Review of Methodologies for Predicting Future Vessel Traffic in the Northern Shelf Bioregion, British Columbia for Transport Canada's Cumulative Effects of Marine Shipping (CEMS) Initiative

Final Report

November 22, 2021

Contact:

Ronald Pelot, Ph.D., P.Eng.

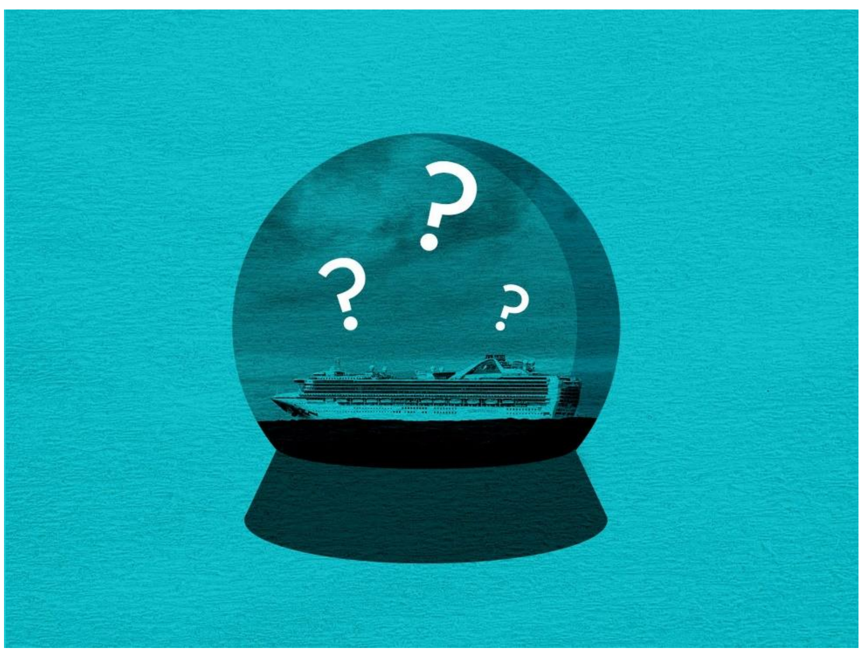
Ronald.Pelot@gmail.com

(902) 489-0804

Prepared for Transport Canada under Contract: T8080-200645

“***Prediction is very difficult, especially of the future***”

Niels Bohr



Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank Jessica Cucinelli who assisted with the search for relevant literature and summarized key elements for each one. She also helped with the reference formatting. I appreciate the contributions from all of the people that I interviewed, who gave willingly of their time, knowledge and opinions regarding marine traffic predictions. I am very grateful to Sam Mansfield from Transport Canada who conducted a detailed review of the draft report resulting in many useful suggestions and edits.

Glossary

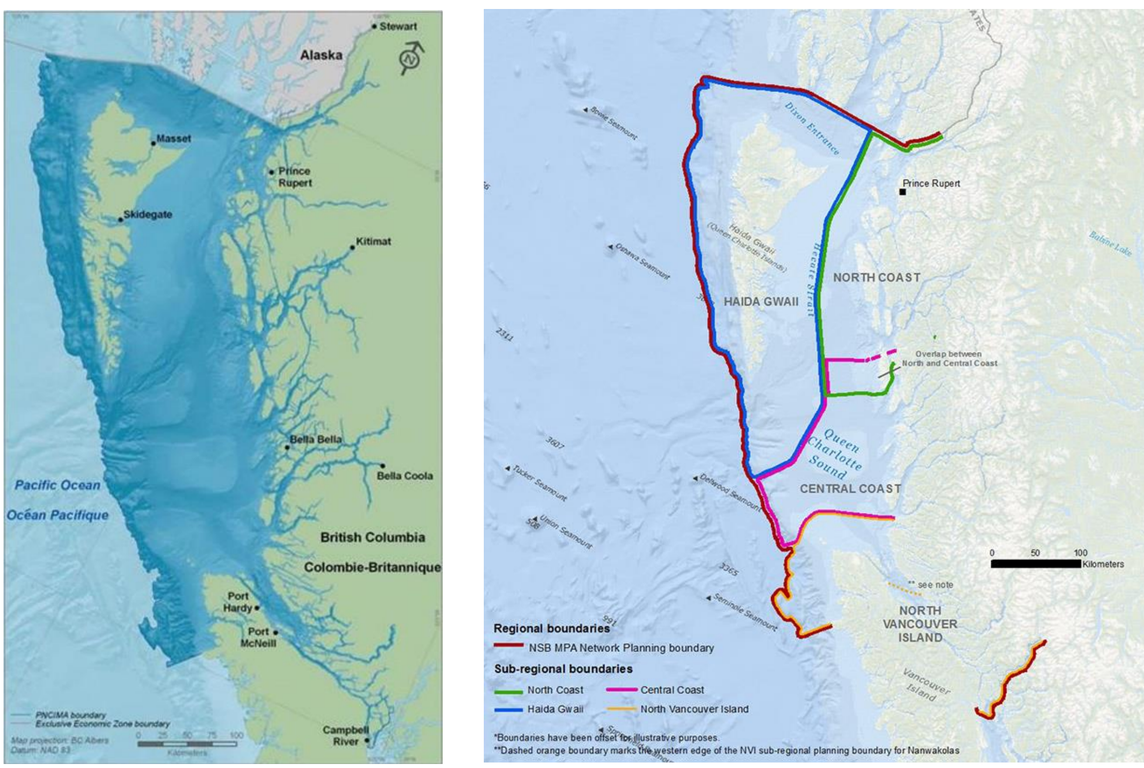
AIS	Automatic Identification System
ANN	Artificial Neural Network
AOPS	Arctic Offshore Patrol Vessels
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
BC	British Columbia
BPNN	Back propagation Neural Network
CAGR	Compound Annual Growth Rate
CCG	Canadian Coast Guard
DFO	Fisheries and Oceans Canada
DRDC	Defence Research and Development Canada
EU	European Union
GDP	Gross Domestic Product
GHG	Greenhouse gas
GIS	Geographic Information System
GSA	Gravitational Search Algorithm
GT	Gross tonnage
IMO	International Maritime Organization
ISIC	International Standard Industrial Classification
LOA	Length Overall
LOH	Length of hull
LRIT	Long-Range Identification and Tracking
ILO	International Labour Organization
IMF	International Monetary Fund
IMO	International Maritime Organization
IPCC	Intergovernmental Panel on Climate Change
LSSVR	Least Squares Support Vector Regression
MAE	Mean Absolute Error

MAPE	Mean Absolute Percentage Error
MARCS	Marine Accident Risk Calculation System
MC	Monte Carlo
MCTS	Marine Communications and Traffic Services
MPA	Marine Protected Area
MSE	Mean Squared Error
MSP	Marine Spatial Planning
NASP	National Aerial Surveillance Program
Nmi	Nautical mile
NOAA	National Oceanic and Atmospheric Administration
NSB	Northern Shelf Bioregion
OD	Origin-destination
OPRF	Ocean Policy Research Foundation
PA	Port Authority
PMV	Port Metro Vancouver (now VFPA)
PNW	Pacific Northwest
QRA	Quantitative Risk Assessment
RBF	Radial Basis Function (ANN)
RBT2	Roberts Bank Terminal 2
RMSE	Root mean squared error
ROMS	Regional Ocean Modeling System
RO-RO	Roll-on/roll-off (ship)
SARIMA	Seasonal autoregressive integrated moving average
SDM	System Dynamic Modelling
SOLAS	Safety of Life at Sea
SQD	Search Query Data
SRES	Special Report on Emissions Scenarios
SSP	Shared Socio-Economic Pathways
STARMA	Space-time autoregressive moving average
STL	Seasonal and Trend decomposition using Loess
TEU	Twenty-foot equivalent unit
TSB	Transportation Safety Board
VC	Valued components
VFPA	Vancouver Fraser port Authority
VTOSS	Vessel Traffic Operation Support System
WTO	World Trade Organization
ZOI	Zones of Interest

1. Introduction

1.1. Purpose

The aim of this project is to describe various approaches to predicting marine shipping traffic for different contexts and objectives, and in particular to assess which method(s) is most suitable to consider future shipping activity in the Northern Shelf Bioregion (NSB) of British Columbia (Figure 1). This work was completed for the Cumulative Effects of Marine Shipping (Transport Canada, 2020) initiative under Transport Canada's Oceans Protection Plan. The CEMS initiative consists of six pilot sites across Canada's three coasts, including the NSB, where work is guided and completed in partnership with interested Indigenous governments. The review provided in this study will serve to guide the methodology applied to predict future shipping patterns, which in turn will help inform policy, management, and infrastructure development in the region. While it is unavoidable to discuss mathematical concepts when presenting prediction methods, this report is primarily aimed at those interested in learning about the range and nature of available forecasting methods, but without most of the details that a modeller would need (which are usually available in the original articles and reports).



(a) (b)

Figure 1. The Northern Shelf Bioregion: (a) overall boundary (source: CPAWS British Columbia, n.d.); (b) Northern Shelf Sub-Regions (source: MPA Network, n.d.)

1.2. Background

Marine vessels inevitably affect their environment in a wide variety of ways. For example, depending on the vessel type and activity, they can cause noise, air emissions, spills, disturbance to marine life, interference with coastal activities of Indigenous communities, etc. Cumulative Effects Analysis (CEA) is a "systematic process of identifying, analyzing, and evaluating changes in environmental, cultural, health, social and economic conditions, caused by multiple interactions among human activities and natural processes, which accumulate across time and space." (Transport Canada, 2020) Consequently, the initiative is examining the impacts of shipping on various environmental, social, and/or cultural valued components (VCs) prioritized by First Nations involved in the NSB pilot site work. 'Valued' implies importance with regard to the ecosystem and also to the value that people place on the affected environment, whether for social, cultural, economic, historical, archaeological, or aesthetic reasons.

Examining shipping impacts on VCs involves many elements: shipping traffic measurement and/or modelling by type, location and time; the potential stressors that derive from vessels depending on several characteristics, such as ship size, age, or operating patterns; the location and sensitivity of the VCs in the vicinity of the shipping activity; and the pathways through which the vessels can adversely affect the VCs. Furthermore, 'cumulative' implies over time, over all vessel activities, and across combined stressor types (noise, pollution, etc.). Finally, since a CEMS outcome is to help inform multiple stakeholders and rightsholders by identifying and where possible quantifying shipping impacts, part of the purpose of examining shipping impacts is to guide future actions which would mitigate the impacts. Thus it is important to predict the shipping activity and impacts in each study region well into the future for two reasons: (i) to determine the rate at which certain key effects will accumulate, which will help prioritize mitigation efforts; and (ii) some actions (such as regulatory changes or infrastructure developments) take many years to implement, so it is preferable to consider the effects of future traffic, not just present patterns.

As shown in this review, there are several prediction methods that have been applied to forecast shipping traffic, and their applicability depends on many factors such as the aim of the forecast, vessel type(s), time horizon, data availability, and other factors.

1.3. Approach for this review

This review aims to provide a broad review of shipping forecasting studies, with a focus on the methods that were used to perform the predictions. It is meant to be a general introduction to shipping traffic forecasting with emphasis on the applications, and an overview of methods and how/where they were applied, but not geared towards modellers/statisticians. That being said, some formulas are included periodically to give a sense of how factors were combined to achieve a particular aim. Furthermore, the principal focus of this report is on forecasting methodologies, not outcomes of particular studies, although these latter are included in some of the descriptions (section 2) and case studies (section 4) to illustrate the nature of the solutions to the prediction problem.

The information gathering for this review was a three-part process: search for articles in the academic literature; acquisition of government and consultants' reports; and interviews with a dozen knowledgeable individuals from various sectors in Canada. As shipping-related data, modeling tools, and modeling expertise have advanced significantly in recent years, most of the literature was restricted to the last decade unless a particularly useful article was found that predates that.

The literature is reported in 4 ways: a brief summary of each report/article that was included; summary tables of all the articles; a few extended cases studies; and the reference list.

The following guidelines were followed in preparing this review:

- Provide sufficient background, definitions, and descriptions of methods in the summaries to allow readers to appreciate the approaches covered;
- The literature review is broad but not comprehensive as there is a large volume of literature on this topic, with some overlap of methods;
- It is mostly limited to articles/reports in the last decade;
- Case studies are included to provide more detailed examples of forecasting methods, and in particular some applications to the west coast of Canada. The case studies provide:
    - High-level overview of the assessment (e.g. project details, objectives for forecasting vessel traffic);
    - Intended geographic scope (i.e. global, national, regional, and/or local scope);
    - Intended temporal scope (e.g. 5, 10, or 20 year+ forecasts);
    - Vessel type(s) considered;
    - Identification of key drivers of vessel activity that affect marine traffic in the assessment area and were considered in the forecasting (e.g. trade routes; wealth in source and destination regions; specific commodity demands);
    - The data sources that were used (e.g. historical vessel traffic information, commodity trends) as well as highlight any uncertainties with data inputs or known data gaps;
    - Description of models and/or simulation tools used;
    - Output format in textual, tabular or visual form (e.g. % increase/decrease, forecasted nautical miles sailed, qualitative forecasts [i.e. low, medium, and high confidence levels], etc.);
    - Future iterations or ongoing efforts to improve the described methodology.

The report is organized as follows:

- To position the review, some fundamentals of marine shipping prediction are presented, including the purpose of such exercises, and some essential concepts and terminology.
- Then each of the methods to developing predictions are explained in sufficient detail for the reader to be able to follow the subsequent review of the literature.
- The literature review provides a synopsis of each article/report, grouped according to the type of problem being addressed as that is often a key determinant in the method(s) that is applied.
- The summary table then provides key features of all the reviewed articles in a concise format.
- Five case studies are provided, with more elaboration on how the prediction methods were applied and that the outcome was.
- The summary provides some final comments and insights gleaned from the review.

1.4. Fundamentals of marine shipping prediction

The general concept of predicting marine traffic is fairly self-evident, but there is a large number of details and nuances that are important to consider when selecting the prediction method, setting the parameters, making the assumptions, and defining the most useful output variables and format. The following sections describe the relevant elements in sufficient detail to be able to understand and appreciate the methods and applications that are subsequently presented from the literature.

1.4.1. Purpose of traffic predictions

Fundamentally, there are two main purposes for predicting marine traffic: to understand and/or to influence. We seek to understand what the characteristics of future shipping traffic may be as measured by a broad set of possible attributes such as how much traffic, where, when, what vessel type, possible impacts, etc. Many of the articles described in the literature review fall into this category. Of course, it is implied that by creating a better understanding, such insights could be used by those who may have some say about managing the outcomes. An example would be to envisage what the vessel traffic in the Canadian Arctic might look like 10-20 years hence, without a specific goal but rather to inform all those who have a vested interest, whether government, industry, or Indigenous peoples (see Case study C – Future marine traffic activity in the Canadian Arctic: A Scenario-based approach). However, many articles are also aimed at supporting specific decisions or policies. For example, a large portion of the container ship traffic models are targeted at port capacity expansion planning, while others, such as those focused on greenhouse gas (GHG) emissions from ships, lay the foundation for mitigation measures through regulatory changes. In no particular order, some of the applications of ship traffic prediction models are: efficiency of maritime supply chain; port infrastructure development; ship-building investments; marine policy planning; managing environmental impacts; safe navigation; fisheries management; and emergency response resource planning. Articles and reports in all of these areas are included in the review.

One useful perspective is to characterize the issues according to the types of risks to be addressed: risks TO ships, risks FROM ships, and risks ON ships. Shipping traffic studies and predictions can serve to assess current or future hazards TO shipping, such as the odds of ships becoming icebound in the north as the uncertainty of navigation in that environment grows. Many publications on shipping forecasts address the impacts FROM ships that may affect their environment, which is the focus of this review. Finally, shipping traffic patterns can provide a basis for exploring spatio-temporal occurrences of problems ON ships involving equipment failure or crew health and safety concerns (i.e. where and when they may occur).

Another informative viewpoint is to consider the nature of the decision which a prediction modelling exercise is meant to inform. A forecast can be developed to assist with exploring alternative futures, commonly known as 'what-if analysis'. However, there are in fact two types of what-if analysis depending on whether it is referring to things that are within our control or not. For problems that are within our control (i.e. decisions), the forecasts are used to explore issues such as whether to proceed, when to proceed, how much to proceed, what the future implications are (ex. costs/benefits). A port capacity expansion is a good example of this. The other type of 'what-if' considers the relationships between shipping traffic and unintended consequences, such as invasive marine species being transferred from one region to another via vessels, which can change in the future due to shipping volumes, route changes, and changing environmental conditions. Another type of circumstance for which shipping predictions can be useful involves the precautionary principle whereby if the harm that may occur if an action is taken cannot be assessed and may lead to grave consequences, then that action should be avoided. A prediction model can show whether the consequences can be foreseen with reasonable accuracy and reliability, and if not perhaps the decision is made to not proceed with the intended plan.

Models in general, including forecasting, can sometimes help assess whether a risk is understood and manageable or not. For example, potential future shipping traffic growth in an area may engender risks that are socially unacceptable, and quantitative analysis may help ascertain the actual degree of risk posed by the new activity. Another interesting application when considering the future is for backcasting. It's a method devised by Robinson (2003) in the 1980s to specify a desired scenario in the future, to help planning exercises to achieve that outcome. For example, a forecast model of shipping in the arctic might predict the amount of GHG emissions at some point in the future, barring any technological or regulatory changes, and if the policymakers would like that level of emissions to be only 10% of the predicted values, then the backcasting process can examine options to achieve that result.

The apparent aim of applying a prediction model to shipping traffic is to consider the future outcomes for the vessel activity, denoted as element (A) in Figure 2. As described in section 1.4.4, if the shipping activity is extrapolated from the historical patterns, then the concept is simple (even if the math models are complicated). Often, the future activity will depend on 1 or more drivers (B in the figure), such as container ship traffic being proportional to the Gross Domestic Product (GDP) at the source and destination regions. Then we are in fact predicting the drivers (B) over time and determining the activity level (A) as a function of the drivers through their relationships, denoted as multipliers (C) in the figure. Sometimes however, one can also predict that the relationship between the drivers (B) and the activity (A) are also subject to change over time, so we also need to predict variations in (C) into the future. For example, if the issue at hand is the impact of shipping on the environment (or VCs), then we may model future shipping patterns but assume that the "environment" that it is impacting is unchanged. However, sometimes the environment could change, and that has to be predicted as well. For example, the impact of shipping on cetaceans (through disturbance or strikes) in the future could depend a lot on changes in cetacean populations and/or spatial distribution that are different than the present. Hence, we can consider predicting changes in (E) as well. Sometimes the pathway or mechanism by which the hazard associated with the shipping impacts the environment may be expected to change as well (D). For example, if invasive species are transferred through ballast water exchanges, and the future technology or regulations change such that this process risk is largely mitigated, then we may want to include a prediction about changes to that pathway (D) in the future. Finally, any prediction assumes that a large number or "external" factors (F) other than those that impact the activity (B) directly are relatively unchanged over the forecast horizon. However, longer term predictions, such as climate change scenarios, must also project any potential significant changes in some of these "external" factors (F) to see how it affects any or all of (A), (B), (C), (D) or €. Thus 'shipping prediction' must often take a holistic view of all relevant options, particularly as the timeframe lengthens. All of the elements in Figure 2 are found scattered across the papers reviewed in this report.

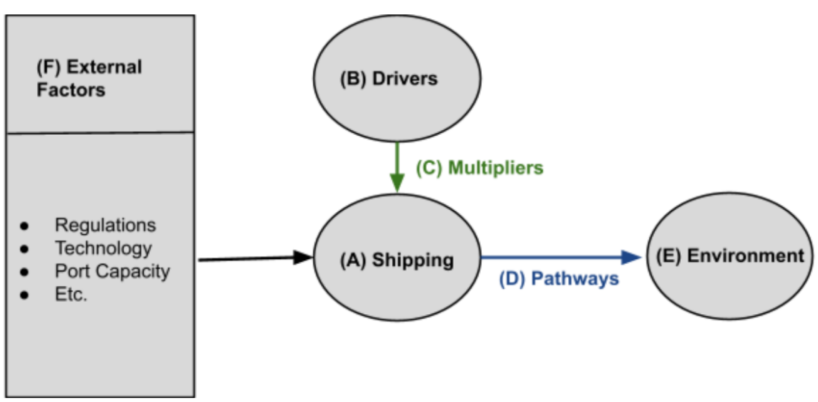


Figure 2. Complete model elements for future shipping impacts

1.4.2. Essentials of prediction approaches

The following terms are sometimes used interchangeably in general parlance, but they must be distinguished when generating models about the future.

a) A ***prediction*** is a calculation about a potential future outcome based on current data and relationships between the important factors (also called variables) underlying the activity (also referred to as a causal forecasting). A prediction usually assumes that any future changes in such relationships are small enough to not significantly change the modelling outcome. For example, if the number of container ships to a given region have been found to be proportional to the Gross Domestic Product (GDP) of the region, then estimating the number of container ships calling at a given port there in 10 years could be based on how many ships typically have been using that port in recent years and what the GDP is expected to be in a decade. The actual relationships for this situation can be more complicated than this, as illustrated in the methodology section and case studies below, but this clarifies that a prediction is essentially a model based on data and relationships that are known today. It is also possible to predict future changes in the relationship itself between the causal variable(s) and the predicted variable.

b) A ***forecast*** relies on a series of historical data points on a particular phenomenon (such as ships traversing a given area annually) and predicts the future based on trends and patterns in that data (ex. see Figure 3). Trends could be upwards or downwards, it could be a linear change every year or a percent change for example; seasonality could be driven by seasonal activity (literally, as in summer vs winter), periodic conditions (ex. annual fisheries openings), or weekly patterns (ex. weekends versus weekdays). This time series forecasting is therefore a particular subset of prediction methods which extrapolates the historical activity into the future, but does not explicitly try to determine the underlying factors driving the activity itself, nor changes to them.

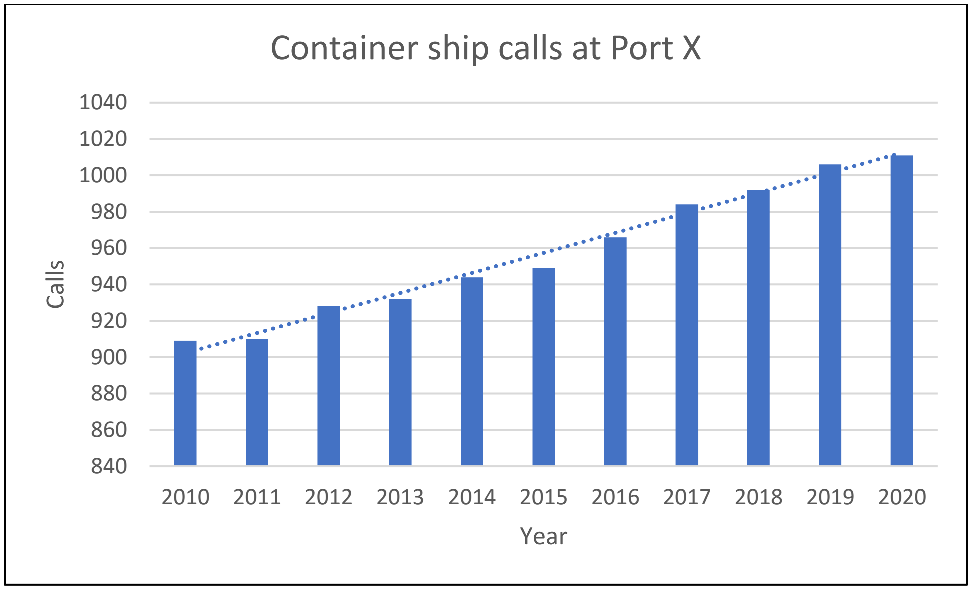


Figure 3. Hypothetical example of container vessel traffic trend at a given port

c) A ***projection*** is a broader concept about envisaging what the possible future could be for an assumed set of conditions. Often this is done through scenario development, where each scenario corresponds to an estimated or chosen set of possible future conditions, and multiple scenarios can then be compared and contrasted for planning purposes. For example, the volume and nature of shipping in the Canadian Arctic in fifty years depends on many highly uncertain factors, including the diminishing presence of sea ice due to global warming, permits being issued for major resource extraction, shipping governance arrangements, etc. Various combinations of such factors can be considered to help plan for marine infrastructure development or risk mitigation measures, for example. The most common use of projections is to assess the implications for several plausible scenarios, but this approach can also be used to contemplate more extreme outcomes, or "worst case" scenarios. One particular version of scenario-building is to detail future shipping traffic outcomes for a specified development, such as a planned port expansion. Thus it is essentially assuming that the new port call capacity would be reached at some point, and the modelling serves to investigate potential impacts of the associated traffic growth.

Predictions (a) and forecasts (b), both using data and models, generate the most likely outcomes for the given assumptions, and thus are special cases of projections (c). A projection is sometimes produced through a combination of numerical modelling and expert prognostication. However, it is more often generated by several experts and/or stakeholders jointly conceptualizing what the future may hold, anchored in current and historical phenomena and perhaps potential new driving factors that can affect the outcome. These hypothetical factors could be deliberate, such as considering the impact on shipping of a new policy, or alternatively speculating on unexpected shifts in the conditions, such as a possible collapse of a fishery, or a global pandemic affecting shipping patterns. Figure 4 shows in simple terms how these three methods are related. Forecasting is based simply on the historical data for the shipping activity, while predictions are based on that as well as the time series of one or more causal factors, and projections can encompass prediction or forecasting models as well as expert-based scenario generation. These methods will be explained in greater detail in section 1.4.4. While projections based on scenario-building is addressed in this report, the main methodological focus is on model-based predictions, including forecasting.

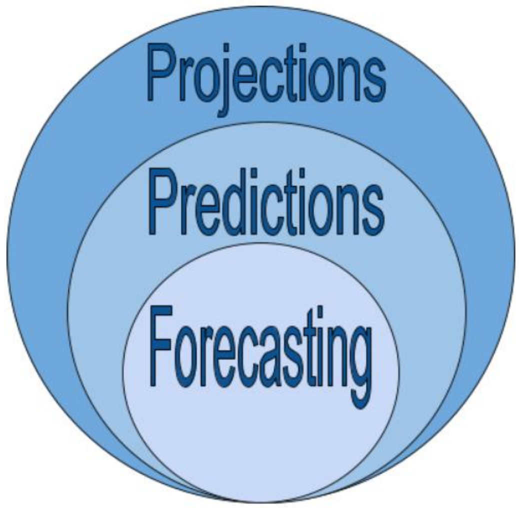


Figure 4. Relationship between projecting, predicting, and forecasting

1.4.3. Elements required to perform shipping predictions

A forecasting system should follow these basic steps.

1. Determine the use of the forecast
2. Select the items to be forecast
3. Determine the time and effort needed to make a forecast
4. Select the appropriate forecasting method (Note: the remainder of this section describes the elements needed to complete this step in more detail)
5. Gather the data and parameters needed to make the forecast
6. Make the forecast
7. Validate the results
8. Implement the solution(s)

The choice of a prediction method, as well as decisions about the inputs and outputs to the process, depend on many aspects, as described next.

- ***Time horizon***: One of the most important aspects is the choice of endpoint for which we need the projection, often referred to as the time horizon. For the purposes of this report, they will be (somewhat arbitrarily) categorized as follows, with examples provided:

  a. *Immediate predictions*, ranging from minutes or hours ahead to days. This applies to real-time tracking of vessels, typically by the military or for navigation, to detect any unusual patterns relative to the anticipated behaviour. It can also be used for dynamic routing decisions to deal with hazards such as icebergs, storms or other vessels. This type of modelling falls outside the scope of this report but it has some elements in common with the longer time horizons, including data needs and types of prediction models.

  b. *Short term predictions* ranging from a few weeks to a year. Such models are mostly aimed at operational planning, such as expected traffic at a port or through a constrained waterway for a given time period. Other short term assessments dependent on such information could be estimating the cumulative effects of shipping in an area over a coming season or year, or prepositioning Coast Guard assets for more effective emergency response.

  c. *Medium term predictions*, for roughly 5 to 20 years ahead. This is the most common time horizon for shipping prediction modelling because it is needed for most infrastructure development projects and policy development. New terminals, upgraded facilities at ports, expansion of land-based links for multi-modal supply chain expansions, are examples of medium-term planning that rely on good future shipping traffic estimates.

  d. *Long term predictions*, on the order of decades. A very distant perspective lends itself to scenario-based projections as numerical predictions would have very high uncertainty as the timeframe lengthens, and many new factors would come into play from future technological, economic, environmental and/or governance changes that would invalidate the current predictive models. Long term projections are largely reserved for two applications: (i) major strategic decisions about infrastructure development such as expanding the Panama canal, planning for significant shipping traffic through the Arctic, building a new shipyard, etc. (ii) studying the continual impacts to or from ships over the long term, such as climate change impacts on shipping, or the cumulative effects of shipping on air emissions or underwater noise levels over decades.

Other aspects of the modelling that depend on the duration include the period length of historical data used in the model, and the degree of uncertainty as the projection's timeline lengthens. Such issues are discussed separately below.

- ***Categorization of vessel activity or type***: The aim of some shipping traffic predictions is to capture only one type of vessel or activity. For example, many studies focus on container shipping for the purpose of terminal capacity expansion planning (ex. see Case study A – VFPA – RBT2 Container traffic). For other applications, such as vessel noise impacts in an area, the intent would be to include all ships, and perhaps smaller vessels as well (fishing, recreational) in the model. In this case, prediction models would have to be developed for each sub-group of vessels because the drivers and parameters for activity levels are different for distinct vessel types or activities, and often their respective impacts must be estimated separately. For example, the noise generated by ships depends on many vessel characteristics, but studies have shown that these can be aggregated according to several broad groupings such as vessel type (ex. container, bulk carrier, ro-ro, ferry, tanker), a measure of size (length overall, or gross tonnage), and so on (McKenna, 2013). Consequently, for example, a prediction for shipping noise in a given region would have to forecast container ship traffic by distinct size groupings because of changes in the relative abundance of container ships per size grouping (i.e. an increase in the abundance of large container ships may correspond with a decreased abundance of smaller container ships).

  The principle here is that if the forecasting objective involves more than one ship type or relevant vessel characteristic, then multiple prediction models must be applied and their outputs combined as needed to produce the overall impact.

  Some key ship characteristics that are often used to group the data are described here.

