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

Deliverable 9.1

A review of existing benchmarking techniques and practices

DOCUMENT ID
PORTOPIA|D|9.1|DT|2014.07.11|V|FINAL

DUE DATE OF DELIVERABLE
2014-06-30

ACTUAL SUBMISSION DATE
2014-07-11

DISSEMINATION LEVEL
PU (Public)

DELIVERABLE 9.1

A review of existing benchmarking techniques and practices

AUTHORSHIP

Author(s)	Vonck, I. & Notteboom, T.
Beneficiary Partner	ITMMA – University of Antwerp (UA)
Issue Date	2014-07-11
Revision	De Schepper S.
Status	Final
Contributors	Marintek
Pages	38
Figures	13
Tables	3
Annexes	3

SIGNATURES

Author(s)	
Coordinator	

DELIVERABLE 9.1

A review of existing benchmarking techniques and practices

REVISION

Revision Date	Reviewer Name	Reviewer Signature
2014-12-05	De Schepper, S.	

DELIVERABLE 9.1

A review of existing benchmarking techniques and practices

Summary Report

The aim of WP9 is to develop benchmarking and weighing techniques that allow individual sea and inland ports to compare their activities and operations with the EU average and relevant peer groups, and with ports in other important regions like Asia and the Americas (for seaports).

In its most literal definition, benchmarking is a continuous process of evaluation of products, services and practices with respect to those of the strongest competitors or of the enterprises recognized as leaders. In a direct way, the benchmarking is a process of evaluation and improvement of performance. Therefore it can be used for assessing the competitive situation of units or even systems as a whole. Within the port, and maritime industry, benchmarking is an important used by the majority of companies and organizations in order to improve competitive practices, functions and products.

Benchmarking originated from the geodesy, where measurements were done compared to a point of reference or a benchmark. It is a method to determine the function of activity of a process or KPI of an organization. A second aspect is to improve the investigated aspect within the company. In essence are two phases, one of comparison and one of improvement. There are multiple goals to benchmarking; internal objective image competition, best in class etc. best for us I an added value for the strategic level of the cluster.

Although benchmarking is traditionally a one on one comparison of 2 specific units of analysis, a majority of clustering techniques exist on comparing a specific unit to a chosen aggregate. A good example for the financial market can be found in stock analyses where the debt/equity from any specific company is often compared to that of the market in general, or a sub category (like top 10/15 companies, companies in the same geographical area etc.). In further deliverables this will become one of the major challenges to achieve. Deliverable 9.1 however only deals with a summation and theoretical background on the large variety of techniques currently in existence.

The structure of this deliverable will be as follows: in the first part we will outline the importance of a complete benchmarking tool in the maritime industry. The second part will focus on the existing techniques, levels and possibilities of benchmarking. The third part offers an overview of existing work done in sectors, which offer large similarities to the port industry (like the aviation industry, shipping industry, city development and chemical industry). Fourthly a discussion is presented on what should be included in the benchmark engine and which issues must be tackled. Finally a conclusion will be offered with an interim assessment on preferred benchmarking techniques for the PORTOPIA program.

DELIVERABLE 9.1

A review of existing benchmarking techniques and practices

TABLE OF CONTENT

DELIVERABLE 9.1

A review of existing benchmarking techniques and practices

LIST OF TABLES

Deliverable 9.1

A review of existing benchmarking techniques and practices

LIST OF FIGURES

1 POSITIONING OF THIS REPORT IN WP9

The aim of WP9 is to develop benchmarking and weighing techniques that allow individual sea and inland ports to compare their activities and operations with the EU average and relevant peer groups, and with ports in other important regions like Asia and the Americas (for seaports). The specific objectives of the work package include:

- Review of existing benchmarking techniques and practices;
- Designing methodology to aggregate results of individual ports;
- Selection and customisation of benchmarking techniques;
- International benchmarking against non-European port systems.

Deliverable 9.1 (D9.1) is focused on the review of existing benchmarking techniques and practices.

2 THE IMPORTANCE OF BENCHMARKING IN THE MARITIME CLUSTER

Seaport functions are diverse in scope and nature and have evolved over time. During the past decades port entities have undergone consistent change on the macro- and micro economical level. The roles and functions each port undertakes can be identified through political geographical (urban and spatial), economic and social perspectives. This is off course important when outlining an approach for benchmarking purposes. There are three major ways of looking at ports and investigating the impact of all these functions. The three approaches are directly related to ways in which a port can be assessed and compared to its peers.

From a macro-analytical and public policy perspective ports are viewed as economic catalysts for the regions they serve where the aggregation of services and activities generates benefits and socio-economic wealth. Ports create direct and indirect value-added and employment. Consequently, they receive a lot of attention as part of national or supranational maritime cluster policy and industrial policy. From this macro point of view we therefore find a large array of indicators linked to regional added value creation. Also port development is often associated with urban planning. The emphasis here lies on the port-city interface and on waterfront redevelopment and other initiatives to re-establish the link between port area and city. Sadly enough, until now, planners often fail to place the link between port and city, it is however of large importance to look at indicators affecting both. From an environmental perspective, port planning and management should ensure sustainable development. Environmental sustainability of port projects has become as important as economic and financial viability. Ports often form an integral part of coastal management policies. As such ports have adopted a real environmental role and function. Benchmarking any form of sustainability measures from both macro and micro perspective will be of great added value.

Another way to look at ports follows a micro-economic perspective. This is benchmarking in its purest form. Port operations are usually oriented towards the two traditional components of ships and cargo. Services to ships include those performed at the sea or waterways side (dredging, pilotage, mooring/unmooring, etc.) and at the ship/shore interface (berthing, repair and maintenance, supply and bunkering, etc.). Services to cargo can be divided into those performed at the ship/shore interface (stowing, loading, discharging, etc.) and those entirely performed in land-side areas such as consolidation, storage and distribution. Key in the micro-perspective approach are the concepts of efficiency/performance and sustainability at the operational level (i.e. a company or terminal).

A third approach consists of a hybrid perspective; the wave of corporatization of port authorities reflects this increasing business and market-oriented approach to port management. Ports are part of a wider logistics and production system and increasingly performing teleport functions. Port functions are extended to trade, logistics and production centres with an extensive portfolio of operations spanning across production, trade and service industries. This approach is best suited when analysing risk in a port setting. It allows for a full assessment of causes and effects and looks at the full integration of a port system from the single entity company involved in land-water transshipment to the node in global supply chains. Benchmarking tools would focus more upon the cooperation and connections of the ports to the surrounding network.

A full list of indicators linked to the different aspects of ports can be found in the fact-finding and EcoPorts reports; these results are based on the previous PPRISM documents. One could argue that port benchmarking is an exercise that has to be made on micro- economical level as explained in the previous part. That it only matters for companies striving for a competitive advantage, be it industrial powerhouses or the more governmental port authorities. However the necessity of a good benchmarking framework lies within many factors and changes within the industry.

2.1 A changing industry drives benchmarking needs

The need for benchmarks is driven by a growth in uncertainty both from within and outside the maritime and port industry. Companies and corporations have always had to accept risk-taking as an inherent part of business. Capitalism revolves around the tenant that the more risks you take, the more you stand to win or lose. This foundation for doing business hasn't changed. What has changed however is: the quotient of risk needed to deliver the target reward and the way in which the risk is structured, perceived and managed. Benchmarking is an important tool used by managers in the quest for reduced uncertainty. All the forms of uncertainty affecting ports can be measured and defined in a variety of ways. It is however no the main goal of the deliverable to measure uncertainty but to provide a clear overview of possible benchmarking techniques. However, a list of major changes affecting the performance and viability or sustainability of ports must be included to assess the necessity of certain indicators in later stages.

Ships are getting bigger and are putting increased stress on port infrastructure. This trend affects ports in more ways than just draft restrictions, which is in itself an important port indicator. The entire infrastructure had to be adapted to accommodate these giants and the sheer amounts of volume they bring with them. Comparisons of ports aimed at mapping the best performances in this frame are not only interesting to the authorities themselves but also to the companies both in- and outside the port. Providing them with an overview of possibilities within the same geographical area. Indicators allowing an overview of the situation include: channel depth, dock depth, quay length, average dwell time, rail infrastructure, crane size (bulk), etc.

Within the entire port setting the power balance is shifting at an alarming rate. This continuous evolution and shifting invokes the need for benchmarks on different levels. One of the common examples which is very popular these days is the changing position of traders. The traditional picture of the landscape with a small number of professional traders and the abundance of customers is under pressure with an increased number of actors acquiring trading skills. While there is still a significant group of pure traders, quite a few traders have vertically integrated their activities and became asset-based, even controlling key storage facilities and assets. A benchmarking exercise showing best available lands, most activities surrounding certain cargo's and best surrounding added value activities provides valuable insights for traders and port managers alike. The big question remains, what will the future bring? Against the rising tide of unstable and rising prices for many of core commodities in food and industrial supplies, the entire industry is on the verge of change. Traders are not the only group affecting ports at their core. Other 'adjacent' activities to the maritime industry like the presence of insurance, brokerage, education etc. are to be included in the picture and become important benchmarking aspects.

The entire sustainability framework drives continuous innovation and pushes existing infrastructure to its limits. NOx and SOx emissions are only the tip of the proverbial iceberg when discussing sustainable ports. Sustainability measures are increasingly

important when applied to port choice. A 'green' port is often considered 'superior' to a heavily polluting counterpart. A set of indicators allowing for a qualitative and quantitative baseline will be a major asset to both ports and companies in their attempts to minimize their global (carbon) footprint.

Another important issue is that ports are privatizing and or corporatizing all over the world, leaving the protective governmental environment as they go. It will therefore become ever more important to look at clear indicators showing performance in order to both continuously improve and assess the situation. Competition is increasing, pushing the need for benchmarks on all levels. For ports as a whole the need is driver by the wave of privatization. However, also the companies themselves should see merit in port wide benchmarks. Allowing them to look at possible alternatives on the market and optimize their location and structure. Since the best way of coping with uncertainty within logistics networks and EDC location conundrums is betting on top locations (Vonck and Notteboom, 2013), port selection is a major aspect for these entities.

Finally the uncertainty on many levels, which is inherent to not only the port but also the entire macro-economic maritime sector, drives the need for benchmarking ever upwards. For one, the classic view on supply chain engineering used to be very linear in the past. In recent years, logistics networks have become far more complex leading to additional risks and vulnerabilities. The connectivity dashboard of PORTOPIA will provide valuable insights on this matter. Macro-economic changes like globalization and containerization have contributed to the emergence of reverse logistics, the decoupling of order and delivery and the creation of transport networks instead of chains. Add to this the shift process in corridors, which have become the main arteries of world trade, and the picture becomes even more blurry, as strategic points along maritime corridors such as the Panama Canal, Suez Canal, the Straits of Malacca and the Straits of Gibraltar function as important turntables in extensive hub-and-spoke and relay/interlining activities. Cargo flows are more uncertain and footloose than ever before with new possibilities peaking and the need for a clear overview driven by a powerful benchmarking engine rising.

