European Common Energy Data Space Framework Enabling Data Sharing - Driven Across – and Beyond – Energy Services

enershare.eu

Enershare has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

D6.1 Federated learning, data-driven services, data visualisation and Digital Twins

Alpha version

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Publication details

Grant Agreement Number	101069831
Acronym	ENERSHARE
Full Title	European Common Energy Data Space Framework Enabling Data Sharing-Driven Across — and Beyond — Energy Services
Topic	HORIZON-CL5-2021-D3-01-01 'Establish the grounds for a common European energy data space'
Funding scheme	HORIZON-IA: Innovation Action
Start Date	Jul 1, 2022
Duration	36 months
Project URL	enershare.eu
Project Coordinator	Engineering
Deliverable	D6.1 – Federated learning, data-driven services, data visualisation and Digital Twins (Alpha version)
Work Package	WP6 – Cross-value chain AI-based data-driven services
Delivery Month (DoA)	M12
Version	1.0
Actual Delivery Date	30 June, 2023
Nature	Report
Dissemination Level	PU
Lead Beneficiary	INESC TEC

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Authors	Ricardo Bessa (INESC TEC), José Villar (INESC TEC), Pedro Macedo (INESC TEC), Armando Moreno (INESC TEC), Ricardo Silva (INESC TEC), José Paulos (INESC TEC), Carla Gonçalves (INESC TEC), João Viana (INESC TEC), Ricardo Andrade (INESC TEC), Elena Lazovik (TNO) David Campo (FIWARE), Alberto Abella (FIWARE), Leonardo Carreras (RWTH), Oudom Kem (ENGIE), Weiqin XU (ENGIE), Andrej Čampa (COMS), Blaž Bertalanič (COMS), Ioannis Dimanidis (FORTUM), Rui Martins (SEL), Adelaide Ambrósio (SEL), Vassilis Sakas (ED), Apostolos Kapetanios (ED), Konstantinos Kotsalos (ED)
Quality Reviewer(s)	Paride D'Ostilio (ASM Terni) Clement Villeviere (TNO)
Keywords	Federated learning, data-driven energy services, cross-sector services, visual analytics, digital twin

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Document History

Ver.	Date	Description	Author	Partner
0.1	Feb 27, 2023	ToC	Ricardo Bessa	INESC TEC
0.2	Jun 07, 2023	Complete version with comments	Ricardo Bessa et al.	All WP6 partners
0.3	Jun 12, 2023	Revised version	Ricardo Bessa et al.	All WP6 partners
0.4	Jun 19, 2023	Final version for internal review	Ricardo Bessa et al.	All WP6 partners
1.0	Jun 30, 2023	Fina version after internal review	Ricardo Bessa et al.	All WP6 partners

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Table of Contents

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

List of Figures

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

List of Tables

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

List of Acronyms

Acronym	Description
AE	Autoencoder
AI	Artificial Intelligence
ALPG	Artificial Load Profile Generator
API	Application Programming Interface
BN	Bayesian Network
Cbc	Coin-or branch and cut
CEC	Citizens Energy Community
CM	Condition Monitoring
COP	Cross-sector Operators’ Portal
CPO	Charging Point Operators
DER	Distributed Energy Resources
DIHN	Directly Influenced Hypernodes
DNN	Deep Neural Network
DP	Differential Privacy
DSO	Distribution System Operator
DT	Digital Twin
EV	Electric Vehicle
EWH	Thermoelectric Eater Heaters
FAR	Flexibility Analytics and Register
FL	Federated Learning

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

IDS	International Data Space
IEGSA	Interoperable European Grid Services Architecture
KG	Knowledge Graph
KPI	Key Performance Indicators
LSTM	Long Short-term Memory
MILP	Mixed-Integer Linear Problem
ML	Machine Learning
MMR	Mid-market-rate
MQTT	MQ Telemetry Transport
O&M	Operation and Maintenance
P2G	Power-to-gas
P2P	Peer-to-peer
PaaS	Platform-as-a-Service
PMSG	Permanent Magnet Synchronous Generator
POOL	Pre- / Post-delivery pool
REC	Renewable Energy Community
RES	Renewable Energy Sources
RPC	Remote Procedure Call
SDR	Supply and demand ratio
SM	Smart Meters
SQL	Structured Query Language
TCN	Temporal Convolutional Networks
TRL	Technology Readiness Level

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

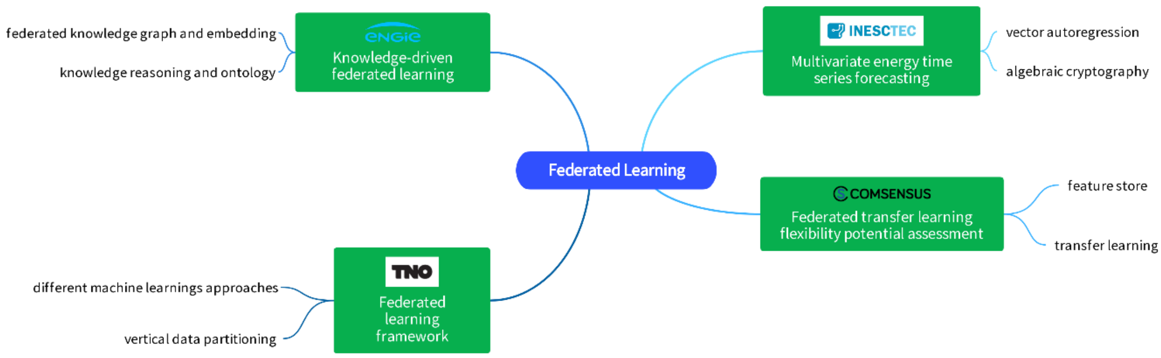
TSO	Transmission System Operator
VAR	Vector Autoregressive
WT	Wind Turbine

Executive summary

The ENERSHARE project aims to incorporate and customize the Data Space paradigm within the energy sector, bringing together data, storage, and computing infrastructures to create interoperable resources. By leveraging privacy-preserving federated learning, the project developed in WP6 data-driven services and System-of-System Digital Twin (DT) applications across multiple value chains. These cutting-edge services address various energy sector challenges, including load flexibility estimation and optimizing renewable energy sources. Additionally, the project enables cross-sector services, such as wellness alarms, energy poverty monitoring, and electric vehicle (EV) management. System-level DTs are developed to enhance interoperability and support complex energy planning and real-time operations. Overall, the project harnesses energy data and advanced technologies to drive efficiency, sustainability, and well-being across domains.

Federated learning methods

The following federated learning (FL) methods were implemented in ENERSHARE project:



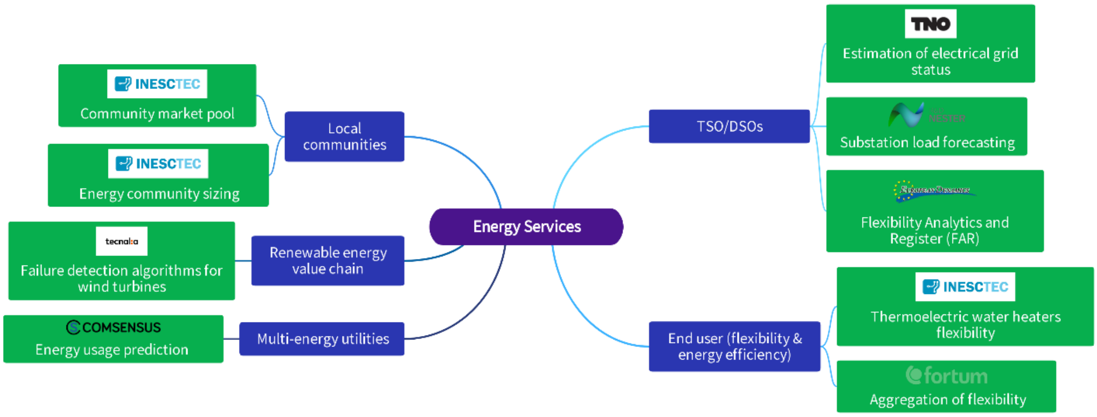
- **Knowledge-driven federated learning (ENGIE):** platform composed of two main components: 1) reasoning engine takes as an input a knowledge graph (KG) describing a given use case such as metadata of available devices and a set of rules describing the type of models; 2) FL engine provides the infrastructure for training FL local and global models as well as for handling their communications. The main innovation is that the platform allows useful knowledge such as use case knowledge from domain experts and ML model knowledge from data experts to be represented as a KG, which can be reasoned upon by the reasoning engine to extract the necessary information to configure and automate the entire FL training process for a given use case. The application is load and generation forecasting, and the current implementation of the knowledge-driven federated learning platform is at the technology readiness level (TRL) 5.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- **Federated learning framework (TNO):** Generic FL platform that enables training decentralized data across multiple devices, allowing seamless aggregation of models trained on local data while promoting knowledge sharing. The added value consists of adapting and customizing the flower Framework to solve challenges within the energy domain while using time series data from various households. It also includes FL anonymization techniques library to choose suitable techniques for the application to different use cases. The use cases are operational predictions related to energy use and possible faults in the physical equipment of the energy grid. The generic FL platform to run any local and global models is on TRL 5.
- **Multivariate energy time series forecasting (INESC TEC):** Vector AutoRegressive (VAR) model, which has been widely used to forecast time series that may have different data owners. In the energy sector, the VAR model is particularly useful for updating very short-term forecasts (e.g., from 15 minutes to 6 hours ahead) using real-time data from geographically distributed sensors, wind turbines, and solar power inverters. ENERSHARE enhances the previous solution from the H2020 Smart4RES project by exploring the transformation of variables using spline and potentially kernel functions. In addition, it will implement a Data Space compatible API that facilitates collaboration among data owners, considering the algorithmic developments. The use case is load forecasting at the consumer and substation levels. So far, a proof of concept with real and large-scale data has been done with a Python script that simulates collaboration between multiple agents using power data from wind farms. However, the integration with the ENERSHARE Data Space was not yet done, so we should assume TRL 4 in what concerns the integration. Nevertheless, for this report, all the components necessary to build an API were identified and specified, so it is expected that by the second deliverable (D6.2) there will be a first API version implemented and tested with real wind power data from Norway, corresponding to TRL5. The target TRL (TRL 7) will be reached by D6.3.
- **Federated transfer learning flexibility potential assessment (COMS):** In FL, the training process occurs on distributed devices or servers, each with its own local dataset. Instead of training a model from scratch on each device, transfer learning allows for the initialization of the models using a pre-trained model, which provides a more effective starting point for learning on each local device or server. These models will be capable of detecting the flexibility potential of the grid by detecting whether PVs, electric vehicles, or heat pumps are present. The innovation within this project is the combination of federated transfer learning and both services in flexibility potential assessment and issues detection in the grid.

<u>**Data-driven energy services**</u>

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831



- **Energy community sizing with assets sharing (INESC TEC):** Sizing of a renewable energy community by computing a Mixed-Integer Linear Problem that uses data available in the Data Space and considers assets sharing. This service demonstrates innovation by facilitating the comprehensive integration of several inputs and functions to enhance the optimization of a renewable energy community operation and sizing, namely: offering the capability to integrate both shared assets and individual assets, including members that may not yet own them, considering their preferences; the possibility to select from a range of business models, considering the limited development and practical implementation of such models. For this phase of the project, a TRL5 is already implemented, considering that, apart from the shared batteries and the practical implementation of a range of business models, all the other functions have already been implemented, tested, and approved with success.
- **Flexibility modelling of thermoelectric water heaters (INESC TEC):** Estimates periods where it is possible to provide flexibility, by optimizing the operation of the thermoelectric water heaters (EWH), using only the appliance consumption data as input. This service can be integrated with the former one for flexibility assessment at the community level. The main innovation is the capacity to handle cases with data scarcity (i.e., only the monitored electrical energy consumption of the EWH is available), but at the same time detailed modelling of the physical load. This service is currently at TRL5 level. Its core operation is fully implemented and functional. It is prepared for its execution in a block, i.e., for estimating flexibility in fixed periods as a whole. For it to be integrable on a shared energy community platform, the implementation must be adjusted to be compatible with running in a sliding window.
- **Community market pool to estimate energy price of internal transactions (INESC TEC):** Estimates a price for the internal transactions within a Renewable Energy Community (REC) and Citizens Energy Communities (CEC) that promote self-sufficiency and a reduction of the costs for each of its members. A particular use case for this service consists of encouraging vulnerable consumers to be actively enrolled in the energy

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

transition by being offered more competitive prices thanks to their participation in an energy community. This service must be understood in combination with the service "Energy community sizing with assets sharing" where similar business models can be considered at the planning and sizing phase of the REC. This tool is currently at TRL6. The methodology has been implemented and tested on real data from members of a CER that is being established in Portugal.

- **Data-driven failure detection algorithms for wind turbine components (TECNALIA, HINE, ENGIE):** Anomaly detection of wind turbine drive train components is covered by this service, and includes the fault detection of one component: gearbox, electric generator, and hydraulic pitch system. The main innovations are: a) an original approach based on a two complementary model is developed: the autoencoder and a graphical probabilistic model (Bayesian network); b) application of the methodology described in (Ainhoa et al., 2023) to a different technology, in this case to generators with permanent magnets (the methodology only was applied to double fed induction generator doubly fed induction generator - DFIG); c) calculating statistics of the relationships between accumulator pressure variations, proportional valve commands, and cylinder motion speeds, where the advantage is that it is avoided having to define faults in terms of pitch system transfer function parameters and the inherent integration of observers. Currently, it is at TRL5, as this service has been validated in a relevant environment.
- **Substation load forecasting tool (NESTER):** Provides day-ahead net-load forecasts, and it can be applied to nationwide aggregate load prediction, all the way down to domestic-level consumption forecasting. It does this by leveraging on a Stacking Ensemble of machine learning models, which, using past net-load data and features extracted from said data, is able to negate the biases of each of the stack members and make an improved forecast than any of the member models. In ENERSHARE, it is expected that the inclusion of the Pattern Sequence Forecasting model in the stack will further improve the accuracy of the model. The current TRL of the Load Forecasting tool is TRL 6. It has been demonstrated, both in past projects and in its current form used in ENERSHARE, to provide accurate forecasts using real data. It is intended for the tool to be integrated with the ENERSHARE Data Space, and providing the resulting forecasts and error metrics, while maintaining the input data's confidentiality.
- **Energy usage prediction (COMS):** Data-driven models to predict energy consumption, with a particular focus on smart meters of users and heat meters of larger buildings for which hourly data are available. The main idea of the service is to support grid planning of the energy mix with the models obtained from real measurements. Based on this, user profiles can be created and segmented, solving privacy issues by creating typical user clusters. The service provides the ability to model different types of users and their electric or heating power profiles. In addition, the models allow finer and more detailed

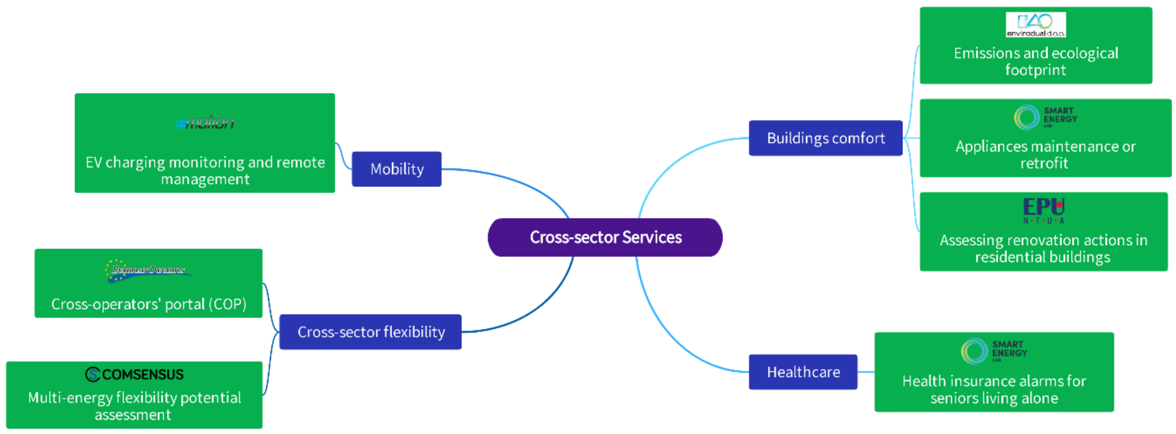
ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

profiles, which are necessary in the case of electric profiles, as the commonly used 15-minute time slices may not be sufficient in the case of electric appliances. Another innovation is that training the models can also allow automatic partial disaggregation of the most interesting flexible loads only from smart metering measurements. The service is currently at TRL 4, with the first test of the tool done on the real data, and the first model created based on this data.

- **Service local load forecasts and estimation of electrical grid status (TNO):** Service that provides local load forecasts and status updates on the energy grid that takes the form of a simulation of the energy grid and provides an answer on whether the specific grid transformer will be overloaded in the next 24 hours. The main innovation is the integration of the FL platform, predictive models, and the global grid simulator into the Data Spaces by putting it as a service running on International Data Space (IDS) connector. We aim to create special clients within the connector of IDS which can be put as a plug-in there to run. The global model containing the energy grid simulator is on TRL 6 since it was successfully used for the specific use case as an operational demonstrator, but it is not yet deployed and tested in actual operational settings.
- **Flexibility Analytics and Register (FAR) service (ED):** Support flexible service providers at any level – from lower-level residential users up to aggregators – positioning themselves in the flexibility and ancillary grid services. It considers the implementations and general guidelines of IEGSA, developed on INTERRFACE project, advancing them to accommodate useful information for sub-metered resources that belong to residential users. It provides open features for residential users or delegated representative market parties, e.g., aggregators, to register and catalog their flexible sources that could potentially provide grid (electricity and thermal) services. The Alpha version of the FAR described in this document is considered a TRL5, the updated version that will be described in D6.2 and D6.3 will evolve on TRL6 and TRL7, accordingly, considering the interaction with demo partners in Pilot 3.
- **Aggregation of flexibility from end users (FORTUM):** This service is developed by Fortum, on behalf of Hiven (Fortum startup), which aims to offer a complete solution for residential energy management. As of now, Hiven offers a Smart Charging application, which users with compatible chargers and vehicles can connect with, and let it schedule the charging of their EV during the cheapest hours of the day. In ENERSHARE, the service was augmented with the option to perform Demand Response, which will attenuate the charging of each individual EV in the Hiven network according to the live grid frequency. The current TRL is 5 and a great deal of the work involved in this was not on the algorithm development side but on putting the groundwork for having a software foundation upon which to build further.

Data-driven cross-sector services

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831



- **Multi-energy flexibility potential assessment (COMS):** The service is based on the existing electric flexibility assessment service. To expand the scope of the service, it was extended to support multiple energy vectors, in our case the heating vector. In addition, when zones are created, data-driven profiles can be used instead of statistically generated user profiles, allowing the service to be tailored to more realistic local communities. The added value of the enhanced service is that it supports data-driven models for user profiling, rather than just statistically-based models of the household. This improves the accuracy of the results by considering the specifics of the community obtained from the data-driven profiles. The service is currently at TRL 4, with the first test of the tool done on the real data, and the first model created based on this data.
- **Cross-operators' portal (COP) (ED):** Bridge the gaps and challenges between different energy sector operators leveraging heterogeneous data sources. Therefore, considering the developed Flexibility Register service (FAR) that catalogs and reports the availability of connected flexibility in the heating and electricity sector, the module of the cross-sector operators' portal (COP) allows the coordinated assessment and representation of common merit order lists among grid operators. Providing inputs from the Weather forecast services, the grid analysis services, and the flexibility assessment service. A joint merit order list is presented to support the coordinated operation of cross-sector operators enabling towards to share and reserve flexibility to address technical challenges. The Alpha version of the Cross-operators' portal described in this document is considered a TRL4, updated version that will be described in D6.2 and D6.3 will evolve on TRL5 and TRL7, accordingly, considering the interaction with demo partners in Pilot 3.
- **Emissions and ecological footprint service (ENVI):** Provides a calculation of nine emission values for: CH_4, CO, CO_2, CXHY, N_2O, NOX, PM10, PM2.5, and SO_2 along with the ecological footprint. This estimation is based on the useful energy of a building in conjunction with the type of fuel it uses for heating. Calculating emissions from energy

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

consumption together with fuel used is not new however estimation of emissions from buildings' energy consumption for heating using machine learning is what this service offers as innovation. Service can also be used to see changes in emissions if one or more parts of the building are renovated like the rooftop, façade, and/or windows. The service is at TRL 4 and was validated and tested in Envirodual's internal server environment.

- **EV charging monitoring & remote management (EMOT):** EV charging monitoring & remote management service was recently established by EMOT for Electricity Grid VS Electric Mobility Cross-Sector operation and it is based on IoT charging stations developed by EMOT. Since energy meters usually only report 15-minute average data, it was installed a single-board computer inside EMOT charging stations enables real-time data collection with one-second sample rate. These fine-grained data are used to train a ML model and obtain an accurate prediction for a highly efficient charging session management to balance a grid in a condition of high penetration from variable renewable energy plants bringing benefits to DSO, CPO, and EV users. According to real-time DSO needs, based on data collected from distributed smart meters, charging sessions can be started/stopped remotely and power output can be modulated remotely using EMOT APIs. The service is at TRL 5, validated in EMOT electric mobility environment.
- **ML-based models for assessing renovation actions in residential buildings (NTUA):** Consists of two Machine Learning models for implementing two different tasks related to the domain of building retrofitting and energy autonomy on the residential scale. The first model is tailored to assessing specific actions at the building level, while the second model aims at assessing the potential of installing rooftop solar panels in residential buildings. The end users of this service include but are not limited to, building owners, financing institutions, investment bodies, and policy specialists. The innovation of both components lies in the AI-based, data-driven manner used for predicting the types of renovations required, as well as the power generation potential. As of our knowledge, most of today's approaches use exhaustive methodologies which require in-place monitoring of a building to propose the most well-fitting retrofitting actions in order to improve the energy class (e.g., energy audits). The service is currently at TRL4, where the methodological framework and algorithms for Assessing Retrofitting Actions and for Predicting the Generation of Rooftop Solar Panels were designed and tested with sample data provided from Pilot 7.
- **Health insurance alarms for senior living alone (SEL):** Monitor energy consumption in households with seniors living alone, identify regular usage patterns, and trigger alarms when there are significant changes from these patterns. The innovation of this use case lies in the data-driven approach to evaluating the need for assistance that seniors living alone may have while maintaining their independence. The service takes advantage of

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

consumption disaggregation algorithms to understand exactly which appliances are being used and when they are being used. It will also be able to identify the regular deviations in behaviour, creating a normal deviation range. Service at TRL 4 with service defined and data collection period started.

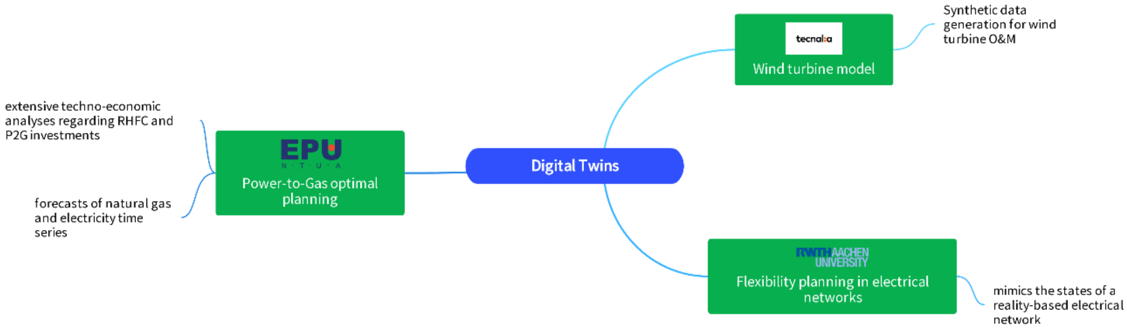
- **Appliances maintenance or retrofit (SEL):** Aims to improve the quality of living and energy consumption in households by detecting higher energy consumption of appliances early on and increasing energy efficiency by suggesting maintenance or renewal of appliances. The information will be shared with consumers and housing providers, and the implementation will involve several actors, including consumers, housing providers, market information aggregators, energy service companies, maintenance and appliance retailers, and appliance producers. The innovation of this service lies in the disaggregation algorithm capacity starting from a learning methodology based on gathered data of main electricity consumption and specific consumption data from the appliance plugs. Service at TRL 4 with service validated and algorithms test starting.

Data visualisation layer

The Data Visualisation Layer objective is to provide all the needed functionalities that will allow the optimal presentation, exploitation, and understanding of the data produced and collected within the various ENERSHARE activities. In essence, this layer provides an advanced visualisation dashboard, with a user-friendly and intuitive interface, through which the end users are able to explore, query and analyse the aforementioned datasets. These functionalities allow a deeper understanding of these datasets and, ultimately, more useful and meaningful conclusions and data storytelling. For the realisation of the above-mentioned objectives, the main technology that was selected is Apache Superset, which is an open-source data exploration and visualisation platform, allowing the users to create visualisations and execute queries without writing any code, thus empowering them to analyse and present data in a user-friendly and interactive manner. The main focus was put on presenting an overview of Apache Superset functionalities, as well as the envisioned way of integrating the tool with the various datasets deriving from the project's activities and other implementation details related to the deployment approach and the overall solution that will be offered.

System-of-system Integrated Digital Twins

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831



Describes in a non-exhaustive way the main concepts of the DTs, the system-of-system DTs, and the approach that the ENERSHARE project is taking regarding the concepts. There is no unified view in the industry now due to the novelty of the topic, and the innovations provided by ENERSHARE could contribute to the discussion. One of the crucial points in the process of developing a DT is the data interoperability and the use of data spaces, then based on the experience of the FIWARE Smart Data Models we discuss the key challenges and ways to address them. Finally, the developed DTs' services in ENERSHARE are:

- **Digital Twin for optimal data-driven Power-to-Gas optimal planning (NTUA):** Digital simulation and optimization platform, called TwinP2G, which platform integrates the national transmission and distribution networks of natural gas and electrical power managed by DESFA and IPTO, respectively. It employs a DT architecture to enable multi-resolution simulations involving power-to-Gas technologies and regenerative hydrogen fuel cells. The objective is to optimize the use of surplus renewable energy for green hydrogen production through electrolysis. Currently, it is at TRL 3. Initial simulations on dummy datasets have been already created and demonstrated.
- **Digital Twin-based O&M algorithms and generation of synthetic failures data (TECNALIA):** Focused on the improvement of the O&M of wind turbines. For it, it integrates anomaly detection functionality for the gearbox, electric generator, and the hydraulic pitch system, based on simulation models. Currently, it is at TRL 5.
- **Digital Twin for flexible energy networks (RWTH):** The DT concept for electrical networks is based on the simulation tool DPSim ("Real-time Simulator for Power Systems") and the acquisition of measurement points from a real electrical network. The idea behind this is to replicate the behaviour of the network in real-time settings, with the use of the most up-to-date status information that becomes available to make the calculations. It is currently at TRL 3, with the setup of the toolchain that acts as the core enablers of the DT functions.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

1 Introduction

The primary objective of the ENERSHARE project is to incorporate and customize the Data Space paradigm within the energy sector. This involves bringing together data, storage, and computing infrastructures with commonly used services and tools for data analysis and sharing, with the aim of creating interoperable resources. In addition to this concept and infrastructure, the project is also developing data-driven services and System-of-System Digital Twin (DT) applications in WP6 ("Cross-value chain AI-based data-driven services") that span multiple value chains. This is achieved by leveraging privacy-preserving federated learning, enabling energy stakeholders to expand their business offerings across and beyond traditional energy services.

A set of innovative digital services and applications, spanning across different energy sectors and value chains, was specified, and developed as part of this project in Month M12. These services build upon the advancements made in previous and ongoing projects such as H2020 BD4NRG, OneNet, PLATOON, BRIGHT, Smart4RES, and I-NERGY. They will be seamlessly integrated into the Energy Data Space (in WP8), leveraging the Data Value (WP7), Trust (WP4), and Interoperability layers (WP3). These cutting-edge services rely on validated data-driven algorithms, carefully designed to prioritize privacy and security. By utilizing federated learning models, these algorithms minimize the need for sharing sensitive data while still ensuring the generation of accurate insights and predictions.

This encompasses a wide range of services that leverage advanced techniques, including federated learning, to address various challenges in the energy sector. For instance, load flexibility estimation considers scenarios where private data is limited or scarce, ensuring privacy and accuracy. Additionally, in district heat systems, aggregated data from multiple stakeholders is utilized to offer planning services to the electrical grid, optimizing energy usage. The project also focuses on enhancing renewable energy sources (RES) by developing a fault detection and classification service specifically designed for wind turbines. This ensures early identification of issues, promoting efficient and reliable operations.

Furthermore, the project extends its impact beyond the energy sector by enabling cross-sector services through the utilization of energy data. These services include learning energy utilization patterns to identify potential problems and trigger alarms for wellness services, monitoring energy poverty and unhealthy environments, facilitating green financing for energy efficiency, remote monitoring, and management of electric vehicles (EVs), as well as assessing emissions and ecological footprints. In summary, this project encompasses a diverse range of services that harness energy data and cutting-edge technologies to drive efficiency, sustainability, and well-being across various domains.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

System-level DTs were also developed to capitalize on enhanced interoperability and foster a culture of trusted and sovereign data sharing. These DTs play a pivotal role in facilitating intricate cross-stakeholder system-level energy planning and enabling near real-time operations customized for optimal flexibility planning, Power-to-Gas green hydrogen planning, and efficient operation and maintenance (O&M) of RES, particularly wind-based generation.

1.1 About this document

This deliverable will report about the alpha version of the federated learning (Task 6.1), data-driven services (Tasks 6.2 and 6.3), data visualisation (Task 6.4) and DTs (Task 6.5). The objective of this document is to describe these developments and software, as well as its level of maturity and readiness. The software components described in this document completed the first technology development cycle, which means that have well defined software functions and API, and the next steps are the integration with the Data Space components and data visualisation layer.

1.2 Intended audience

The intended audience for this deliverable is:

- Data providers and consumers of the Data Space, willing to know the potential of the data-centric energy and cross-services, as well as DT capabilities to generate synthetic data.
- Potential end-users of the services, namely: consumers, energy communities, transmission system operator (TSO), distribution system operator (DSO), multi-energy utilities, O&M companies, RES power plants developers, non-energy service providers (i.e., mobility, healthcare, buildings).
- Data Space operators interested in creating value from data by enabling data-centric service provision.

1.3 Reading recommendations

This document is divided into 7 chapters. Chapter 1 is this introduction. Chapter 2 describes the enhancement and adaption of Federated Learning (FL) methods from previous projects, capable of considering these strong constraints when it relates to private or confidential data in the energy domain. The FL methods cover use cases with different requirements in terms of statistical learning model type (linear, non-linear, etc.) and data ownership, such as data partitioned by examples (cross-device setting) or partitioned by features (cross-silo setting).

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Chapter 3 describes the developed energy services, covering the following stakeholders group: local communities, RES value chain, multi-energy utilities, TSO, DSO, end-user/consumer. Chapter 4 describes the services that integrate energy-related data with other sector data spaces (mobility, healthcare, finance, buildings, cross-sector flexibility), considering cross-stakeholder and cross-border data/service/resource exchange and sharing. Chapter 5 describes the developed novel visualisations engine capable of identifying, manipulating, formatting, and presenting data in an efficient way to optimally communicate their meaning and create interpretable knowledge for the services described in Chapters 3 and 4. Chapter 6 describes the system-level DTs covering wind turbine model, power-to-gas optimal planning and flexibility planning in electrical networks. Finally, Chapter 7 presents final remarks, in particular the level of fulfilment of the key performance indicators (KPI) defined in the ENERSHARE objectives for WP6.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2 Federated Learning Methods

Privacy and confidentiality of data are of significant concerns in the energy domain. These constraints dictate the ability and willingness of grid operators and stakeholders to share their data. Federated Learning (FL) makes it possible to train statistical models over remote devices or siloed data sources, while keeping data localised. By moving models or algorithms to train locally, this approach mitigates the risk of leaking sensitive information and enables private learning between devices without violating any legal, administrative, or ethical constraints that require data to remain local.

Task 6.1 aims at leveraging FL to propose innovate learning approaches that are, on the one hand, respecting privacy constraints on data and, on the other hand, tailored to specific requirements of the pilots and use cases of ENERSHARE. In this section, we describe in detail the four different privacy-preserving FL approaches proposed within the Task 6.1, their current state and expected development over the course of the project, and their future integration with ENERSHARE's data space.

2.1 Knowledge-driven federated learning

2.1.1 Description of the FL approach

In ENERSHARE, ENGIE aims at applying FL approach to predict energy consumption or generation that requires training models on data from multiple distributed sources and that imposes constraints on data privacy, ownership, and locality. To this end, we propose a knowledge-driven federated learning platform that leverages both the data and domain knowledge. The FL platform is composed of two main components, namely reasoning engine and FL engine, as depicted in Figure 1.

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

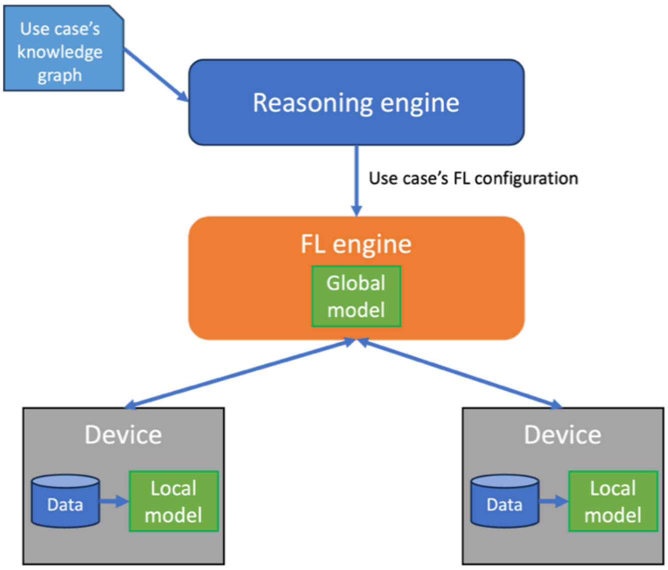


Figure 1: Architecture of knowledge-driven federated learning platform

Reasoning engine

The reasoning engine takes as an input a knowledge graph (KG) describing a given use case such as metadata of available devices (e.g., available data for training, device API) and a set of rules describing the type of models to be used for FL; it performs inference to determine the target device(s) for each model. The rules used for extracting knowledge can be customised as needed and are based on Datalog, a rule language that supports rich semantics. Beyond the traditional Datalog engine, we make several semantic extensions to make it more powerful and applicable in a wider range of use cases. Concretely speaking, the reasoning engine supports Negation as Failure, Bind operator, Comparison operator and Aggregation operator. Furthermore, this reasoning engine supports incremental data and rules insertion/deletion, which makes it possible to handle dynamic cases, where the RDF graph and common knowledge evolve overtime, with minimum computation.

In traditional Datalog engines, stratification is a paradigm that is used to determine the reasoning execution plan. Instead of stratification, our reasoning engine uses hypergraph to make the reasoning execution more fine-grained; this approach is especially efficient when handling rule-incremental cases, which is crucial for managing different FL cases as some of the common knowledge may be modified, while others remain unchanged. The overall idea of this approach is to find the directly influenced hypernodes (DIHNs) in the hypergraph of the rule set according to the changes; these DIHNs will trigger a chain reaction according to the topological order starting from DIHNs in the hypergraph. By applying this approach, we can avoid unnecessary execution of some certain rules that will not be impacted by these changes and maintain the result's correctness.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

As for the execution of each Datalog rule, we inspire from the Volcano Model in relational database architecture that consumes little memory space and gives the system good extensibility. Another crucial aspect that may impact much the reasoning engine performance is the join order of each rule execution. Thus, in our reasoning engine, there exists a join order optimiser to handle specifically the optimisation of the join order. This optimiser is designed to use a heuristic and cost-based strategy to generate a left deep join tree that makes a balance of the time consumption between join order determination and rule execution.

Federated learning engine

The FL engine provides the infrastructure for training FL local and global models as well as for handling their communications. The engine is designed to be generic and adaptable so that it can execute any models as required for each given use case. Built upon Flower[1], which is an open-sourced federated learning framework, the FL engine enables the user to federate any workload, ML models, or framework, using any programming language.

As illustrated in Figure 1, the FL engine receives a use case's FL configuration, which essentially contains, but not limited to, the information regarding the target devices to include in the FL process, from the reasoning engine. With this information, FL engine can automate FL training process by creating an on-the-fly setup of FL from connecting to training devices until aggregation. This feature reduces the efforts required when facing dynamics such as changes in the use case setting, devices, or models, or when applying to a new use case. By simply modifying the use case KG to reflect the changes, the platform can automate the entire process with the new configuration.

Models

In the federated setting, local models are distributed and trained with local data over different connected devices selected to participate in the training. After local training, model updates are sent back to FL server to update the global model. The type of models to use depends on each individual use case. For testing the current version of our FL approach, we use the platform to train a model to predict power generation of wind turbines in a wind farm; 4 local models are trained with local data of 4 different wind turbines and aggregated to update the global model. Local and global models are fully connected neural network (with 3 linear layers and 2 activation functions). This choice of the models was made for simplicity in testing the approach and the platform. When applying to a use case of ENERSHARE, the details and specificity of the use case will dictate the type of model to use (e.g., LSTM or knowledge graph embedding).

[1] https://flower.dev/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2.1.2 Innovation

The FL platform allows useful knowledge such use case knowledge from domain experts and ML model knowledge from data experts to be represented as a KG, which can be reasoned upon by the reasoning engine to extract the necessary information to configure and automate the entire FL training process for a given use case. The adaptation of the training process to respond to changes is performed by simply updating the use case KG and relaunching the process, thereby reducing human intervention. The FL engine is designed as a generic tool that supports any model such as classic ML models and KG embedding models[2] taking inputs as time series data and knowledge graphs, respectively.

2.1.3 Functions

Using the FL platform, the process of implementing a forecasting model requires 4 main steps, namely use case configuration, FL setup, federated model training, and forecasting.

1. Use case configuration: The objective of this step is to determine the configuration of the federated learning required to build the forecasting model for a given use case.
    a. Define the use case KG and rules for device selection;
    b. Reasoning engine makes inferences using the use case KG and the defined rules to select the compatible devices for training the model and extract the required information for communicating with the devices (e.g., IP address and port).
2. FL setup: This phase initialises the FL process based on the use case configuration received from step 1.
    a. FL engine initialises the global model, launches a FL server, and sends a request to each of the selected devices to participate in the training along with local model specifications;
    b. Upon receiving the request, each device launches a FL client and establishes the connection with FL server.
3. Federated model training: Once the FL server is connected to all the required FL clients, the training process starts.
    a. Each FL client starts their local training using local data. To enhance the privacy of the local data used for training, differential privacy methods (e.g., differentially-private stochastic gradient descent) can be applied during the training at each FL client. After local training, the FL clients send model updates (e.g., model parameters) back to the FL server;
    b. FL server aggregates the received model updates from the clients to update the global model. The training process can repeat until the global model converges.

[2] The development and testing with KG embedding are in progress and expected to be ready for the next release.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

4. Forecasting: After the training completes, we have the parameters of the global model that contain the learnings from data of the clients. With these parameters, a model can be instantiated and used for performing the forecast.

2.1.4 Input and Output Data Format

Data exchanges required for different steps of developing FL forecasting are described below. The summary of input and output data and their format are presented in Table 1.

Table 1: Input and output data format for knowledge-driven federated learning

Component		Input/Output description	Variable type
Reasoning engine	INPUT	• Use case knowledge graph (N-triples) • Rules (Datalog)	N-triples format data Engine specified Datalog expression
	OUTPUT	• Devices' ID • Devices' IP address • Devices' ports	List of strings List of strings List of integers
FL engine	INPUT	• Devices' ID • Devices' IP address • Devices' ports	List of strings List of strings List of integers
	OUTPUT	• Trained global model parameters	Matrix of floats
Forecasting	INPUT	• Trained global model parameters	Matrix of floats
	OUTPUT	• Forecast	Float

Use case configuration

The use case KG contains the knowledge in triples and serialised using N-triples format. Rules are based on Datalog and stored in a file (*see example below*). Reasoning engine takes them as input from local files, without any data transfer.

```
<TargetDeviceForModel1>(X) :- <WindTurbine>(X) and <belongsTo>(X, <windFarm_1>) and
<hasBaAverageMeasurement>(X, M1) and <hasWs1AverageMeasurement>(X, M2) and
<hasWs2AverageMeasurement>(X, M3) and <hasCmAverageMeasurement>(X, M4) and
<hasDsAverageMeasurement>(X, M5) and <hasRmAverageMeasurement>(X, M6) .
<hasBaAverageMeasurement>(X, M) :- <hasMeasurement>(X, M) and <Ba_avg>(M) .
<hasWsAverageMeasurement>(X, M) :- <hasMeasurement>(X, M) and <Ws_avg>(M) .
```

