot adding private data:

Add 10 private data sources from different modalities in open format



Data sources:

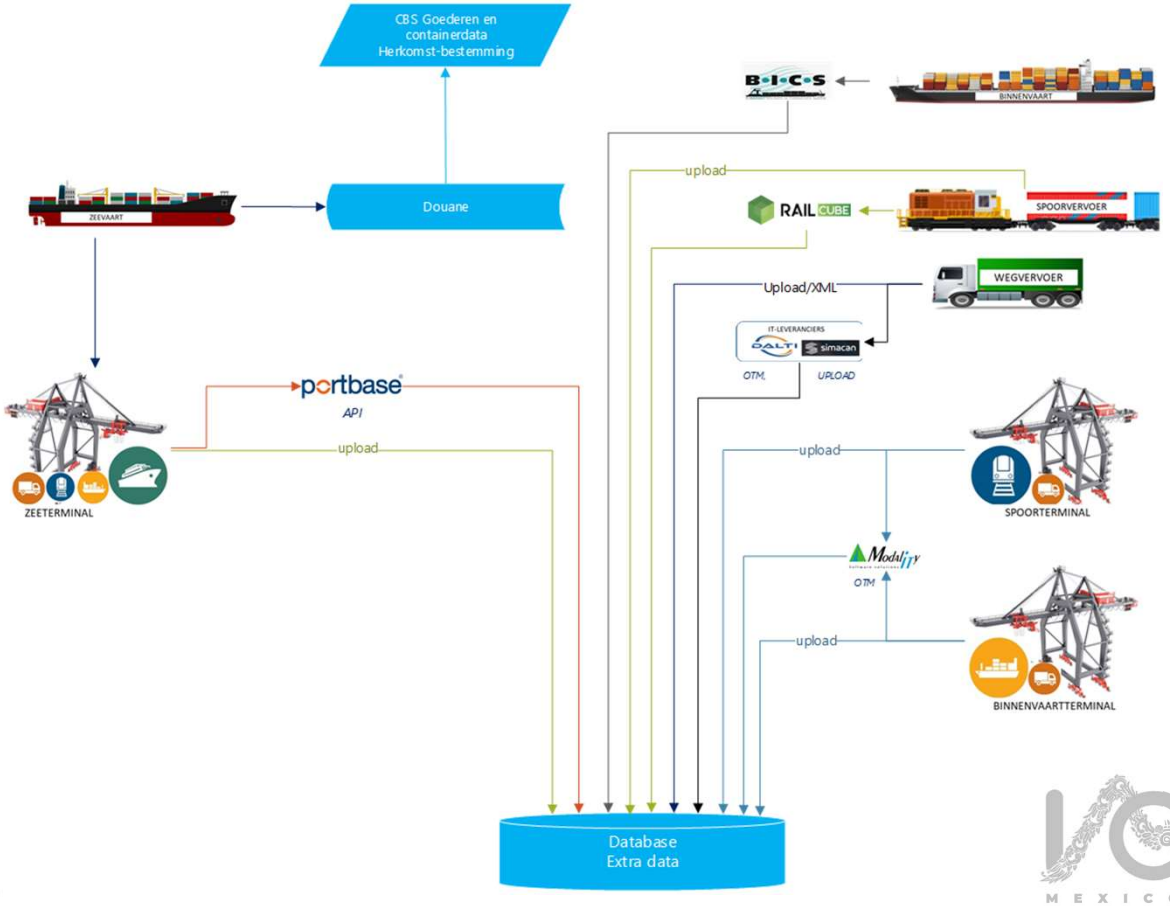
- Sea terminal
- Inland container terminals
- Rail terminal

Data collection and preparation feasible; added value!