2.2 Growing academic interest

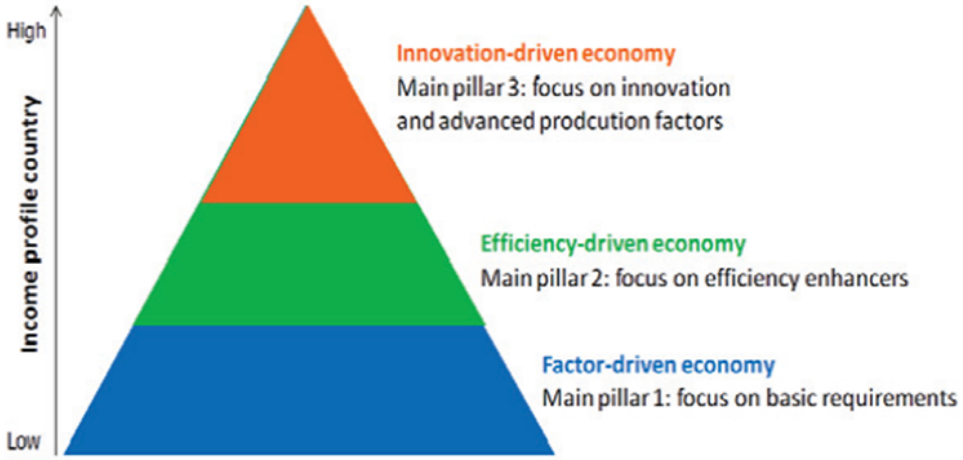
Benchmarking is one of those rare subjects, which was first introduced in an industry environment before being accepted by researchers as an attention-grabbing topic. Today however, we see an increased interest in these types of studies. Before benchmarking became the norm, other techniques were used when analysing strengths and weaknesses, for example the competitive analysis method and quality function deployment (QFD). The first technique was used by organizations as a means of collection data and measures regarding markets sales products production costs or budgets of competitors. The main difference with benchmarking lies in the fact that it does not provide insights as to how the competitors achieved this position. QFD interprets costumer needs and expectations and states them in terms of technical requirements. Like with the competitive analysis approach lies more on the "what" and not the "how". Even though there are not the full benchmarking we have today the literature was filled with case studies and illustrations of successful initiatives. These days, the benchmarking concept is understood to be an act of imitating or copying. But in reality this proves to be a concept that helps in innovation rather than imitation, as stated by Thompson and Cox (1997). Many authors have contributed to the literature on benchmarking resulting in more than 350 publications as of June 2002. With the specific trend of academics growing interest.

2.3 Benchmarking affects all layers of the industry

We already mentioned performance and increased competition in the previous section. Figure 1 adapted from Van Den Bosch et al. (2011) shows the evolution a country or business undergoes in order to remain competitive. When trying to create a competitive economy a country goes through multiple phases. For high-income countries like region of Western EU, innovation and advanced production factors become essential in order to remain competitive. Ports themselves are centres of innovation but a constant pressure on rates and prices still forces the efficiency parameters to remain within their competitive boundaries. The tools used in benchmarking are tailored for this purpose. A more hands on approach in visualizing this concept can be provided by splitting up the benchmark target groups in the three levels of management linked to the different types of economy.

The evolution most port systems have undergone results in an added value of benchmarking on different levels within the port organization. Ports grew from basic operational units where the only activity was transfer from land to sea or a pure operational level system. The second level or efficiency driven economy arose with the introduction of production and efficiency goals. Management has always determined these goals. Since we are currently in an innovation driven economy the strategic department or in our case top level management has also much to gain with a high level benchmark. On an operational level a pure factor level analysis comes in to play where internal and external benchmarks allow port companies to objectively seek the best port location. On a management level efficiently elements are more important opening the doors to best practice benchmarking and benchmark exercises linked to efficiency goals.

Figure 1 Different phases of economic development



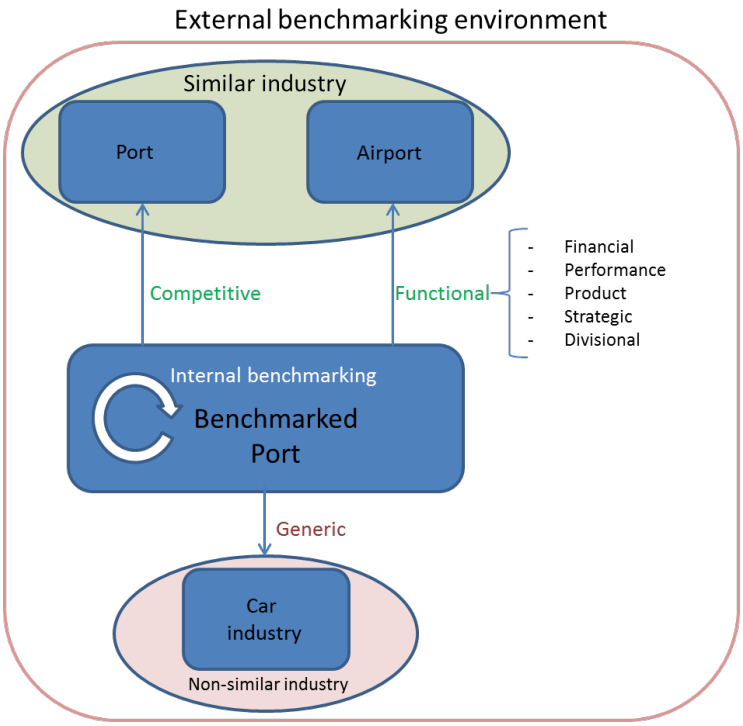
Source: Van Den Bosch et al. (2011)

3 DIFFERENT BENCHMARKING METHODOLOGIES

3.1 General benchmarking guidelines

As with many strategic management tools there is an abundance of possibilities when addressing benchmarking. Figure 2 below gives an overview of the different classifications within this complex subject.

Figure 2 Overview of benchmark types



Benchmarking can take place in two environments, namely internal and external. The latter can be divided in three sub types namely:

- Competitive: a benchmark process between a unit and a direct competitor (port-port);
- Functional: a benchmark process between a unit and a comparable entity (port-airport);
- Generic: a benchmark process between a unit and a different entity (port-university).

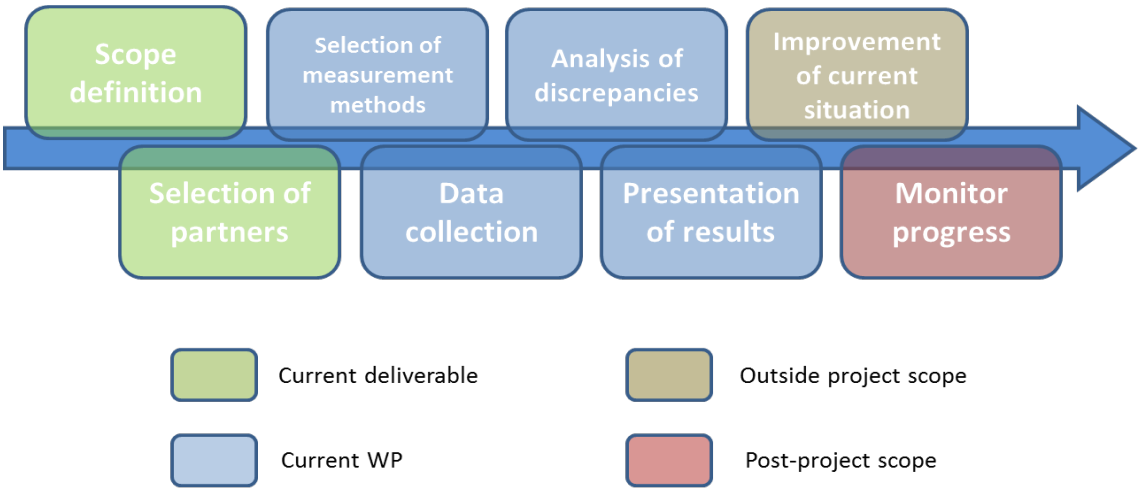
Next to the four main types of benchmarks, there are six sub categories grouping the specific indicators which are measured.

- Process benchmarking: focus on observation and investigation of business processes with a goal of identifying and observing the best practices from one or more benchmark firms.
- Financial benchmarking: performing a financial analysis and comparing the results in an effort to assess your overall competitiveness;
- Performance benchmarking: assessing the competitive position by comparing products and services with those of target firms;

- Product benchmarking: designing new products or upgrades to current ones.
- Strategic benchmarking: involves observing how others compete. This type is usually not industry specific meaning it is best to look at other industries;
- Departmental benchmarking: a company will focus its benchmarking on a single function in order to improve the operation of that particular function. Complex functions such as Human Resources, Finance and Accounting and Information and Communication Technology are unlikely to be directly comparable in cost and efficiency terms and may need to be disaggregated into processes to make valid comparison.

The benchmarking process itself is divided in eight phases. The first one being scope definition, followed by the selection of partners, choice of measurement methods, data collection, analysis of discrepancies, presentation or results, improvement of current situation and monitoring of progress and on-going benchmarks. Of these phases the first two are direct goals of this deliverable as can be seen in Figure 3.

Figure 3 Benchmarking phases and timing



3.2 Institutions dealing with transport statistics

Three public institutions in Europe deal with the collection and publication of comprehensive transport statistics on a supranational scale. The following overview provides these institutions with major publications (Romereskirchen, 2000):

- Eurostat
  - Transport: annual statistics
  - EU transport in figures
- Unctad
  - Review of maritime transport
- ECMT
  - Statistical trends in, transport
  - Trends in the transport sector
- UNECE
  - ANNUAL BULLETIN OF TRANSPORT STATISTICS FOR Europe and North America

Furthermore there are many nationally organized and unofficial organisations that produce or publish transport statistics. International road federation, international civil aviation organisation, union of railways, German economic research institute, etc.

3.3 Tools and techniques

3.3.1 Academic tools

The following rules are adapted from Benchmarking European Transport (Deiss, 2000). These rules outline the basic economic principles behind transport benchmarking (an overview is given in Table 1).

