This project has received funding from the European's Union Horizon 2020 research innovation program under grant agreement No. 769355

PEI Definition and Algorithms v1

Deliverable No.	D5.2	**Due Date**	25-OCT-2019
Type	Report	**Dissemination Level**	Public
Version	1.0	**Status**	Final
Description	This deliverable describes the methodology for PEI computation including the eKPIs to be used, the statistical algorithms as well as a description of the data sources which will be used.		
Work Package	WP5		

,

Authors

Name	Partner	e-mail
Carlos E. Palau	P01 - UPV	cpalau@dcom.upv.es
Ignacio Lacalle	P01 – UPV	iglaub@upv.es
Benjamin Molina	P01 – UPV	benmomo@upvnet.upv.es
Andreu Belsa Pellicer	P01 – UPV	anbelpel@upv.es
Charles Garnier	P05 – CATIE	c.garnier@catie.fr
Erwan Simon	P05 – CATIE	e.simon@catie.fr
Olivier Le Brun	P07 – CREO	lebrun@creocean.fr
Deborah Mille	P07 – CREO	mille@creocean.fr
Luka Traven	P08 – MEDRI	luka.traven@medri.uniri.hr
Igor Kegalj	P08 – MEDRI	igor.kegalj@medri.uniri.hr
Stjepan Piličić	P08 – MEDRI	stjepan.pilicic@medri.uniri.hr
Teodora Milošević	P08 – MEDRI	teodora.milosevic@medri.uniri.hr
Leonidas Pitsikas	P14 – IPEOPLE	lpitsikas@peoplegroup.gr

History

Date	Version	Change
06-Aug-2019	0.1	ToC and task assignments
11-Oct-2019	0.2	Deliver to Internal Reviewers
25-Oct- 2019	0.3	Submission to PC
30-Oct-2019	1.0	Submission to Participant Portal

Key Data

Keywords	Port Environmental Index (PEI), environmental impact, statistical algorithms, environmental Key Performance Indicators (eKPIs), data sources, Internet of Things (IoT)
Lead Editor	Luka Traven, P08 MEDRI
Internal Reviewer(s)	Ignacio Lacalle, P01-UPV Leonidas Pitsikas, P14-IPEOPLE

Abstract

The scope of this deliverable is to provide an overall description of how the PEI will be developed. The following sections are included in the document: (a) the baseline logic behind PEI and the usage of composite indicators including the advantages and drawbacks of their use, (b) a statistical toolbox to be tested in order to decide on the most robust approach for PEI calculation, (c) a list of eKPIs which will be used for PEI calculation including the methodological approaches of obtaining those data, whether through proxy or direct measurements, (d) a discussion on the data sources to be used including a procedure for addressing their reliability and quantifying their compliance with the IoT paradigm, (e) PEI visualization approaches in the PIXEL dashboard and, (f) links to previous and subsequent WPs for clarification on how the PEI fits in the overall PIXEL infrastructure.

Statement of originality

This document contains material, which is the copyright of certain PIXEL consortium parties, and may not be reproduced or copied without permission. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgment of previously published material and of the work of others has been made through appropriate citation, quotation or both.

The information contained in this document is the proprietary confidential information of the PIXEL consortium (including the Commission Services) and may not be disclosed except in accordance with the consortium agreement.

The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the project consortium as a whole nor a certain party of the consortium warrant that the information contained in this document is capable of use, nor that use of the information is free from risk and accepts no liability for loss or damage suffered by any person using this information.

The information in this document is subject to change without notice.

The content of this report reflects only the authors’ view. The Innovation and Networks Executive Agency (INEA) is not responsible for any use that may be made of the information it contains.

Table of contents

List of tables

List of figures

List of acronyms

Acronym	Explanation
AHP	Analytic hierarchy process
BAL	Budget allocation
BOD	Benefit of the doubt
BWM	Ballast water management
CH_4	Methane
CI	Composite indicators
CO	Carbon monoxide
CO_2	Carbon dioxide
COIN	Composite Indicators and Scoreboards
DAL	Data Acquisition Layer
DEA	Data envelopment analysis
EF	Evaluation Features
EIAP	Environmental Impact Assessments of Port
eKPIs	Environmental Key Performance Indicators
FA	Factor analysis
FAL	Facilitation of International Maritime Traffic
GHG	Greenhouse gas
GPMB	Grand Port Maritime de Bordeaux - Port of Bordeaux
GWP	Global warming potential
ICT	Information and Communications Technology
IH	Information Hub
IoT	Internet of Things
IT	Information Technology
KPIs	Key Performance Indicators
L_{DEN}	Overall day-evening-night noise level
LNG	Liquefied natural gas
L_{night}	23:00 - 7:00hrs noise level
MAR	Missing at random
MARPOL	Maritime Pollution
MCAR	Missing completely at random
MCDM	Multi-criteria decision making
MVP	Minimum Valuable Product
N_2O	Dinitrogen oxide

NGSI	Next-Generation Sensors Initiative
NMAR	Not missing at random
NMVOCs	Non-methane volatile organic compounds
NO_x	Nitrogen oxides
OMI	Ocean Monitoring Indicators
OT	Operational Tools
PAS	Port Activity Scenario
PCA	Principal components analysis
PEI	Port Environmental Index
PIXEL	Port IoT for Environmental Leverage
PM	Particulate matter
QoE	Quality of Experience
RR	Reliability Rating
RTG	Rubber-tired gantry
SA	Sensitivity analysis
SaaS	Software as a Service
SO_2	Sulphur Dioxide
SO_x	Sulfur oxides
TEU	Twenty-foot Equivalent Unit
UA	Uncertainty analysis
UI	User Interface
VOCs	Volatile organic compounds
WHO	World Health Organization
WP	Work Package

1. About this document

This document describes the theoretical framework used to conceptualize the Port Environmental Index (PEI) and describes the mathematical algorithms that will be tested in order to choose the most robust approach for its computation. The final algorithm will be described in PEI definition and algorithms v2. PEI is a composite index which aims at integrating all the environmental aspects of ports operations and their respective indicators into a single metric. The idea behind PEI is for it to be used as a metric that the small and medium-sized ports will use to address their own environmental performance. In addition, the consortium has the aim to standardize the methodology. When this action is achieved, the resulting methodology will allow comparisons of environmental performance between the ports. PEI is built upon significant environmental aspects of port operations which have been identified in D5.1 Environmental aspects and mapping to pilots and this document thus is a logical sequel to that deliverable. This deliverable presents the toolbox for building the CI using a sequence of steps which are comprised of different data acquisition methods as well as statistical methods for data manipulation. Thus the deliverable describes possible data collection methods and the infrastructure needed for their deployment (IoT sensors, web forms, etc) as well as statistical functions which will be used for manipulating the data which include the following: data imputation, data normalization, weighing and aggregation Finally, a methodological approach for testing the robustness of the algorithms and sensitivity to input data is also presented. The mathematical algorithms presented in this document are several and include different methodologies which are described in detail. Based on the proposed methodologies and algorithms analyses a final set will be chosen and described in the D.5.3 PEI definition and algorithms v2. The follow-up deliverable will also include an executable PEI computation code for integration in the overall PIXEL IT infrastructure. Finally, the document includes different visualization approaches of PEI including its sub-indices and indicators with the aim of providing the most relevant information to port operators as well as a description of the links to WP4 Modelling, process analysis, and predictive algorithms and WP6 Enabling ICT infrastructure framework.

1.1. Deliverable context

Keywords	Lead Editor
Objectives	The main objectives of the deliverable are the following: (a) describe the theoretical framework upon which PEI is built; (b) describe the data sources, the methodology, and infrastructure for PEI computation, and (c) present the statistical methods to be deployed and tested. In addition, the document provides the links to other WPs and describes the visualization of the PEI in the PIXEL dashboard.
Exploitable results	The results presented in this deliverable will be exploited in its sequel D5.3 PEI definition and algorithms v2 where based on the analysis performed in this deliverable the exact statistical steps and mathematical approaches will be decided upon and coded in an executable that will compute the PEI. In addition, WP6 will exploit the results of this deliverable for the development of a PEI dashboard to be included in the PIXEL platform.
Work plan	This deliverable is the result of the work performed in M7 to M18 and is related to the task 5.3 PEI development.
Milestones	This deliverable is a partial verification of the milestone MS6 under WP5 PEI “Development completed”. The milestone will be completely verified in D5.3 PEI definition and algorithms v2.
Deliverables	This deliverable builds on the previous deliverable D5.1 Environmental aspects and mapping to pilots. In the deliverable D5.1, the main environmental aspects of port operations have been identified and the

	relevant eKPIs have been recognized. Based on this work the deliverable D5.2 will address the data retrieval procedure and sources, select a minimum list of the eKPIs to be included in PEI calculations and provide a description of the different approaches to visualizing the PEI in the PIXEL dashboard. Based on the proposed statistical tools and methodologies D5.3 PEI definition and algorithms v2 will provide the final methodology and an executable code for PEI computation.
Risks	WP5#10. Data availability – the needed data/KPIs for computing PEI will not be available (for pilot ports). WP5#11 Data standardization and interoperability– different pilot ports will have different types of data, or the same type of data measured with different methods which makes comparisons difficult. This deliverable will partially address the risks related to the data acquisition procedure, especially related to the automated data collection procedure through IoT. Although the main objective of this deliverable is the methodological approach for PEI calculation the lack of data availability as well as data standardization and interoperability will be discussed. WP6#12 KPIs weighing - weighing environmental indicators is hard The deliverable will discuss different weighting approaches with the aim of minimizing the inherent biases which are a part of the weighing procedure.

1.2. The rationale behind the structure

This report describes the work done under WP5 T5.3 PEI development. The sections describe and follow a logical sequence for building composite indicators: data acquisition, data imputation, and normalization, weighing and aggregation. Finally, the uncertainty and sensitivity analysis will be described.

1. Theoretical framework
2. eKPIs used for PEI computation
3. Links to WP4 and WP6
4. Statistical toolbox for PEI computation
5. Data imputation
6. Data normalization
7. Weighing of eKPIs
8. Aggregation methods
9. Uncertainty and sensitivity analysis
10. PEI visualization and dashboard

1.3. Version-specific notes

This deliverable presents an overview of the work done under T 5.3 PEI development during M7 – M18. The final version of this deliverable D5.3 PEI definition and algorithms v2 will be finalized in M18 and will provide a full description of the data sources and algorithms for PEI calculation including the executable code to be implemented in the PIXEL IT infrastructure.

2. The Port Environmental Index: theoretical framework

The port sector is a branch of economy whose activities have a direct impact on the quality of environment. As such, it should implement the environmental standards and regulations in order to reduce and/or prevent the adverse environmental effects. In addition, many specific and risky processes which a part of a complex port system are being continuously monitored by public stakeholders which expect the port operators to address and mitigate potential environmental side effects of port operations.

The lack of commitment of ports management towards environmental issues is being increasingly recognized as important and port authorities will have to be more proactive in implementing environmental protection initiatives in all segments of their operations[1]. However, so far appropriate and evidence-based quantitative tools for addressing the ports environmental performance have been lacking, despite several attempts at devising methodologies for addressing the environmental performance of port environments (Kegalj 2016).

Although the port sector and the shipping industry share many challenges and have common environmental consequences of their operations, the environmental responsibility of these sectors has traditionally been considered separately. While maritime activities have been under increased scrutiny not only by the regulators, but also by the increasing number of stakeholders interested in environmental quality, traditionally the environmental policies in the port sector have so far consisted of mostly voluntary actions and initiatives (Kegalj 2016). Today, the increasing development and deployment of environmental directives and regulations is growing while renewable energy and their carbon footprint are becoming priority issues for ports. Moreover, the ports have to show that they conform to the regulations: they have to keep improving themselves and prove their progress on the basis of evidence based quantitative measurement. Thus, an increasing number of ports are implementing environmental management systems (EMAS) to show their commitment to the environment and reduce operational costs and risks.

In the last few years, the growth of global trade resulted in the fast growth of the amount of cargo transported by sea and transhipped in ports. This resulted in an increase of pollutant emissions from port areas in the environment (Joint Research Centre-European Commission 2008). The most significant sources of pollution include incoming ships, loading/discharge operations, bunkering, cargo storage and storage and industrial facilities located within the port area.

Despite all of the above a standardized evidence-based and quantitative measure of port environmental performance is currently lacking. Therefore, current estimate of the port's environmental impacts are not homogeneous and they don't allow for addressing trends in environmental performance nor interport comparisons. Thus, one of the most important goals of PIXEL is to develop a quantitative, transparent and standardized Port Environmental Index (PEI). PEI may serve as basis for monitoring the environmental progress of a port and comparisons against a benchmark. Besides the PEI may assist port management by providing insight into ports environmental issues. It also represents a tool for communication of environmental performance towards stakeholders and thus has the possibility to serve as a marketing tool.

The main idea behind the PEI is based on the analysis and integration of all aspects of port processes which impact the environment. All the relevant environmental indicators (eKPIs) reflecting the impact of port processes on the environment will be merged into subindices and finally aggregated into one unique port environmental impact metric.

Individual indicators of port processes impact on the environment will be selected on the basis of available resources considering their significance, measurability, and representativeness. The main steps when building the PEI are represented in *Figure 2.1*.

[1] Here it is noteworthy that, particularly, ports participating in PIXEL project (GPMB, ASPM, PPA and THPA) are on the list of ports that have tacked, one way or another, environment protective actions. For more information, please refer to deliverable D3.3. of PIXEL. One clear example: LNG dredging action from the Grand Maritime Port du Bordeaux.

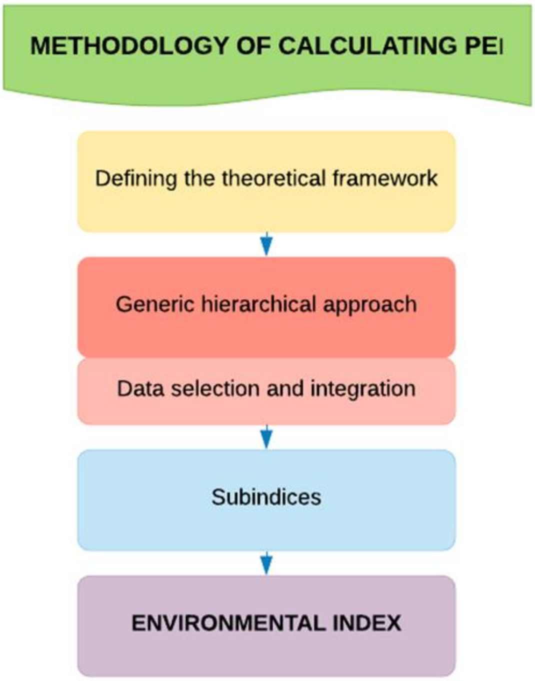


Figure 2.1 PEI methodology

For the purpose of integration, both qualitative and quantitative statistical methods will be used. Overall the PEI should:

- offer a representative information of the port's overall environmental performance;
- be simple for interpretation;
- be based on international standards and offer grounds for international comparison;
- be adequately documented with procedures set in place for ensuring data quality;
- be regularly updated in accordance with transparent procedures and algorithms.

Considering the above, the results obtained by calculating PEI are applicable to:

- indexing ports areas according to their environmental performance,
- evaluate the current environmental condition,
- identify environmental trends of a certain port system,
- estimate long-term environmental sustainability of a port system,
- tackling environmental/green initiatives in ports due to the PEI analysis
- assess the impacts of those initiatives
- benchmark/compare port performance during time

The PEI algorithm is based upon the composite index methodology, which includes various statistical methods for data processing and integration (Nardo et al. 2005). The deployment of PEI will establish a proactive port orientation towards sustainability.

Besides the many advantages, the methodology of composite index has certain shortcomings. The main characteristic of the composite index is its ability to present the results of an integrated analytical framework. A well-designed, strictly implemented composite index may be informative, with potential to provide a "wider image", i.e. the multi-dimensional nature of complex systems, as well as offer a concise statistical analysis which communicates the status and the trends of the system to relevant stakeholders in various ways (Booysen 2002; Hahn 2008; Zhou and Ang 2009; Balica 2012; Ravallion 2010). Maggino and Zumbo (2012) claim that the potential advantage of composite index development is that it may assist in overcoming the issues regarding accurateness, reliability, precision, and validity related with the use of certain indicators, i.e. the variable that is not directly discerned through individual indicator may require an integration of multiple indicators, each of which corresponding to a certain aspect of the variable. Composite index designs may and should be developed in time.

Some of the major shortcomings of the composite index refer to the process of its construction which includes a subjective evaluation in several stages, like weighting and providing missing values. Another disadvantage of the composite index may arise if the construction of the composite index carries an error or the indicators were wrongly interpreted, consequently yielding the wrong result to the decision-makers. The criticism related to the composite index emphasize that creating a group of corresponding indicators is sufficient to describe a certain issue and there is no need for the development of a composite index. Also, some say that the composite index is a waste or that it requires a lot of effort to collect and arrange the data into one unique value. Furthermore, it was pointed out that composite indicators may be misused to advocate the desired policy, while the process of construction is not transparent, and it does not have clear statistical conceptual principles. It is considered that the composite index may result in inappropriate rules unless dimensions and performances that are difficult to measure are taken into consideration. Kaufmann and Kraay (2007) claim that uniting the indicators may reduce the impact of errors in measurement related to any individual indicator. On the other hand, others warn that uniting the indicators increases the impact of error in measurement and that the above problems related to individual indicators amplify in the aggregation process.

Therefore, it is important to point out that uniting individual indicators into a composite index to obtain concise statistics results in the loss of specific traits and may conceal important information about certain indicators (Molle and Mollinga 2003; Abson et al. 2012; Kenney et al. 2012). The composite index may not be able to record the connection of indicators, may neglect important dimensions that are difficult to measure and may conceal the weakness of some components (Molle and Mollinga 2003; Zhou and Ang 2009; Abson et al. 2012).

Considering these and other related issues, the complex index may potentially provide wrong guidelines in policy and practice if they are applied in a discriminatory way or if the results are misinterpreted, misrepresented or overrated. Therefore, care should be taken to avoid such risks. As it was already said, it is important to be aware that the variables that can be measured easily or that are readily available do not necessarily constitute analytically correct or valid indicators. As Barnett et al. (2008) stated, indicators are sometimes selected not because the data reflect important elements of the model, but because the data are relatively available and are easily manipulated. Care should be taken to avoid that trap. Error in measuring input data is the source of insecurity in index output (Tate 2013). A combination of various data sources may increase the impact of measurement error and thus provide biased final results. Other types of errors in measurement related to surveys include the error in sampling, problems in survey distribution, not answering the survey, unclear questions or answers, difference in opinions among the subjects and mistakes in data treatment (Kaufmann & Kraay 2007; Joint Research Centre-European Commission 2008). Other disadvantages include missing values and mistakes incurred during data or formula revision (Wolff et al. 2011).

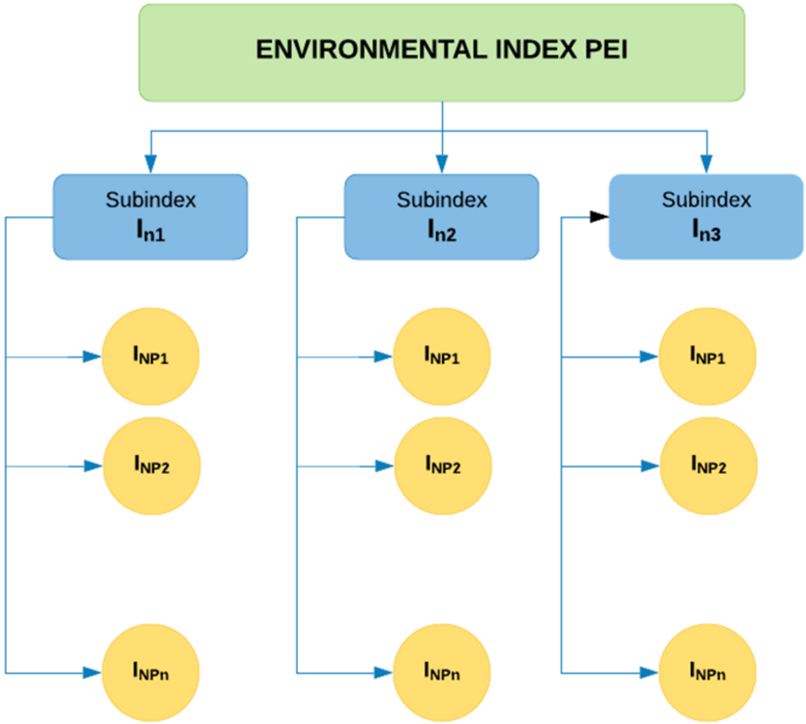


Figure 2.2 Generic linear hierarchical model for the integration of port environmental impacts into a composite unique environmental metric

The PEI resulting from the model should be:

- representative, i.e. to include all aspects of port impact on the environment,
- straightforward and simple for interpretation,
- able to be periodically calculated and updated,
- serve as a benchmark for to the ports and port authorities with the purpose of optimizing the technological processes, reduce costs and improve operations,
- provide basis for possible ranking of port terminals of the same type and purpose.

In the process of developing the theoretical model, it is necessary to determine the stages – procedures in forming the environmental index *(*Figure 2.3*)*.

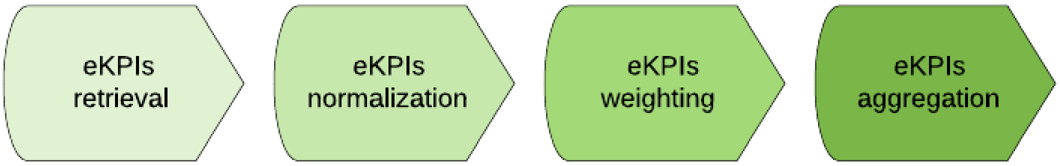


Figure 2.3 A generalized procedure for PEI calculation. From eKPIs retrieval through IoT data sources to the final procedure of their integration into the final PEI metric

In the first step, it is necessary to define the most relevant indicators of port processes impact on the environment. In the selection of indicators, account should be taken for granted that they are universally applicable. As stated before, relevant indicators are selected on the basis of their significance, measurability, and representativeness, as well as their mutual relation, i.e. correlation. Some environmental eKPIs will be directly measured whereas others will be obtained through proxy data.

3. Environmental aspects of port operations

Assessment of port environmental performance is the main objective of the PEI. In order to make the assessment as precise and complete as possible, it is important to choose valid environmental aspects to be covered. Numerous sources describe different port environmental aspects. After a thorough examination of in D5.1 Environmental aspects and mapping to pilots, the following ones were decided to be considered in the PEI:

- Emissions to the atmosphere
- Wastewater emissions
- Production of waste
- Light pollution
- Odor
- Noise pollution

It needs to be pointed out that other environmental aspects are sometimes regarded as significant, most notably resource (energy) consumption. The reason for the omission of the aspect is that it is already covered with "emissions to air" (air pollution) aspect which results from energy resources consumption; estimating this consumption separately would consequently lead to double counting and overestimating of that aspect's impact.

3.1. Emissions to air

Emissions to air are considered to be one of the most important (if not the most) environmental aspects of port operations. Those emissions are especially significant in cases of unfavorable atmospheric conditions, such as low wind speeds and stable stratification (Bachvarova et al. 2018). There are several different pollutants described in the literature (Bachvarova et al. 2018; Bailey and Solomon 2004):

- carbon monoxide (CO)
- volatile organic compounds (VOCs)
- nitrogen oxides (NO_x)
- sulfur oxides (SO_x)
- particulate matter (PM)
- ozone (a secondary type of pollutant)

Those pollutants are described as presenting a threat to human health, mainly increasing risk of lung cancer, asthma, and respiratory and cardiovascular diseases, as well as increasing risk of bronchitis in children (Bachvarova et al. 2018; Bailey and Solomon 2004). In addition to these major pollutants, there are also other air pollutants, such as formaldehyde, heavy metals, dioxins and pesticides (Bailey and Solomon 2004).

Also, there is an issue of GHG (greenhouse gas) emission. Those are the gases that contribute to the greenhouse effect, which results in excessive heating of the Earth's surface. In works such as Styhre (2017), their reduction is described as a "major challenge". The main greenhouse gases resulting from ships in ports were listed:

- carbon dioxide (CO_2)
- methane (CH_4)
- dinitrogen oxide (N_2O)

The same gases were also listed in Saharidis and Konstantzos (2018), which deals with truck operations in ports.

Atmospheric emissions are caused by several activities. They are usually divided into the following categories (Bailey and Solomon 2004; Casazza et al. 2019; Zheng et al. 2017):

- vessels (excluding harbor craft)
- harbor craft
- cargo-handling equipment
- traffic (road and railway)

Although ratio between the emissions resulting from different activities might differ significantly from port to port, to put things into perspective, data for the Port of Los Angeles in 2011 is provided. Port users (vessels, trucks, and trains) were responsible for around 80% of the total emissions, while cargo-handling activities, consisting of emissions caused by harbor craft vessels and port cargo-handling equipment, contributed with the remaining 20%. Those 20% are also the activities most easily controlled by the port operators (Zheng et al. 2017).

In Bailey and Solomon (2004), no allocation of pollution between sources is provided, but large ships (cargo, tanker and cruise ships) are listed as the main sources of air pollution, with the influence of tugboats and towboats being described as not negligible. Also, most of the diesel port equipment used for loading, unloading, and transportation of cargo is seen as contributing significantly to air pollution. That equipment includes gentry cranes, yard trucks, "top-picks", "side-pick", forklifts and others (Bailey and Solomon 2004).

The activities also have different impact on the quantity of emissions of different pollutants. Ships are seen as the source of the vast majority of the sulfur oxides pollution, as the ship engines use heavy fuel oil, which has 2700 times higher content of Sulphur than the fuel used by road vehicles. While the ships are also one of the main sources of nitrogen oxides, other sources (such as construction machinery, railway, and road traffic) are also significant contributors. Similarly, particulate matter pollution is caused by all types of diesel and heavy fuel engines.

Considering the ship pollution, distinction should be made between berthed ships that turn their main engines off, using only auxiliary engines for heating, cooling and/or electricity, and those that use main engines while at berth. The latter represent much higher risk for the environment than the former (Viana et al. 2014). The emissions generated while at berth are called "hotelling emissions" and are a significant risk for local population (Han 2010).

In Chen et al. (2017), different types of ships were ranked according to their influence on SO_2 emissions, with the container ships, fishing boats, oil tankers, and bulk carriers being the main sources of pollution. All of the ship types, with their respective shares of SO_2 pollution, can be seen in *Figure 3.1*.

.

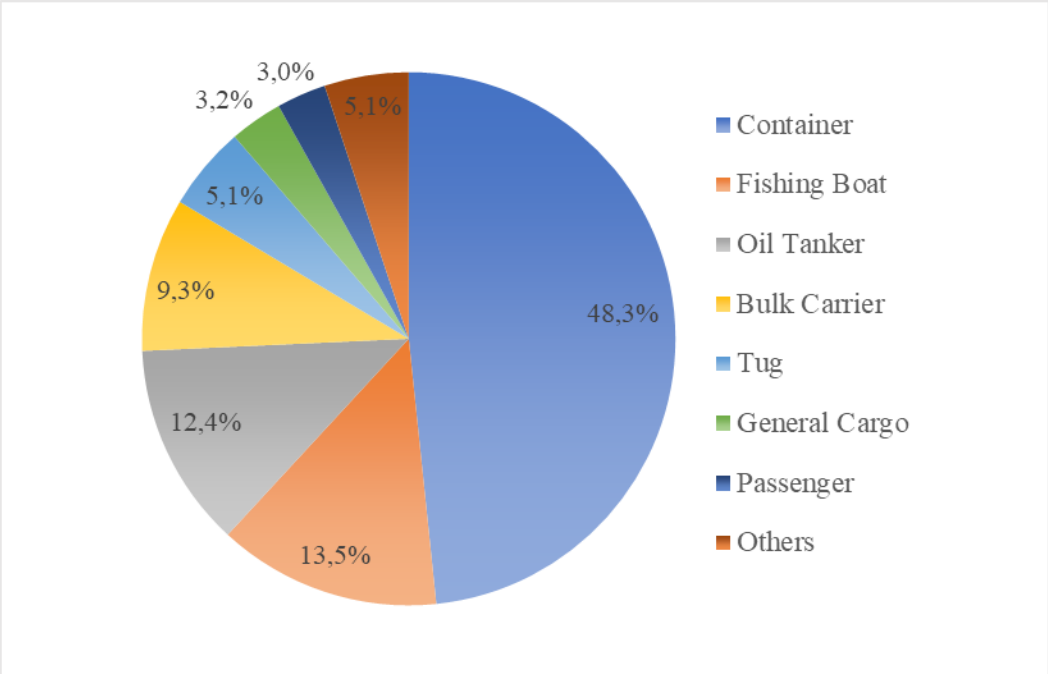


Figure 3.1 Shares of SO_2 pollution by ship type (Chen et al. 2017)

Another study covering the impact of ships on the air pollution addressed the NO_x and SO_2 emissions of different ships in the Port of Copenhagen, making difference between the maneuvering activities and time at dock (Saxe and Larsen 2004). The results can be seen in the *Figure 3.2*.

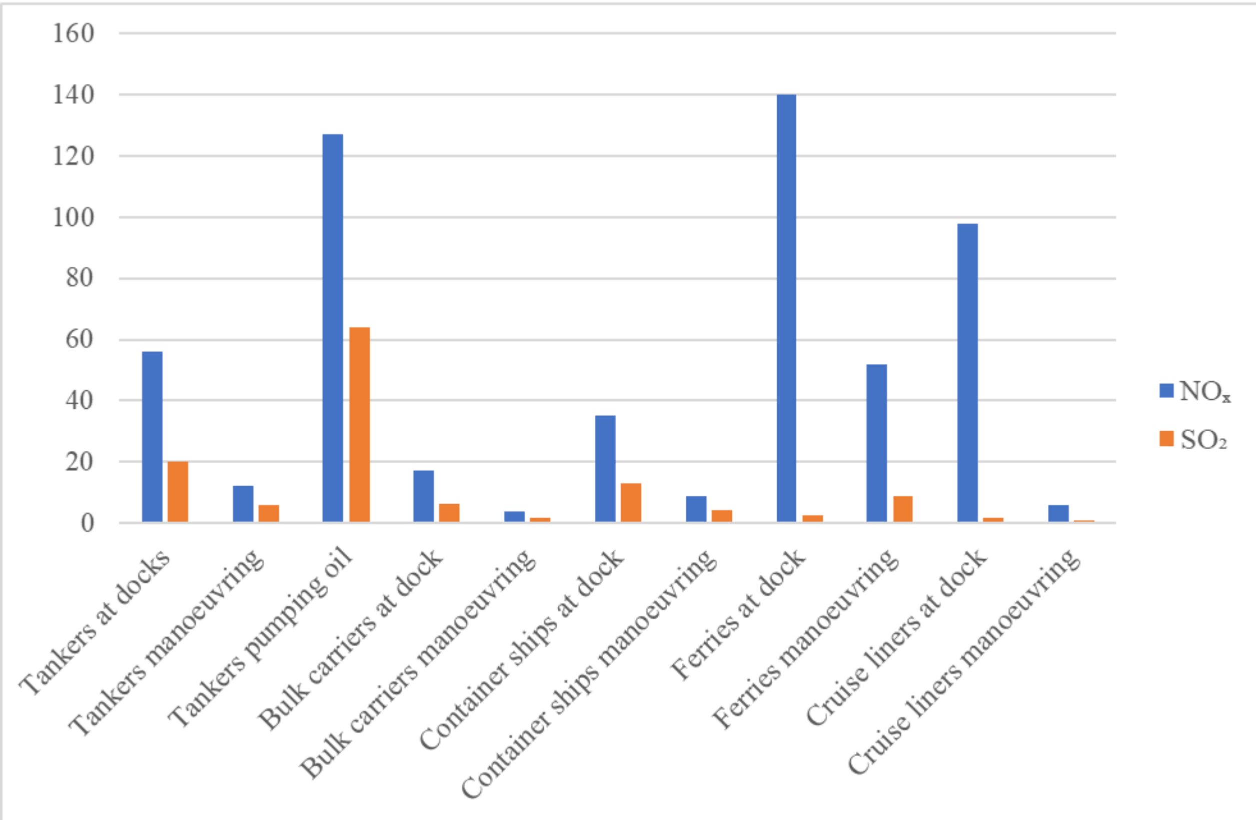


Figure 3.2 SO_2 and NO_x emissions by ship type and activity (Saxe and Larsen 2004)

More detailed representations of the emission ratios of ship activities (in the Port of Gothenburg) were provided in (Winnes et al. 2014), with the results being shown in the *Figure 3.3*.

