Maritime and Port-Related Big Data Developments

SmartPort Lunch Meeting

Jan van Dalen
jdalen@rsm.nl

RSM Erasmus University Rotterdam

February 21st, 2017
The Erasmus Center for Data Science and Business Analytics is a joint initiative of various research groups within Erasmus School of Economics (ESE) and Rotterdam School of Management, Erasmus University (RSM). The center combines the available expertise from different scientific disciplines including business information management, marketing, econometrics, innovation management, operations research, and business strategy. Application domains of the center include marketing analytics, e-commerce, business (network) operations, supply chain management, transport & mobility, and HR & learning analytics. Moreover, the center bundles knowledge on methodologies, management and legal aspects. These research groups are currently active in the center:

- Department of Technology and Operations Management (RSM);
- Econometric Institute (ESE);
- Marketing group within the Department of Business Economics (ESE);
- Department of Marketing Management (RSM).

Some examples of projects:

- Analyze smart card data to forecast peaks in customer travel behavior and align these with infrastructure capacity (NS: the Dutch railway operator);
- Develop learning agents to optimize auction design by predicting customer bidding behavior (FloraHolland);
- Detect safety and security risks using large data sets to improve Customs' risk analyses and improve supply chain performance (EU Cassandra, Dutch and UK Customs);
- Predictive analyses of demand patterns based on (online) transaction data, making possible personalized product recommendations and contextual marketing;
- Probabilistic customer base analysis models to understand customer heterogeneity and predict customer behavior;
- Recognize financial events in news messages, and apply these in algorithmic trading (NWO: The Netherlands Organisation for Scientific Research);
- Sensor-based traffic flow prediction (CBS: Statistics Netherlands);
- Nowcasting using online search queries;
- New media analytics: design and effectiveness of (on-line) communication;
- Influence of choice architecture on online market places;
- Assessing supplier default risk from supply chain data.

Links:

http://www.erim.eur.nl/centres/future-energy-business/
http://www.kenniswerkplaats-urbanbigdata.nl/
Past analysis based on proprietary data

Recent studies based on open data

Opportunities by combining data sources

Big data and port performance

The Hype
Gartner 2015 Hype Cycle: Big Data is Out, Machine Learning is in
Big Data: the hype

- Sensor data
- Social media
- Biomedical data

- Unstructured
- Semi-structured
- Structured

- Volume
- Variety
- Velocity
- Veracity

- Data in flow
- Rate of generation and analysis
- Untrusted
- Uncleansed

Maritime and Port-Related Big Data Developments
February 21st, 2017
Many interesting underlying issues

- Business opportunities of data solutions enormous
- Organizations, companies and governments alike, struggle with data-driven decision making
- Interplay between data and strategy requires specific mindset
- Serious issues concerning ethics, privacy, organizational image

Maritime and Port-related (Big) data applications

- Past research using proprietary data
- Recent data using open AIS data
- Opportunities by combining data sources
- Big data and port performance
Past analysis based on proprietary data

Recent studies based on open data

Opportunities by combining data sources

Big data and port performance
Past analysis based on proprietary data

Large-scale collection by many parties
<table>
<thead>
<tr>
<th>Transactions</th>
<th>Data collection</th>
<th>Marketing</th>
<th>Products</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Charter market</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• O/D</td>
<td>• Volume, value, O/D</td>
<td>• Volume, O/D</td>
<td></td>
<td>Development studies, trade analysis, transport demand</td>
</tr>
<tr>
<td>• $</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Ship id</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Cargo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Oil traders</strong></td>
<td>• Ship id, port of call</td>
<td>• LMIU ship movements database</td>
<td>• Daily quotations of bunker costs</td>
<td></td>
</tr>
<tr>
<td>(arr./dep.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Bunker $</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Brokers</strong></td>
<td>• Ship id, cargo</td>
<td>• MR Inc, LMIU, Brokers</td>
<td>• MR Inc: indices;</td>
<td></td>
</tr>
<tr>
<td>volume, O/D, $</td>
<td>Ships: only brokers</td>
<td>LSE: indices, avg</td>
<td>Brokers: indices,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rates</td>
<td>avg rates, earnings;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ships: recent quot.</td>
<td></td>
</tr>
<tr>
<td><strong>Owners</strong></td>
<td>• Ship details</td>
<td>• Baltic Exchange</td>
<td>• Daily quotations from broker panels (avg</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rates by route)</td>
<td></td>
</tr>
<tr>
<td><strong>Classification societies</strong></td>
<td>• Ship details, flag,</td>
<td>• LR Fairplay, ship register,</td>
<td>• Orderbook, Deliveries, Scrappings Current fleet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>owner</td>
<td>various brokers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ship market**
- • $ |
- • Ship id |
- New built, secondhand, scrapping