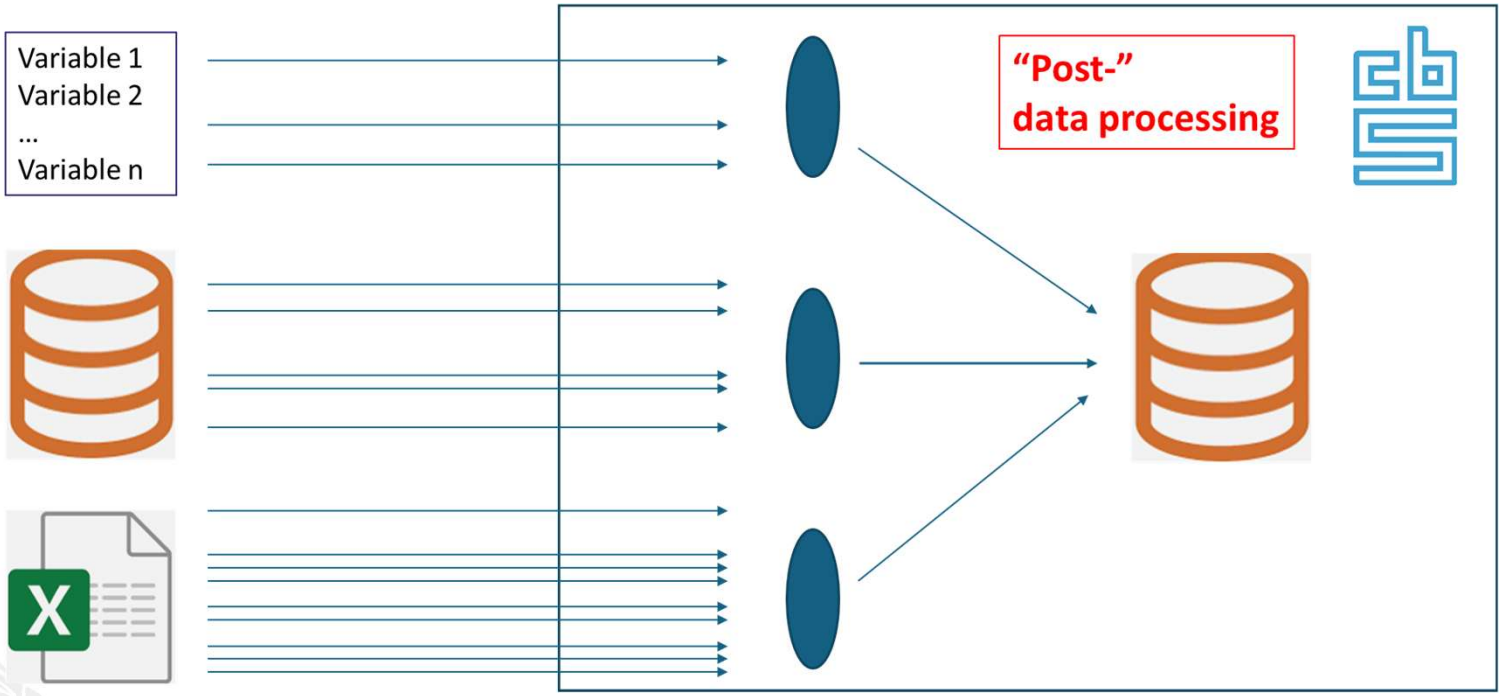
History container project

2022

Scale up



Data collection strategy



Innovative data collection

- Private data
  - Voluntary
- Open data format
  - Extract from traffic management systems
- Different formats
  - JSON
  - XML
  - Excel / csv
  - API
  - PDF

