MAGPIE

SMART GREEN PORTS

Data Models and Data Analytics for Green Ports

Funded by the European Union

Data Models and Data Analytics for Green Ports D4.4

Grant Agreement No.	101036594
Start date of Project	1 October 2021
Duration of the Project	60 months
Deliverable Number	D4.4
Deliverable Leader	CEA
Dissemination Level	CO
Status	V1.0
Submission Date	30-01-2024
Author	Ariane Piel, CEA, ariane.piel@cea.fr Boris Dartiguepeyrou, CEA, boris.dartiguepeyrou@cea.fr Yves-Marie Bourien, CEA, yves-marie.bourien@cea.fr Eric Francois, CEA, eric.francois@cea.fr Muriel Andurand, CIRCOE, m.andurand@circoe.com Baptiste Olivier, CIRCOE, b.olivier@circoe.com André Lisboa, EDP, andre.lisboa@edp.pt Gloria Goncalves, EDP, gloria.goncalves@edp.pt Gonçalo Calado, EDP, goncalo.calado@edp.pt João Megre, EDP, joao.megre@edp.pt Giovanni de Nunzio, IFPEN, giovanni.de-nunzio@ifpen.fr Zenaida Mourão, INESC TEC, zenaida.mourao@inesctec.pt

	Karol B. Gonçalves, INESC TEC, karol.b.goncalves@inesctec.pt Tomás Rocha, INESC TEC, tomas.rocha@inesctec.pt Adrian Galvez, INESC TEC, adrian.c.galvez@inesctec.pt Ilia Ponomarev, INESC TEC, ilia.ponomarev@inesctec.pt Jorrit Harmsen, TNO, jorrit.harmsen@tno.nl

This project has received funding from the European Union's Horizon 2020 (MFF 2014-2020) research and innovation programme under Grant Agreement 101036594.

Modification Control

VERSION #	DATE	AUTHOR	ORGANISATION
V0.1	27-06-2023	Piel	CEA
V0.2	20-07-2023	Piel, G. de Nunzio, Z. Mourão	CEA, IFPEN, INESC TEC
V0.3	10-08-2023	M. Andurand	CIRCOE
V0.4	07-09-2023	Lisboa, B. Dartiguepeyrou, M. Andurand	EDP, CEA, CIRCOE
V0.5	14-09-2023	J. Harmsen	TNO
V1.0	22-09-2023	All authors	All partners
V2.0	23-01-2024	Piel, B. Dartiguepeyrou, G. de Nunzio, Z. Mourão	CEA, IFPEN, INESC TEC

Release Approval

NAME	ROLE	DATE
Z. Mourão	WP Leader	11-10-2023
L. van der Lugt	Peer reviewer	30-10-2023
J. Pruyn	Peer reviewer	29-10-2023
R. Will	Project Office	29-01-2024
M.B. Flikkema	Scientific Coordinator	30-01-2024
A.J. Polman	Project Coordinator	30-01-2024

Table of content

Executive Summary

This deliverable reports on the work carried out within T4.4 of WP4. The aim within T4.4 is to provide, through modelling and simulation tools, a prediction capability at the service of the tools developed in T4.5. This deliverable describes the back-end models and tools necessary to fulfil this prediction capability. The interactions between these models and tools are represented in Figure 1.

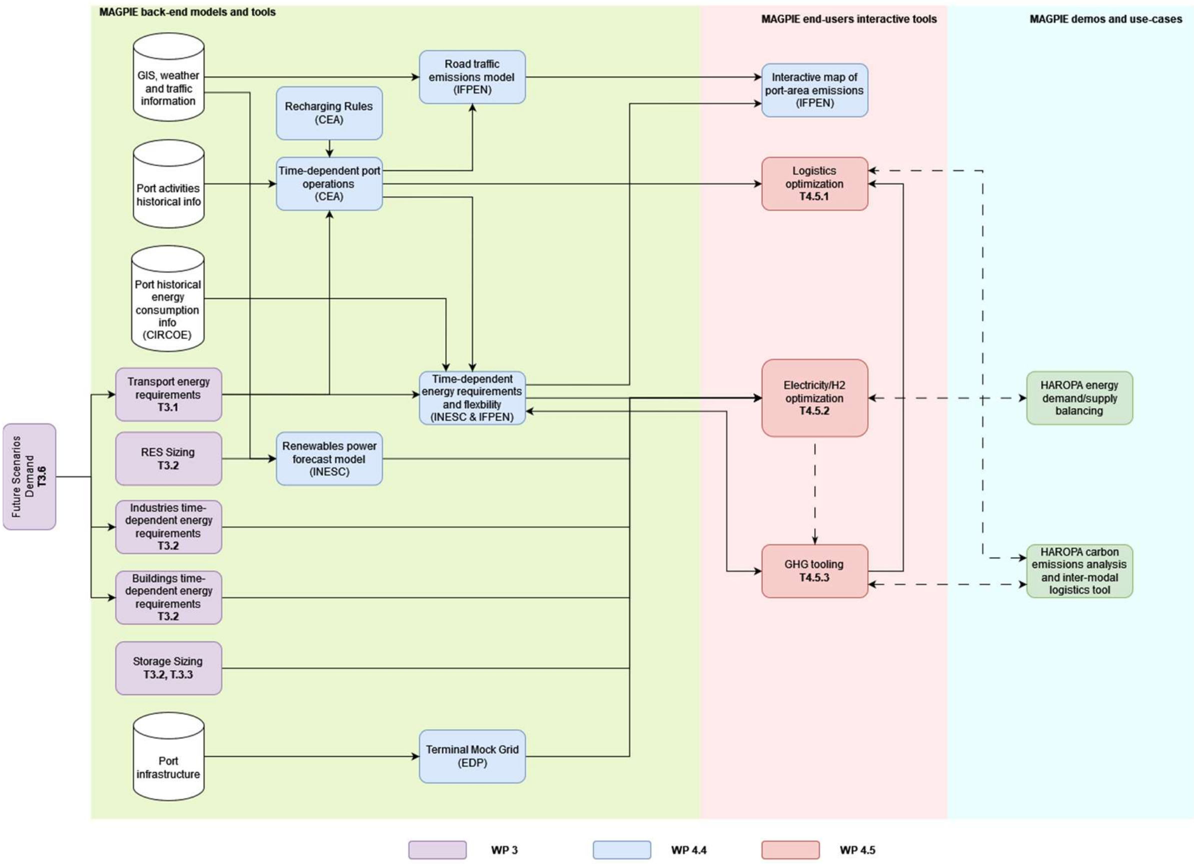

Figure 1 Overall MAGPIE back-end models and tools architecture

In order to generate the time-series describing the logistics, energy and emissions associated with port operations, a multi-dimensional digital port simulator was developed to provide synthetic data of a prototype port terminal in the form of time series. These logistics time-series are then used by the backend models developed in T4.4 to generate the input on demand and supply that will be processed by T4.5 tools.

As shown in Figure 1 time-dependent port operations will be simulated by a terminal simulator, "Proto Port", which is augmented with an energy dimension, relying on results from T3.2 and T3.3, to provide time-dependent energy requirements (time series of port operations with associated energy demand). Complementary models reflect energy forecast and status with a mock grid and renewables power and forecast. In addition, a large-scale dynamic emission model and map discloses the road and port CO_2 and pollutant emissions per unit time on every road link, relying on their traffic volume model (number of road or port vehicles per unit time on every road link) and the aforementioned time-dependent energy requirements.

1. Introduction

According to the aim of the task (ANNEX 1) several models were developed in T4.4. that describe the main infrastructural, logistics and operations domains, providing physical and data models that represent dynamically logistics, energy use and emissions for different systems in ports – e.g., energy use of vessels at berth, allocation of cargo handling equipment, movements of assets in a terminal associated with loading and unoading of cargo, rules for attribution of charging infrastructure, energy use of cranes, emissions of trucks involved in cargo handling activities from the port, impact of electrification of assets in terminals.

The physical models were developed from literature and by enhancing pre-existing models. The dynamic time-series describing the energy and emissions of the logistics operations would in principle rely on available logistics data and detailed characterization of all infrastructure in the port, which , as explained previously in deliverables D4.1, D4.2 and D4.3, is difficult to obtain for all relevant port operations. Thus, to overcome this issue, the MAGPIE partners involved in T4.4 developed a multi-dimensional digital port simulator, which can provide synthetic data of a prototype port terminal in the form of time series. Starting from these time series of terminal operations, the backend models can be used to forecast energy demand and associated emissions of different port assets which are needed as input for tools in T4.5. For local energy supply and storage , a set of backend models have been developed that rely on data that can be generated through physically based models and on data generated by the terminal simulator.

Even though only the application of the models to container terminals is described in detail in this deliverable, all of the backend models developed in T4.4. are flexible and modular and can be applied to other types of terminals. Some of the models that can be used to calculate the energy demand of assets used in operations in other types of terminals are described in ANNEX 2.

This deliverable reports on the work carried out within T4.4 of WP4. The deliverable fulfils the aim of T4.4 to provide, through modelling and simulation tools, a prediction capability at the service of the tools developed in T4.5.

1.1. Context and objectives

The implementation of the digital tools in MAGPIE is planned as a 3-tiered approach. The digital sharing infrastructure described in deliverables D4.2 and D4.3, constitutes the first tier, providing the main infrastructure that supports the collection and storage of data, produces meaningful insights for decision making, creates the protocols to share data between stakeholders and tools and standardizes the way in which information is shared and used. The second tier includes the intelligence and analysis layer that takes raw data to produce the inputs and models needed for the third tier digital tools that will be developed in T4.5 to enable sustainable port energy use and operations, and a greener transport supply chain. It is precisely the layer of back-end models that constitute the "Data models and analytics" layer that are being developed as part of Task 4.4 and which are described in this document.

As set out in deliverables D4.2 and D4.3, in order to overcome the difficulty in obtaining port and terminal specific data needed for the development of the MAGPIE digital tools, a general framework of a "Proto Port" is being used as the basis for the development of the data sharing models and ontology (T4.2, T4.3), the back-end models and tools (as per Figure 2) (T4.4) and, the digital tools to be developed in T4.5. The "Proto Port" is a proxy structure that represents a typical container terminal, where synthetic data of logistics operations at the level of the terminal assets are generated by a terminal simulator tool, developed by

CEA in T4.4. Building from this basis, the work in task 4.4 had the following main goals (as per original description available in ANNEX 1):

1. Development of a multi-dimensional digital model of ports that provides a holistic, human-centric and simulation-based model for matching energy demand and supply,
2. Development of physical and data-driven models of the different systems in ports allowing their dynamic representation in time,
3. Developing and applying ML and AI models to predict the use of resources and power generation in ports, to support the optimization and decision-making tools developed within task T4.5,
4. Developing models of energy consumption and emissions of the main operations/assets/vehicles/vessels linked to the movement of goods and people within the port area,
5. Enhancing the DeCaMod and DeCaMod2 projects to map emissions related with international shipping, operation of hubs and terminals and other relevant areas.

1.2. Work Package Dependencies

The activities carried out generally within WP4 interact with several activities and outputs in other MAGPIE work packages. For the work developed in Task 4.4, the following links are relevant:

- WP4 is closely connected with WP3, as mapped in Figure 2. Particularly,
  - The work carried out in Task 3.1 that quantified the current and future energy demand of transport modalities can provide inputs on emission factors of different technologies for ocean going vessels, inland water shipping, trains and road freight which can be used as input data for the tools that quantify GHG emissions of transport supply-chains. As deliverable D3.1 covered also future technology shifts, the data can be used for strategic planning and testing decarbonisation scenarios.
  - The work carried out in tasks T3.2-T3.3, that quantified the electricity, and hydrogen supply and demand of transport modalities, buildings, and industries in port areas will be used within WP4. This work has been described in deliverables D3.2 and D3.3. The models and data generated within WP3 will be used to generate time-dependent demand models for buildings, industries, and transport modalities, that complement the terminal oriented back-end models under development in T4.4. One of the main differences is that while the energy supply models in WP3 focus mainly on planning and sizing of future low-emissions and renewable energy production, the digital tools, and back-end models under development in WP4 focus also on operational decision-making and optimization. This means, for example, that the renewable energy forecast models and the energy demand models described in this deliverable focus on the forecast for periods of days or weeks, with high temporal resolution and not years as do the modelling of supply and demand in WP3.
  - Models for sizing of future renewable capacity (e.g., PV, wind) may be used as input for the renewable generation forecast, especially, the physical based models proposed in D3.2 may be used to generate synthetic data for training of the data-based forecast models.
  - The future demand scenarios under development in T3.6 provide information on the decarbonization pathways of transport supply chains, and of the activities in ports, which can be used to run additional scenarios in the terminal simulator tool to then generate the time dependent demand profiles,

emissions, and flexibility potential with and without EES. These would then be used as input for the digital tools under development in T4.5.
    - Finally, the WP3 demos, especially demo 3 (Shore Peak Power shaving) and demo 2 (Smart Energy Systems) could provide data and model parameters relevant for some of the back-end tools under development in T4.4 -e.g., time-dependent demand profiles, flexibility modelling (with and without EES), and the terminal (mock) grid.
- With demos and models under development in WP5 and WP6 especially for ongoing work in demo 7 (Green energy container), demo 9 (Green Connected Trucking) and demo 10 (Spreading road traffic) which could provide use cases of the implementation of the back-end models of T4.4, but also data to feed the models. Additionally, it would be beneficial to compare some of the models from WP5 and WP6 with the ones in development in T4.4 and T4.5, to ensure consistency in the use of input parameters and mathematical formulation of e.g., storage, energy demand and energy supply of specific systems/areas – e.g., trucks and charging in demo 9, battery storage in IWT and demand of IWT in demo 7, strategies for reduction of emissions through management of road freight in demo 10.
- In terms of the non-technical barriers work of WP7 strategies for differentiated and dynamic tariffs for electricity for OPS systems, or to encourage shifting consumption to fit local renewable generation could be relevant to the uptake and modelling of energy demand and supply in the T4.4. back-end models. Similarly, additional use cases and scenarios resulting from the outputs of WP7 could be run in the logistics, energy, and emissions back-end models of T4.4.
- Some of the data being collected and used for KPI monitoring in WP8 could be used as input to the back-end models under development in T4.4. Additionally, some of the KPIs could be used as additional analytics implemented in the post-processing of the terminal simulator tool. Some preliminary post processing results and generic KPIs are shown in Chapter 3.
- The results of the post-processing of outputs of the terminal simulator tool, could be used to test the efficacy of different scenarios for decarbonization of ports and transport supply chains may be relevant for the Master Plan under development in WP9. Additionally, some of the backend models under development in T4.4 could inform the current work on the Vision Elements, in particular those in Groups "New World Energy", "Smart and Efficient", "Future Proof Business Models", and "Nature Positive".
- Finally, the continuing relationship with WP10 in the exchange of practices and exploration of synergies with other sister projects in terms of the contribution of digitalisation of ports to the decarbonizations of its operations.

1.3. Outline

This document is structured following the overall architecture detailing the interactions between the different underlying back-end models and tools (Figure 2):

- Results from WP3 (purple boxes in Figure 2) contribute to the development of the Renewable power forecast models and time-dependant energy requirements, as further detailed in Section 1.2.
- Section 2.1 Terminal simulation tool (Proto Port) generates the synthetic data simulating port terminal activity (Time-dependant port operations).
- Section 2.2 Time dependent energy requirements describes the mathematical formulation that will be used to model the demand time-series for different terminal assets, building from the time-dependant port operations from Proto Port. These time

series constitute the main input data that will be processed by the tools developed in T4.5 (red boxes in Figure 2).

- Section 2.3 CO_2 and emissions mapping constitute a road-traffic emissions model that feeds the interactive mapping tool of port-area emissions described in Section 4.1.
- Section 2.4 Flexibility modelling describes the models that can be used to model the flexibility potential of different terminal assets, without energy storage systems (ESS), building from the time-dependent demand models presented in Section 2.2.
- Section 2.5 Charging rules are necessary to configure part of Proto Port with specific rules concerning recharging strategies. Additionally, quantification of flexibility with EES is also described in this section.
- Section 2.6 Renewable electricity production forecast models describes the models that will be used for renewable power forecast, while Section 2.7 Terminal Grid Modelling describes the terminal grid model. These two models (Renewables Power Forecast Models and Terminal Mock Grid) will provide renewable energy supply and network context to the energy matching and GHG tools developed in T4.5.
- Examples of the implementation of the energy demand and emissions mapping models are shown in Sections 3.1 and 3.2, respectively, where outputs from the Terminal simulator tool are used to estimate energy demand and emissions from OPS systems, from equipment supporting horizontal and vertical movement of cargo in the terminal, and trucks.
- Section 4.1 Interactive dashboard for a comprehensive port-area emissions map showcases the interactive map of port-area emissions that implements the road-traffic emissions model described in Section 2.3.
- Section 4.2 Greenhouse Gas tool details how the model underlying the GHG tool from T4.5 will extend the Decarbonisation Model (Decamod).
- Section 4.3 Carbon simulator tool for ports details data collected from three terminals of HAROPA Port along with a carbon simulator allowing to assess which port operations have the most important CO2 impact. HAROPA data is considered generic and used as an input for the time-dependant energy requirements.

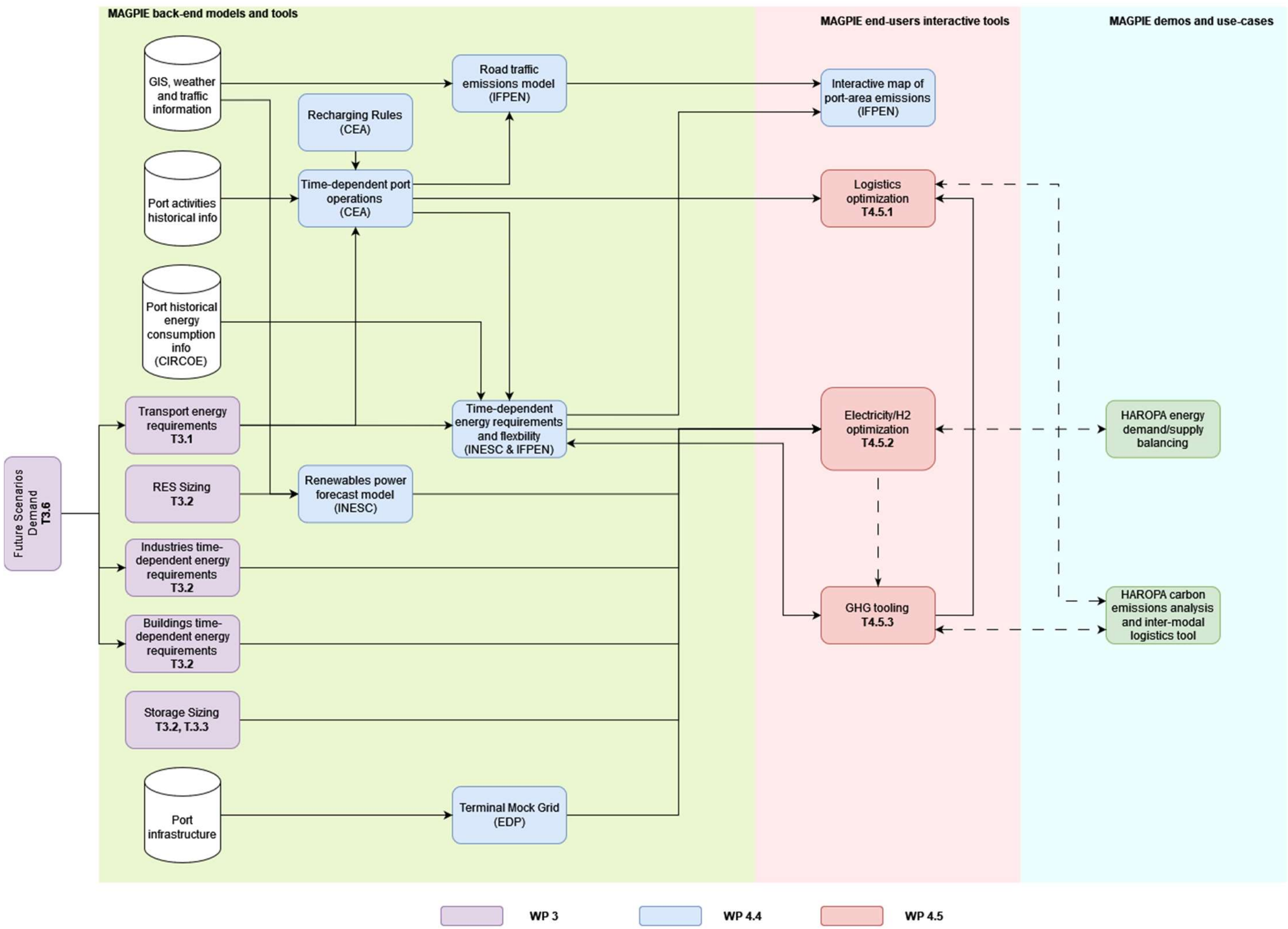

Figure 2 Overall MAGPIE back-end models and tools architecture

2. Modelling and intelligence for Green Ports Digital Twin platforms and services

This chapter describes the main back-end models developed in Task 4.4, as shown in Figure 2. Each section explains the main goals and scope of the models and presents the main mathematical formulation for the model implementation.

2.1. Terminal simulation tool (Proto Port)

As part of a previous project called SONARIS, CEA has developed a model, simulation, and visualization tool of a synthetic, yet generic and representative container terminal.

Participating in the MAGPIE project, we quickly realized that in order to develop energy tools as part of the WP3 and WP4 historical detailed data from the ports was needed. More specifically the tool developers needed quite detailed data such as hourly consumptions or time series of the movements of containers on the port to test their optimization algorithms.

Such data revealed itself to be complex to obtain and none of the port partners was able to provide the required data. CEA, therefore proposed to use the container terminal model (Figure 3) and simulator as a Proto Port to generate the data needed to meet MAGPIE objectives. Indeed, even if the data generated by the simulator is not "real", it is detailed and representative and allows the development of tools that could be then adapted to tackle real data.

This means that two types of data are use, the ones used to configure the model (input data) and the ones generated by the simulation (output data). The first type can be easily compared to real data to ensure the input dataset are representative of the reality. The second type (output data) can be compared to real data with a high level of confidence only of the first type (input data) of data is close enough to real data.

Proto Port is a discreet event simulator that can be used from a web interface.

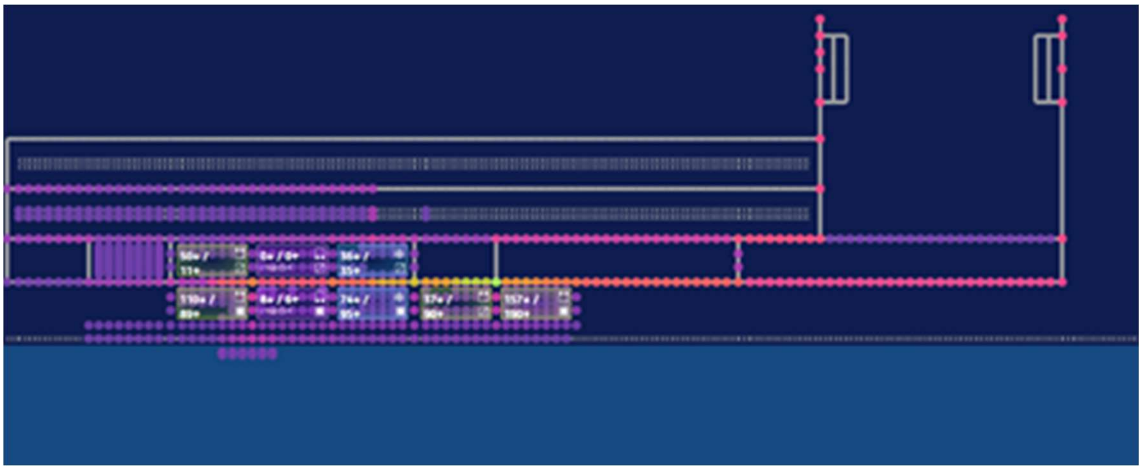

Figure 3 – Visual interface of the terminal simulator tool developed by CEA showing a representation of a container terminal.

It enables the user to load input data for a simulation (e.g., including maps, resources, process information), run a simulation, and visualize the results on various interfaces such as dashboards, maps, and timelines (Figure 4).

The model is detailed enough to simulate all the movements of the different resources (e.g., Gantry cranes, mafis, trucks, ReachStakers) present on a container terminal together with the containers including various types (e.g., reefers, sea containers) and sizes (20" and 40") and transports (trains, railcars, trucks, boats, barges).

Included in the simulation are all the storage strategies on the terminal (height of storage, dedicated storage areas for a given container type), the traffic exceptions (i.e.. some areas allow only terminal equipment traffic; some other parts allow external truck movement). The model also manages the logic related to the mission (jobs) allocation to the resources, this means that the simulation will manage the different priorities given to the different movements depending on many parameters such as resources organization, storage strategy, time of departure of the transport, and type of transport.

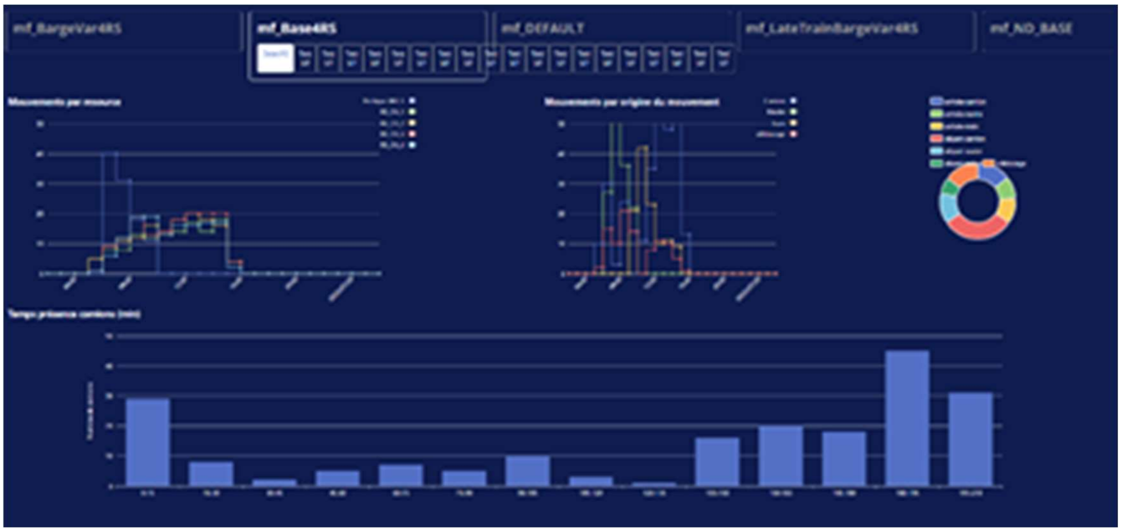

Figure 4 – Example of output as visualized in the user interface of the Proto Port simulator.

Every change in position status is logged into an influx database, and thanks to an ETL (Extract Transform Load) is analyzed for dynamic display on a web interface.

Based on this previous achievement, CEA brought to the MAGPIE project a means to generate synthetic data as if collected from a real terminal through Proto Port.

The current tool architecture can be described in Figure 5.

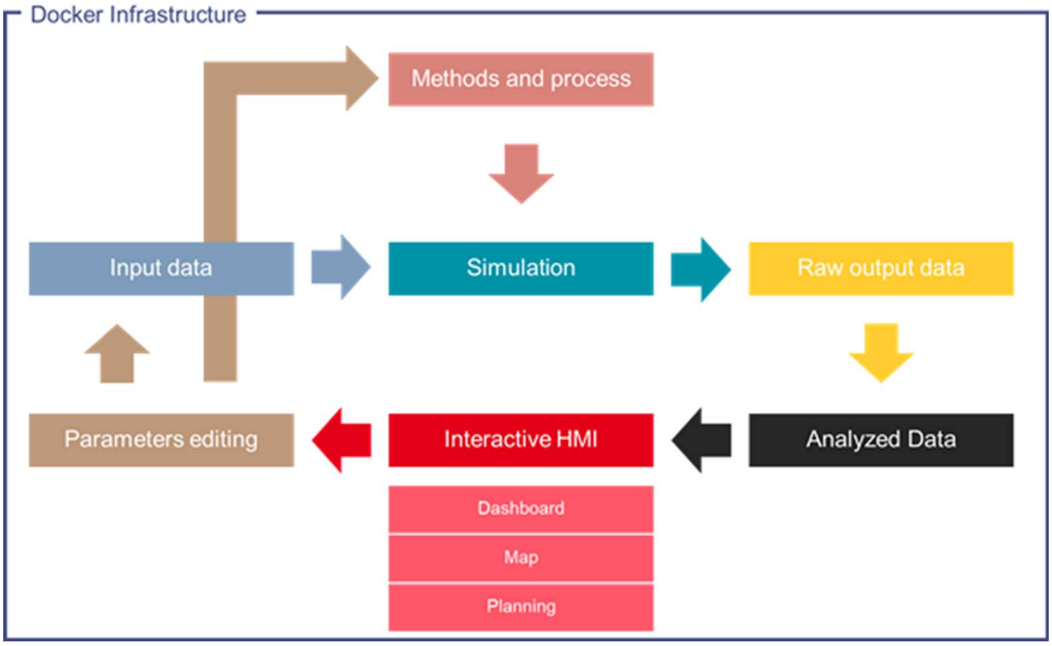

Figure 5 Proto Port tool architecture

As part of MAGPIE, the analyzed data, the interactive HMI and parameters setting are not used.

The input data corresponds to an SQL database where all the information related to the terminal is stored (i.e., Map, list of resources, transport schedule, storage strategy).

The methods and process are related to the rules applied in the model (i.e., Priority parameters, mission patterns).

The simulation is the simulator itself executing the model.

Raw output data is an InfluxDB Database hosting the results of the simulation.

Nevertheless, the tool developed initially in SONARIS required improvements and additional features to meet the requirements expressed by the MAGPIE consortium mainly linked to the integration of the energy consumption parameters as well as the addition of logic linked to the use of electric vehicles. On top of that, additional complexity such as new resource types and mission types had to be introduced.

The following sections describes the work identified to meet those requirements.

Generic Data Generator

CEA, based on the inputs from the different teams involved in the MAGPIE Project, has started developing a new generic Data Generator to be able to generate the input data necessary for the simulator to operate.

More specifically, it appeared to be necessary to be able to modify the activity levels (transport flow) and resource types easily to test multiple scenarios.

Currently the input data is generated manually or can be extracted from an existing infrastructure but considering the need in MAGPIE to generate multiple datasets to be used for the WP3 tools. We started the development of a stand-alone data generator able to populate the input data database.

The generator will be set using high-level parameters such as yearly volume, monthly schedules, container types and category split and seasonality. It will generate a detailed transport schedule and associated lists linking a given container to its arrival and departure transport.

The user of the generator will be able to specify the simulation period.
We noticed that some users requiring a relatively long period to be simulated (1 year) vs. other users requiring only hours or days.
This will be independent from the granularity of the simulation (event based, meaning that the data output doesn't follow a fix time scan but logs information only when something happens) this approach is currently generating data points for every status change in the terminal (in case of high activity can be several data points per second).

12/2023 update: the standalone generator for container terminal has been developed and improved and is currently at v9.

Simulator

The need for a wider variety of resources on the container terminal model was identified, and the teams are currently making modifications to meet this additional complexity.

Specifically, the need to have different types of resources such as straddle carriers or train gantry cranes.

In addition, the need to be able to extract more easily the information related to the energy consumption as well as the carbon emissions, CEA started to implement ways to calculate and extract this information from the models during simulation rather than by post-processing.

We planned to introduce a first evaluation of the energy consumption of the equipment using a polynomial formula based on time, weight of the load and speed of the carrier.

This formula will be using parameters accessible from the input data.

Once the consumption implemented, we will be able to assign for each equipment one or several energy containers (to cover hybrid systems) and subtract during simulation runtime the energy consumed.

Finally, we will have to implement heuristics to handle situations when the energy container is empty or close to empty and trigger recharging actions that can be handled with minor modifications by the current model.

Similarly, another trigger will be implemented to stop the re-charging of the equipment based on the amount of work requiring the resource and level of the battery.

This additional development is technically very accessible with the current model.

Input Data base and data Server

Based on the discussion of WP4, specifically around the ontology, CEA has started the refactoring of the data server of the SONARIS tool. The objective is to reach a point where the integration of the data flow within the MAGPIE dataspace would be easier. More specifically CEA switched to a JPA data server to populate and extract the information from the input database in a more efficient and seamless way. MAGPIE project, as mentioned above will be introducing complexity on the attributes of the equipment moving in the terminal. Each resource will now have additional parameters enabling the user to modify the consumption profile, and the recharging rules. We expect that the number of parameters will need to be adapted several times during the development of the WP4 tools hence we decided to move to this new technology to facilitate the refactoring of the input database structure and consequently make the whole process faster.

Output data structure

The tool output data is stored in an influxDB database. Extraction of the data is accessible via API documented by the influxDB. The analysis of the output data can be complex if the data structure and input data context is not well understood by the user. The objective is to reduce the complexity of data extraction and understanding. We launched an improvement action on the data flow out of the simulator aimed at reducing the number of cross references needed across multiple sources (input and output) to get to a specific information.

Specifically, for MAGPIE we are planning on changing the way the logs are structured to simplify data extraction and understanding, for instance adding more information to the equipment logs (including the energy aspects and more background information) which most probably will be the most used during the MAGPIE project.

It means that we will be using less numbers to describe types for example and rather use literal phrasing. We will also regroup all the output flow used in MAGPIE in a dedicated Influx DB measurement.

Portal for external connections

CEA is also putting in place a web portal enabling the partners of the project to extract the simulation logs via API rather than using export files.

The portal is scheduled to be accessible at the end of 2023 or the beginning of 2024.

Detailed Output data structure

Find below the detailed information of the output data currently available that (as mentioned previously) is being reworked. Currently we are using 3 types of measurements in the influxDB database:

Activity time series

It is a log of the state changes of the missions of the simulator (missions are the jobs executed by the resources moving containers)

Table 1 - Activity time series output of the ProtoPort simulator

Column	description	Unit/format
Time	simulation Time stamp	LocalDateTime YYYY-MM-DD HH:MM:SS
Executant	Reference of the machine executing the mission	String
Type_executant	Type of the executant	String (Reach Stacker, Mafi, Portique, Camion)*
Ref_transport	Reference of the transport related to the mission	String (starting with NA_, TR_, CA_)**
Type_mission	Type of mission	String (ARRIVAL, DEPARTURE, UNBLOCK)***

Ref_produit	Reference of the container	String
Ref_mission	Reference of the mission	String (MIS_ //Ref_produit// _ //Ref_Transport// _ // Type_mission//_ mission index ****
Status	Status of the mission	String (ALLOCATED, TRANSFERT_START, FINISHED, BLOCKED, DELETED, BLOCKED_AT_LOADING, RESTARTED, CANCELLED)

Position Time Series

It is a log of the position changes of all the equipment moving the containers.

Table 2 - Position time series output of the ProtoPort simulator

Column	description	Unit/format
Time	simulation Time stamp	LocalDateTime YYYY-MM-DD HH:MM:SS
ref	Reference of the machine	String
type	Type of the machine	String (Reach Stacker, Mafi, Portique, Camion)*
Coord_X	Y coordinate on map	Number
Coord_Y	Y coordinate on map	Number
noeud	Reference of the node	String

Occupation time series

It is a log of the container stock variations on the yard.

Table 3 - Occupation time series output of the ProtoPort simulator

Column	description	Unit/format
Time	simulation Time stamp	LocalDateTime YYYY-MM-DD HH:MM:SS
emplacement	Reference of the storage spot	String
Noeud	Reference of the node	String
pile	Stack number	Number (1,2,3)
Position_pile	Height in stack	Number (integer)
produit	Reference of the container	String
categorie	Category of the container	String (Caisse Maritime, Caisse Mobile, Reefer, Dangereux)*****
mouvement	Movement direction	+1 (placing container on storage) or -1 (removing container)

*: Reach Stacker = ReachStacker, Mafi = terminal truck, Portique = gantry Crane, Camion = outside truck

**: NA_ = NAvire = Ship, TR_ = TRain = Train, CA_ = CAmion=Truck

***: Arrival = container coming to the terminal, Departure = container leaving terminal, Unblock = Shifting mission (movement to free up a container in a stack)

****: mission index: to take a container to/from a transport from/to a storage spot, machines are working together, and the missions follow one another (ex: crane → truck → RS, in that example, the mission indexes will respectively be 0, 1 and 2)

*****: Caisse Maritime = sea container, Caisse Mobile = mobile box, Reefer = reefer, Dangereux = dangerous goods.

2.2. Time dependent energy requirements

Existing studies of energy consumption and modelling in ports typically focus on specific terminals – e.g., container terminals, groups of equipment or systems supporting logistics operations – e.g., cargo handling equipment in container terminals, or individual systems – e.g., the modelling energy demand of quay cranes. Additionally, port energy demands are highly dynamic and stochastic [1], [2], and influenced by many uncertain factors, such as the port's daily routine, activity handling, and environmental variables (e.g., weather, temperature, and maritime conditions). Furthermore, the specific layout of each port (e.g., type of terminals, handling equipment, building and facilities, services provided) will also influence the energy consumption profile of the port. For instance, reefer containers require constant refrigeration to maintain the quality of the products, consuming a significant amount of energy; therefore, ports which usually handle perishable goods with specific temperature and humidity requirements will typically have higher electricity consumption during warmer seasons. In some terminals, refrigerated containers can account for half of the total electricity consumption by storage yards [3].

Thus, a detailed quantification – via direct measurement or estimation - of the energy consumption of the port, and by equipment and system, is an important stage towards the identification of energy efficiency improvement and decarbonisation measures suitable to the different needs and characteristics of ports. Additionally, there is a need to improve the integration of energy management and real-time operational planning, including a more detailed description of the relationship between the total working time, the actual idle time, and the energy consumption at the level of cargo handling equipment (CHE), and other vehicles and systems supporting terminal operations[4], [5].

This section describes possible approaches to model energy demand in port terminals, with a more detailed description of container terminals, in agreement with the scope of the Proto-port use case and the ProtoPort tool that will be used as the basis for the implementation of the Energy Matching tool. For other types of terminals, a brief description of the characterisation of energy demand is also provided in ANNEX 2.

Modelling the energy demand of different port assets and systems

Considering the scope of the work in T4.4 and the contributions to the tools in T4.5, the models presented in this section cover the energy use of equipment and systems related with activities that provide power to vessels during berth and the activities related with the vertical and horizontal movement and storage of cargo within container terminals. The modelling of the other demand loads such as those for buildings and industry are not covered here as these will be obtained from models developed in WP3, as previously noted. These have been described in the deliverables D3.2 (for electricity) and D3.3 (for hydrogen).

2.2.1. Energy demand of vessels at berth

This section describes the models that can be used to estimate the energy demand for vessels, particularly for their consumption at berth. The energy demand of vessels at the different stages of the port call (e.g., cruising, berthing) will depend on the fuel used, the specific fuel consumption, the engine load and activity patterns for the vessel. To model the consumption of vessels approaching, mooring and berthing at the terminal, existing models can be used [6], [7], [8], [9] and are the basis for equation (1):

$$E = \sum_{i \in I} \sum_{j \in J} P_j l_{ij} t_i, \tag{1}$$

With, E- Vessel energy consumption [in kWh]; I - set of vessels status (activities): cruising, maneuvering, etc.; J - set of engines in the vessels (including main and auxiliaries); P_j - nominal power of engine j [in kW]; l_{ij} - load factor corresponding to engine j during vessel's activity i. Note that l_{ij} = 0 if engine j is not used during activity i; t_i- duration of activity i [in units of time, e.g., hours, minutes].

The nominal power of the engines can be obtained from the manufacturer, while typical/average [10], estimated/predicted [11], simulated [12] or AIS-based [13] values can be used for load factors and the duration of each activity for common vessel types. Several power models for vessels (Propeller Law, Admiralty Law, Holtrop & Mennen method, and Kristensen method) can be used to estimate the load factor, and consequently the energy consumed by vessels. The most used is the Propeller Law [14]:

$$l = \left(\frac{V_{ins}}{V_{max}}\right)^3, \tag{2}$$

where V_{ins} is the ship speed and V_{max} is the maximum speed.

Typical load factors have been proposed for different kinds of vessels. Table 4 shows some examples of these values [9] for ocean going vessels (OGVs) and harbour crafts that can be used for the estimation of the energy demand of vessels, in the absence of real time energy consumption data.

Table 4 - Typical Load Factors (LF) for vessels ([a]ME - Main Engine; [b]AE - Auxiliary Engine; [c]For Tankers) [9].

Vessel Type	Operational mode	[a] LF_{ME}	[b] LF_{AE}
Ocean Going Vessels	Cruising	0.80	0.30
	Manoeuvring	0.20	0.50
	Hotelling	0.20	0.40; 0.60[c]
Assist tug	Not Applicable	0.31	0.43
Commercial fishing		0.27	0.43
Crew boat		0.38	0.32
Excursion		0.42	0.43
Ferry		0.42	0.43
Government		0.51	0.43
Ocean tug		0.68	0.43
Tugboat		0.31	0.43
Work boat		0.38	0.32

The calculation of energy demand profiles associated with the different operations of vessels can thus be carried out using equations (1) and (2). For the particular use case of the Proto Port only the operations of the vessels from the moment of berth are considered. Accordingly, the focus of the rest of this section is the methodology to generate the demand profiles of vessels at berth, as this is specifically considered in the Proto Port use case, and in the Energy Matching Tool (EMT). In particular, the formulation of the consumption of vessels while connected to an Onshore power supply (OPS) is presented below. The use of other types of fuels, e.g., hydrogen, ammonia, is not described here, but could be easily adapted from equations (1) and (2).

Energy demand profiles of vessels connected to an OPS

The Fit for 55 legislative initiative and in particular the FuelEU Maritime [15], set forth ambitious objectives for the use of OPS for all electricity needs of cargo and passenger vessels above 5000 gross tonnage calling at EU ports from 2030. This underscores the pivotal role that OPS technologies are expected to play in the maritime sector and ports, and thus the need to model these loads as part of port renewable energy systems. By supplying power to ships at berth, OPS can lead to reduction in GHG and ambient pollutant emissions, by replacing the use of on-board diesel combustion engines for the needs of the vessels while at berth.

In order to calculate the energy consumption of vessels at berth connected to an OPS (E_{OPS}), equation (1) can be adapted as shown in equation (3) [16], considering the load, rated power of auxiliary engines, berthing time, and the loss rate of the electrical grid (η).

$$E_{OPS} = \sum_{j \in J} P_j l_j t_i (1 + \eta) \tag{3}$$

Accordingly, in order to model and predict the total energy demand for all the vessels connected to the OPS systems in a terminal in a given time interval (e.g., the next 24 to 48 hours), the number of vessels at berth connected to the OPS for a given period of time (t_i) and the (auxiliary) engine characteristics of the vessels must be known. In future ports, data sharing infrastructure (as set out in deliverable D4.2) can be set up to provide this type of information as part of the digital twin of the terminal. This data could be shared directly with any terminal or port energy management tool, such as the EMT (T4.5). When this information is not openly or easily available, two approaches can be implemented to estimate the rated power of the vessels connected to the OPS systems:

1. Considering average power values for the auxiliary engines, according to vessel characteristics,
2. Considering the power of the main engines, based on the vessel's characteristics, and using an auxiliary to main power ratio.