  - *Vessel size*: There are many ways to represent a vessel's size, so only a sampling is provided here to illustrate the variety. Length Overall (LOA) is a common measure which encompasses any protrusions from the bow or stern of a vessel, versus Length of Hull (or LOH). The dimensions of a vessel (in meters) are also contained in AIS (Automatic Identification System) transmissions from ships. Gross Tonnage (GT) is based on a nonlinear function of a ship's total interior volume. A container ship's capacity (and hence indirectly its size) is measured in in twenty-foot equivalent units (TEU), which is an approximate measure of how many containers it can carry. The draft of a ship is an important measure for many reasons, in the short term for navigation purposes and in the long term for channel design or dredging.

  - *Vessel type*: As mentioned above, most shipping predictions are performed for a given vessel category. However, another distinction to be made is what should be encompassed within a general category. For some studies, smaller categories may be necessary, subject to the general limitations described below. For example, a prediction of commercial fishing vessel activity in a specified region in several years could be based on current traffic levels and trends of all fisheries in the area, or perhaps the expected changes in bottom trawling, scallop fishing, and lobster fishing are expected to be quite different and should be forecast separately. Similarly, 'growth in recreational boating' does not distinguish between yachts, cruisers, or those intended for watersports or sports fishing, each of which may have different impacts on various VCs.

- *Vessel age:* While age is not a causal factor for shipping impacts, it is a commonly used surrogate measure of ship design and/or condition. As an easily available measure for most vessels, some prediction models may use age to distinguish between older vessels that may be more prone to incidents due to outmoded technology or poor maintenance, and thus for example likely to cause greater impacts with respect to pollution.

- *Vessel equipment and design:* Many environmental impacts of shipping are conditional on the specific vessel equipment and design. For example, several characteristics of ship design, including hull, engine, and exhaust system design, can significantly reduce various noxious gases from any air emissions. While most studies conducted on the broader issue of shipping traffic do not delve into such details, for longer term forecasting if significant changes are expected in a relevant ship characteristic, then this aspect should be accounted for in the prediction. For example, limitations on sulphur content in shipping fuel and trends towards adopting more closed-loop scrubber systems may change the nature of the pollution impacts in the future.

- *Vessel operating area:* A vessel that operates in international waters would be subject to some different regulations than one with strictly domestic movements. This could have ramifications for the likelihood of adverse incidents or the nature of the consequences therefrom.

- *Other considerations wrt ship grouping:* Ships may be grouped by many other attributes depending on the study's aims. Examples are:
    - Cargo/commodity Type
    - Direction of movement: export, import, regional transit and internal transport
    - Ship flag
    - Motorized vs non-motorized (for pleasure craft)

To summarize, the grouping of vessels by one or more characteristics to perform meaningful predictions depends on many factors, including:

- The aim of the prediction;
- Data availability (historic or current traffic levels; vessel characteristics, information on activity drivers (economics; port capacity, etc.));
- The ability to adequately model relationships between input variables and the desired outputs for each grouping;
- The time horizon (in general, precision is less relevant for longer forecasting horizons (see the discussion of Uncertainty below));
- A suitable trade-off between model complexity and accuracy (typically, more detailed models with more specific categories could be more individually accurate, but lead to a more complex model overall which may have disadvantages).

- ***Other modelling "elements".*** There are many other factors to consider when formulating a prediction model, some of which are mentioned here.
  - Surrogate (or indicator) variables: oftentimes, a surrogate variable is used because it is more readily measured than the actual characteristics that underly the phenomenon. For example, age of ship is often used as a proxy when modelling incidents at sea because older ships are not as reliable or have older equipment designs.
  - Number of groupings: aside from the reasons to group ships discussed above, and even if plentiful high quality data were available, it often does not make sense to have too many groups because it is hard to grasp the results, or translate them into actionable plans. If a model has 5 types of vessels and 10 size categories, then we are predicting the outcomes for 50 groups, which may not be practicable.
  - Data availability/quality: perhaps monthly time series would be best for a given problem, but only annual data are available.
  - An attribute that is forecast to change: if the aim is to consider the future impacts which depends on a particular vessel characteristic that may change significantly, such as container vessel sizes, then that data must be included (see Figure 2).

- ***Exposure measures***: In the context of ship traffic modelling, the exposure is a measure of the level of activity that relates to the impact of concern. There are many ways to measure shipping activity, and the choice depends on which is more suitable for the nature of the problem, and/or which data are available. Some examples are:

  - Counts (movements, visits, trips)
  - Time at sea
  - Distance travelled
  - Cargo flow (in volume or weight)
  - Traffic density in prescribed grid squares

  The exposure measure is often used as an indicator variable for a chosen dependant variable. For example, we can state ship air emissions as a function of distance travelled, or fishing vessel incidents relative to the number of trips taken. Such proportionality is important for forecasting, because when it holds one can assert that doubling the number of vessels transiting an area per unit time would double the cumulative noise generated for example. Often, exposure measures are just assumed to be adequate for the given problem being studied, but this may need to be tested.

- ***Uncertainty***: By definition, the future is uncertain. In fact the fundamental purpose of prediction is to reduce the uncertainty through a methodical approach that attempts to narrow our understanding of the potential path that may occur relative to what guesswork would do. Nevertheless, the uncertainty persists. There are many sources of uncertainty, but some of the key ones when forecasting are:
  - Intrinsic uncertainty: any activity has discernable variations which are detectable if it is measured closely enough. For example, even if vessels have a planned travel speed, it will fluctuate a bit due to the propulsion system.
  - External uncertainty: external factors that fluctuate can also cause uncertainty in the phenomenon of interest. Variability in weather can affect ship speed at sea.
  - Measurement uncertainty: few measurements are perfect, whether due to instrument variability or calibration, errors, transmission problems, etc. A common issue in this regard for shipping modelling is the use of Automatic Identification System (AIS) data, which must be carefully inspected and generally "cleaned" before use.
  - Data quality: databases suffer from many irregularities such as missing information, erroneous information, conflicting information, etc. This limitation also applies to AIS databases.
  - Modelling errors: as described in section 1.4.4.5, this does not mean a mistake, but to the fact that a model does not fit reality perfectly; there is always some unexplained residual that the model does not predict. In fact, prediction model choice, and the choice of parameters are usually done to minimise the error between the fitted model and actual data based on historical activity.
  - Biases: subconscious (or deliberate) biases are often introduced to prediction modelling when subjective inputs are required.

There are several ways to characterize, capture and deal with “uncertainty”. Some are:

- There are many measures to quantify the variation in an activity if not captured by the fitted prediction model (i.e. the error), which gives an idea of how much uncertainty remains when applying the model. Some of these are described in section 1.4.4.5.
- Many prediction models reviewed below include base case results, as well as high and low (or worst and best) cases to present a realistic range of outcomes depending on the assumptions.
- Extreme cases are sometimes considered to prepare for even the most uncertain outcomes. For example, transarctic shipping across the north pole may commence before mid-century enabled by the reducing ice cap, but few governments are actively addressing this issue (The Maritime Executive, 2019).
- There are many sources of uncertainty as described above, but a modeller should differentiate between uncertainty around current operations/activity versus the uncertainty generated by predicting future outcomes.
- Monte Carlo analysis (described in section 1.4.4.4) is an approach to add probability distributions to quantities that are uncertain, then this variability gets incorporated into the prediction process to yield the probability distribution of the future activity.
- Sensitivity analysis: there are many ways to conduct sensitivity analyses, but essentially it involves altering a quantity or parameter that we are unsure of, and to see whether that alters the output of the modelling and, if so, by how much. If the result hardly changes, then it is not very sensitive to that parameter, but if it changes a lot, then we must pay great attention to the implications of that input quantity.

1.4.4. Basic introduction to forecasting methods

Prediction methods can be grouped into a few main categories, as shown in Figure 5. The quantitative approaches encompass causal prediction models, as well as time-series (forecasting) models. Many quantitative models fall into these two categories, and an overview of the key ones is presented below. In a general sense, qualitative models rely on expert opinion about potential future scenarios without explicitly relying on quantitative modelling, but realistically those opinions are generally anchored by direct or indirect knowledge of existing or historical activity levels and relationships with various drivers. These concepts are also considered below. The aim of this section is to give the reader an introduction to the various approaches to prediction and related terminology, rather than a detailed explanation of the underlying mathematics which are best applied by expert modellers. Such models are presented in many textbooks, such as the classic one by Makridakis (1998) or Hyndman (2018).

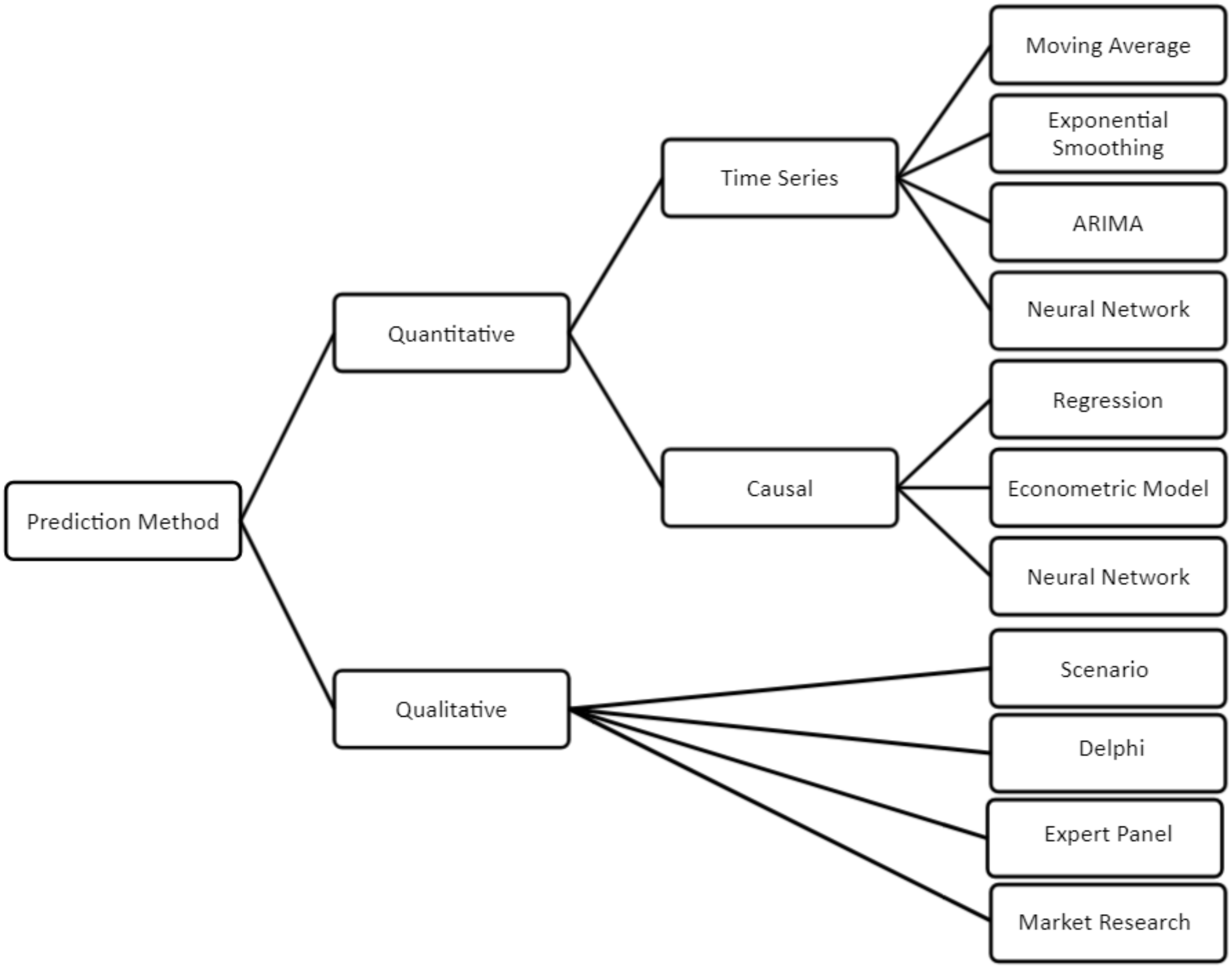


Figure 5. Categorization of prediction methods (adapted from Brillio, 2018)

1.4.4.1. Time series models

A time series refers to a sequence of observations over a historical period. Usually, the observations are at regular intervals. For example, having a record of the number of container ships coming through a given port every month for the last 5 years is a time series. If there are sufficient observations in every period, and enough periods in the data set, then patterns can be reliably measured using various analyses. Common patterns include:

- *Stationary*: where there are fluctuations across periods, but the average remains fairly constant over time;

- *Trends*: this is usually a linear increase or decrease in the activity levels over time; however it could also be non-linear (such as exponential growth);
- *Seasonal*: when the activity being observed is affected by factors that vary periodically. The period could be quarterly, monthly or weekly variations. For example, recreational boating in an area might be significantly higher on the weekend than during the week;
- *Cyclical*: when there are patterns of ups and downs in the activity over time, but not at regular periods, such as the stock market or the business cycle.

In Figure 6, a time series of quarterly measurements is shown over 10 years in the right-hand figure. The left-hand charts show how the series has been decomposed into the 4 underlying patterns of (from top to bottom) a trend, a cycle, a seasonal component, and the remainder which is a stationary series (with irregularities, but constant average). Many prediction models rely on breaking down the historical information in this way. It is difficult, and often impossible, to discern the underlying patterns visually without performing such decomposition. For example, the seasonal and irregular components are easily confounded. Similarly, the cyclic component is not readily apparent in the right-hand series because its scale is relatively small, ranging from -5 to +5 units, compared with the entire series which shows fluctuations in the scale of 50+ units.

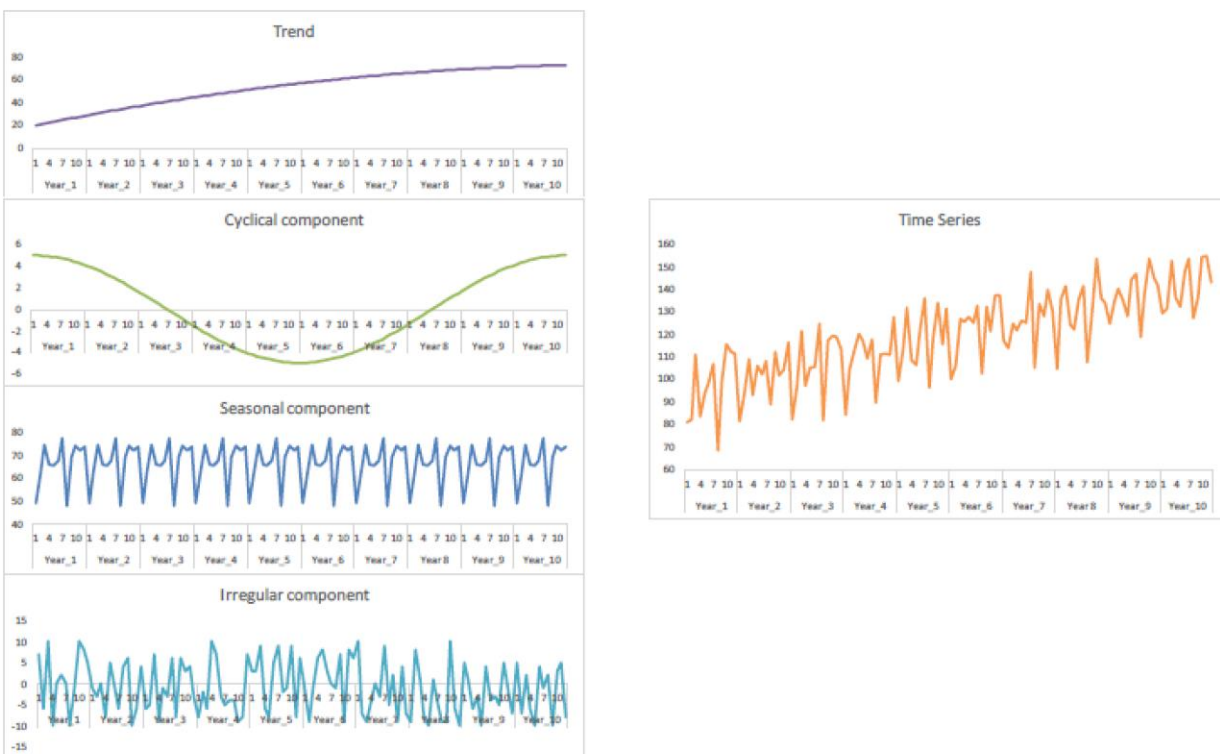


Figure 6. Additive decomposition of a time series (source: Indra, 2015)

Aside from gaining insights through visual inspection, models can be fit to the historical data and then tested for goodness of fit to capture the trends and/or seasonality if they exist. Normally, there are fluctuations in the data (irregularities) from period to period on top of the larger underlying patterns, and this can also be measured using various statistics such as the standard deviation.

Time series analysis methods can be applied to forecast the activity in question if these criteria are satisfied:

- the amount of historic data is sufficient to be representative of the activity and to perform robust analyses;
- the underlying patterns (and possibly statistic variations) can be adequately captured by the model used;
- the future conditions under which the activity occurs will be the same as, or very similar to, the historical conditions;
- and, related to the last point, the time series forecasting is not applied too far into the future. The farther out we look, the more likely it is that one or more underlying conditions that affect the activity could change significantly and/or that there are shocks to the system that engender major changes in the activity. For example, if a new port is opened in an area, then the past level of shipping traffic in that area will not be expected to be the same in the future. Other examples of major disruptions that appear in the literature include the impact on global shipping of COVID-19 or of the Global Financial Crisis of 2007-2008.

Some of the key measures and methods are briefly described here:

- Moving Average Smoothing (MA):

  The MA is a simple method that also forms the basis for more advanced prediction methods. It involves choosing a set number of consecutive historical periods and then averaging the data over those periods. Then we move forward in time by adding the next period to the set and removing the oldest period, and then recalculating the average, and then keep repeating this process. For example, if we have monthly shipping data in an area for a year, and we choose a 3-month period to apply the MA, then we average Jan-Mar which gives the 1st MA, then we average Feb-Apr which gives the 2nd MA, then we average Mar-May, which gives the 3rd MA, etc. The MA calculation helps smooth out fluctuations in the data, which allows us to then identify trends and/or seasonality in the data.

  One way that the MA approach can be used for simple forecasting is to predict one period ahead. So in the above example, the shipping activity for Sept-Nov. can be averaged to predict the December traffic. Another use is that once the historic data has been smoothed using MA, a trend line and/or a curve that reflects seasonality can be fit to the MA values, and then these curves can be used to project out several periods. For example, if the ship traffic has been smoothed with MA over the past year, a trend line may show that the number of ships has been increasing by say 10 ships each month. Thus for next year, January should be about 10 more ships than the preceding December, February about 10 more ships than January, etc. That is, we just extrapolate the trend into the future. There are many variations to the MA model, such as weighting the more recent periods more than earlier periods in the set, but the concept is the same.

- Exponential Smoothing (ES):

  In this model, the forecast puts more weight on the most recent data points and decreasing amounts of weight as the data points get older. Compared with the MA model, the ES model can be more responsive to changes in the activity since more importance is placed on the recent observations, while still leaving the older observations in the estimate but with much lower weight. There are many variations of this model on how to choose the best weights, and to test which approach fits best using historical records for validation. A common more sophisticated version called Holt-Winters incorporates three smoothing equations for the activity level, the trend, and the seasonal component respectively.

- ARIMA models (AutoRegressive Integrated Moving Average):

  The autoregressive model uses observations from previous time periods as input to a regression equation to predict the value at the next step. A regression equation defines the relationship between a dependent variable and one or more independent (predictor) variables. The "auto" in ARIMA refers to the fact that the predictor variables are earlier observations of the same variable being predicted. For example, to predict future monthly shipping traffic levels, the regression is done on historic data points that measure the same variable (i.e. shipping traffic levels). So, if X_t is the shipping volume in an area in month t, then:

  $$X_t = b_0 + b_1 \cdot X_{t-1} + b_2 \cdot X_{t-2} + \ldots + b_k \cdot X_{t-k}$$

  where X_{t-1} is the volume in the previous month, X_{t-2} is the volume two months ago, etc., and the 'b' values are constants determined mathematically when fitting the model to the data.

  While the aim here is not to delve into the mathematics, this equation neatly demonstrates that the future value (X_t) is assumed to be based only on the previous activity levels, and not on other factors (see causal models in section 1.4.4.2). There are many assumptions when using such models, including that the historic observation series is stationary, but there are also many variations to the model that can account for other pattern types, such as SARIMA (seasonal ARIMA). One commonly used procedure called X11 serves to make seasonal adjustments to time series through additive or multiplicative operations, usually on monthly or quarterly data. Note that the moving average appearing in the ARIMA acronym refers to the moving average of the random fluctuations in the observations that can occur in any period (called white noise), not the MA of the actual observations that was described previously.

1.4.4.2. Causal models

As discussed in section 1.4.2, longer term prediction models incorporate causal factors into the mathematical modelling. As opposed to time series forecasting where we are not trying to capture the underlying causes of the activity levels in the model, in many cases maritime traffic levels can be best foreseen by understanding its main drivers and how they may evolve over time. Many of the articles reviewed below involve causal models such as: how the gross domestic product (GDP) is a main determinant of the amount of container shipping in a region; how the tanker traffic between regions is a function of many factors relating to oil supply and demand; how Northern Canada community resupply depends on the population in the region. Casual models are also referred to as explanatory models.

Some of the key methods are briefly described here:

- *Regression*: Most causal models are based on some form of regression. The item that we want to predict is called the forecast variable, or the dependent variable. The causal factors that we include are called predictor variables, explanatory variables, or independent variables. The simplest version is a regression assuming a linear relationship between the forecast variable y and a single predictor variable x:

  $$y_t = a_0 + a_1 \cdot x_t + \varepsilon_t$$

  The coefficient a_1, the slope, represents the predicted change in y for a unit change in x (at time t), while a_0, the intercept, gives the value of y when x is 0. ε_t is random "error", not in the sense of a mistake but rather deviations in the observations from the fitted straight line. Thus ε reflects anything that may affect the forecast variable y other than the predictor variable x. These error terms are also referred to as residuals. Note that a regression model itself does not establish causation, only a correlation between the variables. However, the predictor variables are often assumed to be underlying causal factors for the phenomenon in question.

Of course, many regression models involving maritime traffic use more than one predictor variable, each with their respective coefficients reflecting how they affect the forecast variable. These coefficients are not known in advance, but are determined by using a long time series of historical data and modelling how y in those previous periods relate to the x's, using a method like least squares estimation. Many of the articles reviewed below test how well a given regression model fits the historic data, while other authors assume a relationship between the causal and forecast variables based either on others' modelling work or on expert opinion. For example, most models of container shipping volume in a defined geographic space are assumed to relate to GDP (at the destination region at least, and sometimes also at the originating country and perhaps in other regions). Some authors though contest that using solely GDP as the predictor variable of container shipping may be insufficient.

A great number of regression models have been developed and incorporated into commercial statistical software, including: multiple regression models where there are 2 or more predictor variables; inclusion of variables in the equation to account for trends and/or seasonality; nonlinear regression which is sometimes a good fit when linear relationships do not exist, etc.

- *Econometric models*: Many factors in the economy depend on other measurable factors, and econometric models capture the statistical relationships between causal factors and one or more dependent variables. While econometric models can be built using a wide variety of functions, linear regression is one of the most common bases for these representations. Such models arise in the shipping forecast literature because many predictions of shipping activity are intrinsically driven by economic factors, in the local economy, in the global economy, in the shipping industry, in the oil industry, etc.

1.4.4.3. Neural Network models

Artificial Neural Networks (ANN) involve a prediction approach that is based on a simple model of how neurons in the brain interact. The predictor variables form the input layer, the forecast variable(s) form the output layer, and typically there are intermediate, or "hidden", layers in between (Figure 7). In the most common ANN, each layer of neurons (more commonly referred to as nodes) produces outputs which are then fed into the next layer of neurons. These inputs are combined using a weighted linear combination. This is referred to as a multilayer feedforward network. Various weighted combinations and nonlinear functions can be applied to the information flowing through the network. The parameters for these functions are "learned" from being fed data with known inputs and their related historical outputs. Then the model can be used for forecasting based on new input data to predict what the outputs would be. ANN is a very valuable tool because, as opposed to the time series and causal models described above, it does not presuppose relationships between the predictor and forecast variables. It is meant to discover how complex mixtures of the input factors can be combined effectively to reliably predict the output variables, thus it is a black box approach where one cannot specify the relationships explicitly. It is a very active area of research to develop and test new ANN model structures that are well-suited to various problem domains.

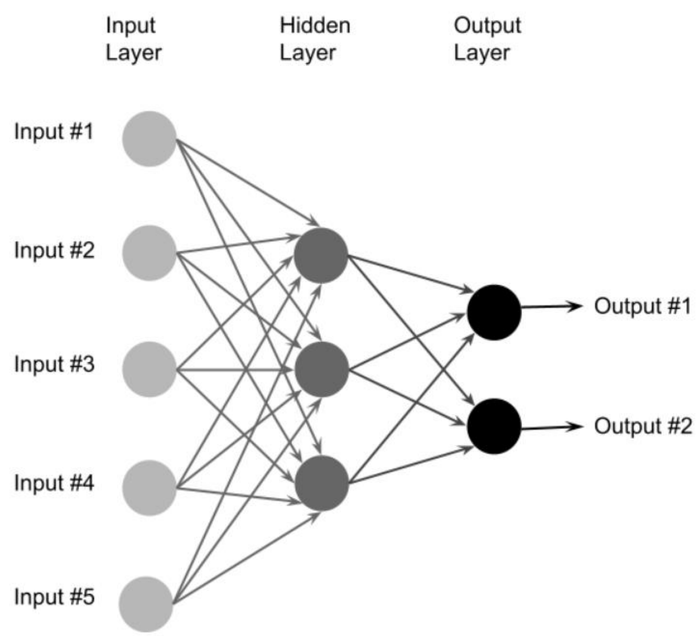


Figure 7. A neural network with one hidden layer and five inputs.

Neural networks can be applied to autoregression problems as described in section 1.4.4.1, but it has been argued that they do not necessarily perform better than existing time series models. More importantly, they are often applied successfully to causal forecasting problems, sometimes when none of the traditional models are adequate to represent the complex interaction of factors, and sometimes because they simply perform better than the traditional models. In the shipping forecast literature, ANNs have been used under both of those conditions, as being the only suitable approach, or perhaps performing better in comparison with conventional models. Sometimes, they have been used as part of blended models which combine ANN and regression equations in novel ways (see section 1.4.4.7). For completeness, note that ANNs are a specific type of 'Machine Learning', which in turn is a subset of 'Artificial Intelligence'. While ANNs are a fairly recent development in the modelling world, they have already been proven to be quite accurate and very valuable in a wide range of applications.

1.4.4.4. Monte Carlo Simulation

Monte Carlo (MC) is not a forecasting tool, but rather a method that is included in certain forecasting models. One way to address uncertainty in the resulting values of predicted variables in the future is to apply a statistical distribution to a value, instead of assuming that it will be a deterministic quantity. For example, a container shipping forecast model for a port may predict that the number of ships calling is proportional to the GDP growth in the region, so if the GDP is predicted for a decade ahead then we can calculate the container traffic using the appropriate multiplier factor from the regression model. However, this year-10 traffic level could instead be represented by a range of values by assigning a probability distribution to what the GDP value might be in a decade, and run the model (i.e. simulation) hundreds or thousands of times, each run representing one instance from the probability distribution. This will then produce a distribution of what the traffic would likely be in 10 years, rather than just a point estimate. MC is an even more valuable tool for forecasting when several factors in the model have inherent uncertainty, and we represent each of them with a suitable probability distribution respectively, and then run the entire model which will aggregate the combined estimates of each of these uncertain factors. That is, the variability of the forecast variable will result from the possible random choice of each of the input factors collectively.

1.4.4.5. Model error measurement

A very important stage in developing a prediction model is to evaluate the accuracy of the forecast by measuring the difference between the predicted quantity and the actual outcome, which is referred to as the error. This is usually done by dividing the historical time series of input data into two parts: the older data being the 'training data', and the newer part reserved as the 'test data'. Only the training data is used to develop the model and establish its parameters, whether it's an ARIMA model or a causal model, or another type. Then when the model is set, it can be used to "predict" what the forecast variable series should be during the "test data" period. Then those forecast values can be compared to the actual outcomes during that period, yielding a set of error terms.

These errors can serve to assess if the prediction model is performing well enough for its intended purpose, or to readjust the model parameters to minimize the errors. There are many ways to measure and track these error terms, and some of the most common ones are described here:

- Mean Absolute Error: $MAE = \frac{1}{n}\sum_{t=1}^{n}|x_t - \widetilde{x_t}|$
- Mean Squared Error: $MSE = \frac{1}{n}\sum_{t=1}^{n}(x_t - \widetilde{x_t})^2$
- Root mean square error: $RMSE = \sqrt{\frac{\sum_{t=1}^{n}(x_t - \widetilde{x_t})^2}{n}}$
- Mean Absolute Percentage Error: $MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{x_t - \widetilde{x_t}}{x_t}$

The *n* represents the number of periods over which we are calculating the error, the Σ sign adds up the errors over these periods, the x_t is the actual value at time t, and $\tilde{x}_t$ is the predicted value at time t. MAE simply averages the absolute value of the errors (so that the negative deviations don't cancel out the positive deviations). MSE squares the errors so that larger deviations get weighted much more, which is useful when large errors are particularly undesirable. The RMSE, the square root of the MSE, expresses the error in the same units as the original variable. MAPE adds up the percent error in each period. There are many other metrics to assess the fit of the prediction models, but these common ones above appear in the literature on shipping forecasting.

1.4.4.6. Qualitative predictions

Qualitative predictions are needed when either there is insufficient data to perform a quantitative prediction, or to complement a quantitative forecast. In the former case, data insufficiency can occur in several circumstances:

- we are forecasting a new activity (e.g. a new annual intercity sailing regatta);
- it's an existing activity but little or no data have been collected (e.g. small vessel traffic in the north);
- the future conditions are expected to be significantly different than historical circumstances, so existing data are not applicable (e.g. fishing vessel patterns in the years following the cod moratorium on the East Coast).
- The future is highly uncertain, which is particularly common when the forecast time horizon is very long.

There are many approaches to generating potential futures based on judgment, and key ones that appears in the shipping forecasting literature are described here.