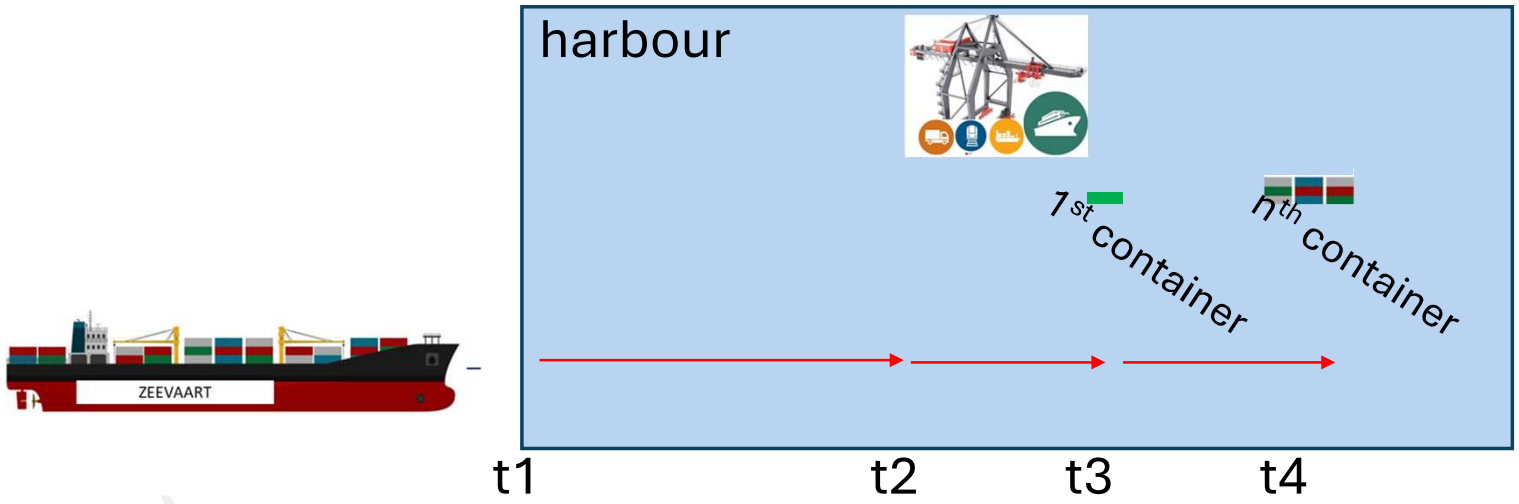
DATA PROCESSING

Challenges with input data

- Data from carriers and from terminals
- Different input data formats
- Different variables/columns
- Different level of detail

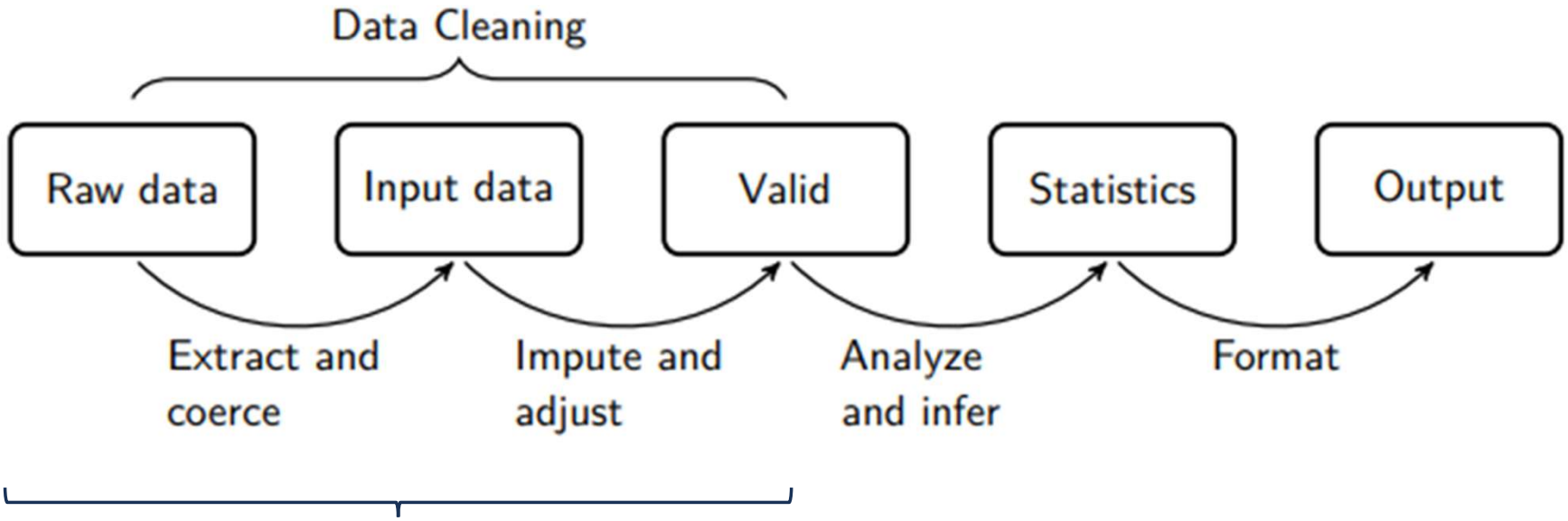
```
1  terminal,equipment_number,size,timestamp
```

Challenges with definitions



actual time of arrival?