The first approach considers average power ranges for different types of vessels, and can be prone to significant error, as the ranges are quite extensive, as shown in Table 5, which presents some the values considered in the literature. As an example of a reliable source of this information, a comprehensive overview of the loads according to vessel type and size for OPS systems can be found in a recent report by the EPA [17].

Table 5. Average power requirement of vessels at berth [2], [18].

Vessel type	Power
Container	1 MW- 4 MW
Cruise	7 MW
Reefer	2 MW - 5 MW
RoRo	700 kW
Tanker	5 MW - 6 MW
Bulk and General Cargo	300 kW - 6 MW

The second approach uses proportionality factors for different types of vessels, with some of the values found in literature shown in Table 6 [19], [20]. These can be used to calculate the power of the auxiliary engines (P_j) by using the power of the main engines (P_{ME}) and the proportionality ratios (ρ_α) in line with equation (4):

$$P_j = \rho_\alpha P_{ME} \tag{4}$$

Table 6. Numerical values of the auxiliary to main ratio (ρ_α) [19], [20].

Vessel type	ρ_α
Bulk Carrier	0.26
Container	0.22
General Cargo	0.22
Miscellaneous	0.27
Passengers	0.19
RO-RO	0.26
Reefer	0.25
Tanker	0.40
Yacht	0.21

2.2.2. Energy demand of Terminal oriented activities

The energy consumption of terminal activities includes all the activities related to the horizontal and vertical movement of cargo: Quay loading and unloading, quay to storage movements of cargo, storage of cargo, and, finally, receipt-delivery operations. Table 7 provides information on the types of systems and equipment involved in these activities and the type of terminals where these would be used. This is by no means an exhaustive list and is used as an illustration of the type of systems that would have to be modelled to estimate the energy loads associated with cargo handling activities within terminals. When digital twins of ports, representing energy and emissions of different assets, become available, real time energy consumption of all the different equipment and systems, as well as relevant logistics information will be available and tools such as the EMT, or GHG emissions tool would be able to use this information to calculate the best way of matching renewable supply to demand within the port, or to estimate the GHG emissions related with the operations within the port. In the meantime, while real-time monitoring of cargo handling systems and equipment is not ubiquitous, proxy or measured data from existing studies and demos in ports can be used to forecast demand profiles of these operations, using information from logistics (e.g., scheduling). This section describes the modelling of the energy demand of cargo handling activities as presented in Table 7, with particular focus on container terminals.

Table 7. Examples of terminal activities and associated vehicles/equipment/systems.

Activity	Systems/Equipment	Type of Terminal
Quay Loading/Unloading	Quay Crane	Container, General, Break Bulk
	Ship Loader/Unloader	Dry Bulk
	Hopper	
	Payloader	
	Loading Arms	Liquid Bulk
Quay to Storage	Internal Trucks	Container
	Conveyor Systems	Dry Bulk
	Pump Stations	Liquid Bulk
Storage	Refeers	Container
	Yard Cranes	
	Reach Stacker	
	Empty Container Handler	
	Forklift	General, Break Bulk
	Stacker/Reclaimer	Dry Bulk

	Payloader	
	Dust Control Systems	
	Silos	
	Tanks	Liquid Bulk
Receipt-Delivery	Yard Cranes	Container
	Reach Stacker	
	Empty Container Handler	
	Forklift	General, Break Bulk
	Stacker/Reclaimer	Dry Bulk
	Payloader	
	Hopper	
	Silos	
	Tanks	Liquid Bulk
	Shunt Locomotive	All
	Train Terminal Crane	
	External Trucks	
	Barge	

Energy Demand of Terminal Oriented activities in Container Terminals

Two main approaches have been previously used to analyse energy consumption trends of cargo handling systems and equipment during successive periods, e.g., years in container terminals. The approaches are represented in Figure 6, where the selected throughput indicator (TEU or containers) per unit of time is used to estimate the energy or the time needed for processing one throughput unit:

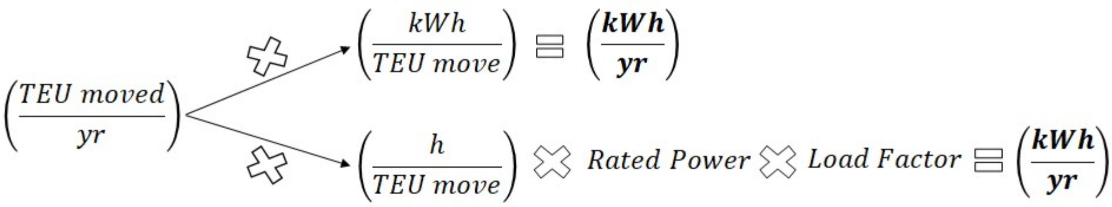

Figure 6. Approaches used in the estimation of energy consumed in container Terminal Oriented activities

Within these approaches is important to consider the correct definition of the movement of a container in the terminal (to avoid double counting of energy) and, how to include the energy consumed in idling. Using as a starting point the classification of activities presented in Table 7, the energy loads and profiles of each system, vehicle and equipment can be estimated according to the methodologies presented below.

Quay Loading/Unloading

The loading/unloading process of the containers (specified in the stowage plan) begins once vessels are at berth. During the quay loading and unloading, the main energy consumers are the committed quay cranes, which move the containers between the quay and the vessel. The time dependent energy consumption of a quay cranes (QC) can be estimated using equations (5) to (7) which are based on existing models [21]. According to these models, the energy consumption of quay cranes can be estimated considering the following factors: working energy consumption, non-working energy consumption and moving energy consumption:

$$E_w = \mu \times V, \quad (5)$$

$$E_{nw} = t \times \varepsilon + \frac{V}{p}\varepsilon, \quad (6)$$

$$E_d = \tau \times d, \quad (7)$$

where:

E_w, E_{nw}, E_d - working, non-working and moving energy consumption;
μ - working energy consumption per move [kWh/move];
V - handling volume [move];
τ - moving energy consumption per unit distance [kWh/m];
d - moving distance [m];
t - moving time [h];
ε - non-working energy consumption per unit time [kWh/h];
p- handling efficiency [move/h].

Equation (7) can be used to estimate the energy consumed to position the committed QCs alongside the berth; while Equation (6) can be used to estimate the energy consumption from air conditioning, lighting and other auxiliary equipment of the QCs, which are running during the QC operation. In general, typical values of power and time required during an entire cycle (one move/one hour/one meter) should be available; to link the operational/logistic data to the energy consumption.

Another approach to estimate the energy consumption of QCs, considers the loading and unloading of the containers from vessels to the quay in six different general movements, each its own energy consumption specification [22]:

1. Moving spreader horizontally from quay (idle position) to ship,
2. Lowering spreader above ship to get a container,
3. Lifting spreader and container from ship,
4. Moving spreader and container horizontally from ship to quay side,
5. Lowering spreader and container to terminal truck on quay side,
6. Lifting spreader from quay (to idle position).

The main concern in this approach is that, for all the cranes, the precise time at which a movement is executed must be known. The authors proposed mathematical model is the following:

$$E = \int_{t=0}^{T} \left(\sum_{i=1}^{I=6} m_i \times e_{il} + a \right) dt, \quad (8)$$

where:

E – energy consumed by a specific QC [kWh];
i - type of movement;
I - number of all movements:
m_i - executing particular type of movement (binary: 0 if negative, 1 if positive);
e_{il} - energy consumption for particular type of movement and container load [kW/s];
a - auxiliary energy consumption for the crane.

For each movement, data concerning the processing time of a container and corresponding energy consumption specifications should be available. These values can also be dependent on the weight of the container being handled. Each container load (0-100%) will have its own pre-defined energy specifications (kW/s). Multiplied by the operation times per sub-movement, the total energy consumption for a container can be determined and compared with real data.

The profile of the power demand, even for a fixed crane handling identical containers, will vary depending on other factors such as the container's position on the ship, wind and other aleatory variables. Figure 7 shows the load profiles of different cranes obtained in [23] (QC) and [24] (YC). Although obvious differences will exist (peak power, cycle duration, average energy consumption, etc.), distinctive characteristics can be identified: the existence of two power peaks, associated with two lifting movements (with and without container).

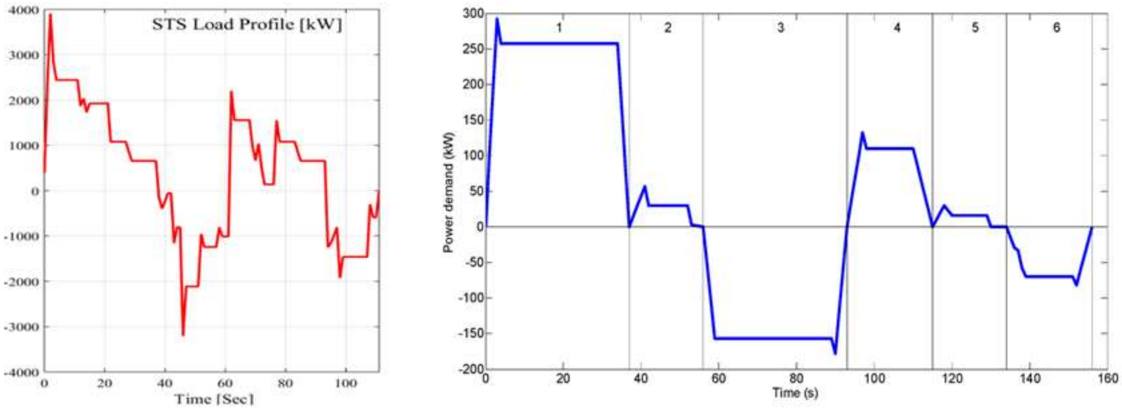

Figure 7. Load profile in a crane. Adapted from [23], [24].

As an approach to simplify the analysis process, these profiles could be smoothed out by considering the maximum power required for each movement and the average time needed for that movement, or by simply reducing the number of movements considered.

A third, and simpler, approach that can be used to calculate the energy consumption of quay cranes [24] (E_w)[in kWh] is given in Equation (9):

$$E_w = C \times \gamma, \tag{9}$$

where C [TEU] is the handling capacity of the QC and γ [kWh/TEU] is the energy consumption rate, which can be considered as 5.23 kWh/TEU.

Alternatively, the electricity consumption of QCs can be calculated by Equation (10) [18]:

$$E_w = I \times V \times \rho \times T \times (1 + \eta), \tag{10}$$

where:

E_w – QCs electricity consumption [kWh];
I - energy consumption for average operation capacity [kWh/TEU];
V – QCs efficiency [TEU/h],;
ρ - utilization ratio of quay cranes in port;
T - working hours [h];
η – line loss rate.

Finally, the energy consumption of QCs can also be estimated from Equation (1) [25], [26], using the load factors considered in Table 8. As we can observe, the load factors that have been considered are very different, so a first step should be the selection of the most appropriate value according to the specific operational context. In this case, conducting surveys among the assets operators could provide helpful information, since by simple extrapolating values of similar assets being used in other ports could introduce several uncertainties.

Quay to Storage

For the estimation of the energy consumption of the horizontal movement of the containers from the quay to the yard, several alternatives equipment should be considered, e.g., AGV, Straddle Carriers, Terminal Tractors. A possible approach [27] to estimate the energy consumed during these activities considers the travelling energy consumption (E_t) and the waiting energy consumption (E_w):

$$E_t = \alpha \times d, \tag{11}$$
$$E_w = \beta \times w, \tag{12}$$

Where:

α - the traveling energy consumption [kWh/m];
d - moving distance [m];
β - the energy consumption per unit of time when waiting [kWh/h];
w - waiting time [h].

Another approach [24] estimates the diesel consumption of the internal trucks (E_t)[kg] differentiating if the vehicles are loaded or not:

$$E_t = R^s \times v^s \times t^s, \tag{13}$$

where:

$s \in \{0,1\}$ – state variable (vehicle loaded or not);
R^s - diesel consumption efficiency of vehicles at the status of s [kg/km];
v^s - speed [km/h]
t^s - working time [h].

As horizontal movement of cargo to storage is not exclusive to ports, models and approaches commonly used for estimating the energy demand of heavy-duty vehicles for similar activities can be easily adapted to the terminal's specific layout. As already pointed out [28], the total workloads of the internal trucks can be computed based on the handling volume, the planned berthing time and the planned departure time of vessels. The total workload of a container terminal within a period can be obtained based on:

- The total unloading volume of all vessels at berth within a period.
- The total loading volume of all vessels at berth within a period.
- The overflowed workload from the previous period.
- The total re-arrangement movements due to fails in the stowage plan or other cases within a period.
- Additionally, the waiting time, e.g. queuing time in a block, should be included in the transportation time of an internal truck for a task.

Another aspect to consider is that the energy demand of the internal trucks (as well other mobile assets) is decoupled from the container movement (logistic side). Therefore, in the case of the battery electric vehicles, some load flexibility is introduced.

Table 8. CHE load factors found in literature.

CHE Type	LF[25]	LF[26]	LF[29]	LF[30]
RTG	0.2	0.42	0.43	0.20
Forklift	0.3	0.44	-	0.3
Container handler	0.59	-	-	
Reach stacker	0.51	-	0.59	
Container trailer	0.39	-	0.39	
ITV	-	0.32	-	
Straddle Carrier				0.2
Empty Container Handler	-	0.42	0.59	
Small Forklift	-	0.36	-	
Forklift Truck			0.30	
Crane	-	0.18	-	0.43
STS Crane	-	-	0.43	
Sweeper	-	-	-	0.68
Loader	-	-	-	0.55
Yard Tractor	-	-	0.39	0.51

Storage

Once the container has been moved from the quay, it will be stored in the container yard, with the support of the yard cranes (YCs). Additionally, the storage activity also includes the energy consumption of the reefers that are plugged to the electrical grid in the yard. For the YCs, any of the approaches considered for the QCs remain valid; however, considering typical energy consumption for moving one container introduces bigger uncertainties in this case. For QCs, one container move is compounded always by the same sequence of movements. In the case of YCs however, the energy consumed for positioning one container will depend also on the position of the crane before initiating the task, if any reshuffle operation is needed and the specific position where the container will be located. Thus, a more detailed model such as the one proposed in Equation (8) should be used to estimate the energy consumption of YCs, after adaptation to include the gantry movement.

In the case of the reefers, estimating the energy demand based on specific models for each of the containers handled in the terminals is a complex task. The models used should consider the refrigeration systems and its components; the larger thermal capacities in the system, such as the metal in the heat exchangers, the air in the container and the cargo. For instance, fruits and vegetables are generally quite sensitive to atmospheric and temperature variations, which means that the cargo temperature and air composition in the container must be kept within strict limits; whereas for frozen goods, the rules that must be followed to preserve cargo quality are more lenient than those for chilled goods [31].

Due to the complexity of obtaining specific models for each container based on thermal and thermodynamic analysis, some authors have been using typical values reported elsewhere or based on specific experimental conditions. For instance in [32] a mean refrigerated container energy consumption rate of 2.7 kW/TEU is assumed; whereas in [33] it is considered that the average power consumption of a reefer is 3.6 kW per TEU. This value of 3.6 kW per TEU was obtained, according to [34], for a very broad average value for all container types, ambient conditions, and cargo types. A 20' container tends to be closer to 4 kW and a 40' container tends towards 7 kW. As a result of new developments and the associated improvements in the efficiency of the containers, this value is dropping.

Other methodologies have estimated the average energy consumption per TEU of the reefers from experimental data, e.g. in [32], where the mean rate of power consumption of six frozen containers was estimated at 16 kW, which equates to a mean of 2.7 kW/TEU; and the mean rate of power consumption of ten frozen (6) and chilled (4) containers was 44 kW, a mean of 4.4 kW/TEU. The variability on the consumption of reefers [35], under the effect of solar radiation of a 40 feet high cube refrigerated has been measured at an initial power consumption of 7.3 kW, with the maximum power consumption reaching 7.5 kW at noon; whereas on rainy days the average power consumption is around 7.3 kW, and the trend line tends to be constant. Therefore, estimating the energy consumed by the reefers at a particular terminal is a process that should account for the specific characteristic of the terminal and the dynamic of the reefers received in the terminal.

Alternative modelling approaches specifically developed for reefer containers, have been used to estimate the internal temperature of the reefers considering the increase of the internal temperature of the reefers once it is switched-off for time t_{off} [s] [36], [37], [38]:

$$T(t + t_{off}) - T(t) = \Delta T_{amb}\left(1 - e^{-\frac{Ak}{mc_p}t_{off}}\right) \tag{14}$$

where:

$T(t)$ – is the internal temperature of the reefer at time t [°C];
ΔT_{amb} - is the difference between the ambient and the internal temperature [°C];
A - is the surface area of the container $[m^2]$
k - is the heat transition coefficient of the container $[W/m^2 \cdot K]$;
m - is the mass of reefer's content $[kg]$;
c_p - is the specific heat capacity of reefer's content $[kJ/kg \cdot K]$.

As noted in Equation (14), for a fixed reefer, several variable factors affect the temperature variation, e.g., the ambient temperature, cargo weight and thermal characteristics. If the temperature is so high that there is a risk of damage to the cargo due to overheating, the reefer needs to be brought back to its set point temperature, and so the reefer is rapidly cooled to decrease the internal temperature. During this process, in addition to the usual auxiliary power, a maximum amount of cooling power is applied. The applied cooling power can be obtained from:

$$T(t + t_{on}) - T(t) = -\frac{P_R t_{on}}{mc_p}, \tag{15}$$

where:
t_{on}- is cooling time $[s]$;
P_R - is the refrigerating power provided by the reefer $[kW]$;

Note that in Equation (15), the cooling time can be obtained from the desired cargo temperature $(T(t + t_{on}))$ and the reefer's refrigerating power (P_R),. This cooling (t_{on}) time and the switched-off time (t_{off}) can then be used to estimate an average electrical demand (D) $[kW]$ for each container at the yard, according to the following equation [39]:

$$D = \frac{P_R t_{on} + P_{aux}(t_{on} + t_{off})}{t_{on} + t_{off}}, \tag{16}$$

where P_{aux} stands for the auxiliary loads (e.g., lights, controller, fan).

Table 9 and Table 10 show examples of reefer thermodynamic and geometric parameters and electric power demand found in literature [40].

Table 9. Reefer parameters

	Reefer Length		Reefer Parameters	
	20	40	$k\ [W/(m^2K)]$	0.4
Surface Area: $A\ [m^2]$	73.56	135.26	$c_p\ [kJ/kgK]$	1.46-4.06
Cargo: $m\ [kg]$	18000-21000	22500-26500	Cargo temperature range [°C]	-25/18

Table 10. Reefer electric power demand and cooling

T [°C]	Cooling Power $P_R\ [kW]$		Electric Power Consumption $P_{e,R}\ [kW]$	
	Refrigeration System Type			
	1	2	1	2
21	13.7	13.5	10.7	12.0
2	9.9	11.4	9.1	10.5
-18	5.9	5.6	6.1	6.0
-29	3.8	3.02	5.7	4.5

Equation (14) can be modified to introduce a dimensionless factor that considers the reefer's exposure to the sun (S) [41]:

$$T(t + t_{off}) - T(t) = \Delta T_{amb}\left(1 - e^{-\frac{Ak(1+S)}{mc_p}t_{off}}\right) \tag{17}$$

In this case the power is estimated in pulses (peak power) due to the combined use of auxiliary and cooling power. The pulse is applied until the temperature has reached the set point. After this, the reefer operates in its usual on/off mode, originating a fluctuating pattern.

An important limitation of the equations (14) and (17) is that these models apply only for non-perishable goods. Another aspect is that these models should be used solely for short-term forecasting of the internal temperature and the power demand of reefers.

An alternative to the physically based models presented above, would be to develop a data-based model if a good quality historical consumption data set is available. In this approach important factors influencing the energy demand of the reefers at the yard can be identified, to create a bespoke forecast model for the reefer power demand. These types of data-driven approaches have been described in the literature [42], where a multiple regression analysis, a simple linear model, was used to predict the energy consumption of the reefers at a terminal. The authors concluded that for the specific use case:

- The model is able to explain variations in energy consumption very well;
- The regression analysis shows that the number of arriving reefers effects the total energy consumption the most;
- The plug-in temperature, thermal, insulation of the reefers, and the cargo type have a negligible impact on energy consumption;

- The temperature set-point, offline time, weight, ambient temperature, and sun-hours are found to be non-significant.

Other models are available in the literature, which show accurate results, but need detailed technical information for the reefers, which may not be easily accessible [43], [44], [45], [46].

Receipt-Delivery

During this activity, the main consumption is due to the YCs involved in receiving (delivering) a specific container to (from) the yard from (to) an external truck. Therefore, we refer back to any of the models represented in Equation (1) and Equations from (5) to (10).

2.3. CO_2 and emissions mapping

In light of the numerous discussions with representatives of the port authorities, partners of the MAGPIE project, a tool capable of contextualizing CO_2 emissions produced by port activities with the ones produced by the surrounding areas appeared as a valuable decision-support and monitoring instrument that nowadays is somewhat lacking.

In order to succeed in this contextualization, one must accurately assess not only the CO_2 emissions produced by the vehicles operating within the port boundaries, but also the emissions produced by human activities and mobility outside the port.

The final goal of this tool, which will take the form of an interactive dashboard, is therefore to estimate and show at any given time during a day the contribution to total CO_2 emissions coming from the port-related activities and from the surrounding areas. This will allow port authorities to understand at what time of day the port is a preponderant CO_2 emitter, and therefore to understand what measures to take to reduce such a contribution.

In this section, we provide a brief introduction about the scientific literature and the typical challenges when estimating road traffic emissions. This will provide a broader context and understanding of the data produced in this deliverable, enabling readers to better comprehend the displayed and analysed information within the interactive dashboard.

In general, estimating road traffic emissions involves three different modelling blocks: road vehicles flow model, pollutant emissions model and vehicle fleet composition model.

Real-world vehicle pollutant emissions depend not only on the vehicle technology but also on driving style, road infrastructure, and traffic management measures. However, in many cases, changes in road infrastructure and traffic management prioritize factors such as capacity, congestion, and user safety, with little consideration given to their impact on pollutant emissions [47], [48]. Nevertheless, studies have shown that modifications in road infrastructure and traffic management can significantly affect driving conditions and vehicle emissions [49].

The advent of Intelligent Transportation Systems (ITS) and Traffic Information Systems (TIS) has made traffic data increasingly accessible. This accessibility has created opportunities for innovative models and methods for traffic-related predictions. Accurately estimating driving behaviour and dynamic speed profiles, accounting for time-varying traffic conditions, is particularly valuable.

Traditionally, speed estimation methods have relied on obtaining a single mean value estimation per road link. These methods utilize real-world traffic data collected from probe vehicles or monitoring systems such as Loop Detectors (LP) and Floating Car Data (FCD) [50], [51], [52], [53]. However, the usage and accuracy of these methods depend on the availability of measurement data.

Traffic models offer an alternative approach to estimate road link speed by utilizing the fundamental diagram theory [54], [55]. These models describe the deterministic relationship between flow speed and density (number of vehicles per unit length). Macroscopic models consider the aggregate behaviour of traffic flow on road links and estimate the mean speed of a road link. On the other hand, microscopic models [56], [57] consider each vehicle separately, primarily relying on car-following and lane-changing theories. However, these models often require complex calibration and extensive inputs (e.g., Origin/Destination matrix) that may not be available for all road networks.

Driving cycles, constructed from a history of real trips using stochastic approaches [58], [59], are also employed to generate realistic speed trajectories. However, these driving cycles are typically used for long distances and do not fully account for the impact of topology and infrastructure at a high resolution.

IFPEN has developed an approach that enables a better overall understanding of traffic-related pollutant emissions and their underlying factors at a high spatial and temporal resolution. This approach includes extensive validation and comparison, contributing to improved traffic safety, reduced emissions, and energy consumption.

To accurately estimate pollutant emissions based on predicted driving behaviour, an adapted microscopic vehicle and emissions model is necessary. Driving behaviour significantly influences pollutant emissions levels, regardless of the vehicle or its technologies. Therefore, understanding and monitoring vehicle usage patterns can have a dual benefit: at the driver scale, it can directly lead to decreased emissions through improved driving behaviour and habits, while at the regulatory scale, it can assist in the development of future standards and infrastructure.

Regarding microscopic emissions modelling, the environmental impact of vehicles has traditionally been evaluated through dynamometer emissions tests. However, data derived from such tests may not be representative of real-world driving conditions [60]. To address this issue, the Portable Emissions Measurement System (PEMS) has been developed since the 1990s [61]. While these systems are suitable for measurements on specific vehicles, conducting large-scale studies of real driving emissions (RDE) is often not feasible due to cost constraints. As a result, limited knowledge is available regarding the impact of real-world conditions on emissions, with recent studies starting to shed light on this subject [62], [63]. An alternative method for indirectly measuring real traffic emissions is by using air quality sensors. However, large-scale diffusion of such sensors is limited, making it difficult to directly attribute pollution to its specific cause. Emission factors coupled with real GPS data have been used to estimate vehicle emissions [64]. However, emission factors only consider average vehicles and average driving styles, making them suitable for estimating average emissions on long trips but not for real traffic emissions that need to consider the local impact of driving style and slope [65]. To account for these phenomena, a finer level of modelling known as microscopic modelling is necessary, with the input typically being a 1 Hz vehicle speed profile. While several microscopic models already exist, they are primarily designed for offline studies [66], [67], [68].

Unfortunately, microscopic models require significant computational resources that do not align with the requirements of the current project, which necessitates emissions estimations on a large scale and for extended durations. Therefore, IFPEN has developed a novel approach that significantly reduces computational load while continuing to estimate emission and noise levels at the road segment level, capturing vehicle dynamics and topological information. This model is referred to as mesoscopic, in contrast to the microscopic models described above that require high-frequency data, and the macroscopic models of the state of the art such as COPERT or HBEFA.

An illustration of the proposed workflow to estimate road traffic emissions in the port surrounding areas is provided in Figure 8.

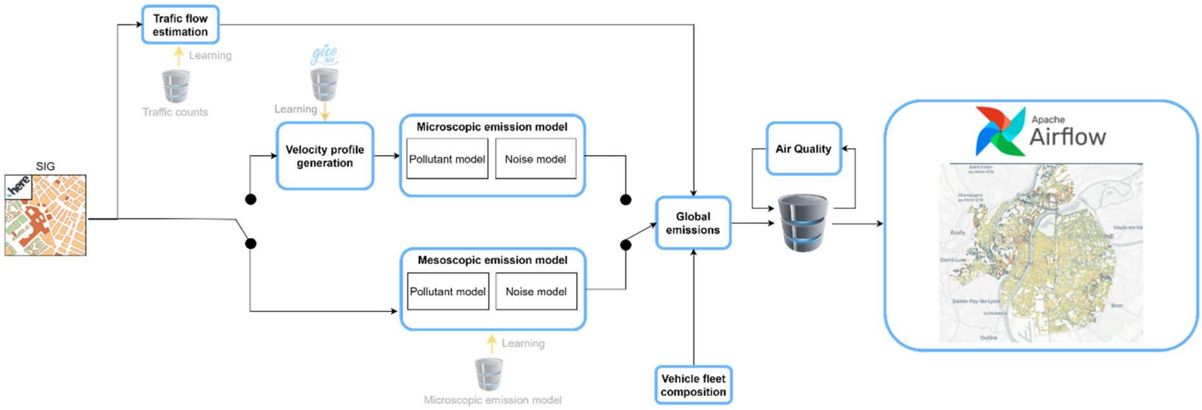

Figure 8 Illustration of the workflow proposed by IFPEN to compute road traffic emissions and project them on a map.

Road traffic emissions modelling

2.3.1. Road traffic flow estimation

The vehicle flow estimation model used in this work is based on some previous IFPEN developments presented in detail in [69]. In this section, a short presentation of the model is provided. Interested reader can refer to the original paper.

The proposed data-based traffic flow model estimates daily traffic flow on any road-link of a road network, even those not equipped with sensors. It takes as inputs data available everywhere to make the predictions. As illustrated in Figure 9, the overall approach is comprised of several sub-models, each sub-model construction will be detailed in this section.

For the training part, we first consider road-links with available traffic flow measurements. In our case, it covers about 800 road-links with loop-detectors in Lyon, France, every 6 minutes in 2018. For those road-links, the model learns the correlation between traffic counts data and the selected input data.

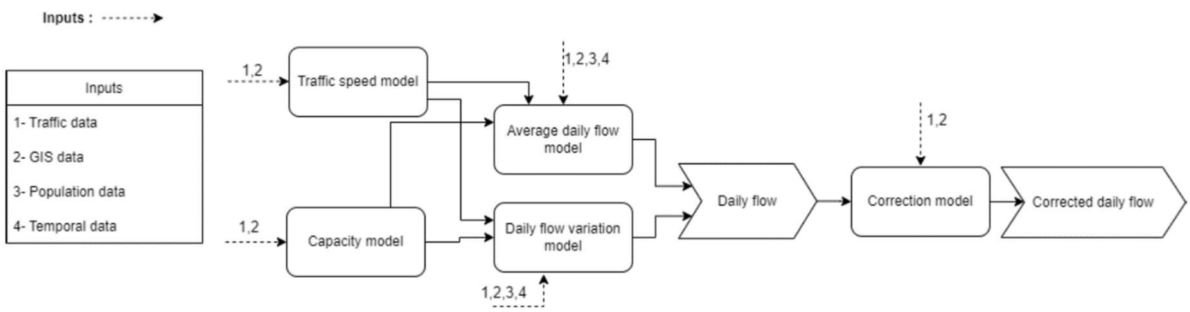

Figure 9 Illustration of the overall road traffic flow estimation approach.

<u>Traffic data:</u> These data mainly include the traffic flow and the traffic speed as measured by loop detectors on certain road-links.

<u>Geographic Information System (GIS) data:</u> These data can be provided through any GIS (Here Maps, Google Maps, OpenStreetMap, etc.). Each GIS provides a decomposition of the

road network into road-links. The selected GIS data used as input features will be detailed for each sub-model.

Population data: It includes mainly open access data of population statistics according to a certain geographical decomposition in zones. Each road-link is associated with a zone, so it is easy to link these data to road-links. These data can include population densities, socio-professional categories, and the overall attractiveness level (industries, schools, etc.) of the area, and are introduced in order to improve the extrapolation capabilities of the model to various road networks. In this work, we limit ourselves to considering the population density per zone.

Temporal data: This step consists in defining several temporal descriptors to distinguish and predict the temporal evolution of the traffic flow for different days with different characteristics. For example, traffic flow has a different daily evolution on a working day of the week as compared to a weekend day. The evolution is also different according to the season. In this work, we consider the days of the week (Monday…, Sunday), months, school and public holidays.

The proposed modelling approach is made up of several sub-models.

2.3..2. Traffic speed model

This model takes as input GIS data in order to estimate a synthetic traffic speed. It is calibrated in road-links with available traffic speed measurements from loop-detectors. The general idea of this module is to improve accuracy of the GIS data by learning the relationship between such data and the speed ground-truth. The traffic speed is highly correlated to the congestion level. Its estimation can therefore be used to improve traffic flow prediction.

A multi-layer perceptron Neural Network (NN) is used as the regression algorithm:

$$V_{te} = f_v(V_l, V_n, V_t)$$

Where V_{te} is the estimated traffic speed, V_l, V_n and V_t are respectively the speed limit, the free-flow speed and the traffic speed retrieved from the GIS.

2.3..3. Capacity model

The capacity reflects the maximum traffic flow level that a road-link can reach. This data can either be provided by a microscopic traffic model or estimated from counting loop measurements. Such a model can therefore be used to set the maximum traffic flow that a road-link can reach and thus improve the prediction capabilities of the overall flow estimation model by imposing a physical upper bound.

As with the traffic speed model, a multi-layer perceptron Neural Network (NN) is used to learn the correlation between GIS data and capacity:

$$Q_e = f_c(N_l, FC, V_l, V_n, L_e)$$

Where Q_e is the estimated capacity, N_l is the number of lanes, FC is the functional class, which represents a hierarchy of road-links and categorizes them according to their functions, L_e represents the road-link length.

2.3..4. Average daily flow model

This model uses the traffic speed and capacity estimations along with other GIS, population and temporal data, in order to predict the daily traffic flow mean and standard deviation of any road-link.

To do so, the measured traffic flow data is first processed for each road-link and decomposed as follows:

$$\Phi(t_d) = \sigma_f(t_d)\Phi_n(t_d) \; + \mu_f(t_d)$$

Where $\Phi(t_d)$, that is the traffic flow of day t_d, is a vector comprised of 6-minute sampled measurements. σ_f and μ_f are respectively the standard deviation and mean of the measured daily traffic flow. Finally, Φ_n is the measured normalized daily flow variation.

At this point, the idea of the approach consists in estimating the daily traffic flow $\hat{\Phi}(t_d)$ by estimating the three components $\hat{\sigma}_f$, $\hat{\mu}_f$ and $\hat{\Phi}_n$, as:

$$\hat{\Phi}(t_d) = \hat{\sigma}_f(t_d)\,\hat{\Phi}_n(t_d) + \hat{\mu}_f(t_d)$$

Therefore, a random forest (RF) regression algorithm is first used to learn the mean and standard deviation of the daily traffic flow on road-links and time periods with available measurements, based on selected input data as follows:

$$[\hat{\sigma}_f, \hat{\mu}_f] = \; f_r(Q_e, V_{te}, T_d, P_d, V_l, FC, N_l, L_e)$$

Where T_d and P_d are respectively the temporal and population data.

Once the daily traffic flow mean and standard deviation are estimated, the next step is to calibrate a normalized daily flow variation model.

2.3..5. Daily flow variation model

This model predicts the traffic flow daily evolution for a given day and road-link. It is calibrated with the measured normalized daily traffic flow (i.e. Φ_n).

The first step is to identify the most representative daily variation profiles, as observed in the measurement data. A time-series K-means algorithm is used to cluster the normalized daily traffic flow. It uses the Dynamic Time Warping (DTW) metric to measure the similarity between two temporal sequences.

Once the clusters are identified, a classification approach is used to estimate the probabilities of belonging to each cluster. A random forest classifier is used along with a K-fold cross-validation to optimize its structure.

$$P_c \; = \; f_p(Q_e, V_{te}, T_d, P_d, V_l, FC, N_l, L_e)$$

Where P_c is a vector representing the probabilities of belonging to each cluster. Once those probabilities are estimated, the normalized daily traffic flow can then be constructed for a given day and road-link:

$$\hat{\Phi}_n \; = \Sigma_{i=1}^{K} P_c(i) C(i)$$

Where C are the clusters' centroids representing the most representative traffic flow daily profiles, and K is the chosen number of clusters. Figure 10 shows the selected clustering with the 20 most representative traffic flow profiles. Some clusters present a large traffic flow variation, especially during the evening peak hour. This corresponds to a highly congested traffic state. Other clusters have an almost constant traffic flow evolution.

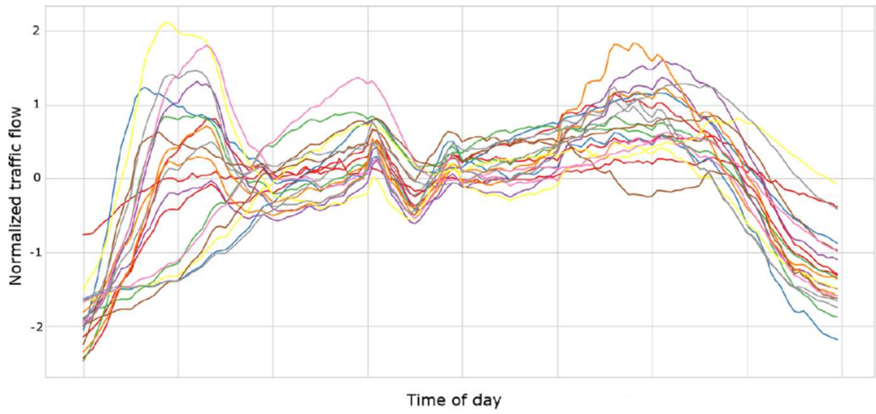

Figure 10 Clustering with the 20 most representative normalized daily traffic flows.

At this stage of the estimation model, the daily traffic flow can be predicted on any road-link and day by combining the average and variation daily flow models.

2.3.6. Correction model

The daily traffic flow evolution can be predicted on each day and each road-link of a given road network. However, this estimation is done independently of the hierarchy and connectivity of the road network. In other words, two neighbouring road segments can have different traffic flow estimations and no spatial continuity is ensured.

A correction step is thus introduced to ensure spatial traffic flow continuity. For each time step, this correction is formulated as an optimization problem with constraints on the connectivity.

$$\min_{\widehat{\Phi}_o} J = \sum_{i=1}^{N} \left(M(i) \left(\widehat{\Phi}(i) - \widehat{\Phi}_0(i) \right)^2 \right)$$

subject to the following constraints:

$$0 \leq \widehat{\Phi}_o(i)$$

for each road-link i, and

$$\sum_{j \in k_{in}} \widehat{\Phi}_0(j) \leq \sum_{j \in k_{out}} \widehat{\Phi}_0(j)$$

for each unique node k connecting road-links.

N is the number of road-links in the considered road network. $\widehat{\Phi}_0$ is the optimization variable consisting of all the corrected traffic flow on a road network at a given time step. $\widehat{\Phi}$ is the estimated traffic flow.

M is a weight equal to $\lambda \gg 1$ when a loop-detector is present on a road-link and equal to 1 otherwise. Accordingly, in the case of existing flow measurements on a road-link, the traffic flow estimation in the previous step is considered more precise.

Finally, k_{in} is the subset of road-links upstream from node k and k_{out} is the subset of road-links downstream from node k. The last constraint considers the possible traffic congested state. In the case of traffic congestion, the upstream traffic flow is considered inferior to the downstream traffic flow. In free-flow conditions, traffic flows are equal.

An interior-point method with IPOPT has been used to solve this optimization problem. It is known to be efficient and very fast for complex problems (large road networks).

Once the optimization is solved, a traffic flow with consideration of connectivity and spatial continuity is estimated for each road-link and each time step of a considered day.

2.3..7. Use of the model and error metrics

The online use of the proposed model can be summarized as follows:

- Decomposition of the road network by road-links according to the considered GIS and extraction of the GIS data on each road-link.
- Extraction of the population data and relevant temporal variables according to the considered area and time period, as defined in the model construction section.
- The overall model then takes all these data as input to estimate the daily flow on each road-link of the road network considered. This global model combines the traffic speed, capacity, average and variation daily flow models. All these models have been calibrated in the model construction part. The estimated average daily flows are then taken as input to the optimization part in order to take into account the connectivity of the road network.

To assess the extrapolation capabilities of the model, various validation scenarios are defined:

- Temporal extrapolation: Learning from all road-links with loop-detectors (about 800) in Lyon, France, in 2018 and validating the traffic flow predictions in 2019 for the same road-links.
- First spatial extrapolation: Learning from 80% of the road-links with loop-detectors in Lyon, France, in 2018 and 2019 and validating the traffic flow predictions for the remaining 20% of the road-links for the same period.
- Second spatial extrapolation: Learning from all road-links with loop-detectors in Lyon, France, in 2019 and validating the traffic flow predictions in the same year for road-links with loop-detectors (about 700) in Paris, France.

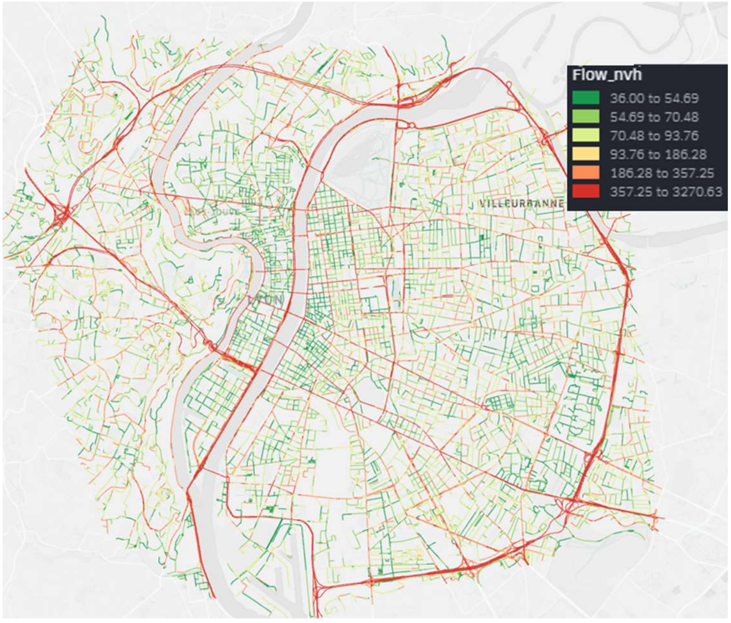

Figure 11 Example of traffic flow prediction (number of vehicles per hour) on 20/05/2019 at 14:30 in Lyon, France.

In order to compare the performance of the proposed model with standard traffic simulation tools, we evaluate the GEH [70] statistic, which is very common in the field of traffic engineering and traffic forecasting to compare two sets of traffic volumes. Traffic volumes in a road network vary over a wide range, the GEH is therefore preferred to the average of the relative error, which could give more misleading results for low traffic flow points. In our framework, the GEH metric can be defined as:

$$GEH = \sqrt{\frac{2(\hat{y}_j - y_j)^2}{\hat{y}_j + y_j}}$$

where y_j is the measured traffic flow at time j, $\hat{y}_j$ is the predicted one, and j represents now an hourly interval. In general, a GEH of less than 5 is considered a good match between the modelled and observed hourly volumes and 85% of the volumes in a traffic model should have a GEH less than 5. GEHs in the range of 5 to 10 may require further investigation. Finally, if the GEH is greater than 10, there is a high probability that there is a problem with model calibration or the data itself. A summary of the error metrics for the proposed model is provided in Table 11.

Validation scenario	SMAPE	Median GEH	% of estimates with GEH<5
Temporal extrapolation	11%	2.46	79%
1st spatial extrapolation	23%	6.88	41%
2nd spatial extrapolation	25%	9.11	33%

Table 11 Summary of the proposed traffic flow model's error metrics.

Considering the open-access nature of the input data required to run the model, as well as the minimum calibration effort needed to design it, we deem these estimation results to be comparable with more complex traffic simulators. For instance, a recent work [71] shows that a commercial macroscopic traffic simulator with time-consuming calibration effort and necessary simplification of the simulated road network also fails to meet the target criteria of 85% of all counts to have a GEH of less than 5. In fact, the model reaches a 75%-80% performance with a specific focus and calibration on the day peak hours, and the validation was mainly carried out for temporal extrapolation.

From microscopic to mesoscopic emissions model

The proposed mesoscopic emissions model aims to improve current macroscopic models based on emission factors by introducing a new mesoscopic emissions model with additional input parameters (in addition to average speed) that have the greatest impact on traffic-related pollutant emissions, even at the road segment level. Specifically, the additional input variables considered in the model aim to characterize the impact of congestion, topographical features, and road signage on emissions. This model represents a significant improvement over the state of the art by making the impact of road infrastructure on emissions explainable, visible, and easily accessible, even for non-experts.

Figure 12 provides an overview of the workflow of the enhanced macroscopic emissions model. It takes as input macroscopic data, which can be obtained from any road segment through various Geographic Information Systems (GIS) such as HERE Maps, Google Maps, OpenStreetMap, etc. Each GIS provides a decomposition of the road network into road segments. Each segment is defined as a basic link between two network nodes without any variation in segment-related attributes (once the segment is defined), such as segment length, number of lanes, signage, speed limits, etc. Attributes that can vary within a segment include curvature, road slope, and traffic conditions.