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
<hasCmAverageMeasurement>(X, M) :- <hasMeasurement>(X, M) and <Cm_avg>(M) .
<hasDsAverageMeasurement>(X, M) :- <hasMeasurement>(X, M) and <Ds_avg>(M) .
<hasRmAverageMeasurement>(X, M) :- <hasMeasurement>(X, M) and <Rm_avg>(M) .
<TargetIpForModel1>(Y) :- <TargetDeviceForModel1>(X) and <hasIpAddress>(X, Y) .
```

FL setup

The reasoning engine and FL engine are hosted on the same server. The exchanges between them are done locally, without any data transfer. The FL devices provide a REST API allowing the FL engine to send requests for them to start FL training process locally (i.e., launch FL client) and if needed to send local model specification for the client to instantiate.

Federated model training

FL server and FL clients are based on Flower's server and client. Their communications, which include exchanging local and global model updates/parameters, are done via Remote Procedure Call (RPC) protocol, implemented in Flower's framework. For this version of the implementation, we employed the platform to develop a forecasting model to predict wind turbine power generation. The dataset used for the test contains time-series data with 10mn granularity generated from 4 different wind turbines in a wind farm. The dataset has the following features: pitch angle, wind speed, torque, and power, which is the target variable to be predicted.

2.1.5 Implementation details

The current implementation of the knowledge-driven federated learning platform is at TRL 5; the reasoning engine and the FL engine have been implemented and tested; a proof of concept has been developed and model trained using open data. For D6.2 (M19), we aim to reach TRL6 by adapting for and testing with real pilot data and use case(s); we expect to extend the approach with knowledge-based models such as knowledge graph embedding or knowledge-infused learning, but the feasibility and applicability of which will depend on the available data and the use case(s). For D6.3 (M28), we target TRL7 by integrating and deploying the service. The platform is implemented using Python (version 3.9), and Table 2 presents the list of python libraries used in the development.

Table 2: Third-party software and libraries

Library	Description	License
pandas	Analysis and preprocessing the data	3-clause BSD
torch	Deep learning library for python	Open source

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

pickle	Storing and loading data	MIT License
threading	Thread management	MIT License
sklearn	Data processing, training	Open source
opacus	Library for differential privacy	Apache License 2.0
flask	Micro web framework for developing web applications	3-clause BSD
flower	Federated learning library	Open source

2.1.6 Integration with ENERSHARE Data Space

With respect to the ENERSHARE data space, the forecasting model will be exposed via an API to be integrated in the data space as a service or an app in the App store. To execute FL training process, the FL client will need to retrieve data via the data space. In the case where the model must be sent to the device (i.e., local data cannot be made available in the data space), the FL client can be run on a local machine at the data source.

2.2 Federated learning framework

2.2.1 Description of the FL Approach

In ENERSHARE project, TNO concentrated on using FL in complex settings of energy market where there are different organisations and businesses as well as local communities which use energy, playing different roles in energy provision process, and where some data, owned by these stakeholders, is sensitive (for example, privacy data on households, or business confidential information of companies). FL could be applied in such settings for controlled information exchange between parties with different roles (see the Figure 2 below).

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

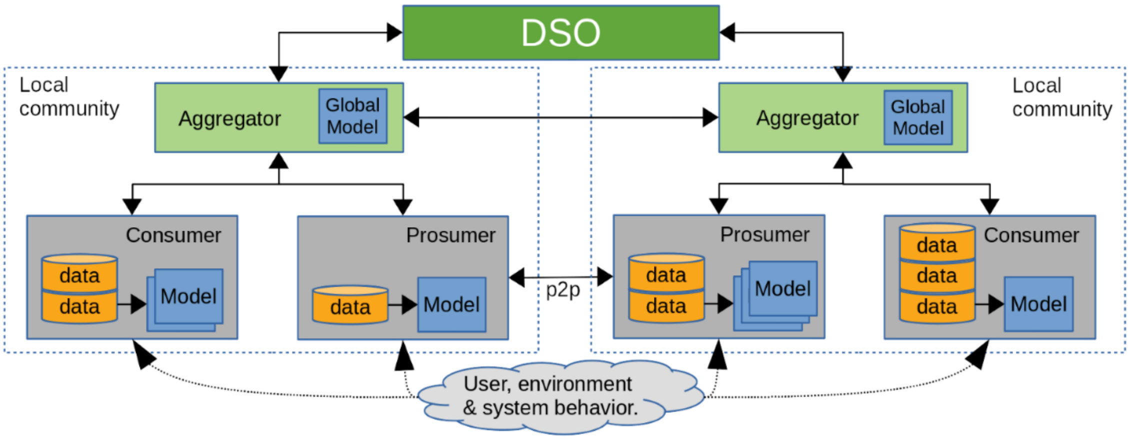


Figure 2: Architecture of local-predictions driven federated learning platform

In these settings we concentrate on operational predictions related to energy use and possible faults in the physical equipment of the energy grid. Since different parties are involved, we need to create a platform for FL, and 3 types of models:

- Local Model for the prediction of the energy consumption for next 24 hours per household (is implemented with the use of LSTM machine learning technique).
- Pseudonymisation semi-global model (allowing guarantee on not using actual privacy data).
- Global model with the simulation of the energy grid to predict the load on the energy transformers. Anonymisation instead

All these models run on the generic Federated learning platform which we develop as well.

2.2.2 Innovation

First innovation: FL anonymization techniques library to choose suitable technique for the application to different use-cases. The library includes multiple data obfuscation methods and differential privacy implementations with the idea that the user can choose the most suitable technique to apply according to the use-case demands.

Second innovation: generic Federated Learning platform that enables training decentralized data across multiple devices, allowing for seamless aggregation of models trained on local data while promoting knowledge sharing. Our added value consists of adapting and customizing the flower Framework to solve challenges within the energy domain while using time series data from various households.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2.2.3 Functions

- **Multivariate_data function** – to parse the input data this is multivariate time series data in csv format and pre-process the data for the model development. The data is scaled using the MinMaxScaler().
- **Model training function** – The data was split into the train and validation using the from_tensor_slices function. Prepared data is then used in model training. General model is downloaded and locally saved for the basis in federated transfer learning approach. The said general model is fine-tuned and retrained to build a new model adapted to input data.
- **Model Fit function** – The LSTM model (see local models description) was trained using batch size = 32, 25 epochs and patience level was set at 5. **Model Fit function** – LSTM model was trained using batch size = 32, 25 epochs and patience level was set at 5. This returns the updated model parameters and results after training it using the provided hyperparameters and locally held dataset.
- **Evaluate Function** – Evaluates the loss of the model the provided weights using the locally held dataset. The performance of the model is evaluated using the Mean Absolute Error (MAE) as an evaluation metric. The MAE is obtained by taking the absolute difference between the target variable, active power's true and predicted values.

FL Functions

Using the FL platform, the process of implementing a forecasting model requires the following main steps: FL setup, federated model training, and local forecasting.

FL setup: This phase initialises the FL process

a. The bashscript is run to initialize the Flower server, the strategy then is instantiated using the provided configuration, and the global model is set up. The server starts and listens for requests from participating devices.
b. Once the server is fully launched, the clients are started via the same script and they send requests to the server to participate in the training along with local model specifications.
c. Upon receiving the request, the server establishes the connection and instantiates a flower client for each connected device.

FL model training: Once the FL server is connected to all the required FL clients, the training process starts. Each FL client uses local data to begin their local training. Following local training, the FL clients communicate model updates (such as model parameters) to the FL server, which

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

aggregates the updates and updates the global model. Until the global model converges, the training process can be repeated.

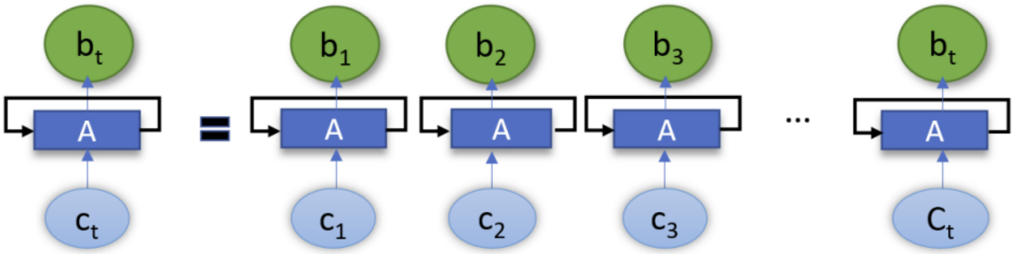


Figure 3: LSTM Network Overview

Local Models Description

In our proposed solution, the local models utilize the Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) to develop a local predictive ML model providing the forecasting for the next 24 hours for the household energy usage. The model weights afterwards are aggregated in the semi-global model, which is a novel approach to address the challenge of data privacy. RNNs, which are specialized in processing sequential input, have been investigated for this task. This family of networks is capable of processing variable-length sequences and sharing parameters with other model components. However, the problem of vanishing or exploding gradients presents a challenge for RNNs when learning long-term dependencies. To address this issue, we extended the LSTM model, which includes a self-loop weight that is controlled by another hidden unit and is significant for the gating of these models. By dynamically adjusting the time scale of integration based on the input sequence, the model becomes more adaptable to a wider range of application scenarios. As shown in Figure 2, cells are constantly connected to one another, and the input characteristic is computed by a standard artificial neuron unit, which may accumulate its value into the state if the sigmoidal input gate permits it. The linear self-loop weight of the state unit is controlled by the forget gate, and the output of the cell can be turned off via the output gate. All of the gating units have sigmoidal nonlinearities, while the input unit may have any squashing nonlinearity. Additionally, the state unit may provide an additional input to the gating units.

Anonymization:

In the ENERSHARE project one of the innovation is developing a more secure layer on top of federated learning framework as an additional security layer. The choice of selection of the anonymization technique and implementing is still a work in progress. The idea is to use Differential privacy or the similar approach. Differential Privacy is a mathematical framework that sets a limit on an individual's influence on the outcome of a computation, such as the parameters of a ML model. This is accomplished by bounding the contribution of any individual

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

user and adding noise during the training process to produce a probability distribution over output models. DP comes with a parameter (ε) that quantifies how much the distribution could change when adding or removing the training examples of any individual user (the smaller the better).

Overview of the Federated Learning Platform:

FL platform allows to run any FL local and global models and guarantees the communication between those global and local models.

The EnerShare FL platform solution is based on the open-source Flower platform. Flower is a framework that has been built by a community of researchers and engineers for implementing Federated learning systems. Flower allows collaborative machine learning by enabling diverse devices to collaborate on training models without sharing raw data. Its design was based on providing a highly customizable framework which can be used in various use cases, extendable, allowing to build new state-of-the-art systems and Framework-agnostic, supporting different Deep Learning frameworks (e.g., TensorFlow, PyTorch, Keras and so on).

Flower also offers abstracts of its functions and server – client-side logic which facilitates customizing it to any use case. This allowed us to build our FL solution based on Flower architecture prototype and adapt it to solving the challenge with the overload on the grid. Flower's architecture consists of two main parts as shown by Figure 4. The Flower server part represents the heart of the Federated Learning loop (via the Flower Strategy Object). To send and receive messages between clients and the server, the Flower communication protocol is implemented on top of bidirectional gRPC (Foundation) streams, ensuring a lightweight way of communication. Remote Procedure Call (RPC) is known for its efficient binary serialization format, allowing exchange of multiple messages without having to re-establish a connection for every request/response pair. In addition, RPC can also run on any operating system.

Flower also offers possibilities of anonymization via differential privacy wrappers. Flower also provides differential privacy (DP) wrapper classes for the integration of the central Differential privacy into training pipelines defined in any of the various ML frameworks that Flower is compatible with. This wrapper-based approach has the advantage when compared with inheritance like approaches of being easily composable with other wrappers that can be contributed to the Flower library in the future, e.g., for secure aggregation.

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

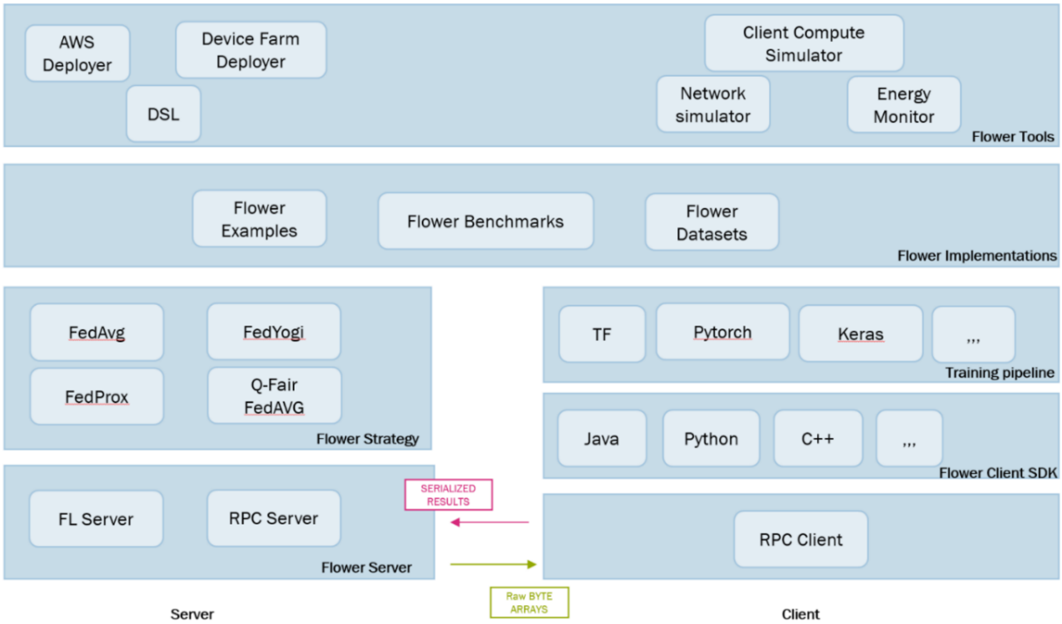


Figure 4: Flower Framework architecture

2.2.4 Input and Output Data Format

The table below describes the expected input and output data format of each component developed within our solution.

Table 3: Input and output data format for Federated Learning solution

Component		Input/Output description	Variable type
Local model	INPUT	• Timeseries dataset	Timestamp, float
	OUTPUT	• Predicted target valuable	Float
Semi-global model	INPUT	• Timestamps • Local data from every involved household	Float Float
	OUTPUT	• Predicted target variable (anonymised)	Float

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Global Model (Grid simulator)	INPUT	• Predicted output from semi global model	Float Float (positive value)
	OUTPUT	• Probability of overload in transformer	String

The dataset used for the first implementation version is the time-series data generated from heat pumps and partitioned across the connected households. The dataset features consist of active power, reactive power, and apparent power. The target variable to be predicted is the active power. The data is resampled on an interval of 15 minutes. The active power is predicted for the next 24 hours. Below you can see the snapshot of the data being used.

1	measuredat	pmactivepowerw	pmapparentpowerw	pmreactivepowerw
2	3-5-2021 15:00	0.2100	31.183	-0.42886584
3	3-5-2021 15:15	0.2200	31.183	-0.42886585
4	3-5-2021 15:30	0.2300	32.183	-0.42886585
5	3-5-2021 15:45	0.2301	33.183	-0.42886585
6	3-5-2021 16:00	0.2302	34.183	-0.42886585
7	3-5-2021 16:15	0.2303	35.183	-0.42886585
8	3-5-2021 16:30	0.2304	36.183	-0.42886585
9	3-5-2021 16:45	0.2305	37.183	-0.42886585
10	3-5-2021 17:00	0.2306	38.183	-0.42886585
11	3-5-2021 17:15	0.2307	39.183	-0.42886585
12	3-5-2021 17:30	0.2308	40.183	-0.42886585
13	3-5-2021 17:45	0.2309	41.183	-0.42886585
14	3-5-2021 18:00	0.2310	42.183	-0.42886585
15	3-5-2021 18:15	0.2311	43.183	-0.42886585
16	3-5-2021 18:30	0.2312	44.183	-0.42886585
17	3-5-2021 18:45	0.2313	45.183	-0.42886585
18	3-5-2021 19:00	0.2314	46.183	-0.42886585
19	3-5-2021 19:15	0.2315	47.183	-0.42886585
20	3-5-2021 19:30	0.2316	48.183	-0.42886585
21	3-5-2021 19:45	0.2317	49.183	-0.42886585
22	3-5-2021 20:00	0.2318	50.183	-0.42886585
23	3-5-2021 20:15	0.2319	51.183	-0.42886585

Figure 5: Local model input dataset

The Screenshot of the output from semi-global model is provided in the Table beneath.

Table 4: Output of the semi-global model

Number of clients	Training time	MAE
2	73.76	0.053
5	96.38	0.051
8	125.64	0.049
10	144.42	0.047

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

In this task anonymization is the more secure layer on top of the federated learning model which is a semi-global model. This is done so that DSOs cannot trace back from which household the overload is coming. The input to the anonymization model would be the output from the local and semi global models. The output from the anonymization model would be then send to the grid simulator to predict the probability of overload in a transformer line. The choice of which anonymization technique to apply is still a work in progress. Using differential privacy or similar approach could be one of the solutions for the anonymization.

APIs of the FL learning solution are based on Flower framework APIs.

In terms of API, the Table 5 presents a list of Flower APIs that were used to build the solution platform.

Table 5: Flower Federated Learning API

API	Description
Client	Base class for Flower clients. Once instantiated, it enables to start a Flower client node which connects to a flower server.
Server	Enables to start a Flower server using a gRPC transport layer.
Common	Allows for common components shared between server and client, such as the model parameters, client status and evaluate instructions for a client.

Transfer through gRPC:

The flower client contains local data and the flower client object. Messages exchanged between clients and the server are ensured via bidirectional gRPC (Foundation) streams. The Flower communication protocol is implemented on top of bi-directional gRPC (Foundation).

2.2.5 Implementation details

The generic FL platform to run any local and global models is on TRL 5: Technology demonstration. We have already deployed the platform and run the predictive local models and the global models on it.

The local predictive models based on LSTM ML technique are on TRL 5. We have developed the models and trained them on the actual historical dataset from the heat pumps. However, it is too early to say that it is on TRL 6 since the models need to be trained on different types of devices. For now, it is an operational demonstration. The semi-global model is at TRL 3. The implementation of the obfuscation data techniques is in progress, and with the first versions of the implementations we demonstrate that the solution is feasible under certain assumptions.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Table 6: Software libraries used

Library	Description	URL	License
flwr	Flower federated learning framework	https://flower.dev/docs/index.html	Open-source Apache 2.0
multiprocessing	Process-based parallelism	https://docs.python.org/3/library/multiprocessing.html	Open Source, Apache 2.0
NumPy	Scientific computing	https://numpy.org/	3-Clause BSD
os	Miscellaneous operating system interfaces	https://docs.python.org/3/library/os.html	Open Source
pandas	High-performance, easy-to-use data structures and data analysis tools	https://pandas.pydata.org/	3-Clause BSD
scipy	Fundamental algorithms for scientific computing in Python	https://scipy.org/	3-Clause BSD
tensorflow	Machine Learning library for data automation, model tracking, performance monitoring, and model retraining	https://www.tensorflow.org/	Open-Source
Sklearn	Machine Learning library and data analysis	https://scikit-learn.org/stable/	Open source
datetime	Time and dates types	https://docs.python.org/3/library/datetime.html	Open Source

2.2.6 Integration with ENERSHARE Data Space

FL solution from TNO will become an integral optional part of the IDS connector. There will be the possibility to include the special client for local, semi-global and global models into the basis IDS connector developed by TNO. The source-code of the TNO IDS connector can be found here: TNO Security Gateway Documentation (https://tno-tsg.gitlab.io). Every type of client will be integrated separately and will be using the communication basis from IDS connector.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2.3 Multivariate energy time series forecasting

2.3.1 Description of the FL approach

Various statistical methods have been developed to integrate time series data from multiple parties. One such method is the Vector AutoRegressive (VAR) model, which has been widely used to forecast time series that may have different data owners. In the energy sector, the VAR model is particularly useful for updating very short-term forecasts (e.g., from 15 minutes to 6 hours ahead) using real-time data from geographically distributed sensors (such as anemometers, pyranometers), wind turbines, and solar power inverters. A VAR model describes the dynamic between multiple time series by assuming that each time series in the system depends on its own past values and the past values of all the other time series in the system.

Mathematically, a VAR model can be described as follows. Let $\mathbf{y}_t = (y_{1,t}, y_{2,t}, \dots, y_{N,t})$ be an N-dimensional multivariate time series, where N is the number of data owners, and $y_{i,t}$ denote the time series from the i-th data owner at time t. A VAR model of order p can be written as:

$$\mathbf{y}_t = \boldsymbol{\eta} + \sum_{\ell=1}^{p} \mathbf{y}_{t-\ell}\mathbf{B}^{(\ell)} + \boldsymbol{\varepsilon}_t,$$

where

- $\boldsymbol{\eta} = [\eta_1, \dots, \eta_N]$ is the constant intercept (row) vector;
- $\mathbf{B}^{(\ell)}$ is the coefficient matrix at lag $\ell \in \{1, \dots, p\}$, and the coefficient associated with lag ℓ of time series i to estimate time series j is positioned at (i, j) of $\mathbf{B}^{(\ell)}$, $i, j \in \{1, \dots, N\}$;
- $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \dots, \varepsilon_{N,t})$ is a white noise vector.

Figure 6 depicts a visual representation of a VAR model making clear the natural division of data and coefficients by agents. The main issue is how to estimate the coefficients without sharing sensible information – a brief discussion is provided in the next paragraphs.

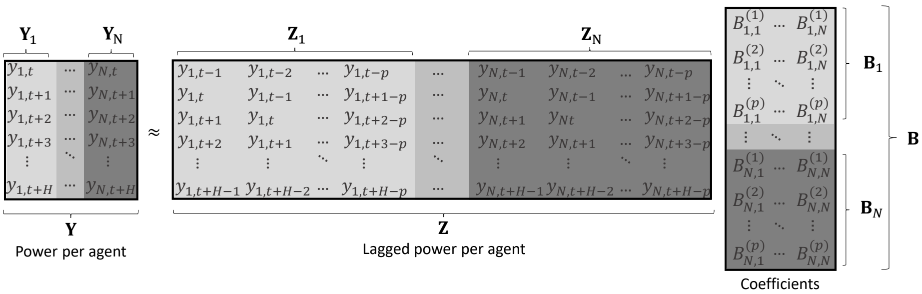


Figure 6: VAR model representation (1 timestamp ahead, using p lags on power time series).

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

As the number of data owners increases, as well as the number of lags, regularization techniques must be used to perform automatic variable selection and remove redundant data. The Least Absolute Shrinkage and Selection Operator (LASSO) is convenient to use when handling high-dimensional data since it shrinks some of the coefficients to zero, performing variable selection. Instead of assuming that all lagged multivariate time series are contributing to the model, LASSO extracts, with a small computational effort, the predictors with the strongest contribution to forecast the target variable. In the LASSO-VAR approach, the coefficients are estimated by solving the minimization problem:

$$\widehat{\mathbf{B}}_{\text{LASSO}} = \arg\min_{\mathbf{B}} \|\mathbf{Y} - \mathbf{ZB}\|_2^2 + \lambda \|\mathbf{B}\|_1,$$

where $\|.\|_r$ represents both vector and matrix L_r norms, and $\lambda > 0$ is a scalar penalty parameter to be tuned. To solve this optimization problem without compromising data privacy, a critical literature review has been made in (Gonçalves et al., 2021) showing that existing privacy-preserving techniques are unsatisfactory. As a result, a novel solution has been put forth to encrypt the data is such a way that data owners can solve the LASSO-VAR model (Gonçalves et al., 2021a). In summary:

1. Collaboration among data owners is established to encrypt their respective data using a collectively unknown key $\mathbf{M}$ ($\mathbf{M} = \mathbf{M}_1\mathbf{M}_2 \ldots \mathbf{M}_N$, where only $\mathbf{M}_i$ is known by i-th data owner). Additionally, covariates ($\mathbf{Z}_i$) are locally transformed using a secret key ($\mathbf{Z}_i\mathbf{Q}_i$), crucial for safeguarding the coefficient values. An overview of the data encryption protocol is provided in Figure 7. Both covariates ($\mathbf{Z}_i\mathbf{Q}_i$) and target ($\mathbf{Y}_i$) are encrypted using the protocol depicted in Figure 7.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Algorithm for Encryption of a private data $\mathbf{X}_i$

Input from the i^{th} **agent:**

- data to encrypt: $\mathbf{X}_i \in \mathbb{R}^{T \times s}$
- integer $r \in]\sqrt{Ts - u}, T[$, where u is the number of unique values in $\mathbf{X}_i$
- random private matrices: $\mathbf{M}_i \in \mathbb{R}^{T \times T}$, $\mathbf{C}_i \in \mathbb{R}^{T \times (r-s)}$, $\mathbf{D}_i \in \mathbb{R}^{r \times r}$

Input from the j^{th} **agent** ($j \neq i$)**:** $\mathbf{M}_j \in \mathbb{R}^{T \times T}$

Output: $\mathbf{MX}_i = \mathbf{M}_1 \dots \mathbf{M}_N \mathbf{X}_i$ (encrypted data)

Initialization: Agent i computes and shares $\mathbf{W}_i \in \mathbb{R}^{T \times r}$ with the N^{th} agent,

$$\mathbf{W}_i = [\mathbf{X}_i, \mathbf{C}_i]\mathbf{D}_i.$$

Agent N receives $\mathbf{W}_i, \forall i$.
Agent N shares $\mathbf{M}_N\mathbf{W}_i$ with the $(N-1)^{\text{th}}$ agent.
for agent $j = N-1, \dots, 1$ **do**
  Agent j receives $\left(\prod_{k=j+1}^{N} \mathbf{M}_k\right) \mathbf{W}_i$, and
  if $j > 1$ **then**
    shares $\mathbf{M}_j \left(\prod_{k=j+1}^{N} \mathbf{M}_{A_k}\right) \mathbf{W}_i$ with agent $(j-1)$
  else
    shares $\mathbf{M}_j \left(\prod_{k=j+1}^{N} \mathbf{M}_{A_k}\right) \mathbf{W}_i$ with agent i
  end if
end for
Agent i receives $\mathbf{MW}_i$ from the 1^{st} agent and recovers $\mathbf{MX}_i$,

$$[\mathbf{MX}_i, \mathbf{MC}_i] = \mathbf{MW}_i\mathbf{D}_i^{-1}.$$

Figure 7: Data encryption.

2. Data owners estimate the coefficients by sharing the encrypted version of their local models, as depicted in Figure 8. A main advantage of this algorithm is that global model's coefficients are not affected or deteriorated as is the case with differential privacy techniques.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Algorithm for Synchronous Privacy-preserving LASSO-VAR.

Input: Encrypted data from all (encryption algorithm) $\mathbf{MZ}_i\mathbf{Q}_i$, $\mathbf{MY}_i$, $\mathbf{Q}_i^\top \mathbf{Z}_i^\top \mathbf{M}^{-1}$
Output: Encrypted coefficients $\mathbf{B}'_i$, i=1, . . . , N ($\mathbf{B}'_i$=$\mathbf{Q}_i\mathbf{B}_i$)

Initialization: $\mathbf{B'}_i^0$, $\overline{\mathbf{H}}^0$, $\mathbf{U}^0 = \mathbf{0}$, $\lambda, \rho \in \mathbb{R}^+$, $k = 0$
for agent $i = 1, \ldots, N$ **do**
 $\mathbf{P}_i = \left((\mathbf{Z}_i\mathbf{Q}_i)^\top(\mathbf{Z}_i\mathbf{Q}_i) + \rho \mathbf{Q}_i^\top \mathbf{Q}_i\right)^{-1}$
end for
while $\exists\ i$ such that $\frac{\|\mathbf{B}_i^{k+1} - \mathbf{B}_i^k\|_2}{\max(1, \min(\|\mathbf{B}_i^{k+1}\|_1, \|\mathbf{B}_i^k\|_1))} > tol$ **do**
 for agent $i = 1, \ldots, n$ **do**
  Initialization: $\widetilde{\mathbf{B}}_i^0$, $\widetilde{\overline{\mathbf{H}}}^0$, $\widetilde{\mathbf{U}}^0 = \mathbf{0}$, $j = 0$

$$\mathbf{K}_i = \mathbf{MZ}_i\mathbf{Q}_i\mathbf{B'}_i^k + \overline{\mathbf{H}}^k - \overline{\mathbf{MZQB'}}^k - \mathbf{U}^k$$

  while $\frac{\|\widetilde{\mathbf{B}}_i^{j+1} - \widetilde{\mathbf{B}}_i^j\|_2}{\max(1, \min(\|\widetilde{\mathbf{B}}_i^{j+1}\|_1, \|\widetilde{\mathbf{B}}_i^j\|_1))} > tol$ **do**
   $\widetilde{\mathbf{B}}_i^{j+1} = \mathbf{P}_i \left(\mathbf{Q}_i^\top \mathbf{Z}_i^\top \mathbf{M}^{-1}\mathbf{K}_i + \rho \left(\widetilde{\overline{\mathbf{H}}}^j - \widetilde{\mathbf{U}}^j\right)\right)$
   $\widetilde{\overline{\mathbf{H}}}^{j+1} = S_{\lambda/\rho^2}\left(\mathbf{Q}_i\widetilde{\mathbf{B}}_i^{j+1} + \widetilde{\mathbf{U}}^j\right)$ # soft threshold operator
   $\widetilde{\mathbf{U}}^{j+1} = \widetilde{\mathbf{U}}^j + \mathbf{Q}_i\widetilde{\mathbf{B}}_i^{j+1} - \widetilde{\overline{\mathbf{H}}}^{j+1}$
   $j = j + 1$
  end while
  $\mathbf{B'}_i^{k+1} = \widetilde{\mathbf{B}}_i^j$
 end for
 $\mathbf{MZ}_i\mathbf{Q}_i\mathbf{B'}_i^k$ is shared with peers or central node, who computes:

$$\overline{\mathbf{MZQB'}}^k = \frac{1}{N}\sum_i \mathbf{MZ}_i\mathbf{Q}_i\mathbf{B'}_i^k$$

 if $\|\mathbf{MY} - \overline{\mathbf{MZQB'}}^k\|_2 - \|\mathbf{MY} - \overline{\mathbf{MZQB'}}^{k-1}\|_2 > tol$ **then**
  Analyze individual errors, ban malicious agent(s) and re-start
 end if

$$\overline{\mathbf{H}}^{k+1} = \frac{1}{N+\rho}\left(\mathbf{MY} + \overline{\mathbf{MZQB'}}^k + \rho\mathbf{U}^k\right)$$

$$\mathbf{U}^{k+1} = \mathbf{U}^k + \overline{\mathbf{MZQB'}}^{k+1} - \overline{\mathbf{H}}^{k+1}$$

 $k = k + 1$
end while

Figure 8: Estimation of LASSO-VAR coefficients.

2.3.2 Innovation

While VAR models can be very effective for very short-term forecasting, their performance for larger forecasting horizons is poor as weather forecasts need to be incorporated, and the relationship between weather forecasts and power is nonlinear. To overcome this challenge, ENERSHARE proposes to enhance the previous solution from the H2020 Smart4RES project by

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

exploring the transformation of variables using spline and potentially kernel functions. Initial experiments have been conducted using two families of splines (natural cubic and B splines), demonstrating the promising potential of such data transformation techniques.

In addition, ENERSHARE will implement an API that facilitates collaboration among data owners, considering the algorithmic developments. Next sections focus on such implementation and describe the main functions (e.g., data encryption, model fitting and prediction), input/output data formats and implementation details.

2.3.3 Functions

The implementation of the multivariate time series forecasting approach described in the preceding subsections involves four main steps: initialization, data encryption, collaborative model fitting, and forecasting. Each component, and the set of required functions, are described below:

1. **Initialization**: This phase is crucial for identifying the participants and determining communication channels. In a conventional method, a seed node is introduced to initiate communication with all involved agents and record pertinent information or metadata, enabling peer-to-peer (P2P) communication.
   ENERSHARE Data Space will simplify these steps by using connectors in liaison with the metadata broker (more details below in Section 2.3.6).
2. **Data encryption** (data owners only): data owners cooperate to encrypt their data, according to the algorithm in Figure 7, and the following functions are needed:
   a. **Local data encryption** (local to each data owner), see line "1:" of Figure 8
      i. Generate random matrices $(\mathbf{C}_i, \mathbf{D}_i)$
      ii. Encrypt local data: perform the operation $\mathbf{W}_i$ = $(Data_i|\mathbf{C}_i) * \mathbf{D}_i$
      iii. Create a JSON file with dates and corresponding $\mathbf{W}_i$
      iv. Transmit JSON file (traditionally, to the first IP in the IP address book).
   b. **Global data encryption** (sequential), see lines "2:" to "11:" of Figure 7
      i. Generate random matrix $\mathbf{M}_i$
      ii. Operations with $\mathbf{M}_i$
      iii. Transmit JSON file.
   c. **Recover final encryption** (local to each data owner) see line "12:" of Figure 8
      i. Invert matrix $\mathbf{D}_i$
      ii. Operations with inverse of $\mathbf{D}_i$

3. **Collaborative model fitting** (P2P or Centralized): Once the data has been encrypted using the previous functionality, data owners can proceed with estimating their coefficients. This process involves performing local model estimation and aggregating the local models to create a global model, as depicted in Figure 8. However, there are

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

two important challenges that need to be tackled during the aggregation phase: asynchronous communication and the detection of malicious noise. In essence, the following functions must be considered:

  d. **Initialization** (data owners locally)
  e. **Local model update** (data owners locally)
  f. **Information exchange**, with two possible schemes:
    i. P2P: all agents exchange their local contributions with peers, without relying on a centralized authority.
    ii. Centralized: all agents involved exchange their local contributions with a central entity. This central entity acts as a coordinator or aggregator.
  g. **Convergence analysis** to detect malicious noise injection
    i. P2P: all agents locally analyse the local contributions they received to identify any instances of malicious noise injection originating from their peers.
    ii. Centralized: central entity analyses the received local contributions to identify any instances of malicious noise injection originating from agents.
  h. **Asynchronous communication**
    i. P2P: all agents locally consider the latest contributions from each peer.
    ii. Centralized: central entity considers the latest contributions shared by each agent.
  i. **Global aggregation**
    i. P2P: all agents locally aggregate the latest contributions from each peer.
    ii. Centralized: central entity aggregates the latest contributions submitted by each agent.

4. **Forecast** (P2P): The function mentioned earlier provides the encrypted coefficients for each agent, which are necessary for the subsequent prediction process. After obtaining these encrypted coefficients, the final step considers the trained model to make predictions. To achieve this, each agent performs calculations and shares its contribution with every other agent. Once all the contributions have been received, each agent can then obtain its forecast result.
  j. **Local contribution** (data owners locally)
  k. **Exchange of local contributions** (data owner > data owner)
  l. **Aggregation of local contributions** (data owners locally)

The sequence diagrams presented below depict the components 2, 3, and 4, represented by Figure 9, Figure 10, and Figure 11, respectively. These diagrams illustrate the sequential flow of processes and information within each component, providing a clearer understanding of their operations.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

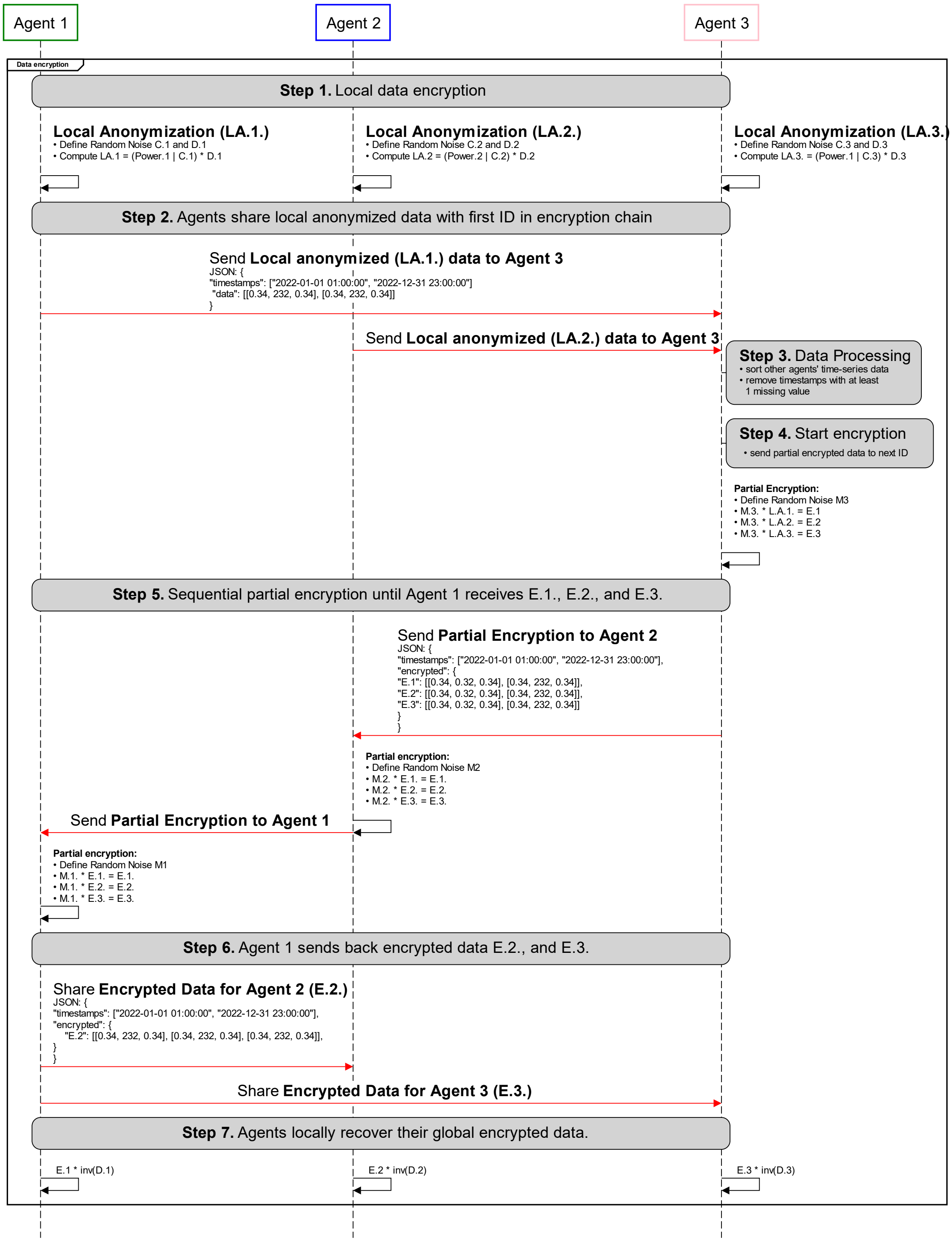


Figure 9: Sequence diagram of the Data Encryption function for the Federated Learning method.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

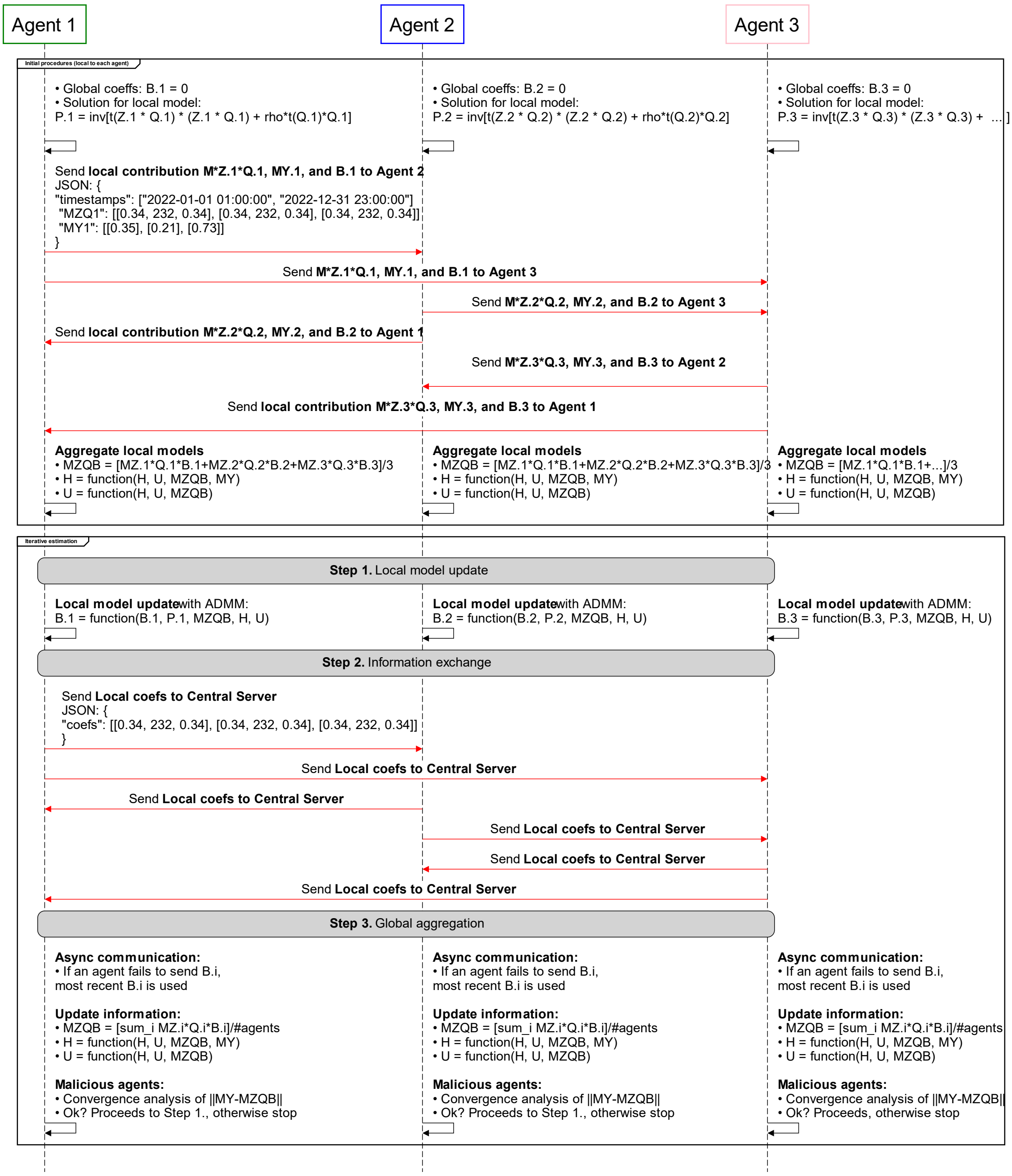


Figure 10: Sequence diagram of the Collaborative model fitting function for the Federated Learning method.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

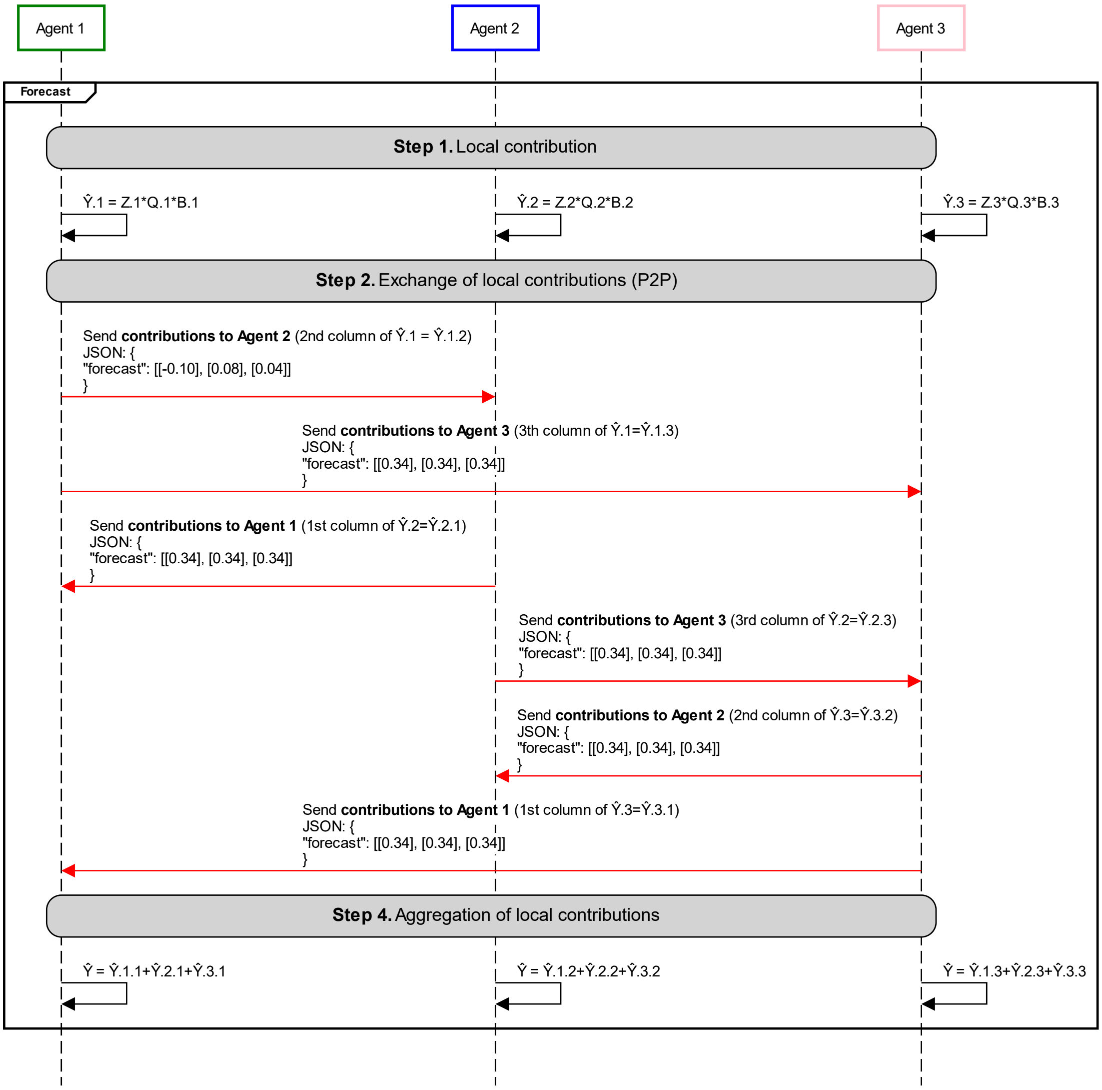


Figure 11: Sequence diagram of the Forecast function for the Federated Learning method.

2.3.4 Input and Output Data Format

Table 7: Input and Output Data Format for FL in Multivariate Time Series Forecasting

Functionality		Input/Output description	Variable type
Initialization	INPUT	• Peer API URL • API version	String String
	OUTPUT	• Error message • Ledger log	String List of strings

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Data Encryption	INPUT	• Timestamps • Local data anonymized • Random matrices	Integers Float Float
	OUTPUT	• Timestamps • Encrypted data	Integers Float
Coefficients	INPUT	• Encrypted data • Rho and lambda	Float Float (positive value)
	OUTPUT	• Encrypted coefficients	Float
Forecast	INPUT	• Decrypted coefficients and local features	Float
	OUTPUT	• Forecast	Float

2.3.5 Implementation details

So far, a proof of concept with real and large-scale data has been done with a Python script that simulates collaboration between multiple agents using power data from wind farms. However, the integration with the ENERSHARE Data Space was not yet done, so we should assume TRL 4 in what concerns the integration. Nevertheless, for this report, all the components necessary to build an API were identified and specified, so it is expected that by the second deliverable (D6.2) there will be a first API version implemented and tested with real wind power data from Norway, corresponding to TRL5. The target TRL (TRL 7) will be reached by D6.3, and demonstrated with data from Pilot 2.

The API will be developed using Python programming language (version 3.10) and distributed using a stack of Docker containers, i.e., a set of services that assures the multiple functionalities of the privacy-preserving forecast.

In terms of required libraries, Table 8 presents a non-exhaustive list of expected Python dependencies for this API.

Table 8. Python expected dependencies for the Federated Learning API.

Library	Description	URL	License
datetime	Basic date and time types	https://docs.python.org/3/library/datetime.html	Open Source
dotenv	Read .env file and set environment variables	https://pypi.org/project/python-dotenv/	3-Clause BSD

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

http	HTTP modules	https://docs.python.org/3/library/http.html	Open Source
itertools	Functions creating iterators for efficient looping	https://docs.python.org/3/library/itertools.html	Open Source
joblib	Pipelining tools	https://joblib.readthedocs.io/en/latest/	3-Clause BSD
json	JSON encoder and decoder	https://docs.python.org/3/library/json.html	Open Source
math	Mathematical functions	https://docs.python.org/3/library/math.html	Open Source
multiprocessing	Process-based parallelism	https://docs.python.org/3/library/multiprocessing.html	Open Source
NumPy	Scientific computing	https://numpy.org/	3-Clause BSD
os	Miscellaneous operating system interfaces	https://docs.python.org/3/library/os.html	Open Source
pandas	High-performance, easy-to-use data structures and data analysis tools	https://pandas.pydata.org/	3-Clause BSD
scipy	Fundamental algorithms for scientific computing in Python	https://scipy.org/	3-Clause BSD
statsmodels	Statistical models, Hypothesis tests, and Data exploration	https://www.statsmodels.org/	3-Clause BSD

2.3.6 Integration with ENERSHARE Data Space

Figure 12 depicts the use of data spaces for the implementation of the FL method discussed earlier. Under the ENERSHARE Data Space, this service and its API can be distributed as an App in the App Store, Figure 13 offers a more detailed representation. Different agents can, this way, participate in the forecasting process using their private data. By deploying this app in a data space connector, each agent starts communicating with the wider network of participants. The Metadata Broker in the Data Space would facilitate the initialization of the forecasting process by taking on the tasks (that would traditionally be assigned to a seed node) described in the Initialization function in the Functions section. Having guaranteed the preconditions for

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

the forecasting to be computed with control and privacy, the rest of the communication would occur in a peer-to-peer logic, and connectors between any two agents would provide the necessary infrastructure for data and information exchange. The IDS Reference Testbed under development provides offers various features such as the Connector, Metadata broker, CA (Certificate Authority), and DAPS (Dynamic Attribute Provisioning Service). These components work together to ensure secure and trusted data exchange between the Agents and the Data Space.

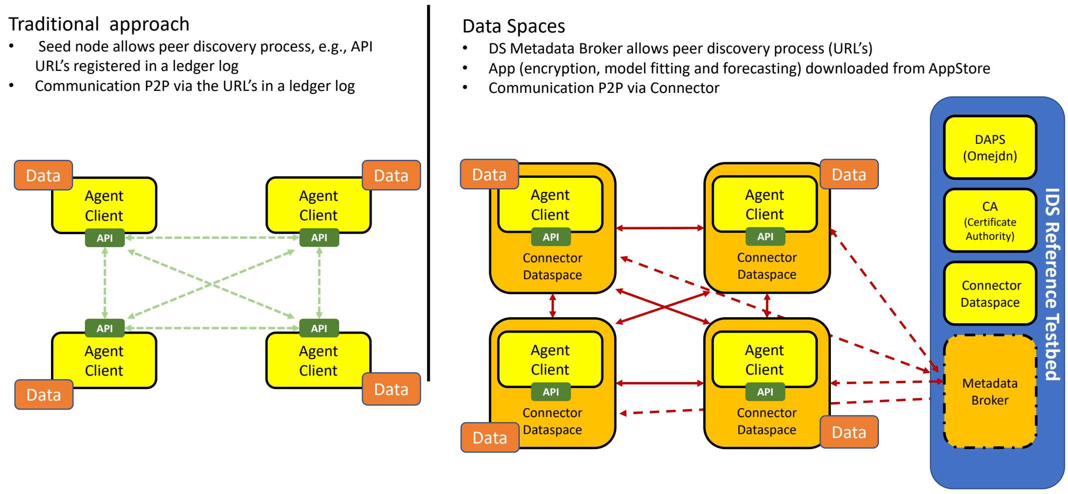


Figure 12: Architecture planned for the interaction between Agents and the ENERSHARE Data Space.

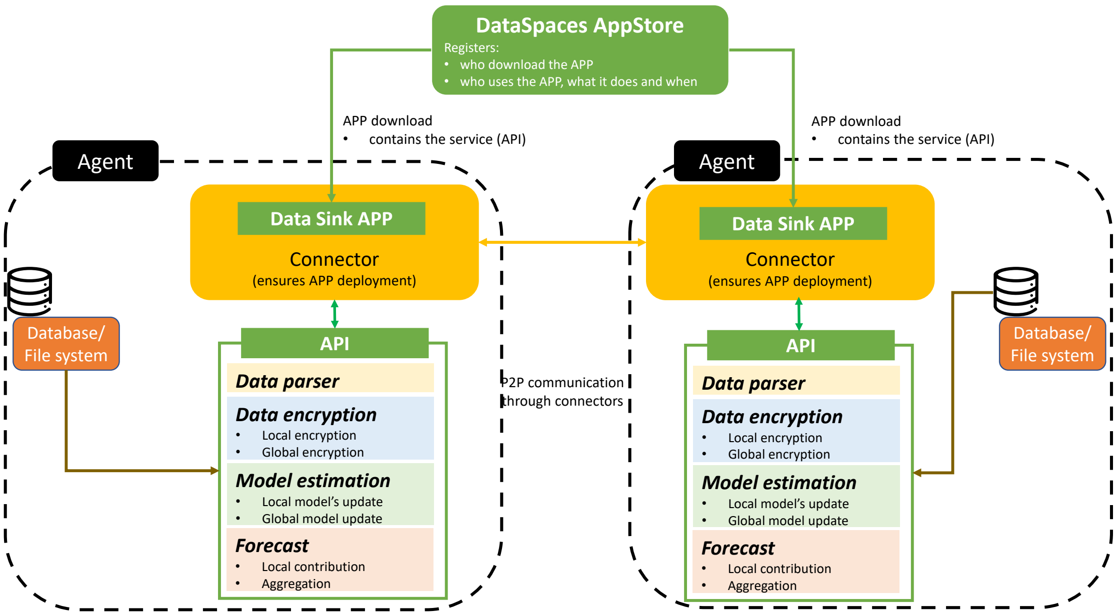


Figure 13: Agent interactions within the context of multivariate time series forecasting using Data Spaces.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2.4 Federated transfer learning flexibility potential assessment

2.4.1 Description of the FL approach

Federated learning is a machine learning approach that allows for training of models across multiple decentralized devices or servers without the need to transfer raw data to a central location. In federated learning, data remains distributed across multiple devices or data sources, and the model is trained locally on each device using its respective data. In this project we focus on decentralized data as federated learning approach. Since the data remains on the local devices or servers, it alleviates concerns about exposing sensitive or private information during the training process. Since majority of the time local machines do not possess enough data to train a suitably performing models, we utilize a transfer learning approach to build models. In federated learning, the training process occurs on distributed devices or servers, each with its own local dataset. Instead of training a model from scratch on each device, transfer learning allows for the initialization of the models using a pre-trained model, which provides a more effective starting point for learning on each local device or server. As can be seen in figure bellow, we build one general model through reference data, analyse the data, and transform it into features through feature store, and then train a model with the data and save it into model registry. To build our federated model, we take the same analysis scripts from feature store and apply them to the private/confidential data and load a general model from the model registry. The general model is then fine-tuned with private data to create a new tailored model that is served to the needed service.

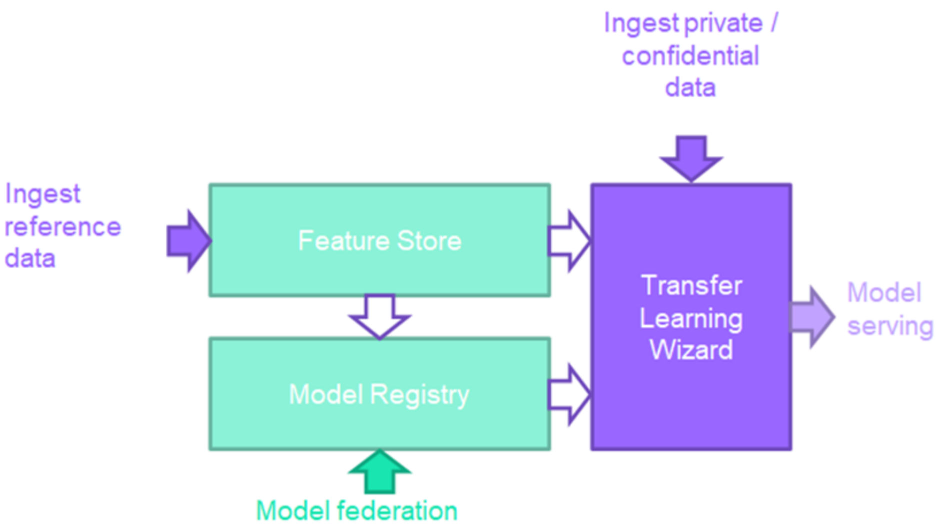


Figure 14: Federated Transfer Learning concept

Within the project we are building classification models on sub-station level. These models will be capable of detecting flexibility potential of the grid by detecting whether PVs, electric

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

vehicles, or heat pumps are present. The goal of the models is to be able to detect which parts of the grid can be included in mitigating potential congestions and enabling resilience in the grid. Additionally, the classification models will also be able to detect issues on the grid, such as voltage fluctuation. The classification is done using a deep learning model based on Long Short-Term Memory cells, convolutional neural networks, and multilayer perceptron, that can be easily trained in federated transfer learning method. Each substation within the project collects its own data and the models are trained locally on fixed time intervals.

2.4.2 Innovation

The innovation within this project is the combination of federated transfer learning and both services in flexibility potential assessment and issues detection in the grid. The said approach leverages the data privacy, while at the same time enables us to develop models that will enable more stable operation of the electrical grid system. The innovation is also a GUI application that can be easily run on any local machine, or at the end user, that can build their own model and subsequent service for their own home. Although the application runs in a browser, whole model training is done locally, while developed models and used data are never shared with anyone or transferred to any server. GUI application is an important part of democratising artificial intelligence (AI), an effort of making accessible and available to a wider range of individuals and organizations, beyond a select few who possess specialized knowledge or resources.

2.4.3 Functions

- **Input parsing function** – to parse the inputted data in csv, excel, or parquet format and pre-process the data for the model development.
- **Analysis function –** function that analyses the pre-processed data for errors, correlation coefficient and does mitigations to fix potential issues that would prevent the successful model training. It also automatically drops the columns without data, or if all the inputs in the column are equal to zero.
- **Model training function –** prepared data is then used in model training. The general model is downloaded and locally saved for the basis of the federated transfer learning approach. The said general model is fine-tuned and retrained to build a new model adapted to input data.
- **Export function** – trained model is then exported in a desired format (e.g., JSON, XML, MLEM) and can be used in the user's application.
- **Deployment function** – In case the user wants an automated process in an already built service, the deployment function is used instead of the export function. This function deploys the trained model into a dedicated service and tests its performance.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2.4.4 Input and Output Data Format

Service is run in the browser, but model training is done on the local machine. The input data into the service can be in csv, parquet, excel, or pickle format. The output data model is in either json, xml, mlem, or .pt format. The table contains the needed features for developing the flexibility potential assessment and gird issues detection. DateTime is used as an index for the data, Power, Voltage L1-L3 are used as features for the model, while Activity Label contains the array of labels whether the EV, PV, HeatPump, or similar are active at the certain DateTime in the time series. Similar is for the Issues label, where the label contains a binary label at a certain DateTime in the time series of Voltage L1-L3 if an issue is present. Additionally, user can add additional available features that are not listed in the table, that might help the final model prediction.

Table 9: Federated Transfer learning model development input

Variable	Unit	Type	Description
DateTime	-	datetime	Datetime of time series data
Power	W	float	Power flow collected at the substation meter
Voltage L1	V	Float	L1 voltage at the substation meter
Voltage L2	V	Float	L2 voltage at the substation meter
Voltage L3	V	Float	L3 voltage at the substation meter
Activity label	-	Array{int}	Array of labeled activities in TS data, that can be detected for flexibility potential assessment
Issues label (power)	-	int	Binary label of the interval of issues observed within the time series data
Issues label (voltage L1)	-	int	Binary label of the interval of issues observed within the time series data
Issues label (voltage L2)	-	int	Binary label of the interval of issues observed within the time series data
Issues label (voltage L3)	-	int	Binary label of the interval of issues observed within the time series data

2.4.5 Implementation details

The first version of the application is written using open-source Python libraries such as PyTorch, mlem, pandas, numpy, and scipy. The whole GUI part of the federated transfer learning model development application is written with the help of the open-source Dash Python library. The application is then dockerized for deployment on the server and can be accessed through a web browser. In Figure 15 we can see the first version of the application. In the first step, we upload the data into the application, which is then displayed to the user for his inspection. The application automatically also calculates the correlation map between the uploaded features. This can help the user to remove potential highly correlated features from their dataset. Features with high correlation are more linearly dependent and hence have

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

almost the same effect on the dependent variable. So, when two features have a high correlation, we can drop one of the two features and in return speeding the training process and robustness.

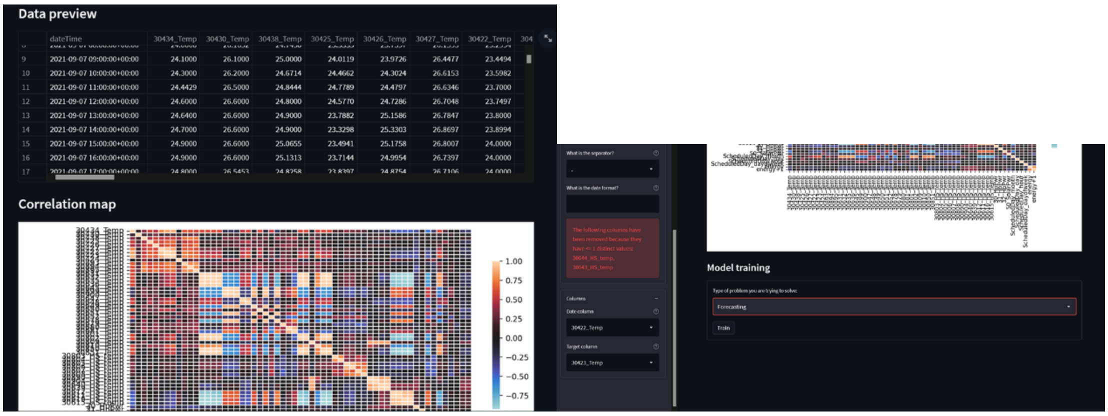


Figure 15: Federated transfer learning model development service

Similar to the model development application, also the flexibility potential assessment and issue detection services are written in Python with the same list of Python libraries. This application is also dockerized and will be deployed on local machines in the Pilot grid substations. Future steps in the development of the service will be to integrate the federated transfer learning model development directly into the service for automatic model development on fixed schedule. As it can be seen from the bottom figure, we have predicted behaviour of the grid in orange, while the anomalous readings are seen in red. In the initial application we are detecting issues in the power measurements, while the models for the voltage issue detection are still being developed.

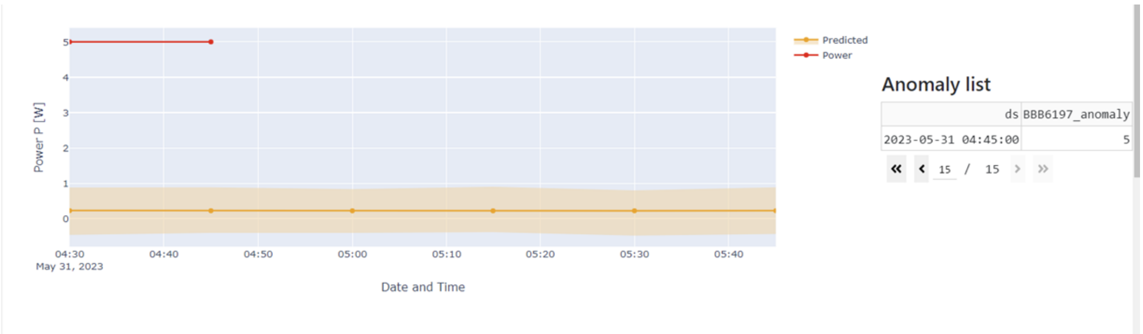


Figure 16: Grid issues detection service

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2.4.6 Integration with ENERSHARE Data Space

Services will integrate with data space as data consumers. Services adopt a data space compliant connector that will serve as interface to build a general model for federated transfer learning. Additionally, the developed generic models will be shared (transferred) between users through the data space to enable easier federated transfer learning and model fine-tuning.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3 Data-driven Energy Services

This Chapter outlines the data-driven energy services provided by ENERSHARE, targeting different stakeholders, namely:

- **Local communities:** data-based evaluation of the economic feasibility of sharing resources/assets business models; price estimation of local energy communities.
- **TSOs/DSOs:** improvement of net-load forecasting with crowdsensing; monitoring the growth of flexible resources, estimation and forecasting of electrical grid status.
- **Multi-energy utilities:** vector energy services (electricity, heat, and gas) by leveraging behind-the-meter/local communities' data.
- **RES value chain:** Data-driven power drive O&M algorithms; benchmarking of wind turbines or wind power plant behaviours.
- **End-user/consumer:** proxy models for studying energy efficiency actions; aggregation of flexibility for market purposes.

These services will be provided by energy stakeholders, who may collaborate with non-energy stakeholders such as facility operators, municipalities, social services, and original equipment manufacturers. Citizens can choose to share their personal data in exchange for new services or compensation. It is essential to note that ENERSHARE ensures data remains at the generation point, employing federated learning algorithms in some services. Data owners retain full control over their data, emphasizing trust and sovereignty as per WP4 guidelines. Integration with the Data Space (explained in a dedicated subsection) is a key aspect of these services.

3.1 Energy community sizing with assets sharing

3.1.1 Description of the Service

The deployment of Renewable Energy Communities (REC) enables energy sharing between its members, which reduces the energy exchanged between each member and its energy retailer and, consequently, in economic savings for the members. How the energy and resulting savings are shared between the members will depend on the chosen business model. The business model can also include the possibility for other stakeholders to invest and participate in the economic and social benefits obtained by the deployment and operation of the community.

This service, developed by INESC TEC, is based on a tool that allows the sizing of a renewable energy community by computing a Mixed-Integer Linear Problem (MILP) considering the following inputs:

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

1. The characteristics of each potential participating member: consumption profiles, photovoltaic, battery and other production/storage systems owned and their technical characteristics, retailer and grid access tariffs.
2. The chosen business model that dictates the objective function (e.g., maximize the self-consumed energy in the community, maximize the overall economic benefits, etc.) and the constraints of the problem.
3. The technologies characteristics and costs of the assets that will eventually be shared among the members and/or individually installed.

The results of this optimization are:

1. The size of each shared and individual asset, as well as its operation (production profile and charge/discharge for production and storage systems, respectively).
2. The energy shared between the members (or within the community), which includes the results of the market/contracts that allow for this sharing.
3. The members' energy bills.

Figure 17 provides an overview of the tool.

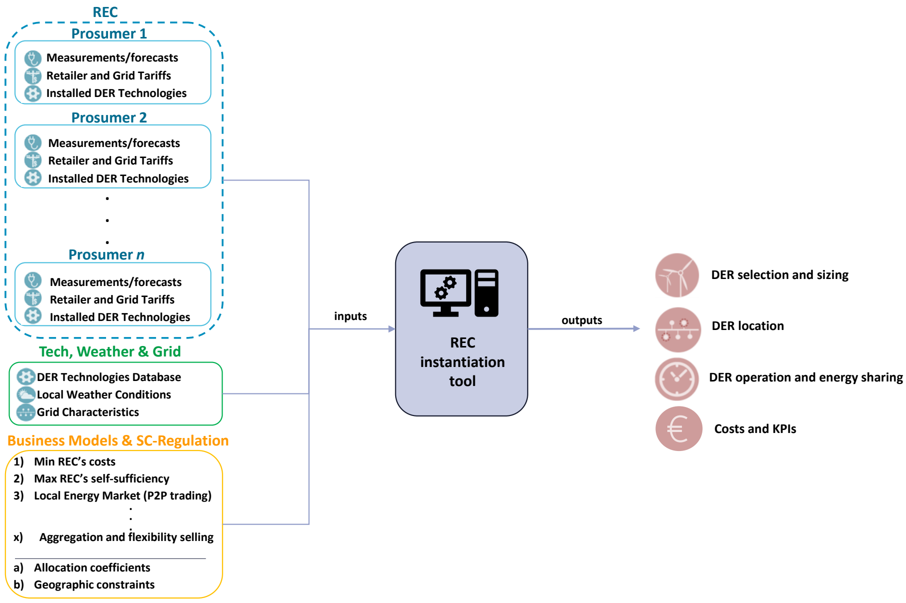


Figure 17: Energy community sizing service

3.1.2 Innovation

The tool demonstrates innovation by facilitating the comprehensive integration of several inputs and functions to enhance the optimization of a renewable energy community operation and sizing, namely: offering the capability to integrate both shared assets and individual assets,

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

including members that may not yet own them, considering their preferences; the possibility to select from a range of business models, considering the limited development and practical implementation of such models. These models enable the implementation of a broad range of characteristics that will dictate the community operation (and consequently, its sizing) considering not only members' (i.e., participating self-consumers) but also other stakeholders' preferences, as it is the case, for example, of a small and medium enterprise that wishes to manage and invest in community assets.

3.1.3 Functions

The following functions, described by the order in which they are executed, are essential to execute the whole process of sizing a community, considering the specificities of this project.

1. **Processing and storage of technology data** – in this first step, the data regarding the technologies (PV and storage) technical and cost-related characteristics are read from CSV/Excel files and stored in a DB.
2. **Read inputs from the database** – retrieving the data stored in the database: technology related, as described in the first step, but also other inputs data (see 3.1.4 for more info regarding this topic).
3. **Read service parameters via API** – request the user to fill in its service parameters (see 3.1.4 for more info regarding this topic) and read those parameters.
4. **Inputs Adjustment** – consists of executing calculations to adjust (e.g., normalization) some inputs so that they present the desired order of magnitude, units, etc.
5. **Model definition** – decision variables definition, mixed integer linear programming (MILP) instance creation, and respective formulation and insertion of constraints.
6. **Model execution** – saves and solves the proposed model.
7. **Export Results** – saves optimization outputs in optimal formats.

The MILP problem, which formulation is described in detail in Section 9.1, is, in short, composed by:

1. An objective function, to maximize or minimize some calculation (e.g., minimize the sum of the energy exchanged between each member and its retailer).
2. The restrictions:
    a. Energy balances, namely the members' net consumption, their individual exchanges with the grid/retailer as well as the physical and/or market energy exchanges between themselves.
    b. PV constraints, for both individual and shared systems, namely the energy output concerning its installed capacity and the maximum power capacity that can be installed.
    c. Storage/battery constraints for both individual and shared systems, including the maximum energy capacity that can be installed, the maximum and

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

minimum charge and discharge power, the maximum and minimum state-of-charge and one equation that computes the state-of-charge for each period.

The objective function, the physical and/or market energy exchange and the maximum power capacity that can be installed are highly dependent on the chosen business model, which, in turn, is very dependent on the members' preferences.

3.1.4 Input and Output Data Format

The inputs for this tool, shown in Table 10, can be divided into two main categories: service parameters and inputs data. Service parameters consist of inputs defined by the tool user for each time it runs it. Examples are the total time horizon of the simulation and chosen business model (not yet included in the table below for the reasons presented in 3.1.5). So, these parameters can be used to see the different outcomes presented by the tool. Hence, these parameters must be fulfilled every time the user wants to run the sizing problem. On the other hand, input data are fixed and do not change every time we want to run the sizing problem.

Table 10: Energy community sizing service: Input data specifications

Description	Units	Variable Type	Format Type	Type of input
Total time horizon of the simulation (service parameters)	days	Integer	JSON	Service Parameter
Duration of each timestep	minutes	integer	JSON	Service Parameter
ID of each installation	-	string	JSON	Input Data
Power limit for transactions at the Point of Common Coupling	kW	float	JSON	Input Data
Information provided by the retailer for each prosumer: (1) Buy and sell energy tariffs (input data) (2) Contracted power tariff (3) Contracted power (4) Voltage level	(1) €/kWh (2) €/(kW.day) (3) kW (4) -	(1) time series(float) (2) float (3) float (4) string	JSON	Input Data
Energy consumption diagram (for each time step)	kWh	time series(float)	JSON	Input Data
Storage technologies specifications: (1) Maximum and minimum state-of-charge (2) Maximum charge and discharge power (inverter) (3) Charge and discharge efficiency (4) lifespan (5) investment costs (6) members' ownership	(1) % (2) kW (3) % (4) years (5) €/kWh (6) %/member	float	Sourced from CSV/Excel file. Stored in a DB	Input Data

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

PV technologies specifications: (1) efficiency (2) yearly degradation (3) generation profile (4) lifespan (5) investment costs (6) members' ownership	(1) % (2) % (3) kWh/kW$_{installed}$ (4) years (5) €/kW (6) %/member	float	Sourced from CSV/Excel file. Stored in a DB	Input Data
(1) pre-installed, maximum and minimum storage to be installed (2) pre-installed, maximum and minimum PV to be installed	(1) kWh (2) kW	float	JSON	Input Data
Grid access tariffs (provided by DSO)	€/kWh	time series(float)	JSON	Input Data

All the outputs are presented in JSON. The following table presents their characterization.

Table 11: Energy community sizing service: Output data specifications

Description	Units	Variable Type	Format Type
Total cost of REC	€	float	JSON
Batteries initial energy content	kWh	list(float)	JSON
Batteries energy content	kWh	list of time series(float)	JSON
Batteries charged and discharged energy	kWh	list of time series(float)	JSON
Net consumption behind-the-meter	kWh	list of time series(float)	JSON
PV generation diagram	kWh	list of time series(float)	JSON
Imported and exported energy from the retailer	kWh	list of time series(float)	JSON
To be installed PV power	kW	list(float)	JSON
To be installed Battery capacity	kWh	list(float)	JSON
Contracted power (maximum metered power flow)	kW	list(float)	JSON
Total Installed PV power	kW	list(float)	JSON
Total Installed Battery capacity	kWh	list(float)	JSON
Allocated generation diagram from shared PV	kWh	list of time series(float)	JSON
Self-consumption energy	kWh	list of time series(float)	JSON
Energy consumed from the community	kWh	list of time series(float)	JSON
Energy purchased from the community	kWh	list of time series(float)	JSON
Energy sold to the community	kWh	list of time series(float)	JSON
Pool price of the community or price of the bilateral transaction (dependent on the business model)	€/kWh	time series(float) or list of time series(float)	JSON

3.1.5 Implementation details

For this phase of the project, a TRL5 is already implemented, considering that, apart from the shared batteries and the practical implementation of a range of business models, all the other functions have already been implemented, tested, and approved with success. Although the shared batteries have not been tested yet with the current model formulation, they were

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

previously tested, before the formulation evolved to its current state. Furthermore, this tool has already been used to formulate some real-world problems in our workspace, meaning different business models and configurations (that impact the constraints of the problem), which gives us a high level of confidence that its full implementation and testing will be done in a few weeks.

For the following deliverables, improvements on the practical utilization of the tool will be presented, namely the capacity to choose from a range of predefined business models and also the capacity to assess the inputs received, that is, the compliance with certain criteria (e.g., retailers' tariffs being higher than zero; otherwise, a warning/error is raised). Given this perspective, the next phase will introduce a tool characterized by a more refined and professional look. Also, improvements in the formulations may be presented, as the authors continue their work on the mathematical formulation.

When or if PV generation profiles are not available in the data provided by SEL, the European tool PVGis is used. An open-source API for sending and retrieving the desired PV data is already available PVGIS[3].

The technical and cost data for PV and storage assets is sourced from the Open Energy Data Initiative[4], available in CSV/Excel files, and stored in a DB.

This tool was developed in Python 3.9 using several internal and external libraries, as shown in Table 12, using *PuLP* as the framework for optimization modelling. For this work, the model was tested with the CBC (COIN-OR Branch and Cut solver pre-built in Pulp), under the license 3-Clause BSD.

Table 12: Energy community sizing tool open-source Python libraries

Library	Description	URL	License
copy	Shallow and deep copy operations	https://docs.python.org/3/library/copy.html	Open Source
datetime	Basic date and time types	https://docs.python.org/3/library/datetime.html	Open Source
itertools	Functions creating iterators for efficient looping	https://docs.python.org/3/library/itertools.html	Open Source
json	JSON encoder and decoder	https://docs.python.org/3/library/json.html	Open Source
loguru	Logging library	https://loguru.readthedocs.io/en/stable/	MIT License

[3] Hub, EU Science. PVGIS Online Tool. [Online] European Commission, 01 March 2022. https://joint-research-centre.ec.europa.eu/pvgis-online-tool_en

[4] (OEDI), Open Energy Data Initiative. 2022 Annual Technology Baseline (ATB) Cost and Performance Data for Electricity Generation Technologies. [Online] National Renewable Energy Laboratory (NREL), 20 February 2023. https://data.openei.org/submissions/5716.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

math	Mathematical functions	https://docs.python.org/3/library/math.html	Open Source
Matplotlib	Creating static, animated, and interactive visualisations	https://matplotlib.org/	3-Clause BSD
multiprocessing	Process-based parallelism	https://docs.python.org/3/library/multiprocessing.html	Open Source
NumPy	Scientific computing	https://numpy.org/	3-Clause BSD
os	Miscellaneous operating system interfaces	https://docs.python.org/3/library/os.html	Open Source
pandas	High-performance, easy-to-use data structures and data analysis tools	https://pandas.pydata.org/	3-Clause BSD
PuLP	Linear Programming (LP) modeler	https://pypi.org/project/PuLP/	3-Clause BSD
re	Regular expression (REGEX) operations	https://docs.python.org/3/library/re.html	Open Source
time	Time access and conversions	https://docs.python.org/3/library/time.html#module-time	Open Source

This tool is prepared to extract graphical outputs that demonstrate the results in a user-friendly way (e.g., plots). Additionally, it is believed that the tool is incorporated into a platform featuring a graphical user interface.

3.1.6 Integration with ENERSHARE Data Space

Services as fully qualified applications (i.e., with their own private data sources, workflows, decision processes and interfaces) relate with a data space as data producers, data consumers, or both. Services adopt a data space compliant connector that will serve as interface. With each service equipped with a data space connector, service to service interoperable data exchange is enabled.

From the perspective of services, they should consider and adopt a data space compliant connector that is linked with the minimum viable data space components. Services should then ensure a series of steps to join the data space, namely:

- Identify the interfaces and data that is expected to be made available through the data space.
- Link the identified data interfaces and the data concepts they represent considering the ontologies in used in ENERSHARE (WP3).
- Identify, select and apply the relevant usage policies that can bound the limits to data usage (WP4).
- Interface the legacy interfaces of services with a data space connector by creating a data app acting as a data sink or deploy a data app enclosing the service logic – smart data app.

Beyond the components required to deploy minimum viable data space, there are optional components such as the App Store (to distribute available apps within the data space and

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

sponsor reusage of service capabilities) or the Clearing House to monitor and log usage patterns to actively (e.g., block) or passively (i.e., after exchange non-repudiation guarantees) that should be considered in the next releases. After services integrate their interfaces and become data ready, they are required to explore the data space, in order to find the counter-party services they wish to interact with, by checking the data available and establishing a link between connectors.

3.2 Flexibility modelling of thermoelectric water heaters

3.2.1 Description of the Service

Through a very reduced dataset, it is possible to carry out processes of optimization thermoelectric water heaters (EWH) functioning calendar. In this work, the objective of this data-driven service developed by INESC TEC is to estimate periods where it is possible to provide flexibility, by optimizing the operation of the EWH, using only the appliance consumption data as input. For this purpose, only with EWH electrical consumption data, it is possible to determine the periods in which the EWH should start operating, guaranteeing (with some restrictions), comfort levels defined by the consumer, while minimizing the cost of operation or just the total consumption.

Although it is possible to use only the EWH consumption data, for accuracy purposes it is suggested that a historical calendar of uses and durations be used, so that the estimated usage is as close as possible to reality. Using only consumption data allows an estimate calendar to be extracted from this historical diagram, however, as the EWH remains in operation even after the end of the use of hot water (to restore the original temperature), it is not possible, in a direct way, to establish a link between the end of use and the appliance load diagram. In this regard, there are three options:

1. Conversion of the EWH load diagram to a binary usage calendar through typical mapping of the duration of showers and other usages. Its internal process defines a typical duration for baths and singular uses and applies this duration starting on every EWH activation.
2. Conversion of the EWH load diagram to a binary usage calendar based on survey data relating to personal usage, by user. Its internal process uses survey data of quantity, duration, and placement of average usages, and applies this duration starting on every EWH activation.
3. Direct supply of binary calendar of hot water usage.

In addition to this basic information, there is still an additional set of data that must be provided. Regarding estimates, a time series dataset that contains different information for each time

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

instant should be used. In addition, various technical data relating to the specific model under study are required. Here, concerning the first, the model is fed with a dataset that it must include as input:

1. **Inlet/network water temperature**. If it is not possible to estimate/measure the inlet water value, and if it is necessary to use a typical value, this variable must be included as a constant (or as defined by the data provider/collector) for all timestamps.
2. **Hot water usage**. This variable must represent, in binary form, the schedule of hot water usage during the considered period. This variable could be sourced in three ways, as explained above. It could be extracted through the historical load diagram provided as input to the problem. Its internal process defines a typical duration for baths and singular uses, and applies a reference for the number of usages, both based on questionnaires carried out with the user (only if not possible to extract this variable directly).
3. **Pricing**. Even if the model is prepared to perform optimization of operation respecting comfort, minimizing consumption of the device, the fundamental objective will be to minimize costs, which is why a variable with the price of energy for each instant of time must also be associated.

Regarding technical data, the following data is required:

- Heating power of the EWH (kW)
- Height of the EWH (cm)
- Internal water capacity of the EWH (l)
- Maximum allowable water temperature of the EWH (°C)
- Minimum allowable water temperature of the EWH (°C)
- Overall heat transfer coefficient of the EWH (kW/(m^{2}K))
- EWH room ambient temperature (estimate in °C)

Finally, there is a need to define the comfort temperature limit, which must be done through a user survey. If no value is provided, a typical value (40°C) is used.

3.2.1.1 Model Operational Framework

The formulation of the model is described in detail in section 9.2, but its general operation framework is based on the following process:

1. For a given use of hot water that occurs during periods t and $t+1$, it must be ensured that at $t+2$ the temperature of the water stored in the EWH ($temp_t^{EWH}$) is equal to or greater than the comfort temperature ($\widehat{temp}^{set}$). In this sense, it is automatically guaranteed that during use $[t, t+1]$, the temperature is never lower than the user-defined threshold.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2. For this to be respected, there is a need to estimate the minimum accumulated energy required, equivalent to the comfort temperature ($W^{comfort}$).
3. In addition, an expression that calculates the total equivalent energy accumulated in the EWH after a given use of hot water (W_t^{mix}), must be created. To this end, it determines the total energy accumulated after using hot water, through a fluid mixing formula ($\widehat{flow}^{Total} = flow_t^{Inlet} + flow_t^{EWH}$), since the hot water used is usually the result of mixing the water from an EWH and the "cold" water from the network, in order to minimize losses. There is only a "direct connection" between the EWH and the actual output, in cases where the user's hot water temperature requirement/comfort ($\widehat{temp}^{set}$), is above the stored/available temperature ($temp_t^{EWH}$). This expression is dependent on the values of the inlet and internal water temperature of the EWH, at each instant.
4. To guarantee comfort, the value of the final energy accumulated after use (W_t^{mix}), must be equal to or greater than the minimum value of equivalent comfort energy ($W^{comfort}$).
5. Even wanting to guarantee comfort, based on the user defined limits, it is not always possible to do so. In cases where there is a substantial and inconsequential use of hot water, it becomes virtually impossible to guarantee that there is hot water for the entire horizon of use. In this sense, to ensure that the optimization model always presents a valid solution, a penalty variable ($penalty_t^{comf}$) is included here, which, in turn, is added with a positive sign in the objective function. This variable aims to penalize excessive temperature reductions, thus allowing the model to operate beyond the established threshold.
6. In addition to energy losses by usage (represented in a grand total by W_t^{mix}), and gains by heating (W_t^{in}), the model also considers the value of energy losses by convection (W_t^{loss}), through its surface area.
7. Finally, the total operating cost is and minimized through the proportion of EWH heating usage (W_t^{in}), and respective price ($cost_t^{use}$).

In general, it is possible to simplify and represent the presented model as follows:

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

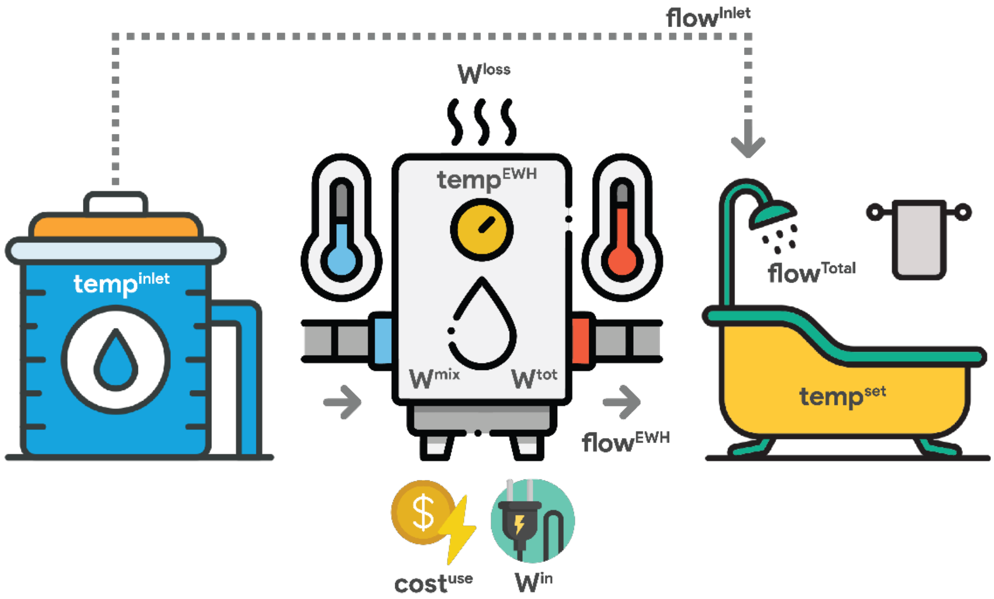


Figure 18: EWH Flexibility Optimization Model Framework

3.2.2 Innovation

This model presents a substantial benefit for use in conjunction with other energy sectors, namely the possible placement in energy communities. The EWH, due to their operational specifications, are capable of functioning as one-way batteries, that is, they can be used as reception points for excess energy that has been produced elsewhere in a cooperative setting. Through the temporal mapping of the periods where it is possible to provide flexibility, the EWH can be used as an energy recipient, heating the water in a period prior to the optimum, but ensuring that the excess energy is not wasted. In these cases, since the water would be being heated before the optimal period (i.e., immediately before its use, minimizing thermal losses), it will have to be heated above the theoretical upper limit, increasing the total energy consumption, but reducing the operational cost, since that it is being supplied by the community pricing.

This model also includes a lot of detail in its formulation, including tools for estimating the accumulated internal energy through fluid mixing formulas, calculating equivalent accumulated energy resulting from the mixture of hot water with the inlet water. In addition, heat losses are also formulated considering the surface area of the EWH, its thermal characteristics, and ambient temperature, and not just considering the thermal losses of the water in open air.

3.2.3 Functions

This subsection provides a description of the developed features and respective approaches in Python programming language.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

8. **Data reading and processing** – several reading and handling tasks are carried out, namely the construction of various auxiliary variables for information processing and conversion into optimal formats.
9. **Input parameters** – defines input parameters/variables, namely constants and user-defined values, and that are not included in external files read in (1).
10. **Regressors formulation** – linearization tasks related to non-linear constraints are performed, namely through linear regression models based on several variables and inputs obtainable via (1-2).
11. **Model definition** – decision variables definition, mixed integer linear programming (MILP) instance creation, and respective formulation and insertion of constraints.
12. **Model execution** – saves and solves the proposed model.
13. **Export Results** – saves optimization outputs in optimal formats.

3.2.4 Input and Output Data Format

For this work, the model requires an input dataset and a DB/CSV file with several time series. Regarding the former, data from the EWH specifications are needed, as well as user-defined data regarding the expected level of comfort. The DB/CSV file must contain data on the temperature of the inlet water, the expected usage schedule calendar, and the energy pricing.

Table 13: EWH optimization model input data specifications

Variable	Source	Value	Default Value	Unit	Type	Description
flow_rate_min	Input	-	8	kg/min	float	Water usage flow rate (fixed/constant value).
ewh_max_temp	Input	-	90	°C	float	EWH internal water maximum allowed temperature
ewh_height	Input	-	100	cm	float	EWH Height
ewh_capacity	Input	-	100	l	float	EWH Capacity/Volume
ewh_power	Input	-	2.5	kW	float	EWH Functioning Power
heatTransferCoeff	Input	-	0.00125	kW/(m^2.K)	float	EWH overall heat transfer coefficient
ambTemp	Input	-	20	°C	float	EWH room ambient temperature
tempSet	Input	-	40	°C	float	User-defined comfort temperature for hot water usage
bigNumber	Input	-	1E+05	-	integer	Sufficiently big number (e.g., 1E5 if EWH's nominal capacity is in the tens of kWh)
dataset	DB/CSV	-	-	-	dataframe	Input dataset containing 4+n variables (timestamp, temp_inlet, delta_use, price[1-n])
timestamp	DB/CSV	-	-	Y-m-d H:M:S	datetime/ list	Timestamp of the specific period
temp_inlet	DB/CSV	-	15	°C	float/list	Measured inlet network water temperature

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

delta_use	DB/CSV	[0,1]	-	-	binary/list	Binary parameter indicating periods of hot water usage (1 = using)
price[n]	DB/CSV	-	-	€	float/list	Energy Prices for each observation

Regarding the output data, there are two possibilities:

1. If the estimation/optimization of flexibility is carried out in a block, a CSV file is extracted with time series for the whole period, that include values of:
    a. EWH internal water temperature (°C)
    b. EWH equivalent EWH total energy (kWh)
    c. EWH equivalent mixture energy after usage (kWh)
    d. EWH heating power used (kWh)
    e. EWH equivalent thermal losses (kWh)
    f. EWH heating module operation status [0-1]
    g. EWH available flexibility calendar [0-1]
2. If the optimization problem is carried out in a sliding window, iteratively and with the flexibility schedule updated (with external incentive), the flexibility calendar is calculated for the entire period under study, in block, and then being updated at each iteration, according to the external incentive needs. In the end the extracted data is the same.

3.2.5 Implementation details

This service is currently at TRL5 level. Its core operation is fully implemented and functional. It is prepared for its execution in block, i.e., for estimating flexibility in fixed periods in whole. For it to be integrable on a shared energy community platform, the implementation must be adjusted to be compatible with running in a sliding window. For this purpose, it is necessary to include in its preparation and execution a certain set of external incentives, (variable prices, excess energy from the community, etc.), provided by a third party, so that the estimation of flexibility can be done iteratively, updating the schedule of the remainder flexibility in a sliding way.

It is also estimated that the tool is included in a platform with a graphical user interface and that it can extract graphical outputs that present the results in a visibly more convenient and comprehensive way (e.g., plots).

This model was developed in Python 3.10 using several internal and external libraries, as shown in Table 14.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Table 14: EWH optimization model open-source Python libraries

Library	Description	URL	License
re	Regular expression (REGEX) operations	https://docs.python.org/3/library/re.html	Open Source
datetime	Basic date and time types	https://docs.python.org/3/library/datetime.html	Open Source
math	Mathematical functions	https://docs.python.org/3/library/math.html	Open Source
NumPy	Scientific computing	https://numpy.org/	3-Clause BSD
pandas	High-performance, easy-to-use data structures and data analysis tools	https://pandas.pydata.org/	3-Clause BSD
Matplotlib	Creating static, animated, and interactive visualisations	https://matplotlib.org/	3-Clause BSD
scikit-learn	Tools for predictive data analysis	https://scikit-learn.org/	3-Clause BSD
PuLP	Linear Programming (LP) modeler	https://pypi.org/project/PuLP/	3-Clause BSD

This EWH optimization model has been developed in Python, mainly using *PuLP* as the main framework for optimization modelling. *PuLP* is a widespread open-source linear programming library that presents a user-friendly interface to define and solve linear programming problems. For this work, paired with this library, several solvers were tested, to guarantee that the problem runs smoothly, and in the most efficient way possible. Thus, the model was tested with the CBC (COIN-OR Branch and Cut) and *HiGHS* solvers (Huangfu and Hall, 2018), which are open source, and with the IBM ILOG CPLEX solver, that requires a commercial license. For more details, check Table 15.

Table 15: EWH optimization model tested solvers

Solver	Description	URL	License
CBC	COIN-OR branch and cut solver pre-built in PuLP	https://coin-or.github.io/	3-Clause BSD
HiGHS	Open-source serial and parallel solvers for large-scale sparse linear programming (LP), mixed-integer programming (MIP), and quadratic programming (QP) models	https://highs.dev/	MIT License/Open-Source
IBM ILOG CPLEX	High-performance optimization solver for linear, mixed-integer and quadratic programming	https://docs.python.org/3/library/math.html	Commercial

3.2.6 Integration with ENERSHARE Data Space

The integration with the Data Space will follow the same strategy described in Section 3.1.6.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.3 Community market pool to estimate energy price of internal transactions

3.3.1 Description of the Service

This service, developed by INESC TEC, aims to estimate a price for the internal transactions within a Renewable Energy Community (REC) and Citizens Energy Communities (CEC) that promotes self-sufficiency and a reduction of the costs for each of its members. A particular use case for this tool consists of encouraging vulnerable consumers to be actively enrolled in the energy transition by being offered more competitive prices thanks to their participation in an energy community. On the other hand, the tool can be used to support the study of emerging business models. The optimal price is computed for every step of a given horizon based on the establishment of a market pool between members/shared resources with their own meter and energy contract with a retailer. Several market pool mechanisms will be considered and studied.

To establish the community market pool, regardless of the mechanism for price computation chosen, members/resources that are expected to have a positive net load (i.e., a forecasted consumption from the main grid) are considered “buyers”, and members/resources with a negative net load (i.e., a forecasted injection to the main grid) are considered “sellers”. Each participant can either be a member with behind-the-meter load and/or generation and/or storage assets, or a shared resource, as long as it has its own meter and an energy contract with a retailer. The net load and opportunity costs with the respective retailers can be used as the buying offers of the “buyers” and selling offers of the “sellers”, respectively. The opportunity costs can be represented by the peers’ contracted buy prices and sell tariffs with their respective retailers, or by any custom, pre-defined function based on their expected profit margins.

The service can either be used on a pre-delivery timeframe (e.g., day-ahead) or on a post-delivery timeframe. In the first case, the net loads are still not completely known, since on one hand they are based on generation and demand forecasts but, more importantly perhaps, the controllable assets, i.e., storage assets, can still be scheduled. On both timeframes, the service will calculate the optimal schedules and internal transactions that can be potentially established, either by indicating how much each member could buy/sell at the local energy pool market, or by establishing direct peer-to-peer (P2P) bilateral transactions that could take place. Notice that the objective of this tool is not to provide optimal set points for REC or CEC management, but to support the study of business models and the integration of vulnerable consumers in the communities. Table 16 resumes the applicability of the service to both timeframes.

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

Table 16: Applicability of the energy prices estimation tool to different timeframes and market paradigms.

Time frame	Market paradigm	Application
Pre-delivery	Pool	• Storage assets' scheduling; • Schedule of internal transactions per peer (i.e., total energy to buy and sell at the local energy market that can minimize REC costs) – implies that public grid usage tariffs are socialized and can't be minimized; • Compute the optimal energy price for those local transactions.
	Bilateral contracts	• Storage assets' scheduling; • Schedule of P2P bilateral contracts (i.e., the scheduled energy transactions between peers that can minimize REC costs) – public grid usage tariffs are considered and can be minimized; • Compute the optimal energy price for those local transactions.
Post-delivery	Pool	• Definition of internal transactions per peer that minimize REC costs based on their metered net load (i.e., total energy bought and sold at the local energy market) – implies that public grid usage tariffs are socialized and can't be minimized; • Compute the optimal energy price for those local transactions.
	Bilateral contracts	• Definition of P2P bilateral contracts that minimize REC costs based on their metered net load (i.e., the effectively established energy transactions between peers that can minimize REC costs) – public grid usage tariffs are considered and can be minimized; • Compute the optimal energy price for those local transactions.

The service will consider several market pricing mechanisms:

- **Mid-market-rate (MMR)**
  - A mechanism that considerers the maximum selling ($max_n(\lambda_{n,t}^{sell})$) and buying prices ($min_n(\lambda_{n,t}^{buy})$) for each instant t among all participating members n to compute the internal transactions' price λ_t:

$$\lambda_t = \frac{max_n(\lambda_{n,t}^{sell}) + min_n(\lambda_{n,t}^{buy})}{2}$$

    A variation of this equation will also be made available (**PARTIAL_MMR**), where only the prices of members with accepted offers in a traditional pool are considered.

- **Supply and demand ratio (SDR)**
  - A mechanism that considers a ratio between total supply $\sum_n(E_{n,t}^G)$ and total demand $\sum_n(E_{n,t}^L)$ within the REC, to which the following logic is applied:

$$if \quad 0 < \frac{\sum_n(E_{n,t}^G)}{\sum_n(E_{n,t}^L)} < 1 \; then:$$

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$$\lambda_t = \frac{min_n(\lambda_{n,t}^{buy}) \cdot \left(max_n\left(\lambda_{n,t}^{sell}\right) + \lambda^{comp}\right)}{\left(min_n(\lambda_{n,t}^{buy}) - max_n\left(\lambda_{n,t}^{sell}\right)\right) \cdot \frac{\sum_n\left(E_{n,t}^{G}\right)}{\sum_n\left(E_{n,t}^{L}\right)} + max_n\left(\lambda_{n,t}^{sell}\right)}$$

$$if \;\; sdr_h \geq 1 \quad then:$$

$$\lambda_t = max_n\left(\lambda_{n,t}^{sell}\right)$$

The possibility to consider a compensation factor, $0 \leq \lambda^{comp} \leq min_n\left(\lambda_{n,t}^{buy}\right) - max_n(\lambda_{n,t}^{sell})$ will be added (a variant called **compensated SDR** or **SDRC**).
Also here, a variation of this equation will be made available (**PARTIAL_SDR/ PARTIAL_SDRC**), where only the prices of members with accepted offers in a traditional pool are considered.

- **Pre- / Post-delivery pool (POOL)**
  - Classical pool equilibrium computed by intersecting the aggregated buying and selling curves built from the buying and selling bids of the REC. Note that this mechanism can be applied to either market paradigm considered, i.e., pool or bilateral contracts, since it is only used to compute the prices and does not define the market paradigm itself.

Further information on these pricing mechanisms applied on the context of REC can be found in (Mello et al., 2023).

The service is meant to be used for estimating prices on REC where members may have storage assets. On a post-delivery timeframe, the usage of those assets is reflected on the metered net load of the respective peers. On the other hand, on a pre-delivery timeframe, the usage of such assets needs to be the subject of optimal scheduling. For that reason, an overarching iterative algorithm can be implemented where prices are computed with any chosen pricing mechanism after an optimal energy management MILP is run. The formulation of this MILP considers an adaptation of the first two stages approach previously published by INESC TEC researchers in (Rocha et al., 2023) for the optimal REC management. The formulation is here adapted to consider a pool-like internal market instead of peer-to-peer exchanges. In the two-stage MILP approach, the first stage is run to compute the individual results that are then used as baselines on a second stage corresponding to the collective optimization of the REC, to guarantee that the cost of all members of a REC is not impaired by their participation on the local market. The MILP will consider all members and resources, and their respective net loads and opportunity costs, and will minimize the total REC costs, dispatching the available individual and collective storage assets, to determine how much energy is traded locally based on the internal

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

transaction prices previously computed. By changing the schedules of the storage assets, new net loads are computed for each REC member, leading to a new expected pool price. Each time a new pool price is computed, a new MILP can be run to recompute the storage assets schedules, until a stopping criterion is met such as price stabilization or maximum number of iterations. The flowchart for the overarching algorithm used is illustrated in Figure 19.

Figure 19 also highlights an alternative approach based on the collective REC optimization, where the internal price is computed as the dual variable of the internal market equilibrium constraint of the MILP problem. Indeed, the internal market equilibrium constraint states that the sum of all internal sales $E_{n,m,t}^{SALE}$ must equal the sum of all internally purchases $E_{n,m,t}^{PUR}$, for each time step t (also called interval settlement period or ISP), and is formulated as follows:

$$\sum_{n,m} \left(E_{n,m,t}^{PUR}\right) - \sum_{n,m} \left(E_{n,m,t}^{SALE}\right) = 0$$

Through this approach, which we call "**POOL_DUAL**", the iterative algorithm is no longer needed, and a single-stage MILP approach is used. However, this approach does not allow to impose the constraint from the first stage to guarantee that no member sees a lower benefit than it would see by optimizing its energy behaviour individually, since the prices are no longer a variable in the problem. Although several tests already performed (Silva et al., 2023), show a good performance of this approach in terms of fair share of the collective benefits, with this approach it cannot be guaranteed that the first stage constraint always holds. Also note that this approach, unlike the iterative one, cannot be used under a market paradigm where bilateral contracts are established since under that market paradigm, the equation above is substituted by another:

$$E_{n,m,t}^{PUR} - E_{m,n,t}^{SALE} = 0$$

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

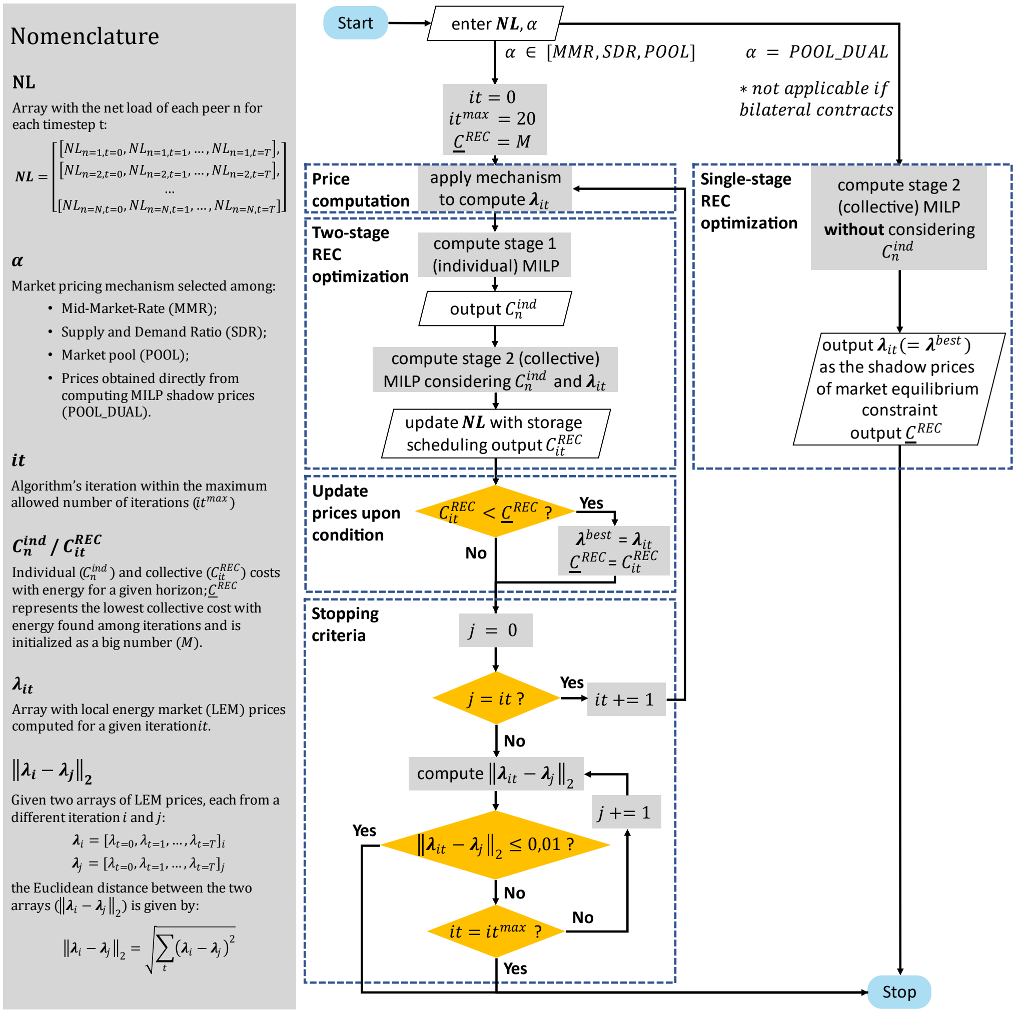


Figure 19: Overarching algorithm for the price estimation tool on a pre-delivery timeframe.

Figure 20 illustrates the same overarching algorithm adapted to a post-delivery timeframe. Since the batteries scheduling is no longer a decision variable, the algorithm is simpler when considering market mechanisms MMR, SDR or POOL, and only a single iteration is required to compute the best prices and define the lowest REC operation cost.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

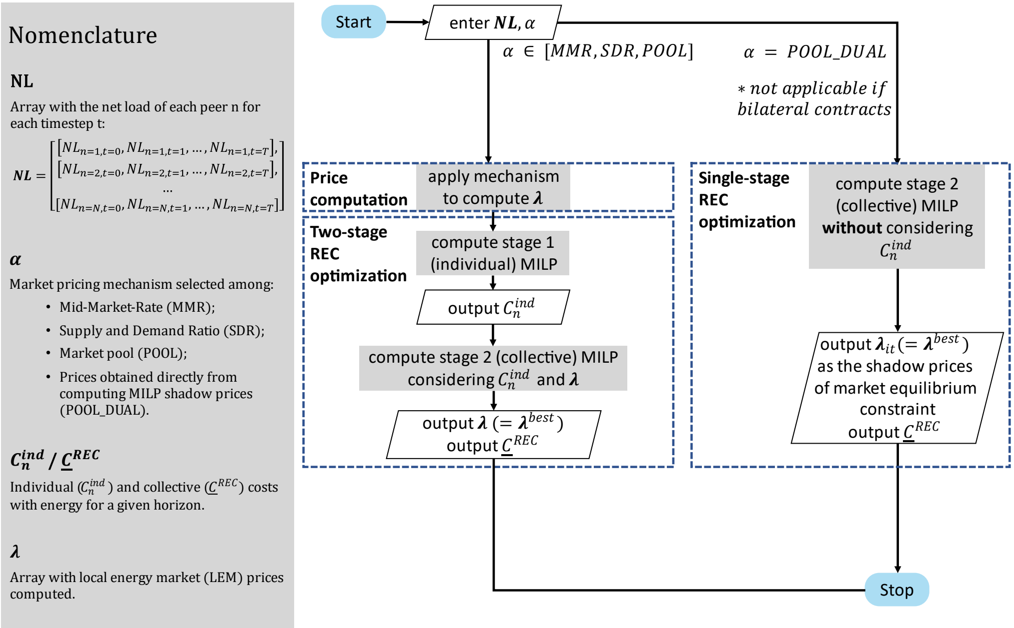


Figure 20: Overarching algorithm for the price estimation tool on a post-delivery timeframe.

3.3.2 Innovation

This service proposes the use of an optimization approach based on different pricing mechanisms to determine the schedules of the flexible assets of the REC members, the internal transactions, and the resulting energy prices. Both pre-defined pricing strategies, compared to marginal pricing strategies can be selected and compared, to assess the fairness of sharing of the REC collective benefits. These different alternatives will be further explored and compared to draw some conclusions on the best approaches to:

- promote RECs' self-sufficiency,
- minimize RECs' energy costs,
- distribute costs and revenues in the fairest way possible, given each member's/resource's contribution to the first two objectives.
- Test shared assets business models

This service must be understood in combination with the service "Energy community sizing with assets sharing" where similar business models can be considered at the planning and sizing phase of the REC.

3.3.3 Functions

The service has been implemented as an API in a containerized environment, entirely coded in Python programming language. The most relevant functions implemented are the described next:

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- **Input parsing function** – to parse the inputted data in raw format, namely the REC structure (i.e., definition of the members and their propriety on each resource), the forecasted data for each REC member, resources parameters (e.g., storage assets' power and energy characteristics) and contracted prices information for a chosen horizon.
- **Main function** – to orchestrate the whole overarching algorithm starting from calling the input parsing function, running each iteration of the algorithm until one of the stopping criteria is met and finally calling the outputs exportation function.
- **Price computation functions** – three separate functions to compute the internal transactions prices based on the forecasted net load of each member and shared resource computed at each iteration of the overarching algorithm, one for each price mechanism (MMR, POOL and SDR).
- **MILP class of functions** – implemented as a class of several functions, to define, solve and export the results of the two-stage or single-stage MILP procedure computed to define the optimal scheduling of controllable assets (i.e., storage) and internal transactions, provided the forecasted net load of each member and shared resource, and the computed internal transaction prices.
- **Stopping criteria function** – to verify if one of the stopping criteria has been met at the overarching algorithm, i.e., by computing the Euclidean distance between the most recent internal transactions' prices array and all the previous ones and checking if it is inferior to a certain threshold (e.g., 0.01), and by verifying if it^{max} iterations have already passed.
- **Outputs exportation function** – to compile the most relevant outputs, namely the internal transactions' prices final array and several results from the MILP solution originated with those prices (e.g., storage schedule, total energy bought and sold in the local energy market per member and shared resource) and to construct the outputs JSON file requested.

3.3.4 Input and Output Data Format

Apart from some configuration parameters, the inputs for this service will be mainly linked to the MILP procedure since it requires the definition of the REC structure, i.e., its members and resources and the relationship between them, load and renewable generation forecasts, and contracted prices per meter. Table 17 discretizes and characterizes the inputs required for this service.

Table 17: Input data specifications for the internal transactions' energy price estimation module.

Variable	Options/ Range	Default Value	Unit	Type	Description
start_ time	-	-	-	datetime	Starting datetime for the optimization horizon.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

end_ time	-	-	-	datetime	Ending datetime for the optimization horizon.
delta_t	{5, 15, 60}	-	min.	integer	Optimization step, in minutes.
iter_max	[0 - 20]	20	-	integer	Maximum number of iterations allowed to the overarching algorithm.
euclidean_ threshold	[0 - 1]	0.01	€/kWh	float	Minimum threshold for the Euclidean distance measurement between two computed price arrays for the overarching algorithm to stop.
peers	-	-	-	array[string]	Array with the unique identification of all peers in the REC.
meters	-	-	-	array[string]	Array with the unique identification of all meters in the REC.
rec_ configuration	-	-	-	array[{string: bool}]	Array indicating if a transaction between each pair of meters, uses or not the public utility power grid.
meters_ ownership	-	-	-	array[{string: float}]	Array provided for each peer, with that peers' ownership over each meters' resources.
meter_ net_load	-	-	kWh	array[{datetime: float}]	Array with the forecasted net load for each meter, for each datetime step of the optimization horizon.
meter_ buy_prices	-	-	€/kWh	array[{datetime: float}]	Array with the contracted energy buying prices for each meter, for each datetime step of the optimization horizon.
meter_ sell_tariffs	-	-	€/kWh	array[{datetime: float}]	Array with the contracted energy sell tariffs for each meter, for each datetime step of the optimization horizon.
meter_ max_p	-	1E3	kW	float	Maximum power flow (can be the contracted power with the retailer), admissible at each meter.
self_ consumption_ tariffs	-	-	€/kWh	array[{datetime: float}]	Tariffs to be paid (usually) by the consumer on a REC transaction, for each datetime step of the optimization horizon.
meters_ storage	-	-	-	array[{string: string, string: float}]	Array with the storage assets parameters for each meter, identified with an unique storage identifier, and several parameters,

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

					listed in the following rows of this table.
• nominal_ capacity	-	-	kWh	float	Storage asset's current nominal capacity.
• charge_ efficiency	-	100	%	float	Storage asset's charge efficiency.
• discharge_ efficiency	-	100	%	float	Storage asset's discharge efficiency.
• initial_ capacity	-	(*)	kWh	float	Estimated energy content at the storage asset at the beginning of the optimization horizon. (*) When not provided will equal the minimum capacity defined by **min_soc**.
• max_ power	-	1E3	kW	float	Maximum power setpoint at which the storage asset can charge/discharge.
• max_soc	[0, 100]	100	%	float	Maximum admissible state of charge (SoC) of the storage asset.
• min_soc	[0, 100]	0	%	float	Minimum admissible SoC of the storage asset.

Table 18 discretizes and characterizes the outputs provided by this service.

Table 18: Output data specifications provided by the internal transactions' energy price estimation module.

Variable	**Options/ Range**	**Unit**	**Type**	**Description**
stopping_ criterion	{'iter_max', 'euclidean'}	-	string	Indicates by which of the two stopping criteria has the overarching algorithm stopped.
iteration_ number	[0 – **iter_max**]	-	integer	Iteration number of the overarching algorithm at the moment it stopped.
minimum_ euclidean_ distance	-	€/kWh	float	Minimum Euclidean distance found between the computed arrays of internal transactions' prices.
meters_ bought_locally	-	kWh	array[{datetime: float}]	Array with the computed energy bought locally by each meter, for each datetime step of the optimization horizon.
meters_ sold_locally	-	kWh	array[{datetime: float}]	Array with the computed energy sold locally by each meter, for each datetime step of the optimization horizon.
meters_ bought_retailer	-	kWh	array[{datetime: float}]	Array with the computed energy bought to the retailer by each meter, for each datetime step of the optimization horizon.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

meters_ sold_retailer	-	kWh	array[{datetime: float}]	Array with the computed energy sold to the retailer by each meter, for each datetime step of the optimization horizon.
meters_ soc	[0 - 100]	%	array[{datetime: float}]	Array with the computed storage SoC for each meter with storage, for each datetime step of the optimization horizon.
meters_ charge_power	-	kW	array[{datetime: float}]	Array with scheduled charge setpoints for each meter with storage, for each datetime step of the optimization horizon.
meters_ discharge_ power	-	kW	array[{datetime: float}]	Array with scheduled discharge setpoints for each meter with storage, for each datetime step of the optimization horizon.
peers_ bought_locally	-	kWh	array[{datetime: float}]	Array with the computed energy bought locally by each peer, for each datetime step of the optimization horizon.
peers_ sold_locally	-	kWh	array[{datetime: float}]	Array with the computed energy sold locally by each peer, for each datetime step of the optimization horizon.
peers_ bought_retailer	-	kWh	array[{datetime: float}]	Array with the computed energy bought to the retailer by each peer, for each datetime step of the optimization horizon.
peers_ sold_retailer	-	kWh	array[{datetime: float}]	Array with the computed energy sold to the retailer by each peer, for each datetime step of the optimization horizon.
internal_ transactions_ prices	-	€/kWh	array[{datetime: float}]	Array with the computed internal transactions' prices, for each datetime step of the optimization horizon, for each iteration.
best_internal_ transactions_ prices	-	€/kWh	array[{datetime: float}]	Array with the computed internal transactions' prices, for each datetime step of the optimization horizon, for the iteration that achieved the lowest operation cost for the whole REC (**best_rec_cost**).
peers_costs	-	€	array[{string: float}]	Array with the computed individual operation costs, for each peer and for each iteration.
best_peers_ costs	-	€	array[{string: float}]	Array with the computed individual operation costs, for each peer and for the iteration that achieved the lowest operation cost for the whole REC (**best_rec_cost**).
rec_costs	-	€	array[{integer: float}]	Array with the computed REC operation costs, for each iteration.
best_rec_costs	-	€	float	Lowest computed REC operation cost among all iterations.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.3.5 Implementation details

This tool is currently at TRL6. The methodology has been implemented and tested on real data from members of a CER that is being established in Portugal. Although with a relatively simple underlying business model, that does not consider shared resources, the results obtained proved the efficacy of the methodology, showing a substantial increase on the community's self-consumption rates and a corresponding reduction of the costs with energy. These came as a result of the members' participation in a local energy market, with prices set by this tool. An interesting conclusion from this use case was that an equilibrium between generating members and consuming members is paramount for achieving the best KPIs, which is a good indicator that communities involving vulnerable consumers, that can only contribute with their consumption, when paired with mainly generating members, can be successful.

Much like other tools described previously, this tool is implemented in Python 3.10. The API will be containerized in Docker containers. The main dependencies used are listed in Table 19.

Table 19: Internal transactions' energy price estimation module Python dependencies

Library	Description	URL	License
apispec	Pluggable API specification generator	https://apispec.readthedocs.io/en/latest/	MIT
apispec_ webframeworks	apispec plugins for integrating with various web frameworks.	https://pypi.org/project/apispec-webframeworks/	MIT
base64	Base16, Base32, Base64, Base85 Data Encodings	https://docs.python.org/3/library/base64.html	Open Source
datetime	Basic date and time types	https://docs.python.org/3/library/datetime.html	Open Source
dotenv	Read .env file and set environment variables	https://pypi.org/project/python-dotenv/	BSD
flask	Python web framework.	https://flask.palletsprojects.com/en/2.3.x/	BSD
flask_mail	For setting an SMTP with a Flask application and sending mails	https://pypi.org/project/Flask-Mail/	BSD
flask_migrate	Handles SQLAlchemy database migrations	https://pypi.org/project/Flask-Migrate/	MIT
flask_sqlalchemy	Support for SQLAlchemy	https://pypi.org/project/Flask-SQLAlchemy/	BSD
flask_swagger_ui	For adding Swagger UI to the API.	https://pypi.org/project/flask-swagger-ui/	MIT
functools	Higher-order functions and operations on callable objects	https://docs.python.org/3/library/functools.html	Open Source

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

hashlib	Secure hashes and message digests	https://docs.python.org/3/library/hashlib.html	Open Source
http	HTTP modules	https://docs.python.org/3/library/http.html	Open Source
itertools	Functions creating iterators for efficient looping	https://docs.python.org/3/library/itertools.html	Open Source
itsdangerous	To cryptographically sign data (namely tokens)	https://pypi.org/project/itsdangerous/	BSD
joblib	Pipelining tools	https://joblib.readthedocs.io/en/latest/	BSD
json	JSON encoder and decoder	https://docs.python.org/3/library/json.html	Open Source
jwt	To encode and decode JSON Web Tokens (JWT)	https://pypi.org/project/jwt/	Apache License
loguru	Logging library	https://loguru.readthedocs.io/en/stable/	MIT License
math	Mathematical functions	https://docs.python.org/3/library/math.html	Open Source
multiprocessing	Process-based parallelism	https://docs.python.org/3/library/multiprocessing.html	Open Source
NumPy	Scientific computing	https://numpy.org/	BSD
os	Miscellaneous operating system interfaces	https://docs.python.org/3/library/os.html	Open Source
pandas	High-performance, easy-to-use data structures and data analysis tools	https://pandas.pydata.org/	BSD
PuLP	Linear Programming (LP) modeler	https://pypi.org/project/PuLP/	BSD
re	Regular expression (REGEX) operations	https://docs.python.org/3/library/re.html	Open Source
requests	HTTP library	https://requests.readthedocs.io/en/latest/	Apache License
secrets	Generate secure random numbers for managing secrets	https://docs.python.org/3/library/secrets.html	Open Source
sqlalchemy	SQL toolkit	https://pypi.org/project/SQLAlchemy/	MIT
sys	System-specific parameters and functions	https://docs.python.org/3/library/sys.html	Open Source
threading	Thread-based parallelism	https://docs.python.org/3/library/threading.html	Open Source
time	Time access and conversions	https://docs.python.org/3/library/time.html#module-time	Open Source
urllib3	HTTP client	https://urllib3.readthedocs.io/en/stable/	MIT
uuid	UUID objects according to RFC 4122	https://docs.python.org/3/library/uuid.html	Open Source
werkzeug	WSGI web application library	https://pypi.org/project/Werkzeug/	BSD

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

The MILP is computed using the CBC solver. The tool will require the provision of configuration parameters that characterize the REC in the pilot, as well as access to load and renewable generation forecasts for the members' resources, as well as access to the contracted prices and feed-in tariffs of each member. Since the tool is expected to run daily, requests to the API must be done by a third-party platform that can access and compile all necessary day-ahead data. Furthermore, the same third-party platform, or another that receives the response of this tool, should consider a graphical user interface to better convey the provided results.

3.3.6 Integration with ENERSHARE Data Space

The integration with the Data Space will follow the same strategy described in Section 3.1.6.

3.4 Data-driven failure detection algorithms for wind turbine components

3.4.1 Description of the Service

The installation of wind turbines presents several challenges and difficulties for operators. In fact, wind turbines are often subject to harsh operating conditions that increase the risk of various faults. The wind turbine, although quite adapted to its particular task and environment, it remains very brittle to extreme climate conditions. Unscheduled maintenance caused by unexpected faults can be very costly due to the maintenance support and the lost production time. According to (Wang, 2016), the operating and maintenance (O&M) costs lie between 20% and 25% of the overall levelized cost of electricity of current wind power systems. To address these issues, a service for anomaly detection of wind turbine drivetrain components is proposed in this work. There are three approaches for facing this service, and each approach will cover the fault detection of one component: gearbox, electric generator and hydraulic pitch system.

3.4.1.1 Anomaly detection of gearbox

First approach faces the anomaly detection of the gearbox, although it could be applied to other components as electric generator o power converter.

The gearbox is a critical component of a wind turbine, playing a crucial role in converting the low-speed rotation of the turbine's blades into high-speed rotation suitable for generating electricity. Situated between the rotor hub and the generator (see Figure 21), the gearbox serves as the mechanical interface that transmits power efficiently. It comprises a complex arrangement of gears, shafts, and bearings that work together to step up the rotational speed while maintaining a consistent torque.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

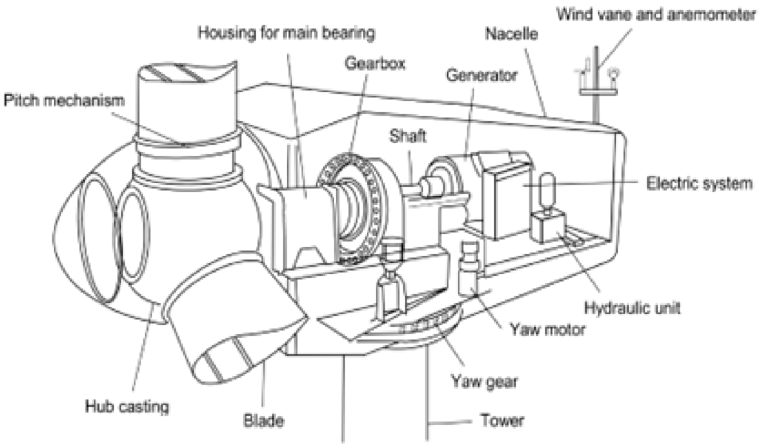


Figure 21: The different components of the nacelle

Due to the high stresses and demanding operating conditions it faces, the gearbox requires a regular maintenance to ensure reliability and enhance the overall performance of wind energy systems. In this project, we propose a two-phase approach for gearbox anomaly detection:

- The first phase aims to train an autoencoder model using only safe SCADA data (without gearbox anomalies). The gearbox anomaly is detected in the real-time SCADA observation by computing the reconstruction error between input and reconstructed data.
- In the second phase, we propose an algorithm based on the Bayesian network (BN) model to interpret and explain the anomaly results returned by the autoencoder. In this context, beyond the wish of using the BN to identify the main parameters deviation responsible of the anomaly, there is also the will to understand and validate the already detected failure (by the autoencoder) in the gearbox component. In Figure 22, we summarize the phases of the approach.

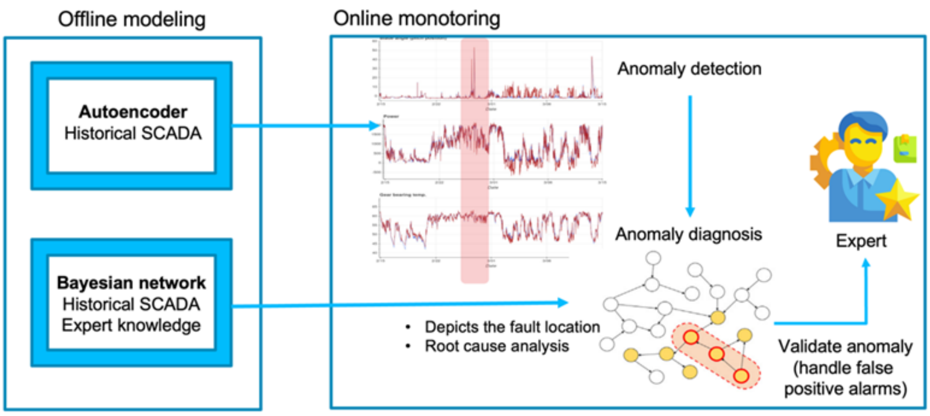


Figure 22: Gearbox anomaly detection using autoencoder and Bayesian network models

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

3.4.1.2 Anomaly detection of electric generator

Second approach faces the anomaly detection of electric generator, based in a hybrid model. This means that, in addition to real operation data, a physics-based model that contains mathematical relationships is used. It allows knowledge to be acquired from real operation data while preserving physical design relationships, helping in the detection and classification of different failure conditions. Followed methodology is detailed in (Pujana et al., 2023) and summarized in Figure 23.

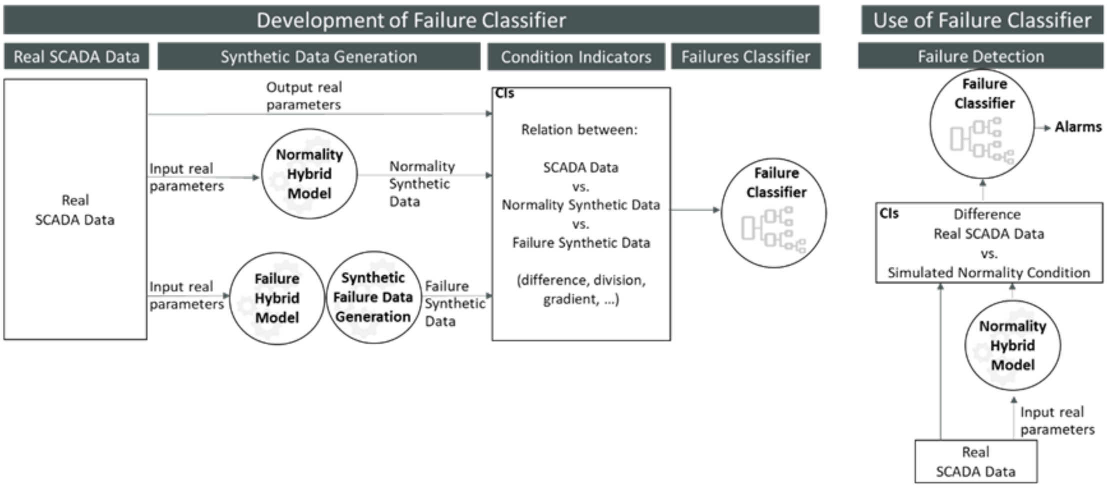


Figure 23: Methodology for developing and use of Failure Classifier

First, a normality hybrid model of a permanent magnet synchronous generator is developed. This normality hybrid model is based in a Simulink model that represents the behaviour of the Permanent Magnet Synchronous Generator (PMSG) in normal conditions (Figure 24) and is trained with real operation data using the Simulink Design Optimization. This training data must be related to a period without anomalies.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

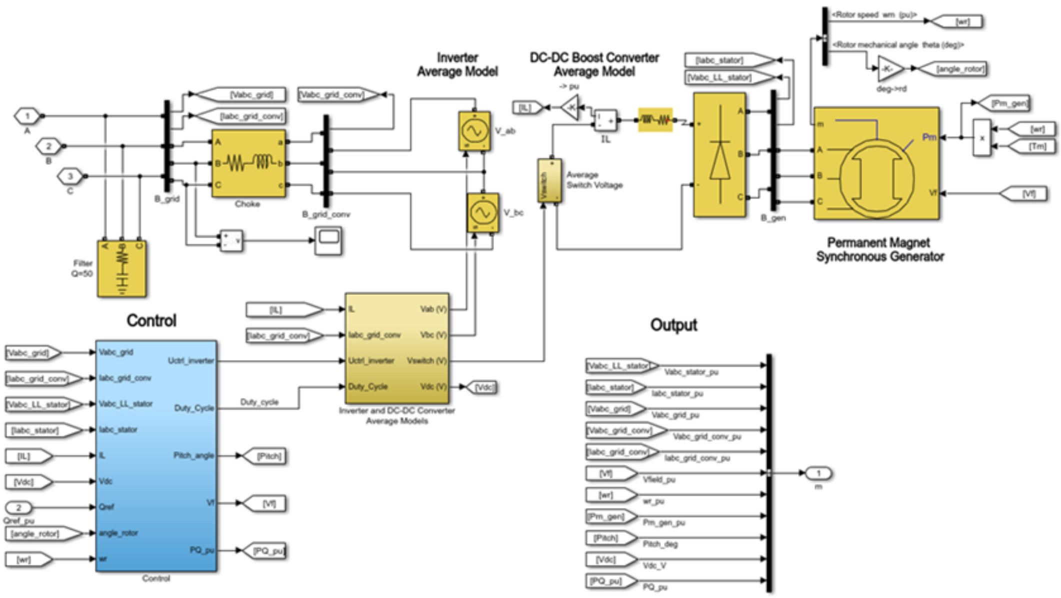


Figure 24: Model in Simulink of a Permanent Magnet synchronous generator connected to grid through a power converter

In turn, selected failure modes are modelled also in Simulink, and calibrated with real data in the same way as the normality hybrid model. This results in a Failure Hybrid Model that allows the generation of failure synthetic data. Taking as input several condition indicators calculated from operation and synthetic data, a failure classifier is developed.

A state-of-the-art analysis has determined that a data acquisition frequency of at least 1Hz is needed to detect main electric failure modes (short-circuits in stator winding).

3.4.1.3 Anomaly detection of hydraulic pitch system

Third approach faces the anomaly detection of hydraulic pitch system, through the generation of a model to perform condition monitoring (CM) analysis, focusing on the main failure modes extracted from a prior preformed FMECA. It has been tackled in two steps. First, a model of the Wind Turbine (WT) has been developed to simulate its behaviour. Data generated from this model (synthetic data) is the fuel to feed the generation of a mathematical model that will distinguish between normal and abnormal behaviour.

10-minute records of cylinder positions, proportional valve commands, accumulator pressures and pump activation signals are collected, with a 10 ms sampling interval. They are processed on the edge and 8 real values from each 10-minute record are produced. Those 8 real values seem to have good correlation with several pitch system faults, according to generated synthetic data. Synthetic data are stored as comma-separated values, one file per 10-minute record, with 10 ms intervals between samples. They are read into memory via numpy function

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

genfromtxt. The data corresponding to accumulator pressure are selected and filtered with a moving-window linear fitting algorithm (function movav, developed by us) which gives an estimation of the pressure variation rate and a smoothed pressure reading. The pressure variation rate is divided by the pressure squared to get a value proportional to an estimation of the oil flow rate into the accumulator. The data corresponding to cylinder position are selected and filtered similarly to obtain an estimation of the cylinder velocity, which is proportional to the oil flow rate from the accumulator to the cylinder. Finally, the pump activation signals are processed to estimate the oil flow rate from the pumps to the accumulators, by assigning the nominal flow rate to instants when the pumps are activated and zero to those when they are not. The resulting flow rate estimations are linearly fitted via successive calls to LinearRegression.fit from scikit-learn. The resulting parameters are plotted and good clustering by type of fault is observed.

3.4.2 Innovation

In the last decade, numerous efforts have been devoted to overcome the limitation of current maintenance strategies, which are mainly based on preventive and corrective maintenance. These approaches allow real-time monitoring rather than costly time interval intervention and performance tune-ups. Feedback from various industries shows that condition monitoring techniques can detect anomalies before they turn into system-critical faults, allowing therefore maintenance to be well scheduled. The wind industry nowadays employs various condition-monitoring techniques to detect a failure at a sufficiently early stage. These techniques include acoustic monitoring (Bouno et al., 2005), oil monitoring (Zhu et al., 2013), thermal monitoring (Guo et al., 2011), etc. However, the widespread implementation of these methods is often infeasible due to the high cost of installing custom sensors and the complex communication infrastructure and protocols needed to manage their data.

To address these limitations, numerous data-driven approaches based on SCADA data have been provided. In (Encalada-Dávila et al., 2021), an unsupervised main bearing fault prognosis approach is presented. In this approach, only healthy SCADA is collected and an artificial neural network is used to predict the low-speed shaft temperature. An anomaly detection threshold is then established based on the residuals between the predicted and real values. Authors in (Tutivén et al., 2022) proposed a semi-supervised method based on a one-class support vector machine classifier. The method detects anomalies in the main bearing by using a decision function that categorizes real-time data as similar or dissimilar to the safe data. Other unsupervised anomaly detection approaches have been proposed in (Campoverde-Vilela et al., 2023; Li et al., 2020; Zhao et al., 2018). While these methods are efficient, they cannot point out the sensor observations responsible for the failure. Note that this information is required in the planning of diagnosis and maintenance actions. In (Elasha et al., 2019) a supervised machine learning approach is proposed to detect failures on the gearbox component. Unfortunately, the training process of the anomaly model needs historical faulty data to be

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

tagged, which is a time-consuming and error-prone task. Additionally, the highly imbalanced nature of gearbox fault observations in the SCADA data may lead to other performance issues.

3.4.2.1 Anomaly detection of gearbox

An original approach based on a two complementary model is developed: the autoencoder and a graphical probabilistic model (Bayesian network). In this case, we rely on the effectiveness of deep autoencoder network to quantify the deviation between normal and current behaviour of the gearbox signal. In the next phase, a Bayesian network is used to confirm and explain the anomaly of the gearbox by pointing out the list of sensors (or input parameters) responsible of the observed deviation.

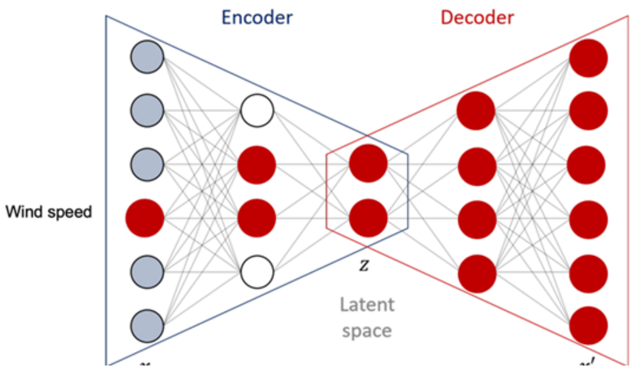


Figure 25: Problem of error propagation in the autoencoder model

3.4.2.2 Anomaly detection of electric generator

The main innovation lies in the application of the methodology described in (Ainhoa et al., 2023) to a different technology, in this case to generators with permanent magnets (the methodology only was applied to double fed induction generator DFIG). Other innovation is that the failure diagnosis is geared towards electric failure modes, mechanical failures are more investigated with new technologies, but researches about electrical failures are hard to find in the state-of-the-art.

3.4.2.3 Anomaly detection of hydraulic pitch system

The main contribution lies in calculating statistics of the relationships between accumulator pressure variations, proportional valve commands and cylinder motion speeds. The advantage is that it is avoided having to define faults in terms of pitch system transfer function parameters and the inherent integration of observers. It also allows to use fragmented data.

The advantage over other approaches, and the novelty, is in that the software does not have to run in parallel to the turbine control, or in real time, and that it can easily deal with missing data, because the data are processed as described above. Additionally, in order to produce the

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

synthetic data, we have produced a high-fidelity electric-hydraulic-mechanic Simscape model of the pitch system and coupled it with an aero-elastic model of a wind turbine.

3.4.3 Functions

3.4.3.1 Anomaly detection of gearbox

This approach involves four key steps (see figure 3). The first step is called **data preprocess**. The SCADA dataset used in this study includes many parameters with varying outliers, incomplete observations, and mismatches in time and date stamps. Therefore, preprocessing operations are crucial to obtain a clean dataset that is suitable for machine learning algorithms. The first step is to select the set of variables to be included in the model training phase. It is important to choose the most important features that provide valuable insights into the gearbox behaviours.

According to domain experts, the mean values of environmental measurements, rotor speed, and the temperatures of the gearbox should be exploited. To justify the preference for exogenous variables, it is important to note that in wind turbine systems, component states are strongly correlated. As a result, selecting temperature features that are related to other components rather than just the one under investigation (the gearbox) can reveal additional failures in components that are closely linked to the variables used. In such situations, the model may lose its ability to differentiate between the failures of interest and those occurring in other related components. For each variable, we generate the lagged (t-1) and (t-2) variables to represent the temporal dimension during model learning. After the feature selection process, a data cleaning pre-processing based on the quantile approach is executed to eliminate outliers and sensor measurement errors in the selected features. These abnormal observations do not contain valuable guidance in detecting anomalies. After identifying the outliers, they are treated as missing values and filled using a cubic Hermite interpolating polynomial (CHP) (Liang et al., 2020). Finally, we deal with the very different magnitudes of the selected features. For a fair training process, every variable X is scaled using the min-max normalization process:

$$x'_i = (x_i - min(X))/(max(X) - min(X)).$$

The second step consists in **building an anomaly detection model** using deep autoencoder model. The autoencoder (AE), also referred to as autoassociator, is a highly effective type of artificial neural network for diagnostic tasks that has been applied in various real-world applications. Unlike supervised predictive models, autoencoders offer a prominent flexible model that can represent complex functions without requiring labeled data during the learning process. An AE is an unsupervised and symmetrical neural network that consists of two primary components: an encoder $f : R^n \to R^e$, which compresses the input data ($\mathbf{x}_i$) into a lower latent representation ($\mathbf{h}$), and a decoder $g : R^e \to R^n$, which maps the latent representation back to the original data. Then, the desired output is the input itself. The reconstruction error between

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

input and output is therefore used to detect anomalies on the gearbox. In our case, we employ a training approach that focuses exclusively on data containing normal gearbox behaviour, deliberately excluding any instances of anomalies (using the information contained in the maintenance interventions file). This approach allows the model to deeply understand the intricate patterns and structures associated with normal operations, enabling it to accurately distinguish between normal and abnormal gearbox behaviour. By training exclusively on normal data, we equip the autoencoder to effectively identify and flag any deviations or anomalies that may occur during real-time operations.

In addition to training the autoencoder using only normal gearbox behaviour, setting an appropriate threshold for anomaly detection is crucial. Once the autoencoder is trained, it can reconstruct the input data and compare it to the original samples. The difference between the reconstructed and original data serves as an indication of anomalies. By establishing a suitable threshold, we can determine the level of deviation from normal behaviour that should trigger an alert. The threshold ϵ is determined based on the distribution of the different RE_i that are computed using the validation dataset. In our case, observations are considered anomalous if the reconstruction error associated with an input set $\mathbf{x}_i$ is greater than three times the standard deviation above the mean.

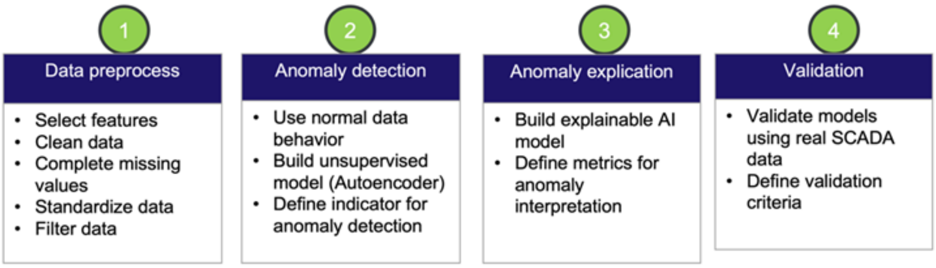


Figure 26: The pipeline of the anomaly detection and explanation approach

Although the AE is often effective in detecting anomalies in SCADA data, it is nevertheless unable to reveal which features are the root causes of the anomaly. Roughly speaking, the deviation of only one of the considered input features is enough to cause the propagation of the error through the AE network, resulting in a significant reconstruction error in most other features as well (see figure 2). To address this limitation, we introduce a novel method based on the Bayesian network to help us figure out which features are responsible for the anomaly.

In the third phase, we use the training data to build a complementary model called Bayesian network. We propose an anomaly explanation model based on the Bayesian network model (BN) (Heckerman et al., 1995). Unlike autoencoder, the BN is a type of probabilistic graphical model that uses a graph-based representation to compactly represent a high dimensional distribution over complex systems.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Definition 1 (Bayesian network): A BN is a pair *(G, Θ)* where G = (***V***,*A*) is a directed acyclic graph (DAG), V represents a set of random variables, A is a set of arcs, and $\Theta = \{\theta_{X_i} | Pa(X_i)\}_{X_i \in V}$ is the set of the conditional probability distributions (CPD) of the nodes / random variables X_i in G given their parents $Pa(X_i)$, i.e. $\theta_{X_i} | Pa(Xi) = P(X_i | Pa(X_i))$. The BN encodes the joint probability over V as:

By their graphical structure, BNs encode an independence model, i.e., a set of conditional independencies between variables, characterized by the d-separation property [13]. Several approaches for learning the BN graph have been proposed in the literature. These algorithms can be divided into 3 classes: i) the search-based approaches that focus on optimizing the scoring function of the structure [13]; ii) the constraint-based approaches that exploit statistical independence tests to find the best structure (Spirtes and Glymour, 1991); iii) the hybrid methods that exploit a combination of both [15]. Although these approaches are effective, they often fail to identify all causal directions in the graph. To cope with this problem, we decide to use an alternative approach which consists in exploiting experts' knowledge together with a score-based approach to optimize the causal discovery. In our situation, experts provide their knowledge about some causal relations between variables. This knowledge is about the existence or absence of arcs and is expressed as hard structural constraints, i.e., the BN learning algorithm is not allowed to modify these relations. Expert knowledge is also used after learning to modify, add or reverse mistaken causal relations. We now describe the proposed method to detect and explain the anomaly using AE and BN models.

Actually, our main focus revolves around discovering the optimal probabilistic inference query of the Bayesian network that enable us to understand the source of the anomaly. This query plays a vital role in determining the responsible sensor(s) for the detected anomalies within the system. Through Bayesian inference, we can effectively assess the probabilities associated with various sensors and their potential involvement in the observed anomalies. Details about this approach will be discussed in the next delivery.

The final phase consists in validating both developed models. We demonstrate the efficiency of our approach by testing it on normal and anomalous gearbox observations. We investigate the importance of the use of the explainable BN model to enhance the understanding of the anomaly and eliminate false positive alarms. To do so, we will use the SCADA data that contains normal and anormal data of the gearbox.

This step has not been addressed yet, as we are currently awaiting the development and validation of the Bayesian network model, as well as the availability of testing data that contains anomaly on the gearbox.

3.4.3.2 Anomaly detection of electric generator

The first function to be used in a **data preprocess**, in order to deal with outliers, incomplete observations, and mismatches in time and date stamps. Therefore, preprocessing operations

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

are crucial to obtain a clean dataset that is suitable for machine learning algorithms. The following steps are included:

- Step 1: Load and parse data: The data is expected to be obtained from text-based format files, in JavaScript Object Notation (JSON) syntax, and UTF-8 encoding. It allows the storage of a wide range of data (date times, categorical, etc), and complex nesting data structures. Data is converted into a specific format, date strings into a date format, categorical variables into a consistent numerical format, and the rest of the values into floating-point arithmetic.
- Step 2: Check duplicated samples: the task of examining duplicated samples is executed by checking the unique date-time value for each sample. In case of duplicated samples are found, one duplicate sample is kept, and the rest are deleted.
- Step 3: Check range signal data: Detecting wrong signal values that are outside of the expected range requires human expert knowledge to define the maximum and minimum values for each signal. The range of the variables will be determined once the rate of the turbines under study are defined.
- Step 4: Check outliers by behaviour: the application of domain expertise is essential here to determine if an observation that lies an abnormal distance from other values corresponds to an specific operation mode or not.

The second function to be used will **Split data.** Selecting the appropriate datasets for training, validation, and testing is a crucial step in developing an accurate normality model for any component. Therefore, it is essential to consider all operational and environmental conditions that the wind turbine may encounter, such as different wind velocities and seasonal conditions. To train and validate the healthy models, only data without failure are used. Data containing anomalous observation are used to test the model performance, and also for calibration the failure hybrid model.

Once data are ready to be used, the develop a physics-based model of the PMSG is needed. This model is developed in Simulink, and is trained with real data and using the Simulink Design Optimization to **develop the Normality Hybrid Model,** which will allow the generation of normality synthetic data, as indicated in Figure XXX. In the same way, the **development of the Failure Hybrid Model,** being these two components clue point to develop the failure classifier.

For the **development of the failure classifier**, both unsupervised and clustering algorithms are used, in order to make a benchmarking of the capacities in the failure detection. Algorithms as isolation forest, random forest and neural networks have been used.

3.4.3.3 Anomaly detection of hydraulic pitch system

Instantaneous oil consumption is estimated from the accumulator pressure variation and, independently, from the cylinder motion as well. This is done, as described before, by filtering

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

the pressure and positions data with function movav to get the pressure variation rate and the cylinder velocity, then dividing the former by the pressure squared, then multiplying by accumulator and cylinder parameters (quantity of nitrogen and effective areas, respectively). Linear models are fitted to the resulting estimations, as described before, via **LinearRegression.fit**. Besides, non-linear models are fitted to the instantaneous proportional valve commands and cylinder velocities, via **GaussianProcessRegressor** with an **RBF+constant kernel**. That gaussian process is used to predict the cylinder velocity at 25% proportional valve opening, via the **GaussianProcessRegressor.predict** functio

3.4.4 Input and Output Data Format

There are two types of data: the SCADA data and the history of ordinary and extraordinary maintenance interventions.

The SCADA data provides a rich source of continuous-time observations (10 min / 1s) about several of condition variables and is expected to be obtained from text-based format files, in JavaScript Object Notation (JSON) syntax, and UTF-8 encoding. The datasets consist of environmental, electrical, component temperature, hydraulic, and control variables. The environmental related variables include ambient temperature (◦C), nacelle temperature (◦C), wind speed (m/s), and turbulence index. These variables are highly correlated with the electrical and component temperature variables. The electrical variables group encompasses active power (kW), phase voltage (V), power factor, reactive power (kW), and electric network frequency (Hz). These variables describe the power generated by the wind turbine before it enters the distribution grid. The active power curve is very sensitive to small variations in the wind speed output of the nacelle anemometer. Electrical network frequency and phase voltage measurements are obtained to control potential fluctuations, while reactive power and power factor measurements provide insight into the efficiency of electric power utilization. The goal for wind turbine control is to maintain a power factor as close to 1 as possible.

The component temperature variables correspond to temperature signals (in ◦C) from several key locations in the nacelle (see Figure 21), such as the gearbox, generator, and main bearing. Component temperatures are one of the main condition parameters that are closely related to the performance of the wind turbine. Hydraulic variables describe observations of the general accumulator, brake pressure, general accumulated pressure of the blades, and hydraulic group pressure. These parameters are generally used to control the pitch, yaw, and braking systems of the wind turbine. The pitch cylinder for each blade is actuated using a hydraulic accumulator for different operations, such as blade engagement and blade safety position. Control variables are related to all control systems that guarantee safe operation, optimize power output, and ensure the long structural life of the wind turbine. The wind turbine is equipped with blade pitch control, which allows to maintain the optimum blade angle to achieve certain rotor speeds. The yaw controller is another important control system that is responsible for the rotation of the

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

entire wind turbine. It ensures that the wind turbine is constantly facing the wind to fully capture the incoming wind power. Additionally, the rotor and generator speeds are two key parameters that must be controlled for power limitation and optimization.

In addition to the SCADA data, we can access extra information that keeps track of the ordinary and extraordinary maintenance interventions required by each wind turbine over its lifetime. These details are typically stored in Excel files, where a description of the repair actions, together with a timestamp of the date on which it happened, are reported.

3.4.5 Implementation details

3.4.5.1 Anomaly detection of gearbox

Current TRL is TRL5, as this service has been validated in relevant environment. At final stage, it is expected to achieve TRL7 (demonstration in operational environment).

This study will leverage a range of advanced technologies to facilitate our research objectives. Firstly, we will utilize **TensorFlow**, a powerful open-source machine learning framework that enables efficient development and deployment of the deep autoencoder model. TensorFlow provides a comprehensive ecosystem of tools, libraries, and resources that empower us to tackle complex data analysis tasks and train sophisticated neural networks.

In addition, we will employ **Agrum**, a versatile and user-friendly library for probabilistic graphical models. Agrum offers a wide array of algorithms and functionalities, allowing us to construct, manipulate, and reason about probabilistic models effectively. With Agrum, we can explore uncertain relationships within our data, perform probabilistic inference, and gain valuable insights from complex systems.

To ensure seamless and reproducible experimentation, we will containerize our research environment using **Docker**. Docker allows us to encapsulate our entire software stack, including dependencies, into portable containers. This approach ensures that our experiments can be easily replicated and shared across different platforms and environments, eliminating the hassle of configuring software dependencies.

The programming language of choice for our study will be **Python**. Python offers a rich ecosystem of libraries and frameworks, making it ideal for scientific computing, data manipulation, and machine learning tasks. Its intuitive syntax, extensive community support, and vast array of available tools, such as **NumPy** and **Pandas**, enable us to efficiently handle and analyze large SCADA datasets.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.4.5.2 Anomaly detection of electric generator

Current TRL is TRL5, as this service has been validated in relevant environment. At final stage, it is expected to achieve TRL7 (demonstration in operational environment).

In this case, the normality hybrid model developed in **Simulink-Matlab** will be containerized using **Docker**, with the advantages explained in the previous section. The programming language will be **Python**, making use also of its libraries and frameworks.

3.4.5.3 Anomaly detection of hydraulic pitch system

Current TRL is TRL5, as this service has been validated in relevant environment. At final stage, it is expected to achieve TRL7 (demonstration in operational environment).

The data is processed with **Python** via some scikit-learn data fitting functions, as mentioned above, and some filters of our own, also mentioned above. The synthetic data are produced with an electric-hydraulic-mechanical model implemented via **Simscape** and coupled with DTU's 10 MW reference wind turbine model running on OpenFAST.

3.4.6 Integration with ENERHARE Data Space

The preliminary schema of the integration of this service in ENERSHARE Data Space is shown in Figure 27. IDS connectors will be used for the exchange of data. ENGIE, TECNALIA and HINE will implement an IDS pipeline formed of one connector provider and consumer in each side to be able to execute the service properly.

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

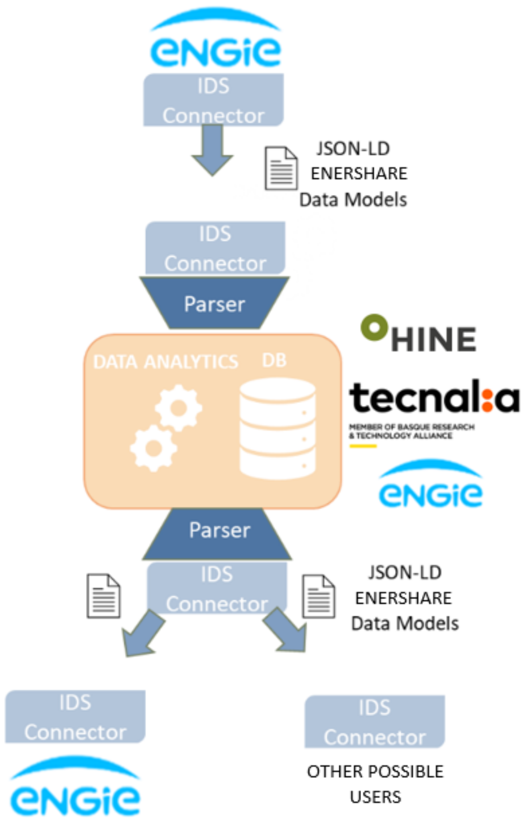


Figure 27: Integration of Failure diagnosis service in ENERSHARE Data Space

These IDS connectors will be integrated with the rest of ENERSHARE Common IDS components (Metadata Broker, Identity Manager, Vocabulary Hub...). ENGIE has built a data lake in their DARWIN platform where all the data from different windfarms and different heterogeneous data sources is available. This data lake contains historical data and also records new streaming data. ENGIE will develop an adapter to create a JSON-LD directly from the data stored in their data lake to a JSON-LD format harmonized according to the ENERSHARE Common Data Models. In the same way, TECNALIA will create a wrapper to transform the input data from the JSON-LD sent by ENGIE to the internal data model used within the data analytics tools. TECNALIA will define also the mapping rules to convert the outputs of the data analytics tools to send the results to ENGIE in a JSON-LD format harmonized according to the common ENERSHARE Data Models. This created JSON-LD files will be send in the payload as an IDS artifact. An orchestrator will be developed that will be in charge of calling the different components within the data analytic tools with the appropriate data coming from the database and processing and parsing the outputs. All the components will be dockerized.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.5 Substation Load forecasting tool

3.5.1 Description of the Service

The Substation Load Forecasting Tool (AI Forecasting) developed by R&D Nester provides day-ahead net-load forecasts. It can be applied to nation-wide aggregate load prediction, to all the way down to domestic level consumption forecasting. It does this by leveraging on a Stacking Ensemble of ML models, which, using past net-load data and features extracted from said data, is able to negate the biases of each of the stack members and make an improved forecast than any of the member models.

In the current implementation of AI Forecasting, the data used for training and forecasting – which pertains to substation net-load – is manually extracted from REN's (Portuguese TSO) internal statistics web page and stored in csv (Comma Separated Values) format. This file is then used as input for the tool, which starts with a Data Pre-Processing stage, in which the data is restructured and checked for inconsistencies such as duplicate or missing time steps, and these are handled through data imputations. The time zone used is also converted to UTC to fix the missing and repeated time steps during day light savings time transitions. The load data is also scaled to an interval of [0; 1].

After this, the Feature Engineering process starts in which relevant and explanatory variables for net-load forecasting are extracted from the net-load historical time series data. These are lagged features to represent the importance of past net-load into future ones. These features and the observed net load are separated into a training vectors dataframe, **X** and into a target values series, **y**.

Using these data sets, it is possible to train the ML model which will generate net-load forecasts at the transmission substation level. The ML model is a Stacking Ensemble Regression model with two layers – one layer with different Base Learner models and one with the Meta-Learning model. The working principle behind the Stacking Ensemble is to individually train different forecasting models with the same training features and targets, whose forecasts will then be used as training features for the final estimator, which acts as a Meta-Learner, generating predictions for the target value from data derived from an estimation made on the original training data. The models used in the ENERSHARE implementation of AI Forecasting are the following: Lasso, XGBoost, Decision Tree and SVR with rbf kernel for the Base Learners and an SVR with linear kernel for the final estimator. The base learners are trained on the entire training dataset, while the final estimator is fitted using cross-validated predictions of the base estimators. The number of folds used in the cross-validation is 5.

In Figure 28, the high-level overall architecture of the Load Forecasting Tool is displayed. The blocks that make up this high-level description are detailed below.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

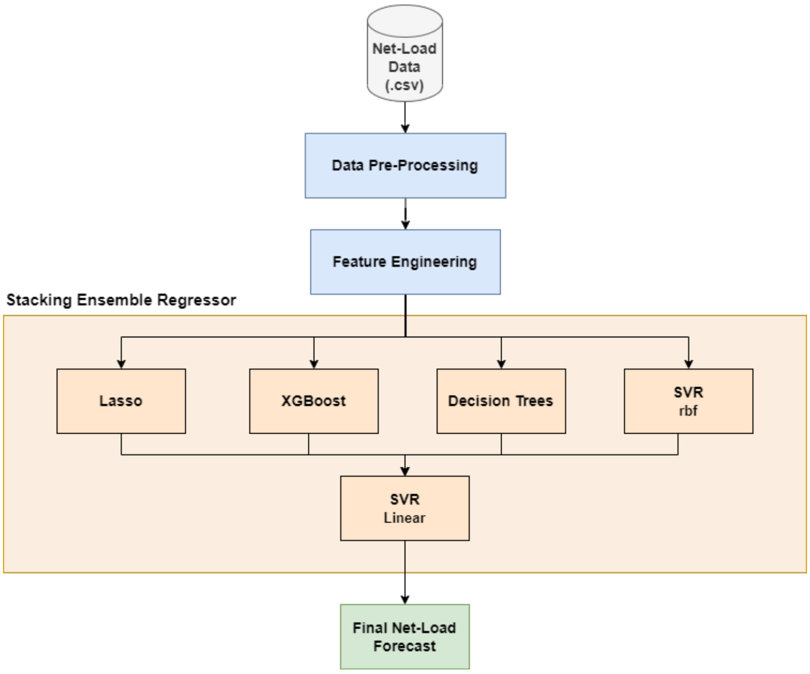


Figure 28: High Level Architecture of the Load Forecasting Tool

Data Pre-Processing

The net-load data, contained in time series stored in csv format, is read by the tool into a Pandas' dataframe. The first step of data pre-processing is setting the time steps as the index and formatting it in datetime format. This index is then sorted, to ensure that the time series is ordered temporally. The data is then check for inconsistencies. This is done by grouping the load data by day and counting how many data entries there are per day. If the number is different from 96 (i.e., a full day with 15 minute time steps, the granularity used for substation data) then an inconsistency has been found. These are either duplicate or missing data. Duplicate data is handled by keeping the first record and dropping the duplicates. Missing data is imputed by filling the missing time steps with the corresponding ones from the previous day. The data is then normalized to a [0; 1] interval using '`MinMaxScaler`' from sklearn's preprocessing module.

Feature Engineering

The Feature Engineering section of the tool is handled by the internally defined '`preprocess_lagged_data`' function. This function has as input arguments the dataframe with the data from which the features shall be extracted, as well as an integer which will define the maximum number of day the data shall be lagged (i.e., how many days in the past we will look, in order to make a day-ahead forecast). The lagging of data is done using a '`for`' loop, which iterates from 1 to argument 'n_`days`' and shifts the original load data by the number of

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

days corresponding to the iteration number ‘i’, appending to the dataframe a column for each iteration. By the end of this loop, the dataframe shall have the original load data plus ‘n_days’ number of columns, whose data is lagged by 1 day from each other. This process will result in the generation of ‘NaN’ entries at the beginning of the dataset (no past beyond the first entry) and these are dropped. Then the dataframe is split into two variables, **X**, which will contain the features for training and then prediction, and **y**, which will contain the targets (i.e., the original load data). This is done by simply dropping the column with the original load data to create **X** and making **y** equal to this column. The function then returns the **X**, **y** tuple. The schema of these dataframes is shown in Table 20.

Table 20: Features (X) and target (y) dataframes

X

index	‘Load 1 days’	…	‘Load n days’
DatetimeIndex	float64	…	float64

y

index	‘Load’
DatetimeIndex	float64

After running this function, the tool uses the function ‘train_test_split’ from sklearn, to get training and testing data sets, with a 0.9/0.1 split and no shuffling done on the data.

Training and Forecasting

With these datasets it is now possible to train our model, which will fit the training data containing the relevant features to the target data that the model will attempt to predict. As stated before, the tool uses a Stacked Ensemble of ML models to forecast the net-load data at substation data. The sklearn library already contains a ‘StackingRegressor’ class in sklearn.ensemble module, in which the base estimators are passed as an argument inside a list. The final estimator can also be chosen by passing the desired model class as an argument. To implement the stack, with the desired estimators (SVR, XGBoost, Decision Trees and Lasso as base estimators and Linear SVR as the final estimator), the class should be called as such:

```
stack = StackingRegressor(estimators=[
        ('svr', SVR()),
        ('dt', DecisionTreeRegressor()),
        ('lasso', Lasso()),
        ('xgb', XGBRegressor())],
        final_estimator=LinearSVR(), cv=5)
```

This configuration uses default hyperparameters. It is required to import the base and final estimator classes from their respective modules. In order to train the model stack to predict net load data, the ‘fit’ method is called with the training split of X and y passed as arguments. With this method, the base estimators are fitted to the entire training dataset, while the final

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

estimator is trained using cross-validated predictions of the base estimators using `cross_val_predict`. After this is complete, the '`predict`' method is called on the test split of the training data (`xtest`) and the result is compared with the target test data (`ytest`), calculating the error of the estimation using the Normalized Root Mean Square Error (nRMSE) and the Symmetric Mean Absolute Percentage Error (SMAPE). The forecasted load is then stored in a csv file.

3.5.2 Innovation

The current ENERSHARE version of AI Forecasting is still, at this stage, quite similar to implementation in previous projects, such as GIFT. The models used in the stack differ slightly, with the inclusion of Lasso and Decision Trees as base estimators improving the accuracy of the model as whole. It is also expected the inclusion of the Pattern Sequence Forecasting model in the stack will further improve the accuracy of the model. This will come in the next release of the tool in D6.2.

In this version of the tool, the model is running on the TSO's substation load data which is retrieved manually. An improvement to the tool which will also be occurring in future releases of the tool is the automation of this process.

3.5.3 Functions

- **Main function:** where the whole code resides and where the other functions are called. Small data manipulations and the data input and output reside here.
- **Data Imputations function:** data treatment function, which drops duplicate time stamps (keeping the first entry) and fills missing time stamps with the corresponding ones from one day prior (missing data from day D at H:M is filled with data from D-1 H:M).
- **Feature Engineering function:** creates the dataset with the features vectors by lagging the original data and generating columns, which are lagged by one day from each other. Each row from this data set represents one timestamp, the observed value at that particular time stamp and the observed values at that time in previous consecutive n days. The function takes as argument the original data as a dataframe and the number of days in the past to use as features. The outputs are the features dataframe and the target series.
- **Ensemble model function:** core of the tool, where the model is trained and forecasts are made. Using the training data, both the features and targets, the Stacking Ensemble model is fitted to the purpose of forecasting day-ahead net-load data. It outputs a prediction made for the day-ahead of the training data.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- **Error metric computation functions:** compute the NRMSE and SMAPE error metrics given two NumPy arrays, one with predicted data and another with actual, observed data. Outputs a float, which is the corresponding error metric value in percentage.
- **Plotting function:** produces overlapping plots of two data arrays or series, intended to represent the forecasted data and the real observed data, providing a visual comparison between the two.

3.5.4 Input and Output Data Format

The data required as input for the load forecasting tool is a net-load time series, with the historical data from a given transmission substation. The final result of the tool will be a forecasted net-load time series, with the historical data included as a column, with NaN values for future time stamps, where metering is not possible. The specifications for the input and output data tables are shown in Table 21 and Table 22.

Table 21: Input data format specifications for the load forecasting tool

Column	Unit	Format	Type	Description
Load	MWh	-	float64	Metered net load at transmission substation level
Date	Minute	%d/%m/%Y %H:%M	datetime	Timestamps of the metered load

Table 22: Output data format from the load forecasting tool

Column	Unit	Format	Type	Description
Load_pred	MWh	-	float64	Forecasted net load at transmission substation level
Load	MWh	-	float64	Metered net load at transmission substation level
Date	Minute	%Y-%m-%d %H:%M	datetime	Timestamps of the forecasted and metered load

3.5.5 Implementation details

The current TRL of the Load Forecasting tool is TRL 6. It has been demonstrated, both in past projects and in its current form used in ENERSHARE, to provide accurate forecasts using real data. It is intended for the tool to be integrated with the ENERSHARE Data Space, running inside a connector and providing the resulting forecasts and error metrics, while maintaining the inputs data’s confidentiality. The will be used in the Portuguese pilot as a benchmark for the

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

federated learning forecast tool. This integration will lead to a demonstration in an operational environment, making the tool reach TRL 7. This is expected to happen by D6.3.

The tool was developed in Python 3.9.12 and the packages required to run the tool's code are listed in Table 23.

Table 23: List of Python dependencies used to run the load forecasting tool

Name	Version	License	Description
Pillow	9.5.0	Historical Permission Notice and Disclaimer (HPND)	Python Imaging Library (Fork)
Pygments	2.15.1	BSD License	Pygments is a syntax highlighting package written in Python.
asttokens	2.2.1	Apache 2.0	Annotate AST trees with source code positions
backcall	0.2.0	BSD License	Specifications for callback functions passed in to an API
colorama	0.4.6	BSD License	Cross-platform colored terminal text.
comm	0.1.3	BSD License	Jupyter Python Comm implementation, for usage in ipykernel, xeus-python etc.
contourpy	1.0.7	BSD License	Python library for calculating contours of 2D quadrilateral grids
cycler	0.11.0	BSD License	Composable style cycles
debugpy	1.6.7	Eclipse Public License 2.0 (EPL-2.0)	debugpy is an implementation of the Debug Adapter Protocol for Python.
decorator	5.1.1	BSD License	Decorators for Humans
executing	1.2.0	MIT License	Get the currently executing AST node of a frame, and other information
fonttools	4.39.4	MIT License	Tools to manipulate font files
importlib-metadata	6.6.0	Apache Software License	Read metadata from Python packages
importlib-resources	5.12.0	Apache Software License	Read resources from Python packages
ipykernel	6.23.1	BSD License	IPython Kernel for Jupyter
ipython	8.13.2	BSD License	IPython: Productive Interactive Computing
jedi	0.18.2	MIT License	An autocompletion tool for Python that can be used for text editors.
joblib	1.2.0	BSD License	Lightweight pipelining with Python functions
jupyter_client	8.2.0	BSD License	Jupyter protocol implementation and client libraries
jupyter_core	5.3.0	BSD License	Jupyter core package. A base package on which Jupyter projects rely.
kiwisolver	1.4.4	BSD License	A fast implementation of the Cassowary constraint solver
matplotlib	3.7.1	Python Software Foundation License	Python plotting package

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

matplotlib-inline	0.1.6	BSD 3-Clause	Inline Matplotlib backend for Jupyter
nest-asyncio	1.5.6	BSD License	Patch asyncio to allow nested event loops
numpy	1.24.3	BSD License	Fundamental package for array computing in Python
packaging	23.1	Apache Software License	Reusable core utilities for various Python Packaging interoperability specifications.
pandas	2.0.2	BSD License	Powerful data structures for data analysis, time series, and statistics
parso	0.8.3	MIT License	A Python Parser
patsy	0.5.3	BSD License	A Python package for describing statistical models and for building design matrices.
pickleshare	0.7.5	MIT License	Tiny 'shelve'-like database with concurrency support
platformdirs	3.5.1	MIT License	A small Python package for determining appropriate platform-specific dirs, e.g. a "user data dir".
prompt-toolkit	3.0.38	BSD License	Library for building powerful interactive command lines in Python
psutil	5.9.5	BSD License	Cross-platform lib for process and system monitoring in Python.
pure-eval	0.2.2	MIT License	Safely evaluate AST nodes without side effects
pyparsing	3.0.9	MIT License	pyparsing module - Classes and methods to define and execute parsing grammars
python-dateutil	2.8.2	Apache Software License	The dateutil module provides powerful extensions to the standard datetime module, available in Python.
pytz	2023.3	MIT License	World timezone definitions, modern and historical
pywin32	306	Python Software Foundation License	Python for Window Extensions
pyzmq	25.1.0	BSD License	This package contains Python bindings for ZeroMQ. ØMQ is a lightweight and fast messaging implementation.
scikit-learn	1.2.2	BSD License	A set of python modules for machine learning and data mining
scipy	1.10.1	BSD License	Fundamental algorithms for scientific computing in Python
seaborn	0.12.2	BSD License	Statistical data visualisation
six	1.16.0	MIT License	Python 2 and 3 compatibility utilities
stack-data	0.6.2	MIT License	Extract data from python stack frames and tracebacks for informative displays
statsmodels	0.14.0	BSD License	Statistical computations and models for Python
threadpoolctl	3.1.0	BSD License	threadpoolctl
tornado	6.3.2	Apache Software License	Tornado is a Python web framework and asynchronous networking library, originally developed at FriendFeed.
traitlets	5.9.0	BSD License	Traitlets Python configuration system
typing_extensions	4.6.3	Python Software Foundation License	Backported and Experimental Type Hints for Python 3.7+
tzdata	2023.3	Apache Software License	Provider of IANA time zone data

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

xgboost	1.7.5	Apache Software License	XGBoost Python Package
zipp	3.15.0	MIT License	Backport of pathlib-compatible object wrapper for zip files

3.5.6 Integration with ENERSHARE Data Space

As of this moment, and for ENERSHARE's Alpha release, the load forecasting tool's integration with the ENERSHARE Data Space is not yet been established. As stated above, the tool will be connected to the Data Space via connector, in the context of providing a benchmark for a federated learning algorithm that takes advantage of consumer level data (from the Data Space) to make substation level forecasts.

3.6 Energy Usage Prediction Service

3.6.1 Description of the Service

The main goal of the service is to build data-driven models to predict energy consumption. We are mainly interested in smart meters (SM) of users and heat meters of larger buildings for which hourly data are available. The main idea of the service is to support grid planning of the energy mix with the models obtained from real measurements. Based on this, user profiles can be created and segmented, solving privacy issues by creating typical user clusters. This model will serve as additional input for the service in Section 4.4. The service consists of two tools: the model training tool and the load forecasting tool. Since the clustering of users by certain characteristics is done manually, we developed a simple tool that performs the training using the real data obtained from the SMs of the EKC grid of Pilot 3, Figure 29. This part is not shared with other services, but the models resulting from training are exposed for use in other services. Once the model is trained, it is available in the forecasting service. In this part, the only input needed to obtain a typical profile is the user's location and the selection of the weather forecast service to be used: a) PVGIS or b) open-meteo. PVGIS is better suited for long-term planning because its datasets contain typical meteorological data and are better suited for creating typical long-term studies that require typical profiles. By using the historical data or weather forecast from open-meteo.com, more detailed user profiles can be created and used for short-term planning. It should be noted that the tool retrieves all weather data automatically, minimizing the input required by the end user of the service.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

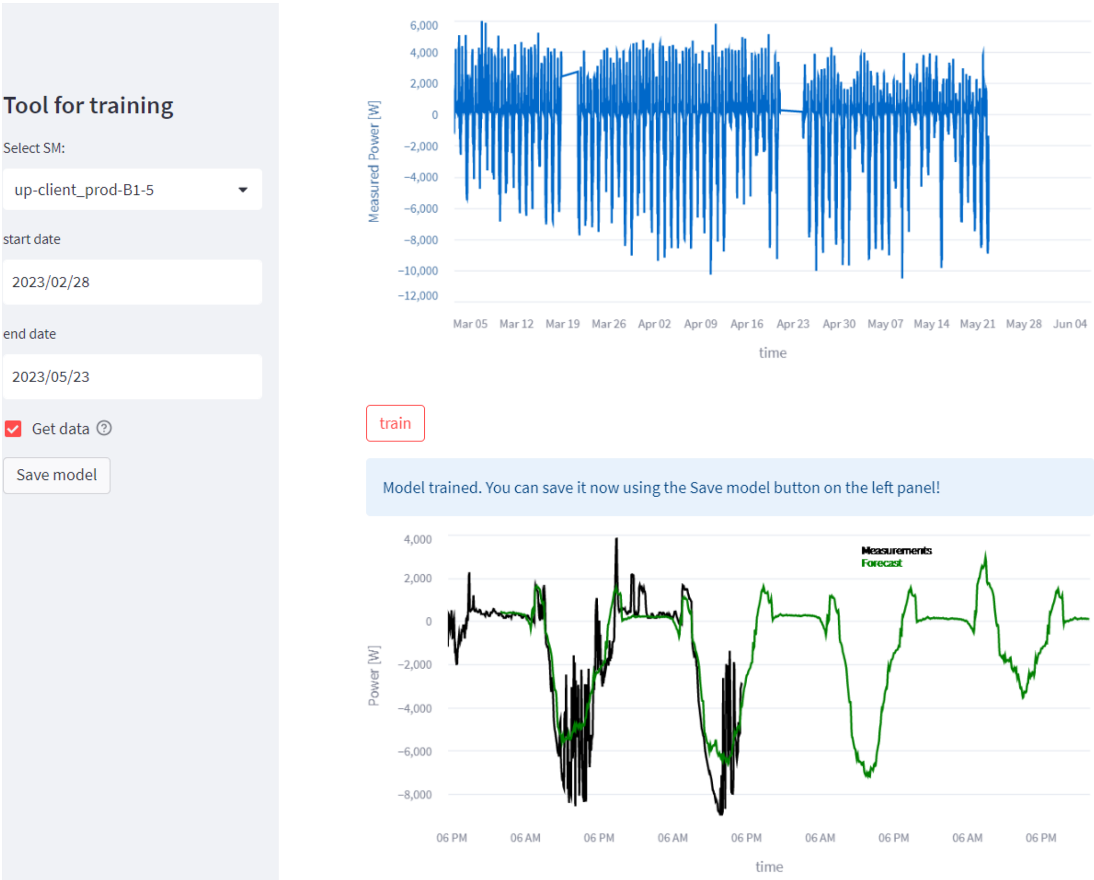


Figure 29: UI for training the SM forecasting models with the connection to EKC users DB.

3.6.2 Innovation

The service provides the ability to model different types of users and their electric or heating power profiles. In addition, the models allow finer and more detailed profiles, which are necessary in the case of electric profiles, as the commonly used 15-minute time slices may not be sufficient in the case of electric appliances. In the case of electrical user profiles, 5-minute intervals are used, providing an optimal balance between data granularity and quantity.

Another innovation is that training the models can also allow automatic partial disaggregation of the most interesting flexible loads only from smart metering measurements. The models created provide disaggregated data if available.

3.6.3 Functions

The service consists of the training module that connects to DB with SM measurements and creates the prediction model based on the historical SM meter data and weather data obtained

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

from https://open-meteo.com API, upper part of Figure 30. For ease of use, the training is automated and uses the UI, which was created using the Streamlit library. However, only the forecast part is provided to other services via fastAPI or directly as a Python library. Two forecasting models are created, referred to as Model A and Model B. Model A is only used to create SM forecasts, its output is used to label larger loads such as EV or HVAC. The new model is trained with additional labels as an additional regressor in the Prophet model.

To expose the functionality of disaggregation in the forecasting model, Model A first predicts the load profile of SM, from which an additional label for Model B is obtained. By setting different regressors to predicted values or zero, we can obtain different profiles with or without the labelled data. By subtracting the profiles with and without the label, we essentially obtain the load profile. For example, if we set the solar insolation to zero, we get the forecast SM without PV. If we subtract this forecast value from the SM forecast with insolation from the weather forecast, we get the PV profile. The same idea is used to obtain other loads.

Note that this is an additional feature. In most cases we will only use the forecast from SM, but if it is needed in the future, the disaggregated data is already available.

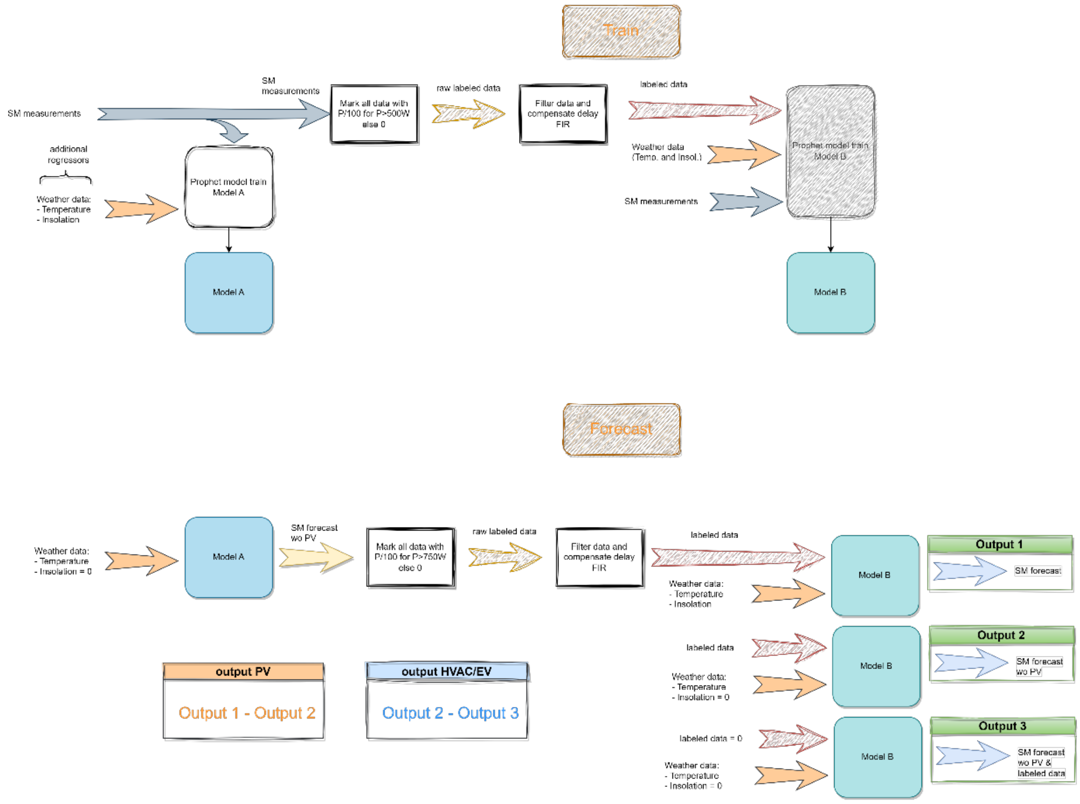


Figure 30: Program flow for training and forecasting data-driven user profiles.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.6.4 Input and Output Data Format

The input and data format are specified in Table 24. Some functions share the same variables. There are three calls that have also been implemented in fastAPI for testing purposes. One call is used to check which data-driven models are available. Then there are two calls for the historical power profile and forecasting power profile of a household. Both need as input the longitude, latitude, weather forecaster (PVGIS or open-meteo) and model name. In addition, the historical model also requires the date, while the forecast call requires the type of time format (ISO or epoch). An example of the output of the implemented API call for smart meter forecasting is (example request URL http://127.0.0.1:8000/forecasting?long=46.23&lat=15.26&pvgis=false&model=user_home_PV_HVAC&TimeFormat=ISO):

Table 24: Energy Usage Prediction Models data specifications