Statistical value chain



- Harmonize data formats
- Standardize variables

Standardizations

- Standardize all variables, column names and data types.

- Locations:
  - Use geo-coding software to get lat/lon and UN/LO codes.
  - Use routing software to get the travelled distance.

Input location	Standardized
Mexico city	coordinates = [-99.07, 19.43] UN/LO = MXMEX

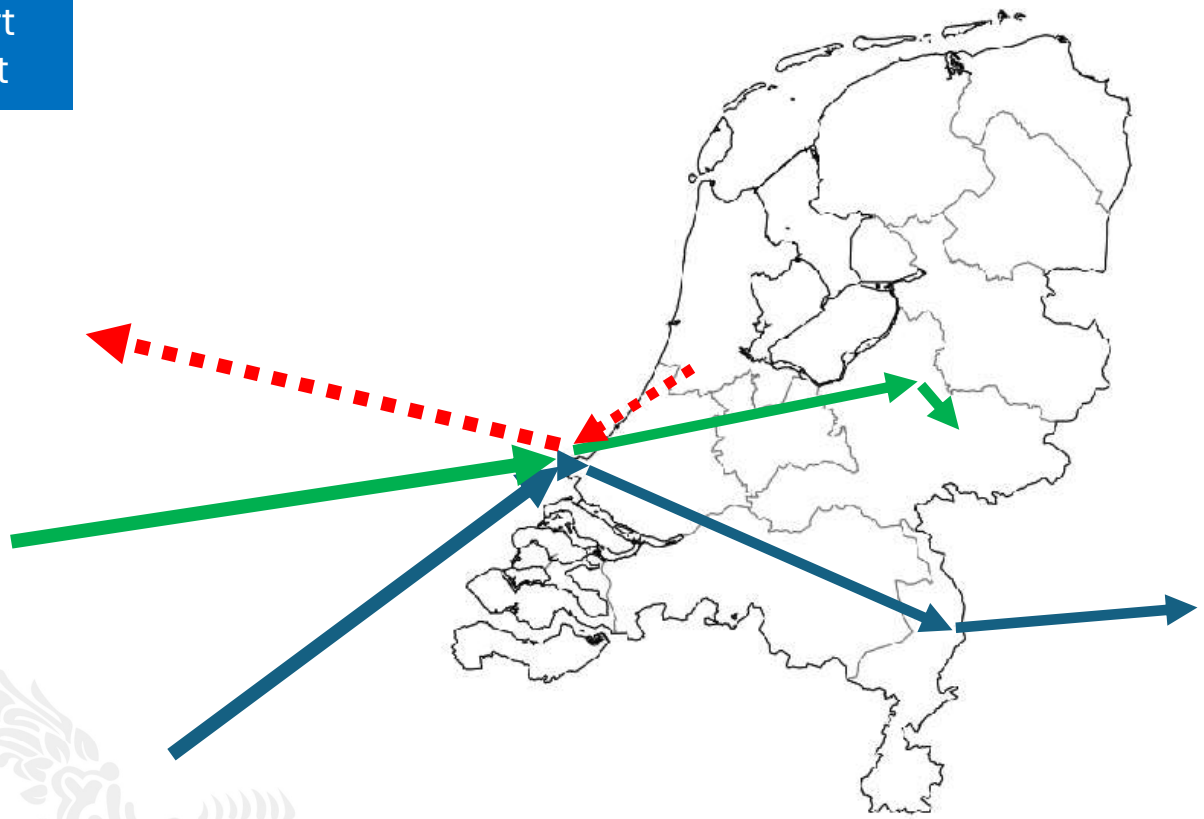
Standardizations (2)

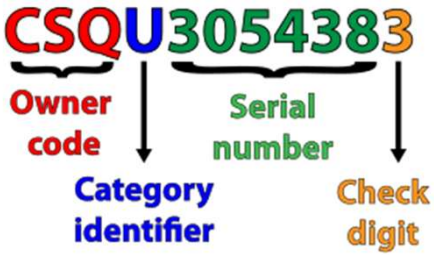
- Good descriptions:
  - Use text classification with cosine similarity to get NST2007 classification.

