7th Framework Programme
SST.2013.6-2. Towards a competitive and resource efficient port transport system
Collaborative Project
Grant Agreement no. 605176

Deliverable 1.3
Port Traffic Forecasting Tool
DELIVERABLE 1.3

Port Traffic Forecasting Tool

AUTHORSHIP

Author(s)  Vonck Indra, Prof. dr. Theo Notteboom, Dr. Francesco Parola, Dr. Giovanni Satta, Dr. Luca Persico
Beneficiary Partner  UA-ITMMA
Issue Date  2015-11-13
Revision  M. Langenus
Status  Final draft version

Contributors  UNIGE
Pages  46
Figures  12
Tables  12
Annexes  0

SIGNATURES

Author(s)  Vonck Indra, Theo Notteboom, Francesco Parola, Giovanni Satta, Luca Persico
Coordinator  Michaël Dooms

Disclaimer

The information contained in this report is subject to change without notice and should not be construed as a commitment by any members of the PORTOPIA Consortium or the authors. In the event of any software or algorithms being described in this report, the PORTOPIA Consortium assumes no responsibility for the use or inability to use any of its software or algorithms. The information is provided without any warranty of any kind and the PORTOPIA Consortium expressly disclaims all implied warranties, including but not limited to the implied warranties of merchantability and fitness for a particular use.
(c) COPYRIGHT 2013 The PORTOPIA Consortium
DELIVERABLE 1.3
Port Traffic Forecasting Tool

REVISION

<table>
<thead>
<tr>
<th>Revision Date</th>
<th>Reviewer Name</th>
<th>Reviewer Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-11-13</td>
<td>M. Langenus</td>
<td>M. Langenus</td>
</tr>
</tbody>
</table>

HISTORY

<table>
<thead>
<tr>
<th>Version Date</th>
<th>Version Id</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013-12-31</td>
<td>V0.0</td>
<td>[DRAFT]</td>
<td>[BRIEF DESCRIPTION OF THE REVISION]</td>
</tr>
<tr>
<td>TBD</td>
<td>V1.0</td>
<td>[FINISH]</td>
<td>[BRIEF DESCRIPTION OF THE REVISION]</td>
</tr>
</tbody>
</table>
Deliverable 1.3 (D1.3) is embedded in Task 1.3: Estimation of future traffic flows in the European port system (in the short, medium and long term). This task focuses on the development and implementation of a range of methods to increase insight in the expected future (traffic) outlook for the European port system in the short, medium and long term.

We look at three things:

- A first sub-task explores and applies a range of time-series based forecasting techniques in order to generate aggregated and top-down forecasts on the maritime traffic evolution of the entire European port system and significant parts thereof.

- A second sub-task aims to develop a forecasting meta-system which groups data and expert information on medium (1 to 5 years) and long term (> 5 years) developments in port activities in Europe.

- The third sub-task is focused on the short-term development of the European port system. It involves the drafting of a survey that will measure short-term traffic expectations (next quarters up to a year) of relevant stakeholders in the European port business, with a main focus on port authorities.
DELIVERABLE 1.3
Port Traffic Forecasting Tool

1 LIST OF CONTENTS

2 Positioning of this report in WP1 ................................................................. 6
3 Forecasting freight transport ................................................................. 7
3.1 GDP approach .................................................................................. 7
   3.1.1 Desk and field research approach ............................................. 8
3.2 General forecasting techniques ....................................................... 8
   3.2.1 Qualitative forecasting techniques ....................................... 8
   3.2.2 Time series analysis ......................................................... 10
   3.2.3 Causal models ................................................................... 16
   3.2.4 Overview ......................................................................... 19
3.3 Academic research in port traffic forecasting .............................. 20
3.4 Assessment of techniques in port data forecasting ...................... 26
3.5 Combining port development plans with forecasting ..................... 27
4 Forecasting examples in similar sectors .......................................... 32
  4.1 Air Transport ............................................................................. 32
   4.1.1 Boeing and Airbus forecasts ..................................................... 32
   4.1.2 The International Air Transport Association (IATA) Airline Industry Forecast 2014-2018 ....... 32
   4.1.3 Lessons from the air transport forecasting ................................ 33
  4.2 Rail transport ........................................................................... 33
   4.2.1 Railway Demand Forecasting and Service Planning Processes (Rail Freight Service Review) .... 33
   4.2.2 Transnet freight demand forecast ............................................ 35
   4.2.3 Lessons from the rail transport industry ........................................ 36
4.3 Traffic flow forecasting ................................................................. 36
   4.3.1 Traffic flow: Scenario, Traffic Forecast and Analysis of Traffic on the TEN-T, Taking into
          Consideration the External Dimension of the Union. .................... 36
4.4 Academic research .................................................................... 37
5 Forecasting in Portopia ................................................................... 37
  5.1 Methodology ............................................................................ 38
   5.1.1 Short term ......................................................................... 38
   5.1.2 Medium term ................................................................... 40
   5.1.3 Long term ......................................................................... 41
   5.1.4 Implementation .................................................................. 42
5.2 References ................................................................................ 43
2 POSITIONING OF THIS REPORT IN WP1

The aim of WP1 within the PORTOPIA project is to further develop the PPRISM indicators on market trends and structure and to seek meaningful expansions. The specific objectives of the work package include:

- Improving data availability and comparability of PPRISM indicators
- Collecting and presenting data at a more disaggregated level in terms of goods types and time periods
- Developing new indicators (ratios and indexes)
- Develop forecasts on short, medium and long term developments in port activities in Europe using a combination of techniques (modelling, meta-analysis and survey)
- Incorporation in a European Port Observatory (EPO) with link between indicators and specific policy targets in the EU transport policy.

Deliverable 1.3 (D1.3) is embedded in Task 1.3: Estimation of future traffic flows in the European port system (in the short, medium and long term). This task focuses on the development and implementation of a range of methods to increase insight in the expected future (traffic) outlook for the European port system in the short, medium and long term.

We look at three things:

- A first sub-task explores and applies a range of time-series based forecasting techniques in order to generate aggregated and top-down forecasts on the maritime traffic evolution of the entire European port system and significant parts thereof
- A second sub-task aims to develop a forecasting meta-system which groups data and expert information on medium (1 to 5 years) and long term (> 5 years) developments in port activities in Europe.
- The third sub-task is focused on the short-term development of the European port system. It involves the drafting of a survey that will measure short-term traffic expectations (next quarters up to a year) of relevant stakeholders in the European port business, with a main focus on port authorities.
3 FORECASTING FREIGHT TRANSPORT

According to Graham Cox (Director, Maritime Traffic Forecasts Ltd, Hothfield, Kent, UK), there are two traditional methodologies that have been used to forecast port traffic volumes; one based on GDP and the other on desk and field research.

3.1 GDP approach

The GDP approach is based on the assumption that most economic drivers determine demand for cargo. In practice this principle is correct but has been oversimplified to apply only the GDP as causality to traffic and impact on ports.

This method works with multipliers (for e.g. cargo grows 1.5 times for each point of GDP growth and then forecasted to assess future cargo growth. This method is often combined with other forms of qualitative forecasting. For example analysts sometimes apply extra multipliers based on market assessments or econometric modelling. This practice introduces considerable subjectivity and is hard to audit.

This simplification of forecasting in the maritime industry may have occurred because of data problems in general and because the GDP approach has been an appropriate simplification for containerized trade route traffic; given that there is no information about the nature of the cargo within the boxes. As is shown in deliverable 9.1 data issues are prominent in maritime data isd therefore quality issues often arise. Unfortunately, analysis based predominantly on GDP trends has become the norm for all cargo types and right down to the port level, although that some cargo’s do not have a clear link with GDP evolution (dichotomy between containers and container content). The large majority of forecasts produced for the liquid and dry bulk trades follow the same simple GDP methodology. This has happened despite the availability of data on other economic drivers and more accurate information about the volume and make-up of port trade; even in some cases for containerized traffic.

It is important to note that there is no fundamental statistical or economics justification for analysing trade simply in terms of a relationship with GDP. Richer statistical methodologies reveal underlying relationships with the drivers of this type of traffic that GDP cannot capture. It is potentially deficient, if for no other reason that it ignores the effect on traffic flows that different sectors of the economy can exert: such as construction or the retail trade.

An example is that certain sectors move in different paths than others. Agriculture could be failing in a region but steel industry booming. This development will have a materially different effect on traffic flows from that indicated by GDP if that sector’s imports or exports constitute a significant proportion of cargoes. Use of a sole GDP driver will always create a large weighing error with different kinds of maritime traffic. These different trends in sectors can persist into the medium and long-term and are hence important for port studies, both in terms of determining the historic relationships and in particular how they are projected.

A second major issue with this method is that the standard use of fixed relationships: i.e. the assumption that the multiplier was always the same value and will remain so. An examination of historical data indicates that such relationships are not static and are unstable for a multitude of reasons. If there has been structural change during the period that is used to establish the relationship between GDP, sectors and/or any other drivers that affect port traffic and this change is not recognized, the multipliers used to project cargo volumes in the forecast period will be wrong. So in essence this methodology is flawed at detecting turning points in individual cargo trends.
3.1.1 Desk and field research approach

This approach is rather different and comparable to expert opinions and the Delphi method. It involves interacting with individuals and organizations in the industrial sectors that import and export to ascertain their views and knowledge of developments. According to Cox this principle is sound since the supply and demand side are often not distinguishable in statistics. Dangers of this approach include the fact that interviewees may have a too narrow view of the market, rely themselves on other forecasts or include external variables which are not relevant.

For short-term forecasting this technique has its benefits. Since it is generally based on the level of orders of the main actors in the sector. However useful this approach is for the short-term, it is not suitable for a port forecast that needs medium- and long-term insight.

A possible alternative to these classic methods could be a multi-factoral methodology. Namely by conducting an in-depth statistical or econometric analysis, using long runs of data to identify the key sector drivers for each kind of freight, and how they and other disaggregated data modify the relationship of cargo type with GDP and hence alter traffic projections. Secondly, a time series analysis appraising the impact of past shocks to the economy on economic relationships. Thirdly, the examination of the structure of the market of export and import industries. For example, the competitiveness of industries in international markets relevant to the traffic of a port needs to be examined. Changes in government policy and new technology can have significant impacts. Desk and field research or expert advice in these areas complements statistical analysis to make the forecast as close to reality as possible.