- Future Scenario building:
  As described in section 1.4.2, building scenarios provides a structured approach to develop a set of contrasting, plausible set of circumstances that are relevant for the predicted variable. The generation of such hypothetical futures for planning purposes is increasingly common in many domains, but it comes with several caveats. First of all, many authors state that scenarios are not predictions of what may happen, but simply some realistic outcomes which can give insight into possible future conditions which may have to be dealt with. Furthermore, users state that scenarios are supposed to be considered equally likely, as the assumptions are so open-ended that it would be hard to opine about which outcomes are more likely. Note that the term "scenarios" is not used consistently throughout the literature. While the meaning in this report is as described above, the term can also be used in the sense of trying different parameters in a model, or testing the impact of various events or decisions on outcomes.

  The Climate Change domain relies heavily on generating scenarios to consider future conditions and impacts, due to the very high uncertainty in many of the factors, the system's complexity with so many interacting elements (natural, political, technological, etc.) and the long time horizon. In greenhouse gas studies by the International Maritime Organization (IMO, 2009), these statements provide more insight on scenario planning:
  - Scenario planning is a common tool for researchers evaluating uncertain futures. Some of the definitions of scenario planning, include:
    - "[An] internally consistent view of what the future might turn out to be – not a forecast, but one possible future outcome".
    - "[A] tool for ordering one's perceptions about alternative future environments in which one's decisions might be played out".
    - "[A] disciplined methodology for imagining possible futures in which organizational decisions may be played out".
  - Scenarios help us envision a future in order to develop robust decisions and test how these decisions play out in possible future worlds.

  The IPCC (Intergovernmental Panel on Climate Change) also uses the following terminology for its scenarios (IMO, 2009):

  Storyline: a narrative description of a scenario (or a family of scenarios), highlighting the main scenario characteristics and dynamics, and the relationships between key driving forces.
  Scenario: projections of a potential future, based on a clear logic and a quantified storyline.

Scenario family: one or more scenarios that have the same demographic, politico-societal, economic and technological storyline.

- The Delphi Method:
  The Delphi Method, which was developed decades ago, serves to generate a forecast through consensus of a set of experts. The experts do not meet and they remain anonymous to all except to the facilitator who is leading the forecasting initiative. The aim and scope of the forecast is described thoroughly and clearly by the facilitator. Sometimes, this initial plan is circulated to the experts before their judgment is invoked, to solicit clarifications or suggestions. Once the plan is finalized, each expert creates their forecast individually, with detailed reporting of the forecast characteristics, assumptions, and limitations. The facilitator collects all of the experts' outputs and prepares a single summary document which includes summary statistics of the collective forecasts (average levels, standard deviation if applicable, high and low cases for the forecast variable) as well as an outline of the main assumptions and justifications that the experts applied. Significant differences in some estimates and outliers may be highlighted. This summary document is provided to all of the experts, who then use this aggregate information coming from others to update their individual forecasts, as well as commenting on the overall nature of the responses to date. This process is repeated until (most of) the experts' reach a reasonable consensus about the forecast.

  While this is usually a prolonged process (2-3 iterations may take weeks to months depending partly on the size of the expert group, often a dozen or more people), it serves to blend the knowledge of many experts without the potential detrimental impacts of group dynamics.

- Expert panels:
  Expert panels can be convened for broad scenarios generation as described previously, but also for more narrow forecast initiatives as described in the Delphi section above. The advantage of joint discussions (relative to the Delphi method) is that it can generally proceed much more quickly, and benefits from group brainstorming which can effectively generate a plentiful set of ideas about relevant factors, assumptions, possible future circumstances, etc. Several well-known group dynamics aspects may impede achieving the best outcome though, such as more vocal or powerful people dominating the discussion, information overload, etc. An example using such a process is found in a report on shipping risk by the Council of Canadian Academies (CCA, 2016).

- Market research

Market research through surveys, interviews or focus groups serves to identify potential futures based on users' preferences and perspectives. This technique is used extensively in domains such as business and politics. However, it can also be applied to gauge the public's intentions in other activities which may influence potential future directions, or perhaps guide decisions by those involved with the activity in question which will then shape the future reality. For example, people's opinions on pleasure boating may be gathered through a survey to determine whether interest is growing in this activity, if more people intend to acquire boats in the foreseeable future, etc. This could give insight into the future amount of recreational boating traffic in a region. As another example, if a large company is planning to establish a new aquaculture operation in a coastal area with its associated operations boating traffic, public opinion may indicate stiff opposition to the point where a decision is made not to proceed.

- Indigenous Knowledge
  Since European contact, Indigenous Peoples have witnessed changes and trends in marine vessel traffic in their traditional waters, including remote areas that may lack adequate vessel monitoring capability. Weaving Indigenous knowledge into a vessel prediction model may be useful at many stages of the analysis, including: providing estimates of traffic levels to complement existing data; suggesting causal factors that may be relevant determinants of activity; proposing future scenarios that should be explored to assess impacts of interest; incorporating elements into the models which they deem to be important which may have been overlooked by the modellers; and validating prediction model outcomes against the marine context with which they are very familiar.

To summarize, qualitative models play a significant and increasing role in predicting or envisaging situations for the phenomenon that we want to forecast. However, the general principle is that if there are data available, they should be used to the full extent possible as long as they satisfy the necessary conditions listed earlier. This may be to lay the groundwork for subjective assessments, or to run parallel quantitative and qualitative forecasting exercises, and then to somehow combine the outcomes. There is a large body of literature examining the cognitive biases that people have when making judgments, decisions, or projections, including subconscious influences due to psychological processes such as anchoring or recency effect (Kahneman, 2011), all the way to deliberate self-serving motives to make the prediction support a desired course of action. This latter behaviour can be mitigated somewhat by having the experts doing the predictions being distinct from the end-users of the forecasts.

1.4.4.7. Blended Models

While most shipping forecast reports rely on a single type of model, a significant number of studies combine one or more prediction approaches. For example, it is possible to combine ARMA and ANN into a hybrid model, where the former deals with linearity in the historical data and the latter deals with the nonlinear residuals resulting from the ARMA application (ex. Wang, 2017). van Dorsser (2012) generated a very long-term forecast of 90 years based on a combination of System Dynamic Modelling (SDM), Judgement, and Causal Relations. SDM captures the complex nonlinear interactions between variables in a system including feedback loops, allowing exploration of cause and effect, and impacts of changes in the system. Combining quantitative and qualitative models to generate a prediction is usually referred to as a blended model, of which several are presented in the literature review in section 2. Good examples can be found in Walsh (2019) and Indra (2015). As part of the PORTOPIA project, which is aimed at the development of key performance data for European ports, Indra (2015) also provides a description of a wide assortment of prediction techniques for port traffic forecasts, culminating in a tabular summary of the applicability of the methods for their purposes (Table 1). It has also been shown in some studies though that more sophisticated complex models do not always produce better results than simple quantitative models (ex. Peng, 2009).

Table 1. Overview of forecasting methods considered for the PORTOPIA port traffic forecasting study (source: Indra, 2015)

Qualitative

	Delphy	Market research	Panel Conscensus	Visionary forecast	Historical analogy
Accuracy					
Short Term	Fair to good	Excellent	Poor to fair	Poor	Poor
Medium term	Fair to good	Good	Poor to fair	Poor	Fair to good
Long Term	Fair to good	Fair to good	Poor	Poor	Fair to good
Identification of turning points	Fair to good	Fair to good	Poor to fair	Poor	Poor to fair
Data required	Questionnaires	Market data	Expert pannel	Scenario set	Raw data
Possible within Portopia forecast	Yes	Yes	Yes	Yes	Yes

Time series

	Moving average	Exponential Smoothing	Box-Jenkins	X11	Trend projections
Accuracy					
Short Term	Poor to good	Fair to very good	Excellent	Excellent	Very good
Medium term	Poor	Poor to good	Poor to good	Good	Good
Long Term	Very poor	Very poor	Poor to good	Good	Good
Identification of turning points	Poor	Poor	Very poor	Very poor	Good
Data required	2 year raw data	2 year raw data	2 year raw data	3 year raw data	5 year raw data
Possible within Portopia forecast	Yes	Yes	Yes	Yes	Yes

Causal methods

	Regression	Econometric model	Intention to buy	IO model	Economic IO model
Accuracy					
Short Term	Good to very good	Good to very good	Poor to good	NA	NA
Medium term	Good to very good	Very good to excellent	Poor to good	Good to very good	Good to very good
Long Term	Good to very good	Good	Very poor	Good to very good	Excellent
Identification of turning points	Very good	Excellent	Good	Fair	Good to very good
Data required	Quarterly data	Quarterly data	raw data	raw data	raw data
Possible within Portopia forecast	Yes	Yes	No	No	No

	Diffusion index	Leading indicator	Life cycle analysis
Accuracy			
Short Term	Poor to good	Poor to good	Very poor
Medium term	Poor to good	Poor to good	Poor to good
Long Term	Very poor	Very poor	Poor to good
Identification of turning points	Good	Good	Poor to good
Data required	raw data	raw data	raw data
Possible within Portopia forecast	Yes	Yes	No

2. Overview of key literature

Over 60 articles and reports are covered in this review. The aim of the review is not to focus on the results of the studies, but to summarize the prediction methods used and comment on their applicability to various conditions such as vessel type, data availability, time horizon for a forecast, etc. While not comprehensive, this set of articles provides a collection of different modelling techniques applied to a variety of traffic prediction problems. Key elements of each study are summarized in section 3 – Compendium of reviewed articles and reports. This section is primarily organized according to the type of problem that a prediction study is addressing, as the methods that are used in that context can be compared and contrasted.

2.1. Ports and Container ships

Given the scale, cost and longevity of port developments, it is not surprising that the bulk of the shipping forecast literature addresses this issue. Furthermore, a disproportionate fraction of those studies model container shipping, perhaps because that transportation mode is so extensive with rapid growth over the years and, in general, large volumes of high-quality data are available for analysis.

2.1.1. General studies on port forecasting

- A comprehensive introduction to this topic can be found in Parola (2020). Their article addresses traffic forecasting for port planning, and an innovative feature is that they do an extensive literature review and also contrast how academics have approached this problem versus port authorities (PA) based on examining 28 EU ports. Focussing on container and cargo traffic, the 22 articles that they reviewed comprise a mix of qualitative, quantitative, and mixed statistical methodologies, with a detailed summary table to compare them. The methods used across these studies include all of the methods described in section 1.4.4. The results are too extensive to report on here in detail, but a few key observations are made. It is noted that purely subjective predictions are not sufficient for PAs to make port management or investment decisions, but quantitative or hybrid models suffice. They also point out that several authors compare quantitative techniques to determine which produces the best results for their given circumstances, such as neural network versus regression (Gosasang, 2011; Lam, 2004), or Peng and Chu (2009) who compared six univariate methods for container throughput volumes in Taiwanese ports. It is pointed out that traditionally most port planning modelling has been based on trend models which disregards significant shifts in underlying drivers that could be captured in causal models, as well as scenario planning. This tendency however is reversing in more recent years through the use of hybrid models. The academic studies are contrasted with actual Port

Authority planning processes which consider among other things: global economic trends; inter-port competitions; port size (smaller ports do not conduct such sophisticated predictions); private vs public port ownership; economic, social, political, nature-related or health shocks to the system; etc.. Some port planning processes are very structured and transparent, while some medium to small ports have a more fragmented ad hoc approach. The authors conclude that blended models are the best approach, and that PAs should increasingly include sophisticated models in their planning, and be more transparent about their planning processes and associated data.

- Another comprehensive study on port traffic forecasting by Indra et al. (2015), is done in the context of the PORTOPIA project on European port planning. This paper provides a detailed overview of key forecasting methods (see Table 1) and prediction studies conducted using various methods across different ports. Their aim is to determine which methods are best suited for estimating future traffic flows in the European port system in the short (0-1 year), medium (1-5 years) and long term (beyond 5 years). Aside from the technical merits of the various mathematical models, their recommendations also depend on other factors (Indra et al., 2015, p.37):

    - The cost of developing the forecasting model
    - How complicated are the relationships that are being forecasted?
    - Is it for short-run or long-run purposes?
    - How much accuracy is desired?
    - Is there a minimum tolerance level of errors?
    - How many data points are available?

  They state a caveat on the exclusive use of scenarios, which are useful but subject to inconsistencies and vagueness, instead opting for having them as a complementary technique to quantitative forecasting, especially in the longer term. They also caution that historic reliance on GDP as the main causal factor for ship volume predictions can be unreliable as it is too coarse to capture significant changes in the markets for individual commodities or it ignores the role that other causal factors play. However, they concede that it's often relied on due to data availability and desired model simplicity. Cox (2010) also produced a memo describing the limitations of GDP as a reliable indicator of shipping trade volume, and suggests some alternatives.

  Indra et al. (2015) list a large set of indicator variables in four categories (market trends and structure; socio-economic; environmental and safety; and governance) denoting which can be forecast over which timeframes, and which indicators are most suitable for short, medium, and long term ship traffic forecasts. Ultimately, based on the criteria noted above, their recommendation is:

- Short term forecasting: an automated system, based on moving average or exponential smoothing;
- Medium term forecasting: market research, historical analogy, X11 ARIMA approach, trend projections, regression and econometric models;
- Long term forecasting: econometric model, regression analysis, or input-output model.

- Gargari et al. (2019) compares an ARIMA model to hybrid models that are based on ARIMA plus a Neural Network to forecast short-term container traffic in a port. ARIMA is applied to the linear component of the historical data, and NN to the (non-linear) residuals (Figure 8). One hybrid model examined is additive and the other is multiplicative.

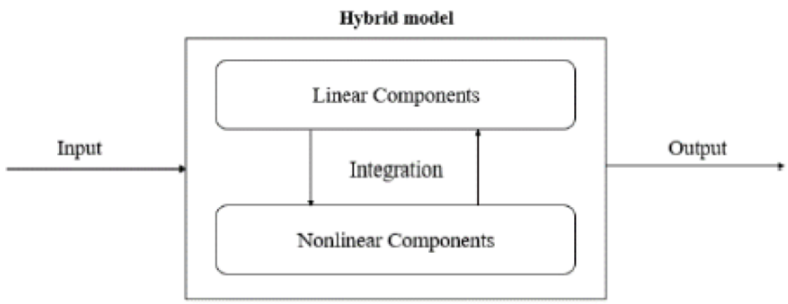


Figure 8. Block diagram of the neural network and ARIMA hybrid model (source: Gargari, 2019)

Traffic data from the Port of Rajaee, Iran, yields 100,139 visits which are then aggregated monthly over the 13 years from 2005-2018. These data are used to fit the models, and then to test their respective accuracies using four measures: R (correlation coefficient), MAD, MSE and MAPE. They compare the models through tabular results of their accuracy measures as well as plots of their respective forecasts relative to the actual historical outcomes, with the NN model scoring best on most measures.

- Peng and Chu (2009) compared 6 time series short-term (1-2 years) forecasting methods for container throughput volumes (in TEU) at three Taiwanese ports using monthly data from 2003 to 2006, split into training and testing sets. The methods were: classical decomposition (based on a multiplicative model, with trend, cyclical, seasonal and irregular factors), a trigonometric model, seasonal dummy variables, grey forecast, hybrid grey forecast, and SARIMA Seasonal ARIMA). Using three measures of model accuracy (MAE, MAPE, and RMSE) they determine that the classical decomposition model produces the best results (with SARIMA having a slight edge in one port). This demonstrates that more complex models do not always produce more accurate results.

2.1.2. Prediction models applied to specific ports

The models in this section are focussed on specific ports, either because data are available there to develop and test their prediction models, or because they are incorporating characteristics that are particular to their examined port. (Note that Canadian West Coast ports are covered in section 2.1.3).

- Zhang et al. (2013) developed a model to forecast cargo throughput (measured in Mt) in a large (unidentified) port in southern China, using annual data from 2002-2011. They fit the data using two methods:
    - Grey prediction model (GM): this is a time series forecasting model which can be used when there is very little data available, down to 4 observations. It includes differential equations to incorporate the rate of change of the variable over time.
    - Logistic growth curve model: this function, well know in the biology realm, appears like an elongated S-curve. The initial growth stage is approximately exponential, then the growth slows to linear at saturation, and then eventually levels off.

  They chose the first method due to the paucity of data. They chose the second method because it represented the developmental patterns of cargo through Chinese ports across the preceding years. They ran the two models independently, and also used them jointly. The combined model assigns a weight variable to each of the two forecast methods (with the sum of those weights equal to 1), and then uses a non-linear optimization on the sum of the weighted MSE of each model's forecast in each period to determine the two weights that makes the blended model fit the data best. They tabulate the results and show that the combined model performs the best. It is then used to predict the throughput 4 years later.

- Jugović et al. (2011) undertook to forecast container traffic out to 2040 for the port of Rijeka, Croatia. Based on a decade of data to 2008, he compared simple regression models based on linear, parabolic and exponential for Rijeka and two competitor ports (Koper and Trieste) respectively, then assuming low, medium and high cases for the market share of Rijeka relative to the other two ports. They also developed a causal regression model of container traffic (in TEU) versus projected GDP growth in the country. Finally, they developed another forecast of shipping as a function of anticipated import/export value until 2040. While all of these predictions produce different outcomes, the authors assert that the range gives a sense of the direction things will go, and how management plans and port infrastructure development must accommodate such expectations.

- van Dorsser et al. (2012) developed a very long term prediction up to 2100 for the total port throughout in the Le Havre – Hamburg Range to help with the planning of the Dutch inland waterways where 60% of the barge loads come from or to the ports in that region. Noting that long term projections using scenarios (i.e. storylines) are not predictions and thus do not address the relative likelihood of occurrence, they state that basing scenarios instead on probabilistic forecasts can produce scenarios that are somewhat equally likely and thus more useful for policy development. After reviewing the literature on long-term forecasting, they propose a three-stage blended method to achieve their aims. Acknowledging that GDP is the key predictor for port throughput, the question then becomes how to best predict the (Dutch) GDP in the long run. Considering that the GDP is a function of these four factors:

- the population of the working age class
- the labour participation fraction of the working class
- the annual number of hours worked per employee
- the development of the GDP output per hour worked

they develop a probabilistic forecast of the first one (population) using a system dynamic model (see section 1.4.4.7), and apply expert judgement to assess the remaining three factors. A causal model then links this GDP forecast to the port throughput. They then explore in depth what is a suitable relationship between GDP and throughput for long term forecasting. They examine the following three regression models in great detail based on statistical properties, theoretical considerations for each model, and testing using ex-post forecasting back to 1970.

linear (levels): $PT_t = \alpha + \beta \cdot GDP_t + \varepsilon_t$

logarithmic: $\ln(PT_t) = \alpha + \beta \cdot \ln(GDP_t) + \varepsilon_t$

linear (differences): $\Delta PT_t = \alpha + \beta \cdot \Delta GDP_t + \varepsilon_t$

where:

α : Intercept value

β : Linear coefficient

PT_t : Port Throughput level in year t

GDP_t : GDP index level in year t

ΔPT_t : Difference in Port Throughput between year t and year t-1

ΔGDP_t : Difference in GDP index between year t and year t-1

ε_t : Error term in year t

They determine that, with some minor assumptions, a combination of the first and third models should produce good forecasts of long-term PT/GDP behaviour. Then using Monte Carlo simulation they generate the long term forecasts, as shown in Figure 9.

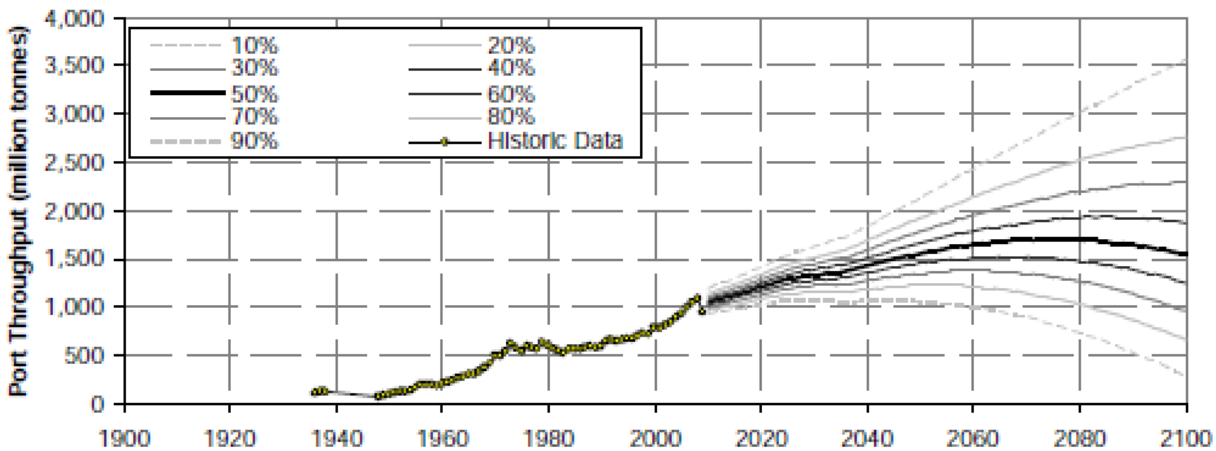


Figure 9. Final Forecast based on a Combined Levels and Differences Approach (source: van Dorsser, 2012)

Despite the wide potential variation in throughput by 2100, anticipating the decreasing and stabilizing throughput volumes will help infrastructure planners to consider the type of capacity expansions that will suit future demand for some 20 to 30 years while keeping in mind the ultimate long-term capacity needs.

- Hajbi's (2011) focuses primarily on how the decentralization of Moroccan ports in the preceding decade provokes a need for more consistent port traffic prediction which should also involve all levels of port stakeholders. The author proposes a three-stage process:
    - *Operational level*: port operators each produce 5-year forecasts
    - *Functional level*: analysis consolidating and correcting the previous level's outputs, and adding more inputs such as port statistics, generate 10-year forecasts
    - *Expert level*: create a 30-year forecast using quantitative methods considering micro and macroeconomic data, as well as qualitative methods (Delphi and other) to incorporate social, political and commercial factors.

  Hajbi (2011) also demonstrates a good fit for GDP as a predictor of container traffic at Moroccan ports using linear regression.

- Wang et al. (2017) uses AIS data to fit regression and neural network models for short term forecasting of traffic in the port of Shanghai, which can help with operations planning. Details are provided in Case study D – Spatio-temporal traffic forecasting based on AIS.

2.1.3. Canadian West Coast Ports and Container Traffic

- The Vancouver Fraser Port Authority (VFPA) has commissioned many studies over the years to examine shipping patterns, trends, and predictions. Other studies address the Prince Rupert Port or British Columbia's port strategy overall. For example, the Ministry of Small Business and Economic Development & Ministry of Transportation (2005) produced a British Columbia Ports Strategy document presenting predictions for all commercial ship types for 2020. While the forecast method is not specified, the drivers and major underlying assumptions are discussed. Depending on the sector, increases of 50-300% traffic were expected over the 15 year horizon. The aim was to position the province's ports to maximize their growth opportunities through capacity expansion, strategies to improve competitiveness, and marketing.

  Below in this section, summaries are provided for more recent prediction methods used for BC regional port analyses.

  - In 2011, Worley Parsons completed a projection of shipping traffic at Roberts Bank for the Deltaport container terminal and Westshore Coal Terminal to 2030. The predictions were made on an annual basis, with 5-year interval reporting in the document. For Deltaport, the actual total container projections were prepared earlier by PMV based on the Vancouver port share of the Canadian market, and selecting their "High projection" estimate. Another factor considered is the potential increases in container vessel size, estimated from 2010 to 2030 in 5-year increment, expressed as the percent of ships expected at Deltaport in each of 10 size categories (in TEU). The Worley Parsons (2011) study considered the following three cases:
    - Case 1: High "Direct" container traffic projection. Deltaport has a sustainable capacity of 2.4 million TEU. Container vessel sizes are as per forecasted size increases. Deltaport has a maximum capacity of 3.0 million TEU in interim years of high demand. Maximum Westshore throughput is 35 million tonnes of coal.
    - Case 2: High "Direct" container traffic projection. Deltaport has a sustainable capacity of 3.0 million TEU. Container vessel sizes are as per forecasted size increases. Maximum Westshore throughput is again 35 million tonnes of coal.
    - Case 3: High "Direct" container traffic projection. Deltaport has a sustainable capacity of 3.0 million TEU. Container vessel call size remains at 2010 level of 6,250 TEU per ship call. Maximum Westshore throughput is again 35 million tonnes of coal.

    Based on detailed analysis of historic container traffic in the area, historic and anticipated growth in container ship sizes, and the current and expected increases in

TEU per ship call, they predicted the Deltaport container volumes (in Million TEU) for the 3 cases out to 2030.

For the Westport exports, they consider the world demand for metallurgical and thermal coal, the availability of coal in Western Canada and the western U.S. for export, and the capacity of the coal terminal. Based on an Australian study that projected regional metallurgical coal trade to 2016 and assuming a constant demand growth per year after that to 2030, along with expert opinion on the other factors (including Westport planned capacity expansion) they produce their annual predictions. This ship traffic is then combined with the projected container ship traffic to Deltaport to produce the total number of vessel movements per year.

- Ocean Shipping Consultants (OSC, 2016) conducted an extensive study on container traffic at the Port of Vancouver (PoV), with the aim of generating a forecast out to 2050. The actual forecasting method is based on straightforward regression of container volume based on GDP, but the estimation of the GDP by year (different for the 1st decade to 2025, and then by 5-year increments to 2050), by market regions, and by Vancouver port share relative to competitors is extensively researched and justified. For the medium term (10 years), base-, high-, low-case circumstances with all of their attending factors are postulated. For the longer forecast (2025-2050), three macro-economic scenarios are developed: continuing free trade; a partially protectionist world; a new economic and trade paradigm. Under each of these, estimates are made for the relevant regional GDP values, for the "multipliers" that relate GDP to container TEU in the extrapolations, and for the relative market share that PoV may hold. Figure 10, reproduced from the OSC (2016) report, displays the wide range of factors considered and the rich interplay in combining them to yield the final desired forecast. Similar approaches taken in other studies are explained in more detail in Case study A – VFPA – RBT2 Container traffic.

  Having generated the potential container throughput for PoV for 35 years, the study noted that the actual realization also depends on many other factors including:
  - The overall capacity available at the Port of Vancouver to meet potential demand;
  - Shifts in the development of deep-sea containerisation – vessel sizes and market issues (such as increasing types of commodities being containerised);
  - The competitive position of container terminal facilities in terms of marine accessibility;
  - Including such items as water depth and berth lengths;
  - The relative costs and capacity of intermodal links to/from the broader hinterland in contrast to other port options able to serve the same hinterland regions.

Table 2. Main container demand drivers for the Continuing Free Trade' Scenario 2026-2050 (source: OSC, 2016)

	2026-2035	2036-2050
GDP		
Canada	Declines to 2.2% pa	Declines to 1.8% pa
West Canada	Declines to 2.5% pa	Declines to 2.0% pa
USA	Declines to 2.2% pa	Declines to 1.8% pa
Other Major Asia	Grows to 6.0% pa	Consolidates to 4.5% pa
China	Reduces to 6.0% pa	Consolidates to 4.5% pa
Multiplier		
North America - imports	Declines to 1.1	Declines to 1.0
Asia - exports to	Remains at 0.6	Remains at 0.6
Market Share Pacific Northwest - imports	Remains at 15.5%	Remains at 15.5%
Regional Distribution - within N.America		
Imports	Current distribution maintained	Current distribution maintained
Exports	Current distribution maintained	Current distribution maintained
Proportional Importance of Asia - % of total port demand		
Imports	Current distribution maintained	Current distribution maintained
Exports	Current distribution maintained	Current distribution maintained

- A follow-on study to the OSC (2016) container forecasts was conducted by Mercator (2018) to translate the OSC outputs which were expressed as volume (TEU) into ship calls. Their mandate was to "forecast how many separate vessel deployments are likely to be operated in four separate timeframes (2020, 2025, 2030 and 2035) in each of four distinct service-types (Asia – Salish Sea; Asia – Prince Rupert/California; Europe – PNW/California; and ANZ – PNW/California) – and what average sizes of ships are likely to be used in each forecasted deployment (where PNW = Pacific Northwest, and ANZ = Australia/New Zealand). There is no new forecast methodology introduced, but many factors are explored in depth that affect the conversion of container volume in TEU to port calls.

- The most recent studies on container shipping forecasting for VFPA in anticipation of the RBT2 expansion are by Drewry (2020) and WSP (2020), both of which are explained in some detail in Case study A – VFPA – RBT2 Container traffic below. Furthermore, VFPA reported on the usefulness of both of the studies, and although their respective results differed given the complexity of the task and the long timeframe, VFPA intends to base their capacity planning on values between the two studies' predictions (Port of Vancouver, 2021)

- Two more reports on the future of west coast commercial shipping warrant mentioning. In 2018, PoV produced a report entitled: "Roberts Bank Terminal 2 Project Overview and Rationale" (Port of Vancouver, 2018). Therein, they present at length many facets of the proposed RBT2 expansion, including graphing the high, base, and low case forecasts for the west coast container traffic generated by the OSC report presented above (Ocean Shipping Consultants, 2016), as a backdrop for the necessity of capacity expansion in the region. The Prince Rupert Port Authority (2020) also produced a report on their Land Use Plan, which incorporates projections of future growth broken down by terminal to 2029 (the prediction method is not specified).