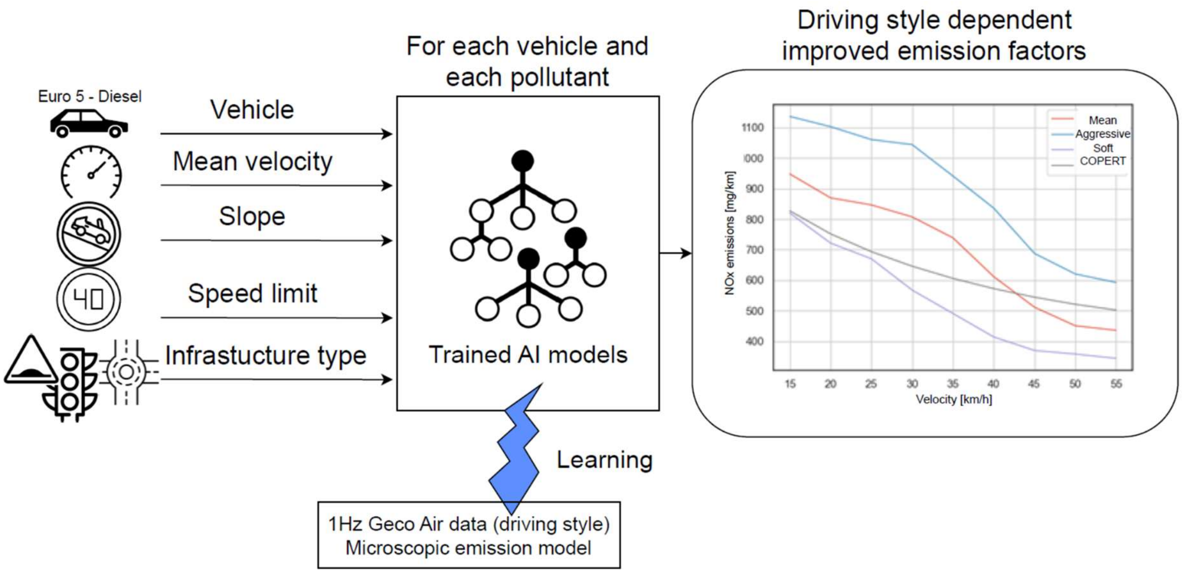

Figure 12 Functional diagram of the mesoscopic emission model.

The model is calibrated using emission estimates derived from the microscopic emissions model, incorporating trajectories from the anonymous Geco Air database, which spans 100 million kilometres. Geco Air is a free smartphone application designed for the general public, empowering individuals to actively reduce their mobility-related pollutant emissions. By capturing GPS signals at a frequency of 1 Hz, Geco Air acquires data on speed, acceleration, and altitude, which, when combined with detailed modelling of the vehicle, engine, and after-treatment system, enables accurate estimations of pollutant emissions. This valuable information empowers users to make informed decisions about their environmental impact and take steps towards reducing emissions.

In addition, the model leverages Geco Air trajectories to estimate improved emission factors for vehicles across the vehicle fleet, encompassing various engine types and conforming to the latest European standards. It goes beyond merely estimating the median emission factor, which represents the average driving style, and provides emission factors associated with two additional driving styles: gentle and aggressive. These additional factors are calculated by estimating the 25th and 75th percentiles of microscopic emissions derived from Geco air trajectories for the corresponding road segment.

As mentioned below, the enhanced macroscopic emissions model considers several macroscopic inputs that have a significant impact on emissions:

- Infrastructure Type: The idea is to categorize the road segments that make up one or more road networks. The objective of this categorization is to define a set of typical and representative road infrastructure cases within a road network. Macroscopic GIS data is used as input for this step. Several macroscopic descriptors can be used to define the categories, such as the presence and type of signage, speed limits, etc. provides an example of categorization with seven defined categories. Each category exhibits different emission levels, resulting in the definition of multiple infrastructure categories as output from this step.

- Slope Level: Slope has a significant impact on fuel consumption and emissions. Therefore, slope level is considered as an input to the mesoscopic emissions model. The idea is to define different slope levels (-5%, -2.5%, -1%, 0%, 1%, 2.5%, 5%) and calculate an emission factor for each level.

- Average Speed: Like most mesoscopic emissions models, the average traffic speed is also considered as an input in our model.

- Speed Limit: For the same average speed, emission levels and congestion can vary greatly depending on different speed limits. Therefore, an additional input is proposed by defining five speed limit levels: 30, 50, 70, 90, and 110 km/h.

Once all these inputs are defined, the goal is to learn the enhanced emission factor model. Such a model allows us to calculate emission levels for other road segments or networks outside the training database. As mentioned earlier, this training database is derived from the microscopic emissions model associated with trajectories recorded from Geco Air and projected onto the cartographic reference of the considered GIS.

An initial database of one million trips in Lyon, France was utilized. This database was subsequently reduced by removing redundant data, reducing training time and model complexity. The approach involves defining the range of variation for all input variables and discretizing them through multidimensional binning. For each multidimensional bin (e.g., BinX=[Infrastructure=1, Slope=-5%, Average Speed=[20-25] km/h, Speed Limit=30 km/h]), the corresponding Geco Air sub-trajectories within that bin are retrieved from the initial database. Emission calculations are performed for all sub-trajectories in that bin. Finally, only the sub-trajectories corresponding to the median, 75th percentile, and 25th percentile (for aggressive and gentle driving modes) of emissions within that bin are considered in the reduced database. This process is repeated for each multidimensional bin.

The model identifies, for each type of emission (NO_x, CO_2, HC, CO, PM, and noise), and vehicle category, the link and correlation between the emission factor, infrastructure category, speed limit, slope, and average speed. This identification is represented as follows:

$$[\overline{EF}, STD] = f(Cat, slope, V_m, flow, EMS, veh)$$

Where $\overline{EF}, STD$ represents the average emission factor and corresponding standard deviation (allowing estimation of factors associated with gentle and aggressive driving styles Cat, V_m, V_l, $slope$, EMS, and Veh represent the infrastructure category, average speed, speed limit, slope, emission type, and vehicle category, respectively.

Without loss of generality, the function f can be learned from a supervised learning algorithm, such as multi-variable linear regression, a neural network, a random forest, or a combination thereof. The choice of the model and its structure was guided by cross-validation methods (such as "k-fold") to reduce overfitting issues and improve accuracy.

Enhanced emission mesoscopic model has been recalibrated for each vehicle of the considered fleet (including light vehicles (LV), light commercial vehicles (LCV), heavy-duty vehicles (HDV) and two-wheelers (2WD)), for different euro norms, weights, motorization type (gasoline/diesel), and emission line post-processing type. Then for each vehicle and for each pollutant in NO_x, CO_2, HC, CO, PMe (exhaust) and noise, a dedicated model is defined.

To represent the French vehicle fleet, 50 real vehicles are considered. They represent 33 LV, 7 LCV, 5 HDV and 1 2WD. Some cars have the same Euro standard and the same engine but different sizes, all sizes are grouped to obtain the following table considering 30 average vehicles.

LV and 2WDs are discriminated according to their Euro standard and their motorization. The LCVs are discriminated along their Euro standards and based on weight ranges. All trucks have the same motorization and Euro standard, but they are discriminated according to their weight but also to their emission line post-processing type (EGR or SCR).

The Figure 13 shows NO_x emissions comparison between the mesoscopic emission model (in blue) and the microscopic emission model (in red) for a diesel Euro5 vehicle. Both models represent the same emission evolution in function of the mean speed. The standard deviation and the distribution of emissions by speed bin are also greatly reproduced by the mesoscopic model. Same observation is shown in Figure 14. Both models have similar evolution in function of the slope.

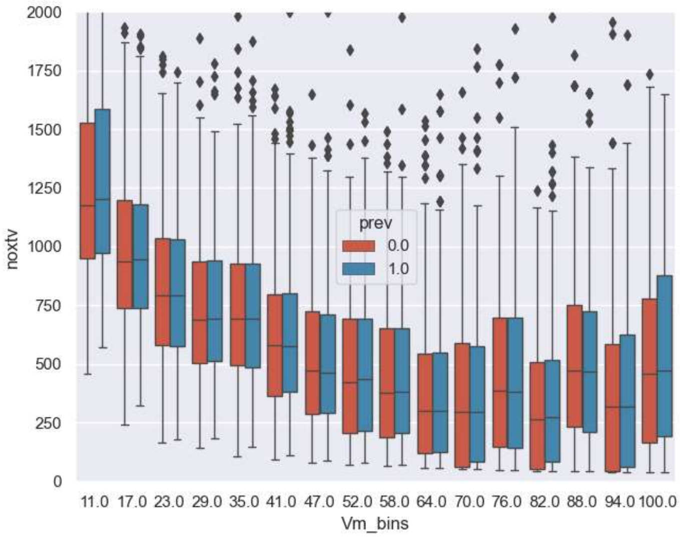

Figure 13 Assessment of NOx emissions evolution in function of the mean speed for a Diesel Euro5 vehicle. Comparison between predicted values (in blue) and data from microscopic emissions model (in red).

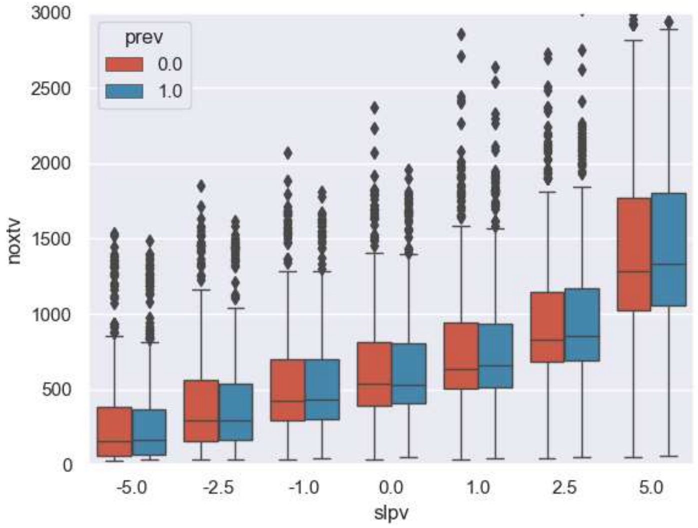

Figure 14 Assessment of NOx emissions evolution in function of the for a Diesel Euro5 vehicle. Comparison between predicted values (in blue) and data from microscopic emissions model (in red).

Quantitatively, Table 12 summarizes the mean absolute error (MAE) evolution for each pollutant (including the noise). For this assessment, the full dataset is divided into a training set (75% of the dataset) and a validation set (25% of the dataset). For all the pollutants, the error is comparable between the train and the validation set, which assesses the absence of overfitting. Error magnitude varies in function of the pollutant. For Diesel Euro5 vehicles NO_x emissions level is greater than other pollutants (except CO_2) with an average value of 650mg/km. Thus, relative error for this pollutant is 13%.

Pollutant	MAE (training)	MAE (validation)

NO_x [mg/km]	84	94
CO_2 [g/km]	17	21
PMe [mg/km]	0.5	0.6
HC [mg/km]	0.4	0.4
CO [mg/km]	19	23
Bruit [db]	0.2	0.2

Table 12 MAE for each pollutant with training and validation set.

As exercise, we propose to apply the enhanced emission mesoscopic model for different scenarios:

- Scenario 1: Road link without signage, without slope (0%) and with a speed limitation of 110km/h
- Scenario 2: Road link with a slope of 2.5% and a speed limitation of 110km/h.
- Scenario3: Road link with signage (traffic light), without slope and with a speed limitation of 50km/h.

For each scenario, enhanced emissions factors levels are compared to the state-of-the-art emissions factors. The state-of-the-art COPERT is the European standard for calculating vehicle emissions. COPERT tables are derived from a combination of sources, including laboratory tests, on-road measurements, and modelling. The emission factors in the tables are regularly updated to reflect improvements in the understanding of vehicle emissions. The development of COPERT is coordinated by the European Environment Agency. The version of COPERT used is that of October 2021. The figure below shows the NO_x emissions comparison between enhanced emissions factors and from COPERT for these three scenarios and for a range of mean speed. The same Diesel Euro5 vehicle is considered for this application. Scenario 1 is a road infrastructure nominal case without signage and without slope. For this case, estimations given by the proposed model are comparable to ones given by COPERT, considering the mean driving style. In addition, the proposed model estimates two others emission levels, considering a soft and an aggressive driving style. Emissions with these driving styles fit the lower and upper boundary of emissions with COPERT.

For the scenario 2 the slope level passes from 0 to 2.5%. COPERT does not consider the slope for the computations for LV. Thus, in this case, it underestimates emissions level and COPERT curve is closer to soft driving style curve. Finally, for the scenario 3, we consider an urban case with a traffic light and without slope. The traffic light leads to stops and accelerations that induces higher emissions level than for a nominal case. COPERT does not consider the signage because it means everything, thus it underestimates NO_x emissions.

Finally, a complete benchmark to compare the estimated emissions against the state of the art (COPERT) was realised. For this purpose, the model-based approach of the IFPEN solution allows to simulate exhaustively the estimation cases that can be encountered and thus to virtualize the road segments by simulating all possible cases.

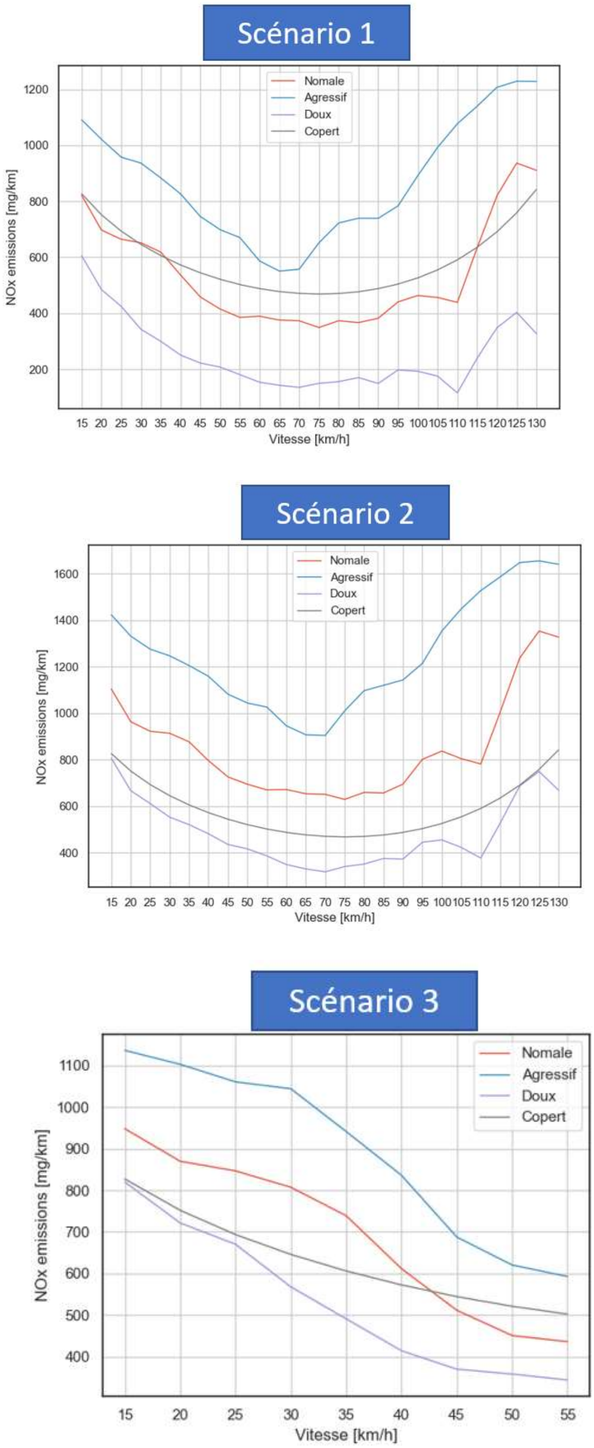

Figure 15 Comparison between the mesoscopic model and COPERT for NOx pollutant emissions. The simulated vehicle is a Diesel EURO 5 for both models and for three road infrastructure scenarios.

The simulations skim the whole possible space of values for traffic speed, speed limit, infrastructure case and slope and for two pollutants CO_2 and NO_x. The scales are from 5km/h to 130km/h every 5km/h for the traffic speed, from 10km/h to 130km/h every 10km/h for the speed limit, from -15% to 15% every 2% for the slope. All infrastructure cases are considered. In total, 37856 virtual links are created to perform this benchmark. The computation time of the IFPEN model emissions factors is about 8 min for all the virtual links. The computation time of COPERT emissions factors (using an IFPEN implementation) is nearly 2 hours.

To facilitate the reading, the results are only presented for two types of vehicles, Euro 5 diesel cars and Euro 5 diesel EGR trucks with a weight of 14T. For cars, COPERT only considers the average speed. For trucks, the slope is also considered. As expected, the dispersion associated with the emissions estimated by the model is higher than the emissions estimated by COPERT. Indeed, the IFPEN model takes more parameters as input.

Figure 16 below shows the impact of the slope on the model in comparison with COPERT:

CO_2	NO_x

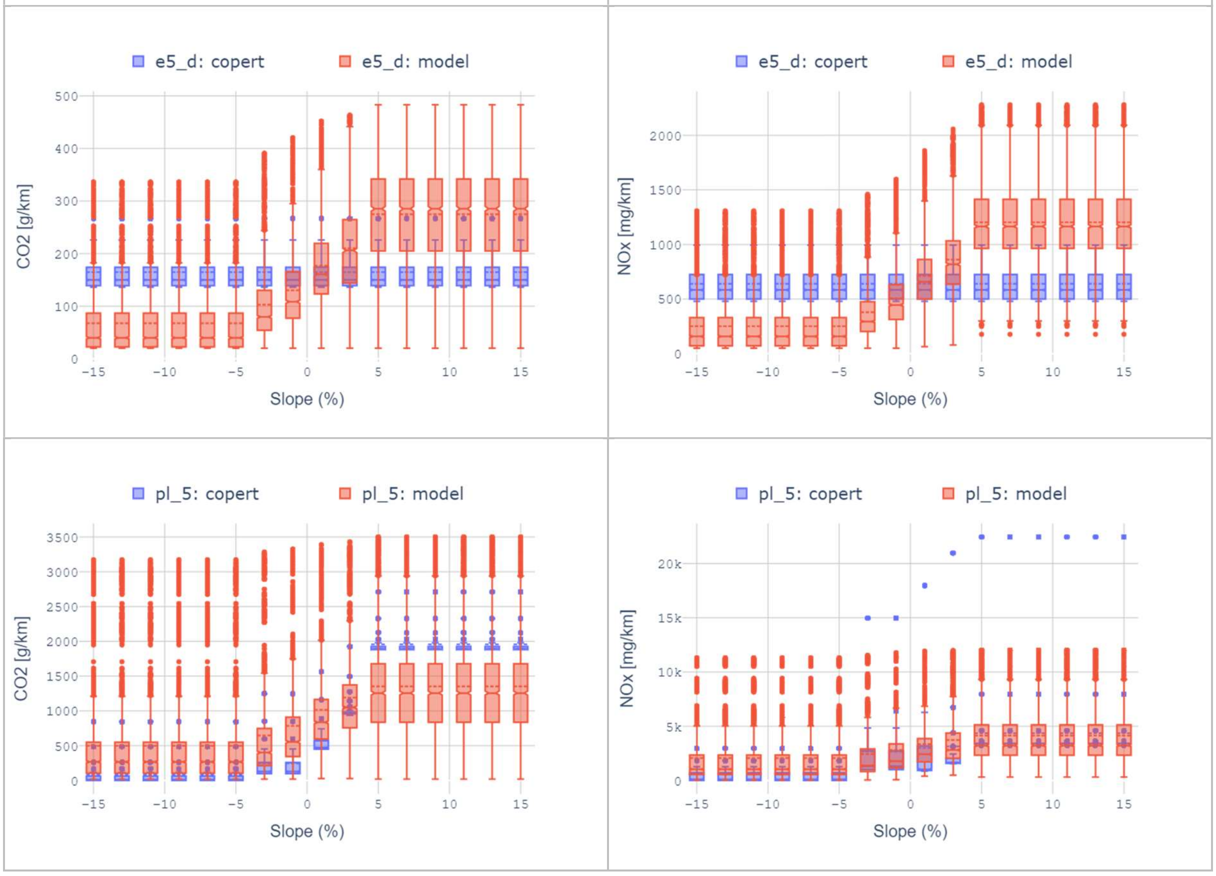

Figure 16 Graphs of the impact of slope on emissions, compared with COPERT.

As the model is learned for slope values equal to -5, -2.5, -1, 0, 1, 2.5, 5, a larger sample of slope values is chosen from -15 to 15 each 2%.

Since COPERT does not consider slope for LVs, emissions from negative slopes are overestimated and emissions from positive slopes are underestimated by COPERT. For emissions related to slopes close to 0, the model overlaps COPERT.

In the case of HDVs, the slope is considered by COPERT. The model overestimates the emissions related to negative slopes, although COPERT remains within the possible values of the model. The model underestimates the emissions related to positive slopes. For CO_2 COPERT emissions are in the first or third quartile of model emissions. For NO_x the average

COPERT emissions are within the model emissions, but the high slope low speed emissions are found above the model emissions possible values. This can be explained by the fact that part of the validation domain of the benchmark represents situations that are difficult to achieve for HDVs. The biggest errors are with low speed and high slope.

The Figure 17 below shows the impact of traffic speed on the model in comparison to COPERT:

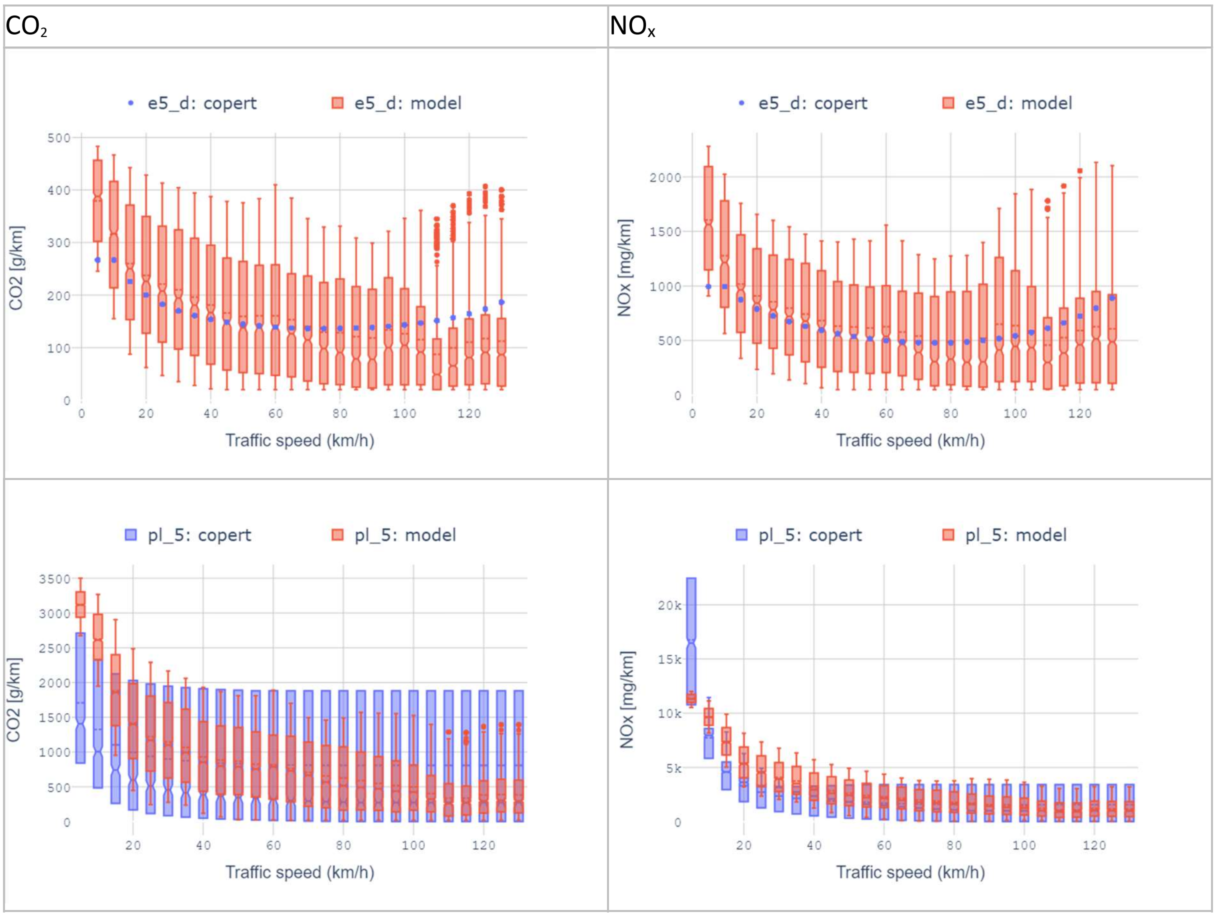

Figure 17 Graphs of the impact of traffic speed on emissions, compared with COPERT.

The values taken for the traffic speed for the benchmark are from 5km/h to 130km/h every 5km/h.

For cars, the model shows a higher dispersion. For CO_2, between 20km/h and 70km/h, the average emissions of COPERT and the model are overlapping or very close. For lower speeds, the model overestimates the emissions and for higher speeds it underestimates them. The average relative error is 1.3%. For NO_x, from 10km/h to 70km/h the average emissions of COPERT and the model are superimposed or very close.

The average relative error is 9.3%. For HDVs, the model overestimates the emissions at low speed and underestimates them at high speed. For CO_2, despite the large dispersion of COPERT below 10km/h the model emissions are higher than COPERT emissions. Above 10km/h for CO_2 and 40km/h for NO_x, the average emissions of both are close. The average relative errors are 2.6% for CO_2 and 13.5% for NO_x.

Thus, the IFPEN model is relevant compared to the state of the art, COPERT. It shows more dispersion for cars by considering the slope. This gives more accurate emissions and considers more parameters than only the average speed. The average values of COPERT and the model are rather close with differences of less than 30% in all cases.

The enhanced emission mesoscopic model has been recalibrated for each vehicle representative of a light vehicle fleet (for each Euro norm and engine). It also has been recalibrated for the most representative vehicles of the light commercial vehicles fleet and heavy-duty vehicles fleet. This model is really easy to use and was provided to several of IFPEN partners (i.e. public agencies for the air quality surveillance) who vocalized a significant need to correct emissions factors considering different level of slope, speed limitation and infrastructure type.

Vehicle fleet composition model

In order to finally estimate emission on a road-link where vehicle flow has been estimated and mesoscopic models gave average emission per vehicle for each vehicle type, it remains to set an important variable: vehicle fleet composition. What is the part of each main vehicle category in the fleet: Light Vehicle (LV), Light Commercial Vehicle (LCV), Heavy Duty Vehicle (HDV) and two-wheelers (2WD)? Of course, emissions are specific for each category due to weight and engine differences. Moreover, for a vehicle category it is important to know the part of electric, gasoline and diesel powertrain and to distinguish penetration rate of the different euro norm version.

2.3..1. Data source for vehicle fleet composition

Accurate vehicle fleet composition has to stick to national singularity when focusing on a port area (or any territory): electrical penetration rate, diesel/gasoline balance differ from a country to another. If fixing a vehicle fleet composition is useful to study the present or the past, it is interesting to be able to estimate its evolution in next years/decade. Since new vehicle aims at reducing individual emission, vehicle fleet composition scenarios allow to evaluate traffic emission in ten or twenty years, for example with iso traffic flow. Future fleet estimation relies on national or international studies.

Data source for vehicle fleet composition are multiple. Basically, each country may provide statistic on car license plate: number of vehicles by category, powertrain/fuel type, and even distribution over the different euro norm version. We can also catch sell estimation for next years, following different scenarios. For instance, if public authorities promote low emission vehicle (purchase bonuses, old vehicle maluses) the electric or hybrid vehicle sales will be greater than without any national incentive.

Nevertheless, it will be an error to assume that the national vehicle stock (i.e. the sum of the vehicle owned by citizens and companies) is representative of vehicles encountered on the road. Indeed, license plate, sales statistics only give an overview of the vehicle stock, which is not exactly the fleet composition we are looking for. A vehicle type may be used more frequently and for longer journeys than other. In Table 13, we give public statistics from CITEPA about French vehicle fleet. In the second column, number of vehicles are given, and the corresponding proportion of the total vehicle stock appears in the third one. In the fourth column we use the vehicles x km notion which can be presented by the total amount of km covered by the vehicle of the corresponding category. In other words, it corresponds to the multiplication of the vehicle fleet by the average annual number of km covered by a vehicle of this type. This notion is very important when we aim at estimating traffic emission because it fits better to the probability to find each vehicle type in a traffic flow. If we compare percentage given in the second and the fourth column, we see 2WD outnumber HDV four to one if we only focus to the vehicle stock. But on the road, HDV outnumber 2WD more than two to one.

Vehicle Category	Number of vehicles (thousands)	Part of the vehicle stock	Billon vehicles x km	Part of the fleet on road
Light Vehicle (LV)	38847.0	75%	452.8	76%
Light Commercial Vehicle(LCV)	8222.0	16%	98.5	17%
Heavy Duty Vehicle (HDV)	912.1	2%	32.0	5%
Two-wheelers (2WD)	3668.0	7%	13.4	2%

Table 13 Vehicle fleet composition (France) [72], [73].

If vehicle stock statistics can be obtained from administration, the total amount of km covered by a vehicle category is more difficult to know. It mainly relies on fixed point measurement on the territory which collects vehicle plate license and then build statistics about the categories of the observed vehicles. Updates are less frequent and local differences are not handled. Sometimes territorial studies exist but may not handle full vehicle type characteristic: for example, vehicle counting station often distinguish HCV from LV. But LCV and LV are not separated, and powertrain information (fuel type, euro norm) is missing.

For MAGPIE we focused on the following points:

- Handle the country specificities of the 3 European ports in the project,
- Distinguish powertrain types and corresponding euro norm,
- Fit road fleet rather than vehicle stock.

In order to handle country specificities, we based our vehicle fleet compositions on IEA (International Energy Agency) studies. IEA studies give vehicle stock with LV, LCV, HDV and 2WD distinction and the part of each powertrain type in each vehicle category. Data are available every 5 years from 2010 to 2040. 2025 and after are available for 2 distinct scenarios:

- STEPS: Stated Policies Scenario (+ 2,6 °C)
- SDS: Sustainable Development Scenario (+ 1,7 °C)

Annual sales for each category are given too.

In order to fit our needs, we have to update IEA data in two ways:

- refine vehicle category, mainly with euro norm distinction and weight categories for HDV,
- transform vehicle stock to vehicle fleet on the road

We refine vehicle category using CITEPA and SDES French data

1. part or vehicle corresponding to each euro norm for LV with gasoline / diesel distinction
2. part of vehicle corresponding to each euro norm and weight distinction for LCV
3. part of HCV in each weight category and part of HCV without Selective Catalytic Reduction system in each weight category (this system reduces tailpipe emissions of NOx, mandatory since euro 4)

Last process we apply to data is the projection from vehicle stock to vehicle fleet. One more time we use CITEPA information at vehicle macro category level (LV, LCV, HDV, 2WD) to fix the mix of each macro category in the road fleet.

Figure 18 gives an overview of French Light Vehicle mix evolution from 2020 to 2040 following the IEA Stated Policies Scenario.

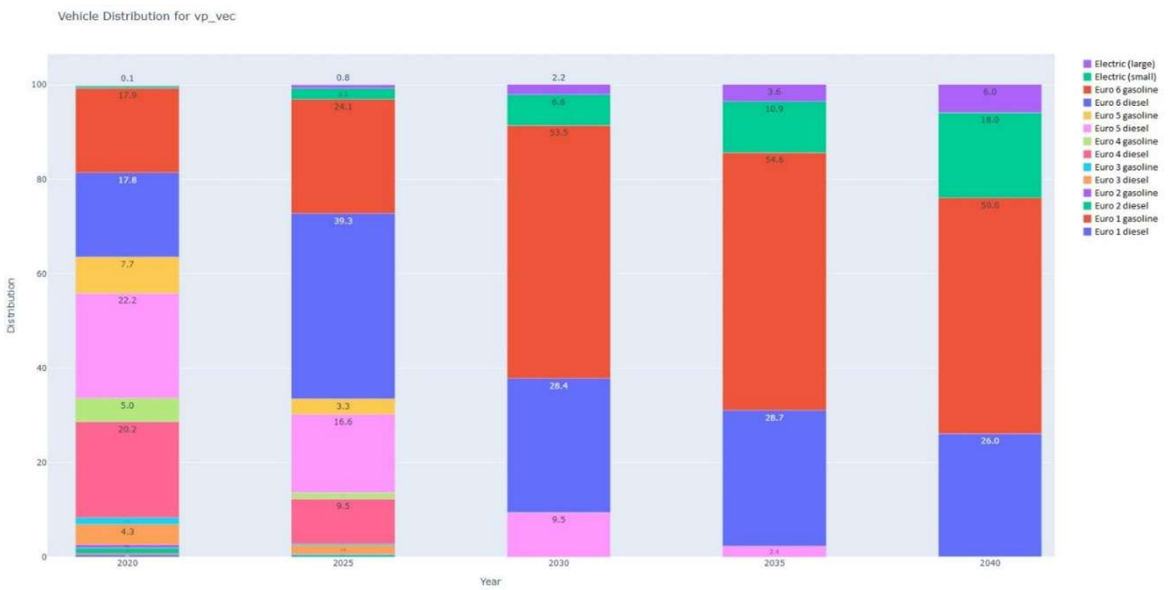

Figure 18 Light Vehicle powertrain type and euro norm distribution in France from 2020 to 2040 following IEA STEP scenario.

2.4. Flexibility modelling

Section 2.2 presented the models that can be used to generate the demand profiles for the vehicles, equipment, vessels, and other systems that support the terminal operations. Building from the models presented, this section will explore the possibility of changing the load curves for different assets to fit specific optimization objectives of the EMT, e.g., maximization of self-consumption, reduction of the curtailment of local renewable generation, or provision of grid services. The first step in the development of the flexibility models for port assets is the definition of the types of flexibility that could be available within a port, and more specifically within a terminal. As an example, the following four types of flexibility have been proposed in the literature [74]:

a) *Load Curtailment:* reduction of total electricity usage without shifting the designated load to any other time period.
b) *Load Shifting:* reschedule and shift of electricity usage from one time period to other time periods.
c) *Utilizing Onsite Generation:* reduction of load by turning on an onsite or backup generator to supply some or all the electricity loads.

d) *Utilizing Energy Storage System (ESS):* energy storage systems are used to supply some or all the electricity needs.

Onsite generation and ESS are not directly related with the load flexibility of port assets because the load itself is unchanged. Therefore, these types of flexibility sources will not be discussed in this section. This section will discuss the two first sources of flexibility, without considering energy storage systems. The flexibility modeling enabled by the ESS is discussed and modelled in section 2.5.

The development of the flexibility models for load shifting or load curtailment starts with modeling the maximum power reduction (also called downward flexibility) for each time step and price of load change. Additionally, for load shifting, a maximum power increase, also called upward flexibility, must be found. The second step of the development of the flexibility models consists in the definition of the constraints that must be considered in EMT.

In the next sub-sections, load flexibility models for different types of seaport load will be described. Within this work, only port specific loads were considered. Also, types of loads that affect logistic of a port were not included because additional estimation of the effect on other assets is required.

Flexibility of OPS systems

As mentioned before, Onshore Power Supply (OPS) or Cold Ironing is an effective measure to decrease emissions. However, currently there are significant economic and technical barriers to the widespread deployment of OPS system. Potentially, this could mean that stakeholders such as policymakers, port authorities, or terminal owners will have to deploy e.g., incentives, subsidies, taxes, to incentivize the installation of these systems in ports. Policymakers have three board types of instruments available to promote emission reductions in the energy activities of ports [75]:

a) *Regulatory approach* – e.g., establishes standards of technological processes. For instance, FuelEU Maritime mandates vessels above 5000 gross tonnes to connect to OPS during berth.
b) *Economic incentive* - i.e., market-based policies, including emission taxes, fees, and subsidies for OPS installation or vessels retrofit.
c) *Hybrid approaches* – consisting of a combination of regulatory and economic incentives.

The choice of instrument to be implemented will affect the flexibility modeling. Additionally, the berth plan is an essential input to model the flexibility of OPS systems. If we define the berth plan by binary variable $B_{s,b,t}$, the value of the variable is equal to 1 if the vessel $s \in \mathbb{S}$ is in berth $b \in \mathbb{B}$ at time $t \in \mathbb{T}$, where $\mathbb{S}$ is the set of vessels that will be allocated with berthing slots in the port, $\mathbb{B}$ is the set of berths in the port, and $\mathbb{T}$ is the set of all operation periods. Other parameters that are necessary for modeling are $P_{b,t}^{sup,ma}$ - the maximum power of onshore power supplier located at berth b in time t; $P_{s,t}^{ves}$ – the power consumption of vessel s in time t; and $P_{b,t}^{OPS,sc}$ - the scheduled power consumption of OPS at berth b in time t. The OPS power is found based on the OPS scheduling algorithm. If the electricity price is known, an optimization algorithm is applied to find an optimal schedule of power consumption. Otherwise, the algorithm can choose any feasible values that satisfy constraints.

2.4.1. Estimation OPS flexibility potential under regulatory policies

Regulatory policies set specific goals for use of OPS systems for all or part of the vessels or terminals. Therefore, to estimate load flexibility, information about the minimum energy consumption requirements from OPS for vessel s (E_s^{min}) and minimum power consumed by vessel s (P_s^{min}) should be provided. These parameters can be provided directly from the policy or derived from requirements on maximum greenhouse gas emissions. In the second case, these minimum requirements could be derived by other tools that are being developed within the MAGPIE project, such as the GHG tool.

In case of regulatory policies, the total scheduled energy consumption of a vessel from the OPS is assumed to be equal to the minimum required energy consumption from the OPS system because there are no economic incentives to increase the consumption:

$$\sum_{t\in\mathbb{T}}\sum_{b\in\mathbb{B}} P_{b,t}^{OPS,sch} \cdot B_{s,b,t} = E_s^{min}, \forall s \in \mathbb{S} \tag{18}$$

Another constraint is related with the minimum and maximum power of the OPS:

$$\sum_{s\in\mathbb{S}} P_s^{min} \cdot B_{s,b,t} \leq P_{b,t}^{OPS,sch} \leq P_{b,t}^{sup,max}, \forall b \in \mathbb{B}, t \in \mathbb{T} \tag{19}$$

Additionally, the OPS power cannot exceed the power consumption of a vessel:

$$P_{b,t}^{OPS,sch} \leq \sum_{s\in\mathbb{S}} P_{s,t}^{ves} \cdot B_{s,b,t}, \forall b \in \mathbb{B}, t \in \mathbb{T} \tag{20}$$

It should be noted that the load flexibility could be achieved only if the following condition is met:

$$\sum_{t\in\mathbb{T}} P_s^{min} < E_s^{min}, \forall s \in \mathbb{S} \tag{21}$$

If this condition is not satisfied, the power consumed from OPS is always equal to the minimum power set by the policy. In case of rescheduling load to provide flexibility, the following conditions must be met for a new OPS power ($P_{b,t}^{OPS}$):

$$P_{b,t}^{OPS} = P_{b,t}^{OPS,sch} + P_{b,t}^{OPS,fl} \tag{22}$$

where $P_{b,t}^{OPS,fl}$ is a change in power consumption from OPS at berth b in time t. If $P_{b,t}^{OPS,fl}$ is negative, the load is reduced. If $P_{s,t}^{OPS,fl}$ is positive, the load is increased. Maximum allowed upward and downward provided flexibility at each time step should be identified by substitution equation (22) with equations (*19*) and (*20*):

$$P_{b,t}^{OPS,fl\,up} = \min\{\left(P_{b,t}^{sup,max} - P_{b,t}^{OPS,sch}\right); \left(\sum_{s\in\mathbb{S}} P_{s,t}^{ves} \cdot B_{s,b,t} - P_{b,t}^{OPS,sch}\right)\} \tag{23}$$

$$P_{b,t}^{OPS,fl\,down} = P_{b,t}^{OPS,sch} - \sum_{s\in\mathbb{S}} P_s^{min} \cdot B_{s,b,t} \tag{24}$$

The cost of these changes is assumed to be zero.

The types of regulatory policy described above allow only load shifting, in accordance with equation (18). To reschedule OPS operation for flexibility provision, the following constraints must be included in EMT:

$$-P_{b,t}^{OPS,fl\,down} \leq P_{b,t}^{OPS,fl} \leq P_{b,t}^{OPS,fl\,up}, \forall b \in \mathbb{B} \tag{25}$$

$$\sum_{t\in\mathbb{T}}\sum_{b\in\mathbb{B}} P_{b,t}^{OPS,fl} \cdot B_{s,b,t} = 0, \forall s \in \mathbb{S} \tag{26}$$

2.4..2. Estimation OPS flexibility potential under market-based policies.

Under market-based policies, there are no constraints on minimum energy and power consumption from OPS. Therefore, equation (18) does not exist, and in equation (19), $P_{b,t}^{OPS}$ has only an upper limit. However, the OPS power cannot be less than 0. Therefore, upward and downward load flexibility are estimated as follows:

$$P_{b,t}^{OPS,fl\,up} = \min\{\left(P_{b,t}^{sup,max} - P_{b,t}^{OPS,sch}\right); \left(\sum_{s\in\mathbb{S}} P_{s,t}^{ves} \cdot B_{s,b,t} - P_{b,t}^{OPS,sch}\right)\} \tag{27}$$

$$P_{b,t}^{OPS,fl\,down} = P_{b,t}^{OPS,sch} \tag{28}$$

Due to absence of any constraints on required energy consumption, there is an opportunity for load curtailment. Therefore, flexibility in EMT could be represented in the following way:

$$P_{b,t}^{OPS,shift} \leq P_{b,t}^{OPS,fl\,up} \tag{29}$$

$$P_{b,t}^{OPS,curt} - P_{b,t}^{OPS,shift} \leq P_{b,t}^{OPS,fl\,down} \tag{30}$$

$$\sum_{t\in\mathbb{T}} P_{b,t}^{OPS,shift} = 0 \tag{31}$$

Where $P_{b,t}^{OPS,shift}$ is shifted load, $P_{b,t}^{OPS,curt}$ is curtailed load.

It is assumed that scheduled OPS power consumption is the most cost effective. Therefore, under market-based policies, changing of OPS power leads to increased expenses on fees and taxes that are associated with GHG emissions. Taxes are paid either for increase in instantaneous emissions or for increase in total emissions for a certain period. If the first approach is used, the cost of flexibility could be calculated by the following price function:

$$C_{b,t}^{flex,inst} = C\left(P_{b,t}^{OPS,curt} - P_{b,t}^{OPS,shift}\right) \tag{32}$$

The function can be linear or a step-function. Note that the cost can be negative, meaning that the solution decreases paid taxes. However, under assumption of most cost-effective solution, the savings are less than electricity cost.

If the second approach is used, load shifting does not affect the cost. Even though the taxes are defined for a certain period, the cost could be identified for each time step.

2.4..3. Estimation OPS flexibility potential under regulatory and market-based policies.