Raw text	Cleaned text	Classification
1764 CARTONS PALLETIZED WITH 26460 KG NET WEIGHT OF FROZEN HALF CHICKEN BREAST BONELESS SKINLESS WITHOUT INNERFILLET SALTED	FROZEN HALF CHICKEN BREAST BONELESS SKINLESS WITHOUT INNERFILLET SALTED	NST2007 = 04.1 (Meat, raw hides and skins and meat products)

Container number essential

import
export





→ Unique 'tracking' id

How to create container chains?

- A chain starts most of the time abroad (maritime/rail) and ends with road transport to the consumption location in NL or abroad (or vice versa).
- Merge all different standardized data sources together.
- Select one container number and sort by date and time:
  - If two actions are consecutive and both in the data, then this is part of the chain.
  - Missing actions can sometimes be imputed:
    - Within X hours
    - And use same UN/LO code

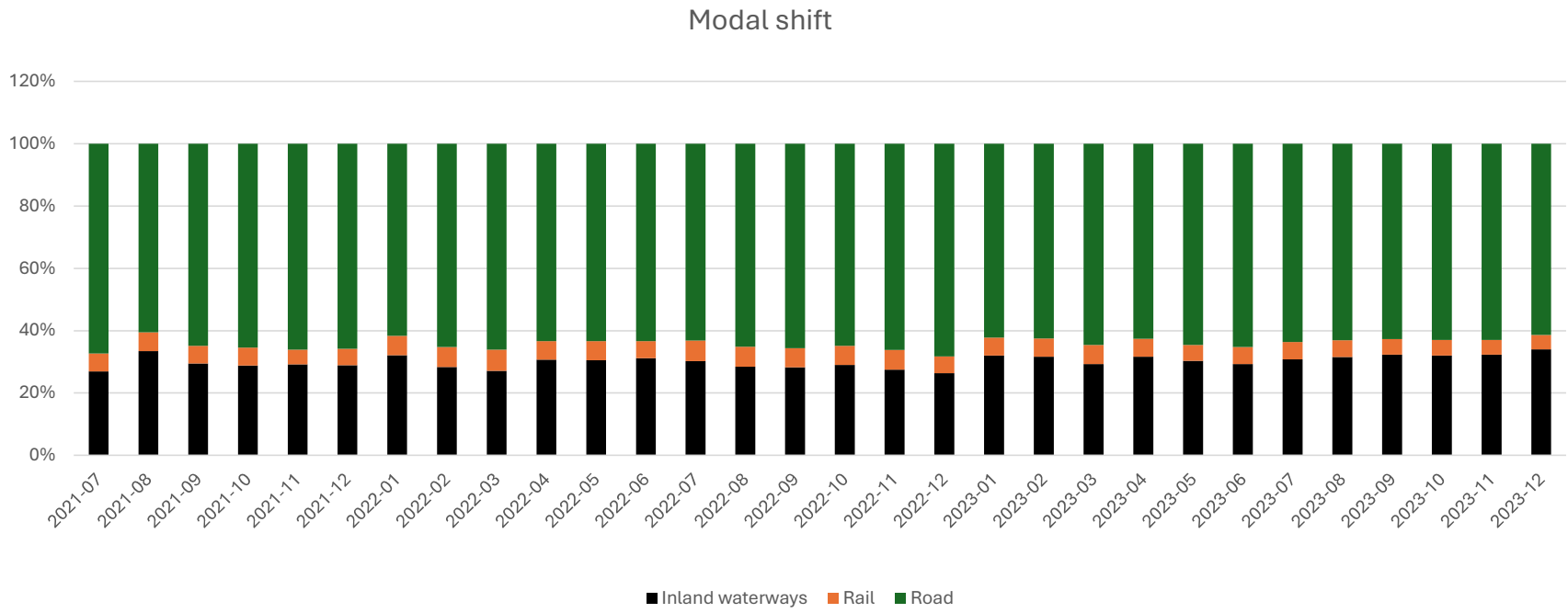
Challenges

- What to do with inconsistencies between two data sources?