```
{
  "DateTime": [
    "2023-05-29T00:00:00",
    "2023-05-29T00:05:00",
    …,
    "2023-06-01T23:55:00"
  ],
  "SM": [
    480.1311002399249,
    493.59140460963425,
    …,
    354.1510012011453
  ],
  "PV": [
    0,
    0,
    …,
    0
  ],
  "HVAC": [
    0,
    0,
    …,
    0
  ],
  "Background": [
    480.1311002399249,
    493.59140460963425,
    …,
    354.1510012011453
  ]
}
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Variable	Source	Default Value	Unit	Type	Description
lon	Input	46.23	-	float	Longitude of household/Building
lat	Input	15.26	-	float	Latitude of household/Building
PVGIS	Input	False	-	bool	If True PVGIS is used, otherwise Open-meteo
model	Input	user_home_PV_HVAC	-	string	Name of the model to use, you can query the list of models with
TimeFormat	Input	ISO	-	string	ISO or epoch
date_v	Input	2023-03-15	-	date	EWH overall heat transfer coefficient
models	output	-	-	String/list	List of available models
Datetime	output	-	Y-m-d H:M:S	datetime/list	Timestamp of the specific period
SM	output	-	W	float/list	Modeled historical or forecasted smart meter power profile for selected day
Flow_meter	output	-	W	float/list	Modeled historical or forecasted energy flow meter power profile for selected day
PV	output	-	W	float/list	Modeled historical or forecasted PV power profile for selected day. Optional if available.
HVAC	output	-	W	float/list	Modeled historical or forecasted HVAC power profile for selected day. Optional if available
EV	output	-	W	float/list	Modeled historical or forecasted EV power profile for selected day. Optional if available.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

background	output	-	W	float/list	Modeled historical or forecasted non labeled profile for selected day. Optional if available.

3.6.5 Implementation details

Deliverable (Month)	TRL level	Comment
D6.1 (M12)	TRL 4	First test of the tool done on the real data, first model created based on this data.
D6.2 (M19)	TRL 5	Interconnect with other tools. A basic test using the Data Space connector would be made, if available.
D6.3 (M28)	TRL 6	Fully integrated in the pilot. If available, a connection using the Energy Data Space will be also tested.

The service was developed using Python v3 and its comprehensive set of standard libraries. Additionally, we leveraged the Prophet tool, licensed under MIT, to create accurate models. To improve accessibility and integration with other tools, the service will be containerized and deployed on a publicly accessible server.

3.6.6 Integration with ENERSHARE Data Space

As of this moment, and for ENERSHARE's Alpha release, the tools integration with the ENERSHARE Data Space is not yet been established. IDS connectors will be used for the exchange of data between services, daily load/generation profiles.

3.7 Service local load forecasts and estimation of electrical grid status

3.7.1 Description of the Service

The main goal of this service is to provide a functionality for grid estimation, based on inputs from decentralized local models. The service, called the grid simulator, aims to provide an alternative to a centralized grid state estimation approach, so that an insight into grid state can

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

be gained when a centralized approach does not exist or when it is not feasible due to data privacy and data scarcity concerns. This, for example, is the case with low-voltage (LV) grids, where grid observability is low due to a lack of measuring devices. In such a case, a bottom-up approach can be taken, where nodes in the grid (e.g., households) can locally forecast their consumption and send their forecasts to the grid simulator. The nodes (households) are agnostic of the grid topology and of other nodes' (households') forecasts. The grid simulator receives the grid topology from an external source (e.g., from the DSO), and local model forecasts are provided by the models described in Section 0. The grid simulator itself can be hosted on the cloud or at a third party's location (e.g. an aggregator). With such an approach, the DSO is agnostic about the individual loads on the nodes, but the grid observability is increased, as the DSO can get information on an overall state of the grid.

The grid simulator service is a service that takes as an input electricity local load forecasts and grid topology, based on which it determines the state of the grid in terms of transformer loads. The load on the transformer is estimated based on the forecasts for the coming 24 hours, and the service determines whether there is a chance for transformer overloading. Currently, only the load on the transformer is estimated, but parameters such as voltages, currents and loads on the lines can also be calculated.

3.7.2 Innovation

The service increases observability of electricity grids where a centralized approach does not exist, or it is unavailable due to data privacy or data scarcity concerns. It does so by using a decentralized, privacy-preserving Machine Learning technique, Federated Learning and external sources to retrieve a grid topology.

3.7.3 Functions

The grid simulator is in fact the global model for the Federated Learning platform (see in Section 0). The service calculates power flows and the state of the grid, and it consists of two building blocks: an OpenDSS component and a DSO Agent component. Figure 31 shows the grid simulator service and its components within the entire platform. The service and its components are indicated by the dashed line.

OpenDSS is a comprehensive electrical power system simulation tool primarily for electric utility power distribution systems. It can be used to perform complex analyses using a flexible, customizable, and easy to use platform intended specifically to meet current and future distribution system challenges and provides a foundation for understanding and integrating new technologies and resources. Using grid topology (including transformer and line specifications) and powers at nodes, OpenDSS can calculate the state of the grid, determining, for example, loads of lines and transformers. The DSO component is a part of the grid simulator that checks for transformer overloading. It should be noted that this component does not

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

represent a DSO, nor is it hosted at the DSO's location, but it simply performs the grid state analysis to determine if overloading exists.

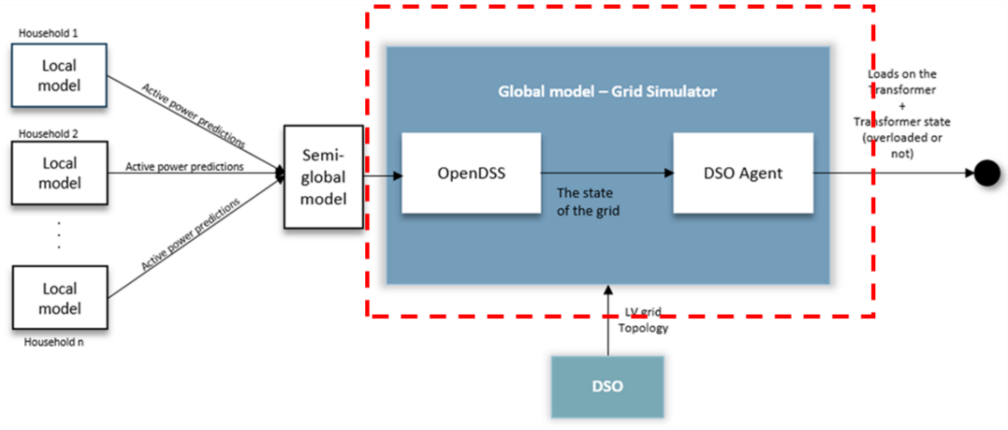


Figure 31: Grid simulator service - within dashed lines

3.7.4 Input and Output Data Format

As an input, the grid simulator service (more specifically, its OpenDSS component) receives active power predictions from local models (for a number of timesteps) in the form of CSV files, as well as a detailed LV grid topology, with grid component specifications (e.g., transformer capacity and overload threshold). Using this information, OpenDSS component determines the state of the grid per prediction timestep. At each timestep, OpenDSS asks the DSO Agent to perform its calculation and check the output of OpenDSS against transformer specifications (contained in the grid topology).

Table 25 lists the descriptions of input variables to the grid simulator that come from the local models. Note that the active power predictions from the heat pump come from every household, where "<<n>>" is replaced by the household number. Example local model inputs for two households are given in Table 26.

Table 25 Input variable descriptions – active power predictions from local models

Variable name	Variable type	Variable unit	Variable description
timestamp	Datetime	-	Date and time of prediction
USER<<*n*>>EDEMAND	Double	kW	Active heat pump power for a specific household consumer

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Table 26 Input variables example – active power predictions from local models

timestamp	USER4EDEMAND	USER11EDEMAND
31-12-2016 23:00	0.020173255	0.023264322
31-12-2016 23:15	0.024205938	0.024615599
31-12-2016 23:30	0.026022611	0.02609534
31-12-2016 23:45	0.027786981	0.02572422
1-1-2017 00:00	0.01573997	0.012853722
…	0.016865972	0.016718347
1-1-2017 22:30	0.019217351	0.018988341
1-1-2017 22:45	0.013275342	0.013188589

Besides the local model predictions, the grid simulator service receives an LV grid topology description from an external source (e.g., a DSO). The topology is described using *Energy System Description Language* (ESDL)[5] to get information about energy assets, carriers and commodities, infrastructure etc. ESDL is an open-source modelling language created for modelling the components in an energy system and their relations to each other. It is an XML schema definition (XSD) for formally describing energy assets in an XML format. This allows ESDL to be used as a formal specification of an energy grid for unambiguous interpretation by experts and tools. An example topology ESDL file can be found in Appendix 9.3.

The output of the grid simulator is the list of loads on the transformer, including an indication whether a transformer is overloaded. Table 27 shows the output variables of the grid simulator, their types and units.

Table 27 Output variables description

Variable name	Variable type	Variable unit
timestamp	Datetime	-
transformer_load	Double	kW
transformer_overloaded	Boolean	-

[5] Energy System Description Language

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

overload_threshold	Double	kW

Figure 32 shows the output of running the grid simulator service for an LV grid of 10 households. Power flow is calculated for every time step, and OpenDSS asks the DSO Agent to perform its calculation and check the output of OpenDSS against transformer specifications (contained in the grid topology). The DSO Agent checks for overloading and notifies the interested parties. The results of the grid state calculations are written into InfluxDB. The final output, including transformer loads for every time step and whether it is overloaded or not, is shown in Table 28.

```
Solving power flow. Time-step: 2017-01-01 13:30:00
Calculating the state of the grid. Time-step: 2017-01-01 13:30:00
Sending to DSO Agent...
DSO: Done processing DFs
...Execution Finished.
Writing to Influxdb
Asking DSO Agent to perform calculation
Transformer is overloaded
Solving power flow. Time-step: 2017-01-01 13:45:00
Calculating the state of the grid. Time-step: 2017-01-01 13:45:00
Sending to DSO Agent...
DSO: Done processing DFs
...Execution Finished.
Writing to Influxdb
Asking DSO Agent to perform calculation
Transformer is not overloaded
Solving power flow. Time-step: 2017-01-01 14:00:00
Calculating the state of the grid. Time-step: 2017-01-01 14:00:00
Sending to DSO Agent...
DSO: Done processing DFs
...Execution Finished.
Writing to Influxdb
Asking DSO Agent to perform calculation
Transformer is not overloaded
```

Figure 32: Grid simulator output

Table 28 Grid simulator output

timestamp	transformer_load	transformer_overloaded	overload_threshold
31-12-2016 23:00	0.06655793410920938	FALSE	0.07
31-12-2016 23:15	0.06785210269619803	FALSE	0.07
31-12-2016 23:30	0.08234543563808853	TRUE	0.07

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

31-12-2016 23:45	0.03865532853642178	FALSE	0.07
1-1-2017 00:00	0.05038579242859747	FALSE	0.07
…	…	…	…
1-1-2017 22:30	0.04758799035497654	FALSE	0.07
1-1-2017 22:45	0.05099188852674355	FALSE	0.07

Figure 33 below shows the plot of the transformer loads per time step. The red dashed line indicates the overload threshold, while the red dots indicate the times at which the transformer is overloaded.

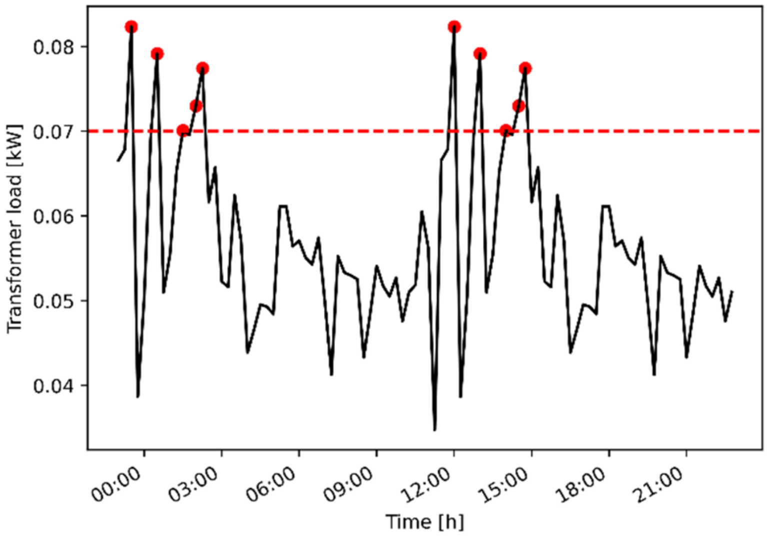


Figure 33: Estimated transformer loads and overloads

3.7.5 Implementation details

The global model containing energy grid simulator is on TRL 6 since it was successfully used for the specific use-case as an operational demonstrator, but it is not yet deployed and tested in actual operational settings. The grid simulator service is implemented in Python 3.7 and uses InfluxDB to store the data. The solution is packed in Docker container by using Docker version 22.0. The Docker image with run instructions to use as a service will be provided. To run the grid simulator.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.7.6 Integration with ENERSHARE Data Space

TNO's grid simulator service will become an integral optional part of the IDS connector. There will be the possibility to include the special client for local, semi-global and global models into the basis IDS connector developed by TNO. The source-code of the TNO IDS connector is to find here: TNO Security Gateway Documentation (https://tno-tsg.gitlab.io). Every type of client will be integrated separately and will be using the communication basis from IDS connector.

3.8 Flexibility Analytics and Register (FAR) service

3.8.1 Description of the Service

The developed service aims to support flexible service providers at any level, from lower-level residential users up to aggregators, positioning themselves in the flexibility and ancillary grid services. The provided service allows the individual service providers to register their flexible assets in a secured data base. The Flexibility Analytics and Register service (FAR) empowers Flexibility Service Providers (FSPs) to bridge their flexibility resources to markets. The Flexibility Register contains fundamental information on the technical characteristics and the location of connection points that can provide flexibility services to system operators. It aims to gather and share relevant information on potential sources of flexibility. It acts as a central repository, including technical information on resources that are connected to the joint area of responsibility of different System Operators (i.e., not limited to electricity system operators). The idea of this integrated data middleware platform is to bundle multiple flexibility actors connected in along the grid and enable thereafter the extraction of aggregated analytics that could support the understanding of flexibility providers on flexibility needs in the grid across timeline. The integrated data in the FAR aims to facilitate cross-sector data towards the potential flexibility sharing among sectors. Within this instance heating data will be ingested to the FAR allowing for integrated analytics for electricity and heating sector.

The FAR provides proper interface for Market Operator(s) to define/configure flexibility product/services specifying the eligibility requirements. The following Figure presents a conceptual approach of the provided analytics, the expected inputs and the retrieval of the analytics from the Flexibility Register.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

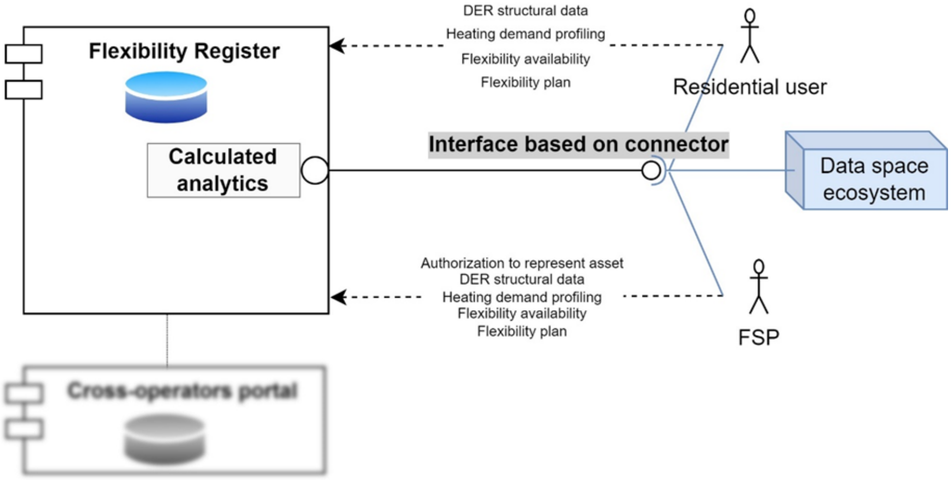


Figure 34: Logic view of Flexibility Register service.

The **FAR core processes** that are performed within this module include:

- user management,
- DER (including heating details) registration,
- product definition (for Market Operator),
- calculated analytics for flexibility providers.

3.8.2 Innovation

The proposed FAR services considers the implementations and general guidelines of IEGSA (Interoperable pan-European Grid Services Architecture), developed on INTERRFACE project, advancing them to accommodate useful information for sub-metered resources that belong to residential users. More in particular, the Flexibility Register will be providing open features for residential users or delegated representative market parties, e.g., aggregator, to register and catalogue their flexible sources that could potentially provide grid (electricity and thermal) services. For this purpose, analytical information of sub-metered flexibility availability as well as a registry of the flexible devices will be reported. The functions of IEGSA pertaining resource and product qualification will be adapted as per the demo needs.

3.8.3 Functions

The main process flow of the Flexibility Register service for any flexibility service provider is:

- **User authentication:**

The Auth request requires the username and the password of the FSP. The current implementation for the Alpha version is based on jwt authentication and the subsequent

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

generated token per user; on the next stage for complete integration of this service with other functions, it will be proposed in the pilot level to adopt Keycloak as a horizontal authentication service. The structure of the request is as shown:

```
POST /api/users/auth HTTP/1.1

Host: {{application url}}:{{port}}

Content-Type: application/json

Content-Length: 42

Body

{

"username":"{{username}}",

"password":"{{password}}"

}
```

If the username and password are valid the response will contain a JWT token (that will be used to grand access to the next requests) and a user DTO entity containing the users' details.

```
{
    "jwt": "{{Jwt token response}}",
    "userDTO": {
        "id": 1,
        "createdOn": "2021-02-08T00:00:00Z",
        "createdBy": "admin",
        "username": "admin",
        "password": null,
        "role": "****",
        "email": "****",
        "country": "****",
        "eicCode": "****"
    }
}
```

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- **Create resource entries**

Each residential user will be able to register information in the FAR database information on the flexible/controllable devices. Those can be accordingly, uploaded utilizing the API call POST /resources/

The expected body shall have the following attributes:

```
{
  {
  "resource_id": "EV1",
  "baseInfo: {
    "Name": "****",
    "Description": "Residential EV",
    "MeteringPointID": "EIC/GSRN",
        "VoltageLevel": "admin",
        "ResourceOwnerEmail": "..",
  }
     "locationInfo":{
    "Country": "****",
    "Area": "****",
     "PostalCode": "****"
 "DERtype": "EV/mG/BESS/CL",
    "StreetAddress": "****",
 }
     "Details":{
    "NominalPower": "[kW]",
    "NominalCapacity": "[min]",
     "RampingPeriod[min]": "[min]"
 "FullActivationTime": "[min]",
    "DeactivationLength": "[min]",
```

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
"MinimumDurationLen": "[min]",
    "MaximumDurationLen": "[min]",
    "MaxReboundAfterUp": "[min]"
"MaxReboundAfterDwn": "[min]",
    "MaxUpPower": "[kW]",
    "MaxDwnPower": "[kW]"
"ActualUp": "[kW]",
    "ActualDwn": "[kW]",
}
}
```

Accordingly, within resources the heating assets specification for residential user will be able to be stored in the FAR's database. Nonetheless, the expected dataset is under specification.

A given resource can be updated by a user utilizing the PUT /resources/

- **Creating a resource group**

A user/FSP can select a set of resources, group them and make them available for particular services (e.g., mFRR services), using the POST / resource-groups/flexibility-register

```
{
    "name": "Res Group 5",
    "description": "mixed DER",
    "resourceDTOs": [
        { "id": ID1, ID2}
    ],
    "productDTOs": [
        { "type": "mFRR",  }
    ]
}
```

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- **Flexibility availability**

Important steps are the actualization of the flexibility availability in the FR database by uploading the available actual up/down power on the POST /api/resource/a14-actual-regulation-power

The expected data format for is the following:


```
<PlannedResourceSchedule_MarketDocument>xmlns="urn:iec62325.351:tc57wg16:451-7:plannedresourcescheduledocument:6:0">

  <mRID>3715c5f3-557e-4384-9969-91b1006b000</mRID>

  <revisionNumber>1</revisionNumber>

  <type>A14</type> <!--A document providing the schedules for resource objects submitted by a resource provider-->

  <process.processType>A19</process.processType><!--Reguesting from Entso-e -->

<sender_MarketParticipant.mRID codingScheme="A01">42F4652A9381D904</sender_MarketParticipant.mRID>
<sender_MarketParticipant.marketRole.type>A27</sender_MarketParticipant.marketRole.type>

<receiver_MarketParticipant.mRID codingScheme="A01">10V1001C--00239S</receiver_MarketParticipant.mRID><!--Platform EIC code-->
<receiver_MarketParticipant.marketRole.type>A35</receiver_MarketParticipant.marketRole.type><!--Platform role-->

  <createdDateTime>2021-12-21T06:00:00Z</createdDateTime>

  <schedule_Period.timeInterval>

    <start>2021-12-21T07:00Z</start>

    <end>2021-12-21T15:00Z</end>

  </schedule_Period.timeInterval>

  <domain.mRID codingScheme="A01">10YFI-1--------U</domain.mRID> <!--EIC code of region. Estonia in this case-->

  <subject_MarketParticipant.mRID codingScheme="A01">XXXXXX</subject_MarketParticipant.mRID>
<subject_MarketParticipant.marketRole.type>A27</subject_MarketParticipant.marketRole.type>

  <!--Zero or more repetitions:-->

  <PlannedResource_TimeSeries>

    <mRID>5698201254879633</mRID> <!--Unique identification of the timeseries-->

    <businessType>A61</businessType> <!--maximum available A61-->
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
    <flowDirection.direction>A01</flowDirection.direction><!--A01 = UP A02 = DOWN-->

    <product>8716867000146</product>

    <connecting_Domain.mRID codingScheme="A01">10YFI-1--------U</connecting_Domain.mRID> <!--
Specific measurement point EIC code-->

    <registeredResource.mRID    codingScheme="A01">0OL8IGU7L</registeredResource.mRID>    <!--
Resource identification. Poissibliy group EIC-->

<resourceProvider_MarketParticipant.mRID
codingScheme="A01">42F4652A9381D904</resourceProvider_MarketParticipant.mRID><!--EIC  code  of
region. Estonia in this case-->

    <acquiring_Domain.mRID codingScheme="NLU">10YFI-1--------U</acquiring_Domain.mRID>

    <marketAgreement.type>A01</marketAgreement.type><!--The     type     of     the     market
agreement-->