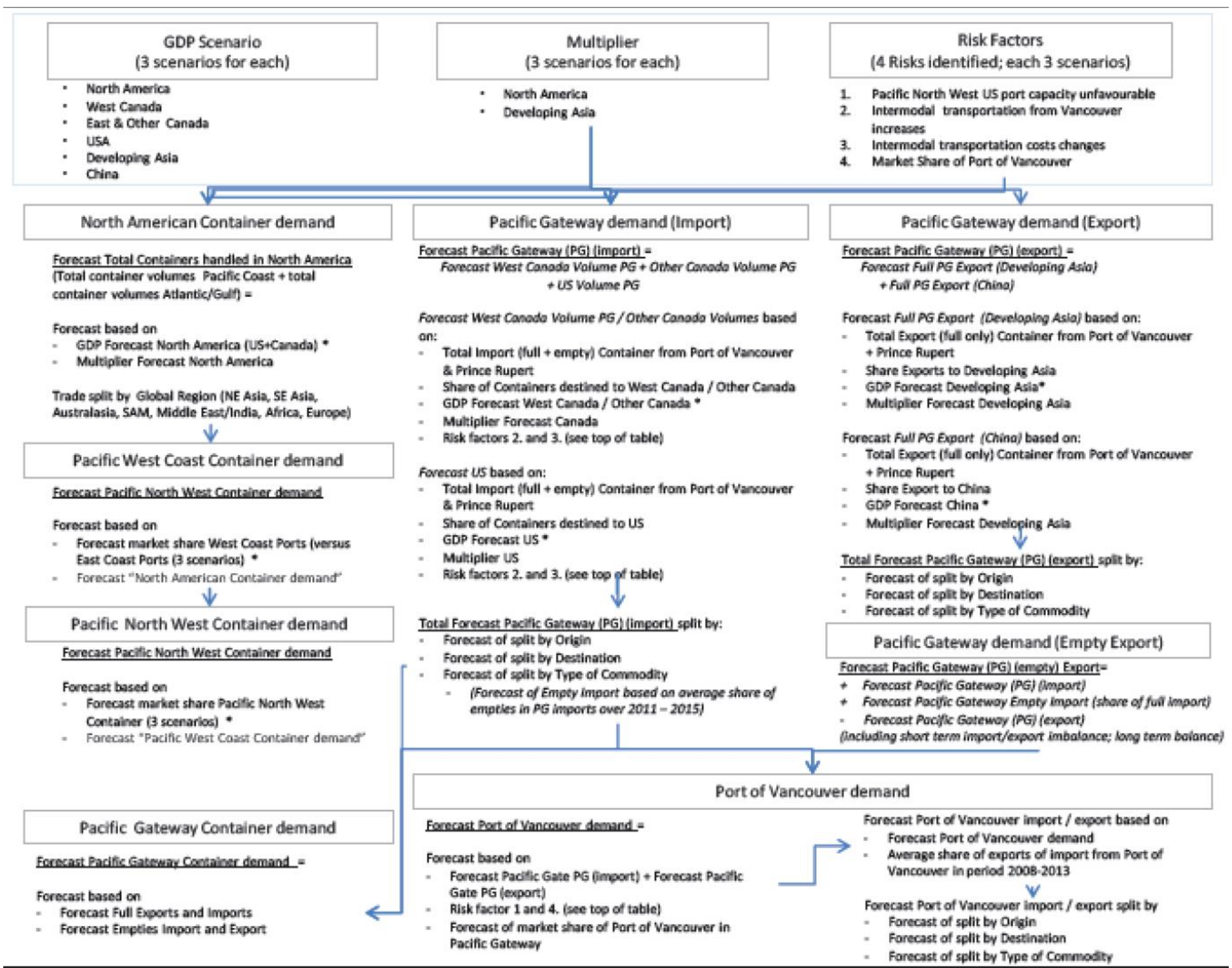


Figure 10. OSC Forecast Demand Methodology for POV Container throughput (source: OSC, 2016)

2.2. All shipping types

Some studies that develop predictions for maritime traffic cover multiple vessel types. In this section, several such reports are reviewed, with the first set involving global shipping, followed by a few that consider traffic in a large region.

2.2.1. Global shipping

- The International Maritime Organization (IMO) commissioned a study to forecast greenhouse gas emissions from ships, which was conducted by an international consortium (IMO, 2009). Fundamentally, the exercise required predicting the global shipping traffic to 2050 (in tonne-miles), while also considering other factors that would determine the GHG emissions as a function of traffic levels, based on the economy, transport efficiency, and energy. The scenarios that were developed for the GHG study can be considered as a detailing of shipping and seaborne trade within possible futures that were outlined by IPCC SRES[1] (2008) scenarios for 2050. It is instructive to consider the 4 relevant storylines from IPCC (IMO, 2009):
  1. Storyline A1: a future world of very rapid economic growth, global population that peaks in mid-century and declines thereafter, and rapid introduction of new and more efficient technologies. Major underlying themes are economic and cultural convergence and capacity building, with a substantial reduction in regional differences in per capita income. In this world, people pursue personal wealth rather than environmental quality.
  2. Storyline A2: a very heterogeneous world with continuously increasing global population and regionally oriented economic growth that is more fragmented and slower than in other storylines.
  3. Storyline B1: a convergent world with the same global population as in the A1 storyline but with rapid changes in economic structures toward a service and information economy, with reductions in material intensity, and the introduction of clean and resource-efficient technologies.
  4. Storyline B2: a world in which the emphasis is on local solutions to economic, social, and environmental sustainability, with continuously increasing population (lower than A2) and intermediate economic development.

  These storylines provided the foundation for groups of modelling experts to generate 40 (non-ranked) scenarios, from which the IMO study adopted six

[1] Intergovernmental Panel on Climate Change – Special Report on Emissions Scenarios

groups of scenarios; 3 in the first storyline (A1F1, A1B, A1T) and one from each of the other storylines (A2, B1, B2). A Delphi approach comprising shipping experts from around the world was then applied to estimate values for all of the factors that would affect the future traffic levels for the three types of ships under consideration: coastwise shipping (smaller), ocean-going shipping (larger), and container ships (all sizes).

Two methods were considered for the seaborne trade prediction models: (i) GDP-based regression for all three ship types; (ii) a concurrent study by OPRF[2] (2008) that also used IPCC SRES scenarios, and that based container traffic projections on GDP but the cargo transport on other factors such as total population and primary energy use. This latter study produced lower estimates in 2050 than the purely GDP model. For the GHG study, the analysts averaged the outcomes from these two methods. They also interpolated exponentially from the 2050 values to estimate the demand for marine transport in 2020. Finally, using the range of values produced through the two models, that generated an upper bound and lower bound for each ship group, for each scenario, and for the two forecast horizons (2020 and 2050). A sample output (base case) is shown in Table 3, where the values are relative to the seaborne trade in 2007 which is assigned a score of 100. So, under A2 for example, the container global maritime flow (tonne-miles) is projected to be 6.45 times higher in 2050 than in 2007.

Table 3. Projections of tonne-miles used in the IMO-GHG study (2007 = 100) (source: IMO, 2009)

	A1B	A1F	A1T	A2	B1	B2
2050						
Ocean-going shipping	245	245	245	190	185	155
Coastwise shipping	245	250	245	215	185	185
Container	900	875	905	645	615	525
Average, all ships	402	397	403	302	288	247
2020						
Ocean-going shipping	131	131	131	121	120	114
Coastwise shipping	131	132	131	126	120	120
Container	194	193	195	176	173	165
Average, all ships	146	146	146	135	133	127

- An earlier study by Eyring et al. (2005) follows a similar process to the IMO-GHG study described above to forecast global ship emissions to 2020 and 2050. It is also based on four IPCC SRES scenarios (from 2000), to predict the seaborne transport

[2] Ocean Policy Research Foundation, Japan

and related emissions depending on technological improvements, fuel changes, ship size, etc. However, it is worth mentioning this paper because the regression terms are explicitly stated. Using historical time series from 1985 to 2001 to fit the model parameters, the future world seaborne trade in terms of volume in millions tons ($TST_{f,DS}$) for a specific ship traffic scenario DS in a future year *f* can be estimated from the scenario's future $GDP_{f,DS}$ as follows:

$$TST_{f,DS} = -1081 + 0.76 * GDP_{f,DS}$$

To translate this projected annual trade volume to the number of ships, another regression was fit to 1985-2001 data based on the correlation between the number of ships and trade volume. So, the number of ships $N_{f,DS}$ can be related to the total seaborne trade TST in any future year *f* as follows:

$$N_{f,DS} = 74{,}983 + 2.62 * TST_{f,DS}$$

- Sardain (2019) developed a model to examine future spread of invasive species through the global shipping network, up to 2050. The non-indigenous species get transported through ballast water, biofouling of the hull, and for land-based organisms through the cargo itself (ex. wood). Acknowledging that traditionally shipping traffic is viewed as depending greatly on the GDP of the origin and destination regions as well as the inter-region distance, they postulate that other factors could also be play a role in the prediction model.

  They divided the world's marine-coastal countries into 15 socio-ecoregions (SERs) for the shipping traffic source/destination locations, which are regions displaying marine biogeographic and ecological similarity that also incorporate socioeconomic regionalization. They draw their global development scenarios on 5 SSPs (shared socio-economic pathways) from IPCC's Fifth Assessment Report (2014) which provides much of the exogenous input data needed for their regression model. They group the shipping vessels into 7 categories: bulk carrier, chemical tanker, container ship, crude oil/oil products tanker, general cargo ship, LNG tanker, ro-ro cargo ship.

  They developed and tested 5 alternative multiple regression models, selecting the one with the best statistical validation, fitting the model to data for years 2006-2009 and forecasting ship movements for 2014. The resulting choice, named RAUG (residual-adjusted unconstrained gravity) model, is defined as follows:

$$\log[X_{ijst}] = \beta_{s0} + \beta_{s1}\log[GDP_{it}] + \beta_{s2}\log[GDP_{jt}] + \beta_{s3}\log[pop_{it}] + \beta_{s4}\log[pop_{jt}] + \beta_{s5}\log[dist_{ijt}] + \beta_{s6}CL_{ijt} + \beta_{s7}CB_{ijt} + \beta_{s8}CCH_{ijt} + \beta_{s9}RTA_{ijt} + \varepsilon$$

where:

- s : subscript denoting ship category
- i : subscript denoting source region of ship traffic
- j : subscript denoting destination region of ship traffic
- t : year
- β : coefficient for each term in the equation (9 of them)
- X_{ijst} : ship movements of ship type s from source SER i to destination SER j in year t
- GDP : gross domestic product in region i or j, in year t
- pop : population in region i or j, in year t
- dist : great-circle distance between SERs, based on the mean latitude and longitude of all countries within an SER, weighted by each country's population
- CL_{ij} : common language between SER_i and SER_j
- CB_{ij} : common border between SER_i and SER_j
- CCH_{ij} : common colonial history between SER_i and SER_j
- RTA_{ij} : regional trade agreement between SER_i and SER_j

In all cases, source GDP, destination GDP, common language and regional trade agreements had positive relationships with vessel traffic, while distance had a negative relationship. Other variables were only predictive for some ship types. Where these variables were significant, shipping traffic had a positive relationship with common colonial history and an ambivalent relationship with source population, destination population and contiguity. RAUG predicted 90% of the variation when predicting for 2014 for all ship types, and between 59% and 93% for the distinct ship types. They integrated the RAUG model with global socioeconomic development scenarios, IPCC global climate change projections and shipping-mediated marine invasion models to forecast traffic and marine invasions across socio-ecoregions to 2050.

2.2.2. Regional shipping

- Zhang et al. (2019) hypothesized that modelling the cargo shipping traffic in a very active region would be best done by using a spatiotemporal method. The study area in the South China Sea covers about 4.6 million square kilometers and encompasses 60 ports that see at least 10 voyages per month on average (Figure 11). The authors introduce the STARMA (space-time autoregressive moving average) time series model with a spatial element. It serves to characterizes the spatial-temporal dependence between different regions by deriving the interaction between adjacent regions within the system. Based on the time series of traffic flows between ports using AIS data, they group the ports into a hierarchy which characterizes their degree of adjacency to other ports (Figure 12). They develop a spatial weight matrix between the "port nodes" based on a gravity model, an approach commonly used in transportation network models that reflects the amount of traffic between nodes. They test whether the monthly traffic levels between each feasible pairs of origin-destination is autocorrelated in time and/or spatially correlated with the traffic on other routes in the adjacent region.

Using monthly AIS data from 2011-2016 they develop the STARMA parameters based on these functional relationships, using 2011-2015 as the training data and 2016 as the test sample. They also fit a (non-spatial) STL regression model (Seasonal and Trend decomposition using Loess) for comparison. According to three indicators to gauge the forecasting accuracy (Pearson's correlation coefficient (R), root mean square error (RMSE), and mean absolute percentage error (MAPE)), the STARMA model outperforms the STL model showing that the spatial-temporal model can improve the accuracy of the forecast.

- Transport Canada (2021, personal communication) has recently been developing forecasts of vessel traffic in the Salish Sea, in southern British Columbia. The vessels are categorized into twelve types (cargo, container, dry bulk, ferry/ro-ro, fishing, government/research, passenger, pleasure vessels, special ships, tanker, tugs/port, and other), and 15 size categories (measured in Gross Tonnage). Using satellite and land-based AIS data for 2015-2019, combined with other sources that provide additional vessel characteristics, an ARIMA model was generated for each vessel type based on 4 years' of data, with the last year reserved for validation. The units used were nautical miles (Nmi) sailed, as another common measure that counts transits was deemed to be inadequate for larger areas of interest (too many difficulties calculating accurate transit numbers). The ARIMA models are used to project traffic in the area, by vessel type, to 2030. Additional estimates are done to predict routes and volumes associated with likely additional traffic in the coming decade due to increased trade and capacity expansion, which is then added to the base case to determine the overall future traffic.

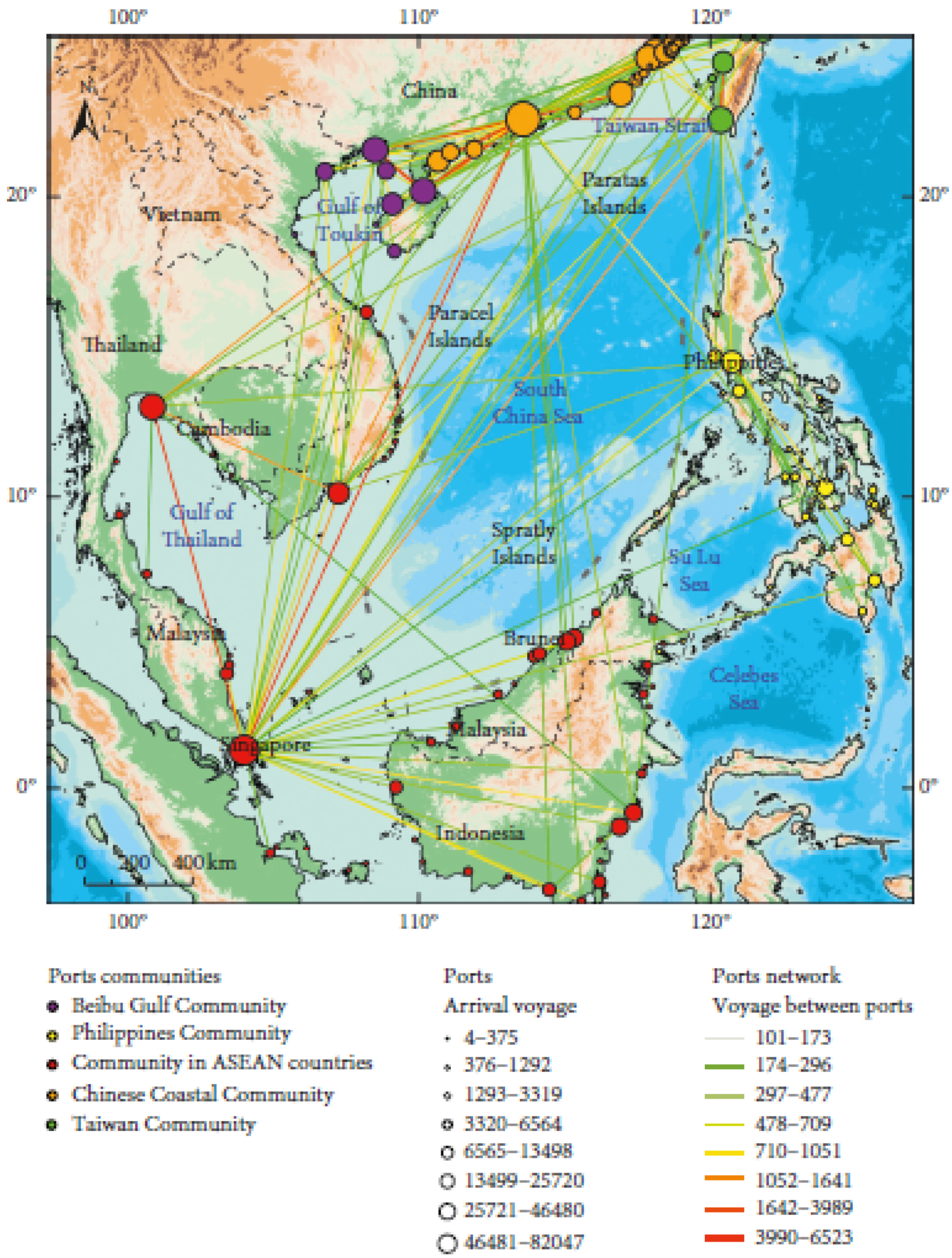


Figure 11. Port node connectivity of the shipping network in the South China Sea (source: Zhang, 2019)

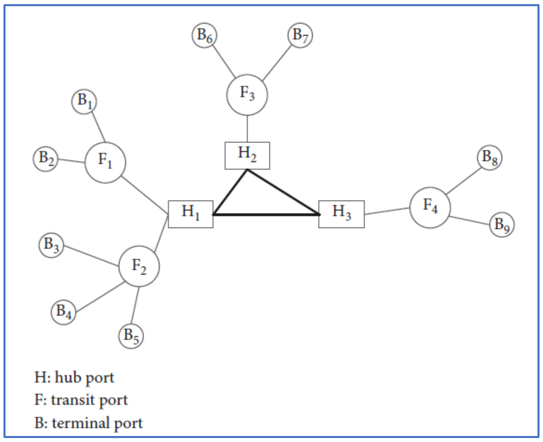


Figure 12. Diagram for port correlation in ship traffic flow network (source: Zhang, 2019)

- Hodgson at al. (2013) used a scenario-based approach explore plausible futures 15 years ahead for marine transportation in the Canadian Arctic. Details are provided in Case study C – Future marine traffic activity in the Canadian Arctic: A Scenario-based approach

- A team at Dalhousie University conducted a large-scale study on maritime traffic patterns around Canada, including in the Arctic. For the northern region, traffic projections were made for 10 years ahead using causal models and Monte Carlo analysis. Several reports were generated on various aspects of the study. Details are provided in Case study B – Future marine traffic activity in the Canadian Arctic: A Monte Carlo Simulation Approach.

2.2.3. Fishing vessels

Soykan et al. (2014) developed a model to predict fishing effort location up to 12 months ahead. Their main objective was to provide predictions that would help law enforcement develop mitigation strategies to reduce illegal bycatch, especially of species at risk, with foreknowledge of where fleets might congregate. The authors developed and tested their model on two Pacific fisheries: drift gillnet (DGN) and albacore troll (AT). For both fisheries, they gridded the respective areas of interest, and acquired data on historical monthly fishing efforts (1981 to 2001 for DGN, and 1991 to 2010 for AT). Data sources include fishing logbooks, satellite data, and NOAA datasets on ocean conditions.

The causal model includes the predictor variables shown in Table 4. Other models on fishing effort had included only static variables, whereas these authors include dynamic variables which they assume might affect fish distribution, hence fishing vessel locations. Similarly, other papers exist on species distribution modelling (SDM) but not related to fishing effort prediction.

The model applied is a Boosted Regression Tree (BRT) which is a machine learning model. A BRT is similar to a Neural Network in that it ingests a lot of data on a set of (independent) predictor variables, applies algorithms that serve to yield the dependant variable, and then the model iterates over time to improve the fit (i.e. it learns), with or without human intervention. Advantages of BRT over other regression models is that they do not require prior knowledge of the system in order to build a model, and they can handle collinear variables, interactions between variables, nonlinear relationships between predictor and response variables, and missing data.

They run the model, reserving 1 year of data for validation, then apply four metrics to assess the model performance. Using a Variable Importance (VI) measure to rate each variable relative to the others in how much it contributes to the explanatory model, they determine:

- that the dynamic variables are about equally important as the static ones;
- a ranking of the variables in descending importance for each of the fisheries, with some static ones being rated highest for the DGN fishery, and some dynamic ones having more importance within the AT fishery;
- that the model can predict well up to 12 months ahead where the fishing efforts will occur with fairly good reliability, which can then serve to help in bycatch enforcement planning.
- that the BRT approach is particularly useful as some of the satellite data sets had a lot of missing data, which can be handled by this type of model.

Finally, they noted that adding vessel-related and socio-economic predictor variables would likely improve the model, which they had begun testing. They also discovered that using a shorter historical time series also gave quite good results, which could be particularly useful for applying this model to emerging fisheries without a long track record.

Table 4. List of predictor variables used in modeling fishing effort (source: Soykan, 2014)

Variable		Abbreviation
Temporally static		
	Latitude	Latitude
	Longitude	Longitude
	Distance to coast	DIstCoast
	Distance to nearest port	NearPort
	Mean depth	Depth
Temporally dynamic		
	Fishing year	Year
	Month	Month
	Northern oscillation index	NOI
	Pacific decadal oscillation index	PDO
	Sea surface temperature	SST
	Sea surface height	SSH
	Sea surface height variability	SSHV
	East-west currents	UGEO
	North-south currents	VGEO
	Eddy kinetic energy	EKE
	Chlorophyll *a* index	CHLA
	Primary productivity	PP
Random number		RN

2.2.4. Cruise ships

Forecasting cruise ship traffic is more akin to forecasting other tourism sectors than other ship type categories. Cruise Industry News (2021) produces a wide variety of reports for sale on cruise activities and trends according to region, types of cruises, etc. Historical data are amassed to characterize the activity by passengers' home countries, cruise destinations, demographics, product offerings (i.e., nature of cruise), and other characteristics. Data are available on every cruise ship worldwide, as well as those on order. Similar high-level information is published by the Cruise Lines International Association (CLIA, 2019) that reports year-over-year growth (in percent) in the market by various factors such as passenger home country. Notwithstanding shocks to the industry such as the Global Financial Crisis (2007- 2008) and COVID-19, most of the prognostication appears to be established based on these historical growth trends as well as available and planned capacity. That is, the global industry continues to grow its fleet, and it assumes a certain vessel fill rate, which is a self-fulfilling prophecy as the companies adjust prices dynamically to help entice more passengers as needed. This of course is predicated on a plentiful supply of passengers, but experience has shown that if the price is right, historically the demand materializes.

The industry, and individual cruise lines, conduct multiple market studies to gauge ongoing interest in their products broken down by multiple categories of users and interests. There are shipping consulting companies that provide analyses and forecasts to many sectors of maritime shipping. MSI (2021), for example, builds market forecasts for cruise ships and other vessel categories based on econometric models (supply and demand) and expert opinion. Such forecasts could include cruise berth demand, fleet supply, earnings, prices, etc., but such predictions would have to be translated into vessel traffic volumes as needed for a vessel traffic forecast. Such predictions serve many aspects of planning for cruises, such as port development, tourism development at destinations, new destinations, and marketing the cruise products. Aside from global industry trends, government and association sources provide historical and projected growth rates by country or region. For example, the CLIA (2019) Northwest and Canada section notes the 36% growth in cruise activity in Canada from 2016 to 2019. The BC Ministry of Small Business and Economic Development and the Ministry of Transportation, in their BC Port Strategy report (2005), predicted a growth in cruise traffic of 50-150% from 2005 to 2020. The Prince Rupert Port Authority Land Use Plan (2020) plays a role as a port of call on the Seattle-Alaska cruise traffic route, and is therefore subject to the vagaries of that market. Several analytical methods found in the literature are presented next.

- Kollwitz and Papathanassis (2011) conducted a review of cruise industry forecasting approaches to date, noting that they often incorporated unverified assumptions, and were typically accepted and used with little questioning of their methods nor validity. They led a Delphi exercise with 10 experts from diverse cruise industry stakeholders to critique the prediction approaches. The main finding was that the methods essentially were based on (cruise trip) supply driven models, as the demand from tourists was implicitly expected to fill the available capacity based on historical patterns in this rapidly growing sector, in conjunction with the cruise ship companies' practice of offering discounts to fill the berths as needed. This modus operandi fails to consider other exogenous factors such as socio-economic drivers, natural or economic disasters that can significantly impede the industry, or competition from other tourism sectors.

- Pavlić (2013) developed a forecast for cruise passengers stopping at Dubrovnik, Croatia, for the purpose of tourism development at the historic city, but also for consideration of vessel traffic impacts. Using a SARIMA forecasting model to account for the pronounced seasonality of the cruise traffic, the author applies it to monthly cruise passenger data from 1998 to 2011 to fit the model parameters (after some tests and transformations to render the time series stationary), and then forecasts the expected arrivals for 2012-2014 inclusive. The author specifies that this 3-year forecast's validity depends on the underlying structural factors remining constant.

- Çuhadar at al. (2014) compared various Artificial Neural Network forecasting models to predict the cruise tourism demand to Izmir, Turkey for one year ahead. Figure 13 clearly illustrates the seasonal nature of the activity as well as the steady growth over the years, which is typical of many cruise destinations. The authors provide a detailed description of three alternative ANN models, Multi-Layer Perceptron (MLP), Radial Basis Function (RBF), and Generalized Regression neural network (GRNN), and points out their structural differences. He then applies the models to the time series of the monthly cruise passenger data in Izmir for 2005 to 2013. They reserve the last 2 years' of data for testing model performance and, based on the Mean Absolute Percentage Error (MAPE), they determine that the RBF model fits the data best, which they then use to predict the traffic for the following year (2014).

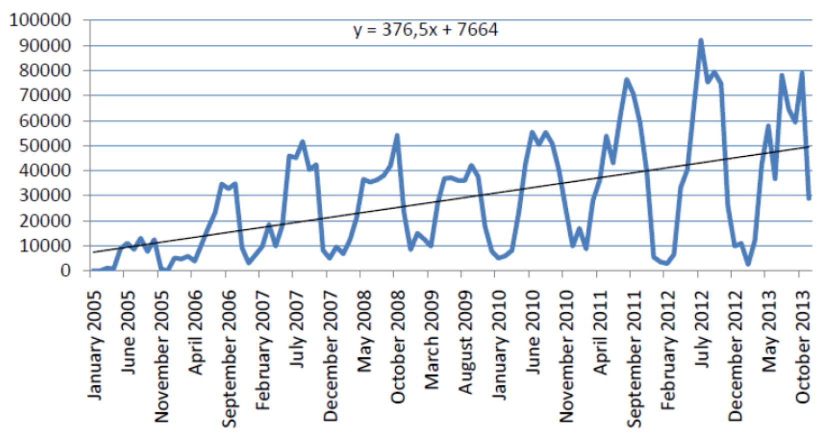


Figure 13. Time Series chart of cruise tourist arrivals in Izmir (2005-2013) (source: Çuhadar, 2014)

- Xie et al. (2021) reiterate that predicting cruise passenger volumes is akin to forecasting in other areas of tourism. They summarize many recent forecasting models applied to tourism (Table 5) that have focussed on "big data" as predictor variables, sometimes combined with other data sources. SQD in the table refers to Search Query Data, which is an increasingly popular means to scrape data from the web through search engines, and combine them into databases on features of interest. In regards to tourism, researchers typically use Google or Baidu to determine what consumers' interests may be focussing on using keywords such as a destination city, famous sites (ex. Forbidden

City), or terms like ‘hotels’ and ‘flights’ in combinations with particular tourist destinations, etc. Tracking cell phone positioning data for tourists also reveals a lot about their preferences and patterns. This type of big data research is an extension to traditional market studies.

The authors apply and compare five forecasting models for weekly cruise tourism demand forecasting: ARIMAX (ARIMA with exogenous variables), BPNN (back propagation neural network), RBF (radial basis function), LSSVR-CV and LSSVR-GSA. The latter two are explained in more detail here as they end up being the best predictors. An LSSVR model (Least Squares Support Vector Regression) serves to simplify very large-scale data sets by mapping them into a higher-level feature space using a nonlinear kernel function. The basic method is to estimate the parameters using a cross-validation approach (hence the LSVRR-CV label). In a variation, they also invoke GSA (Gravitational Search Algorithm) to establish the LSSVR parameters before the training step.

The authors combine three types of data as the predictor variables for their regression models: SQD, tourism volume, and Chinese economic indices (purchasing manager index; consumer confidence index; real effective exchange rate index). To acquire the consumer info, they established a set of keywords (see Table 6) and using the search engine Baidu. They examined the correlation between the Chinese weekly tourist volume and each keyword for 0-12 lag periods (i.e. 0 to 12 weeks ahead of the corresponding tourism), and chose those that had the best explanatory power. For various combinations of input databases, they calculated the performance of the 5 models according to 3 measures: RMSE, MAPE, and WIA (Willmott’s index of agreement (Table 7). They demonstrate that the LSSVR-GSA outperforms all of the other models in both predictive accuracy (on all measures) and generalizability. They conclude that the keyword search results provide a valuable input for cruise predictions, as do the economic indexes.

Table 5. Typical works on tourism demand forecasting with big data (source: Xie, 2021)

Author(s)	Data source(s)	Variable(s)	Forecasting model(s)
Bangwayo-Skeete and Skeete (2015)	SQD	Keywords	AR-MIDAS
Choi and Varian (2012)	SQD	Keywords	Linear regression
Huang et al. (2017)	SQD	Keywords	Linear regression
Li et al. (2017)	SQD	GDFM and PCA-based indexes	Linear regression
Li et al. (2018)	SQD	PCA-based indexes	BPNN
Lv et al. (2018)	SQD	Keywords	SAEN
Park et al. (2017)	SQD	Averaging keywords	ARIMAX
Rivera (2016)	SQD	Summation-based index	DLM
Sun et al. (2019)	SQD	SAS-based index	KELM
Yang et al. (2015)	SQD	SAS-based index	Linear regression
Pan and Yang (2017)	SQD, website traffic, and weather information	Keywords, website traffic volume, and snowy data	ARIMAX
Raun et al. (2016)	Mobile tracking data	Space-time tracking data	Logistic regression
Yang et al. (2014)	Web traffic	Web traffic volume	ARIMAX

Note: autoregressive integrated moving average with exogenous variables (ARIMAX); autoregressive mixed-data sampling (AR-MIDAS); back propagation neural network (BPNN); dynamic linear model (DLM); generalized dynamic factor model (GDFM); kernel extreme learning machine (KELM); principal component analysis (PCA); shift and summation (SAS); stacked autoencoder with echo-state regression (SAEN).

Table 6. Search queries from Baidu (source: Xie, 2021)

Cruise	cruise	Tourism	cruise tourism
	cruise company		cruise strategy
	cruise website		cruise travel notes
	cruise lines		cruise tourism strategy
	Costa Cruise		cruise tour price
	Costa Cruise website		ship tourism
	Costa Serena		ship travel
	Royal Caribbean	Cruise sites	cruise home port
	Royal Caribbean International		cruise terminal
	Royal Caribbean International website		Sanya cruise
	Star cruises		Shanghai cruise
	Viking cruises		

Table 7 The forecasting performance of the models (source: Xie, 2021).