Table 1 Different benchmarking techniques

	Explanation	Advantages	Disadvantages
PPM	Ratios showing efficiency based on a partial analysis of one output over one input	- Simple - Low data requirement - Easily interpreted	- No global image - No substitution or combination possibility between input and output
DEA	Non-parametric method using linear programing which creates efficiency border	- No assumption on CP functions - Low data requirement	- Prone to data errors - Measures inefficiency but does not explain it
SFA	DEA based method including a inefficiency variable caused by a lack or mistake in data	- Explains inefficiency - Explains outliers - Extra statistical tests possible	- Insufficient data for specification of variables functions and errors
TFP	Ratios showing efficiency based on an weighed aggregate of all inputs and outputs	- Global overview - Shift of input variables possible	- Requires information on index pricing - Extensive data requirement
VFP	TPF with a constant cost level for input variables	- In inconsistent data for cost factors	- Does not include all costs
MPI	Measures productivity change between two separate moments in time	- No cost for input output required	- Extensive data requirement
CI	An index composed out of sub-indexes, each based on different variables	- Simple model	- Extensive data requirement - Difficult annual evolution

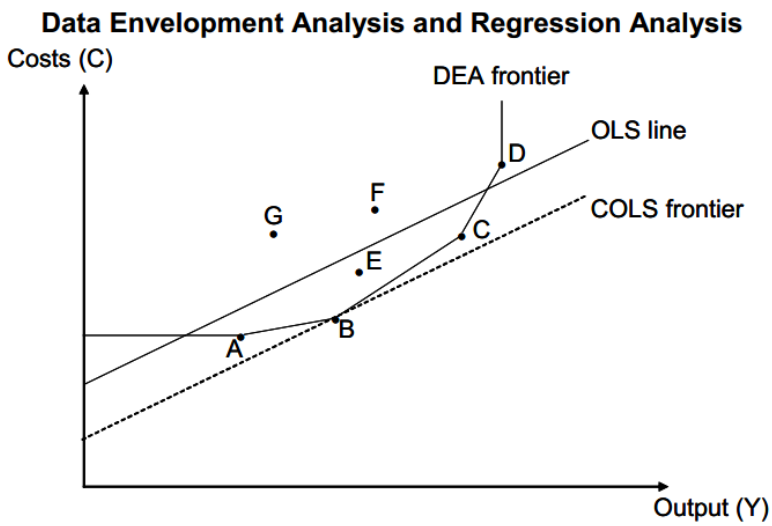
Source: Airport Benchmarking 2011

Within the academic sector port or terminal benchmarking is often done with specific tools like DEA (Data Envelopment Analysis) and SFA (Stochastic Frontier Analysis) analysis. These quantitative techniques focus on the production function of ports or terminals, thereby making relative comparisons between outputs vs. inputs to measure relative efficiency. Another approach is to use the cost function as a basis. There are parametric techniques (which include SFA, Corrected Ordinary Least Squares (COLS) and others), which assume a particular specification of the relationship between a firm's costs and a set of cost drivers, which might include, for example, the outputs produced, input prices and a range of exogenous factors. Econometric analysis is then used to estimate the parameters of that relationship. Having estimated a cost function, inefficiency is one of the factors (alongside others such as omitted variables, measurement errors and so on) that can explain the differences between the observed

level of costs for a particular firm and the level of cost predicted by the estimated cost function.

Data envelopment analysis (DEA) is a nonparametric method in operations research and economics for the estimation of production frontiers based on linear programming techniques. Since it is one of the most frequently used techniques we will delve a little bit further into the subject. It is used to empirically measure productive efficiency of decision-making units (or DMUs). Although DEA has a strong link to production theory in economics, the tool is also used for benchmarking in operations management, where a set of measures is selected to benchmark the performance of manufacturing and service operations. Utilizing the selected variables, such as unit cost and output, DEA software searches for the points with the lowest unit cost for any given output, connecting those points to form the efficiency frontier. In the circumstance of benchmarking, the efficient DMUs, as defined by DEA, may not necessarily form a "production frontier", but rather lead to a "best-practice frontier" (Cook, Tone and Zhu, 2014). One can also combine the relative strengths from each of these approaches in a hybrid method (Tofallis, 2001) where the frontier units are first identified by DEA and then a smooth surface is fitted to these. This allows a best-practice relationship between multiple outputs and multiple inputs to be estimated. A classic DEA representation is provided in Figure 4.

Figure 4 Classic DEA representation



Source: NERA Economic Consulting

Some of the advantages of DEA are:

- No need to explicitly specify a mathematical form for the production function;
- Proven to be useful in uncovering relationships that remain hidden for other methodologies;
- Capable of handling multiple inputs and outputs;
- Capable of being used with any input-output measurement;
- The sources of inefficiency can be analysed and quantified for every evaluated unit.

Some of the disadvantages of DEA are:

- Results are sensitive to the selection of inputs and outputs (Berg 2010);

- You cannot test for the best specification (Berg 2010);
- The number of efficient firms on the frontier tends to increase with the number of inputs and output variables (Berg 2010).

3.3.2 Tools from practice

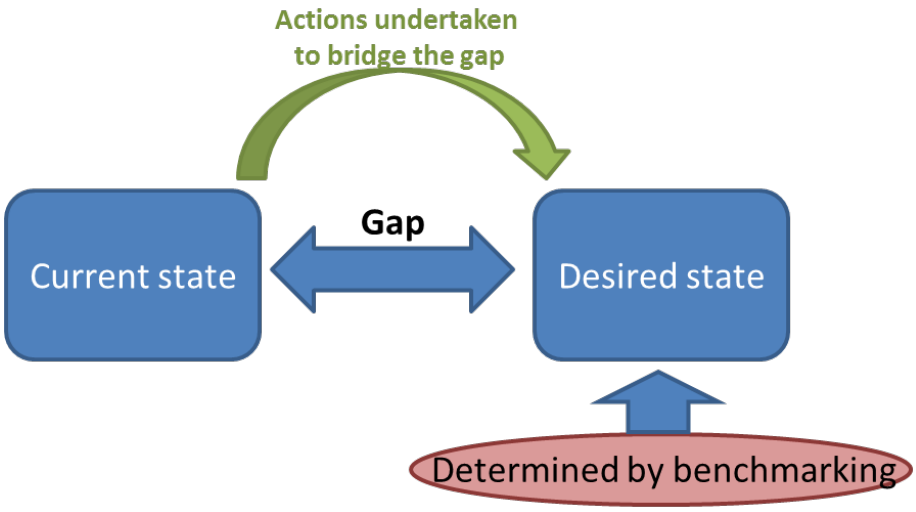
Academics are not the only ones interested in the applications of benchmarking. The anecdote "you can't manage what you can't measure" has become a cliché in the business world. Within the PORTOPIA project the use of web-based information "dashboards" has a central role. This method is already used within many industries as a means of gaining greater insight into the effectiveness of business operations. Forrester (2010) reviewed the state of the art, including surveys of 22 early-adopter companies that were using computer-based information dashboards. An important lesson to take away from previous work is that, if dashboards aren't connected to the people who "own" the processes they are evaluating, then the information does not become actionable.

The industry uses a wide range of methods and tools in order to manage the benchmarking process. A non-exhaustive list can be found in the Appendix. We will discuss three of the most frequently used methods below.

1. Gap Analysis

A Gap Analysis is a basic method used to determine the steps needed to get from a current state to a desired future state. By conducting a Gap Analysis, institutions can identify what they need to do to "bridge the gap" and make a project a success. Although this may seem to differ from the definition of benchmarking Figure 5 shows the link with the benchmarking process. A gap analysis in itself is impossible without some form of benchmarking at its source. A 'future desired state' is always based on variables obtained from either an internal or external benchmarking action.

Figure 5 Gap analysis in benchmarking



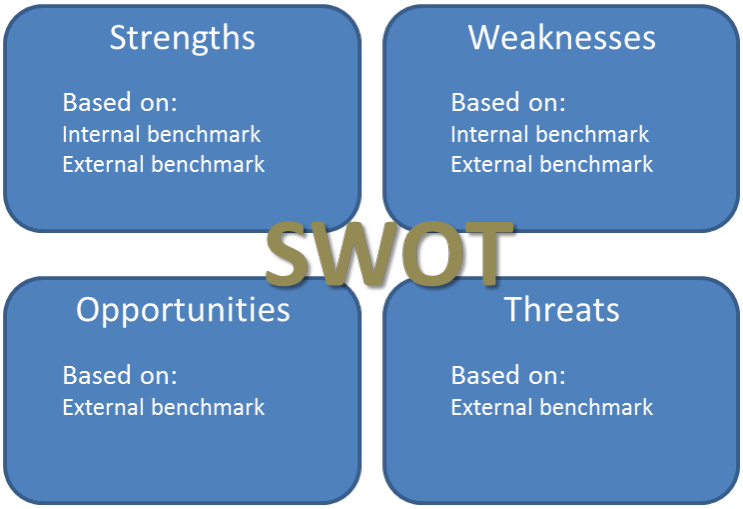
Within the port industry gap analysis are used in the following instances:

- Analysis the requirements for access canal dredging;
- IT system analysis;
- Resilience and vulnerability analysis;
- Project management.

2. SWOT Analysis

As is the case with the gap analysis, the SWOT analysis is also an indirect tool to benchmark an organization (example of a SWOT tool is given in Figure 6). It allows a company to assess a changing environment and respond proactively by mapping strengths, weaknesses, opportunities and threats. By using the tool correctly, and combining it with a benchmark, an institution can map where change is possible and refine plans mid-course. As stated before it is not the means to benchmark but more the result of a combined benchmark or the way in which to present benchmark results.

Figure 6 Benchmark SWOT tool



Within the port sector SWOT analyses are primarily used for the benchmarking and assessment of either individual ports or entire regions. Most applications reside within the strategic planning department. The benchmarking emphasis of SWOT in seaports exists mostly in the threats segment where ports seem to focus their attention on their competitors and the specific abilities they have.

3. Use of ratios

An important aspect of benchmarking, especially within the PORTOPIA setting are ratios. Initially used primarily in financial benchmarking, they are an important tool for industry-based analysis. A ratio is basically a number representing a comparison between two things. Financial ratios in particular, allow for comparisons and, therefore, are intertwined with the process of benchmarking, comparing one's business to that of relevant others or of the same company at a different point in time processes on a specific indicator or series of indicators.

An added bonus when using ratios is that, most of the time, they are well-established calculations derived from verified data. The most useful comparisons can be made when metrics definitions are common and consistent between compared units and over time. Benchmarking using ratio analysis can be useful to a wide variety of audiences. For example, for an investor, benchmarking can involve comparing a company to peer companies that can be considered a viable alternative investment. In this process, the investor may compare the focus company to others in the peer group is we shift the focal point towards ports, possible direct investments can be managed and shifted due to constructive benchmarking. From a management perspective however, benchmarking with ratio analysis may be a way for a manager to compare their company to peers using externally recognizable, quantitative data. Within the port

industry this reflects both managers on an aggregate port level as on a smaller inter-port company level.

In the appendix an example is given of how ratios can be visualized within a dashboard like system. The presentation uses financial ratios but can be a useful guideline in further development of the benchmarking dashboard of PORTOPIA. The information is from the FINTELOnline Benchmarking Report on 01/07/2014.

Ratios which can be used within PORTOPIA were already developed in the PPRISM project and will be discussed further in the next section of this document.