3.2 General forecasting techniques

In virtually every decision they make, executives today consider some kind of forecast. This is no different for managers in the maritime or port industry. Sound predictions of demands and trends are no longer luxury items, but a necessity. In order to cope with seasonality and volatility or large swings of the economy. Forecasting is a useful tool to help with these issues. This document provides an overview for the most used methodologies and different approaches and distils the most suitable techniques for the Portopia project.

There are three basic types of forecasting techniques namely qualitative models, time series analysis and causal models (both quantitative models). The first uses qualitative data (expert opinion, for example) and information about special events of the kind already mentioned, and may or may not take the past into consideration. The second, on the other hand, focuses entirely on patterns and pattern changes, and thus relies entirely on historical data. The third uses highly refined and specific information about relationships between system elements, and is powerful enough to take special events formally into account. As with time series analysis and projection techniques, the past is important to causal models.

Each of the different techniques has its advantages and disadvantages and suits certain situations. Often they have alternating levels of success in the difference between short medium and long term forecasting.

3.2.1 Qualitative forecasting techniques

A qualitative forecasting technique is an Estimating method that relies on expert human judgment combined with a rating scale, instead of on hard (measurable and verifiable) data.

- used when data are scarce or ambiguous
- frequently used in new-technology areas
Advantages

- Flexible data needs allow for inclusion of non-numerical data
- Ambiguous or non-complete data can be processes
- Allows experience and judgment of senior executives and outside experts.

Disadvantages

- Not always accurate
- If the opinion of one person, whose view prevails, is incorrect, the forecast is incorrect
- Difficult to eliminate the forecaster’s personal bias

DELPHI METHOD

This is a group technique in which a panel of experts is questioned individually about their perceptions of future events. The experts do not meet as a group, in order to reduce the possibility that consensus is reached because of dominant personality factors. Instead, the forecasts and accompanying arguments are summarized by an outside party and returned to the experts along with further questions. This process continues until a consensus is reached.

Advantages:

- Quite effective for long-range forecasting.
- Eliminates the disadvantages of group think.
- There is no committee or debate.
- The experts are not influenced by peer pressure

Disadvantages:

- Low reliability
- Lack of consensus from the returns.

MARKET RESEARCH FORECAST

Testing hypotheses based on market trends and general economic knowledge. The best results of this technique arise when combined with other qualitative forecasts. A first attempt of this approach was made in deliverable 1.1 ‘market trends’.

PANEL CONSENSUS FORECAST

This technique uses open meetings where all participants provide or exchange ideas. In addition, instead of only relying on historical data, the consensus method goes one step further. The historical data is enhanced with the current market trends and events to ensure that the demand uncertainty is reduced.

VISIONARY FORECAST

Personal insights opinion and facts to predict possible scenarios. This is a non-scientific method and often prone to errors with a high degree of uncertainty. Delphy method or panel consensus are preferable alternatives. This technique is the most radical and consists of a prophecy of the future
based on personal insight, judgement, and, when available, historical analogies that can be extrapolated into possible future forecasts. Indeed, it is characterized by subjective guesswork and imagination and, in general, the methods used are non-scientific and non-quantitative (Ross, 2011).

HISTORICAL ANALOGY FORECAST

This is a forecasting method that assumes that two different kinds of phenomena share the same model of behavior. For example, one way to predict the sales of a new product is to choose an existing product which "looks like" the new product in terms of the expected demand pattern for sales of the product. Hence, historical analogy forecast constitutes a judgmental forecasting technique based on identifying a sales history that is analogous to a present situation, such as the sales history of a similar product, and using that past pattern to predict future sales.

3.2.2 Time series analysis

The second subgroup of analysis is the time series analysis. This technique is used when several years’ data for a product or product line are available. It measures the rate and changes rate of change. Time series analysis has the main objective of:

- describing data and obtain simple measures to summarize the main properties of the time series. The simple analysis of the graph can unveil the existence of trend, seasonality, outliers and turning points;
- indentifying the data generating process that generates the random variable of which a sequence of observations are available;
- make forecast estimations of future values. This is based on the principle that the behaviour of the phenomenon in the past is maintained in the future also.

To practically conduct time series analysis two main approaches are known: i) classical approach to time series analysis and ii) Box-Jenkins approach or modern time series analysis.

CLASSICAL APPROACH TO TIME SERIES ANALYSIS

It is based on the error model: \( x_t = f(t) + \varepsilon_t \quad t = 1, 2, ..., n \)

In this case the series \( x_t \) is defined as the resultant of an explicit analytical function \( f(t) \) plus a random error term \( \varepsilon_t \) which must comply with the typical assumptions of the regression model: \( E(\varepsilon_t) = 0; E(\varepsilon_t^2) = \sigma^2 < +\infty; E(\varepsilon_t \varepsilon_s) = 0 \quad \forall t \neq s \). This model is particularly appropriate for regular phenomena.

According to the classical approach, every time series \( x_t \) can be considered as the result of a combination of different unobserved factors (stochastic and non stochastic), called the components of a time series:

- Trend component \( T_t \): can be defined as long-term change in the mean level. Trend component comprise all cyclic components whose wave length exceeds the length of the observed time series.
- Cyclic component \( C_t \): refers to variations exhibited by the time series at fixed period that cannot be considered as seasonality. For example, economic data are often affected by business cycle of about 5 years.
- Seasonality \( S_t \): consists of periodic, repetitive, and generally regular and predictable patterns in the levels of a time series. Seasonality can repeat on a daily, weekly, monthly or quarterly basis, these periods of time are structured and occur in a length of time less than a year. Seasonal fluctuations in a time series can be contrasted with cyclical patterns.
• Irregular fluctuations $\varepsilon_t$ : after trend, cycle and seasonal variation have been removed something is left in the residuals that is not predictable from the past history. Making sure that the $\varepsilon_t$ component is truly random is a good guarantee that the decomposition of the series into components is correct.

The classical methods of analysis of time series are concerned with the decomposition of the series into the four above mentioned components. It is worth noting that the decomposition is, in general, not unique unless certain assumptions are made.

The most common decomposition models used are:

- additive decomposition model: $x_t = T_t + C_t + S_t + \varepsilon_t$ (this model assumes that components have the same unit of measurement of $x_t$ and are independent each other; see Figure 1)
- multiplicative decomposition: $x_t = T_t * C_t * S_t * \varepsilon_t$ (this model assumes that only the trend component has the same unit of measurement of $x_t$, the remaining components are dimensionless number proportional to the trend component). Applying a log transformation it is easy to pass from multiplicative to additive model.
- mixed model: $x_t = (T_t * C_t) + S_t + \varepsilon_t$ (with additive seasonal component).

![Figure 1 Example of additive decomposition of a time series according to the classical approach](image)

Source: own elaboration.

The classical approach to time series has a number of advantages:

1. It is based on rather intuitive concepts
2. It is exploitable also when the series length is rather short
3. It is very useful when the series plot reveals that the variability is strongly dominated by a component, for example trend or seasonality.

It also has a number of disadvantages:

1. The decomposition is not unique
2. The stochastic part is only in the error term

Several methods to estimate trends and seasonal components have been developed.

**Trend Projection**

The estimation of the trend component (see Figure 2) of an observed time series $x_t$ is based on regression functions. For example, suppose to have a trend that can be classified as polynomial, hence we assume that the trend dynamics follows this form: $f(t) = a_0 + a_1t + a_2t^2 + \cdots + a_k t^k$. Using least squares estimation methods, it is possible to obtain the estimates of the coefficients. The polynomial that results can be a good guide to interpret the phenomenon. Much more caution must be used, instead, when the aim is to make prediction since there is no warrant that the functional form, supposed to hold over the sample period, does so also outside the time span of the observed time series.

*Figure 2 Estimation of linear trend in a monthly time series by using trend projection*

Source: own elaboration.

**Seasonality Estimation**

The regression method used for the estimation of the trend component can be adopted also for the estimation of the seasonal component, representable by a periodic function $h(t)$. A first approach to study $h(t)$ is to assume that $h(t)$ is a sum of dummy variables $h(t) = \sum_j a_j D_{jt}$ where $D_{jt}$ is 1 in the $j$-th period and 0 otherwise, $a_j$ indicates the level of the phenomenon at the $j$-th period. A second possibility is to assume that the seasonal component is given by a sum of harmonic functions (see Figure 3). Generally a sum of 2, 3 harmonic waves can be a reasonable option to represent a sufficiently complex seasonal dynamics.

**Moving Average Approach**

When the analytic description of the phenomenon is not possible, for example because the dynamic exhibited by the phenomenon is extremely irregular, the moving averages are a good tool to filter
the time series and indentify the underlying non-linear trend. Moving averages try to capture the dynamics of the phenomenon without defining an analytical form that holds over the entire time span as a variation law. Using moving averages it is possible to estimate the trend, remove seasonality from series and reduce the random component. For example, in order to emphasize the trend component in a daily time series (see Figure 4) one needs an average over 12 subsequent values to produce a deseasonalized series not influenced by seasonal component.

**Figure 3** Estimation of seasonal component in a monthly time series by using the harmonic function approach

Source: own elaboration.

**Figure 4** Estimation of seasonal component in a monthly time series by using the harmonic function approach

Source: own elaboration.

**EXPONENTIAL SMOOTHING APPROACH**

Moving averages can be seen as filters. When a trend is estimated and other fluctuations are smoothed, this is like filtering out the desired component and exclude the non-desired components. The simple moving average model described above treats the observations on which it is calculated equally and completely ignores all preceding observations. However, in some context,
a asymmetric filter is preferable in order to assign higher weights to more recent observations (i.e. the most recent observation should get a little more weight than 2nd most recent, and the 2nd most recent should get a little more weight than the 3rd most recent, and so on). A popular technique applied to get smoothed values, is the exponential smoothing (see Figure 5): \( x_t = \sum_{j=0}^{\alpha} (1 - \alpha)^j x_{t-j} \) where the weight applied to any observations \( \alpha (1 - \alpha)^j \) decrease geometrically.