Under both policies, the upward and downward flexibility are calculated according to equations (23) and (24). There are two scenarios that define a flexibility potential for the considered hybrid approach:

1) If condition (18) is met, there is only possibility for load shifting. Accordingly, in the energy management tool, equation (25) and (26) should be considered, and the flexibility cost is identified as follows:

$$C_{b,t}^{flex,inst} = C\left(-P_{b,t}^{OPS,shift}\right) \tag{33}$$

2) For the second scenario, the total scheduled energy consumption of OPS for all berths must be more than the sum of minimum energy requirements for all ships:

$$\sum_{t\in\mathbb{T}}\sum_{b\in\mathbb{B}} P_{b,t}^{OPS} > \sum_{s\in\mathbb{S}} E_s^{min} \tag{34}$$

In this case, there is an opportunity for both, load shifting and load curtailment. Within the energy management tool, conditions (28) to (30) must be included. Costs of flexibility are defined by equations (31) and (32).

Flexibility modelling of Reefer containers

Another type of load that could be changed without disrupting the port logistics are reefer containers. Due to thermal inertia, reefer containers could have a significant potential for flexibility [76]. One of the approaches that could be used for the estimation of this flexibility is based on the internal temperature estimation of reefers. This approach has been previously used to model a real time control to maintain the temperature of containers within necessary limits, and to reduce operational cost. First, the control algorithm estimates the internal temperature for the next time step. Then, it identifies a flexibility of each container. And finally, aggregates all containers by descending flexibility and identify optimal control strategy using fuzzy logic. To use similar approach for flexibility modelling, information about control algorithm is necessary. Suppose the algorithm provide information about power consumption ($P_{i,t}^{cool,sch}$) and internal temperature ($T_{i,t}$) of reefers $i \in \mathbb{I}$ for each time step within selected interval $t \in \mathbb{T}$. The temperature is defined based on the following equation:

$$T_{i,t+1} = T_{i,t} - \frac{P_{i,t}^{cool,sch}\Delta t}{m_i c_i} + \Delta T_{amb}\left(1 - e^{-\frac{Ak}{m_i c_i}\Delta t}\right) \tag{35}$$

where: $StR_{i,t}$ is the binary variable that defines reefer's operational state ($StR_{i,t} = 1$ if reefer the reefer is switched on); Δt is a time step; m_i is the mass of reefer's content (in kg); c_i is the specific heat capacity of reefer's content (in J/kg·K); ΔT_{am} is the difference between reefer's internal temperature and ambient temperature (in °C); A is reefer's surface area (in m²); k is the heat transition coefficient of reefer's content (in W/m²·K).

The cooling power can be divided in two components:

$$P_{i,t}^{cool,sch} = P_{i,t}^{cool,base} + P_{i,t}^{cool,ch} \tag{36}$$

where $P_{i,t}^{cool,base}$ is the base power consumption of container, i.e., the power needed for maintenance of the current temperature, and $P_{i,t}^{cool,ch}$- is the power used to change the internal temperature of the container.

To find the base power, in equation (36) we need to set $T_{i,t+1} = T_{i,t}$:

$$0 = -\frac{P_{i,t}^{cool,base}\Delta t}{m_i c_i} + \Delta T_{amb}(1 - e^{-\frac{Ak}{m_i c_i}\Delta t}) \tag{37}$$

And derive base power from the following equation:

$$P_{i,t}^{cool,base} = \left(T_{amb,t} - T_{i,t}\right) \cdot \left(1 - e^{-\frac{Ak}{m_i c_i}\Delta t}\right) \cdot \frac{m_i c_i}{\Delta t} \tag{38}$$

For the predefined time interval:

$$P_{i,t}^{cool,base} = \left(T_{amb,t} - T_{i,t}\right) \cdot C_1 \tag{39}$$

Where $C_1 = \left(1 - e^{-\frac{Ak}{m_i c_i}\Delta t}\right) \cdot \frac{m_i c_i}{\Delta t}$ is constant for all time intervals.

To find the second part of cooling power, we need to move $T_{i,t}$ to the left part in equation (35) and multiply the equation by $\frac{m_i c_i}{\Delta t}$:

$$(T_{i,t+1} - T_{i,t}) \cdot \frac{m_i c_i}{\Delta t} = -\left(P_{i,t}^{cool,base} + P_{i,t}^{cool,ch}\right) + \Delta T_{amb} \cdot (1 - e^{-\frac{Ak}{m_i c_i}\Delta t}) \cdot \frac{m_i c_i}{\Delta t} \tag{40}$$

$\Delta T_{amb} \cdot (1 - e^{-\frac{Ak}{m_i c_i}\Delta t}) \cdot \frac{m_i c_i}{\Delta t}$ is the base power defined in equation (*35*). Therefore:

$$P_{i,t}^{cool,ch} = (T_{i,t} - T_{i,t+1}) \cdot \frac{m_i c_i}{\Delta t} \tag{41}$$

For the predefined time interval:

$$P_{i,t}^{cool,ch} = (T_{i,t} - T_{i,t+1}) \cdot C_2 \tag{42}$$

Where $C_2 = \frac{m_i c_i}{\Delta t}$ is a constant for all time intervals.

In the proposed flexibility model, the power used for changing temperature is a result of the implementation of an optimization algorithm. The base power is computed for the next time step according to the following equation:

$$P_{i,t}^{cool,base} = \left(T_{amb,t} - T_{i,0} + \sum_{\tau=0}^{t} \frac{P_{i,\tau}^{cool,ch}}{C_2}\right) \cdot C_1 \tag{43}$$

This model requires that information about ambient temperature in all time intervals and internal temperature of reefers is known. The total power consumption is bounded by a capacity of a power supplier:

$$P_{i,t}^{cool,base} + P_{i,t}^{cool,ch} \leq P_{max} \tag{44}$$

Also, it cannot be less than 0:

$$P_{i,t}^{cool,base} + P_{i,t}^{cool,ch} \geq 0 \tag{45}$$

The following constraints ensure that internal temperature of the reefers is maintained within predefined bounds:

$$P_{i,t}^{cool,base} \geq \left(T_{amb,t} - T_i^{max}\right) \cdot C_1 \tag{46}$$

$$P_{i,t}^{cool,base} \leq \left(T_{amb,t} - T_i^{min}\right) \cdot C_1 \tag{47}$$

Note that the implementation of the proposed model can be bounded by the error of temperature estimation. The error increases with the increasing of a prediction horizon. To define applicability of the proposed model, the estimation of the error should be investigated.

Flexibility modelling of cranes

Another possible source of load flexibility in a terminal are cranes. The scheduling of quay and yard cranes are based on the arrival time of vessels. The optimization algorithms in the literature aim to minimize a time of crane operation to reduce the waiting time of vehicles. When the number of vessels at berth is low, there is a time interval to perform a certain job, which could be used to provide flexibility for the electrical network. This section describes in detail the flexibility modelling of yard cranes but could be easily adapted for quay cranes.

Operation of yard cranes consists of three steps: upward movement, horizontal movement (translation), and downward movement, as represented in Figure 19:

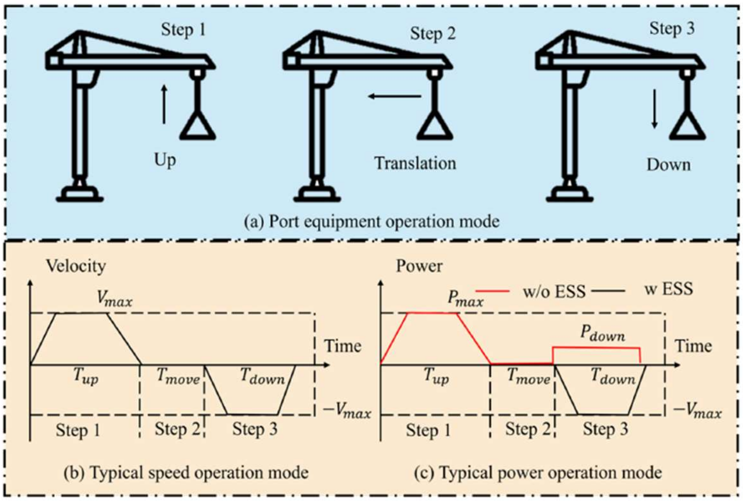

Figure 19 - Power model of port cranes [36].

The power consumption of cranes is negligible and assumed to be zero during the horizontal movement. In cases without EES, the power consumption for the downward movement is constant but can be neglected as well. Some literature [77] propose a trapezoidal model of port equipment power with constant acceleration. This section describes the the general case with ESS but can be easily implemented for the first step (upward movement) and the third step (downward movement) if ESS is not used. The load shifting described by the model is represented in Figure 20.

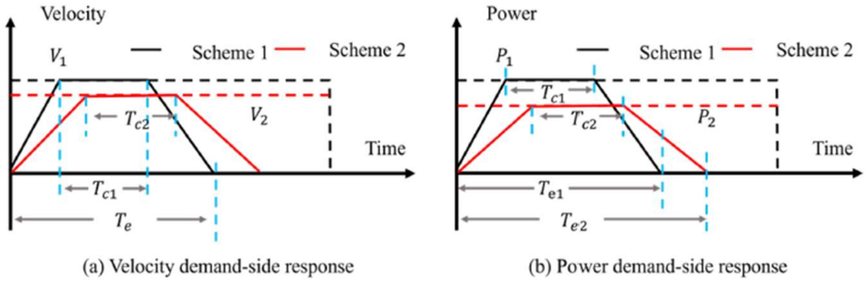

Figure 20 - Demand response model of port cranes[57].

In this model, one possible source of flexibility results from the reduction in the power consumption of cranes for the upward movement due to the reduction of the lifting velocity which leads to increased time to complete the job. The total energy consumed for this action is assumed to be constant. Therefore, the following equation is used to find the power consumption of the crane:

$$\frac{1}{2}P^{t}_{crane}(T^{t}_{c} + T^{t}_{e}) = E_{up}, \tag{48}$$

Using this model, one can shift the crane's power consumption by increasing time of operation. Another option is to shift the job start time. Both options for rescheduling power consumption will be considered in one problem. For the proposed model, the acceleration time $(T_e - T_c)$ can be neglected because the chosen time step would be much higher. Therefore, for simplicity we can use "rectangular" model with modified notations:

$$P^{up}_{y,j} \cdot \sum_{t \in \mathbb{T}} B^{up}_{y,j,t} = E^{up}_{y,j}, \qquad \forall y \in \mathbb{Y}, j \in \mathbb{J} \tag{49}$$

$$P^{down}_{y,j} \cdot \sum_{t \in \mathbb{T}} B^{down}_{y,j,t} = E^{down}_{y,j}, \qquad \forall y \in \mathbb{Y}, j \in \mathbb{J} \tag{50}$$

Where, the indices are defined as: y is the index of yard cranes from set of cranes $\mathbb{Y}$; j is the index of jobs to be performed from set of jobs $\mathbb{J}$, and t is the index of time from set $\mathbb{T}$. For the constants: $E^{up}_{y,j}$ is the energy that is needed for upward movement within job j by crane y; $E^{down}_{y,j}$ is the energy that is stored in ESS during the downward movement within job j by crane y, which would be set to zero is no EES is available. For the variables: $P^{up}_{y,j}$ is the power consumption of crane y during job j; $P^{down}_{y,j}$ is the power production of crane y during job j; $B^{up}_{y,j,t}$ is the binary variable that is equal to 1 if the crane y is making an upward movement during job j at time t; $B^{down}_{y,j,t}$ is the binary variable that is equal to 1 if crane y is making a downward movement during job j at time t. Additionally, a binary variable is also introduced to represents the translation movement, $B^{move}_{y,j,t}$, which is a binary variable that is equal to 1 if crane y moves horizontally during job j at time t. In the proposed model the operation of the cranes also includes idling because as power consumption is assumed to be 0. To satisfy the sequence of operations represented in Figure 20, the following variables are introduced:

- $f_{y,j,t}^{up}, f_{y,j,t}^{down}, f_{y,j,t}^{move}$ starting indicators of upward/downward and horizontal movement; equal to 1 if the crane y starts an upward/downward movement within job j at time t,
- $l_{y,j,t}^{up}, l_{y,j,t}^{down}, l_{y,j,t}^{move}$ stopping indicators of upward/downward and horizontal movement; equal to 1 if the crane y starts an upward/downward movement within job j at time t.

The variables for upward movement are found in equations (51) and (52), the variables for downward movement are (53), (54) and for horizontal movement (55), (56).

$$f_{y,j,t}^{up} - l_{y,j,t}^{up} = B_{y,j,t}^{up}, -B_{y,j,t-1}^{up}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{51}$$

$$f_{y,j,t}^{up} + l_{y,j,t}^{up} \leq 1, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{52}$$

$$f_{y,j,t}^{down} - l_{y,j,t}^{down} = B_{y,j,t}^{down}, -B_{y,j,t-1}^{down}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{53}$$

$$f_{y,j,t}^{down} + l_{y,j,t}^{down} \leq 1, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{54}$$

$$f_{y,j,t}^{move} - l_{y,j,t}^{move} = B_{y,j,t}^{move}, -B_{y,j,t-1}^{move}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{55}$$

$$f_{y,j,t}^{move} + l_{y,j,t}^{move} \leq 1, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{56}$$

Within one job, the following constraints must be satisfied to ensure the proposed sequence:

$$l_{y,j,t}^{up} = f_{y,j,t}^{move}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{57}$$

$$l_{y,j,t}^{down} = f_{y,j,t}^{down}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{58}$$

To ensure that after one job, another is immediately started, we need to introduce the following constraint:

$$f_{y,j,t}^{up} = l_{y,j-1,t}^{down}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}, \mathrm{t} \in \mathbb{T} \tag{59}$$

The job of the cranes consists of loading/unloading of vehicles and rearranging containers within the storage area. The rearranging operation is a flexible job because rescheduling of these actions does not affect the logistic part, while the loading and unloading operations cannot be shifted. Therefore, some operations must be fixed:

$$f_{y,j,t}^{up} = const, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}_{load}, \mathrm{t} \in \mathbb{T} \tag{60}$$

$$l_{y,j,t}^{down} = const, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J}_{unload}, \mathrm{t} \in \mathbb{T} \tag{61}$$

Another set of constraints defines the minimum time for performing each action:

$$\sum_{t\in\mathbb{T}} B_{y,j,t}^{up} = T_{y,j}^{up\ min}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J} \tag{62}$$

$$\sum_{t\in\mathbb{T}} B_{y,j,t}^{move} = T_{y,j}^{up\ move}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J} \tag{63}$$

$$\sum_{t\in\mathbb{T}} B_{y,j,t}^{down} = T_{y,j}^{up\ down}, \forall y \in \mathbb{Y}, \mathrm{j} \in \mathbb{J} \tag{64}$$

The minimum time to perform the action by a crane can be found based on the maximum speed of the movement of the crane and distance (in the case of yard cranes). The last step is to define power consumption of each crane in each time step based on the provided information:

$$P_{y,t} = \sum_{\mathrm{j}\in\mathbb{J}} P_{y,j}^{up} \cdot B_{y,j,t}^{up} - \sum_{\mathrm{j}\in\mathbb{J}} P_{y,j}^{down} \cdot B_{y,j,t}^{down}, \forall y \in \mathbb{Y}, t \in \mathbb{T} \tag{65}$$

Accordingly, the part of the flexible load can be found by the following equation:

$$P_{y,t}^{fl} = P_{y,t}^{sch} - P_{y,t}, \forall y \in \mathbb{Y}, t \in \mathbb{T} \tag{66}$$

Considering all constraints together will result in the most cost-effective way of operation of yard cranes. However, the problem might be hard to solve within EMT because of the additional non-linearities and a lot of constraints for binary variables and may require simplification in the next stages of implementation.

2.5. Charging rules for battery-based vessels, terminal vehicles and equipment

Many vehicles are operating on a port terminal, from gantry cranes, straddle carriers, reach stacker to Mafi or terminal tractors. Nowadays diesel engines power many of them and few of them are powered with hybrid electric-diesel engines (for instance straddle carriers) or are electrical (mainly the gantry cranes).

The port terminal is also an interface with many other land-based vehicles as trains and trucks, currently mainly powered by diesel engines as explained in the Deliverable 3.1.

All the vessels (e.g., cargo, barge, container ships) within the port area also consume energy for moving but also when goods and people are charged and discharged due to auxiliary loads. Up to now, diesel engines produce this energy. As mentioned before, in a near future, vessels at the docks have to get an electrical connection to power supply the auxiliary loads from the port electrical grid and without using diesel engines. This requirement of an electrical power supply for auxiliaries will be also applied for vessels waiting for a slot at the terminal's docks in the surroundings of the port; offshore power supply (OPS) or a better schedule of available docks' timeslot some days before vessel time arrival will be necessary. Whatever the **vessel** type is, the required electrical power for the auxiliaries can be forecasted with good accuracy, even before the vessel's arrival, and is likely not flexible. No charging rules are useful, and power is supplied at the required value for each time step. If the vessel has a battery (for propulsion or other purposes), the charge of the battery from the port electrical grid is flexible, and charging rules are required. The battery could be used as a flexible source to power supply the auxiliaries too. The following parameters are needed to integrate the charging of the vessel's battery within the Magpie energy matching tool.

$t_{i,vess\_arrival}$, time of vessel *i* arrival. It is known from port terminal schedule.

$t_{i,vess\_departure}$, time of vessel *i* departure. It is known from port terminal schedule.
$D_{i,vess\_connect}$, time for electrical connection of vessel *i*
$D_{i,vess\_disconnect}$, time for electrical disconnection of vessel *i*
$t_{i,vess\_init}$, time for starting electrical supply of vessel *i*. $t_{i,vess\_init} = t_{i,vess\_arrival} + D_{i,vess\_connect}$.
$t_{i,vess\_end}$, time for stopping electrical supply of vessel *i*. $t_{i,vess\_end} = t_{i,vess\_departure} - D_{i,vess\_disconnect}$.
$Enom_{i,vess\_batt}$, the nominal energy of the battery of vessel *i*.
$SOC_{i,vess\_batt\_init}$, the initial *i* vessel's battery state of charge (SOC), when it is electrically connected at dock. Is equal to 0 if battery is empty and 1 if battery is full. As an assumption, the value is unknown (no forecast) before the vessel arrives to the charging station of the port terminal.
$SOC_{i,vess\_batt\_end}$, the required *i* vessel's battery state of charge (SOC), when it is electrically disconnected from the dock
$SOC_{i,vess\_batt\_min}$, the minimal *i* vessel's battery state of charge (SOC)
$SOC_{i,vess\_batt\_max}$, the maximal *i* vessel's battery state of charge (SOC)
$Pch_{i,vess\_batt\_min}$, the minimal *i* vessel's power for charging battery
$Pch_{i,vess\_batt\_max}$, the maximal *i* vessel's power for charging battery
$Pdch_{i,vess\_batt\_min}$, the minimal *i* vessel's power for discharging battery
$Pdch_{i,vess\_batt\_max}$, the maximal *i* vessel's power for discharging battery
$\eta_{i,vess\_batt}$, the mean efficiency to charge and discharge the *i* vessel's battery. If value is not available, 1 is the default value.
P_{j,ch_point_min} , the minimal *j* charging point's power
P_{j,ch_point_max} , the maximal *j* charging point's power
η_{j,ch_point} , the mean efficiency of the charging point *j*. If value is not available, 1 is the default value.

At each time *t*, the power applied for the vessel *i* at charging point *j* for charging vessel's battery respects

In charge, $-\min(Pch_{i,vess\_batt\_max}, P_{j,ch\_point\_max}) \leq Pvess\_batt_{i,j}(t) \leq -\max(Pch_{i,vess\_batt\_min}, P_{j,ch\_point\_max}) \leq 0$

In discharge, $0 \geq \max(Pdch_{i,vess\_batt\_min}, P_{j,ch\_point\_max}) \geq Pvess\_batt_{i,j}(t) \geq \min(Pdch_{i,vess\_batt\_max}, P_{j,ch\_point\_max})$

And the *i* vessel's battery SOC at time *t* is calculated as $SOC_i(t) = SOC_i(t-1) - Pvess\_batt_{i,j}(t) \times \eta_{i,vess\_batt} \times \eta_{j,ch\_point} \times \Delta t$

And, this SOC must always respect $SOC_{i,vess\_batt\_min} \leq SOC_i(t) \leq SOC_{i,vess\_batt\_max}$

As the electrical connection to the dock is shared between power for vessel's auxiliaries $Pvess_{aux_{i,j}}$ and charging / discharging power of the vessel's battery $Pvess\_batt_{i,j}(t)$, if any, the power applied $P_{i,j}(t)$ at charging point *j* for the vessel *i* must respect at any time

$$P_{j,ch\_point\_min} \leq |P_{i,j}(t)| \leq P_{j,ch\_point\_max} \text{ with } P_{i,j}(t) = Pvess_{aux_{i,j}}(t) + Pvess\_batt_{i,j}(t)$$

At the time departure $t_{i,vess\_end}$ of the vessel *i* from the port, vessel's battery SOC must be equal to $SOC_i(t_{i,vess\_end}) = SOC_{i,vess\_batt\_end}$

Beyond the electrical needs for the vessels (auxiliaries and possible battery), it is likely having an electrification of some of the trucks coming in and out from the port terminal and of the

vehicles operating within the terminal (from gantry cranes, straddle carriers, reach stacker and mafi).
Magpie project aims to evaluate the impact on electricity needs within port and terminal framework of such electrification scenario.

For the **trucks coming in and out from the port terminal**, a similar approach of charging parameters / rules than for the vessels can be proposed for the energy matching tool. As for the vessel, the time arrival and the time departure of the truck is driven by logistics optimization. The main difference is there is no need for auxiliary power supply for the trucks. It is unlikely the trucks coming in and out of the port terminal will have Vehicle-to-Grid (V2G, capable to charge and to discharge its battery from/to the electrical grid) capability or will, it is assumed that the trucks will only require charging power.

The following parameters and equations can be proposed for trucks coming in and out of the port terminal.

- $t_{i,truck\_init}$, time of truck *i* arrival in port terminal and could start its charge. Unknown (no forecast).
- $t_{i,truck\_end}$, time of vessel *i* departure from port terminal and charge must be finished.
- $Enom_{i,truck}$, the nominal energy of the battery of truck *i*.
- $SOC_{i,truck\_init}$, the initial *i* truck's battery state of charge (SOC), when it arrives at port terminal. Is equal to 0 if battery is empty and 1 if battery is full. As an assumption, the value is unknown (no forecast) before the truck arrives to the charging station of the port terminal.
- $SOC_{i,truck\_end}$, the required *i* truck's battery state of charge (SOC), when it will leave the port terminal
- $SOC_{i,truck\_max}$, the maximal *i* truck's battery state of charge (SOC)
- $P_{i,truck\_min}$, the minimal *i* truck's power for charging battery
- $P_{i,truck\_max}$, the maximal *i* truck's power for charging battery. It can be automatically calculated by the charging station, if using the most recent standard.
- $\eta_{i,truck}$, the mean efficiency to charge and discharge the *i* truck's battery. If value is not available, 1 is the default value.
- P_{j,ch_point_min} , the minimal *j* charging point's power
- P_{j,ch_point_max} , the maximal *j* charging point's power
- η_{j,ch_point} , the mean efficiency of the charging point *j*. If value is not available, 1 is the default value.

At each time *t*, the power applied for the truck *i* at charging point *j* for charging truck's battery respects

$$-\min\left(P_{i,truck\_max}, P_{j,ch\_point\_max}\right) \leq Ptruck_{i,j}(t) \leq -\max\left(P_{i,truck\_min}, P_{j,ch\_point\_max}\right) \leq 0$$

And the *i* truck's battery SOC at time *t* is calculated as $SOC_i(t) = SOC_i(t-1) - Ptruck_{i,j}(t) \times \eta_{i,truck} \times \eta_{j,ch\_point} \times \Delta t$

And, this SOC must always respect $SOC_i(t) \leq SOC_{i,truck\_max}$

At the time departure $t_{i,truck\_end}$ of the truck *i* from the port terminal, truck's battery SOC must be equal to $SOC_i(t_{i,truck\_end}) = SOC_{i,truck\_end}$

A third category of **electrical vehicles** is the ones that **operate within the port terminal but are always connected to the electrical grid**. It is the case for the gantry cranes for instance. They can have a small battery charged when a container goes down, helping for moving up the next container. This battery is only for local power optimization and cannot be considered as a flexible source by the energy matching tool.
For these electrical loads, no charging rules are required. The power consumption of these

vehicles is directly linked to loading and unloading activities, i.e., logistic and supply chain planning. Needed power needs to be supplied without flexibility, but it can be forecasted based on planned activities and vehicle power characteristics (for instance values explained in section 4.3..2). Calculating the power and energy needs for these vehicles considering the logistics moves generated by the terminal port simulator is the role of the '*time-dependent energy requirements and flexibility*' tool (see Figure 2 in section 1.3).

Finally, the last category of vehicles operating for the port terminal activities is the **moving vehicles within the port terminal but without permanent electrical connection**. Mafi, straddle carriers and reach stackers are examples of this type of vehicles. All the power consumptions and the characteristics of the vehicles are known are they are part of the fleet to operate the terminal. In case they are electrical, they have a battery permitting operating for several hours before going to the charging station; as for gantry cranes, some of them could charge their battery when container moves down.

In the framework of the tools developed for Magpie, **it is assumed that the timeslot for charging such terminal vehicles are given by the logistic /supply chain optimization and the terminal simulator**; these timeslot are the time periods when the vehicle is not used for loading and unloading the containers. Hence, the logistic planning is not optimized to charge the vehicles at the best time (it could be done by integrating charging requirements within the logistic / supply chain optimization, but it is not the case here), but the vehicles are charged when it is possible from a logistical schedule.

As explained in section 2.2., the '*time-dependent energy requirements and flexibility*' tool developed for Magpie will be able to calculate the power and energy needs of these vehicles based on the planned activities and their power characteristics (e.g., power consumption, efficiency). It will be able to forecast the power and energy needs for the vehicles after the charging timeslot and before the next one thanks to the same type of information. Hence, the battery SOC when it arrives at charging station and the required SOC for leaving the charging station can be forecasted.

First charging rule for terminal moving vehicles is the value of the minimal SOC, $SOC_{veh\_op\_min}$, under which the vehicle must be charged and is removed temporary from the terminal operating fleet in order to be charged. This situation should be avoided as much as possible with a fleet large enough and with a sufficient autonomy (linked to the nominal energy of each vehicle *i*, $Enom_{i,veh}$) to operate properly the logistic moves in the given time; otherwise it will directly impact the logistics' operations and the planned activities will not be performed if no substitution vehicle is available. Depending on the nominal energy of the vehicle's battery $Enom_{i,veh}$, the energy consumption of the vehicle for moving and the maximal distance of the charging station, $SOC_{veh\_op\_min}$, the value can be determined as close to zero as possible (for instance 0.01 or 0.02 above the minimal SOC value for the battery of the vehicle). A fast charging flag $FCh_{i,veh}$ could be activated to organize the fastest possible charge of the vehicle *i* (1 if fast charging is required, 0 else).

Another charging rule could be to set a SOC value alarm $SOC_{ve\ \_alarm}$ to activate a flag for the energy matching tool indicating the vehicle will soon have the battery empty. It could be the SOC value allowing a 15 or 30 minutes time period of operation, in order to let the energy matching tool checking if all the foreseen operations are feasible or if a solution with another vehicle or a charge of this vehicle *i* should be proposed.

Apart from these two SOC thresholds, the charge of such vehicles operating within the port terminal is free in the available period when the vehicle is not used for loading and unloading the containers; it could be between two vessels loading/unloading periods, or when workforce conductor team is changing or when the vehicles is waiting in the lane to get a container. As for vessels and trucks, several parameters are needed to plan the charging of the fleet of this kind of vehicles.

$t_{i,tve\ \_init}$, time of vehicle *i* availability for charging as it is not used for logistics or it is empty and must be charged. Forecasted by port terminal simulator and '*time-dependent energy requirements and flexibility*' tool. Duration for connecting the vehicles for charging purposes could be decreased by using autonomous charging plug as the ones proposed by Rocsys, a member of Magpie project.

t_{i,veh_end}, time of vehicle *i* end of availability for charging as it must participate to logistics' tasks.

$Enom_{i,veh}$, the nominal energy of the battery of vehicle *i*.

$SOC_{i,ve\ \_init}$, the *i* vehicle's battery state of charge (SOC), when it arrives at charging point. Is equal to 0 if battery is empty and 1 if battery is full. The value can be forecasted by the '*time-dependent energy requirements and flexibility*' tool based on the vehicle's logistic activities (from port terminal simulator) and its electrical characteristics (consumption for different tasks and moves, nominal energy, last SOC …).

SOC_{i,veh_end} , the required *i* vehicle's battery state of charge (SOC), when it will leave the charging point. The value is calculated by the '*time-dependent energy requirements and flexibility*' tool based on the vehicle's future activities before next charging period and its electrical characteristics (consumption for different tasks and moves, nominal energy, last SOC …).

SOC_{i,veh_max} , the maximal *i* vehicle's battery state of charge (SOC)

SOC_{i,veh_min} , the minimal *i* vehicle's battery state of charge (SOC)

P_{i,veh_min} , the minimal *i* vehicle's power for charging battery

P_{i,veh_max} , the maximal *i* vehicle's power for charging battery. It can be automatically calculated by the charging station, if using the most recent standard.

$\eta_{i,veh}$, the mean efficiency to charge and discharge the *i* vehicle's battery. If value is not available, 1 is the default value.

$FCh_{i,veh}$,indicator for showing a charge as fast as possible is needed for this *i* vehicle

P_{j,ch_point_min} , the minimal *j* charging point's power

P_{j,ch_point_max} , the maximal *j* charging point's power

η_{j,ch_point} , the mean efficiency of the charging point *j*. If value is not available, 1 is the default value.

At each time *t*, the power applied for the vehicle *i* at charging point *j* for charging vehicle's battery respects

$$-\min\left(P_{i,veh\_max}, P_{j,ch\_point\_max}\right) \leq Pveh_{i,j}(t) \leq -\max\left(P_{i,ve\ \_min}, P_{j,ch\_point\_max}\right) \leq 0$$

And the *i* vehicle's battery SOC at time *t* is calculated as $SOC_i(t) = SOC_i(t-1) - Pveh_{i,j}(t) \times \eta_{i,veh} \times \eta_{j,ch\_point} \times \Delta t$

And, this SOC must always respect $SOC_i(t) \leq SOC_{i,veh\_max}$

At the time departure t_{end} of the vehicle *i* from the charging station, truck's battery SOC must be equal to $SOC_i\left(t_{i,veh\_end}\right) = SOC_{i,ve\ \_end}$

If the charge of the whole fleet of port terminal vehicles is not optimized by the energy matching tool, all the vehicles will request a charge at maximum full power (of the vehicle and of the charging point) as soon as they connect to the charging point. It could generate a high peak of power supply on the electrical grid, leading to possible congestion issues. It could also lead to charging the vehicles not during the lowest price possible timeslot or the possible timeslot with the highest renewable electricity share. In this case some additional parameters could be useful to manage the charging station.

$Type_{j,ch\_point}$, the type of charging point j (could be DC plug of different types and AC plug of different types)
$Phase_{j,ch\_point}$, the phase of the electrical grid used by the j charging point
$Avail_{j,ch\_point}$, time series of the availability of the j charging point

For practical reason, the parameters $t_{i,tv\ \_init}$ and $t_{i,ve\ \_end}$ can be gathered within a common time series of vehicles availability at charging station under time: 1 if the vehicle is available at charging station and 0 if not.

An example of such charging optimization of an electrical fleet is given hereafter. The optimal fleet management has to: satisfy operator needs, i.e. giving enough energy before vehicle departure at $t_{i,ve\ \_end}$, respect the electrical limitations of the charging infrastructure of the charging station, and to achieve specific objectives regarding vehicles charging (at the lowest price, as fast as possible, highest use or renewable energy, …) Once realistic data will be available for a fleet of one of the Magpie port or from the port terminal simulator (based on data given by port operators), it could be implemented directly into the energy matching tool or in a dedicated tool. It integrates the electrical characteristics (electrical models) of the charging points and the whole charging station (made of several charging points); it permits considering the operating limitations as power limitations and balancing between phases for the optimization. Having optimal electrical fleet management integrated in the energy matching tool permits to get the optimal energy management of the area at a whole but it requires that the manager of the energy matching tool is also the charge point operator (CPO). Having it as an external tool communicating to the energy matching tool would maybe lead to a less optimal energy management as it is first optimized for the optimal fleet management within the terminal, and then optimized as a whole in the terminal considering the optimal electrical profile for charging the fleet as an input.

Between the different charge points and the charging station management tool, managed by the CPO, the communication standard is OCPP. An OCPI, Open charge point interface, is necessary to have the gateway between CPO and energy optimal management tool, part of the energy matching tool or not. Figure 21 illustrates this. Optimally all the data and models are stored in the port digital twin information system with dedicated databases.

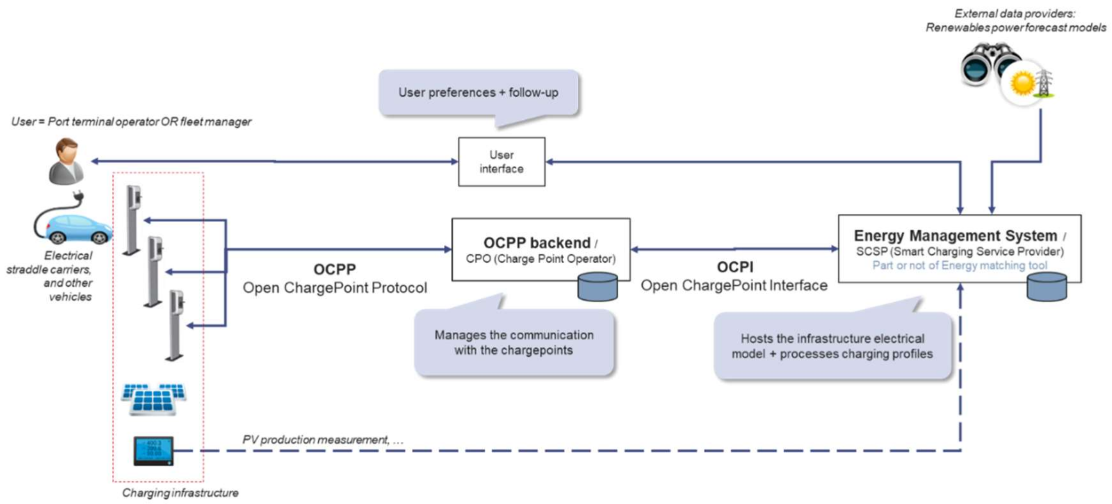

Figure 21 - Overview of data exchange for optimal management of a charging station with several charging points for straddle carriers or terminal vehicles considering renewables electricity production or not.

For each charging session, an ID is allocated and different steps are performed before starting the charge. The arrival state of charge, $SOC_{i,ve\ \_init}$, can be communicated or modified, if already forecasted. The departure time, $t_{i,ve\ \_end}$, and the required SOC at the end of the charging station, SOC_{i,veh_end}, is given by user or by the '*time-dependent energy requirements and flexibility*' tool and associated to the charging session ID. The maximal and minimal powers can be calculated automatically if the charging point protocol is OCPP 1.6 or more recent OCPP.

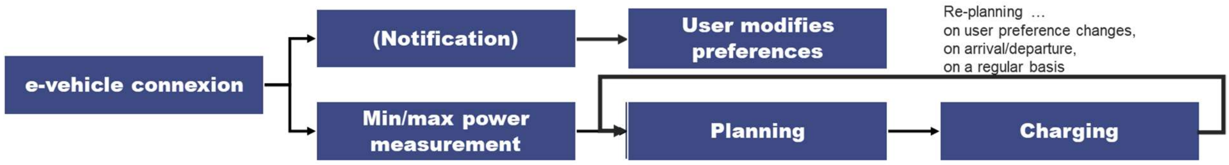

Figure 22 – Typical charging session steps

A charging profile (as a time series) for each connected vehicles is generated as an output of the optimal energy charging fleet tool taking into account the different optimization goals (giving the right amount of required energy for each vehicles, respecting the electrical infrastructure limitations, lowest price of recharging fleet for CPO, higher amount of renewable electricity for charging fleet …). The planning for each electrical vehicle operating on port terminal and the whole fleet charging planning can be updated at a regular time step, or each time a new event occurs (new vehicle, deviation between required energy and charged energy, deviation between renewable electricity production forecast and real production …). An example of a fleet charging profile (one color is one vehicle) without optimization (charging as soon as it is connected to charging point) and with optimization to maximize PV (photovoltaic) self-consumption and self-production is given by following picture. In this specific example, there are 18 charging sessions planned and the PV self-consumption ratio is about 66 % without optimization and 95 % with optimization.

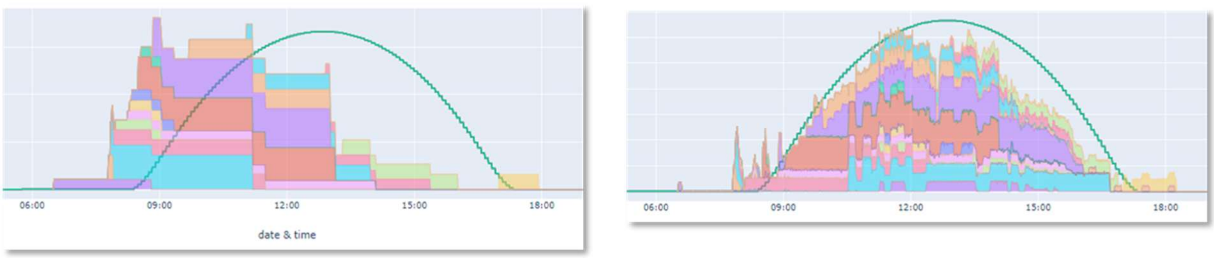

Figure 23 – Two daily charging profile for a fleet of electrical vehicles with 18 charging sessions: not optimized (left), and optimized to maximize PV self-consumption (right). The green bell curve is the PV production forecast

The following picture sum-up the possible charging rules defined for the 4 kinds of electrical vehicles in this part: vessels with or without embedded, e-trucks, port terminal electrical vehicles connected to electrical grid, port terminal electrical vehicles with battery and not connected to electrical grid except at charging station.

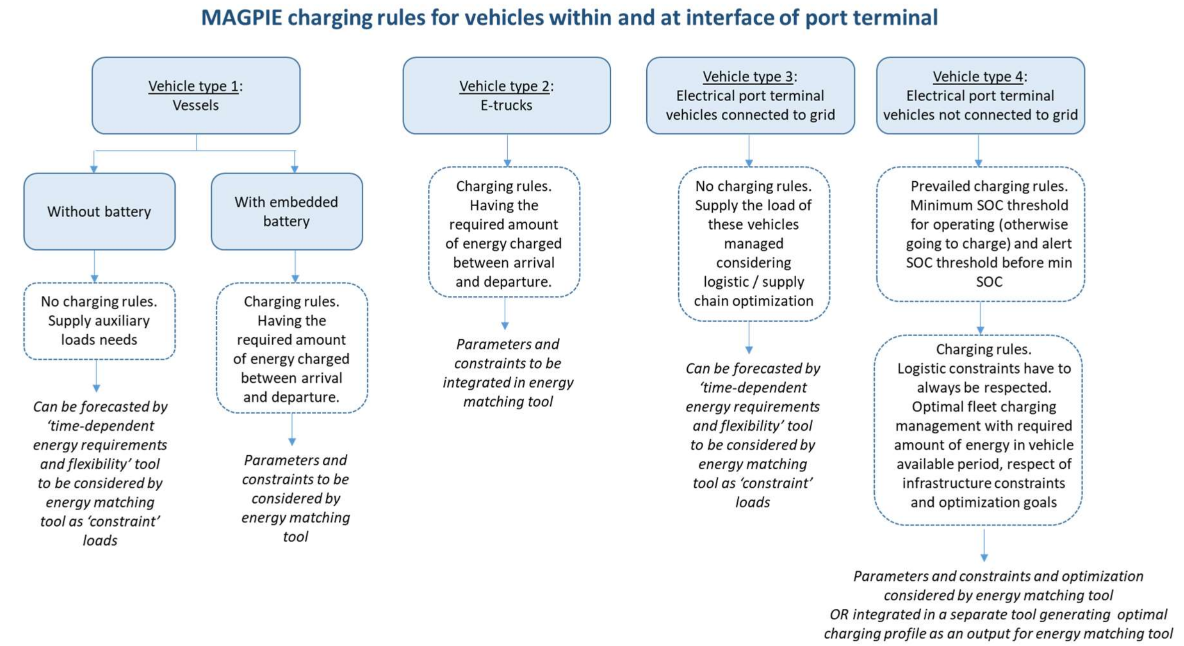

Figure 24 – Magpie charging rules proposal for the electrical vehicles operating within a port terminal or at its interfaces

2.6. Renewable electricity production forecast models

This section describes the models that will be developed to forecast the electricity production from renewable resources. Specifically for the first implementation in the suite of models being developed under T4.4, and T4.5, the forecast will only cover solar PV and wind resources in a first instance. The models presented here are preliminary and will likely need to be adapted in the upcoming months, as data availability may constrain some of the developments. As solar and wind energy become increasingly integral to the port energy matrix, its inherent variability, shaped by sunlight intensity or wind speed, cloud cover, and temperature, demands accurate forecasting. Especially, as accurate forecasting of renewable electricity production enables grid operators, utilities, and individual consumers to predict and manage energy availability. Moreover, in a decentralized smart grid environment, where consumers double as energy producers, accurate forecasts are essential for better use of energy storage, consumption management, and even grid feedback.

The problem of renewable electricity (RE) forecasting generically consists in determining the optimal parameters (θ) of a model (g) to estimate future power output ($\hat{p}_{t+k|t}$) [78]:

$$\hat{p}_{t+k|t} = g\left(p_t, p_{t-1}, \dots, p_{t-l}, x_t, x_{t-1}, \dots, x_{t-l}, \hat{x}_{t+k|t}, \theta\right) \tag{67}$$

The model is run at time t for a given lead time k. In this generic formulation, p_t, p_{t-1} and x_t, x_{t-1} are the present and past observations of the predicted and predictor variables, respectively, and $\hat{x}_{t+k|t}$ are the forecasts of the predictors along the forecast horizon. To this end, the development of RE forecasting models from raw data to the forecasting itself usually follows a systemic process [79], [80], [81], [82], [83] as represented in Figure 25.

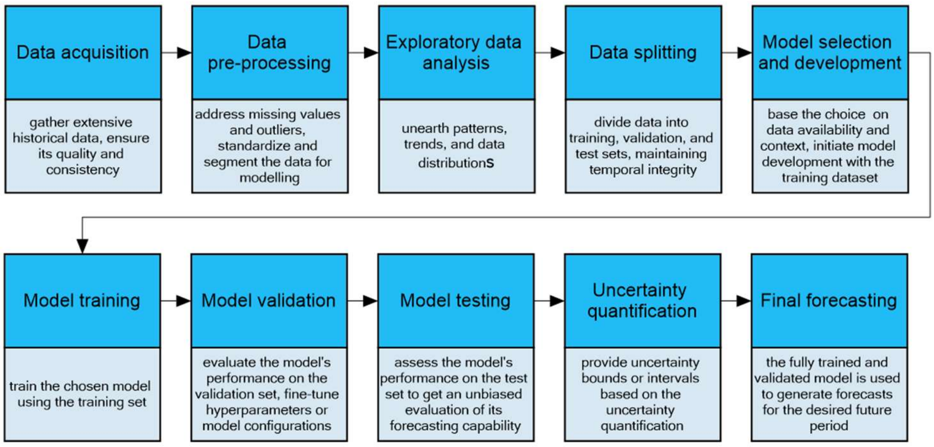

Figure 25 – Stages in the development of RE forecasting models.