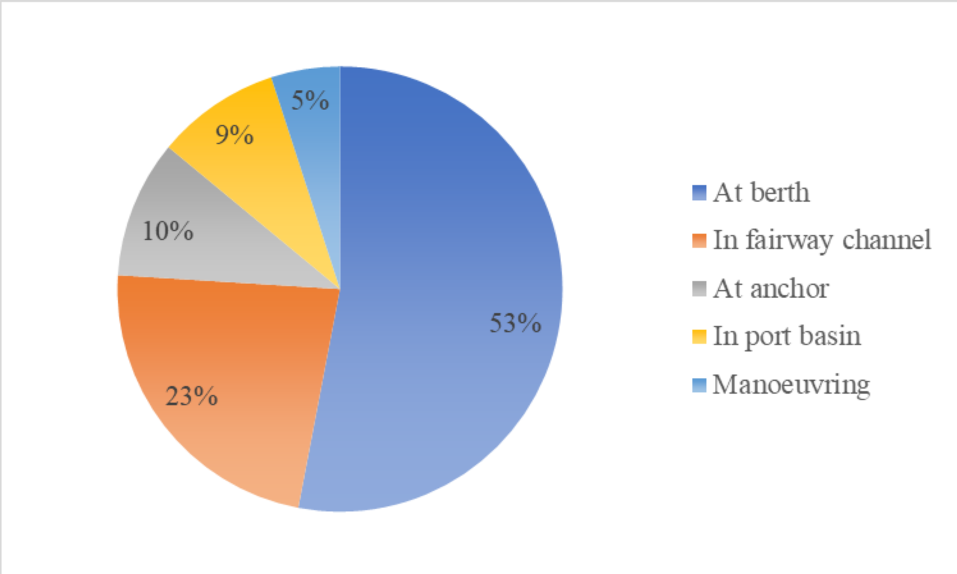


Figure 3.3 Emission ratios of ship activities in the Port of Gothenburg (Winnes et al. 2014)

Similar to the studies covering impact of different ship types, a research was conducted and presented in Martínez-Moya et al. (2019), describing the contributions of various sources of pollution used for port operations in Noatum Container Terminal Valencia to the total CO_2 emissions. RTG (rubber-tired gantry) cranes and yard tractors were discovered to be the main pollutant, while the emission share of other equipment can be seen in *Figure 3.4*.

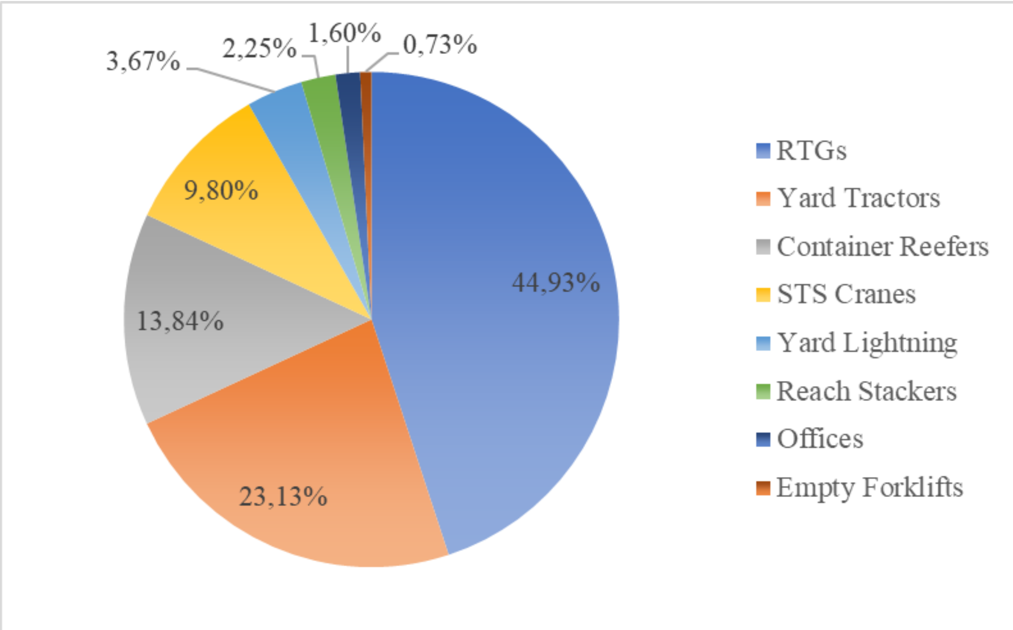


Figure 3.4 CO_2 Emission shares of port equipment in Noatum Container Terminal Valencia

In addition to the papers dealing only with the influence that port have on the environment, some strategies and technologies for the reduction of air pollution, as well as its monitoring, in the ports are presented in works such as Bjerkan and Seter (2019), Casazza et al. (2019), Dulebenets (2018) and Han (2010).

3.2. Water pollution

Water pollution in seaports can be a result of numerous different emissions. In Gómez et al. (2015), many of the activities that could potentially represent a risk for the water quality are listed. Those activities are as follows:

- management wastes activities (collection, transportation, and storage of wastes)
- vessel/port interface (includes, but not limited to, abandoned vessels, mooring, vessel repairs, oil-supplying installations and manipulation of dangerous cargos)
- cargo terminals activities (exterior and interior storage and distribution, residues of cargos)
- passenger terminals (parking/land traffic)
- fishery installations (organic residues, shellfishing residues, sewage, and process waters)
- urban activities (storage urban residues, construction, and demolition, sewage waters)
- industrial activities (storage industrial residues and raw materials, contaminated surfaces and sediments, waters used in refrigeration systems, sewage and process waters)
- agricultural activities (residues from livestock farming, pesticides)
- maritime activities (antifouling activities, ballast water, waste from ships, invasive species in hulls, extractive activities, biofouling in hulls, permanent anchorages)
- other (all other activities that don't belong in those categories

Similarly, de la Lanza Espino et al. (2010) list vessel construction, maintenance and repair, food and fish processing, seafood sale, maritime equipment sales, marine product preservation and property leasing as the main runoff sources.

Contaminant sources can be divided by the method of discharge (Gómez et al. 2015):

- point – contamination through some predefined, fixed point (i.e. channeled run-off, storm relief, and sewage)
- diffuse – unchannelled source of contamination (i.e. dredging and filtrations)

Although the main point of this section are environmental aspects of port operations, it should be noted origins of the contamination can be external or internal. Even when considering only internal sources of contamination, there is also contamination caused by the foreign activities and a clear distinction between those should be made. It is not uncommon for external and foreign activities to have a share in the pollution of 20%-40% (Gómez et al. 2015). In the same paper, several ports have listed contamination sources, both internal and external, with the number of them in most cases exceeding 40.

In Ondiviela et al. (2012), few examples of the discharge sources have been given. In the Port of Gijon, main point sources of pollution are shellfish hatcheries and urban areas, with most of the urban discharges resulting from concessionaries activities (urban effluent). Most of the diffuse sources originate from loading of solid bulk and liquids, fuel supply and handling of the containers. It is interesting to note that most (43%) of the diffuse sources come from port activities and only 29% from the concessionary's activities, while concessionaires' activities are responsible for 65% of point sources and port activities for none. The rest of the sources are a result of external activities. Also, bacteriological and organic contaminants are seen as low hazard, while the maximum risk was given to the small sewage effluents (effluents resulting from a smaller number of people, such as port employees).

In Puig and Darbra (2019), oil chemical spills and the spreading of invasive species by the exchange of ballast waters are listed as the main environmental impacts in the realm of water pollution. In the same source, water pollution is described as an increasingly significant environmental aspect, rising from the 10th place in 2013 to 4th place in 2017. As seen in Olson (1994), the oil spills are sometimes included in "waste" category and not "water pollution". Considering oil spills, it is very important to have in mind that around 80% of all oil spills occur inside the port and harbor (Ball 1999).

3.3. Production of waste

In a port, the waste production can be caused by different activities: administrative and planning activities of the Port Authorities (garbage), cargo handling operations, port industry, shipbuilding and repair, cruise ships or ferries garbage, etc. (Darbra et al. 2005). The handling of waste is composed of two phases – collection and treatment. The main difference in the conduction of those phases is that the first phase must be conducted on every ship and in ports, while the second one is done only partially on ships and in ports. A significant issue is the reception capacity in various ports – a ship must either wait until it arrives in a port where the waste can be disposed according to the environmental regulations or the crew will discharge it illegally (Olson 1994).

There are several types of waste generated by the ships and ports (Olson 1994):

- oily waste – mostly connected with oil terminals and tank farms. The difference should be made between accidental oil spills, usually caused by overflows during the loading of various oil tanks and by the bursting of pipes, and operational (predictable) oil spills, which happen during the cleaning of, for example, land tanks, and oil separators;
- bulk chemical waste – generated during the handling of oil products in bulk in the ports and terminals and during the release of ballast and tank wash water in chemical tankers. This type of waste is considered to be very harmful both to humans and to the environment;
- noxious substances in packaged form (dangerous goods) – those substances can differ by their origin (ships or ports). The spills usually happen because of the defective or inadequate packaging of the substances and occur during the handling in the terminals. The same can happen on the ships, although badly secured cargo can also contribute to this type of pollution;
- sewage – there is both port- and ship-generated sewage. Ships should discharge sewage to shore reception facilities in order to minimize negative impact;
- garbage – the definition encompasses all kinds of waste (not counting fresh fish) that is generated during the usual ship operation. It can include everything from empty boxes and bottles to engine room waste and discarded medicines.

Although not all of it happens in the port areas, it is interesting to note that ships are responsible for the majority of oil in the sea, as well as for the most garbage that is thrown into it (Pérez et al. 2017). The same source also provides the data on how much the amount of garbage has increased, at least in Spanish ports, and the increase is very significant (over two and a half times from 2003 until 2014). It is also estimated that around ¼ of all ships bring waste to ports, with the amount of oily waste being a function of the age of the ship and the amount of solid waste depending on the number of people on the ship.

Another interesting fact is that cruise ships generate around ¼ of total waste generated by merchant vessels, despite having a share of less than 1% in the total global merchant fleet (Butt 2007). This also means that the ports that have a larger number of cruise ships should be much more worried about this type of issues.

In Wang et al. (2018), five types of waste resulting from cruise ship activities were described:

- sewage (“black water”) – human body wastes;
- greywater – wastewater from sources such as baths and laundry;
- oily bilge water – mixture of water, oil, and lubricants;
- solid waste – food, garbage, refuse, trash, rubbish, and sludge;
- hazardous waste

As per PEI, the first three types of waste will be covered under the wastewater emission. Solid waste and hazardous waste generation are two categories that will be considered as a separate environmental aspect during the PEI development. To put things into perspective, Wang et al. (2018) give the following daily pollution loads for a large cruise ship: “4,000 gallons of sewage, 249,000 gallons of greywater, 5300 gallons of bilge water, 50 ton of garbage, 12,000 bottles and 12,000 cans, and 10 tons of hazardous waste.”

Sources of industrial waste in the ports are listed in Ball (1999) and are as follows (only the sources of solid waste are listed here):

- Ship maintenance
- Dredging operations
- Waste from vessels not used for domestic purposes
- Removed parts of the ships

Although most of this subsection is dedicated to the waste produced by the ships, there are two more categories to consider – waste generated in the port (on the land) and waste that enters ports by rivers and streams. As stated in Mohee et al. (2012), in their study of the port in Mauritius, most of the solid wastes generated on the landside belong to a category of "green wastes" and food wastes, with paper also contributing significantly (*Figure 3.5*). For the second waste category, the one resulting from outside sources, it is harder to make general assumptions about it. In the case described in the paper, main issue was domestic waste, but, depending on the port, it might also include industrial waste.

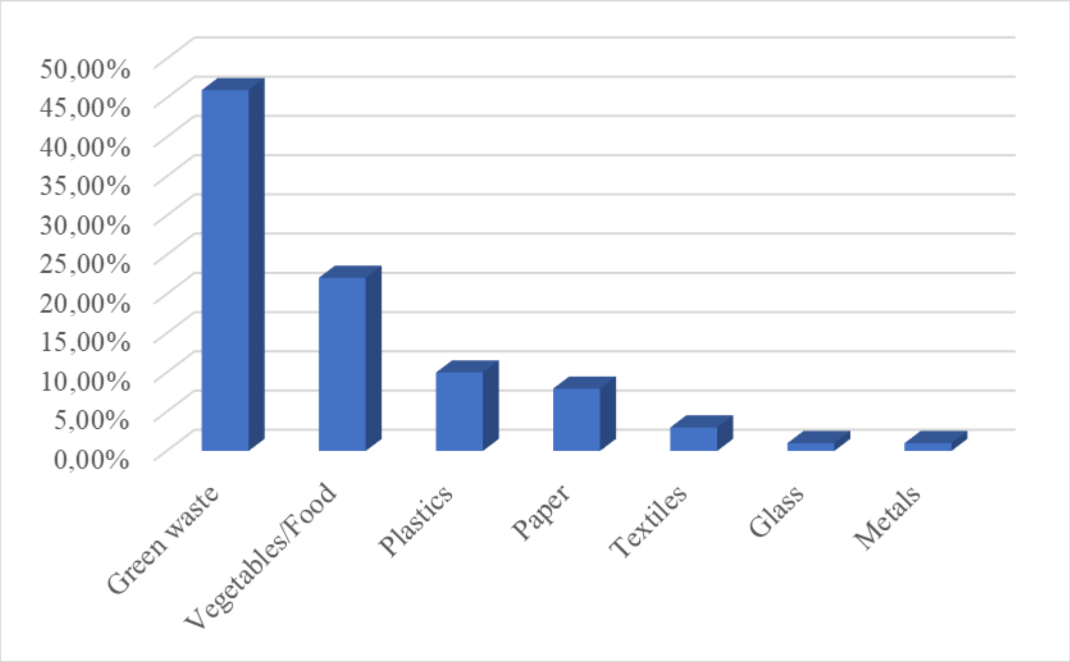


Figure 3.5 Distribution of the waste in the Port of Mauritius (Mohee et al. 2012)

3.4. Light pollution

Light pollution can be defined as "the brightening of the night sky caused by streetlights and other man-made sources that hinder the observation of stars and planets" or "any adverse effect of artificial light" (Elsahragty and Kim 2015). Although the ports sometimes do not weight it as very a significant environmental aspect, it is nevertheless described in the same source as the major environmental aspect that is "affecting our ecosystems, our health and the animals' health which requires from us facing it together". In the study of light pollution in Constanta by Pocora et al. (2016), the port area is mentioned as the most polluted area (considering light pollution) and as the main reason for the sky-glow in Constanta.

However, some of the ports recognize the light pollution as a significant environmental issue, such as the Port of Livorno. The port in question listed the activities it recognizes as having most impact on this kind of pollution (Autorità Portuale di Livorno 2012):

- light towers;
- dredging;
- maintenance of public areas;

- realization of new works;
- industrial activities on the port area;
- traffic (freight ships, passenger ships, and traffic of dangerous goods are considered to be most significant).

Light polluting devices in the Port of Tasmania are sorted into one of the four categories (TasPorts 2019):

- Fixed operational and safety/security lighting
- Vessel operational and maritime safety lighting
- On-water and landside construction lighting
- Project/event lighting

Elsahragty and Kim (2015) also mentioned parking areas, street lighting, as well as exterior building and sign lighting as contributing factors to the lighting pollution in ports. The most important of these issues is claimed to be exterior building lighting. Among the cargo handling machinery, flashing lights of straddle carriers and forklifts are highlighted as being a large nuisance to the neighboring population in Bailey et al. (2004). Also, in Jiang et al. (2018), it was stated that lighting in lifting areas is noticeably higher than in the container areas.

Although the main issue of the study was energy efficiency, most important activities with the direct influence on the light pollution were described in Hippinen and Federley (2014) and are as follows:

- indoor and warehousing lighting (of less significance for the PEI as it doesn't affect outside light levels significantly);
- outdoor lighting (high mast lighting systems, lighting of heavy industrial equipment and port container cranes lighting).

Additional port activities, not mentioned above, but which contribute to the light pollution are (PEMA 2016):

- Gate technologies
- Truck lanes lighting
- Perimeter security
- Workshops
- Near berth navigation
- Waterfront walkways lighting

The previous paragraphs provide the conclusion of insufficient tackling of light pollution by ports in global terms[2]. One of the reasons might be the same one as for the odor pollution - there is a difference in perception between port authorities and general population. For example, light pollution has little to no impact on the people working in the port, as they either don't work at night or, if they do, the outdoor lighting doesn't affect their work. However, for general population, it can be a cause of serious sleep disturbances during the night. Another one might be that it mostly influences animals and fauna, so it is not as discussed as some other aspects. Like the below-described odor pollution, light pollution is also considered by PIXEL to be an aspect that needs to be addressed during the development of PEI.

3.5. Odor

Odor pollution is considered to be a significant environmental aspect more often than light pollution. Among the studies dealing with it in more detail is Peris-Mora et al. (2005), written as part of the INDAPORT research. Different types of ports can have different operations causing odor. As it was mentioned in this subsection, fishing ports have different main source of the bad smell – fish handling and transportation, as was described in Chirmata and Ichou (2016). More specifically, the odor arises from fish loading and unloading zones, as well

[2] Note that this sentence aims to be generic for European ports. There are cases, such as GPMB, that take light and odour pollution into serious considerations and have included it when the Green Marine initial contact was established.

as outdoor storage places for fresh fish. Similar conclusions were done in Wibowo et al. (2017), also covering an odor pollution in a fish port.

The INDAPORT research also provides the review of the possible sources of odor pollution. It lists the following:

- odors from handling and transforming perishable bulk solids;
- odors from MARPOL V waste treatment;
- odours from fish handling;
- odours from water purifiers;

Paipai (1999) considers odor pollution and presents several activities that have an influence on it and also considers some legislative regulations. Among the listed sources are:

- bank disposal (in case when dredged material has significant amount of organic matter)
- demolition works (depending on the past use of the building/structure)
- cargo handling (decomposition of spilled fish)
- handling of chemicals (some chemicals have bad odor)

Additionally, Port of Gothenburg considers the loading of bunker oil as the main odor-producing activity, as it causes the release of substances that generate odors (Gothenburg Port Authority 2017). In the Doraleh Container Terminal, most of them before mentioned activities are not listed as the main source of odor. Instead, only the dredging activities (dredges materials) are seen as its main causes (African Development Fund 2008). Diesel exhaust caused by the port equipment and ships, as well as vapor resulting from liquid bulk transport are as well considered to be a potential problem (Corson and Fisher 2009)

Another important factor contributing to the level of odor is livestock transportation. The odour is mostly produced by the "biological degradation of the organic material within the pad" (McCarthy 2003). Degradation process results in a large number of volatile organic compounds, which are responsible for the smell. Despite the animals themselves are responsible for the part of the odor, the manure is the main source. It is interesting to note that the smell from the recently loaded vessels is less intensive than in those loaded more previous time point (McCarthy 2003). In addition to cattle transportation, loading of molasses during sugar seasons is seen as the main root of odor in (Townsville Ocean Terminal 2007).

In UNESCAP (2009), pollutants causing odor were divided in few categories. The first one is bottom contamination (contamination of bottom sediments) by various hazardous materials. Similarly, urban and industrial effluents can cause degradation of stagnant port water, thus causing significant smell. Inadequate waste handling and various gas leakages during the transportation of the liquid cargo can also lead to significant problems with odor pollution.

3.6. Noise pollution

Noise pollution is widely considered to be among the main environmental aspects, second only to air pollution. Probably the main contributing factor to the importance of noise pollution in ports is that ports are usually located in the close proximity to urban areas (Van Breemen 2008). When talking about noise pollution in the ports, environmental (air) noise is usually seen as the main issue. However, in works such as Wilhelmsson et al. (2013), underwater noise is also regarded as a significant environmental issue, despite its main impact being on marine species and not on human population.

Noise pollution is very clearly described in (Van Breemen 2008) and are divided in these groups:

- industrial noise sources;
- traffic noise sources;

Industrial noise sources are mainly related to cargo handling and activities related with the maintenance of ships and the port machinery. The sources include (Van Breemen 2008):

- port services and facilities;
- terminals (cargo handling and warehousing);
- industrial areas;
- machinery;
- workshop;
- vessel repair and/or maintenance;
- shunting yards;
- berthed ships.

Port activities covered with "cargo handling and warehousing" and "machinery" include several different noise sources. Among those covered in Van Breemen (2008) are:

- container handling
- other cargo handling
- cranes
- vehicles
- auxiliary equipment

Traffic noise sources division follows the usual division of traffic, with the first two being the most prevalent in ports (Van Breemen 2008):

- road traffic
- railways
- air traffic

In the consideration of noise sources, industrial noise sources are almost always considered in the noise assessment process, while there is an always-ongoing debate whether road traffic sources should be omitted or included. Main reason for their omission arises from the fact that port authorities are only partially responsible for that kind of noise pollutants. In Van Breemen (2008), it was decided that only the sources within the port area will be included in the assessment process. Although there are outside traffic sources dedicated only to the ports' needs, their inclusion would lead to a significant broadening of the scope and would require consideration of other (non-port related) sources in their proximity. The traffic is considered for evaluation once the vehicles enter the port and until they exit the port, including the entrance/exit.

4. Environmental Key Performance Indicators (eKPIs)

In the Deliverable 5.1 Environmental aspects and mapping to pilots, a list of all existing eKPIs has been compiled based on the available technical and scientific literature (Trozzi & Vaccaro 2000; Peris-Mora et al. 2005; Darbra et al. 2009; ESPO 2012; Puig et al. 2014, 2017; Puig Durán 2016; Saeedi Pash et al. 2017; Roos & Kliemann Neto 2017; Di Vaio et al. 2018; ESPO 2018; Kegalj et al. 2018). These eKPIs must be correlated with port activities and grouped according to different environmental aspects. The environmental aspects that have been used for grouping eKPIs include the following: emission to the atmosphere, wastewater emissions, noise emissions, waste production, odors and light emissions.

In this part, the aim is to provide environmental and toxicological contexts of each identified eKPI and provide chemical characteristics and existing environmental limit values to assess environmental and health risks. In the next deliverable (D5.3), we will specify how these eKPIs can be obtained by manipulating individual values in a certain manner.

4.1. Emissions to the atmosphere

4.1.1. Carbon dioxide

Carbon dioxide emissions of port include the total amount of CO_2 emissions that are directly and indirectly caused by port activities. The CO_2 is one of the most emitted greenhouse gases contributing to global climate change and warming. This indicator is used as a reference against which to rate the global warming potential (GWP) of other greenhouse gases. At environmental concentrations in air (0.04%), CO_2 has no impacts on human health.

The CO_2 concentration will not be measured directly because the methodology of PEI calculation is based on emissions from ports only. Direct measurements would require the installation of sensors on all the machines of the supply chain, which is not feasible. However, GHG emissions in general and CO_2 emissions, in particular, can be derived using emission factors based on proxy data which includes fuel consumption, fuel type and technical specifications of the engine. This is the method that will be used in the project and the methodology has been proposed in the WP4 Deliverable 4.2 PIXEL models v2.

4.1.2. Particulate Matter (PM)

"**Particulate matter**" (PM) is the general term used to describe solid particles and liquid droplets found in the air. The composition and size of these airborne particles and droplets vary. Some particles are large enough to be seen as dust or dirt, while others are too small to be visible to the naked eye. Two size ranges, known as PM_{10} and $PM_{2.5}$, are widely monitored, both at major emissions sources and in ambient air. PM_{10} includes particles that have aerodynamic diameters less than or equal to 10 micrometers (µm), approximately equal to one-seventh the diameter of human hair. $PM_{2.5}$ is the subset of PM_{10} particles that have aerodynamic diameters less than or equal to 2.5 µm.
According to the WHO (World Health Organization, 2013), chronic exposure to these particles increases respiratory and cardiovascular morbidity, and mortality from cardiovascular and respiratory diseases (like lung cancer), because particles are small enough to penetrate in the respiratory system.
This indicator can be measured directly by sensors to obtain PM concentration, or estimated using proxy indicators.

4.1.3. NO_x and SO_x emissions

NO_x and SO_x are combustion products that are emitted into the environment from the ships in the form of smoke. In the atmosphere, emissions of sulfur and nitrogen compounds are transformed into acidifying substances such as sulphuric and nitric acid and are causing environmental damage on forest and aquatic life. These compounds are associated with adverse effects on human health because high levels in the air can cause respiratory illness.

SO_x and NO_x emission can be derived using proxy data i.e. using emission factors.

4.1.4. Non-Methane volatile organic compounds emissions (NMVOC)

Non-methane volatile organic compounds (NMVOCs) are a collection of organic compounds that differ widely in their chemical composition, but they display similar behavior in the atmosphere. NMVOCs are emitted from a large number of sources including combustion activities, solvent use, and production processes. NMVOCs contribute to the formation of ground-level (tropospheric) ozone, and certain species such as benzene and 1,3 butadiene are directly hazardous to human health.

4.1.5. Synthesis of atmospheric emissions eKPIs

Synthesis of atmospheric emissions eKPIs can be seen in *Table 4.1:*

Table 4.1 Synthesis of atmospheric emissions eKPIs

Matrix	***eKPI***	***Description***	***Methodology of acquisition***
AIR	***CO_2 emissions***	Measure or calculation of the total amount of CO_2 emissions that is directly and indirectly caused by an activity	Estimation using proxy data
AIR	**NO_x and SO_x emissions**	Measure or estimation of NO_x and SO_x emissions	Estimation using proxy data
AIR	**Non-Methane volatile organic compounds emissions (NMVOC)**	Total emission of non-methane volatile organic compounds in ports	Estimation using proxy data
AIR	**Particulate Matter (PM) emissions**	Measure or estimation of the total amount of particulate matter emissions	Estimation using proxy data

4.2. Wastewater discharge to marine environment

The PIXEL project and the PEI (to be developed and deployed under WP5) aim at evaluating the port's overall environmental performance and associated impacts. This evaluation must be based only on real emissions done by the port to the environment. The indicators to be used for computing the PEI must be directly related to port operational activities or to the port treatment systems efficiency which are limiting the discharges of pollutants into seawater.

4.2.1. Total water consumption

This indicator allows to identify and to quantify the total volume of water taken for the activities. This indicator can be determined, for example, on the basis of water consumption invoices for the different operators in port (port authorities, terminal and ships), or directly measured using connected flow meters positioned on the water distribution systems.

4.2.2. Sanitary wastewater

This indicator is used to determine the amount of wastewater generated by the various activities on the port and collected by the sanitation network. This indicator can be determined on the basis of the flow from the different operators in port (port authorities, terminal, and ships) to the wastewater treatment plant if it exist. Or can be directly measured using IoT connected flow meters positioned on the sanitation network.

If some of these wastewaters are not collected and transferred to treatment plant, this information needs to be known. Untreated wastewaters have higher impact if they are directly dumped in seawater.

4.2.3. Dirty ballast water recuperation from ships

For ships stability when navigating, ballast water must be taken on board. The seawater pumped can contain marine species that are carried across the seas and released at the ships port destination. This transfer could potentially introduce marine species in another environment, and they can become an invasive one with devasting consequences for the local ecosystem.

The International Convention for the Control and Management of Ships' Ballast Water and Sediments (BWM Convention) was adopted in 2004 to introduce global regulations to control the transfer of potentially invasive species. The BWM Convention entered into force on September 2017. Under the Convention, all ships in international traffic are required to manage their ballast water and sediments to a certain standard, according to a ship-specific ballast water management plan. These standards are given in the annex – Section D Standards for Ballast Water Management and rely on ballast water exchange standard and on ballast water performance standard. All ships have to carry a ballast water record book and an international ballast water management certificate. The ballast water management standards are being phased in time: currently, new ships need to install an on-board ballast water treatment system whereas existing ships should exchange ballast water mid-ocean (OMI). The transition will be done in time and ports can propose facilities to collect and treat these ballast waters.

This indicator allows to determine the total volume of ballast water collected by port and not discharged in port waters. Of course, it depends on the practices on the different ports. Although shipowners are obliged to treat their ballast water, ports provide recovery or treatment services. The data can, therefore, be obtained either directly from the port authorities that manage the service or transmitted as a form by the service provider.

4.2.3.1. Grey and black wastewater recuperation from ships

Water is important for ships being used for cleaning and cooling of machinery and, of course, providing the necessities of life for the crew and passengers. The latter produces two distinct wastewater streams commonly referred to as black and greywater. Blackwater is the sewage while greywater is water that has been used for cooking and cleaning activities.

This indicator quantifies the wastewater emissions from boats in the port. This data is quite simple to collect since the recovery and processing services are offered by port authority or terminals. Either the volumes can be collected directly by the port authority or they must be transmitted by the service provider.

4.2.3.2. Stormwater network on port

In ports as in urban areas, rainwater flows on hard and impermeable surface.

Stormwater run-off collects and concentrates nutrients and pollutants dispersed by activities. If these stormwaters are not collected and treated, they are discharged directly in port waters. The main pollutants of concern in stormwater are suspended solids, nutrients, but also heavy metals, hydrocarbons or fecal bacteria.

This indicator is used to calculate the area of the port equipped with a stormwater collection network. The larger the proportion of networks reduce and minimize the emissions of pollutants into the aquatic or marine environment. The percentage (%) of the port area that has a system for the collection or/and treatment of rainwater can be obtained by processing the aerial photos of the harbor and the map of the rainwater collection network. For the estimation of this type of emissions, it’s important to know what facilities are present on the harbor: is the stormwater just collected and discharged into the natural environment or are these rainwaters treated?

4.2.3.3. Accidental leakages or spills

In port and in its neighborhood, there are different sources of soil or seawater pollution like operations on terminals and fuel deposits (accidental discharge of oil in the soil, loss from deposit tankers and pipeline) or spill from the bulk handling device (oil, rubber, etc.) and dust spread during the handling (transports between quay and storage area).

4.2.4. Synthesis of wastewater environmental indicators

Synthesis of wastewater environmental indicators is shown in *Table 4.2.*

Table 4.2 Synthesis of wastewater environmental indicators

Matrix	*eKPI*	*Description*	*Methodology of acquisition*
WATER	**Ballast water recuperation from ships (m^3 per unit cargo)**	Total volume of ballast water collected by port	Measurement
WATER	**Grey and black wastewater recuperation (m^3 per unit cargo)**	Total volume of grey and black wastewaters collected by port	Measurement
WATER	**Stormwater network (%)**	Percentage of the port area that has a system for the collection and treatment of rainwater	Measurement or estimation using proxy data
WATER	**Sanitary wastewater (m^3 per unit cargo)**	Sanitary wastewater produced by port activities	Measurement
WATER	**Total water consumption (m^3 per unit cargo)**	Total volume of water withdrawn for activities	Measurement
WATER	**Accidental leakage or spill (per unit cargo)**	Number of accidental leakages or spills for chemicals products based on environmental management	Estimation using proxy data

4.3. Noise emissions

4.3.1. Noise emissions indicators and monitoring

4.3.1.1. Compliance with limits at day, evening and night-time

This indicator is based on direct measurements of noise emissions in the ports. It allows to determine the number of overruns of the legal limits for noise. We compare noise emissions with the values given by the legislation to determine the number of overruns over a given period: either day, evening or night according to the existing regulations. The emission thresholds are generally based on the risks of hearing impairments on humans or animals and thus make it possible to consider the impact of the nuisance on health and fauna. However, one of the objectives of PIXEL is to reduce general noise nuisance well below the hearing impairment levels.