**Dr. J. van Dalen (RSM/EUR)**

Maritime and Port-Related Big Data Developments

February 21st, 2017
Fixture parties

Charterer (Cargill, Krupp, etc)

Broker (MRI, Clarkson, SSY, etc)

Owner (Bergenson, Cosco, etc)
Examples of fixtures

**Erini (1982):** Coal, Hroads-Antwerp; 125000-10%. Juli2030-FIO; 3,25 days
25000tshinc (SwissMarin). 8.75$

Bulk carrier spot contract for owner Erini with a ship built in 1982. The shipment involves coal, transported from Hampton Roads (USA) to Antwerp. Total tonnage is 125000± 10%. The ship is expected in Hampton Road between 20 and 30 juli, free in and out. Loading may take 3.25 days including Sundays and holidays, at a rate of 25000 tons per day. Charterer SwissMarin agreed on a fixture price of 8.75$/ton.

**Habil Duckling (1981):** Hvy grains, River Plate-China; 55000-5%. Juli1525-FIO; 5000t (Wilmar). 34.75$

Bulk spot contract between owner Habil Duckling and charterer Wilmar to ship 55000± 5% tons heavy grain from River Plate (Argentina) to China at a rate of 34.75$/ton. Loading starts between 15 and 25 July.
Past analysis based on proprietary data

Ocean freight indices
Past analysis based on proprietary data

Ocean freight indices widely diverge.

Published freight rate indices (LSE, MRI) widely diverge$^1$

Background existing indices roughly known (size-weighted unit-value indices), but not in detail (e.g. information used)

---

Past analysis based on proprietary data

Ocean freight indices

Initial matched model (ship) indices showed very unusual patterns, prone to time window considered.

Takeaways:
- re-consideration notion of a 'product' and its associated 'price', choice of aggregation basis (route, type of contract) matters, indices diverge in periods with sparse data, role of expert opinion not always clear
Past analysis based on proprietary data

Fuel and speed in time charter contracts
Past analysis based on proprietary data

Fuel and speed in time charter contracts

- Speed loss ($\Delta S$): warranted minus actual speed
- Excess fuel consumption ($\Delta C$): actual minus warranted consumption
- Charterers can claim the owner, if fuel consumption is more than warranted, or speed is less than warranted
- A substantial amount of variation is within ship

\[ \Delta C \]
\[ \Delta S = 0 \]

\[ \Delta C = 0 \]

\[ \Delta S \]

---

Recent studies based on open data

Past analysis based on proprietary data

Opportunities by combining data sources

Big data and port performance

Recent studies based on open data
Recent studies based on open data

**Source:** https://www.marinetraffic.com/

- **AIS (Automatic Identification System)** data generated by ships using transponders, both ocean-going inland-water ships
- At scale collected and sold by, e.g. MarineTraffic, AISHub, VT Explorer
Together with Teqplay and two master students we explored the potential use of AIS data for: (i) measuring emissions in the port area as a result of inland shipping; and (ii) supporting loading decisions depending on predicted water depth.
Recent studies based on open data

Low water levels
Recent years have seen sustained periods of low water levels in inland waterways. These low water levels seriously affect transportation.

Source: http://www.thelocal.de/20161013/low-water-levels-on-rhine-river-causing-ships-to-run-aground
Apart from water supply, water levels are determined by sediment transport.

This dynamically changes water levels in the sense that underwater dunes may form at short notice.

Source: http://www.personal.kent.edu/~sclement/dynamics/rivers/rivers.htm
Recent studies based on open data

Low water levels

Low water levels affect a range of stakeholders in addition to barge operators: shippers, suppliers, customers, insurance companies.

Water levels immediately impact loading decisions of barge operators as well as the route selected for sailing.

Accurate information about water levels, or rather water depth is therefore highly relevant.

**Stakeholders in inland water way transportation (source: Cindy Zwaan, 2016)**
Recent studies based on open data

Low water levels

Water depth $D$ can be measured in different ways:

- Ship-related:
  \[ D = d_A + c_A \]

- Riverside-related:
  \[ D = H_w - H_b \]

The latter is relatively static, and does not considering changing water bed conditions.
The idea\(^3\)

- If one knows a ship’s draft \(d_A\) (which will be relatively stable while sailing) and measures keel clearance (e.g., via echo), then water depth can be measured dynamically.
- Project Covadem actually based on this idea, but has relatively small scale, and based on echo sounders.
- Instead, water depth may be measured using AIS data by exploiting the physical relation between draft, speed and water depth.
- Some problems: measuring draft, finding the precise relation between draft, speed and depth, and doing the empirical analysis.