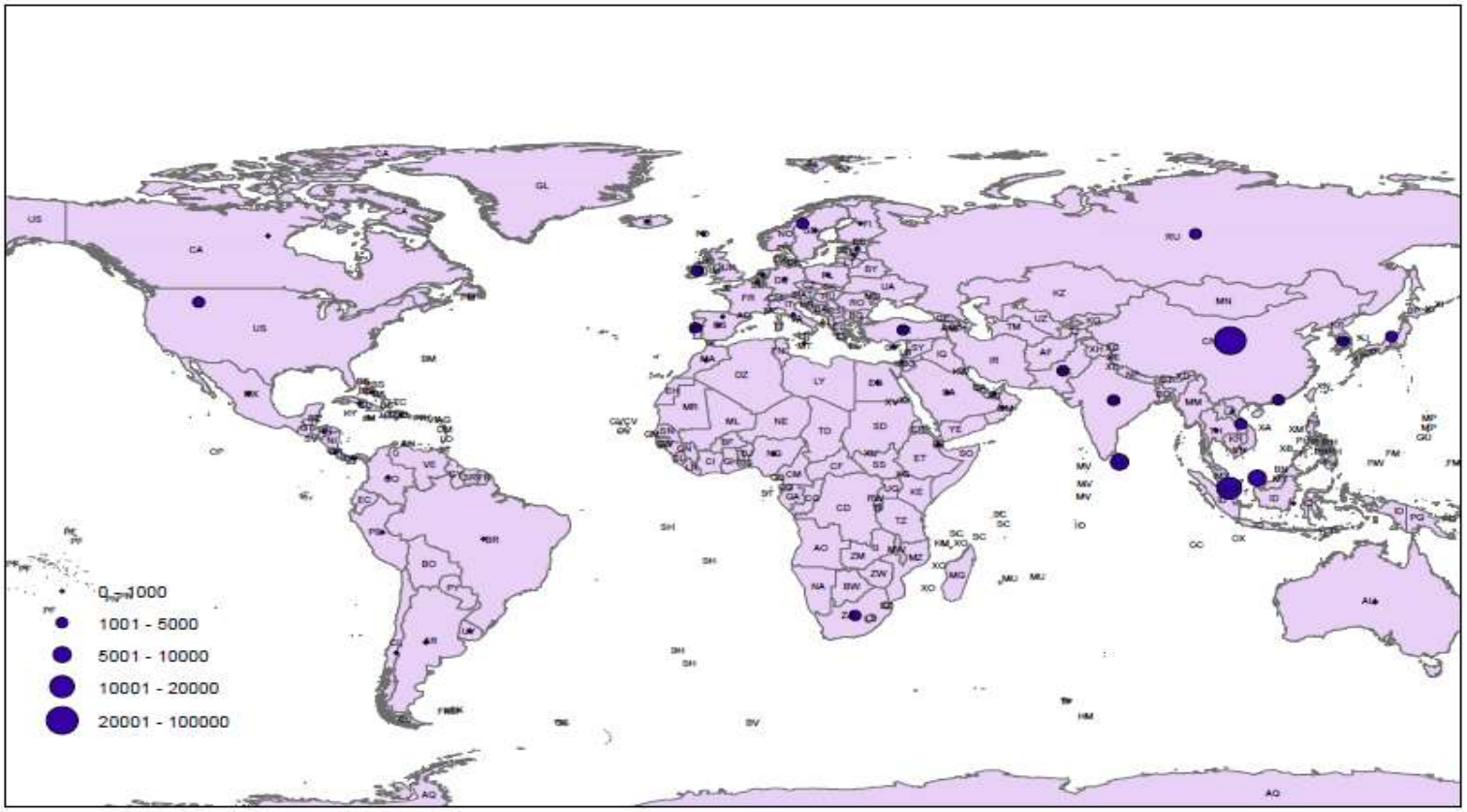
- How to create an estimation method?
  - Correct for missing input data
  - Impute missing parts of the chain
  - Using statistics per transportation mode as total
  - Work in progress