3.4 Benchmarking as part of the management of the transport sector

As mentioned before, commercial companies use benchmarking in order to improve their performance and gain a heads up on their competitors. The transport sector is no different on a local and competitive level. However at EU and Member state level benchmarking implies measuring the performance of the transport sector and comparing results with fixed points or standards (OECD, 2000).

Both academic and private sector benchmarking tools exist. However since the public sector works on a higher level the dimensions of the problem increase and the used theories require modification. Transport benchmarking is prevalent in the railway sector, waterway and air transport system an id thee linkage of logistical nodes or telecommunication systems. For an entity like a port this offers a wide array of underlying indicators due to the inherent complexity. Its productivity and efficiency depend on geographical and external factors, the demographic structure of hinterland and local industry and differences in management culture and national policies.

Examples of existing benchmarks on country level are: the Netherlands general efficiency of the freight transport sector, efficient utilization of infrastructure, messenger intermodality, promotion of cycling, port hinterland traffic, inland navigation), Austria (modal share public transport, modal share railway passenger transport, infrastructure pricing, CO2 emission from transport); Sweden (road safety transport, port hinterland container traffic by rail, modal share rail freight transport), the UK (road transport safety, efficiency of air transport sector and Denmark (infrastructure quality).

4 BENCHMARKIGN IN SIMILAR SECTORS

4.1 The airline sector

The benchmarking of airports is quite comparable with sea- or inland ports. However since airports are often private institutions more emphasis is aimed at OPEX and other financial indicators than with seaports. Overall in the last 15 years privatization and commercialization of an increasing number airports have contributed to a substantial number of benchmarking studies. The majority of the focus is on a national level. However international studies are increasingly common. Next to the financial aspect of airports the capacity utilization is also one of the major subjects of analysis. A third aspect surrounds the need for better public policy assessment and aims at the main question if privatized airports operate more efficiently than publicly owned airports. This last point can be translated to the port setting, which we are attempting to benchmark.

Due to the wide variety and number of benchmarking studies out there we will provide a short literature review. It will be aimed at showing the different challenges and lessons to take away from the analysis of this sector without going in too much detail.

4.1.1 Air Transport Research Society: Airport Benchmarking Report

The purpose of the ATRS study is to benchmark airports worldwide and to categorize them into three different geographic regions. These regions consist of North America, Europe and the Asia-Pacific Region. The studies measure productivity and efficiency; unit cost and cost competitiveness; and financial performance. Measurements of Factor Productivity (variable and total) include some factors that are beyond managerial control and also influence productivity and efficiency measurements. The ATRS study is one example of a more academic approach with the underlying framework being a regression analysis, which differentiates between the factor-productivities.

The study focuses on three different parts:

- Productivity and operating efficiency;
- Unit cost competitiveness;
- Airport user charges.

The underlying database contains historical information since 2002 on financial data, traffic, capacity, major airports and airports groups. The data is furthermore segregated into different indicators including but not limited to: traffic statistics and composition, characteristics of the airport (e.g., runway, and terminal). The underlying methodology is based on a multilateral output (input) index procedure, which in essence is a composite index aggregation method.

4.1.2 BOOZ European Airport Benchmarking Study 2012

This document was commissioned by BAA to benchmark OPEX performance in particular of one of the major airports in Europe against a variety of European comparators. The study uses a selection process to identify several European airports that could be considered comparable to Heathrow on a specific set of infrastructure indicators. The study is particular in the way that it does not include any Asian or American airports and is therefore restricted to a specific range of competing airports. This is an important aspect to take away for the benchmarking process in ports where competition is also partially based on a geographical and infrastructural scale.

Throughout the study there are several adjustment in order to make the OPEX and CAPEX comparable for all investigated entities. This is a problem we will also face when comparing seaports. For example in order to allow for the rail and NATS utility costs at Heathrow an econometric model has been developed to account for these differences. The model furthermore uses a "residual' approach whereby costs are separated into controllable and non-controllable.

The conclusions of the study go further than pure benchmarking results. Each analysis links the results with ongoing macro-economic changes or country-specific variables (the uncontrollable variables mentioned earlier). A special note is included in the assessment about sample size, which is smaller than other benchmark studies, but according to the authors more relevant due to the geographical range scope. The results of the exercise are also compared to other studies and datasets in order to validate aggregated data. Within the seaport setting this can for example be done between the PORTOPIA RES system and Eurostat data.

4.1.3 Lessons from airline benchmarking

- Classification into valid geographical regions is a necessity;
- Next to geographical ranges attention must be directed to the directly competing entities;
- A clear set of rules is needed to describe entities (size, capabilities, ...);
- Aggregate analyses are common and useful, not everything can be benchmarked one on one.

4.2 Shipping benchmarking

For benchmarking in the shipping sector we turn to both the inland part of shipping and the seaborne part. Shipping its true maritime context means the transport of goods by boats. However on a more supply chain level also intermodal transport and road transport can be considered as relevant information for the term 'shipping'.

4.2.1 OECD intermodal benchmarking report

The Organisation for Economic Co-operation and Development (OECD) has an ongoing commitment to freight-related transport research. The Programme of Research on Road Transport and Intermodal Linkages (RTR), which is involved with multimodal transport strategies, economic performance of transport infrastructure and sustainable development focuses on, amongst other things, benchmarking and system performance measures of intermodal traffic. The overall emphasis of the work on benchmarking is centred on comparing the relative efficiency of modes, modal combinations and modal interfaces.

An important lesson to take away from the OECD intermodal benchmarking exercise is the different goals of benchmarking with regards to the different stakeholders in the benchmarked entity. If we transpose this to the required benchmarking system of PORTOPIA we have to deduce that the benchmarking had to be tailored to the main user, being ports.

Consumers' perspective (better services at lower prices):

- Evaluate reduction in logistics costs, taking into account consumer service levels

Shippers' and logistics service providers' perspective (better services at lower cost):

- Accurate evaluation of reductions in logistics costs in relation to the quality of logistics services and transportation;
- Estimation of total logistics costs, including production, sales, collection, storage, transportation and data processing;
- Selection of the most appropriate transportation service(s) from origin to destination by assessing all the factors, including time, cost, reliability and flexibility.

Government perspective (balance between efficiency and environmental friendliness of logistics, improvement in transport safety):

- Efficiency: efficient infrastructure development projects in terms of the advancement of logistics;
- Deregulation and standardisation to enable improvements in logistics efficiency;
- Environmentally friendly logistics: comprehensive assessment of external costs of logistics systems (i.e. air and noise pollution, congestion), and improvement in intermodal transport systems;
- Improvement in transport safety.

The first aspect the OECD study covers is the selection of indicators, in their methodology they look at previously developed indicators. Here we already see an example of the wide variety of possible indicators with over 300 possibilities for transport related benchmarking. Once the selection of indicators is done some issues appear. Many measurement issues are primarily a consequence of the availability and reliability of appropriate performance indicators. It is only when measurement issues are resolved that meaningful comparisons can be made. Availability of performance indicators is influenced by such factors as:

- Confidentiality constraints;
- Information is not collected;
- Information is not kept over time;
- Only partial information is collected and/or maintained over time;
- Differences in definitions of like concepts across jurisdictions or sectors;
- Changes in thresholds over time or across sectors;
- Jurisdictional constraints;
- Budget constraints.

Since this benchmark is in a sector that is relatively close to the port industry, being general transport, there are a lot of parallels that can be drawn with our benchmark. The main things to remember are the lessons to take away and the selection and validation of indicators. The underlying framework is also a valuable comparison point. It compares national benchmark results but is always themed around the central topics of infrastructure performance asset performance and service-level performance. For each of these categories a set of KPI's is chosen and analyzed. Examples include Empty running, Time utilization and fuel efficiency, all KPI's which, in theory, could also be tailored to ports. The final part of the document attempts to provide a uniform approach for benchmarking multiple different modes. The focus is more on labor or capital productivity with a classical input output model. Important is that other measures are also considered but not expanded upon. Once again these issues are close to the port setting Environment Safety Service quality and infrastructure.

4.2.2 Global Shipping Benchmarking Analysis PWC

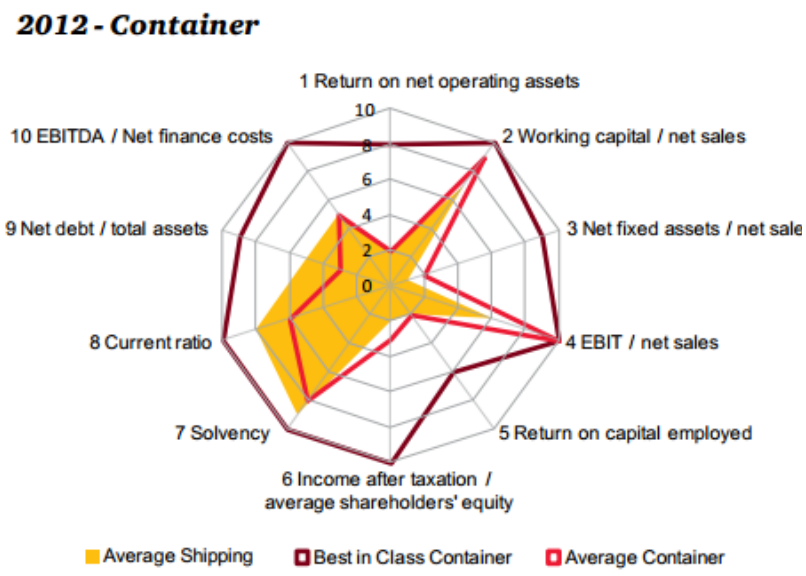
The PWC global shipping benchmarking analysis is now in its fifth year, in this report PWC provides an overview of the factors that impacted the shipping industry in the previous year and analyse how these have been reported by a large number of shipping companies from around the world. The document outlines four main topics: market developments, sustainability, financial performance and the coverage of the study. We won't go too deep into the results of the study but try to assess the added value for PORTOPIA and possible adaptions to the port sector.

The first part of the document links external market developments to the concept of shipping in order to create an overview of the situation. Indicators like global growth, the fleet order book, financial markets and supply & demand are investigated. This top down analysis allows in a later stage to analyse and explain the situation and performance of the benchmarked companies. The general document only provides an aggregate benchmark of the industry over time. Companies can ask for a tailored benchmark against payment.

Further in the document a more detailed analysis is made. Here the document introduces indicators on a more segmented level. Average Earnings, for example, are split up per cargo group and size. Further on they are linked to a forecast and analysed in combination with external data inputs like number of new buildings vs. demolitions, vessel value, shipping finance and CO2 sustainability.

The final part of the study tackles the financial and key performance indicators. Here the use of ratios is the main focus. Amongst others Working Capital/net sales, Net fixed assets/net sales are used and aggregated in order to create a complete overview. The ratios are then mapped according to these aggregates per sub sector (Average score overall shipping industry 2012, Average score subsector 2012, Best in class in subsector 2012) as can be seen in Figure 7.