Figure 5 Estimation of non-linear trend in a monthly time series by using exponential smoothing technique

Box and Jenkins (1970): The Modern Approach to Time Series Analysis

According to the modern approach to time series, the observed time series is conceived as a finite realization of a stochastic process. The stochastic part coincides with the systematic part of the data generating process and not simply the error term. This is formally described by the stochastic model: \( x_t = g(\varepsilon_t, \varepsilon_{t-1}, ...) \)

Compared to the classical approach, here the \( f(t) \) component is assumed to be absent or preliminarily removed and the focus is on \( g(\varepsilon_t, \varepsilon_{t-1}, ...) \), a correlated component process that governs the entire data generating mechanism. The modern approach is based on the Box-Jenkins procedure (Box and Jenkins, 1970). Box - Jenkins Analysis refers to a systematic method of identifying, fitting, checking and using integrated autoregressive, moving average (ARIMA) time series models. The method is appropriate for time series of medium to long length (at least 50 observations). ARIMA models are the most general class of models for forecasting a time series which can be made to be stationary\(^1\) by differencing, in conjunction with nonlinear transformations such as logging or deflating.

According to this approach, a time series random variable is usually viewed as a combination of signal and noise and the signal could be a pattern of fast or slow mean reversion or sinusoidal oscillation or rapid alternation in sign and it could also have a seasonal component. An ARIMA model therefore is a filter that separate the signal from the noise. As a result, the signal is extrapolated into the future to obtain forecasts. The ARIMA forecasting equation for a stationary time series is a linear equation where the predictors consist of lags of the dependent variable and/or lags of the forecast errors. The forecasted value of the dependent variable (X) is equal to a constant and/or a weighted sum of one or more recent values of X (autoregressive component) and/or a weighted sum of one or more recent values of the errors (moving average component). The acronym ARIMA stands for “Auto-Regressive Integrated Moving Average”. Lags of the

\(^1\) A time series random variable is stationary if its statistical characteristics are all constant over time (it has no trend, its variations around its mean have a constant amplitude and it wiggles in a consistent pattern).
stationarized series in the forecasting equation are called “autoregressive” terms, lags of the forecast errors, conversely, are called “moving average” terms. A time series which needs to be differenced to be made stationary is said to be an “integrated” version of a stationary series.

SARIMA MODELS

In order to taking into account the seasonal component, extant literature developed SARIMA models. The seasonal part of an ARIMA model has the same structure as the non-seasonal part: it may have an AR factor, an MA factor and/or an order of differencing. In the seasonal part of the model, all of these factors operate across multiples of lags (the number of periods in a season). A seasonal ARIMA model (SARIMA) is classified as an ARIMA(p,d,q)x(P,D,Q) model, where P=number of seasonal autoregressive (SAR) terms, D=number of seasonal differences, Q=number of seasonal moving average (SMA) terms (see Figure 6)

Figure 6 Monthly time series forecasting according to SARIMA model: ARIMA (0,0,2)(0,1,1) [seasonality=12]

Source: own elaboration.

Therefore, to sum up prior suggestions, time series analysis is one of the most well-known statistical techniques to make predictions under the assumption that existing patterns will continue into the future. It helps to identify and explain:

- Any regularity or systematic variation in the series of data which is due to seasonality—the “seasonals”;
- Cyclical patterns that repeat any two or three years or more;
- Trends in the data;
- Growth rates of these trends.

Advantages:

- The time series method is a useful tool to measure both financial and endogenous growth
- The time series method of forecasting is the most reliable when the data represents a broad time period
- Allows to distil trend, seasonality and clutter disadvantages
Deliverable 1.3
Port Traffic Forecasting Tool

- Over correcting for trends
- Lacks of precision
- Time delays during turning point assessment

Disadvantages:

- It is not quite suitable for identifying turning points (see Figure 7).

Figure 7 Example of unidentified turning point of a time series analysis prediction.
Source: own elaboration.

X11

Is a filter based method based on the Box Jenkins methodology. It provides seasonal adjustment and is often known as X11 style methods. These are based on the ‘ratio to moving average’ procedure described in 1931 by Fredrick R. Macaulay, of the National Bureau of Economic Research in the US. The procedure consists of the following steps:

1) Estimate the trend by a moving average
2) Remove the trend leaving the seasonal and irregular components
3) Estimate the seasonal component using moving averages to smooth out the irregulars.

Seasonality generally cannot be identified until the trend is known, however a good estimate of the trend cannot be made until the series has been seasonally adjusted. Therefore X11 uses an iterative approach to estimate the components of a time series. As a default, it assumes a multiplicative model. The model tests for the presence of identifiable seasonality; moving seasonality; residual seasonality in addition it automatically removes regression effects before ARIMA modelling and automatically selects the moving average range.

3.2.3 Causal models

These models are the most sophisticated kind of forecasting tool and express mathematically the relevant causal relationships. They incorporate the results of a time series analysis whilst taking into account everything known of the dynamics of the flow system.
Advantages

• explanatory power it allows for a understanding of the relationships among variables
• The lack of consistent superior performance of econometric models is disturbing, particularly as the extremely simple models perform nearly as well

Disadvantages

• Variables with insignificant coefficients have to be rejected according to the principles of econometrics
• Change in the relationships between variables increases with the projection period
• Substantial cost and time requirements

REGRESSION

Regression analysis is widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning. This methodology is also used to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships. In restricted circumstances, regression analysis can be used to infer causal relationships between the independent and dependent variables. However this can lead to illusions or false relationships, so caution is advisable.

ECONOMETRIC MODEL

Under this category of models we can find a lot of variations of the basic ordinary least square (OLS) regression model. Among them panel data regression model emerges as available solution for forecasting complex phenomena

Panel data (longitudinal time-series data) is a dataset in which the behavior of statical units are observed across time. Panel data allows to control for variables that cannot be observed or measure like cultural factors or difference in business practices across companies; or variables that change over time but not across entities (i.e. national policies, federal regulations, international agreements, etc.). Therefore panel data models account for individual heterogeneity. With panel data it’s possible to introduce variables at different levels of analysis (i.e. students, schools, districts, states) suitable for multilevel or hierarchical modeling. Some drawbacks are data collection issues (i.e. sampling design, coverage), non-response in the case of micro panels or cross-country dependency in the case of macro panels (i.e. correlation between countries). There are two approaches in panel data analysis:

Fixed-effects model. Fixed-effects model explores the relationship between predictor and outcome variables within an entity (country, person, company, etc.). Each entity has its own individual characteristics that may or may not influence the predictor variables (for example, being a male or female could influence the opinion toward certain issue; or the political system of a particular country could have some effect on trade or GDP; or the business practices of a company may influence its stock price). In this approach, it is assumed that something within the individual may impact or bias the predictors or the outcome variable. Fixed-effects model remove the effect of those time-invariant characteristics so it’s possible to assess the net effect of the predictors on the outcome variable. Fixed-effects will not work well with data for which within-cluster variation is minimal or for slow changing variables over time.
Random-effects model. Random effects model, unlike the fixed effects model, assumes that the variation across entities are random and uncorrelated with the predictor or independent variables included in the model. An advantage of random effects is that it’s possible to include time invariant variables (i.e. country of origin) that in fixed effects models are absorbed by the intercept. Random effects assume that the entity’s error term is not correlated with the predictors which allows for time-invariant variables to play a role as explanatory variables. In random-effects is crucial to specify all the individual characteristics that may influence the predictor variables. However, sometimes some of those variables may not be available therefore leading to omitted variable bias in the model.

**INTENTION TO BUY MODEL**

An intention to buy model is a combination of surveys performed on the possible clients. And trend mapping of projected sales.

**IO MODEL**

This model is a quantitative economic technique that represents the interdependencies between different branches of a national economy or different regional economies. Because the input–output model is fundamentally linear in nature, it lends itself to rapid computation as well as flexibility in computing the effects of changes in demand. It is also used to identify economically related industry clusters and also so-called "key" or "target" industries (industries that are most likely to enhance the internal coherence of a specified economy).

<table>
<thead>
<tr>
<th>Industry demand</th>
<th>Final demand</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td>Region R</td>
<td></td>
</tr>
<tr>
<td>Z^1i</td>
<td>Z^1R</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region R</td>
<td>F^1i</td>
<td>F^1R</td>
</tr>
<tr>
<td>Z^Ri</td>
<td>Z^RR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign</td>
<td>Foreign</td>
<td></td>
</tr>
<tr>
<td>imports</td>
<td>exports</td>
<td></td>
</tr>
<tr>
<td>Value added</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y^1</td>
<td>Y^R</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>x^1</td>
<td>x^R</td>
</tr>
<tr>
<td></td>
<td>z^e</td>
<td>z^R</td>
</tr>
<tr>
<td></td>
<td>c^1</td>
<td>c^R</td>
</tr>
</tbody>
</table>

*Figure 8 Example of IO model*

**ECONOMETRIC IO MODEL**

This technique is a combination of a IO model and an econometric model.

**DIFFUSION INDEX**

Diffusion indexes measure the proportion of the components that contribute positively to the index. The first step in computing the diffusion indexes is to calculate if a component increased, decreased, or had no change. Components that rise more than 0.05 percent are given a value of 1, components that change less than 0.05 percent are given a value of 0.5, and components that fall more than 0.05 percent are given a value of 0. Next, sum the values of the components. Third, divide by the number of components. Finally, multiply by 100.
LEADING INDICATOR

Leading indicators are events or measures that are a prelude to, and can thus help predict, another event or measure. If you have ever experienced a thunderstorm you know that a bolt of lightning always precedes a clap of thunder. Thunder is the sound made by lightning, but due to the relative speeds of light and sound, lightning is a leading indicator of thunder. This relationship is a good analogy for how leading indicators are useful in business. Many underlying economic conditions move through value chains and geographies at different speeds. Understanding and measuring the fastest moving conditions will give you better insight into those measures that are important to you, but do not move at the same rates. As an example, large computer hardware manufacturers such as IBM and HP are keenly interested in monthly measures of corporate IT spending, as these serve as indicators for how much hardware they will sell in future periods.

LIFE CYCLE ANALYSIS

Is a technique to assess the environmental aspects and potential impacts associated with a product, process, or service, by: Compiling an inventory of relevant energy and material inputs and environmental releases. Evaluating the potential environmental impacts associated with identified inputs and releases.

Figure 9 Life cycle model

3.2.4 Overview

Table 1 below provides an overview of the forecast methods available. Accuracy depicts the usefulness of the tools on short (0-1y) medium (1-5y) and long term forecasts (+5y). Identification of turning points expresses the ability of the methodology to indicate trend reversals. Data required offers an overview of the necessary data.
Academic research in port traffic forecasting

Over the last two decades, both academics and practitioners awarded a growing attention to the study of forecasting techniques in the port domain.