Depending on the time horizon of the prediction, renewable generation forecast may be categorized as i) very short-term (a few seconds to a few minutes), ii) short-term (a few minutes to a few hours), iii) medium-term (a few hours to a few days), or iv) long-term (a few days to a few weeks, months, or years). Also, the type of information incorporated into the model dictates the nature of the forecast tool [84].

Several approaches are possible, typically categorized into white-box, black-box and grey-box models. The white-box models (also known as physical models) consider a detailed description of the generation unit (e.g., layout of a wind farm and wind turbine model; orientation and tilt of the panels in a solar farm) and of the surrounding terrain (e.g., orography, obstacles, roughness) – as shown in Figure 26. This information is used to downscale the data generated by Numerical Weather Predictions (NWP) into a site-level weather forecast, to which a weather-to-power model is applied to generate the forecasts. A more detailed description of physical models can be found in deliverable D3.2.

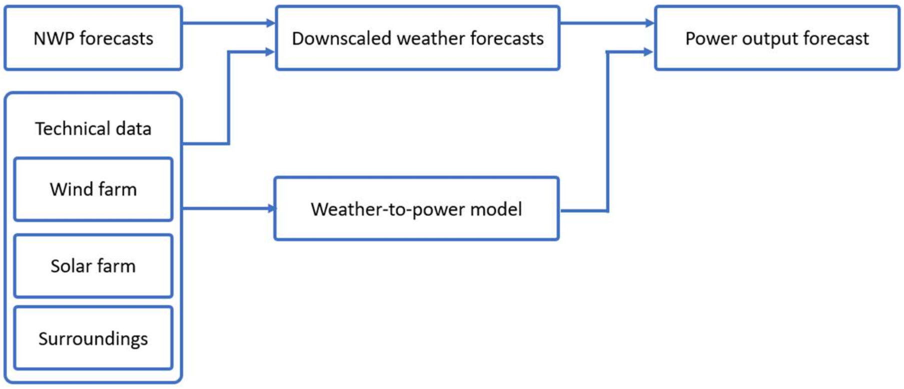

Figure 26 - Generic workflow of a white-box RE forecasting model.

Alternatively, black-box models (also known as statistical models) delve into the operational data of the generation unit to derive a statistical relationship between power output and other predictor variables, as shown in Figure 27. Black-box models inherently capture the specificities of the renewable generation unit (e.g., orographic effects across a wind farm, shading in a solar plant) without requiring an extensive technical characterization, which is an advantage over physical models. However, they need historical data for training.

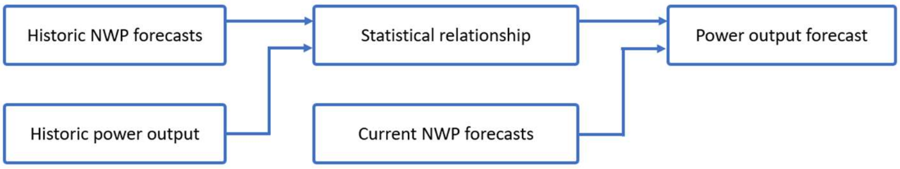

Figure 27 - Generic workflow of a black-box RE forecasting model.

Finally, grey-box models leverage on the benefits of both white- and black-box models, combining the physical grounding and interpretability of the first with the flexibility and practicality of the latter. Further information on the characteristics of the different types of models, their suitability and the reasons for selecting specific models for the forecast of PV and wind generation are presented in ANNEX .

Solar power forecasting

This section describes the forecast model for solar PV electricity production. The proposed model is a black-box model, which assumes availability of data, as described below. The future data-sharing infrastructure built according to the methodologies described in deliverables D4.2 and D4.3 could provide this data.

When there's an abundance of both weather and PV production data, black-box models can provide good quality outputs and will probably be the modelling resource most used during the development of the tool. The adequacy of the data, for PV purposes, shall account with model intricacy requirements, data variability and inherent patterns, trends, and seasonality - which can be remarkable in the PV case. Generally, simpler models, like linear regression, require a few years of hourly data (around 17,500 points). In contrast, sophisticated models like deep neural networks need more extensive datasets, preferably spanning several years with hourly granularity (ideally, 35,000+ data points).

Traditional time series models may be employed such as ARIMA (that comprehends a blend of data's trend, seasonality, and noise), and Exponential Smoothing (that encompasses Holt's linear and Holt-Winters' methods, which discern data trends and seasonality). Tree-based models (Random Forest, XGBoost, LightGBM) can also be employed, as these can detect non-linear correlations between weather data and PV output. Lagged PV production and weather data can be employed as features for future PV production prediction. Deep learning paradigms are particularly useful, and examples of methods are Long Short-Term Memory (LSTM) - a recurrent neural network variant, LSTMs excel at long-term data dependency capture; Gated Recurrent Unit (GRU) - another RNN type, GRUs are simpler and can sometimes outpace LSTMs in efficiency, maintaining comparable performance; One-Dimensional Convolutional Neural Networks (1D CNNs) - for local pattern or seasonality detection in time series forecasting; and, Hybrid Models such as Fuse Convolutional Neural Networks (CNNs) with LSTMs/GRUs. The premise is to first identify patterns and then model sequences. Given the data reservoir and a forecasting horizon of 24-48 hours with hourly timesteps, LSTMs are the most suitable models for PV supply forecasting. Their architecture and performance make them especially suitable for time series forecasting, given their capacity to understand relationships between weather variables and PV production intricacies. Details on forecasting methods may be found in [82]. In summary the following process will be implemented to develop the forecast model for electricity production from PV systems:

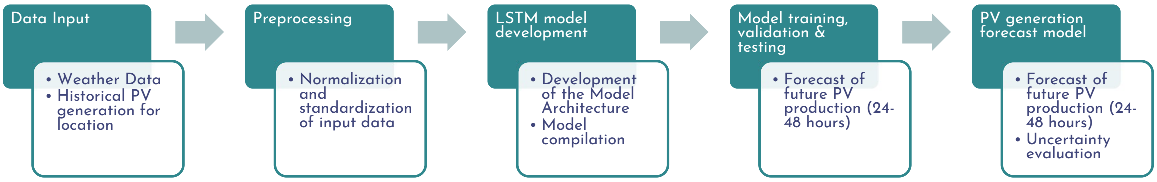

Figure 28 – Stages of development of the PV generation forecast model.

2.6..1. Data input

Several predictors can be influential in the forecast of PV electricity production: Solar Irradiance (Directly influences the amount of energy the PV system can generate), Temperature (PV system efficiency is temperature-dependent), Cloud Cover (impacts solar irradiance and hence PV production), Humidity (can affect the efficiency and lifespan of PV systems), Wind Speed (might influence the cooling of solar panels), Historical PV Production Data (Past PV production data is essential for time series forecasting as it helps the model identify patterns, trends, and seasonality) [85]. For the sake of model complexity and computational cost, it is foreseen that the implemented model will use historical PV production data, selected weather data such as irradiation and ambient temperature. When incorporating lagged variables in the context of LSTM, these become new predictors in your input data, essentially making the LSTM model a type of Nonlinear AutoRegressive model with eXogenous inputs (NARX). The model will use Auto-Regressive Inputs (include lagged values of the PV production as predictors in your input data), Exogenous Inputs (irradiance, temperature) and Input sequencing (each timestep in a sequence would now include the selected lagged values and values of the exogenous variables for that timestep). Lagging the exogenous variables can help capture delayed effects they might have on the target variable.

The selected lagged values of PV production consider the recent behaviour of PV production that can give insights into its short-term future, especially to capture any anomalies or unexpected behaviours. It is recommended to use the last 24 hours. It's beneficial to have a full day's context of actual PV production, as this helps in identifying daily patterns and any deviations from them.

The selected lags for solar irradiance consider the daily pattern of solar irradiance to be essential for predicting PV production. However, the pattern is largely repetitive across days, so we don't necessarily need data from the entire previous day and care should be taken for the night period. The recommended lag is the last 6 hours (this captures the immediate past and is especially relevant for short-term forecasts). The temperature changes have a slower dynamic than solar irradiance and can be essential in determining PV efficiency, and for this reason, it is recommended to consider the last 12 hours (this gives a half-day context, capturing any significant temperature shifts). During the future development, we should highlight that these values may be altered concerning the possible computational cost the model will represent.

2.6..2. Data processing

The pre-processing should consider Normalization/Standardization, as neural networks, including LSTMs, work best when input data is scaled to a standard range, usually between

0 to 1 (normalization) or to have a mean of 0 and a standard deviation of 1 (standardization). Also, in this stage the sequencing should be arranged.

2.6..3. LSTM Model architecture

The input layer will take in sequences of data. The LSTM layers (one or more can be added and number of neurons can vary based on complexity). The LSTM unit consists of three main components (these are inherently part of each LSTM layer and are responsible for regulating the flow of information): Input Gate$(i(t))$ Forget Gate $(f(t))$, Output Gate $(o(t))$. The Input gate controls what information from the current input and the previous hidden state should be added to the current cell state - it computes a candidate cell state based on the current input and the previous hidden state ($C(t)$ is a kind of "memory" for the LSTM, it can store long-term information). The Forget Gate controls what information from the previous cell state should be retained or discarded. It determines which parts of the previous cell state are no longer relevant for the current time step. The Output Gate controls what information from the current cell state should be used to compute the current hidden state and the output. It regulates the flow of information from the cell state to the hidden state and output. Details can be found in [86], [87]. The general equations for the architecture of the LSTM are as follows:

$$f(t) = \sigma\big(W[f] * \big(h(t-1), x(t)\big) + b[f]\big) \tag{68}$$

$$i(t) = \sigma\big(W[i] * \big(h(t-1), x(t)\big) + b[i]\big) \tag{69}$$

$$\tilde{C}(t) = tanh\big(W[C] * \big(h(t-1), x(t)\big) + b[C]\big) \tag{70}$$

$$C(t) = f(t) * C(t-1) + i(t) * \tilde{C}(t) \tag{71}$$

$$o(t) = \sigma\big(W[o] * \big(h(t-1), x(t)\big) + b[o]\big) \tag{72}$$

$$h(t) = o(t) * tanh\big(C(t)\big) \tag{73}$$

Where:

$x(t)$ - This is the input to the LSTM cell at time tt. In our context, it would be a vector containing the current and lagged values of weather variables and PV production.

$h(t-1)$ - This is the hidden state from the previous time step. It carries information from earlier inputs.

$W[f]$,$W[i]$,$W[C]$,$W[o]$ - These are weight matrices for the forget, input, cell, and output gates, respectively. They determine how much importance to give to each component in the input and the previous hidden state.

$b[f]$, $b[i]$, $b[C]$, $b[o]$ - These are the bias terms for the forget, input, cell, and output gates, respectively. They shift the output of each gate.

$o(t)$ - The output gate's output. It decides what parts of the cell state will be in the hidden state h(t)h(t).

$h(t)$ - The hidden state at time tt. It's a function of the cell state, as modulated by the output gate. This will be used as input for the next time step and can also be used as the LSTM's output.

After the implementation of the LSTM layers, one or more dense layers can be added to refine the prediction. This is a standard feedforward neural network layer. Optionally, dropout layers can be added between dense layers to prevent overfitting. It is followed by the output layer, a single neuron that will output the forecasted PV production value for the next timestep.

Once the architecture is defined, the model needs to be compiled. This involves specifying: the 1) Loss Function (common choices for regression tasks are Mean Squared Error (MSE) or Mean Absolute Error (MAE)); the 2) Optimizer (algorithms that minimize (or maximize) the loss function, e.g., Adam, RMSprop, or SGD); and, the 3) Evaluation Metrics (while the optimizer works on the loss function, we often want to track other metrics, like MAE or RMSE).

2.6..4. Model training, validation, and testing

This stage consists in feeding the training data into the model. This involves presenting the model with sequences of predictors and their corresponding target PV production values, and backpropagation through time. In essence, the model computes the gradient of the loss with respect to its weights for every timestep in the sequence, and then it updates the weights. It should also have an early stopping component, to monitor the model's performance on the validation set. If the performance starts to degrade (indicating potential overfitting to the training data), the training process should be interrupted. Once an epoch of training is complete (an epoch is one full pass through the training dataset), the performance of the model should be validated on the validation set (epoch-wise validation), using metrics such as MAE RMSE and MAPE.

After training is complete, i.e., delivering a satisfactory model performance on the validation data, the performance of the model should be evaluated with the test data. This gives an unbiased assessment of how well the model is likely to perform on completely unseen data. Other factors should be carefully tracked and checked during this stage, such as the batch size, learning rate, statefulness, regularization and gradient clipping.

2.6..5. Forecasting & uncertainty evaluation

Once trained, the model can be used to forecast future PV production. For forecasting multiple steps into the future, a rolling-forecast origin can be used, where the model is repeatedly used to predict the next time step, then actual data is added back into the predictors, and the process is repeated.

Lastly, to estimate the uncertainty of the forecasts, a Monte Carlo Dropout, Bayesian LSTMs, quantile Regression or Ensemble of LSTMs could be used. Given the complexities of LSTMs and the NARX structure, starting with Monte Carlo Dropout might be the most pragmatic approach, as it introduces minimal changes to the standard LSTM training and inference processes. Once preliminary results are handled, more sophisticated methods like Bayesian LSTMs or ensemble approaches should be explored [88].

Wind power forecasting

The selection of the modelling approach for the wind power forecasting tool considers several factors such as data availability, accuracy, integration, and forecast horizon. The RE forecasting tools will be integrated in the digital infrastructure of the port, where generation data from different sources (including wind) is expected to be widely available. As such, it will be possible to take full advantage of the benefits of statistical modelling without significantly adding to the data requirements of the digital infrastructure. Moreover, the RE forecasts will be fed into the Energy Matching Tool for an optimum generation/demand balance that minimizes greenhouse gas emissions. For an anticipated forecast horizon of 24 to 48 hours ahead (short- to medium-term forecast), it is necessary to incorporate NWP data into the wind power forecasting model to maintain accuracy. For the same reason, a non-parametric approach is more suitable to modelling forecast uncertainty. Considering these constraints, a statistical model based on Gradient Boosting Trees is proposed for the wind power forecasting tool.

Gradient boosting [89] is a machine learning technique suited for both regression and classification applications. This technique consists of an ensemble of base learners typically in the form of weak prediction models, i.e., simple models that describe the data in a very straightforward way. It performs numerical optimization through gradient-descent minimization by recurrently training the base learners on the residuals of previous model iterations. This ensemble technique thus aims at building a single strong learner through the combination of multiple weak learners. The goal is to minimize a loss function (L) that quantifies the difference between the observed values (y) and the model predictions ($F(x)$). The loss function may take on a variety of formulations depending on the application (discussed further along in this section).

Gradient Boosting Trees (GBT) [90] are a particular formulation of gradient boosting models that employ regression trees as base learners and have been extensively used in literature and practical applications of wind power forecasting tools [91]. One of the advantages of GBT is its non-parametric nature. In addition, the model is suitable for industrial integration due to its scalability for a high number of explanatory variables, allowing the wind power forecasting tool to take full advantage of the extensive amount of historic and operation data expected to be available in the context of the port Digital Twin. Moreover, the selection of an adequate loss function allows for great flexibility in tackling forecast uncertainty.

The modelling algorithm begins with defining a first iteration of the model (F_0) which may be as simple as the average of the target variable:

$$F_0(x) = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{74}$$

The additive training process builds on this initial model formulation and iterates over the following steps until convergence:

1. Considering the loss function (L), the negative gradient ($-g(x_i)$), is computed for each observation in the training sample. The negative gradient points in the direction of the steepest decrease in the loss function:

$$-g(x_i) = -\frac{\partial L[y_i, F(x_i)]}{\partial F(x_i)} = y_i - F(x_i) \tag{75}$$

2. A regression tree with a set of hyperparameters a ($h(x; a)$) is fitted using the negative gradients as the target variable. This stage aims to find a simple model to further minimize the residuals of the previous iteration:

$$-g(x_i) = h(x_i; a) \tag{76}$$

3. The model is updated at each iteration ($F_m(x)$) by adding to the previous model, ($F_{m-1}(x)$) the prediction of the current regression tree scaled by the learning rate (τ):

$$F_m(x) = F_{m-1}(x) + \tau * h(x; a) \tag{77}$$

As with most machine learning algorithms, GBT include two distinct types of model parameters: learnable (or simply parameters) and non-learnable (or hyperparameters). Whereas learnable parameters are calculated during training on a given dataset, non-learnable parameters cannot be inferred from the training data. Instead, they must be defined beforehand, prior to the training stage. Bayesian Optimization improves the search speed using information on past performances to determine the next point to assess in the search space, for which reason it is widely used in the optimization of hyperparameters of machine learning models. Considering a training dataset with a yearly timespan (discussed further along in this section), a 12-fold cross validation is employed to produce monthly validation scenarios as in [90]. The prediction error is estimated for each monthly fold and the 12 monthly values are averaged to obtain the final evaluation metric of the hyperparameter optimization. Table 14 shows the proposed range for the different hyperparameters in the optimization process. It should be noted that the values and bounds for the hyperparameters indicated at this point are merely presented as reference for the GBT model and may be altered in a further stage of implementation.

Table 14 - Bounds for hyperparameters of the Gradient Boosting Tree

Hyperparameters	Comments		Range
Regression tree	Maximum depth	High values increase risk of overfitting Low values limit model accuracy	5 to 9
	Minimum number of samples to split an internal node	High values limit model accuracy Low values increase model complexity and risk of overfitting	150 to 350
	Minimum number of samples required to be at a leaf node	High values limit model accuracy Low values increase model complexity and risk of overfitting	20 to 80
	Maximum number of features	High values decrease bias Low values decrease variance	Square root of total number of features
Boosting process	Learning rate	High values increase risk of overfitting Low values increase computation time	0.01 to 0.05
	Number of boosting iterations	High values increase risk of overfitting Low values limit model accuracy	500 to 800
	Fraction of samples to fit the individual base learners	High values decrease bias Low values decrease variance	0.8

An initial iteration of the forecasting tool is proposed using only information directly available from the NWP data, namely wind speed and wind direction at multiple heights.

The latter is included in the base model due to its important contribution to explaining the variability of wind power output, leading to a significant improvement of forecast performance without an excessive increase in model complexity and data requirements. The variability of the wind resource poses a great challenge for the forecasting tool, as the NWP models are generally unable to capture the relevant atmospheric phenomena across all spatial and temporal scales. If necessary, additional features may be extracted from the raw NWP data as to increase the amount of information input to the forecast model. This increases the precision and robustness of the forecast model, while also decreasing the error and the uncertainty of point and probabilistic forecasts, respectively. These additional features are to be investigated upon at a further stage of implementation [90].

For locations with limited availability of historic wind power data, the applicability of the proposed statistical approach is restricted. In this situation, a physical model may be temporarily used to perform the forecast. Alternatively, the physical model may be used to generate the training data for the statistical model, in which case a synthetic time-series of historic wind power output is computed based on past records of NWP and then fed into the statistical model as a substitute for real operational data. In both scenarios, the statistical model may be trained sufficient training data is collected, if it meets the quality criteria described in the final section.

Having collected all the necessary training data, the model hyperparameters are tuned to ensure optimal model configuration. Afterwards, the training of the model is conducted, and a 12-fold cross validation scheme is employed. In this case, the yearly training dataset is divided into 12 monthly folds and a separate model is trained considering each possible combination of 11 months of training data – the monthly fold left out of the training subset is used for validation and the monthly results are averaged into the final evaluation metrics.

Once a final model is obtained, it may then be used to generate both point and probabilistic forecasts, considering an expected forecast horizon of 24 to 48 hours. For the generation of point forecasts, the absolute and square loss functions may be used, respectively:

$$L[y, F(x)] = \frac{1}{N}\sum_{i=1}^{N}|y_i - F(x_i)| \tag{78}$$

$$L[y, F(x)] = \frac{1}{N}\sum_{i=1}^{N}[y_i - F(x_i)]^2 \tag{79}$$

Where N is the total number of training samples i. Contrarily, the quantile loss function is employed to estimate the probability density of the forecast:

$$L[y_i, F(x_i)] = \begin{cases} \alpha * [y_i - F(x_i)] & , F(x_i) \leq y_i \\ (1-\alpha) * [y_i - F(x_i)] & , F(x_i) > y_i \end{cases} \tag{80}$$

The quantile loss function attributes different weights to the prediction errors depending on the quantile (α) to be estimated, particularly penalizing under-estimations for large quantiles and vice-versa. The probabilistic forecast is represented by a set of individual forecasts run for different quantiles (e.g., from 0.05 to 0.95 with a 0.05 increment).

If it becomes necessary to attest to the model accuracy, an additional dataset on historic power output is required. A physical model may also be used to generate additional synthetic wind power data for model validation. A sliding time window approach is suggested for validation, as in [90]:

- Define the training window with a yearly timespan (e.g., months 1 to 12).
- Define the test window with a 3-month timespan (e.g., months 13 to 15).
- Slide the training and test windows 3 months forward and repeat the procedure until the end of the dataset is reached.

The monthly error metrics for each individual validation subset are then averaged into the final performance indicators of the forecast model. This approach has the advantage of assessing model performance for different operating conditions and seasonal patterns.

Final remarks

For both the PV and wind power forecasting tools, a training dataset of historic power output and NWP data is necessary. This dataset must have a minimum timespan of 1 year to capture seasonality of the renewable resources with sufficient accuracy. However, ideally the training dataset would be longer (e.g., at least 2 years) to allow for the modelling of inter-annual variability that is typical of renewable resources. Longer training datasets have additional advantages such as increased model stability, lower risk of overfitting, and improved generalization. Nonetheless, incomplete yearly datasets (e.g., 1.5 years) may introduce bias in the model and skew the forecasts towards specific seasonal patterns. For this reason, complete yearly datasets (e.g., 1 year, 2 years) shall be preferred for a balanced representation of seasonality.

Prior to training the model, filtering and quality control of the input data – namely historic power output – are necessary to ensure representativeness of normal operating conditions (e.g., eliminating maintenance periods or erroneous power output measurements). Information on any known cause of unavailability of the generation unit is necessary to correct the input data as to reflect normal operation. After this stage, the training data must maintain a high recovery rate (ideally above 90-95%) and the missing values shall be dispersed across the dataset (as to avoid any punctual behaviour being overlooked by the model).

Moreover, near real-time access to NWP data is essential for a regular run of the forecasting model in an operational scenario. The NWP data may be obtained through cooperation with meteorological research institutions and other entities that regularly run the simulations for different parts of the world. Alternatively, weather models may be run locally to generate the necessary time-series of NWP data. In this context, mesoscale models such as the Weather Research and Forecasting (WRF) model [92] are suitable candidates. This model is open-source and its use is unrestricted, however it requires an adequate setup and parameterization and consumes substantial computational resources for an accurate simulation.

Considering the requirements and time resolution of the demand forecast and battery management models and the Energy Matching Tool, different schedules are possible to run the renewable generation forecasts. For example, the model may be run daily for the following 24 to 48 hours. Alternatively, the forecasts may be generated on an hourly basis to maintain a constant time horizon.

Point forecast quality is typically evaluated using conventional error metrics such as the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) [93]. The MAE assesses the magnitude to which the model predictions deviate from the observed values and is defined as follows:

$$MAE = \frac{1}{N}\sum_{k=1}^{N} |y_i - \widehat{y_i}| \tag{81}$$

Where y_i and $\hat{y_i}$ are the observed and forecasted values of power output, respectively, and N is the total number of samples i. Similarly, the RMSE is a measure of the deviation between modelled and observed data, with the added advantage of weighing large deviations more strongly due to its quadratic factor – this is especially relevant in power forecast applications since large forecast errors have a strong negative impact on grid operation and management. It is defined as follows:

$$RMSE = \frac{1}{N}\sqrt{\sum_{k=1}^{N}(y_i - \hat{y_i})^2} \quad (82)$$

Assessing the quality of a probabilistic forecast is more complex. The main required properties of probabilistic forecasting are reliability (i.e., the probability density of the forecasts shall match that of the observations) and sharpness (i.e., the predictive interval shall be as narrow as possible as this conveys more detailed information and facilitates decision-making under a lower level of risk/uncertainty). However, the overall quality of the probabilistic forecast is not strictly determined by any of these properties, rather an additional error metric – the skill score – is necessary to combine the assessment of both reliability and sharpness into a single indicator. The Continuous Ranked Probability Score (CRPS) [94] is a commonly used skill score in power forecasting and compares the cumulative distribution functions of the forecasts ($F(x)$), and the observations ($O(x)$). Compared with other skill scores, the CRPS is more robust in the presence of outliers and extreme events [78].

$$CRPS = \int_{-\infty}^{+\infty}[F(x) - O(x)]^2 dx \quad (83)$$

The error metrics described above are usually normalized to compare the accuracy of different models and provide insight into the relative significance of the forecast errors. In this context, the mean power output or rated power of the generation facility may be used as reference for the normalization of the error metrics. Alternatively, if the model works with the capacity factor rather than the power output, the results may be normalized by the mean capacity factor.

2.7. Terminal Grid Modelling

For the Energy Matching Tool (EMT) to proceed with its optimal energy match to reduce greenhouse gas emissions, not only the production, storage, and demand components, but also the grid has to be modelled. As presented in Figure 2, the majority models focus on characterizing these components' technical constraints and forecasting their operation (e.g., production and demand profiles, as well as flexibility characterization), the last being subject to the grid topology in place. Grid topology is important because its restrictions, e.g., respecting components' capacity or voltage regulation, have an impact on the EMT's energy match decisions. It is not only important to have the optimal energy matching but also to guarantee grid operational constraints' feasibility.

From discussions with MAGPIE's port partners, in the real-world the ports' grid topology setup varies. There can be players within the port either directly connected to the public grid through a main supply line and transformer (if required), or a dedicated port grid, which is connected to the public grid through a substation, that then connects to dedicated transformers at each consumer site, such as a specific terminal or industry. Both options can also coexist. In the ports, it is common to have a significant number of underground utility lines, and that there are low voltage (230V/400V), and one or multiple medium voltage (MV) grids. For instance, in Port of Sines, there are 15kV and 30kV MV grids, while in Deltaport, they can reach up to 40kV. In each consumer's slot, energy is distributed through their distribution box to it's the relevant facilities, with the topology being subject to the loads' energy requirements, e.g., in a container terminal, cranes and building loads require different distribution characteristics.

Given the diversity in port grid configurations and stakeholders, a decision was made to streamline the analysis. Rather than performing an all-encompassing characterization of, for example, the Port of Rotterdam, considering the vast scale of the electrical grid, it was decided to reduce the scope of the analysis and, instead, focus on developing the concept of the EMT Terminal. An EMT terminal acts as a simplified port ecosystem that aims to consider key ports' players, and considers a unique connection to the main grid, as presented in Figure 29. It should be highlighted that the modelling itself is also prepared to receive a wider and more complex grid, i.e., while in MAGPIE the EMT will be applied to a terminal, it will be developed in a way to enable it to be applied to an entire port.

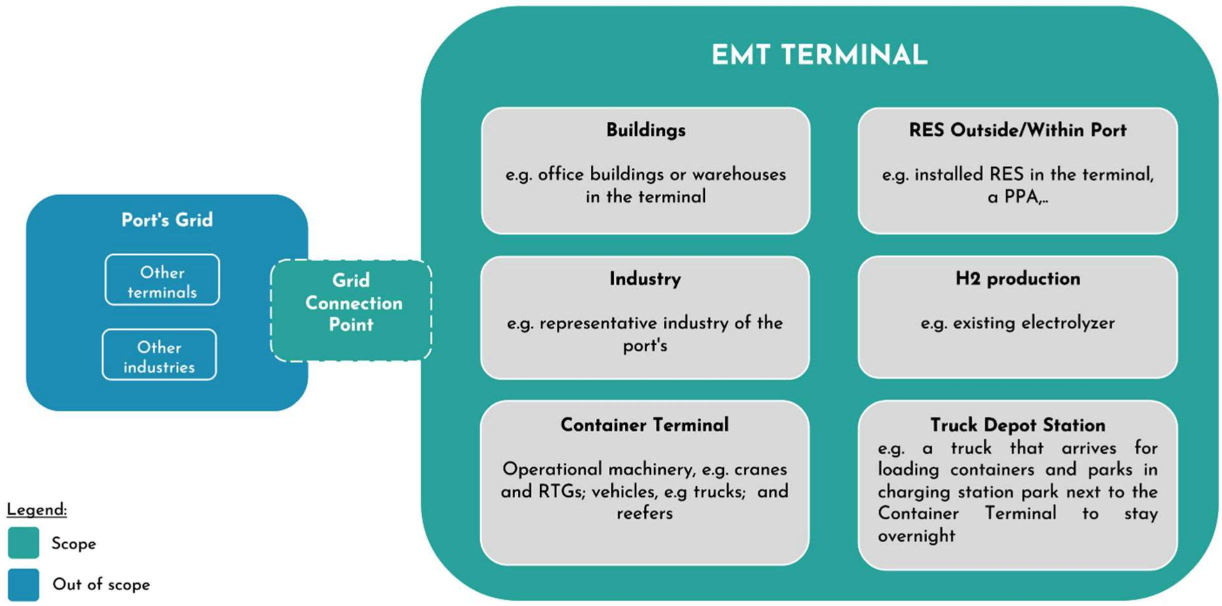

Figure 29 Representation of the EMT Terminal

Taking advantage of the Container Terminal simulator (described in section [CEA Simulator section]) and of having a representative partner within the MAGPIE consortium, a more detailed characterization is expected for the EMT Terminal, similar to the scheme in Figure 29. The core loads of container terminals can be divided into four main elements: buildings (office buildings and warehouses), terminal lighting, operational assets operation (cranes and RTGs), and reefers refrigeration. As exemplified in Figure 29, the loads require different grid requirements in terms of voltage and capacities. Moreover, according to the loads, different architectures can be in place, such as radial or loop grid. For instance, a loop grid configuration can ensure reliability, serving both cranes and reefers, allowing for non-stop crane operations and continuous refrigeration [95].

Looking ahead, with the growing electrification of transportation, particularly in the waterborne transports, onshore power supply – OPS - (also knowns as cold ironing), will necessitate dedicated modifications in the terminal's infrastructure, e.g., larger power supply to terminals and dedicated lines and transformers to OPS. Different type of grid architectures exists to include OPS, such as low voltage supply connection (LVSC) and high voltage supply connection (HVSC) [96]. In line with this future-oriented perspective, terminals must also plan for the integration of recharging stations to support horizontal transport vehicles like terminal trucks. Furthermore, as terminals expand, it's essential to consider the increasing consumption of already electrified loads, like cranes and warehouses climatization, which will significantly impact the terminal's energy and grid requirements.

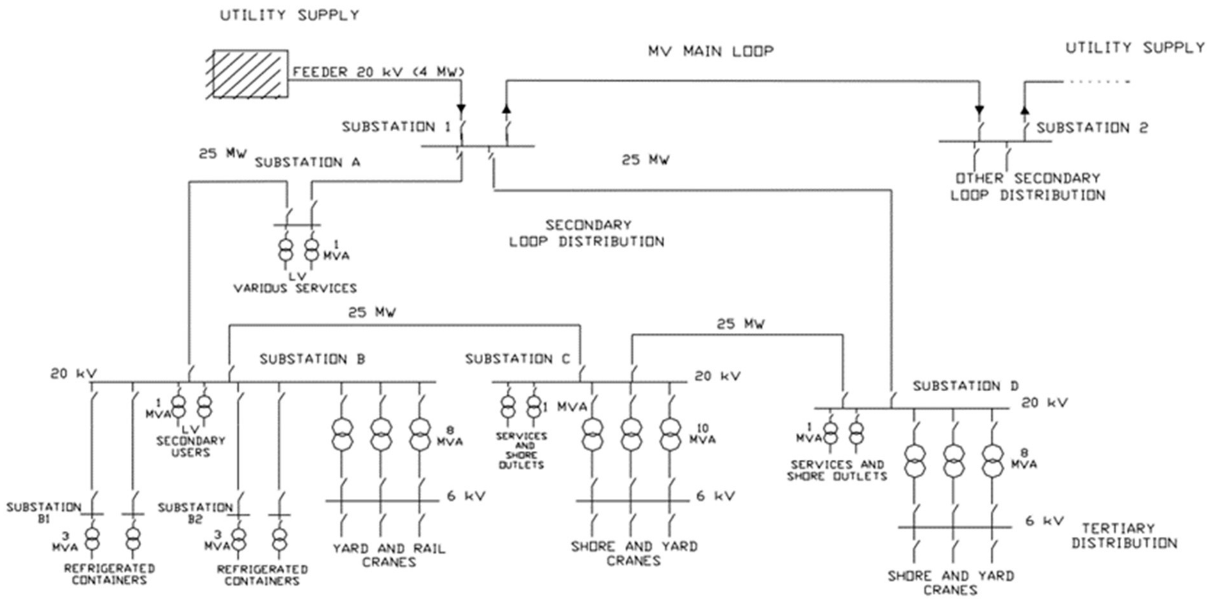

Figure 30 A sample of a global distribution in a port: the main loop at 20 kV, the secondary loop distribution at 20 kV for a terminal container and the tertiary distribution at 6/0.4 kV (from [97])

To model the EMT terminal's grid, PyPSA [98], an open-source python software toolbox for simulating and optimizing modern power and energy systems will be used within the EMT. This toolbox allows not only to analyse exclusively the electrical grid but also to couple other energy vectors, e.g., hydrogen.

Table 15 - Considered PyPSA components to model the grid (description according to PyPSA documentation)

Component	Description
Bus	Electrically fundamental node where x-port objects attach.
Carrier	Energy carrier, such as AC, DC, heat, wind, PV, or coal.
Line	Lines include distribution and transmission lines, overhead lines, and cables.
Transformer	2-winding transformer
Link	Link between two buses with controllable active power - can be used for a transport power flow model OR as a simplified version of point-to-point DC

Load	Active and Reactive power consumer
Generator	Power generator
Storage Unit	Storage unit with fixed nominal-energy-to-nominal-power ratio.
Store	Generic store, whose capacity may be optimised.

For the EMT terminal, the components that are going to be considered to characterize the grid are: buses, lines, transformers, and the external grid connection (as a slack generator). Technical characteristics of these components, e.g., lines apparent power, should be provided by the user as described in PyPSA (PyPSA also provides standard characteristics for different types of components which can be applied). If technical details are not provided, the EMT can still perform the energy matching without considering the grid restrictions. On top of this terminal grid model, components, such as generators, storages, non-flexible and flexible loads, will be created and associated to the respective bus. This is embedded in the EMT, which will consider the operational and technical characteristics of the mentioned components according to the models represented in Figure 2.

The terminal's grid can be modelled in a simplified or complex manner depending on the available data and required detail. The approaches foreseen are: (i) a single bus that directly connects to the main grid and all consumers; (ii) considering multiple consumers for each bus with their respective lines and transformers characteristics; and (iii) also considering the distribution topology at the consumer's level for the container terminal.

With the objective of demonstrating the EMT in a complex grid topology, to prove its concept and implementation, an IEEE network case is going to be considered (still being assessed the exact IEEE network to use) in the EMT terminal. Some of the nodes will be replaced by the EMT Terminal players, e.g., container terminal, office buildings, industry, among others.

3. Post-processing of the terminal simulation data

This chapter describes practical applications of some of the models presented in Chapter 2, including how the PortoPort simulator data may be used to model the energy loads of the terminal assets and in the calculation of CO_2 emissions of port vehicles.

3.1. Energy demand profiles for cargo handling equipment and other terminal systems supporting the logistics operations

Different types of input data are required to implement the models described in section 2.2. This section maps the data needs for the demand modelling of some of the port assets and showcases a practical implementation of the models using the outputs of the ProtoPort simulator. Figure 31 describes the relationship between the required inputs and the some of the possible approaches (model-driven or data-driven) and the feasibility of obtaining such data – green indicates the data available from the ProtoPort simulator, yellow indicates the data that should be available from equipment manufacturers, literature and historical data, and in red historical detailed data for data-driven models, which would have to be collected (or have been collected) through direct measurements at the level of the equipment/vehicle/system or available through the terminal logistics platforms.

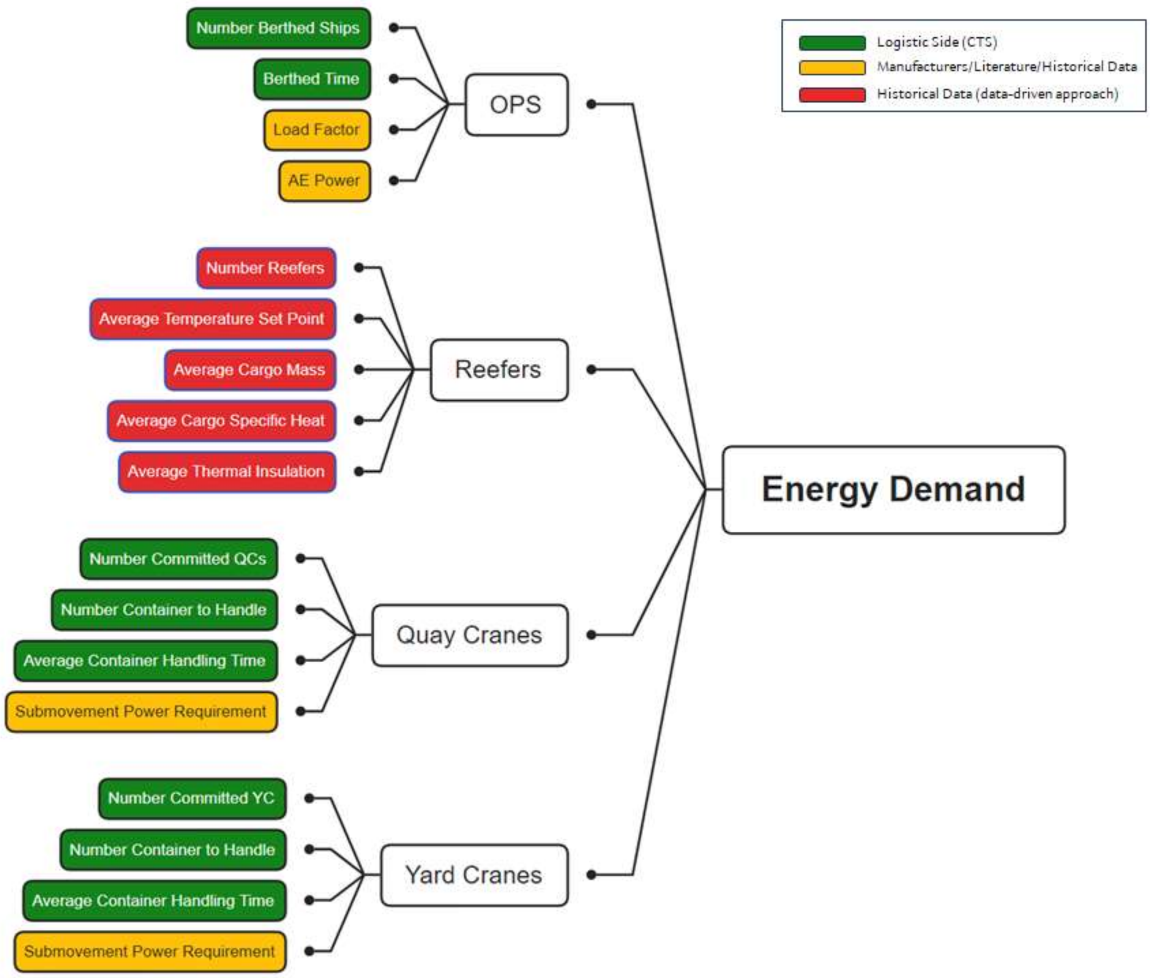

Figure 31 - Data inputs required for the energy demand models of terminal assets and systems.

Energy demand of the OPS systems

In the case of the OPS systems, the first group of inputs should come from the port authority or terminal operator: predicted number of berthed ships and their berthing schedule (ETA and ETD); also, historical data can be used based on the AIS system to obtain average berthing times per call. Once the ships are known, its characteristics (type of vessel, DWT, main/auxiliary engines power) are used for determining the load factors and the auxiliary engines power, if this information is not readily available in some specialized database. In that case, the information can be provided for the port authority and manufacturer, or values from the literature can be used.

For the energy demand estimation, based on the Container Terminal Simulator (CTS) (i.e., the ProtoPort simulator), the berthing time can be obtained. Table 16shows the information for a specific vessel (NA_3) obtained from the ProtoPort output.

Table 16 - An extract from the Container Terminal Simulator Influx DB, concerning ships activity.

_time	ref	type	coord_x	coord_y	noeud
2022-01-02 09:45:45+00:00	NA_3	Navire	386	-145	N1
2022-01-03 02:29:27+00:00	NA_3	Navire	-9999	-9999	V

From here, we can observe that the specific ship was berthed 16h:43m:42s. Therefore, if the information about the characteristic of the ship identified as NA_3 is known (auxiliary engines power), by using Eq. (3) the energy consumption associated to the OPS system can be determined. As instance, suppose that the line loss rate is $\eta = 0.02$ and the auxiliary engine power is 12000 kW, then, considering a typical load factor for containerships equal to 0.6, the energy demand for the OPS can be obtained by:

$$E_{OPS} = \sum_{j \in J} P_j l_j t_i (1+\eta) = 12000\ \mathrm{kW}\ \times 0.6\ \times\ 16.72\mathrm{h}\ \times 1.02 = 122792\ \mathrm{kWh}$$

Energy demand of Reefers

In the case of the reefers, a data-driven approach seems to be a feasible option to estimate the energy demand. In this case average values, about the thermodynamical properties of the container (and its cargo) should be obtained, but also the number of reefers at the terminal's yard should be considered. The thermodynamical properties can be obtained through historical data (based on the Bill of Lading or other cargo description document); whereas the number of reefer plugged-in at the terminals is information that the terminal operator can provide. Also in this case, the Container Terminal Simulator allows us to keep track of the number of reefers at the terminal on each hour. Note that, knowing the initial number of reefers plugged-in at a specific instant of time (R_{t_0}), it is only necessary to identify which reefers where loaded (R^+_{t+1}) or unloaded (R^-_{t+1}) during the period:

$$R_{t+1} = R_{t_0} + R^+_{t+1} - R^-_{t+1} \tag{84}$$

Table 17. An extract from the CTS regarding reefers' movement on the terminal.