Model	RMSE	MAPE	WIA	RMSE	MAPE	WIA
	Time series			Time series + Economic indexes		
ARIMAX	34475.855	6.384	0.9456	31664.27	5.944	0.9542
BPNN	33836.939	6.261	0.9505	32740.62	5.86	0.9542
RBF	32448.538	6.290	0.9536	30384.029	5.842	0.9605
LSSVR	31811.728	5.858	0.9557	29842.032	5.563	0.9561
LSSVR-GSA	27710.239	5.362	0.9633	23380.277	4.305	0.9669
	Time series + PC keywords			Time series + PC keywords + Economic indexes		
ARIMAX	34023.511	6.278	0.9465	30603.692	5.712	0.9567
BPNN	30227.744	6.127	0.9535	29396.441	5.591	0.9561
RBF	28497.638	5.948	0.9591	27977.112	5.459	0.9633
LSSVR	21957.279	5.304	0.9797	20741.608	5.154	0.9801
LSSVR-GSA	20399.558	4.401	0.9810	17174.676	3.861	0.9869
	Time series + Mobile keywords			Time series + Mobile keywords + Economic indexes		
ARIMAX	33256.406	6.055	0.9487	30493.220	5.596	0.9571
BPNN	32926.785	5.898	0.9464	26661.749	4.783	0.9673
RBF	31548.110	5.866	0.9501	25437.250	4.593	0.9697
LSSVR	29275.590	4.766	0.9556	20004.678	4.376	0.9819
LSSVR-GSA	17465.026	3.861	0.9855	14958.891	3.311	0.9896

2.2.5. Forecasting of specific shipping impacts

As discussed in section 1.4.1, forecasting of shipping activities is often done to explore future impacts of the activity, or temporal cumulative impacts. Aside from predicting changes in shipping activities, forecasts are often warranted to examine future changes in other parts of the system affected by shipping (see Figure 2). The papers presented in this section predict the nature of various elements in the shipping realm, addressing a range of impacts that may result from shipping in the future.

- Iacarella et al. (2020a) develop a model to estimate the transfer of aquatic nonindigenous species between Marine Protected Areas (MPA) via ships and boats through ballast water or hull fouling. Using a year's worth of AIS (Automatic Identification System) ship-tracking data on the west coast of Canada, they develop a network representation of vessel traffic amongst 83 MPAs of interest (Figure 14). Focussing on 8 species of interest that appear respectively in some of the MPAs, they calculate the likelihood of introduction and establishment in other MPAs based on vessel traffic connections and environmental suitability (water temperature and salinity). Then they repeat the calculation under future climate conditions (average conditions from 2041-2070) based on the BC Regional Ocean Modeling System (ROMS) projections. Their results show that the spread and establishment of these nonindigenous species is predicted to be much worse under future environmental conditions than it would be now. Note that this study does not consider changes in shipping traffic in the future.

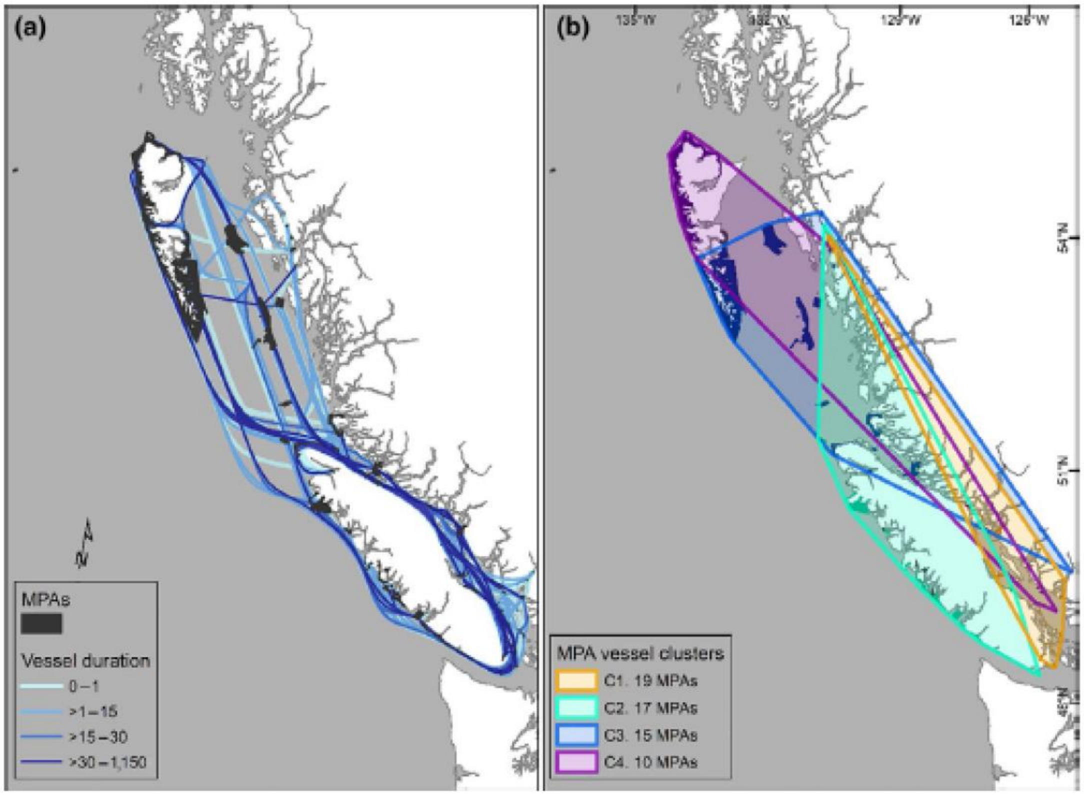


Figure 14. Vessel traffic connections between MPAs are vectors of nonindigenous species across the coast of BC. (a) Connections between MPAs colour-coded by duration (days) within the MPAs and (b) MPAs that are highly linked by vessel traffic are indicated as clusters (source: Iacarella, 2020a)

- Nuka Research & Planning Group (2013) produced a set of three reports for the BC Ministry of Environment for a West Coast Spill Response study. The second report focusses on Vessel Traffic, including forecasts for 5 and 15 years ahead of the baseline year 2012 (i.e. 2017 and 2027). The vessels in the region, recorded through AIS and other databases (ex. for barges), are grouped into 13 categories for ships and other vessels over 20m long, noting that while small vessels are important for oil spill risk, there was insufficient information on them to conduct proper estimates. Rather than relying on port calls, they applied six "passage lines" to capture the traffic spatial distribution (Figure 15). The traffic forecasts are based on ongoing, planned, or conceived capacity expansion of various terminals at ports on the BC Coast, relying on throughput estimates from the port developers. In some cases, the throughput was provided in volume of cargo, which the authors converted to number of vessel transits with some assumptions on vessel sizes. They also include potential growth in the region due to increases in traffic at port in the US Northwest (based on USA sources on projections).

Figure 15. Passage line locations (source: Nuka, 2013)

- One shipping consequence that is increasingly under investigation is the impact of ship noise and vessel strikes on whales. Two reports produced by DFO for the BC coast refer to potential traffic increases. In DFO (2017a), a spatial analysis produces the intersection of ship traffic (as captured by AIS in 2013) and two types of whales at risk, fin and humpback whales, whose distribution was estimated through aerial surveys from 2012-2015. These data permit the calculation of potential risk from whale-ship encounters. Although the data sources and methods are not specified, the authors state that expected increases to maritime traffic (broken down by speed and vessel category) due to fleet and port expansions over one to two decades could result in increases in the relative risk of mortality from a ship strike on the order of 1.2 to 1.8 times for humpback whales, and 1.3 to 2.5 times for fin whales. In 2018, DFO completed a technical review of the potential impact of new tanker traffic associated with the proposed Trans Mountain Pipeline expansion project on Southern Resident Killer Whales. They note that the expansion plan would result in an additional 696 tanker transits per year, with the assumption that the vessel capacity will remain the same as present, and that it will use existing inbound and outbound shipping lanes from the Westridge Terminal. The report explains at length different mitigation options to reduce the risk to the whales, such as speed reductions or ship design improvements.

- Another shipping impact involves grey water release, such as water from sinks, showers, laundry facilities, galleys, etc. VARD and WWF (2018) consider this issue in the Canadian Arctic, for the present (base year is 2016), as well as predicting the situation for 2025 and 2035. Their approach to forecasting is scenario-based, considering current traffic levels by vessel category and drivers for changes to each one based on demand and/or supply capacity. They also add in potential new sources of traffic. A few of the factors underlying potential changes for the vessel categories are listed here as examples:
    - *Tourism sector*: fuel prices; accessibility to Northwest Passage
    - *Mining sector*: planned closure of some operations, and expansion of others
    - *Sealift / community resupply*: increased population; improved shipping facilities in the north
    - *Bulk trade*: possible reestablishment of traffic to the Port of Churchill; increasing bulk shipments through the Northwest Passage
    - *Fisheries*: potential small increases in quotas of commercial fishing in current fishing areas
    - *Government operations*: new icebreakers and AOPS (Arctic Offshore Patrol Vessels)
    - *Research operations*: no increase forecasted
    - *Other operations* (towing, escort, etc.): no increase forecasted.

  Applying the metrics that they developed for potential grey water discharge rates by vessel category, they then produced forecasts of discharges by area for the two time horizons.

- A large-scale study was completed by Herbert Engineering and Environmental Research Consulting (2014) to develop marine vessel incident predictions in light of the proposed Roberts Bank Terminal 2 (RBT2) container facility at Port Metro Vancouver (PMV). The types of incidents considered included allisions, collisions, fire/explosions, drift groundings, powered groundings, founderings and structural failures, vessel failures, and container mishandlings. All types of vessels in the region are considered in conjunction with the anticipated additional container traffic to RBT2. The region was divided into 4 zones to examine differences in traffic mix, volumes, and potential incident rates across them (Figure 16). The principal forecast was conducted for 2030, but with the alternative of 2025 also considered since new regulations on ship fuel tank protection will not have been implemented on many ships by then, but would be by 2035.

The incident predictions (IP) are conducted for five scenarios:

- Base year 2012;
- Increases in vessel traffic due to growth consistent with existing marine facilities, without RBT2;
- Increases in vessel traffic due to growth consistent with existing marine facilities, with RBT2;
- Increases in vessel traffic due to growth consistent with foreseeable new marine facilities, without RBT2;
- Increases in vessel traffic due to growth consistent with foreseeable new marine facilities, with RBT2.

Existing traffic levels are calculated using AIS data and port call data. Current incident rates are estimated using accident and casualty databases (Transportation Safety Board, Canadian Coast Guard, Pacific Pilotage Authority, and Sea-Web), literature, and expert opinions of PMV, stakeholders, and the study authors. For rare incidents, international numbers are used and adjusted to the PMV context. An event is classified as an incident when there is the possibility of damage to life, environment or property. The authors note that different exposure measures apply to various incident types involving container ships: number of ship calls; number of containers throughput in TEU; and time in transit.

They produce tables for each vessel category, for all relevant incident types, for 2025 and 2030, providing detailed explanations of the factors considered for potential changes in traffic levels, port capacity, and incident rates. The incident predictions resulting from this study serve as inputs to subsequent phases of the Quantitative Risk Assessment (QRA) that evaluate the secondary consequences of the events in terms of human and environmental receptors (Figure 17).

Figure 16. Zones within the Vancouver Region (source: Herbert, 2014)

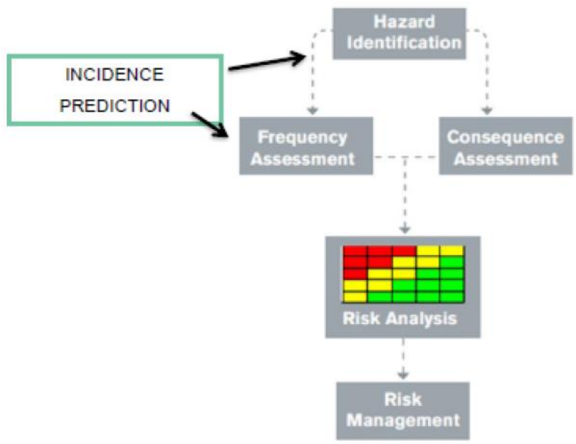


Figure 17. Incidence Prediction in the Risk Management Process (source: Herbert, 2014)

2.2.6. Traffic analysis, but non-forecasting (particularly West Coast and NSB)

The foundation for the quantitative forecasting models (time series and causal) is historical traffic data, amongst other information. This section includes a few reports relevant to gathering or using marine traffic data, with a particular focus on the BC coast. Since 2004, the International Maritime Organization (IMO) has required AIS (Automatic Identification System) use by all vessels >500GT, for any vessel >300GT that is on an "international voyage" and for all passenger vessels with more than 12 passengers. Additional carriage requirements have been added since then, some differing by individual countries. For example, in 2019 Transport Canada expanded its AIS carriage requirements to include any vessel operating outside of sheltered waters that is carrying more than 12 passengers and is eight metres or more in length.

While originally introduced to help with the avoidance of collision and grounding, the extensive databases from collecting AIS signals have since been used for a broad spectrum of applications, including shipping impact studies. Since AIS requires line-of-sight, land-based and ship-based receivers must be within about 40 nautical miles (75 km) of each other and with limited physical obstructions (e.g. fjord systems), so AIS traffic information was only available around coastal zones or between neighbouring ships. With the advent of satellite-captured AIS signals (S-AIS) around 2011, the reach of this monitoring has become global.

- Fournier et al. (2018) published a review paper covering an extensive set of over 200 academic papers on uses of AIS data for various applications (as shown in Figure 18). Aside from the various use contexts, the literature also encompasses articles on modelling, visualization and S-AIS-based traffic simulation. Complementary information can be found through the online AIS handbook prepared by the Task Team on AIS data of the UN Global Working Group on Big Data for Official Statistics (n.d.). This handbook describes methodologies for processing AIS data for various applications, as well as providing a set of case studies. It also provides more effective networking and learning for people involved with this topic.

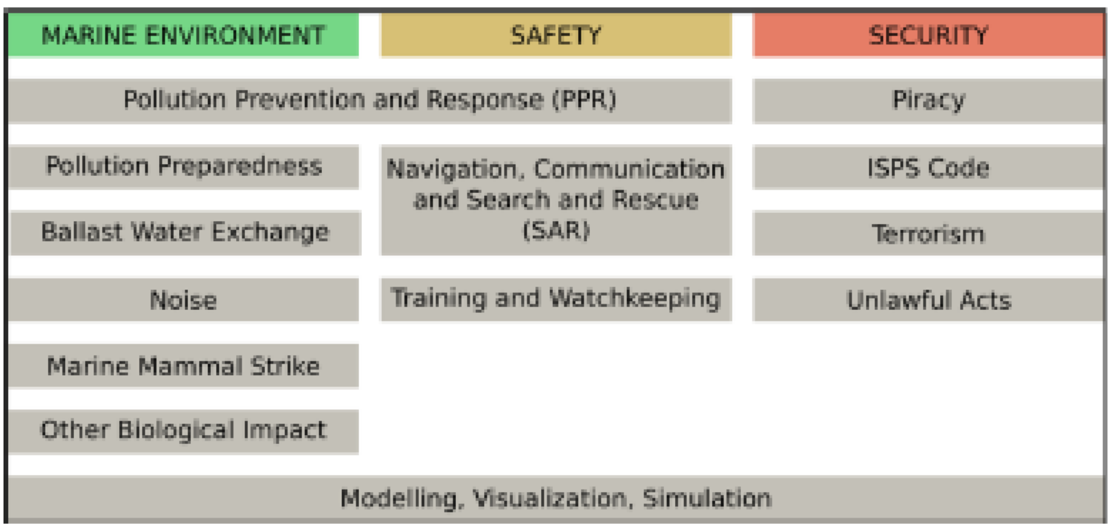


Figure 18. Main topics of the review, using select IMO categories, extended by authors (Fournier, 2018)

- Clear Seas Centre for Responsible Marine Shipping (Clear Seas, 2020) has been undertaking multiple studies to identify and describe risks related to commercial marine shipping in Canada. Hence, they acquired and processed AIS data for three years (2014-2016) for traffic on the West Coast of Canada with the aim of parsing, analysing, and portraying the information in a broad number of ways, which provides very comprehensive insights into the patterns by location, season, cargo type, vessel type, etc. There is a particular focus on the oil contents, whether as fuel or transported (ex. tankers). The commercial traffic included falls into four categories, (broken into some subcategories): cargo, tanker, tug, and cruise ships. Vessel attributes obtained through other databases adds characteristics to the AIS tracks, which allows for many more perspectives on the shipping patterns. The resulting tables, graphs and maps synthesize many characteristics of the vessel activities, which provides useful insights and inputs for subsequent studies. As an example, Figure 19 demonstrates the spatial patterns in the region for four sample months for cargo traffic between Washington State and Alaska (US-AK) in contrast to other US cargo traffic (mostly to Asia) labelled as Innocent Passage.

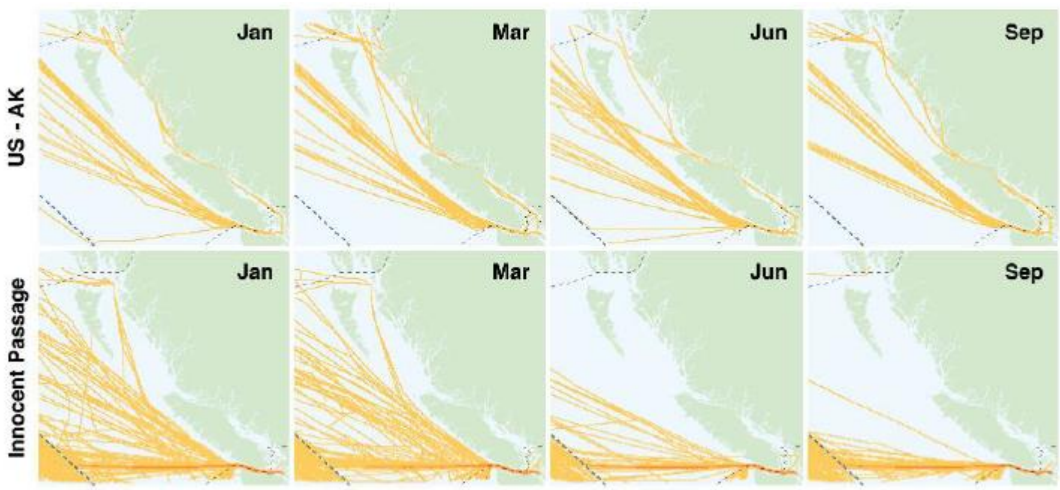


Figure 19. U.S. Traffic for January, March, June, and September 2016 – Cargo Vessels (source: Clear Seas, 2020)

- While the Automatic Identification System (AIS) has been a great boon for monitoring vessel movements for safety, traffic management and capacity planning, it only captures part of the marine traffic since it is only mandatory on certain vessel classes. The mandatory AIS transceivers are dubbed Class-A, but there are also Class-B transceivers that are cheaper and smaller than Class-A, but their range is shorter and their signal transmissions are lower priority if there are competing demands on the communications network from Class-A units. Most boats that adopt an AIS device voluntarily opt for Class-B units, including many fishing vessels and pleasure boats. Thus, while the AIS data provides some information on small vessel traffic patterns, it only provides a small part of the picture for such activities.

  There are other ways to capture small vessel traffic. Serra-Sogas et al. (2018) conducted a detailed study to assess small vessel detection capability in the Salish Sea, BC, via two surveillance methods: areal surveys, and image analysis from satellite data. The National Aerial Surveillance Program (NASP) run by Transport Canada serves to monitor for ship oil spills and coincidentally to serve as a deterrence for potential polluters. Imagery and coincident information such as time and location were processed over 2015 to 2017. The number of vessel detections over the entire area of interest, and also broken down into four subregions, was tabulated. Vessel lengths were estimated as well as vessel speeds. Monthly charts were produced as well as maps to illustrate the temporal and spatial variations of the traffic respectively. Comparing with AIS records for vessels in that same area on NASP surveillance days, it was determined that over 70% of the vessels identified were not transmitting AIS information. Using satellite imagery from Google Earth Pro, 4 images were selected for analysis over the time span from 2014 to 2016, two in the summer and two in the winter. Various procedures were applied to identify and characterize the types of vessels spotted in the images, including smaller recreational vessels.

- Serra-Sogas conducted three studies in 2020 to produce additional information about maritime traffic in the Northern Shelf Bioregion (NSB), two of them based on AIS data, and one on aerial surveys. In one (Serra-Sogas, 2020a), terrestrial AIS and Satellite AIS data sets were fused for traffic in the NSB for 2015-2019. Vessels were grouped into 10 categories: cargo, container, dry bulk, tanker, fishing, pleasure vessels, ferry/Ro-Ro, passenger, government/research, tugs/port. Nautical Miles sailed were plotted monthly for all vessel types combined (with trend lines), then for each type of vessel individually, then for vessel size categories, and finally broken into 4 subregions of the NSB. Another study (Serra-Sogas, 2020b) repeated this exercise, but for 10 smaller focus areas within the NSB and graphing data weekly. Finally, NASP aerial images were processed for 6 areas/routes in the NSB over a 1 to 2 year period (Serra-Sogas, 2020c). For each route, the following outputs were reported: number of surveys conducted; number of vessels spotted, percent of vessels also transmitting AIS; by vessel type (sport fishing, commercial fishing, other, pleasure craft, sailboat, tugboat, passenger vessel, cargo, and tanker); maps of vessel densities (kernel density) for AIS and non-AIS vessels. Interestingly, about 72% of the detected vessels are non-AIS carriers, consistent with the earlier study in the Salish Sea (Serra-Sogas, 2020a).

As shown in this section and throughout this report, AIS is a very rich source of data to examine traffic in an area, for possible applications to a wide range of management, planning or research purposes. However, one of its main deficiencies is the lack of coverage for smaller vessels, which are integral to some types of studies. There are other ways to collect information on small vessels usage such as coastal radar, aerial images, public surveys of boating use, but such data is much harder to collect, analyze, and interpret than AIS data. In any case, such traffic analyses can form the basis for a quantitative predictions of vessel activity.

3. Compendium of reviewed articles and reports

Table 8. Synopsis of reports on marine shipping forecasting.

Author(s)	Year	Article/Report title	Region	Timeline	Vessel type	Data source	Drivers	Method	Outputs	Purpose
Akhtar, J., Bjørnskau, T., & Jean-Hansen, V.	2012	Oil spill risk analysis of routeing heavy ship traffic in Norwegian waters	North Sea & Norwegian Sea	~20 years	chemical tanker, gas tanker, oil tanker, cargo ship >5000 gt (mostly cruise ships), other ships >5000 gt (mostly fishing vessels) and "all other vessels"	AIS	increase in oil transport and gas tanker transport	scenario-building	graphs, # of sailings,	environmental impact (oil spill risk analysis)
Clear Seas	2020	Vessel traffic in Canada's Pacific region	Pacific coast	N/A	cargo, tankers, tugs, and cruise ships	AIS and port calls	existing traffic levels, and potential increases due to port expansions	N/A	density and traffic pattern maps, graphs, various charts	traffic modelling for MSP and risk management (particularly oil spills)
Cox, G.	2010	Forecasting port traffic the safer way: Basing traffic forecasts solely on GDP can be inaccurate – here's an alternative way to more reliably plan for the future	N/A	N/A	all cargo types	GDP and trade trends; infrastructure; structural changes to shipping industry; gvt policy	imports/exports; economics	multi-factoral approach	N/A	discussion piece on forecasting methods and limitations of GDP as sole predictor
Cruise Industry News	2021	2021 Cruise industry news annual report	Global	7 years	cruise	multiple industry, economic, and port statistics	multiple industry, economic, and port statistics	supply/demand scenarios	reports, figures, tables	prospectus for broad variety of end users: cruise companies; tourism companies; governments; ports; etc.

Cruise Lines International Association	2019	2019 Global Market Report	Global	N/A	cruise	multiple industry, economic, and port statistics	multiple industry, economic, and port statistics	expert opinion; historical growth rates	reports, figures, tables	prospectus for broad variety of end users: cruise companies; tourism companies; governments; ports; etc.
Çuhadar, M., Cogurcu, I., & Kukrer, C.	2014	Modelling and forecasting cruise tourism demand to Izmir by different artificial neural network architectures	Izmir port, Turkey	1 year	cruise	port data	N/A	3 artificial neural networks models	# cruise tourists	port city tourism planning
DFO	2017	Assessing the risk of ship strikes to Humpback and Fin whales – west coast of Vancouver Island	Pacific Canada	1 to 2 decades	all compulsory Automatic Identification System (AIS)-reporting ships	AIS for ships and aerial survey data for whales	fleet and port expansion	linear time series modeling	maps/GIS; ships per hour, % increase in accidents	cetacean management/conservation; strikes
DFO	2018	Technical review: Potential effectiveness of mitigation measures to reduce impacts from project-related marine vessels on Southern Resident Killer Whales	BC	if/when TCP (trans-Canada pipeline) develops	tankers	multiple	pipeline expansion project	N/A	maps	impact management for the Southern resident killer whale (contaminants, salmon, oil, noise, strikes)
Dinwoodie, J., Tuck, S., & Rigot-Müller, P.	2013	Maritime oil freight flows to 2050: Delphi perceptions of maritime specialists	global	~40 years	tankers	multiple econometric studies	multiple social, economic, and environmental changes	Delphi, scenario-building	tables	consider impact of curtailing world oil use on its maritime transport
Drewry Shipping Consultants	2020	VFPA long term container forecast: 2020 – 2060 final report	Pacific Canada	40 years	container	multiple	economic growth, GDP	regression on GDP; market share assumptions, scenario building	TEU, graphs	port development (increase capacity to meet future demands)
Engler, C., & Pelot, P.	2013	Analysis of marine traffic along Canada's coasts: Phase 2 – Part 1: Factors influencing Arctic traffic	Canadian Arctic	10 years	all types in the Arctic	multiple gvt and industry reports	demand and supply markets; ice conditions	literature review	drivers and hindrances for changes in traffic levels; reported projections	Monte Carlo forecasting traffic a decade ahead, for safety and security planning

EnviroEmerg Consulting Services	2008	Major marine vessel casualty risk and response preparedness in British Columbia	Pacific coast (Washington, BC, Alaska)	15 years	all commercial ships; oil tankers and barges (>150 GT) and other vessels >400 GT	Vessel Entry and Transit (VEAT) counts for USA/Canada vessel trips, port calls; previous port development studies	port terminal developments, oil pipeline expansions and proposals, and liquid natural gas plants.	driver-based estimates	vessels per year, % growth rate per year, graphs, TEU	policy direction to mitigate environmental damage (e.g. oil spills)
Étienne, L., Pelot, R., & Engler, C.	2013	Analysis of marine traffic along Canada's coasts: Phase 2 – Part 2: A spatio-temporal simulation model for forecasting marine traffic in the Canadian Arctic in 2020.	Canadian Arctic	10 years	all types in the Arctic	multiple gvt and industry reports; Statistics Canada data	demand and supply markets; ice conditions	Causal model, Monte Carlo method and scenario building	spatio-temporal maps, tables and summary statistics for all vessel types	develop basis for forecasting traffic a decade ahead, for safety and security planning
Eyring, V., Köhler, H.W., Lauer, A., & Lemper, B.	2005	Emissions from international shipping: 2. Impact of future technologies on scenarios until 2050	Global	~ 45 years (2005-2020; 2005-2050)	container, tanker, bulk, other (civilian ships >100 GT)	GDP, historic shipping data, AMVER (Automated Mutual-assistance Vessel Rescue system) data	fuel consumption and power installed on shipping fleet	time series regression and scenario building	graphs, # of ships, % growth, TST (total seaborne trade), maps	environmental impact (air emissions)
Fournier, M., Hilliard, R.C., Rezaee, S., & Pelot, R.	2018	Past, present, and future of the satellite-based automatic identification system: areas of applications (2004–2016)	N/A	N/A	all AIS users	Satellite AIS; literature	N/A	N/A	summary of AIS uses by problem area	AIS uses
Gargari, N.S., Akbari, H., & Panahi, R.	2019	Forecasting short-term container vessel traffic volume using hybrid ARIMA-NN model	Ragaee port, Iran	Forecasting review	container	historic traffic observation data	N/A	identifies and compares various models: Neural Network ARIMA model hybrid ARIMA-NN Model	graphs, # of vessels	port planning and capacity management

Hajbi, A.	2011	Traffic forecasting in Moroccan ports	Morocco	~5-30 years	container	import/export data (in tonnage), GDP	economic drivers (coal import)	time series, Delphi modeling	graphs, import/export tonnage	improve efficiency of supply chain
Herbert Engineering Corp. & Environmental Research Consulting	2014	Roberts Bank terminal 2 technical report: Marine vessel incidence prediction inputs to the quantitative risk assessment. File: 2013-032-01 Rev. 2.	BC coast - Strait of Georgia / Southern shore	10 & 20 years	container ships primarily, then all other types of commercial vessels in the area	AIS and port calls; CCG, TSB, IMO, other	economic/port development	scenario building based on potential infrastructure capacity increases	TEU, graphs	environmental impact and safety (oil spills, groundings, vessel collisions, founderings, fires/explosions)
Hilliard, R.C., & Pelot, R.	2012	Analysis of marine traffic along Canada's coasts: Phase 1 final report	Canada's coasts	N/A	all commercial vessels	LRIT & VTOSS	N/A	spatio-temporal mapping of traffic densities	maps, tables	capture the traffic levels and mix in all coastal areas
Hodgson, J.R.F., Russell, W.D., & Megannety, M.	2013	Exploring plausible futures for marine transportation in the Canadian Arctic	Canadian Arctic	~ 15 years	large vessels (no pleasure craft)	N/A	ice melt, tourism, economic development	scenario building	N/A	policy making, shipping governance
Iacarella, J.C., Lyons, D.A., Burke, L., Davidson, I.C., Therriault, T.W., Dunham, A., & DiBacco, C.	2020	Climate change and vessel traffic create networks of invasion in marine protected areas	Pacific Canada, includes NSB	between 2041 and 2070	all	S-AIS 2016 (Marine Traffic)	potential increases in traffic in the region	N/A	connectivity graphs; likelihood of spread of nonindigenous species	environmental impact (invasive species)
Iacarella, J.C., Clyde, G., & Dunham, A.	2020	Vessel Tracking Datasets For Monitoring Canada’s Conservation Effectiveness	BC coast	decades ahead for changes in water conditions	AIS carriers	AIS, DFO, online databases	N/A	shipping patterns constant, but establishment of invasive species changes	maps; percent changes in invasion	transmission of invasive species from one MPA to another through shipping, including under future conditions