The most relevant data for calculating the indices and the PEI should be derived from a simulation based on the different noise emission points and their characteristics. But if the emissions data of different gears (trucks, cranes, etc.) or boats are not available, acquisition can be done with sensors.

4.3.1.2. L_{DEN} (overall day-evening-night noise level)

This indicator is also based on direct noise measurement on port and allows to determine the sound level over a 24-hour period, with penalties of +10 dB(A) for night period and +5 dB(A) for evening period). This data can be estimated on simulations or on measurements done with IoT sensors.

4.3.1.3. L_{night} (23:00 - 7:00hrs noise level)

Like L_{DEN}, this indicator is also based on direct noise measurement on port and encompasses the average sound level during the night (between 23:00 and 7:00). This data can be estimated on simulations or calculated on measurements done with IoT sensors.

4.3.1.4. Noise level monitoring in seawater

This indicator is based on monitoring sounds distinguishing biological from anthropogenic ones. Noise level measurements in seawater are used to quantify the impact of port and ships traffic on marine mammals and fish populations. It is a specific indicator still under development aiming to distinguish between noise from human activities and the biological noise of different marine species.. This indicator will therefore not be included in the PEI calculation but may be integrated as a specific eKPI in the future.

4.3.2. Synthesis for Noise

Table 4.3 shows the synthesis of noise pollution indicators.

Table 4.3 Synthesis of noise pollution indicators

Matrix	*eKPI*	*Description*	*Methodology of acquisition*
NOISE	**Compliance with limits at day, evening and night-time**	Measures of the number of overruns of the legal limits	Measurement
NOISE	**L_{DEN} (overall day-evening-night noise level)**	Measure of the average sound level over a 24-hour period	Measurement
NOISE	**L_{night} (23:00 - 7:00hrs noise level)**	Measure of the average sound level by night	Measurement

4.4. Production of waste

4.4.1. Total waste production

This eKPI was chosen because it allows to consider the quantity of wastes produced by the different operators on the port without distinction of types of waste. it indicates the total amount of waste in tons disposed of by the port authorities and the terminals.

4.4.1.1. Amount or total of waste production

In function of the method of garbage collection on the ports, this data can be collected and transmitted to the port authorities. Generally, waste is collected by municipal services or private companies. Indeed, onboard weighing system exists on some trucks and allow the waste collection service to obtain a weighing of all garbage bins on a site. In the case that the trucks are not equipped with the weighting system, the garbage can be weighted by weighting the entering and exiting trucks – the difference in weight is the weight of the garbage.

These data can be collected in forms or transmitted directly by web services.

4.4.1.2. Generation of hazardous waste, solid urban waste and other

The production of waste by the ports can be declined in different typologies based on the environmental and health impacts of waste, and their mode of management more or less constrained:

- hazardous waste is a waste with properties that makes it dangerous or capable of having a harmful effect on human health or the environment: The annex III of the Waste Framework Directive gives a definition of hazardous waste as a waste that displays one or more of the fifteen hazardous properties listed;
- non-hazardous waste.

Waste designated as hazardous based on Commission notice on technical guidance on the classification of waste (2018/C 124/01) triggered a number of important obligations for instance on labeling and packaging but also in terms of monitoring and treatment. The volume or weight data can, therefore, be extracted from the waste management plan for the ports in tons or in cube meters.

4.4.2. Total garbage from ships

Regulations for the prevention of pollution by garbage from ships are contained in Annex V of MARPOL. MARPOL Annex V seeks to eliminate and reduce the amount of garbage being discharged into the sea from ships. All ships of 100 gross tonnage and above, every ship certified to carry 15 persons or more, and every fixed or floating platform must carry a garbage management plan on board, which includes written procedures for minimizing, collecting, storing, processing and disposing of garbage. And all ships of 400 gross tonnage and above and every ship which is certified to carry 15 persons must provide a Garbage Record Book and to record all disposal and incineration operations. The date, time, position of the ship, description of the garbage and the estimated amount incinerated or discharged must be logged and signed. The Garbage Record Book must be kept for a period of two years after the date of the last entry.

Although the data exists, it may be complicated for port authorities to retrieve them. Their acquisition can, therefore, be done by weighing the waste collected by the service operator performing this recovery.

Waste recycled on port

It is now important to take into account that some wastes produced may be subject to recycling or specific treatment to limit their impact on the environment.

The implementation of a recyclable sorting and waste management policy on the port can be integrated into the calculation of the PEI in the form of the percentage of recycled waste indicator.

Like for total waste production, this data can be collected and transmitted to the port authorities. Indeed, onboard weighing system exists on trucks and allow the waste collection service to obtain a weighing of all garbage bins on a site.

These data can be collected in forms or transmitted directly by web services.

4.4.3. Synthesis for waste indicators

Synthesis for waste indicators can be seen in *Table 4.4.*

Table 4.4 Synthesis of waste indicators

Matrix	*eKPI*	*Description*	*Methodology of acquisition*
WASTE	**Amount or total of waste produced**	Sum of all waste produced by port authorities and terminal operators	Measurement or estimation
WASTE	**Generation of hazardous waste**	Sum of hazardous waste produced by port authorities and terminal operators	Measurement or estimation
WASTE	**Generation of non-hazardous waste**	Sum of all solid urban waste produced by port authorities and terminal operators	Measurement or estimation
WASTE	**Total garbage from ships**	the amount of waste to be landed from ships	Measurement or estimation
WASTE	**Percentage of waste recycled in a port**	Sum of all recycled waste on port and separately collected	Measurement or estimation

4.5. Light emissions

In the past century, the extent and intensity of artificial night lighting has increased. The term of light pollution is applied to the artificial light that alters the alternation of day and night (nychthemeral rhythm).

On ports, light installations are necessary for the safety and security of loading and unloading operations, storage and personnel working on site. But these lights can sometimes be the source of discomfort for people living near the port with substantial effects on human health but also on the biology and ecology of species in the wild.

Light pollution is increasingly taken into account in the impacts of ports on the environment. But like for noise, and air, it’s difficult to quantify this impact with environmental measurement.

At this moment of the project, this indicator cannot be dismissed for PEI calculation and we, therefore, proposed to keep it if measurements are made on the ports. These measurements can be made with sensors from sources of emissions or at the edge of ports.

5. PEI data requirements

The Port Environmental Index is based on the direct impacts of port activities. The environmental values cannot be used as a single metrics without any correlation with the real activities monitoring in port.

To consider the activities of the various ports and their specificities, input data are required. They will make it possible to determine the typologies of activities of each port, based on their traffic and supply chain for cargo. But this data will also help to correlate environmental indicators with the intensity of the port's activities. The aim being to consider the specificity of each port: a port carrying out a large part of passenger traffic will not have the same impact on environment as a port having a heavy traffic of ship carrying liquid chemical compounds.

These minimum data requirements will also be useful for the eKPI values estimation: for example, if one of the eKPIs cannot be directly measured or if this data is not available from port, a proxy can be estimated. This estimation can be done based on the: historical data activities, mapping of supply chain, operational data and machines specification and emission factors. This methodology is used in the WP4 using models (Deliverable 4.2 PIXEL models v2).

These data requirements can be organized by the different operators and sources of emissions in port: Port authority; terminal operator, and shipowner among others.

5.1. For operational information

In the various Environmental Impact Assessments of Port (EIAP) (Saeedi Pash et al. 2017), the environmental indicators are standardized using operational indicators. It allows to take into account the intensity of activities and its impacts in the PEI calculation, these operational KPIs allow to compare the environmental data between different ports although they do not have the same activities. All indicators must to be reported per unit cargo.

5.1.1. Number of ships and ship type

The activity of the port and its emissions in air can be estimated based on simple indicators: the number of ships calls as well as their typology. These ships' data can be obtained by coupling vessels calls and the Lloyd's register. Their typologies of vessels can be classified using the following different categories:

- Container boats – Used for carrying 20, 40 and 45 feet containers of merchandise.
- Bulk vessels – Used for transporting bulk commodity items like iron ore, coal, and wheat, etc.
- Breakbulk vessels –Used for carrying a mixture of different types of cargo including bagged and palletized cargo.
- Ro-Ro vessels – Used for transporting wheeled cargo such as vehicles and excavators.
- Multi-purpose vessels –Used for transporting cargoes of all types mentioned above.
- Tanker vessels –Used for carrying liquid items like chemicals and oil.
- Crude carriers – These vessels only carry crude oil.
- LNG carriers –Used for transporting Liquefied Natural Gas.
- Reefer vessels –Used to carry temperature-sensitive cargo like fish, fruit, and meat.

The passenger vessels can be a specific category includes ships from 10-person foot ferries up to cruise ships able to carry over 6,000 passengers.

5.1.2. Amount of cargo and passenger

These activity indicators can be different related to the activities in port. The useful indicator is the weight or volume of cargo movement, or the number of twenty-foot Equivalent Unit (TEU) in case of containers, and the number of passengers. For more simplicity, the weight of cargo moved will be used.

5.2. Air emissions eKPIs

5.2.1. Fuel type and Fuel quantity consumption

This data gives information on the fuel consumption and fuel type used in all engines used on the port and ships.

Direct measurements are difficult because it is impossible connect a sensor between each engine. But fuel flow meters are existing and can be implemented only on typological supply chain.

If not possible, this parameter can be evaluated directly on bills of fuel consumption or it can be extrapolated starting from the supply chain and referring to the consummation characteristics (type of fuel, power and hours that the machine is operating) of the engines used employing some modeling techniques (see D4.2).

5.2.2. For terminals and port authorities

For the determination and calculation of the indicators for air, the necessary input data are as follows:

- Electricity consumption
- Characteristics of all supply chains used for loading or unloading cargo.

5.2.2.1. Trucks within the port

- Truck no.
- engine power (kW).
- working time in the port (h) of a truck within the supply chain
- fuel consumption (g or l per fuel).

5.2.2.2. Trains within the port

- Train no.
- engine power (kW)
- working time in the port (h) of a truck within the supply chain
- fuel consumption (g or l per fuel)

5.2.2.3. Port machinery

- engine power (kW)
- working time (h)
- fuel consumption (g or l per fuel)

5.2.3. For Ships

For the determination and calculation of the indicators for air, the necessary input data are as follows:

- Main engine power (kW).
- Auxiliary engine power (kW).
- Maneuvering time (h).
- Time at berth (h).
- Load factor of the main engine in maneuvering (%).
- Load factor of the main engine at berth (%).
- Load factor of the auxiliary engine in maneuvering (%).
- Load factor of the auxiliary engine at berth (%).

5.3. Wastewater emission eKPIs

For the determination and calculation of the indicators for wastewater, the necessary input data are as follows:

- Total water consumption: It can be given in cubic meters consumed by all port operators, but also by ships.
- Sanitary wastewater produced (m^3): This wastewater production can also be given in cubic meters, for all operators on port.
- Total Port area (m^2): This data is needed for eKPIs calculation.
- Port area equipped with stormwater collection system (m^2): This data is needed for eKPIs calculation.
- Blackwater recovery from ships (m^3): when this recovery is done in the port.
- Greywater recovery from ships (m^3): when this recovery is done in the port.
- Bilgewater recovery from ships (m^3): when this recovery is done in the port.

5.4. Waste production eKPIs

For the determination and calculation of the indicators for waste, the necessary input data are as follows:

- Municipal solid waste (tons).
- Hazardous waste (tons).
- Inert waste (tons).
- Recycled waste (tons) or waste collected separately.

5.5. Noise emissions eKPIs

The most relevant data for calculating the indices and the PEI should be derived from a simulation-based on the different noise emission points and their characteristics.

But if the emissions data of different gears (trucks, cranes …) or boats are not available, acquisition can be done with noise sensors. For greater accuracy in the use of the data, the results must be accompanied at least by the acquisition date, the location of the acquisition, the method used for the measurement and the unit of the acquisition.

5.6. Light emissions eKPIs

For light, the result of measurement made on the port can be used. Same as for noise measurements and for greater accuracy in the use of the data, the results must be accompanied at least by the acquisition date, the location of the acquisition, the method used for the measurement and the unit of the acquisition.

5.7. Time scale for collecting data

In the framework of the PIXEL project, the calculation of PEI must allow the ports having a simple and global view of their environmental management and its results but also, thanks to the use of numerical models, having efficient decision support system. eKPIs must provide information on the current condition of the environment in the port area. They may help port managers to:

- Better recognize the potential of impacts of all the port activities separating them by administrative entities present on the ports: Port authorities, Terminals operators and ship owners.
- Identify actions/measures that can reduce their environmental impact and will allow them to decrease their budget or justify investments.

The frequency of data acquisition depends on what we want to measure and for which purpose. Most of the data can be collected monthly due to identified seasonal variability. Indeed, passenger traffic activities or goods traffic activities are subject to strong monthly variations that will significantly influence the environmental impacts.

5.8. Synthesis of minimum data needed

The Synthesis of the minimum needed for the PEI calculation is given in *Table 5.1*.

Table 5.1 Synthesis of the minimum data needed for the PEI calculation

Type of origin	Sub-type of data	Data required	Type of data	Unit	Acquisition
Ships	Common	Amount of cargo	Integer	Tons or m3 or TEU or passengers	Web app or URL
Ships	Common	ID	Integer		Web app or URL
Ships	Common	Ship type	Text		Web app or URL
Ships	Common	Type of cargo	Text		Web app or URL
Terminals	Common	Amount of cargo	Integer	Tons or m3 or TEU or passengers	Web app or URL
Terminals	Common	ID	Integer		Web app or URL
Terminals	Common	Type of cargo	Text		Web app or URL
Supply chain	Diesel locomotive	engine power	Decimal value	kW	Forms
Supply chain	Diesel locomotive	fuel consumption	Decimal value	g/kW	Forms
Supply chain	Diesel locomotive	working time in the port (h) of a train within the harbor	Decimal value	h	Forms
Supply chain	Diesel locomotive	Train no.	Integer		Forms
Port authorities	Emission to air	Electricity consumption	Decimal value	kWh	Forms or Web app or URL
Port authorities	Emission to air	Fuel quantity	Decimal value	t	Forms
Port authorities	Emission to air	Fuel type	Text		Forms
Port authorities	Emission to air	Primary energy (type)	Text		Forms
Ships	Emission to air	Auxiliary engine power (kW)	Decimal value	kW	Forms
Ships	Emission to air	Load factor of the auxiliary engine at berth (%)	Decimal value	%	Forms
Ships	Emission to air	Load factor of the auxiliary engine in maneuvering (%)	Decimal value	%	Forms
Ships	Emission to air	Load factor of the main engine at berth (%)	Decimal value	%	Forms
Ships	Emission to air	Load factor of the main engine in maneuvering (%)	Decimal value	%	Forms
Ships	Emission to air	Main engine power (kW)	Decimal value	kW	Forms
Ships	Emission to air	Maneuvering time (hr)	Decimal value	h	Forms
Ships	Emission to air	Time at berth (hr)	Decimal value	h	Forms
Ships	Emission to air	Fuel type	Text		Forms
Terminals	Emission to air	Electricity consumption	Decimal value	kWh	Forms or Web app or URL
Terminals	Emission to air	Fuel quantity	Integer	t	Forms
Terminals	Emission to air	Fuel type	Text		Forms
Terminals	Emission to air	Primary energy (type)	Text		Forms
Port authorities	Water emissions	Sanitary wastewater	Decimal value	m3	Forms
Port authorities	Water emissions	Number of accidental leakages or spills	Integer	nb/month	Forms
Port authorities	Water emissions	Total Water consumption	Decimal value	m3	Forms
Port authorities	Water emissions	Technological wastewater	Decimal value	m3	Forms

Type of origin	Sub-type of data	Data required	Type of data	Unit	Acquisition
Port authorities	Water emissions	Volume of ballast water collected on port	Decimal value	m^3	Forms
Port authorities	Water emissions	Total area of the port	Decimal value	m^2	Forms
Port authorities	Water emissions	Total area of the port where stormwater network is in place	Decimal value	m^2	Forms
Ships	Water emissions	Bilge wastewater	Integer	m^3	Forms
Ships	Water emissions	Blackwater	Integer	m^3	Forms
Ships	Water emissions	Gray water	Integer	m^3	Forms
Terminals	Water emissions	Sanitary wastewater	Decimal value	m^3	Forms
Terminals	Water emissions	Total Water consumption	Decimal value	m^3	Forms
Terminals	Water emissions	Technological wastewater	Decimal value	m^3	Forms
Port authorities	Light	Date	Date		Sensors
Port authorities	Light	Unit	Decimal value		Sensors
Port authorities	Light	Location	Geographical coordinates	Decimal degrees (WGS84)	Sensors
Port authorities	Light	Method	Text		Sensors
Port authorities	Noise measurements	Date	Date		Sensors
Port authorities	Noise measurements	Measure	Decimal value	Db	Sensors
Port authorities	Noise measurements	Location	Geographical coordinates	Decimal degrees (WGS84)	Sensors
Port authorities	Noise measurements	Method	Text		Sensors
Port authorities	Odour measurements	Date	Date		Sensors
Port authorities	Odour measurements	OD	Decimal value		Sensors
Port authorities	Odor measurements	Location	Geographical coordinates	Decimal degrees (WGS84)	Sensors
Port authorities	Odor measurements	Method	Text		Sensors
Port authorities	Trucks	Truck no.	Integer		Forms
Supply chain	Trucks	engine power	Decimal value	kW	Forms
Supply chain	Trucks	fuel consumption	Decimal value	g/kW	Forms
Supply chain	Trucks	working time in the port (h) of a truck within the harbor	Decimal value	h	Forms
Port authorities	Waste	Amount of recycled waste on port	Decimal value	kg	Forms
Ships	Waste	Amount of recycled waste on port	Decimal value	kg	Forms
Terminals	Waste	Amount of recycled waste on port	Decimal value	kg	Forms
Port authorities	Waste	Amount of waste production	Decimal value	kg	Forms
Ships	Waste	Amount of waste production	Decimal value	kg	Forms
Terminals	Waste	Amount of waste production	Decimal value	kg	Forms
Port authorities	Waste	Hazardous waste	Decimal value	kg	Forms
Ships	Waste	Hazardous waste	Decimal value	kg	Forms
Terminals	Waste	Hazardous waste	Decimal value	kg	Forms
Port authorities	Waste	Inert waste	Decimal value	kg	Forms
Ships	Waste	Inert waste	Decimal value	kg	Forms
Terminals	Waste	Inert waste	Decimal value	kg	Forms
Port authorities	Waste	Solid urban waste	Decimal value	kg	Forms

Type of origin	Sub-type of data	Data required	Type of data	Unit	Acquisition
Ships	Waste	Solid urban waste	Decimal value	kg	Forms
Terminals	Waste	Solid urban waste	Decimal value	kg	Forms

The minimum data required are needed to calculate all the components of the Port Environmental Index. These can be grouped into 5 sub-indexes:

- AIR index
- WATER index
- NOISE index
- WASTE index
- LIGHT index

The following figure (*Figure 5.1*) summarizes the links between all inputs data needed, their uses to obtain the nineteen eKPIs values, and their integration in PEI calculation.

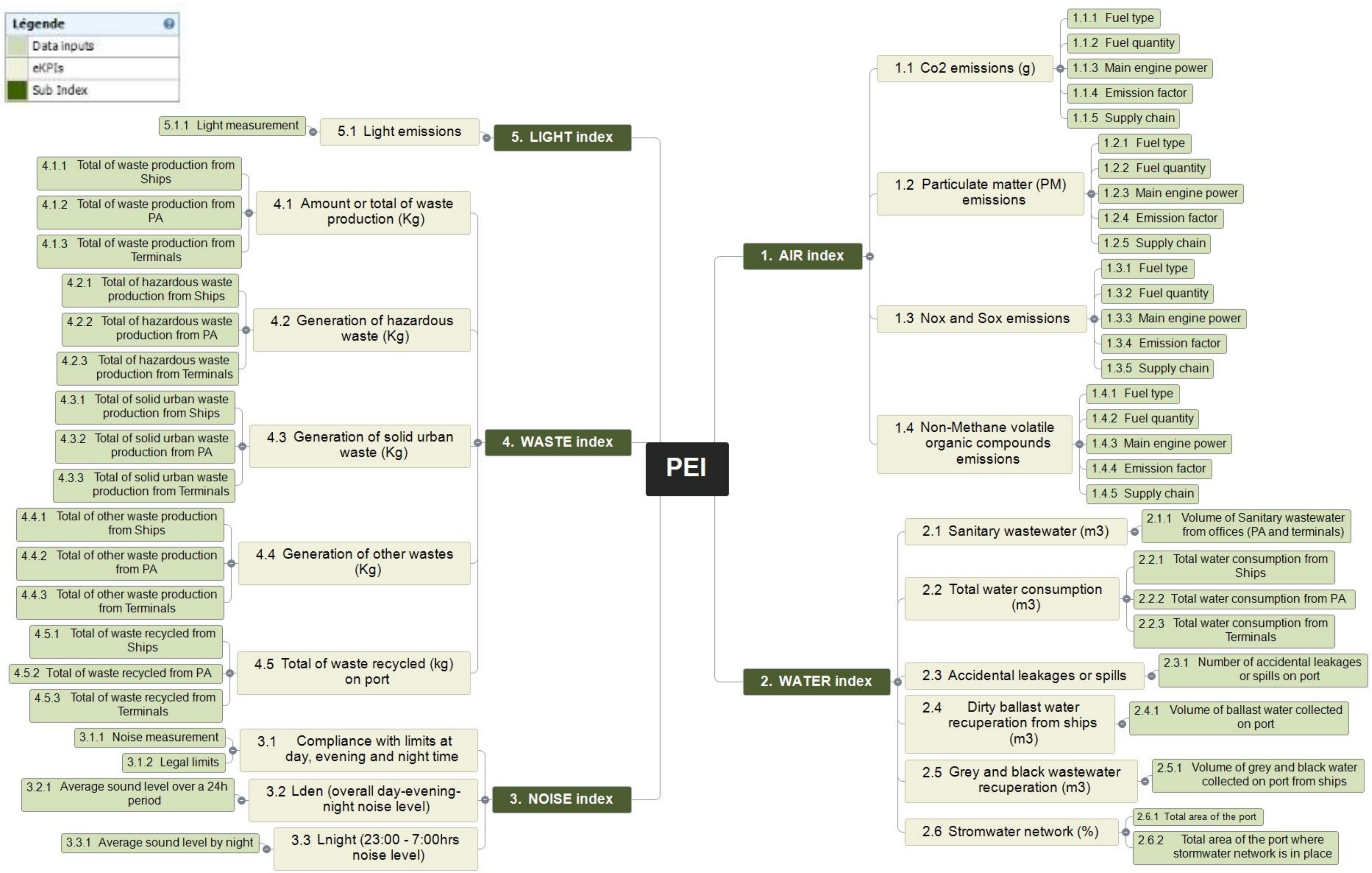


Figure 5.1 Link between data inputs and PEI

6. Building emission inventories and links to WP4

The European Environmental Agency (Van Aalst, et al. 1998) defines emission inventories as a collection of data presenting an emission of pollutant (to air) and related parameters including:

- the location of the emission and the elevation of the release of the pollutant
- the time dependence, i.e. the certain time period during which the emissions are totaled
- the chemical properties of the pollutant
- the cause of the emission and its relation to human activity

Port activities can be divided into two basic categories as shown by Paipai (1999):

- development activities. May be on land and at the land-water interface and concern construction works with the associated transfer stations for construction material and possibly demolition works with debris movement;
- operational activities. They occur both at the land-water interface and on-land. Those are the activities related to moving of cargo between the ships and the hinterland. More specifically they involve loading and unloading of cargo from ships, moving cargo to storage facilities, storage of cargo under specific conditions related to humidity, temperature and lighting, and cargo sorting and moving in and out of the port from the hinterland side. The cargo movement-related land activities of the ports have been modeled in WP 4 - Modelling, process analysis, and predictive algorithms.

6.1. Emissions locations

Ports activities are taking place both at the landside and at the seaside of the ports. At the seaside, there are specified areas for ship maneuvering and sea mooring.

In the framework of the PIXEL project, the areas inside the port can be distinguished based on the operations that are taking place in them or based on the way pollutants are being emitted.

In terms of operations, the following area types are distinguished:

- berthing areas – areas where ships dock and machinery (usually cranes) are used to load/unload cargo;
- storage areas – areas where cargo is stored. These areas can be open (yards) or closed (e.g. warehouses, tanks, silos). Operations taking place in these areas are related to sorting and storing of cargo as well as temperature, humidity, and lighting regulation;
- infrastructure (buildings) – generic areas such as terminal operators' personnel buildings, facilities management warehouses, electrical substations or package units where energy is being consumed;
- land loading/unloading areas – areas where machinery is used to load or unload cargo to hinterland transport vehicles such as various types of trucks or rail wagons;
- gates – port entrances where queues of trucks can potentially be generated due to paperwork involved.

In terms of pollutants emission, the following types of areas are distinguished:

- Point Sources – stationary sources of pollutants concentrated at a very limited space. A crane used for the unloading of one ship can be considered a point source.
- Area Sources – numerous stationary sources of pollutants operating in one area. The berthing area where two cranes are being used to unload a ship of relatively big length can be considered as an area source.
- Line Sources – mobile sources of pollutants operating on predefined routes. A gantry crane unloading a ship of relatively big length can be considered as a line source.
- Volume Sources – three-dimensional sources of pollutant emissions. Such emissions are usually dust or dry bulk cargo fugitive emissions.

As an example, in the following page we are showing the basis of the emission construction approach for Thessaloniki port (PIXEL partner):

In the following *Figure 6.1*, the port of Thessaloniki and its berthing areas (piers) are shown:

Figure 6.1 Thessaloniki Port

In the following example (*Figure 6.2)*, the operation of 11 rail cranes on the 5th Pier at Thessaloniki port is considered as pollution sources. In should be noted the cranes, their position and emission data are purely made up and serve only as an example for emission calculations in the Port of Thessaloniki.

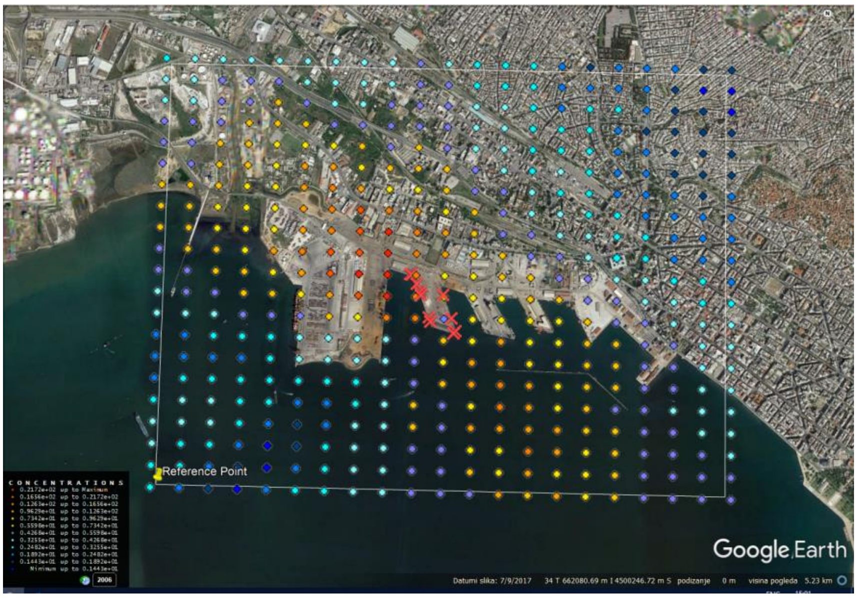


Figure 6.2 Example of SO_2 emissions dispersion from 11 diesel rail cranes

For the calculation of the emissions of SO_2 resulting from the operation of these cranes, each crane is considered a point source operating on the berthing area of the port.

6.2. Dependencies

PIXEL aims to assess the environmental impact of ports operations for the stakeholders to be able to:

1. compare the environmental footprint for the handling of a specified unit of the same cargo between ports;
2. compare the environmental footprint for the handling of a specified unit of the same cargo over time.

PIXEL aggregates the emissions (of each pollutant separately) for one supply chain per unit of a specific cargo type. In the framework of PIXEL WP4, a supply chain describes the type and number of machines, their operating duration and their position in the port when used for handling one unit of a specific type of cargo. As such, PIXEL supply chains contain only the operations that are taking place inside the port. Cargo types can be cereals, rods, wood chips, oils, containers, etc. Furthermore, the user of PIXEL system will be able to aggregate the emissions based on different cargo categories (e.g. dry bulk cargo, liquid bulk cargo, general cargo, etc.).

These types of aggregations make it possible to evaluate the environmental performance of ports of different sizes, for the same type of cargo.

Additionally, and in order for a port's stakeholders to be able to monitor its progress in terms of emissions reductions PIXEL aims to aggregate each pollutant emissions over these timeframes:

- Yearly
- Monthly
- Weekly
- Upon user's request

6.3. Causes of emissions

In WP4 - Modelling, process analysis and predictive algorithms, the ports operations related to the PIXEL Use Cases are analyzed and modeled. In particular, the following models have been considered:

- Port and city environmental management Model – develops the interoperability of energy demand and production models, pollution and transport demand models by identifying the inputs and outputs of each model and taking into account the impact of the results of one model on another. The initial triggering of the model's simulation is resulting from the Ports Activities Scenario which is fed by the vessels calls.
- Energy demand Model – models the port's energy demand and production to provide information about energy availability, reliability, and efficiency.
- Hinterland multimodal transport Model – models the evacuation of cargo to the hinterland focusing on its impact on the wider ports' area traffic.
- Environmental pollution Model – simulates the future emissions of sources related to the energy demand and the Hinterland multimodal transport models.

The interoperability of the models is shown below (*Figure 6.3).* The Port Activity Scenario feeds directly the Energy model which in turn provides directly some data to the Environmental pollution model, while the Hinterland Transport and the Emissions models are also getting input through predictive algorithms related to vessels' calls.

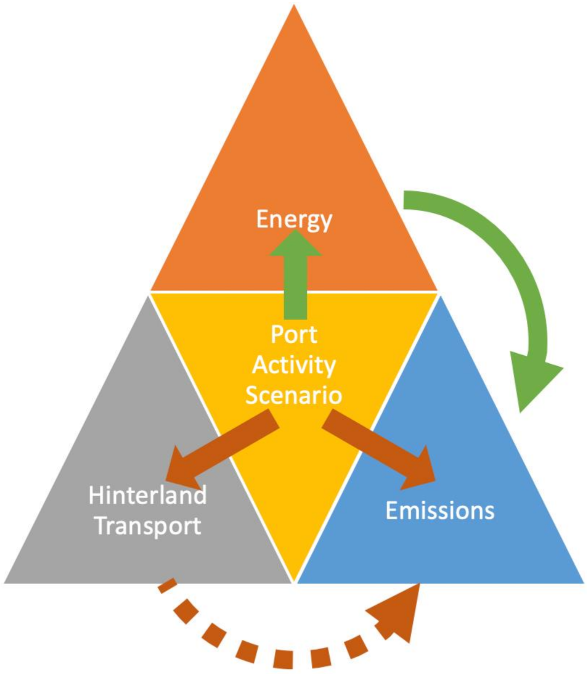


Figure 6.3 Interoperability between PIXEL models

6.3.1. Supply chain and Terminal operations

The direct link between the ports' activities, the energy demands, and the air pollutants emissions within the port area is related to the fact that usually most of the ports' machinery used for cargo handling are using internal combustion engines (usually diesel). When electricity consuming machinery is used, they may operate on electricity generated within the port (through generators or environmentally friendly methods such as photovoltaic panels) or may connect to the national grid through substations and package units.

Examples of machinery used in ports for cargo handling are cranes of any type, gantries, straddle carriers, tractors, hoppers, stackers, pumps, electricity generators, etc.

Regarding the PEI calculation, cargo-related activities are grouped within the following types of origin:

- the activities related to cargo handling (loading/unloading from ships and transfer to storage areas) within the port are generally grouped in the Terminal type of origin.
- the activities related to cargo movement in and out of the ports are generally grouped under Supply Chain.