_time	emplacement	noeud	pile	position_pile	produit	categorie	mouvement
2022-01-02 10:01:04	EZZA30I125	I125	1	1	CNTR_1751	Reefer	1
2022-01-02 10:04:07	EZZA28G118	G118	3	3	CNTR_519	Reefer	1
2022-01-02 10:06:37	EZZA28G118	G118	3	4	CNTR_894	Reefer	1
2022-01-02 10:06:38	EZZA36I74	I74	3	1	CNTR_1999	Reefer	1
2022-01-02 10:07:54	EZZA2G443	G443	3	1	CNTR_1718	Reefer	1
2022-01-02 10:13:28	EZZA36I67	I67	3	1	CNTR_325	Reefer	1
2022-01-02 10:19:02	EZZA2G443	G443	3	2	CNTR_1314	Reefer	1
2022-01-02 10:24:37	EZZA36I75	I75	3	1	CNTR_985	Reefer	1
2022-01-02 10:26:14	EZZA28G117	G117	3	3	CNTR_632	Reefer	1
2022-01-02 10:28:09	EZZA43G149	G149	1	1	CNTR_786	Reefer	-1
2022-01-02 10:32:59	EZZA36I68	I68	3	1	CNTR_1961	Reefer	1
2022-01-02 10:38:33	EZZA36I76	I76	3	1	CNTR_1113	Reefer	1
2022-01-02 10:52:30	EZZA36I69	I69	3	1	CNTR_474	Reefer	1
2022-01-02 10:55:44	EZZA36I69	I69	3	2	CNTR_679	Reefer	1
2022-01-02 11:00:51	EZZA36I77	I77	3	1	CNTR_598	Reefer	1

Table 17 shows extract obtained from the ProtoPort simulator, with the movement of refeers at the terminal during a specific period. Suppose that the initial number of reefers plugged-in at 10:00am of 2022-01-02 it is known (R_{t_0}), then by analysing the outputs in Table 17 the number of reefers loaded ($R_{t+1}^{+} = 14$) and unloaded ($R_{t+1}^{-} = 1$) can be obtained. Therefore, assuming that the reefers are instantaneously connected after arriving to the yard, the number of container plugged-in at 11:00am of 2022-01-02 ($R_{t+1} = R_{t_0} + 13$) can be obtained by simply using equation (84).

Energy demand of cranes

The time dependent energy demand of cranes can be calculated considering the power requirements during each of the type of movement (as described in section 2.2..2), in order to obtain an accurate approximation as proposed in equation (8). This necessary data could in principle be obtained through the equipment manufacturers or based on historical data, if the power of each crane is monitored with enough temporal granularity. Additionally, the average handling time and power requirement of each crane can be used to feed the model. In this section, the model described in section 2.2..2 was fed with data from the container terminal simulator outputs, in addition to literature and manufacturer data. An example of the container terminal simulator data outputs relevant for the calculation are shown in Table 18.

Table 18. An extract from the CTS regarding the movement of a container by a crane.

_time	ref_executant	type_executant	ref_transport	type_mission	ref_produit	ref_mission	status

2022-01-04 22:06:16	Portique _NA_1	Portique	NA_7	DEPART URE	CNTR_1 413	MIS_CNT R_1831_N A_7_DEP _2	ALLOCAT ED
2022-01-04 22:07:00	Portique _NA_1	Portique	NA_7	DEPART URE	CNTR_1 413	MIS_CNT R_1831_N A_7_DEP _2	TRANSFER _START
2022-01-04 22:09:06	Portique _NA_1	Portique	NA_7	DEPART URE	CNTR_1 413	MIS_CNT R_1831_N A_7_DEP _2	FINISHED

The data outputs of the CTS (as per description of the simulator in chapter 2, section 2.1) shown in Table 18 including the commitment of the cranes for moving the containers, can be used to generate relevant energy related KPIs. To calculate these, the maximum estimated power consumption of the cranes can be assumed from the start of a transfer mission and until the mission is over. The reader can refer to the tables Table 1, Table 2 and Table 3; where the structure of the database is explained.

In Figure 32 shows some possible post processing data, using the number of active assets during each hour. From the outputs in Table 18, the analysis shows that the maximum number of ships simultaneously at berth was three (on two different days); and the maximum number of committed cranes was also three. However, most of the time only two or even one crane is needed. In the case of the number of reefers in the yards, the value is variable, oscillating between 476 and 527.

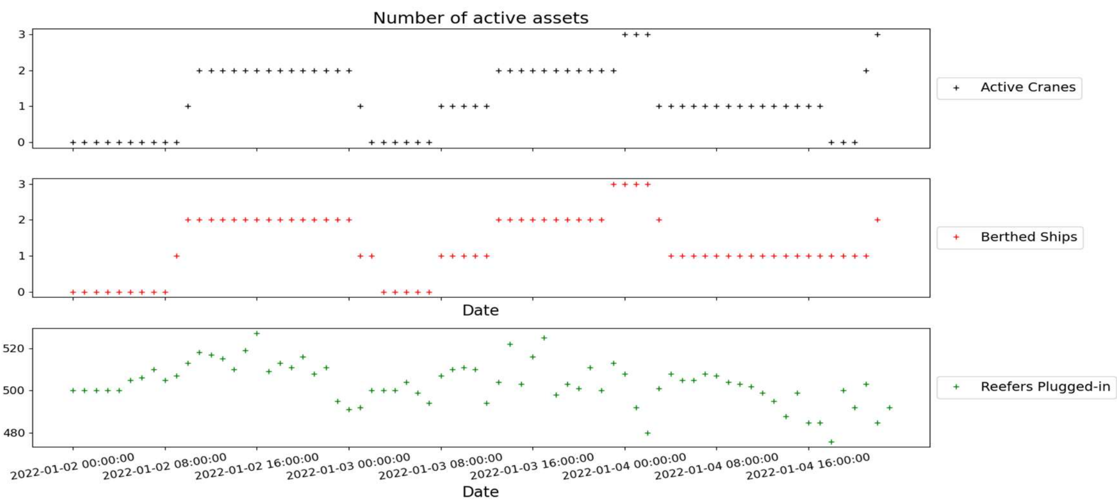

Figure 32. Committed number of assets.

Figure 33 shows some additional energy metrics, e.g., the hourly electricity demand estimation during a 4-day period. This result is based on the outputs from the CTS, but also considering some basic, but important, assumptions. Namely, the operations carried on at the terminal are analysed in an hourly basis and a fixed (maximum) power requirement is assigned to each operation. These power requirement values were obtained from the tables presented in section 2.2..2. These average values can be replaced by more accurate data from the specific terminal, to feed the mathematical models and enhance the accuracy of the results.

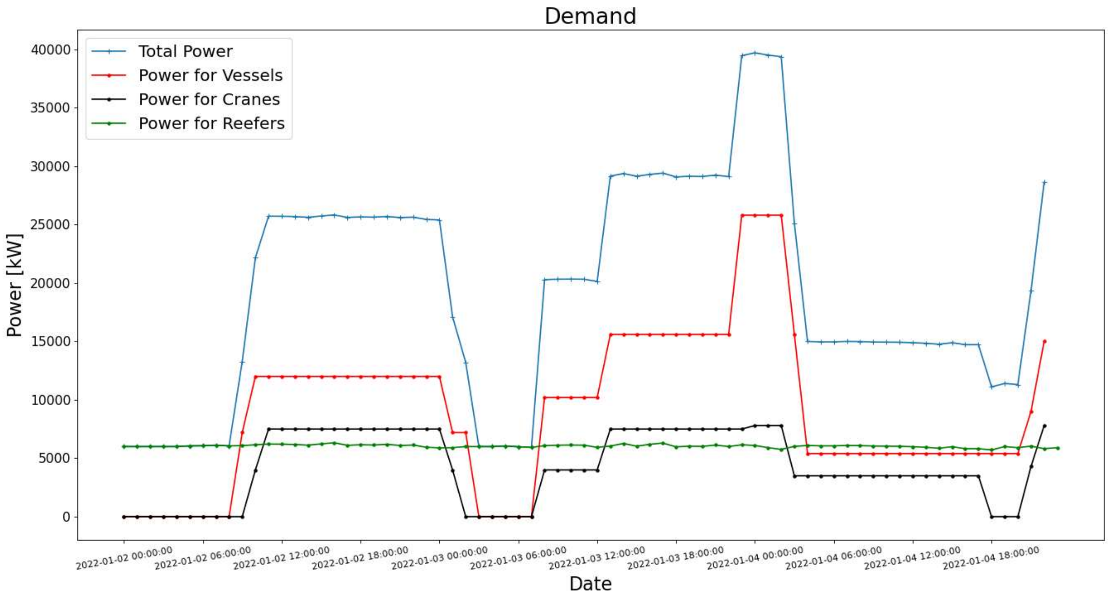

Figure 33. Hourly electricity demand based on the Container Terminal Simulator tool.

Table 19 shows the additional values used as input in the demand models of the selected assets to estimate the hourly electricity demand at the terminal over the 4-day period. The total power is obtained from considering all the cranes that will be handling cargo during each hour (and allowing overlapping of the lifting movement), the number of ships at berth and the number reefers connected to the grid, for each hour in the 4-day period. Different types of cranes, with different power requirements were considered, and the maximum lifting power required is shown in Table 19.

Table 19 – Additional input used in the calculation of the terminal energy demand from port assets.

Type of Asset	Asset ID	Power Requirement [kW]	Comment
Crane	Portique_NA_1	290	Different types of cranes considered. Maximum lifting power considered.
	Portique_NA_3	300	
	Portique_NA_5	4000	
	Portique_NA_6	3500	
Ships	NA_1	17000	This is the rated power for the auxiliary engines. A load factor 0.6 was considered for the OPS electricity demand.
	NA_2	8000	
	NA_3	12000	
	NA_4	9000	
	NA_5	15000	
	NA_6	10000	
	NA_7	17000	
Reefers		12	The same value for all reefers.

As the power needed for the cranes is related to the weight of the container being handled, knowing this information will increase the accuracy of the estimation. In the case of the reefers, the power is related to the temperature set-point and the type of commodity stored. Since this information is not currently available, a maximum consumption of 12 kW and an initial number of 500 reefers (R_{t_0}) connected to the terminal grid were considered. The results represent a worst-case scenario for the reefer demand, i.e., all the reefers at the terminals are running at the same time, and the import reefers are instantaneously connected to the grid, once unloaded from the incoming ship. All these initial assumptions can be revisited when real data is available. For instance, the distribution of the type of cargo being stored on each reefer would be particularly useful, as this would allow a tune-up the models to improve the quality of the estimation.

Additional analysis of energy related outputs are shown in Figure 34 and Figure 35.

The histogram in Figure 34 shows the hourly power demand frequency. Although it is not possible to identify a clear pattern, we must highlight that hourly power demand above 30000 kW, are less frequent.

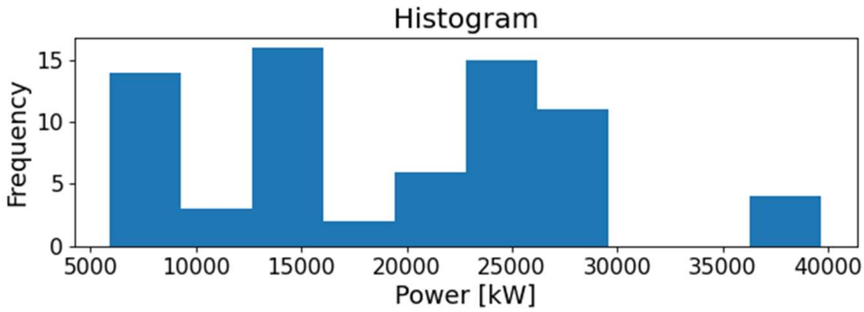

Figure 34. Hourly power demand histogram.

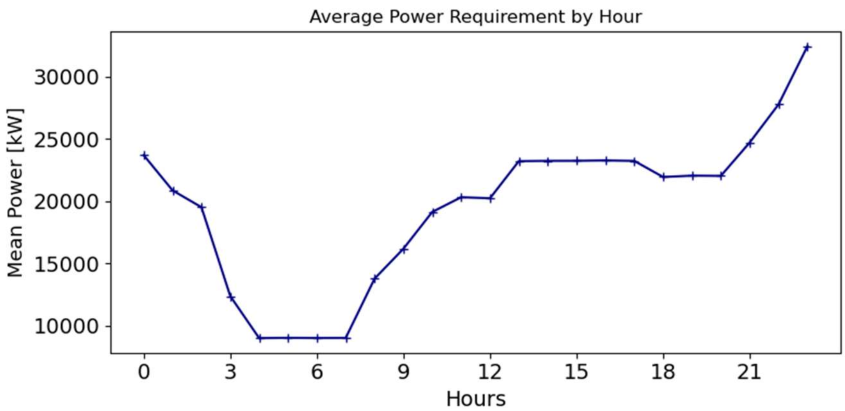

Figure 35. Mean power demand at different hours.

Finally, Figure 35 shows the average power demand at different hours, based on the data collected during the 4-days period. The outputs show that the lowest demand is observed between 04:00 and 07:00; whereas after 07:00 the demand continuously increases until midnight, when demand begins to decrease again.

3.2. Energy demand and CO2 emissions of port vehicles

To estimate CO2 emissions of port vehicles, namely trucks in this section using real data, ports should share information about the fleet composition, the speed, mass and distance travelled by each truck, as well as the total number of trucks acting within the considered area. in the terminal simulation data, we need their travelled distance, the instant speed, and the instant mass variations. However, it is unlikely that ports have access to such detailed information (especially at the scale of each single truck), and it was not possible to obtain such data within a time horizon compatible with the requirements of WP4.4. Therefore, the energy demand for trucks was calculated based on the Proto Port simulation data. From the three databases provided by the terminal simulation (« position », « occupation » and « activity ») we are aiming to reconstruct the instantaneous distance, speed and mass variables for each truck. The elementwise (Δ) travelled distance is calculated as follows:

$$\Delta \ell_{t_i} = \sqrt{\Delta x_{t_i}^2 + \Delta y_{t_i}^2}$$

where x_{t_i} and y_{t_i} are the coordinates of the vehicle i under study at time instant t_i. The instant speed can be calculated by:

$$V_i = \frac{\Delta \ell_{t_i}}{\Delta t_i}$$

The main contribution in the mass variation (Δ_{m_i}) is due to containers loading and unloading from vehicles. Therefore, the containers' movement and exchanges between the vehicles are tracked in the database. In this study we assume each single unit of containers has a predefined fixed amount of mass.

By using the emission factors for each vehicle, the three calculated instantaneous signals (namely $\Delta \ell_{t_i}, V_i, \Delta_{m_i}$) and the engine's model we can reconstruct the time series of emissions of each truck and, after aggregation, of the whole simulated terminal.

The extraction of the interested variables from the simulated model of the port is in fact a post-processing procedure. This post-processing is done basically on three databases generated from simulation. These are position database, activity database, and occupation database. The position database records the position of each vehicle. The activity database registers the container and vehicles interactions. And the container exchanging in the loading/unloading points is stored in the occupation database.

For calculating the mass variations, the so-called activity database should be screened for the vehicle under study. This information can be extracted by looking at the column "ref_transport". This column contains the tag of vehicles in which they are loaded by a container in a specific time. The container loading/unloading action data of a specific vehicle is extracted from there. The database (activity) is grouped with respect to the vehicles in the "ref_transport" column. After that, this data should be merged with the position database of that vehicle. For doing so we have to consider, the law of continuity of the position of each vehicle within the time.

The loading/unloading action is read from the "mouvement" column of the third database which is called "occupation". For doing this action, the container tag should be matched with the container tag in "ref_produit" column in the merged databases of "activite" and "position".

In this study it is assumed that all the trucks are 5-LH category. This assumption is not far from practice, as in [99] this subgroup is reported as the most common vehicle in Europe and in [100] it is mentioned that 61.9 % of the regulated trucks are among this category. These trucks are assumed to comply with Euro VI standard. In Figure 36 the CO2 emission factors for each sub-group of vehicles produced from different manufacturers are plotted. It shows the vehicles within the sub-group 5-LH have emission factors close to 57 g-CO2/ T-Km. By considering these assumptions, the emission factor for this study is assumed to be 57 g-CO2/ T-Km[1].

[1] This factor is adopted from the following:
https://theicct.org/publication/co2-emissions-from-trucks-in-the-eu-an-analysis-of-the-heavy-duty-co2-standards-baseline-data/
and see p. 45 Transport & Environment (2021). *Easy Ride: why the EU truck CO2 targets are unfit for the 2020s*

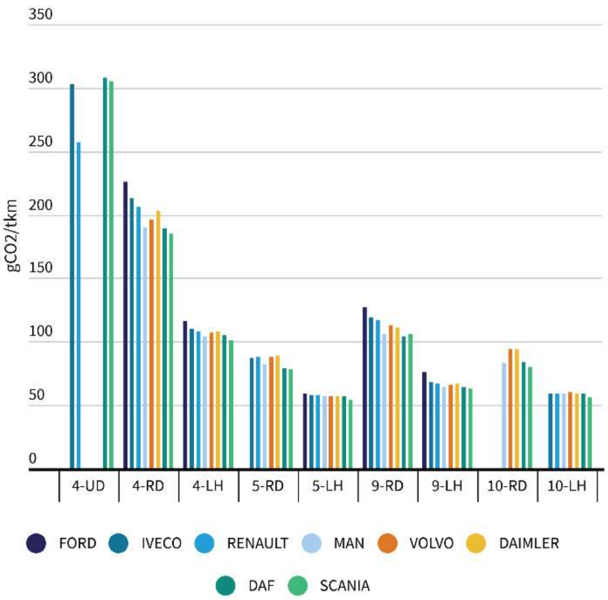

Figure 36 CO2 emissions of heavy-duty vehicles vs. the vehicles sub-groups, analysed by Transport and Environment during the reference period (July 1st 2019 to June 30th 2020), published by EEA in June 2021 [47].

For the mass variations of the vehicles, the mass of the containers is assumed to be 13.2 Tonnes, which is selected from Table 29 in [99]. The mass of the 5-LH vehicle subgroup is selected as the minimum value of the mass in this group. This value is reported in many references as 16 Ton. The curb weight of the 5-LH vehicle is selected as 7.8 Tonnes [101]. It is assumed that the weight of the trailer is 6.5 Tonnes [102]. Therefore, the weight of the transportation vehicle without container is considered to be 14.3 (7.8+6.5) Tonnes.

In order to evaluate the CO_2 emissions of trucking activities within Proto Port, we define the following performance indicator. The following equation is used for calculating the CO_2 emission within time interval of $[T-t_i, T+t_f]$

$$\mathrm{CO_2\ Emission\ within}\ [T-t_i, T+t_f] = E_{CO_2} \sum_{i=1}^{n_{tr}} \sum_{k=T-t_i}^{T+t_f} v_i(k)\, \Delta k\, m_i(k)$$

where Δk is the simulation time step, n_{tr} is the number of active trucks within the time interval, $v_i(k)$ is the speed of the vehicle in that time interval, $m_i(k)$ is the mass of the vehicle at the instant k, and E_{CO_2} is the emission factor of the vehicle. The initial simulation time is $T-t_i$ and the final simulation time is denoted as $T+t_f$. This is a function of time (T). When it is reported within an hour it can be indicated with the unit of g-CO_2/h. There are many insightful indicators that one may extract and analyse from the post-processing of the Proto Port simulation data.

Figure 37 shows the histogram of the trucks unitary CO_2 emissions, that is the statistical distribution of the kilograms of CO_2 emitted by each single truck during its activity time.

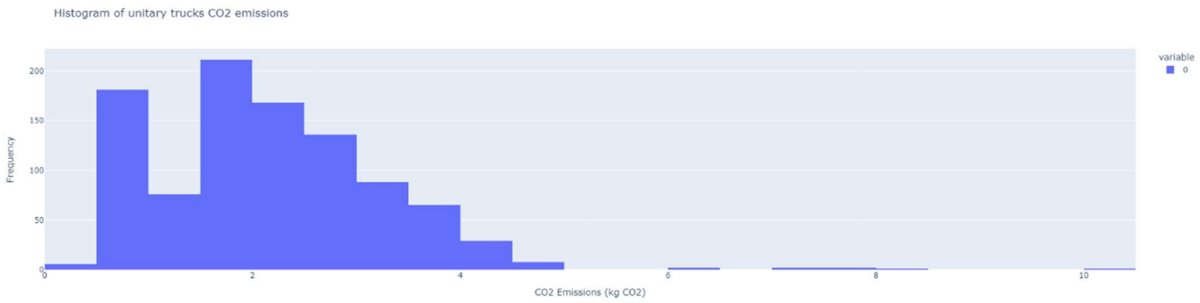

Figure 37 Proto Port simulation data analysis – Trucks unitary CO2 emissions

Figure 38 displays the number of active trucks over the simulation time. This is clearly time-dependent and due to the mission schedule simulated within Proto Port. The number of active trucks is important to compute the global CO_2 emissions of Proto Port as illustrated in the following.

Figure 38 Proto Port simulation data analysis – Number of active trucks over the simulation time

Figure 39 shows the amount of CO_2 emitted every hour by all the trucks active in the simulation. In other words, this can be seen as the hourly CO_2 emission of Proto Port (only due to trucking activities)

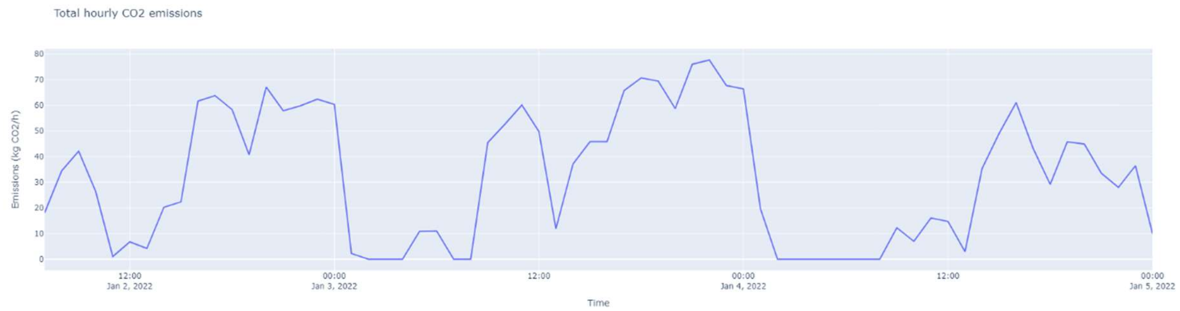

Figure 39 Proto Port simulation data analysis – Trucks hourly CO2 emissions

Finally, Figure 40 displays the instantaneous variables that we were able to extract from the simulation data for each truck, namely speed, position, mass (with accurate consideration of loading/unloading of containers according to the scheduled missions) and CO_2 emissions.

These results in having a second-by-second CO_2 emission for each truck, which can be easily converted in instantaneous energy demand of these vehicles. Instantaneous CO_2 emissions and energy demand of vehicles or aggregated at the terminal scale represent very valuable information for port authorities and their decision-making process. Furthermore, such instantaneous information may be used as an input for MAGPIE tools (WP 4.5) as illustrated in the deliverable structure diagram at the beginning of this document.

Figure 40 Proto Port simulation data analysis – Instantaneous speed, mass and emissions signals for each truck

As a final note, it would be possible to generalize this kind of analysis to any container terminal by, for instance, learning the hourly patterns of trucks movements, or the typical trucks load during the day, or even the number of incoming containers observed in the Proto Port simulated data. The adaptation of the obtained energy demand profiles to a real port could then be achieved by applying a scaling factor depending on aggregated information (easier to obtain) about port activities.

4. Tools for traffic emissions modelling and carbon simulator for ports

As shown in Figure 2, in the scope of T4.4. several back-end models and tools were developed to produce the necessary time dependent logistics, energy and emissions data that will be needed for the tools being develop in T4.5. This chapter describes the tools that were developed or enhanced as part of work carried out in task 4.4:

- IFPEN developed a tool that maps CO_2 and pollutant emissions from port activities and their impacts on areas adjacent to the ports, building from the models for port traffic described in section 2.3,
- CIRCOE developed a carbon simulator that estimates the carbon emissions of a terminal considering different options of technology and fuel substitution, and traffic management,
- This chapter also describes some of the features of the GHG tool, developed by TNO, as some of the inputs and characterization were developed within T4.4, namely how the model underlying the GHG tool will extend the Decarbonisation Model (Decamod).

In addition to the insight that the tools described in this chapter provide, the inputs and outputs of the tools may be used as input for other models and tools under development in WP4 as shown in Figure 2.

4.1. Interactive dashboard for a comprehensive port-area emissions map

An interactive dashboard is defined as a data management tool that enables users to interact with data by tracking, analysing, monitoring, and displaying key metrics. By utilizing an interactive dashboard, users can delve deeper into operational information and apply multiple filters to it. This includes monitoring the impact of different sources on road traffic emissions, comparing trends over specific time periods, and conducting historical comparisons.

Technical aspects and features description

The interactive dashboard is based on the Mobicloud cloud services platform developed and operated by IFPEN (https://mobicloud.ifpen.com/). Mobicloud enables the generation, enrichment, and visualization of data through web-services, automated processing workflows and web visualization tools.

These components are based on standard software stacks to ensure interoperability with other platforms (MongoDB or Postgresql databases, communication between services via REST APIs using the JSON exchange format).

This 100% Cloud platform is deployed by the provider OVHCloud and uses a set of virtual or dedicated servers in the various French Data Centers offered by this provider.

Setting up interactive dashboards to visualize simulations of pollutant emissions or air quality can be divided into three key stages:

- Simulation data generation
- Data consolidation / aggregation
- Visualization / analysis of results

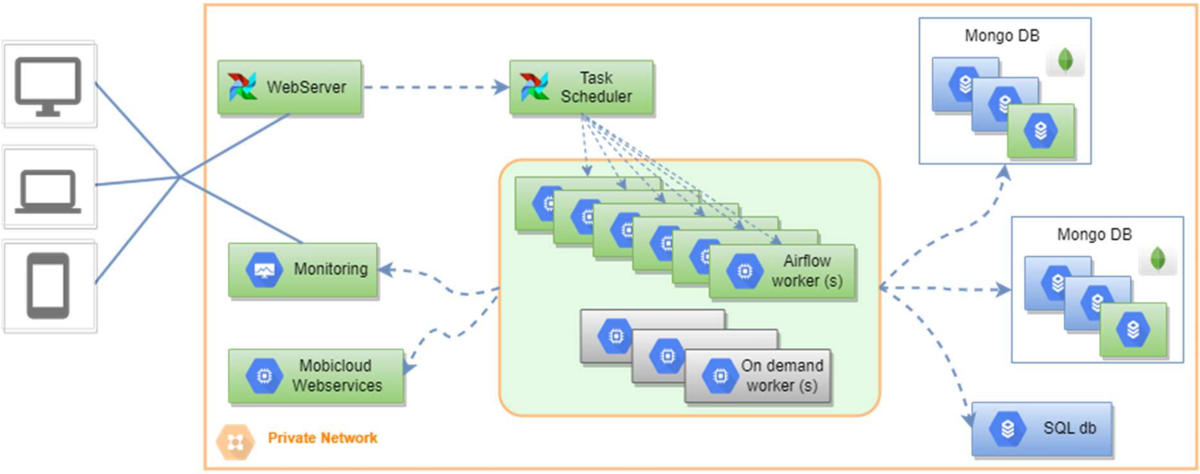

Figure 41 - Software stack for simulation data generation

Simulation data generation is based on the Apache Airflow solution (https://airflow.apache.org/), which orchestrates a set of tasks according to available computing resources. This solution has been deployed on the Mobicloud platform with a set of dynamically configured workers that can evolve to meet demand. The definition of the various tasks and their interdependence is done in Python. The various simulation algorithms running on this platform also use this language.

Workers can access the various Mobicloud databases needed to generate and analyse simulation data in the form of interactive dashboards:

- A Mongodb database containing all raw simulation results. This database (MongoMobicloud) is a Mongo replica-set made up of 3 servers, ensuring continuity of service by replicating data in 2 datacenters. This database is fed by processing workflows executed on Airflow workers.
- A Postgresql + Postgis relational database (DataMobicloud) containing data ready for viewing on interactive dashboards.
- Finally, a Mongodb database is used for dynamic configuration of visualization interfaces and web-services (monitoring).

Simulation data can be generated:

- Directly from the Airflow platform
- By calling Mobicloud web-services (python + Django) interfaced with the Airflow platform, using swagger or scripting tools.
- More simply, via dashboards using these web-services (see dedicated section: Interactive dashboard for simulation requests).

Data consolidation/aggregation

A direct visualization of simulation results was first implemented, with interactive dashboards querying Mobicloud APIs requesting raw results from the MongoMobicloud database. Performance was limited when it came to visualizing simulations on a large number of road segments, aggregating results over the entire simulated territory or comparing several simulations on the same territory.

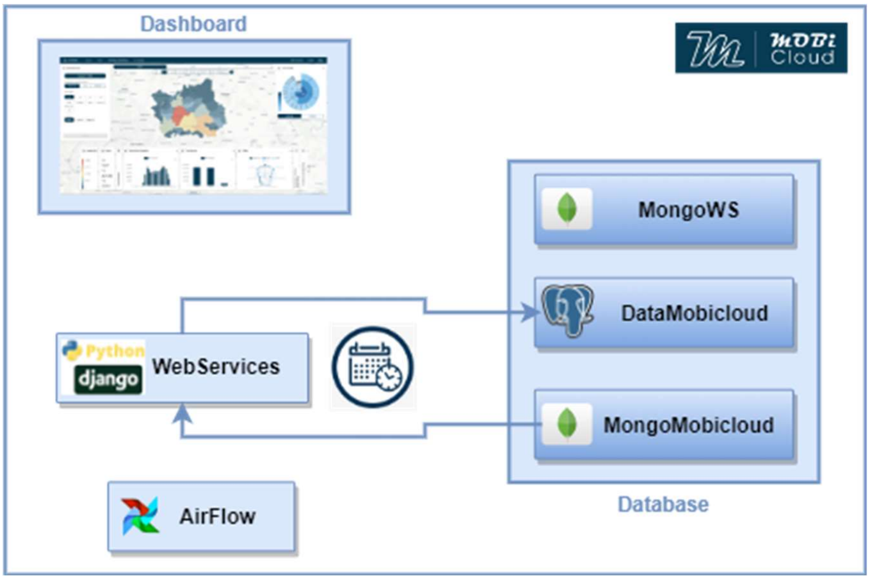

Figure 42 - Software stack used for the dashboard

Cold processing of the results was thus implemented to optimize visualization. Data formats are optimized in the DataMobicloud database (Postgresql + PostGis), and results are also aggregated across different topologies (H3 tiles and Iris zones for French territories). A set of KPIs is pre-calculated for the various simulations (min, max, mean of each data item to quickly analyse changes between simulations, or according to the time of day).

A set of tables has been set up on the DataMobicloud database to store the data available for each topology.

These processes are carried out automatically every night (cron job on a Mobicloud server) for new results to be visualized, or on request.

4.1.1. Visualization / analysis of results

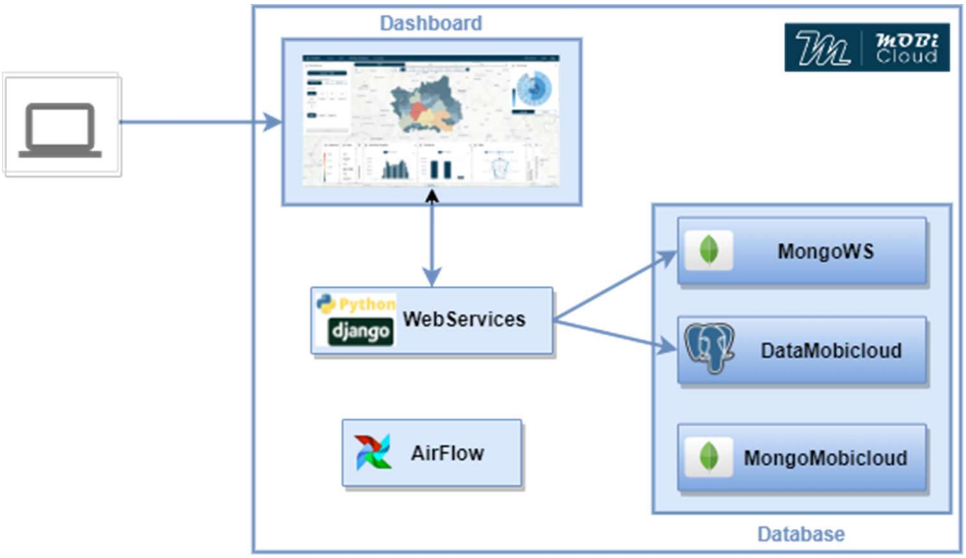

Figure 43 - Software stack used for visualisation

Results can be visualized and analysed using dashboards deployed on Mobicloud, based on ReactJS for the front-end and Python + Django web-services for data access.

The configuration of the interface and visible results is defined in the database so that it can be easily updated when new studies are added (web-services accessing the MongoWS database).

Access to calculation data (spatial data and KPIs) is also via Mobicloud web-services accessing the DataMobicloud consolidated database.

Overview of the interactive dashboard prototype

Once the technical background has been presented, this section focuses on the actual dashboard from the user's perspective.

The first step is to log in to the Mobicloud platform (in the case of an existing account). After successful login, the user is directed to the landing page where available simulations are displayed on the map (Figure 44). At present time, simulations of the port areas of Rotterdam, Le Havre and Sines are available.

Next, the user is taken to the statistics page called "Stats." Here, the user can access an aggregated view of simulation results at the territory scale. Information on the simulation performed, can be retrieved. A small map displaying the territory covered by the simulation is available in the top left corner (Figure 45).

A table presents KPIs. By default, only two scenarios are selected if available, but more can be added for a comprehensive analysis. The table displays indicators related to simulation parameters, such as road length, mean velocity, number of vehicles, population, mean road traffic noise, and more. It also presents yearly extrapolated emission quantities for each pollutant.

Policy makers seeking high-level insights can get a general view from this page. For example, they can quickly identify that with the Low Emission Zone implemented in 2030, NO_x emissions will be reduced by approximately 87%. The "Analyse" button allows users to delve into the data in more detail by switching to the map-based view.

Figure 44 Interactive dashboard – Landing page with the available simulations

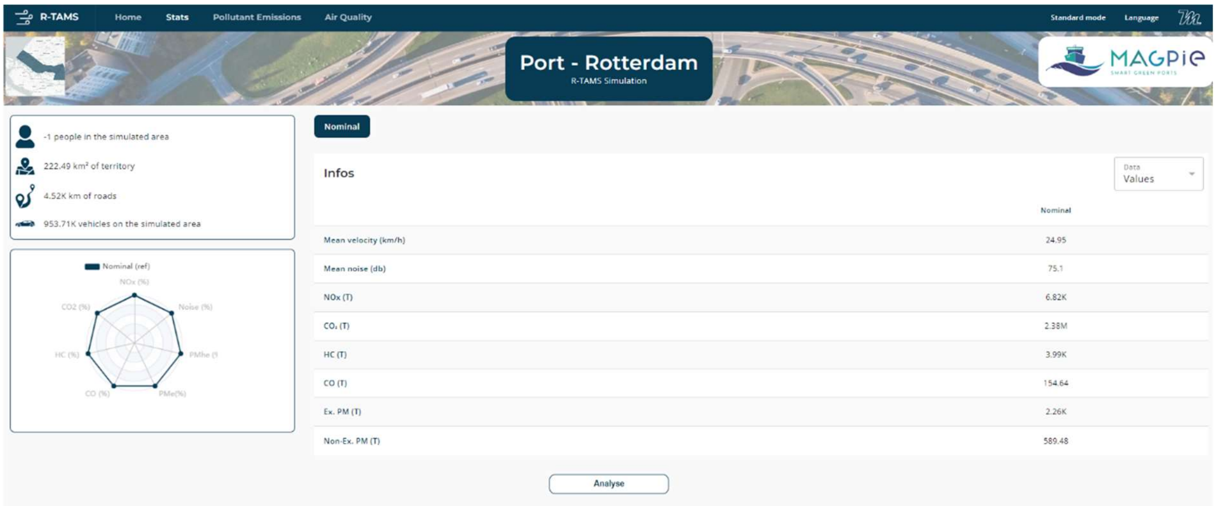

Figure 45 Interactive dashboard - Page with macroscopic simulation statistics

In the pollutant emissions tab, the user starts by choosing the type of visualization, such as by road segment or by tile (hexagons).

Next, the user selects the pollutant to analyse from options determined by the mesoscopic model capabilities, including NO_x, CO_2, HC, CO, exhaust and non-exhaust PM, and noise. For validation purposes, intermediate quantities used to estimate pollutants, such as speed limit, average speed, flow rate, slope, etc., can also be visualized.

Then, the user can choose the representation of the quantity, such as mg/s, mg/km, or specific representations for area aggregations in mg/s.km. Further work will be conducted in the project to refine the choice of representations and calculation methods. Currently, mesoscopic models estimate pollutants in mg/km, which represents emissions per vehicle on an average fleet. Emissions per vehicle are firstly determined by the model, followed by a weighted average aggregation step to get average vehicle emissions depending on the fleet distribution hypothesis considered in the simulation. However, this variable only considers the effect of infrastructure and does not integrate the impact of traffic volume. Another variable in mg/s is determined to account for traffic volume and road segment length. The most representative unit will be determined based on specific needs, such as average emissions in mg/km, sum of emissions in mg/s, normalized quantities by surface area or total road length in the area, etc.

The next setting parameter is the colour scale, which can be linear, quantile, or logarithmic. The user can also select the time period associated with colour scale on an hourly, daily scale, constant across available scenarios or even custom.

Once the data is loaded, the user can observe emissions on the desired territory over a background map. The tool provides information about the assumptions of the vehicle fleet and emissions based on the zoom level: statistics, emission distribution, hourly evolution, or evolution of emissions according to simulated scenarios. Figure 46 shows the pollutant emissions page, perhaps the main analysis tool of the proposed interactive dashboard. In the following, a brief description of the various dashboard elements is given.

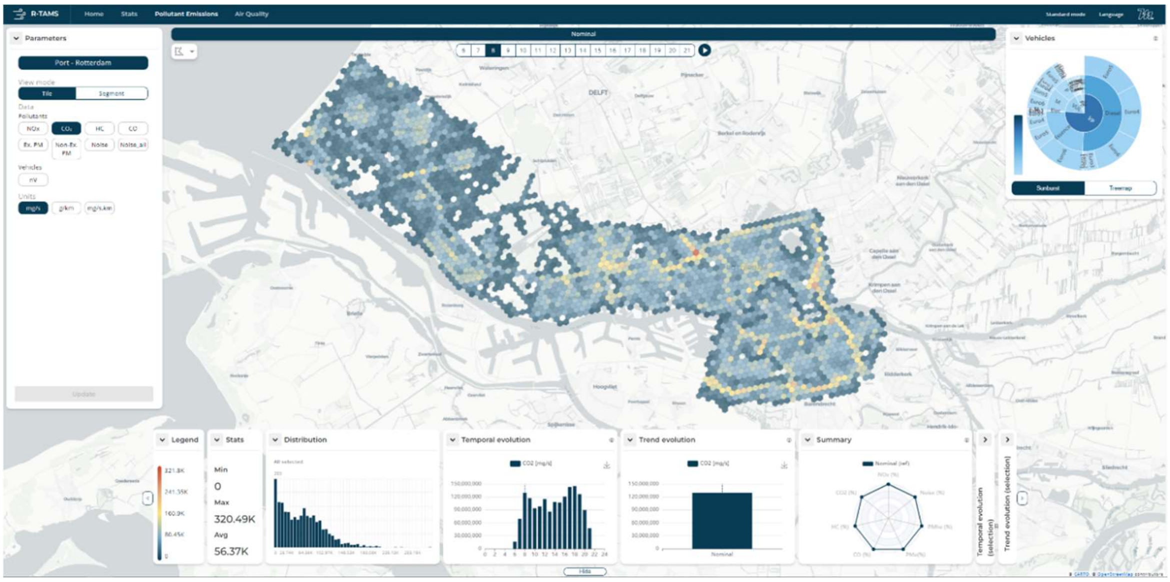

Figure 46 Interactive dashboard – Page with pollutant emissions visualisation on a map (example of the Port of Rotterdam area).

0- Mode and language selection, Mobicloud home:

The dashboard offers two modes: a light mode and an advanced mode. The light mode provides a simplified interface for users seeking a quick overview, while the advanced mode offers more in-depth parameterization options. Additionally, users can select their preferred language (currently available in English and French). Clicking on the Mobicloud logo allows users to navigate back to the Mobicloud home page.

1- Page selection:

Users can switch between different pages within the dashboard. In addition to pollutant emissions, the dashboard provides pages for landing, statistics (Stats), and air quality. This allows users to explore different aspects of the data and gain comprehensive insights.

2- Parameters selection:

This section has been described earlier. The "Compare" button facilitates the visualization of absolute differences between two scenarios, enabling users to perform in-depth comparative analysis.

3- Main map-based visualization:

The primary section of the dashboard presents a visually appealing map where users can visualize emissions data.

4- Time frame of available hours:

Users can select a specific hour to analyse or utilize the play button to automatically cycle through available hours at a regular interval of 5 seconds. This feature enables users to observe how emissions change over time and identify temporal patterns.

5- Scenario selection:

In cases where multiple scenarios have been simulated, users can select the desired scenario from the top of the dashboard. This feature enables users to analyse and compare the impact of different scenarios on air quality.

6- Vehicle fleet distribution:

This graph provides insights into the distribution of the vehicle fleet. Users can analyse the percentage composition of personal cars, heavy goods vehicles, light commercial vehicles, and two-wheeled vehicles. Additionally, the graph displays the proportion of engine characteristics within these vehicle categories, offering a comprehensive understanding of the fleet composition.

7- Data subset selection tool:

This tool allows users to select a specific area on the map. The selected area is then displayed through widgets such as "Stats" and "Distribution," enabling users to focus their analysis on a specific region of interest.

8- Legend

The legend provides a key for interpreting the colours and symbols used on the map.

9- Statistics widget:

The statistics widget offers users an overview of global statistics based on the selected zoom level. It provides essential information such as minimum, maximum, and average values related to the displayed data. This dynamic feature allows users to obtain key insights into the data at a glance.

10- Distribution widget:

The distribution widget presents a histogram that visualizes the distribution of a specific variable of interest. Users can interact with the histogram, selecting specific portions to highlight areas with for example higher emission values. This interactive feature facilitates the identification of patterns and outliers in the data.

11- Temporal evolution widget:

The temporal evolution widget displays how emissions change throughout the day for a selected road segment, or tile. It provides an hour-by-hour analysis, allowing users to observe trends and fluctuations in emissions over time. Additionally, it presents trends for higher, lower, and mean values across the territory for each hour.

12- Trend evolution widget:

Similar to the temporal evolution widget, the trend evolution widget enables users to observe the evolution of emissions across different scenarios. Users can compare and analyse emission trends across various scenarios, providing valuable insights into the impact of different factors on air quality.

13- Additional widgets:

The dashboard is designed to accommodate the addition of new widgets in the future. These widgets will be accessible through a carousel-type button, allowing users to access and explore new functionalities as they become available.

In conclusion, the interactive dashboard provides a user-friendly interface for visualizing and analysing air quality data. Each element of the dashboard is thoughtfully designed to enhance the user experience and facilitate in-depth exploration of emissions, scenarios, and temporal patterns. This dashboard might already assist users in gaining valuable insights and making informed decisions related to air quality management.

4.2. Greenhouse Gas tool

As part of the MAGPIE project, a Greenhouse gas tool will be developed. The main aim of the tool is to analyse transport chains that go via a port and to collect and organise carbon footprinting data to establish GHG emissions along the transport chains on the origin-destination level.

The emission data gathering and analysis process will be approached in a structured way. Transport chains related to the port will be split into uniform transport chain elements, for which the data will be gathered per modality and per cargo type.

The tool will deliver for each transport chain element, GHG emission intensities and per tonne-km or tonne throughput measured in CO_2 equivalents. The presented emissions will be compliant with the ISO 14083 standard (published in March 2023) [103] on quantification and reporting of greenhouse gas emissions arising from the operation of transport chains. The results will furthermore be aligned with the CountEmissionsEU proposal, which will serve as a common framework for quantifying the greenhouse gas emissions of transport services across different modes.

Figure 47 presents a schematic overview of the different chain elements, the required input and the link to ISO 14083.