In particular, significant research efforts have been dedicated to forecast both port cargo throughput and container throughput (Lam et al., 2014). Indeed, cargo throughput has been proved to play a valuable role for each port. First, it constitutes the most relevant production index for evaluating port growth and success; second, it affects future port planning and development (Zhang et al., 2013). Analogously, extant literature recognizes that reliable models capable to forecast container throughput are necessary in order to support decisions on planning, upgrading and upsizing of container terminal facilities. The capacity of a container terminal, in fact, cannot easily be adjusted or increased in the short term by leveraging on inventory, outsourcing or overtime working (Peng and Chu, 2009). Capacity management and long-term investments, indeed, require accurate forecasts. The forecasting of handling volumes may provide terminal operators and port authorities (PAs) with useful data, for taking decisions and implementing new strategies more effectively (Talluri and Ryzin, 2004).

Several scholars investigated the accuracy of several forecasting techniques, as previously defined, in predicting port cargo throughput or container throughput. Both qualitative and quantitative methods have been considered. As concern qualitative methods, for example, Rashid et al. (2015) debated on scenario-based forecasting as a viable technique for predicting container throughput, focusing on the Hamburg - Le Havre range (Port of Antwerp). As uncertainty related to explanatory
variables is a serious concern in long-term forecasting, the Authors suggest scenario-based analysis as a viable tool for investigating the impact of various sources of the economic and transport trends on the container throughput's future trajectories.

Referring to quantitative methods based on time series analysis, Jiang and Lei (2009) compared nonlinear grey model with grey model to forecast cargo throughput while Xu (2011) proposed an autoregressive forecast model to predict cargo throughput. Peng and Chu (2009) investigated which model is capable of generating the most accurate prediction of container throughput. In particular, they tested six univariate models (classical decomposition model, trigonometric regression model, regression model with seasonal dummy variables, grey model, hybrid grey model and SARIMA model) on three major Taiwanese ports (i.e. Keelung, Taichung and Kaohsiung), in the 2003-2006 timeframe. In their study, the classical decomposition model emerges as the best forecasting model since it has the highest values of accuracy. Peng and Chu’s study (2009), also suggest that more sophisticated and complex statistical methods do not necessarily perform better that simpler ones.

A number of prior contributions focus on causal models. Lam et al. (2004), challenging port cargo throughput forecasting, propose and develop a neural network analysis in order to forecast 37 types of freight movements in the Port of Hong Kong. The explanatory factors included in the model are trade value of imports/exports/re-exports at 1990 prices, population, electricity demand, and Hong Kong Gross Domestic Product (GDP). The outputs of the analysis demonstrate that neural network models are more accurate than regression analysis when their results are compared with the observed data, but the reliability of the proposed neural network forecasting model decreases in the long-term. Chou et al. (2008), by focusing on import containers throughput, propose a modified regression model for forecasting the volumes of Taiwan’s import containers. In their study, they demonstrate that within quantitative causal model, the total forecast error for modified regression model is significantly lower than that for traditional regression models, which do not consider and modify the errors produced by the non-stationary contribution coefficient.

Some recent academic contributions tempted to compare or combine various forecasting methods and techniques. Chen and Chen (2010), for example, focus on time series analysis and create an optimal predictive model of volumes of container throughput at ports by using genetic programming (GP), decomposition approach (X-11) and seasonal auto regression integrated moving average (SARIMA). In their analysis performed on Taiwanese ports, within the Jan. 1978 - Dec. 2006 period, they discover that both in short- and long-term forecasting the sequence of container throughputs at sampled ports has a valuable trend and seasonal cycle nature. X-11, SARIMA and GP forecasts provide all accurate predictions, but the GP is proved to be the optimal method for predicting container throughput. Moreover, Gosagang et al. (2011), explore and compare the neural network method and linear regression technique for predicting the container throughput at Bangkok port. By testing these two forecasting techniques in the 1999-2010 timeframe, the Authors demonstrate that the accuracy of linear regression technique is less than neural network approach, using multilayer perceptron. Xiao et al. (2012) propose the application of a hybrid methodology, which includes qualitative and quantitative analysis to the container throughput forecast. In particular, they develop a feed forward neural network based on the improved particle swarm optimization with adaptive genetic operator for forecasting and then adjust the forecast results of this model with the knowledge from port experts. Empirical results on the Tianjin Port suggest that forecasts developed with their hybrid model are significantly more accurate than some other methods. Analogously, de Langen et al. (2012) argue that extant forecasting approaches adopted in port domain rely uniquely on trend forecasts and trend based models. These models, are not capable to capture disruptions of historical patterns, especially for certain types of commodities. As a result, they propose a forecasting approach that uses a model, in combination with expert judgements and commodity specificity research. This approach is applied to forecasts for 2030 volumes of all major commodities handled in the Hamburg - Le
Havre range, in four different scenarios. By incorporating expert-based analysis the proposed methodology provide valuable modifications to the assumption and results of the traditional transport model used in the study. Moreover, qualitative analysis (expert judgements) allows to identify throughput drivers for a number of commodity and likely disruptions of prior growth trajectories are identified. In addition, the paper demonstrated how relevant commodity specific trends are in forecasting overall port cargo throughput. Instead, Zhang et al. (2013) develop a combined model composed of grey-forecast and logistic-growth curve model in order to improve the accuracy of forecast model of cargo throughput for ports. They test this model on Chinese ports in the 2002-2011 timeframe, and demonstrate that combined models can obtain relatively higher forecast accuracy when it is difficult to gather sufficient data through field / desk research. In addition, the Authors state that forecast obtained with the combined model they propose are more accurate than any in individual ones. Finally, Xie et al. (2013) propose three hybrid approaches based on least squared support vector regression (LSSVR) model in order to forecast container throughput in two main Far East ports. Their outcomes demonstrate that the proposed hybrid approaches can achieve better forecasting performance than individual approaches.

Tables 2 and 3 report some of the most recent studies on forecasting technique in the port domain, focusing on port cargo throughput and container throughput. For each contribution key research dimensions are highlighted, including: authors, year of publication, main topic, aims of the paper, geographical coverage, method, techniques/models performed, sample timeframe, temporal perspective, time unit, accuracy measures and main conclusions.

Overall, the review performed on extant port literature facing the issue of forecasting methods and techniques provide some useful insights and recommendations applicable within the PORTOPIA project.

- Both qualitative and quantitative forecasting methods appear viable within the port context.
- Port cargo throughput and container throughput emerge as the predominantly investigated phenomena in extant port literature, due to their relevance for port planning and development.
- Both short-term and long-term perspectives are adopted in forecasting the above-mentioned phenomena. In this perspective, qualitative methods, e.g. Delphi method, market research, scenario-based forecasts, etc., as well as some quantitative techniques, which paved on time series analysis, e.g. classical decomposition model (X-11), SARIMA, etc., are the solutions preferred by academics for short-term predictions in the port domain. Also neural network approach is demonstrated to be an accurate technique but requires more data inputs. Conversely, time series analysis and causal models (traditional regression models, modified regression models, etc.) are commonly used in forecasting both port cargo throughput and container throughput in the long term.
- The results of prior studies on this topic unveil a certain degree of uncertainty in determining the best solution to adopt. In this vein, diverse accuracy measures used to assess the goodness of the statistical models proposed and developed can lead to different conclusions. Therefore, in the PORTOPIA Project, the adoption and the comparison of various accuracy measure is suggested. Root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percent error (MAPE) are the preferred accuracy measures used in similar studies.
- The selection of a specific forecasting technique does not depend only on the chosen temporal perspective (short/medium/long term) but also on the object of the forecast (i.e. the variable investigated in the forecast). So different forecasting methods and ad-hoc forecasting models should be implemented for each type of variable. Of course, this approach is not preferable within the PORTOPIA dashboard, which includes many relevant
Deliverable 1.3
Port Traffic Forecasting Tool

indicators. Hence, we could use (more) time-consuming and expensive techniques for developing some models capable to forecast the most significant variables in the port domain. Port cargo throughput and container throughput seem to be the most valuable variables to assess.

- Even when focusing on a single variable, forecasting methods should consider several aspects related to the investigated variable. If container throughput is estimated, for example, diverse types of traffic, e.g. seaborne traffic and inland waterways traffic (Lam et al., 2004), or import/export traffic flows (Coto-Millán et al., 2005), should be treated in a different manner by the predicting model. In this perspective, a forecasting model which fits well for a gateway port should not be optimal for a transshipment hub. As a result, various alternative specifications are expected to be introduced in order to take this concern into account. For example, de Langen et al. (2012) propose to consider three types of container volumes: direct deepsea, shortsea, and transshipment flows.
## Table 2 Forecasting methods and techniques in the port domain: recent academics contributions on cargo and container throughput

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Topic</th>
<th>Aims of the paper</th>
<th>Geographical coverage</th>
<th>Method</th>
<th>Temporal perspective</th>
<th>Time unit</th>
<th>Accuracy measures used</th>
<th>Main conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lam, Acce, Ng, Seabrooke, Hui</td>
<td>2004</td>
<td>Port cargo throughput</td>
<td>The paper proposes and develops neural network analysis for forecasting 37 types of freight movements in Hong Kong.</td>
<td>Hong Kong</td>
<td>Quantitative (Causality analysis)</td>
<td>Neural Network approach</td>
<td>1983 - 2003</td>
<td>Mean absolute error (MAE); R² (squared multiple correlation or the coefficient of determination).</td>
<td>The analysis demonstrates that neural network (NN) models are more accurate than regression analysis when their results are compared with the observed data. Further robustness checks, moreover, shows that the reliability of the proposed NN forecasting models decreases with the increasing time horizon.</td>
</tr>
<tr>
<td>Chou, Chu and Liang</td>
<td>2008</td>
<td>Import containers throughput</td>
<td>The paper proposes a modified regression model for forecasting the volumes of Taiwan's import containers.</td>
<td>Taiwan</td>
<td>Quantitative (Causality analysis)</td>
<td>Modified regression model</td>
<td>1989 - 2001</td>
<td>Total error.</td>
<td>The outcomes demonstrate that the total forecast errors for the proposed modified regression model is lower than that for traditional regression models, which do not consider and modify the errors produced by the non-stationary contribution coefficient.</td>
</tr>
<tr>
<td>Peng and Chu</td>
<td>2009</td>
<td>Container throughput</td>
<td>The purpose of the study is to search a model that is capable of generating the most accurate prediction of container throughput useful for port authorities.</td>
<td>Taiwan (3 major ports: Keelung, Taichung and Kaohsiung)</td>
<td>Quantitative (Time series analysis)</td>
<td>Six univariate models: classical decomposition model (X-11); trigonometric regression model; regression model with seasonal dummy variables; grey model; hybrid grey model and SARIMA model.</td>
<td>Jan. 2003 - Dec. 2006</td>
<td>Short term</td>
<td>Root mean squared error (RMSE); Mean absolute error (MAE); Mean absolute percent error (MAPE).</td>
</tr>
<tr>
<td>Chen and Chen</td>
<td>2010</td>
<td>Container throughput</td>
<td>The contribution aims at creating an optimal predictive model of volumes of container throughput at ports by using genetic programming (GP); decomposition approach (X-11) and seasonal auto regression integrated moving average (SARIMA).</td>
<td>Taiwan</td>
<td>Quantitative (Time series analysis)</td>
<td>Genetic programming (GP); decomposition approach (X-11); seasonal auto regression integrated moving average (SARIMA)</td>
<td>Jan. 1978 - Dec. 2006</td>
<td>Short term</td>
<td>Mean absolute percent error (MAPE).</td>
</tr>
<tr>
<td>Gooasang, Chandrasprakul, Kiattisin</td>
<td>2011</td>
<td>Container throughput</td>
<td>The manuscript aims to explore and compare the neural networks method and linear regression technique for predicting the container throughput at Bangkok Port.</td>
<td>Thailand (Bangkok Port)</td>
<td>Quantitative (Causality analysis)</td>
<td>Neural Network approach and linear regression technique</td>
<td>Jan. 1999 - Dec. 2010</td>
<td>Long term</td>
<td>Root mean squared error (RMSE); Mean absolute error (MAE).</td>
</tr>
</tbody>
</table>