    <measurement_Unit.name>MAW</measurement_Unit.name>

    <objectAggregation>A01</objectAggregation> <!--Agregated per code-->

    <!--1 or more repetitions:-->

    <Series_Period>

      <timeInterval>

     <start>2021-12-21T07:00Z</start>

         <end>2021-12-21T15:00Z</end>

      </timeInterval>

      <resolution>PT1H</resolution>

      <!--1 or more repetitions:-->

      <Point>

        <position>1</position>

        <quantity>0.7</quantity>

        <!--Zero or more repetitions:-->

        <Reason>

          <code>A95</code>

          <!--Optional:-->

          <text>additional information</text>

        </Reason>

      </Point>
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
    &lt;/Series_Period&gt;
  &lt;/PlannedResource_TimeSeries&gt;
&lt;/PlannedResourceSchedule_MarketDocument&gt;
```

- **Analytics retrieval**

Several analytics are calculated at FAR aiming to share aggregated flexibility availability useful data for multiple actors (flexibility service providers, typical end-users, market operator, System operators). Below can be found a set of available inquiries that can be made to the flexibility register:

1. GET api/analytics/ flexibility-register/electricity_aggregated_up

Calculates the total (nominal) connected aggregated up-regulation power in the grid.

2. GET api/analytics/flexibility-register/electricity_aggregated_up_current

Calculates the actual connected aggregated up-regulation flexible power in the grid for the current Market Time Unit.

3. GET api/analytics/flexibility-register/electricity_aggregated_dwn

Calculates the total connected aggregated down-regulation power in the grid.

4. GET api/analytics/flexibility-register/aggregated_dwn_current

Calculates the actual connected aggregated down-regulation flexible power in the grid for the current Market Time Unit.

5. GET api/analytics/flexibility-register/resource-groups?product-type=CM&date=XXX

This API provides all the aggregated nominal capacity for resources/resource groups that are assigned to provide a specific product or services.

6. GET api/analytics/flexibility-register/flexibility-needs?date=XX

This API shall provide current flexibility need for both electricity and thermal grids. This response type and the availability (i.e., periodicity) of this analytic is upon definition with demo partners.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.8.4 Input and Output Data Format

The input and output data are shown in Figure 34, while in the section of Functions, documenting all the necessary APIs for importing data as well as how to retrieve the calculated analytics. All outputs are designed in JSON formats, in upcoming updates and based on the pilot needs proper modifications will be made.

3.8.5 Implementation details

The Alpha version of the FAR described in this document is consider a TRL5, updated version that will be described in D6.2 and D6.3 will be evolve on TRL6 and TRL7, accordingly, considering the interaction with demo partners in Pilot 3 (i.e., datasets specification, interaction with external services). The next version will consider the integration with data connector for the trusted and secure sharing of FAR's analytics in the ENERSHARE data space ecosystem.

The FAR software and its components will be delivered utilizing the Docker containers functionalities. The service will be available online in four specified co-existing docker images to efficiently distribute the software avoiding the likelihood of collisions. Therefore, there is one image specified to host the backend application (flexibility_reg-backend) as well as one for the frontend application (flexibility_reg -frontend). The other two docker images are created to host the FAR database (db) and the MySQL one for administrative purposes that provides access to DB of FR.

Database

The FAR's DB is being developed with an underlying relational database system to store, manage and safeguard system data. Communication with the backend is being realised through secure REST API calls and data exchange. The database is not available publicly, instead the backend is responsible to communicate with it internally and provide the necessary public interface. This can be utilized for either Human to Machine or Machine to Machine communication.

Internet access

The Backend system is not considered to be accessible through the internet and should communicate only with authorised machines (authorised in terms of pre-defined) following security mechanisms as in a following chapter.

Machine-to-machine communication is being realised through secure REST API calls and data exchange.

FAR's Technology Stack

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

The technologies used for FAR service development are listed in the following paragraphs. All technologies and tools are open source, thus there is no requirement for the demos to acquire certain commercial licenses to deploy them.

Languages

FAR service is being programmed using the following programming languages: Java, HTML, CSS, Typescript.

Frameworks

The frameworks used for FAR development are: Spring, Angular, JUnit 5.

Server Technologies

Server technologies used for development and deployment of FAR are: Docker, HTTPS, Unix shell, Bash, Symmetric/Asymmetric Encryption, CentOS, MariaDB, SQL, nginx.

3.8.6 Integration with ENERSHARE Data Space

The described FAR service will be openly available to the ENERSHARE data space ecosystem through the App Store. Additionally, the retrieval of the calculated analytics will be accessible through connector APIs (NGSI- APIs) for the secured and trusted information exchange.

3.9 Aggregation of flexibility from end users

3.9.1 Description of the Service

3.9.1.1 Introduction

This service is developed by Fortum's Energy Science Team, on behalf of Hiven[6], an internal startup of Fortum, which aims to offer a complete solution for residential energy management. As of now, Hiven offers a Smart Charging application, which users with compatible chargers and vehicles can connect with, and let it schedule the charging of their EV during the cheapest hours of the day. Our goal is to augment this service with the option to perform Demand Response (DR), which will attenuate the charging of each individual EV in the Hiven network according to the live grid frequency. This way, the flexibility of many residential users can be aggregated and be used to participate in markets for Ancillary Services, enabling value creation for both grid operators and vehicle owners.

[6] https://hiven.energy

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

In the Nordic countries, the Primary Reserves Market, also known as Frequency Containment Reserve (FCR), is split into three categories, depending on the grid frequency deviation:

- **FCR-N:** Normal deviation, grid frequency is in the band [49.9, 50.1] Hz.
- **FCR-D Up:** Disturbance deviation, grid frequency is in the band [49.5, 49.9] Hz.
- **FCR-D Down:** Disturbance deviation, grid frequency is in the band [50.1, 50.5] Hz.

In our implementation, we are currently focusing on the Swedish FCR-D Up market, but our approach can be trivially extended to any Nordic FCR-D Up market, as they are subject to almost the same rules. Furthermore, the general logic can also be applied to other FCR markets with minor revisions.

In the case of FCR-D Up, given a frequency deviation $\Delta f_t = f_t - f^{\text{nom}}$, and the placed reserve capacity bid C^{UP}, the market participant needs to modify its baseline consumption P_t^{base} by $\Delta P_t = -C^{\text{UP}}\alpha_t^{\text{UP}}$ where α_t^{UP} is the activation level imposed by the current frequency deviation:

$$\alpha_t^{\text{UP}} = \begin{cases} 0, & 49.9\,\text{Hz} \leq f_t \\ 2.5\,\Delta f_t, & 49.5\,\text{Hz} \leq f_t \leq 49.9\,\text{Hz} \\ 1, & f_t \leq 49.5\,\text{Hz} \end{cases}$$

Therefore, the DR service is responsible for modulating the power of each EV charger, so that, in aggregate, the whole consumption of the charger fleet tracks $P_t^{\text{base}} + \Delta P_t$. This is achieved by disaggregating the reference consumption $P_t^{\text{base}} + \Delta P_t$ into individual commands, which are then submitted to each charger through Hiven's network. In the next subsection, we will describe the software architecture of the DR service.

3.9.1.2 Architecture

The DR service is implemented as a worker that reads the live grid frequency from a queue, in the form of `FrequencyEvent` messages (c.f. Section 3.9.4), queries the Hiven network through REST endpoints for Capacity bid and charger telemetry messages (`DeviceTelemetry`, c.f. Section 3.9.4) of each charger in its network. Then, all this information is utilized in the disaggregation algorithm which produces commands for each individual charger. These commands are submitted to another REST endpoint of Hiven's network, in the form of `ChargerSetCurrentCommand` messages (c.f. Section 3.9.4). This is represented pictorially in Figure 35.

3.9.2 Innovation

We aim to exploit the increasing prevalence of EVs by utilizing their batteries and rather predictable and flexible charging schedule to provide balancing services to help stabilize grid frequency. This way we empower end users behind the meter, with no market access, to

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

participate in Ancillary Services, enabling a win-win scenario for both vehicle owners and grid operators, creating value for both parties. For this to become a reality, we developed a service that can steer and manage distributed flexible resources, such as EV chargers, over unreliable connections, in fast timescales. This requires a bespoke control solution that can deal with the uncertain nature of the problem, while being fast, reliable, and fair to the end users.

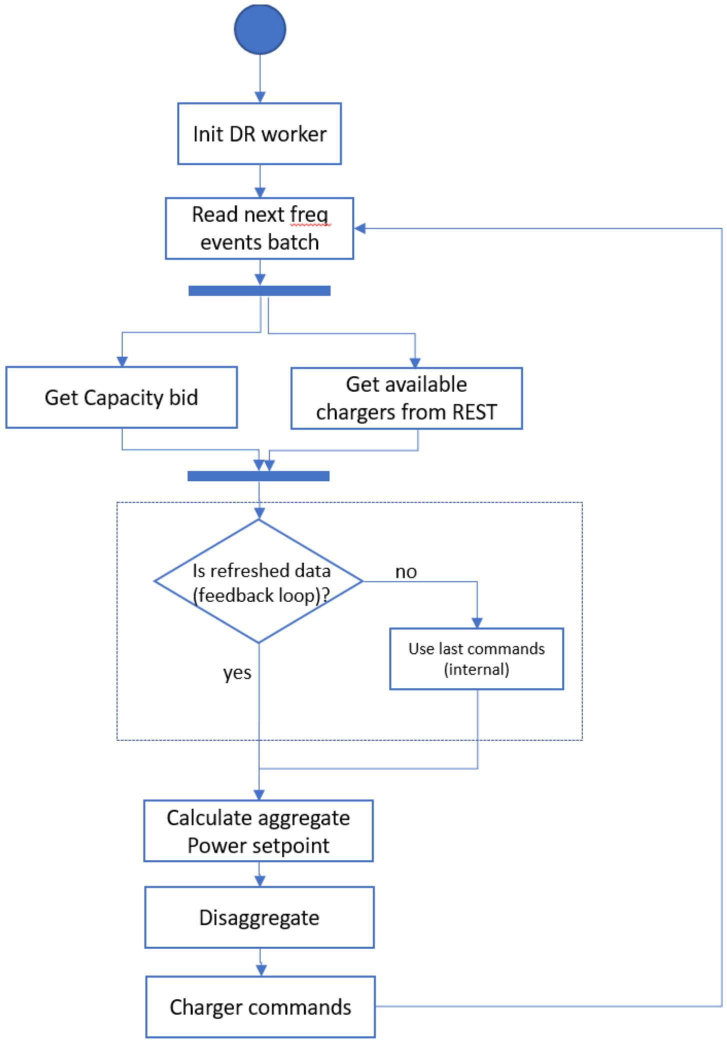


Figure 35: Architectural diagram of the DR service.

3.9.3 Functions

Here we will explain in more detail the individual components of the architecture diagram presented in Figure 35.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.9.3.1 Init DR Worker

Initialize the worker and instance variables, read settings from environment variables, authenticate and connect with appropriate cloud services.

3.9.3.2 Read Frequency Events

Query the frequency FIFO queue for frequency measurements in the form of `FrequencyEvent` messages (c.f. Section 3.9.4, Input and Output Data Format). If there are new measurements, pop them from the queue, and forward them to the next part of the algorithm, which is only triggered upon new measurements. The queue is assumed to be populated with frequency measurements by another, external actor, which will possess the grid frequency measurement device.

3.9.3.3 Get Capacity Bid

Query Hiven's REST endpoints to get the capacity bid we have placed in the day-ahead market for the current hourly block and price area. For now, this is not implemented, and is assumed to be constant, specifically the minimum possible bid, i.e., 100 kW.

3.9.3.4 Get Charger List

Query Hiven's REST endpoints to get a list of chargers that are currently online. Then for each charger, query Hiven's REST endpoints again to get the latest `DeviceTelemetry` for each one (c.f. Section 3.9.4, Input and Output Data Format).

3.9.3.5 Calculate Aggregate Power Setpoint

As described in the introduction, the aggregate power setpoint is calculated as

$$P_t^{\mathrm{set}} = P_t^{\mathrm{base}} + \Delta P_t .$$

whereP_t^{base} represents the baseline aggregate consumption of the whole charger pool, which is calculated as

$$P_t^{\mathrm{base}} = \sum_{c \in C_t} P_c^{\max} ,$$

as it assumed that, under normal operating conditions, every charger should consume its maximum power. In the above, the index set C_t represents all the chargers of the pool that at time instant t are online and charging, and are drawing their configured maximum current. The difference ΔP_t is calculated as described in the Introduction (Section 3.9.1.1).

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.9.3.6 Disaggregation algorithm

A key component of the DR service is the disaggregation algorithm, which assigns commands to each charger, so that the aggregate consumption of the charger fleet tracks the requested setpoint P_t^{set}. Each charger Is assumed to be able to obtain a finite set of distinct power values, and we toggle only between the minimum and maximum power setting.

This could potentially be cast as a change-making problem, but because the DR service needs to act very fast upon receiving a `FrequencyEvent`, we opted for a greedy algorithm that instead attempts to reduce the tracking error at each iteration with as few iterations as possible by choosing the charger with the maximum flexibility offered, i.e., the difference between maximum and minimum consumption.

A pseudocode description of the greedy disaggregation algorithm is presented below:

Algorithm 1 Disaggregration method

1: **Input:** Power setpoint P_t^{set}; import, min, and max power for each charger $\{(P_{c,t}, P_c^{\text{min}}, P_c^{\text{max}})\}_{c \in \mathcal{C}_t}$
2: Calculate total aggregate import power: $P_t^{\text{im}} \leftarrow \sum_{c \in \mathcal{C}_t} P_{c,t}$
3: Initialize the set of chargers that are not to be used: $\mathcal{N}_t \leftarrow \{\emptyset\}$
4: **for** $c \in \mathcal{C}_t$ **do**
5:     **if** $\text{sign}(P_t^{\text{im}} - P_t^{\text{set}}) \cdot (|P_{c,t} - P_c^{\text{min}}| - |P_{c,t} - P_c^{\text{max}}|) < 0$ **then**
6:         Append c to $\mathcal{N}_t$: $\mathcal{N}_t \leftarrow \mathcal{N}_t \cup \{c\}$
7:     **end if**
8: **end for**
9: Initialize the commanded power delta: $\Delta P^{\text{com}} \leftarrow 0$
10: Initialize the set power setpoints of each charger: $\mathcal{U}_t \leftarrow \{\emptyset\}$
11: **while** $|\mathcal{C}_t \backslash N_t| > 0$ **do**
12:     $c^* \leftarrow \text{argmax}_{c \in \mathcal{C}_t \backslash N_t} P_c^{\text{max}} - P_c^{\text{min}}$
13:     Let $P^{\text{new}} \leftarrow P_t^{\text{im}} + \Delta P^{\text{com}} - \text{sign}(P_t^{\text{im}} - P_t^{\text{set}}) \cdot (P_{c^*}^{\text{max}} - P_{c^*}^{\text{min}})$
14:     **if** $|P^{\text{new}} - P_t^{\text{set}}| < |P_t^{\text{im}} + \Delta P^{\text{com}} - P_t^{\text{set}}|$ **then**
15:         Let $\Delta P^{\text{com}} \leftarrow \Delta P^{\text{com}} - \text{sign}(P_t^{\text{im}} - P_t^{\text{set}}) \cdot (P_{c^*}^{\text{max}} - P_{c^*}^{\text{min}})$
16:         Command power setpoint for charger c^*: $P_{c^*}^{\text{com}} \leftarrow P_{c^*}^{\text{min}}$, **if** $\text{sign}(P_t^{\text{im}} - P_t^{\text{set}}) > 0$, **else** $P_{c^*}^{\text{max}}$
17:         Append command to command list: $\mathcal{U}_t \leftarrow \mathcal{U}_t \cup \{P_{c^*}^{\text{com}}\}$
18:     **end if**
19:     Add c^* to the non-usable charger list: $\mathcal{N}_t \leftarrow \mathcal{N}_t \cup \{c^*\}$
20: **end while**
21: **Return:** List of power setpoint commands $\mathcal{U}_t$

To give a more high-level description of the above, Algorithm 1 firstly looks at the difference between the current aggregate consumption and the desired aggregate setpoint, and then removes from consideration all chargers that are already at the appropriate power level for that direction (lines 4-8): for instance, if we need to reduce our consumption, only chargers that currently at their maximum power level are considered. Then, it goes through the charger list in order of potential flexibility, i.e., $P_c^{max} - P_c^{min}$ (line 12), and creates commands so that the chosen chargers flip from maximum to minimum consumption and vice versa (line 16). If a point is reached, where changing the status of a charger yields a tracking error worse than our current

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

one (line 14) we remove that charger from consideration (line 19) and move to the next one until the charger list is empty.

This approach, while not necessarily optimal (in terms of error achieved) will result in an error that, is upper bounded by the maximum flexibility offered by any charger in C_t. Since in our case the Swedish TSO, Svenska Kraftnät (SVK), allows a +10%/-25% allowed deviation from the setpoint[7], for the minimum capacity bid of 100 kW, this is far greater than the flexibility offered by any of the chargers in our fleet. Furthermore, this should present even less of a problem when moving to larger bids, as the contribution of a single charger would be a much smaller percentage of the aggregate consumption.

Practical considerations

As this service is developed with the goal to be deployed to real life system, there are many practical considerations that need to be taken into account.

- **Noisy measurements:** Even though the power levels are assumed to be distinct, the measurements we get through the `Telemetry` messages are often noisy due to fluctuations in both the voltage level provided and the current drawn. This can make it hard to calculate our exact baseline power, which needs to be reported to the TSO. Therefore, when performing such calculations, as in the disaggregation algorithm, we round the received measurements to the nearest distinct power (current, voltage, respectively) level setting.
- **Latency and reliability issues:** As we are dealing with resources that are distributed over a network, and specifically over residential internet connections, we expect a great degree of variability both in latencies and in dropout rates. Furthermore, when a power setpoint command is issued to a charger, it needs to perform a transaction with the plugged-in EV before applying it. Therefore, on the next polling we might receive a `Telemetry` with a charger state that seems to be unaffected by the command. These stochastic actuator latencies can be somewhat mitigated by keeping track of the last issued charger command, and take it into account during disaggregation: if the algorithm were to issue the same command to a charger (line 16), we skip that charger and add it to the non-usable charger list N_t for that instant. This is not reflected in the pseudocode description of the algorithm for the sake of presentation and clarity.
- **Performance considerations:** Implementing the proposed as algorithm as presented would result in a time complexity of $O(n^2)$, but simply sorting the charger list by flexibility, i.e.,$P_c^{max} - P_c^{min}$, would improve the complexity to an average case of $O(n \log n)$.

[7] https://www.svk.se/aktorsportalen/bidra-med-reserver/forkvalificering/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- **Command data model:** As described in Section 3.9.4, the issued commands are in reality current limit settings (`ChargerSetCurrentLimitCommand`) but translating from active power levels to current levels is straightforward. We opted for power setpoints in Algorithm 1 just for the sake of presentation.

3.9.3.7 Submit Charger Commands

The calculated charger commands, which follow the schema of ChargerSetCurrentLimitCommand (c.f. Section 3.9.4, Input and Output Data Format), are then submitted to the Hiven network through a REST endpoint, so that they are forwarded to each charger individually. Furthermore, as mentioned before, upon successful submission a local copy of each command is saved to the worker's memory, to be later incorporated in a future call of Algorithm 1.

3.9.3.8 Simulator

Finally, while not present in the architecture diagram of Section 3.9.1, a simple simulator of a charger fleet was developed so that it can be used for testing and continuous integration purposes, as well as enabling rapid prototyping without involving actual hardware. It conforms to the same data schemas as those described in Section 3.9.4, but its inputs are the outputs of the DR service, and its outputs are simulated charger `DeviceTelemetry` messages. This component currently simulates the application of commands, the noisy nature of the measurements, the stochastic latencies, and the overall asynchronous nature of the control problem.

3.9.4 Input and Output Data Format

In this section we describe the data schemas utilized by the DR service and how it communicates with other parties, in our case Hiven's internal network. As mentioned previously, we communicate with Hiven's network through REST endpoints transferring JSON messages. Our data model is heavily inspired by OCPP 1.6[8], and, for the most part, it can be viewed as an implementation of a subset of it.

[8] https://www.openchargealliance.org/protocols/ocpp-16/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Table 29: DeviceTelemetry data schema

Variable	Description	Type
`timestamp`	Timestamp when message was received by Hiven's network.	`datetime`
`device_id`	Alphanumeric unique identifier.	`str`
`telemetry`	Telemetry message, see Table 3.9.2	`Telemetry`

Table 30: Telemetry data schema

Variable	Description	Type
`connector_status`	Status of connector (referred to `ChargePointStatus` in OCPP 1.6). See Table 3.9.3.	`ChargerConnectorStatus`
`connector_status_timestamp`	Timestamp when connector status was applied.	`datetime`
`measurement`	Measurement message, see Table 3.9.4	`TelemetryMeasurement`

Table 31: ChargerConnectorStatus data schema. For an extended explanation please refer to the OCPP 1.6 standard[9]. This is equivalent to ChargePointStatus.

Value	Description
`AVAILABLE`	Connector is available.
`PREPARING`	Connector is not available but charging has not yet started.
`CHARGING`	Charging is ongoing.
`SUSPENDED_EVSE`	Connected to EV but not offering energy.
`SUSPENDED_EV`	Connected to EV but it is not accepting energy.
`FINISHING`	Charging has finished but connector is still unavailable.

[9] https://www.openchargealliance.org/protocols/ocpp-16/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

RESERVED	Connector is reserved.
UNAVAILABLE	Connector is not available.
FAULTED	Error/fault has occurred.

Table 32: TelemetryMeasurement data schema

Variable	Description	Type	Unit
timestamp	Timestamp of measurement.	datetime	-
power_active_import	Instantaneous active power consumed (imported).	float	W/kW
current_offered	Maximum current limit set by hardware.	int	A
current_imports	Instantaneous current line imports.	Tuple[float, float, float]	A
voltages	Instananeous line voltage measurements.	Tuple[float, float, float]	V

Table 33: FrequencyEvent data schema

Variable	Description	Type	Unit
frequency	Grid frequency measured.	float	Hz
received_at	Timestamp when received.	datetime	-
originated_at	Timestamp when measured.	datetime	-

Table 34: ChargerSetCurrentLimitCommand data schema

Variable	Description	Type	Default value
name	Name/type of command.	str	command:energy:set_current
value	Current limit to be set in A.	int	-

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

3.9.5 Implementation Details

3.9.5.1 Technology Readiness Level

We classify our current work as TRL 5. A great deal of the work involved in this was not on the algorithm development side but on putting the groundwork for having a software foundation upon which to build further upon. Additionally, significant effort went into integrating with Hiven's internal APIs and having the DR service coexist harmoniously with the existing infrastructure. Be that as it may, we currently have a service that is end-to-end operational. This was verified by performing integration tests with real hardware and we have recorded success for the whole pipeline from ingesting (virtual) frequency measurements to modulating the power consumption of chargers accordingly.

The algorithm needs further research and improvement, as it is heavily biased towards steering chargers with large flexibility, which means these will be used much more often than others. Furthermore, performance could potentially be improved by optimizing the code even more towards our use case. Additionally, future work will include adapting to the new SVK rules for FCR-D[10], which are to go into force on September 2023[11].

To showcase the validity of our current approach we present the performance of a simulated charger pool of 67 chargers with different flexibilities. The baseline consumption is 240 kW, and for a capacity bid of 100 kW, we simulate the aggregated response of the pool when subjected to three (simulated) frequency signals. The results are presented in Figure 36-Figure 38.

Now that a foundation for controlling a pool of distributed chargers has been established, further work could be also allocated on forecasting the available flexibility so that it can be nominated on the day ahead market, and determining what would the optimal capacity bid be given such a forecast. The aim for the DR service is to get prequalified by SVK, participate in the Ancillary Services Market, and be a production-grade part of Hiven's ecosystem.

[10] https://www.svk.se/contentassets/4f9ae6d3a8e449c291c9ddda105c7c72/fcr-technical-requirements-14.03.2022.pdf

[11] https://www.svk.se/en/about-us/news/news/implementation-of-new-technical-requirements-for-fcr/

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

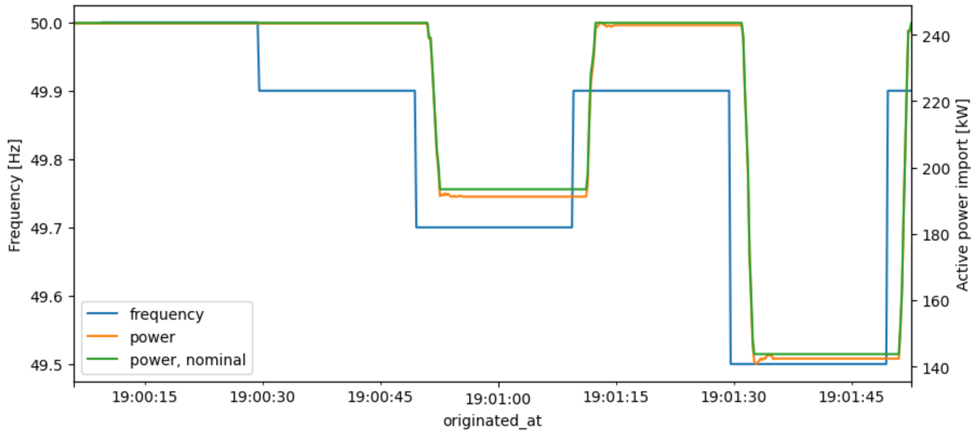


Figure 36: Simulated capacity test of a simulated charger pool with 67 units with baseline consumption of 240 kW and a capacity bid of 100 kW.

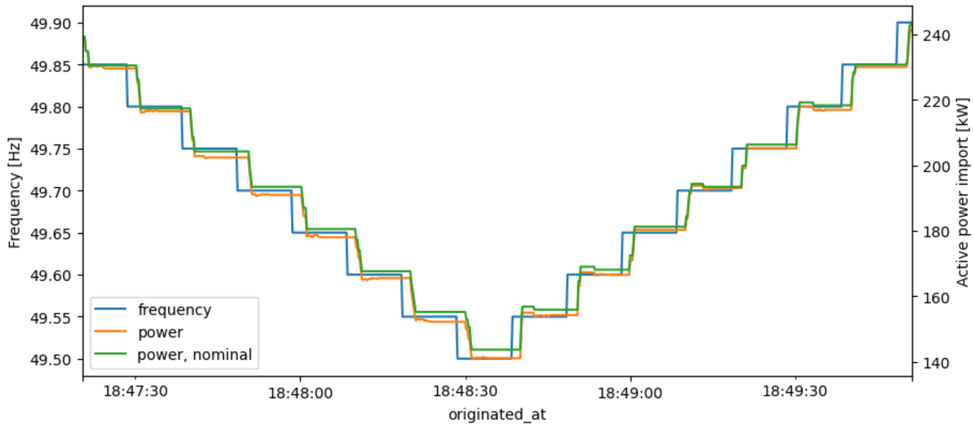
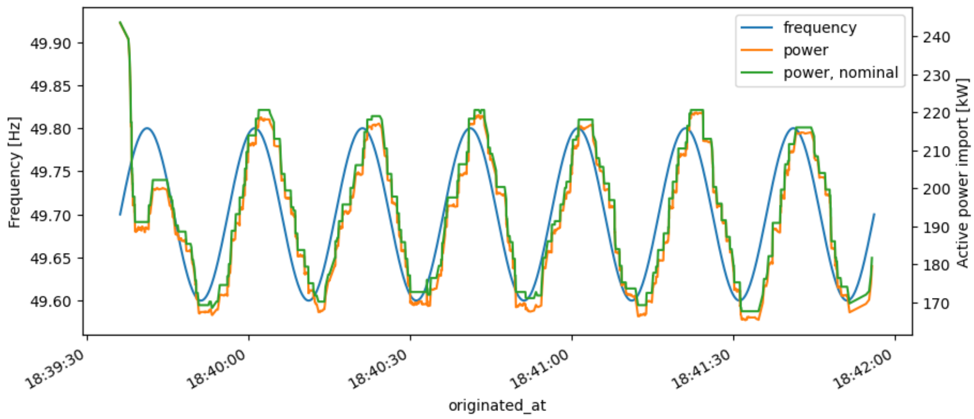


Figure 37: Simulated linearity test of a simulated charger pool with 67 units with baseline consumption of 240 kW and a capacity bid of 100 kW.



ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Figure 38: Tracking a simulated frequency signal generated by sine of period 20s, centered at 49.7 Hz, with an amplitude of 100 mHz. Pool of 67 simulated chargers with baseline consumption of 240 kW and a capacity bid of 100 kW.

Table 35: Future deliverables roadmap

Deliverable	TRL	Description
D6.2 (M18)	TRL 6	• Fairness and performance improvements. • Flexilibity forecasting. • More extended end-to-end tests.
D6.3 (M28)	TRL 7+	• Pilot demonstration • Optimal bidding strategies • Prequalification with SVK

3.9.5.2 Software details

The software is written in Python 3.9[12] and it takes the form of a Web Service, and is hosted as an EC2 instance on the Amazon Web Services Cloud[13]. It containerized using Docker[14], and then deployed to ECS via Terraform[15]. The code is hosted on a private GitHub repository and package management is achieved with poetry[16].

Table 36: List of Python libraries used.

Library	Description	URL	License
`stdlib`	Python standard library.	https://docs.python.org/3.9/library/	Python Software Foundation License (PSFL)
`boto3`	AWS SDK providing an object-oriented API as well as low-level access to AWS services.	https://aws.amazon.com/sdk-for-python/	Apache-2.0 License

[12] https://python.org
[13] https://aws.amazon.com/
[14] https://www.docker.com/
[15] https://www.terraform.io/
[16] https://python-poetry.org/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

`pydantic`	Data validation and settings management using Python type annotations.	https://docs.pydantic.dev/latest/	MIT License
`httpx`	Fully featured HTTP client for Python 3, which provides sync and async APIs, and support for both HTTP/1.1 and HTTP/2.	https://www.python-httpx.org/	BSD-3-Clause license
`Authlib`	The ultimate Python library in building OAuth and OpenID Connect servers.	https://authlib.org/	BSD-3-Clause License
`pyjwt`	JSON Web Token implementation in Python.	https://pyjwt.readthedocs.io	MIT License
`zipp`	A pathlib-compatible Zipfile object wrapper.	https://zipp.readthedocs.io/en/	MIT License
`flake8`	The wrapper which verifies pep8, pyflakes, and circular complexity.	https://flake8.pycqa.org/	MIT License
`mypy`	Static type checker.	https://www.mypy-lang.org/	MIT License
`black`	The uncompromising Python code formatter.	https://black.readthedocs.io/en/	MIT License
`typer`	Typer, build great CLIs. Easy to code. Based on Python type hints.	https://typer.tiangolo.com/	MIT License
`pytest`	Framework that makes it easy to write small tests, yet scales to support complex functional testing.	https://pytest.org/	MIT License
`moto`	A library that allows your tests to easily mock out AWS Services.	http://docs.getmoto.org/	Apache-2.0 License
`scipy`	Fundamental algorithms for scientific computing in Python.	https://scipy.org/	BSD-3-Clause License

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

numpy	The fundamental package for scientific computing with Python.	https://numpy.org/	BSD-3-Clause License
pandas	Flexible and powerful data analysis / manipulation library for Python, providing labeled data structures similar to R data.frame objects, statistical functions, and much more.	https://pandas.pydata.org/	BSD-3-Clause License
jupyterlab	Interactive development environment for notebooks, code, and data.	https://jupyter.org/	BSD-3-Clause License
matplotlib	A comprehensive library for creating static, animated, and interactive visualisations in Python.	https://matplotlib.org/	BSD-3-Clause License

3.9.6 Integration with ENERSHARE Data Space

As of this moment, and for ENERSHARE's Alpha release, the DR service's integration with the ENERSHARE Data Space has not yet been established. Instead, as it is developed for Hiven's behalf, it is integrated with Hiven's proprietary APIs instead.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

4 Data-driven Cross-Sector Services

4.1 Introducing cross-sector conception

This Chapter discusses on the design and implementation of services that integrate data and business processes stemming from the energy sector with other sector services such as water, transport, finance. The benefit of sharing data with other sectors is essentially exploited for the extraction of knowledge that can steer the under-coupling sectors. The services that are described in the following sections will be part of the Enershare App Store environment for their discoverability and accessibility from all potential involved users in the Enershare data space ecosystem.

The provided services could be summarized by the following list:

- **Multi-energy flexibility potential assessment:** leverages users' data to reinforce the security of energy supply for selected community, addressing the potential energy poverty by selecting the most economical heating source for the community,
- **Cross-operators' portal:** leveraging heterogeneous data sources (i.e., flexibility needs, flexibility availability, weather data) joint flexibility utilization for systems' secure operation,
- **Emissions and ecological footprint:** calculating emissions from building's energy consumption for heating along with the feature to see changes in emissions if one or more parts of building are renovated like roof, facade and/or windows,
- **EV charging monitoring & remote management:** cooperation mechanism between DSOs, CPOs and EV users to reduce the power grid upgrade magnitude,
- **ML-based models for assessing renovation actions in residential buildings:** for assessing renovation plans, assessing specific actions in building level as well as at assessing the potential of installing rooftop solar panels in residential buildings; this service includes -but are not limited to- building owners, financing institutions, investment bodies and policy specialists,
- **Health insurance alarms for senior living alone:** through energy consumption habits identifies regular usage patterns and triggers alarms when there are significant changes from these patterns, supporting the end-user's independence and quality of life,
- **Appliances maintenance or retrofit:** aims to improve the quality of living and energy consumption in households by detecting higher energy consumption of appliances early on and increasing energy efficiency by suggesting maintenance or renewal of appliances.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

This Chapter commences with the following sub-section that shares experience from the System Operators’ point of view on information sharing, particularly among energy vectors, whilst the next sub-section 4.3 that discusses on the data management processes that are essential to couple cross-sector data spaces; the remainder sub-section present the development of cross-sector services.

4.2 Experiences from cross-operators information sharing

In this section, some of the cross-sector (more cross-domain) data sharing schemes are presented, in particular from the transmission system operator’s (TSO) perspective. The cross-sector data sharing, in this section, hence refers to the data that is shared among different sectors or industries, which does not necessarily relate directly with energy data. Contrary to the typical energy data that TSO shares with its stakeholders – distribution system operators, market operators, regulatory agencies, and others – which include generation and consumption data, network topology, and other well-known energy/electricity data, the cross-sector data is used to either support the existing energy data or to provide the inputs for stakeholders to evaluate the TSO's performance and track its progress. There is, in fact, a vast variety of beyond-energy data that the TSOs are required to share. Among these, the main sharing schemes can be divided into two key parts:

- The data shared by TSO, which is not energy sector data;
- The data that TSO uses from other stakeholders/companies, which are not a part of the energy sector.

Along with the energy data, such as generation and load profiles, network topology, and operational state of the system (the open data that TSO is obliged to share due to local and EU-level regulations), TSO is obliged to disclose the information related to corporate sustainability in accordance with the EU Non-Financial Reporting Directive (Directive 2014/95/EU). This directive mandates large public interest entities, including TSOs, to provide information on their environmental impact, social and employee-related matters, respect for human rights, anti-corruption, and bribery issues. The directive aims to enhance transparency and comparability of non-financial information, enabling stakeholders to make informed decisions and encouraging companies to integrate sustainability considerations into their business strategies. Thus, TSOs in the EU zone are obliged to share the sustainability reports in accordance with the Environmental, Social, and corporate Governance (ESG) framework, which is designed to be embedded into the organization's strategic planning process, taking into account the interests and methods of creating value for all stakeholders involved, including employees, customers, suppliers, and financiers.

The TSOs’ Sustainability reports consist of three main pillars:

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- Environmental framework under which the TSO discloses the information about the environmental impact assessment processes and studies, the monitored birdlife, sound environment, flora, electromagnetic fields, the energy consumption of TSO's property for electricity, natural gas, other fuels, energy intensity, greenhouse gas (GHG) emissions, SF6 emissions, etc.;
- Social framework under which the TSO discloses the information about the characterisation of human resources, distribution of employees by age group and gender, human resource turnover, salary indicators, training and performance topics, and safety of the employees;
- Corporate Governance framework under which the TSO disclose information about the corporate structure of the company and economic value created and distributed within the company.

In addition to the data related to ESG framework, every TSO, as a public company, needs to share confidential data required by local regulations with state authorities, such as personal data related to employee healthcare, insurance of the TSO's property of all types(including buildings, power substations and other) and other non-energy market related data.

In addition to the non-energy data that TSO shares, it uses the non-energy data from other stakeholders/companies that improve the operational and planning procedures in energy-related activities. As an example, the Portuguese TSO receives real-time meteorological, sea, and atmospheric data from IPMA (Portuguese Institute of Sea and Atmosphere), such as:

- Lightning data in order to correlate this data with the TSO's database of the protection systems and certify the line outage due to lightning, in addition, this data is used for lightning prediction for maintenance procedures;
- Wind speed and direction data for the improvement of the wind forecast for the wind power plants;
- Solar radiation and ambient temperature data for the improvement of the solar forecast for the solar PV power plants;
- Wind speed and direction, solar radiation, and ambient temperature data for the implementation of the DLR – Dynamic Line Ratings for the Very High Voltage overhead lines;
- Satellites data related to the infrared map for fire detection in order to prevent possible line outages;
- Sea level data, waves data, and tides data for the offshore wind power plants (planned to be used in the future when offshore wind power plants will be put into operation).

On top of the aforementioned data exchange from/with external agents, the Portuguese TSO also collects and uses data internally that, although not directly related to energy, supports and serves energy activities' purposes. Namely, at the assets management level, several internal procedures are executed to provide support on automatic and semi-automatic predictive maintenance plans using off-grid data. These include collecting satellite data or LiDAR

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

inspection data (typically provided by contracted third-party services) covering a large extent of power transmission lines to characterize the respective Right-of-Ways. These data support the identification and classification of transmission tower models, detects degradation or missing components, and also identifies possible obstacles and encroachments surrounding these assets (vegetation indexes and other), to eventually trigger maintenance actions.

4.3 On data management processes to enable cross-sector services

Enabling cross-sector services often requires effective data management processes to facilitate the exchange and integration of data between different sectors. Here are some key data management processes that can support the development of cross-sector services which are related to the services being developed within ENERSHARE:

1. Data Governance: Data governance involves establishing policies, procedures, and frameworks to ensure the availability, integrity, quality, and security of data. It includes defining roles and responsibilities, data ownership, data classification, and access controls. Cross-sector services require a clear data governance framework to address issues such as data sharing, consent, privacy, and compliance across different sectors.
2. Data Standardization: Standardizing data formats, structures, and definitions is crucial for interoperability and seamless data exchange between sectors. Adopting common data standards, such as industry-specific data models or international standards like XML or JSON, helps ensure consistency and compatibility when sharing data across different systems.
3. Data Integration: Data integration involves combining data from multiple sources or sectors into a unified view. It requires establishing data pipelines, transformation processes, and data mapping techniques to harmonize and consolidate disparate datasets. Integration techniques may include Extract, Transform, Load (ETL) processes, data warehousing, or application programming interfaces (APIs) for real-time data exchange.
4. Data Sharing Agreements: To enable cross-sector services, data sharing agreements or partnerships need to be established between participating organizations or sectors. These agreements outline the terms, conditions, and protocols for data sharing, including data access, security, usage restrictions, and compliance with privacy regulations. Clear guidelines and legal frameworks ensure transparent and responsible data sharing practices.
5. Data Privacy and Security: Cross-sector services involve sharing sensitive data, and it is crucial to address privacy and security concerns. Data anonymization, encryption, and access controls should be implemented to protect personal or confidential information. Compliance with relevant privacy regulations, such as the General Data Protection

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Regulation (GDPR), is necessary to safeguard individuals' privacy rights and promote trust in data sharing.

6. Interoperability and Metadata Management: Metadata provides descriptive information about the data, including its origin, structure, relationships, and meaning. Establishing interoperability practices ensures that data can be discovered, understood, and effectively utilized across different sectors. Metadata standards, data dictionaries, and cataloguing techniques help in organizing and retrieving relevant data.
7. Data Analytics and Insights: Cross-sector services often involve deriving insights and intelligence from combined datasets. Data analytics techniques, such as data mining, machine learning, and predictive modelling, can be applied to identify patterns, trends, and correlations across sectors. Analytical tools and visualisation techniques enable stakeholders to extract actionable insights and make informed decisions.

By implementing these data management processes, organizations and sectors can create a solid foundation for enabling cross-sector services, enabling seamless integration, and unlocking the value of shared data for improved service delivery, innovation, and societal benefits.

That means that there should be a solid technological basis for the energy-related services which will be used in different sectors of economy to provide the insights on energy grid status, the usage of energy and predicting the future demand on energy. Such a basis is provided by the International Data Spaces. The International Data Spaces (IDS) initiative is a framework, specifically designed to facilitate secure and sovereign data sharing between organizations and sectors. It aims to enable data-driven cross-sector services by providing a trusted infrastructure for data exchange.

The IDS emphasizes *data sovereignty*, which means that data owners retain control over their data and have the authority to decide who can access and use it. This aspect is crucial for cross-sector services as it establishes a trusted environment for organizations to share data while maintaining control over their own data assets.

The IDS also focuses on *ensuring secure and trustworthy data sharing*. It provides mechanisms for authentication, authorization, and encryption, ensuring that data is shared only with authorized parties and remains protected from unauthorized access or tampering. This security framework instils confidence among participants and encourages data sharing across sectors.

The IDS framework incorporates *data governance* principles, providing organizations with the tools and mechanisms to manage data access, usage rights, and consent management. This enables organizations to define and enforce data usage policies, comply with privacy regulations, and respect individual rights. The ability to maintain control over data governance is essential when sharing data across sectors.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

These are crucial points to cover the *data management processes from 1 till 4*. For this reason, all the services implemented for energy sector and cross-sector usage are being integrated with the IDS components, such as IDS Connectors, IDS catalogue, etc. In the end all services are planned to run on IDS infrastructure and using the IDS data governance and secure data sharing mechanisms. Thus, every service describes the steps for integration with IDS.

However, some data management processes are needed to be handled on the level of the individual services as well. This concerns the following processes for energy-related service:

- Data Collection and Acquisition: This process involves capturing energy-related data from various sources, such as smart meters, sensors, energy management systems, weather databases, and other relevant sources. Data collection methods in the variety of the services under development include automated data retrieval, manual entry, and integration with third-party systems. The majority of the services, based on Machine Learning and Federated learning techniques extract data from the EnerShare use-case owners. The automated data retrieval requires from the services the technical side the clear agreements with the data owners, clear interfaces for data requesting and receiving, secure protocols for data retrieval, and clear input and output formats from both sides. The descriptions of the services and their input and output formats are provided per service in this document.
- Data Integration and Fusion: Practically all energy-related services require integrating data from multiple sources to gain comprehensive insights. Data integration involves combining data from different systems, formats, or organizations. Fusion techniques, such as merging datasets, data normalization, and data alignment, help create a unified view of energy-related data for analysis and decision-making. Every service in its functionality describes the techniques used for data fusion according to the goals of service itself.
- Data Analysis and Modelling: Analysing energy-related data is essential to derive meaningful insights, trends, and patterns. Data analysis processes in different services involve statistical analysis, machine learning, data mining, and predictive modelling techniques to identify energy consumption patterns, demand forecasts, energy efficiency opportunities, and anomaly detection. Advanced analytics help optimize energy usage and support informed decision-making. Every service describes the specific data analysis flow, needed to reach the goal of the service.

There are two points which require extra effort in handling them for any service which needs it: if the required data is sensitive: privacy-related or commercial secret; and the introduction of the best interoperability practices for making all internal energy services and cross-sector services interoperable with others. The following sections address these issues specifically.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

4.3.1 Data privacy preserving techniques

When the data used by the service comes from the private households, smart meters within it, or from low population area, such data needs special attention and requires to introduce extra techniques for guaranteeing that there is no violation of the consent from the citizens to use that data, and not being able to trace specific sensitive information about any citizen. Data could become sensitive if it is required for providing energy-related insights, but it represents the commercial secret of company. In this case also data is required to be, for example, anonymized or encrypted before it is used.

Some services, being developed in EnerShare, such as energy use predictions, transformer overloading predictions, load forecasting do use data from smart meters from the private households. Therefore, there is a need in the implementation of the privacy-enhancing techniques above the functionality of the service itself.

Privacy-enhancing techniques refer to a range of methods and technologies designed to protect the privacy of individuals in the digital realm. These techniques aim to minimize the collection, use, and disclosure of personal data, as well as prevent unauthorized access to sensitive information. Here are some common privacy-enhancing techniques which are useful for energy-related services:

- Anonymization: Anonymization involves removing or altering personally identifiable information (PII) from data sets to prevent the identification of individuals. This technique ensures that data cannot be traced back to specific individuals.
- Encryption: Encryption involves converting data into a coded form that can only be accessed with a decryption key. By encrypting data, even if it is intercepted or accessed without authorization, it remains unreadable and unusable.
- Differential Privacy: Differential privacy is a method that adds statistical noise or perturbation to data before releasing it to protect individual privacy. It aims to ensure that the presence or absence of a specific individual's data does not significantly affect the results or insights derived from the dataset.
- Privacy by Design: Privacy by Design is an approach that advocates embedding privacy considerations and protections into the design and architecture of systems, applications, and technologies from the very beginning. It involves proactive measures to ensure that privacy is a core principle and not an afterthought.
- Pseudonymization: Pseudonymization involves replacing direct identifiers (e.g., names) with pseudonyms or codes. The purpose is to make it difficult or impossible to directly identify individuals without additional information.
- Access Controls: Access controls involve implementing mechanisms to restrict access to personal data to authorized individuals or entities. This includes user authentication, role-

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

based access control, and other security measures to limit data access to those with a legitimate need.

The use of the specific Privacy-enhancing technique depends on the use-case and functionality of the service. The specific technique is defined together by the service developer and use-case owner. For example, in the service of loading predictions based on Federated Learning the measurements from the private households are going through the application of the differential privacy (adding specific noise and swapping some parts of information) before that data goes to the DSO which does not have a consent to use such data directly.

The services which need to deal with the privacy-enhancing technique, describe specific technique as a part of its functionality.

4.3.2 Semantic interoperability to achieve sector integration

To provide data-driven cross-sector services it is necessary to deal with several challenges from the point of view of semantic interoperability and data management, as data from different data sources and data providers need to be understood and combined together.

The first challenge is to be able to **understand** the data coming from different data sources or data sets. Syntactic interoperability is about using common data formats and data protocols so that the information can be read, whereas semantic interoperability is about understanding the information that has been read and being able to provide a unique meaning to the piece of data, ensuring consistency of the data across systems regardless of individual data format. To this end, WP3 will provide the Enershare ontology and Open APIs for energy services.

The second challenge is related to **data transformations**. When combining cross-sector information, it may be necessary to perform operations on the data values. Data values can be expressed in a variety of units of measurement, e.g., gas consumption is usually expressed in nm^3 (normal cubic meters), electricity consumption is usually expressed in kWh (kilo watts hour) and energy for mobility in l (liters of gasoil or diesel). Energy utility companies, such as Cepsa, Repsol or Total Energies, that sell different types of energy for the same client, may need to combine these values to offer cross-selling or discounts.

Besides, **data aggregation** or **data fusion** techniques may be required. Data aggregation is the process of gathering data and presenting it in a summarized format, whereas data fusion is the integration of data and knowledge from the same or several different sources.

Let's consider for example, time series data provided by counters. Depending on the device, the values can be provided in different temporal aggregation periods, e.g., every 5 minutes, every hour, on a daily basis, etc. This is the case of the energy consumption example mentioned above. In order to combine the information from two datasets with different aggregation

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

periods, it may be necessary, for example, to calculate average values in a period. Besides, counters not always provide the data in a time series format. In some cases, a timestamp is stored every time a sensor detects a measurement. This implies that in order to transform the information into an aggregated time series format, some calculations must be performed. In addition to temporal aggregations, there are use cases in which spatial aggregations are needed. In these situations, all values within the area of interest need to be added together.

Finally, data fusion techniques may be needed in analytical prediction services to integrate data from different sources such as household historical values, weather forecast data, calendar data (festivity or working day), etc.

The Data Mashup Editor, provided in WP3 (see deliverable D3.1), is a versatile tool that can be used for data harmonization and data transformation.

4.4 Multi-energy flexibility potential assessment

4.4.1 Description of the Service

The service is based on the existing electric flexibility assessment service. To expand the scope of the service, it has been extended to support multiple energy vectors, in our case the heating vector. In addition, when zones are created, the data-driven profiles can be used instead of statistically generated user profiles, allowing the service to be tailored to more realistic local communities. In any case, a mix of data-driven and statistical profiles is possible. The main advantage of such an approach is that community behaviour sometimes differs by geographic location or national level. In the case of statistical models, they should capture and adapt to all these specificities.

The service communicates with Energy Usage Prediction Service via fastAPI and creates various profiles based on data-driven models. These profiles are then used in conjunction with other models to determine the community's electricity and heating needs, as shown in Figure **39**. Together with the flexibility factor calculated from the daily electricity and heat power profiles, the planning and smart usage of different energies will be done. The main objective is to reinforce the security of the energy supply for the selected community and to address the potential energy poverty more efficiently by selecting the most economical heating source for the community.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

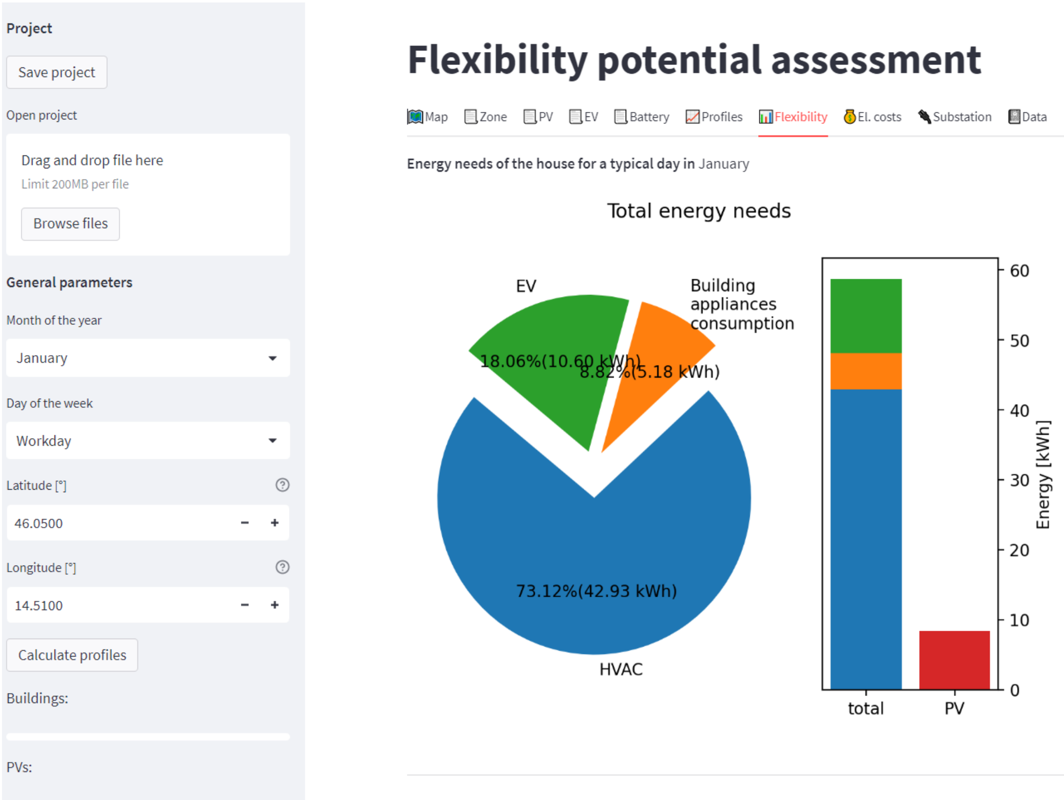


Figure 39: UI of service for calculating flexibility potential of zone

4.4.2 Innovation

The service is based on the service developed within the MATRYCS project. The added value of the enhanced service is that it supports data-driven models for user profiling, rather than just statistically-based models of the household. This improves the accuracy of the results by taking into account the specifics of the community obtained from the data-driven profiles. In addition, the service incorporates another energy vector (heat), as opposed to a flexibility assessment based only on electrical energy. The results will leverage the cross-sectoral domain so that additional cross-sectoral scenarios can be built on top of the multi-energy flexibility assessment, such as energy security and energy poverty.

4.4.3 Functions

The program flow of the service is shown in Figure 40. The main idea of the service is to create the energy profiles of households and commercial buildings. Based on these profiles, the different assets are considered and the potential flexibility is calculated according to these assets. For households, we use external libraries to generate statistical consumption profiles for different types of households in the zone. For this purpose, the artificial load profile generator (ALPG) was used. For the needs of the service, the ALPG library was modified and adapted to

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

our needs. The service was not only extended to work with the statistical models, but it can also import the calculated profiles from the service Energy Usage Prediction Service. This service uses the real historical data to create SM profiles, more is described in Section 3.6. In this case, flexibility is calculated from disaggregated power profiles. Also, in this case, the building does not need to be accurately described in terms of energy efficiency, so we do not need to collect all the building data related to the building envelope (e.g., wall area, heat transfer coefficient, etc.), the appliances in the house, or the PV system details that would otherwise be required.

In the case of creating zones, the PV system is decoupled from the house model, while the EV model is part of the household model but decoupled from commercial buildings. The main reason for this is that with the higher penetration of EVs and PVs in the system, the model can be easily adapted to the needs at the time of development. In addition, the location of the PVs and EVs does not affect the calculation and therefore can be simplified or even similar systems can be combined into one to speed up the calculations.

Once all the profiles are present in the zone, the flexibility is calculated according to the predefined limits, so that the user's comfort is affected as little as possible. In our case, we have preset these limits to some default values. In the case of comfort, the threshold value for the allowed temperature change is 1 °C, so the heating or cooling element must supply or extract the necessary energy in this time window. In the case of EVs, flexibility is defined only for the period when people are at home or after they return from work. In addition, the battery model is kept simple, with limited charging or discharging power and round-trip efficiency.

Output flexibility is used to address various challenges in the community, such as energy security and energy poverty.

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

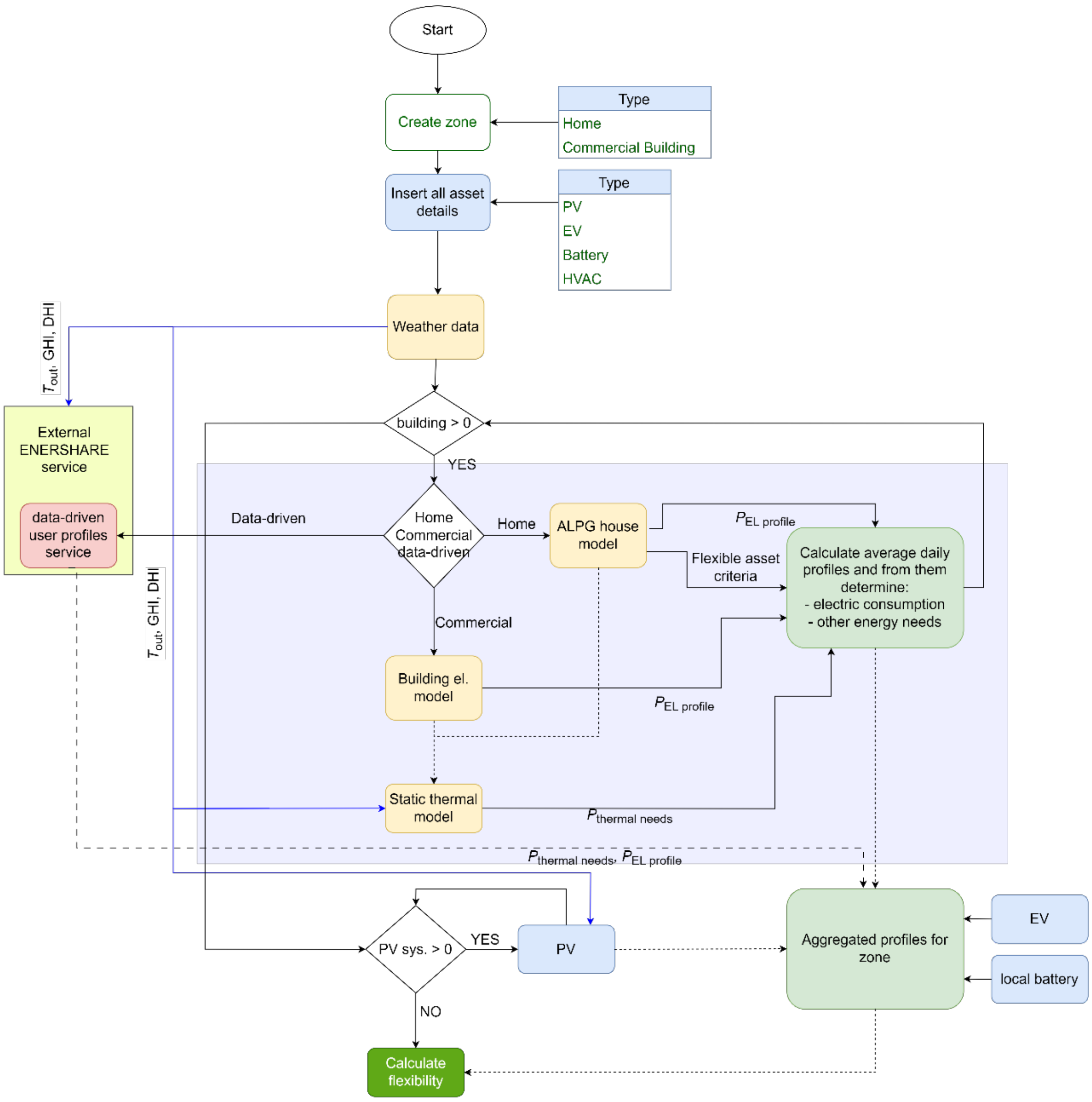


Figure 40: Program flow of flexibility assessment service

4.4.4 Input and Output Data Format

The input and data format are specified in Table 37.

Table 37: Flexibility assessment service data specifications

Variable	Source	Default Value	Unit	Type	Description
Dataframe related to building parameters in the zone					
type_building	Input	“Private house”	-	String/list	Options: “private house”, “commercial building”, “data driven model”
type_of_family	Input	“single worker”	-	String/list	Options: "Single worker", "Single jobless", "Single part-time", "Couple",

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

					"Dual worker", "Family dual parent", "Family dual worker", "Family single parent", "Dual retired", "Single retired"
cooling_type	Input	"Air conditioner"	-	String/list	Options: "Air conditioner", "Other"
cooling_el_P	Input	3000	W	Integer/list	Nominal electric power of a cooling device
heating_type	Input	"Heat pump"	-	String/list	Options: "Heat pump", "Electric heater", "Other"
heat_el_P	Input	2000	W	Integer/list	Nominal electric power of a cooling device
background_P	Input	1000	W	Integer/list	Estimated background consumption of the commercial building
peak_P	Input	5000	W	Integer/list	Estimated peak consumption of commercial building during the working hours
office_start	Input	9	h	Integer/list	Starting of office hours
office_end	Input	17	h	Integer/list	End of office hours
room_param	Input	"Basic"	-	String/list	Options: "Basic", "Advanced", in case of Basic only walls area, U value of façade and room temp. need to be provided, other parameters are estimated
walls	Input	500	m²	Integer/list	Area of a building envelope
U_walls	Input	0.2	W/m²K	Float/list	The thermal resistance of the walls/envelope of the house
T_set	Input	20	°C	Integer/list	Desired room temperature
floor_area	Input	500	m²	Integer/list	[Advanced] Area of floors in the building
volume_building	Input	400	m³	Integer/list	[Advanced]
ach_vent	Input	0.35	-	Float/list	[Advanced] Fraction of air mass exchanged through ventilation
vent_eff	Input	0.6	-	Float/list	[Advanced] Ventilation efficiency
thermal_cap	Input	165000	J/m²K	Integer/list	[Advanced] Thermal capacitance of the room
window_param	Input	"Basic"	-	String/list	Options: "Basic", and "Advanced", in the case of Basic, only window area and U value of windows need to be provided, other parameters are estimated.
window_area	Input	20	m²	Integer/list	Area of glazing on the building
U_window	Input	1.1	W/m²K	Float/list	Thermal resistance of the windows
south_window_area	Input	10	m²	Integer/list	Area of windows oriented to the south, thermal gain from insolation.
south_window_azimuth	Input	0	°	Integer/list	The azimuth, or orientation, is the angle of the windows relative to the direction due South. - 90° is East, 0° is South and 90° is West
windows_tilt	Input	90	°	Integer/list	Angle of the south windows from the horizontal plane
Dataframe related to PV systems in zone					
Pn	Input	5	kWp	float/list	Nominal power of PV system

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Inclination	Input	30	°	Integer/list	The angle of the PV modules from the horizontal plane
Azimuth	Input	0	°	Integer/list	The azimuth, or orientation, is the angle of the PV modules relative to the direction due South. - 90° is East, 0° is South and 90° is West.
EV_capacity	Input	40.0	kWh	float	The typical battery size of EVs at households
EV_power	Input	3.7	kW	float	Charging power at the household
Penetration_EV	Input	50	%	integer	Penetration of EVs into households
commute_distance_EV	Input	25	km	integer	Average one-way commute distance done by EV
number_of_cars	Input	1	-	integer	Total number of all EV vehicles in the commercial fleet, commercial buildings.
distance_EV	Input	20	km	integer	Average daily distance done by EV vehicle for commercial fleet
bat_capacity	Input	15.0	kWh	float	Battery capacity in the Zone
bat_power	Input	5.0	kW	float	Peak (dis)charging power
bat_efficiency	Input	90.0	%	float	The efficiency of battery charging/discharging.
latitude	Input	46.05	-	float	Avg. latitude of the zone
Longitude	Input	14.51	-	float	Avg. longitude of the zone
date_v	Input	2023-03-15	-	date	Date time of simulation
Datetime	output	-	Y-m-d H:M:S	datetime	Timestamp of the simulation
Total_Energy_needs_name	output	-	-	string/list	Total energy needs by asset
Total_Energy_needs	output	-	kWh	float/list	Energy needs by asset
Total_electric_Energy_needs_name	output	-	-	string/list	Total electric energy needs by asset
Total_electric_Energy_needs	output	-	kWh	float/list	Electric Energy needs by asset
Total_heat_Energy_needs_name	output	-	-	string/list	Total heat energy needs by asset
Total_heat_Energy_needs	output	-	kWh	float/list	Heat Energy needs by asset
Flexibility_by_asset	output	-	-	string/list	List of flexible assets
Flexibility_energy_by_asset	output	-	kWh	float/list	Flexible energy by asset

4.4.5 Implementation details

Deliverable (Month)	**TRL level**	**Comment**
D6.1 (M12)	TRL 4	First test of the tool done on the real data.
D6.2 (M19)	TRL 5	Interconnect with other tools (Service 3.6) and expanded to import heat profiles. A basic test

		using the DataSpace connector would be made, if available.
D6.3 (M28)	TRL 6	Fully integrated in the pilot. If available, a connection using the Energy DataSpace will be also tested.

The service was developed using Python v3 and its comprehensive set of standard libraries and Streamlit library. Artificial load profile generator (ALPG) open-source (GPL v3) library was integrated. To improve accessibility and integration with other tools, the service will be containerized and deployed on a publicly accessible server.

4.4.6 Integration with ENERSHARE Data Space

As of this moment, and for ENERSHARE's Alpha release, the tools integration with the ENERSHARE Data Space is not yet been established. IDS connectors will be used for the exchange of data between services, daily load/generation profiles.

4.5 Cross-operators' portal (COP)

4.5.1 Description of the Service

This service is developed particularly to bridge the gaps and challenges between different energy sector operators leveraging heterogeneous data sources. Therefore, considering the developed Flexibility Register service (presented in Chapter 3) that catalogues and reports the availability of connected flexibility in the heating and electricity sector, the module of cross-sector operators' portal (COP) allows the coordinated assessment and representation of common merit order lists among grid operators. Providing inputs from the Weather forecast services, the grid analysis services (developed in Pilot 3), the flexibility assessment service (developed in the context of Pilot). An open API (POST ReserveBid) will be available to allow the ingestion of latest available bids by an Aggregator.

The assessments of Flexibility Register will provide information on the current available upward and downward flexibility, while the Flexibility Assessment tool will provide information cross-energy flexibility.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

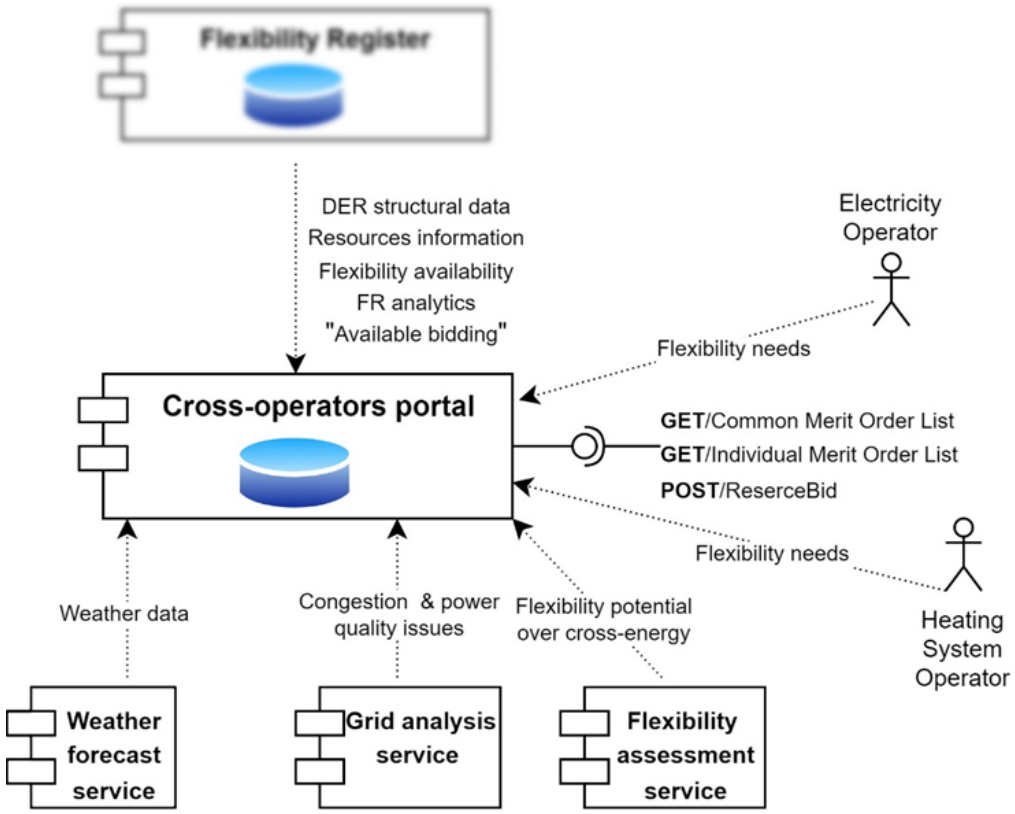


Figure 41: Logic view of cross-operators' portal

4.5.2 Innovation

The proposed Flexibility Register Services considers the implementations and general guidelines of IEGSA, developed on INTERRFACE project, advancing them to accommodate the needs for cross-operators flexibility sharing. Proper data management functions and will be developed to allow the ingestion of data incoming from the heating system operator. Furthermore, a joint merit order list will be presented to support the coordinated operation cross-sector operators enabling towards to share and reserve flexibility to address technical challenges (e.g., congestion management or power quality issues).

4.5.3 Functions

The main process flow of the Cross-operators' portal service, for DSO and the heating system operator, is:

- **User authentication:**

The Auth request requires the username and the password of the operators' user account. The current implementation for the Alpha version is based on jwt authentication and the

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

subsequent generated token per user; on the next stage for complete integration of this service with other functions, it will be proposed in the pilot level to adopt Keycloak as a horizontal authentication service (including the Flexibility Register and any other pilot services). The current structure of the request is as shown:

```
POST /api/users/auth HTTP/1.1

Host: {{application url}}:{{port}}

Content-Type: application/json

Content-Length: 42

Body

{

"username":"{{username}}",

"password":"{{password}}"

}
```

If the username and password are valid the response will contain a JWT token (that will be used to grand access to the next requests) and a user DTO entity containing the user's details.

Subsequent steps upon user's authentication are:

a. The CPO requests updated form from the FR database on DER structural data, Resources information, Flexibility availability, FR analytics, Available bid forms,
b. Weather forecast service: weather data updates are stored at the CPO database (to be used for Flexibility assessment service),
c. Grid analysis service, is a Pilot-specific developed service that provides inputs to CPO for expected congestion and power quality issues.
d. Flexibility assessment service, is a Pilot-specific developed service that provides essential information to the CPO on the projected thermal flexibility that can be shared with the electricity sector based on user/building characteristics (see Section 4.3),
e. Based on the above, input the CPO proposes individual Merit Order Lists (MOLs), one per grid operator, and the Common Merit Order List (CMOL) which assumes the shared flexibility among electricity with thermal grid.
f. Finally, through the CPO service the operators can Reserve bids and mark them for activation.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

4.5.4 Input and Output Data Format

The inputs were reported on the previous section from a) - d).

In regard to a) the specified all the data related to FR's service have been specified in Section 3.7, apart from the resource/resource group details introduced from the FR, where the input is the following:


```
[
{
 "id": 23,
        "createdOn": "2022-01-19T08:39:27Z",
        "createdBy": "FSP1",
        "uuid": "Y5KILYSPD",
        "name": "Test Resource",
        "description": "test",
        "status": "ACTIVE",
        "gridQualification": "QUALIFIED",
        "gridQualificationDescription": null,
        "resourceDTOs": [
            {
                "id": 3,
                "createdOn": "2021-12-22T11:20:16Z",
                "createdBy": "FSP1",
                "uuid": "L8V5JDCSB",
                "name": "Test Node",
                "description": null,
                "status": "ACTIVE",
                "gridQualification": "QUALIFIED",
                "gridQualificationTSO": "QUALIFIED",
                "gridQualificationTSODescription": "",
                "gridQualificationDSO": "QUALIFIED",
                "gridQualificationDSODescription": "",
                "resourceType": "LOAD",
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
            "voltageLevel": 0.4,
            "resourceOwnerEmail": null,
            "resourceOwnerTelephone": null,
            "authorizationToPublishData": true,
            "fspResourceAuthorization": null,
            "responsibleTSODSO": "DSO1",
            "meteringPointID": "5990011",
            "area": null,
            "streetAddress": null,
            "postalCode": null,
            "longitude": null,
            "latitude": null,
            "plantType": null,
            "country": "Slovenia",
            "maxUpRegulationPower": 5.0,
            "maxDownRegulationPower": 5.0,
            "actualUpRegulationPower": 5.0,
            "actualDownRegulationPower": 5.0,
            "nominalPower": 5.0,
            "nominalCapacity": 15.0,
            "rampingPeriod": 15.0,
            "fullActivationTime": 15.0,
            "deactivationPeriodLength": 15.0,
            "minimumDurationDelivery": 15.0,
            "maximumDurationDelivery": 15.0,
            "recoveryTime": null,
            "maxReboundAfterUpRegulation": 15.0,
            "maxReboundAfterDownRegulation": 15.0,
            "resourceGroupDTOs": null
        }
    ],
    "productDTOs": [
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
{
    "id": 48,
    "createdOn": "2022-01-19T08:39:27Z",
    "createdBy": "FSP1",
    "uuid": "FD7LINL6Z",
    "productQualification": "QUALIFIED",
    "productQualificationDescription": null,
    "resourceGroupDTO": {
        "id": 23,
        "createdOn": "2022-01-19T08:39:27Z",
        "createdBy": "FSP1",
        "uuid": "Y5KILYSPD",
        "name": "Test Resource",
        "description": "test",
        "status": "ACTIVE",
        "gridQualification": "QUALIFIED",
        "gridQualificationDescription": null,
        "resourceDTOs": [
            {
                "id": 3,
                "createdOn": "2021-12-22T11:20:16Z",
                "createdBy": "FSP1",
                "uuid": "L8V5JDCSB",
                "name": "Test Node",
                "description": null,
                "status": "ACTIVE",
                "gridQualification": "QUALIFIED",
                "gridQualificationTSO": "QUALIFIED",
                "gridQualificationTSODescription": "",
                "gridQualificationDSO": "QUALIFIED",
                "gridQualificationDSODescription": "",
                "resourceType": "LOAD",
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
                "voltageLevel": 0.4,
                "resourceOwnerEmail": null,
                "resourceOwnerTelephone": null,
                "authorizationToPublishData": true,
                "fspResourceAuthorization": null,
                "responsibleTSODSO": "DSO1",
                "meteringPointID": "5990011",
                "area": null,
                "streetAddress": null,
                "postalCode": null,
                "longitude": null,
                "latitude": null,
                "plantType": null,
                "country": "Slovenia",
                "maxUpRegulationPower": 5.0,
                "maxDownRegulationPower": 5.0,
                "actualUpRegulationPower": 5.0,
                "actualDownRegulationPower": 5.0,
                "nominalPower": 5.0,
                "nominalCapacity": 15.0,
                "rampingPeriod": 15.0,
                "fullActivationTime": 15.0,
                "deactivationPeriodLength": 15.0,
                "minimumDurationDelivery": 15.0,
                "maximumDurationDelivery": 15.0,
                "recoveryTime": null,
                "maxReboundAfterUpRegulation": 15.0,
                "maxReboundAfterDownRegulation": 15.0,
                "resourceGroupDTOs": null
            }
        ],
        "productDTOs": [
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
{
    "id": 48,
    "createdOn": "2022-01-19T08:39:27Z",
    "createdBy": "FSP1",
    "uuid": "FD7LINL6Z",
    "productQualification": "QUALIFIED",
    "productQualificationDescription": null,
    "resourceGroupDTO": null,
    "bidDTOs": null,
    "productDefinitionDTO": {
        "id": 2,
        "createdOn": "2021-12-17T13:57:22Z",
        "createdBy": "MO",
        "uuid": "NW9TF8OPX",
        "name": "mFRR - Slovenia",
        "type": "MFRR",
        "country": "Slovenia",
        "eligibleCountries": "",
        "minimumDeliveryPeriodDuration": 1.0,
        "maximumDeliveryPeriodDuration": 60.0,
        "rampingPeriodLength": 15.0,
        "fullActivationTime": 60.0,
        "minimumQuantity": 1.0,
        "maximumQuantity": 1000.0,
        "deactivationPeriod": 15.0,
        "allowProductRequalification": false,
        "productRequalificationThreshold": null,
        "numberOfQualifiedResourceGroups": null,
        "productDTOs": null,
        "userDTO": null
    }
},
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

}

]

For inputs b) to d), the specified input is upon discussion with Pilot 3 partners.

e) The expected output on individual MOLs and the CMOL is presented below:


```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<MeritOrderList_MarketDocument xmlns="urn:iec62325.351:tc57wg16:451-7:moldocument:7:3">
    <mRID>754824f6-78bf-43cf-8240-d01c0b9a28f0</mRID>
    <revisionNumber>1</revisionNumber>
    <type>A43</type>
    <process.processType>A19</process.processType>
    <sender_MarketParticipant.mRID codingScheme="A01">XXXEICCODE</sender_MarketParticipant.mRID>
    <sender_MarketParticipant.marketRole.type>A35</sender_MarketParticipant.marketRole.type>
    <receiver_MarketParticipant.mRID
codingScheme="A01">XYZCODE</receiver_MarketParticipant.mRID>
   <receiver_MarketParticipant.marketRole.type>A04</receiver_MarketParticipant.marketRole.type>
    <createdDateTime>2021-05-12T07:48:44Z</createdDateTime>
    <period.timeInterval>
        <start>2021-05-11T21:00Z</start>
        <end>2021-05-12T20:59Z</end>
    </period.timeInterval>
    <domain.mRID codingScheme="A01">XYZCODE.mRID>
    <TimeSeries>
        <marketAgreement.mRID>43X-T- -K_123_mFRR_2a</marketAgreement.mRID>
        <marketAgreement.createdDateTime>2021-05-12T05:17:07Z</marketAgreement.createdDateTime>
        <priority>1</priority>
        <resourceProvider_MarketParticipant.mRID
codingScheme="A01">XXX</resourceProvider_MarketParticipant.mRID>
        <registeredResource.mRID codingScheme="A01">169426169</registeredResource.mRID>
        <acquiring_Domain.mRID codingScheme="A01">10YLV-SSS</acquiring_Domain.mRID>
        <connecting_Domain.mRID codingScheme="A01">10YLV-SSS</connecting_Domain.mRID>
        <auction.mRID>MFRR_AUCTION</auction.mRID>
        <auction.paymentTerms>A01</auction.paymentTerms>
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
        <businessType>Z54</businessType>
        <bid_Period.timeInterval>
            <start>2021-05-12T07:00Z</start>
            <end>2021-05-12T09:00Z</end>
        </bid_Period.timeInterval>
        <quantity_Measurement_Unit.name>MWH</quantity_Measurement_Unit.name>
        <currency_Unit.name>EUR</currency_Unit.name>
        <price_Measurement_Unit.name>MWH</price_Measurement_Unit.name>
        <energyPrice_Measurement_Unit.name>MWH</energyPrice_Measurement_Unit.name>
        <direction>A01</direction>
        <minimumActivation_Quantity.quantity>1</minimumActivation_Quantity.quantity>
        <stepIncrement_Quantity.quantity>1</stepIncrement_Quantity.quantity>
        <marketObjectStatus.status>A06</marketObjectStatus.status>
        <Period>
            <timeInterval>
                <start>2021-05-12T07:00Z</start>
                <end>2021-05-12T09:00Z</end>
            </timeInterval>
            <resolution>PT1H</resolution>
            <Point>
                <position>1</position>
                <quantity.quantity>0.0</quantity.quantity>
                <price.amount>12.1</price.amount>
                <energy_Price.amount>12.1</energy_Price.amount>
            </Point>
    </TimeSeries>
    <Reason/>
</MeritOrderList_MarketDocument>
```

f) The reservation of specific flexibility capacities is performed with the following API method:

```
POST /CPO/a14-actual-regulation-power
```

```
<PlannedResourceSchedule_MarketDocumentxmlns="urn:iec62325.351:tc57wg16:451-7:plannedresourcescheduledocument:6:0">
```

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```xml
  <mRID>3715c5f3-557e-4384-9969-91b1006b000</mRID>

  <revisionNumber>1</revisionNumber>

  <type>A14</type> <!--A document providing the schedules for resource objects submitted by a resource provider-->

  <process.processType>A19</process.processType><!--Reguesting from Entso-e -->

  <sender_MarketParticipant.mRID codingScheme="A01">XXXXXX</sender_MarketParticipant.mRID>

  <sender_MarketParticipant.marketRole.type>A27</sender_MarketParticipant.marketRole.type>

  <receiver_MarketParticipant.mRID codingScheme="A01">XXXX</receiver_MarketParticipant.mRID><!--Platform EIC code-->

<receiver_MarketParticipant.marketRole.type>A35</receiver_MarketParticipant.marketRole.type><!--Platform role-->

  <createdDateTime>2021-12-21T06:00:00Z</createdDateTime>

  <schedule_Period.timeInterval>

    <start>2021-12-21T07:00Z</start>

    <end>2021-12-21T15:00Z</end>

  </schedule_Period.timeInterval>

  <domain.mRID codingScheme="A01">10YFI-1--------U</domain.mRID> <!--EIC code of region. Estonia in this case-->

  <subject_MarketParticipant.mRID codingScheme="A01">XXXXX</subject_MarketParticipant.mRID>

  <subject_MarketParticipant.marketRole.type>A27</subject_MarketParticipant.marketRole.type>

  <!--Zero or more repetitions:-->

  <PlannedResource_TimeSeries>

    <mRID>5698201254879633</mRID> <!--Unique identification of the timeseries-->

    <businessType>A61</businessType> <!--maximum available A61-->

    <flowDirection.direction>A01</flowDirection.direction><!--A01 = UP A02 = DOWN-->

    <product>8716867000146</product>

    <connecting_Domain.mRID codingScheme="A01">10YFI-1--------U</connecting_Domain.mRID> <!--Specific measurement point EIC code-->

    <registeredResource.mRID codingScheme="A01">0OL8IGU7L</registeredResource.mRID> <!--Resource identification. Poissibliy group EIC-->

<resourceProvider_MarketParticipant.mRID codingScheme="A01">42F4652A9381D904</resourceProvider_MarketParticipant.mRID><!--EIC code of region. Estonia in this case-->

    <acquiring_Domain.mRID codingScheme="NLU">10YFI-1--------U</acquiring_Domain.mRID>

    <marketAgreement.type>A01</marketAgreement.type><!--The type of the market agreement-->
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```xml
    <measurement_Unit.name>MAW</measurement_Unit.name>
    <objectAggregation>A01</objectAggregation> <!--Agregated per code-->
    <!--1 or more repetitions:-->
    <Series_Period>
      <timeInterval>
         <start>2021-12-21T07:00Z</start>
         <end>2021-12-21T15:00Z</end>
      </timeInterval>
      <resolution>PT1H</resolution>
      <!--1 or more repetitions:-->
      <Point>
        <position>1</position>
        <quantity>0.7</quantity>
        <!--Zero or more repetitions:-->
        <Reason>
          <code>A95</code>
          <!--Optional:-->
          <text>additional information</text>
        </Reason>
      </Point>
    </Series_Period>
  </PlannedResource_TimeSeries>
</PlannedResourceSchedule_MarketDocument>
```


4.5.5 Implementation details

The Alpha version of the Cross-operators' portal described in this document is consider a TRL5, updated version that will be described in D6.2 and D6.3 will be evolve on TRL6 and TRL7, accordingly, considering the interaction with demo partners in Pilot 3 (i.e., datasets specification, interaction with external services). The next version will consider the integration with data connector for the trusted and secure sharing of Cross-operators' portal in the Enershare data space ecosystem. The specific details on technological stack are the same as presented in Section 3.7 for the FR service.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

4.5.6 Integration with ENERSHARE Data Space

The proposed service will be exposed with a connector allowing the retrieval of analytics to the overall Enershare data space ecosystem with connector APIs assuming the secure and trusted information exchange.

4.6 Emissions and ecological footprint

4.6.1 Description of the Service

The transition to carbon neutrality, a more sustainable energy economy and increased use of renewable energy technologies brings forth numerous challenges. In the context of energy use for heating and cooling buildings, the cost of investment in new technologies and energy prices are crucial. However, the emissions of greenhouse gases as well as harmful gases and particulates are equally important. These emissions not only have a significant impact on climate change but also affect human health. Achieving a balance between these factors is a complex task that requires an integrated approach, considering both environmental and economic perspectives.

This service, developed by ENVI, provides a calculation of nine emission values for: CH4, CO, CO2, CXHY, N2O, NOX, PM10, PM2.5 and SO2 along with the ecological footprint. This estimation is based on the useful energy of a building in conjunction with the type of fuel it uses for heating. If the useful energy of a building is unknown the emissions are estimated with the assistance of a deep neural network (DNN) trained on historical real data. This DNN model considers various building attributes such as the heated area, age of the building, type of building and more, as well as monthly average temperatures to estimate the useful energy needed for heating the building.

4.6.2 Innovation

Calculating emissions from energy consumption together with fuel used is not new however estimation of emissions from building's energy consumption for heating using machine learning is what is this service offering as innovation. Service can also be used to see changes in emissions if one or more parts of building are renovated like roof, facade and/or windows.

4.6.3 Functions

The service has been implemented as an API service entirely coded in R programming language and encapsulated in a docker environment. Trained deep neural network is exported from Keras and rewritten as an R code so there is no need for Python or any other dependencies.

Service consists of two relevant functions:

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- **Emissions calculator** – function that takes list of buildings with useful energies and fuels defined and returns list of associated emissions or list of buildings' attributes and returns list of associated estimated emissions along with standard deviation for confidence interval calculation.
- **Fuel change emissions matrix calculator** – function that takes single building's useful energy and returns matrix of emissions for different fuels or single building's attributes and returns matrix of estimated emissions for different fuels along with standard deviation for confidence interval calculation.

4.6.4 Input and Output Data Format

All input and output data exchanged are in JSON format.

4.6.4.1 Emissions calculator

Table 38: Input – useful energy of a building

Variable	Type	Unit	Default value	Options/range	Description
useful_energy	number	kWh/year	-	$\mathbb{R}^+$	Annual useful energy consumption of the building in kWh.
fuel	string	-	-	EE | HHO | HP | HPAW | HPWW | LPG | NG | WOOD	Fuel type used for heating system. EE – electrical energy, HHO – Extra-light heating oil, HP – Heat Pump unknown type, HPAW – HP air to water, HPWW – HP water to water, LPG - Liquefied Petroleum Gas, NG – Natural Gas, WOOD – all types of wood biomass
install_year	integer	-	-	$\mathbb{Z}^+$	Installation year of a heating system

JSON example:


```
[
  {
    "useful_energy":                                    22645,
    "fuel":                                             "NG",
    "install_year":                                     2015
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
  },
  {
    "useful_energy":                                         11674,
    "fuel":                                                 "WOOD",
    "install_year":                                           2001
  }
]
```

Table 39: Input – building attributes

Variable	Type	Unit	Default value	Options/ range	Description
heat_area	number	m²	-	ℝ⁺	Heating area.
building_type	string	-	-	EDUCATION | HOTEL, INDUSTRY | MULTI_RESIDENTIAL | OFFICE | SHOP | SINGLE_FAMILY_HOUSE	Building type.
building_construction	string	-	-	BRICK | COMBO_MATERIALS | CONCRETE | METAL_CONSTRUCTION | PREFAB | STONE | WOOD | UNKNOWN	Building construction material.
build_year	integer	-	-	ℤ⁺	Year of construction of the building
roof_restoration_year	integer	-	build_year	[build_year,∞)	Roof restoration year
facade_restoration_year	integer	-	build_year	build_year,∞)	Facade restoration year
windows_restoration_year	integer	-	build_year	[build_year,∞)	Windows restoration year

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

fuel	string	-	-	EE | HHO | HP | HPAW | HPWW | LPG | NG | WOOD	Fuel type used for heating system. EE – electrical energy, HHO – Extra-light heating oil, HP – Heat Pump unknown type, HPAW – HP air to water, HPWW – HP water to water, LPG - Liquefied Petroleum Gas, NG – Natural Gas, WOOD – all types of wood biomass
install_year	integer	-	build_year	[build_year,∞)	Installation year of a heating system
air_temp	array{number}	°C	-	length(array) = 12	Monthly average air temperature. First element is for January and the last is for December.
air_temp_min	array{number}	°C	-	length(array) = 12	Monthly minimum air temperature. First element is for January and the last is for December.
air_temp_max	array{number}	°C	-	length(array) = 12	Monthly maximum air temperature. First element is for January

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

and the last is for December.

JSON example


```
[
  {
    "heat_area": 142.4,
    "building_type": "SINGLE_FAMILY_HOUSE",
    "building_construction": "CONCRETE",
    "build_year": 1994,
    "roof_restoration_year": 2005,
    "facade_restoration_year": 2007,
    "fuel": "NG",
    "install_year": 2015,
    "air_temp":
[1.1,0.9,5.7,10,15,18.5,22.3,20.7,14.8,9.5,5.4,1.1],
    "air_temp_min": [-11.5,-9.9,-2.6,-
1.4,5.9,7.7,11.7,10,5.4,1,-3.9,-5.2],
    "air_temp_max":
[10,9.8,15.3,22.6,25.4,28.7,33.4,32.2,29.8,19.5,20.2,12.5]
  },
  {
    "heat_area": 103.3,
    "building_type": "SINGLE_FAMILY_HOUSE",
    "building_construction": "BRICK",
    "build_year": 1948,
    "roof_restoration_year": 2006,
    "windows_restoration_year": 1996,
    "fuel": "HHO",
    "install_year": 1995,
    "air_temp":
[0.8,0.5,5.2,9.4,14.4,17.9,21.7,20.2,14.3,9.1,5.2,0.8],
    "air_temp_min": [-11.2,-9.4,-2.7,-
1.7,5.3,7.3,11.2,9.9,5.3,0.4,-5.1,-5.6],
    "air_temp_max":
[9.3,9.1,14.4,21.8,24.8,28,32.6,31.6,29,19,20.2,11.3]
  }
]
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Table 40: Output

Variable	Type	Unit	Options/ range	Description
fuel	string	-	BIOMASS | EE | HHO | LPG | NG	Fuel type.
CH4	number	kg	$\mathbb{R}^+$	Methane emissions
CO	number	kg	$\mathbb{R}^+$	Carbon monoxide emissions
CO2	number	kg	[0,∞)	Carbon dioxide emissions
CXHY	number	kg	$\mathbb{R}^+$	Hydrocarbons emissions
N2O	number	kg	$\mathbb{R}^+$	Dinitrogenoxide emissions
NOX	number	kg	$\mathbb{R}^+$	Nitrogen oxides emissions
PM10	number	kg	$\mathbb{R}^+$	Particulate Matter of 10 microns diameter or less
PM2.5	number	kg	[0,∞)	Particulate Matter of 2.5 microns diameter or less
SO2	number	kg	$\mathbb{R}^+$	Sulfur dioxide emissions
EF	number	global hectares	[0,∞)	Ecological Footprint
std*	number	-	[0,∞)	Multiplicative standard deviation. For example: 95% confidence interval is calculated: *[CO2 / std^Qn(0.975), CO2 * std^Qn(0.975)]* where Qn is quantile function of a normal distribution. In example case Q(0.975) ≈ 1.96. Given formula in this example is equivalent to:*exp(ln(CO2) ± Qn(0.975)*ln(std))*, where *exp* is exponential function and *ln* is natural logarithm

[*] - only available if building's useful energy is estimated using DNN

JSON example

```
[
  {
    "fuel":                                          "NG",
    "CH4":                                         0.0849,
```

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
    "CO":                                                   3.1355,
    "CO2":                                               4953.2107,
    "CXHY":                                                 0.5375,
    "N2O":                                                  0.0085,
    "NOX":                                                  2.6876,
    "PM10":                                                 0.1075,
    "PM2.5":                                                0.1075,
    "SO2":                                                       0,
    "EF":                                                   1.6692,
    "std":                                                    1.12
  },
  {
    "fuel":                                                 "WOOD",
    "CH4":                                                  0.4218,
    "CO":                                                 148.3274,
    "CO2":                                                       0,
    "CXHY":                                                 5.2533,
    "N2O":                                                  0.2109,
    "NOX":                                                  5.2533,
    "PM10":                                                46.9703,
    "PM2.5":                                               45.7343,
    "SO2":                                                  0.6798,
    "EF":                                                        0,
    "std":                                                  1.1305
  }
]
```


4.6.4.2 Fuel change emissions matrix calculator

Table 41: Input – useful energy of a building

Variable	Type	Unit	Defaultvalue	Options/range	Description
useful_energy	number	kWh/year	-	$\mathbb{R}^+$	Annual useful energy consumption of the building in kWh.
year	integer	-	Current year	$\mathbb{Z}^+$	Simulated installation year of a heating system

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

JSON example

```
[
  {
    "useful_energy":                                          16342,
    "year":                                                    2023
  }
]
```

4.6.4.2.1 Input – building attributes

Input is the same as input for Emissions calculator (building attributes) with additional variable:

Variable	Type	Unit	Defaultvalue	Options/range	Description
year	integer	-	Current year		Simulated installation year of a heating system

JSON example

```
[
  {
    "heat_area": 142.4,
    "building_type": "SINGLE_FAMILY_HOUSE",
    "building_construction": "CONCRETE",
    "build_year": 1994,
    "roof_restoration_year": 2005,
    "facade_restoration_year": 2007,
    "fuel": "NG",
    "install_year": 2015,
    "air_temp":
[1.1,0.9,5.7,10,15,18.5,22.3,20.7,14.8,9.5,5.4,1.1],
    "air_temp_min": [-11.5,-9.9,-2.6,-
1.4,5.9,7.7,11.7,10,5.4,1,-3.9,-5.2],
    "air_temp_max":
[10,9.8,15.3,22.6,25.4,28.7,33.4,32.2,29.8,19.5,20.2,12.5],
    "year": 2023
  }
]
```

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

4.6.4.2.2 Output

Output structure is the same as output for Emissions calculator where we always get emissions for all fuels available in the system.

4.6.5 Implementation details

Technology readiness level:

Deliverable (Month)	TRL level	Comment
D6.1 (M12)	TRL 4	Service validated and tested in Envirodual's internal server environment.
D6.2 (M19)	TRL 5	Interconnect with other tools/services if possible. A basic test using the DataSpace connector will be made if available.
D6.3 (M28)	TRL 6	Fully integrated and demonstrated in the pilot 3.

The service is developed using programming language R v4 with its standard packages and Tidyverse collection of R packages. For building and training DNN the Keras and TensorFlow are being used. Service can be deployed as standalone R package or web service encapsulated inside a docker image.

4.6.6 Integration with ENERSHARE Data Space

Service is intended to anyone who wishes to enrich their results with emissions values and is easily integrated in any process. Service will adopt a data space compliant IDS connector that will serve as interface. With each service equipped with a data space connector, service to service interoperable data exchange is enabled. Under the ENERSHARE Data Space this service and its API can be distributed as an application in the App Store. In this way different agents can predict emissions with their own private data.

4.7 EV charging monitoring and remote management

4.7.1 Description of the Service

Transition to electric mobility was initiated by creating new opportunities and new obstacles: by increasing the number of electric vehicles, the amount of electricity that must be supplied

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

increases and, therefore, a necessary strengthening of the power lines follows. However, through a cooperation mechanism between DSOs (Distribution System Operators), CPOs (Charging Point Operators) and EV users, it is possible to reduce the power grid upgrade magnitude by coordinating the electric vehicles charging. DSO monitors the electricity grid and, thanks to accurate forecasting systems, is able to identify how, when and where to charge electric vehicles to avoid congestion problems.

CPO, thanks to EV charging monitoring & remote management service, will thus be able to provide a dynamic charging session based on real-time/forecasted DSO needs, bringing a benefit to DSO who will have a balanced grid and, meanwhile, offering an advantageous charging price in charging stations located in congested areas and attract a greater number of EV users increasing CPO revenue. DSO, CPO and EV data will be hosted in ENERSHARE Cross-Sector Data Space.

4.7.2 Innovation

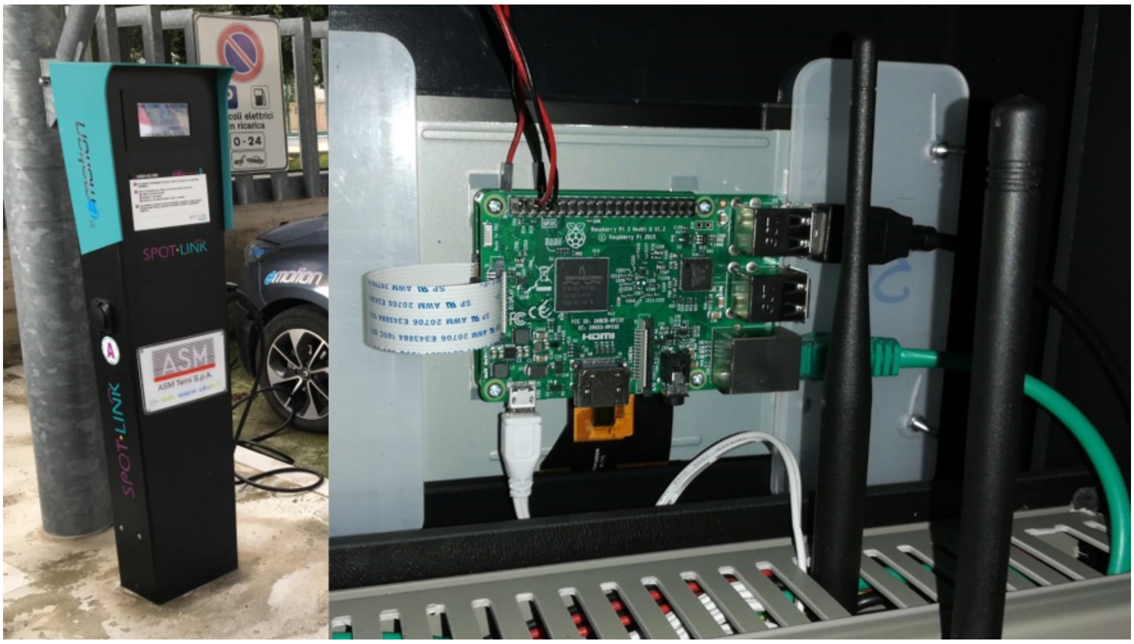


Figure 42: EMOT charging station and its single-board computer

EV charging monitoring & remote management service was recently established by EMOT for Electricity Grid VS Electric Mobility Cross-Sector operation and it is based on IoT charging stations developed by EMOT. Since energy meters usually only report 15-minute average data, we installed a single-board computer inside EMOT charging stations to enable real-time data collection with one second sample rate. These fine-grained data are used to train a ML model and obtain an accurate prediction for a highly efficient charging session management to balance a grid in a condition of high penetration from variable renewable energy plants bringing benefits

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

to DSO, CPO and EV user. According with real-time DSO needs, based on data collected from distributed smart meter, charging sessions can be started/stopped remotely and power output can be modulated remotely using EMOT APIs.

4.7.3 Functions

EV charging monitoring & remote management service functions are:

- Real-time monitoring.
- Remote charging session start & stop.
- Remote charging station power output modulation.

Figure 43 shows real-time data collected each second by EMOT charging stations broadcasted live in a MQ Telemetry Transport (MQTT) broker for data sharing with ENERSHARE Cross-Sector Data Space.