Figure 7 Spider web diagram representation



Source: Global Shipping Benchmarking Analysis (PWC 2013)

4.2.3 Lessons from shipping benchmarking

- A benchmark should be tailored to suit the main user or stakeholder it is intended for in the PORTOPIA case this means ports ;
- In order to select the correct range of indicators it is important to look at previously developed indicators, in this case this would be the PPRISM project and national benchmarking exercises;
- The indicators should be selected based on a precise set of rules allowing for optimal data quality and relevance;
- The benchmark should link to external indicators and macro-economic activities explaining the results of the exercise;
- Since the tool aims at specific ports it could be an added value to select the indicators in such a way that aggregated market reports can be quickly generated;
- The use of ratios seems to be the best approach when benchmarking.

4.3 Railway benchmarking

Benchmarking of railway systems was quite hard until the 1990s, when the European Commission (EC) decided that to allow fair competition on the European railway network. After this moment a separation of essential functions in the sector was required. EC Directives such as 91/440 have far-reaching implications for the management of the railway infrastructure, where manager could mostly focus on internal analysis, now external benchmarking has become a main tool for efficiency measurement. This analysis method requires valid comparisons of performance between different companies, so attaining a reasonable level of comparability is essential. This is in line with ports since railway infrastructure companies are large, complex, and their scopes of operation differ between European countries. Also management systems differ due to funding and geopolitical aspects. A final link is that railway systems are viewed positively by citizens and policymakers around the world because of their impact on mobility, their potential to improve land use and development in urban centres.

The general problems which benchmarking faces in the rail sector are similar to those of the port sector. First of all the multiplicity of outputs is an issue (cf. passenger/km, ton/km total tonnage #trains/vessel calls etc..) A large railway as well as a port has millions of products to offer and each of those carries some relevance to specific stakeholders. In addition the lack of homogenous units of measurement renders the entire exercise even more complex.

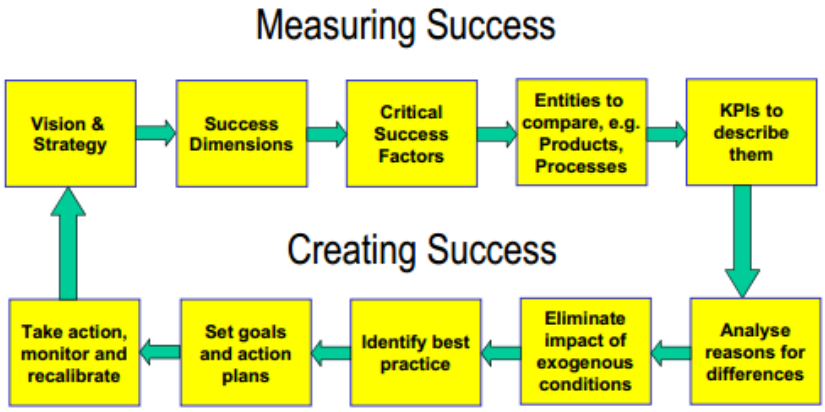
Secondly the complexity of the production process adds to the issue of successful benchmarking. All different forms of labour and capital, the asset management systems in place, government intervention must be taken into account.

Finally the operating environment and government intervention al also similar. They affect a strong influence on both rail and port performance. However geography has other influences as well? Climate, complexity of the logistics network, port depth and weather are all likely to influence costs. Government intervention strongly influences performance in this case.

4.3.1 Improverail Benchmarking concept

The improverail-benchmarking concept is an exercise which stems directly form the need for more external benchmarking as described above. The process they developed starts by defining 'Success Dimensions' which enable the railway to view success not in terms of functional output, but in terms of an integrated aim such as delivering service quality or a safe railway (IMPROVERAIL consortium, 2003). In a second stage factors are defined which are critical to success, and KPIs are developed to describe these. The methodology used can be seen in Figure 8.

Figure 8 The holistic Improverail approach



Source: Improverail Benchmarking concept

A first aspect of the analysis method is applying the KPIs to a specific part of the entity. They could be applied to the whole of the entity as a single unit, but the research proposed a process for splitting the railway into comparable units. This is a similar approach as the one used by the PORTOPIA project. The benchmarking process should continue to understand the reasons for differences in performance, identifying and eliminating external factors affecting performance upon which the railway entity's management has no control. The focus of the benchmarking exercise should be on the comparison of best practices and industry standards, which are under the control of management. As mentioned earlier the resulting KPIs should then be used once again to monitor the effect of change over a prolonged period of time.

The purposes of KPIs developed in IMPROVERAIL include the following:

- To identify priority areas for improvement within each railway;
- To identify the best performing railways in the different success dimensions and critical success factors: these will be the ones to emulate;
- To help determine how much improvement is possible (when normalized).
- To monitor progress over time: if performance has improved, what practices have led to the improvement?;
- To monitor an railways' own progress over time, in order to determine if it is progressing towards its vision and objectives;
- To estimate the effect of differences in processes and practices on actual outcomes;
- To set targets for improvement.

Three process benchmarking pilot case studies were undertaken during the course of the research. In general, the case studies found that good information systems and easy information availability were closely correlated with good punctuality performance.

Keys to successful railway Infraco benchmarking, which can be partially transposed to the port industry, are amongst others that the stakeholders must drive the benchmarking. Also a strong confidentiality agreement is essential to protect commercially is an absolute necessity. The first version of the exercise should be done with a core group of involved target stakeholders. Top management support is a necessity and senior managers should be well involved in the project. While concise questionnaires are important in collecting data, they should be followed up with face-to-face interviews with the relevant managers where appropriate. Understanding and treatment of the context and the environment of the investigated setting is of primordial importance.

4.3.2 The International Transport Forum: 'Railway Efficiency'

The International Transport Forum at the OECD is an intergovernmental organisation with 54 member countries. It acts as a strategic think-tank, with the objective of helping to shape the transport policy agenda on a global level and ensuring that it contributes to economic growth, environmental protection, social inclusion and the preservation of human life and well-being. This short summary is aimed at a document that was developed around Railway efficiency. Which is a topic of interest worldwide for railway managers operating in competitive markets and for fiscally strained governments.

Their methodology firstly analyses external factors influencing efficiency, then it identifies the key cost and revenue drivers and finally tries to describe an efficient railway system. As for the benchmarking aspect, before delving more into the framework they developed, we fist have to highlight an important lesson namely that this methodology is only useful as long as the diverse "raisons d'être" of national railway systems are acknowledged in the analysis (OECD, 2010).

With regards to benchmarking firstly several important points to consider when utilizing benchmarking. Here we find: differences in the goals and roles, differences in network and operations characteristics, differences in railway system structures, differences in railway accounting standards and debt and differences in the corporate status. Most of these differences are also present when analysing seaports.

Following the benchmarking outline a gap analysis is done with different indicators as a base like track and train utilization and staff requirements. These results are finally discussed and analysed and opportunities and possible solutions are offered for improved efficiency.

4.3.3 Lessons from Railway benchmarking

- In order to make valid comparisons, a large single entity should be split up into comparable parts, rather than the whole company;
- The indicators are divided into a set of KPIs that, which, in a perfect scenario, are perfectly comparable and a separate set where not everything is comparable;
- Both internal and external benchmarking techniques have their merit in order to create the complete view of the situation;
- A major issue is the use of categories like 'other' in the source data where discrepancies can exist between different users/data providers of the system.

5 OPTIMAL BENCHMARKING FOR PORTS

5.1 Selection of indicators

When the performance of a sector or entity is being compared to its peers, one of the first tasks is to reach consensus on the choice of indicators. Depending on the target group and the required measurements these indicators can be of an extremely wide variety. As a rule, economists will choose indicators of efficiency/productivity, unit costs or price changes. Financial analysts will use financial returns as we have seen in the general description of benchmarking methodologies. Others will prefer safety or environmental performance for a more governmental perspective. All of these indicators are valid in their own way, and depend on the required investigation. The choice of potential indicators in the port, and by extension, shipping sector can be overwhelming. The first order of business is to ensure that analysts are clear on the objectives of the benchmarking exercise. Second, they must ensure that the indicators are available and are reliable.

From the previous part we have learned that the indicators must uphold a certain set of rules. These rules will be further extended upon in the following part of this section. We base ourselves on part of the literature on data availability and completeness and the work done in PPRISM. For a full list of indicators as of 3/12/2014 see annex.

5.1.1 Rules for data availability and completeness

Missing data and incomplete data takes multiple different forms. As described by Little & Rubin, (2002) there are three main types of missing data. The decision of which type the missing data is a pure judgment call.

- Missing completely at random (do not depend on the variable of interest);
- Missing at random (do not depend on the variable of interest, but are conditional on other variables in the data set);
- Not missing at random (Missing values depend on the values themselves).

In order to deal with this missing data there are three main methods:

- Case deletion;
- Single imputation;
- Multiple imputation.

The first option is simply omission of missing data in the benchmarking analysis but this approach leads to problems when assessing data over time. The second and third option is the imputation of data based on a wide variety of methods. Although this approach seems the most logical there are some drawbacks. In the words of Dempster & Rubin (1983).

The idea of imputation is both seductive and dangerous. It is seductive because it can lull the user into the pleasurable state of believing that the data are complete after all, and it is dangerous because it lumps together situations where the problem is sufficiently minor that it can be legitimately handled in this way and situations where standard estimators applied to real and imputed data have substantial bias.

Currently the project is using an imputation method based on a basic interpolation mechanism (mean/median/mode substitution). A new method is being developed and will be described in the forecasting deliverable. Other possibilities for data imputation

include: regression imputation, hot-and cold-deck imputation, expectation-maximisation imputation, or multiple imputation, e.g. Markov Chain Monte Carlo algorithm (OECD 2008).

The use of imputation models affects the quality of the data. More on this will be provided in the next section of this document.

The selection of data for the benchmarking model should adhere to the highest standards used in the industry. following rules and dimensions are created based on work done by the IMF, Eurostat and OECD:

Relevance

The relevance of data is a measure of the added value of the information. ”Added value” is defined as the degree to which the underlying statistics meet current and potential needs of the users. It depends upon both the coverage of the required topics and the use of appropriate concepts.

In the context of a benchmarking exercise, relevance has to be evaluated with the stakeholder as ultimate user of the system. Careful evaluation and selection of basic data has to be undergone in order to assess the relevance for the user of the benchmark engine.

Accuracy

The accuracy of basic data is the degree to which they correctly estimate or describe the quantity or characteristics that they are designed to measure. Accuracy refers to the closeness between the values provided and the (unknown) true values, discrepancies can occur in reporting methods or incomplete datasets which use data inclusion models. In practice, these attributes are typically measured or described in terms of the error, or the potential significance of error, introduced through individual major sources of error. Within the Portopia project we attribute different ‘levels of trustworthiness’ to different data sets (se deliverable 9.2).