Source: own elaboration.
### Table 3: Forecasting methods and techniques in the port domain: recent academics contributions on cargo and container throughput (continued)

Source: own elaboration.


3.4 Assessment of techniques in port data forecasting

Not all discussed methodologies are applicable in the Portopia forecast. Some issues may be present due to lack of available data, on other cases some methodologies might not allow the necessary precision or achievement of goals.

In order to provide the best possible results we will only take into account the methods which score above “fair” (all green values in the table).

For short term forecasting this would limit the selection (in order of quality) to:

- Market research
- Box Jenkins
- X11
- Trend projections
- Exponential smoothing
- Econometric model
- Regression

At first sight it would seem that market research, a qualitative methodology would be optimally combined with a Box Jenkins methodology. However Box Jenkins and X11 techniques are quite complex and hard to implement in a system. Trend projections and exponential smoothing on the other hand, both time series techniques are a good fit and both good to very good in accuracy. Implementation should be less costly and when combined with a qualitative aspect be perfectly able to forecast 1-3 quarters. The best match for short term forecasting would be a combination of a qualitative techniques with a time series analysis. Market research is already implemented in the project and could be complemented with external validation like the Delphy Method. For time series analysis trend projections or exponential smoothing are advised due to their low complexity and fast possible implementation. A slight advantage is put on trend projections since they allow to identify turning points better than exponential smoothing.

For medium term forecasting:

- Econometric model
- Regression
- Market research
- X11
- Trend projections

Medium term forecasting is best performed using causal models according to the table. Regressions and econometric models are best in accuracy and identification of turning points. This is a valuable option for this type of forecasting, however these methods require a lot of data and indirect indicators. Market research is a valuable option and already performed within the project. For X11 implementation the same issues exists as with short term forecasting. Trend projections are on the other hand an easy to implement guide for medium term forecasting and as an added bonus allow for a good identification of turning points. As best match a more in depth market research is proposed with trend projections allowing for a fast implementation and good accuracy.

For long term forecasting:

- Econometric model
- Regression
• Trend projections

Long term forecasting is often the most complex and sensible in selection of techniques. As with medium term forecasting econometric models and regression analysis are the best match. If data is available these two are an option but require a lot of input and work. Trend projections are also in this case a good match making it a good to very good option for all types of forecasting. The Commission already has a long term forecasting model based on econometric principles, it would seem advisable that this model is maintained and combined with qualitative forecasting techniques or trend projections from the Portopia program in order to validate the results.

3.5 Combining port development plans with forecasting

Cargo demand forecasting is a crucial activity for planning business development and building logistics infrastructures. For this reason, seaports are required to carefully monitor market changes and estimate industry trends before undertaking huge investments in new facilities that notably commit resources in the long-term. Hence, the realization of port infrastructures needs a great amount of financial public and private resources and an endowment of technical and organizational capabilities. Port planning consists in a complex analytical work that should be able to match cargo flow projection and future demand estimation with the setting of a suitable supply of infra- and supra-structures.

In this regard, Port Authority (PAs) is the public management body entitled to take executive decisions on port planning with the aim to enlarge port spaces and achieve a more efficient exploitation of the overall area. In so doing, the PA is expected to attract the interest of private partners for co-financing, developing and running new logistics facilities. The undertaking of sound forecasting activities by PA should unveil future market opportunities in various maritime business segments, facilitating a smooth allocation of technical and financial resources in the new ventures.

The analysis of a number of European Core Ports shows a quite dis-homogeneous behavior of PAs in relation to planning activities (i) and cargo forecasting (ii) (Tables 4 and 5). Overall, 28 PAs located in 16 EU countries are scrutinized, taking into account website contents, and focusing on statistics and historical traffic data, forecasting exercises and key archival documents, e.g., Master Plans, Annual Reports, Green Reports, etc. The degree of sophistication and thematic scope of the available planning documents are heavily affected by the diverse managerial approach of the PA in relation to strategic decisions and development plans. Indeed, in a number of cases, port planning activity is not practically formalized into an overarching document capable to summarize the main contents of the planning activity. Especially in the past, in fact, PAs were used to develop port areas and infrastructures by performing “step by step” expansion programs. These short/mid-term actions resulted into temporally fragmented development projects, drove to forget an unitary and logically-interconnected vision in port growth. In this perspective, PAs frequently undertook investments in specific business areas (and not in other), attracted by some favorable demand forecasts (that, in the end, may materialized in practice or even not). This mode of acting was also partially justified by those imitative choices of PAs that simply followed the mainstream behavior of the closest competitors. This definitely represents a myopic way of conceiving port planning, thus easily leading to inefficient and sub-optimal solutions. In recent years, however, a growing number of PAs was able to realize overarching long-term documents and planning efforts, ensuring economies of scale and scope as well as a more conscious analytical perspective in the modelling of port
expansion programs. In Europe, ports like Rotterdam, Hamburg, Antwerp made pioneering attempts in conjugating forecasting and planning activities, keeping the pace of the quick and profound modifications of the competitive landscape (e.g., containerization, diffusion of ICT, deployment of mega-vessels, growing market concentration, etc.).

As mentioned, the analysis of planning documentation and website contents regarding sample PAs enables to identify relevant differences in their strategic and forecasting behavior. In particular, the main addressed issues are the following.

1) A growing number of PAs develops formalized and well-established planning documents whereas other (medium-sized and minor) ports are acting on an informal and fragmented basis, by approving a (individual) project after the other. In the latter case, port development is not built around a common backbone but it is more the result of spot opportunities or political/market pressures.

2) Where available, port planning documents may have various forms and typologies depending on their role within the overall planning exercise. In particular, such documents may differ in terms of time horizons (mid-term, long-term) and nature (stand-alone vs. updating documents, strategic vs. tactical documents). Quite often the same PA can realize a two-tier planning activity based on the realization of documents addressing diverse temporal ranges (short/mid vs. long-term). Therefore, planning documents have a differentiated structure and a diverse portfolio of contents. Such documents are also characterized by a different degree of flexibility: some documents strictly define future investments and port development objectives, whereas others only sketch broader strategic views and project ideas/proposals. In practice, this intrinsic heterogeneity translates into a diverse naming:

- Business Plan
- Strategic Plan
- Master Plan
- 3-year Operating Plan
- Strategic Development Plan
- Progress Report.

3) In its planning documents, the PA may have a diverse attitude to the explicit disclosure of the contents and outcomes of the forecasting exercise. The release of detailed information regarding forecasting techniques and input data may provide insights on the accuracy and reliability of long-term visioning and strategic objectives. In other words, the disclosure of data and method regarding purpose-made forecasting might reinforce or weaken the arguments in support of investment decisions both in economic and social terms. Generally speaking, PAs are quite reluctant to provide information on this, preferring to keep some information asymmetries between themselves and the community of stakeholders. Apart from a few exceptions, forecasts seem almost “taken for granted” as planning documents are fully concentrated to describe the contents of expansion programs and investment timing and content.

4) Next to this, the PA attitude to publicly disclose on their company website traffic data, expansion programs, forecasting information, and planning documents is also a theme that should deserve more attention by practitioners and policy makers. The scrutiny of PA websites unveils a diverse willingness of PAs to disseminate information about planning and forecasting studies. Some PAs, like Rotterdam and Hamburg, extensively disclose throughput statistics and traffic data time-series,
as well as show their estimations for (long-term) economic and demand growth. On the contrary, other PAs report a minimum amount of historical data and do not provide (on the website) any information regarding planning documents and forecasting exercises.