6.3.2. Other causes

Other sources of air pollutants emissions within the ports are related to:

- the administrative operations of the ports are taking place inside buildings. These operations involve energy consumption from fossil fuel operating machinery like boilers and electricity generators and electricity operating machinery like computer systems, heat pumps, lighting, etc.

  Additionally, storage areas are consuming energy for the storage of cargo. This type of consumption is not directly related to the movement of cargo within the port, e.g. regulated tanks will remain regulated regardless of the quantity of liquid stored inside them, furthermore the amount of energy consumption is also depended on additional factors such as ambient temperature, etc. Regulation (thermal, pressure, humidity, etc.) operates on electricity generated within the port or through the national grid.

  Finally, one other source of energy consumption within the port is lighting. Usually lighting in ports is provided by electricity from the national grid.

  For the PEI calculation, we propose to group these sources under Port Authorities.

- Ships movement within the port involves not only the use of the ships' engines for maneuvering and docking but also the use of additional boats such as tugboats, coast guard and customs vessels, etc. Sea vessels use a variety of energy sources, however most of the emissions originate from the use of navy fuel or diesel.

  These sources of pollutant emissions are grouped under the Ships type.

6.4. Types of pollutants considered

All ports activities are related to air pollutant emissions in general and are also an important source of environmental noise pollution. Furthermore, solid bulk handling has the risk of accidental release of bulk material to the environment and liquid bulk handling has the risk of accidental spills. Also, intensive maritime transport activities are related to abrasion of the sea ground, underwater noise, etc. Finally, administrative activities and ships berthing are also related to waste and wastewater.

6.4.1. Air pollutants

Environmental air pollution is the top 2018 environmental priority (EcoPorts | ESPO 2018) by the European port sector as a whole, followed by energy consumption.

In the first version of this deliverable (Deliverable 5.1 – Environmental factors and mapping to pilots) and in chapter 9.2 / table 9.1, the pollutants emissions identified as eKPIs (environmental Key Performance Indicators) to assess the port's overall environmental performance, and for computing the PEI, were:

- CO_2 emissions
- Fine particles emissions (NO_x, SO_x – linked both to cargo handling and dredging operations)
- Non-Methane volatile organic compounds emissions (NMVOC)
- Nitrogen Oxides (NO_x) emissions
- Carbon monoxide (CO)
- Particulate Matter (PM) emissions
- Total Greenhouse Gas (GHG) Emissions

As mentioned in 6.3.1, the majority of the machinery used for cargo handling are using internal combustion engines running on diesel fuel. Calculations of the emissions of these engines will be based on the emission factors of diesel engines utilizing WP4 models. Furthermore, other types of fuel consumption in ports is done through generators, boilers, etc. that are mostly running on diesel fuel. The main pollutants considered for diesel fuel combustion are:

- Nitrogen Oxide (NO_x)
- Carbon Monoxide (CO)
- Carbon Dioxide (CO_2)
- Methane (CH_4)
- Sulphur Dioxide (SO_2)
- Non-methane volatile organic compounds (NMVOC)
- Particulate Matter 10 micrometers or less in diameter (PM_{10})

7. Data collection methodology and links to WP6

One of the main value propositions behind PEI is that it is being based on quantitative information. Which means, data coming from direct measurements of port activities. Unlike other approaches (as it were explained in Deliverable 5.1), the PEI aims at being fed by information collected from heterogeneous data sources. While this is conceptually crystal-clear, at the moment of implementation it must be based on a robust technological structure in order to be calculated properly. This, altogether with the environmental science rationale underlying the calculations will deliver the single indicator that will become one of the most innovative outcomes of PIXEL.

In this section, the technological structure to support the PEI is explained. Besides, the relation between the PEI calculation and its place within the whole PIXEL infrastructure is thoroughly described. Furthermore, the necessary inputs and outputs from the work on WP5 towards WP6 is depicted and forwarded through this text. The aim is to be able to provide all needed details to the most technical teams in the project (designers, developers, and integrators) in order to completely embed PEI in the whole PIXEL solution.

According to this purpose, the section is structured in a top-down approach. First, the calculation of PEI is overviewed and contextualized. Then, the section explains the ICT infrastructure definition and where and how PEI fits into the PIXEL architecture. Afterward, a series of subsections have been envisaged to include information on the data sources that will be able to introduce data to the PEI computing, their identification, selection and categorization and how they will be managed in PIXEL. The different steps for merging all PEI logic into a logically executable code are as well commented. Finally, some hints on the interpretation of the PEI, the different types of representation and their meanings are analyzed.

In the section 9, the visualization of the PEI will be detailed to complete the technological perspective of the WP5 product included in the global PIXEL framework.

7.1. PEI as a model in PIXEL

The PEI is considered a service for a PIXEL user. Here, we understand service as a set of useful functions provided to the user through an interface drawing from data processing and backend execution.

PIXEL has already defined the modular structure to perform all feature covered by the solution. Apart from data visualization, security through authenticated access and a global dashboard, PIXEL bases its functioning on this concept of service. The idea behind the “services” approach is to provide flexibility and scalability to ports in order to tackle different issues to be potentially solved by PIXEL.

These services will be, in the project, different models that the ports participating in the Consortium have requested in order to test, validate, assess and take advantage of PIXEL using real data and expecting real impact. Some examples of these models are energy consumption and production simulation, road traffic prediction, modeling of the air pollution emitted to the atmosphere, different intermodal and synchromodal data analytics and supply chain modeling. PEI, being based on data collection, imputation, processing, and visualization, must be therefore considered a service equal (from a systematic and software perspective) to the aforementioned.

More information about modeling, services and their introduction on the PIXEL framework can be found at deliverable D4.1 and D4.2 (to be delivered at the same time than D5.2). The description of the models in those documents are useful to understand why the PEI needs to be considered a service comparable to different modeling capabilities that PIXEL will include.

Regarding the integration, PIXEL has devoted a full work package (WP7) for integrating data, models, and services in specific physical implementations to be deployed at the 4 pilot ports. Later, another work package (WP8) will deal with impact assessment and technical and business validation.

Therefore, the PEI is (technologically) considered by WP6, WP7, and WP8 as a “model”.

According to D4.1, PEI will, therefore, need to provide a packed piece of software to be run by the Operational Tools. The idea of Operational Tools and where is PEI exactly incorporated in the PIXEL architecture is detailed in the next sub-section.

7.2. PEI position within PIXEL ICT infrastructure

The global PIXEL architecture was defined in deliverable D6.1 and it has been established and technologically advanced through deliverable D6.3. Despite not being the focus of this text nor this WP, the components composing the reference architecture of PIXEL are described below.

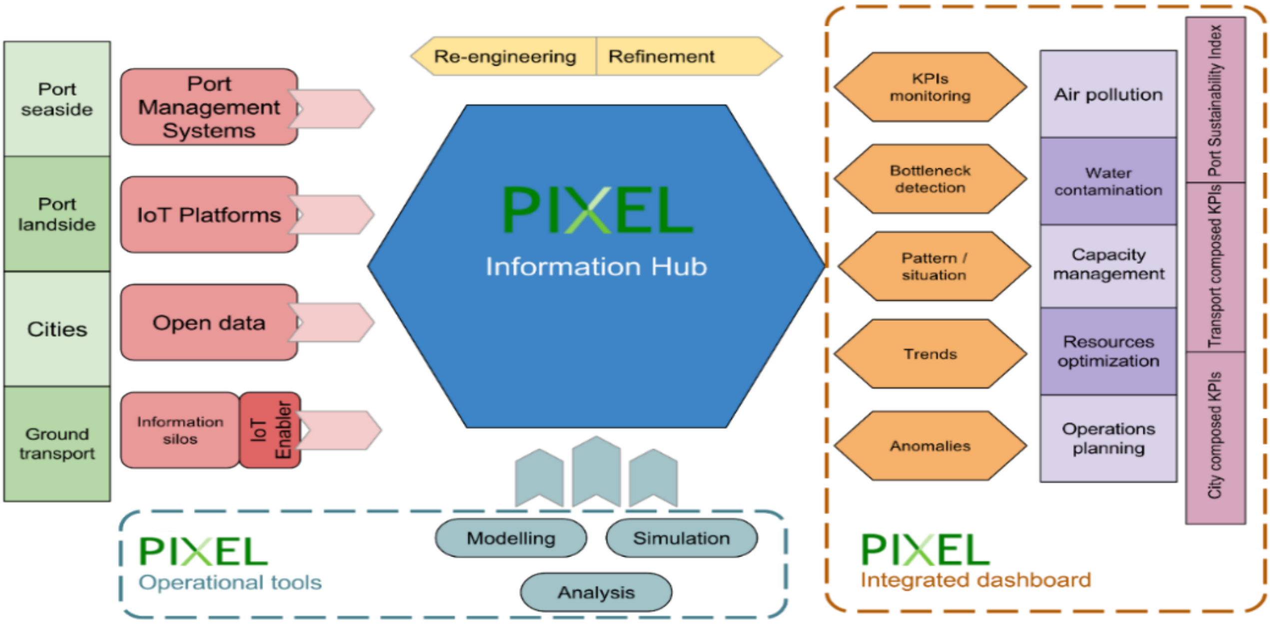


Figure 7.1 Reference basis for PIXEL architecture

Data acquisition (DAL). This module aims at gathering every relevant data of port activities and forwarding it to upper levels of the architecture. The module has been designed and is being developed to collect data from heterogeneous sources. For each data source to be included in PIXEL, there is a methodology established that includes gathering the data from the source in raw format and upstream it in a formalized data model agreed throughout all the platform. This inner component forms the foundation of data management in PIXEL

Information Hub (IH). It is the core element of the architecture, as it constitutes the sink where the different information siloes discharge and store their real-time data. A key innovation potential of this component is underpinned by the fast development of the Internet of Things (IoT) in logistics, environmental and wellbeing sectors. The basic concept behind the Information Hub is its capacity for long-term storage and its role of centralizing element: serving the data to whichever other module needing it.

Operational Tools (OT) are created and executed for applying simulations and run services extracting data from the Information Hub. Thus, the role of this module is to bring closer to the user the predictive algorithms and simulation models laying behind the transport-related innovation developed in the project.

Integrated Dashboard. The top modular component m for monitoring features, KPIs tracking, time evolution of parameters, historical data, reports, forecasting and other capabilities provided by previous layers.

Security. Transversally to the other modules, there is a crucial action implemented in the project for ensuring security and sovereignty of the data throughout all the layers.

Realizing the role of each component, in calculation we depict the execution flow that will take place in the system each time the PEI is requested to be calculated. Conceptually, the Operational Tools (OT) module will act as the orchestrator providing the intelligence to the procedure. OT will analyze the input data required by the PEI software, OT will retrieve it from the proper location in the IH. Afterward, the PEI module must be run in a SaaS (Software as a Service) fashion and its outcome (in a specific format) will also be collected by the OT to feed the Dashboard and particular PEI visualization. The four data items represent the different ways of piece of data can be introduced into the system. Ideally, all data should come from automated sources (sensors and external services), but assuming ports are not ideal, we contemplate alternative data imputation methods. This very topic will be further explained in section 7.4.

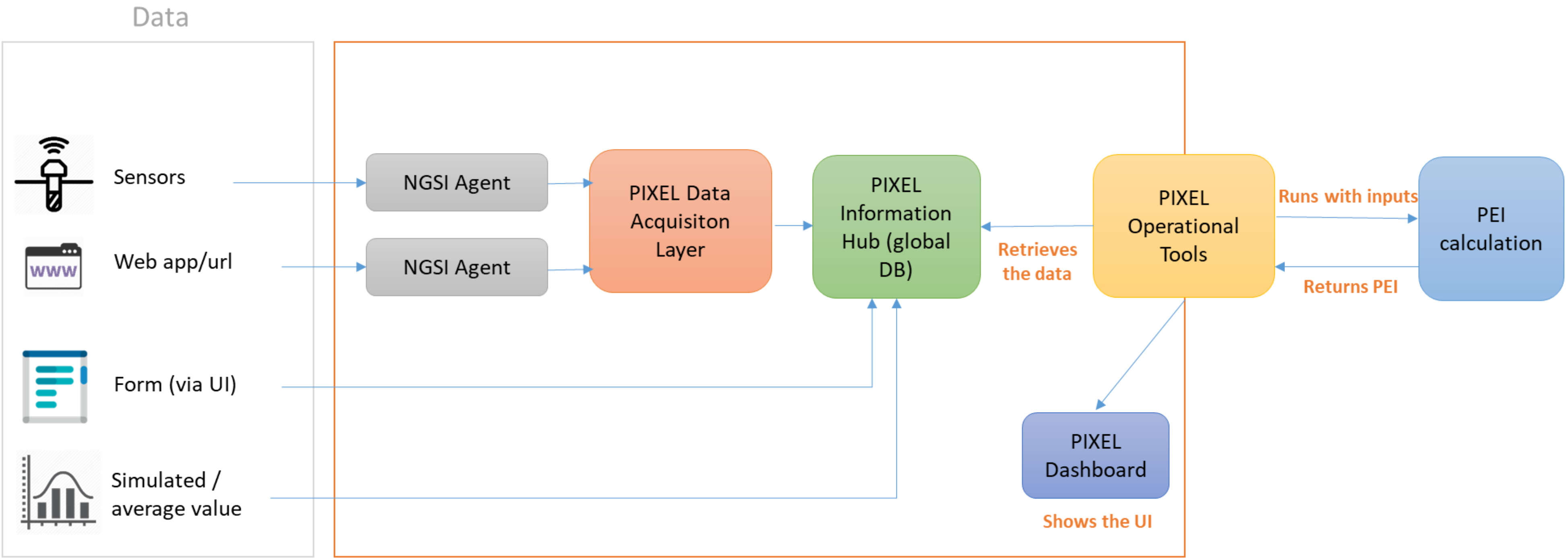


Figure 7.2 Execution flow for PEI calculation

Once the global placement of PIXEL calculation has been presented, the role and composition of the "**PEI calculation**" box in the diagram can be analyzed in more detail.

As mentioned above, PEI must be "*provided as a packed piece of software to be run by the Operational Tools*". This means that it is a duty of WP5 team to develop a program able to transform the conceptual environmental calculations into actionable instructions materialized as code. After this is achieved, some technological actions will need to be performed by WP6 team in order to convert the program into a service that will be hosted in the PIXEL infrastructure and that will be called by the Operational Tools in a classic REST API approach.

Summarising, the composition of the PEI calculation module needs the following:

1. Software to perform mathematical operations to certain inputs and to come out with certain outputs. In this point, it must be assumed that all data will be available and only inner operations must be addressed. The technological tools and languages for this program will be decided during T5.3.

WP5 must provide a "product" compliant to the previous consideration. The next two items will be addressed by WP6.

2. The program must be embedded inside a wider frame. This frame must be prepared to be stored as a server and to be called via API. The design of this API is also task of this WP5.
3. The frame will be containerized for real implementation and deployment. Virtualization and container tools such as Docker will be used in this step.
4. .

The following *Figure 7.3* helps to comprehend the design patterns that the PEI calculation module must follow.

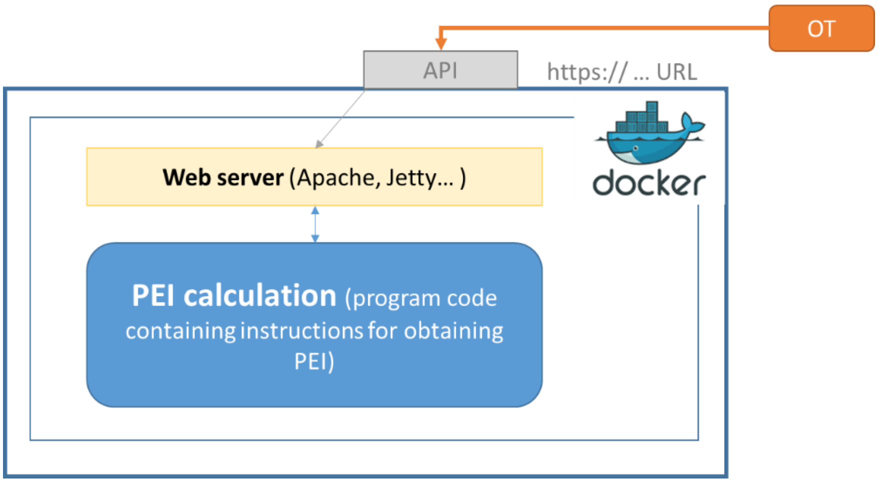


Figure 7.3 PEI as a packed service for execution

Finally, according to the previous, we have compiled a list of points explaining the information that must be provided to the team developing and integrating the whole ICT solution (WP6 and WP7). Better, in the list below, there are the pieces that need to be handed to WP6 and WP7 from WP5 team in order to effectively run the PEI (**as a model**).

1- *To inform T6.2 about which NGSI agents need to be developed for the PEI depending on which sensors/data sources will feed its calculation*

2- *To specify inputs and outputs (JSON) for running the PEI by the Operational Tools*

3- *To design the UI through which the user needs to visualize the information*

4- *To develop and make available the piece of software that calculates the PEI*

The previous steps will be completed during the course of the Task 5.3 that finishes on M24. Deliverable 5.3 will gather all technical specifications and software artefacts needed in WP6 and WP7 for the integration of PEI calculation in PIXEL ICT framework.

7.3. Data for calculating the PEI

Considering the technological foundations set, in this section we make a profound description on one of the key components in PEI: the environmental and operational **data** feeding the calculations. PEI aims at being based on data, therefore a solid strategy and methodology for its acquisition and introduction in the system is needed. Furthermore, establishing certain criteria and distinctions will provide flexibility to the PEI and a clear orientation for its adoption by small and medium ports in Europe.

In tasks T5.1 and T5.2, the team in WP5 launched certain activities in order to understand which data is needed to calculate PEI. First, drawing from previous research and from the environmental expertise by several members, the most important environmental factors in the ports were discovered. Later, a series of interviews, questionnaires, and surveys were fulfilled by the ports in PIXEL consortium. Finally, all the knowledge gained was compiled to elaborate a set of environmental factors that needed to be considered by the PEI. Afterward, in T5.2 (and in compliance with the methodology settled in D5.1), the WP5 team depicted a complex list of **minimum data requirements** that all ports should comply with, in order for their PEI to be calculated. This list was, to summarise, a reference document that has served as the basis for the PEI construction from that moment on.

The idea in task T5.3 is to transform from the minimum data requirements document to **actionable parameters and values to be populated** each time the PEI needs to be calculated for a port. The data needed for its calculation comes from the work done in T5.1 and T5.2. These data will later be proxied and operated to obtain the eKPIs (section 5). This is further explained in section 7.5. Technologically, the process is the one reckoned in 7.2: data will be collected through DAL, stored by IH and used internally in the calculations helped by the Operational Tools.

The *Figure 7.4* shows the structure of the compiled document of minimum data required, simplifying the concept to ease the data collection methodology (section 7.4). This will be analyzed later, but for now the reader must assume that all ports willing to be evaluated by this methodology must commit to providing a reliable value to each single piece of data in one way or another.

Ideally, all data will come from automated and precise sources such as sensors, legacy systems with proper integration or from updated databases or remote services. However, for flexibility and small-medium ports orientation, we must contemplate that not all the ports will have all data obtained in an automatic and sensor-based fashion. Hence, a hybrid approach must be considered. In PIXEL, we want to create a methodology looking at **diverse levels of digitalization** in a port. Usually, bigger ports with more resources are currently at a superior level of digitalization and automation of data acquisition than the small and medium ones. PEI aims at being a flexible and scalable tool able to be acquired and adopted by any port in Europe and abroad.

Therefore, we have envisioned 5 different types of data sources that can feed PEI calculation. These come from the 4-data source type in *Figure 7.2,* but threshing the web services in external web services and the Port Activity Scenario utility (being developed in the scope of T4.1 in PIXEL):

1 – Sensors

2.a -Web services

2.b – PAS

4 – Forms

5 – Averaged/extracted from valid references

Each data source type has its own characteristics and will be evaluated while calculating the PEI in order to obtain a reliability rating of the PEI each time it is calculated. In the five tables below (*Table 7.1* - *Table* 7.5), the conceptual foundations for that evaluations are set. During the execution of tasks T5.3 and T5.4, this reliability rating associated to the type of data source being used will be polished and put down in a formal way.

Table 7.1 Data source type: sensors

Type of data source	Sensors
Valid examples	Noise measurements, light sensor, cameras providing dwell time duration…
Strengths	
• Real-time: Using sensors, the most updated data will be in the IH at every moment, considering that the DAL register of the data source will be refreshing each few seconds (or even less, depending of the particular sensor). This guarantees high level of IoT-orientation for the PEI. However, there are data that are not prone to be collected by sensors, thus real-time cannot be an evaluation feature in those cases. • Precision: Commercial sensors usually come with very detailed datasheet and specifications. Ports should be able to analyze the precision needed to obtain a reliable eKPI associated (directly or proxied) from that data source. Some data are more sensitive of precision while others might not be very critical or its relative weight contribution to the PEI might not be enough to require high precision. • Automation: This is the main strength of this type of data source. Once the sensor is installed and technicians have been able to integrate it in the DAL, the process should be transparent for the PEI owner and final users. Besides, automated values cannot be cheated if the proper security and methods are applied, thus the final result of the PEI will be sensibly more reliable.	
Weaknesses	
• Cost: Industry-validated sensors for port and maritime sector are, normally, quite expensive. Despite the efforts in this WP are focused on looking for mature-enough close-to-open solutions, the devices commercially available can reach considerable amounts of money. • Installation: Once the sensor is acquired, a non-trivial effort must be exerted by technicians for the installation. On-site sessions, availability tests, maintenance, and support are necessary if the PIXEL owner wants the data source to be reliable in the PEI calculation. Again, these tasks must be performed by specialized staff that is not always present in the less-resourced ports. • Availability: Even with good installation, physical sensors are always susceptible of service cuts, battery problems, difficult network reach, dynamic range errors, false measurements or damage suffering from climate conditions or unexpected events. These make the sensors a non-reliable source if certain contingency measures are not taken. • Lack of knowledge: Specialised staff is not always available in a port. Besides, within PIXEL the heterogeneous team conforming the Consortium is covering just specific knowledge domains constrained to the project objectives. Experienced workgroup should be created in each port for a proper coverage of the full spectrum of sensors.	

Table 7.2 Data source type: web services

Type of data source	Web services
Valid examples	Vessel calls, Port Community Systems information…
Strengths	
• Close-to-real-time: Usually, web services (conceiving them under the classic REST API approach) are characterized by high availability and low response time. Typically, network delays, database consultation time and packet forwarding take milliseconds, henceforth the information is provided to the requester in a close-to-real-time fashion. • Automation: As long as the external remote web service is running, it suffices to run a http client in PEI's calculation side to have an automated execution environment. Automation is, same than for sensors, the most important feature that is looked for in the PEI context. This approach can be even better than collection by sensors in some cases where descriptive/complex information is needed, such as in the vessel calls for one week, or for knowing legal limits compliance.	

• Seamless installation: While sensors need time and efforts to be properly installed, deployed and integrated, external web services just need an HTTP connection and the right credentials to access the information. This is a major advantage of this type of data source.
Weaknesses
• Heterogeneity (not standardized): There is not a standardized methodology or consolidated reference on how to access external web services. This is a weakness on this approach considering that each port may need to acquire data from lots of different external raw sources, and no global standard easy-to-create and easy-to-employ technological module can be created. • Dependant on port policies: According to the experience gained in PIXEL, access policies to external and internal services in the ports are quite heterogeneous and need to be analyzed one by one. In some cases, data like vessel calls is property only of the Port Authority and it will not provide external access, in others the terminal is not willing to unveil information of tonnage or goods. • Restrictions of access: Sometimes, the security policies in the port do not permit the entrance of external data, or it must pass through certain software controls. This issue affects the web service data source type as long as the data is not managed by the port itself.

Table 7.3 Data source type: PAS

Type of data source	**PAS**
Valid examples	CO_2 emission factor drawing from Supply Chain definition (task T4.1 tool)
Strengths	
• Automation: The main strength of using the Port Activity Scenario tool is that it can be considered automated. Once the data sources specified as needed (check deliverable D4.2) and the supply chain definition is done, several eKPIs (or data items, depending the case) can be directly extracted from the PAS module execution. • Knowledge of procedure: PAS module is an innovative tool developed within PIXEL. As the most important outcome of task T4.1, this tool is able to provide values such as emission factor, energy consumption or machinery-associated indicators from certain data sources and modeling. WP4 team is cooperating with the rest of WPs, thus the knowledge of its functioning can be leveraged for the PEI.	
Weaknesses	
• Not validated: Being a tool in an early stage of maturity, it has not been validated or assessed in real environments. Therefore, its consistency and accuracy cannot be ensured. This will be a weakness for this kind of data source at the reliability rating evaluation phase. • Dependant on supply chain description: If one port is not able to describe thoroughly its supply chain or if lacks detail (for instance, in machinery specification, timeframes, loading duration, etc.), the error will be dragged through the stages of PEI computation. Furthermore, this is also dependant on the willingness of all agents within the port (terminals, P.A.s, stevedores, etc.) to provide information on their equipment and processes. • Hierarchical dependency on data: As it has been said, the PAS calculation relies both on supply chain definition and on different data (such as vessel calls, previous consumptions, etc.). If this data is not provided or it is defective, the eKPIs generated out of the PAS could be corrupted. • Additional computation needed: Despite considering automated, the truth is that obtaining factors out of PAS will require the Operational Tools doing additional computations. The intelligence of the system should run the PAS processing, adapting the output and preparing the software context before running the PEI module calculation. This may take time (to be analyzed in WP7 and WP8) that could affect accuracy, closeness-to-real-time and seamless visualization towards the final user.	

Table 7.4 Data source type: Forms

Type of data source	Forms
Valid examples	UI web form for knowing average dwell time of a train inside the port…
Strengths	
• Ad-hoc for the user: One of the advantages of using forms for data population is their capacity to be adapted to the final user. Through forms, WP5 team can implement certain techniques such as data boundaries and constraints, data quality check and written information helping uses to fulfill them. • Less availability-failure-prone: Being a piece of software utility developed by PIXEL members in our own environment makes it less prone to availability issues. Furthermore, web forms is a very commonly used technology and it will not mean a problem from the development and integration viewpoint. • Replicable: Once one form has been developed and approved by the ports within WP5 work, it can be replicated with slight modifications for populating other data in PEI. Additionally, changes can be done in a very easy way, so the implementation and adoption of PEI in other ports should not mean a huge effort if the data is to be provided via forms.	
Weaknesses	
• Human factor: The biggest shortcoming of this data imputation type is the human intervention. Obviously, an automated indicator should not emanate from human-typed data. People make mistakes when introducing information, it is not automatically cross-checked and there can be gaps for misuse. However, at some stages of PEI adoption, for certain data, it might be necessary/useful to bring PEI owner the possibility of manual imputation. Also, the data might be manipulated to yield better results • Intervention required: The PEI will not be automatically calculated. It will need (in a previous step or at the moment that it is requested) the intervention from port staff (environmental manager, data owner or even IT department) by the manual introduction of data. This might slow down the process and create diverse pitfalls deriving into not-very reliable single PEI indicator. • Transparent to the Context Broker: The information coming from forms will be stored at the IH just after being typed and accepted by the system. The Data Acquisition Layer will not be aware of its existence, thus no other modules could learn and extract data in an automated way of those forms. • Less control of data refresh: Data might be totally outdated in the IH due to the fact of not being processed by the DAL. Furthermore, it might happen that the proper staff to fulfill the forms is not available at the moment of PEI calculation request, hence a not-reliable value would be obtained. • Needs control flags and data pre-analysis: In order to avoid mistakes, false positives, misimputation, and constraints-compliance, a pre-analysis is needed to set data quality check strategy at the moment of forms completion. This would mean intensive efforts both for WP5 team and for the port personnel for deployment and integration. Besides, it augments the inter-dependency between WP5 and WP6.	

Table 7.5 Data source type: Forms

Type of data source	Averaged/extracted from valid references
Valid examples	Average of total garbage landed from ships, fuel type…
Strengths	
• Good for MVP of PEI: Imputing estimated data is a good approach for the first and middle stages of task T5.3. It will allow the WP5 team to have an intermediate Minimum Valuable Product to work over in regards, especially, to the normalization, weighting, mathematical operations and visualization. Besides, being able to crystallize the conceptual framework into a showable tool (initially based on these simulated values) will be good for ports and technical team to polish all representation aspects. • Good for debugging: Once an initial product has been created, it will be easier for technical team to understand the failures in the technological approach and to iteratively deliver new versions of already	

packed PEI calculation module. This debugging will be useful as it will always work with the same technical skeleton (same stages, same inputs, same outputs) just introducing new knowledge (normalization and weighting methods, for example). • Zero availability-failure-prone: Using average data or reference values extracted from the valid sources (EC reports, previous summaries on ports performance, past data of the same port) will ensure full availability of data in each single item. • Replicable: Same approach of average data selection (see previous item) can be applied to different to maintain the same level of reliability rating. This is considered replicable as this method can be executed by any port regardless its size and resources availability.
Weaknesses
• Lack of accuracy: No matter how trustworthy the selected data source might be, it will not be based on current measurements or actual input, therefore the reliability of the value needs to be put into question. This is one of the main disadvantages of using average or estimated data coming from experience. • Zero digitalization: No digitalization, development nor integration process is involved in this approach. The composite indicator generated out of average data would not depend on any value-generator component (gained by PIXEL). • Not IoT: Data will be plainly stored by hand into the IH without any DAL validation. The information will not comply with any of the IoT concepts. • Requires knowledge: As well as for the sensors installation, the port indicating reference values for their data for PEI need to have large knowledge of the sector. Not only on environmental affairs but on machinery, energy, and global functioning of the port.

7.4. Data collection methodology

Drawing from the previous section, we can create the following corollary:

“Technologically, data stored on the IH will be used by the Operational Tools. This data must have been inserted into PIXEL environment via NGSI agents on the Data Acquisition. Alternatively, data coming from not-automated sources will come directly from the IH without having passed through the DAL”.

Also, after a thorough read of section 7.1, we can summarise:

“PEI must be flexible. PEI must be able to be peer-compared. PEI is different for each port. PEI must be a traceable value that should be improving with time”.

Furthermore, we understand in 7.3 that a sheet of data needed to collect has been already provided:

“*The Figure 7.4 shows the structure of the compiled document of minimum data required, simplifying the concept to ease the data collection methodology*”

Therefore, the only pending job to glue everything with regards to data collection **is to identify which type of data source is appointed to each piece of data needed for PEI**.

For that, in WP5 we have done a 3-fold data source identification approach:

- Ideal data source for each piece of data: this is the “academic” view of the data origin. This field has been completed based on the opinion of experts in environmental and ICT fields within the Consortium.
- Industry average (commercially available) data source: Maximum currently deployable data source type associated to each piece of data needed in PEI. This has entailed work of research in this WP.
- Current availability on PIXEL project ports of each data: Completed by the WP5 team in collaboration with other WPs (taking advantage of previous works like deliverable D3.4) and with the ports.

For this, the tool that has been used **is a spreadsheet** that compiles the minimum data required for PEI calculation as rows. The different identifiers were set as columns that are going to be filled with the collaboration of environmental-specialized, business-oriented and technologic-specialised partners of PIXEL Consortium

<u>Use of the data source appointment datasheet in WP5</u>:

The aim of the Excel file created is, once again, to identify how data can/will be collected.

The data sources will help obtaining eKPIs, as it has been set following the indications in section 5, created by the environmental experts in the project.

Regarding **HOW** each data **COULD BE COLLECTED**, we are using two columns: (i) the *IDEAL D.S.* column related to the **ideal data source** through which the information could be obtained and (ii) the *MAX. D.S.* column related to the feasible data source that one port (considering the state of the art) could have access to nowadays[3]. In this second case, if there is not enough information about which technologies could be available or not, we will use the industry average approach. This will consist on analyzing the most common equipment/method use by ports in the real world and use it as the maximum feasible reference.

Finally, each port using PEI must identify **HOW** the data **IS REALLY BEING COLLECTED** at that moment. This will complete the 3rd important column. This table must be completed for each port using PEI.