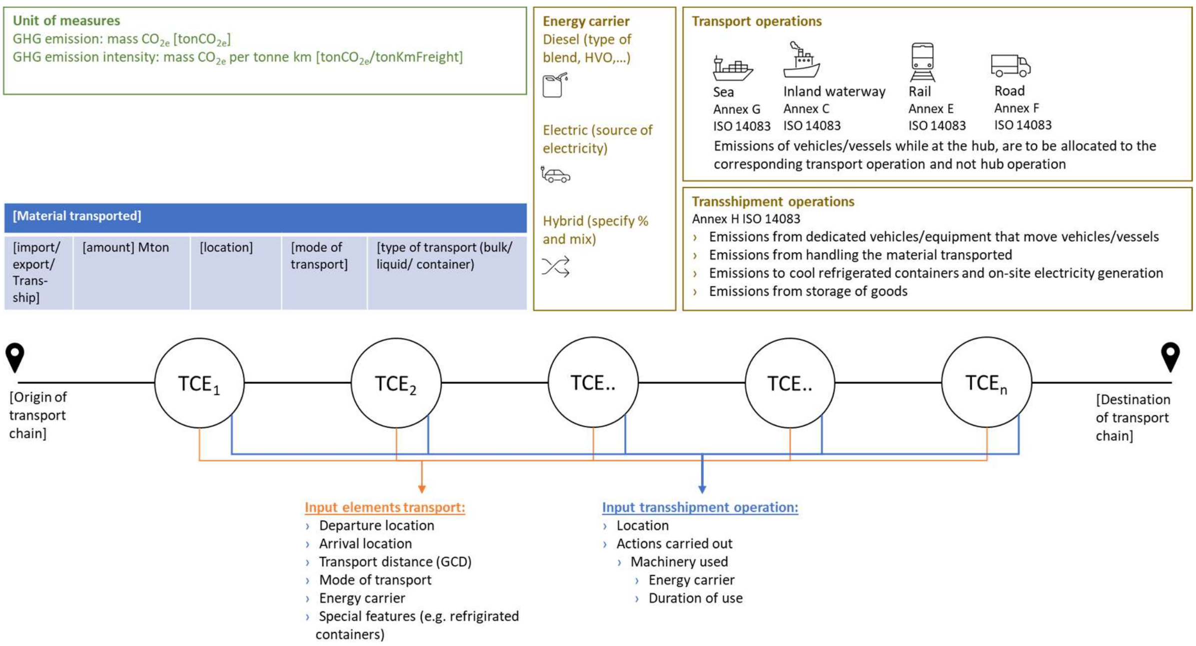

Figure 47 – Overview of the different chain elements considered in the GHG emissions tool.

The scope of the tool will be:

- The tool will provide insight on the emissions for freight transport on the level of :
  - individual supply chains,
  - Freight corridors (both maritime and hinterland corridors), and
  - port as a whole
- All elements of the logistics supply chain linked to the ports will be included:
  - Maritime transport and hinterland transport (road, rail, and IWT) on both sides of the maritime leg.
  - Emissions of transhipment

- All types of goods will be included (e.g. dry bulk, liquid bulk, general cargo, containers, Ro/Ro)
- The tool will give insight in the current GHG emission levels and for forecast scenarios up until 2050
- The tool will furthermore give insight in the effect of GHG reduction measures, including:
    - Application of alternative energy carriers (such as electric, bio-fuels, hydrogen),
    - Other technical options (wind assist for ships, truck platooning)
    - Efficiency and logistics measures (slow steaming, modal shift)
- The tool will first be applied for the Port of Rotterdam. Depending on the data availability, the tool might be extended to the other MAGPIE ports.

The model will extend from the Decarbonisation Model (Decamod) that has been created by TNO. Decamod is a CO_2 accounting model that provides insight in the impact of decarbonisation measures in logistics. The scope of the model is all landbased freight transport (road, rail, and inland shipping) in the Netherlands. The model allows for applying so-called 'what-if' scenarios to provide insight into the effects of various sustainability scenarios. After defining the scenarios to be analysed, Decamod can calculate them relatively quickly. The results support decision-making processes on CO2 reduction measures in logistics with targeted quantitative analyses. As an extension to the current toolset, a maritime module and a transhipment module is being developed [104].

The scope of the Greenhouse Gas tool extends that of currently available Carbon Toolsets and frameworks, such as the GLEC Framework, BigMile tool, and CO_2 calculation tools such as EmissionInsider and Routescanner:

- Many tools primarily focus on container transport. First results of the Decamod study show that this only relates to around 10% of emissions of freight transport on a corridor, and that thus, it is important to also take into account other freight transport [104].
- The tools focus on individual routes and not on the total impact on the level of a corridor or a port.
- The available tools give insight in the current emissions, but do not give insights in the reduction potential of reduction measures.

The Greenhouse Gas tool will both make use of results of other tools and models as it will feed them with information:

- The GHG emission factors that are used in the different toolsets will be coherent and in line with other legislative EU documents.
- The GHG tool will provide insights on trade flows to the Logistics Optimisation tool. Outcomes of logistics measures of the logistics optimisation tool will be taken as input in the GHG tool
- The GHG tool will make use of the information on energy use and emissions of terminal operations of the Terminal simulation tool.

4.3. Carbon simulator tool for ports

Data collection of CO2 emissions on three terminals (Le Havre port)

4.3.1. Le Havre port information

Haropa Port encompasses all the port entities along the Seine; Le Havre, Rouen and Paris. This pooling of facilities allows commercial advantages and allows a larger offer for a wider number of maritime sectors. For example, the port of Rouen is one of the major French ports for bulk cereals export, while the port of Le Havre is the leading French port for the transport of containers.

The port of Le Havre is the first French port and the 5th European port for containers. It is composed of two distinct parts comprising 6 maritime terminals. One part is located on the historic port, gathering 3 terminals and a second port created in the 2000s to accommodate larger container ships.

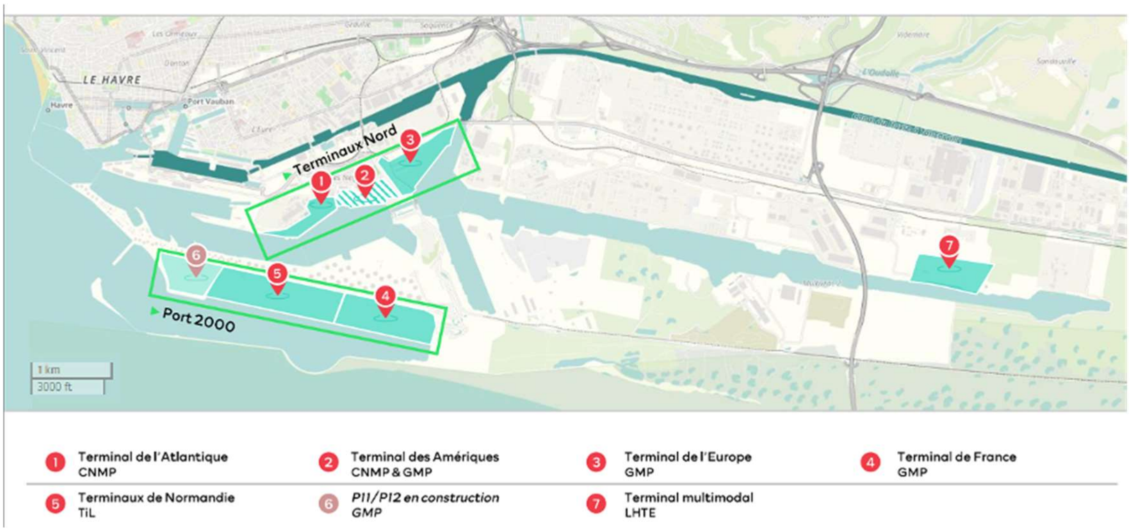

Figure 48 Haropa port's data.

Each terminal has its own equipment to manage container flows. The available equipment at the port is organised as shown in the table below:

Table 20 - Characteristics of the equipment in the container terminals of the Le Havre Haropa port.

	Quai des amériques	Quai de l'atlantique	TILL TN MSC	terminal de France
Dock length	500 m	800m	2100	1400 m ; 4 postes à quai
Draught	Min. 14,50 m, 80% days > 14,80 m	Min. 12,80 m, 80% days > 13,80 m	15,3 - 17,3	15,3 - 16,3
Crane	5 post-panamax	4 post-panamax	9 super post panamax + 2 over Panamax	14 super post panamax + 3 portique ferroviaire
Spout length	18/20 EVP	18 EVP	22/23 EVP	22/23 EVP
lifting height	33/40 m (ZPMC)		40/42 metre	6 : 40m / 4 : 49 / 4: 54 m
Surface	35 ha		1100	72
Straddle Carrier	30		Type 2+1 thermal :10 Type 3+1 thermal : 63 + 1 élec	68
Reefer	300		1100	1183

The port also has a multimodal platform allowing delivery both by barge to Paris up the Seine and by rail to destinations throughout France. This platform aims at boosting modal shift towards transport solutions other than 100% truck transport.

4.3.2. Energy consumption of vehicles in the Le Havre port

With regards to the carbon emissions due to container flows in the port of Le Havre, we focused on the consumption of the various engines used to handle containers on the quays. These include unloading and loading gantries, straddle carriers allowing the movement of containers and gantries used for the loading and unloading of barges and trains for multimodal transport. Data per container is taken into account thus allowing to simulate projections depending on the increase or decrease of traffic. Input data includes the following:

Gantries:

- Consumption of one move (Wh) = 14kWh
- Number of moves per hour = 24

Straddle carriers:

- Consumption in liters of a hybrid straddle carrier (mostly used in the port of Le Havre) = 15 liters
- Average number of movements for one container = 3

4.3.3. Input data

Figure 49 shows the input required for the tool:

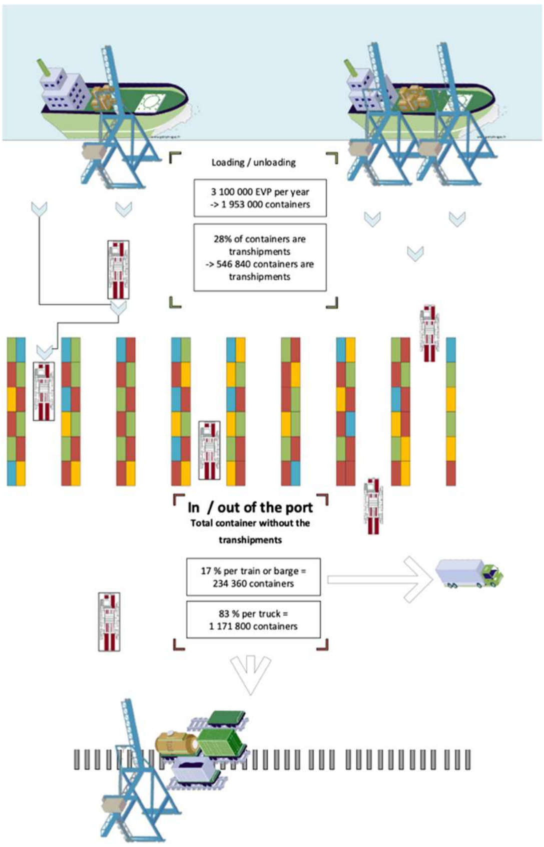

Figure 49 – Input data for the Le Havre port.

4.3..4. Consumption

Table 21 Shows the energy consumption of different equipment at the Le Havre port, calculated from local measurements and data provided by the terminal operators.

Table 21 - Data used in the CIRCOE carbon simulator, as calculated for the Le Havre port from local measurements and terminal information.

Mobile lifting cranes	Straddle carriers	Mobile lifting cranes
Power by move = 14KWh	Consumption in litre per hour = 15 litres	Consumption = 14KWh
24 moves per hour 1 move = 1 container	1 container = 3 moves (in / out and internal move	24 moves per hour 1 move = 1 container
1 953 000 containers x 14 kWh = 27 342 MWh	1 move = 500 metres = 5 min in average	234 360 x 14 kWh = 3 281 MWh
	Per container = 3 moves = 15 min = 3.75 litres	
	Per year = 1 953 000 x 3.75 litres = 7 323 750 litres	

4.3..5. Carbon impact

To calculate the carbon emissions of these flows (Figure 50), we have used the information provided by ADEME (French agency for the ecological transition). For diesel energy, we took into account the emission factor based on CO_2/litre and not CO_2/ton/km because the weight of each container is unknown information. Regarding the electricity emission factor, we have found a low result because it is based on the French energy mix, which is largely composed of nuclear energy and renewable energies.

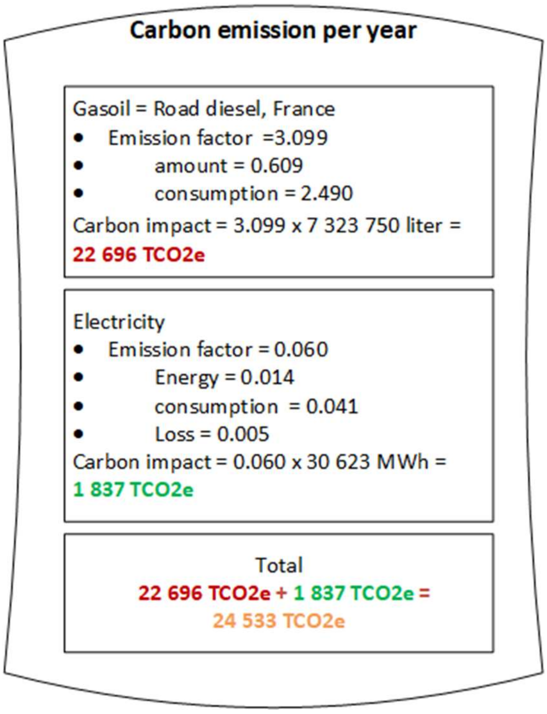

Figure 50 - An example of the output of the carbon simulator: Carbon emissions per year.

4.3.6. Carbon simulator

A simulator was created, taking into account vehicles with carbon emissions and then aiming at analysing what would be the possible evolutions of these emissions if certain parameters were modified. We have played on the following parameters with this tool:

- Modification of port activity
  - Increase / decrease of the number of TEUs
  - Distribution between 20ft and 40ft containers
  - Percentage of transhipments
- Straddle carrier
  - Average distance travelled
  - Truck consumption per hour
- Multimodality
  - Percentage of departures or arrivals by train or barge

These changes of parameters allow to anticipate the evolution of volumes at the port and aim at focusing on the straddle carriers which consume less energy or cover shorter distances, as shown in the two simulations below:

- The increase of 1 million TEUs on the port automatically leads to an increase of the carbon impact:

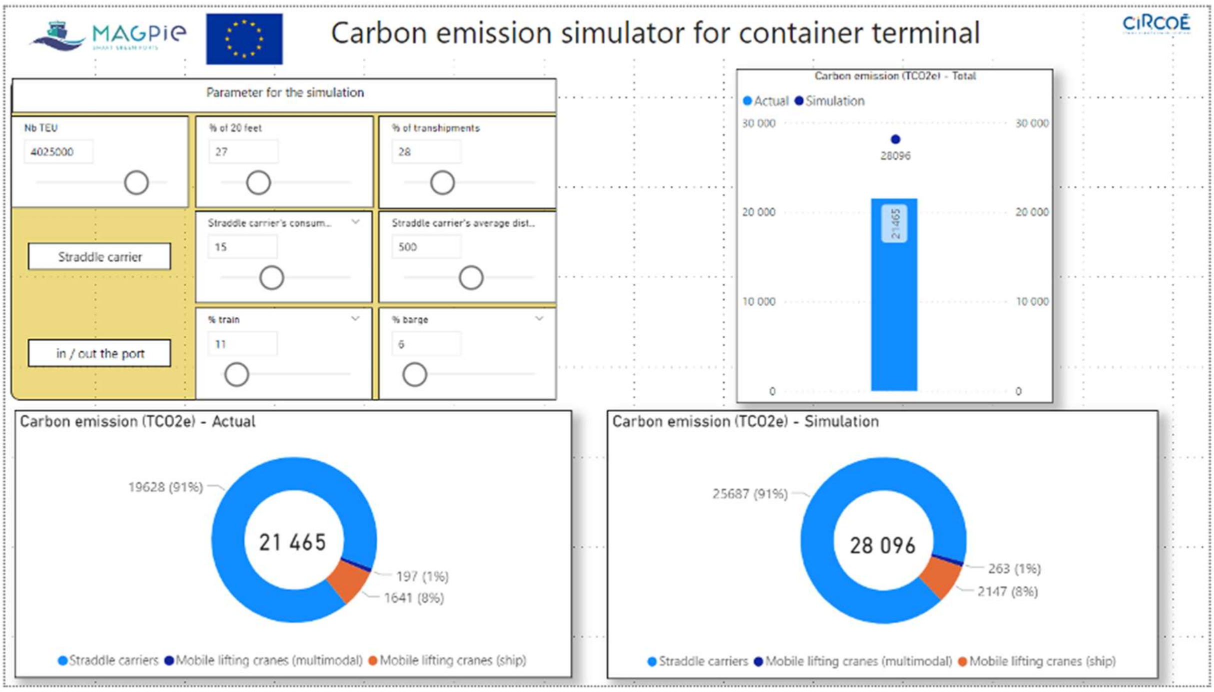

Figure 51 – User interface of the carbon simulator tool – results of an increase in annuals TEU throughput.

Maintaining the same number of TEUs while dividing by two the average distance covered by the straddle riders:

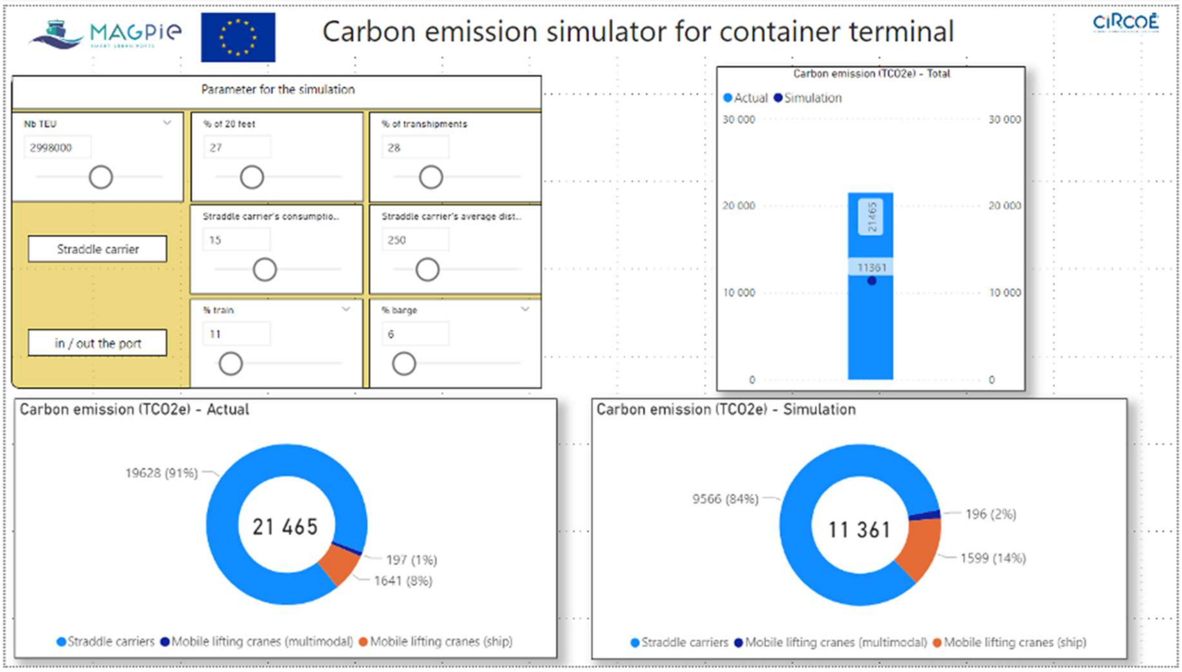

Figure 52 - User interface of the carbon simulator tool – results of halving the distance covered by the straddle carriers.

This tool may be used by MAGPIE partners on demand.

4.3..7. Conclusion

The carbon emission simulator shows that during the movement of containers on the port of Le Havre, the most significant carbon impact comes from the straddle carriers, due to the fact that diesel energy generates a high CO_2 level. As for gantries, the carbon impact remains limited since they run on electricity. In France, the energy mix is strongly based on the use of nuclear power plants as well as renewable energies. This results in a very low emission factor. Possible improvements towards reducing the carbon impact could be to cover smaller distances in the port of Le Havre or to use vehicles that consume less diesel. Another lever would be to replace diesel with an alternative energy, initially B100 or LNG for instance, and at a later date, hydrogen or ammonia.

Data collection of CO2 emissions for DeltaPort

4.3..1. Information about the port

DeltaPort is the association of the Rhine-Lippe Port and the City Port, located within the municipal area of the City of Wesel, plus the Port of Emmelsum, located within the municipal area of the City of Voerde. The unification of these port areas created a unique logistics network for waterway-, rail- and road-based transports on the Lower Rhine that is unrivaled in its form. State-of-the-art superstructures facilitate the handling of dry bulk, break bulk, liquid bulk and heavy cargo, as well as containerized and temperature-controlled goods.

4.3.2. Input data

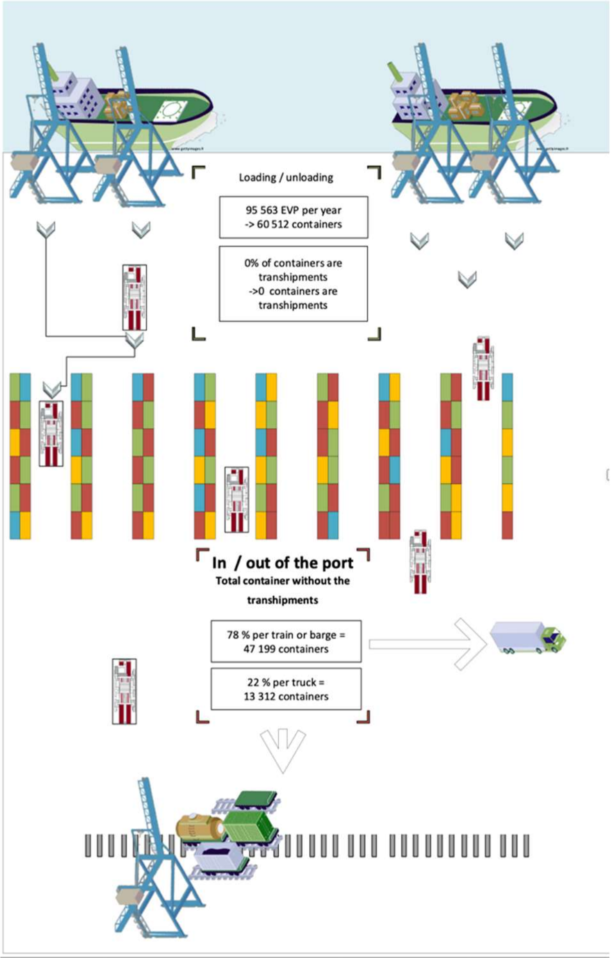

Fig. 53. Input data for Deltaport, without transhipment data

4.3.3. Consumption

Table 21able 22 Shows the energy consumption of different equipment at Deltaport, calculated from local measurements and data provided by the terminal operators.

Table 222- Data used in the CIRCOE carbon simulator, as calculated for the Deltaport from local measurements and terminal information.

Mobile cranes	Straddle carriers	Gantry cranes for barge
Consumption in litre per hour = 8 litres	Consumption in litre per hour = 15 litres	Consumption = 5.89KWh
25 moves per hour 1 move = 1 container	1 container = 3 moves (in / out and internal move	15 moves per hour 1 move = 1 container
60 512 moves = 2 420 hours	1 move = 300 metres = 3 min in average	35 096 x 5.89 kWh = 206 715 kWh
2420 hours = 19 363 litres	Per container = 3 moves = 9 min = 2.25 litres	
	Per year = 60 512 x 2.25 litres = 136 152 litres	

4.3.4. Carbon impact

To calculate the carbon emissions of these flows, we relied on the information transmitted by ADEME to know the emissions factors. For diesel, this is the emission factor combining combustion and manufacturing of the product. For the electricity emission factor, we have a low result due to the fact that this is based on the French energy mix largely composed of nuclear energy and supplemented by renewable energies.

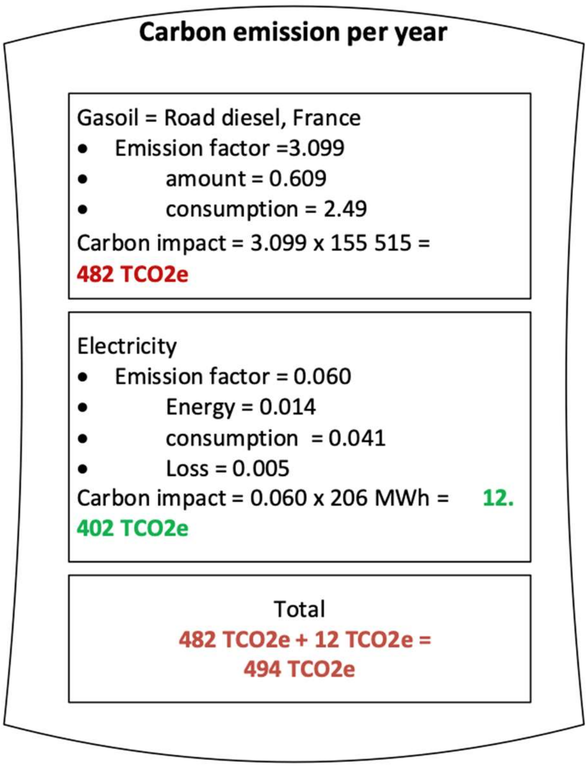

Figure 54- An example of the output of the carbon simulator: Carbon emissions per year.

4.3.5. Carbon simulator

We proceeded in the same way for Delta Port as for Le Havre port (see above), and came up with the following results:

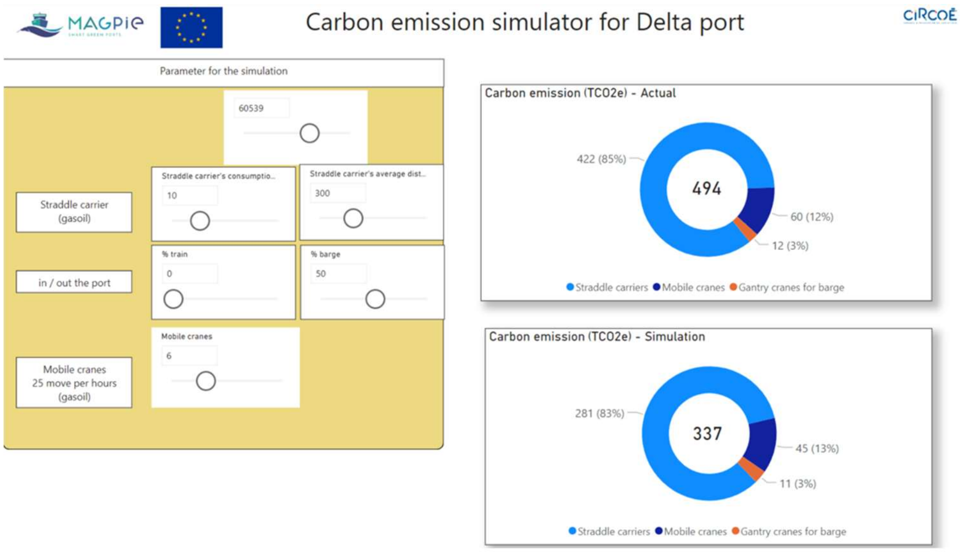

Figure 53- User interface of the carbon simulator tool – decrease of distance by straddle carriers

For the same number of containers, if we reduce the consumption of trolleys and gantries for barges that run on diesel, we will reduce the number of tons of CO2 emitted. A large part of Delta Port's equipment uses diesel, which represents the largest part of the environmental impact. As a result, it is necessary to work towards a decrease used of gasoil or for a reduction of the distance covered by the straddle carriers.

5. Conclusions

In conclusion, T4.4 successfully delivers a multi-dimensional digital port simulator, which provides synthetic data of a prototype port terminal in the form of time series, augmented with the associated energy requirements. These developments resolve a major difficulty faced in the MAGPIE project that is the lack of available data from the ports.

This document has given a thorough description of the backend models and tools, including identifying the scope, the data needs, and the outputs of each model. The connection with other backend models and with the tools under development in T4.5 have also been mapped.

The backend models are being developed and tested already with the terminal simulator outputs. The results presented here for a couple of use cases to map the energy demand of several assets within the terminal for a specific timeline, and to map the CO_2 emissions of trucks within the port area showcase the some of the capabilities and insight that the models will be able to provide once fully developed.

The carbon simulator tool of CIRCOE and the emissions mapping dashboard tool from IFPEN, which have already been applied to MAGPIE ports, further showcase the importance of these tools for green ports and their energy transition. Additionally, with the possibility of mapping impacts beyond the borders of the ports, these tools can also be used to improve interactions of stakeholders within the port and of ports with the hinterland. In particular, the tools can contribute to the ports dialogue with the cities and (sub)urban areas around them, by allowing (co-)development of scenarios of energy transition where the impacts of technological and non-technological (e.g., behavioural, logistics, planning) can be visualised and tested.

The models and tools presented here will continue their development in scope of T4.5, and the first version of the outputs and models will be released in M30, with the first deliverable of T4.5 – The first version of the MAGPIE digital tools. Further work is also in the pipeline to collect data, connect the backend models to the data sharing infrastructure of T4.2, and to make the connections with WP3 and the digital tools under development in T4.5. An update on the mathematical formalism, algorithms and implementation of the models and tools is expected for M30.

ANNEX 1

Task 4.4. is described in the following way in the Grant Agreement:

T4.4: Modelling and Intelligence (prediction capability) (M9-M24) [CEA (24 PM), INESC (15 PM), CIRCOÉ (4 PM), IFPEN (12 PM), TNO (5 PM)]

This task aims to develop a multi-dimensional digital model of the ports, compiling the infrastructural, logistics and operations domains, as well as the energy data-driven models from WP3, to provide a holistic, human-centric and simulation-based model for matching energy demand and supply. Physical and data models for the different systems in ports will be defined, to allow their dynamic representations over time. In addition, Machine Learning and Artificial Intelligence will be applied to predict the use of resources in ports, capable to be used as a service in T4.5, where optimization and decision-support services will be developed. For this, the main operations responsible for emissions and energy consumption in the ports area will be considered, such as mobility of both people and goods; loading/unloading phases at ships; and all operations involving vehicle movements. This task will enrich the models developed in DeCaMod and DeCaMod2 projects, responsible for assessing trends and decarbonisations measures in national transport. These models will be enhanced with new perspectives on emissions related to international shipping, operation of hubs and terminals and other relevant areas, not covered by the existing models.

ANNEX 2

This section describes the energy demand models that can be applied to other types of terminals – this complements section 2.2 which only describes the energy demand models for assets involved in container terminal operations.

Energy demand of Terminal Oriented activities in Dry Bulk Terminals

In the case of dry bulk terminals, the specific equipment and layout will be dependent on the commodities moved. For instance, as pointed out in [105], terminals moving coal and iron ore combine ship loader/unloader systems (equipped with large grabs), conveyor systems and stacker reclaimer systems to connect the yard to the bulk carriers. Whereas grain terminals (designed to handle and store wheat, soy, and other grains and oilseeds) can be based on discontinuous systems (such as those using grabs) and continuous ship unloaders (CSU), including pneumatic chain, screw, or twin-belt machines. Grain terminals could also use grain elevators stockpile or store grains using bucket elevators or pneumatic conveyors that scoop up grain from a lower level and deposit it in a silo or other storage facility. Large grain terminals can have dozens of large silos located next to each other. Many grain terminals offer additional services such as cargo sieving to calibrate the grain, and fumigation. Therefore, estimating the energy demand is a complex task, that should account for the specific conditions of each terminal, mainly the equipment involved and the parameters when operating these assets. During our literature review we were not able to identify approaches and models for estimating the energy demand in bulk terminals, with such a clarity as in the case of container terminals. It seems that there have not been so much academic efforts in this area, or at least the results are not so openly available. As an alternative, we focused on individual models of some of the most common and highly demanding equipment, even if they have not been analysed during its use in ports. We present next some insights about how we understand the energy demand can be estimated.

Conveyor belt systems

According to [106], even when several energy calculation models to design the drive system of belt conveyor (BC) systems exist, the complexity of such models motivates the quest for new approaches. The commonly employed models are derived from some well-known standards or specifications, such as the ISO 5048, DIN 22101 or JIS B 8805, but they adopt many complex equations, needed to describe the contributions of each part of the BC to the total energy consumption, requiring some detailed parameters, which are only suitable for the design rather than the optimization. Two main categories of energy models are described in [107]: resistance based energy models and energy conversion based models. Accordingly, an analytical energy model is proposed to estimate the mechanical power of the BC (P_t):

$$P_t = \frac{V^2 T}{3.6} + \theta_1 T^2 V + \theta_2 V + \theta_3 \frac{T^2}{V} + \theta_4 T, \tag{85}$$

where V is the belt speed $[m/s]$, T is the feed rate $[ton/h]$ and $\theta_1, \theta_2, \theta_3, \theta_4$ are determined by the structural parameters and components of a belt conveyor, by the operation circumstance and by the characteristic of the material handled, therefore, they are relatively constant for a certain belt conveyor. It should be considered that, in practice, maintenance, readjustment, retrofit, abrasion and circumstance change probably make a belt conveyor away from its design condition, consequently, changes of the parameters $\theta_1, \theta_2, \theta_3, \theta_4$. Hence, when (85) is applied to a practical belt conveyor for energy optimization, its four coefficients should be estimated through experiments instead of being derived from design parameters. In [107] the authors also gave insight into how to experimental determine the coefficients in equation (85). They explain that if P_t, V and T are measured on-line or off-line, $\theta_1, \theta_2, \theta_3, \theta_4$ can be estimated from these data to guarantee the accuracy of energy model. It is important to note that P_t is the mechanical power of a belt conveyor, it can hardly be measured directly, however, it can be indirectly obtained from the electric power of the motor by:

$$P_t = \eta P_M, \quad (86)$$

Note also that, once the parameters $\theta_1, \theta_2, \theta_3, \theta_4$ have been obtained and knowing the overall efficiency of the driving system (η), the electric demand can be estimated combining (85) and (86) based on the main parameters of operations: belt speed and feed rate.

In practice, power meters, encoders and electronic belt scales are usually equipped with belt conveyors to obtain P_M, V and T, respectively. For belt conveyors without permanent instruments for P_M, V, and T, the off-line parameter estimation is employed, where temporary instruments, usually portable ones, will be used for necessary experiments. On the other hand, if the belt conveyors are equipped with permanent instruments, the on-line estimation will be carried out. This model is also used in [106], [108]; whereas in [36], a simple approach is used, using a linear model that considers the power consumption approximately as a primary function of the operating speed:

$$P = kV + b, \quad (87)$$

where k and b are parameters determined by the system structure. The study set four standard speeds to investigate the power consumption and acceleration/deceleration dynamics around different standard speeds. Thus, the large-scale linear model was replaced by segmented linear models for simplification. Although a mathematical model based on DIN 22101 was obtained in [109], it is interesting to observe the linear dependency of the electrical power drive and the material flow $[t/h]$. This linearity can be observed on the measurements results that were carried out for different materials and different speeds. Therefore, although the speed and type of cargo will affect the energy consumption, a general model (based instead on the throughput of the conveyor [t/h]) could be first used if sufficient detailed data is not available.

A comparative study on power calculation methods for conveyor belts is conducted in [110]. The authors analysed the performance of the conventional CEMA, DIN and Dunlop–Fenner methods; but also compare some data-driven approaches. In this case, it seems that non-linear regression (NLR) and gene expression programming (GEP) could be appliable for the evaluation of the power consumption. In our opinion, whenever existing data makes it possible, the use of this data-driven approaches could provide satisfactory results. We must consider that, as pointed out in [107], conventional energy models are mostly built under the design conditions. When a belt conveyor operates away from its design condition, inevitably, these models will result in large differences of energy calculation. In practice, most belt conveyors are not working under the design conditions and some of them are working far away from their design conditions. We also identified other methods, see [111] and [112]; however, these seems to be very related to specific sets of characteristics of the systems, for instance, the type of electric motor driving the system.

Ship Loader/Unloader-Stacker/Reclaimer

Information about typical power consumption of these assets was not openly available. In this case, the estimation can be primarily based on the information provided by the manufacturers. However, we must understand that the information provided is related to specific operating conditions that, most certainly, will be not the real operating conditions at the terminal. To overcome this uncertainty, a data-driven approach, based on registering historical energy consumption of these assets under different operating conditions, could be used. Similarly, to the approaches followed for conveyor systems, a set of factors impacting the energy consumption could be first identified and then, a mathematical model could be obtained based on a parameter estimation process (NLR, GEP). Among the parameters that could affect the energy consumption of these assets could be: type of commodity, total cargo weight, vessel's DWT, loading/unloading rate, etc.

Energy demand of Terminal Oriented activities in Liquid Bulk Terminals

In the case of liquid bulk terminals, information on energy consumption was also not so openly available. These types of terminals are, in many cases, owned by private companies with a high demand for some commodities, such as refineries or chemical industries. For the loading/unloading of liquid commodities, the main assets involved are loading arms (for fuels and some chemicals, for instance acrylonitrile, ammonia) and hoses (for other specific chemicals, such as caustic soda or bituminous asphalt). Pump stations and compressors consume energy when moving the liquid between the ships and the storage tanks through dedicated pipelines. Finally, depending on the additional services offered by the liquid terminals, additional energy consumption is required. Among the services that could also provided in liquid terminal we mention:

- Blending;
- Cooling;
- Electrical heating;
- Nitrogen blanketing;
- Steam heating;
- Warm Nitrogen Purging;
- Liquid and Gas transshipment;
- Gassing-up and degassing;
- Surveyor on-site (with fully equipped lab);
- Product treatment;
- Waste water and slops treatment.

This diversity of services makes very complex the analytical modelling of the total energy demand on these terminals. We present next some of the efforts made in this area that were identified during our literature review.

LNG Terminals

During our literature review, most of the information found concerns LNG terminals. These are very complex systems, and the total power consumption will depend on the layout of the terminals, the specific operating parameters and operation mode: holding, unloading and reloading [113]. In Figure 53 we can observe a typical LNG terminal, according to [114]. This layout corresponds to a one-stage recondensation system (among the simplest structures): boil-off gas (BOG) produced in the LNG storage tank is pressurized in the compressor and enters the recondenser. LNG flows into the recondenser after being pressurized by low-pressure LNG pump. LNG and BOG undergo contact heat exchange in the recondenser, and BOG is completely condensed by subcooled LNG. LNG at the outlet of the recondenser is directly pressurized to the pressure of the pipe network by the LNG high-pressure pump. Finally, the LNG is heated by seawater to a specified temperature and delivered to the user [115].

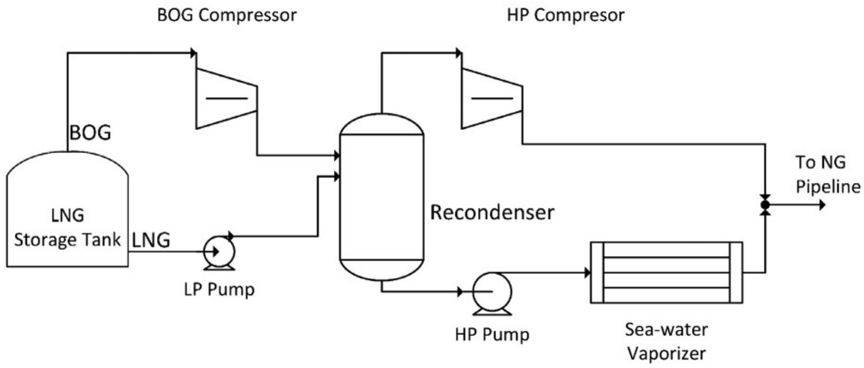

Figure 53. Schematic of a typical LNG regasification terminal [115].

As we can observe, the main power consumption comes from the pumps and compressors, where the sea water pump (to feed the sea-water vaporizer) should also be considered. To estimate the energy consumption, several operating parameters should be considered at once and the first law of thermodynamics can be used [115], [116], [117]:

$$\begin{aligned} W_{sea} &= m_{sea}(h_{sea,out} - h_{sea,in})/\eta_{sea}, \\ W_{LNG} &= m_{LNG}(h_{LNG,out} - h_{LNG,in})/\eta_{LNG}, \\ W_{com} &= m_{com}(h_{com,out} - h_{com,in})/\eta_{com}, \end{aligned} \tag{88}$$

where in the above formulas, m represents the mass flow of the stream $[kg/s]$, h represents the specific enthalpy of the state point $[kJ/kg]$, η represents the isentropic efficiency of the component. The sub-indices stand for the sea pump, the LNG pumps and the compressors. The power consumption of the overall LNG receiving station is then obtained from:

$$W_{PC} = W_{sea} + \sum W_{LNG} + \sum W_{com}, \tag{89}$$

therefore, the electrical demand can be obtained by introducing the efficiency of each component. However, although this is a straightforward approach, the complexity and diversity of terminal's layout complicate the estimation process. We must understand that a problem arising during the entire LNG supply chain is the generation of BOG; and in the specific case of the import/export terminals, the alternative selected for the terminals to manage the BOG, determines the layout of such a terminal and of course, the energy consumption. As an illustrative example, in Figure 54 we show the four alternatives analysed in [115] for the BOG recondensation process in an import terminal.

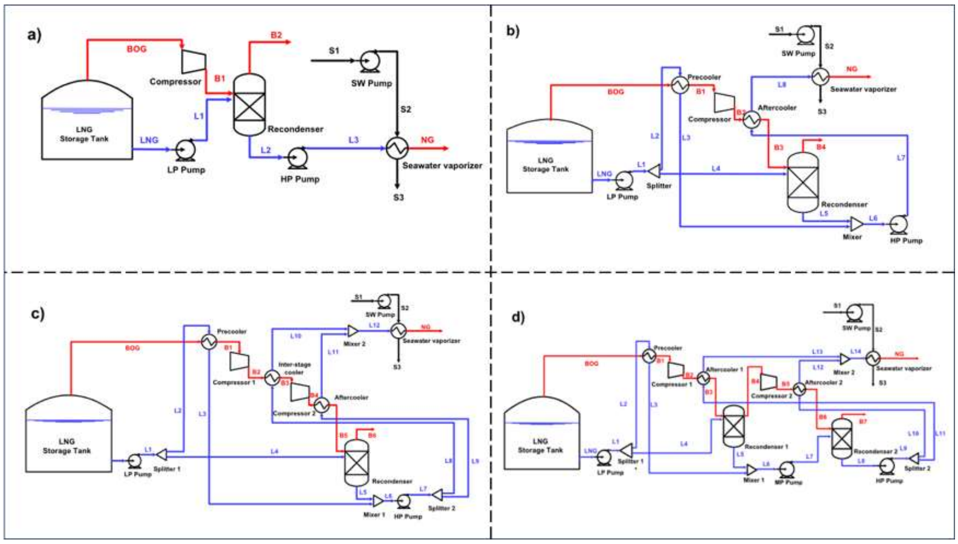

Figure 54. The schematic diagram of BOG recondenser. Adapted from [115].