Indra, V., Notteboom, T., Parola, F., Satta, G., & Persico, L.	2015	Port traffic forecasting tool	Europe	1-20 years	commercial shipping	N/A	economic, socio-economic, environmental and safety, governance	multiple qualitative and quantitative models reviewed	N/A	synthesis of forecasting tools, and recommendations for short, medium and long term predictions for European shipping traffic
International Maritime Organization (IMO)	2009	Second IMO GHG study 2009	Global	~40 years	coastwise (small vessels), ocean-going (large vessels), container	GDP; IPCC scenarios	population, economy, shipping technology, energy Use	scenario building; Delphi; regression	tonne-mile index, graphs, GDP growth rate, Gross Tonnes (GT)	environmental impact (air emissions)
Jugović, A., Hess, S., & Jugović, T.P.	2011	Traffic demand forecasting for port services	Port of Rijeka, Croatia	~ 30 years	cargo/container	historic data of economic trends and commodity flows	economic drivers (import, export, GDP)	time series, regression, and scenario building	graphs, TEU, % increase	port management & capacity planning
Kinder Morgan	2013	Trans Mountain expansion project (Volume 8a - Marine Transportation)	Burnaby, BC, Canada	~ 6 years	tankers, cargo, passenger (large vessels)	AIS (Marine traffic data)	pipeline expansion project	capacity expansion	maps	risk assessment
Kinder Morgan	2013	Trans Mountain expansion project (Volume 8c - TERMPOL reports)	Burnaby, BC, Canada	~ 15 & 25 years	tanker, cargo, passenger, fishing, other	AIS (Marine traffic data); existing port call data	pipeline expansion project	MARCS for oil spills and scenario building for forecasting	growth rate (% per year), sailed NMi, graphs, maps	oil spill risk assessment
Kollwitz, H., & Papathanassis, A.	2011	Evaluating cruise demand forecasting practices: A Delphi approach	Europe	N/A	cruise	others' reports on cruise ship forecasting	cruise ship berth supply; tourism demand	Delphi method	summary of Delphi findings on forecasting assumptions and approaches	cruise tourism demand forecasting in Europe
Lerner, J.	2018	Review of cumulative effect management concepts and international frameworks	Canadian and global case studies	N/A	N/A	N/A	N/A	N/A	N/A	cumulative effects management
Living Ocean Society	2011	Shipping on the British Columbia coast: Current status, projected trends, potential casualties, and our ability to respond: A briefing report	Pacific Canada	~ 10-15 years	container, bulk, cruise ships	others' projections	terminal expansion and potential pipeline project from Alberta	driver/capacity based modeling	% increase in vessel traffic, inconsistent metrics, graphs	incident response; noise; strikes

Maritime Strategies International	2021	Maritime Forecasting and Strategic Advisory	Global	N/A	multiple vessel types	multiple	multiple	economic outlooks; market research	reports, figures, tables	studies for commercial clients
Maritime Traffic Forecasts Ltd.	n.d.	Key elements of our traffic forecasting approach	as required	as required	as required	as required	as required	historical data; economic drivers	as required	studies for clients on shipping predictions
Mercator Int'l	2018	Roberts Bank Terminal 2 container vessel call forecast study	Pacific Canada (NSB & NWSA)	~ 2, 7, 12, & 18 years	container	container volume forecasts (TEU) from OSC (2016) study; historic data; many industry and government reports	increase in trade, Asia-PNW corridor, new vessel technology, restructuring of carrier alliances	scenario building	TEU, port calls; graphs, maps	port development & terminal capacity
Metro Vancouver	2015	Short sea shipping in Metro Vancouver	Vancouver	no forecast, but considers Vancouver's 2040 planning horizon	RO-RO, barges	N/A	alleviate land transport; cost-effectiveness; transit times	N/A	discussion of options and tradeoffs	alleviating some land transport
Ministry of Small Business and Economic Development & Ministry of Transportation	2005	BC port strategy	Pacific Canada	~ 15 years	container, cargo, cruise	Port Aauthority traffic profile (tonnes), paying cruise passengers	economic growth (Asia-Pacific trade)	driver based estimates	graphs, TEU, tonnage, % change, # of passengers (for cruises)	port planning for sustainable economic growth
Minton, G., Folegot, T., Cosandey-Godin, A., Jacob, T., Lancaster, M., & Ushio, M.	2021	Shipping and cetaceans: A Review of impacts and mitigation options for policymakers and other stakeholders	Global	N/A	commercial and fishing	AIS and VMS	N/A	N/A	statistics; charts	Environmental (Cetacean) management, risk assessments and modeling; noise; strikes

Nuka Research and Planning Group	2013	West coast spill response study Volume 2: Vessel traffic study	Northern and Southern ports in BC	5 and 15 years	13 vessel categories, with a focus on tankers	AIS, MCTS, passage lines	port development (economic) and updating port infrastructure (capacity)	driver based modeling	maps, graphs, vessel transits by passage lines, # vessels per year	environmental impact (oil spills)
Ocean Policy Research Foundation (OPRF)	2008	Research study: The world's changing maritime industry and a vision for Japan	Global	~ 60 years	container, LNG, bulk, tanker	GDP	population, economy, technology, energy, land-use, and agriculture	scenario building	graphs, cargo/trade volumes, # of vessels, TEU	environmental impact (emissions)
Ocean Shipping Consultants	2016	Container traffic forecast study – Port of Vancouver, 2016	Pacific Canada - Port of Vancouver	annually for 10 years, then every 5 years for another 25 years	container	historical data, economic growth, market share data	market share, regional GDP, import and export commodities	regression, scenario building	graphs and tables of container volume (TEU)	capacity planning for the port
Parola, F., Satta, G., Notteboom, T., & Persico, L.	2020	Revisiting traffic forecasting by port authorities in the context of port planning and development	EU + International	~ 5-30 years	mostly cargo and container	port traffic data	N/A	synthesis of multiple methods	N/A	synthesis of forecasting methods
Pavlić, I	2013	Cruise tourism demand forecasting - The case of Dubrovnik	Dubrovnik port, Croatia	3 years	cruise	historic cruise passenger numbers	N/A	seasonal ARIMA model	tabulated forecasts of cruise tourism	management, policy adjustments
Peng, W. & Chu, C.	2009	A comparison of univariate methods for forecasting container throughput volumes	Taiwan	1-2 years	Container throughput (TEU)	Port data	N/A	classical decomposition, trigonometric regression model, regression model with seasonal dummy variables, grey forecast, hybrid grey model, SARIMA	TEU	short term (1-2 years) operations planning for the ports

Port of Vancouver	2017	ECHO Program: Estimating the effects of noise from commercial vessels and whale watch boats on Southern Resident Killer Whales	Pacific Canada - Port of Vancouver	N/A	N/A	N/A	N/A	N/A	N/A	Environmental (cetacean) management, impacts from noise
Port of Vancouver	2018	Roberts Bank terminal 2 project overview and rationale	Pacific Canada	~ 20 years	container	historic data; others' forecasts	increase in trade, port development	reports on others' studies: driver-based modeling & scenario building	TEU, graphs	port planning and management
Port of Vancouver	2021	Vancouver Fraser Port Authority view on container forecast volumes	Pacific Canada	40 years	container		Roberts Bank expansion	hybrid model	TEU, graphs	port development (increase capacity to meet future demands)
Prince Rupert Port Authority	2020	Land Use Plan	Pacific Canada-Prince Rupert (NSB)	20 yr plan (shows 10 yr traffic forecasting)	container/ liquid bulk and passenger (cruise) vessels	unknown	trade with the Asian market (supply/demand), port growth	N/A	TEU, graphs	land use/port development
Sánchez, R.J., Perrotti, D.E., & Fort, A.G.P.	2021	Looking into the future ten years later: Big full containerships and their arrival to South American ports	Global & South America	back 10 years	container ships	multiple	Infrastructure, economics, technology, environment	linear predictive models	tables	general port capacity to handle large ships
Sardain, A., Sardain, E., & Leung, B.	2019	Global forecasts of shipping traffic and biological invasions to 2050	Global	~ 35 years	bulk carriers, chemical tankers, container ship, oil tankers, cargo ship, LNG tanker, ro-ro cargo,	AIS; IPCC scenarios	increase in wealth and global population further increase trade and demand for goods. socio-economic drivers, GDP	multiple regression; global development scenario building	graphs, # of vessel movements	environmental impact (invasive species)
Seaport Consultants Canada Inc.	2013	Projections of vessel movements	Pacific Canada	~ 20 years	mainly cargo, dry/liquid bulk & container although other vessel types are discussed (tugs, tankers, gvt, fishing, passenger)	historic AIS data (Marine Traffic)	traffic trends and potential capacity expansion	percent annual growth rates by type of vessel	graphs, % growth	Trans Mountain expansion project proposal

Serra-Sogas, N.	2020	Analysis of vessel traffic trends within the Northern Shelf Bioregion	Pacific Canada (NSB)	4 years	all	AIS and S-AIS	N/A	time-series with regression	bar charts and line graphs of time series by vessel type and vessel size	illustrating spatio-temporal vessel patterns in NSB for 5 years
Serra-Sogas, N.	2020	NASP aerial vessel surveys in NSB	Pacific Canada (NSB)	N/A	sport fishing, fishing, pleasure, sailboat, tug, passenger, cargo, tanker & other	NASP + AIS	N/A	N/A	monthly histograms, maps	illustrating spatio-temporal vessel patterns in NSB for 5 years
Serra-Sogas, N.	2020	Vessel traffic analysis for focus areas within the NSB	Pacific Canada (NSB)	5 years	all	AIS and S-AIS	N/A	time-series with regression	Line graphs of time series by vessel type and vessel size	illustrating spatio-temporal vessel patterns in NSB for 5 years
Serra-Sogas, N., Canessa, R., O'Hara, P., Smallshaw, L., & Warrior, M.	2018	Small vessel traffic study in the Salish Sea and Southern Resident Killer Whale critical habitat	Pacific Canada - Salish Sea	N/A	pleasure craft (includes those without AIS)	NASP aerial surveys, AIS and satellite imagery	seasonality, sport fishing	N/A	maps	ecosystem management; noise
Soykan, C.U., Eguchi, T., Kohin, S., & Dewar, H.	2014	Prediction of fishing effort distributions using boosted regression trees	International	~ 6-12 month	fishing vessels	satellite/occurrence data; historic logbook data; NOAA ocean data	fishing pressure	boosted regression trees	maps, % change	fisheries management & conservation; bycatch intervention
Tarber, J.	2013	Termpol 3.2–Origin, destination & marine traffic volume survey Trans Mountain Expansion Project	Pacific Canada - Strait of Georgia	~ 20 years	tankers, cargo, bulk & container, although other vessel types are discussed (tugs, gov, fishing, passenger)	historic AIS data (Marine Traffic); port call info; planned expansions	economic drivers (supply/demand), terminal expansions, Trans Mountain pipeline expansion	Percent annual growth rates by type of vessel	graphs, ship calls, % growth, TEU, sailed NMi	Trans Mountain expansion project proposal
The Tioga Group & Hackett Associates	2020	2019-2050 Bay area seaport forecast	USA	~ 30 years	cargo vessels (containerised, Ro-Ro, dry bulk, break bulk and liquid bulk)	import rates	economic drivers, demand for certain commodities, terminal expansions	scenario building	graphs, TEU	port development

van Dorsser, C., Wolters, M., & van Wee, B.	2012	A very long term forecast of the port throughput in the Le Havre – Hamburg Range up to 2100	Europe	~ 90 years	N/A	GDP, population	demographic, economic, or industrial developments	causal regression; Monte Carlo, expert opinion	graphs	sustainable port/infrastructure development
Vard Marine Inc. & WWF-Canada	2018	Canadian Arctic greywater report: Estimates, forecasts, and treatment technologies	Canadian Arctic	10 & 20	all vessels	S-AIS and NORDREG vessel data	resource development, tourism, ice melt, capacity development	scenario building	graphs, maps, GIS, % increase of vessel traffic	environmental impact (marine pollution)
Walsh, C., Lazaroub, N-J., Traut, M., Price, J., Raucci, C., Sharmina, M., & Agnolucci, P., Mander, S., Gilbert, P., Anderson, K., Larkin, A., Smith, T.	2019	Trade and trade-offs: Shipping in changing climates.	16 regions of the world, decomposable by country	30 years	container, dry bulk, wet bulk	commodity trade volumes and routes; IPCC scenarios;	others' climate change scenarios, and socio-economic pathways	scenarios, causal and econometric models; stakeholder engagement	line charts for the three vessel categories for all scenarios	environmental sustainability of shipping sector
Wang, S., Wang, S., Gao, S., & Wenguo, Y.	2017	Daily ship traffic volume statistics and prediction based on Automatic Identification System data	Shanghai port, China	1 month	cargo, passenger, tanker	AIS	N/A	time series and hybrid modeling	graphs, # of vessels	port/operational management
WorleyParsons Canada Services Ltd.	2011	Projections of vessel calls and movements at Deltaport and Westshore terminals	Pacific Canada - Strait of Georgia (Deltaport and Westshore Terminals)	20 years	containers and cargo (coal)	ship movement and ship call data	coal market trends; container ship size trends	scenario building, assuming growth rates of various drivers	TEU, % growth rate in ship calls, graphs	capacity expansion consideration
WSP UK Inc.	2020	Long-term container traffic forecast, 2020-2060	Pacific Canada & USA	40 years	container	historical Canadian west coast port container data; GDP; many other	regional GDP, market share;	regression on GDP; Monte Carlo; scenario building	tables & graphs (TEU)	port development (increase capacity to meet future demands)

Xie, G., Qian, Y., & Wang, W.	2021	Forecasting Chinese cruise tourism demand with big data: An optimized machine learning approach	China	N/A	cruise	Search Query Data; economic databases; historical tourism volume	market demand; economic conditions	multiple regression and ANN models	tables of models' performance	tourism and government planning
Zhang, C., Huang, L., & Zhao, Z.	2013	Research on combination forecast of port cargo throughput based on time series and causality analysis	a southern port in China	4 years	N/A	cargo throughput (Mt)	foreign trade system	Grey forecasting; logistics growth curve	predicted average annual increase in trade (%), graphs	predict trade and port development
Zhang, X., Chen, G., Wang, J., Li, M., & Cheng, L.	2019	A GIS-based spatial-temporal autoregressive model for forecasting marine traffic volume of a shipping network	South China Sea	short-term	cargo	AIS	port and traffic management	spatial-temporal regression	GIS, maps, Graphs	safe navigation and the sustainable development of shipping

4. Case studies

This section includes five case studies, which describe in more detail elements of predictions done using different methods and in quite different contexts. Recall that the aim of this report is not to give a comprehensive review of each study with respect to its aims, assumptions, outputs, etc., but to provide sufficient information to properly describe the prediction methods that were used for each problem type.

4.1. Case study A – VFPA – RBT2 Container traffic

Two recent studies commissioned by the Vancouver Fraser Port Authority (VFPA) include forecasts of container traffic in the port out to 2060 as part of the estimation for future terminal capacity expansion needs, in particular for Roberts Bank Terminal 2 (RBT2). As both of the reports were undertaken to address the same aim, they are both included in this case study to present a number of aspects and approaches to conducting these predictions. Some of the sections below are common to both studies, while others explain the details specific to each report.

i. ***Study title(s) and author(s):***

WSP UK Ltd. (2020, October). Long-term container traffic forecast, 2020-2060 final report.

Drewry Shipping Consultants Ltd. (2020, October). VFPA long term container forecast: 2020 – 2060 Final report.

ii. ***Aim of the study:*** The objective of both studies is to forecast the container flows through the terminals at the VFPA from 2020-2060, while also highlighting anticipated changes at other regional competitive ports including at Prince Rupert and the Northwest USA.

iii. ***Geographic scope:*** Both studies are predicated on considering trade factors at a range of geographic scopes as follows in increasing order of specificity:
- Global: A global perspective is required for several reasons, including: consideration of future worldwide economic growth (driving trade); major regional shifts in production; global changes in the shipping industry (ex. vessel size); global disruptions to shipping (ex. pandemic).
- Continent-wide: The major ports of North America (NA) compete in many respects for some common markets, such as the American Midwest. Thus the

studies consider the GDP forecasts in different regions of the continent, along with multimodal infrastructure, to assess how changes in that balance might affect the VFPA container throughput.

- Regional: The relative desirability and capacity of other ports in the region, notably in Prince Rupert and the West Coast of the USA, bear directly on the anticipated demand for container flow through VFPA.
- Local: The studies' outputs are of course focussed on the VFPA activities now and in the future, but other local factors include actual and potential container handling capacity and multi-modal connections.

iv. ***Time horizon:*** The forecasts for container traffic at VFPA are from 2020-2060. Some parts of the analyses are shorter term, some go to 2050, and others to 2076.

v. ***Vessel type(s):*** Container ships.

vi. ***Method:*** The main method at the core of these studies is that of making predictions based on the relationship between container volume at a port or in a region and economic growth as measured by GDP. These primary findings are then explored using a variety of sensitivity analyses to establish reasonable upper and lower bounds to the predictions.

The WSP approach: (see Figure 20 and Figure 21)

- Based on historical data, use regression analysis to establish that the import (inbound) and export (outbound) through VFPA are highly correlated with the GDP in various regions of North America. The ratio of TEU to GDP over a given time period is referred to as a multiplier.
- Use GDP forecasts from external sources for each of these three major domestic regions (Western Canada; other Canada; USA) to determine compound annual growth rates (CAGR) over the prediction time horizon.
- To account for growing uncertainty over longer timeframes, assume that the multipliers for each origin/destination (OD) are reduced in the future based on expert opinion, to generate a more conservative TEU prediction in the short term (to 5 years), medium term (to 15 years) and long term (to 40 years), This is referred to as 'softening' the multipliers.
- For the base case, assumptions are made in regards to how the market share of container volume held by VFPA may change relative to the present over the forecast horizon. This is broken down by loaded inbound flows, loaded outbound flows, empty flows, and relative to competing ports.
- Using the Monte Carlo method, perform sensitivity analyses on the two key variables (GDP, and TEU/GDP multipliers) for each year for the six main

loaded flow OD combinations (2 directions X 3 regions). A PERT[3] distribution is used for each flow; for its inputs, the most likely value corresponds to the model base case values, and the low and high values are calculated based on historical fluctuations in each parameter. The output for each of the 6 flows is a probability distribution of the CAGR over a 30-year horizon.

- The forecast TEU for each flow is then presented over the 30 years with the most likely value (median) and upper (P90) and lower (P10) ranges. These are then combined to give the overall stochastic (i.e. probabilistic) range on the estimate for VFPA container traffic.
- In the first decade (2020-2030), continued globalization and Asian demand for both imports and exports are assumed. Subsequently, three scenarios are then considered to explore the impact of larger macro-economic changes that may affect container flows through VFPA in the long run (from 2030-2050). The Scenario-adjusted High, Base, and Low case are based respectively on different combinations of these seven factors (which are each described in depth and estimated in the report): economic growth outlook; trade propensity outlook; pandemic recurrence; global financial crisis recurrence; trade agreement cessation; local labour disruption; shifting manufacturing trends away from China; strategic market focus of the Port of Prince Rupert. Each case is translated in annual TEU container volumes through VFPA to 2076.

[3] PERT is a probability distribution based on maximum, minimum and most likely values, and it is a particular case of a BETA distribution. It appears like a skewed bell curve.

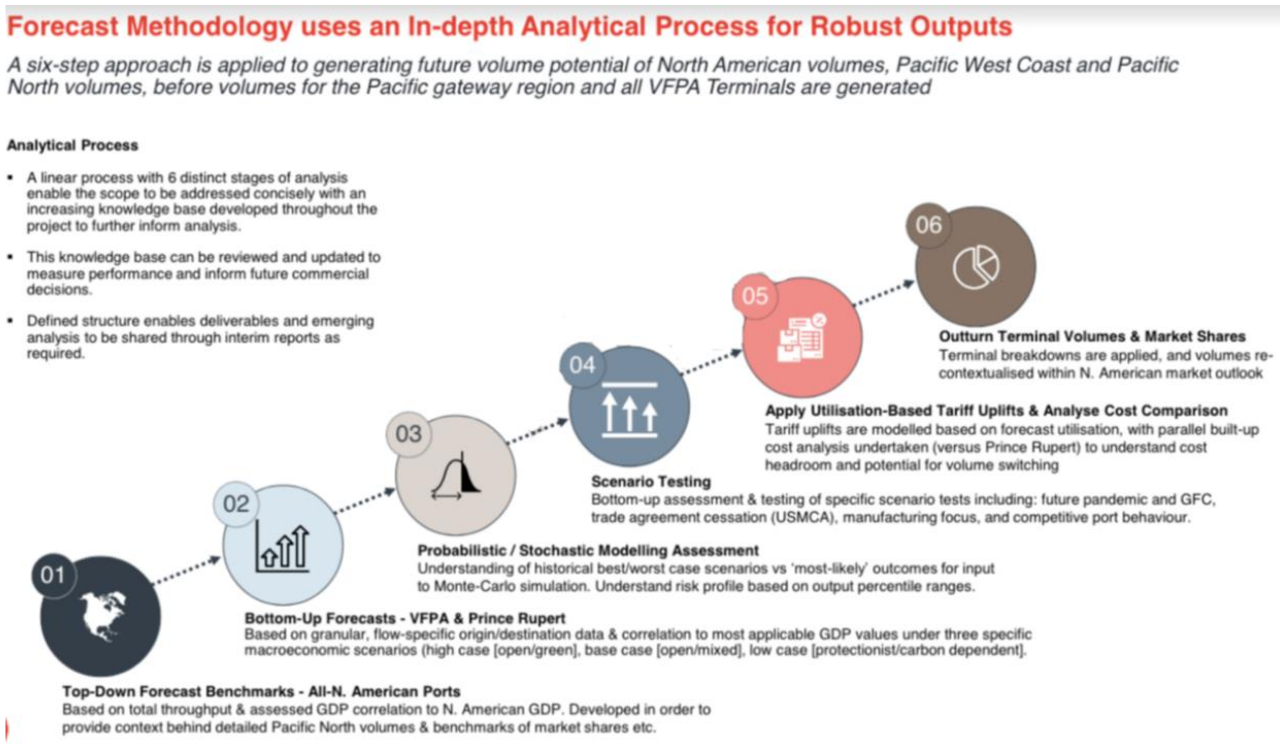


Figure 20. WSP forecasting methodologies applied in several stages (source: WSP, 2020)

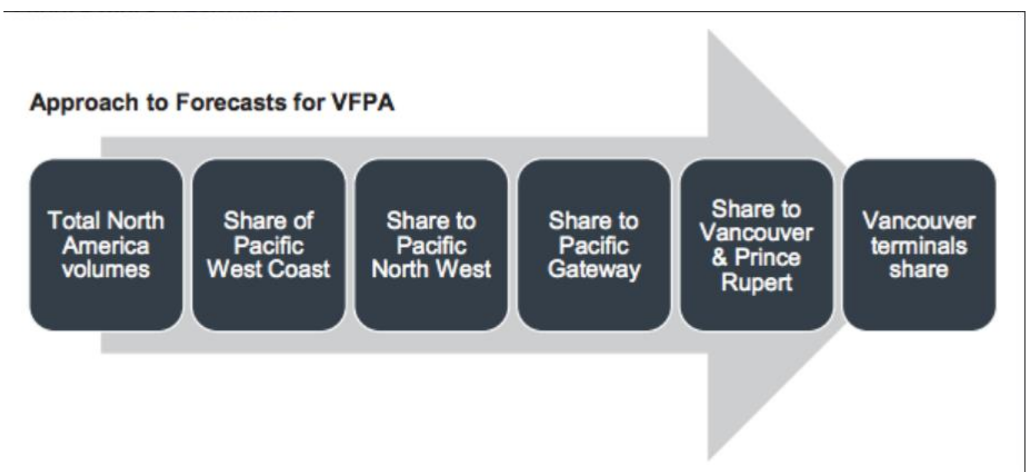


Figure 21. WSP: Top-down market share split for containers (source: WSP, 2020)

The Drewry approach:

- Regression analysis is used to establish the historical patterns for the following variables: shipped container volumes relative to GDP; Canadian trade for specific commodities that are shipped in containers; USA containerised volume in tonnes.
- Distinct rates are calculated by flow direction (import/export) and trade routes (ex. Asia, Europe, Oceania).

- The base case forecast starts at the high level with economics forecasts for Canada and the USA (CAGR for 2019-2025, and for 2025-2060), then projections by international trade routes, with overall container throughput broken down into coastal region shares for Canada (East and West Coast) and the USA (5 coastal regions), which is then split into port shares for the Pacific Northwest region (VFPA, Prince Rupert, and Seattle/Tacoma).
- Distinct forecasts for containers are produced for full/empty and import/export TEUs, with the empty ones being based on a percentage of the loaded ones.
- Scenario analysis is then conducted to incorporate large-scale risks or opportunities that could have a long-term impact on VFPA container volumes. Ten important factors are derived from six possible areas of focus as shown in Figure 22. Base, low, and high cases are then defined through various choices of these ten factors through the process illustrated in Figure 23. Adjustments are made accordingly to the CAGRs, and the resulting forecasts are graphed to 2060.

Figure 22. Areas of focus for long-term scenarios development (source: Drewry, 2020)

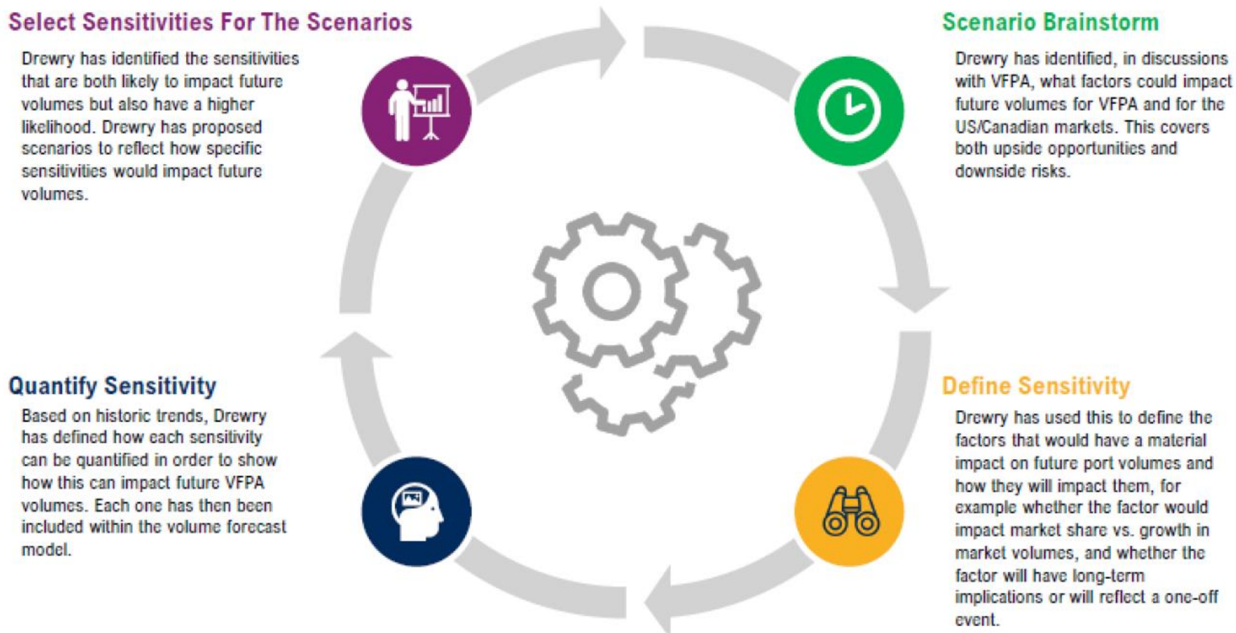


Figure 23. Process for converting scenario analysis into predicted impacts on VFPA container volumes (source: Drewry, 2020)

vii. ***Data sources***:

Given the extensive scope of these studies, there is a very large number of data sources included, and listing them all here would not be useful. Some key ones however are provided to illustrate the variety of sources:

WSP:

- IMF (forecasted GDP growth values)
- Oxford Economics (world GDP growth)
- Statistics Canada, World Bank, Port of Vancouver, WSP (containerised exports by commodity)
- US Bureau of Economic Analysis (container port demand growth)
- VFPA (for container throughput by market)

Drewry:

- IHS Markit (for economic assumptions and trade statistics)
- Oxford Economics and IMF (for economic assumptions)
- VFPA (for container throughput by market)
- information published on port authority websites (such as the Northwest Seaport Alliance)

- Drewry's internal databases (container throughput for the US and Canada markets; TEU)

viii. ***Output format***:

The main output of the two reports are tables and graphs of projected annual container volume through VFPA in TEU over the forecast horizon. Low case and high case projections are also provided. Ancillary outputs include forecasts by imports, exports, full vs empty containers, predicted throughput of VFPA by market region, and relative to other ports. WSP commented on total ship calls, which is not directly proportional to TEU throughput as ship sizes are expected to increase (which they also project).

ix. ***Comments***:

These in-depth studies both relied on a variety of techniques to project container throughput across several decades. Fundamentally, they are causal models with the container traffic being predicated on the GDP in North American regions, considering apportionment across trade routes and competing ports. Thus forecasting both the germane GDP values, and the ratios of TEU/GDP (i.e. multipliers), provides the base outputs for VFPA prediction needs. This approach is nicely complemented by sensitivity analysis, whether by Monte Carlo to incorporate variability in these factors, or by scenario analysis to consider larger shifts in the economic, trade and operating environments.

It is notable that VFPA produced a bulletin (Port of Vancouver, 2021) stating that although the two studies generated somewhat different outputs, they both used sound methodologies and given the complexity and long forecast horizon it is not unexpected to have some variance. Ultimately, the VFPA plans to use a compromise forecast that lies between the two studies' projections.

4.2. Case study B – Future marine traffic activity in the Canadian Arctic: A Monte Carlo Simulation Approach

i. ***Study title(s) and author(s):***

This forecasting model is included within a set of reports prepared for DRDC (Defence Research and Development Canada) on the "Analysis of Marine Traffic" (AMT), of which the first report listed here modelled all types of marine traffic on Canada's three coasts, the second one presents the background and data required as inputs to the forecasting model, and the third report describes the prediction model and results.