Here below is the structure that has been used:

:

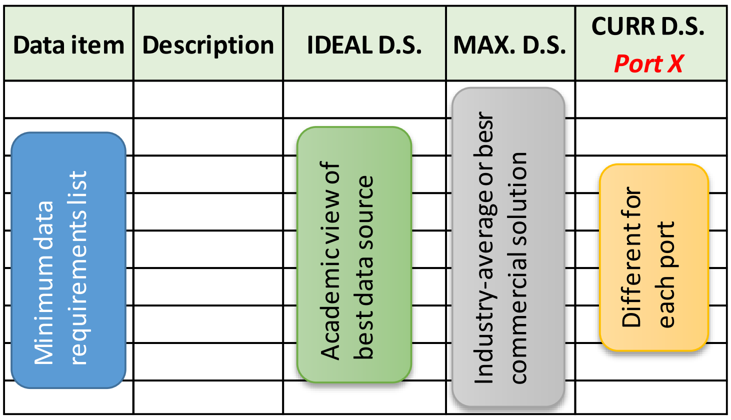


Figure 7.4 Data collection tool structure

The full result of this data collection methodology **will be provided in the next document of the WP5 (deliverable D5.3).**

To clarify, an example of two full rows completed would be the following:

Table 7.6 Example of data collection

Data item	Description	Ideal DS	Max. practical DS	Current DS
Noise	Measurement in dB	Sensor	Sensor	*Depending the port*
Ship type	Text (type of vessel)	Web service	Web service (vessel calls)	*Depending the port*
Truck w. time	Working time (h)	Sensor	Web service	*Depending the port*

The information contained in the outcome of this piece of work will be used three-fold:

1. For the reliability rating evaluation
2. To inform WP6 about the types of data sources needed to be implemented in the DAL in each pilot port in PIXEL (*Current data source column*).
3. To provide a value for weighting and normalization in the first step of the PEI calculation (this is extended in section 7.5).

[3] Here we are referring to current ways that ports are currently using to measure certain things. E.g.: Port of Rijeka – Check paper with DOI: 10.1007/s10661-018-6965-z

7.5. Reliability rating and process summary

In the *Figure 7.5*, an overall view of the process is depicted. In this sub-section each step is analyzed and properly referred:

Mapping according to availability

This step has already been realized and it has consisted on fulfilling the datasheet described in 7.4. The result of that research work is having the means for data-sources benchmarking. Knowing where a port is positioned in regard to digitalization compared to the ideal PEI and the industry average technology adoption will be a main asset to exploit in T5.4 and T5.5. As a matter of fact, this sheet will be the basis for the reliability rating calculation and for the PEI comparison.

Reliability rating calculation

The PEI is a single indicator obtained after a series of calculations over certain incoming data. The values are retrieved every time PEI is run but not always with the same type of data sources as origin. Sensors might not be available; the port might be scarce in resources or different options of fulfillment might exist. This flexibility is crystallized and evaluated in the PEI with the so-called **Reliability Rating (RR)**.

The rationale behind this concept is to have a tool for informing that a PEI number must be understood altogether with a **trustworthiness margin**. It is not the same when one piece of data is retrieved from an on-site sensor than when it is taken from past data or literature references. In this section, we are explaining the methodology that will be followed in PIXEL to obtain and attach a RR to a single PEI calculated.

The foundations of this RR calculation are the following:

- Each piece of data has got an "optimal way of retrieval". The idea is to calculate "how far" is the current mechanism from this optimal way.
- The "optimal way" will be the one coming from the analysis explained in 7.4 (column *industry average*).
- As thoroughly described through tables in section 7.3, the different data source types have both strengths and weaknesses. The reliability rating aims at formalizing this **"evaluation"** only from a technological point of view.
    - This is because we conceive the PEI as IoT and ICT-based. Therefore, we want to create a methodology to rate how close is one way of data retrieval in a port to be the most technologically-oriented.
- Thus, a set of **Evaluation Features (EF)** are identified in order to classify "how close" is the current data from the optimal. Those indicators are hybrid: technical (such as accuracy on the measurement, real-time, availability) and operational (such as compliance with regulations, failure-prone, human intervention). These features are selected by PIXEL partners participating in WP5.
    - These EF will be different per each conceptual origin of data. This means respecting the environmental experts' separation of the data (section 5.2, deliverable D5.1 and Minimum data requirements sheet). Henceforth, we will elaborate **three** sets of evaluation features (one per data coming from ships, another for data coming from Terminals and another related to Port Authorities' information).
- The "optimal way" is cataloged as 100%, and the other type of data retrieval options will be diminished in percentile depending on its rating obtained from the feature's evaluation. This means that the ideal data source will have a 100% and, depending on the data source used by the port, it will be decreased in certain margin trying to show how close it is from the ideal.
- The PEI will be accompanied by 3 percentiles. The total RR associated to data sources related to (i) ships, (ii) terminals and (iii) port authorities. This is aligned with the environmental separation and with the different EFs to be used.
- In a first stage, the evaluation will be based on a **reduced set of features**. These features and its associated values compared to the "optimal way" will be fulfilled following the ***budget allocation***

weighting method (see section 8). This **reduced set** will be **completed** for the data sources **by experts** on technological and operational fields **within PIXEL Consortium**.

- In a second stage, the evaluation will rely on an extended set of features. These features and its associated values compared to the "optimal way" will be fulfilled following the ***budget allocation weighting method*** (see section 8). This **extended set** will be **completed** for the data sources **by experts** on technological and operational fields **external to PIXEL Consortium**.

In the table below, there is an extract of a hypothetical RR calculation. As it is observed, every data source is analyzed and a RR is associated. Then, the percentiles are N for ships, terminals, and port authorities. This would correspond to the X%, Y% and Z% values in *Figure 7.5*.

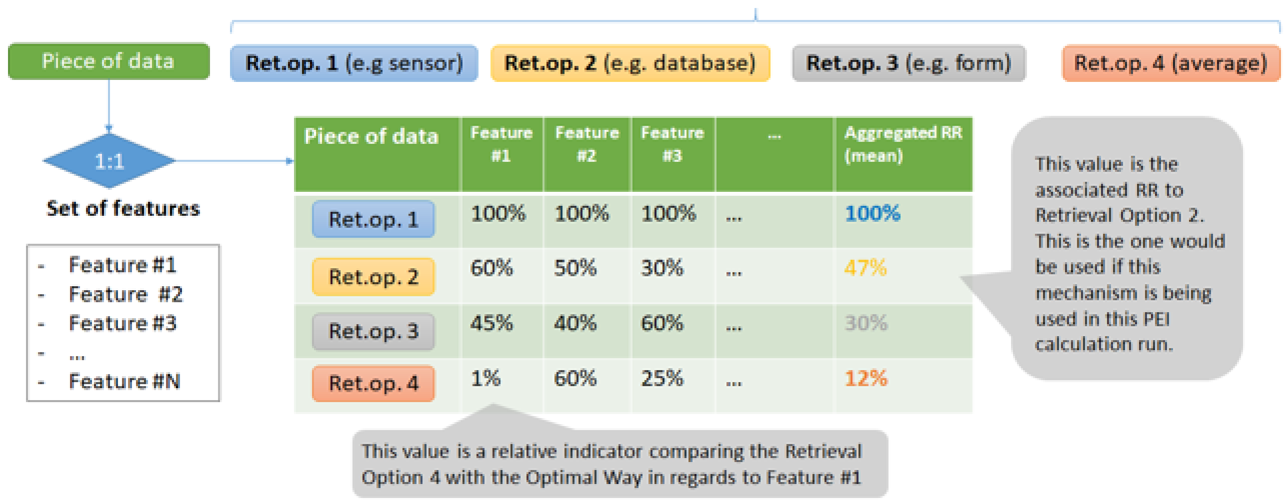
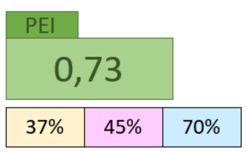


Figure 7.5 Reliability rating calculation of single piece of data overview

Table 7.7 Reliability rating calculation of single piece of data overview

Data origin	**Data**	**Optimal way**	**Current way**	**Reliability rating**	**Aggregated RR**
Ships	…	..	…	- %	
Ships	Ship type	Web Service	Form	25%	37%
Ships	Main engine power (kW)	Datasheet	Datasheet	100%	
Terminal	…	…	…	- %	
Terminal	Electricity consumption	Sensor	Historical data	10 %	45%
Terminal	Amount of cargo	Web Service	FAL forms	80%	
…	…	…	…	- %	
Port authority	Total water consumption	Sensor	Web service	60%	70%
Port authority	Noise (dB)	Sensor	Sensor	100%	



With the previous result after RR calculation, the PEI would result as in the figure at left. A single indicator with a value between 0 a 1 accompanied by 3 sub-indices indicating the RR of the PEI considering the percentiles associated to ships data sources, terminals data sources and port authority data sources.

eKPIs, indices and PEI value construction

Later, the data retrieved from the IH, altogether with their associated RR, will pass through a set of operations (details in section 8) in order to obtain the final single value of PEI.

According to the PEI methodology (deliverable D5.1), the program will need to perform a 3-step mathematical operation procedure:

- **From data single items to eKPIs**: eKPIs are composed either by direct assignment from data items (1:1) or proxied from various data (N:1). This is detailed in section 6.
- **From eKPIs to indices**: Six indices have been selected: Air, water, noise, waste, odors and light. In this step, the operations will be aggregation, normalization and weighting.
- **From indices to PEI**: Finally, a single composite indicator will be created. PEI value will come from aggregation.

These steps are just included here for reference. A more detailed explanation on the meaning of each operation can be found in sections 5 and 8 of this document.

Joint representation for integral understanding

Finally, the result of PEI calculation and the RRs obtained will be shown to the user through a web interface. At this point, all previous considerations about comparison capabilities, closeness to ideal data source and the evolution of the PEI with the time will be depicted via graphic components.

PEI visualization means are detailed in the one of the following sections of this document (9). Data representation and the different options for visualization are analyzed and selected. Finally, a sketch design of the final interface is also provided. This point will be polished and made definitive during the execution of task *T5.4 - Guidelines for PEI adoption and recommendations.*

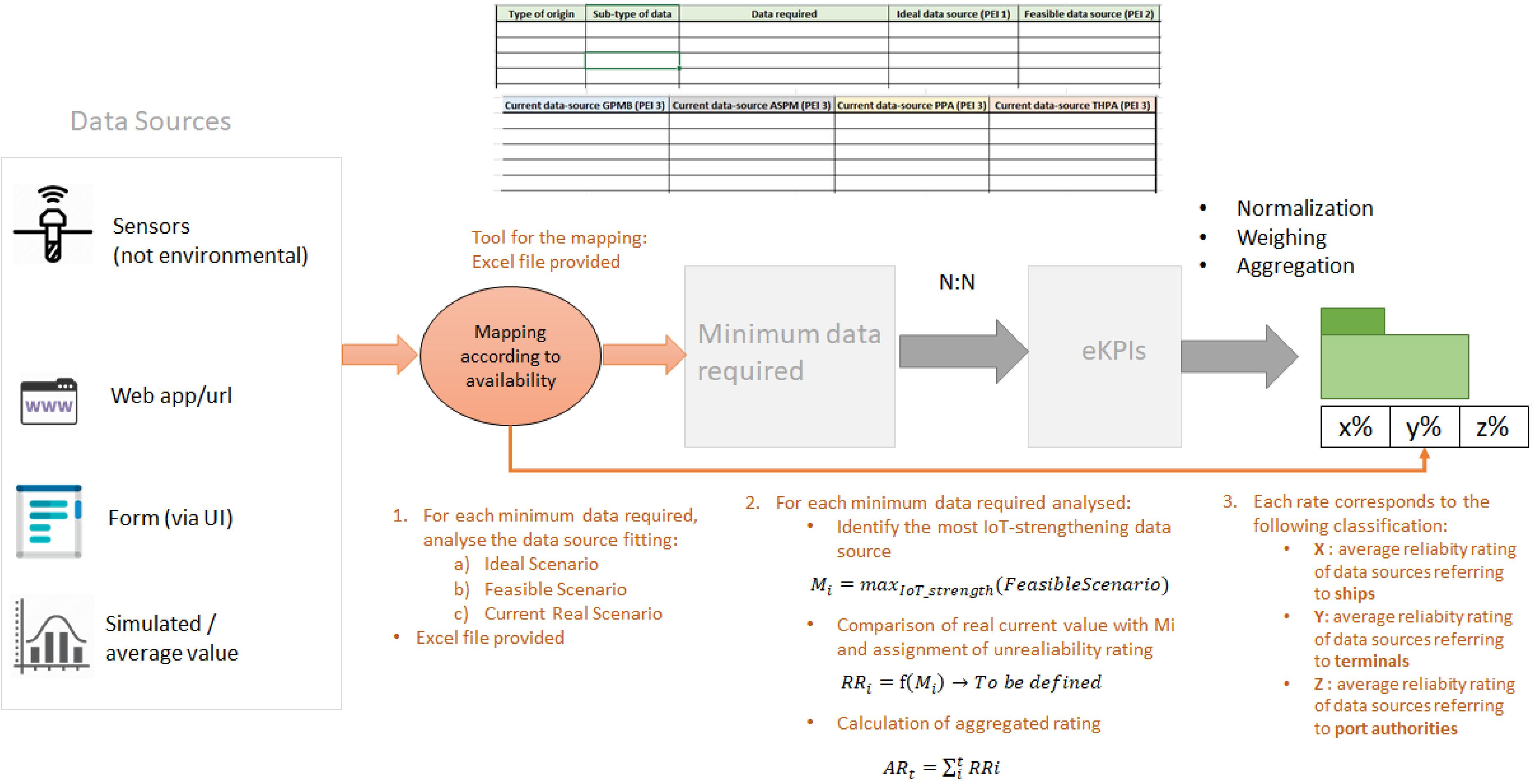


Figure 7.6 Reliability rating calculation

8. Toolbox for constructing the PEI

8.1. Introduction to composite indicator

A composite indicator is based on the mathematical combination (or aggregation) of individual indicators that represent different dimensions of a concept. According to the European Commission (2002): "composite indicators are based on sub-indicators that have no common meaningful unit of measurement and there is no obvious way of weighting these sub-indicators". In PIXEL, a list of environmental indicators (eKPIs) have been identified and the Port Environmental Index (PEI) aims to be a summary of those indicators. This list of environmental indicators has been described in section 5.

Composite indicators are more and more recognized as useful both for policy-making and communication. It is much easier to understand them than analyze many indicators at the same time. In many fields composite indicators have been used to provide useful information. However, if a composite indicator is not well constructed, it can send a misleading and non-robust message. A lot of studies in the literature describe the mathematical background for building a composite indicator with their pros and cons. In Sharpe (2004) we can read "The aggregators believe there are two major reason that there is value in combining indicators in some manner to produce a bottom line. They believe that such a summary statistic can indeed capture reality and is meaningful and that stressing the bottom line is extremely useful in garnering media interest and hence the attention of policymakers. The second school, the non-aggregators, believe one should stop once an appropriate set of indicators has been created and not go the further step of producing a composite index. Their key objection to aggregation is what they see as the arbitrary nature of weighting process by which the variables are combined". To complement this, we can add what is written in Saisana (2004): "it is hard to imagine that debate on the use of composite indicators will ever be settled […] official statisticians may tend to resent composite indicators, whereby a lot of work in data collection and editing is "wasted" or "hidden" behind a single number of dubious significance. On the other hand, the temptation of stakeholders and practitioners to summarise complex and sometimes elusive processes (e.g. sustainability, single market, etc.) into a single figure to benchmark country [or port in PIXEL context] performance for policy consumption seems likewise irresistible". Other description of the controversy of composite indicators can be read in Cherchye et al. (2007). There is a lot of literature related with composite indicator and new methodological approaches are being continuously published.

Based on Nardo et al. (2005) and Saisana (2004), *Table 8.1* sums up the pros and cons for the deployment of a composite indicator.

Table 8.1 Pros and cons of composite indicators

Pros	Cons
Summarise complex or multi-dimensional issues	May send misleading policy messages
Easier to interpret than trying to find a trend in many separate indicators	May invite drawing simplistic policy conclusion
Facilitate the tasks of ranking on complex issues	Selection of indicators and weights could be the target of political challenge
Assess progress over time on complex issues	May disguise serious failings in some dimension of the phenomenon.
Reduce the size of a set of indicators	Increase the quantity of data needed because data are required for all the sub-indicators.
Facilitate communication	Dragged error
Provide the big picture	Loss of information
Easier benchmarking	Hidden mistakes not realized

Many composite indices have been already developed in different fields: the Human Development Index (Jahan 2017), the Sustainable Society Index (Van de Kerk and Manuel 2008), the Financial Secrecy Index (Tax Justice Network 2013) and the Environmental Performance Index (Hsu et al. 2016). In PIXEL, we aim to develop and use a composite indicator in order to "model" the environmental impacts of ports to help decision making. The objective of the following section is to present a state of the art about "how to build a composite indicator" in order to identify the sources of subjective or imprecise assessment of the PEI. Readers have to keep in mind that PEI as a composite indicator will be a "presentation and comparison of performance" in port areas "to be used as starting points for further analysis and discussion (Saisana 2004).

Readers must also notice that there are a lot of challenges in composite indicators development as reported in Saltelli et al. (2007): "From a purely mathematical point of view, the aggregation convention used for composite indicators deal with the classical conflictual situation tackled in multi-criteria evaluation. Thus, the use of a multi-criterion framework for composite indicators in general, and for sustainability and well-being indices in particular, is relevant and desirable. […] However, the so-called "multi-criterion problem" can be solved by means of a variety of mathematical approaches, all of them plausible. This situation is due to Arrow's impossibility theorem (Arrow, 1963), which proves that it is impossible to develop a "perfect" mathematical aggregation convention. This implies that it is desirable to have mathematical algorithms that may be recommended in some respect to different aggregation procedures. This makes sensitivity analysis a fundamental step during the development of any composite indicator. Other authors like Munda and Nardo (2005) have highlighted the importance of the quality of the aggregation convention. This quality depends on the context and composite indicator uncertainties coming from both technical and socio-political issues.

In PIXEL, we will try to provide a confidence interval in the PEI in order to be transparent as possible on this uncertainty. In order to do so, it means that we must be able to know or estimate the uncertainty of the input data. Indeed, a good quality composite indicator value depends not only of the methodology used but primarily on the quality of the framework and the data used. Besides this, PEI will address very carefully the data sensitivity and quality not only with a confidence interval associated to "composite indicator" building but with a reliability rating of the data source that populates the raw values (see more in section 7).

8.2. Construction of a composite indicator

According to the Competence Centre on Composite Indicators and Scoreboards (COIN) (https://ec.europa.eu/jrc/en/coin/10-step-guide/overview) and to the Handbook on Constructing Composite Indicators (Joint Research Centre-European Commission 2008), the following steps have to be followed when a composite indicator is built. All these steps are iterative by nature.

1. Establish the theoretical framework: What is badly defined is likely to be badly measured
    a. Clear understanding and definition of the multidimensional phenomenon to be measured
    b. Discuss the added value of the composite indicator
    c. Nested structure of the various sub-groups of the phenomenon
    d. List of selection criteria for the underlying variables (input, output, process, etc.)
2. Select the data: The quality of composite indicator value depends largely on the quality of indicators
    a. Quality assessment of the available indicators
        i. Policy relevance: Can the indicator be associated with one or several issues around which key policies are formulated?
        ii. Simplicity: Can the information be presented in an easily understandable, appealing way to the target audience?
        iii. Validity: Is the indicator a true reflection of the facts? Were the data collected using scientifically defensible measurement techniques? Is the indicator verifiable and reproducible?

    iv. Time series data: Is time series data available, reflecting the trend of the indicator over time?
    v. Availability of affordable data: Is good quality data available at a reasonable cost or is it feasible to initiate a monitoring process that will make it available soon?
    vi. Sensitivity: Can the indicator detect a small change in the system?
    vii. Reliability: Will the same result be obtained by making two or more measurements of the same indicator? Would two different researchers arrive at the same conclusions?
  b. Discuss strengths and weakness of each selected indicator
  c. Summary table on data characteristics
    i. Availability (across port, time)
    ii. Source
    iii. Type (hard, soft or input, output, process)
    iv. Descriptive statistics (mean, median, skewness, min, max, variance, etc.)

3. Perform a correlation analysis: indicators are often chosen with little attention paid to the interrelationships between them, but this will directly impact the quality and the usefulness of the composite indicator
4. Data imputation: The idea of imputation is both seductive and dangerous
5. Preliminary data treatment:
    a. Confidence interval for each imputed value
    b. Discuss and treat outliers
    c. Make scale adjustments
6. Data normalization: Avoid mixing apples and pears
    a. Directional adjustment: higher values correspond to better performance
    b. Select a suitable normalization method that respects conceptual framework and the data properties
7. Data weighting and data aggregation: What matters more ...weights more...
    a. Discuss whether compensability among indicators should be allowed and up to which level of aggregation. This means that the ability of indicators with very low scores to be fully compensated for by indicators with high scores is limited.
    b. Discuss whether correlation among indicators should be considered during data weighting
    c. Select a suitable weighting and aggregation method
8. Robustness and sensitivity tests: Sensitivity analysis can dissipate some of the controversies surrounding composite indicators
    a. Consider different methodological paths to build the index and if any different conceptual frameworks
    b. Identify the source of uncertainty and provide the composite index with confidence intervals
    c. Conduct sensitivity analysis to show what sources of uncertainty are more influential
9. Back to the data: This step is useful in order to be sure to have an iterative process when building a composite indicator.
10. Visualization: A well-designed graph can speak louder than words…
    a. Identify suitable presentational tools
    b. Select visualization technique which communicates the most important information without hiding vital information
    c. Present the results in a clear, easy to interpret and accurate manner.

In this document, and based on the scientific literature, we define the full methodology to follow in PIXEL project in order to build the Port Environmental Index with a clear methodology and mathematical background (mainly based and adapted from Tax Justice Network (2013)).

Table 8.2 Steps for the construction of a composite indicator

Steps	**Objectives**	**Proposition of Responsible partners for building PEI**
Step 1: Theoretical Framework Provides the basis for the selection and combination of variables into a meaningful composite indicator (involvement of experts and stakeholders is envisaged at this step).	A clear understanding and definition of the multi-dimensional phenomenon to be measured. Definition of eKPIs. A nested structure of the various sub-groups of the phenomenon if needed. A list of selection criteria for the underlying variables, e.g. input, output, process. Clear documentation of the above.	MEDRI, CREOCEAN
Step 2: Selecting variables Based on the analytical soundness, measurability, and relevance of the indicators to the phenomenon being measured and related to each other. The use of proxy variables should be considered when data are scarce. For selecting variables experts and stakeholders can be involved.	Check the quality of the available indicators (eKPIs). Discuss the strengths and weaknesses of each selected indicator. Obtain or estimate confident value for each input variables or each eKPIs Create a summary table on data characteristics, e.g. availability (across port, time), source, type (hard, soft or input, output, process).	MEDRI, CREOCEAN
Step 3: Imputation of missing data Provides a complete dataset	Estimate missing data. A complete data set without missing values. A measure of the reliability of each imputed value so as to explore the impact of imputation on the composite indicator. Discuss the presence of outliers in the dataset Document and explain the selected imputation procedures and results.	CATIE, UPV, MEDRI, CREOCEAN
Step 4: Preliminary data treatment Studies the overall structure of the dataset, assess its suitability, and guide subsequent methodological choices (e.g., weighting, aggregation).	Check the underlying structure of the data along the main dimensions (e.g., principal components analysis, cluster analysis). Identify groups of indicators or groups of indices that are highly correlated and provide an interpretation of the results. Compare the statistically determined structure of the data set to the theoretical framework.	CATIE, UPV, MEDRI, CREOCEAN
Step 5: Normalization Renders the variables comparable	Select suitable normalization procedure(s) that respect both the theoretical framework and the data properties.	CATIE, UPV, MEDRI, CREOCEAN

	Discuss the presence of outliers in the dataset as they may become unintended benchmarks. Indeed, some normalization introduce a distortion effect since the extreme values (max and min) could be unreliable outliers. This is the case with re-scaling. Make scale adjustments, if necessary. Transform highly skewed indicators, if necessary.	
Step 6: Weighting and aggregation Done along the lines of the underlying theoretical framework.	Discuss whether correlation issues among indicators should be accounted for. Discuss whether compensability among indicators should be allowed. Compensability is the ability of indicators with very low scores to be fully compensated for by indicators with high scores. Select appropriate weighting and aggregation procedure(s) that respect both the theoretical framework and the data properties.	CATIE, UPV, MEDRI, CREOCEAN
Step 7: Uncertainty and sensitivity analysis Assess the robustness of the composite indicator in terms of e.g., the mechanism for including or excluding an indicator, the normalization scheme, the imputation of missing data, the choice of weights, the aggregation method.	Consider multi-modeling to build the composite indicator, and if available, alternative conceptual scenarios for the selection of the underlying indicators. Identify all possible sources of uncertainty in the development of the composite indicator and accompany the composite scores and ranks with estimates of uncertainty. Conduct sensitivity analysis of the inference (assumptions) and determine what sources of uncertainty are more influential in the scores and/or ranks.	CATIE, UPV, MEDRI, CREOCEAN
Step 8: Back to the data and links to other indicators Reveals the main drivers for an overall good or bad performance. Transparency is essential for good analysis and policymaking.	Check for correlation and causality (if possible). Identify if the composite indicator results are overly dominated by few indicators and to explain the relative importance of the sub-components of the composite indicator. Correlate the composite indicator with other relevant measures, taking into consideration the results of sensitivity analysis.	MEDRI, CREOCEAN
Step 9: Visualization of the results Should receive proper attention, given that the visualization can influence (or help to enhance) interpretability.	Identify a coherent set of presentational tools for the targeted audience. Select the visualization technique which communicates the most information. Present the composite indicator results in a clear and accurate manner.	UPV, MEDRI, CREOCEAN

This document focuses on the steps 3 to 7 and aims at providing the state-of-the-art and the mathematical framework to build the PEI. Based on the theoretical framework about the PEI that are available in 2 and on the objectives of the PEI and PIXEL, we provide a first guidance and insights of how the PEI should be built from a mathematical and statistical point of view.

8.3. Mathematical methods for imputing missing data

8.3.1. Definition of a missing value

A missing value occurs when no observations can be made for a variable in the data set. Missing data can appear in a random or non-random way. According to literature, three main patterns can be identified for missing data:

- *Missing completely at random* (MCAR). A MCAR data is missing in a complete random way if the probability of missing is the same for all observations. This probability only depends on external parameters.
- *Missing at random* (MAR). A MAR data is missing at random if the probability of missing is dependent on other variables in the data set. For example, the missing values in CO_2 emissions would be MAR if the probability of missing data on CO_2 emissions depends on type of vessels but, within each category of vessel, the probability of missing CO_2 emissions is unrelated to the value of CO_2 emissions.
- *Not missing at random* (NMAR). The missing values depend on the values themselves. For example, high energy consumers are less likely to report their CO_2 emissions.

It is worth mentioning that:

1. It is not possible to know, only with the observed data, if the missing values are missing at random or not. A link with the context and the way the data are gathered is needed.
2. No imputation method is free of assumptions and we must check their statistical properties and impacts on the data set.
3. Imputed data should be defined with a related uncertainty reflected by variances estimates.
4. If data are NMAR, there is no model for imputing them.

8.3.2. Analysis of missing values in PIXEL

The first step to do in order to deal with missing data is to have a clear analysis of the structure of missing values. By analysis the data set of all the sub-indices that will be used to build the PEI and for the 4 ports of the PIXEL project, we must first identify:

- Which values are missing;
- How many values are missing;
- With which frequency the data are missing;
- If the data are missing completely at random, at random or not at random;
- A statistical description of missing values (mean, median, variance, …).

With this simple analysis, we will have qualified information to impute missing data. The objective is to be able to answer the following questions:

- What do we know about the mechanism behind the missing value?
- Do the missing values contain information?
- What happens if we ignore the missing values?

Knowing all of this, a method for imputing missing data will be chosen. In the next section, we provide a state-of-the-art about imputing missing data and we discuss the different methods that can be applied in the context of PEI and PIXEL project.

8.3.3. State of the art and description of the different approaches

Three main methods can be identified for dealing with missing data (according to Joint Research Centre-European Commission (2008)):

1. Case deletion: simply omits the missing records from the analysis. This approach ignores possible systematic differences between complete and incomplete data set. It produces unbiased results only if we delete value with MCAR assumptions.

The two other approaches consider that the missing value are part of the analysis:

2. Single imputation: mean/median/mode substitution, regression imputation hot-and-cold-deck imputation, expectation-maximisation imputation.
3. Multiple imputation: Markov Chain, Monte Carlo algorithm.

8.3.3.1. Single imputation

The objective with single imputation is to impute value using a predictive distribution that is generated by using observed data trough implicit or explicit modeling as indicated by Little and Rubin (2019). An important limitation of this imputation approach is that it underestimates the variance of the imputed missing value. Then, the robustness of the composite index derived from the imputed data set is not assessed.

1. Implicit modeling

We use imputation algorithms based on underlying implicit assumptions that require to be checked if they are reasonable and appropriate to the issue under consideration. The main risk of this type of missing data modeling is the tendency to consider the resulting dataset as complete, forgetting that an imputation has been made.

1.1. Hot deck imputation

The missing value are completed using a “similar” value that is already registered in the data set. For example, missing values for CO_2 emissions of one vessel may be replaced with CO_2 emissions of another vessel with similar characteristics available in the internal database of port (e.g. engine, length, type of fuel, …).

1.2. Substitution

The missing value is replaced by a value that was not selected before. For example, the energy consumption of a specific building is replaced by the energy consumption of another building that what was not selected before.

1.3. Cold deck imputation

The missing value is replaced with a constant value from an external source. This value can come from literature, expert knowledge or similar studies. For example, a missing value for fuel consumption of crane in the GPMB can be replaced by the fuel consumption of a crane considering the manufacturer value.

2. Explicit modeling

For this type of imputation, the predictive distribution of the missing value is based on a formal statistical model where assumptions are made explicit.

2.1. Unconditional mean/median/mode

The missing value is replaced by the mean/median of recorded values. If a port is not able to provide its lighting consumption, we can use the mean of lighting consumption observed for the other ports. This approach leads to underestimation of the true variance of the missing value. Thus, the uncertainty in the PEI will be underestimated.

2.2. Regression

The missing value is replaced by the predicted value obtained after a regression process. The dependent variable of the regression is the sub-indicator hosting the missing value and the regressor(s) is(are) the sub-indicator(s) showing a strong relationship with the dependent variable (usually a high degree of correlation). For example, if a value of NO_x emissions is missing for truck but we know what the distance cover by the truck is, we can use distance cover as the regressor and NO_x emissions as the dependent variable since there is high degree of correlation between these variables.

2.3. Expectation Maximisation

This technique iteratively goes through the data while still preserving the covariance structure of the data. Each iteration consists of an E step and an M step. Assuming that data are MAR or MCAR, the E step finds the conditional expectation of the "missing" data, given the observed values and current estimates of the parameters. These expectations are then substituted for the "missing" data. In the M step, maximum likelihood estimates of the parameters are computed as though the missing data had been filled in. "Missing" is enclosed in quotation marks because the missing values are not being directly filled in. Instead, functions of them are used in the log-likelihood. According to Joint Research Centre-European Commission (2008): “the advantage of the EM is its broadness (it can be used for a broad range of problems, e.g. variance component estimation or factor analysis), its simplicity (EM algorithm are often easy to construct conceptually and practically), and that each step has a statistical interpretation and convergence is reliable. The main drawback is that in some cases, with a large fraction of missing information, convergence may be very slow. The user should also care that the maximum found is indeed a global maximum and not a local one.”

