Marine Telematics
Gerard Walls

Agenda

• Marine telematics – not pricing per se, but the first step
• Only features and hypothesised/potential risk factors and how to define them – not proven in any way
• This is a demonstration of what could be done, not about conclusions to be drawn – especially as this comes from a relatively small data set (1200 or 1300 vessels, but in some parts only on 100)
• Artistic licence taken in labelling some graphs for simplicity (without corruption of message)
Executive summary

Market Context

• Longest written class
• Most data?
• Weaknesses in rating models
• Very soft market
• Underwriter behaviour in cycle
• Global premium ~$30bn*

BIG PRIZE AT STAKE

High Level Approach

Steps followed:
- Positioning, port call and vessel metadata
- Combine datasets, cleanse and enrich
- Engineer factors which could differentiate risk
- Analyse results
- Tweak/Change engineering steps & iterate analysis
- Create visuals or outputs which are appropriate for the users

Mapping of analyses conducted to insurance perils

<table>
<thead>
<tr>
<th>Line of Business</th>
<th>Perils</th>
<th>Potential risk drivers</th>
<th>Data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hull</td>
<td>Fire</td>
<td>High mileage</td>
<td>Vessel positions</td>
</tr>
<tr>
<td></td>
<td>Wreck</td>
<td>Journey profile</td>
<td>Port calls</td>
</tr>
<tr>
<td></td>
<td>Collision / Allision</td>
<td>Coastal proximity</td>
<td>Vessel profile</td>
</tr>
<tr>
<td></td>
<td>Breakdown</td>
<td>Excessive speed</td>
<td>Incidents</td>
</tr>
<tr>
<td></td>
<td>Piracy</td>
<td>Ports visited and concentration</td>
<td>Weather history</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transparency</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Poor fuel quality</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>New / risky crew</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Passage across high-risk waters</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stationary at sea</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abandoned journeys</td>
<td></td>
</tr>
</tbody>
</table>

Note: Bold indicates coverage of PwC analysis
Results
What we analysed

More detail on approach
We have explored how data and analytics could inform marine underwriting pricing by combining and analysing three sources of data.

Data Acquisition ➔ Transformation ➔ Analysis ➔ Visualisation

Data Acquisition

- AIS
  Positional data (13 months) for:
  - 7 fleets
  - c. 1,200 other vessels
  Note 1: the 1,200 other vessels provide general comparison but were not purposefully selected and therefore may not be representative of their fleet operator

- Port Call
  (for the same set of vessels)

- Vessel Meta
  (gross tonnage, year of build and vessel type)

Transformation

- Combined and Enriched Dataset
  E.g. converting positional data to distance and speed, identifying and excluding spurious data points.

Analysis

- Mileage
- Journey Profile
- Speed
- Coastal Proximity
- Port Visit

Visualisation
Feature Engineering

**Measured Movement:**
- 1000km
- 12.5 knots

*Not actual size*

Feature Engineering

**True Movement:**
- 1200km
- 15 knots

*Not actual size*
Likely to be less differentiation in the mileage and tonnage data for general cargo ships

Different vessel types have very different profiles
Mileage

My client’s older vessels appear to be covering lower mileages

07 June 2017

Journey profile

- Loose correlation between range and mileage
- These low range vessels covering a lot of mileage. Is this higher risk?

07 June 2017
Journey profile

• More clarity on correlation when splitting by type
• These low range vessels covering a lot of mileage appear even more differential

Journey profile

• We can tell exactly where a vessel has been and how long it stayed in a single position
• We can also identify when a vessel has switched off its transponder
Further Feature Engineering

Once we have the movements across a range of vessels we can develop further features for each vessel and in relation to other vessels eg:

- Speed – average, extreme dangerous, variability
- Proximity – to shore and other vessels
- Areas visited – which ports, number of visits, length of stop, number of vessels in same port at the same time (aggregations), times in risky areas eg war zones, heavily pirated areas or busy shipping lanes
- Timing – of movements made and in what circumstances (departure/arrival/holding), fleet utilisation

Can be combined with other data – eg weather

Distinct ports visited

- Higher number of ports could mean more risk
- But need to consider mileage or distances between ports
Distance between ports visited and number visited

- Vessels with lower average journey lengths and lower average port visits may be lower risk (hypothesis)

Ports visited v journey length heatmap

- Could compare heatmaps for individual fleets v overall portfolio
- Could be used judgementally by underwriter or numerical methods could be used to create a metric
Ports concentration

- A small fleet’s ship presence at the port of Jurong Island, Singapore between January and June 2016
- Other key concentrations
- Remember Tianjin - >$3bn – some insurers hit by costs >$200m

Speed over time

- Not immediately insightful
- Indicates occasional data issue (or more likely transformation issues)
Speed over time – individual ship

- Makes more sense for individual ship
- Can identify stationary periods - or even indicate lack of utilisation (link to claims)
- Need to go deeper