In this case, the four analysed alternatives are: a) one-stage recondensation system; b) one-stage recondensation system with pre-cooling and after-cooling; c) one-stage recondensation system with two-stage compression; and d) two-stage recondensation system. In here we can observe how different equipment is used on those cases, making the energy estimation process more complicated. But this could become more complex when other BOG management alternatives are included, such as integrating BOG recondensation and LNG cold energy power generation system or combining BOG fuelled gas turbine and LNG cold energy power generation system. In these cases, additional equipment consuming or generating electricity are included, such as turbines and air compressors, see for instance [117].

We consider that, as in the case of bulk terminals, data-driven approaches could be useful to obtaining energy demand estimates. The important factors affecting the energy consumption on a specific terminal (depending on the layout could be the inlet/outlet fluids' pressure, temperature, and flow rates) could be first identified and then, parametric estimation methods could be used for obtaining models able to capture this relationship. We consider that these models should also consider the terminal mode, i.e., holding status or handling a tanker.

ANNEX 3

This Annex describes in more detail the different data-based modelling approaches that are briefly mentioned in section 2.6. Additionally, further justification is provided for some of the methodological choices made for the renewable generation forecasting models.

Table 23 summarizes the white-box and black-box approaches to RE forecasting. Further details may be found in [84], [118]. Grey-box models leverage on the benefits of both white- and black-box models, combining the physical grounding and interpretability of the first with the flexibility and practicality of the latter.

Modellling approach	Physical	Statistical	
		Time-series	Regression
Forecast horizon	Short- and medium-term	Very short-term	Short- and medium-term
Data requirements	NWP forecasts Detailed characterization of generation unit for downscaling of NWP forecasts and design of weather-to-power module	Historical data on power output and other explanatory variables (e.g., wind speed, irradiation)	NWP forecasts Historical data on power output, NWP forecasts and other explanatory variables (e.g., wind speed, irradiation)
Strengths	Historical data is not required	Simple formulation and low computational cost Specificities of generation unit are inherently captured by historic data Competitive for very short- and term forecasting due to inertia of atmospheric processes	Specificities of generation unit are inherently captured by historic data Modelling of complex relationships between input and output data Competitive for short- and medium-term forecasting due to inclusion of NWP data Suitable for large volumes of historic data
Limitations	Difficulty in acquiring the necessary physical data Resolution of the physical input data limits model's accuracy Not competitive for very short-term forecasting due to high temporal resolution required for NWP data	Historic data required for model training Not competitive for longer forecast horizons due to limitations in modelling non-linear and stochastic nature of renewable resources across longer time scales	Historic data required for model training Higher computational cost for large training datasets Not competitive for very short-term forecasting due to high temporal resolution required for NWP data

Models	Prediktor Previento LocalPred eWind EPREV	Auto-Regressive (AR) Auto-Regressive Moving Average (ARMA) Auto-Regressive Integrated Moving Average (ARIMA) Auto-Regressive Moving-Average with eXogenous inputs (ARMAX) Box-Jenkins	Artificial Neural Networks (ANN) Fuzzy Logic Other machine learning techniques

Table 23 - Summary of white- and black-box approaches to RE forecasting.

The rationale for deciding to opt for a white-box, grey-box, or black-box model is multifaceted, and should account for factors such as data availability, accuracy vs. interpretability, integration and adaptability, and prediction horizon, for example, besides the objectives of the energy management system [119].

Concerning data availability, white-box models, grounded in physical principles, are favoured when historical production data is limited but there is a robust understanding of the system's physics. Conversely, black-box models thrive on abundant historical data, extracting patterns without predefined physical constraints. Grey-box models serve as the intermediary, blending both realms—ideal when there is moderate production data and a desire for physical understanding. Concerning accuracy vs. interpretability, white-box models offer insights into production factors due to their physics-based nature. Black-box models, like deep neural networks, might prioritize accuracy over interpretability. Grey-box models attempt to strike a balance. Regarding integration and adaptability, the model's role within a smart energy management system is crucial. For real-time, rapid forecasts, a computationally efficient black-box model might be preferable. But for insights and adaptability rooted in physics, white or grey-box models could be more fitting. Finally, concerning prediction horizons, short /medium-term horizons (24-48 hours) can experience swift weather shifts, making black-box models potentially more adept at capturing these nuances.

Moreover, forecasting renewable generation requires not only estimating the point forecast (i.e., the amount of wind or solar power expected sometime in the future) but also the expected uncertainty (i.e., the likelihood that the real power output deviates from the point forecast). Uncertainty forecasts may be produced through ensembles of deterministic models (either white-, grey-, or black-box). In this context, different NWP datasets are fed as input to the forecast model to generate individual RE forecasts, which are then combined into a single forecast in an additional modelling stage. Alternatively, the ensemble forecasts may be generated by running different models on the same NWP data. When using a grey- or black-box approach, it is also possible to conduct probabilistic forecasting, in which case the uncertainty forecasts are derived from a deterministic weather data input using a single forecast model, thus not requiring ensemble techniques. This increases the model's dependence on a particular set of weather data from which it inherits its underlying bias and limitations. However, it also allows for a very significant reduction of computational costs in data acquisition and processing and in model training and validation. More details on uncertainty assessment may be found in [120].

Contrary to simple point estimates, probabilistic forecasting provides additional information on the uncertainty of the estimate within the model itself – with this complementary information, it is possible to assess the risk associated with the point forecast. In the context of renewable generation forecast, probabilistic forecasting usually means estimating the probability density function (PDF) for future power output and may be achieved using either i) parametric or ii) non-parametric approaches. Parametric approaches assume a certain distribution for the forecast uncertainty (e.g., Gaussian or Beta distributions). However, it may be unreasonable to assume a given shape for the PDF as it may not follow a known distribution, or it may change over time. In this case, the use of a non-parametric model is more suitable. Table 24 summarizes the parametric and non-parametric approaches to probabilistic forecasting. More information on probabilistic forecasting is found in [78].

Modelling approach	Remarks	Models
Parametric	Assumption of distribution shape PDF described with a few parameters Competitive for very short-term forecasting	Homoscedastic time-series Heteroscedastic time-series Machine learning
Non-parametric	No assumption of distribution shape PDF described as a finite number of points (more computationally expensive) Captures stochastic nature and geographic variability of renewable resources Competitive for short- and medium-term forecasting	Quantile regression Kernel density estimation Ensemble forecasting Machine learning

Table 24 - Summary of parametric and non-parametric approaches to RE forecasting.

PV forecast modelling

As previously mentioned, the selection of the modelling approach will need to take into consideration the different data characteristics of individual ports under consideration. For this reason, this document reports different possible approaches, shown in *Table 25*.

Proposed approach	White-box	Grey-box	Black-box
Method	PV Performance Modelling Approach [121]	PV Performance Modelling Approach + error consideration	Long Short Term Memory model in a NARX fashion [122], [123]
Characteristics / Reasoning	potent candidate for two primary purposes: to be used as Direct Forecasting by integrating the model directly with weather forecasts to predict PV output for the desired horizon, and as Dataset Generation for Black-Box Models, by employing the white-box model to curate a comprehensive dataset for black-box models. it is rooted in PV operation physics. It is essentially very adaptable, as the model can be tailored to specific PV system characteristics and locations. PV	considers the physical foundation, usually harnessing a white-box model that ensures that forecasts are rooted in the inherent physics of PV systems, thereby understanding how climatic variables impact PV output. Then, considers error compensation because it captures nuances not addressed by the physical model, potentially arising from non-linear behaviors, overlooked influencing factors, or anomalies. Also, if well developed, the model can seamlessly incorporate future data or models having good adaptability. It can	Utilizing Long Short-Term Memory (LSTM) networks in a Nonlinear AutoRegressive with eXogenous inputs (NARX) fashion is especially apt for PV supply prediction. LSTMs inherently excel at capturing long-term dependencies in time series data, making them adept at understanding patterns spanning different temporal scales in PV production. The NARX approach further enhances this by incorporating lagged values of both the target variable

	Performance Modelling Approach harnesses fundamental PV cell principles in tandem with meteorological insights to predict power yield. It considers Solar Irradiance to PV Current, Temperature Implications, Cell Temperature Estimation, Module/System Losses, Weather Data Integration (Feed the model with forecasted solar irradiation and temperature to anticipate PV production for the subsequent 24-48 hours).	account with a data optimization segment, leveraging the data at hand, and synthesizing detailed weather information with sparse PV production data, while demonstrating robustness by amalgamating the strengths of both modeling approaches, grey-box models can yield robust and precise predictions.	(PV production) and exogenous predictors (like weather conditions). This combination ensures that the model not only understands the sequential nature of PV production but also considers the impact of past weather conditions and PV outputs on future production, providing a robust and comprehensive forecasting framework.
Validation	demand distinct validation techniques. It should prioritize physical consistency, aligning closely with real-world values. Residual analysis is recommended, as it can offer insights into potential model refinements. Residual Analysis scrutinizes the residuals (differences between predictions and actuals), as any residual patterns could signify overlooked physical aspects or temporal influencers. Also, provides a comparison with Ground Truth, for a short prediction horizon (24-48 hours), historical instances under similar conditions can serve as	Validating a grey-box model necessitates a holistic approach, ensuring both its components (physical and empirical) harmoniously and accurately function. Residual Scrutiny can be used by analyzing residuals (forecasted versus actual outputs) so it can unveil patterns, ensuring the data-driven component effectively bridges the physical model's gaps. Then, Component-specific Validation can also be performed to independently validate both the physical and empirical components, ensuring the former abides by established principles and the latter effectively harnesses empirical patterns. Ground Truth Comparison is also an option to complement, due to the short 24-48	To validate the LSTM NARX-fashion PV supply forecast model, first, chronologically split the data into training, validation, and test sets. After each training epoch, evaluate the model on the validation set using metrics like MAE, RMSE, and MAPE to gauge prediction accuracy. Implement early stopping by monitoring the validation performance, halting training if the model starts to overfit. Analyze residuals to ensure no patterns are missed and compare the model's performance with simpler baseline models. Finally, after hyperparameter tuning and validation, assess the model on

	reference points to gauge model accuracy. Physical Consistency should also be considered, to ensure predictions align with known physical bounds, accounting for rapid meteorological shifts in short horizons. Sensitivity Analysis can also be considered during the validation process, as while typically beneficial, its relevance heightens with abrupt input parameter changes in short prediction windows.	hour forecasting horizon, prior similar instances can act as benchmarks to assess model fidelity. Especially critical for the empirical component, cross-validation also ensures robust generalizability across timeframes. To this end, predictions must always remain physically plausible, particularly vital for short horizons prone to abrupt variations, and in this specific segment of models, sensitivity analysis is essential for grey-box models to ascertain neither component disproportionately influences the overall forecast.	the test set for an unbiased performance evaluation.
Uncertainty assessment	Uncertainty in modeling is indispensable, especially for applications like PV forecasting where predictions inform critical decisions. A recommended strategy for this case is to prioritize parameter uncertainty in white-box models (utilize methods like bootstrapping to sample the dataset, fit the model, and gauge parameter variability for uncertainty quantification). Another approach may be model form uncertainty assessment, addressed through sensitivity analysis, which pertains to discrepancies between the chosen	Uncertainty assessment may be considered through possible strategies as combined uncertainty quantification and ensemble techniques (construct various grey-box models with minor variations, using the ensemble to gauge uncertainty based on prediction variances). Given the hybrid nature of grey-box models, ensemble techniques are especially apt. By generating a spectrum of potential model realities, it can create a prediction distribution for each timestep. This distribution's spread then offers an uncertainty estimate.	Monte Carlo Dropout can be used, as well as Bayesian LSTMs, quantile regression and ensemble of LSTMs.

	model and the actual physical process.		

Table 25 – Summary of the proposed approaches.

Bibliography

[1] W. Wang, Y. Peng, X. Li, Q. Qi, P. Feng, and Y. Zhang, "A two-stage framework for the optimal design of a hybrid renewable energy system for port application," *Ocean Engineering*, vol. 191, p. 106555, Nov. 2019, doi: 10.1016/j.oceaneng.2019.106555.

[2] N. N. Bakar *et al.*, "Optimal Configuration and Sizing of Seaport Microgrids including Renewable Energy and Cold Ironing—The Port of Aalborg Case Study," *Energies*, vol. 15, no. 2, Art. no. 2, 2022, doi: 10.3390/en15020431.

[3] M. A. Budiyanto, Nasruddin, and F. Zhafari, "Simulation study using building-design energy analysis to estimate energy consumption of refrigerated container," *Energy Procedia*, vol. 156, pp. 207–211, Jan. 2019, doi: 10.1016/j.egypro.2018.11.129.

[4] Ç. Iris and J. S. L. Lam, "A review of energy efficiency in ports: Operational strategies, technologies and energy management systems," *Renewable and Sustainable Energy Reviews*, vol. 112, pp. 170–182, Sep. 2019, doi: 10.1016/j.rser.2019.04.069.

[5] T. Spengler and G. Wilmsmeier, "Chapter 7 - Sustainable Performance and Benchmarking in Container Terminals—The Energy Dimension," in *Green Ports*, R. Bergqvist and J. Monios, Eds., Elsevier, 2019, pp. 125–154. doi: 10.1016/B978-0-12-814054-3.00007-4.

[6] G. Villalba and E. D. Gemechu, "Estimating GHG emissions of marine ports—the case of Barcelona," *Energy Policy*, vol. 39, no. 3, Art. no. 3, Mar. 2011, doi: 10.1016/j.enpol.2010.12.008.

[7] K.-I. Kim and K. M. Lee, "Dynamic Programming-Based Vessel Speed Adjustment for Energy Saving and Emission Reduction," *Energies*, vol. 11, no. 5, Art. no. 5, 2018, doi: 10.3390/en11051273.

[8] A. Misra, K. Panchabikesan, S. K. Gowrishankar, E. Ayyasamy, and V. Ramalingam, "GHG emission accounting and mitigation strategies to reduce the carbon footprint in conventional port activities – a case of the Port of Chennai," *Carbon Management*, vol. 8, no. 1, Art. no. 1, Jan. 2017, doi: 10.1080/17583004.2016.1275815.

[9] Y. Peng, X. Li, W. Wang, K. Liu, X. Bing, and X. Song, "A Method for Determining the Required Power Capacity of an On-Shore Power System Considering Uncertainties of Arriving Ships," *Sustainability*, vol. 10, no. 12, Art. no. 12, 2018, doi: 10.3390/su10124524.

[10] R. A. O. Nunes, M. C. M. Alvim-Ferraz, F. G. Martins, and S. I. V. Sousa, "Assessment of shipping emissions on four ports of Portugal," *Environmental Pollution*, vol. 231, pp. 1370–1379, Dec. 2017, doi: 10.1016/j.envpol.2017.08.112.

[11] Y. Peng, H. Liu, X. Li, J. Huang, and W. Wang, "Machine learning method for energy consumption prediction of ships in port considering green ports," *Journal of Cleaner Production*, vol. 264, p. 121564, Aug. 2020, doi: 10.1016/j.jclepro.2020.121564.

[12] Z. Wan *et al.*, "Evaluation of emission reduction strategies for berthing containerships: A case study of the Shekou Container Terminal," *Journal of Cleaner Production*, vol. 299, p. 126820, May 2021, doi: 10.1016/j.jclepro.2021.126820.

[13] D. Chen *et al.*, "Estimating ship emissions based on AIS data for port of Tianjin, China," *Atmospheric Environment*, vol. 145, pp. 10–18, Nov. 2016, doi: 10.1016/j.atmosenv.2016.08.086.

[14] N. B. Binti Ahamad, J. M. Guerrero, C.-L. Su, J. C. Vasquez, and X. Zhaoxia, "Microgrids Technologies in Future Seaports," in *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, Jun. 2018, pp. 1–6. doi: 10.1109/EEEIC.2018.8494428.

[15] *Regulation on the use of renewable and low carbon fuels in maritime transport*. 2023. [Online]. Available: https://data.consilium.europa.eu/doc/document/PE-26-2023-INIT/en/pdf

[16] P. E. N. G. Yun, L. I. Xiangda, W. A. N. G. Wenyuan, L. I. U. Ke, and L. I. Chuan, "A simulation-based research on carbon emission mitigation strategies for green container terminals," *Ocean Engineering*, vol. 163, pp. 288–298, Sep. 2018, doi: 10.1016/j.oceaneng.2018.05.054.

[17] EPA, "Shore Power Technology Assessment at U.S. Ports. 2022 Update," EPA-420-R-22-037, 2022. Accessed: Sep. 06, 2023. [Online]. Available: https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P1016C86.pdf
[18] T. G. Papoutsoglou, "A cold ironing study on modern ports, implementation and benefits thriving for worldwide ports," National Technical University of Athens, 2012. [Online]. Available: http://dx.doi.org/10.26240/heal.ntua.12134
[19] F. D'Agostino, G. P. Schiapparelli, S. Dallas, D. Spathis, V. Georgiou, and J. Prousalidis, "On Estimating the Port Power Demands for Cold Ironing Applications," in *2021 IEEE Electric Ship Technologies Symposium (ESTS)*, Aug. 2021, pp. 1–5. doi: 10.1109/ESTS49166.2021.9512359.
[20] F. D'Agostino, D. Kaza, G. P. Schiapparelli, F. Silvestro, C. L. Bossi, and F. Colzi, "Assessment of the Potential Shore to Ship Load Demand: the Italian Scenario," in *2021 IEEE Power & Energy Society General Meeting (PESGM)*, Jul. 2021, pp. 1–5. doi: 10.1109/PESGM46819.2021.9638000.
[21] J. He, Y. Huang, and W. Yan, "Yard crane scheduling in a container terminal for the trade-off between efficiency and energy consumption," *Advanced Engineering Informatics*, vol. 29, no. 1, Art. no. 1, Jan. 2015, doi: 10.1016/j.aei.2014.09.003.
[22] H. Geerlings, R. Heij, and R. van Duin, "Opportunities for peak shaving the energy demand of ship-to-shore quay cranes at container terminals," *Journal of Shipping and Trade*, vol. 3, no. 1, Art. no. 1, Feb. 2018, doi: 10.1186/s41072-018-0029-y.
[23] M. Kermani, G. Parise, B. Chavdarian, and L. Martirano, "Ultracapacitors for Port Crane Applications: Sizing and Techno-Economic Analysis," *Energies*, vol. 13, no. 8, Art. no. 8, 2020, doi: 10.3390/en13082091.
[24] D. Chen, W. Niu, W. Gu, and N. Schofield, "Game-Based Energy Management Method for Hybrid RTG Cranes," *Energies*, vol. 12, no. 18, Art. no. 18, Jan. 2019, doi: 10.3390/en12183589.
[25] Y. Zhang, Y. Peng, W. Wang, J. Gu, X. Wu, and X. Feng, "Air emission inventory of container ports' cargo handling equipment with activity-based 'bottom-up' method," *Advances in Mechanical Engineering*, vol. 9, no. 7, Art. no. 7, Jul. 2017, doi: 10.1177/1687814017711389.
[26] Y. Mamatok and C. Jin, "An integrated framework for carbon footprinting at container seaports: the case study of a Chinese port," *Maritime Policy & Management*, vol. 44, no. 2, Art. no. 2, Feb. 2017, doi: 10.1080/03088839.2016.1262077.
[27] J. He, Y. Huang, W. Yan, and S. Wang, "Integrated internal truck, yard crane and quay crane scheduling in a container terminal considering energy consumption," *Expert Systems with Applications*, vol. 42, no. 5, Art. no. 5, Apr. 2015, doi: 10.1016/j.eswa.2014.11.016.
[28] J. He, W. Zhang, Y. Huang, and W. Yan, "A simulation optimization method for internal trucks sharing assignment among multiple container terminals," *Advanced Engineering Informatics*, vol. 27, no. 4, Art. no. 4, Oct. 2013, doi: 10.1016/j.aei.2013.08.001.
[29] O. Okşaş, "Carbon emission strategies for container handling equipment using the activity-based method: A case study of Ambarlı container port in Turkiye," *Marine Policy*, vol. 149, p. 105480, Mar. 2023, doi: 10.1016/j.marpol.2023.105480.
[30] Starcrest Consulting Group, "San Pedro Bay Ports Emissions Inventory Methodology Report," Version 3a 2022, 2022.
[31] K. K. Sørensen, J. Stoustrup, and T. Bak, "Adaptive MPC for a reefer container," *Control Engineering Practice*, vol. 44, pp. 55–64, Nov. 2015, doi: 10.1016/j.conengprac.2015.05.012.
[32] W. B. Fitzgerald, O. J. A. Howitt, I. J. Smith, and A. Hume, "Energy use of integral refrigerated containers in maritime transportation," *Energy Policy*, vol. 39, no. 4, Art. no. 4, Apr. 2011, doi: 10.1016/j.enpol.2010.12.015.
[33] K. K. Sørensen, J. D. Nielsen, and J. Stoustrup, "Modular simulation of reefer container dynamics," *SIMULATION*, vol. 90, no. 3, pp. 249–264, Mar. 2014, doi: 10.1177/0037549713515542.
[34] Gesamtverband der Deutschen Versicherungswirtschaft E.V., *Container handbook*. 2017.
[35] M. A. Budiyanto and T. Shinoda, "The effect of solar radiation on the energy consumption of refrigerated container," *Case Studies in Thermal Engineering*, vol. 12, pp. 687–695, Sep. 2018, doi: 10.1016/j.csite.2018.09.005.

[36] Y. Lu, S. Fang, G. Chen, T. Niu, and R. Liao, "Cyber-Physical Integration for Future Green Seaports: Challenges, State of the Art and Future Prospects," *IEEE Transactions on Industrial Cyber-Physical Systems*, vol. 1, pp. 21–43, 2023, doi: 10.1109/TICPS.2023.3283234.

[37] F. D. Kanellos, "Real-Time Control Based on Multi-Agent Systems for the Operation of Large Ports as Prosumer Microgrids," *IEEE Access*, vol. 5, pp. 9439–9452, 2017, doi: 10.1109/ACCESS.2017.2706091.

[38] F. D. Kanellos, E.-S. M. Volanis, and N. D. Hatziargyriou, "Power Management Method for Large Ports With Multi-Agent Systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1259–1268, Mar. 2019, doi: 10.1109/TSG.2017.2762001.

[39] T. K. Tran, "Study of electrical usage and demand at the container terminal," Ph.D., Deakin University, 2012. Accessed: Aug. 30, 2023. [Online]. Available: https://dro.deakin.edu.au/articles/thesis/Study_of_electrical_usage_and_demand_at_the_container_terminal/21101890/1

[40] J. Prousalidis *et al.*, "Optimizing the operation of port energy systems," in *2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, Jun. 2019, pp. 1–6. doi: 10.1109/EEEIC.2019.8783476.

[41] J. H. R. (Ron) van Duin, H. (Harry) Geerlings, A. (Alexander) Verbraeck, and T. (Tushar) Nafde, "Cooling down: A simulation approach to reduce energy peaks of reefers at terminals," *Journal of Cleaner Production*, vol. 193, pp. 72–86, Aug. 2018, doi: 10.1016/j.jclepro.2018.04.258.

[42] J. H. R. van Duin, H. Geerlings, L. A. Tavasszy, and D. L. Bank, "Factors causing peak energy consumption of reefers at container terminals," *Journal of Shipping and Trade*, vol. 4, no. 1, p. 1, Apr. 2019, doi: 10.1186/s41072-019-0040-y.

[43] Z. O'Neill and S. Narayanan, "Model-based estimation of cold room temperatures in a supermarket refrigeration system," *Applied Thermal Engineering*, vol. 73, no. 1, pp. 819–830, Dec. 2014, doi: 10.1016/j.applthermaleng.2014.08.038.

[44] L. Filina and S. Filin, "An analysis of influence of lack of the electricity supply to reefer containers serviced at sea ports on storing conditions of cargoes contained in them," *Polish Maritime Research*, vol. 15, no. 4, pp. 96–102, 2009, doi: 10.2478/v10012-008-0002-z.

[45] K. K. Sørensen, M. J. Skovrup, L. M. Jessen, and J. Stoustrup, "Modular modeling of a refrigeration container," *International Journal of Refrigeration*, vol. 55, pp. 17–29, Jul. 2015, doi: 10.1016/j.ijrefrig.2015.03.017.

[46] R. Pei, J. Xie, H. Zhang, K. Sun, Z. Wu, and S. Zhou, "Robust Multi-Layer Energy Management and Control Methodologies for Reefer Container Park in Port Terminal," *Energies*, vol. 14, no. 15, 2021, doi: 10.3390/en14154456.

[47] CITEPA, "Gaz à effet de Serre et Polluants Atmosphériques. Bilan des émissions en France de 1990 à 2018," France, 2020.

[48] M. Acciaro, J. M. Vassallo, G. Dalton, and E. de la Peña, "Transport infrastructure," Publications Office of the European Union, Belgium, 2017. Accessed: Nov. 27, 2023. [Online]. Available: https://data.europa.eu/doi/10.2777/76533

[49] P. G. Höglund and J. Niittymäki, "Estimating vehicle emissions and air pollution related to driving patterns and traffic calming," presented at the Conference Urban Transport Systems, Sweden, 1999. Accessed: Nov. 27, 2023. [Online]. Available: https://trid.trb.org/view/715493

[50] J. W. C. van Lint and N. J. van der Zijpp, "Improving a Travel-Time Estimation Algorithm by Using Dual Loop Detectors," *Transportation Research Record*, vol. 1855, no. 1, pp. 41–48, Jan. 2003, doi: 10.3141/1855-05.

[51] M. Treiber and D. Helbing, "An Adaptive Smoothing Method for Traffic State Identification from Incomplete Information," in *Interface and Transport Dynamics*, H. Emmerich, B. Nestler, and M. Schreckenberg, Eds., in Lecture Notes in Computational Science and Engineering. Berlin, Heidelberg: Springer, 2003, pp. 343–360. doi: 10.1007/978-3-662-07969-0_33.

[52] L. Chu, J.-S. Oh, and W. Recker, "Adaptive Kalman filter based freeway travel time estimation," presented at the 84th TRB Annual Meeting, Washington, USA, Jan. 2005.

[53] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 606–616, Aug. 2011, doi: 10.1016/j.trc.2010.10.002.
[54] G. F. Newell, "A simplified theory of kinematic waves in highway traffic, part I: General theory," *Transportation Research Part B: Methodological*, vol. 27, no. 4, pp. 281–287, Aug. 1993, doi: 10.1016/0191-2615(93)90038-C.
[55] S. Smulders, "Control of freeway traffic flow by variable speed signs," *Transportation Research Part B: Methodological*, vol. 24, no. 2, pp. 111–132, Apr. 1990, doi: 10.1016/0191-2615(90)90023-R.
[56] M. Fellendorf and P. Vortisch, "Microscopic Traffic Flow Simulator VISSIM," in *Fundamentals of Traffic Simulation*, J. Barceló, Ed., in International Series in Operations Research & Management Science. , New York, NY: Springer, 2010, pp. 63–93. doi: 10.1007/978-1-4419-6142-6_2.
[57] Q. Yang and H. N. Koutsopoulos, "A Microscopic Traffic Simulator for evaluation of dynamic traffic management systems," *Transportation Research Part C: Emerging Technologies*, vol. 4, no. 3, pp. 113–129, Jun. 1996, doi: 10.1016/S0968-090X(96)00006-X.
[58] R. C. Effa and L. Larsen, "Development of real-world driving cycles for estimating facility-specific emissions from light-duty vehicles.," presented at the Air and Waste Management Association Annual Specialty Conference Emission Inventory: Perception and Reality, California, USA, 1993.
[59] J. Lin and D. A. Niemeier, "Estimating Regional Air Quality Vehicle Emission Inventories: Constructing Robust Driving Cycles," *Transportation Science*, vol. 37, no. 3, pp. 330–346, 2003.
[60] L. Pelkmans and P. Debal, "Comparison of on-road emissions with emissions measured on chassis dynamometer test cycles," *Transportation Research Part D: Transport and Environment*, vol. 11, no. 4, pp. 233–241, Jul. 2006, doi: 10.1016/j.trd.2006.04.001.
[61] L. A. G. Breton, "Real-time on-road vehicle exhaust gas modular flowmeter and emissions reporting system," US6148656A, Nov. 21, 2000 Accessed: Nov. 27, 2023. [Online]. Available: https://patents.google.com/patent/US6148656A/en
[62] J. Gao, G. Tian, A. Sorniotti, A. E. Karci, and R. Di Palo, "Review of thermal management of catalytic converters to decrease engine emissions during cold start and warm up," *Applied Thermal Engineering*, vol. 147, pp. 177–187, Jan. 2019, doi: 10.1016/j.applthermaleng.2018.10.037.
[63] J. Gao, H. Chen, K. Dave, J. Chen, and D. Jia, "Fuel economy and exhaust emissions of a diesel vehicle under real traffic conditions," *Energy Science & Engineering*, vol. 8, no. 5, pp. 1781–1792, 2020, doi: 10.1002/ese3.632.
[64] H. Liu, X. Chen, Y. Wang, and S. Han, "Vehicle Emission and Near-Road Air Quality Modeling for Shanghai, China: Based on Global Positioning System Data from Taxis and Revised MOVES Emission Inventory," *Transportation Research Record*, vol. 2340, no. 1, pp. 38–48, Jan. 2013, doi: 10.3141/2340-05.
[65] K. M. Sentoff, L. Aultman-Hall, and B. A. Holmén, "Implications of driving style and road grade for accurate vehicle activity data and emissions estimates," *Transportation Research Part D: Transport and Environment*, vol. 35, pp. 175–188, Mar. 2015, doi: 10.1016/j.trd.2014.11.021.
[66] X. Ma, W. Lei, I. Andréasson, and H. Chen, "An Evaluation of Microscopic Emission Models for Traffic Pollution Simulation Using On-board Measurement," *Environ Model Assess*, vol. 17, no. 4, pp. 375–387, Aug. 2012, doi: 10.1007/s10666-011-9296-9.
[67] K. Chen and L. Yu, "Microscopic Traffic-Emission Simulation and Case Study for Evaluation of Traffic Control Strategies," *Journal of Transportation Systems Engineering and Information Technology*, vol. 7, no. 1, pp. 93–99, Feb. 2007, doi: 10.1016/S1570-6672(07)60011-7.
[68] A.-L. Badin *et al.*, "Combining microscopic traffic modelling and 3 pollutant emission modellings to assess modifications of traffic supply and demand," *Journal of Earth Sciences and Geotechnical Engineering*, vol. 7, no. 1, pp. 175–190, 2017.
[69] M. Laraki, G. De Nunzio, and B. Othman, "A Large-Scale and Data-Based Road Traffic Flow Estimation Method Leveraging Topography Information and Population

Statistics," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, Oct. 2022, pp. 2343–2349. doi: 10.1109/ITSC55140.2022.9921799.
[70] O. Feldman, "The GEH measure and quality of the highway assignment models," presented at the ETC Conference Papers 2012, Association for European Transport and Contributors, 2012. [Online]. Available: https://aetransport.org/public/downloads/V7AGa/5664-5218a2370407f.pdf
[71] P. Richards, "Visum Sub-Area Validation and Calibration Report. Suburban Rail Loop East," Australia, SRL-AJM-RAI-PWD-REP-XTE-NAP-0001503, 2022. [Online]. Available: https://bigbuild.vic.gov.au/__data/assets/pdf_file/0018/641340/SRL-TN01-Att-N-Visum-Sub-Area-Validation-and-Calibration-Report_compressed.pdf
[72] CITEPA, "Rapport OMINEA – 16ème édition," 2019.
[73] CITEPA, "Inventaire des émissions de polluants atmosphériques et de gaz à effet de serre en France – Format Secten," 2021.
[74] M. Parvania, M. Fotuhi-Firuzabad, and M. Shahidehpour, "Optimal Demand Response Aggregation in Wholesale Electricity Markets," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 1957–1965, Dec. 2013, doi: 10.1109/TSG.2013.2257894.
[75] D. Steenken, S. Voß, and R. Stahlbock, "Container terminal operation and operations research - a classification and literature review," *OR Spectrum*, vol. 26, no. 1, pp. 3–49, Jan. 2004, doi: 10.1007/s00291-003-0157-z.
[76] S. G. Gennitsaris and F. D. Kanellos, "Emission-Aware and Cost-Effective Distributed Demand Response System for Extensively Electrified Large Ports," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4341–4351, Nov. 2019, doi: 10.1109/TPWRS.2019.2919949.
[77] S. Fang and H. Wang, *Optimization-Based Energy Management for Multi-energy Maritime Grid*. in Springer Series on Naval Architecture, Marine Engineering, Shipbuilding and Shipping. Springer Singapore, 2021. [Online]. Available: https://doi.org/10.1007/978-981-33-6734-0
[78] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, Apr. 2014, doi: 10.1016/j.rser.2014.01.033.
[79] C. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*. in Information Science and Statistics. New York, USA: Springer, 2006. Accessed: Nov. 28, 2023. [Online]. Available: https://link.springer.com/book/9780387310732
[80] A. F. Psaros, X. Meng, Z. Zou, L. Guo, and G. E. Karniadakis, "Uncertainty quantification in scientific machine learning: Methods, metrics, and comparisons," *Journal of Computational Physics*, vol. 477, p. 111902, Mar. 2023, doi: 10.1016/j.jcp.2022.111902.
[81] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*. MIT Press, 2018. Accessed: Nov. 28, 2023. [Online]. Available: https://mitpress.mit.edu/9780262039406/foundations-of-machine-learning/
[82] J.-P. Lai, Y.-M. Chang, C.-H. Chen, and P.-F. Pai, "A Survey of Machine Learning Models in Renewable Energy Predictions," *Applied Sciences*, vol. 10, no. 17, Art. no. 17, Jan. 2020, doi: 10.3390/app10175975.
[83] F. Hartwig and B. E. Dearling, *Exploratory Data Analysis*. in Quantitative Applications in the Social Sciences. SAGE Publications Inc., 1979. doi: 10.4135/9781412984232.
[84] J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting," *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 762–777, Mar. 2014, doi: 10.1016/j.rser.2013.12.054.
[85] K. Bot, A. Ruano, and M. da G. Ruano, "Short-Term Forecasting Photovoltaic Solar Power for Home Energy Management Systems," *Inventions*, vol. 6, no. 1, Art. no. 1, Mar. 2021, doi: 10.3390/inventions6010012.
[86] A. Y. Appiah, X. Zhang, B. B. K. Ayawli, and F. Kyeremeh, "Long Short-Term Memory Networks Based Automatic Feature Extraction for Photovoltaic Array Fault Diagnosis," *IEEE Access*, vol. 7, pp. 30089–30101, 2019, doi: 10.1109/ACCESS.2019.2902949.
[87] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[88] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," in *Proceedings of The 33rd International Conference on Machine Learning*, PMLR, Jun. 2016, pp. 1050–1059. Accessed: Nov. 28, 2023. [Online]. Available: https://proceedings.mlr.press/v48/gal16.html
[89] J. H. Friedman, "Greedy function approximation: A gradient boosting machine.," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, doi: 10.1214/aos/1013203451.
[90] J. R. Andrade and R. J. Bessa, "Improving Renewable Energy Forecasting With a Grid of Numerical Weather Predictions," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1571–1580, Oct. 2017, doi: 10.1109/TSTE.2017.2694340.
[91] G. I. Nagy, G. Barta, S. Kazi, G. Borbély, and G. Simon, "GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1087–1093, Jul. 2016, doi: 10.1016/j.ijforecast.2015.11.013.
[92] "Weather Research & Forecasting Model (WRF) | Mesoscale & Microscale Meteorology Laboratory." Accessed: Sep. 11, 2023. [Online]. Available: https://www.mmm.ucar.edu/models/wrf
[93] E. B. Ssekulima, M. B. Anwar, A. Al Hinai, and M. S. El Moursi, "Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: a review," *IET Renewable Power Generation*, vol. 10, no. 7, pp. 885–989, 2016, doi: 10.1049/iet-rpg.2015.0477.
[94] J. E. Matheson and R. L. Winkler, "Scoring Rules for Continuous Probability Distributions," *Management Science*, vol. 22, no. 10, pp. 1087–1096, 1976.
[95] G. Parise, L. Parise, F. M. Pepe, S. Ricci, C. L. Su, and P. Chavdarian, "Innovations in a container terminal area and electrical power distribution for the service continuity," in *2016 IEEE/IAS 52nd Industrial and Commercial Power Systems Technical Conference (I&CPS)*, USA, May 2016, pp. 1–6. doi: 10.1109/ICPS.2016.7490238.
[96] N. N. Abu Bakar, N. Bazmohammadi, J. C. Vasquez, and J. M. Guerrero, "Electrification of onshore power systems in maritime transportation towards decarbonization of ports: A review of the cold ironing technology," *Renewable and Sustainable Energy Reviews*, vol. 178, p. 113243, May 2023, doi: 10.1016/j.rser.2023.113243.
[97] G. Parise, L. Parise, L. Martirano, P. B. Chavdarian, C. -L. Su, and A. Ferrante, "Wise Port and Business Energy Management: Port Facilities, Electrical Power Distribution," *IEEE Transactions on Industry Applications*, vol. 52, no. 1, Art. no. 1, Feb. 2016, doi: 10.1109/TIA.2015.2461176.
[98] "https://pypsa.org/."
[99] Magpie Smart Green Ports, "DELIVERABLE D3.1 TRANSPORT ENERGY REQUIREMENT," 2022. [Online]. Available: https://www.magpie-ports.eu/wp-content/uploads/2023/01/Deliverable-D3.1-Transport-energy-requirements.pdf
[100] Transport & Environment, "Easy Ride: why the EU truck CO2 targets are unfit for the 2020s," 2021. [Online]. Available: https://www.transportenvironment.org/wp-content/uploads/2021/10/202108_truck_CO2_report_final.pdf
[101] E. Mulholland, P.-L. Ragon, and F. Rodríguez, "CO2 emissions from trucks in the European Union: An analysis of the 2020 reporting period," International Council on Clean Transportation, WORKING PAPER 2023-14, 2023. [Online]. Available: https://theicct.org/wp-content/uploads/2023/07/hdv-co2-emissions-eu-2020-reporting-2-jul23.pdf
[102] À. Freixas and P. Mentink, "Aerodynamic and Flexible Trucks for Next Generation of Long Distance Road Transport," EUROPEAN COMMISSION Horizon 2020, AEROFLEX 6.3, 2019. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c412450e&appId=PPGMS
[103] "ISO 14083:2023." Jan. 20, 2023. Accessed: Nov. 28, 2023. [Online]. Available: https://www.iso.org/standard/78864.html
[104] A. Rondaij and J. S. Spreen, "Decamod: toolbox voor rekenen aan CO2-reductie in transport en logistiek," TNO, Netherland, TNO 2020 R11938, 2020. Accessed: Nov. 28, 2023. [Online]. Available: http://resolver.tudelft.nl/uuid:d14b3031-cab3-4a13-a0db-a970ef188143

[105] T. Notteboom, A. Pallis, and J.-P. Rodrigue, *Port Economics, Management and Policy*. London: Routledge, 2022. doi: 10.4324/9780429318184.
[106] Y. Mu *et al.*, "Optimal scheduling method for belt conveyor system in coal mine considering silo virtual energy storage," *Applied Energy*, vol. 275, p. 115368, Oct. 2020, doi: 10.1016/j.apenergy.2020.115368.
[107] S. Zhang and X. Xia, "Modeling and energy efficiency optimization of belt conveyors," *Applied Energy*, vol. 88, no. 9, pp. 3061–3071, Sep. 2011, doi: 10.1016/j.apenergy.2011.03.015.
[108] S. Zhang and X. Xia, "Optimal control of operation efficiency of belt conveyor systems," *Applied Energy*, vol. 87, no. 6, pp. 1929–1937, Jun. 2010, doi: 10.1016/j.apenergy.2010.01.006.
[109] J. Hiltermann, G. Lodewijks, D. L. Schott, J. C. Rijsenbrij, J. A. J. M. Dekkers, and Y. Pang, "A Methodology to Predict Power Savings of Troughed Belt Conveyors by Speed Control," *Particulate Science and Technology*, vol. 29, no. 1, pp. 14–27, Jan. 2011, doi: 10.1080/02726351.2010.491105.
[110] E. Köken, A. I. Lawal, M. Onifade, and A. Özarslan, "A COMPARATIVE STUDY ON POWER CALCULATION METHODS FOR CONVEYOR BELTS IN MINING INDUSTRY," *International Journal of Mining, Reclamation and Environment*, vol. 36, no. 1, pp. 26–45, Jan. 2022, doi: 10.1080/17480930.2021.1949859.
[111] T. Mathaba and X. Xia, "Optimal and energy efficient operation of conveyor belt systems with downhill conveyors," *Energy Efficiency*, vol. 10, no. 2, pp. 405–417, Apr. 2017, doi: 10.1007/s12053-016-9461-8.
[112] C. Yang, L. Bu, and B. Chen, "Energy modeling and online parameter identification for permanent magnet synchronous motor driven belt conveyors," *Measurement*, vol. 178, p. 109342, Jun. 2021, doi: 10.1016/j.measurement.2021.109342.
[113] M. S. Khan, S. Effendy, I. A. Karimi, and A. Wazwaz, "Improving design and operation at LNG regasification terminals through a corrected storage tank model," *Applied Thermal Engineering*, vol. 149, pp. 344–353, Feb. 2019, doi: 10.1016/j.applthermaleng.2018.12.060.
[114] H. Nagesh Rao and I. A. Karimi, "Optimal design of boil-off gas reliquefaction process in LNG regasification terminals," *Computers & Chemical Engineering*, vol. 117, pp. 171–190, Sep. 2018, doi: 10.1016/j.compchemeng.2018.06.003.
[115] T. Yuan, C. Song, J. Bao, N. Zhang, X. Zhang, and G. He, "Minimizing power consumption of boil off gas (BOG) recondensation process by power generation using cold energy in liquefied natural gas (LNG) regasification process," *Journal of Cleaner Production*, vol. 238, p. 117949, Nov. 2019, doi: 10.1016/j.jclepro.2019.117949.
[116] S. Le, J.-Y. Lee, and C.-L. Chen, "Waste cold energy recovery from liquefied natural gas (LNG) regasification including pressure and thermal energy," *Energy*, vol. 152, pp. 770–787, Jun. 2018, doi: 10.1016/j.energy.2018.03.076.
[117] J. Bao, T. Yuan, L. Zhang, N. Zhang, and X. Zhang, "Comparative study of three boil-off gas treatment schemes: From an economic perspective," *Energy Conversion and Management*, vol. 201, p. 112185, Dec. 2019, doi: 10.1016/j.enconman.2019.112185.
[118] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. John Wiley & Sons, 2015.
[119] G. Allaire, *Numerical Analysis and Optimization: An Introduction to Mathematical Modelling and Numerical Simulation*. in Numerical Mathematics and Scientific Computation. Oxford, New York: Oxford University Press, 2007.
[120] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*, 3rd ed. Melbourne, Australia: OTexts, 2021. [Online]. Available: https://otexts.com/fpp3/
[121] A. Luque and S. Hegedus, Eds., *Handbook of Photovoltaic Science and Engineering*. John Wiley & Sons, Ltd, 2003. doi: 10.1002/0470014008.index.
[122] K. P. Murphy, *Machine Learning: a Probabilistic Perspective*. MIT Press, 2012.
[123] J. M. P. Menezes and G. A. Barreto, "Long-term time series prediction with the NARX network: An empirical evaluation," *Neurocomputing*, vol. 71, no. 16, pp. 3335–3343, Oct. 2008, doi: 10.1016/j.neucom.2008.01.030.