```
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:28'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:28'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:28'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:29'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:29'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:29'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:30'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:30'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:30'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:31'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:31'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:31'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:32'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:32'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:32'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:33'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:33'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:33'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:34'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:34'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:34'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:35'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:35'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:35'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:36'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:36'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:36'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:37'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:37'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:37'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:38'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:38'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:38'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:39'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:39'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:39'}]}
{'towers': [({'tower_id': 18, 'plug1': 30, 'power1': 0, 'plug2': 187, 'power2': 0, 'timestamp': '2023-05-24 11:02:40'},), ({'tower_id': 24, 'plug1': 37, 'power1': 0, 'plug2': 38, 'power2': 12.0, 'timestamp': '2023-05-24 11:02:40'},), {'tower_id': 25, 'plug1': 39, 'power1': 0, 'plug2': 40, 'power2': 0, 'timestamp': '2023-05-24 11:02:40'}]}
```


Figure 43: EMOT charging stations real-time data published on MQTT broker

Figure 44-Figure 46 shows EMOT remote charging station management APIs implemented for ENERSHARE demonstration activities.

```
fildev@backdev
1 ❯ wget --post-data="tower_id=18&plug_number=2&user_auth=***" https://emotion-projects.eu/api/actions/start-charge/
--2023-03-29 11:32:39--  https://emotion-projects.eu/api/actions/start-charge/
Loaded CA certificate '/etc/ssl/certs/ca-certificates.crt'
Resolving emotion-projects.eu (emotion-projects.eu)... 82.165.124.120
Connecting to emotion-projects.eu (emotion-projects.eu)|82.165.124.120|:443... connected.
HTTP request sent, awaiting response... 204 No Content
2023-03-29 11:32:39 (0.00 B/s) - 'index.html' saved [0]
```

Figure 44: EMOT remote charging session start API

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
fildev@backdev
❯ wget --post-data="tower_id=18&transaction_id=65c0ad74-a91d-4367-ad86-f29f9ca4aeba" https://emotion-projects.eu/api/actions/stop-charge/
--2023-03-29 11:35:20--  https://emotion-projects.eu/api/actions/stop-charge/
Loaded CA certificate '/etc/ssl/certs/ca-certificates.crt'
Resolving emotion-projects.eu (emotion-projects.eu)... 82.165.124.120
Connecting to emotion-projects.eu (emotion-projects.eu)|82.165.124.120|:443... connected.
HTTP request sent, awaiting response... 204 No Content
2023-03-29 11:35:20 (0.00 B/s) - 'index.html' saved [0]
```

Figure 45: EMOT remote charging session stop API

```
fildev@backdev
❯ wget --post-data="tower_id=18&hw_type=dc&power=200" https://emotion-projects.eu/api/actions/set-power/
--2023-03-29 11:59:21--  https://emotion-projects.eu/api/actions/set-power/
Loaded CA certificate '/etc/ssl/certs/ca-certificates.crt'
Resolving emotion-projects.eu (emotion-projects.eu)... 82.165.124.120
Connecting to emotion-projects.eu (emotion-projects.eu)|82.165.124.120|:443... connected.
HTTP request sent, awaiting response... 204 No Content
2023-03-29 11:59:21 (0.00 B/s) - 'index.html' saved [0]
```

Figure 46: EMOT remote charging station power output modulation API

4.7.4 Input and Output Data Format

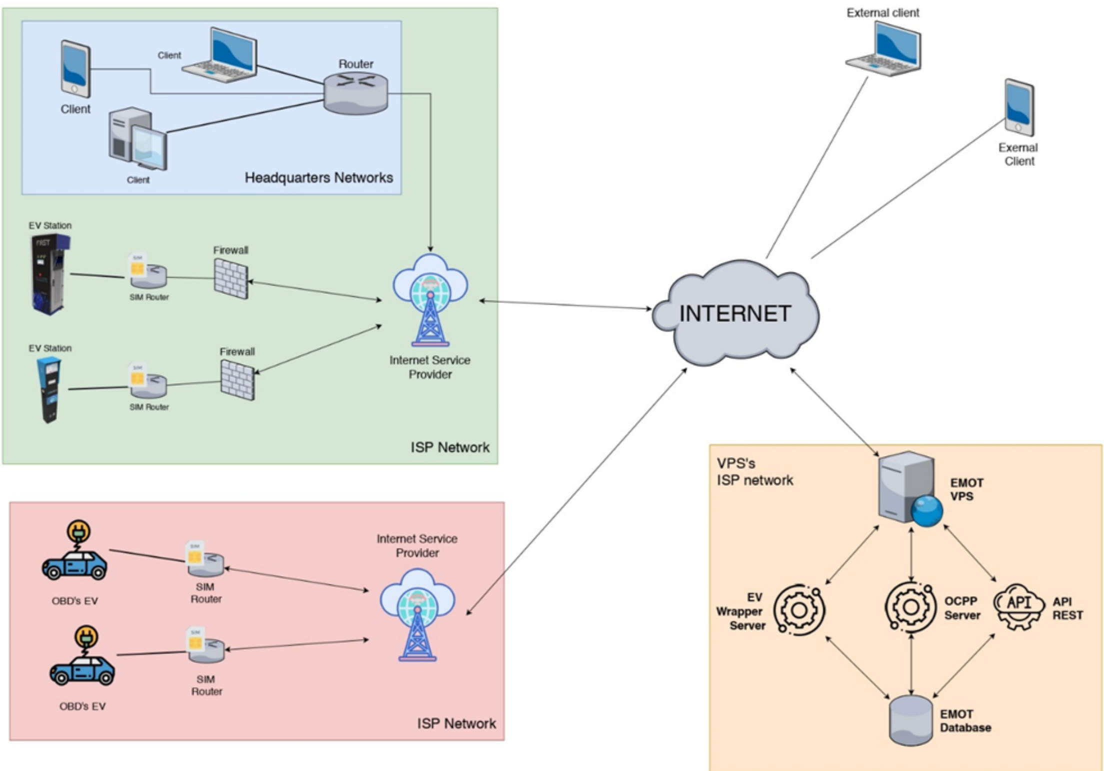


Figure 47: EMOT network topology

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

EMOT charging stations exchange data through a modem connected to a single-board computer; charging station protocols are OCPP, the application protocol for communication between charging stations and EMOT central management system, and WebSocket, the computer communications protocol, providing full-duplex communication channels over a single TCP connection. EMOT OCPP server accepts communications and data exchange only with the client program that is installed in the charging station computer; OCPP server accepts the connection by the client only and exclusively if a valid authentication key is used at the time of the request. Charging station input/output data format is JSON and the sampling rate is one second.

4.7.5 Implementation details

Deliverable (Month)	TRL level	Comment
D6.1 (M12)	TRL 5	Technology validated in EMOT electric mobility environment.
D6.2 (M19)	TRL 6	IDS connectors development and integration with other services related to grid congestion management for an intermediate Terni pilot technology demonstration
D6.3 (M28)	TRL 7	Fully integrated and demonstrated in Terni pilot

The service is developed via Django web framework (v4.1.7) using Python (v3.10) as programming language and OCPP library (v2.0.1).

4.7.6 Integration with ENERSHARE Data Space

Exchange of data and commands between ENERSHARE Data Space and EMOT system will be granted by IDS connectors, to be integrated with the rest of ENERSHARE Common IDS. Data from charging stations and electric vehicles will be available in EMOT electric mobility platform data lake, containing both historical and real-time data. EMOT data will be wrapped and provided in JSON format according to the common ENERSHARE Data Models. These created JSON files will be sent as an IDS artifact and an orchestrator will be developed to call the different service components that will be dockerized.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

4.8 ML-based models for assessing renovation actions in residential buildings

4.8.1 Description of the Service

The aim of the service developed by NTUA is to provide a solid methodological framework for assessing renovation actions in residential buildings. This service consists of two Machine Learning models for implementing two different tasks related to the domain of building retrofitting and energy autonomy in the residential scale. The first model is tailored for assessing specific actions in building level, while the second model aims at assessing the potential of installing rooftop solar panels in residential buildings. The end users of this service include -but are not limited to- building owners, financing institutions, investment bodies and policy specialists.

Regarding the list of functions of this service, it includes two different prediction tasks which are tackled with Machine Learning models and the use of novel techniques such as meta learning and data augmentation, aiming to address the problem of data scarcity and to improve the accuracy of the models. In this respect, the functions of this service are described as follows:

- A supervised Machine Learning classification model for addressing the problem of predicting the optimal retrofitting actions in order to improve the energy class of the building.
- A supervised Machine Learning regression model for predicting the annual potential of rooftop solar panel in terms of energy production and the level of energy autonomy of the residential building after the installation.

More details on both models are provided in Section 4.8.3. The internal architectural schema of both functions which are incorporated in this service are provided in the following Figure 48 for the classification framework for proposing retrofitting actions, and in Figure 49 for the regression framework for predicting the potential of rooftop solar panels and the energy autonomy of the building.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

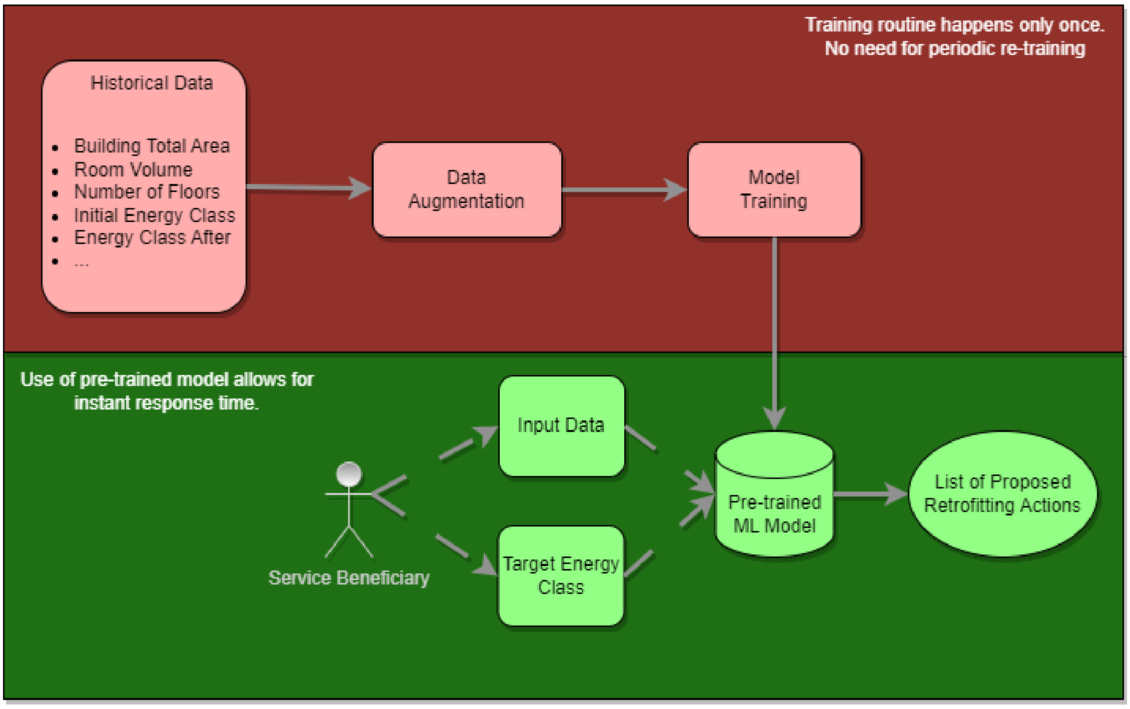


Figure 48: Classification framework for assessing retrofitting actions.

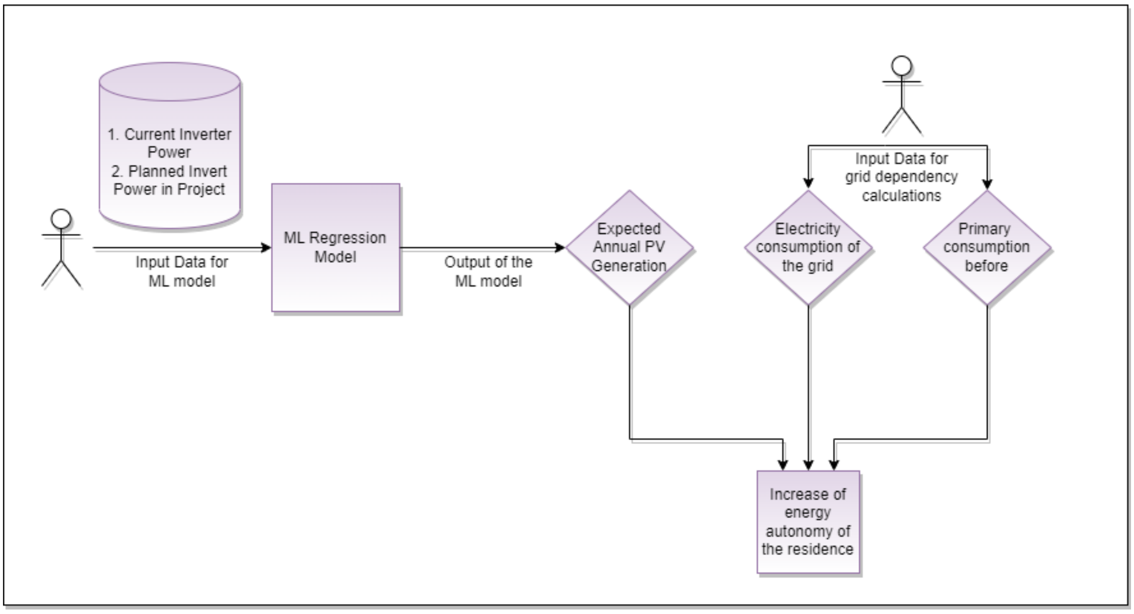


Figure 49: Framework for predicting the potential of rooftop photovoltaic in residential buildings.

As seen in Figure 48 and Figure 49, the architecture of both functions is based on pre-trained models and thus does not require periodic retraining due to the static nature and the slow

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

update rate of input data. This, in turn, allows for instant response time upon request of prediction for a new instance.

Regarding data management, in this initial stage of testing the ML models, data is fed to the models in a static fashion through .CSV files. Data management is conducted with respect to FAIR principles: (F) The naming convention has been agreed internally among the organizations developing (NTUA) and using (LEIF) the service, (A) The available data are used internally in NTUA shared drive, password protected with internal user credential, (I) the data format is in Comma-separated values (CSV) files and (R) regarding re-use, the data will only be available for access by partners in the consortium associated with the development of the service. In the final form of the service, the intention is to store data in a relevant database from which it will be fed when re-training is required. This database will not be used, though, during the operational phase of the service, as only the pre-trained model will be required.

4.8.2 Innovation

The innovation of both components lies to the AI-based, data-driven manner used for predicting the types of renovations required, as well as the power generation potential. As of our knowledge, most of today's approaches use exhaustive methodologies which require in-place monitoring of a building in order to propose the most well-fitting retrofitting actions in order to improve the energy class (e.g., energy audits).

In this framework, these services take advantage of the numerous historical data collected from similar retrofitting actions and renovation projects and attempt to substitute these traditional methods by novel data-driven approaches, which are based on the similarity between different retrofitting activities in similar buildings. Thus, the models proposed within this service use key outcomes of past projects, such as BD4NRG's meta-learning classification model for supporting decisions on energy efficiency investments. Thus, the novelty in this service is that it directly proposes which measures must be taken for a single residential house and not just predict the energy class after the renovation.

Also, another novelty of the service lies in the use of data augmentation techniques to deal with the limited number of available data. Data augmentation techniques play a crucial role in addressing the challenge of limited available data within the context of the proposed service. By augmenting the existing dataset, these techniques effectively increase the diversity and quantity of training samples, which in turn improves the accuracy and robustness of the AI models.

4.8.3 Functions

Model for Assessing Retrofitting Actions: The functions of this service include the data preparation function, the data augmentation function (which is not yet implemented in this first development cycle), the model fitting function and the prediction function which will serve as the final output of the service.

- *Data Preparation:* This function involves collecting and preprocessing historical data related to retrofitting actions and renovation projects. The data is cleaned, organized, and prepared for training the Machine Learning model.
- *Data Augmentation:* To overcome the challenge of limited available data, data augmentation techniques are applied in this function. The existing dataset is expanded by generating synthetic data points through transformations and variations. This augmented data enhances the model's robustness and improves its ability to handle diverse building characteristics, leading to more accurate predictions.
- *Model Fitting:* In this function, the Machine Learning classification model is trained using the prepared dataset. The model learns from the historical data and extracts patterns, correlations, and similarities between different retrofitting activities and their impact on improving the energy class of residential buildings. The model is based on tree-based classifiers, namely the Random Forest implementation of SkLearn, the XGBoost and the LightGBM. Following, ensemble of these three models produces the final prediction.
- *Prediction:* Once the model is trained and augmented, it can be used to predict the optimal retrofitting actions for a given residential building. The model takes into account the specific characteristics of the building and provides recommendations on the actions that would result in the most effective improvement of its energy class. There are four possible actions that the model will advise for or against each (the output of the recommendation can also be given in terms of possibilities for each class of renovation actions):
    - Carrying out construction works
    - Reconstruction of engineering systems
    - Water heating system installation
    - Heat installation

Model for Predicting the Generation of Rooftop Solar Panels: The functions of this service include the data preparation function, the model fitting function and the prediction function which will serve as the final output of the service. Data augmentation is not needed for this model as there is enough historical data.

- *Data Collection:* Historical data related to rooftop solar panel installations and their associated energy production is collected for training the Machine Learning model. Factors which are assessed include the electricity consumption from the grid before the project, the inverter set power in the project and the amount of electricity produced by the solar panels among others.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- *Model Training*: In this function, the Machine Learning regression model is trained using the collected data. The model learns the patterns and relationships between the various factors to accurately estimate the annual energy production of rooftop solar panels and the level of energy autonomy achieved by the residential building, by comparing with the status of the residential building before the implementation of the action. Similarly, to the previous model, tree-based classifiers will be used, namely the Random Forest implementation of SkLearn, the XGBoost and the LightGBM, and they will be combined using an ensemble of these three models with the use of meta-learning.
- *Prediction:* Once the model is trained and augmented, it can be used to make predictions on the potential energy production and energy autonomy of a residential building after installing rooftop solar panels. The model takes into account the specific characteristics of the building and provides estimates of the generated power and the increase in energy autonomy levels.

4.8.4 Input and Output Data Format

The input data used for developing the renovation planning model include data from private houses (one- and two-apartment houses) that have increased their energy efficiency with public funding in Latvia (Latvia). In the following table, we give a detailed description of the potential input data used in this model. It must be noted that in the second development cycle the input features will be updated and finalized according by conducting feature importance and explainability tools such as Shap.

Table 42: The input features of the for Assessing Retrofitting Actions.

Feature ID	Name	Description	Unit	Data Type
1	Building Total Area	Total area of the building. It is determined by summing up the entire room, including the basement floor, plinth floor, technical floor, attic floor	m^2	real
2	Reference area	Reference floor area according to standard LVS EN ISO 52000-1:2020 9.4.3. point	m^2	real
3	Above-ground floors	Number of above-ground floors.	1: one floor 2: two floors 3: three floors	integer

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

4	Underground floor	Information on whether the house has an underground floor.	0: there is no underground floor 1: there is underground floor	boolean
5	Energy consumption before	Total energy consumption before the project.	kWh/m^2	real
6	Carrying out construction works	Carrying out construction works in the enclosing structures during the project (to increase the energy efficiency of the house)	0: no construction work 1: construction works have been carried out	boolean
7	Reconstruction of engineering systems	Reconstruction of engineering systems (ventilation, recuperation) to increase the energy efficiency of the house (during the project).	0: engineering systems have not been restored 1: restoration of engineering systems has been carried out	boolean
8	Heat installation	Installation of heat installations to ensure the production of heat from renewable energy sources	0: no installation 1: installation present	boolean
9	Water heating system	Installation of a new water heating system (during the project).	0: no installation 1: installation present	boolean
10	Initial energy class	The initial energy efficiency class of the house ranges from A+ to F class. Each house has one class, which is determined in accordance with the regulations of the Cabinet of Ministers of April 8, 2021, No. 222 "Methods for calculating the energy efficiency of buildings and rules for energy certification of buildings" (https://likumi.lv/ta/id/322436)	A+ to F	text
11	Energy class after	Energy efficiency class according to energy audit (after the renovation of the	A+ to F	text

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

building). Energy efficiency class of the house ranges from A+ to F class.

The output of the model will be the proposed category of renovation actions in the building in order to achieve the transition from the initial energy class (e.g., D) to the targeted energy class after the project (e.g., B). These probabilities will be in the range of [0,1], accordingly. The output data will be provided by an API in tabular format for all the buildings inserted to the model.

Regarding the model for Predicting the Generation of Rooftop Solar Panels the input of the model is provided in the following table.

Table 43: The input features for the Model for Predicting the Generation of Rooftop Solar Panels.

Feature ID	Name	Description	Unit	Data Type
1	Electricity consumption of the grid	Electricity consumption from the grid before the project.	MWh per year	real
2	Primary energy consumption before	Primary energy consumption before the installation of the solar panel system during the project.	kW	real
3	Current inverter set power	Current inverter set power - inverter power that was already installed before the project.	kW	real
4	Inverter power in project	Inverter set power in project - in addition to the existing inverter.	kW	real
5	Electricity produced by solar panels	The amount of electricity produced by the solar panels, which are installed in the project.	MWh per year	real
6	Primary energy consumption after	Primary energy consumption after installing the solar panel system.	MWh per year	real
7	Reduction of primary energy	The reduction of primary energy after the installation in annual basis.	MWh per year	real
8	CO_2 emissions reduction	The reduction of CO_2 emissions in annual basis.	t CO_2 eq/year	real

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

9	Granted support	Granted support after installation of solar panels.	EUR	integer

The output of this model will also be provided via an API, and will be twofold: First, the predicted annual electricity produced by the solar panels will be estimated and given as output. Secondly, the increase of energy autonomy and the dependence of the grid after the installation will be calculated.

4.8.5 Implementation details

Deliverable (Month)	**TRL level**	**Comment**
D6.1 (M12)	TRL 4	Design of the methodological framework and algorithms for Assessing Retrofitting Actions and for Predicting the Generation of Rooftop Solar Panels and test of the models with sample data provided from Pilot 7.
D6.2 (M19)	TRL 6	Design of the full-stack technical solution of the service which will include: ➢ Database Design ➢ Interactive User Interface Implementation ➢ Creation of the Data Sharing Infrastructure
D6.3 (M28)	TRL 7	Full integration and demonstration in the LEIF Pilot (Pilot 7) and interconnection with Data Spaces.

The following programming tools and software will be used:

- Programming Languages: Python v3.0
- ML Libraries: SkLearn, XGBoost, LightGBM (if another algorithm is added, it will be reported in D6.2)
- Tool for Model Training: Jupyter Notebook
- Containerization: Docker
- Web Application Framework: Django

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

4.8.6 Integration with ENERSHARE Data Space and Services

The renovation service mainly relies on the two quasi-static input sources as described in Table 42and Table 43 for each subservice. The final outputs of the sub-services will be as abovementioned: a) the proposed category of renovation actions in the building, b) benefits by the installation of PV panels.

A summary of the combination of inputs/outputs will be available for integration and connection with the ENERSHARE Data Space, and for more user-friendly interactions via a graphical dashboard or interface. Such a summary will be extremely useful for stakeholders (outside the pilot) that are interested in implementing similar renovations (end-users) or similar service (energy agencies) at a pan-European level. This is expected to be an iterative process that will incorporate the suggestions and validation of all the stakeholders, specifically focusing on the pilot requirements and needs.

4.9 Health insurance alarms for senior living alone

4.9.1 Description of the Service

The aim of this serviced developed by SEL is to monitor energy consumption in households with seniors living alone. It is assumed that seniors have daily routines that can be determined through their energy consumption habits.

The main purpose of this service is to identify these regular usage patterns and trigger alarms when there are significant changes from these patterns. By analysing the energy consumption, the service can support the end-user's independence and improve their quality of life. Additionally, it provides reassurance to their family and/or caregivers about the senior's safety.

Initially, consumption data is analysed, and a pattern for each person will be defined. This service considers deviations beyond a certain degree to be abnormal, so the alarms will be triggered when energy consumption is not within these bounds. After the pattern is established, the service will continuously analyse the consumption and compare it to the regular behaviour. The service will then run autonomously.

4.9.2 Innovation

The innovation of this use case lies in the data-driven approach to evaluate the need for assistance that seniors living alone may have while maintaining their independence. The service takes advantage of consumption disaggregation algorithms to understand exactly which appliances are being used and when they are being used. It will also be able to identify the regular deviations in behaviour, creating a normal deviation range. After enough training,

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

the model will provide an individualised pattern of usage and live data will be given to it to compare to this standard behaviour. When variations happen, alarms are triggered to family or caregiving facilities to inform them.

4.9.3 Functions

The service implementation is according to the following functions:

- Data Preparation: This function involves collecting and preprocessing all historical data from the mains meters of the senior citizens. The data is then cleaned, organized and prepared for training the Machine Learning algorithm.
- Machine Learning Algorithm: In this function, the Random Forest algorithm for pattern recognition is trained using the prepared datasets creating a model. The model then learns from the historical data and extracts patterns from the datasets, detecting the normal consumptions at any given time of day, from all the explorer cases.
- Alarmistic definition: A normal deviation is then included in the recognized patterns allowing the definition of a normal margin of electricity consumption. The model has the function of deploying an alarm every time the consumptions do not follow the pattern (including deviation) for over an hour.
- Identification: Once the model is trained, it is able to recognize the pattern of consumption from the main meter data. Secondly the model will be able to send an alarm message to the citizen's caretaker identifying the specific deviation and the time of occurrence.

4.9.4 Input and Output Data Format

The input data consists of energy consumption data from seniors living alone. This input data is considered personal and cannot be shared.

Feature ID	Name	Description	Unit	Data Type
1	Electricity consumption	Electricity consumption from the meter	kWh	real
2	Temperature	Temperature sensor installed in each house	ºC	real
3	Humidity	Humidity sensor installed in each house	%	real

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

The output of the model will be alarms. The output data can be made public after being anonymised.

Feature ID	Name	Description	Unit	Data Type
1	Alarm	Notification system for caretakers	Report	sms
2	Energy consumption pattern	Consumption pattern extracted from the mains meter visualized in the model dashboard	kWh	real
3	Energy consumption	Energy consumptions curve visualized in the model dashboard	kWh	real
4	Temperature	Temperature measurements visualized in the model dashboard	ºC	real
5	Humidity	Humidity measurements visualized in the model dashboard	%	real

4.9.5 Implementation details

Deliverable (Month)	TRL level	Comment
D6.1 (M12)	TRL 4	Model developed, sensors installed and data collection period started. Library datasets created.
D6.2 (M19)	TRL 5-6	First results of pattern recognition and alarmistic model delivered. Sensitive case analysis carried out with stakeholders to fine-tune the model and integrate the service. Service offtakers involved in the definition of requirements for the integration of the service and components.
D6.3 (M28)	TRL 6-7	Fully integrated and demonstrated in the pilot 2. Service implemented in the

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

		Dataspace service library. Service offtakers testing the integration in their own environment.

4.9.6 Integration with ENERSHARE Data Space

As of this moment, the integration of the service with ENERSHARE's Data Space has not yet been established. The proposed service will be exposed to the Data Space with connector APIs allowing private and secure data exchange.

4.10 Appliances maintenance or retrofit

4.10.1 Description of the Service

This service, developed by SEL, aims to improve the quality of living and energy consumption in households by detecting higher energy consumption of appliances early on and increasing energy efficiency by suggesting maintenance or renewal of appliances.

The service proposes using energy consumption data to suggest appliance maintenance or renewal. The information will be shared with consumers and housing providers, and the implementation will involve several actors, including consumers, housing providers, market information aggregators, energy service companies, maintenance and appliance retailers, and appliance producers.

The use case assumes the availability of an mains meter and specific appliance metering plugs as a prerequisite, and the technical details include a list of actors, such as the EMS of relevant households, EMS Connector, Market Information Aggregator Connector, Maintenance and appliance retailer's connector, Appliance producer connector, and Housing provider connector. This is part of a digital service offering for preventive maintenance for home appliances with value creation pathways for different actors.

4.10.2 Innovation

The innovation of this services lies on the disaggregation algorithm capacity starting from a learning methodology based on gathered data of main electricity consumption and specific consumption data from the appliance plugs. To this day there is now similar service in place that can provide predictive maintenance for household appliances based on a non-intrusive methodology.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

In this framework, this service takes advantage of the numerous historical data collected from SEL's Living Energy explorers households and leverages on previous SEL's experience on developing intelligence on non-intrusive load monitoring

Regarding the service clients it is also innovative approach because the service can be addressed to 3 different client types: homeowner, appliance retailer/maintenance provider and electricity utilities.

4.10.3 Functions

- Data Preparation: This function involves collecting and preprocessing all historical data from the mains meter and from the specific appliance plugs from the defined explorers. The data is then organized and prepared for training the Machine Learning model.
- Machine Learning Algorithm: In this function, the Machine Learning classification model is trained using the prepared datasets. The model then learns from the historical data and extracts patterns, correlations, and similarities between the 2 different datasets (mains meter and appliance plug) from all the explorer cases. The model is based on Logistic Regression classifiers, best suited for supervised classification, which allows the recognition of the specific electric signatures of each appliance. A second interaction, to be defined, is the ability to detect irregularities among those signatures, which can lead to the identification of failures or malfunctions on the appliances.
- Identification: Once the model is trained, it can be able to firstly recognize the specific signature profile of each individual appliance from the main meter data. Secondly the algorithm will be able to identify major malfunctions or maintenance needs that the appliance may require by determining deviations of the normal functioning signature part for that type of appliance.

4.10.4 Input and Output Data Format

The input data consists of energy consumption data from Living Energy explorers and their appliances. This input data is considered personal and cannot be shared.

Feature ID	Name	Description	Unit	Data Type
1	Electricity consumption	Electricity consumption from the main meter.	kWh	real

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

2	Electricity consumption	Electricity consumption from the chosen appliances.	kWh	real
3	Board power	Installed capacity of the main board.	kVA	real
4	Temperature	Temperature sensor installed in each house	ºC	real
5	Humidity	Humidity sensor installed in each house	%	real

The output of the model will the specific appliance electricity consumption signatures and the result of the malfunctioning analysis. The output data can be made public after being anonymised.

Feature ID	Name	Description	Unit	Data Type
1	Energy consumption signature	Electricity consumption signature from the chosen appliances driven out of the main meter data	kWh	real
2	Malfuntioning notification	Notification system for the client regarding the malfunctioning analysis	report	sms
3	Main electricity consumption	Electricity consumption of the main meter. This data is available on each explorer dashboard with high degree of detail and visualisation customization.	kWh	real
4	Appliance consumption	Electricity consumption of each individual appliance. This data is available on each explorer dashboard with high degree of	kWh	real

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

		detail and visualisation customization.		
5	Temperature	Temperature measurements visualized in the model dashboard	ºC	real
6	Humidity	Humidity measurements visualized in the model dashboard	%	real

4.10.5 Implementation Details

Technology readiness level:

Deliverable (Month)	TRL level	Comment
D6.1 (M12)	TRL 4	Model developed, extra sensors installed and data collection period started. Library datasets created.
D6.2 (M19)	TRL 5-6	Validation phase with supervised classification and forced failure tests. Sensitive case analysis carried out with stakeholders to fine-tune the model and integrate the service. Service offtakers involved in the definition of requirements for the integration of the service and components.
D6.3 (M28)	TRL 6-7	Fully integrated and demonstrated in the pilot 2. Service implemented in the Dataspace service library. Service offtakers testing the integration in their own environment.

4.10.6 Integration with ENERSHARE Data Space

As of this moment, the integration of the service with ENERSHARE's Data Space has not yet been established. The proposed service will be exposed to the Data Space with connector APIs allowing private and secure data exchange.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

5 Data Visualisation

This chapter aims to describe the Data Visualisation Layer, which is developed in the context of Task 6.4 – “Toolkit for interactive data visualisation services”. For the realisation of this Task, Apache Superset has been selected to serve as the basic visualisation dashboard and reporting tool. Thus, in the following subsections, main focus will be put on presenting an overview Apache Superset functionality, as well as the envisioned way of integrating the tool with the various datasets deriving from the project’s activities and other implementation details related to the deployment approach and the overall solution that will be offered.

5.1 Description of the Data Visualisation Layer

The Data Visualisation Layer objective is to provide all the needed functionalities that will allow the optimal presentation, exploitation and understanding of the data produced and collected within the various ENERSHARE activities. In essence, this layer provides an advanced visualisation dashboard, with a user-friendly and intuitive interface, through which the end users are able to explore, query and analyse the aforementioned datasets. These functionalities allow a deeper understanding of these datasets and, ultimately, more useful and meaningful conclusions and data storytelling.

For the realisation of the above-mentioned objectives, the main technology that was selected is Apache Superset[17], which is an open-source data exploration and visualisation platform, allowing the users to create visualisations and execute queries without writing any code, thus empowering them to analyse and present data in a user-friendly and interactive manner. Apache Superset is designed to be user-friendly; it is fast and lightweight, highly customisable and scalable, while it is also continually evolving, due to a vibrant community of developers and users.

Its wide range of functionalities includes a variety of visualisation capabilities, from simple pie charts to highly detailed geospatial charts (as reported also in the tools’ official website), as well as quick, easy and well-documented integration with several databases such as Structured Query Language (SQL) based data sources through SQLAlchemy[18], including also no SQL ones.

[17] https://sup erset.apache.org/

[18] https://www.sqlalchemy.org/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

5.2 Capabilities

In this subsection, we intend to display an overview of the capabilities and the different functionalities offered by Apache Superset. We will explain in detail the input required by the end users and we will also include screenshots and explanations of the steps followed for each procedure. For the following explanations and the demonstration of the tool's capabilities, we will use a dataset provided by the project's pilot 7. In fact, the dataset consists of several columns which represent various information about energy consumption, renewable energy production, CO2 Emissions reduction, and other valuable indicators, about some renewable energy projects in a number of regions in Latvia.

As also mentioned in the previous section, Apache Superset offers a vast variety of visualisation charts alternatives, which vary from plain and usual kinds of charts, like pie charts and bar charts, up to more complex ones, like heatmaps or map visualisations. The end users can form queries over the stored data via an interactive and easy-to-use interface which allows them to choose between the possible values of a number of parameters, as well as the kind of visualisation which will be produced. Figure 50 displays this interface.

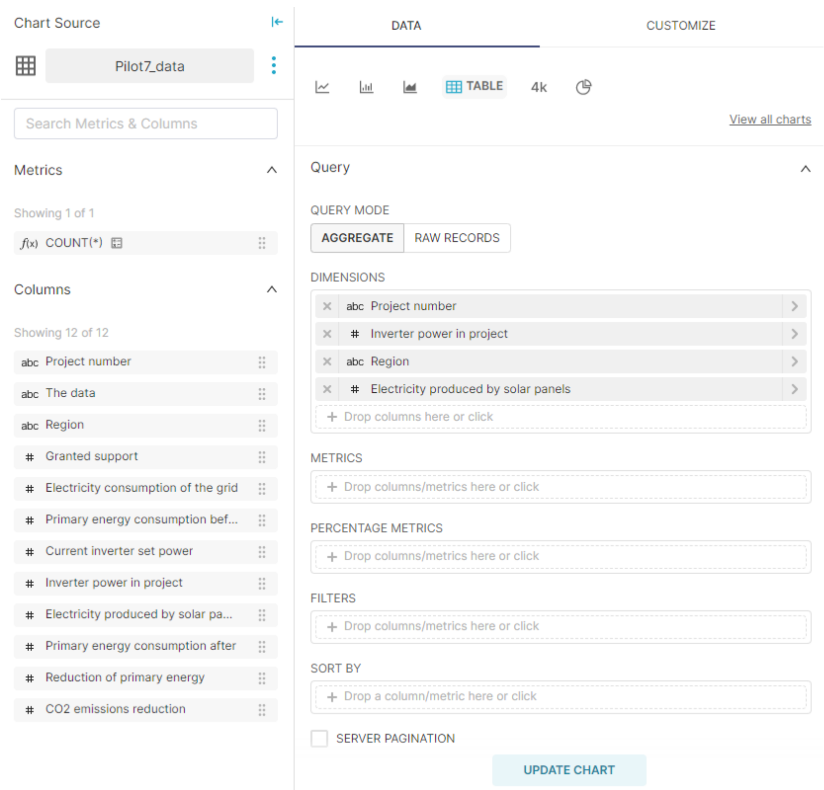


Figure 50: Apache Superset interface for building a query

The outcome of the above displayed query is displayed in Figure 51.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Project number	Inverter power in project	Region	Electricity produced by solar panels
PME2-942	10	Kurzeme	10
PME2-99	11	Rīga	11
PME2-794	6	Rīga	5.49
PME2-76	10	Rīga	6.81
PME2-1820	10	Rīga	10.72
PME2-583	8	Rīga	8.03
PME2-1458	10	Rīga	10.8
PME2-406	5.55	Kurzeme	4.68
PME2-1643	8.66	Rīga	6.6
PME2-1548	10	Rīga	11.22
PME2-502	5	Kurzeme	5.25
PME2-1195	10	Rīga	8.3
PME2-661	4	Vidzeme	3.85
PME2-917	4	Rīga	3.91
PME2-1032	8	Vidzeme	8.25
PME2-861	10	Rīga	10.42
PME2-497	8.4	Rīga	7.58
PME2-950	10.92	Rīga	10
PME2-617	10	Rīga	9
PME2-1283	11.1	Rīga	10.8
PME2-1598	8	Rīga	5.16
PME2-1919	3.68	Zemgale	4.05

Figure 51: Apache Superset table representation of the requested data

As depicted in the figures above, the end users can fully parameterise their queries, by editing the fields that handle the columns that will be shown, as well as adding any desired filters or applying any restrictions related to a variety of parameters, such as the limit of the rows to be displayed. Moreover, the table that is displayed offers some interaction alternatives, such as toggling the order (ascending or descending) of the columns that contain numerical values.

Apart from the graphical interface presented above, Apache Superset also offers the capability of performing queries by writing plain SQL code. This option is available via the SQL Lab tab that is placed on the top bar of the tool. An example query is illustrated in Figure 52.

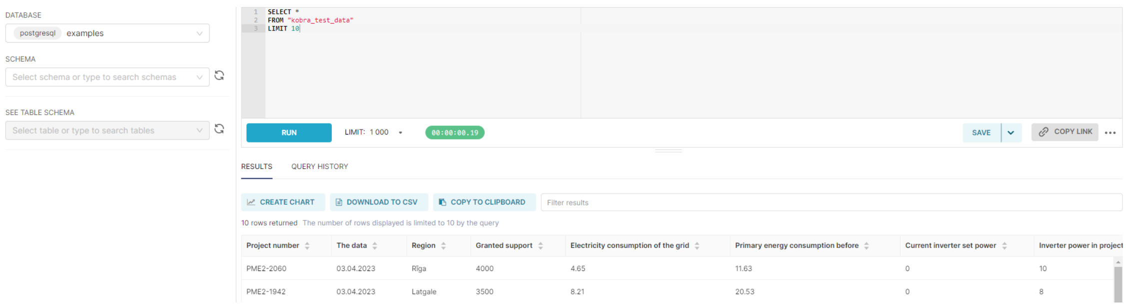


Figure 52: Apache Superset interface for SQL Queries

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Moreover, using Apache Superset, the end users can create a wide range of chart visualisations on the provided datasets. Figure 53 displays the outcome of a pie chart visualisation, in which the production of solar panels is displayed, grouped by the region the data are coming from.

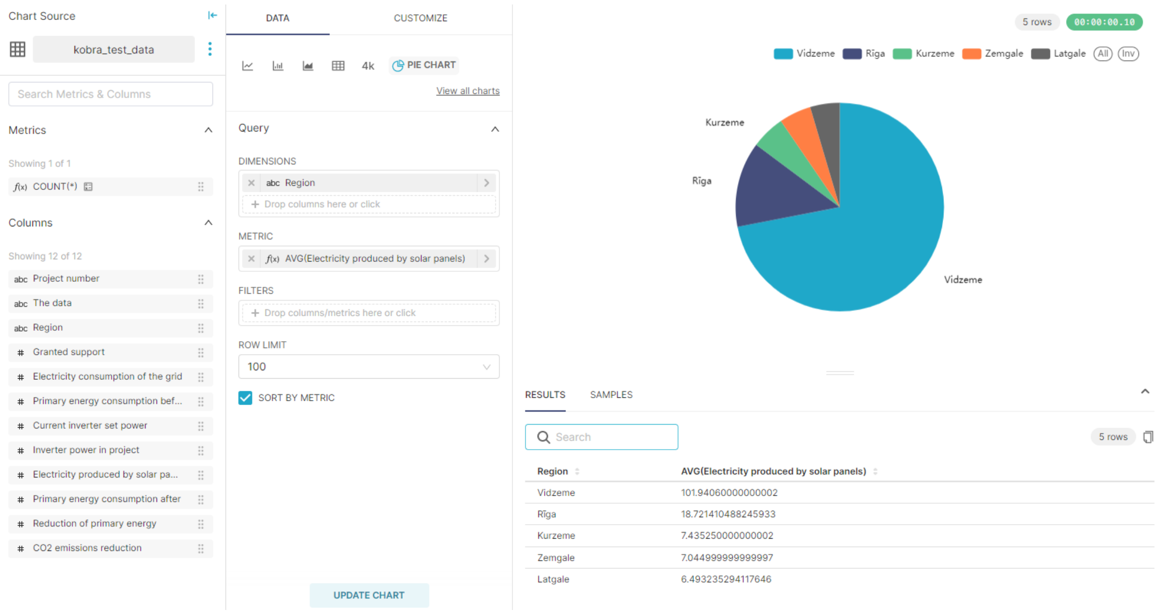


Figure 53: Apache Superset Pie chart visualisation

It is worth noting that the created chart is interactive, meaning that upon hovering on a specific slice of the chart, a tooltip explaining the result in more detail appears. The same characteristic applies to all the available chart visualisations. The users can also inspect the results of the query in the form of tables, by just clicking the corresponding button in the right bottom of the window, as shown in Figure 54. Furthermore, the looks of the charts can be tailored according to the end users' needs and preferences, since a customisation window is also available, allowing them to adjust any aspects like the colouring, the spacing and several other formatting parameters. Apache Superset also has the capability of caching some results in order to minimise the amount of resources it consumes and it also offers the possibility of saving the results of the chart visualisations, so that the end users have easy access to them, via the "Charts" tab which is available in the top bar of the tool.

Figure 54 below shows a Bar chart visualisation, in which the average of two columns, electricity produced by solar panels and CO2 emissions reduction, are displayed, group by region.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

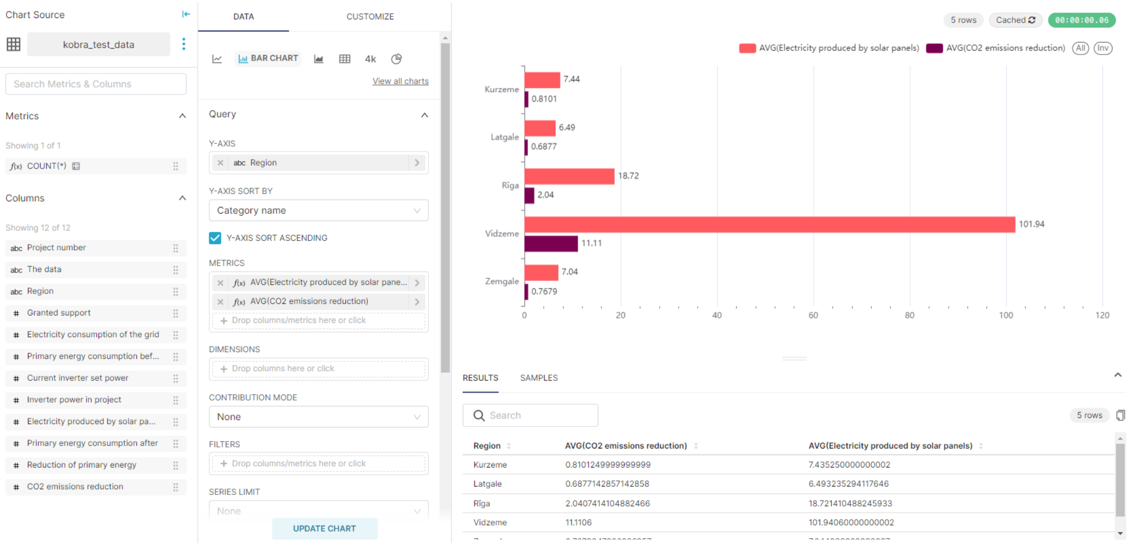


Figure 54: Apache Superset Bar Chart visualisation

Moreover, one of the most valuable features provided by Apache Superset, is the creation of dashboards. The end users are able to save their chart visualisations and gather them into customisable dashboards, in order to have quick access to them. This feature also allows the end users to see an analytical overview of the saved visualisations, assisting them to infer valuable information and insights by displaying an overall view of all the data that have been already visualised. Each visualisation can have a separate title and it is displayed in a dedicated card, through which the end user can proceed to any customisations needed, refresh them and quickly visualise the presented outcomes into reactive tables. Figure 55 below displays an example dashboard created, including the above-described charts and an additional one, which illustrates the average current inverted set power by region.

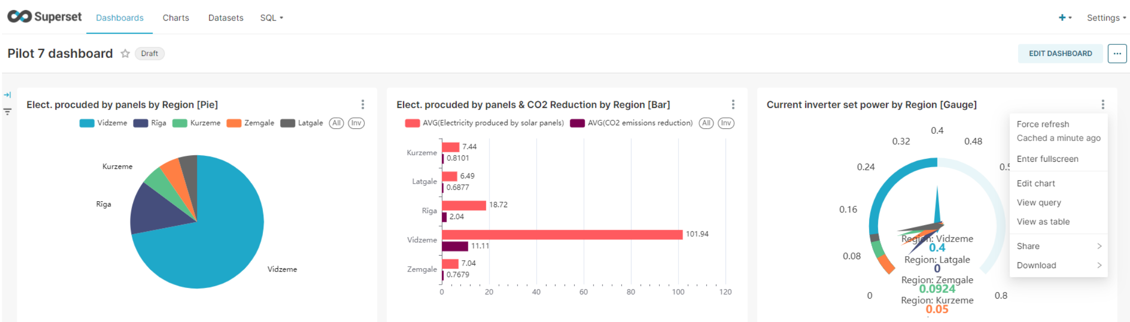


Figure 55: Apache Superset dashboard

Last but not least, since the datasets of the project can contain sensitive data, it is of crucial importance to ensure that fine-grained authentication and authorisation methods are implemented so as to prevent unauthorised access to them.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Apache Superset has a default authentication system that provides basic authentication and role-based access control. It supports various methods like username and password and OpenID Connect. The system allows users to create accounts, log in, and manage their own dashboards and datasets. Administrators can define roles and permissions to control access to specific features and resources.

However, it can also be integrated with external authentication tools like Keycloak[19] and Keyrock[20]. These are popular tools that offer advanced authentication and authorization capabilities, including Single Sign-On, Multi-Factor Authentication, and centralised user management. Integration allows users to authenticate and access Apache Superset using their existing credentials from the external tool. It also centralises user management, roles, and permissions across multiple applications, improving security and reducing administrative overhead.

During the course of the project, and more specifically in the next technology releases, based on the needs that will arise, and the decisions made in the responsible WPs, we will decide which identity management and access control technology we will integrate with the Apache Superset instances EnerShare.

5.3 Integration with ENERSHARE Data Space and Services

As already mentioned, Apache Superset is a lightweight, yet powerful tool that allows advanced data exploration and visualisation. One of the key features that has boosted its usage is the fact that it can seamlessly connect with a diverse range of databases, offering high flexibility for users that require access and analysis of their data. Thus, Apache Superset is an optimal choice for the EnerShare needs, since it can establish connections not only with popular relational databases, such as MySQL[21], PostgreSQL[22] and SQL Server[23], but also with NoSQL databases like MongoDB[24] or Cassandra[25]. Additionally, Apache Superset supports cloud-based databases like Amazon Redshift[26], Google BigQuery[27] and Apache Druid[28]. Apache Superset offers an easy and understandable way of connecting with any database, by filling in a form with the needed fields required to establish the connection.

[19] https://www.keycloak.org/
[20] https://fiware-idm.readthedocs.io/en/latest/
[21] https://www.mysql.com/
[22] https://www.postgresql.org/
[23] https://www.sqlservertutorial.net/
[24] https://www.mongodb.com/
[25] https://cassandra.apache.org/_/index.html
[26] https://aws.amazon.com/redshift/
[27] https://cloud.google.com/bigquery
[28] https://druid.apache.org/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

As for the needs of ENERSHARE, our intention is to deploy one instance of Apache Superset in the premises of each pilot and, additionally, to have a central deployment for the whole project. In this way, each pilot will be able to visualise and process their data, having secure and exclusive access to them, enabling exploration and analysis in an intuitive and user-friendly fashion. Moreover, the central deployment of Apache Superset will allow to upload any project data that are not classified and have a central point with dashboards that demonstrate the functionalities of the tool but also some useful insights for the available datasets.

5.4 Implementation details

Regarding the implementation details, Apache Superset is an open-source platform that has been under active development since 2015, formerly known as Caravel. Apache Superset does not have an official Technology Readiness Level (TRL) assigned to it. TRL is a measurement system used to assess the maturity and readiness of a technology or product for deployment and commercialisation. However, having gained significant popularity, it has a vast community of contributors. The license under this software is Apache License 2.0[29].

Apache Superset is primarily written in Python, a popular and widely used programming language which provides a robust ecosystem for data manipulation, analytics, and web development. Due to that fact, it is a suitable choice for building such an interface. At the same time, Apache Superset also relies on several third-party software components, each of which has its own license. Flask[30], SQLAlchemy[31], Pandas[32] and NumPy[33] are some of the key dependencies of Apache Superset.

Regarding its deployment, Apache Superset can be easily deployed using Docker[34] and Docker Compose[35], thus allowing Apache Superset to run in a containerised environment. Docker allows encapsulating Apache Superset and its dependencies into a portable and isolated container, ensuring consistency across different deployment environments. Docker Compose users can define and manage multi-container applications, and the combination of Docker and Docker Compose makes the platform's deployment more flexible, scalable, reproductible, and efficient.

Official documentation on how to deploy it on the website of Apache Superset, so the process followed for the needs of ENERSHARE is the most safe and reliable one.

[29] https://www.apache.org/licenses/LICENSE-2.0
[30] https://flask.palletsprojects.com/en/2.3.x/
[31] https://www.sqlalchemy.org/
[32] https://pandas.pydata.org/
[33] https://numpy.org/
[34] https://www.docker.com/
[35] https://docs.docker.com/compose/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

6 System-of-systems Integrated Digital Twins

This chapter describes in a non-exhaustive way the main concepts of the digital twins, the system-of-systems DTs and the approach that the ENERSHARE project is taking regarding the concepts. There is no unified view in the industry now due to the novelty of the topic, and the innovations provided by ENERSHARE could contribute to the discussion. One of the crucial points in the process of developing a DT is the data interoperability and the use of data spaces, then based in the experience of the FIWARE Smart Data Models we discuss the key challenges and ways to address them.

Finally, the description of the digital twins developed as services in ENERSHARE follows:

- Digital Twin for optimal data-driven Power-to-Gas optimal planning
- Digital Twin based O&M algorithms and generation of synthetic failures data
- Digital Twin for flexible energy networks

Each section discusses the architectural decisions, the overall description and sets of functions and the details and roadmap for the implementation. Please notice that the descriptions here are non-binding and subject to changes, as the parallel developments of services from the diverse technical WPs may be incorporated, as well as the technological decisions and the improvements during the successive iterations of the development stages of the DT services.

6.1 System-of-system Digital Twins: ENERSHARE Approach

A digital twin is a virtual representation of an object or entity (most of the times physical, but can also be virtual) where the data flows and couples the relevant information for a particular application in a period of the lifecycle of the object or entity.

As consequence, a set of very important processes need to be accurately done and the challenges originated solved, to make sure that the existing systems and the incoming information connectors and interfaces from diverse sources can interoperate with the DT and provide a complete solution to the stakeholders.

With the data previously mentioned the service provider creates a complex clone of the system, by using simulations or data-driven solutions, generating the data output for the specific goal that it is designed for: mimic a particular set of characteristics that reproduce the expected variables of interest of the object or entity twinned. This coupling, called mirroring or

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

shadowing, allows the creation of an instance of a DT of the entity that fulfils a particular goal, i.e., testing of a future product, forecasting the behaviour, evaluating the performance, etc.

A DT instance then can additionally be composed of the aggregation of different other DTs or their outputs, combining different representations to be able to operate in a multi-physics or a multi-objective domain.

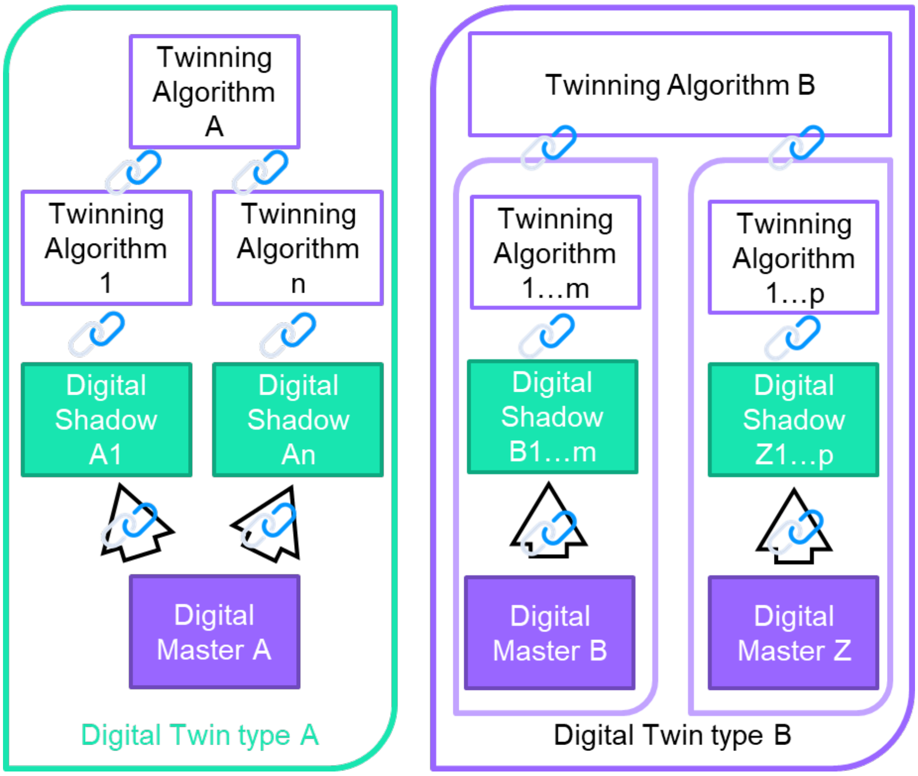


Figure 56: Types of DTs as system-of-systems

When considering concepts of system-of-systems, the DTs are particularly interesting due to their dual condition in that respect:

DTs are per-se system-of-systems: as the individual parts are separately twinning parts or behaviours of the system (i.e., a model-based twin and a data driven twin using the same data). This case is shown in the Figure 56 as type A. DTs can use other DTs as data sources: two or more different DT systems are used to generate an output that is relevant to the user (i.e., Multiphysics systems like wind turbines, electricity networks with renewables). This is shown in the Figure 56 as type B.

In this deliverable, the concrete implementation of the DTs will show at least one of the concepts described. The next iteration of the deliverables will elaborate more on the topic when

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

a more detailed architectural concept is made available, as the solutions get tested and concrete in their approach.

In a concept of system-of-systems, a focus in the interoperability is crucial. As intermediate and end goal, it could be achieved by using the Dataspace to enable the cross-sector and cross-domain operations in a standardized way. Hence, the data models and their efficient use becomes one of the pillars for the DTs and their implementation. The experience of the future developed solutions will enable ENERSHARE to enrich and collaborate with the initiatives and models adopted.

Data Models and Formats are related to the interpretation of data within Data Spaces. Initiated by the FIWARE Foundation, the Smart Data Models (SDM) initiative aims to provide a collection of data models accompanied by descriptions and representations in various data formats. Specifically, the JSON and JSON-LD formats are compatible with the NGSIv2 and NGSI-LD APIs, but also available in csv, geojsonfeatures, DTDL, etc, as well as other RESTful interfaces. When available, the data models published under this initiative align with widely accepted standards, either across domains (e.g., schema.org) or within specific domains (e.g., IEC CIM for the Energy domain). To address the absence of standard data models, a community-driven approach is employed, wherein multiple organizations contribute and collaboratively maintain the models they have developed, implemented, and tested in real-world projects. The SDM initiative places great emphasis on agility, leading to significant growth in the number of covered data models and contributing organizations over the past years. This agile approach, see manifesto, allows to produce a data model in less than a week if information is provided and less than a day for updates and extensions. At the time of document publication, the initiative had published resources linked to over 1000 data models.

One of the key challenges addressed by the initiative is the variability in mapping a given data model specification to JSON or JSON-LD, each with its own valid approach. By relying on the published mappings into JSON/JSON-LD that are compatible with NGSIv2/NGSI-LD APIs or other OpenAPIs, developers can avoid interoperability issues that arise from alternative mappings.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

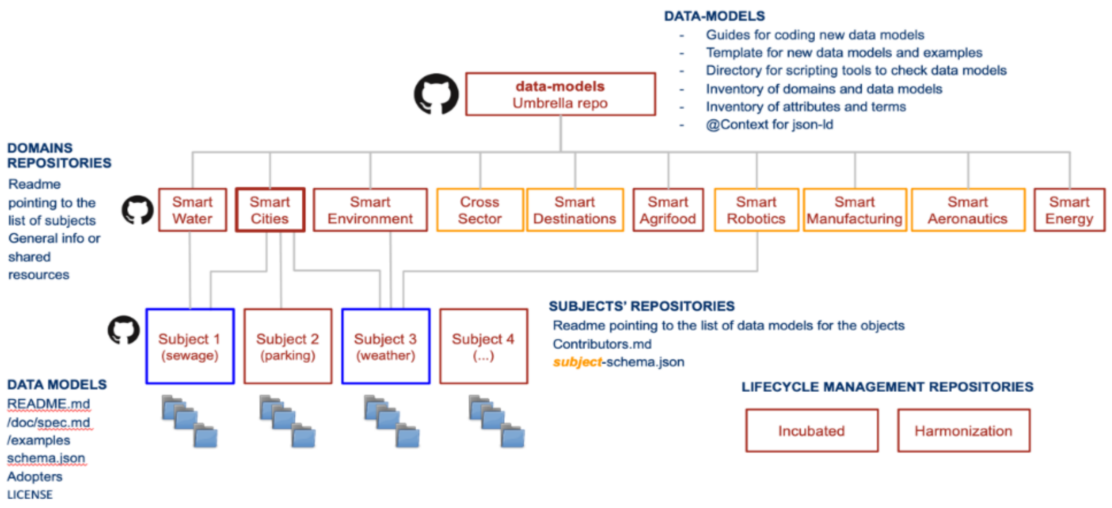


Figure 57: Smart Data Model organization on GitHub

Figure 57 illustrates the organizational structure of resources within the Smart Data Models initiative on GitHub. Data models are categorized under "subjects" such as energy, weather, transportation, and environment, which are in turn associated with repositories dedicated to various application domains, including Smart Energy, Smart Cities, Smart Agrifood, Smart Manufacturing, and Smart Water, etc (13 domains available). It is worth noting that some subjects are specific to application domains (e.g., "smart parking" within smart cities and communities), while others have relevance across multiple domains (e.g., "weather" applicable to nearly every domain or "sewage" relevant to both Smart Energy, Smart Cities and Smart Water domains).

The Smart Data Models initiative follows an open governance model that defines the lifecycle of data models, encompassing the incubation of new models and the curation of existing ones through the harmonization of diverse contributions. The management processes and procedures adhere to best practices established by open-source communities and are guided by principles of transparency and meritocracy. Recognized organizations such as TM Forum, OASC, and IUDX collaborate with the FIWARE Foundation to support this open governance model and more than 100 companies contribute to the data models.

6.2 Digital Twin for optimal data-driven Power-to-Gas optimal planning

Power-to-gas (P2G) systems are being developed as an innovative technology for energy storage and conversion of environmentally friendly energy sources into gaseous fuels. P2G systems utilize excess energy production from renewable sources such as wind and solar energy to produce hydrogen or methane. These fuels can be stored and used later for electricity generation or as fuels for vehicles.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

6.2.1 Description of the DT

In the context of ENERSHARE pilot 4 (DEPA) we aim to develop a digital simulation and optimization platform called TwinP2G. The platform integrates the national transmission and distribution networks of natural gas and electrical power managed by DESFA[36] and IPTO[37], respectively. TwinP2G employs a DT architecture to enable multi-resolution simulations involving power-to-Gas (P2G) technologies and regenerative hydrogen fuel cells. The objective is to optimize the use of surplus renewable energy for green hydrogen production through electrolysis.

The high-level architecture of TwinP2G, as shown in Figure 58, follows a Platform-as-a-Service (PaaS) design meant for multiple user roles. It comprises various subcomponents that utilize state-of-the-art technologies. The architecture and functionalities have been also described in (Pelekis, et al., 2023), however a brief and updated version is included in the current document as well. Note here that the descriptions there are not binding for the implementation and several things have changed. According to the latest architecture version, TWINP2G comprises of several components as follows.

- **Data warehouse and ingestion mechanism** (fully conceptualized – currently under initial development stages)
- **Simulation and optimization** (fully conceptualized – functionalities demonstrated in section 6.2.3)
- **Forecasting toolkit** (developed in I-NERGY project – pending integration with the DT)
- **Front-end** (to be developed)
- **Identity and access management** (to be developed)

These components are further analysed in the next section. Note here that the components do not share the same maturity level as above mentioned.