Hilliard, R.C. & Pelot, R. (2012). *Analysis of Marine Traffic along Canada's Coasts: Phase 1 Final Report.*

Engler, C., & Pelot, R. (2013). *Analysis of marine traffic along Canada's Coasts Phase 2 – Part 1: Factors influencing Arctic traffic.*

Étienne, L, Pelot, R., & Engler, C. (2013). *Analysis of marine traffic along Canada's Coast Phase 2 – Part 2: A spatio-temporal simulation model for forecasting marine traffic in the Canadian Arctic in 2020.*

ii. ***Aim of the study:*** Enhancing safety and security in Canadian Arctic waters involves monitoring maritime activities and improving preparedness to respond to unexpected events. Both of these strategies have long lead times, therefore it is important to not only consider current maritime activity levels in the North, but to produce shipping traffic forecasts so that resources for monitoring traffic and managing incidents can be effectively deployed over time. To this end, projections of traffic densities were conducted for 7 years ahead (i.e. for the year 2020).

iii. ***Geographic scope:*** The Area of Interest (AOI) corresponds to waters north of 60° latitude, extending up to and including 300 nautical miles (nmi) from the shoreline of Canadian territory, and also encompasses Hudson Bay and James Bay.

iv. ***Time horizon:*** The analysis was conducted in 2013, using data and reports up to then, for a prediction of shipping traffic in 2020.

v. ***Vessel type(s):*** The study prepares distinct forecasts for eight vessel categories: merchant/bulk/cargo; tankers; fishing; cruise ships; pleasure craft; government (non-military); research and exploration; tugs and service vessels.

vi. ***Method***:

Starting with a baseline of current (2013) traffic types, traffic levels, locations and activities in the Canadian Arctic, scenarios are built 7 years in the future using Monte Carlo simulation considering changes in the activity levels (based on relevant drivers) and expected ice conditions (for spatio-temporal feasibility). The model development process is as follows.

a. *Mapping the current levels of Arctic shipping traffic*: The data derived from the Coast Guard managed Long Range Identification and Tracking (LRIT) database provides time-stamped position information about large ships. Smaller vessel traffic (pleasure craft and fishing vessels) estimates were generated separately using incident data and population.
b. *Identifying the drivers of Arctic shipping traffic*: The specific activities that determine the demand for each type of vessel traffic were developed. For example, a mine operation in the North would engender a certain number of bulk carrier trips per month.
c. *Projecting changes in the identified drivers*: The potential changes in activity levels for each of the drivers over a 7-year horizon were estimated from publicly available sources. For example, a mine expansion or a new mine would increase the shipping traffic.
d. *Projecting changes in Arctic shipping traffic*: The estimated impact on maritime traffic for each driver was determined from the literature review. The specific parameters used as the basis of the simulations required setting a range (minimum and maximum values) for each factor, and a most likely value (mode) somewhere between these extremes.
e. *Determining the spatial distribution of current and projected traffic*: To allow the maximum flexibility in ship routing in the Arctic, a network model (graph) was constructed which allows potential traffic to travel between Zones of Interest (ZOI). ZOI represent destinations of the traffic (such as a community or mine), origins of the traffic (such as “gateways” into the arctic area of interest on the East and West sides), or areas where activity takes place (such as a fishing zone). Traffic can then be simulated on this network, as a function of the demand for services, and the feasibility or the routes. (see Figure 24)
f. *Considering current ice coverage and projected changes due to climate change*: The present-day ice conditions and projections were acquired from Canadian Ice Service data. The model requires information on the spatial extent, temporal variation throughout the year, and ice type in order to determine accessibility of various ZOI.
g. *Considering feasibility of passage*: The feasibility of passage through any given area at a specified month of the year is determined not only by the ice conditions, but also by the bathymetry and the ice class of the vessel. Thus scenarios for 2020 traffic levels are run for ships with different ice classes.
h. *Monthly Traffic Spreading*: As the Arctic ice melts, it is expected that maritime traffic will spread out somewhat over the longer summer season. For example, 10 vessels/month for 2 months could translate into 5 vessels/month for 4 months, but this reassignment across periods is done probabilistically. The Ice Numeral predictions (for ice conditions) for 2020 were used to compute which links (graph edges) would be navigable for each ship type in each month. If there is some traffic on a given link during one month, then this traffic is expected to spread out across previous and subsequent months (up to two months before and two after), as long as the link is also navigable these preceding and following

months. However, this simulation exercise does not increase the total amount of traffic on the link, as changes in the overall traffic levels is done in the prediction stage.

i. *Monte Carlo (MC) simulation*: As described in Section 1.4.4.4, a Monte Carlo simulation was run for each vessel category, for each month, for each feasible route, for the year 2020. The potential change from the current traffic level was either based on a multiplicative factor for an existing activity, or an additive factor for a new activity or new route. The high, mode, and low range for each driver variable were used to define a triangular distribution from which the random selection was made for the possible future traffic outcomes. Running the model thousands of times by sampling from the distributions for each input factor generates overall estimated traffic patterns throughout the north, with the output itself represented as the most likely value (mode) by ship type, location and time (month), as well as the statistical range for each of these output values.

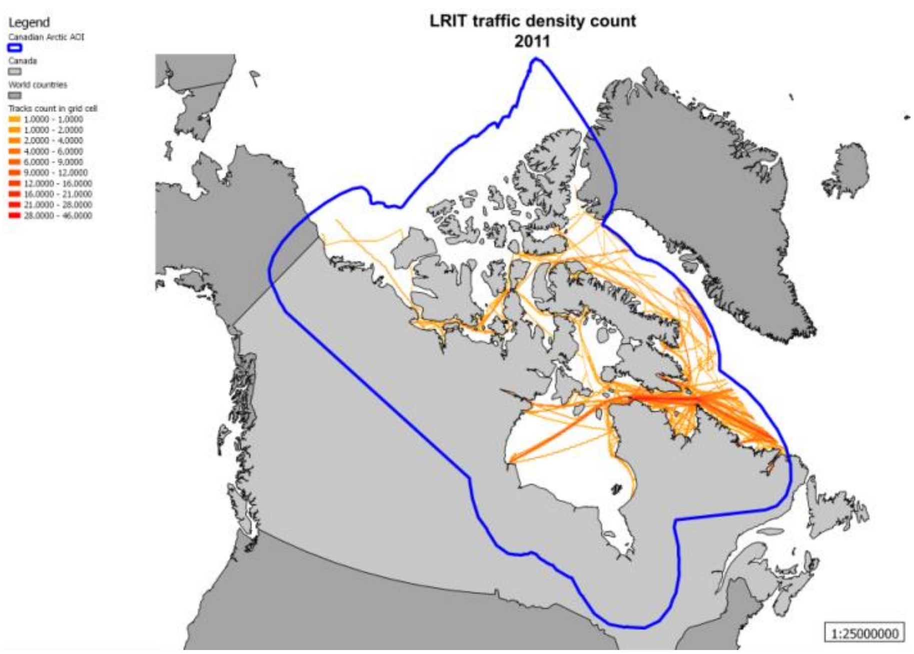


Figure 24. Northern traffic routes based on historical data (source: Étienne, 2013)

vii. ***Data sources***: A broad range of databases and documents were referenced to produce this model, including:
- LRIT (long range identification and tracking) data supplied by the CCG provides time-stamped position information vessels that fall under SOLAS
- Literature (academic, government, consulting, and industry reports)
- Information from Statistics Canada

- Ice data from Canadian Ice Service
- SISAR incident data from the CCG

viii. ***Output format***:

Many output types were generated due to the richness and complexity of the resulting forecasts for 2020. Three of the main products are shown below:

- The feasible Canadian Arctic regions that can be traversed in a specified month for a given types of ice-class vessels (see
- Figure 25);
- The monthly distribution of traffic in 2020 expressed as a frequency distribution of the number of tracks from the MC simulation (see Figure 26);
- Spatial traffic density maps for 2020 (see Figure 27).

All of these output formats can be customized by vessel type, time period, or location.

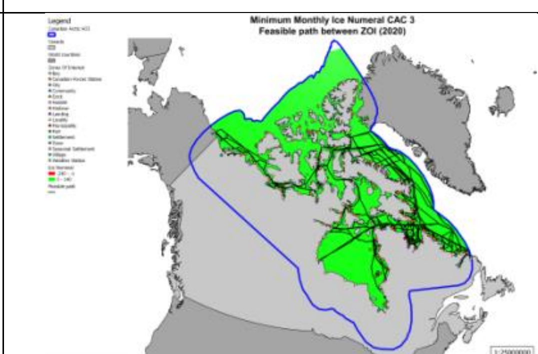
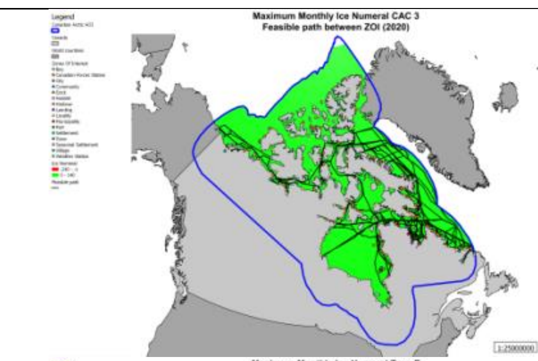
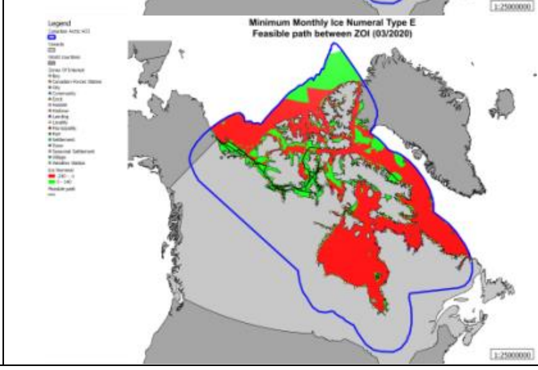
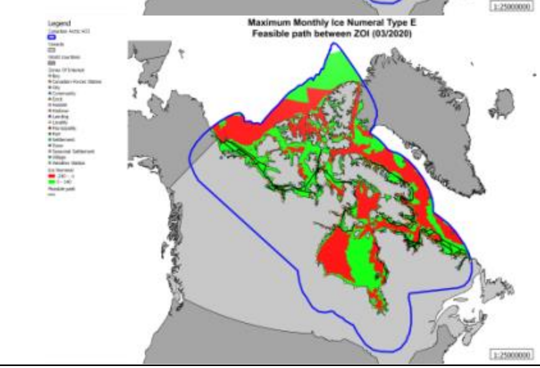
March 2020	**Minimum Ice Numeral**	**Maximum Ice Numeral**
CAC3		
Type E		

Figure 25. March 2020 feasible path comparison for Type E and CAC 3 ships (source: Étienne, 2013)

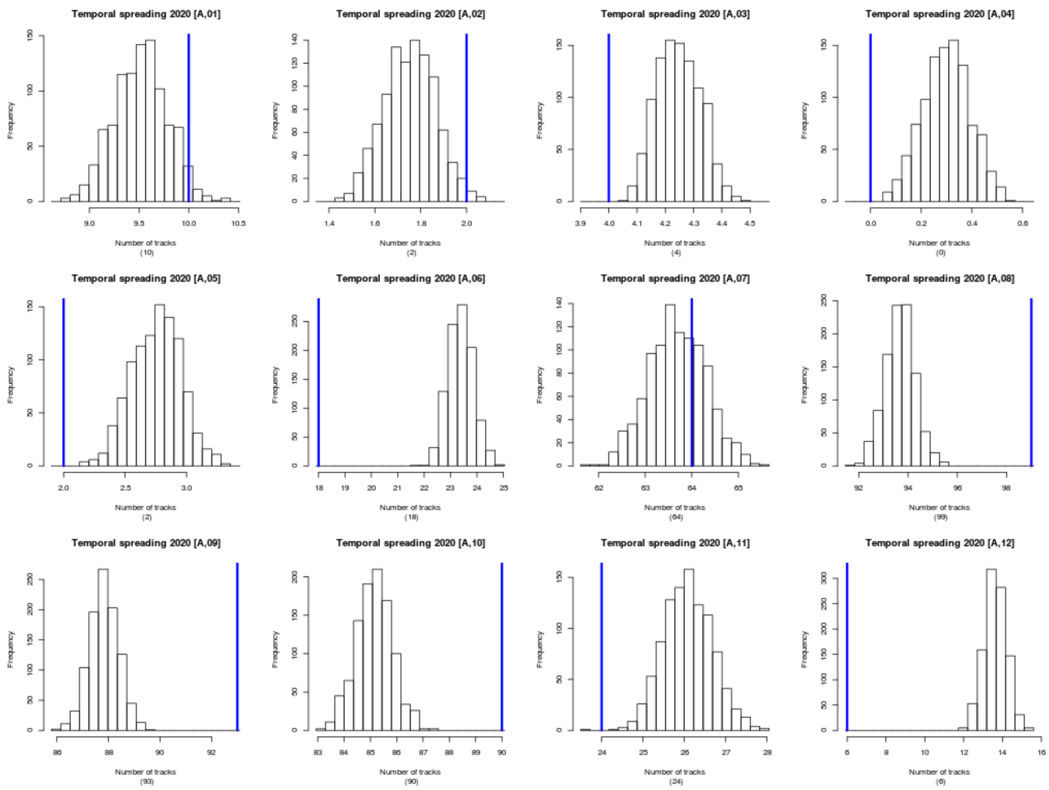


Figure 26. Monthly spatio-temporal MC traffic simulation histograms (All ships) (source: Étienne, 2013)

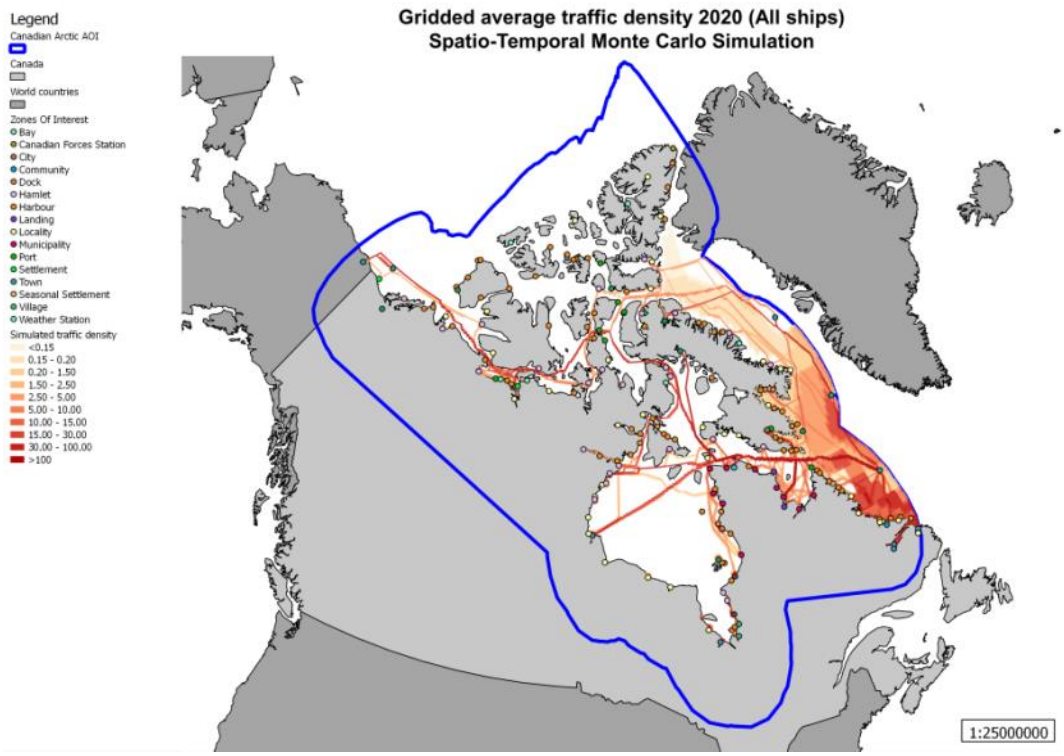


Figure 27. Average yearly gridded simulated traffic density for 2020 (source: Étienne, 2013)

ix. ***Comments***:

This causal model provides a very good paradigm for basing future traffic characteristics on current patterns, altered by potential changes in the main drivers for each activity type. The Monte Carlo simulation approach has several benefits including being able to incorporate several drivers for each activity type as needed, allowing for uncertainty in any of the factors through a probabilistic distribution, and being able to easily combine the outputs for all traffic types.

There are, however, several limitations to this study. The relationship between an activity level (and associated vessel transits) and its driver(s) is conjectured based on literature and experience. For example, the dependency between population growth in northern communities and the number of vessel visits required for resupply is hard to estimate. Such estimates could be strengthened though consultation with industry-specific experts using the Delphi method for example (see Section 1.4.4.6). Furthermore, the model parameters can easily be adjusted to run what-if scenarios to explore the effect of changes to any factor included in the model. This study assumes that there are no significant changes in 2020 traffic due to infrastructure developments, technological changes, nor regulatory evolution. For a medium-term prediction of 7 years, without any advance notice of such imminent changes this is probably an acceptable assumption.

4.3. Case study C – Future marine traffic activity in the Canadian Arctic: A Scenario-based approach

i. ***Study title(s) and author(s):***

Hodgson, J.R.F., Russel, W.D., & Megannety, M. (2013). “Exploring plausible futures for marine transportation in the Canadian Arctic a scenarios’ based approach”.

ii. ***Aim of the study:*** The aim of the study is to envisage potential maritime traffic in the Arctic generally, and the Canadian Arctic in particular in a medium-term timeframe. It employs a scenario-based approach to incorporate possible directions for multiple driving factors, globally and regionally, that may collectively affect the various plausible scenarios considered. Thus there are two complementary goals to this work: using the generated future views to help stakeholders improve preparedness such as policy directions, infrastructure development, northern marine management planning; and to demonstrate the utility of this scenario-based approach and assess its potential usefulness for continued application.

iii. ***Geographic scope:*** The geographic focus of the generated scenarios is the Canadian Arctic, but consideration of the entire Arctic and global factors related to maritime transportation are also referenced.

iv. ***Time horizon:*** The report states that it was looking 15 years ahead to 2025, so presumably the projection was done a couple of years before the report was finalized in 2013.

v. ***Vessel type(s):*** Although the vessel types are not explicitly stated, this study considers all forms of commercial maritime transportation including for goods and passengers (cruise ships). A few ship characteristics, such as vessel age and size, are brought up to provide insight on future changes in the global fleet composition and the potential implications. A passing reference is made to fishing fleets.

vi. ***Method:*** The scenario-based approach adopted in this study is predicated on the notion that future shipping in the Arctic will be subject to a broad range of significant changes in coming years with respect to multiple “forces” encompassing market motivation, governance and public interest, and thus it cannot be adequately predicted using traditional forecasting methods. The alternative is to generate future scenarios which are ‘plausible, challenging and contrasting’ considering potential directions in these three large “forces” at the global, entire Arctic, and Canadian Arctic scales (Figure 28).

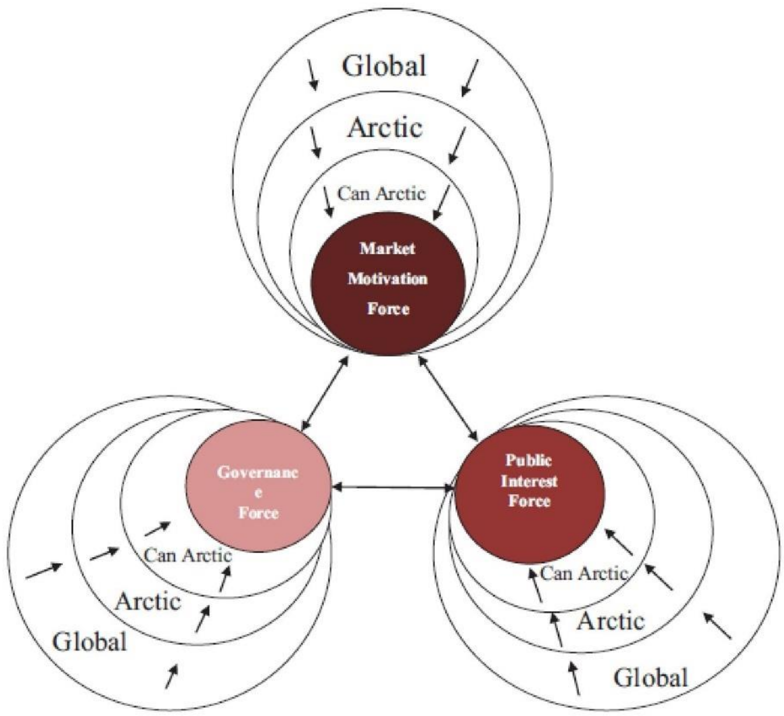


Figure 28. Interaction between the three “Forces” (source: Hodgson et al., 2013)

The “market force” representing the demand for transportation services includes five main variables: the health of the world economy; the amount of annual seaborne commodity trades; the average haul per vessel; the impact of political events affecting transportation routes and infrastructure; and transport costs, including the price of fuel oil. The shipping services supply side encompasses size of the world fleet, level of fleet productivity, rate of ship building production, number of vessels scrapped or lost, and the freight rates for various cargo types.

The “public interest force” includes the following five key considerations: climate change; environmental protection and response; safety of life; security and sovereignty; and northern communities. Each of these is broken down into its major elements for more in-depth explanations which will then serve to consider how future associated changes in each factor may affect shipping volumes, activities or management practices. For example, the environmental aspect covers vessel strikes, noise disturbance, invasive species, air and marine pollution, and impacts specific to icebreaker operations.

The “governance force” presentation explores shipping-related agencies at the international, national, and regional levels, including the roles of States, bodies such as the IMO, ILO and WTO, and other germane organizations. Since shipping governance is particularly complex for many reasons, including the interplay between different levels of government, significant changes can occur in the constraints on, or conversely the support for, shipping operations in the Arctic in coming years. Notably, the governance overview also addresses the implications for the adoption of rapidly developing marine technologies.

All three of these forces are examined, through a broad review of the literature and consultation with subject matter experts, from a global perspective, then concentrating on Arctic-specific aspects, and finally focussing on the Canadian Arctic context (Figure 28). Woven throughout the chapters is the essential identification of which factors pose major uncertainties as they pertain to shipping activities in the north.

The crux of the method is to then develop several scenarios 15 years into the future, considering the relative dominance of the three main forces. Figure 29 shows where each of the four defined scenarios sits in relation to the importance of the three forces. For example, Scenario 1 assumes that in 15 years the maritime governance structure continues to function well and improve (such as the effectiveness of the Arctic Council and the cooperation between governments and industry), the market force strengthens through for example northern resource extraction, resupply and cruise business, while the focus on public interest issues wanes for a variety of reasons. Each of the scenario assumptions is logically argued, and the potential conditions are envisaged considering the uncertainties associated with each factor. This leads to a set of scenarios that are each plausible and that are sufficiently different to offer a range of insights to help inform current planning processes. A tenet of the method is not to prioritize any of the scenarios, as it’s not meant to be a prediction tool of what the future *should* look like as much as what it *could* look like.

The scenario-based method and outcomes of the study were then validated with a group of experts. The method itself was well received, and the outputs were judged to be informative. Nevertheless, it was not clear how such speculations could be translated into more concrete actions in practice. Suggestions were made to consider specific government policies or other levers that could be applied to help direct the shipping activities to more desirable states and hedge against the forces and factors that might be disruptive in the future.

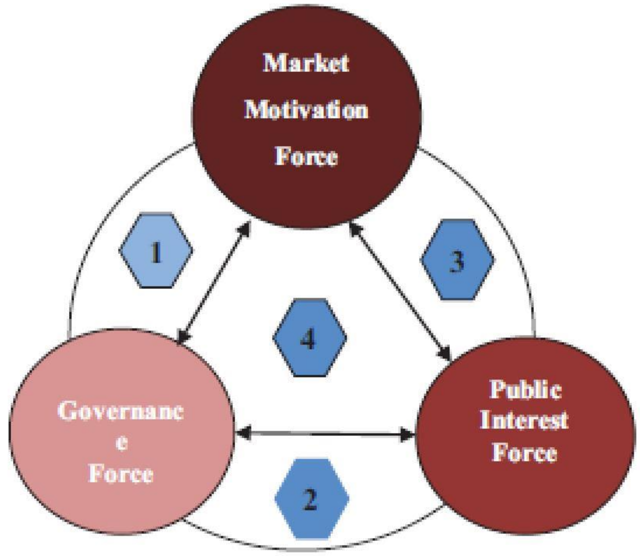


Figure 29. Four possible scenario variations (source : Hodgson et al., 2013)

vii. ***Data sources:*** No quantitative data was used per se as modelling was not used. However, the literature review yielded copious amounts of information on the status and trends of many of the underlying factors references in the "forces" shaping the scenarios.

viii. ***Output format:*** The outputs from the study include descriptions of the various scenarios that were developed and the implications for policy development and planning purposes.

ix. ***Comments:*** In the report's summary, this is stated: "More particularly, it has hopefully been demonstrated that simple forecasting is not the only, nor necessarily the best, way of looking at the future. A further hope is that it has been clearly shown that there is value to be gained by providing insights into the various 'forces' at play with a given context and how the identification of the uncertainties associated with each of those forces can lead to a productive debate regarding plausible scenarios."

While it is indeed very valuable to consider in any projection exercise what the implications would be of significant changes to such fundamental "forces" and their interplay on the activity of interest, this does not obviate the need for and utility of prediction-based modelling, especially in the short to medium term. In fact, these approaches are complementary as the latter assumes that the current activities, relationships and drivers would be substantially the same for many years into the future while accounting for some variation, as demonstrated in Case Study 4.2.

Individual factors that are expected to arise or change substantially, such as traffic transiting the Canadian Arctic in the future can also be incorporated into such models. However, such mathematical models largely rely on the current and foreseeable underlying relevant "forces" remaining steady, which may not be the case, especially as we project further into the future. This is where the scenario-based method can offer additional insights for a much broader range of underpinning conditions which may significantly affect the activity of interest. This is discussed further in Section 5.

Another challenging aspect of the scenario-based approach is our limited human information-processing capability. Given the very large number of factors and uncertainties that were incorporated into each of the four proposed scenarios, it is very difficult for us to understand the implications of the confluence of these variables. Furthermore, while there is a deliberate attempt to generate a set of plausible scenarios that are neither deliberately positive nor negative, the question remains why these four scenarios, when there is almost an infinite set that could be considered, some of which are no doubt much more likely than others, a feature that is not included in this method. However, in terms of forcing our thinking into considering many possible independent futures, it serves a good purpose in reflecting on our preparedness for multiple eventualities.

4.4. Case study D – Spatio-temporal traffic forecasting based on AIS

i. ***Study title(s) and author(s):***

Wang, S., Wang, S., Gao, S., & Wenguo, Y. (2017). "Daily ship traffic volume statistics and prediction based on Automatic Identification System data."

ii. ***Aim of the study:***

This study demonstrates that a traffic prediction for a 1-month period in a localized area with a significant amount of AIS-tracked vessels can be done reliably using various statistical methods. Furthermore, the authors compare which analysis method works best using standard prediction model performance metrics. The authors state that predicting daily ship traffic volumes over the short term is beneficial for traffic management.

iii. ***Geographic scope:*** Shanghai Port, China

iv. ***Time horizon:*** 30-day prediction

v. ***Vessel type(s):*** Passenger ships, cargo ships, tankers

vi. ***Method:*** Taking a 5-month AIS data set (2016.9.27-2017.3.02) in the region of Shanghai, extract the data points for the 3 chosen vessel types that fall within a prescribed bounding box around the port area. Using the required AIS fields, determine which vessels drove into the area each day (i.e. within a 24-hour period), which ones drove out, and which stayed within the port area.

Apply these three modelling approaches to the first 4 months of data, to predict the 5th month's traffic volume:

- Auto-Regressive Moving Average (ARMA)
- Artificial Neural Network (ANN)
- A hybrid approach in which the ARMA model deals with linearity and the ANN model deals with the nonlinear residuals generated by ARMA

The performance of each of the 3 methods is calculated using the 30-days prediction with the actual traffic on those days, according to these three metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). The authors conclude that the hybrid forecast model is more accurate for cargo ships and tankers, but the passenger ship data were insufficient to reliably model.

vii. ***Data sources***: AIS data

viii. ***Output format***:
The prediction results are shown relative to the actual outcomes via line graphs of daily traffic volume over a 1-month period, for two types of ships and for the three analysis methods: ARMA, ANN and the Hybrid model. The performance of the forecasting methods are tabulated.

ix. ***Comments***:
The methods are interesting, and useful as a template for other studies where AIS data are available, the traffic is relatively dense, and the aim is to produce a short term forecast.

4.5. Case study E – Future Scenarios for Global Shipping Traffic

i. ***Study title(s) and author(s):***

Walsh, C., Lazaroub, N.-J., Traut, M., Price, J., Raucci, C., Sharmina, M., Agnolucci, P., Mander, S., Gilbert, P., Anderson, K., Larkin, A., & Smith, T. (2019). "Trade and trade-offs: Shipping in changing climates."

ii. ***Aim of the study:***

This study generates a set of four scenarios of potential commercial traffic levels globally for each of three decades into the future (until 2050) dividing the world trade into 16 regions. This perspective can help the maritime sector become more environmentally sustainable through international policy initiatives and green technology developments.

iii. ***Geographic scope:*** The study encompasses worldwide commercial shipping based on 16 regions, but the results can also be broken down by country.

iv. ***Time horizon:*** The scenario-based projections are conducted for each decade from 2020-2050, with 2010 as the base year.

v. ***Vessel type(s):*** The study covers all commercial shipping, broken down by three main vessel types: dry bulk, wet bulk, and containers.

vi. ***Method:***

- The study generates a set of four projections of global shipping traffic by region (16), country, commodity, and decade (2020-2050) based on future climate change (CC) and socio-economic development hypotheses.
- For global warming, two Representative Concentration Pathways scenarios are used, representing temperature increases by the end of the 21st century of 2°C and 4°C respectively. These are combined with different Shared Socio-Economic Pathways (SSP) scenarios, to generate the 4 scenarios used in this study, namely: Green Road (GR) and Middle Road (MR2C) at 2°C, and High Road (HR) and Middle Road (MR4C) at 4°C.
- The crux is that the long-term environmental, energy and socio-economic pathways can provide the basis for how wealth and trade may change over the regions of the world, and by extension how this will affect maritime trade volumes and routes. The translation from driving factors into trade volumes is

done by considering the anticipated flows of different commodities as categorized by the International Standard Industrial Classification (ISIC) of All Economic Activities.

- The shipping is divided into three categories, for two main reasons: (i) because the impact of these distinct future scenarios are quite different on categories of commodities and hence on the ship types required; and (ii) the projections of shipping capacity needs are based on different models for the three ship types. The three shipping categories are for energy commodities, containerised commodities, and non-energy non-containerised commodities.
- The future shipping scenarios in this study result from a blend of quantitative and non-quantitative projections. To give a sense of the models, energy commodities are projected using the TIAM-UCL global energy system model; containerised goods' projections are computed using the correlation between countries' growth rates of trade, GDP and number and weight of containers containing such goods; the balance of tradeable commodities, namely non-energy, non-containerised goods are projected using elasticities obtained from econometric estimations.
- Extensive stakeholder engagement served to help frame and generate the scenarios, provide expert opinion, information and datasets, and to help validate the models.
- The elements of the scenarios generation are shown in Figure 30, reproduced from the Walsh et al (2019) paper.