8.3.3.2. Multiple imputation

In the multiple imputation method, the imputation of missing data is done assuming a random process reflecting uncertainty. The objective is to create “n” complete data set doing “n” imputation. Any imputation method can be used in multiple imputation. According to Joint Research Centre-European Commission (2008), “Multiple Imputation method imputes several values (N) for each missing value (from the predictive distribution of the missing data), to represent the uncertainty about which values to impute. The N versions of completed data sets are analyzed by standard complete data methods and the results are combined using simple rules to yield single combined estimates (e.g., MSE, regression coefficients), standard errors, p-values, that formally incorporate missing data uncertainty“.

8.3.4. How to deal with missing data in PIXEL

Knowing the theoretical background of PEI and the difficulty to obtain data with associated uncertainty, we think that the following method will be suitable to be used in PIXEL (of course this can change after analysis of the missing data):

- Case deletion
- Hot deck imputation
- Substitution
- Cold deck imputation
- Unconditional mean/median/mode
- Regression

Moreover, as described in deliverable D4.1 “*Data is a key point in order to have useful models and data analysis and data collection is an intensive phase. Models and data analysis can be developed using different levels of details regarding data. Indeed, depending of the purpose of modeling and the expected precision, inputs data could be less or more detailed. As defined in GloMEEP and IAPH, 2018, we also consider the following types of data: scaled data, screening data, and comprehensive data*”. The same approach could be used for imputing missing data with a cold deck imputation method.

Scaled data

These data use approximations to obtain an order of magnitude corresponding to the types of eKPIs considered. These data are obtained by integrating external data produced by a port having a similar activity and representative of the real activity. The use of this type of data assumes that port activities are similar and should be adapted, for example, to port traffic. The results obtained using this type of data will also be approximate. Therefore, the results of the models based on this type of data have the following features:

- Average and non-specific input data for model port activity.
- Results based on an adaptation according to scale parameter (for example maritime traffic).
- Results are highly uncertain.

This type of data is useful when no information is available and when we are just interested by having an order of magnitude.

Screening data

These data are more detailed than scaled data and use data more specific to the port activity considered. These data use local sources on, for example, the description of the supply chain but use external data for emission factors. The results of the models based on this type of data have the following features:

- Some local input data and external data for model port activity.
- Results based on a simplification of some inputs (for example average energy consumption of cranes and trucks).
- Results are uncertain

Comprehensive data

These data are based on the detailed description of the port activity to be modeled with a complete knowledge of the supply chain, machine specifications, emission factors, etc. These data can come from sensors, administrative documents or expert knowledge. These data will yield much less uncertain results but require a lengthy data collection procedure. The results of the models based on this type of data have the following features:

- A detailed model of the port activities.
- Results can be verified and validated using measurements.
- Results have low uncertainty.

8.4. Mathematical methods for data normalization

The objective of the normalization step is to bring the different indicators that can be used to build the composite indicator to the same standard. Normalization avoids adding up apples and oranges (Tarantola and Saltelli 2007) since indicators could have different order of magnitude and units. In this step indicators are transformed in pure and dimensionless numbers.

We should note that there are a lot of different normalization methods and each one will give a different result for the composite indicator. That is why a robustness test should be done to assess the impact of the normalization method on the composite indicator. In the following text, the main methodologies are presented (according to Munda and Nardo (2005)) and some suggestions for normalization of eKPIs in the context of PEI are provided.

8.4.1. State of the art, description, and comparison of the different approaches

The *Table 8.3* shows the main normalization approaches that can be used for data normalization (Joint Research Centre-European Commission 2008). In the following sections these methods are discussed.

Table 8.3 Different normalization approaches in the context of composite indicator

Method	Equation
Ranking	$I_{qc}^{t} = Rank(x_{qc}^{t})$
Standardization (or z-scores)	$I_{qc}^{t} = \frac{x_{qc}^{t} - x_{qc=\bar{c}}^{t}}{\sigma_{qc=\bar{c}}^{t}}$
Re-scaling	$I_{qc}^{t} = \frac{x_{qc}^{t} - min_c(x_q^{t_0})}{max_c(x_q^{t_0}) - min_c(x_q^{t_0})}$
Distance to a reference port	$I_{qc}^{t} = \frac{x_{qc}^{t}}{x_{qc=\bar{c}}^{t_0}}$ or $I_{qc}^{t} = \frac{x_{qc}^{t} - x_{qc=\bar{c}}^{t_0}}{x_{qc=\bar{c}}^{t_0}}$
Logarithmic transformation	$I_{qc}^{t} = \ln(x_{qc}^{t})$
Categorical scales	If x_{qc}^{t}in the upper 5th percentile then y_{qc}^{t}= 100 If x_{qc}^{t} in the upper 15th percentile then y_{qc}^{t}= 80 Ifx_{qc}^{t} in the upper 35th percentile then y_{qc}^{t}= 60
Indicators above or below the mean	$if \frac{x_{qc}^{t}}{x_{qc=\bar{c}}^{t_0}} > (1+p), \quad then\ I_{qc}^{t} = 1$ $if \frac{x_{qc}^{t}}{x_{qc=\bar{c}}^{t_0}} < (1-p), \quad then\ I_{qc}^{t} = -1$ $if\ (1-p) < \frac{x_{qc}^{t}}{x_{qc=\bar{c}}^{t_0}} < (1+p), \quad then\ I_{qc}^{t} = 0$
Percentage of annual differences over consecutive years	$I_{qc}^{t} = \frac{x_{qc}^{t} - x_{qc}^{t-1}}{x_{qc}^{t}}$

8.4.1.1. Scale transformations

The scale effect is the fact that the different measurement units in which an indicator can be expressed before its normalization. Normalization methods can be invariant to change in measurement units (they provide the same value whatever is the measurement unit) but some normalization methods are not invariant. A clear example is provided in Munda and Nardo (2005). This should be taken into account when choosing the normalization method for a specific problem.

Since the PEI is based on eKPIs with different measurement units and order of magnitude, we must be careful to not have a scale effect when building it. This means that we should take attention that the normalization method does not (i.e.) increase the effect on the composite indicator of indicators with very small ranges.

8.4.1.2. Ranking of indicators

This is the simplest method. It consists of ranking each indicator across ports. This method is independent to outliers but there is a loss of information about absolute values.

If the data are time-dependent, the ranking is done for each point in time. Thus, it is possible to know the relative position of ports for each indicator in time, but we cannot follow the absolute performance of the ports (if port A reduce its CO_2 emissions but port B reduces it owns faster, port A's ranking will be deteriorated).

Despite the simplicity of the ranking methods, it seems that ranking of indicators is not a good method to use in PIXEL project since it will not reflect absolute performance and improvement of performance.

8.4.1.3. Standardization (Z-scores)

In this method for each environmental indicators (eKPI) the average and the standard deviation across ports are calculated. The normalization formula is:

$$Nomalised\ value = \frac{[Observed\ value - Average\ value]}{Standard\ deviation}$$

The min and max of the normalized value depend on each eKPI. This method is the most used because all indicators have the same scale (average of zero and standard deviation of one). An average of zero avoid introducing distortions in the aggregation step due to difference of indicator means.

If the data are time-dependent, the average and the standard deviation are calculated for the desired reference time (if we want to compare port on a monthly basis, average and standard deviation must be calculated on a monthly basis).

8.4.1.4. Re-scaling

The objective here is to obtain a normalized indicator with value between 0 (worst performance) and 1 (best performance). Each indicator is transformed using the following formulae:

$$Normalised\ value\ (t) = \frac{[Observed\ value\ (t) - \min value(t)]}{[\max value - \min value]},$$

where min and max values are the minimum and the maximum values across all ports at a specific time.

The method could introduce a distortion effect since the extreme values (max and min) could be unreliable outliers. It could also increase the effect on the composite indicator of indicators with very small ranges. The re-scaling method could be adapted to time-dependent studies but will not be stable if new values are included.

8.4.1.5. Distance to a reference port

In this approach, the normalized value takes the ratio between the value of an eKPIs and the value of the eKPIs of a reference port. A good method is to consider that the reference port is a target to be reached in a given time frame. Indicators that are higher than 1 show port with less performance. The reference port could also be the group leader. It is also possible to consider that the reference point is the initial performance of the port. This approach is used in environmental problems and allows following the evolution of performance in time. For this reason, the distance to a reference port method could be applied for the PEI.

8.4.1.6. Categorical scales

The objective here is to assign a categorical score. The first step is to design and select the categories (they can be numerical or qualitative). The second step is to assign a score, which in fact could be arbitrary, to each category. The most commonly used method is based on the percentiles of the distribution of the eKPIs across port.

Categorical scales have some drawbacks: i) a large part of information about the variance is omitted, ii) a small variation of indicators might not be reflected.

8.4.1.7. Indicators above or below the mean

This method considers the indicators that are above and below an arbitrary defined threshold around the mean. This method it very simple and not affected by outliers. However, the threshold level is arbitrary and there is the omission of absolute levels.

8.4.1.8. Percentage of annual differences over consecutive years/months

In this approach, each eKPI is transformed in order to represent the percentage growth with respect to the previous years/months. The transformation can be used only when the indicators are available for a number of years/months.

8.4.2. Normalization and the PEI

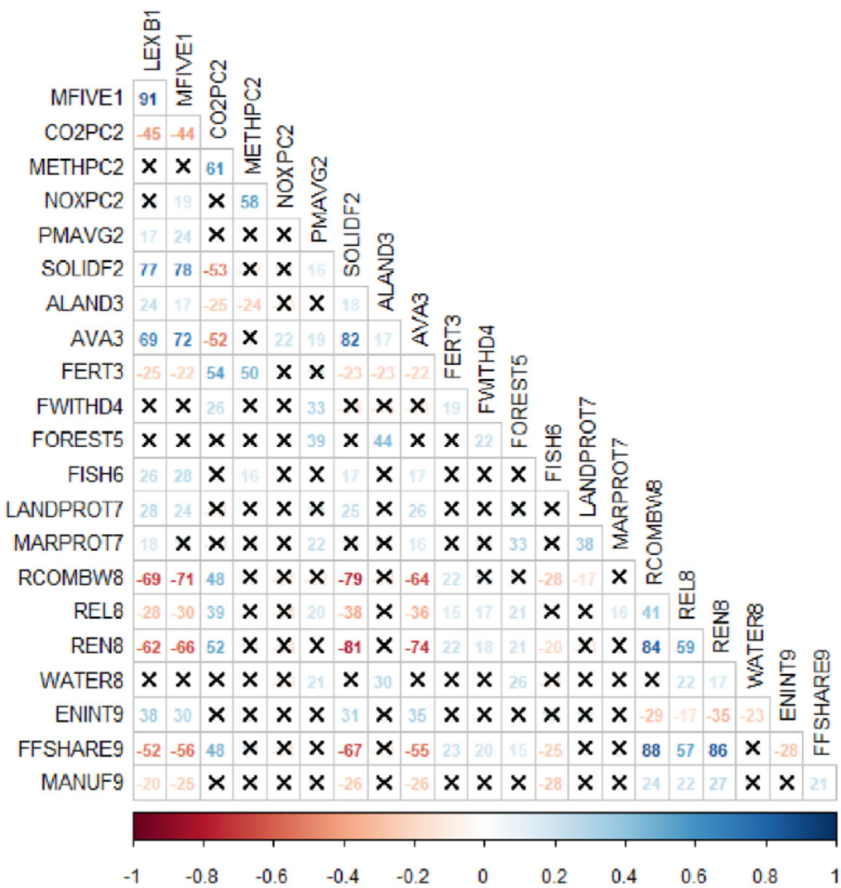
In the Port Environmental Index, the eKPIs are expressed in different standards with different measurement units, so there is a need for normalization. We have also highlighted that we should be careful about the scale effect and take it into account in the normalization process. The following method seems to be well-adapted for the PEI in the context of PIXEL project:

- Standardization
- Re-scaling
- Distance to a reference port or initial time

8.5. Mathematical methods for analyzing data structure

In order to gain insight into the structure of datasets, statistical tools can be used. The first basic step is to calculate mean, median and standard deviation for each of the eKPIs.

Then, a correlation analysis should be carried out in order to help select the most relevant variable to include into the composite index. If some eKPIs are highly correlated a double-counting issue can appear. As shown in literature, indicators are often chosen with little attention paid to the interrelationships between them.



Legend:
LEXB1: Life expectancy at birth;
MFIVE1: Child mortality rate;
METHPC2: Methane emissions;
CO2PC2: CO2 emissions;
NOXPC2: Nitrous oxide emissions;
PMAVG2: Air pollution;
SOLIDF2: Populations using solid fuels;
ALAND3: Share of agricultural land;
AVA3: Share of Agriculture value added;
FERT3: Fertilizer consumption;
FWITHD4: Freshwater withdrawals;
FOREST5: Forest area;
FISH6: Share of capture fisheries;
MARPROT7: Marine protected areas;
LANDPROT7: Terrestrial protected areas;
REL8: Share of Renewable electricity;
REN8: Share of Renewable energy consumption;
RWATER8: Renewable internal freshwater per capita;
RCOMBW8: Share of Combustible renewables and waste;
ENINT9: Energy intensity level;
FFSHARE9: Fossil fuel energy consumption;
MANUF9: Share of manufacturing value added.

Source: authors' calculations.
Notes: X denotes correlations insignificant at a 5% level

Figure 8.1 An example of a correlation analysis in order to build a composite indicator

As listed in Joint Research Centre-European Commission (2008), other methods like principal components analysis, factor analysis or Cronbach coefficient alpha can be used in order to group information on eKPIs. We can also group information by port using cluster analysis or factorial k-means analysis. All of the information obtained can then be used to guide weighting and aggregating processes for which it is important to have a good statistical description of the data to choose the best approach.

8.6. Mathematical methods for data weighting

In literature, there is no agreement on a methodology for weighting individual indicators and no consensus will ever exist (Joint Research Centre-European Commission 2008). As expressed by Gan et al. (2017), the process of indicator integration is an inherently subjective procedure (Morse et al., 2001), selecting appropriate weighting and aggregation methods is challenging (Saisana and Tarantola 2002; Wilson and Wu 2017). This highlights the danger of presenting a composite indicator as "objective". The weighing procedure will also be applied for the Port Environmental Index. Therefore, weighting's assumptions should be crystal clear and fully transparent.

As described by Gan et al. (2017), who provides a review of the methods used for 96 sustainability indicators, in most cases equal weighting of components is used (46.88%), with analytical methods (principal component analysis, benefit-of-doubt, regression analysis, etc.) coming second (30.21%), and opinion-based methods third (15.62%). It is also highlighted in Joint Research Centre-European Commission (2008) that equal weighting is used when there are no statistical or empirical grounds for choosing a different scheme. Literature also described that equal weighting does not mean no weighting since it implies an implicit judgment that weights are equals.

The weighting process is also influenced by the statistical quality of data. For example, a choice can be made to have a higher weight assigned to statistically reliable data so that data with high percentages of missing value do not have a great impact on the composite indicator.

8.6.1. State of the art and description of the different approaches

In the following, we focus on the main methods used in literature as shown in the *Figure 8.2*. These methods are the most commonly used. Regression approach, unobserved components models, conjoint analysis and analytic hierarchy require a large amount of data to produce good results and are the most complex ones. These methods are not described because they seem not applicable at all in the PIXEL context.

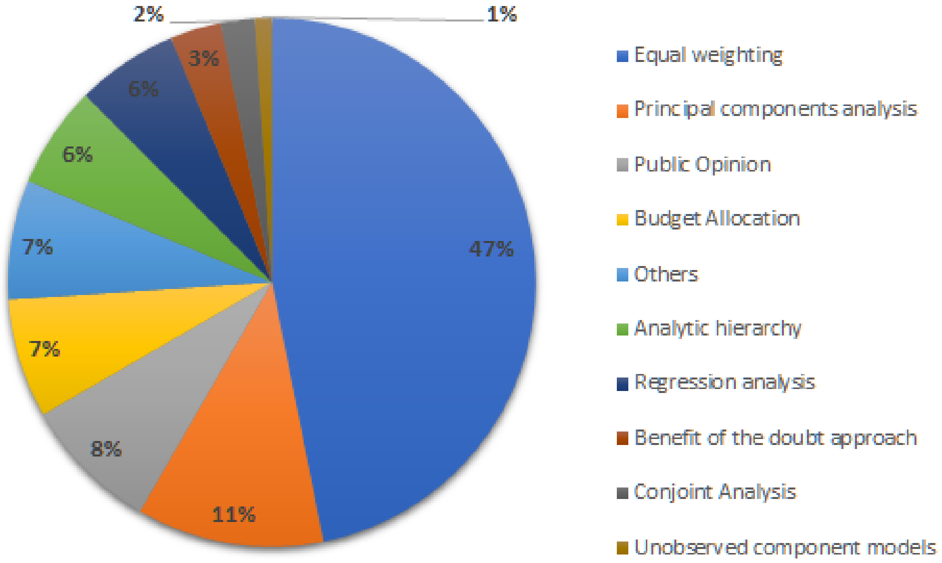


Figure 8.2 Proportions of methods used for indicator weighting according to Gan et al. (2017)

8.6.1.1. Equal Weighting

As already said, this approach is the most commonly used. This method can be used if we considered that all the indicators are equally important or if we have no statistical or empirical evidence supports a different scheme (Joint Research Centre-European Commission 2008). This strategy is also recognized as the simplest one and is easily replicable (Land 2006).

Equal weighting can lead to combine variables with high degree of correlation introducing an element of double counting. To avoid this double-counting, we can define a threshold beyond which the correlation is a symptom of double counting.

According to Joint Research Centre-European Commission (2008), this approach works well if all dimensions (water pollution, air pollution, biodiversity, …) are represented in the composite indicator with the same number of sub-indicators. If this is not the case, it will imply a higher weight to the dimension represented with the higher number.

In PIXEL and for building the PEI, the approach can be used as it is simple to use and can serve as “a basis” for studying the sensitivity of the PEI to the weighting approach.

8.6.1.2. Principal components analysis or factor analysis

Principal components analysis (PCA) and factor analysis (FA) aim at reducing the dimension of the data while minimizing information loss. This approach allows to group together indicators that appear as collinear (indicators must be expressed in the same measurement unit). According to Joint Research Centre-European Commission (2008), PCA/FA are useful because they solve the double-counting problem, but they can only be used with correlated indicators and are sensitive to modifications of data set and small-sample problem. Moreover, because they reduce the dimensionality of the data set, FCA and FA are used and more valuable when many indicators are considered.

This approach seems not applicable for the PEI since the data set is small and number of indicators to weigh relatively low.

8.6.1.3. Benefit of the doubt approach

The benefit of the doubt (BOD) method is a direct application of the data envelopment analysis (DEA) and fully described in Cherchye et al. (2007). This method has the following main advantages: i) the composite index will be sensitive to policy priorities because the weights are defined by the observed performance, ii) it could be "incentive generating" rather than "punishing". The mains drawbacks of this method are that weights will be port-specific and no comparison between ports could be done. One objective of PIXEL project and the PEI is to develop a composite index that can be used to compare environmental performance of small and medium ports across the EU, which is why this weighting approach cannot be used in the PIXEL context.

8.6.1.4. Budget allocation

Budget allocation (BAL) approach or expert opinion is a method where experts with extensive knowledge and experience are joined together to distribute a budget of "n" points over the indicators. Based on experts' judgment, indicators that are judged to be more important are given a larger proportion of the budget. Then the weighting is done according to the budget distribution. The method follows four steps:

1. Selection of the experts for the valuation
2. Allocation of budget to the indicators (eKPIs in PIXEL)
3. Calculation of the weights
4. Iteration of the budget allocation until convergence is reached (optional)

As described in Joint Research Centre-European Commission (2008), BAL is useful for its transparency and explicitness, but the weighting process could reflect local specific conditions and not transferable from one area to another. Moreover, it could measure the urgency of the situation or need of political intervention rather than measure the importance of each indicator (e.g. more weight on wastewater emission if the expert considers that nothing has been done to reduce them).

This type of approach is used when it is essential to bring experts with a wide spectrum of knowledge and experience and is optimal for a maximum number of sub-indicators equal to 10-12. If the number of sub-indicators is higher, a cognitive stress can lead to a biased allocation.

In PIXEL, we will test this weighting process for its transparency and explicitness.

8.6.1.5. Public Opinion

In the public opinion approach, the weights of the indicators are not determined by experts but by a general public. As expressed by Parker (1991), this approach is easy and inexpensive. Stakeholders are also allowed to give their opinion and preference (van Haaster et al. 2017). As shown by Munda and Nardo (2005), this method is useful for multi-criteria processes and can make the process transparent and participative. Indicators receiving more concern are allocated with a higher weight.

The main drawback of the public opinion approach is that weights are based on public concern rather than importance of an indicator. For example, people may pay less attention to biodiversity issue than air pollution.

In PIXEL, this approach can also be used in order to include the port's stakeholders in the Port Environmental Index. It could also be interesting to compare the PEI results using equal weighting, budget allocation, and public opinion approaches.

8.6.2. Weighting and PEI

According to literature (the *Figure 8.3* sums up the different weighting approach according to Gan et al. (2017)) and the theoretical background of the PEI, the following weighting methods seem applicable for the PEI and must be discussed (at least for testing the robustness of the PEI):

- Equal weighting
- Budget allocation
- Public opinion

Common methods for indicator weighting (primarily based on Nardo et al. (2005), OECD (2008), Hermans et al. (2008), and Mikulić et al. (2015)).

Method Name	Type	Examples	Formulas	Benefits	Drawbacks
Equal weighting	Equal weighting	Human Development Index (UNDP, 1990) Genuine Savings (WorldBank, 1999)	$\omega_i = \omega$, $i = 1, \ldots, m$, where ω_i is the weight of the i^{th} indicator and ω a constant used as the weights for all the indicators	Simple, replicable and straightforward.	No insights into indicator relationships; risk of double weighting.
PCA/FA	Statistic-based	Environmental Sustainability Index (Sands and Podmore 2000) The 2006 European e-Business Readiness Index (Pennoni et al., 2006)	$\omega_i = r_j({l_{ij}}^2/E_j)$ $i = 1,\ldots,m$; $j = 1,\ldots,n$ where r_j is the proportion of the explained variance of factor j (or the intermediate composite j) in the data set, l_{ij} the factor loading of the i^{th} indicator on factor j and E_j the variance explained by the factor j	Reduces the risk of double weighting, classifying ungrouped indicators.	Dimensions of sustainability are unpredictable, and weights may differ from reality.
Benefit of the doubt approach (BOD)	Statistic-based	Meta-index of Sustainable Development (Cherchye and Kuosmanen 2004) Macro-economic performance evaluation (Melyn and Moesen 1991)	$\omega_c = \arg\max_{\omega_{c,i}} \frac{\sum_{i=1}^{m} \omega_{c,i} I_{c,i}}{\max_{y_{ij} \in \{\text{studied units}\}} \sum_{i=1}^{m} \omega_{c,i} I_{j,i}}$ $s.t. \sum_{i=1}^{m} \omega_{c,i} I_{j,i} \leq 1$, $\omega_{c,i} \geq 0$ $\forall\, i = 1, \ldots, m$; $\forall\, j = 1, \ldots, n$ where ω_c is the weight vector of unit c, $\omega_{c,i}$ the weight of the i^{th} indicator of unit c, $I_{c,i}$ the normalized score of the i^{th} indicator of unit c, and $I_{j,i}$ the normalized score of the i^{th} indicator of the j^{th} unit	The processes of weighting, aggregation, and index construction are efficiently integrated. Weights are selected to maximize the index for each unit.	Results may not be comparable and lack transparency. A multiplicity of solutions exists.
Regression analysis (RA)	Statistic-based	National Innovative Capacity (Porter and Stern 2001)	$\omega_i = \beta_i$, $i = 1,\ldots,m$ where β_i is the regression coefficient of the i^{th} indicator	Results can be used for updating or validating weights.	Either multi-collinearity among indicators or an improper dependent variable may lead to poor results.
Unobserved component models (UCM)	Statistic-based	The aggregate governance indicators (Kaufmann et al., 1999)	$\omega_i = \frac{\delta_i^{-2}}{1 + \sum_{i=1}^{m} \delta_i^{-2}}$ $i = 1\ldots m$ where δ_i is the variance of the i^{th} indicator	The processes of weighting, aggregation, and index construction are efficiently integrated. Statistical significance can be expressed when conducting comparisons.	Results are sensitive to outliers. Problems of identification may occur if indicators are highly correlated. Reliability and robustness of the model may be lost when adequate data are not available.
Budget allocation (BAL)	Public/Expert opinion-based	The Eco-indicator 99 (Goedkoop and Spriensma, 2001) Overall Health System Attainment (Murray et al., 2000)	–	Transparent and explicit.	Measuring urgency instead of importance; region-specific.
Public opinion (PO)	Public/Expert opinion-based	Concern about environmental problems Index (Parker, 1991)	–	Transparent and participatory.	Measuring concern instead of importance; region-specific.
Analytic hierarchy process (AHP)	Public/Expert opinion-based	Composite sustainability performance index (Singh et al., 2007) Index of Environmental Friendliness (Puolamaa et al., 1996)	$A\omega = \lambda\omega$ where A is the comparison matrix, λ the largest eigenvalue of A, and ω the weight vector as well as the eigenvector corresponding to λ	Has a hierarchical structure that is in line with the structure of sustainability frameworks. Simple and flexible. Providing consistent verification operation. Available for both quantitative and qualitative data.	Requirement of a high number of pairwise comparisons. Inconsistency and cognitive stress may exist if there are too many indicators in each cluster.
Conjoint analysis (CA)	Public/Expert opinion-based	Indicator of quality of life in the city of Istanbul (Ülengin et al., 2001)	$\omega_i = \frac{\partial P(I_1, \ldots, I_m)}{\partial I_i}$ where $P(I_1,\ldots,I_m)$ is the preference function defined by researchers and I_i the i^{th} indicator	Results can be easily used for making sustainability plans. Available for both quantitative and qualitative data.	Requires a large sample of respondents. Has complicated estimation process.

Figure 8.3 Summary of weighting method according to Gan et al. (2017)

8.7. Mathematical methods for data aggregation

8.7.1. State of the art and description of the different approaches

In literature, several examples of aggregation techniques are described for composite indicators. According to Joint Research Centre-European Commission (2008), the most commonly used are the additive aggregation methods. Gan et al. (2017) show that there are three main approaches for aggregation: additive aggregation, geometric aggregation, and non-compensatory aggregation. In the following sub-chapter, we focus on these approaches.

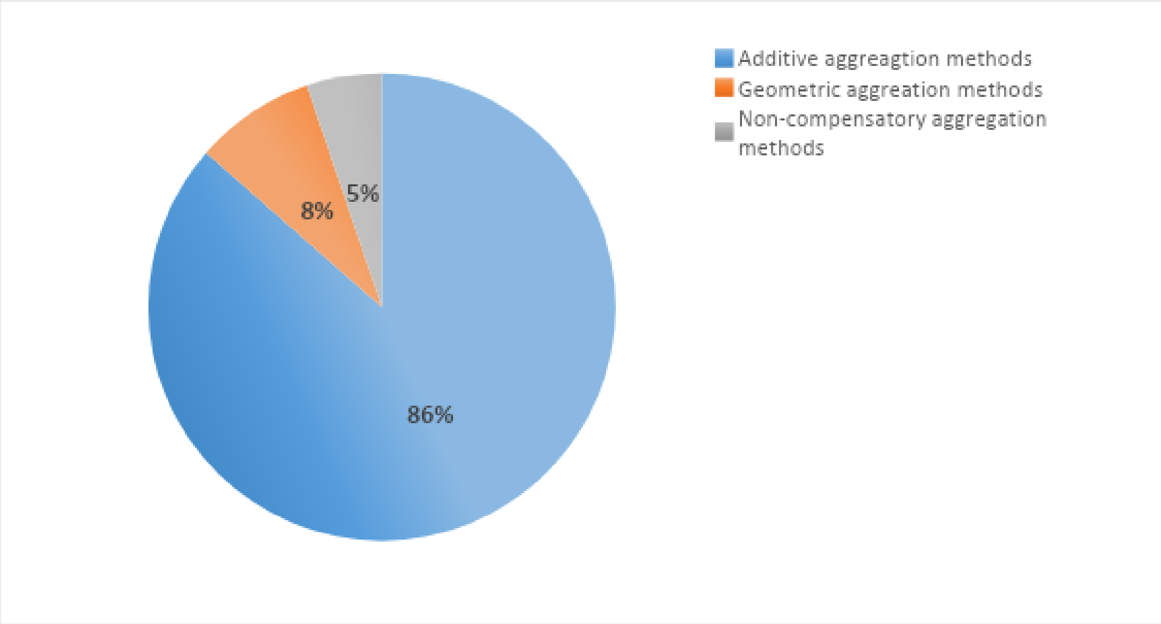


Figure 8.4 Portions of methods used for indicator weighting according to Gan et al. (2017)

8.7.1.1. Additive aggregation methods

In all additive aggregation methods, the normalized values of indicators are summed up using a specific function. The most used function is the weighted arithmetic mean: the normalized values of indicators are summed up using their respective weight. Other additive methods are presented in Joint Research Centre-European Commission (2008).

As noted by Gan et al. (2017), additive aggregation methods should be used carefully since these methods imply two main features:

1. Preferential independence: Indicators must be independent, meaning that the contribution of all indicators can be added together implying that no synergy or conflicts exist among different indicators. As shown by Booysen (2002) and Joint Research Centre-European Commission (2008) this assumption seems unrealistic in many situations. If the assumption is not respected this will result in a biased composite indicator in which the dimension and the direction of the error will be difficult to determine.
2. If there is a substantial interaction between indicators, additive methods should not be used since these methods intrinsically imply a compensatory logic. As demonstrated in Joint Research Centre-European Commission (2008), weights in additive methods have the meaning of substitution rates and do not indicate the importance of the indicator associated.

This approach is useful when sub-indicators have the same measurement units and when scale effect are neutralized. It also has trade-offs between sub-indicators meaning that the deficit in one indicator can be neutralized by the surplus of another one. In additive (linear) aggregation method the compensability is linear.

8.7.1.2. Geometric aggregation methods

Geometric aggregation methods use multiplicative functions. The most commonly used function is the weighted geometric mean. As explained by Booysen (2002) and Bullen (2013), geometric mean-based methods only allow compensability between indicators within certain limitations. This means that the ability of indicators with very low scores to be fully compensated for by indicators with high scores is limited. In geometric aggregation method, the compensability is lower when the composite indicators contain indicators with low values.

When using geometric aggregation methods, the measurement scale must be the same for all indicators in order to remove the scale effects. The normalization method should take this into account.

There are also limitations to such approaches. First, geometric methods are not fully non-compensatory techniques and like the additive, method are preferentially dependent. Second, it is not possible to analyze sensitivity and uncertainty quantification using measurement errors of indicators (Booysen 2002).

8.7.1.3. Non-compensatory methods

Both additive and geometric methods accept compensation among indicators. As explained by Gan et al. (2017), this can be a subject of bias for a lot of composite indicators. If substitution between indicators (eKPIs in PIXEL) is unacceptable a non-compensatory method has to be used.

The non-compensatory methods are based on two main point of view:

- the perspective of a multi-criteria decision making (MCDM),
- the properties of aggregation function.

This approach is focused on decision-maker preference and cantered around the fact that a general objective is to create rankings (Gan et al. 2017). Thus, the process of non-compensatory methods is divided in two phases (Joint Research Centre-European Commission 2008):

1. pair-wise comparison of ports according to the whole set of eKPIs
2. ranking of ports in a complete pre-order.

Munda (2005) provides a full description of multi-criteria decision analysis applied in sustainable development and shows that this approach is an adequate approach for dealing with sustainability conflicts at both micro and macro levels of analysis. This has to be taken into account when building the Port Environmental Index since we are also dealing with sustainable development.

Lee and Anderson (2009) compare the results of a frugal and fast non-compensatory decision-making strategy with a more cognitively intense, attribute-based compensatory strategy and hypothesize that both will generate similar results.