5) In the PAs showing an attitude to the systematic realization of demand forecasts, it is relevant to understand the “developer” of the forecasting elaborations and outcomes (e.g., *ad-hoc* task force, internal bureau, business consultants, etc.). Typically, within their own organization, major PAs dedicate a specific function to port planning. In this regard, a team of experts might be dedicated to forecasting activities for supporting investment decisions. In other port contexts, instead, where the traffic size is rather small or the experience of managers in forecasting is weak, PA normally resorts to external advisors for the definition of some predicting variables and the set-up of an appropriate forecasting methodology.
## Deliverable 1.3
### Port Traffic Forecasting Tool

<table>
<thead>
<tr>
<th>Country</th>
<th>Core Ports</th>
<th>Statistics and historical data</th>
<th>Annual Reports (on-line)</th>
<th>Website contents and other (on-line) documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE</td>
<td>ANTWERP</td>
<td>Large amount of data and statistics are available</td>
<td>From 2006 onwards</td>
<td>&quot;2020 Master Plan&quot; is approved and includes some forecasting exercise (the full version is not available on-line) (LF)</td>
</tr>
<tr>
<td>BE</td>
<td>ZEEBRUGGE</td>
<td>A detailed Section on traffic statistics per segment is provided (from 2004 onwards)</td>
<td>From 2008 onwards</td>
<td>&quot;Strategic Port Infrastructure Project 2013-2022&quot; is under development. Full documentation is not provided on-line. Some forecasting is included in this long-term project (LF)</td>
</tr>
<tr>
<td>BE</td>
<td>GHENT</td>
<td>Some data and statistics</td>
<td>From 2009 onwards</td>
<td>&quot;Strategic Plan 2010-2020&quot; was made public in 2010 and it is available on-line. Simplified traffic forecast estimations are shown in the plan (LF)</td>
</tr>
<tr>
<td>BU</td>
<td>BURGAS</td>
<td>Only a few operational information available</td>
<td>Not available</td>
<td>No information about forecasting</td>
</tr>
<tr>
<td>CRO</td>
<td>RIJEKA</td>
<td>Some operational statistics on port operations and throughput (from 2009 onwards)</td>
<td>Not available</td>
<td>In the website there is Section called &quot;Plans&quot; talking about environmental protection programmes in the mid-term (2010-2013); in addition, there is a Section called &quot;Business potential&quot; dealing with the development and improvement of infrastructures in container and passenger segments. Some rough long-term forecasts are included (LF)</td>
</tr>
<tr>
<td>DE</td>
<td>HAMBURG</td>
<td>Long-term historical database on port throughput per traffic segment. Large empirical base.</td>
<td>From 2008 onwards. Some data and forecasting information are included (LF)</td>
<td>&quot;Port Development Plan to 2025&quot; (on-line). Some forecast estimations are included in long-term planning documents (LF). A rather sophisticated analysis is performed in relation to forecasting. In the long-term planning documentation some methodological notes and forecasting outcomes are disclosed (per business segment) (LF)</td>
</tr>
<tr>
<td>DE</td>
<td>BREMEN PORTS</td>
<td>Detailed data and statistics per commodity (from 2005 onwards).</td>
<td>From 2011 onwards</td>
<td>No relevant information about forecasting</td>
</tr>
<tr>
<td>DK</td>
<td>COPENHAGEN MALMO</td>
<td>Some basic statistics are available on the HTML pages of the website. Deeper contents are included in the Annual Report.</td>
<td>From 2001 onwards</td>
<td>No relevant information about forecasting</td>
</tr>
<tr>
<td>ESP</td>
<td>BARCELONA</td>
<td>Detailed data and statistics per commodity (from 1996 onwards).</td>
<td>From 2006 onwards</td>
<td>&quot;II Strategic Plan 2015-2020&quot; is approved. The contents of the Master Plan have been disclosed by the Port Authority President on February 2015. The full (electronic) version is not available yet. However, some long-term forecast estimations have been released (LF)</td>
</tr>
<tr>
<td>ESP</td>
<td>VALENCIA</td>
<td>Detailed traffic statistics and information (2013 onwards) are disclosed</td>
<td>From 2005 onwards</td>
<td>&quot;Strategic Plan for 2020&quot; is approved and currently under development. The full version is not available on-line. Some fragmented information discloses figures on future traffic estimations (LF)</td>
</tr>
<tr>
<td>ESP</td>
<td>ALGECIRAS</td>
<td>Some traffic statistics from 1999 onwards</td>
<td>From 2013 onwards</td>
<td>No relevant information about forecasting</td>
</tr>
<tr>
<td>ESP</td>
<td>BILBAO</td>
<td>Some historical data and information are available</td>
<td>From 2003 onwards</td>
<td>No relevant information about forecasting</td>
</tr>
<tr>
<td>EST</td>
<td>TALLINN</td>
<td>Some data and information are available in a statistical Section, but limited in time (2013 onwards). Most data are included in the Annual Report.</td>
<td>From 2000 onwards</td>
<td>A section on &quot;Performance result analysis&quot; is provided (2007 onwards), but no relevant information about forecasting is disclosed</td>
</tr>
</tbody>
</table>

**Notes:**
- **SF**: Short-term Forecast (0-1 year)
- **MF**: Mid-term Forecast (1-5 years)
- **LF**: Long-term Forecast (beyond 5 years)

---

**Table 4** The disclosure of forecasting contents in Port Authority’s website, planning documents and annual report.
<table>
<thead>
<tr>
<th>Country</th>
<th>Core Ports</th>
<th>Statistics and historical data</th>
<th>Annual Reports (on-line)</th>
<th>Website contents and other (on-line) documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR</td>
<td>MARSEILLE-FOS</td>
<td>Basic traffic statistics and Key Performance Indicators (KPIs) are available (from 2011 onwards)</td>
<td>From 2012 onwards</td>
<td>No relevant information about forecasting</td>
</tr>
<tr>
<td>FR</td>
<td>LE HAVRE</td>
<td>Some traffic statistics are available (from 2010 onwards) per business segment</td>
<td>From 2011 onwards</td>
<td>The integrated strategic plan “HAROPA 2030” aims to develop a long-term development of the following ports: Le Havre, Rouen and Paris. The documentation is available on-line and contains long-term forecasting and trend analysis from various perspectives (LF)</td>
</tr>
<tr>
<td>GR</td>
<td>PIRAEUS</td>
<td>Some historical data and information are available (from 2008 onwards)</td>
<td>From 2006 onwards. Some short-term forecasts is included (SF)</td>
<td>“Business Plan 2014-2018” has been approved (available on-line). Some mid-term forecasts are provided per traffic segment (MF)</td>
</tr>
<tr>
<td>IRL</td>
<td>DUBLIN</td>
<td>Some traffic statistics are available (from 2007 onwards) per business segment</td>
<td>From 2006 onwards</td>
<td>“The Masterplan 2012-2040” has been approved (available on-line). It contains an-hoc Section “Forecasts” which includes a rather sophisticated long-term analysis regarding economic projections and expected trade volumes (LF)</td>
</tr>
<tr>
<td>IT</td>
<td>GIOIA TAURO</td>
<td>Some basic statistics and traffic data are available</td>
<td>From 2013 onwards (basic)</td>
<td>“3-year Operating Plan” is approved but no relevant information about forecasting is added</td>
</tr>
<tr>
<td>IT</td>
<td>GENOVA</td>
<td>Some historical data and information are available (from 2008 onwards)</td>
<td>From 2009 onwards</td>
<td>In the website there is an ad-hoc Section on “Plans and forecasts” including some basic traffic expectations per cargo segment (mid-term (MF); “3-year Operating Plan” (2015-2017) is approved but no relevant information about forecasting is included; the new “Master plan” has been informally circulated among stakeholders: it contains long-term forecasts (LF)</td>
</tr>
<tr>
<td>IT</td>
<td>LA SPEZIA</td>
<td>Some statistics are available in different sections of the website (more systematically from 2011 onwards)</td>
<td>Not available on the PA Authority website. An abstract of the Annual Report in included in the Annual Report on Italian Ports released by the Ministry of Transport and Infrastructures</td>
<td>“3-year Operating Plan” has been issued but no relevant information about forecasting is added; in addition, the “Master Plan” approved in 2006 (available on-line) does not include any substantial information about traffic forecast</td>
</tr>
<tr>
<td>IT</td>
<td>NAPOLI</td>
<td>Detailed data and statistics per commodity (from 2001 onwards)</td>
<td>From 2013 onwards (basic)</td>
<td>Master plan has not been approved despite the first draft was issued by the Port Committee in 2000; there is no official information about forecasting on the PA website.</td>
</tr>
<tr>
<td>IT</td>
<td>VENEZIA</td>
<td>Some historical data and information are available (from 2006 onwards)</td>
<td>From 2007 onwards</td>
<td>“3-year Operating Plan” is approved (2013-2015) but no relevant information about forecasting is added</td>
</tr>
<tr>
<td>NL</td>
<td>ROTTERDAM</td>
<td>Large amount of data and statistics are available in a sophisticated Section of the website. Large empirical base.</td>
<td>From 2000 onwards</td>
<td>“Business Plan 2011-2015” includes mid-term forecasting (MF); in the “Port Vision 2030” plan some long-term traffic forecasting is provided. Estimated figures are regularly updates in the “Progress Reports” (LF)</td>
</tr>
<tr>
<td>NL</td>
<td>AMSTERDAM</td>
<td>Large database on statistics and historical traffic data (from early-1990s onwards)</td>
<td>From 2008 onwards</td>
<td>“Port Vision 2008-2030” plan is under development and includes long-term forecasting estimates; a “Port Vision 2030” document has been also drafted: it updates traffic forecasts (LF)</td>
</tr>
<tr>
<td>PL</td>
<td>GDANSK</td>
<td>Some historical data and traffic information are available (from 2004 onwards)</td>
<td>Not available</td>
<td>“Strategy 2027” plan has been issued in 2013. It contains long-term economic and traffic forecasts (LF)</td>
</tr>
<tr>
<td>PT</td>
<td>LISBON</td>
<td>Traffic statistics are available but contents are irregular and fragmented (from 1997 onwards)</td>
<td>From 2006 onwards</td>
<td>“Strategic Development Plan (2025)” has been issued by PA; the full document is not available on-line. Only a short brochure is provided and this does not contain any reference to long-term traffic forecasting.</td>
</tr>
<tr>
<td>RO</td>
<td>CONSTANTA</td>
<td>Some historical data and traffic information are available per commodity (from 2006 onwards)</td>
<td>From 2001 onwards</td>
<td>“Master Plan 2014-2020” has been issued but documentation is not available online; the PA developed some long-term forecasting but information are fragmented (LF)</td>
</tr>
<tr>
<td>SW</td>
<td>GOTENBURG</td>
<td>Large database on statistics and historical traffic data (from early-1970s onwards)</td>
<td>Not available</td>
<td>“Master Plan 2035” was adopted in 2014 by the PA; full documentation is not available online, but some news on this unveil the presence of forecasting activity inside the planning document (LF)</td>
</tr>
</tbody>
</table>

Notes:
- **SF** Short-term Forecast (0-1 year)
- **MF** Mid-term Forecast (1-5 years)
- **LF** Long-term Forecast (beyond 5 years)

Table 5 The disclosure of forecasting contents in Port Authority’s website, planning documents and annual report (continued).
4 FORECASTING EXAMPLES IN SIMILAR SECTORS

4.1 Air Transport

4.1.1 Boeing and Airbus forecasts

The most frequently consulted literature in air cargo business on forecasting air cargo are the market forecasts by Airbus and Boeing. These largest airplane manufacturers issue yearly reports on the state of air cargo and the expected future demand. The forecasts in both manufacturers’ reports are created using basic econometric techniques, such as trend analysis. The expected demand is discussed for continental trade flows, such as Asia to North America or Europe to Africa without a distinction in transported commodities. An issue with these two forecasts is that there is often a bias present for ‘positive growth results’ due to the fact that the authors are airplane manufacturers and therefore benefit from growing cargo flows.