First experimental results

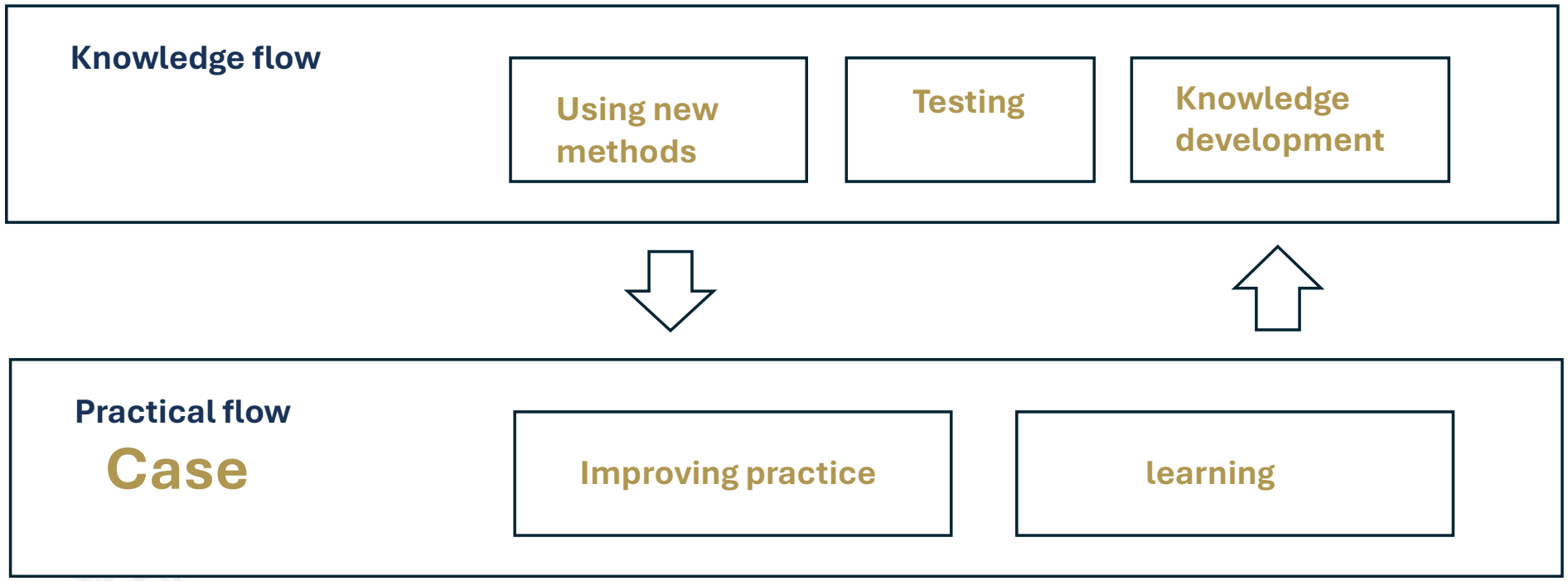
- Modal shift for containers per month from the port of Rotterdam to the inland



Future output



Developing method descriptions for the use of new and innovative data



More information: “20240516-StatisticsNetherlands-Data-Collection-of-Container-Chain-Data

Contact details

For detailed questions on the method, please contact our researchers:

- MultimodalTransport@cbs.nl

Acknowledgement

Co-f*unded by EU Grant related to call SMP-ESS-2022-TRANSPORT-STATS-IBA —New transport statistics*

Co-funded by the
European Union

IOS-ISI 2024 MEXICO CONFERENCE

IOS-ISI 2024
MEXICO CONFERENCE

Thank you

INEGI

International Statistical Institute

IAOS
IMPROVING OFFICIAL STATISTICS