\(^3\)Cindy Zwaan, 2016, *The effect of low water levels on loading decisions in the Northwest European inland waterways network*, Master thesis Rotterdam School of Management, Erasmus University.
Recent studies based on open data on low water levels have shown that problems partly solved by physical between speed in deep and shallow water on the one hand and ship and water characteristics on the other hand:\(^4\):

\[
\frac{V(\infty) - V(h)}{V(\infty)} = 0.1242 \left( \frac{A}{h^2} - 0.05 \right) + 1 - \tanh \left( \frac{A}{V(\infty)^2} \right)^{0.5}
\]

- Unknowns: ship’s speed in deep water \(V(\infty)\), ship’s speed in shallow water \(V(h)\), submerged midsection \(A\) and water depth \(h\)
- Once \(V(\infty), V(h),\) and \(h\) are obtained, loading decisions can be supported using the identity \(A = w_A \cdot d_A\)

---

Recent studies based on open data

Low water levels

Average speed losses of container ships and tankers by designated area (polygon), March 6th 2016

Clearly, the speed losses by tankers tend to be more volatile than the speed losses by container ships.
By way of impression, the maps below show the average speed and average speed loss of container ships on the Rhine, March 6th, 2016.
Recent studies based on open data

Environmental effects of port activity
Opportunities by combining data sources

Past analysis based on proprietary data

Recent studies based on open data

Opportunities by combining data sources

Big data and port performance
Opportunities by combining data sources

Unplanned tram stops and bridge openings
 Opportunities by combining data sources

Unplanned tram stops and bridge openings

- Unplanned tram stops
- Data about unplanned tram stops and bridge openings
Stop duration (in seconds) varies with opening duration (minutes)

For example: every opening of Erasmusbrug comes with a fixed time loss of 484 seconds (or 8 minutes) plus a bridge-opening dependent loss of 70 seconds per opening minute.

The variable costs alone have been estimated €25k per year.

<table>
<thead>
<tr>
<th>Bridge</th>
<th>Constant</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coolhavenbrug</td>
<td>1870.8</td>
<td>9.2</td>
</tr>
<tr>
<td>Parkhavenbrug</td>
<td>2115.8</td>
<td>6.5</td>
</tr>
<tr>
<td>Erasmusbrug</td>
<td>483.9</td>
<td>70.2</td>
</tr>
<tr>
<td>Lage Erfbrug</td>
<td>288.1</td>
<td>10.7</td>
</tr>
</tbody>
</table>
Opportunities by combining data sources

Emissions, traffic and the weather
Opportunities by combining data sources

- Emissions, traffic and the weather (Dublin open data)
- Emission sensors (PM10, SO2, NOx) and nearby traffic sensors
Opportunities by combining data sources

Emissions, traffic and the weather (Dublin open data)

Emission sensors (PM10, SO2, NOx) and nearby traffic sensors

Example: increases in particle matter come with increased emission gasses (not the other way around); rain reduces observed emissions, wind increases measured emissions
Past analysis based on proprietary data

Recent studies based on open data

Opportunities by combining data sources

Big data and port performance
Big data and port performance

Forecasting dwell times of inland shipping in the Port of Rotterdam

- Barge performance monitor: comparison of dwell time of individual ships with the average dwell time of the same ship over the past year
- Dwell times based on AIS data

---

Forecasting dwell times of inland shipping in the Port of Rotterdam\textsuperscript{6}

- Existing barge performance operator based on a pilot, and is backward looking
- Interest is in dwell time forecasts to support transport decisions by individual shippers (long, avoidable dwell times are a waste of time and resources)
- These dwell time forecasts are preferably based on open data
- Objective of the TNO proposal is to develop such a forecast model and assess the added forecast value of open data sources; the project’s outcome should be an easy to communicate future dwell time indicator

\textsuperscript{6}Based on recently accepted TNO proposal with the same title.
Big data and port performance

Type: ● Terminal ● Depot
### Sea side:
- number of sea side terminals
- sailing schedule
- call size
- weather conditions

### Port activity:
- number of ships at quays, sea-going ships and barges
- number of ships waiting in port area
- observed current and past dwell times
- terminal maintenance schedule

### Hinterland:
- number and types of ships approaching
- opening times hinterland terminals
- maintenance schedule hinterland terminals
- seasonal patterns
- data road haulage
Some Challenges

- Collecting the data into coherent framework may prove to be less than straightforward.
- Uncertainties about destinations and travel times of ocean-going ships and barges.
- Dwell time not a linear relation of the number of ships. Interdependence of ships’ activities in the port area may adversely affect the quality of econometric estimates (calling, instead, for simulation-based approaches).