Average speed over year

- Shows 95th percentile and average speed by individual ships
- Some vessels upper speeds are very close to the average (continuity of speed) whilst others show large variation
- There is a link between vessel size and speed (not really evident in this example), but important to split by vessel type
Coastal proximity and speed

1 – Lots of time near the coast (25%) but almost never at high speed

2 – Not a lot of time spent near the coast (4%), but when it does it tends to travel at speed

Proximity to other vessels and speed

Proximity event analysis determines where two ships cross paths with each other. Proximity events within a 50km radius of a port have been excluded. We have highlighted the proximity events which seem to be higher risk:

- One or both ships are traveling at high speeds of more than 20 knots (risk of an accident occurring) or whether neither are travelling fast
- Proximity can be defined at a number of distances (so not disclosed here)
- Proximity events can only be defined within the context of the database of ships you have – larger database means more can be identified. Our analysis was limited in this regard
Proximity to other vessels and speed - analysis

The top diagram shows the number of proximity events by ship type whilst the bottom one shows the number of proximity events by fleet.

Proximity to other vessels and speed - locations

Over time and with more data, it will become easier to identify unusual proximity events.
“Transparency index” - definition

The "Transparency Index" analysis shows where ships spend a period of 8 hours or more “in the dark” without transmitting an AIS message.

- This is normal practice for ships in dangerous zones, or when in dock for longer periods
- We have looked at the frequency of “dark” periods, the time spent “in the dark”, the distance travelled during the “dark” periods, and the average velocity during the “dark” periods.

“Transparency index” – gap numbers

- The rate of transparency appears similar for each of the fleets after allowing for fleet size.
“Transparency index” – time, distance, speed

Whilst Fleet D spends less time in the dark, its vessels tend to travel at a faster speed – likely down to types of vessel.

“Transparency index” – locations – single ship
**Applications**

1. **Fleet risk management**
   Position the insurer as a risk prevention partner

2. **Underwriting insight**
   Provide individual vessel risk profiles for underwriter review and interpretation
   Improve loss ratios through more informed risk taking based on vessel behavioural indicators.
   E.g., analysis of historical speed profile over the last 12 months as against similar ship types.

3. **Pricing models**
   Improve pricing models through machine learning & behavioural risk indicators
   Blend new behavioural risk indicators with traditional static ship data to create more accurate pricing models.

4. **Claims investigations**
   Challenge the legitimacy of potentially dubious claims
   Review historical vessel activity over the last year to determine vessel utilisation (example below). 
   See events immediately preceding an incident to review movements against intended destination.

5. **Non Insurance**
   Government, law enforcement & regulatory
   Use the tracking data to identify trips to sanctioned countries, or to identify potential movements of contraband, human trafficking or other criminal behaviour.

**Static ship data**
- Tonnage, age etc

**Ship activity data**
- AIS positional data, port call data

**Situational data**
- Weather, bathymetry

**Accidents data**
- Claims, details in public domain

---

**“Transparency index” – locations - fleet**

Original - Start of dark period

07 June 2017
Large datasets

Whilst this was a limited analysis on a limited data set, it was still a significantly hard exercise

• Large volumes of data – breadth and depth
• Different data sets could be added (bathymetry and weather)
• Many different ways of constructing features
• Many ways to dissect the data for analytics purposes
Machine learning

This area will be an important development for Marine Analytics

- Supervised learning
  - makes analytics easier
  - significantly reduces time
- Unsupervised learning
  - for identifying features and hence feature engineering
  - To identify trends that had not been conceived of

A machine learning approach coupled with use of powerful data will result in a significant market advantage

Changes to skills and team structures

The manipulation of large datasets is a distinct skillset. To allow a large data team to work well in delivering to pricing they need to be:

- Able to manipulate large datasets
- Flexible in response to changing demands from the pricing team
- Accurate

PwC fortunate to have diverse skills within the organisation (forensics team heavily utilised). As the role of large datasets in pricing becomes more established you may need to consider where these skills already exist in your organisation or whether you need to seek them externally
Identifying / adapting data sources

The world is creating vast amounts of data

• Most of it is not created for the purpose of insurance pricing (aggregated for sanctions compliance, created for operational purposes)
• ‘Incidental’ data can be far cheaper than data explicitly generated at an insurers request
• The owners of the data may not understand insurance
• The owners and technicians may not understand the value of the data they have created for insurers
• There is a clear and very important role here for the actuary to bridge the gap between data, technicians and business problems

Thanks

• for providing the data
• and his Forensic Analysis team for analysing the data
• and her team for spending the time to work with us on this
Next steps
Discussion around future ambitions

Questions

Comments

The views expressed in this presentation are those of invited contributors and not necessarily those of the IFoA. The IFoA do not endorse any of the views stated, nor any claims or representations made in this [publication/presentation] and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this presentation.

The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this presentation be reproduced without the written permission of the IFoA or authors.