In the case of the PORTOPIA dataset, the major sources of error include coverage, input by user, non-response, faulty response and processing errors. The current IT system tries to deal with these issues but continuous data completeness and accuracy analysis is required.

Timeliness

The timeliness of data products reflects the length of time between their availability and the event or phenomenon they describe, but considered in the context of the time period that permits the information to be of value and to be acted upon. It is basically a measure of punctuality, which implies the existence of a publication schedule and reflects the degree to which data are released in accordance with it.

In the context of benchmarking, timeliness is especially important to minimise the need for the estimation of missing data or for revisions of previously published data. As individual basic data sources (RES, ECOPORTS, CEF, FF) establish their optimal trade-off between accuracy and timeliness, taking into account institutional, organisational and resource constraints, data covering different domains are often released at different points of time. Therefore special attention must be paid to the overall coherence of the use of data over time when benchmarking ports.

Accessibility

Accessibility relates to the degree in which the user (creator of the system, in this case the consortium) has access to the data and metadata used for construction of the benchmarking engine. It can affect the cost of creation and accuracy and timeliness parameters. It does not involve the users of the system themselves, however it could affect thrust levels of users if the results of the benchmarking exercise can not be reproduced outside the benchmarking environment.

In the context of benchmarking, accessibility of basic data can affect the overall cost of production and updating of the indicator over time. As mentioned before the credibility of the benchmark results is an issue if data accessibility is too poor. Therefore, the selection of the source should not always give preference to the most accessible source, but should also take other quality dimensions into account.

Interpretability

The interpretability of data products reflects the ease with which the user (consortium) may understand and properly use and analyse the data. The adequacy of the definitions of concepts, target populations, variables and terminology underlying the data and of the information describing the limitations of the data, if any, largely determines the degree of interpretability. The use of the group 'others' is a good example of bad interpretability.

In the context of benchmarking, the availability of definitions and classifications used to produce basic data is essential to assess the comparability of data over time and across ports. This document also provides a first outline for segmentation and definitions in order to increase interpretability of the data (see 5.2).

Coherence

The coherence of data products reflects the degree to which they are logically connected and mutually consistent, i.e. the adequacy of the data to be reliably combined in different ways and for various uses. Coherence implies that the same term should not be used without explanation for different concepts or data items; that different terms should not be used for the same concept or data item without explanation; and that variations in methodology that might affect data values should not be made without explanation.

In the context of benchmarks, two aspects of coherence are especially important: coherence over time and across ports. Coherence over time implies that the data are based on common concepts, definitions and methodology over time, or that any differences are explained and can be allowed for. Incoherence over time refers to breaks in a series resulting from changes in concepts, definitions, or methodology. Coherence across countries implies that from port to port the data is based on common concepts, definitions, classifications and methodology, or that any differences are explained and can be allowed for. The data availability deliverable in WP 1 deals with this issue.

5.2 Benchmark framework

5.2.1 Rules for defining the benchmarked entity (ports)

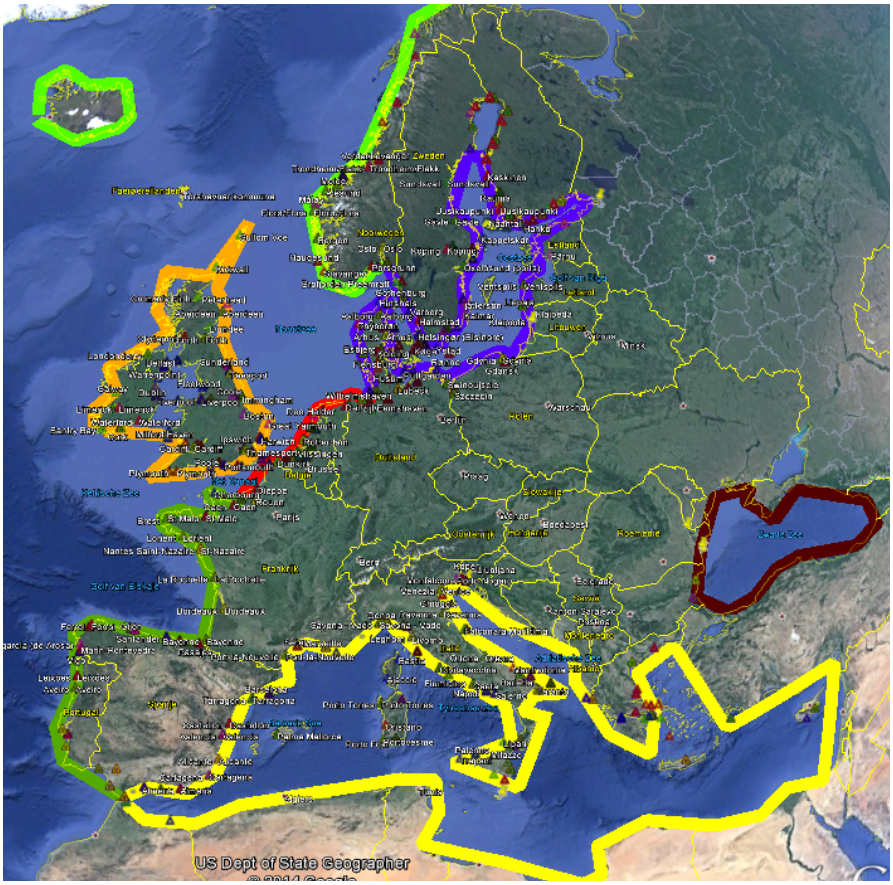
One of the main lessons we can distil from the benchmarking in alternate sectors is the need for segmentation and clear definitions of data and metadata. This section tries to

outline some basic rules for the metadata concepts. However this is not the full list of beta data benchmarking rules that will be provided in deliverable [9.3]

Need for analysis in relevant regions

The need or geographical segmentation is quite logical. However, in the port industry a classical 'radius' approach is not sufficient for this type of breakdown. In order to provide a relevant framework for benchmarking competing entities must be grouped together. For ports within Europe these competitive groups are labelled as ranges, as illustrated in Figure 9. During the last workshop these ranges have been specified and introduced within the system, this is a first version prone to change after peer assessment by stakeholders. The following distribution has been developed:

Figure 9 Ranges



- Atlantic range (Green)
  - Norway + Iceland +South of Le Havre – Gibraltar
- Baltic range (Blue)
  - North of Hamburg - border of Norway
- UK range (Orange)
  - UK
- Hamburg-Le Havre range (Red)
  - South of Hamburg – North of Le Havre
- Mediterranean range (Yellow)
  - East of Gibraltar Mediterranean area
- Black sea (Brown)
  - Black sea region

Need for analysis in relevant types

When comparing ports it is necessary to segment ports into specific types. There is such a large variety of ports that in order to do a successful port-aggregate benchmarking exercise the aggregate must be correctly defined and compiled. The different types of ports are numerous. Table 2 lists a non-exhaustive list of port specializations and types and the optimal way of determination.

Table 2 Different types of ports

Port specialization	
By cargo type	Statistical analysis RES
By main activity	Statistical analysis RES
By land use	Fact finding, statistical analysis
Port infrastructure	
Location on river	Fact finding
Use of equipment	Fact finding
Port management	
Ownership type	Fact finding
Land use system	Fact finding

Need for analysis in relevant sizes

A quite difficult issue is the classification of ports based on size. For one the timeliness and accuracy is a problem since ports can shift size groups over time rendering benchmarking impossible. For now the size classification is based on the fact finding figures.

- Small: *<100 000 tonnes*
- Medium: *100 000 – 500 000 tonnes*
- Large: *>500 000 tonnes*

In a next deliverable and post workshops with stakeholders, we will propose that over time a more complex system is introduced based on relative size in the geographical range with a moving average method allowing for a more consistent and long term classification method.

Most of these metadata concepts have to be included in the fact finding reports of the PORTOPIA project.

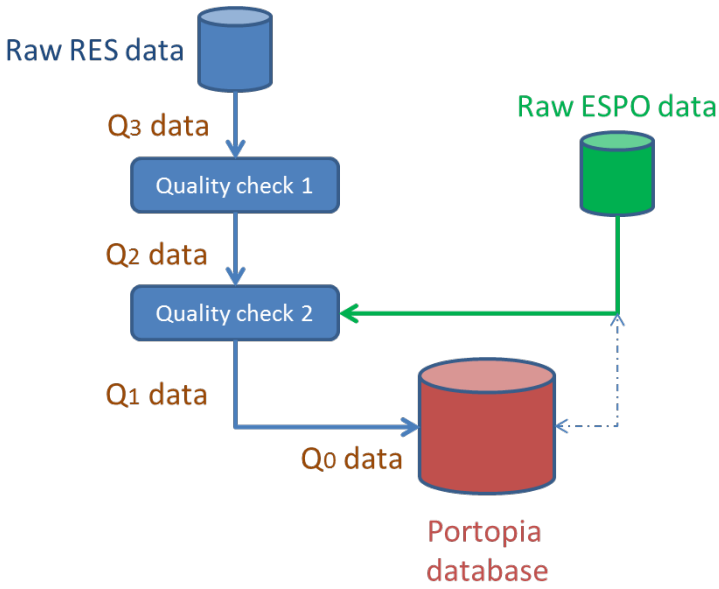
5.2.2 Making viable comparisons

Different levels of data quality

During the last workshop with the IT partner of the project the issue of data quality was already discussed. The proposed solution was working with different quality assurance system. In this stage of the project data is mostly obtained through the RES system. There are some issues specific to RES undermining the data quality. Missing inputs, non-matching inputs, wrong inputs, etc. the first quality check should deal with the majority of these issues. Currently there is an algorithm in development for complex interpolation and backwards forecasting. We will deal with this algorithm in the forecasting and upcoming benchmarking deliverables. After quality check 1, all

data should be complete and fairly accurate, this level of data will be noted as Q2 data (as illustrated in Figure 10*).*

Figure 10 Data quality check



A second quality check was then proposed linking the Q2 data to the ESPO database which contains similar data points of most ports. During quality check 2 these data points will be compared in order to assure consistency and create a double check. This type of data Q1 is the main data with which we work. Finally a continuous review system was proposed in order to let stakeholder check historical data and assure the highest possible quality being Q0.

Long term bucketing benchmarking

A major issue with the bucketing and segmenting of data aspects is that while attempting long term benchmarking ports can shift from one bucket to the other. A good example uses port size where a port can switch between a small and medium size port on a yearly basis. This renders long-term analysis and comparisons of groups and individuals extremely complex. We will therefore propose a benchmarking engine which deals with this issue.