The companies release demand forecasts with compound annual growth rate (CAGR) for aggregated trade flows, such as Europe to North-America. The forecast approach is very similar for both companies and includes:

- Econometric modeling, a tool in determining the overall importance of economic factors
- Explicitly named techniques are not available in the report, but additional correspondence with an analyst from Boeing clarified the econometric modeling to be trend analysis
- Explicitly named economic factors are GDP, transportation costs, exchange rates, real income, investments, export/imports, industrial production and relative prices
- Judgmental modifications to account for changes in growth factors. Examples include air service agreements, trade quotas, restrictions on night operations, increase in capacity, route re-structuring

4.1.2 The International Air Transport Association (IATA) Airline Industry Forecast 2014-2018

The IATA forecast also works with an annual growth rate (CAGR) for the calculation of its 5 year based forecasts. (total market growth prediction is currently of 4.1% over the next five years). Emerging economies, particularly in the Middle East and Africa, will be the fastest-growing markets. The Airline Industry Forecast Is (according to the authors) the most comprehensive tool for analyzing the future freight traffic demand, as well as the evolution of global air travel market.

The report provides detailed five-year traffic forecasts for over 3,000 individual country-pairs, plus aggregate results at regional and global levels. The forecast is derived from the results of a survey done to the industry’s major airlines, civil aviation and airport authorities. The aggregation rules for forecasting include: over 1,000 unduplicated international country pairs, including aggregated values for 6 world regions, 17 world sub regions, and over 900 country to sub region forecasts.

Following indicators are provided by the IATA platform:

- Traffic volume, number of tonnes of freight
- Compound Annual Growth Rate (CAGR)
• Traffic growth in volumes

A summary report by IATA Economics provides analysis of the macroeconomic events shaping the industry and their direct implications for airlines’ projections. It also identifies key challenges and opportunities across top routes and regions. The users get access to annual reports and gain the fundamental tools and data that are required for basic forecasting. This report allows them to understand the market shape and identify the most important new emerging markets as well as where to maximize business investment and build effective strategy plans with intelligence directly from the industry's leading experts.

4.1.3 Lessons from the air transport forecasting

The air business is an interesting sector since it is closely related to the maritime industry. Notable links are:

• High demand for freight and passenger transport
• Forecasting mostly done for fleet development
• Fleets are also high in capital investment
• Often commissioned by liner companies in order to assess current order book evolution

Of course it is not possible to compare all aspects of the airline industry with the maritime sector. Main differences include:

• Cargo is often time-sensitive and high value: perishables
• Cargo is often process critical; medicine or machine parts
• Cargo is of very high value; high-tech or precious metals

Keeping these similarities and differences in mind the following lessons can be derived from the existing forecasts.

• There are many freight rates: for different cargoes, different vessels, and different contract types. But these different freight rates are often surprisingly strongly correlated over time, when seen in a multi-year perspective. This is mainly because of the strong degree of substitutability of cargoes among vessels across routes: most cargoes can be transported on a number of different ships, so conditions spread quickly from trade to trade. Secondly, investors and financial institutions are undifferentiated as to specific trades and thus provide additional substitutability.
• Aggregation should be available on a multitude of levels with a maximum of flexibility
• Objectivity should be a major parameter avoiding the skewedness created by industry forecasters
• The IT system should allow for a maximum of user freedom and interpretation but the added value of an annual comprehensive report is substantial as is proven by the IATA example

4.2 Rail transport

4.2.1 Railway Demand Forecasting and Service Planning Processes (Rail Freight Service Review)
These reports are based on extensive interviews with CN and CP subject matter experts directly involved in and responsible for the development of the railways' demand forecasting and service planning processes. QGI has also relied on the experience and background of the team members that participated in the interviews and prepared this report. There are two types of forecasting included, demand forecasting for medium-long and short term planning. Many railway planning processes including financial, asset, capital investment and service planning are predicated on the railways’ estimate of the freight volumes (demand) to be handled within a given time period. For this information to be meaningful and support the railways’ planning processes demand forecasts must not only provide an estimate of the total volumes but also identify the commodities, timing, and the origins and destinations of the traffic. Medium to long term planning activities provide the railways with a highly aggregated view of demand for time series forecasting.

![Figure 10 Transnet level of detail](image)

Forecast information can be input at any level of detail and the planning system will automatically aggregate up or disaggregate down. Disaggregation of forecast data to the lowest level of detail is done on the basis of historical traffic movements. Detailed traffic forecasts, including those at the individual customer level, can subsequently be adjusted manually to introduce more current customer or market information.

While shippers have expressed the view that they believe railway demand forecasts do not always reflect the true demand they communicate to the railways the railways argue that they must weigh the input of shippers against the expected performance of broader markets and consider potential volatility in demand. QGI’s interviews with railway representatives revealed that many believe that some of their customers have difficulty providing accurate medium to longer term demand forecasts. Railway staff identified many reasons for this including:

- Inefficiencies in customers’ logistics management processes;
- unrealistic or inaccurate market forecasts and market share projections;
• an unwillingness of customers to reveal their own marketing projections.

At an aggregate level railway forecasts of demand are accurate within approximately 10% over the course of a year

4.2.2 Transnet freight demand forecast

The transnet forecasting methodology has been included for its completeness and high level of complexity. The model encompasses multiple layers and is used for all strategic matters ranging from allocation of resources to the creation of a long term capacity planning framework.

<table>
<thead>
<tr>
<th>Planning layer</th>
<th>Source</th>
<th>Forecast period</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational Demand</td>
<td>Contracts and customer needs, resource planning and scheduling.</td>
<td>Daily/weekly/monthly/annually</td>
<td>To allocate resources according to contracted demand.</td>
</tr>
<tr>
<td>MDS Demand (Corporate Plan)</td>
<td>Market intelligence, commercial agreements, 30-year Demand Book, capacity planning</td>
<td>7 – 10 years</td>
<td>To plan and fund business objectives against confirmed and anticipated customer demand.</td>
</tr>
<tr>
<td>LTPF Demand</td>
<td>Modelled macroeconomic growth, industry sector growth projections, historic data, 30-year Demand Book.</td>
<td>30 years</td>
<td>To provide a planning framework for capacity expansions.</td>
</tr>
</tbody>
</table>

Table 6 The transnet forecasting timeframe

It is a combination of 4 models namely:

• Freight Demand Model (FDM): which investigates the sources of supply and demand in the economy, disaggregated to 352 districts and 72 commodities. This model essentially translates economic activity in the form of currency (Rand) into production and consumption of goods in the form of tons. The forecasts generated by this model are based on macroeconomic growth scenarios

• Liquid Fuels Demand Model (LFDM): The LFDM uses supply-side capacity intelligence and matches the fuel production forecasts with demand and then supplies the projected shortfalls with fuel imports.

• Transportation Model (TTM) The TTM uses gravity flow methodology, a well-established technique to model the flow of goods, people, and so forth. The route that freight will follow is determined by the least total resistance of the connecting links from origin to destination.

• Market Share Model (MSM): is created in order to calculate the rail addressable market and to determine rail targets over the longer term. By comparing planned seven-year volumes with what is available in the market, it enables the planners to check the realism of short term targets and make informed longer-term projections for each commodity and on each route.
The input drivers for freight forecasting are a combination of factors which ultimately are converted to sector and commodity-specific compound annual growth rates. Macroeconomic inputs include, import/export forecasts, industry trends, Government expenditure and global trends both regionally and worldwide.

4.2.3 Lessons from the rail transport industry

- Always aggregate based on commodities and geographical area
- Keep a clear split between long, medium and short term
- Look at proxies and drivers and make sure that the right drivers are linked to the right underlying commodities

4.3 Traffic flow forecasting

4.3.1 Traffic flow: Scenario, Traffic Forecast and Analysis of Traffic on the TEN-T, Taking into Consideration the External Dimension of the Union.

This forecast was included since it has a direct link with the Portopia work due to its focus on the European transport sector and the TENT network. Since it was published in 2009 we will focus on the methodology and not the results.

The study uses the TRANS-TOOLS model and has provided a coherent and reliable forecast for the future traffic flows in EU and the neighboring countries for 2020 and 2030 with particular focus on flows between old and new Member States, between new Member States and between EU and the neighboring countries. The analysis has covered all transport modes, both freight and passenger traffic on links and through nodes with a focus on the TEN-T network.

The study uses proxies as drivers in order to estimate the future traffic flows and creates 2 scenarios for the situation in 2020 and 2030 a Baseline scenario and a scenario termed “Sustainable Economic Development”. Generally speaking, there are two types of drivers respectively increasing or decreasing transport demand. Examples of the first type could be economic growth or population growth. An example of the second type could be increasing energy prices.
The baseline is a “Business as Usual” scenario including already agreed infrastructure and policy measures and in line with the current trends. The Sustainable Economic Development scenario describes a faster economic and demographic development, higher fuel costs and provide for an intensive development of the road and rail networks.

The Baseline and the Sustainable Economic Development scenarios have been illustrated by short descriptions of the possible futures and by parameters corresponding to the input for the TRANS-TOOLS model. On this basis forecasts for all modes for 2020 and 2030 in the Baseline and Sustainable Economic Development scenarios for both freight and passenger transport have been produced. Subsequently, these forecasts have then been applied for identifying the most important infrastructure axes for cohesion, development of the internal market and the relations to neighboring countries.

4.4 Academic research

The airline research More academic literature uses, amongst other techniques: The Weighted Majority Algorithm (WMA) lets a set of macroeconomic variables make individual forecasts which assigns normalized weights to the predictors based on their in-sample forecasting accuracy, variations on the Box-Jenkins method described in previous sections (Faraway and Chatfield, 1997), A self-organized, five-layer neuro-fuzzy model is developed to model the dynamics and forecast air passengers (Chen, 2012), A system dynamics simulation (Suryani et al., 2012) and one of the more recent approaches uses a support vector machine to test a different number of input nodes to optimize the predicted cargo own from Beijing to Shanghai (Heng, 2013).

Academic attempts will not be discussed in detail in this deliverable since they often require too much time or resource investment and are too complex to implement.