8.7.2. Aggregation and PEI

Summary of aggregation methods is presented in the *Table 8.4.*

Table 8.4 Summary of aggregation method (Gan et al. 2017)

Common methods for aggregation	Examples	Formulas	Benefits	Drawbacks
Additive aggregation	Environmental Performance Index (Esty et al. 2006)	$SI = \;_1I_1 + \;_2I_2 + \cdots + \;_mI_m = {}^{m}_{i-1}\;_iI_i$ where: *SI* is the sustainability index, $_i$ the weight of the i^{th} indicator, and I_i the normalized score of the i^{th} indicator.	Transparent and simple. Easy to execute sensitivity analysis and uncertainty quantification	Rigorous prerequisites exist, such as mutually preferentially independence.

Geometric aggregation	Living Planet Index (Loh et al. 1998; Loh et al. 2005)	$SI = I_1{}^{1} I_2{}^{2} \ldots I_m{}^{m} = \prod_{i=1}^{m} I_i^{\omega_i}$ where: *SI* is the sustainability index, $_i$ the weight of the i^{th} indicator, and I_i the normalized score of the i^{th} indicator.	Transparent and simple. Can be used for all kinds of ratio-scale variables	Rigorous prerequisites exist, such as mutually preferentially independence.
Non-compensatory aggregation methods	Index for "Social Multi-Criteria evaluation" (Munda 2004)	Rank (*Unit*) $s.t. \varphi_* = max \sum e_{jk}$ $i = 1, \ldots, n$ where: Rank ($Unit_i$) is overall ranking of the *n* researched units, φ the corresponding score of the final ranking of the researched units, and e_{jk} the generic element of the outranking matrix.	No ad hoc restrictions.	Computational problems may be caused by the increasing number of units or indicators. Losing information on the intensity of sustainability.

As described in the previous section, one of the main questions before setting up an aggregation method is to know if compensability is allowed or not. Indeed, in order to use linear or geometric aggregation, the assumption of absence of synergy or conflicts effects among the eKPIs is a necessary condition. We will have to check this assumption in the context of PIXEL and the eKPIs used to build the Port Environmental Index. If it appears that the analyses and the theoretical framework of the PEI and the associated eKPIs show that the different goals are equally legitimate and important, then a non-compensatory logic may be necessary for building PEI. As described in Joint Research Centre-European Commission (2008), this is usually the case of environmental indexes.

8.8. Mathematical methods for uncertainty and sensitivity analysis

As explained in the literature, a composite indicator may send misleading, non-robust policy messages if it is poorly constructed and misinterpreted. We should have in mind that building a composite indicator involves stages where judgments are made (Joint Research Centre-European Commission 2008): the selection of sub-indicators, the choice of a conceptual model, the weighting of indicators, the treatment of missing values, etc. In order to assess the robustness of the composite indicator and increase its transparency a combination of uncertainty and sensitivity analysis must be performed.

Uncertainty analysis (UA) studies how uncertainty in the input factors (eKPIs or inputs for calculate eKPIs for the PEI) propagates through the structure of the composite indicator. Sensitivity analysis (SA) focuses on how much each individual source of uncertainty contributes to the variance of the composite indicator. According to literature, the following steps should be subject on this kind of analysis:

- selection of eKPIs: effect of inclusion/exclusion of sub-indicators. This can be done by excluding one indicator at a time. It is also possible to neglect an indicator by assigning a very small weight;
- data quality: modeling of data error for each input based of available information on variance estimation;
- data imputation: effect of different data imputation approaches;
- data normalization: effect of different data normalization approaches. As previously described, several normalization methods can be used;
- weighting scheme: using different weighting method. For example, two methods in the participatory family (budget allocation and public opinion) and equal weighting;
- weights' value;
- composite indicator formula: compare additive, geometric and multicriteria approaches.

It is worth noticing that UA and SA are influenced by each step and that we cannot just analyze each step on its own. Thus, many combinations and calculations will be performed to obtain a valid uncertainty and sensitivity analysis.

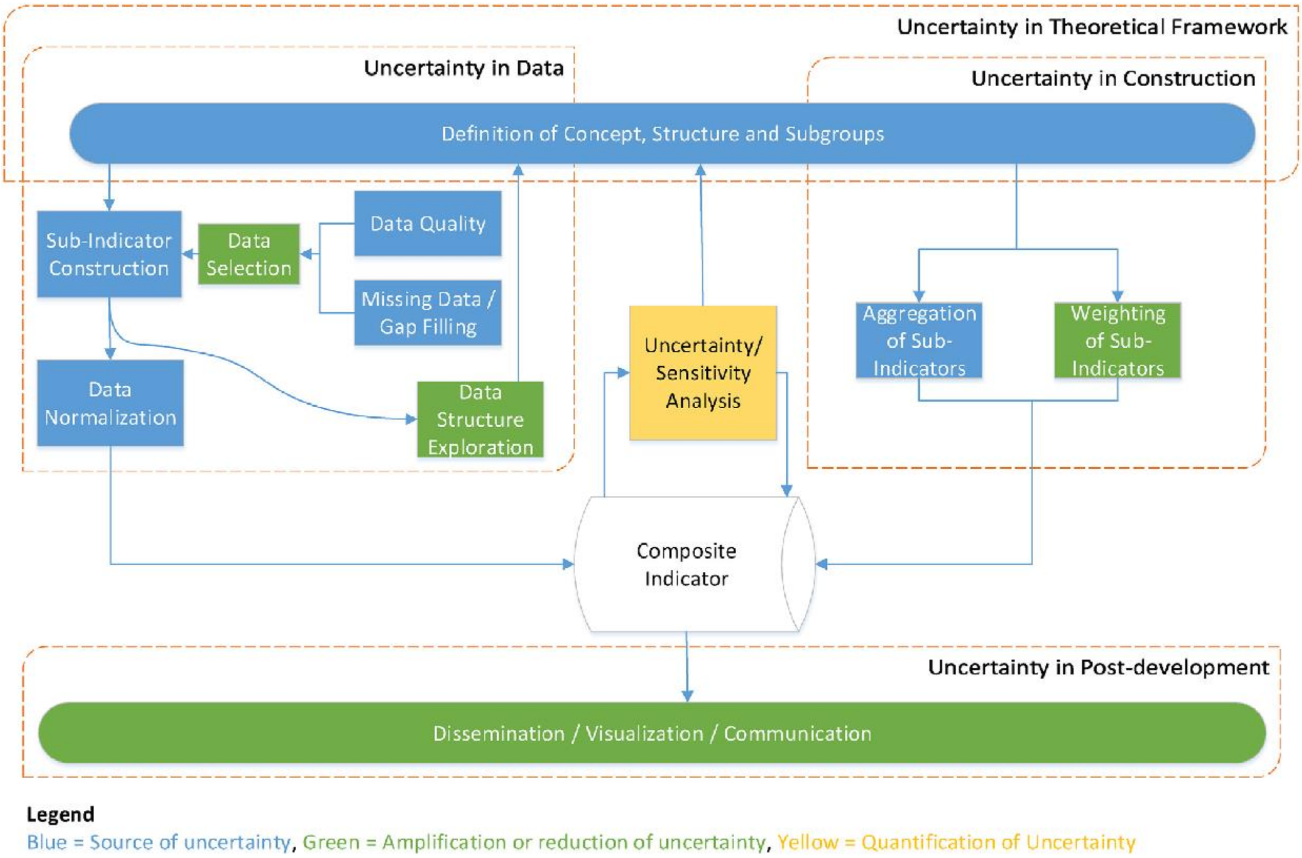


Figure 8.5 Conceptual model of uncertainty flow through a composite indicator (Burgass et al. 2017)

The objective is to be able to answer the following questions (Joint Research Centre-European Commission 2008):

- Does the use of one strategy versus another in building PEI provide a biased picture of ports' performance?
- To what extent do the uncertain input factors (eKPIs or value to calculate eKPIs) affect the ports' rank?

Burgass et al. (2017) provide a good review of uncertainty in environmental composite indicators and shows an interesting conceptual model of uncertainty flow through a composite indicator (see *Figure 8.5*). This demonstrates that there are many different type and sources of uncertainty in the different stages of the composite indicator construction.

Range of uncertainties in composite indicators and how to treat them is shown below.

Table 8.5 Uncertainty in composite indicator (Kaufmann et al. 2011)

Source of Uncertainty	Issue	Reason for Issue	Potential solution
Theoretical Framework	Is theoretical framework representative of the system?	• No systematic process • Subjective	• Systems Modelling • Systematic expert judgment/ stakeholder engagement

		• Lack of transparency and repeatability	• Transparency and iterative improvement
Data	Accuracy of data	• Data quality rarely assessed and therefore not really considered	• Data scoring/pedigree matrices • Systematic expert judgment/ stakeholder engagement • Uncertainty analysis
	Amount of missing data	• Unclear where data gaps are and number of them • Gap-filling methods are subjective	• Transparency and iterative improvement • Uncertainty/sensitivity analysis • Advanced monte-Carlo gap-filling methods
	Is indicator an accurate and desired representation of the system	• Led by data availability, stakeholder or constructor values therefore subjective • Unclear how indicators relate to system	• Systems modeling
	Representation vs quality	• Trade-off between data accuracy and missing data vs how well it represents the system • Subjective	• Transparency and iterative improvement • Systematic expert judgment/ stakeholder engagement
Data Normalization	Different methods	• Subjective	• Transparency and iterative improvement • Uncertainty/sensitivity analysis
Weighting Aggregation	Arbitrary weighting	• Unclear how weights were assigned • "Neutral" weighting still a weighting decision • Subjective	• Systems modelling • Systematic expert judgement/ stakeholder engagement
	Implicit weights may be different to assigned weights	• Statistical properties mean assigned weights don't always work as intended	• Correlation analysis • Uncertainty/sensitivity analysis
	Different methods	• Subjective	• Transparency and iterative improvement • Uncertainty/sensitivity analysis
Communication	Different interested parties	• How to communicate to public/policymakers/scientists	• Transparency and iterative improvement • Multi-layered approach of engagement/analysis

As explained in literature, the following requirement about data are the minimum ones and have to be checked before being selected for inclusion (Joint Research Centre-European Commission 2008):

- Data relevance
- Data accuracy
- Frequency of acquisition
- Accessibility
- Interpretability
- Coherence

For example, for the Living Planet Index (Collen et al. 2009) a score is assigned to data based on their source, methodology, and whether a measure of variation was included. This score can be used to represent uncertainty in PIXEL. The best approach is to describe input data with an average value and an associated standard deviation that model's errors.

As explained previously, when imputing missing data, it is also very important to quantify the degree of errors related to the imputation. Some methods, like the multiple imputation method, allow for obtaining missing values with a statistical description but could be time-consuming and requiring a lot of expertise. Literature highlights the fact that the PEI documentation should be open about which data have been imputed or deleted so that the uncertainty can be easily identified. Ideally eKPIs considered for building PEI should be selected based on their relevance for the PEI objective. However, it often appears that the data sets required to calculate eKPIs are not available, of they are of bad quality with a lot of missing data. This could lead to discarding some eKPIs and consider only eKPIs with reliable data. The *Table 8.5* (Kaufmann et al. 2011) sums up the different issues that can appear when building a composite indicator and provides solutions to solve them.

As explained by Saisana et al. (2005), a Monte Carlo approach can be followed in order to perform an uncertainty analysis. The different steps are described by Saisana et al. (2005). This approach is a well-known strategy for uncertainty analysis and could be easily used in the PEI context.

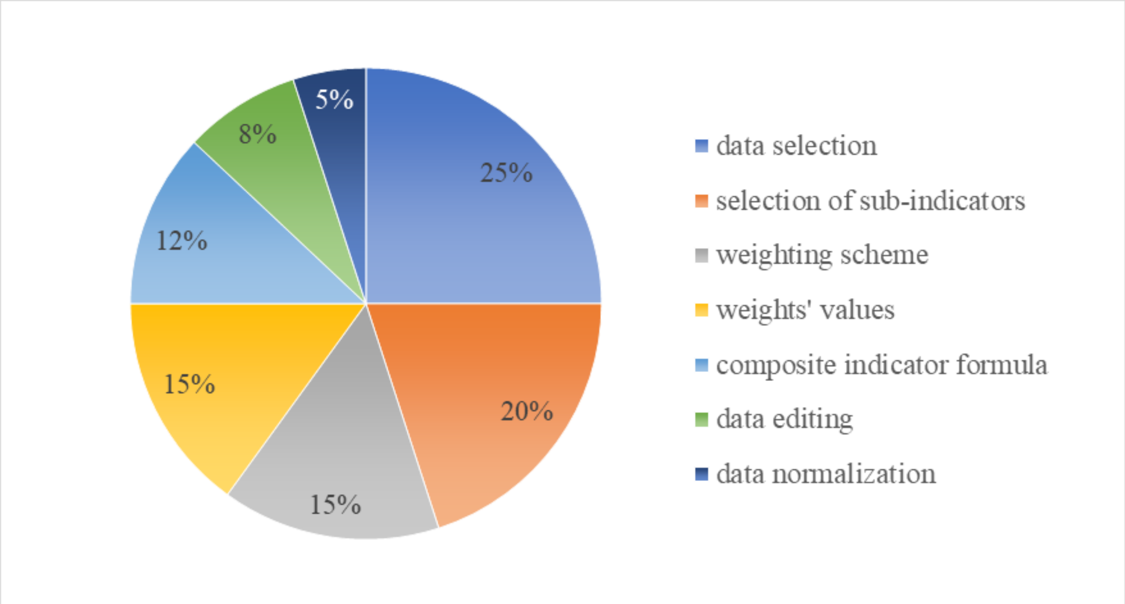


Figure 8.6 Example of sensitivity analysis results (source JRC group, Saisana (2004))

For sensitivity analysis, we can first use the Morris method (Morris 1991; Campolongo et al. 2007) in order to identify the factor that influence the most the variability of the composite indicator. This approach only gives qualitative results as a global sensitivity analysis but is very useful in order to understand which factors have the largest influence on the composite indicator. Furthermore, it is economic in computational time in the sense that it requires several model evaluations that is linear in the number of model parameters. The method can be regarded as global as the final measure is obtained by averaging a number of local measures (the elementary effects), computed at different points of the input space. This approach can be combined with the Sobol method, which is a variance-based sensitivity analysis. It decomposes the variance of the output of the composite indicators into fractions which can be attributed to inputs or sets of inputs (eKPIs). As explained by Saisana et al. (2005), this approach has several interesting features:

- model independence;
- exploration of the whole range of variation in the input factors;
- consideration of interaction effects between the indicators (eKPIs);

A complete sensitivity analysis of the 2008 Environmental Performance Index is available in Saisana and Saltelli (2008). That document can be used as a very good and useful starting point to carry out a sensitivity analysis of the Port Environmental Index.

9. PEI visualization

In this section, the visualization of the PEI is detailed in order to provide a complete technological perspective of the WP5 product (PEI) included in the global PIXEL framework. By using this platform, port authorities and other decision-makers should easily observe their advancement towards a more environmentally friendly port. Furthermore, the partners of the project are considering the future possibility to use it as ideally benchmark of their performance against other ports.

The section 7.2 of this deliverable has offered an overview of PIXEL architecture and has depicted the execution flow that take place each time the PEI is requested to be calculated. It can be seen that PIXEL Operational Tools (OT) are in charge to feed the Dashboard, in this specific case to PEI visualization. In order to offer the complete vision of the process, this chapter is focused on PEI visualization drawing from the information that has been decided to be offered through the Dashboard.

The potential users are going to interact with the PEI elements through the visual component. This visualization element will offer them the results and the work behind the PEI calculation. This Dashboard component aims at guaranteeing that a user is attracted in a first instance by the PEI solution. Therefore, the objective is to provide a user interface facilitating a relevant and satisfactory user experience (Quality of Experience – QoE). For this, designers and developers are taking into account the visualization and usability aspects and the users learning curve. The user experience is going to start from visualization elements based on represented values in a simple visual way (basic visualization elements), to finally offer elements (advanced visualization elements) providing' capabilities of interaction with the data and graphs. This increase of complexity will also appear in each step related with the development of the solution and the developed versions of the product. This is logic and compliant with the Grant Agreement considering that WP5 still needs to define recommendations, guidelines and particular usability features of the PEI (tasks T5.4 and T5.5). The first versions will be focused on the value of the raw data, indices, eKPIs and associated graphs. The final versions will provide more complex graphs and offer possibilities of interaction.

Regarding the structure of this section, the section 9.1 is going to provide a brief overview about the visualization techniques to present and visualize Composite Indicators. We are using this term ("composite") drawing from the definition of the PEI in deliverable D5.1. The section 9.2 is going to detail the data involved in the PEI Visualization and describe how to visualize that data obtained from PEI calculation, in a visual and understandable way.

9.1. Visualization of Composite Indicators

Composite indicators (CI) such as PEI can be presented and visualized in a number of different ways which greatly affects their interpretation. Therefore, it is of uttermost importance to select the visualization technique which will convey the information in an accurate and clear manner.

Usually, CIs can be expressed using the following:

- tabular presentations;
- bar charts;
- line chars;
- time-series charts;
- and more.

In this subsection tentative, PEI visualization techniques will be shown and briefly introduced and discussed. It should be noted that at this moment the PEI's exact algorithms are still to be fully addressed and shown visualizations are provisional versions of the CI.

9.1.1. Tabular presentations

Composite indicators could require a presentation of data where the information is simply presented in the form of a table with rows and columns. Although tables contain complete information, they can sometimes hide sensitive information that could be immediately apparent when shown in graphical format. Therefore, it is important to decide whether to use tables, graphs or both despite the fact that the data representation will be redundant. Although redundancy should be generally avoided, it if helps the end-user get a better insight into the data it can be allowed for.

A possible complex example could be the ports' ranking over a given period. For that case (*Table 9.1*), a ranking of several ports for the years 2018 and 2019 has been shown. The table also shows which ports maintained their ranking, which have improved their ranking or decreased their environmental performance compared to their ranking the previous year. The fact of comparing PEIs among ports is tricky. This topic is being addressed in task T9.4 Exploitation (for more information, consult D9.7 whenever it is ready). The real implementation of that table could be considered in the future only if the ports were able to share this information or display it anonymously.

Table 9.1 PEI tabular visualization

Performance indicators		PEI ranking 2018	PEI ranking 2019	PEI growth (progress)
1.	Port 1	1	1	●
2.	Port 2	2	2	●
3.	Port 3	3	3	●
4.	Port 4	4	4	●
5.	Port 5	5	6	▼
6.	Port 6	6	5	▲
7.	Port 7	7	8	▼

9.1.2. Bar charts presentations

Composite indicators can also be expressed with simple bar charts. Colors can make the chart more visually appealing, highlighting important information. Regarding its integration in PEI Visualization, in section 9.2.1.3, there is included a basic bar chart representation of the data obtained from the calculation of the PEI. Also, for future PEI visualization developments other advanced possibilities are considered such as the port's progress during a year compared to a baseline year. As a generic example bar charts representation can be used for visualizing individual eKPIs as shown in Figure 9.1***¡Error! No se encuentra el origen de la referencia.***. One can observe if different indicators composing the PEI, and the PEI itself, have improved or worsen over a time period.

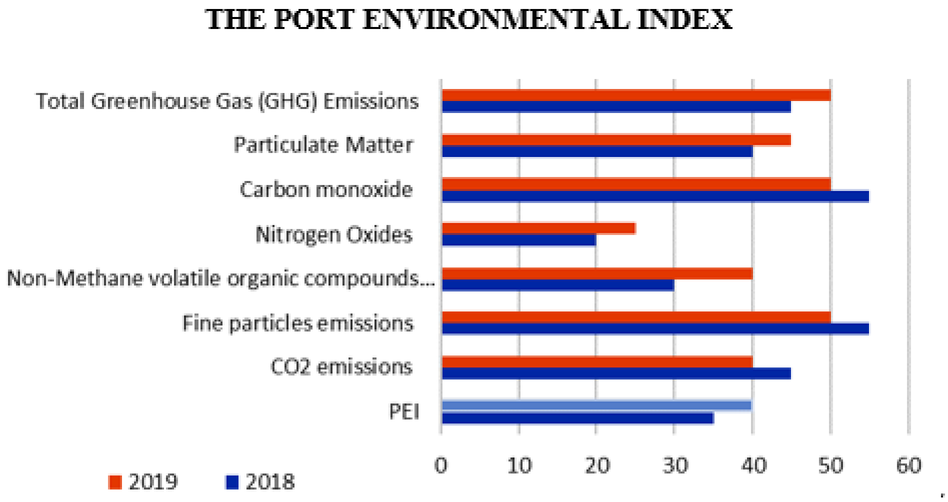


Figure 9.1 Bar chart presentation of the PEI and different eKPIs related to air for the years 2018 and 2019

9.1.3. **Line chart presentations**

Line charts can be used to represent time-series data. Indicators can be displayed using:

- absolute values;
- growth rates, e.g. in percentage points from the previous year or above in previous years;
- indexed levels;
- indexed growth rates.

Line charts can be very useful for the visualization of PEI and the different eKPIs as show in section 9.2.1.3. A user can easily monitor their changing over a given time, for example, monthly evolution of one indicator or index. Also, line charts can be used to observe the environmental performance for different ports at the same time over the years, as in the previous case, only if the ports could share its information (*Figure 9.2*).

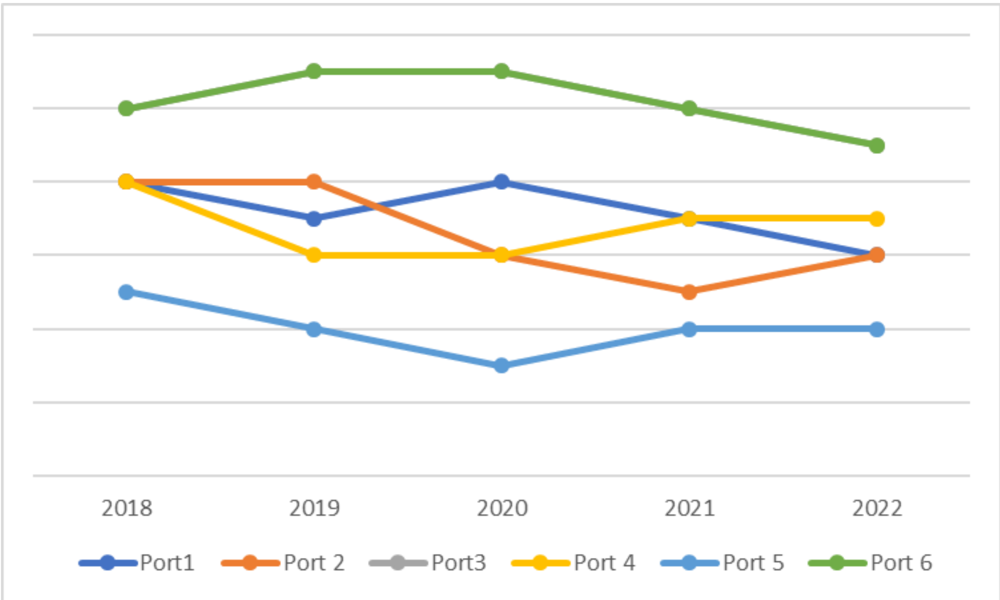


Figure 9.2 Line chart presentation to several ports

9.1.4. Trend diagrams presentations

Composite indicators can also be presented through trend diagrams. Those diagrams can be a sound way for comparing different ports' environmental performance in relation to each other and they can show if a given port is "moving ahead" or "falling behind" compared their previous results. The EU Summary Innovation Index which is used to track the performance of European countries regarding innovations is a great example of trend diagram presentations. In the graph shown below the X-axis represents the EU average and the Y-axis shows the EU trend. The graph area is divided into four quadrants and one can easily observe which country has positive progress and which has a negative compared to their performance during a previous time point.

Trend diagrams are not currently considered as a PEI visualization basic element. But in the future, it could be used to compare eKPI values and their evolution (from a previous time point and now). In this case, each port will have their own trend diagram and they would not be obliged to share this with other ports.

Another possible example would be to compare port's current raking, without real number. Then a trend diagram could be a great tool for the PEI visualization. Ports can easily observe their position in regard to a previous PEI values and their current standing compared to other ports.

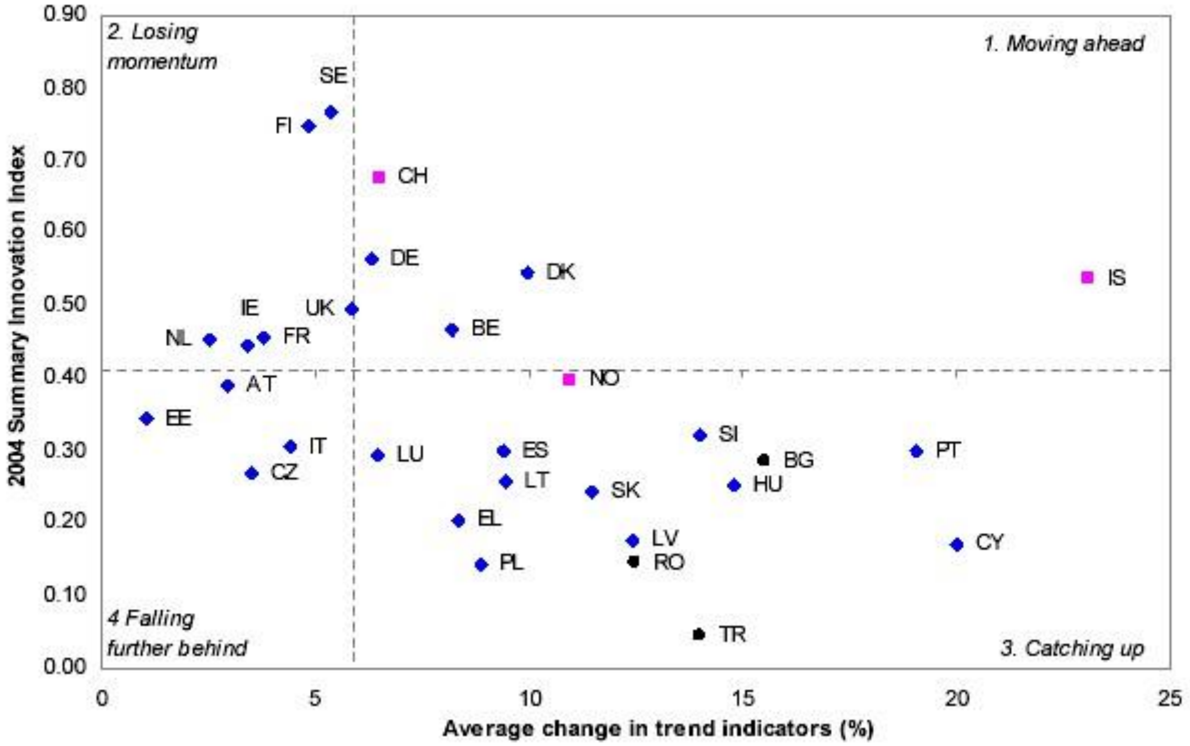


Figure 9.3 The EU Summary Innovation Index represented with a trend diagram

9.1.5. Other types of visualization

There are also available more complicated visualization techniques such as dynamic presentations which can be divided in more sections. For example, sunburst or circumflex charts are a great way to represent different eKPI in a clear manner. The circumflex chart could be divided in more eKPI categories and show their value in each of them. Or related with reliability, show the kind of data input source to obtain its information and make clearer the way of obtaining the reliability rating.

Figure 9.4 Generic example of circumflex chart

The section 9.2.2 include real examples and explanation about other advanced types of representation, such as meter gauges (for the reliability rating in *Figure 9.6*) or flow-directed trees (as in *Figure 9.7*). These elements are going to be used as basic elements of visual PEI visualization.

9.2. Elements for PEI visualization

As indicated in deliverable D5.1 the Port Environmental Index (PEI) is a composite index which integrates all the relevant environmental metrics (KPIs) into a single, easily understandable metric. Composite indices are mathematical aggregations of a set of indicators. The PEI combines different environmental indices, the so-called environmental Key Performance Indicators (eKPIs), into a single metric using an underlying mathematical algorithm.

The section 0 of this deliverable shows a synthesis of the minimum data input necessary to calculate the PEI and the section 5 offers a more complete explanation about its relation with the eKPIs. These eKPIs must be correlated with the impacts of the activities of each port and can be classed by environmental aspects: Emission to the air; Discharges to water; Noise emissions; Production of waste, Energy consumption, and Light emissions. The *Figure 9.5* summarises the link between data inputs, eKPIs, sub-indices, and PEI. These indices are the entire information desired to be transmitted to the user in a visual way through the PEI visualization elements.

Together with the previous explanation, the section 7.5 shows the real reliability/trustworthiness of the value of the PEI due to the current status of digitalization/automation in one port. It exposes that all PEI minimum data input requirements are technically available via:

- Current installed sensors.
- Data existing from web services and applications.
- Data subject to be introduced by forms.
- Simulated/average data not feasible to introduce or measure properly.

That section introduces how these reliability rating will be calculated. In a theoretical way, the general visualization of the PEI would be shown as indicated in *Figure 9.5*. The following sections explain how the visualization tools will represent this rating. The *Figure 9.5* and next figures in this subsection use a range of values from 0 to 100, in contrast with other sections of the deliverable. This decision has been taken in order to provide a more visible values in the initial visualization examples.

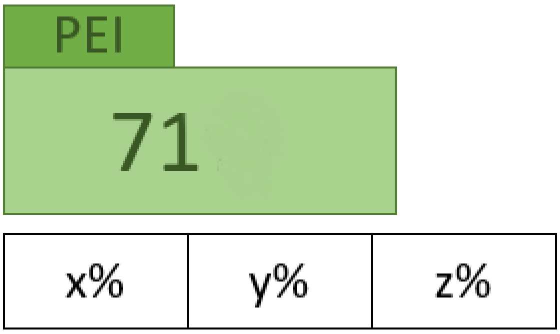


Figure 9.5 Current PEI representation

9.2.1. Essential elements for PEI visualization

This section provides a draft and a tentative plan to describe those elements which bring the first version of the PEI visualization. To address the first version of the graphical interface, the purpose is to offer this content to potential users in a simple and attractive way. This will be done through basic or essential display elements. These visualizations are focused in representing the information related with current and historical values of PEI, sub-indices, eKPIs and reliability rating. These elements cover the visual representation of the main information obtained from the PEI calculation. The next deliverable (D5.3) will include a more detailed description of the interface and new complex and interactive visualization elements.

The following subsections show a brief summary of each basic visualization element.

9.2.1.1. PEI Overview

This design (later to be converted in a webpage running under the umbrella of Dashboard layer of the PIXEL architecture) offers the main information, that is to say, the value of the last PEI execution. Located to its right appears the reliability rating. The sub-indices are located at the bottom. By clicking on each sub-index, it should be possible to visualize the eKPIs associated with it.

The meter gauge shows the value of the PEI, but at the same time indicates the status of the value within a certain range of values. The activity gauge allows to compare the reliability rating. Bar charts offer the information of eKPIs and sub-indices and also have the possibility of including ranges for them.

PEI overview (*Figure 9.6*) is the first visualization that a user will see. The purpose is to provide a sufficient and not excessive amount of information at a first glance. Despite this, all the main information is contained in this screen and users intuitively can access to it.