[36] Desfa. (2022). DESFA S.A. - desfa.gr. https://www.desfa.gr/en/
[37] IPTO. (2022). Independent Power Transmission Operator | IPTO. https://www.admie.gr/en

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

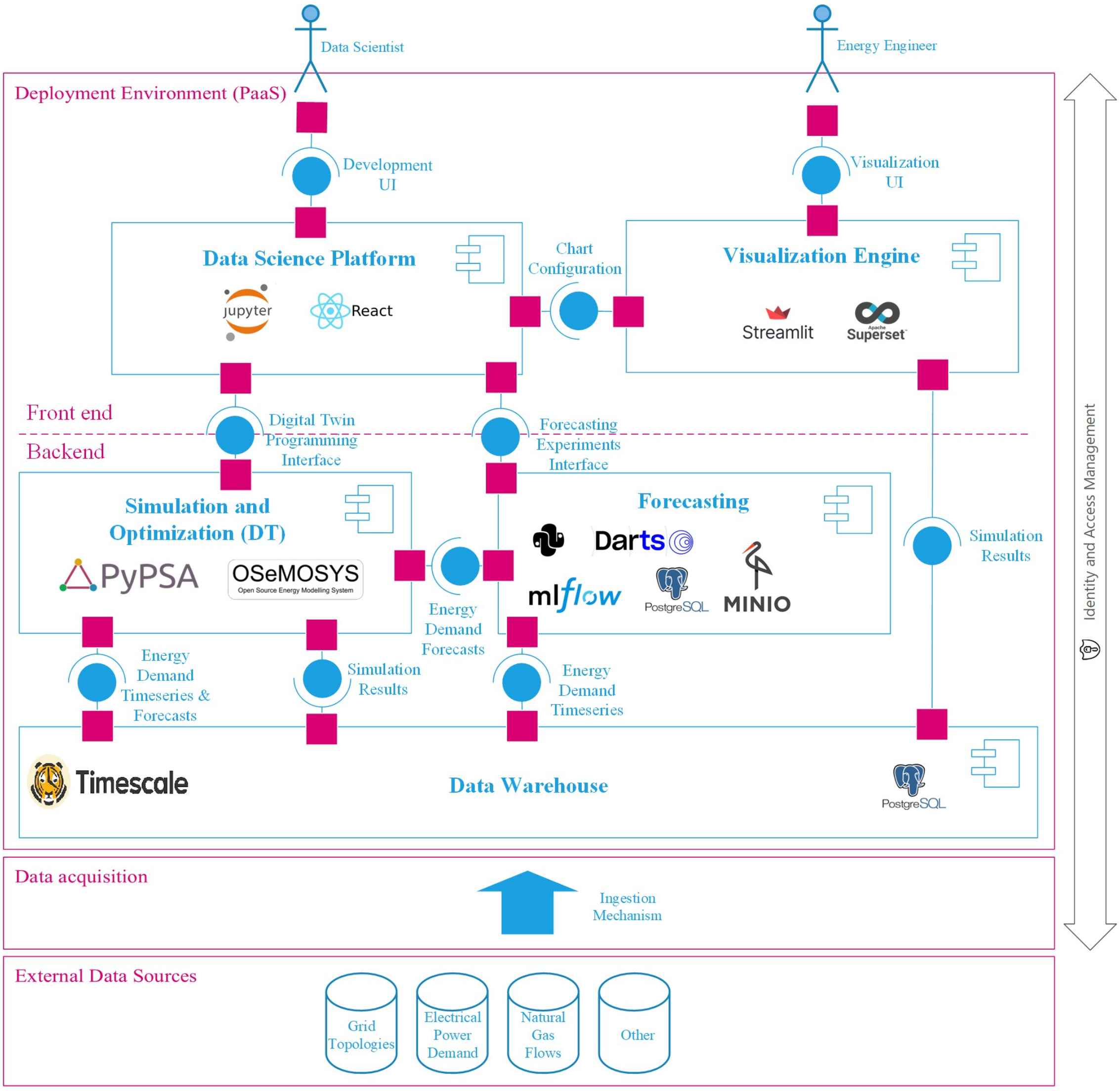


Figure 58: The high-level architecture of TwinP2G

6.2.2 Functions

6.2.2.1 Data warehouse and data ingestion mechanism

Data ingestion from various sources is a crucial process in developing a DT application, as it enriches the local data warehouse with new datasets that are necessary as simulation inputs. Such a database system needs to orchestrate a significant number of data pipelines that will run periodically, consuming raw data from the various APIs and processing them appropriately before storing them in the centralized database. It is important for the system to be scalable and maintainable, allowing for adaptation to changes or issues with the data sources, as well as

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

the addition of new sources. A more focused view of the current approach regarding data ingestion alongside the respective technologies is illustrated in Figure 59.

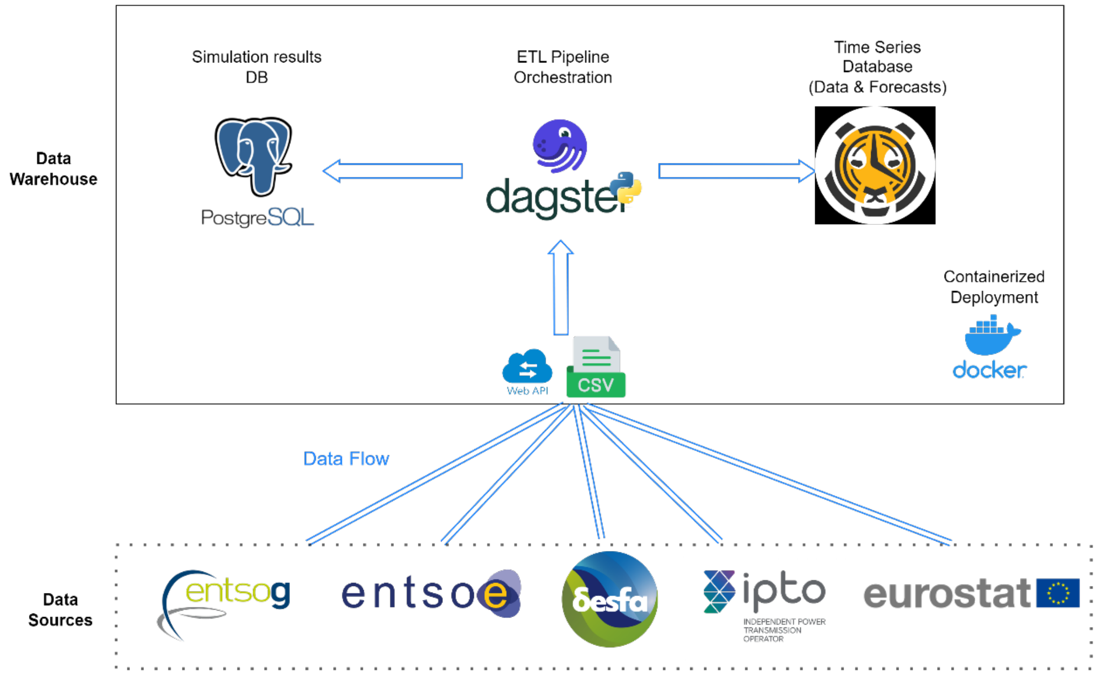


Figure 59: The focused architecture view of the data warehouse and ingestion mechanisms of TwinP2G.

Workflow management

Regarding data pipeline orchestration, monitoring, and error handling Dagster will be used in collaboration with T8.2. Dagster[38] offers a good balance between ease of implementation and adaptability to the task at hand. The logic of the pipelines is written in Python and is easily customizable to the data flow model of Dagster. Thus, Dagster serves as the framework of the system that enriches the database with accurate and up-to-date data. It provides flexibility in code development and deployment, as it can run locally or in a Docker container, and even in a Kubernetes cluster.

Database

The final destination of the data is a time series database. Similarly, there are several available databases with different implementations, internal data models, and query languages, such as InfluxDB, TimescaleDB (a PostgreSQL extension), Prometheus, and others. We opted for TimescaleDB[39] as it is a mature and specialized time series database, also providing good

[38] https://dagster.io/
[39] https://www.timescale.com/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

interoperability with the pure PostgreSQL that will be used for storing the simulation outputs, KPIs, evaluation metrics of the digital twin. More specifically, TimescaleDB is an ideal choice for time series data due to its unique architecture and features. It is built as an extension to PostgreSQL, providing the reliability and scalability of a mature database system. With its hypertable concept, TimescaleDB seamlessly handles large volumes of time-stamped data, enabling efficient data storage, compression, and partitioning. It offers powerful time-series-specific functions and optimizations, making it easy to analyze and query data across time intervals. Moreover, its compatibility with PostgreSQL allows leveraging existing SQL skills and ecosystem, making TimescaleDB a robust, flexible, and high-performance solution for time series data management.

Data sources

Regarding the data sources to be considered in the DT, there are both national and European sources of historical energy data. The domestic sources to be examined are IPTO[40], the Greek electricity transmission system operator, and DESFA[41], the natural gas network operator. These organizations are members of the pan-European organizations ENTSO-E and ENTSO-G, respectively, which will also serve as data sources. Finally, Eurostat provides some energy data, which are examined especially regarding long term projection of macroscopical quantities.

Regarding the national providers, they exhibit significant heterogeneity internally in terms of data access and formatting. IPTO offers an API and provides data, but not in a standardized format. Its API returns a number of .xls files for each type of data, depending on the requested time range. These files have different formats for each data type, requiring specialized processing. However, this automated data consumption will be attempted. IPTO provides several types of data, including system operation data such as RES Generation, System Load, and Energy Balance.

DESFA does not offer a centralized method to access data nor an API. Some types of historical data, such as daily Declared Allocations/Deliveries of Natural Gas and Hourly Receipts/Withdrawals of natural gas, are available as .xls files on its website, which are updated daily. These files can be consumed but will require continuous maintenance since the links change at least on an annual basis. Other types of data, such as Hourly Quality Data for Interconnection Points, are accessed through DESFA's SCADA system web portal, which does not appear to be functional as it never returns data.

[40] https://www.admie.gr/
[41] https://www.desfa.gr/

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

The pan-European entities ENTSO-E[42] and ENTSO-G[43] have Transparency Platforms[44] and provide access to their data through both their websites and APIs. Both entities' APIs are extensive, but the documentation is partial (Hirth, Mühlenpfordt, & Bulkeley, 2018). However, there are available third-party open-source libraries in Python that act as wrappers around the APIs, which could simplify data consumption[45].

Eurostat provides limited energy data, which is accessible both through an API and its website. The data is often integrated with a delay of several months and is aggregated, with most data reported annually and some reported monthly.

Although there is overlap in some of the data from different entities, we will attempt to record them separately in the database as the numbers may differ.

6.2.2.2 Simulation and optimization core

The simulation core of TwinP2G involves physics- and data-driven simulation and optimization. It utilizes the PyPSA open-source power system modelling tool for short-term/mid-term simulations and optimizations. PyPSA performs optimal power flow simulations based on network equations, security constraints, and least-cost optimization. PyPSA will be also considered for long-term projections and optimization, however in case it fails to cover the long-term needs of the pilot, the OSeMOSYS open-source modelling system will fill this gap. OSeMOSYS enables long-run integrated assessment, energy planning, and crisis modelling.

The main functionality of the "Simulation and optimization" component is the processing of historical time series of renewable generation, power and gas demands alongside forecasts produced by the "Forecasting" component. Based on these inputs it enables the study of local grid topologies in Greece, including P2G component investments, and using linear programming methods to cost optimize DTs that are modelled as electrical networks including buses, lines, links, generators and loads. Section 6.2.3 provides a demonstration of the simulation and optimization component.

6.2.2.3 Forecasting

The forecasting component of TwinP2G uses an MLOps framework for time series forecasting developed within the I-NERGY H2020 project (Karakolis, et al., 2022) that is currently under further development for the needs of TwinP2G. The said tool enables the experimentation with various machine learning and deep learning algorithms, such as XGBoost, Random Forest, NBEATS, Temporal Convolutional Networks (TCN), and LSTM networks. The forecasting

[42] https://www.entsoe.eu/
[43] https://www.entsog.eu/
[44] https://transparency.entsoe.eu/, https://transparency.entsog.eu/
[45] https://github.com/EnergieID/entsoe-py

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

platform is based on Python programming language and utilizes MLflow as the MLops platform, Darts as the time series forecasting framework, MinIO as artifact storage, PostgreSQL as the logging database, FastAPI for API development, and JavaScript React for the front-end. The forecasting platform handles various time series data, integrates new datasets with ease, and provides forecasts of different time horizons (short-term, mid-term, long-term). The forecasting component is designed to interface with simulation core by storing the produced forecasts to the TimescaleDB where they can be easily accessed by the latter. A short overview of the load forecasting component can be sought in YouTube[46], while the extended technical descriptions are provided in the related DeepTSF publication which is currently under preparation. Currently, the TwinP2G data ingestion mechanism is designed in a fashion that will enable harmonized interfacing with the load forecasting subcomponent.

6.2.2.4 Front-end

TwinP2G will be able to serve two main user roles or personas. The first persona is the "Data Scientist," who is an experienced user with a scientific background and coding skills, as well as modeling capabilities in P2G use cases. This user can access TwinP2G's "Data Science Platform" to interact with the "Simulation and Optimization" component. They can develop P2G experiments and visualize the results, including simulation and optimization outcomes, forecasting accuracies, and more. Currently, Jupyter notebooks form the core of the data science platform, however a more refined approach will be investigated in the future based on Javascript[47] React or Streamlit[48].

The second persona is the "Energy Engineer", who is an end-user with knowledge and understanding of energy systems but possesses limited coding and modeling skills. Such users will only be interested in observing the behaviour and the current status, and at most having some high-level information to gain insights for decision support concerning future P2G investments. This persona utilizes the "Visualisation Engine" component to monitor simulation and forecast results, as well as relevant metrics. The latter component will rely on a strong interconnection of TWINP2G with the Enershare visualisation engine developed within T6.4.

6.2.2.5 Identity and access management

TwinP2G is planned to provide an identity/access management mechanism that encompasses end-to-end processes, approaches, and technologies for user identification, authentication, and authorization. This mechanism ensures that the respective personas (user roles) are granted

[46] https://youtu.be/NuhHNYefbB8
[47] https://react.dev/
[48] https://streamlit.io/

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

access to the appropriate resources. Keycloak[49] has been considered as the technology for fulfilling these security requirements.

6.2.3 Demonstration – Case study

Currently, the implementation of the service has been focused on the Simulation and Optimization component of the architecture. In this direction, we have thoroughly experimented with PyPSA and tried to create realistic power system and power-to-gas (P2G) DT scenarios, also exploring the functionalities and capabilities of the library.

6.2.3.1 Experimental setup for network simulation

For the purposes of the simulation, we decided to experiment with a small arbitrary region of approximately 5000 inhabitants in Greece. In this direction, we downloaded the "time_series_60min_singleindex.csv" file from ENTSO-E data platform[50] which was used for feeding the energy generation and consumption data to the simulator. The dataset includes the total national photovoltaic production as well as the total energy demand in European countries and covers a period of 5 years and 9 months at an hourly interval from 01-01-2015 to 30-09-2020. After isolating the columns referring to Greece, the total consumption was divided by 2000, and the total PV generation was normalized to a maximum nominal power of 5MW, hence assuming that it would be able to cover the nominal load of the region (5MW as well) during maximum production. In the scenario, there is also a 5MW diesel generator that operates as a backup during periods of low or zero renewable energy production.

The PyPSA library in Python was used to configure the network and then optimize it using linear programming techniques. We configure three different scenarios in order to assess the economical feasibility of investing in an electrolyzer / fuel cell technology to cover the needs of said region. The experiments were executed on a personal computer with an Intel Core i5-7200U CPU @ 2.50GHz and 8 GB RAM.

6.2.3.1.1 Scenario 0 – Baseline

Setup

In the baseline scenario, the network consists of two buses, one being the reference bus ("Diesel") where energy is produced by a diesel generator ("Diesel Gen"), and the other being the production bus ("Solar") where the PV park ("PV Plant") and the load ("load 1") are located. A transmission line ("Line") connecting the reference bus to the production bus. The load, which is determined based on the aforementioned "data_load" values. It is worth noting that the capital costs were set to zero as we assume that this system already exists. The marginal costs

[49] https://www.keycloak.org/
[50] https://data.open-power-system-data.org/time_series/2020-10-06

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

were calculated for the PV park as operational expenses in €/MWh, and for the diesel generator, they include operational expenses and CO2 emissions tax. Figure 60 illustrates the simulated topology of the baseline scenario DT.

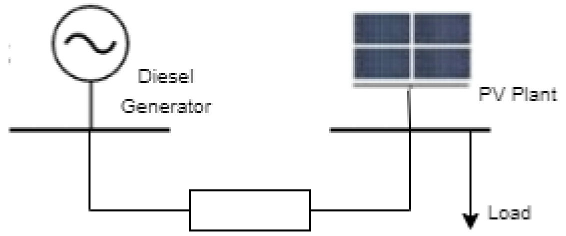


Figure 60: Baseline DT topology for simulation in PyPSA

Simulation results

The simulation was executed for the baseline scenario and the optimization took 0.6 seconds to complete. The cost of this scenario within the predefined simulation horizon was calculated at 35977814.19 €.

From Figure 61 (entire simulation horizon) and Figure 62 (specific zoomed-in period), it becomes evident that the generator operates during the nights when there is no PV production in order to meet the load, while during midday, the generator operates at 30% of its nominal power for stability reasons, and the PV park generates energy. The renewable energy is preferred as its total operational cost is far lower.

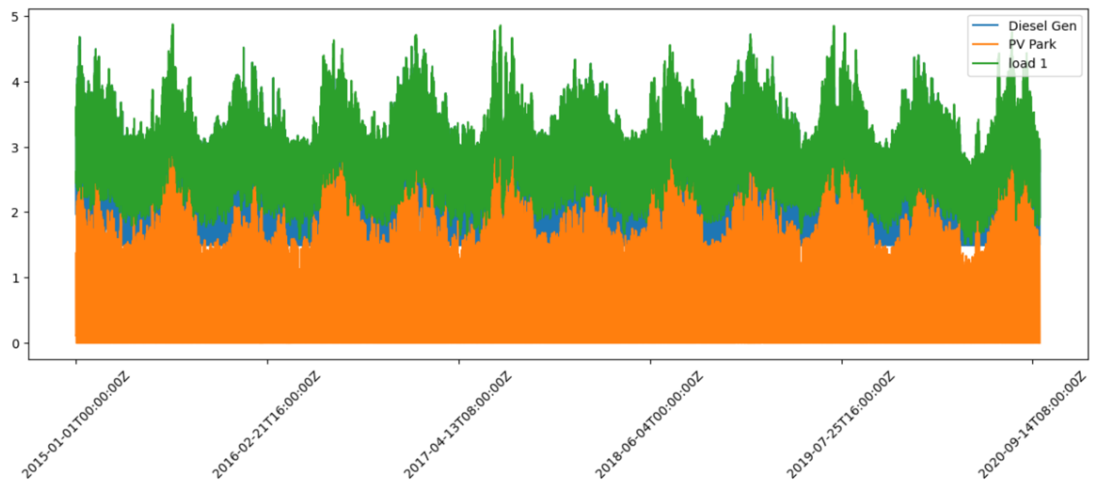


Figure 61: Load and generation active power (MW) of the grid components in scenario 0 for the entire simulation horizon (20150101-20200930)

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

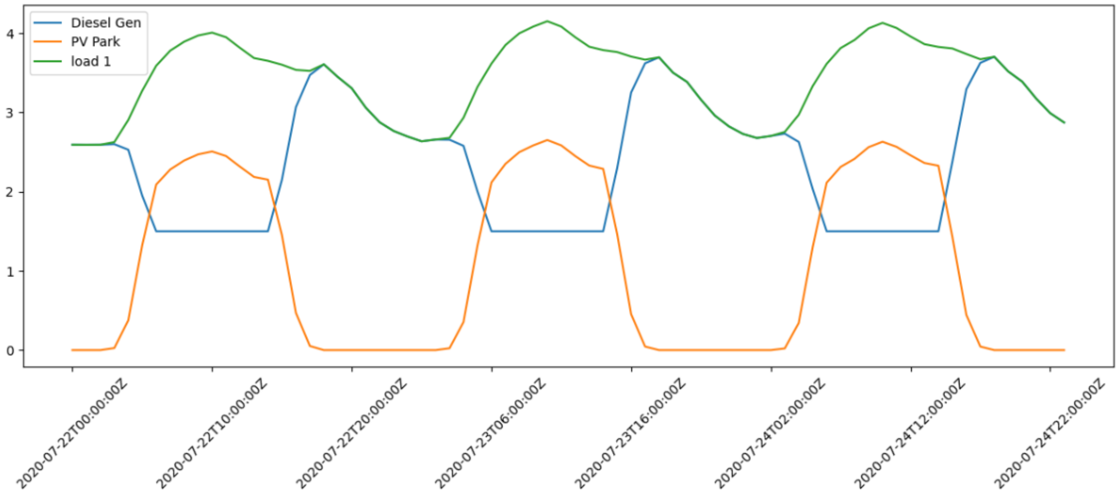


Figure 62: A zoomed in view (20200722-20200724) of the active power (MW) of the power system components for scenario 0.

6.2.3.1.2 Scenario 1 - P2G simulation

Setup

In this scenario we introduce the P2G system on top of the baseline topology. The P2G setup consists of a PEM electrolyzer with a rated power of P=2MW, efficiency of 0.7, initial cost of 1000€/kWh, and operational cost of 0.11€/kWh. A fuel cell with a rated power of P=1.2MW, efficiency of 0.6, initial cost of 1000€/kWh, and operational cost of 0.11€/kWh. A hydrogen store with a capacity of 15MWh, initial cost of 40000€/MWh, and operational cost of 10€/MWh. The "Hydrogen" bus to which all the above components are connected. The electrolyzer (power to H_2) and the fuel cell (H_2 to power) were implemented using "Links" that allow the connection of buses and the conversion of energy from one form to another. Additionally, since the investment has a duration of 25 years while the available data is approximately 5 years, the capital costs were divided by 5. Forecasting is expected to solve this issue in future versions of this deliverable. Figure 63 illustrates the simulated topology of the DT related to the P2G investment.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

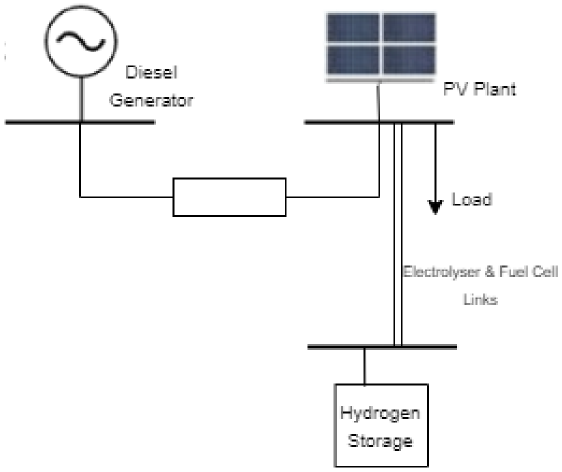


Figure 63: DT topology with P2G for simulation in PyPSA (also applicable to the scenario with reduced equipment costs)

Results

The simulation was executed for the first P2G scenario and the optimization took 1286.1 seconds to complete. The cost of this scenario within the predefined simulation horizon was calculated at 35117815.02 €.

Figure 64 depicts the response of the system during the entire simulation period (5 years and 9 months) and we notice that the P2G system is not operating. This is because, based on the current parameter values, the system utilizes energy from the diesel generator, even though its operational cost is higher. This is because the capital costs of the electrolyzer and fuel cell are quite high and the investment is excluded by the optimization software.

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

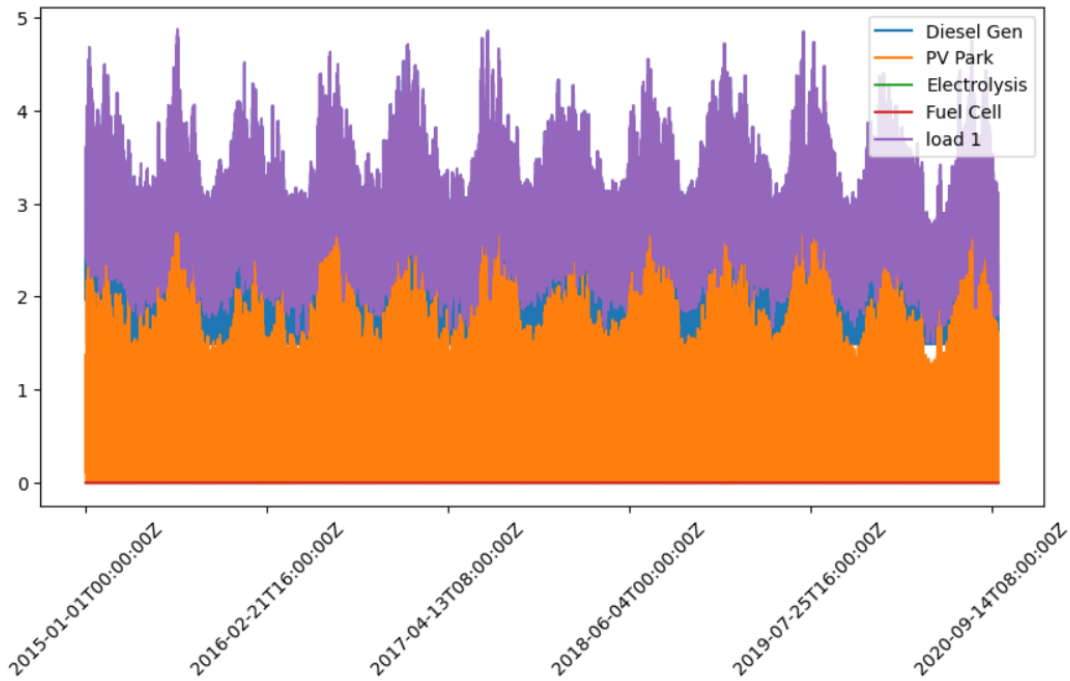


Figure 64: Load, generation, electrolysis, and fuel cell active power (MW) within the power system of scenario 1 for the entire simulation horizon (20150101-20200930).

Figure 65, illustrates the system response for the specific period '2020-07-22' to '2020-07-24'. It is evident that the generator operates during the nights when there is no PV production to meet the load. However, during the afternoons, the generator operates at 30% of its nominal capacity for stability reasons, while the PV park generates energy, and this energy is preferred as its operational cost is lower.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

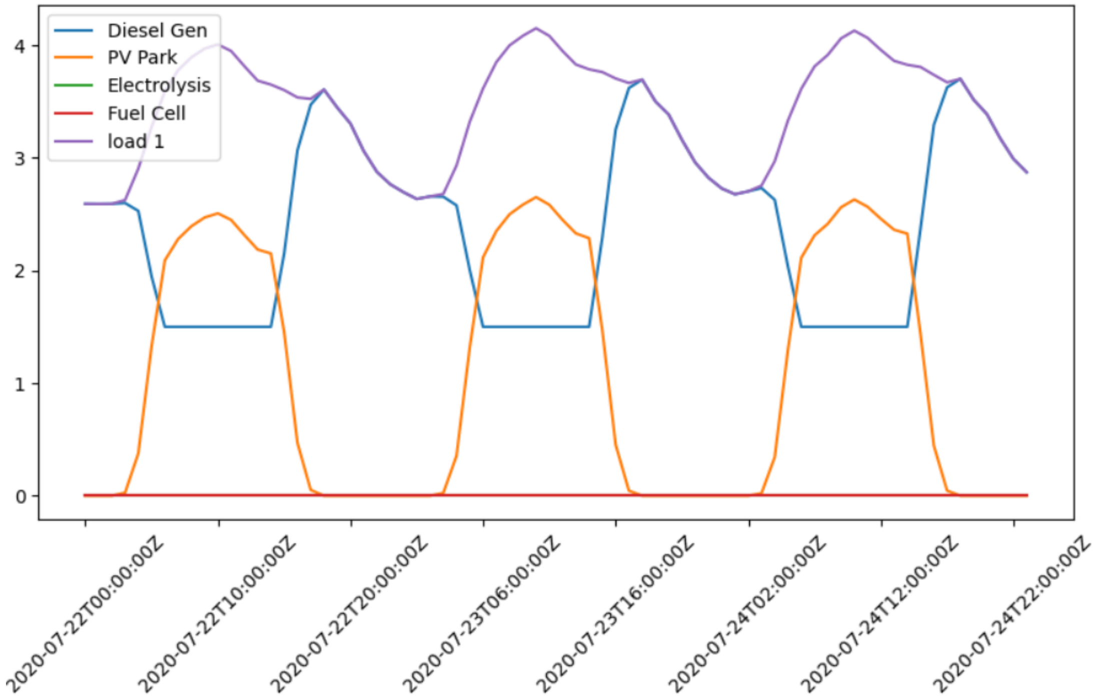


Figure 65: A zoomed in view (20200722-20200724) of the active power (MW) of the power system components for scenario 1

6.2.3.1.3 Scenario 2 – P2G simulation with reduced equipment costs

Setup

As the first P2G scenario, which was based on naïve cost-related assumptions, was proved unsustainable regarding the investment in P2G, we create a second scenario assuming lower prices of the P2G equipment. In 2030, it is speculated that the initial costs of the P2G components will have been reduced by up to 70%[51]. Therefore, the following changes can be applied to the network compared to scenario 2: The PEM electrolyzer is now selected with a nominal power of P=2MW, efficiency=0.7, initial cost of 350€/kWh, operational cost of 0.03€/kWh. The fuel cell with a nominal power of P=1.2MW, efficiency=0.6, initial cost of 350€/kWh, and an operational cost of 0.03€/kWh. The hydrogen storage with a capacity of 15MWh, initial cost of 30000€/MWh, operational cost of 10€/MWh. The operational cost of the diesel generator is increased to 500€/MWh assuming a significant increase of CO_2 emission taxes (Gorre, Ortloff, & van Leeuwen, 2019). Figure 66 equally illustrates the simulated topology for scenario 2.

Results

[51] https://www.iea.org/reports/global-hydrogen-review-2022/executive-summary

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

The simulation was executed for the first P2G scenario and the optimization took 2046.9 seconds to complete. The cost of this scenario within the predefined simulation horizon was calculated at 56041938.02 €.

Figure 66 depicts the response of the system during the entire simulation period (5 years and 9 months) and we notice that the P2G system is operating. This is because, we significantly reduced the capital costs for the investment alongside increasing the operational costs for the diesel generator in a realistic fashion.

Figure 66: Load, generation, electrolysis, and fuel cell active power (MW) within the power system of scenario 2 for the simulation horizon (20150101-20200930)

Figure 67 depicts the system's response for the short period of '2020-07-22' to '2020-07-24'. During the daytime, we can see that the generator (blue) operates at its minimum power output, while the PV park (orange) generates at its maximum capacity. Additionally, a portion of the surplus energy is used for electrolysis (green). From the second graph, it is evident that as the electrolyzer operates, the hydrogen energy stored in the store increases.

During the nights, we observe that there is obviously no generation from the PV plant, and the fuel cell is used to supply energy to the grid. As a result, the hydrogen stock in the store decreases. The remaining energy required to meet the load is generated by the diesel generator.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

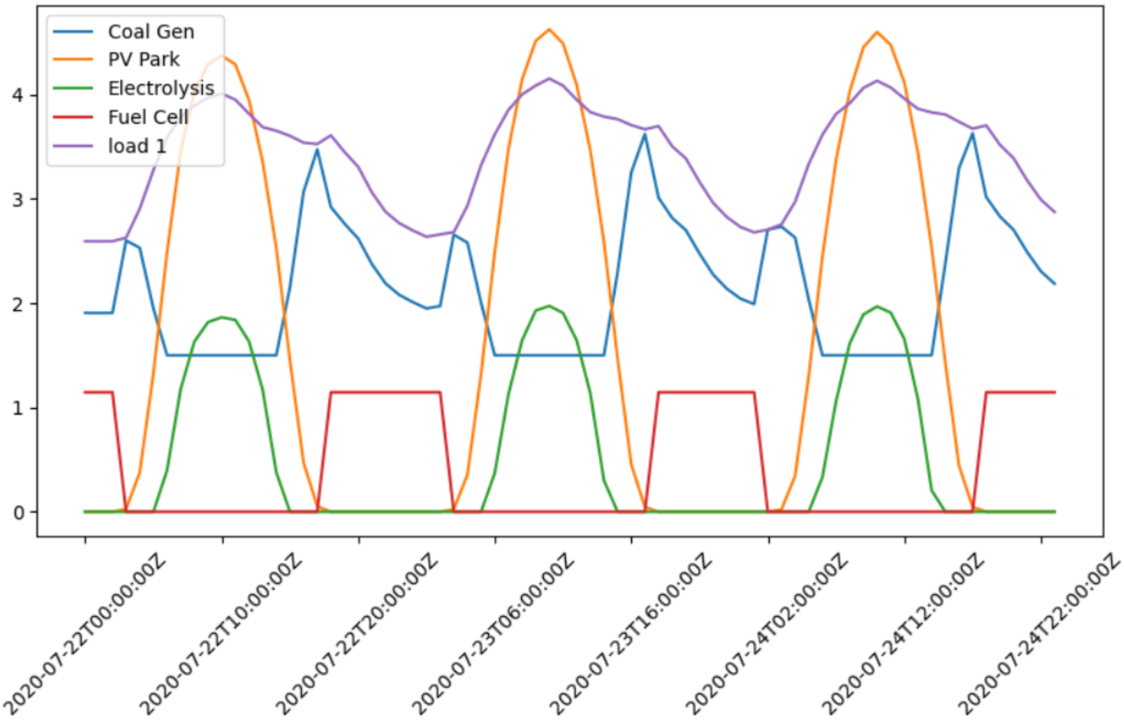


Figure 67: A zoomed in view (20200722-20200724) of the active power (MW) of the power system components for scenario 2

6.2.3.1.4 Conclusions and future work

As from the simulated scenarios in the DT we can observe that current P2G equipment are still quite expensive to be easily integrated into existing power grid setups with already installed RES. However, the simulation horizon in our case was pretty limited due to the limited availability of electrical energy time series data. Additionally, the parameterization of the power grid and the costs of its components was based mainly on empirical and vague estimations. Our focus from now on will be:

- to refine those scenarios with more accurate and long-term datasets (data ingestion and forecasting)
- create more scenarios according to the pilot needs also incorporating natural gas demands in the simulation
- incorporate the costs of specific existing P2G components based on official market prices
- integrate the simulator with the forecasting component alongside with TwinP2G's database in order to be able to consume time series, related forecasts and also publish the simulation results for later consumption by the end-user.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

6.2.4 Implementation details

Regarding the current status and the planning for future work within TwinP2G, Table 44 provides the necessary details.

Table 44: Implementation roadmap for TwinP2G

Deliverable	TRL level	Comment
D6.1	TRL3	Description of the architecture of the toolchain that acts as the core enabler of the DT functions. Initial simulations on dummy datasets have been already created and demonstrated. Data sources have been identified and ingestions specifications are designed.
D6.2	TRL 5	Fully interconnect and test a baseline DT configuration with real input data coming from the database.
D6.3	TRL 6-7	Fully interconnect and test a final DT configuration. The pipeline will start with input data coming from the database. Output data will be stored to the relational output database while the connection to the Energy DataSpace will be tested for sharing those results.

Regarding third party software, the main tools have also been demonstrated in Figure 59, however a detailed listing is also providing here as follows.

- **Languages:** Python, React (for front-end if necessary)
- **Frameworks / tools (all software is open-source):**
    - PyPSA: simulation toolkit
    - OSEMoSYS: for long-term simulation if necessary
    - Dagster: workflow orchestration in collaboration with WP8 (see section 6.2.2.1)
    - TimescaleDB: timeseries database (input and forecasts)
    - PostgreSQL: output/results database
    - Keycloak: identity and access management)
    - Jupyter notebooks: DT programming and configuration
    - Apache Superset: visualisations in collaboration with T6.4
    - Streamlit: for visualisations if necessary
    - Deep-TSF: forecasting interface
- **Containerization:** Docker

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

6.2.5 Integration with ENERSHARE Data Space and Services

The DT mainly relies on two input sources:

- one quasi-static, which represents the topological connections between the electrical components and the values of their parameters, and the initial conditions for the case of dynamic simulations. The most likely source in the project will be manual input.
- The TimescaleDB database entries that have been above described. The DT will be integrated with the database so that timeseries ingested from IPTO, ENTSOe, Eurostat etc. can be given as input to the simulation core to initiate the simulations on the predefined network topology.

The final output of the DT consists of an application layer that could provide contextualized meaningful data, by giving an interpretation of the output. In this direction, the DT will provide several outputs regarding the simulation results storing them to the Postgres database. Several examples follow:

- Total cost during the simulation period
- CAPEX
- OPEX
- Payback period

These outputs are available for integration and connection with the EnerShare DataSpace, and for more user-friendly interactions via a graphical dashboard or interface. This is expected to be an iterative process that will incorporate the suggestions and validation of all the stakeholders, specifically focusing on the pilot requirements and needs.

6.3 Digital Twin based O&M algorithms and generation of synthetic failures data

6.3.1 Description of the DT

The present DT will be focused on the improvement the O&M of wind turbines. For it, it will integrate anomaly detection functionality for the gearbox, electric generator and the hydraulic pitch system, based in simulation models. The individual services are detailed in part 3.4, and the DT developed here will integrate and implement these services as a whole in a DT.

Current TRL is TRL5, as a service of similar characteristics has been validated in relevant environment. At final stage, it is expected to achieve TRL7 (demonstration in operational environment).

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

6.3.2 Functions

This DT will include the functions described in Section 3.4.3 for pre-processing, processing and use of data. All these functions will be dockerized inside the DT.

6.3.3 Integration with ENERSHARE Data Space and Services

In a similar way as made in the data driven services (see Section 3.4.6) and expecting to merge this DT with the service detailed in part 3.4., the preliminary schema of the integration of this service in ENERSHARE Data Space is shown in Figure 68. IDS connectors will be used for the exchange of data. ENGIE, TECNALIA and HINE will implement an IDS pipeline formed of one connector provider and consumer in each side to be able to execute the service properly.

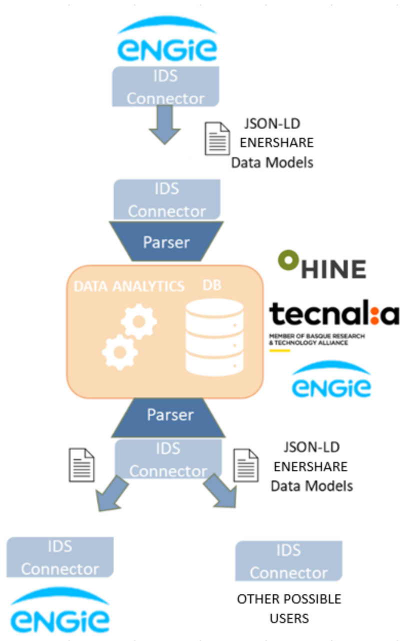


Figure 68: Integration of Failure diagnosis service in ENERSHARE Data Space

These IDS connectors will be integrated with the rest of ENERSHARE Common IDS components (Metadata Broker, Identity Manager, Vocabulary Hub...). ENGIE has built a data lake in their DARWIN platform where all the data from different windfarms and different heterogeneous data sources is available. This data lake contains historical data and also records new streaming data. ENGIE will develop an adapter to create a JSON-LD directly from the data stored in their data lake to a JSON-LD format harmonized according to the ENERSHARE Common Data Models. In the same way, TECNALIA will create a wrapper to transform the input data from the JSON-LD sent by ENGIE to the internal data model used within the data analytics tools. TECNALIA will define also the mapping rules to convert the outputs of the data analytics tools to send the results to ENGIE in a JSON-LD format harmonized according to the common ENERSHARE Data

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Models. These created JSON-LD files will be sent in the payload as an IDS artifact. An orchestrator will be developed that will be in charge of calling the different components within the data analytic tools with the appropriate data coming from the database and processing and parsing the outputs. All the components will be dockerized.

6.4 Digital Twin for flexible energy networks

6.4.1 Description of the DT

The DT concept for electrical networks is based on the simulation tool DPSim and the acquisition of measurement points from a real electrical network. The idea behind is to replicate the behaviour of the network in real time settings, with the use of the most up-to-date status information that becomes available to make the calculations.

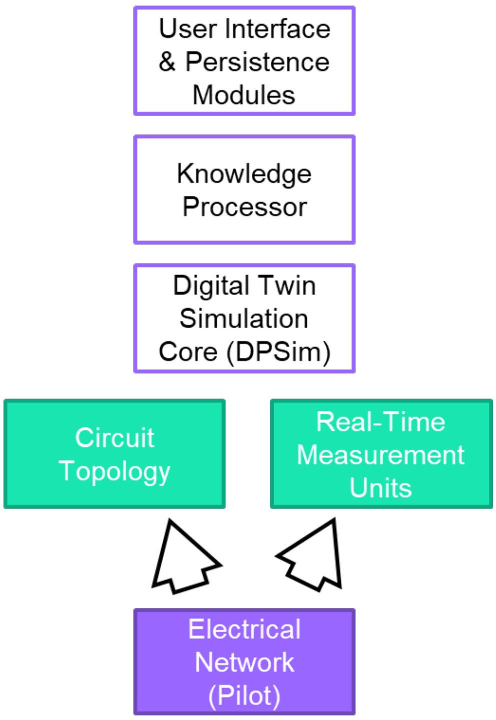


Figure 69: High-level architecture of the DT for flexibility in electrical networks

In the Figure 69 the proposed architecture for the DT with applications in the field of flexibility for electrical networks is depicted. The central role in the application is performed by DPSim, a

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

software designed by the RWTH Aachen University that is going to be adapted and extended to provide real-time outputs.

The modules in the figure will perform in a layered structure, and the architecture is going to be subsequently updated iteratively to include the advances in the Dataspace both for interfaces related to the acquisition of the data and for the interaction with the user and the persistence services.

The Electrical Network (Pilot) is the object that is being twinned. It consists of a real electrical network with the following characteristics:

a. known data about the topology, their components, applicable modelling techniques and the parameters to be adopted.

b. measurement devices in place, that can acquire the values of a set of variables that allow the simulation to run

c. infrastructure that enables the data transfer in a deterministic way for the measurement devices (and to report any changes to the topology)

The Circuit Topology is the representation of the electrical network in a level of abstraction that preserves and represent the behaviour of interest for the Use Case. This includes the connections, the individual components, the model representations to be assigned to them and the parametrization that is related to such model.

The Real-Time Measurement Units are the sensors that are part of the electrical network, with their data and communication links. They are the ones that allow to tell apart a simulation from a DT, by providing the needed linkage to the reality by updating the values and keeping them matched with the reality.

The DT Simulation Core is one of the key components of the simulation services that the DT provides. The inputs from the Circuit Topology and Real-Time Measurement Unit provide the necessary data to make the simulations. The set of voltages, currents, active and reactive powers can be provided upstream to be analysed and used to draw conclusions.

The Knowledge Processor is the other core component of the service, and it will be developed by using the experience on previous EU Projects like i-NERGY. The goal is to give context to the information and help to have a more informed decision-making by providing insights to the human user. The expert human oversight will still be required, but a planning on how to allocate the resources/power available in the network or how to plan the best state for the network could be enabled by the technologies that use a hybrid approach based both on models and on novel techniques like data-driven, machine learning or artificial intelligence.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

The last layer is the frontend to the user, and it will be designed in a cooperative way to cover the needs of the pilot. It should present the information needed by each type of user, according to their expertise. Additionally, it should be capable to save the data or generate simple reports that contain the information displayed.

6.4.2 Functions

The functions in the first version of the integration are designed in a way that can become part of a more user-friendly interface. Some users will only be interested in observing the behaviour and the current status, and at most having a high-level information. However, specialized knowledge regarding the networks will be needed to modify the parameters of the current circuit or the constraints that define the status and the way it is evaluated.

Two different sets of Jupyter notebooks will be created to support the envisioned roles: Electrical Engineer or Technician User (EETU) and Visualisation User(ViU).

The expected set for EETU functions will be:

- Load electrical circuit topological data
- Modify loaded circuit data
- Configure simulation parameters (timestep, duration)
- Configure measurement source and parameters
- Configure ViU status display constraints
- Enable or disable ViU usage
- Start and stop simulation

On the other hand, the expected set of ViU functions are simpler:

- Connect to a simulation instance
- Enable/disable/filter for the information showed regarding the status of the network and devices

6.4.3 Implementation details

Deliverable	TRL level	Comment
D6.1	TRL4	Setup of the toolchain that acts as the core enablers of the DT functions
D6.2	TRL 5	Interconnect and test the basic circuit configuration

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

		and parameters and a minimal set of measurements. A basic test using the Dataspace connector would be made, if available.
D6.3	TRL 6-7	Interconnect and test the final circuit configuration and parameters and the full set of measurements. If available, a connection using the Energy Dataspace will be also tested.

The software will be licensed under MPL 2.0 following the open-source approach of the RWTH Aachen University in regards to the contributions to LFE SOGNO, of which DPSim is part. The programming tools and strategies envisioned are the following:

- Languages: C++, Python v3, Markdown
- Compiler: CMake for Linux/Fedora
- Tools: Jupyter Notebooks, VisualStudio Code
- Containerization: possible with Docker/Rancher
- Job scheduling: for the simulation is managed internally by DPsim with a graph-based approach to calculate the interdependency between the tasks. The ViU does not need to manage this, neither the EETU most of the cases.

6.4.4 Integration with the ENERSHARE Data Space and Services

The input formats depend on the intended purpose of the input. In a DT with electrical applications there are two main input sources:

- one quasi-static, which represents the topological connections between the electrical components and the values of their parameters, and the initial conditions for the case of dynamic simulations. The most likely sources in the project are going to be in MATPOWER format, CIM/CGMES or manual input (mostly for final adjustments).

- one that can be updated, which represents the conditions to be simulated as values of the electrical magnitudes in the circuits. These can be single points or streams of data, and can come from measurements or from predictions of future values or a case study for a condition to be analysed. Most of the times this is represented by timeseries coming from an input interface.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

The output of the DT simulation core, in raw format, is mainly timeseries. Two main sources can be obtained: the electrical magnitudes and values, and the status of the components.

The final output of the DT consists of an application layer that could provide contextualized meaningful data, by giving an interpretation of the output. This meaning could be achieved, for example, by listing the values and their comparison with recommended status or thresholds.

This aforementioned final output is available for integration and connection with the Enershare Dataspace, and for more user-friendly interactions via a graphical dashboard or interface. This is going to be an iterative process and will incorporate the views and suggestions of all the stakeholders of the process.

7 Final Remarks

The ENERSHARE project integrates the Data Space paradigm in the energy sector, developing data-driven services and System-of-System DT applications. The current service ecosystem addresses challenges such as load flexibility estimation, local energy communities' operation, TSO/DSO functions, and optimising RES O&M. The project also enables cross-sector services, including wellness alarms, building comfort monitoring, and EV management. System-level DTs facilitate complex energy planning and real-time operations of electrical networks, wind energy systems and power-to-gas applications.

The following KPIs for ENERSHARE services were defined in the objectives of the Description of Action:

- ≥ 10 energy services demonstrating long-lasting data-centric business models. **In D6.1: 9 energy services were described and available in its alpha version, plus the 4 federated learning implementations for energy services.**
- ≥ 6 non-energy services enabled by the Energy Data Space. **In D6.1: 7 cross-sector services were described and available in its alpha version.**
- ≥ 15 digital services fully integrated with the Energy Data Space. 20 data-driven services described in this the alpha version (D6.1). **For the beta version (D6.2), we expect that at least 12 energy services are integrated with the Energy Data Space.**
- ≥ 3 Digital Twins. **In D6.1: 3 digital were described and available in its alpha version.**

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

8 References

Bisschop, J. (2006). AIMMS optimization modeling. Lulu. com.

Bouno, T., Yuji, T., Hamada, T., & Hideaki, T. (2005). Failure forecast diagnosis of small wind turbine using acoustic emission sensor. KIEE International Transactions on Electrical Machinery and Energy Conversion Systems, 5(1), 78-83.

Campoverde-Vilela, L., Feijóo, M. D. C., Vidal, Y., Sampietro, J., & Tutivén, C. (2023). Anomaly-based fault detection in wind turbine main bearings. Wind Energy Science Discussions, 2023, 1-29.

Elasha, F., Shanbr, S., Li, X., & Mba, D. (2019). Prognosis of a wind turbine gearbox bearing using supervised machine learning. Sensors, 19(14), 3092.

Encalada-Dávila, Á., Puruncajas, B., Tutivén, C., & Vidal, Y. (2021). Wind turbine main bearing fault prognosis based solely on scada data. Sensors, 21(6), 2228.

Gonçalves, C., Bessa, R. J., Pinson, P. (2021). A critical overview of privacy-preserving approaches for collaborative forecasting. International journal of Forecasting, 37(1), 322-342.

Goncalves, C., Bessa, R. J., Pinson, P. (2021). Privacy-preserving distributed learning for renewable energy forecasting. IEEE Transactions on Sustainable Energy, 12(3), 1777-1787.

Guo, P., Infield, D., & Yang, X. (2011). Wind turbine generator condition-monitoring using temperature trend analysis. IEEE Transactions on sustainable energy, 3(1), 124-133.

Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. Machine Learning, 20, 197-243.

Huangfu, Q., Hall, J. J. (2018). Parallelizing the dual revised simplex method. Mathematical Programming Computation, 10(1), 119-142.

Li, M., Wang, S., Fang, S., & Zhao, J. (2020). Anomaly detection of wind turbines based on deep small-world neural network. Applied Sciences, 10(4), 1243.

Liang, G., Su, Y., Chen, F., Long, H., Song, Z., & Gan, Y. (2020). Wind power curve data cleaning by image thresholding based on class uncertainty and shape dissimilarity. IEEE Transactions on Sustainable Energy, 12(2), 1383-1393.

Mello, J., de Lorenzo, C., Campos, F. A., Villar, J. (2023). Pricing and Simulating Energy Transactions in Energy Communities. Energies, 16(4), 1949.

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

Pujana, A., Esteras, M., Perea, E., Maqueda, E., & Calvez, P. (2023). Hybrid-Model-Based Digital Twin of the Drivetrain of a Wind Turbine and Its Application for Failure Synthetic Data Generation. Energies, 16(2), 861.

Rocha, R., Silva, R., Mello, J., Faria, S., Retorta, F., Gouveia, C., Villar, J. (2023). A Three-Stage Model to Manage Energy Communities, Share Benefits and Provide Local Grid Services. Energies, 16(3), 1143.

Silva, R., Faria, S., Moreno, A., Retorta, F., Mello, J., Villar, J. Villar (2023). Impact of transaction pricing mechanisms on energy community benefits sharing. 18th European Energy Market Conference.

Spirtes, P., & Glymour, C. (1991). An algorithm for fast recovery of sparse causal graphs. Social science computer review, 9(1), 62-72.

Tsamardinos, I., Brown, L. E., & Aliferis, C. F. (2006). The max-min hill-climbing Bayesian network structure learning algorithm. Machine Learning, 65, 31-78.

Tutivén, C., Vidal, Y., Insuasty, A., Campoverde-Vilela, L., & Achicanoy, W. (2022). Early fault diagnosis strategy for WT main bearings based on SCADA data and one-class SVM. Energies, 15(12), 4381.

Wang, C. (2016). Health Monitoring and Fault Diagnostics of Wind Turbines. Aalborg Universitetsforlag. Ph.d.-serien for Det Teknisk-Naturvidenskabelige Fakultet, Aalborg Universitet.

Zhao, H., Liu, H., Hu, W., & Yan, X. (2018). Anomaly detection and fault analysis of wind turbine components based on deep learning network. Renewable Energy, 127, 825-834.

Zhu, J., Yoon, J., He, D., Qiu, B., & Bechhoefer, E. (2013, June). Online condition monitoring and remaining useful life prediction of particle contaminated lubrication oil. In 2013 IEEE Conference on Prognostics and Health Management (PHM) (pp. 1-14). IEEE.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

9 Appendix

9.1 Energy community sizing with assets sharing

This appendix shows the current mathematical formulation to optimize the sizing of a Renewable Energy Community (REC). Briefly, at this phase, the model is able to compute:

- The total cost of the REC
- Energy trades among each member and its retailer
- Energy trades among REC members
- Net consumption behind-the-meter (to each installation/member)
- Self-consumption energy (to each installation/member)
- The size of each individual and shared PV system and its operation[52]
- The size of each individual storage system and its operation[1]
- Pool price of the REC or bilateral transactions price (dependent on the business model)

For the coming features, it is expected the implementation of:

- Trades between private grids and different voltages levels (considering dependent fees)
- Shared batteries and their respective business models
- Assets degradation and their curves linearization
- REC business models (e.g., add external investors)
- Economies of scale linearization for the assets (PV and storage systems)
- Installations without any associated members may be considered, with the aim of installing shared batteries and PV systems. These installations have a contract with the retailer with all associated costs (similar to other installations with associated contracted power, tariffs…), allowing the energy exchange with the main grid (retailer).

9.1.1 Assumptions

The formulation of the design optimization requires a previous definition of a set of assumptions related to the CER:

- REC has N members, called prosumers, who have a consumption profile and can also have PV generation and individual batteries. It allows the measurement of the optimal capacity to install individual and shared assets (PV systems and batteries)

[52] Production diagram, energy charge and discharge and its SOC respectively.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- Allowed access to the consumption profiles of all consumers, the production profiles of the panels, and the characteristics of the technology of the assets (e.g., pre-installed, the minimum and the maximum capacity to be installed)
- REC members can have different suppliers and tariffs. It is considered that the values of the tariffs of each member are known for the optimization
- In case of the REC generation being higher than REC consumption, it is always possible to sell extra energy to the main grid and in the opposite way it is always possible to buy energy from the main grid (retailer)
- There are no technical limitations on the electrical grid
- Electricity demand (consumption) is inelastic (it is a parameter). However, batteries can provide flexibility and change the net consumption profile
- There is a minimum amount of PV power and battery capacity that consumers must install. This value can be initially set to 0 and then to be adjusted according to the minimum values available on the market and/or according to the technological specs of the assets
- There is a maximum amount of PV power and battery capacity that prosumers can install, which may reflect a limit on installation areas and/ or budget.
- The total capacity of the assets is defined by the pre-installed capacity plus the capacity to be installed defined by the optimization model, and it is dependent on the chosen business
- Economies of scale are not considered in investment costs (example of economy of scale: the larger the panel size, the smaller the €/W invested), as they give rise to a non-linear problem
- The charging and discharging power of the batteries is limited to the specifications of the chosen inverter

9.1.2 Variables and Inputs

The following table shows the indices, the service parameters, input data and the decision variables required for this optimization problem.

Indices and sets

$t \in T$	Set of time intervals/slots over the operation period	-
$n \in N$	Set of REC prosumers	-
$m \in N \nexists n$	Set of REC prosumers excluding prosumer *n*	-
$k \in K$	Set of shared storage systems	-
$w \in W$	Set of shared PV systems	-

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Service parameters		
Δt	Length of each timestep	minutes
horizon	Total time horizon of the simulation	days
$P_n^{PV,min}$	Minimum power of PV system that must be installed by prosumer *n*	kW
$P_n^{PV,max}$	Maximum power of PV system to be installed by prosumer *n*	kW
$P_w^{PV,min}$	Minimum power of the shared PV system *w* that must be installed	kW
$P_w^{PV,max}$	Maximum power of the shared PV system *w* to be installed	kW
$E_n^{BN,min}$	Minimum nominal capacity of the battery that must be installed by prosumer *n*	kWh
$E_n^{BN,max}$	Maximum nominal capacity of the battery that must be installed by prosumer *n*	kWh
$E_k^{BN,min}$	Minimum nominal capacity of the shared battery *k* that must be installed	kWh
$E_k^{BN,max}$	Maximum nominal capacity of the shared battery *k* to be installed	kWh
Input data		
$P_n^{cc,max}$	Power limit for transactions at the Point of Common Coupling for prosumer *n*	kW
λ_n^{CONT}	Contracted power tariff for prosumer *n*	€/(kW.day)
$\lambda_{n,t}^{buy}$	Buying energy tariff for prosumer *n* during the interval *t*	€/kWh
$\lambda_{n,t}^{sell}$	Selling energy tariff for prosumer *n* during the interval *t*	€/kWh
$E_{n,t}^{C}$	Energy consumption diagram for prosumer *n* during the interval *t*	kWh
$CI_n^{PV,TOTAL}$	Total investment costs for PV system for prosumer *n* (includes maintenance costs and residual value)	€/kW
CI_n^{PV}	Investment costs for PV system for the prosumer *n* adjusted to one-time interval	€/kW
$P_n^{PVN,pre}$	Pre-installed power of PV system for prosumer *n*	kW
$f_{n,t}^{PV}$	Generation profile factor of PV system per prosumer *n* during *t* (depends on efficiency, solar irradiance, temperature, etc.)	%
$P_w^{PV,pre}$	Pre-installed power of the shared PV system *w*	kW
$f_{w,t}^{PV}$	Generation profile factor of the shared PV system *w* during *t* (depends on efficiency, solar irradiance, temperature, etc.)	-

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$CI_w^{PV,Total}$	Total investment costs for the shared PV system *w* (includes maintenance costs and residual value)	€/kW
CI_w^{PV}	Investment costs for the shared PV system *w* adjusted to one-time interval	€/kW
$CI_n^{B,TOTAL}$	Total investment costs for storage system for prosumer *n* (includes maintenance costs and residual value)	€/kWh
CI_n^{B}	Investment costs for storage system for prosumer *n* adjusted to one-time interval	€/kWh
$E_n^{BN,pre}$	Pre-installed nominal capacity of the battery for the prosumer *n*	kWh
$\widehat{SOC}_n^{min}$	Minimum *State-of-charge* for the battery of prosumer *n*	%
$\widehat{SOC}_n^{max}$	Maximum *State-of-charge* for the battery of prosumer *n*	%
$P_n^{BC,max}$	Maximum charging power for the battery of prosumer *n*	kW
$P_n^{BD,max}$	Maximum discharging power for the battery of prosumer *n*	kW
η_n^{BC}	Charge efficiency for the battery of prosumer *n*	%
η_n^{BD}	Discharge efficiency for the battery of prosumer *n*	%
$P_k^{BC,\,max}$	Maximum charging power for the shared battery *k*	kW
$P_k^{BD,\,max}$	Maximum discharging power for the shared battery *k*	kW
$E_k^{BN,pre}$	Pre-installed nominal capacity of shared battery *k*	kWh
SOC_k^{min}	Minimum *State-of-charge* for the shared battery *k*	%
SOC_k^{max}	Maximum *State-of-charge* for the shared battery *k*	%
$CI_k^{B,Total}$	Total investment costs for the shared storage system *k* (includes maintenance costs and residual value)	€/kWh
CI_k^{B}	Investment costs for the shared storage system *k* adjusted to one-time interval	€/kWh
$\beta_{n,w}^{PV}$	Ownership that prosumer *n* has of the shared PV system *w*	%
$\beta_{n,k}^{B}$	Ownership that prosumer *n* has of the shared storage system *k*	%
λ_n^{Grid}	Grid access tariffs (provided by DSO)	€/kWh
M	Big number (e.g., 1E4)	-
Decision Variables		
$E_{n,init}^{B}$	Energy stored in the battery of the consumer *n* at the beginning of the operation	kWh
$E_{n,t}^{B}$	Energy stored in the battery of the consumer *n* during *t*	kWh
P_n^{CONT}	Contracted power of prosumer *n*	kW

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$P_n^{PVN,total}$	Total installed power of the PV system for prosumer *n*	kW
P_n^{PVN}	PV system power installed for prosumer *n*	kW
$E_{n,t}^{PV}$	PV generation power by prosumer *n* during *t*	kWh
$E_n^{BN,total}$	Total nominal capacity of the storage system for prosumer *n*	kWh
E_n^{BN}	Nominal capacity installed of storage system for prosumer *n*	kWh
$SOC_{n,t}$	Battery *State-of-charge* of prosumer *n* during *t*	%
$E_{n,t}^{BC}$	Energy charged in the battery of prosumer *n* during *t*	kWh
$E_{n,t}^{BD}$	Energy discharged of the battery of prosumer *n* during *t*	kWh
$E_{n,t}^{SUP}$	Energy that prosumer *n* buys from the retailer during the interval *t*	kWh
$E_{n,t}^{SUR}$	Energy that prosumer *n* sells to the retailer during the interval *t*	kWh
$E_{n,t}^{CMET}$	Net consumption behind-the-meter of the prosumer *n* during *t*	kWh
$E_{n,t}^{CMET,+}$	Auxiliary variable to calculate $E_{n,t}^{SLC}$ $\begin{cases} E_{n,t}^{CMET,+} = 0, & E_{n,t}^{CMET} < 0 \\ E_{n,t}^{CMET,+} = E_{n,t}^{CMET}, & E_{n,t}^{CMET} \geq 0 \end{cases}$	kWh
$E_{w,t}^{PV}$	Energy produced by the shared PV system *w* during *t*	kWh
$E_{w,n,t}^{PV}$	Allocated energy from PV system *w* to prosumer *n* during *t*	kWh
$P_w^{PVN,Total}$	Total installed power of the shared PV system *w*	kW
P_w^{PVN}	PV system power installed for *w*	kW
$E_{k,init}^{B}$	Energy stored in the shared battery *k* at the beginning of the operation	kWh
$E_{n,k,init}^{B}$	Initial energy stored in the shared battery *k* available to prosumer *n*	kWh
$E_{k,t}^{BC}$	Charged energy in the battery *k* during *t*	kWh
$E_{k,t}^{BD}$	Discharged energy of the battery *k* during *t*	kWh
$E_{n,k,t}^{BC}$	Charged energy in the shared battery *k* by prosumer *n* during *t*	kWh
$E_{n,k,t}^{BD}$	Discharged energy of the shared battery *k* by prosumer *n* during *t*	kWh
$E_k^{BN,Total}$	Total nominal capacity of the shared storage system *k*	kWh
E_k^{BN}	Storage system capacity installed for *k*	kWh
$E_{k,t}^{B}$	Available energy in the shared battery *k* during *t*	kWh
$E_{n,k,t}^{B}$	Available energy in the shared battery *k* for prosumer *n* during *t*	kWh

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$SOC_{k,t}$	*State-of-charge* of the storage system *k* during *t*	%
$E_{n,t}^{SLC}$	Self-consumption energy by prosumer *n* during *t*	kWh
$E_{n,m,t}^{PUR}$	Energy that prosumer *n* purchases from prosumer *m* during *t*	kWh
$E_{n,m,t}^{SALE}$	Energy that prosumer *n* sales to prosumer *m* during *t*	kWh
$E_{n,m,t}^{INT}$	Auxiliary variable to calculate $E_{n,t}^{SLC}$ $\begin{cases} E_{n,m,t}^{INT} = 0, & E_{n,m,t}^{PUR} - E_{n,m,t}^{SALE} < 0 \\ E_{n,m,t}^{INT} = E_{n,m,t}^{PUR} - E_{n,m,t}^{SALE}, & E_{n,m,t}^{PUR} - E_{n,m,t}^{SALE} \geq 0 \end{cases}$	kWh
δ_t^{SLC}	Binary auxiliary variable for self-consumption energy	-
δ_t^{coef+}	Binary auxiliary variable for positive allocation coefficients	-

9.1.3 Objective Function

The objective function (OF) of this REC sizing optimization problem involves minimizing the global costs of the REC, which are the following:

- buying and selling energy by the prosumers from and to the retailers
- using networks for self-consumed energy coming from the shared PV and storage systems, and from trades in the internal markets (pool or bilateral trades, dependent on the chosen business model)
- Contracted power (the costs for consumed or injected power) of each prosumer based on existing power tariffs. The contracted power is a decision variable of the problem and its calculation depends on the established regulation of the country
- Investment (grouping financing, installation, operation, replacement and maintenance costs) in batteries and panels, depending on installed capacities. At a later stage, additional costs may be considered for the degradation resulting from the way resources are operated, for example, in the case of batteries

$$\min \sum_{t=1}^{T} \left(\sum_n \left[\lambda_{n,t}^{buy} \times E_{n,t}^{SUP} - \lambda_{n,t}^{sell} \times E_{n,t}^{SUR} + \lambda_t^{Grid} \times E_{n,t}^{SC} \right] \right) + \sum_n \left(\lambda_n^{CONT} \times P_n^{CONT} \times T \right) +$$
$$\sum_n \left(CI_n^{PV} \times P_n^{PVN} + CI_n^{B} \times E_n^{BN} \right) \times T + \left(\sum_w \left(CI_w^{PV} \times P_w^{PVN} \right) + \sum_k \left(CI_k^{B} \times E_k^{BN} \right) \right) \times T$$

This OF can be divided into four parts:

(1) Costs of buying and selling energy between the prosumers and their retailers, plus the cost of using networks for self-consumption through shared PV and storage systems, and trades into the REC market

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$$\sum_{t=1}^{T}\left(\sum_{n}[\lambda_{n,t}^{buy} \times E_{n,t}^{SUP} - \lambda_{n,t}^{sell} \times E_{n,t}^{SUR} + \lambda_{t}^{Grid} \times E_{n,t}^{SC}]\right)$$

(2) Contracted power costs for the time horizon under the analysis (*T*)

$$\sum_{n}(\lambda_{n}^{CONT} \times P_{n}^{CONT} \times T)$$

(3) Investment costs of the assets (PV and storage systems) that must be installed for the time horizon under the analysis (*T*)

$$\sum_{n}(CI_{n}^{PV} \times P_{n}^{PVN} + CI_{n}^{B} \times E_{n}^{BN}) \times T$$

(4) Investment costs of shared assets (PV and storage systems) that must be installed for the time horizon under the analysis (*T*)

$$\left(\sum_{w}(CI_{w}^{PV} \times P_{w}^{PVN}) + \sum_{k}(CI_{k}^{B} \times E_{k}^{BN})\right) \times T$$

9.1.4 Constraints

The following table shows the constraints under the REC sizing optimization problem.