Assume the following set of ports P which we would like to benchmark at two specific points in time (for example 1980 vs 2013):

$$P = \{p_1, p_2, \dots, p_{10}\}$$

P contains ports which belong to a certain bucket (for e.g. small ports) EITHER at T1 of benchmarking or T2 of benchmarking. Within this set there are 3 different subsets worth comparing, namely ports which are part of the bucket BOTH at T1 and T2, ports which are ONLY part of the bucket in T1 and finally only part of the bucket in T2.

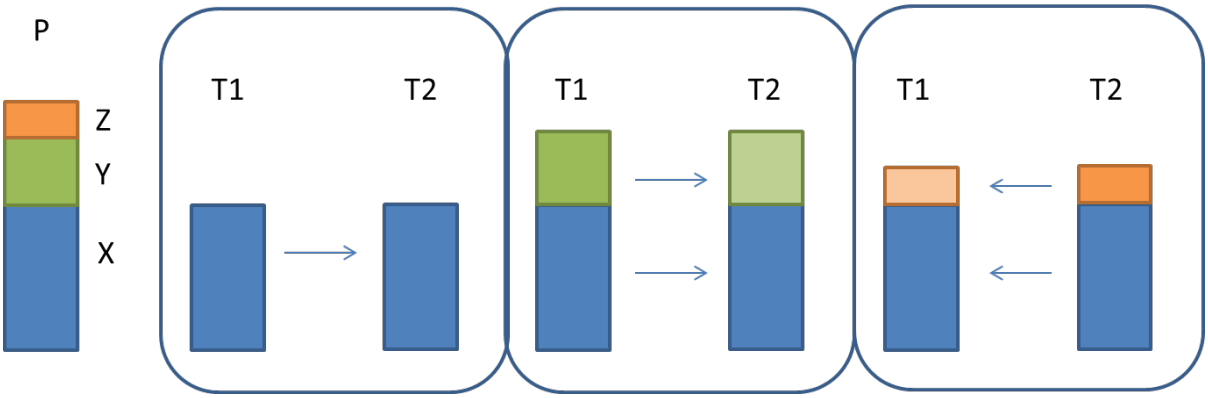
$$X = \{p_1, p_3, , p_7\} \quad \Leftrightarrow \quad X \in T_1 \,\&\, T_2$$

$$Y = \{p_2, p_4, , p_5\} \quad \Leftrightarrow \quad X \in T_1 \,\&\, \overline{T_2}$$

$$Z = \{p_6, p_8, p_9, p_{10}\} \quad \Leftrightarrow \quad X \in \overline{T_1} \,\&\, T_2$$

This basically allows for three different types of comparisons (see Figure 11). The first one being only the data points which stay in the bucket at both times. The second one both the X group and the group of data points which shifts in a later stage and the final one being the data points which are present in the final stage but not the initial phase.

Figure 11 Comparison of data buckets



Internal, external, system analysis

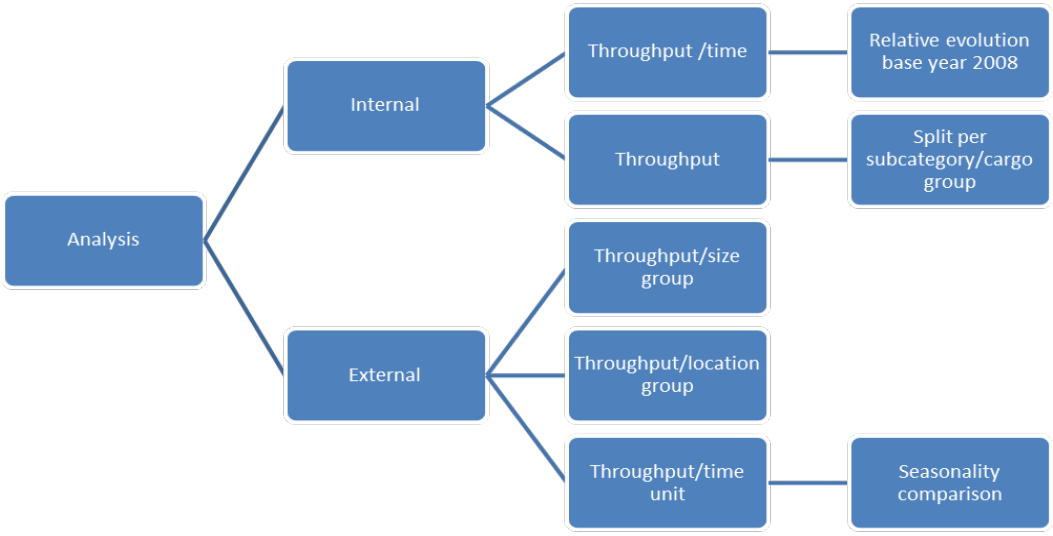
The benchmark has to be done on different levels in order to guarantee optimal relevance for the user. Figure 12 and Figure 13 show the input of the first dashboard for the IT tool. This is not a pure benchmarking exercise using the benchmarking engine which is still under development but more a semi benchmark overview screen.

Figure 12 The use of internal and external benchmarking for the dashboard

		Parameters										Representation form			
		Throughput						Location			Time				
		BB	DB	LB	RR	CT	All	EU	Range	km²		Timeline	Pie Chart	Integer	String
Internal	Total Throughput evolution baseyear 2008	X	X	X	X	X	X				X	X			
	Split of cargo sub-categories %	X	X	X	X	X	X						X		
	Contenerisation rate	X				X					X	X	X		
	Share of Major Bulks in data		X	X							X	X	X		
	General indicators													X	X
External	Share of port throughput	X	X	X	X	X	X	X	X	X	X		X		
	Evolution of traffic	X	X	X	X	X	X	X	X	X	X	X			
	Evolution of growth	X	X	X	X	X	X	X	X	X	X	X			
	# of ports that have grown relative on size	X	X	X	X	X	X				X		X		
	Seasonal comparison within portset	X	X	X	X	X	X	X	X	X	X			X	X
	Industry volatitility	X	X	X	X	X	X	X	X		X			X	X
RES	% of ports included in data														
	% of ports included in RES														

Standard information
Customizable option
Analysis option

Figure 13 Overview of internal and external benchmarking with indicators



We already proposed to make the split between internal and external analysis on dashboard level so we will do the same for the benchmarking engine allowing users to benchmark certain variables within their own port or compared to a set of other ports.

5.2.3 Weighing and aggregation methods

When used in a benchmarking framework, weights can have a significant effect on the overall results of the exercise, be it a composite indicator or a 1-on-peergroup analysis. A number of weighting techniques exist (Table 3), these were originally summarized by the OECD (2008) in their Handbook on Constructing Composite Indicators. Some are derived from statistical models, which we already discussed, such as SFA, DEA and UCM. Each of these methods has its pros and cons but regardless of which method is used, the attributed weights are essentially value judgments. One of the major divisions in the field is the selection of weights based on pure statistical methods whilst others prefer more subjective methods with relative weights based on, for e.g., expert opinion.

Table 3 Compatibility between aggregation and weighting methods

Weighting methods	Aggregation methods		
	Linear[4]	Geometric[4]	Multi-criteria
EW	Yes	Yes	Yes
PCA/FA	Yes	Yes	Yes
BOD	Yes[1]	No[2]	No[2]
UCM	Yes	No[2]	No[2]
BAP	Yes	Yes	Yes
AHP	Yes	Yes	No[3]
CA	Yes	Yes	No[3]

. Normalized with the Min-Max method.

. BOD requires additive aggregation, similar arguments apply to UCM.

. At least with the multi-criteria methods requiring weights as importance coefficients.

. With both linear and geometric aggregations weights are trade-offs and not "importance" coefficients.

Source: OECD (2008)

Amongst the wide variety of weighing techniques, equal weighting (EW) is one of the most commonly used due to its simplicity and visibility. Basically all variables are given the same weight, which implies that all variables are of the same importance. However, as stated by the OECD, it could also disguise the absence of a statistical or an empirical basis, e.g. when there is insufficient knowledge of causal relationships or a lack of consensus on the alternative. In any case, equal weighting does not mean "no weights", but implicitly implies that the weights are equal.

For a port setting statistical weight attributions can be useful. However participatory methods that incorporate various stakeholders may be more in line with the end goals of the project. Experts, ports and various other entities – can be used to assign weights in eventual benchmarking exercises. No matter which method is chosen one important aspect remains namely the time element, keeping weights unchanged across time is a necessity if the researcher is attempting to analyse the evolution of a certain number of variables, indicators or ratios.

6 CONCLUSIONS

Seaport functions are diverse in scope and nature and have evolved over time. During the past decades port entities have undergone consistent change on the macro- and micro economical level. The roles and functions each port undertakes can be identified through political geographical (urban and spatial), economic and social perspectives. The diversity of the seaports lies at the basis of the necessity of a complex benchmarking engine supporting the PORTOPIA programme. Other factors both macro- and micro-economical add to this urgency on a continuous level. Amongst these reasons we find growing complexity in business operations, the pressure for other parameter follow up like sustainability and resilience and the increased privatization. All this is supported by a growing academic interest that pushes the boundaries of benchmarking possibilities ever further.

There is an entire set of benchmarking methodologies currently in use by both academics and professionals alike. Benchmarking itself can take place in 2 environments, namely internal and external. The benchmarking process itself is divided in 8 phases. The first one being scope definition, followed by the selection of partners, choice of measurement methods, data collection, analysis of discrepancies, presentation or results, improvement of current situation and monitoring of progress and ongoing benchmarks. The tools used vary between the users. Within the academic sector benchmarking is often done with specific tools like DEA and SFA analysis. Also The industry uses a wide range of methods and tools in order to manage the benchmarking process. We discussed GAP, SWOT and ratios as possible tools for effective benchmarking. For the Portopia project the use of ratios seems most appropriate for benchmarking. The results could eventually be shown as a gap analysis but this is only in the final phase of the project.

In order to assess what has been done in similar sectors, we investigated the airline, railroad and shipping industry. For each industry two benchmarking studies were analysed and there important lessons distilled. Within the airline sector the focus remains mostly costs and CAPEX. Amongst the most important lessons we found the use of aggregation and segmentation of investigated entities. Within shipping benchmarking the tailoring of the exercise to stakeholders was the main takeaway together with the use of ratios as preferred mechanism. Also it was noted that many measurement issues are primarily a consequence of the availability and reliability of appropriate performance indicators. Finally the rail benchmarking outlined that the importance of benchmarking is to help determine how much improvement is possible and to create a tool which allows you to monitor the investigated entities over time. The main lesson from railway analysis remained that in order to make valid comparisons; a large single entity should be split up into comparable parts, rather than the whole company both on an internal and external benchmarking level.