Figure 9.6 PEI Overview Dashboard

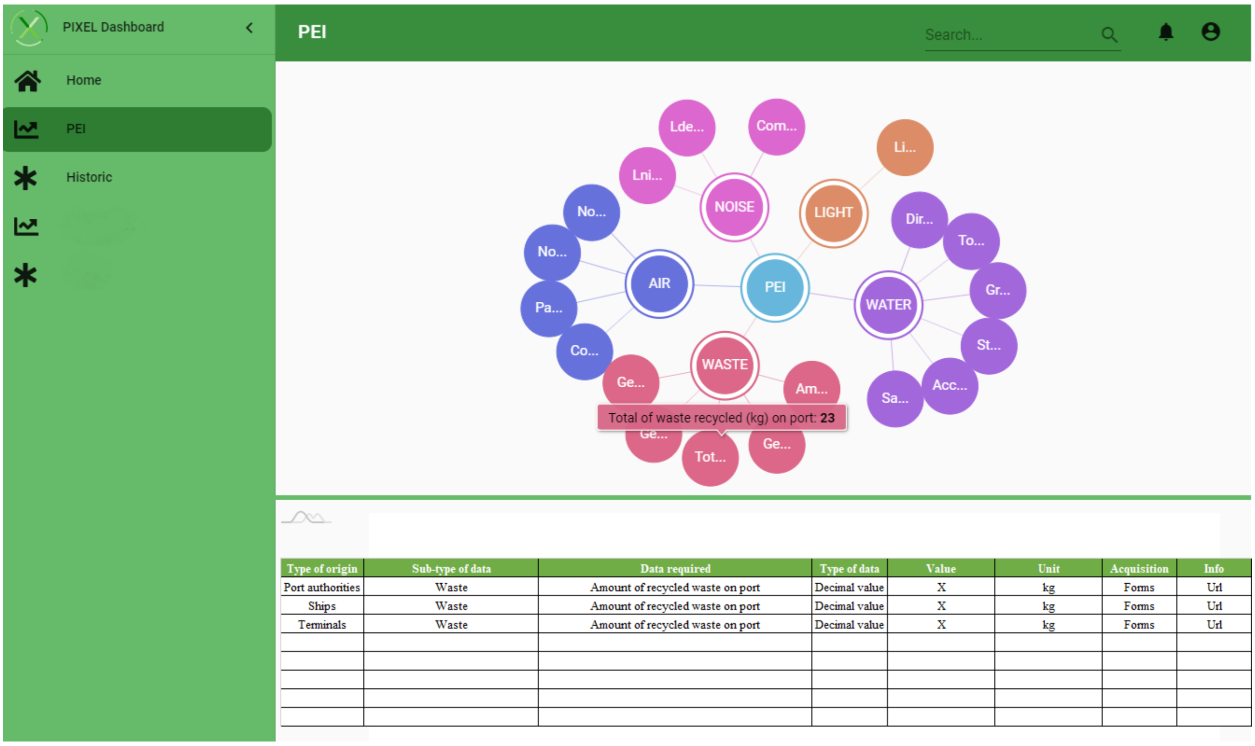
9.2.1.2. Interactive PEI

The PEI interactive version mainly offers the information about the value of the PEI, sub-indices, and eKPIs at a glance. The pointer of the users can be placed on each node and the value will be highlighted (*Figure 9.7*). If they click below, the sub-indices or eKPIs will appear like a bar chart.

Figure 9.7 eKPIs representation through Interactive PEI dashboard

The leaves (*Figure 9.8*) correspond to the information of the eKPI, clicking in on it, there appears a table with the information of each minimum data input value used to calculate this eKPI. The information corresponds to the type of origin, subtype of data, data required, type of data, value, unit and the way of acquisition of that input data.

The user can navigate interactively, discovering the connections between the sub-indices and the eKPIs. They can select the information regarding those nodes that is considered interesting. There are three levels in the data flow, the main or central one corresponds to the PEI, the second level corresponds to the sub-indices, the third level corresponds to the eKPIs. The fourth level that would be the minimum data input to PEI calculation, appears when a user is clicking on the sheets. Drawing minimum data input in the interactive flow would imply to visually overloading the dashboard, so it has been decided to place this last level on a table. This will be easier to be understood when a running mock-up of this visualization would be ready.



Type of origin	Sub-type of data	Data required	Type of data	Value	Unit	Acquisition	Info
Port authorities	Waste	Amount of recycled waste on port	Decimal value	X	kg	Forms	Url
Ships	Waste	Amount of recycled waste on port	Decimal value	X	kg	Forms	Url
Terminals	Waste	Amount of recycled waste on port	Decimal value	X	kg	Forms	Url

Figure 9.8 Minimum data inputs representation through Interactive PEI dashboard

The main objective of this graph is to visually demonstrate that the PEI is a composite index. Users can understand the connections between each of the elements that have been used for their calculation. Advanced options of this visualization will allow features such as modifying the indices values and run PEI calculation with simulated results.

9.2.1.3. Historic value of PEI/sub-indices/eKPIs

Through this visualization, the users obtain access a drop-down menu to choose an option between the PEI, sub-indices, eKPIs, or reliability rating. Once the option is selected, the periodic evolution of the value will be offered over a period of time selected by the user. By clicking on the specific period, the information at that moment of the sub-indices or eKPIs that led to the calculation of the value is shown below. This visualization allows to recover, analyze and compare the historical data of previous executions of the PEI.

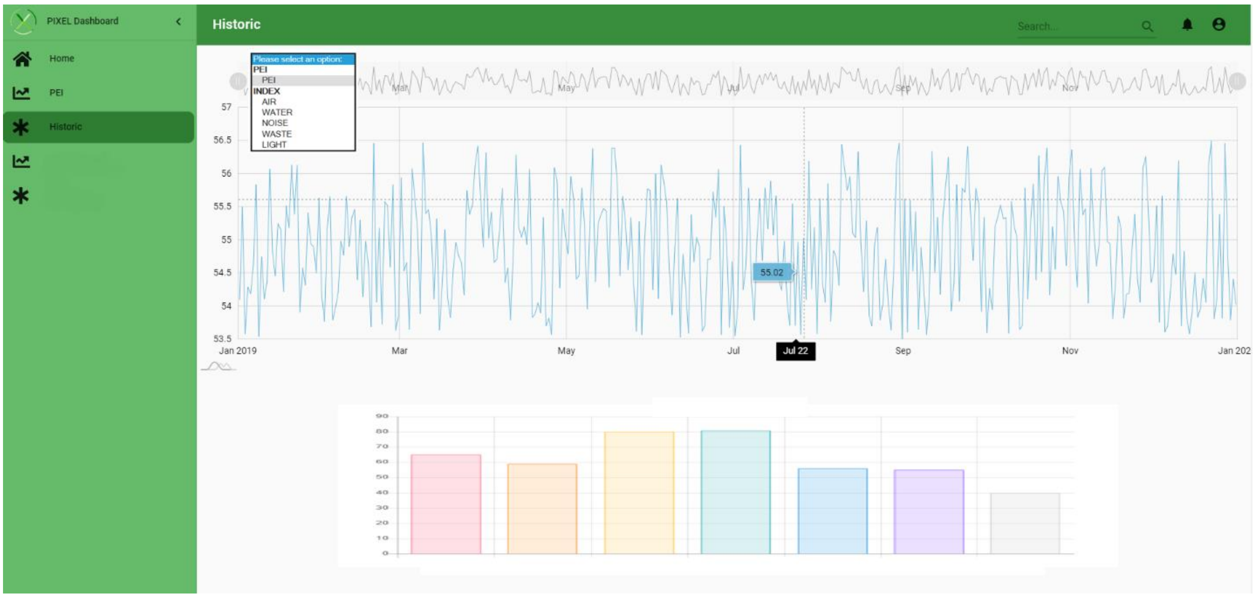


Figure 9.9 PEI Historical visualization

9.2.1.4. PEI executions

This screen shows a summary of current and previous PEI executions. It provides a log register with information about the identifier of the execution, who executed it, by when and whether the event was scheduled or not. It also offers the possibility of launching a PEI calculation.

Through this screen, users can check when the PEI has been executed to take into account the periodicity in which this information has been obtained. The visualization in the form of a table allows, among other things, sort by dates or users. This screen will be modified in the future with new features indicated in section 9.2.2.

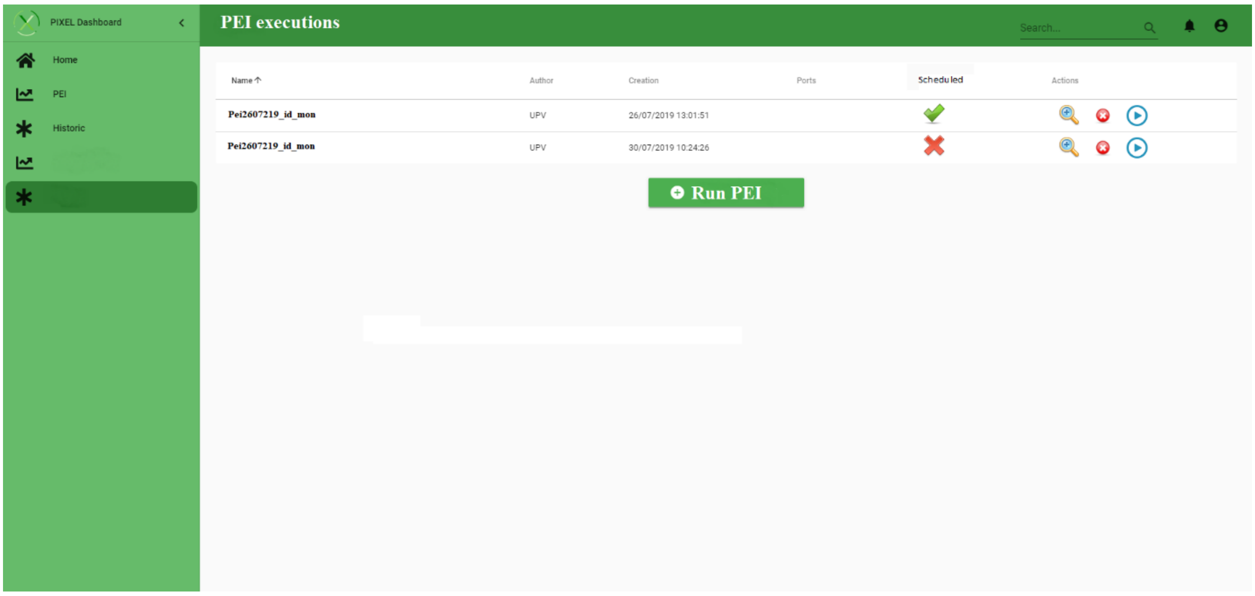


Figure 9.10 PEI executions dashboard

9.2.2. Advanced elements for PEI visualization

These visualizations add extra value to the data obtained from PEI calculation. They provide more complex technical visual features than the basic elements. Their purpose is not only showing the data but to provide extra information and to enable users to interact with all the available information in order to obtain through the visualization elements new flows of information. This will allow users to interact with the visual processing of the information. Users can take into account the information of these advanced elements in order to make decisions or obtain a more complete analysis of the current or simulated situation of a port.

9.2.2.1. Advanced PEI features

These advanced elements are going to be listed in this subsection, but in the next deliverable (D5.3) they will to be extended and fully explained.

- Historical Comparison of values: It allows comparing the PEI, eKPIs, sub-indices or reliability rating between different time periods. The information will be able to be displayed through different graphics or overlapped in the same graphic. It makes easier to analyze and compare the results. Historical Comparison of values improves basic historical visualization because a potential user could interact with different periods of time and data at the same time offering more complex graphs.
- Load a previous PEI status: It allows loading a previous temporary configuration of the PEI, so the users can view the menus as if they were in that previous temporary period. It is an interesting option because it enables the interactive visualization options with the values of a previous PEI execution.
- Reliability: It will offer a complete table with the input data sources. This will show the comparison between the type of source from which the data was obtained and the ideal source from which it could be obtained. Basically, it will show a visual summary of how the calculation of the reliability rating has been carried out, which has been explained in section 7.4. In addition, it enables to access to specific and detailed information about the data inputs for each eKPI. A user could access to it through the interactive PEI menu, by clicking on the desired eKPI.
- Creation of hypothetical scenario: There will be available several forms to modify sub-indices, eKPIs, and data inputs values, in order to be able to launch a new execution of the PEI with this new data creating a new scenario. From the interactive PEI screen by clicking on the desired node, the user will have the possibility to access the forms and enter the new values. This execution will be available in the PEI executions menu; through this table a user can load the PEI status with the values of that simulated execution.
- What if: It shows the graphs resulting after the creation of a hypothetical scenario A potential user could analyze the results and be able to compare with the real values of the PEI. This functionality offers visualization tools to observe at a glance what would happen if the PEI had the values indicated, compared to its current values.
- Alerts: Alerts will be displayed in the home page through a list with a log format. This list will show when what item and why a user has been notified. It is possible to define ranges in the visualization graphs to know how close the value of a data is to trigger an alert.
- Recommendations: These advanced elements are in charge of formalise and list guidelines and recommendations to operate PIXEL properly, gathering valuable knowledge about port environmental indicators suitable to every interested stakeholder.

10. Conclusion / Future Work

In this deliverable, the baseline logic for using a composite indicator for assessing the environmental performance of small and medium-sized port has been described including the eKPIs and a statistical toolbox which will be used for building the computation algorithm. In addition, a general procedure for linking the PEI algorithms to IoT has been explained and detailed. Our future work in D5.3 PEI and algorithms v2 will include defining the exact statistical approaches to be used. In addition, in D5.3 the exact procedure for PEI data retrieval through IoT sources will be explained including an executable for PEI computation to be included in the PIXEL ICT infrastructure.

References

Abson DJ, Dougill AJ, Stringer LC. Spatial mapping of socio-ecological vulnerability to environmental change in Southern Africa. Sustainability Research Institute, School of Earth and Environment, The University of Leeds, Leeds, UK. 2012 Jun.

African Development Fund. Djibouti - Doraleh Container Terminal Port Project (Environmental and Social Impact Assessment Summary (Esias)). Private Sector Department Infrastructure and PPP Division (OPSM.3). 2008 May

Anton Paar. Viscosity of Bunker Oil – Marine Fuel Oil. Anton Paar Wiki. 2019 [cited 2019 Oct 10]. Available from: https://wiki.anton-paar.com/en/bunker-oil-marine-fuel-oil/

Autorità Portuale di Livorno. Dichiarazione Ambientale 2012 – 2015. Autorità Portuale di Livorno. 2012

Bachvarova E, Spasova T, Marinski J. Air Pollution and Specific Meteorological Conditions at the Adjacent Areas of Sea Ports. IFAC-PapersOnLine. 2018 Jan 1;51(30):378-83.

Bailey D, Plenys T, Solomon GM, Campbell TR, Feuer GR, Masters J, Tonkonogy B. Harboring pollution: The dirty truth about US ports. 2004

Bailey D, Solomon G. Pollution prevention at ports: clearing the air. Environmental impact assessment review. 2004 Oct 1;24(7-8):749-74.

Balica S. Approaches of understanding developments of vulnerability indices for natural disasters. Environmental Engineering and Management Journal. 2012 May 1;11(5):963-74.

Ball I. Port waste reception facilities in UK ports Iwan Ball. Marine Policy. 1999 Jul 1;23(4-5):307-27.

Barnett J, Lambert S, Fry I. The hazards of indicators: insights from the environmental vulnerability index. Annals of the Association of American Geographers. 2008 Feb 5;98(1):102-19.

Bjerkan KY, Seter H. Reviewing tools and technologies for sustainable ports: Does research enable decision making in ports?. Transportation Research Part D: Transport and Environment. 2019 Jul 1;72:243-60.

Booysen F. An overview and evaluation of composite indices of development. Social indicators research. 2002 Aug 1;59(2):115-51.

Bullen PS. Handbook of means and their inequalities. Springer Science & Business Media; 2013 Apr 17.

Burgass MJ, Halpern BS, Nicholson E, Milner-Gulland EJ. Navigating uncertainty in environmental composite indicators. Ecological indicators. 2017 Apr 1;75:268-78.

Burgherr P. In-depth analysis of accidental oil spills from tankers in the context of global spill trends from all sources. Journal of hazardous materials. 2007 Feb 9;140(1-2):245-56.

Butt N. The impact of cruise ship generated waste on home ports and ports of call: A study of Southampton. Marine Policy. 2007 Sep 1;31(5):591-8.

Campolongo F, Cariboni J, Saltelli A. An effective screening design for sensitivity analysis of large models. Environmental modelling & software. 2007 Oct 1;22(10):1509-18.

Casazza M, Lega M, Jannelli E, Minutillo M, Jaffe D, Severino V, Ulgiati S. 3D monitoring and modelling of air quality for sustainable urban port planning: Review and perspectives. Journal of Cleaner Production. 2019 May 28.

Chen D, Wang X, Nelson P, Li Y, Zhao N, Zhao Y, Lang J, Zhou Y, Guo X. Ship emission inventory and its impact on the PM2. 5 air pollution in Qingdao Port, North China. Atmospheric Environment. 2017 Oct 1;166:351-61.

Cherchye L, Moesen W, Rogge N, Van Puyenbroeck T. An introduction to 'benefit of the doubt'composite indicators. Social indicators research. 2007 May 1;82(1):111-45.

Chirmata A, Ichou IA. Odour Impact Assessment by Means of Dispersion Modeling, Dynamic Olfactometry and Mobile Electronic Nose around Agadir Fishing Port in Morocco. Journal of Environmental Protection. 2016 Nov 7;7(12):1745.

Collen BE, Loh J, Whitmee S, McRAE LO, Amin R, Baillie JE. Monitoring change in vertebrate abundance: the Living Planet Index. Conservation Biology. 2009 Apr;23(2):317-27.

Corson LA, Fisher SA. Manual of best management practices for port operations and model environmental management system. Purdue University. 2009

Darbra RM, Pittam N, Royston KA, Darbra JP, Journee H. Survey on environmental monitoring requirements of European ports. Journal of Environmental management. 2009 Mar 1;90(3):1396-403.

Darbra RM, Ronza A, Stojanovic TA, Wooldridge C, Casal J. A procedure for identifying significant environmental aspects in sea ports. Marine pollution bulletin. 2005 Aug 1;50(8):866-74.

de la Lanza Espino G, Rodríguez IP, Czitrom SP. Water quality of a port in NW Mexico and its rehabilitation with swell energy. Marine pollution bulletin. 2010 Jan 1;60(1):123-30.

Di Vaio A, Varriale L, Alvino F. Key performance indicators for developing environmentally sustainable and energy efficient ports: Evidence from Italy. Energy policy. 2018 Nov 1;122:229-40.

Dulebenets MA. The green vessel scheduling problem with transit time requirements in a liner shipping route with Emission Control Areas. Alexandria engineering journal. 2018 Mar 1;57(1):331-42.

Elsahragty M, Kim JL. Assessment and strategies to reduce light pollution using geographic information systems. Procedia engineering. 2015 Jan 1;118:479-88.

EcoPorts | ESPO. Top 10 Environmental Priorities of EU Ports 2018 - EcoPorts | ESPO. 2018 [cited 2019 Oct 10]. Available from: https://www.ecoports.com/publications/top-10-environmental-priorities-of-eu-ports-2018

ESPO. Port Performance I Port Performance Indicators: Selection and Measurement indicators. European Sea Ports Organization (ESPO). 2012

ESPO. Environmental Report 2018. EcoPortsinSights 2018. European Sea Ports Organization (ESPO). 2018

Esty DC, Levy MA, Srebotnjak T, de Sherbinin A, Kim CH, Anderson B. Pilot 2006 environmental performance index. New Haven: Yale Center for Environmental Law & Policy. 2006:356.

European Commission. Note on composite indicators. Brussels. 2002

Gan X, Fernandez IC, Guo J, Wilson M, Zhao Y, Zhou B, Wu J. When to use what: Methods for weighting and aggregating sustainability indicators. Ecological indicators. 2017 Oct 1;81:491-502.

Gómez AG, Ondiviela B, Puente A, Juanes JA. Environmental risk assessment of water quality in harbor areas: a new methodology applied to European ports. Journal of environmental management. 2015 May 15;155:77-88.

Gothenburg Port Authority. Sustainability Report 2017. Gothenburg Port Authority. 2017

Grote M, Mazurek N, Gräbsch C, Zeilinger J, Le Floch S, Wahrendorf DS, Höfer T. Dry bulk cargo shipping—An overlooked threat to the marine environment?. Marine pollution bulletin. 2016 Sep 15;110(1):511-9.

Hahn, MB. The livelihood vulnerability index: A pragmatic approach to assessing risks from climate variability and change (thesis). Master of Public Health. Rollins School of Public Health. Emory University Atlanta. Georgia. 2008.

Han CH. Strategies to reduce air pollution in shipping industry. The Asian Journal of Shipping and Logistics. 2010 Jun 1;26(1):7-29.

Hippinen I, Federley J. Fact-finding study on opportunities to enhance the energy efficiency and environmental impacts of ports in the Baltic Sea Region. Motiva Services Ltd. 2014.

Hsu A, Esty D, Levy M, de Sherbinin A. The 2016 environmental performance Index report. Technical report, New Haven, CT: Yale Center for Environmental Law and Policy. 2016.

ITOPF. Oil tanker spill statistics 2016. The International Tanker Owners Pollution Federation Limited. 2017.

Jahan S. Human Development Report 2016-Human Development for Everyone. 2017 Aug.

Jiang W, He G, Long T, Guo H, Yin R, Leng W, Liu H, Wang G. Potentiality of using Luojia 1-01 nighttime light imagery to investigate artificial light pollution. Sensors. 2018 Sep;18(9):2900.

Joint Research Centre-European Commission. Handbook on constructing composite indicators: methodology and user guide. OECD publishing; 2008 Aug 22.

Kaufmann D, Kraay A. Governance indicators: Where are we, where should we be going?. The World Bank; 2007 Nov 9.

Kaufmann D, Kraay A, Mastruzzi M. The worldwide governance indicators: methodology and analytical issues. Hague Journal on the Rule of Law. 2011 Sep;3(2):220-46.

Kegalj I. The Evaluation of the Influence of Harbour Processes on the Environment by Forming an Environmental Index (doctoral dissertation). Tehnički fakultet Sveučilišta u Rijeci. Croatia. 2016.

Kegalj I, Traven L, Bukša J. Model of calculating a composite environmental index for assessing the impact of port processes on environment: a case study of container terminal. Environmental monitoring and assessment. 2018 Oct 1;190(10):591.

Kenney MA, Maldonado J, Chen RS, Quattrochi D. Climate change impacts and responses: Societal indicators for the national climate assessment. 2012.

Land KC. The Foundation for Child Development Child and Youth Well-Being Index (CWI), 1975-2004, with projections for 2005. Durham, NC: Duke University. 2006;318.

Lee L, Anderson R. A comparison of compensatory and non-compensatory decision making strategies in IT project portfolio management. International Research Workshop on IT Project Management 2009. 2009 Dec 14 (p. 9).

Little RJ, Rubin DB. Statistical analysis with missing data. John Wiley & Sons; 2019 Apr 23.

Loh J, Green RE, Ricketts T, Lamoreux J, Jenkins M, Kapos V, Randers J. The Living Planet Index: using species population time series to track trends in biodiversity. Philosophical Transactions of the Royal Society B: Biological Sciences. 2005 Feb 28;360(1454):289-95.

Loh J, Randers J, MacGillivray A, Kapos V, Jenkins M, Groombridge B. Living Planet Report, 1998. Gland, Switzerland: WWF International, New Economics Foundation. World Conservation Monitoring Centre. 1998.

Maggino F, Zumbo BD. Measuring the quality of life and the construction of social indicators. Handbook of social indicators and quality of life research 2012 (pp. 201-238). Springer, Dordrecht.

Martínez-Moya J, Vazquez-Paja B, Maldonado JA. Energy efficiency and CO2 emissions of port container terminal equipment: Evidence from the Port of Valencia. Energy Policy. 2019 Aug 1;131:312-9.

McCarthy M. Investigation into reducing odour emissions from partly loaded sheep vessels whilst in port. Project LIVE. 2003;213.

Mohee R, Surroop D, Mudhoo A, Rughooputh BK. Inventory of waste streams in an industrial port and planning for a port waste management system as per ISO14001. Ocean & coastal management. 2012 Jun 1;61:10-9.

Molle F, Mollinga P. Water poverty indicators: conceptual problems and policy issues. Water policy. 2003 Oct 1;5(5-6):529-44.

Morris MD. Factorial sampling plans for preliminary computational experiments. Technometrics. 1991 May 1;33(2):161-74.

Munda G. Multiple criteria decision analysis and sustainable development. Multiple criteria decision analysis: State of the art surveys 2005 (pp. 953-986). Springer, New York, NY. 2005.

Munda G. Social multi-criteria evaluation: Methodological foundations and operational consequences. European journal of operational research. 2004 Nov 1;158(3):662-77.

Munda G, Nardo M. Constructing consistent composite indicators: the issue of weights. EUR 21834 EN. 2005.

Nardo M, Saisana M, Saltelli A, Tarantola S. Tools for composite indicators building. European Comission, Ispra. 2005 Jun 24;15:19-20.

Olson PH. Handling of waste in ports. Marine pollution bulletin. 1994 Jan 1;29(6-12):284-95.

Ondiviela B, Juanes JA, Gómez AG, Sámano ML, Revilla JA. Methodological procedure for water quality management in port areas at the EU level. Ecological indicators. 2012 Feb 1;13(1):117-28.

Oţoiu A, Grădinaru G. Proposing a composite environmental index to account for the actual state and changes in environmental dimensions, as a critique to EPI. Ecological indicators. 2018 Oct 1;93:1209-21.

Paipai E. Guidelines for port environmental management. HR Wallingford. 1999

Parker JD. Environmental reporting and environmental indices (doctoral dissertation). University of Cambridge. 1991

PEMA. Lighting Technologies in Ports and Terminals: A PEMA Information Paper. PEMA. 2016 May

Pérez I, González MM, Jiménez JL. Size matters? Evaluating the drivers of waste from ships at ports in Europe. Transportation Research Part D: Transport and Environment. 2017 Dec 1;57:403-12.

Peris-Mora E, Orejas JD, Subirats A, Ibáñez S, Alvarez P. Development of a system of indicators for sustainable port management. Marine Pollution Bulletin. 2005 Dec 1;50(12):1649-60.

Pocora A, Lupu S, Toma A, Katona C. Light Pollution in Constanta: A Case Study. “Mircea cel Batran” Naval Academy Scientific Bulletin. 2015;19(2):71-75.

Puig M, Darbra RM. The Role of Ports in a Global Economy, Issues of Relevance and Environmental Initiatives. World Seas: an Environmental Evaluation 2019 Jan 1 (pp. 593-611). Academic Press.

Puig M, Pla A, Seguí X, Darbra RM. Tool for the identification and implementation of Environmental Indicators in Ports (TEIP). Ocean & coastal management. 2017 May 1;140:34-45.

Puig M, Wooldridge C, Darbra RM. Identification and selection of environmental performance indicators for sustainable port development. Marine pollution bulletin. 2014 Apr 15;81(1):124-30.

Puig Durán M. Methodology for the selection and implementation of environmental aspects and performance indicators in ports (doctoral thesis). Universitat Politècnica de Catalunya. 2016

Ravallion M. Mashup indices of development. The World Bank; 2010 Sep 29.

Roos EC, Neto Kliemann FJ. Tools for evaluating environmental performance at Brazilian public ports: Analysis and proposal. Marine pollution bulletin. 2017 Feb 15;115(1-2):211-6.

Saeedi Pash H, Ebadi T, Pourahmadi A, Parhizkar YR. Analysis of Most Important Indices in Environmental Impacts Assessment of Ports. Civil Engineering Journal. 2017 Nov 4;3(10):868-80.

Saharidis GK, Konstantzos GE. Critical overview of emission calculation models in order to evaluate their potential use in estimation of Greenhouse Gas emissions from in port truck operations. Journal of cleaner production. 2018 Jun 1;185:1024-31.

Saisana M. Composite indicators: a review. Second workshop on composite indicators of country performance, OECD, Paris 2004 Feb.

Saisana M, Saltelli A. Sensitivity Analysis for the 2008 Environmental Performance Index. European Commission: Fermi. 2008.

Saisana M, Saltelli A, Tarantola S. Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2005 Mar;168(2):307-23.

Saisana M, Tarantola S. State-of-the-art report on current methodologies and practices for composite indicator development. Ispra, Italy: European Commission, Joint Research Centre, Institute for the Protection and the Security of the Citizen, Technological and Economic Risk Management Unit; 2002 Jan.

Saltelli A, Jesinghaus J, Munda G. Well being stories. Beyond GDP Conference, Experts Workshop 2007 Nov (pp. 19-20).

Saxe H, Larsen T. Air pollution from ships in three Danish ports. Atmospheric environment. 2004 Aug 1;38(24):4057-67.

Sharpe A. Literature review of frameworks for macro-indicators. Ottawa: Centre for the study of Living Standards; 2004 Feb

Styhre L, Winnes H, Black J, Lee J, Le-Griffin H. Greenhouse gas emissions from ships in ports–Case studies in four continents. Transportation Research Part D: Transport and Environment. 2017 Jul 1;54:212-24.

Tarantola S, Saltelli A. Composite indicators: the art of mixing apples and oranges. Composite Indicators-Boon or Bane. Statistisches Bundesamt, Germany. 2007.

TasPorts. Environmental StTownsville Ocean Terminaandard – Light Pollution. TasPorts. 2019 May

Tate E. Uncertainty analysis for a social vulnerability index. Annals of the association of American geographers. 2013 May 1;103(3):526-43.

Tax Justice Network. Financial Secrecy Index 2013. London: Tax Justice Network. 2013.

Townsville Ocean Terminal. Environment Impact Assessment (EIS). Townsville Ocean Terminal Project. 2007 Nov

Trozzi C, Vaccaro R. Environmental impact of port activities. WIT Transactions on The Built Environment. 2000 Sep 7;51.

UNESCAP. Environmental impacts of port development. United Nations: Bangkok. 2009 van Aalst R, Edwards L, Pulles T, De Saeger E, Tombrou M, Tonnesen D. Guidance report on preliminary assessment under EC air quality directives. Copenhagen: European Environment Agency. 1998 Jan.

Van Breemen T. Good Practice Guide on Port Area Noise Mapping and management, Noise Management in European Ports (NoMEPorts) project. LIFE05 ENV/NL/000018; 2008.

Van de Kerk G, Manuel AR. A comprehensive index for a sustainable society: The SSI—the Sustainable Society Index. Ecological Economics. 2008 Jun 15;66(2-3):228-42.

van Haaster B, Ciroth A, Fontes J, Wood R, Ramirez A. Development of a methodological framework for social life-cycle assessment of novel technologies. The International Journal of Life Cycle Assessment. 2017 Mar 1;22(3):423-40.

Viana M, Hammingh P, Colette A, Querol X, Degraeuwe B, de Vlieger I, Van Aardenne J. Impact of maritime transport emissions on coastal air quality in Europe. Atmospheric Environment. 2014 Jun 1;90:96-105.

Walker TR, Adebambo O, Feijoo MC, Elhaimer E, Hossain T, Edwards SJ, Morrison CE, Romo J, Sharma N, Taylor S, Zomorodi S. Environmental effects of marine transportation. World Seas: An Environmental Evaluation 2019 Jan 1 (pp. 505-530). Academic Press.

Wang S, Zhen L, Zhuge D. Dynamic programming algorithms for selection of waste disposal ports in cruise shipping. Transportation Research Part B: Methodological. 2018 Feb 1;108:235-48.

Wibowo TS, Purwanto P, Yulianto B. Analysis of public perception about impaired odor in center fishery processing industry Tegalsari Tegal city. AIP Conference Proceedings 2017 Mar 9 (Vol. 1818, No. 1, p. 020066). AIP Publishing.

Wilhelmsson D, Thompson RC, Holmström K, Lindén O, Eriksson-Hägg H. Marine pollution. Managing ocean environments in a changing climate: sustainability and economic perspectives. Elsevier, Burlington. 2013 Jun 29:127-57.

Wilson MC, Wu J. The problems of weak sustainability and associated indicators. International Journal of Sustainable Development & World Ecology. 2017 Jan 2;24(1):44-51.

Winnes H, Styhre L, Fridell E. Reducing GHG emissions from ships in port areas. Research in Transportation Business & Management. 2015 Dec 1;17:73-82.

Wolff H, Chong H, Auffhammer M. Classification, detection and consequences of data error: evidence from the human development index. The Economic Journal. 2011 Jan 21;121(553):843-70.

World Health Organization. Health effects of particulate matter. Policy implications for countries in Eastern Europe. Caucasus and central Asia. World Health Organization Regional Office for Europe, Copenhagen. 2013.

Zheng S, Ge YE, Fu X, Nie YM, Xie C. Modeling collusion-proof port emission regulation of cargo-handling activities under incomplete information. Transportation Research Part B: Methodological. 2017 Oct 1;104:543-67.

Zhou P, Ang BW. Comparing MCDA aggregation methods in constructing composite indicators using the Shannon-Spearman measure. Social Indicators Research. 2009 Oct 1;94(1):83-96.