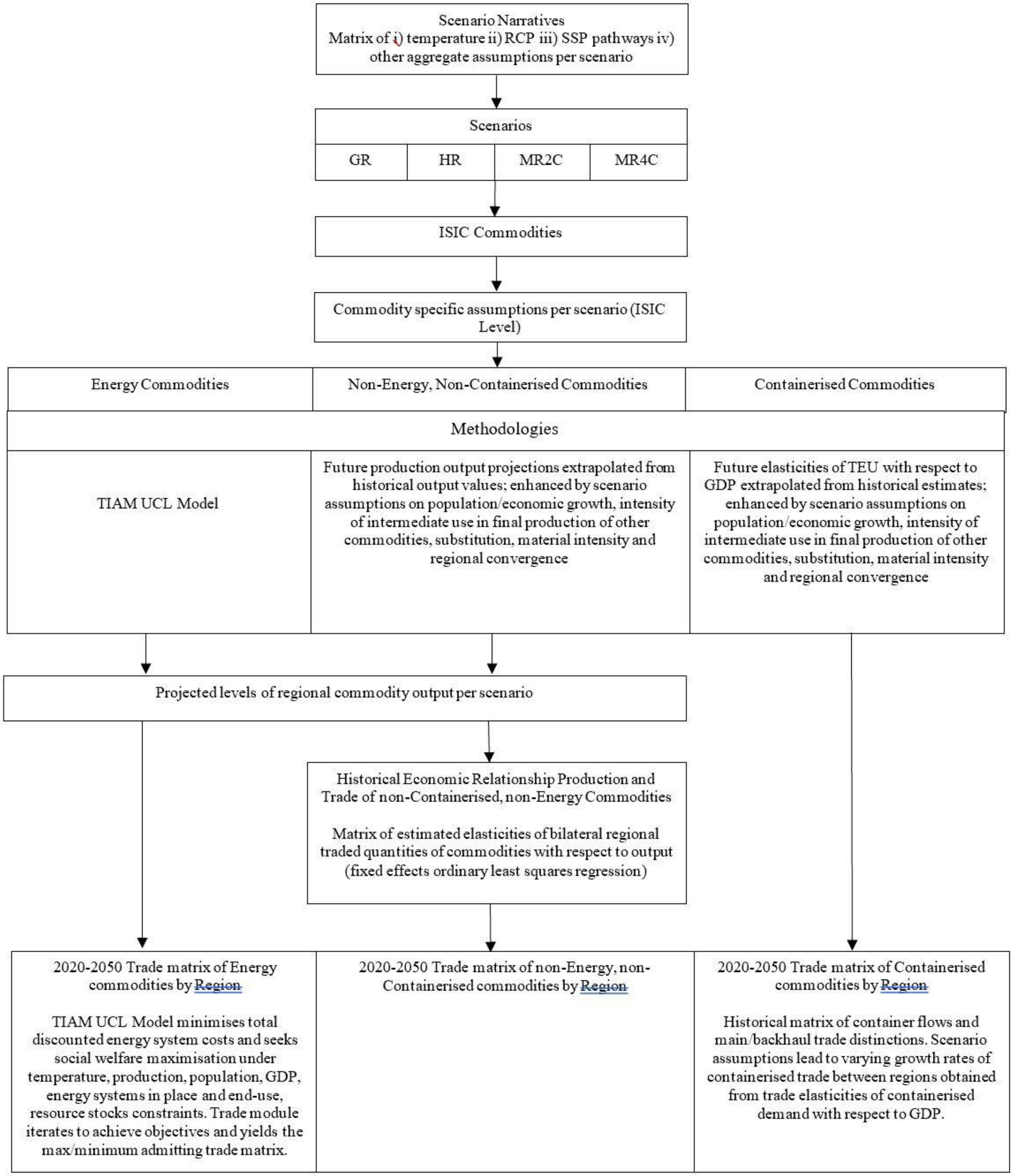


Figure 30. Suite of methods to generate global shipping scenarios for 2025 (source: Walsh et al, 2019)

- It is instructive to note that the authors recognize the significant uncertainty that is inherent in this study due to the complex nature of growth and trends, the long time horizon for the projections, the extensive data sources with their own uncertainty ranges, the complexity of the various models applied, and the inevitable range of estimates when soliciting expert opinion. They present some summary comments on how the scenario-generation process mitigates some of the uncertainty, as illustrated in Table 9.

Table 9. Summary of measures to mitigate uncertainty within the scenario process (source: Walsh et al., 2019)

Error Type	Description	Comment on uncertainty mitigation activities
Context and framing	Appropriate boundaries and agreed description of the system.	Scenario framing is influenced heavily by SSPs which have been widely applied with narratives developed in conjunction with stakeholders. Choice of geographic and temporal boundaries (regional groups as opposed to countries and ending projections at 2050) due to the difficulty in accurately reflecting individual countries across multiple commodities, particularly the scale of climate impact by 2100. Energy system model encompasses all major economic sectors.
Input uncertainty	Uncertainty related to external driving forces or dependent variables that influence model outcomes.	Economic and population projections are taken from SSPs. Baseline (2010) trade data provided by NEA and MSI and validated against UNCTAD data. Production inputs based on official datasets (such as UNIDO, FAO, World bank etc.) See point below on regional production models. Energy system model inputs are taken from external sources (e.g. SSP GDP estimates), and reflect fundamental constraints (e.g. Bio-mass availably) which are influenced by both literature sources and both internal and external review.
Model structure uncertainty	Simplified descriptions of modelled processes.	The choice of three methods to model the global trade system reflects its inherent complexity. Econometric models are compared with existing studies on the relationship between trade and output. Where a statistical relationship between regional output and trade is not available, projection is based on the econometric relationship between trade and GDP and SSP GDP projections. Commodity specific production models (which serve as inputs to the econometric projection) are simplified representations, entailing a source of uncertainty. This allows fundamental constraints i.e. regional limits to production to be readily and explicitly reflected and linked across related commodities. The levels of growth in commodity trade by 2020 are compared against estimates for 2015 ensure a sudden unrealistic change in the economic system is not reflected.
Parameter uncertainty	Uncertainty in the variables associated with models.	All variables used in the quantification and projection of production are taken from literature sources or established datasets. For all such commodities and regions the production model outputs are validated against baseline (2010) data, whilst future projections reviewed in the context of historic levels of growth. The energy system model has been applied extensively including within the IPCC scenario database.

vii. ***Data sources***:

The sources of data are far too extensive to enumerate here. Historic shipping volumes and patterns form the basis of the projections, and the future scenarios under the various environmental and socio-economic conditions serve to estimate the potential supply and demand changes in terms of amounts of various commodities, but also origin and destination for the goods which will drive the shipping traffic levels and routes. Each of the 'external' scenario-driving models relies on extensive variable sets from multiple sources, and the causal models used in this study (econometric, trade elasticities ($\Delta TEU/\Delta GDP$), etc.) are also based on numerous datasets. The article provides detailed references.

viii. ***Output format***:

For each of the 4 scenarios, global trade projections are provided through stacked line charts to 2050 for the three vessel categories. Line charts are also produced for commodity category forecasts. While not presented explicitly in the articles, it is clear that the results have been tabulated by region, country, commodity and decade for in-depth examination of future trade flows.

ix. ***Comments***:

This study epitomizes the blending of causal quantitative prediction models with expert-based scenario generation. It also extends from shorter term projections over a decade to several decades ahead. Like most other shipping prediction approaches, it models distinct shipping sectors separately as the drivers are not the same across the industry.

5. Summary

While hindsight may be 20-20, and eyesight can be 20-20, foresight is never 20-20. The uncertainty about future conditions and events is something that we must all contend with when making decisions or preparing for eventualities. In many sectors, including marine traffic management, predictions serve to lessen the uncertainty and help guide good decisions. The forecasting of marine activities can address many different issues as demonstrated through this literature review, including:

- Impact assessment for risk mitigation planning
- Traffic management planning
- Capacity expansion (ports, fleets, tourism, etc.)
- Marine Spatial Planning
- Marine security (short term forecasting)
- etc.

The reports and articles reviewed portray a broad range of methods for predictions and projections, but essentially they are distilled into three categories: time series modelling based on extrapolating historical data patterns (albeit with very complex formulas sometimes); causal models, where the forecasts are applied to the explanatory variable(s) which will then determine the future marine activity of interest; and scenario-based which is particularly well-suited when there is a data paucity and/or for longer-term projections. As discussed in this report, there is also the option of combining some of these methods into a blended approach. While scenario development clearly involves subjective judgment, it is worth pointing out that all of these methods have elements of judgment, as the quantitative approaches require decisions about what elements to include in the model(s), what specific method(s) to use, how to choose model parameters, and other factors. There is an old truism that if you give the same problem definition and data set to two independent statisticians, they will likely come up with two different outcomes. This does not undermine the analytic process itself (not the expertise of the analysts), but highlights the fact that there is some subjectivity in model application.

Key considerations for choosing the best model depends on many factors, including:

- Type of activity
- Purpose of the prediction
- Forecast horizon
- Size of the area of interest
- Data type, reliability and availability
- How fast the activity is changing
- Level of precision required

- Ease of use of the technique, the cost of producing it and the speed at which it can be produced.

An important concept is the assumption of what relevant factors are expected to remain (relatively) unchanged over the time horizon of the forecast. For example, if container throughputs (in TEU) in a port is forecasted to double in 20 years, then one may assume that the number of ship calls will double by then. However, this would not hold true if the ships' capacities over those years is also increasing. Either the changing underlying factor and its influence on the marine activity levels must also be forecast (Figure 2), or at least what-if analyses carried out to explore variations in factors of interest. Note that sensitivity analysis can be done on any forecast to try out variations on any input factors. Note that "what-if" analyses can explore the effect of factors within our control (such as the scale and timing of potential port expansions), or factors outside out control (such as the effect of a pandemic on shipping traffic).

Finally, as shown throughout these publications, there are many ways to represent the prediction's outputs, such as tables, traffic density maps, dynamic simulations, percent annual changes, and other formats. This is dictated by the nature of the prediction, the type of problem being addressed, and the end-users' preferences.

Despite the challenges of developing predictions, if done well the benefits will generally outweigh the costs. Furthermore, the exercise of modelling itself, done methodically and transparently, already helps decision-makers to clarify the situation and what the potential outcomes may be.

6. References

Akhtar, J., Bjørnskau, T., & Jean-Hansen, V. (2012). Oil spill risk analysis of routeing heavy ship traffic in Norwegian waters. *WMU Journal of Maritime Affairs, 11***,** 233–247. https://doi.org/10.1007/s13437-011-0016-8

Brillio, 2018. *Choosing the Right Forecasting Technique*. https://www.brillio.com/insights/choosing-the-right-forecasting-technique/

Canadian Parks and Wilderness Society [CPAWS]. (n.d.). *The Northern Shelf Bioregion.* https://cpawsbc.org/northern-shelf-bioregion/

Clear Seas Centre for Responsible Marine Shipping [Clear Seas]. (2020, December). *Vessel traffic in Canada's Pacific Region.* https://clearseas.org/wp-content/uploads/2021/02/VTA-Pacific-Final-Report-EN.pdf

Council of Canadian Academies [CCA]. (2016). Commercial marine shipping accidents: Understanding the risks in Canada. Ottawa (ON); Workshop Report.

Cox, G. (2010). Forecasting port traffic the safer way: Basing traffic forecasts solely on GDP can be inaccurate – here's an alternative way to more reliably plan for the future. *Port Technology International.* http://www.maricasts.com/uploads/4/0/3/6/4036321/port_technology_article_september_2010.pdf

Cruise Industry News (2021). 2021 Cruise Industry News Annual Report. https://www.cruiseindustrynews.com/store/product/annual-reports/2021-cruise-industry-news-annual-report/

Cruise Lines International Association [CLIA] (2019). 2019 Global Market Report. https://cruising.org/-/media/research-updates/research/2019-year-end/updated/2019-global-market-report.ashx

Çuhadar, M., Cogurcu, I., & Kukrer, C. (2014). Modelling and forecasting cruise tourism demand to Izmir by different artificial neural network architectures. *International Journal of Business and Social Research, 4*(3), 12-28.

Dalili, N., Ushio, M., Cosandey-Godin, A., & Alidina, H. (2020). *Mitigating shipping impacts on Cetaceans in Canada: Lessons learned and best practices.* World Wildlife Fund Canada. https://wwf.ca/wp-content/uploads/2021/02/WWF-MPA-5-Impacts-Cetaceans-v4.pdf

Department of Fisheries and Oceans Canada [DFO]. (2017a). *Assessing the risk of ship strikes to humpback (Megaptera novaeangliae) and fin (Balaenoptera physalus) whales off the west coast of Vancouver Island, Canada.* DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2017/038. https://waves-vagues.dfo-mpo.gc.ca/Library/40619709.pdf

Department of Fisheries and Oceans Canada [DFO]. (2017b). *Technical review of Roberts Bank Terminal 2 environmental impact statement and marine shipping supplemental report: Effects on marine mammals.* https://waves-vagues.dfo-mpo.gc.ca/Library/40622861.pdf

Department of Fisheries and Oceans Canada [DFO]. (2018). *Technical review: Potential effectiveness of mitigation measures to reduce impacts from project-related marine vessels on Southern Resident Killer Whales.* https://waves-vagues.dfo-mpo.gc.ca/Library/4073982x.pdf

Det Norske Veritas Inc. [DNV]. (2013). *Termpol 3.15 General risk analysis and intended methods of reducing risks: Trans Mountain Pipeline Expansion Project.* pp. 1143-1246. https://transmountain.s3.amazonaws.com/application14/V8C_TERMPOL_REPORTS/1143.html

Dinwoodie, J., Tuck, S., & Rigot-Müller, P. (2013). Maritime oil freight flows to 2050: Delphi perceptions of maritime specialists. *Energy Policy, 63,* 553-561. https://doi.org/10.1016/j.enpol.2013.08.068

Drewry Shipping Consultants Ltd. (2020, October). *VFPA long term container forecast: 2020 – 2060 Final report.* [PowerPoint slides]. https://www.robertsbankterminal2.com/wp-content/uploads/Drewry-container-forecast-report-final.pdf

Engler, C., & Pelot, R. (2013). *Analysis of marine traffic along Canada's Coasts Phase 2 – Part 1: Factors influencing Arctic traffic.* Report prepared for Defence Research and Development Canada. https://cradpdf.drdc-rddc.gc.ca/PDFS/unc125/p537704_A1b.pdf

EnviroEmerg Consulting Services. (2008). *Major marine vessel casualty risk and response preparedness in British Columbia.* https://www.ceaa.gc.ca/050/documents_staticpost/cearref_21799/86129/Major_Marine_Vessel_Casualty.pdf

Étienne, L, Pelot, R., & Engler, C. (2013). *Analysis of marine traffic along Canada's Coast Phase 2– Part 2: A spatio-temporal simulation model for forecasting marine traffic in the Canadian Arctic in 2020.* Report prepared for Defence Research and Development Canada. https://cradpdf.drdc-rddc.gc.ca/PDFS/unc137/p538451_A1b.pdf

Eyring, V., Köhler, H.W., Lauer, A., & Lemper, B. (2005). Emissions from international shipping: 2. Impact of future technologies on scenarios until 2050. *Journal of Geophysical Research: Atmospheres, 110*(D17), 1-15. https://doi.org/10.1029/2004JD005 620

Fournier, M., Hilliard, R.C., Rezaee, S., & Pelot, R. (2018). Past, present, and future of the satellite-based automatic identification system: areas of applications (2004–2016). *WMU Journal of Maritime Affairs, 17*, 311-345. https://doi.org/10.1007/s13437-018-0151-6

Gargari, N.S., Akbari, H., & Panahi, R. (2019). Forecasting short-term container vessel traffic volume using Hybrid ARIMA-NN Model. *International Journal of Coastal & Offshore Engineering, 3*(3), 47-52. doi: 10.29252/ijcoe.3.3.47

Gosasang, V., Chandraprakaikul, W., & Kiattisin, S. (2011). A comparison of traditional and neural networks forecasting techniques for container throughput at Bangkok Port. *The Asian Journal of Shipping and Logistics*, *27*(3), 463–482.

Hajbi, A. (2011). Traffic forecasting in Moroccan Ports. *Supply Chain Forum: An International Journal, 12*(4), 26-35. https://doi.org/10.1080/16258312.2011.11517278

Herbert Engineering Corp., & Environmental Research Consulting. (2014). *Roberts Bank Terminal 2 technical report marine vessel incidence prediction inputs to the quantitative risk assessment*. https://iaac-aeic.gc.ca/050/documents/p80054/101341E.pdf

Hilliard, R.C. & Pelot, R. (2012). Analysis of Marine Traffic along Canada's Coasts: Phase 1 Final Report, DRDC CORA #CR2012-060.

Hodgson, J.R.F., Russel, W.D., & Megannety, M. (2013). *Exploring plausible futures for marine transportation in the Canadian Arctic a scenarios' based approach.* Hodgson and Associates. https://cdn.dal.ca/content/dam/dalhousie/pdf/faculty/science/marine-affairs-program/Marine%20Transportation%20in%20Canadian%20Arctic%20Report%202014.pdf

Hyndman, R.J., & Athanasopoulos, G. (2018). Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on 2021-08-04.

Iacarella, J.C., Lyons, D.A., Burke, L., Davidson, I.C., Therriault, T.W., Dunham, A., & DiBacco, C. (2020a). Climate change and vessel traffic create networks of invasion in marine protected areas. *Journal of Applied Ecology, 57*(9), 1793-1805. https://doi.org/10.1111/1365-2664.13652

Iacarella, J.C., Clyde, G., & Dunham, A. (2020b). *Vessel Tracking Datasets For Monitoring Canada's Conservation Effectiveness*. Canadian Technical Report of Fisheries and Aquatic Sciences 3387, viii + 31 p.

Indra, V., Notteboom, T., Parola, F., Satta, G., & Persico, L. (2015). *Port traffic forecasting tool.* Portopia. http://www.portopia.eu/wp-content/uploads/2016/07/D1.3-Port-Traffic-Forecasting-Tool.pdf

International Maritime Organization [IMO]. (2009). *Second IMO GHG study 2009.* https://wwwcdn.imo.org/localresources/en/OurWork/Environment/Documents/SecondIMOGHGStudy2009.pdf

International Maritime Organization [IMO]. (2014). *Third IMO greenhouse gas study 2014.* https://www.researchgate.net/publication/281242722_Third_IMO_GHG_Study_2014_Executive_Summary_and_Final_Report

Jugović, A., Hess, S., & Jugović, T.P. (2011). Traffic demand forecasting for Port Services. *Scientific Journal on Traffic and Transportation Research, 23*(1), 59-69. https://doi.org/10.7307/ptt.v23i1.149

Kahneman, D. (2011). Thinking, Fast and Slow. New York: Farrar, Straus and Giroux. ISBN 978-0-374-27563-1.

Kinder Morgan Canada Ltd. (2013a). Trans Mountain Expansion Project: An application pursuant to Section 52 of the National Energy Board Act, Volume 8a: Marine transportation. Trans Mountain Pipeline ULC. http://transmountain.s3.amazonaws.com/application/V8A_ 1_of_4_1_TO_4.2.9_MAR_TRANS_ASSESS.pdf

Kinder Morgan Canada Ltd. (2013b). Trans Mountain Expansion Project: An application pursuant to Section 52 of the National Energy Board Act, Volume 8c: TERMPOL reports. Trans Mountain Pipeline ULC. https://transmountain.s3.amazonaws.com/application14/V8C_TERMPOL_REPORTS/index.html

Kollwitz, H., & Papathanassis, A. (2011). Evaluating cruise demand forecasting practices: A Delphi approach. In: Gibson P., Papathanassis A., Milde P. (eds) Cruise Sector Challenges. Gabler Verlag. https://doi.org/10.1007/978-3-8349-6871-5_3

Lam, W.H.K., Ng, P.L.P, Seabrooke, W., & Hui. E.C.M. (2004). Forecasts and reliability analysis of port cargo throughput in Hong Kong. *Journal of Urban Planning and Development, 130*(3), 133–144.

Lerner, J. (2018). *Review of cumulative effects management concepts and international frameworks.* Transport Canada. https://tc.canada.ca/en/corporate-services/review-cumulative-effects-management-concepts-international-frameworks

Living Oceans Society. (2011). *Shipping on the British Columbia Coast: Current status, projected trends, potential casualties, and our ability to respond: A briefing report*. https://www.acee-ceaa.gc.ca/050/documents_staticpost/cearref_21799/86129/Shipping_ on_the_BC_Coast.pdf

Makridakis, S.G., Wheelwright, S.C., & Hyndman, R.J. (1998). Forecasting: Methods and Applications. John Wiley and Sons, Toronto.

Maritime Strategies International [MSI] (2021). *Maritime Forecasting and Strategic Advisory*. https://www.msiltd.com/#aboutus

Maritime Traffic Forecasts Ltd. (n.d.) *Key elements of our traffic forecasting approach.* Maricasts. http://www.maricasts.com/forecast-methodology.html

McKenna, M.F., Wiggins, S.M., & Hildebrand, J.A. (2013). Relationship between container ship underwater noise levels and ship design, operational and oceanographic conditions. Scientific Reports, *3.* DOI: 10.1038/srep01760

Mercator. (2018, November 1). *Roberts Bank Terminal 2 container vessel call forecast study.* [PowerPoint slides]. Port of Vancouver. https://www.robertsbankterminal2.com/wp-content/uploads/RBT2-container-vessel-call-forecast-study-Mercator-November-2018.pdf

Metro Vancouver. (2015, June 8). *Short sea shipping in Metro Vancouver.* [PowerPoint slides]. http://www.metrovancouver.org/services/regional-planning/PlanningPublications/Short_Sea_Shipping_Facts_in_Focus.pdf

Ministry of Small Business and Economic Development & Ministry of Transportation. (2005). *British Columbia Ports Strategy.* http://www.llbc.leg.bc.ca/public/pubdocs/bcdocs/375137 /sbed_mar_18_05.pdf

Minton, G., Folegot, T., Cosandey-Godin, A., Jacob, T., Lancaster, M., & Ushio, M. (2021). *Shipping and cetaceans: A Review of impacts and mitigation options for policymakers and other stakeholders.* WWF. https://www.wwf.nl/globalassets/pdf/nieuws/wwf-shipping-and-cetaceans-report-2021.pdf

MPA Network of the Northern Shelf Bioregion. (n.d.). *The Region.* https://mpanetwork.ca/bcnorthernshelf/the-region/

Nichol, L.M., Wright, B., O'Hara, P., & Ford, J.K.B. (2017). *Assessing the risk of lethal ship strikes to humpback (Megaptera novaeangliae) and fin (Balaenoptera physalus) whales off the west coast of Vancouver Island, Canada.* Department of Fisheries and Oceans Canada [DFO]. https://waves-vagues.dfo-mpo.gc.ca/Library/40616642.pdf

Nuka Research & Planning Group, LLC. (2013). *West coast spill response study, Volume 2: Vessel traffic analysis.* https://www2.gov.bc.ca/assets/gov/environment/air-land-water/spills-and-environmental-emergencies/docs/westcoastspillresponse_vol2 _vesseltrafficstudy_130722.pdf

Ocean Policy Research Foundation [OPRF]. (2008). *Research study: The world's changing maritime industry and a vision for Japan.* https://www.spf.org/en/_opri_media/pdf/200805_ISBN978_4_88404_212_7.pdf

Ocean Shipping Consultants. (2016). *Container traffic forecast study—Port of Vancouver, 2016.* https://www.portvancouver.com/wp-content/uploads/2015/05/2016-Container-Traffic-Forecast-Study-Ocean-Shipping-Consultants.pdf

Parola, F., Satta, G., Notteboom, T., & Persico, L. (2020). Revisiting traffic forecasting by port authorities in the context of port planning and development. *Maritime Economics & Logistics.* https://doi.org/10.1057/s41278-020-00170-7

Pavlić, I. (2013). Cruise tourism demand forecasting - The case of Dubrovnik. *Tourism and Hospitality Management. 19*(1), 125-142.

Peng, W.Y., & Chu, C.W. (2009). A comparison of univariate methods for forecasting container throughput volumes. *Mathematical and Computer Modelling. 50*(7-8), 1045-1057.

Port of Vancouver. (2017). *ECHO Program study summary.* https://www.portvancouver.com/wp-content/uploads/2017/01/2017-07-ECHO-Program-Estimating-the-effects-of-noise-from-commercial-vessels-and-whale-watch-boats-on-SRKW.pdf

Port of Vancouver. (2018). *Roberts Bank Terminal 2 project overview and rationale.* https://www.ceaa-acee.gc.ca/050/documents/p80054/125977E.pdf

Port of Vancouver. (2021). *Vancouver Fraser Port Authority view on container forecast volumes 2020-2060 container forecasts.* https://www.robertsbankterminal2.com/wp-content/uploads/2021-01-19-PDF-Vancouver-Fraser-Port-Authority-view-on-container-forecast-volumes.pdf

Prince Rupert Port Authority [PRPA]. (2020). *Land use plan.* https://www.rupertport.com/land-use-plan/

Robinson, J. (2003). Future subjunctive: backcasting as social learning. *Futures. 35*, 839-856.

Sánchez, R.J., Perrotti, D.E. & Fort, A.G.P. (2021). Looking into the future ten years later: big full containerships and their arrival to south American ports. *J. of Shipping and Trade. 6(2).* https://doi.org/10.1186/s41072-021-00083

Sardain, A., Sardain, E., & Leung, B. (2019). Global forecasts of shipping traffic and biological invasions to 2050. *Nature Sustainability, 2*, 274-282. https://doi.org/10.1038/s41893-019-0245-y

Seaport Consultants Canada Inc. [Seaport]. (2013). *Projections of vessel movements draft report.* https://transmountain.s3.amazonaws.com/application14/V8C_TERMPOL_REPORTS/0142.html

Serra-Sogas, N. (2020a, September). *Analysis of vessel traffic trends within the Northern Shelf Bioregion.* [PowerPoint slides]. Transport Canada.

Serra-Sogas, N. (2020b, September). *Vessel traffic analysis for focus areas within NSB.* [PowerPoint slides]. Transport Canada.

Serra-Sogas, N. (2020c, November 23). *NASP Aerial surveys in Northern Shelf Bioregion Pacific Region.* [PowerPoint slides].

Serra-Soras, N., Canessa, R., O'Hara, P., Smallshaw, L., & Warrior, M. (2018). *Small vessel traffic study in the Salish Sea and Southern Resident Killer Whale critical habitat.* Coastal and Ocean Resource Analysis (CORAL) Group. https://meopar.files.wordpress.com/2018/05/smallvesseltrafficstudy_report_uvic2018.pdf

Soykan, C., Eguchi, T., Kohin, S., & Dewar, H. (2014). Prediction of fishing effort distributions using boosted regression trees. *Ecological Applications, 24*(1), 71-83. http://www.jstor.org/stable/23596804

Tarber, J. (2013). *Termpol 3.2–Origin, destination & marine traffic volume survey Trans Mountain Expansion Project.* Moffatt and Nichol. https://transmountain.s3.amazonaws.com/application14/V8C_TERMPOL_REPORTS/0070.html

The Maritime Executive (2019). *The Arctic Shipping Route No One's Talking About*. https://www.maritime-executive.com/editorials/the-arctic-shipping-route-no-one-s-talking-about.

The Tioga Group., & Hackett Associates. (2020). *2019-2050 Bay area seaport forecast.* http://50.62.26.103/seaport/2019-2050-Bay-Area-Seaport-Forecast.pdf

Transport Canada (2020) Draft — *National framework for assessing the cumulative effects of marine shipping*. https://tc.canada.ca/en/marine-transportation/marine-pollution-environmental-response/cumulative-effects-marine-shipping

UN Global Working Group on Big Data for Official Statistics: Task Team on AIS (n.d.) Key elements of our traffic forecasting approach. AIS Handbook. https://unstats.un.org/wiki/display/AIS/AIS+Handbook+Outline

van Dorsser, C., Wolters, M., & van Wee, B. (2012). A very long term forecast of the port throughput in the Le Havre – Hamburg range up to 2100. *European Journal of Transport and Infrastructure Research, 12*(1), 88-110. https://doi.org/10.18757/ejtir.2012.12.1.2951

Vard Marine Inc., & World Wildlife Fund Canada. (2018). *Canadian Arctic greywater report: Estimates, forecasts, and treatment technologies.*

https://arcticwwf.org/site/assets/files/1851/wwf_canada_grey_water_report_2018_1.pdf

Walsh, C., Lazaroub, N.-J., Traut, M., Price, J., Raucci, C., Sharmina, M., Agnolucci, P., Mander, S., Gilbert, P., Anderson, K., Larkin, A., & Smith, T. (2019). Trade and trade-offs: Shipping in changing climates. *Marine Policy, 106*, https://doi.org/10.1016/j.marpol.2019.103537

Wang, S., Wang, S., Gao, S., & Wenguo, Y. (2017). *Daily ship traffic volume statistics and prediction based on Automatic Identification System data.* 2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics. doi: 10.1109/IHMSC.2017.149

WorleyParsons Canada Services Ltd. (2011). *Projections of vessel calls and movements at Deltaport and Westshore terminals Deltaport Terminal Road and Rail Improvement Project (DTRRIP).* https://www.robertsbankterminal2.com/wp-content/uploads/Projections-of-Vessel-Calls-and-Movements-at-Deltaport-and-Westshore-Terminals.pdf

WSP UK Ltd. (2020, October). *Long-term container traffic forecast, 2020-2060 final report*. [PowerPoint slides]. https://www.robertsbankterminal2.com/wp-content/uploads/WSP-container-forecast-final-report.pdf

Xie, G., Qian, Y., & Wang, S. (2021). Forecasting Chinese cruise tourism demand with big data: An optimized machine learning approach. *Tourism Management, 82*. https://doi.org/10.1016/j.tourman.2020.104208

Zhang, C., Huang, L., & Zhao, S. (2013). Research on combination forecast of port cargo throughput based on time series and causality analysis. *Journal of Industrial Engineering and Management, 6*(1), 124-134. http://dx.doi.org/10.3926/jiem.687

Zhang, X., Chen, G., Wang, J., Li, M., & Cheng, L. (2019). A GIS-based spatial-temporal autoregressive model for forecasting marine traffic volume of a shipping network. *Scientific Programming, 2019,* 1-14. https://doi.org/10.1155/2019/2345450