The final part of the deliverable deals with the optimal benchmarking for ports. When the performance of a sector or entity is being compared to its peers, one of the first tasks is to reach consensus on the choice of indicators. These indicators must uphold a certain set of rules. Missing data and incomplete data, currently the project is using an imputation method based on a basic interpolation mechanism (mean/median/mode substitution). A new method is being developed and will be described in the forecasting deliverable. The selection of data for the benchmarking model should adhere to the highest standards used in the industry. Following rules and dimensions are created based on work done by the IMF, Eurostat and OECD: t*imeliness, accuracy, relevance, accessibility, interpretability ,coherence.*

Within the benchmarking engine the necessity for segmentation exists, which in its turn renders the exercise far more complex. There is a need for analysis in relevant regions, we dubbed these ranges based on the industry standard. This same bucketing is also required for size and type of ports. In order to make viable comparisons between these buckets we in traduced different levels of data quality based on multiple tests and database analyses. Also a first attempt to introduce the long term bucketing engine has been included using three different types of classifications for entities allowing comparisons throughout time. In order to keep all this organized we already proposed to make the split between internal and external analysis on dashboard level so we will do the same for the benchmarking engine allowing users to benchmark certain variables within their own port or compared to a set of other ports.

7 REFERENCES

Babovic, R. C., 2012. Benchmarking as a function of competitiveness and efficiency in business. Economics of agriculture, Volume 1.
Benchmarking: Intermodal Benchmarking report OECD 2013 Paris
Van den Bosch., 2011. De strategische waarde van het Haven- en industriecomplex Rotterdam voor het internationale concurrentievermogen van Nederland. Rotterdam: Erasmus universiteit.
CAA airport operating expenditure benchmarking report 2012 , civil aviation authority 2013 London Regulatory policy group
Cook, W.D., Tone, K., and Zhu, J., Data envelopment analysis: Prior to choosing a model, OMEGA, 2014, Vol. 44, 1-4.
Dattakumar, R. & Jagadeesh, R., 2010. A review of literature on benchmarking. Mysore: Department of mechanical engineering.
Deiss & R, 2000. Transport Benchmarking Methodologies, applications and Data needs. paris: s.n.
Dempster A.P. and Rubin D.B. (1983) Introduction pp.3-10, in Incomplete Data in Sample Surveys (vol. 2): Theory and Bibliography (W.G. Madow, I. Olkin and D.B. Rubin eds.) New York: Academic Press.
Evan, M., 2008. Action oriented benchmarking: concepts and tools, s.l.: Lawrence Berkeley National Laboratory.
FINTELOnline Benchmarking Report on 01/07/2014. Fintel.us
Forrester. 2003. “Making Dashboards Actionable”.
Gap analysis: concepts, methods, and recent results* MD Jennings - Landscape ecology, 2000 - Springer
OECD, 2008. Handbook on Constructing Composite Indicators. Paris: s.n.
OECD, 2000. Transport Benchmarking Methodologies, Applications & Data Needs. Paris: s.n.
OECD, 2014. Report of the International Transport Forum. Paris: s.n.
Humphrey, Albert (December 2005). "SWOT Analysis for Management Consulting". SRI Alumni Newsletter (SRI International).
IMPROVEd tools for RAILway capacity and access management 2003 COMPETITIVE AND SUSTAINABLE GROWTH (GROWTH) PROGRAMME
Jaiswal, S., 2012. A case study on quality function deployment. IOSR Journal of Mehanial and Civil Engineering, 3(6), pp. 24-35.
Key Findings of 2012 ATRS Global Airport Performance Benchmarking project Tae Hoon Oum, Yapyin Choo, Chunyan Yu, 2012
Liebert, V. P., 2011. Airport Benchmarking: An Efficiency Analysis of European Airports from an economic and managerial perspective, Bremen: Jacobs Universtiy .
Little, J., & Rubin, D. (2002). Statistical analysis with missing data (Second ed.). New York: Wiley.
Navigating in stormy waters Special feature: Sustainability reporting in Shipping, 2012 PricewaterhouseCoopers B.V.
Notteboom, T. & Vonk, I., 2011. Economic analysis of the Thine-Scheldt delta port region. Antwerp: University of Antwerp-ING.
Rommerskirchen, S., 2000. Quality of European Transport statistis- requirements results and prospects. Paris: European conference of ministers of transport (ed) in: Transport Benchmarking methodologies appliations and data needs.
Tofallis, C. (2001) "Combining two approaches to efficiency assessment." Journal of the Operational Research Society 52 (11), 1225–1231.

8 ANNEX

Non-exhaustive list of benchmarking techniques:

- Matrix technology
- Comparison tables
- Graphs: Pie chart, Bar chart / Histogram
- SWOT analysis
- Potential/resources-analysis
- Price/performance ratio
- Potential analysis
- Life cycle analysis
- market growth/market share portfolio
- market attractiveness/competitive strength portfolio
- Portfolio attractiveness customer/supplier position
- Technology/resource strength
- Market position/technology position portfolio
- Contribution margin/cost development portfolio
- Price/customer satisfaction portfolio
- Revenue share/revenue portfolio
- Spider web diagram

Asset Efficiency Financial Ratio Analysis

Collection Period

	Industry	Company
Bottom 10%	71.81	
25%	51.16	
Median	**43.19**	**51.17**
25%	35.54	
Top 10%	5.24	

Bottom ← → Top

10% 25% Median 25% 10%

Your company is waiting longer to collect payment after sales have been made than the average firm in your industry. You should determine if you can collect faster through better invoicing and collections as you may be tying up cash needlessly in your accounts receivable balance.

Inventory Turnover

	Industry	Company
Bottom 10%	8.78	
25%	10.40	
Median	**17.81**	**4.94**
25%	22.57	
Top 10%	44.54	

Bottom ← → Top

10% 25% Median 25% 10%

Your company is generating significantly fewer sales dollars on your investment in inventory than the average firm in your industry. You should determine if steps can be taken to decrease your company's inventory as you may have excessive cash tied up in your inventory investment.

Assets to Sales

	Industry	Company
Bottom 10%	0.75	
25%	0.56	
Median	**0.43**	**0.55**
25%	0.30	
Top 10%	0.21	

Bottom ← → Top

10% 25% Median 25% 10%

Your company is generating about as many sales dollars on your investment in total assets as the average firm in your industry. You should try to determine if even additional sales can be generated on your asset investment.

Sales to Net Working Capital

	Industry	Company
Bottom 10%	3.64	
25%	5.16	
Median	**11.10**	**2.27**
25%	21.09	
Top 10%	120.73	

Bottom ← → Top

10% 25% Median 25% 10%

Your company is generating considerably fewer sales dollars on the working capital in your business than the average firm in your industry. You should carefully review your capital structure (debt to equity proportion) and consider ways to use your investment in working capital more efficiently.

Accounts Payable to Sales

	Industry	Company
Bottom 10%	11.53 %	
25%	8.21 %	
Median	**6.28 %**	**6.57 %**
25%	2.91 %	
Top 10%	0.81 %	

Bottom ← → Top

10% 25% Median 25% 10%

Your company has close to the average accounts payable as a proportion of sales within your industry. You should try to determine if accounts payable can provide your company with more cash liquidity.

1. Market Trends & Structure Indicators	Included in PPRISM	Active in portopia
Maritime traffic	Yes	Yes
Hinterland Traffic	No	No
Vehicle Traffic	No	No
HHI	Yes	No
n-port Concentration Ratio Index	No	No
Location quotient (LQ)	No	No
Entropy Index (H)	No	No
Gini Coefficient (G)	No	No
Vessel Traffic	Yes	Yes
Market Share	Yes	Yes
Commodity Diversity Index (CDI)	No	No
Commodity Specialisation Index (CSI)	No	No
Traffic Balance (TB)	No	No
Terminal Market Share [S(T)]	No	No
Terminal Operator Market Share [S(TO)]	No	No
t-terminal Concentration Ratio ()	No	No
Herfindahl-Hirschman Index (HHI)	No	No
Baltic Exchange Indexes	No	No
Fleet statistics	No	No
Material cost Index	No	No
Car production Index	No	No
Steel-vessel index	No	No
Load Rate	Yes	No
Gateway Degree	No	No
Container dependency	Yes	Yes
Liner Shipping Connectivity Index (LSCI)	No	No
Port Connectivity Index	No	No
Call size	Yes	Yes
Modal Split	Yes	Yes

2. Socio-economic Indicators	Included in PPRISM	Active in portopia
Employment (Direct & Indirect)	Yes	No
Direct Employment per tonne	No	No
Direct Employment per hectare	No	No
Added value (Direct & Indirect)	Yes	No
Direct Gross Added Value per tonne	No	No
Direct Gross Added value per hectare	No	No
Direct Gross added value per FTE	Yes	No
Financial health	Yes	No
Labour Cost	No	No
Labour Turnover	No	No
Training per FTE	Yes	No
Gender	No	No
Flowback to Treasury	No	No
Investment	Yes	No
Trade Flows	No	No

4. Logistic Chain and Operational performance Indica	Included in PPRISM	Active in portopia
Maritime connectivity	Yes	No
Maritime Ro-Ro connectivity	No	No
Intermodal connectivity	Yes	No
On-time performance (Sea-going)	Yes	No
On-time performance (Inland waterways, Rail, Road)	Yes	No
Mean-time customs clearance	Yes	No
Availability of Port Community Systems	Yes	No
Ship turnaround time	Yes	No
Quality of port community system	No	No

3. Environmental Indicators	Included in PPRISM	Active in portopia
Total energy consumed	Yes	No
Ratio of renewable energy per total energy consume	No	No
Carbon footprint	Yes	No
Total Greenhouse Gas Emissions	No	No
NOx, SOx, PM10, VOCs CO, O	No	No
Lden (overall day-evening-night noise level)	No	No
Lnight (23:00 - 7:00hrs noise level)	No	No
Total water consumption	Yes	No
Amount of waste	Yes	No
Oxygenation conditions	No	No
Salinity	No	No
Nutrient condition	No	No
EMS standard	Yes	No
Turbidity / Transparency	No	No
Existence of Aspects inventory	Yes	No
Self Diagnostics results (SDM)	No	No
Compliance	No	No
Number of prosecutions for non-compliance	No	No
Existence of monitoring programme	Yes	No

5. Governance Indicators	Included in PPRISM	Active in portopia
Integration port cluster	Yes	No
Extent of performance management	Yes	No
Existence of Performance Measurement	Yes	No
Corporate Responsibility	No	No
Formal reporting CSR	Yes	No
Market openness	Yes	No
Accountability and Transparency	No	No
R&D	No	No
Port authority investment	Yes	No
PAs relations with other seaports/dry ports/inland p	No	No
Information systems indicators	No	No
Customs related indicators	No	No
Safety/Security	Yes	No
Port authority employee productivity	Yes	No
Autonomous management	Yes	No
Ownership status of the PA	Yes	No
Legal Status of the PA	No	No
Operating and non operating functions of the PA	No	No
Strategic Objectives	No	No
Jurisdictions of the PA	No	No
Responsibility for appointing the most senior manag	No	No
Port Planning	No	No