5 FORECASTING IN PORTOPIA

To handle the increasing variety and complexity of managerial forecasting problems, many forecasting techniques have been developed in recent years. Each has its special use, and care must be taken to select the correct technique for a particular application. The selection of a method depends on many factors—the context of the forecast, the relevance and availability of historical data, the degree of accuracy desirable, the time period to be forecast, the cost/benefit (or value) of the forecast to the company, and the time available for making the analysis.

The factors below have been taken into account with the selection of forecasting techniques:

What is the cost associated with developing the forecasting model, compared with potential gains resulting from its use?

- 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 much data points are available? Techniques vary in the amount of data they require.
5.1 Methodology

The goal of this deliverable is to determine the forecasting technique to be used in the system which is applicable to all developed indicators. Each indicator is investigated based on its compatibility with the proposed methodologies. The forecasting will be split up in short- medium and long term periods for maximum relevance and should be implementable into the IT system supporting the Portopia platform.

5.1.1 Short term

As can be seen from the previous sections the short term forecasting timelines are often between a few days and a couple of months in length. The less complex models often give quite accurate matches limiting allocation of resources.

Figure 9

Figure 12 variability on timeframes

Time frame

In order to create a relevant forecasting methodology a difference has to be made between forecasting timeframes. As a rule of thumb the closer the forecasting frontier the more reliable the result. Therefore different methodologies have to be applied to different timeframes.

Within the system the short term forecasting will focus on a period of 1 to 4 quarters. The quarter (3 months) is used as a basis since it is the smallest possible timeframe within the Portopia data reporting system for all proposed indicators. This means that the system should allow for a quarterly update of the indicator forecast.

Applicable methodologies

For the short term forecasting system it is important that all proposed methodologies are totally automated and can run with minimum input from system administrators. In line with section [nr] we will only use the selected methodologies which are applicable to the project and compare them to the different possible indicators.
Deliverable 1.3
Port Traffic Forecasting Tool

### Table 7 Market trends and structure indicators forecast overview

The market trends and structure indicators are best suited for forecast applications. Due to their high data availability and comparability forecast can be made on multiple levels and multiple forecast periods.

<table>
<thead>
<tr>
<th>Market trends and structure indicators</th>
<th>Forecast possible</th>
<th>Smallest timeframe</th>
<th>Added value of forecast</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call size</td>
<td>Yes</td>
<td>Annually</td>
<td>High</td>
<td>M</td>
</tr>
<tr>
<td>GDP growth vs tonne per type of cargo</td>
<td>Yes</td>
<td>Annually</td>
<td>High</td>
<td>M L</td>
</tr>
<tr>
<td>Intra-European traffic incidence in Eu ports</td>
<td>Yes</td>
<td>Quarterly</td>
<td>High</td>
<td>S M L</td>
</tr>
<tr>
<td>Maritime traffic</td>
<td>Yes</td>
<td>Quarterly</td>
<td>High</td>
<td>S M L</td>
</tr>
<tr>
<td>Market share</td>
<td>Yes</td>
<td>Quarterly</td>
<td>High</td>
<td>S M L</td>
</tr>
<tr>
<td>Modal split</td>
<td>Yes</td>
<td>Annually</td>
<td>High</td>
<td>L</td>
</tr>
<tr>
<td>Traffic Growth</td>
<td>Yes</td>
<td>Quarterly</td>
<td>High</td>
<td>S M L</td>
</tr>
<tr>
<td>Transhipment incidence in EU ports</td>
<td>Yes</td>
<td>Quarterly</td>
<td>High</td>
<td>S M L</td>
</tr>
</tbody>
</table>

### Table 8 Socio economic indicators forecast overview

Socio economic indicators are suited for medium and long term forecasts. The complexity involved renders them difficult to incorporate into the IT systems.

<table>
<thead>
<tr>
<th>Socio economic indicators</th>
<th>Forecast possible</th>
<th>Smallest timeframe</th>
<th>Added value of forecast</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct employment</td>
<td>Yes</td>
<td>Annually</td>
<td>High</td>
<td>M L</td>
</tr>
<tr>
<td>Direct value added</td>
<td>Yes</td>
<td>Annually</td>
<td>High</td>
<td>M L</td>
</tr>
</tbody>
</table>

### Table 9 Environmental indicators forecast overview

<table>
<thead>
<tr>
<th>Environmental and safety indicators</th>
<th>Forecast possible</th>
<th>Smallest timeframe</th>
<th>Added value of forecast</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air quality</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Carbon footprint</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Days lost</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Definition of objectives and targets for environmental improvement</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differentiated fees for clean shipping</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy consumption</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>Environmental management index</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environmental policy makes reference to espo guideline documents</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environmental responsibilities of key personnel documented</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environmental training program for port employees</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Existence of an environmental management system</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Existence of an environmental policy</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatal work accidents</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Inventory of relevant environmental legislation</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment in protection</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquefied natural gas bunkering</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marine ecosystems</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nautical accidents</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Noise</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>On-shore power supply</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Port security incidents</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Publicaton of a publicly available environmental reports</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sediment quality</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil quality</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terrestrial habitats</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top ten environmental priorities</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waste</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water consumption</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Water quality</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work related accidents</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Logistics chain and operational performance indicators</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average port dues per ton</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average THCs</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landslide congestion</td>
<td>Yes</td>
<td>Annually</td>
<td>Medium</td>
<td>M L</td>
</tr>
<tr>
<td>Maritime connectivity</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maritime fluidity indicator</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

39
Environmental and safety indicators are harder to forecast since a lot of these indicators are binary or qualitative. The indicators that are possible to forecast can only be forecasted in medium to long term forecasts.

<table>
<thead>
<tr>
<th>Governance indicators</th>
<th>Forecast possible</th>
<th>Smallest timeframe</th>
<th>Added value of forecast</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomous management</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extent of performance management</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration of the port cluster</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermodal container connectivity</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market openness</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Port authority employee productivity</td>
<td>Yes</td>
<td>Annually</td>
<td>Low</td>
<td>M L</td>
</tr>
<tr>
<td>Port authority investments</td>
<td>Yes</td>
<td>Annually</td>
<td>Low</td>
<td>M L</td>
</tr>
<tr>
<td>Reporting Corporate social responsibility</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety and security indicator</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10 Governance indicatos forecast overview

Like the socio economic indicators, governance indicators are also difficult to forecast. Only investments and productivity can be forecast but the added value of forecasting these indicators is rather limited.

<table>
<thead>
<tr>
<th>User perception indicators</th>
<th>Forecast possible</th>
<th>Smallest timeframe</th>
<th>Complexity of forecast</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall user perception quality index</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11 User perception indicatos forecast overview

Short term forecast product development

**Time series forecast system**

The time series forecast will be applies to all indicators with sufficient data points and is based on a moving average or exponential smoothing mother due to their good matches with the forecasting timeframe and low resource requirement. They can be incorporated into the dashboards fairly easily and offer a quick overview of the expectations of the coming months.

Once the forecast methodology is further developed other options could be included into the system allowing the user to switch between different forecast techniques.

**Qualitative forecast/ confidence index**

In addition to the basic times series forecast based on the indicators developed during the project we propose the introduction of the confidence index. A forecast technique based on the qualitative market research/Delphi method methodology. Inclusion of this parameter has multiple advantages:

- It can be used as a check for the time series forecast
- A new indicator is created within the Portopia project
- If can be introduced within the system in combination with the quarterly update of data
- Combines the strengths of qualitative and time series forecasting

5.1.2 Medium term

Timeline
The medium term forecasting timelines runs from 1 to 5 years, due to the low update requirement the automation level of this type of forecast is not as high as with the short term forecast tool. This will allow for a comparison to other sectors.

Methodologies

Possible methodologies for allowing medium term forecasting include market research, historical analogy, X11, Trend projections, regression and econometric IO models. The proposed approach for medium term forecasting would be to use proxy forecasting applied in other sectors. This means linking the indicators to certain sectors which drive the underlying proxy.

Medium term forecast product

Annual forecast update

This product would combine the medium and long term forecasting sections. This annual update and extraction from the system will allow an overview of the forecasting data and work performed.

5.1.3 Long term

Timeline

Long term forecasting will deal with a timespan of over 5 years and with a maximum of 20 years comparable to the masterplan forecasting done by ports. This renders the forecast quite complex and with a high degree of uncertainty.

Applicable methodologies

Trend projections is the only time series technique applicable for long term forecasts with acceptable standards. Due to the high degree of complexity the more resource intensive methodologies are better suited for this type of analysis. A full econometric model, regression analysis or input output model would be best suited for this timeframe. The issue is that these type of analyses are not possible within this project since they require too much resources, be it data storage or man hours.

Long term forecasting product

Update of DG move model

The PwC/NEA (2013) and TRANStools and Primes/Tremove models were used for the traffic forecast analysis conducted for the preparation of the White Paper on EU Transport Policy in 2011. The forecasts are based on a scenario that assumes a status quo of existing policies and already planned policy reforms, but at the same time assumes that these policy reforms do not create a level playing field for all 319 TEN-T ports (of which 83 ports are part of the core network). Following the low growth scenario the cargo throughput in the ports of EU27 would grow from 3.6 billion tons in 2011 to 5.8 billion tons in 2030. The EC argues that this growth would cause capacity problems and an unbalanced use of the port system and associated network.
The forecasts were obtained by running a combination of several models namely a GDP growth model PRIMES (which calculated the long term average GDP growth rates in the EU of 1.4%) and the Trans-tool model. As mentioned in the original document these projections must be taken with caution because of the multiple underlying assumptions. Particular attention must be given to new developments, for example the introduction of new or raising trade barriers or further world trade liberalization. The scenario used as a baseline assumes that the current state of affairs will prevail; it does not consider a sensitive analysis about possible trade agreements. However, the forecasting methodology used by PwC/NEA does not take into account competitive cargo shifts between ports. The baseline forecast only allows mapping the geographical distribution of trade to change, but it does not contain any assumptions about competition between ports.

The proposition for the long term forecast is to interpret the results obtained from the DG move model and update them into the annual forecast document provided by the system.

### 5.1.4 Implementation

The next phase of the project will include implementation of the IT system. During the first phase of implementation the forecast system will be created for all compatible indicators which are already online.

During the second phase a more complex forecast system will be integrated into the existing system allowing the users more freedom in analysis. Also the annual forecast update will be developed and made available on the website.
5.2 References


SEA - Study Centre for the Expansion of Antwerp (till 1985). Prognose van de Trafiek van de Haven van Antwerpen, Antwerpen, several editions.

Deliverable 1.3
Port Traffic Forecasting Tool