$E_{n,t}^{CMET} = E_{n,t}^{C} + E_{n,t}^{BC} - E_{n,t}^{PV} - E_{n,t}^{BD}$	$\forall\, n \in N, t \in T$	(1)
$E_{n,t}^{SUP} + \sum_{k \in K} E_{n,w,t}^{PV} + \sum_{k \in K} E_{n,k,t}^{BD} + \sum_{m}(E_{n,m,t}^{PUR} - E_{n,m,t}^{SALE}) = E_{n,t}^{CMET} + E_{n,t}^{SUR} + \sum_{k \in K} E_{n,k,t}^{BC}$	$\forall\, n, m \in N, n \neq m,$ $k \in K, w \in W,$ $t \in T$	(2)
$\frac{E_{n,t}^{SUP}}{\Delta t} \leq P_{n}^{cc,max}$	$\forall\, n \in N, t \in T$	(3)
$\frac{E_{n,t}^{SUR}}{\Delta t} \leq P_{n}^{cc,max}$	$\forall\, n \in N, t \in T$	(4)
$P_{n}^{PVN} = P_{n}^{PVN,Total} - P_{n}^{PVN,pre}$	$\forall\, n \in N$	(5)
$E_{n,t}^{PV} = f_{n,t}^{PV} * P_{n}^{PVN,Total} * \Delta t$	$\forall\, n \in N, t \in T$	(6)
$P_{n}^{PVN,min} \leq P_{n}^{PVN} \leq P_{n}^{PVN,max}$	$\forall\, n \in N$	(7)

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$E_n^{BN} = E_n^{BN,Total} - E_n^{BN,pre}$	$\forall\, n \in N$	(8)
$E_n^{BN,min} \leq E_n^{BN} \leq E_n^{BN,max}$	$\forall\, n \in N$	(9)
$\frac{E_{n,t}^{BC}}{\Delta t} \leq P_n^{BC,max}$	$\forall\, n \in N, t \in T$	(10)
$\frac{E_{n,t}^{BD}}{\Delta t} \leq P_n^{BD,max}$	$\forall\, n \in N, t \in T$	(11)
$E_{n,t}^{B} = E_{n,init}^{B} + \left(E_{n,t}^{BC} \times \eta_n^{BC} - \frac{E_{n,t}^{BD}}{\eta_n^{BD}}\right)$	$\forall\, n \in N, t = 1$	(12)
$E_{n,t}^{B} = E_{n,t-1}^{B} + \left(E_{n,t}^{BC} \times \eta_n^{BC} - \frac{E_{n,t}^{BD}}{\eta_n^{BD}}\right)$	$\forall\, n \in N, t \in]1, T]$	(13)
$SOC_{n,t} = \frac{E_{n,t}^{B}}{E_{n,t}^{BN,Total}} \times 100\%$	$\forall\, n \in N, t \in T$	(14)
$SOC_n^{min} \leq SOC_{n,t} \leq SOC_n^{max}$	$\forall\, n \in N, t \in T$	(15)
$P_n^{CONT} \geq \left\lvert \frac{E_{n,t}^{CMET}}{\Delta t} \right\rvert$	$\forall\, n \in N, t \in T$	(16)
$P_w^{PVN} = P_w^{PVN,Total} - P_w^{PVN,pre}$	$\forall\, w \in W$	(17)
$E_{w,t}^{PV} = f_{w,t}^{PV} * P_w^{PVN,Total} * \Delta t$	$\forall\, w \in W, t \in T$	(18)
$E_{w,n,t}^{PV} = E_{w,t}^{PV} \times \beta_{n,w}^{PV}$	$\forall\, n \in N, w \in W$ $t \in T$	(19)
$P_w^{GN,min} \leq P_w^{GN} \leq P_w^{GN,max}$	$\forall\, w \in W$	(20)
$\frac{E_{k,t}^{BC}}{\Delta t} \leq P_k^{BC,max}$	$\forall\, k \in K, t \in T$	(21)
$\frac{E_{k,t}^{BD}}{\Delta t} \leq P_k^{BD,max}$	$\forall\, k \in K, t \in T$	(22)
$E_k^{BN} = E_k^{BN,Total} - E_k^{BN,pre}$	$\forall\, k \in K$	(23)
$E_k^{BN,min} \leq E_k^{BN} \leq E_k^{BN,max}$	$\forall\, k \in K$	(24)
$E_{k,t}^{B} = E_{k,init}^{B} + \left(E_{k,t}^{BC} \times \eta_k^{BC} - \frac{E_{k,t}^{BD}}{\eta_k^{BD}}\right)$	$\forall\, k \in K, t = 1$	(25)
$E_{k,t}^{B} = E_{k,t-1}^{B} + \left(E_{k,t}^{BC} \times \eta_k^{BC} - \frac{E_{k,t}^{BD}}{\eta_k^{BD}}\right)$	$\forall\, k \in K, t \in]1, T]$	(26)

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$E^{BC}_{k,t} = \sum_{n} E^{BC}_{n,k,t}$	$\forall\, n \in N, t \in T$ $k \in K$	(27)
$E^{BD}_{k,t} = \sum_{n} E^{BD}_{n,k,t}$	$\forall\, n \in N, t \in T$ $k \in K$	(28)
$SOC_{k,t} = \frac{E^{B}_{k,t}}{E^{BN,Total}_{k,t}} \times 100\%$	$\forall\, k \in K, t \in T$	(29)
$SOC^{min}_{k} \leq SOC_{k,t} \leq SOC^{max}_{k}$	$\forall\, k \in K, t \in T$	(30)
$\sum_{n,m} E^{PUR}_{n,m,t} - \sum_{n,m} E^{SALE}_{n,m,t} = 0$	$\forall\, n,m \in N, n \neq m$ $t \in T$	(31)
$E^{PUR}_{n,m,t} - E^{SALE}_{m,n,t} = 0$	$\forall\, n,m \in N, n \neq m$ $t \in T$	(32)
$E^{SLC}_{n,t} = min\left(max\left(E^{CMET}_{n,t}, 0\right), max\left(\sum_{m} \left(E^{PUR}_{n,m,t} - E^{SALE}_{n,m,t}\right) + \sum_{w} E^{PV}_{w,n,t} + \sum_{k} \left(E^{BD}_{n,k,t} - E^{BC}_{n,k,t}\right), 0 \right)\right)$	$\forall\, n,m \in N, n \neq m,$ $t \in T, w \in W, k \in K$	(33)
$E^{CMET,+}_{n,t} \geq E^{CMET}_{n,t}$	$\forall\, n \in N, t \in T$	(33.1)
$E^{INT}_{n,m,t} \geq E^{PUR}_{n,m,t} - E^{SALE}_{m,n,t}$	$\forall\, n,m \in N, n \neq m$ $t \in T$	(33.2)
$E^{SLC}_{n,t} \geq E^{CMET,+}_{n,t} - \widehat{M}\left(1 - \hat{\delta}^{SLC}_{t}\right)$	$\forall\, n \in N, t \in T$	(33.3)
$E^{SLC}_{n,t} \geq \sum_{m} E^{INT}_{n,m,t} + \sum_{w} E^{PV}_{w,n,t} + \sum_{k} E^{BD}_{n,k,t} - \widehat{M}\left(\hat{\delta}^{SLC}_{t}\right)$	$\forall\, n,m \in N, n \neq m,$ $t \in T, w \in W, k \in K$	(33.4)
$\sum_{m} \left(E^{SALE}_{n,m,t} - E^{PUR}_{n,m,t}\right) - \sum_{w} E^{PV}_{w,n,t} - \sum_{k} \left(E^{BD}_{n,k,t} - E^{BC}_{n,k,t}\right) \leq max\left(-E^{CMET}_{n,t}, 0\right)$	$\forall\, n,m \in N, n \neq m,$ $t \in T, w \in W, k \in K$	(34)
$\sum_{m} \left(E^{SALE}_{n,m,t} - E^{PUR}_{n,m,t}\right) - \sum_{w} E^{PV}_{w,n,t} - \sum_{k} \left(E^{BD}_{n,k,t} - E^{BC}_{n,k,t}\right) \leq -E^{CMET}_{n,t} + \widehat{M}\left(\hat{\delta}^{coef+}_{t}\right)$	$\forall\, n,m \in N, n \neq m,$ $t \in T, w \in W, k \in K$	(34.1)
$\sum_{m} \left(E^{SALE}_{n,m,t} - E^{PUR}_{n,m,t}\right) - \sum_{w} E^{PV}_{w,n,t} - \sum_{k} \left(E^{BD}_{n,k,t} - E^{BC}_{n,k,t}\right) \leq \widehat{M}\left(1 - \hat{\delta}^{coef+}_{t}\right)$	$\forall\, n,m \in N, n \neq m,$ $t \in T, w \in W, k \in K$	(34.2)

Where:

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

(1) represents the net consumption behind-the-meter of the prosumer n, during the interval t, which is given by the sum of the consumption and battery charging subtracted by the PV generation and battery discharging. If net consumption is positive, the prosumer is consuming more than it produces in that period. If net consumption is negative, the prosumer has a production surplus.

(2) represents the energy equilibrium for the prosumer n during the interval t. which is established by equalizing the sum of behind-the-meter net consumption, surplus energy and energy charged in the shared battery with the sum of energy supplied by the retailer, allocated energy from the shared PV system, energy discharged from the battery and the energy balance that results from the internal market (energy purchases subtracting the energy sales).

(3), (4) represent the power limit for transactions at the Point of Common Coupling for prosumer n. That states the energy that prosumer n transactions with the retail (buy or sell) during the interval t.

(5) represents the PV system capacity that must be installed by the prosumer n, which is equal to the total power subtracted by the pre-installed power. The total power is the optimal power that each prosumer should have to achieve the minimization stated in the objective function.

(6) represents the energy produced by the PV system of the prosumer n, which is equal to the total installed power multiplied by the production factor and the duration of the time interval.

(7) establishes the range between the minimum and maximum capacity for the PV system that prosumer n must install. These minimum and maximum limits depend on the chosen business model and can, for example, take the value of 0 or reflect the maximum available area for PV system installation.

(8) represents the battery capacity that must be installed by the prosumer n, which is equal to the total capacity subtracted by the pre-installed capacity. The total capacity is the optimal capacity that each prosumer should have to achieve the minimization stated in the objective function.

(9) establishes the range between the minimum and maximum for the battery capacity that the prosumer n must install. These minimum and maximum limits depend on the chosen business model and can, for example, take the

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

value of 0 or reflect the maximum available area for the storage system installation.

(10), (11) represent the charging and discharging power limit of the battery of prosumer n, allowed by the maximum power of the inverter.

(12) establishes that the energy stored in the battery of the prosumer *n* during the first time period *t* is equal to the energy stored in the initial period (t = 0) plus the energy charged subtracted by the energy discharged during this time interval.

(13) establishes that the energy stored in the battery of prosumer *n* during the period *t* is equal to the energy stored in the previous period (t-1) plus the energy charged subtracted by the energy discharged during this time interval.

(14) represents the state-of-charge of the battery of prosumer *n* during period *t*, which is equal to the energy in the battery during that period divided by the total capacity of the battery.

(15) establishes the range between the minimum and maximum state-of-charge for the battery of the prosumer *n* to ensure a safe battery operation.

(16) establishes that the contracted power of the prosumer *n* must be higher or equal to the absolute value of the net consumption during the time horizon under the analysis.

(17) represents the capacity of the shared PV system *w* that must be installed, which is equal to the total power subtracted by the pre-installed power. The total power is the optimal power to achieve the minimization stated in the objective function.

(18) represents the energy produced by the shared PV system *w*, which is equal to the total installed power multiplied by the production factor and the duration of the time interval.

(19) represents the energy produced by the shared PV system *w* allocated to prosumer *n*, which is established by the ownership that prosumer *n* has

(20) establishes the range between the minimum and maximum capacity for the shared PV system *w* that must be installed. These minimum and maximum limits depend on the chosen business model and can, for example, take the value of 0 or reflect the maximum available area for PV system installation.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

(21), (22) represent the charging and discharging power limit of the shared battery *k*, allowed by the maximum power of the inverter.

(23) represents the capacity of the shared battery *k* that must be installed, which is equal to the total capacity subtracted by the pre-installed capacity. The total capacity is the optimal capacity that each prosumer should have to achieve the minimization stated in the objective function.

(24) establishes the range between the minimum and maximum capacity for the shared battery *k* that must be installed. These minimum and maximum limits depend on the chosen business model and can, for example, take the value of 0 or reflect the maximum available area for the storage system installation.

(25) establishes that the energy stored in the shared battery *k* during the first time period *t* is equal to the energy stored in the initial period (t=0) plus the energy charged subtracted by the energy discharged during this time interval.

(26) establishes that the energy stored in the shared battery *k* during the period *t* is equal to the energy stored in the previous period (t-1) plus the energy charged subtracted by the energy discharged during this time interval.

(27) represents the energy charged in the shared battery *k* during the period *t* that is equal to the sum of the energy charged by each prosumer *n* during this time interval

(28) represents the energy discharged from the shared battery *k* during the period *t* that is equal to the sum of the energy discharged by each prosumer *n* during this time interval.

(29) represents the energy charged in the shared battery *k* during the period *t*, which is equal to the sum of the energy charged by each prosumer *n* during this time interval

(30) establishes the range between the minimum and maximum state-of-charge for the shared battery *k* of the prosumer *n* to ensure a safe battery operation

(31) represents the REC pool market, where the sum of all purchases must equal to the sum of all energy sales transacted internally among the prosumers during the period *t*.

(32) represents the bilateral transactions market of REC, where the energy purchased by prosumer *n* from *m* must be equal to the energy sold by prosumer *m* to *n* during period *t*.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

(33) represents the self-consumption energy of the prosumer *n* during the period *t*, which is the minimum between its net consumption, when positive, and its allocated energy, when positive.

(33.1) establishes the value for the auxiliary variable that helps to calculate $\boldsymbol{E_{n,t}^{SLC}}$ described in (33), which is equal to zero if its net consumption is negative or takes the value of the net consumption if it is equal to zero or higher during *t*.

(33.2) establishes the value for the auxiliary variable that helps to calculate $\boldsymbol{E_{n,t}^{SLC}}$ described in (33), which is equal to zero if its sales are higher than its purchases or equal to the value of its purchases subtracted by its sales during *t*.

(33.3), (33.4) these constraints allow to solve the problems with the utilization of minimum and maximum when implementing the equation (33) in Python.

(34) represents the imposition of the positive allocation coefficients for the prosumer *n* during the period *t*, which establishes that is not allowed to sell more energy than what it has allocated

(34.1), (34.2) these constraints allow to solve the problems with the utilization of minimum and maximum when implementing the equation (34) in python

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

9.2 Flexibility modelling of thermoelectric water heaters

9.2.1 Variables and Inputs

Indices and sets

$t \in T$	Set of time intervals/slots over the operation period	-

Input parameters

Δt	Length of time intervals	h
$\widehat{EWH}^{power}$	Heating power of the EWH	kW
$\widehat{EWH}^{height}$	Height of the EWH	cm
$\widehat{EWH}^{cap}$	Internal water capacity of the EWH	l
$\widehat{EWH}^{\max}$	Maximum allowable water temperature of the EWH	°C
$\widehat{EWH}^{min}$	Minimum allowable water temperature of the EWH	°C
$\hat{c}$	Specific heat capacity of water (4.18 J/g°C)	J/g°C
$\widehat{HC}$	Overall heat transfer coefficient of the EWH	kW/(m²K)
$\widehat{temp}^{set}$	User-defined comfort temperature for hot water usage	°C
$\widehat{temp}^{amb}$	EWH room ambient temperature	°C
$\widehat{M}$	Sufficiently big number (e.g., 1E3 if EWH's nominal capacity is in the tens of kWh)	-
$\widehat{flowrate}$	Water usage flow rate (fixed/constant value).	kg/min
$\widehat{tariff}$	Selection between energy price tariffs between simple (1) or bi-hourly (2), etc.	-
$\hat{\delta}_t^{use}$	Binary parameter indicating periods of hot water usage (1 = using)	-

Variable inputs

$\widehat{EWH}^{StartTemp}$	EWH internal water temperature at the beginning	°C
$\widehat{EWH}^{area}$	EWH Surface Area	m²
W^{init}	Total energy balance of the EWH at the beginning	kWh
W^{max}	Maximum energy balance of the EWH	kWh
W^{init}	Total energy balance of the EWH at the beginning	kWh
$reg_m^{belowSet}$	Auxiliary $\boldsymbol{m}$ linearization parameter from *belowSet* regressor	-
$reg_b^{belowSet}$	Auxiliary b linearization parameter from *belowSet* regressor	-

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$\boldsymbol{reg_m^{aboveSet}}$	Auxiliary $\boldsymbol{m}$ linearization parameter from *aboveSet* regressor	-
$\boldsymbol{reg_b^{aboveSet}}$	Auxiliary $\boldsymbol{b}$ linearization parameter from *aboveSet* regressor	-
Variables		
$\boldsymbol{temp_t^{EWH}}$	Temperature of water at EWH outlet at the beginning of time interval $\boldsymbol{t}$.	°C
$\boldsymbol{W_t^{tot}}$	Total energy balance of prosumer's EWH at time interval $\boldsymbol{t}$.	kWh
$\boldsymbol{W_t^{in}}$	Energy into the prosumer's EWH at time interval $\boldsymbol{t}$	kWh
$\boldsymbol{W_t^{loss}}$	Thermal energy losses at time interval $\boldsymbol{t}$	kWh
$\boldsymbol{W_t^{mix}}$	Thermal energy stored in the EWH after usage and mixing with inlet water at time interval $\boldsymbol{t}$	kWh
$\boldsymbol{\delta_t^{in}}$	Binary variable for EWH operation status (1 = ON, 0 = OFF)	-
$\boldsymbol{\delta_t^{aux}}$	Binary Variable for if-else in constraint (7) at time interval $\boldsymbol{t}$	-
$\boldsymbol{penalty_t^{comf}}$	Penalty cost associated with water temperature reaching below comfort at time interval $\boldsymbol{t}$	-
$\boldsymbol{cost_t^{use}}$	Total period cost of that specific energy usage at time interval $\boldsymbol{t}$	€
$\boldsymbol{price_t^{net}}$	Network pricing of energy usage at time interval $\boldsymbol{t}$	€
$\boldsymbol{tariff^m}$	Energy daily tariff for each $\boldsymbol{m}$ price sets	€

9.2.2 Objective Function

$$min \sum_{t}^{T} cost_t^{use} \cdot a_1 + penalty_t^{comf} \cdot a_2$$

The objective of this problem will be to minimize the operating costs of using hot water, but always wanting to ensure that the comfort threshold is respected. However, as described in constraint (6), a penalty is added that allows the operation, even if this limit is not respected. In this sense, this parameter with a positive sign is adding an additional "cost", due to the use of

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

water below comfort. The values of parameters a_1 and a_2 must be adjusted, according to the importance that the user may give to comfort or price (default as $a_1 = 100, a_2 = 100$).

9.2.3 Constraints

$$\begin{cases} W_t^{tot} = W^{init}, & t = 0 \\ W_t^{tot} = W_{t-1}^{water} + W_{t-1}^{in} + W_{t-1}^{loss}, & t > 0 \end{cases} \tag{1}$$

$$W_t^{in} = \widehat{EWH}^{power} \cdot \Delta t \cdot \delta_t^{in} \qquad \forall\, t \in T \tag{2}$$

$$cost_t^{use} = \delta_t^{in} \cdot \widehat{EWH}^{power} \cdot \Delta t \cdot price_t^{use} + tariff^m \cdot \frac{\Delta t}{24} \qquad \forall\, t \in T \tag{3}$$

$$\begin{cases} temp_t^{EWH} = \widehat{EWH}^{StartTemp}, & t = 0 \\ temp_t^{EWH} = \dfrac{W_t^{tot} \cdot 3600}{\widehat{EWH}^{cap} \cdot \hat{c}}, & t > 0 \end{cases} \tag{4}$$

$$W_t^{loss} = \widehat{EWH}^{height} \cdot \widehat{EWH}^{area} \cdot \left(temp_t^{EWH} - \widehat{temp}^{amb}\right) \cdot \Delta t \qquad \forall\, t \in T \tag{5}$$

$$W_{t-1}^{tot} \geq \widehat{temp}^{set} \cdot \widehat{EWH}^{cap} \cdot \frac{\hat{c}}{3600} - penalty_t^{comf} \qquad \hat{\delta}_t^{use} - \hat{\delta}_{t-1}^{use} \tag{6}$$

$$\begin{cases} temp_t^{EWH} \geq \widehat{temp}^{set} - \widehat{M} \cdot (1 - \delta_t^{aux}), & \hat{\delta}_t^{use} = 1 \\ temp_t^{EWH} \leq \widehat{temp}^{set} + \widehat{M} \cdot \delta_t^{aux}, & \hat{\delta}_t^{use} = 1 \\ W_t^{mix} \geq \left(reg_m^{belowSet} \cdot temp_t^{EWH} + reg_b^{belowSet}\right) - \widehat{M} \cdot \delta_t^{aux}, & \hat{\delta}_t^{use} = 1 \\ W_t^{mix} \leq \left(reg_m^{belowSet} \cdot temp_t^{EWH} + reg_b^{belowSet}\right) + \widehat{M} \cdot \delta_t^{aux}, & \hat{\delta}_t^{use} = 1 \\ W_t^{mix} \geq \left(reg_m^{aboveSet} \cdot temp_t^{EWH} + reg_b^{aboveSet}\right) - \widehat{M} \cdot (1 - \delta_t^{aux}), & \hat{\delta}_t^{use} = 1 \\ W_t^{mix} \leq \left(reg_m^{aboveSet} \cdot temp_t^{EWH} + reg_b^{aboveSet}\right) + \widehat{M} \cdot (1 - \delta_t^{aux}), & \hat{\delta}_t^{use} = 1 \\ W_t^{mix} = temp_t^{EWH} \cdot \widehat{EWH}^{cap} \cdot \dfrac{\hat{c}}{3600}, & \hat{\delta}_t^{use} = 0 \end{cases} \tag{7}$$

- (1) represents the total energy stored within the EWH in each period. For the first instant, it is equal to the value stipulated as initial (or equal to energy associated with the tank filled with water at inlet temperature). For the remaining periods, it depends on the total energy resulting from mixing water (7), the energy gain from heating (2), and the energy thermal losses (5).

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

- (2) represents additional energy introduced by the heating element of the EWH. It depends on the power of the element, the length of the time period and if the EWH is operational or not.
- (3) represents the total cost of operating the EWH, depending on the energy cost, the selected daily tariff, the power of the element, the length of the time period and if the EWH is operational or not.
- (4) represents the temperature value of the internal water of the EWH. For the initial instant, it is equivalent to the established input, while for the following instants, it establishes a conversion to equivalent temperature through the value of the accumulated energy (1)
- (5) represents the total thermal losses of the EWH through its surface, depending on the length of the time period.
- (6) represents the need for the internal water temperature of the EWH after use to be equal to or greater than the threshold defined by the user. It uses an auxiliary variable that allows the model to operate below the comfort threshold, but adds a penalty operating cost, which in turn is added to the objective function with a positive term.
- (7) Definition of the variable that represents the accumulated energy after using hot water, and after internal mixing with inlet water. As it is a conditional (if-else) and non-linear constraint, its formulation is built in such a way that it is programmable. Please refer to the **Mathematical Adjustments** section for a detailed description.

9.2.4 Mathematical Adjustments

1. **Mixing Fluids**

The final temperature when mixing fluids is calculated below.

$$t_{out} = \frac{m_1 \cdot c_1 \cdot t_1 + m_2 \cdot c_2 \cdot t_2}{m_1 \cdot c_1 + m_2 \cdot c_2}$$

2. **Temperature to Energy**

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

The final energy (kWh) is calculated below.

$$w_{out} = \frac{t_{out} \cdot m_{out} \cdot c}{3600}$$

A. Estimate the total hot water flow rate.

For a single hot water usage ($\widehat{flow}^{Total}$), the final usage temperature ($\widehat{temp}^{set}$) is given by the mixture of both EWH hot water flow ($flow_t^{EWH}$) at the current EWH temperature ($temp_t^{EWH}$) and Inlet flow ($flow_t^{Inlet}$) at the respective temperature ($\widehat{temp}^{inlet}$).

The objective is to estimate the necessary EWH flow rate that, at the current EWH temperature, should be mixed with inlet water. The inlet water flow should represent the difference between the usage total flow rate and the estimated EWH flow rate.

Assuming,

$$flow_t^{Inlet} = \widehat{flow}^{Total} - flow_t^{EWH}$$

$$c^{EWH} = c^{Inlet} = \hat{c}$$

Using the mixing fluid expression, gives:

$$\widehat{temp}^{set} = \frac{flow_t^{EWH} \cdot \hat{c} \cdot temp_t^{EWH} + \left(\widehat{flow}^{Total} - flow_t^{EWH}\right) \cdot \hat{c} \cdot \widehat{temp}^{inlet}}{flow_t^{EWH} \cdot \hat{c} + \left(\widehat{flow}^{Total} - flow_t^{EWH}\right) \cdot \hat{c}}$$

Solving to $flow_t^{EWH}$:

$$flow_t^{EWH} = \widehat{flow}^{Total} \frac{\widehat{temp}^{set} - \widehat{temp}^{inlet}}{temp_t^{EWH} - \widehat{temp}^{inlet}}$$

B. Estimate the EWH internal temperature after usage.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

After a single hot water usage, the final EWH internal temperature ($temp_t^{mix}$) can be estimated using the fluid mixture expression. In this scenario, the final water mixture should address the remainder of the hot water inside the EWH after the usage ($\widehat{EWH}^{cap} - flow_t^{EWH}$) at the current temperature ($temp_{EWH}$) being mixed with the necessary inlet water to fill the amount that was used ($flow_t^{EWH}$) at inlet temperature ($\widehat{temp}^{inlet}$). Using the mixing fluid expression, gives:

$$temp_t^{mix} = \frac{\left(\widehat{EWH}^{cap} - flow_t^{EWH}\right) \cdot c \cdot temp_t^{EWH} + flow_t^{EWH} \cdot c \cdot \widehat{temp}^{inlet}}{\left(\widehat{EWH}^{cap} - flow_t^{EWH}\right) \cdot c + flow_t^{EWH} \cdot c}$$

Simplifying:

$$temp_t^{mix} = \frac{\left(\widehat{EWH}^{cap} - flow_t^{EWH}\right) \cdot temp_t^{EWH} + flow_t^{EWH} \cdot \widehat{temp}^{inlet}}{\widehat{EWH}^{cap}}$$

C. Estimating the total stored energy after a hot water usage

After estimating an expression to find the final temperature of the water inside the EWH after a single hot water usage, the $flow_{EWH}$ value can be replaced by the expression found in (A), thus resulting in an expression that relies only on the internal EWH temperature to estimate the total stored energy after a single hot water usage. By using the expression that allows to convert temperature to stored energy, the amount of stored energy (kWh), after a hot water usage, is given by:

$$W_t^{mix} = \frac{temp_t^{mix} \cdot \widehat{EWH}^{cap} \cdot \hat{c}}{3600}$$

Replacing the $temp_t^{mix}$ with the respective expression gives:

$$W_t^{mix} = \frac{\left(\frac{\left(\widehat{EWH}^{cap} - flow_t^{EWH}\right) \cdot temp_t^{EWH} + flow_t^{EWH} \cdot \widehat{temp}^{inlet}}{\widehat{EWH}^{cap}}\right) \cdot \widehat{EWH}^{cap} \cdot \hat{c}}{3600}$$

That simplifies to:

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

$$W_t^{mix} = \frac{\left(\left(\widehat{EWH}^{cap} - flow_t^{EWH}\right) \cdot temp_t^{EWH} + flow_t^{EWH} \cdot \widehat{temp}^{inlet}\right) \cdot \hat{c}}{3600}$$

Replacing $flow_t^{EWH}$ with the expression found in (A), when $temp_t^{EWH} \geq \widehat{temp}^{set}$:

$$W_t^{mix\,|\,temp_t^{EWH} \geq \widehat{temp}^{set}} = \frac{\left(\left(\widehat{EWH}^{cap} - \widehat{flow}^{Total} \frac{\widehat{temp}^{set} - \widehat{temp}^{inlet}}{temp_t^{EWH} - \widehat{temp}^{inlet}}\right) \cdot temp_t^{EWH} + \widehat{flow}^{Total} \frac{\widehat{temp}^{set} - \widehat{temp}^{inlet}}{temp_t^{EWH} - \widehat{temp}^{inlet}} \cdot \widehat{temp}^{inlet}\right) \cdot \hat{c}}{3600}$$

However, when $temp_t^{EWH} < \widehat{temp}^{set}$, the EWH flow rate must be equal to the total flow rate, guaranteeing that the output temperature is actual temperature inside of the EWH, and hence:

$$W_t^{mix\,|\,temp_t^{EWH} < \widehat{temp}^{set}} = \frac{\left(\left(\widehat{EWH}^{cap} - \widehat{flow}^{Total}\right) \cdot temp_t^{EWH} + \widehat{flow}^{Total} \cdot \widehat{temp}^{inlet}\right) \cdot \hat{c}}{3600}$$

D. Use-case Scenario & Linearization

Having described the two expressions, the next step involves the creation of two linear expressions, for each of the time intervals described above. To do so, all non-repeating $temp_{Inlet}$ input values are used, individually paired with $temp_t^{EWH}$ values from each of the intervals and for each of the expressions. For example, for a single observation where $\widehat{temp}^{inlet} = 15$°C, assuming $\widehat{temp}^{set} = 45°C$, and $\widehat{temp}^{Max|EWH} = 85$°C, 41 final W_t^{mix}values are calculated (with $temp_t^{EWH}$values from 45 up to and including 85). This procedure is repeated for each non-repeated $temp_{Inlet}$value, for each of the expressions, with the respective ranges of $temp_t^{EWH}$ values).

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

Table 45: Auxiliary dataset for calculating W(mix), for the two temp(EWH) levels

$EWH_{Capacity}$	flowTotal	$temp_{Inlet}$	$temp_{Set}$	$temp_{EWH}$	$flow_{EWH}$	$temp_{EWH\_Final}$	W_{EWH_Final}
100	80	15	45	15	80.0	15.0	**1.744**
100	80	15	45	16	80.0	15.2	**1.767**
...	...	...	...	...	...	...	**...**
100	80	15	45	43	80.0	20.6	**2.395**
100	80	15	45	44	80.0	20.8	**2.419**
100	80	16	45	15	80.0	15.8	**1.837**
100	80	16	45	16	80.0	16.0	**1.860**
...	...	...	...	...	...	...	**...**
100	80	16	45	43	80.0	21.4	**2.488**
100	80	16	45	44	80.0	21.6	**2.512**
100	80	15	45	45	80.0	21.0	**2.442**
100	80	15	45	46	77.4	22.0	**2.558**
...	...	...	...	...	...	...	**...**
100	80	15	45	84	34.8	60.0	**6.977**
100	80	15	45	85	34.3	61.0	**7.093**
100	80	16	45	45	80.0	21.8	**2.535**
100	80	16	45	46	77.3	22.8	**2.651**
...	...	...	...	...	...	...	**...**
100	80	16	45	84	34.1	60.8	**7.070**
100	80	16	45	85	33.6	61.8	**7.186**

Once this auxiliary dataset is created (see Table 45), trend curves must be created for each of the levels (see Figure 70). The expressions of these trend curves will be those associated with the restriction that allows calculating the stored energy value after using hot water. In this way, the two linear parameters (m and b) are calculated for each of the levels.

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

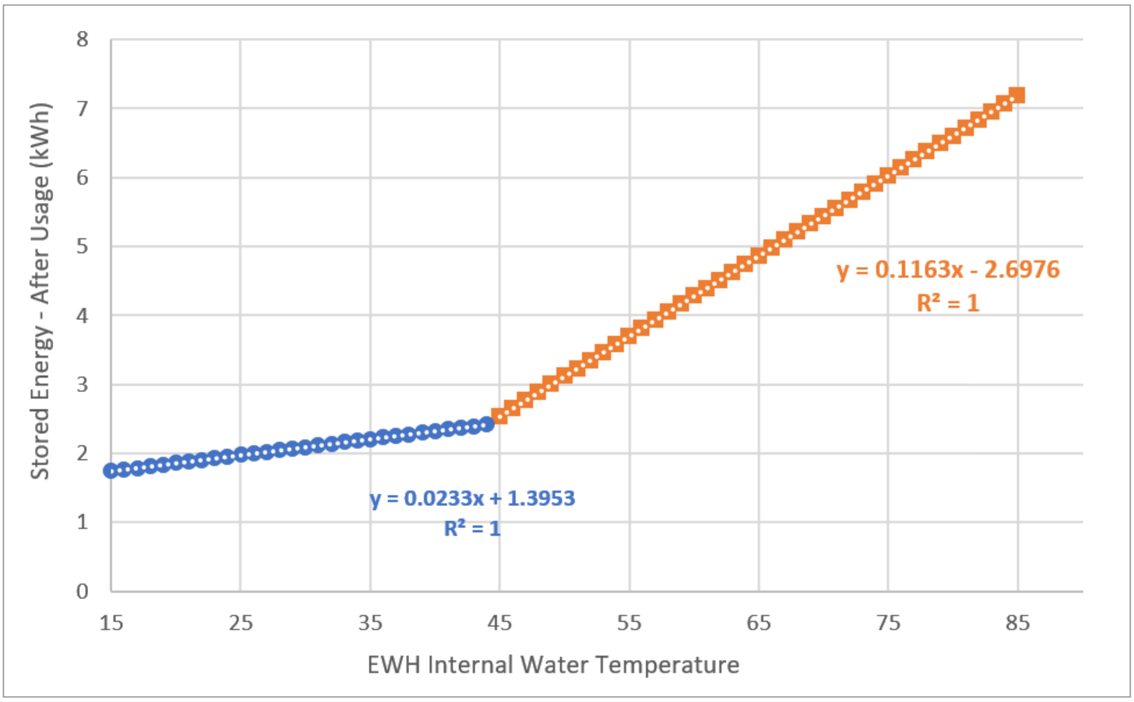


Figure 70: Trendlines for the two EWH temperature levels

With the new values of the parameters (m_1, b_1, m_2, b_2), it is possible to write the final restriction that allows to obtain the stored energy after a single use of hot water ($\hat{\delta}_t^{use} = 1$):

$$If\ \widehat{\delta}_t^{use} = 1:$$

$$If\ temp_t^{EWH} < \widehat{temp}^{set}$$

$$W_t^{mix} = m_1 \cdot temp_t^{EWH} + b_1$$

$$If\ temp_t^{EWH} \geq \widehat{temp}^{set}$$

$$W_t^{mix} = m_2 \cdot temp_t^{EWH} + b_2$$

When there is no hot water usage ($\hat{\delta}_t^{use} = 0$), the constraint is direct, and it uses an equivalent temperature to energy conversion using:

$$If\ \widehat{\delta}_t^{use} = 0:$$

$$W_t^{mix} = temp_t^{EWH} \cdot \widehat{EWH}^{cap} \cdot \frac{\hat{c}}{3600}$$

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

E. **Conditional Constraint**

As shown in (D), the formulation results in a set of conditional constraints. In this type of problem, if the condition is based on an input variable, where it is possible, a priori, to verify its value, it is possible to program the conditional restriction. If the condition is imposed by a decision variable, it is not possible to program the restriction, since, internally, the restrictions, for all instants of time, are written in a file. In this sense, as the conditional variable has not yet been assigned a value, it is not possible to implement it in a conditional constraint. In this problem, we have two degrees of condition: the first one, based on the input variable $\hat{\delta}_t^{use}$, can be introduced directly in the formulation. However, the second condition is based on the $temp_t^{EWH}$ variable, which, being a decision variable, prevents the implementation from being straightforward. In this sense, for this second part, it is necessary to transform the conditional constraint, through the introduction of sufficiently wide tolerance intervals. Through the formulations indicated in (Bisschop, 2006), the restrictions formulated in (D) are then transformed into:

$$If\ \widehat{\delta_t}^{use} = 1:$$

$$
\begin{gathered}
temp_t^{EWH} \geq \widehat{temp}^{set} - \widehat{M} \cdot (1 - \delta_t^{aux}) \\
temp_t^{EWH} \leq \widehat{temp}^{set} + \widehat{M} \cdot \delta_t^{aux} \\
W_t^{mix} \geq (reg_m^{belowSet} \cdot temp_t^{EWH} + reg_b^{belowSet}) - \widehat{M} \cdot \delta_t^{aux} \\
W_t^{mix} \leq (reg_m^{belowSet} \cdot temp_t^{EWH} + reg_b^{belowSet}) + \widehat{M} \cdot \delta_t^{aux} \\
W_t^{mix} \geq (reg_m^{aboveSet} \cdot temp_t^{EWH} + reg_b^{aboveSet}) - \widehat{M} \cdot (1 - \delta_t^{aux}) \\
W_t^{mix} \leq (reg_m^{aboveSet} \cdot temp_t^{EWH} + reg_b^{aboveSet}) + \widehat{M} \cdot (1 - \delta_t^{aux})
\end{gathered}
$$

$$If\ \widehat{\delta_t}^{use} = 0:$$

$$W_t^{mix} = temp_t^{EWH} \cdot \widehat{EWH}^{cap} \cdot \frac{\hat{c}}{3600}$$

F. **Comfort**

Since it is already possible to calculate the volume of energy stored after using hot water, the aim is now to ensure that the comfort temperature is respected. To this end, a restriction is added that indicates that, at the beginning of the period following the use of water, the stored energy of the water in the EWH is at least equal to the energy associated with the comfort

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

temperature. In this sense, the energy value associated with the comfort temperature is calculated using the temperature to energy expression, as follows:

$$W_{confort} = \frac{\widehat{temp}^{set} \cdot \widehat{EWH}^{cap} \cdot \hat{c}}{3600}$$

There is also need to use an auxiliary variable that allows the model to operate below the comfort threshold, but adding a penalty operating cost, which in turn is added to the objective function with a positive term. Hence, the final constraint:

$$If\ \delta_{usage}[t] - \delta_{usage}[t-1] < 0$$

$$W_{t-1}^{tot} \geq W_{confort} - penalty_t^{comf}$$

9.3 LV Grid topology as modelled using ESDL


```xml
<?xml version='1.0' encoding='UTF-8'?>
<esdl:EnergySystem xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:esdl="http://www.tno.nl/esdl" esdlVersion="v2207" version="5" name="dutch_feeder" id="dutch_feeder">
  <instance xsi:type="esdl:Instance" id="dutch_feederInstance" name="dutch_feederInstance">
    <area xsi:type="esdl:Area" id="TestArea" name="Area">
      <asset xsi:type="esdl:Bus" id="Bus1" name="Bus1">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04269403" lon="6.605515616000001"/>
        <port xsi:type="esdl:InPort" name="Bus1Phase1In" id="Bus1Phase1In" connectedTo="transformer1Out" carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus1Phase1Out" id="Bus1Phase1Out" connectedTo="line21Phase1In line15Phase1In" carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line15" assetType="gplkh_4_50_cusvm_4_6" name="line15" length="6.0">
        <port xsi:type="esdl:InPort" name="line15Phase1In" id="line15Phase1In" connectedTo="Bus1Phase1Out" carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line15Phase1Out" id="line15Phase1Out" connectedTo="Bus13Phase1In" carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.0427034" lon="6.60610646"/>
          <point xsi:type="esdl:Point" lat="52.04271974" lon="6.606248617"/>
        </geometry>
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus13" name="Bus13">
        <geometry xsi:type="esdl:Point" CRS="WGS84"
lat="52.042769740000004" lon="6.6061486170000006"/>
        <port xsi:type="esdl:InPort" name="Bus13Phase1In"
id="Bus13Phase1In" connectedTo="line15Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus13Phase1Out"
id="Bus13Phase1Out" connectedTo="line16Phase1In line30Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line16"
assetType="gplkh_4_25_curm_4_2p5" name="line16" length="2.0">
        <port xsi:type="esdl:InPort" name="line16Phase1In"
id="line16Phase1In" connectedTo="Bus13Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line16Phase1Out"
id="line16Phase1Out" connectedTo="Bus1301Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04271974"
lon="6.606248617"/>
          <point xsi:type="esdl:Point" lat="52.04264222"
lon="6.606229842"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus1301" name="Bus1301">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04269222"
lon="6.6061298420000005"/>
        <port xsi:type="esdl:InPort" name="Bus1301Phase1In"
id="Bus1301Phase1In" connectedTo="line16Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus1301Phase2Out"
id="Bus1301Phase2Out" connectedTo="EconnInuser4Phase2In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line21"
assetType="gplkh_4_50_cusvm_4_6" name="line21" length="4.0">
        <port xsi:type="esdl:InPort" name="line21Phase1In"
id="line21Phase1In" connectedTo="Bus1Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line21Phase1Out"
id="line21Phase1Out" connectedTo="Bus16Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.0427034"
lon="6.60610646"/>
          <point xsi:type="esdl:Point" lat="52.04260939"
lon="6.60603404"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus16" name="Bus16">
        <geometry xsi:type="esdl:Point" CRS="WGS84"
lat="52.042659390000004" lon="6.60593404"/>
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
        <port xsi:type="esdl:InPort" name="Bus16Phase1In"
id="Bus16Phase1In" connectedTo="line21Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus16Phase1Out"
id="Bus16Phase1Out" connectedTo="line28Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line28"
assetType="gplkh_4_50_cusvm_4_6" name="line28" length="7.0">
        <port xsi:type="esdl:InPort" name="line28Phase1In"
id="line28Phase1In" connectedTo="Bus16Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line28Phase1Out"
id="line28Phase1Out" connectedTo="Bus20Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04260939"
lon="6.60603404"/>
          <point xsi:type="esdl:Point" lat="52.04254011"
lon="6.605991125"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus20" name="Bus20">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04259011"
lon="6.605891125"/>
        <port xsi:type="esdl:InPort" name="Bus20Phase1In"
id="Bus20Phase1In" connectedTo="line28Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus20Phase1Out"
id="Bus20Phase1Out" connectedTo="line29Phase1In line44Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line29"
assetType="gplkh_4_25_curm_4_2p5" name="line29" length="37.0">
        <port xsi:type="esdl:InPort" name="line29Phase1In"
id="line29Phase1In" connectedTo="Bus20Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line29Phase1Out"
id="line29Phase1Out" connectedTo="Bus2001Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04254011"
lon="6.605991125"/>
          <point xsi:type="esdl:Point" lat="52.04252362"
lon="6.606114507000001"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus2001" name="Bus2001">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04257362"
lon="6.606014507000001"/>
        <port xsi:type="esdl:InPort" name="Bus2001Phase1In"
id="Bus2001Phase1In" connectedTo="line29Phase1Out"
carrier="Electricity"/>
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
        <port xsi:type="esdl:OutPort" name="Bus2001Phase3Out"
id="Bus2001Phase3Out"
connectedTo="EConnInuser10Phase3Inuser10Phase3In"
carrier="Electricity"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:ElectricityCable" id="line30"
assetType="gplkh_4_50_cusvm_4_6" name="line30" length="4.0">
        <port xsi:type="esdl:InPort" name="line30Phase1In"
id="line30Phase1In" connectedTo="Bus13Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line30Phase1Out"
id="line30Phase1Out" connectedTo="Bus21Phase1In"
carrier="Electricity"/>
        &lt;geometry xsi:type="esdl:Line" CRS="WGS84"&gt;
          <point xsi:type="esdl:Point" lat="52.04271974"
lon="6.606248617"/>
          <point xsi:type="esdl:Point" lat="52.04272139"
lon="6.606355906"/>
        &lt;/geometry&gt;
      &lt;/asset&gt;
      &lt;asset xsi:type="esdl:Bus" id="Bus21" name="Bus21"&gt;
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04277139"
lon="6.606255906"/>
        <port xsi:type="esdl:InPort" name="Bus21Phase1In"
id="Bus21Phase1In" connectedTo="line30Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus21Phase1Out"
id="Bus21Phase1Out" connectedTo="line31Phase1In"
carrier="Electricity"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:ElectricityCable" id="line31"
assetType="gplkh_4_6_curm_4_2p5" name="line31" length="3.0">
        <port xsi:type="esdl:InPort" name="line31Phase1In"
id="line31Phase1In" connectedTo="Bus21Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line31Phase1Out"
id="line31Phase1Out" connectedTo="Bus2101Phase1In"
carrier="Electricity"/>
        &lt;geometry xsi:type="esdl:Line" CRS="WGS84"&gt;
          <point xsi:type="esdl:Point" lat="52.04272139"
lon="6.606355906"/>
          <point xsi:type="esdl:Point" lat="52.04264222"
lon="6.6063264010000005"/>
        &lt;/geometry&gt;
      &lt;/asset&gt;
      &lt;asset xsi:type="esdl:Bus" id="Bus2101" name="Bus2101"&gt;
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04269222"
lon="6.606226401000001"/>
        <port xsi:type="esdl:InPort" name="Bus2101Phase1In"
id="Bus2101Phase1In" connectedTo="line31Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus2101Phase2Out"
id="Bus2101Phase2Out"
```

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
connectedTo="EConnInuser11Phase2Inuser11Phase2In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line44"
assetType="gplkh_4_50_cusvm_4_6" name="line44" length="1.0">
        <port xsi:type="esdl:InPort" name="line44Phase1In"
id="line44Phase1In" connectedTo="Bus20Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line44Phase1Out"
id="line44Phase1Out" connectedTo="Bus28Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04254011"
lon="6.605991125"/>
          <point xsi:type="esdl:Point" lat="52.04245765"
lon="6.60597235"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus28" name="Bus28">
        <geometry xsi:type="esdl:Point" CRS="WGS84"
lat="52.042507650000005" lon="6.60587235"/>
        <port xsi:type="esdl:InPort" name="Bus28Phase1In"
id="Bus28Phase1In" connectedTo="line44Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus28Phase1Out"
id="Bus28Phase1Out" connectedTo="line45Phase1In line102Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line45"
assetType="gplkh_4_6_curm_4_2p5" name="line45" length="40.0">
        <port xsi:type="esdl:InPort" name="line45Phase1In"
id="line45Phase1In" connectedTo="Bus28Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line45Phase1Out"
id="line45Phase1Out" connectedTo="Bus2801Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04245765"
lon="6.60597235"/>
          <point xsi:type="esdl:Point" lat="52.04245105"
lon="6.606076956"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus2801" name="Bus2801">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04250105"
lon="6.605976956"/>
        <port xsi:type="esdl:InPort" name="Bus2801Phase1In"
id="Bus2801Phase1In" connectedTo="line45Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus2801Phase3Out"
id="Bus2801Phase3Out"
connectedTo="EConnInuser18Phase3Inuser18Phase3In"
carrier="Electricity"/>
      </asset>
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
      <asset xsi:type="esdl:ElectricityCable" id="line102"
assetType="gplkh_4_50_cusvm_4_6" name="line102" length="3.0">
        <port xsi:type="esdl:InPort" name="line102Phase1In"
id="line102Phase1In" connectedTo="Bus28Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line102Phase1Out"
id="line102Phase1Out" connectedTo="Bus57Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04245765"
lon="6.60597235"/>
          <point xsi:type="esdl:Point" lat="52.04237353"
lon="6.605940163"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus57" name="Bus57">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04242353"
lon="6.605840163"/>
        <port xsi:type="esdl:InPort" name="Bus57Phase1In"
id="Bus57Phase1In" connectedTo="line102Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus57Phase1Out"
id="Bus57Phase1Out" connectedTo="line103Phase1In line128Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line103"
assetType="gplkh_4_10_curm_4_2p5" name="line103" length="4.0">
        <port xsi:type="esdl:InPort" name="line103Phase1In"
id="line103Phase1In" connectedTo="Bus57Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line103Phase1Out"
id="line103Phase1Out" connectedTo="Bus5701Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04237353"
lon="6.605940163"/>
          <point xsi:type="esdl:Point" lat="52.04236199"
lon="6.606060863"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus5701" name="Bus5701">
        <geometry xsi:type="esdl:Point" CRS="WGS84"
lat="52.042411990000005" lon="6.605960863"/>
        <port xsi:type="esdl:InPort" name="Bus5701Phase1In"
id="Bus5701Phase1In" connectedTo="line103Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus5701Phase2Out"
id="Bus5701Phase2Out"
connectedTo="EConnInuser47Phase2Inuser47Phase2In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line128"
assetType="gplkh_4_50_cusvm_4_6" name="line128" length="3.0">
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
        <port xsi:type="esdl:InPort" name="line128Phase1In"
id="line128Phase1In" connectedTo="Bus57Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line128Phase1Out"
id="line128Phase1Out" connectedTo="Bus70Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04237353"
lon="6.605940163"/>
          <point xsi:type="esdl:Point" lat="52.04229106"
lon="6.605918705"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus70" name="Bus70">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04234106"
lon="6.605818705"/>
        <port xsi:type="esdl:InPort" name="Bus70Phase1In"
id="Bus70Phase1In" connectedTo="line128Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus70Phase1Out"
id="Bus70Phase1Out" connectedTo="line129Phase1In line140Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line129"
assetType="gplkh_4_6_curm_4_2p5" name="line129" length="21.0">
        <port xsi:type="esdl:InPort" name="line129Phase1In"
id="line129Phase1In" connectedTo="Bus70Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line129Phase1Out"
id="line129Phase1Out" connectedTo="Bus7001Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04229106"
lon="6.605918705"/>
          <point xsi:type="esdl:Point" lat="52.04228282"
lon="6.6060233120000005"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus7001" name="Bus7001">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04233282"
lon="6.605923312000001"/>
        <port xsi:type="esdl:InPort" name="Bus7001Phase1In"
id="Bus7001Phase1In" connectedTo="line129Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus7001Phase3Out"
id="Bus7001Phase3Out"
connectedTo="EConnInuser60Phase3Inuser60Phase3In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line140"
assetType="gplkh_4_50_cusvm_4_6" name="line140" length="4.0">
        <port xsi:type="esdl:InPort" name="line140Phase1In"
id="line140Phase1In" connectedTo="Bus70Phase1Out"
carrier="Electricity"/>
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
        <port xsi:type="esdl:OutPort" name="line140Phase1Out"
id="line140Phase1Out" connectedTo="Bus76Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04229106"
lon="6.605918705"/>
          <point xsi:type="esdl:Point" lat="52.0422119"
lon="6.60587579"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus76" name="Bus76">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.0422619"
lon="6.60577579"/>
        <port xsi:type="esdl:InPort" name="Bus76Phase1In"
id="Bus76Phase1In" connectedTo="line140Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus76Phase1Out"
id="Bus76Phase1Out" connectedTo="line141Phase1In line146Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line141"
assetType="gplkh_4_10_curm_4_2p5" name="line141" length="28.0">
        <port xsi:type="esdl:InPort" name="line141Phase1In"
id="line141Phase1In" connectedTo="Bus76Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line141Phase1Out"
id="line141Phase1Out" connectedTo="Bus7601Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.0422119"
lon="6.60587579"/>
          <point xsi:type="esdl:Point" lat="52.04221025"
lon="6.605999172000001"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus7601" name="Bus7601">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04226025"
lon="6.605899172000001"/>
        <port xsi:type="esdl:InPort" name="Bus7601Phase1In"
id="Bus7601Phase1In" connectedTo="line141Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus7601Phase1Out"
id="Bus7601Phase1Out"
connectedTo="EConnInuser66Phase1Inuser66Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line146"
assetType="gplkh_4_50_cusvm_4_6" name="line146" length="9.0">
        <port xsi:type="esdl:InPort" name="line146Phase1In"
id="line146Phase1In" connectedTo="Bus76Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line146Phase1Out"
id="line146Phase1Out" connectedTo="Bus79Phase1In"
carrier="Electricity"/>
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
        &lt;geometry xsi:type="esdl:Line" CRS="WGS84"&gt;
          <point xsi:type="esdl:Point" lat="52.0422119"
lon="6.60587579"/>
          <point xsi:type="esdl:Point" lat="52.04213273"
lon="6.605846286"/>
        &lt;/geometry&gt;
      &lt;/asset&gt;
      &lt;asset xsi:type="esdl:Bus" id="Bus79" name="Bus79"&gt;
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04218273"
lon="6.6057462860000005"/>
        <port xsi:type="esdl:InPort" name="Bus79Phase1In"
id="Bus79Phase1In" connectedTo="line146Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus79Phase1Out"
id="Bus79Phase1Out" connectedTo="line147Phase1In line168Phase1In"
carrier="Electricity"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:ElectricityCable" id="line147"
assetType="gplkh_4_10_curm_4_2p5" name="line147" length="2.0">
        <port xsi:type="esdl:InPort" name="line147Phase1In"
id="line147Phase1In" connectedTo="Bus79Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line147Phase1Out"
id="line147Phase1Out" connectedTo="Bus7901Phase1In"
carrier="Electricity"/>
        &lt;geometry xsi:type="esdl:Line" CRS="WGS84"&gt;
          <point xsi:type="esdl:Point" lat="52.04213273"
lon="6.605846286"/>
          <point xsi:type="esdl:Point" lat="52.04213108"
lon="6.605958939"/>
        &lt;/geometry&gt;
      &lt;/asset&gt;
      &lt;asset xsi:type="esdl:Bus" id="Bus7901" name="Bus7901"&gt;
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04218108"
lon="6.605858939"/>
        <port xsi:type="esdl:InPort" name="Bus7901Phase1In"
id="Bus7901Phase1In" connectedTo="line147Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus7901Phase1Out"
id="Bus7901Phase1Out"
connectedTo="EConnInuser69Phase1Inuser69Phase1In"
carrier="Electricity"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:ElectricityCable" id="line168"
assetType="gplkh_4_50_cusvm_4_6" name="line168" length="9.0">
        <port xsi:type="esdl:InPort" name="line168Phase1In"
id="line168Phase1In" connectedTo="Bus79Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line168Phase1Out"
id="line168Phase1Out" connectedTo="Bus90Phase1In"
carrier="Electricity"/>
        &lt;geometry xsi:type="esdl:Line" CRS="WGS84"&gt;
          <point xsi:type="esdl:Point" lat="52.04213273"
lon="6.605846286"/>
```

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
          <point xsi:type="esdl:Point" lat="52.04206181"
lon="6.605814099"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus90" name="Bus90">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04211181"
lon="6.605714099"/>
        <port xsi:type="esdl:InPort" name="Bus90Phase1In"
id="Bus90Phase1In" connectedTo="line168Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus90Phase1Out"
id="Bus90Phase1Out" connectedTo="line169Phase1In line180Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line169"
assetType="gplkh_4_10_curm_4_2p5" name="line169" length="40.0">
        <port xsi:type="esdl:InPort" name="line169Phase1In"
id="line169Phase1In" connectedTo="Bus90Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line169Phase1Out"
id="line169Phase1Out" connectedTo="Bus9001Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04206181"
lon="6.605814099"/>
          <point xsi:type="esdl:Point" lat="52.04205191"
lon="6.605934799"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus9001" name="Bus9001">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04210191"
lon="6.605834799"/>
        <port xsi:type="esdl:InPort" name="Bus9001Phase1In"
id="Bus9001Phase1In" connectedTo="line169Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus9001Phase2Out"
id="Bus9001Phase2Out"
connectedTo="EConnInuser80Phase2Inuser80Phase2In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line180"
assetType="gplkh_4_50_cusvm_4_6" name="line180" length="14.0">
        <port xsi:type="esdl:InPort" name="line180Phase1In"
id="line180Phase1In" connectedTo="Bus90Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line180Phase1Out"
id="line180Phase1Out" connectedTo="Bus96Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04206181"
lon="6.605814099"/>
          <point xsi:type="esdl:Point" lat="52.04197934"
lon="6.605784595"/>
        </geometry>
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus96" name="Bus96">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04202934"
lon="6.6056845950000005"/>
        <port xsi:type="esdl:InPort" name="Bus96Phase1In"
id="Bus96Phase1In" connectedTo="line180Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus96Phase1Out"
id="Bus96Phase1Out" connectedTo="line181Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:ElectricityCable" id="line181"
assetType="gplkh_4_10_curm_4_2p5" name="line181" length="10.0">
        <port xsi:type="esdl:InPort" name="line181Phase1In"
id="line181Phase1In" connectedTo="Bus96Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="line181Phase1Out"
id="line181Phase1Out" connectedTo="Bus9601Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Line" CRS="WGS84">
          <point xsi:type="esdl:Point" lat="52.04197934"
lon="6.605784595"/>
          <point xsi:type="esdl:Point" lat="52.04197439"
lon="6.605916023"/>
        </geometry>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bus9601" name="Bus9601">
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04202439"
lon="6.605816023"/>
        <port xsi:type="esdl:InPort" name="Bus9601Phase1In"
id="Bus9601Phase1In" connectedTo="line181Phase1Out"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="Bus9601Phase1Out"
id="Bus9601Phase1Out"
connectedTo="EConnInuser86Phase1Inuser86Phase1In"
carrier="Electricity"/>
      </asset>
      <asset xsi:type="esdl:Transformer" capacity="800000.0"
id="transformer1" name="transformer1">
        <port xsi:type="esdl:InPort" name="transformer1In"
id="transformer1In" connectedTo="BussourcebusOut"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="transformer1Out"
id="transformer1Out" connectedTo="Bus1Phase1In"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Point" lat="52.04256757283102"
lon="6.60548740628655 7"/>
      </asset>
      <asset xsi:type="esdl:Bus" id="Bussourcebus"
name="Bussourcebus">
        <geometry xsi:type="esdl:Point" lat="52.042455533466146"
lon="6.605462193456334"/>
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
        <port xsi:type="esdl:InPort" name="BussourcebusIn"
id="BussourcebusIn" connectedTo="GenProducerOut"
carrier="Electricity"/>
        <port xsi:type="esdl:OutPort" name="BussourcebusOut"
id="BussourcebusOut" connectedTo="transformer1In"
carrier="Electricity"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:GenericProducer" power="1000000000.0"
prodType="FOSSIL" id="GenericElectricityProducer" name="Electricity
production Provincie">
        &lt;costInformation xsi:type="esdl:CostInformation"&gt;
          <marginalCosts xsi:type="esdl:SingleValue" value="0.7"
id="0fd72c70-fd6e-470e-83d1-f9d9a2f29e7a" name="GenericProducer_f09b-
MarginalCosts"/>
        &lt;/costInformation&gt;
        <port xsi:type="esdl:OutPort" name="GenProducerOut"
id="GenProducerOut" connectedTo="BussourcebusIn"
carrier="Electricity"/>
        <geometry xsi:type="esdl:Point" lat="52.04220026445083"
lon="6.605390310287476"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:Building" id="Building_4"
name="Building_4">
        &lt;asset xsi:type="esdl:EConnection" id="user4" name="Bus1301"&gt;
          <geometry xsi:type="esdl:Point" CRS="Simple"
lat="166.66666666666" lon="125.0"/>
          <port xsi:type="esdl:InPort" name="Inuser4Phase2In"
id="EconnInuser4Phase2In" connectedTo="Bus1301Phase2Out"
carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="Outuser4Phase2In"
id="EconnOutuser4Phase2In" connectedTo="EDemandInuser4Phase2In"
carrier="Electricity"/>
        &lt;/asset&gt;
        <asset xsi:type="esdl:ElectricityDemand" id="user4EDemand"
name="user4EDemand">
          &lt;geometry xsi:type="esdl:Point" lat="164.0" lon="304.0"/&gt;
          <port xsi:type="esdl:InPort" name="Inuser4Phase2In"
id="EDemandInuser4Phase2In" connectedTo="EconnOutuser4Phase2In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user4" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_4"/>
          &lt;/port&gt;
        &lt;/asset&gt;
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04264222"
lon="6.606229842"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:Building" id="Building_10"
name="Building_10">
        &lt;asset xsi:type="esdl:EConnection" id="user10" name="Bus2001"&gt;
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666"/>
```

ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
          <port xsi:type="esdl:InPort"
name="Inuser10Phase3Inuser10Phase3In"
id="EConnInuser10Phase3Inuser10Phase3In"
connectedTo="Bus2001Phase3Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser10Phase3In"
id="EConnOutuser10Phase3In" connectedTo="EDemandInuser10Phase3In"
carrier="Electricity"/>
        &lt;/asset&gt;
        <asset xsi:type="esdl:ElectricityDemand" id="user10EDemand"
name="user10EDemand">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser10Phase3In"
id="EDemandInuser10Phase3In" connectedTo="EConnOutuser10Phase3In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user10" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_10"/>
          &lt;/port&gt;
        &lt;/asset&gt;
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04252362"
lon="6.606114507000001"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:Building" id="Building_11"
name="Building_11">
        &lt;asset xsi:type="esdl:EConnection" id="user11" name="Bus2101"&gt;
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666666"/>
          <port xsi:type="esdl:InPort"
name="Inuser11Phase2Inuser11Phase2In"
id="EConnInuser11Phase2Inuser11Phase2In"
connectedTo="Bus2101Phase2Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser11Phase2In"
id="EConnOutuser11Phase2In" connectedTo="EDemandInuser11Phase2In"
carrier="Electricity"/>
        &lt;/asset&gt;
        <asset xsi:type="esdl:ElectricityDemand" id="user11EDemand"
name="user11EDemand">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser11Phase2In"
id="EDemandInuser11Phase2In" connectedTo="EConnOutuser11Phase2In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user11" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_11"/>
          &lt;/port&gt;
        &lt;/asset&gt;
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04264222"
lon="6.6063264010000005"/>
      &lt;/asset&gt;
```

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
      <asset xsi:type="esdl:Building" id="Building_18"
name="Building_18">
        &lt;asset xsi:type="esdl:EConnection" id="user18" name="Bus2801"&gt;
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666666"/>
          <port xsi:type="esdl:InPort"
name="Inuser18Phase3Inuser18Phase3In"
id="EConnInuser18Phase3Inuser18Phase3In"
connectedTo="Bus2801Phase3Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser18Phase3In"
id="EConnOutuser18Phase3In" connectedTo="EDemandInuser18Phase3In"
carrier="Electricity"/>
        &lt;/asset&gt;
        <asset xsi:type="esdl:ElectricityDemand" id="user18EDemand"
name="user18EDemand">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser18Phase3In"
id="EDemandInuser18Phase3In" connectedTo="EConnOutuser18Phase3In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user18" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_18"/>
          &lt;/port&gt;
        &lt;/asset&gt;
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04245105"
lon="6.606076956"/>
      &lt;/asset&gt;
      <asset xsi:type="esdl:Building" id="Building_47"
name="Building_47">
        &lt;asset xsi:type="esdl:EConnection" id="user47" name="Bus5701"&gt;
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666666"/>
          <port xsi:type="esdl:InPort"
name="Inuser47Phase2Inuser47Phase2In"
id="EConnInuser47Phase2Inuser47Phase2In"
connectedTo="Bus5701Phase2Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser47Phase2In"
id="EConnOutuser47Phase2In" connectedTo="EDemandInuser47Phase2In"
carrier="Electricity"/>
        &lt;/asset&gt;
        <asset xsi:type="esdl:ElectricityDemand" id="user47EDemand"
name="user47EDemand">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser47Phase2In"
id="EDemandInuser47Phase2In" connectedTo="EConnOutuser47Phase2In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user47" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_47"/>
          &lt;/port&gt;
```

ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
        </asset>
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04236199"
lon="6.606060863"/>
      </asset>
      <asset xsi:type="esdl:Building" id="Building_60"
name="Building_60">
        <asset xsi:type="esdl:EConnection" id="user60" name="Bus7001">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666"/>
          <port xsi:type="esdl:InPort"
name="Inuser60Phase3Inuser60Phase3In"
id="EConnInuser60Phase3Inuser60Phase3In"
connectedTo="Bus7001Phase3Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser60Phase3In"
id="EConnOutuser60Phase3In" connectedTo="EDemandInuser60Phase3In"
carrier="Electricity"/>
        </asset>
        <asset xsi:type="esdl:ElectricityDemand" id="user60EDemand"
name="user60EDemand">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser60Phase3In"
id="EDemandInuser60Phase3In" connectedTo="EConnOutuser60Phase3In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user60" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_60"/>
          </port>
        </asset>
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04228282"
lon="6.6060233120000005"/>
      </asset>
      <asset xsi:type="esdl:Building" id="Building_66"
name="Building_66">
        <asset xsi:type="esdl:EConnection" id="user66" name="Bus7601">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666"/>
          <port xsi:type="esdl:InPort"
name="Inuser66Phase1Inuser66Phase1In"
id="EConnInuser66Phase1Inuser66Phase1In"
connectedTo="Bus7601Phase1Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser66Phase1In"
id="EConnOutuser66Phase1In" connectedTo="EDemandInuser66Phase1In"
carrier="Electricity"/>
        </asset>
        <asset xsi:type="esdl:ElectricityDemand" id="user66EDemand"
name="user66EDemand">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser66Phase1In"
id="EDemandInuser66Phase1In" connectedTo="EConnOutuser66Phase1In"
carrier="Electricity">
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user66" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_66"/>
          </port>
        </asset>
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04221025"
lon="6.605999172000001"/>
      </asset>
      <asset xsi:type="esdl:Building" id="Building_69"
name="Building_69">
        <asset xsi:type="esdl:EConnection" id="user69" name="Bus7901">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666666"/>
          <port xsi:type="esdl:InPort"
name="Inuser69Phase1Inuser69Phase1In"
id="EConnInuser69Phase1Inuser69Phase1In"
connectedTo="Bus7901Phase1Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser69Phase1In"
id="EConnOutuser69Phase1In" connectedTo="EDemandInuser69Phase1In"
carrier="Electricity"/>
        </asset>
        <asset xsi:type="esdl:ElectricityDemand" id="user69EDemand"
name="user69EDemand">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser69Phase1In"
id="EDemandInuser69Phase1In" connectedTo="EConnOutuser69Phase1In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user69" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_69"/>
          </port>
        </asset>
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04213108"
lon="6.605958939"/>
      </asset>
      <asset xsi:type="esdl:Building" id="Building_80"
name="Building_80">
        <asset xsi:type="esdl:EConnection" id="user80" name="Bus9001">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666666"/>
          <port xsi:type="esdl:InPort"
name="Inuser80Phase2Inuser80Phase2In"
id="EConnInuser80Phase2Inuser80Phase2In"
connectedTo="Bus9001Phase2Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser80Phase2In"
id="EConnOutuser80Phase2In" connectedTo="EDemandInuser80Phase2In"
carrier="Electricity"/>
        </asset>
        <asset xsi:type="esdl:ElectricityDemand" id="user80EDemand"
name="user80EDemand">
```


ENERSHARE has received funding from European Union's Horizon Europe
Research and Innovation programme under the Grant Agreement No 101069831

```
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser80Phase2In"
id="EDemandInuser80Phase2In" connectedTo="EConnOutuser80Phase2In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user80" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_80"/>
          </port>
        </asset>
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04205191"
lon="6.605934799"/>
      </asset>
      <asset xsi:type="esdl:Building" id="Building_86"
name="Building_86">
        <asset xsi:type="esdl:EConnection" id="user86" name="Bus9601">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="166.66666666666666"/>
          <port xsi:type="esdl:InPort"
name="Inuser86Phase1Inuser86Phase1In"
id="EConnInuser86Phase1Inuser86Phase1In"
connectedTo="Bus9601Phase1Out" carrier="Electricity"/>
          <port xsi:type="esdl:OutPort" name="EConnOutuser86Phase1In"
id="EConnOutuser86Phase1In" connectedTo="EDemandInuser86Phase1In"
carrier="Electricity"/>
        </asset>
        <asset xsi:type="esdl:ElectricityDemand" id="user86EDemand"
name="user86EDemand">
          <geometry xsi:type="esdl:Point" CRS="Simple" lat="250.0"
lon="333.3333333333333"/>
          <port xsi:type="esdl:InPort" name="EDemandInuser86Phase1In"
id="EDemandInuser86Phase1In" connectedTo="EConnOutuser86Phase1In"
carrier="Electricity">
            <profile xsi:type="esdl:InfluxDBProfile"
database="SEMData" multiplier="1.5" filters=""
profileType="ENERGY_IN_J" id="EDemand_user86" port="8086"
measurement="elec_profiles" host="http://10.30.2.1" field="User_86"/>
          </port>
        </asset>
        <geometry xsi:type="esdl:Point" CRS="WGS84" lat="52.04197439"
lon="6.605916023"/>
      </asset>
    </area>
  </instance>
  <services xsi:type="esdl:Services">
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser5CHP" name="DrivenByDemandUser5CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser9CHP" name="DrivenByDemandUser9CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser13CHP" name="DrivenByDemandUser13CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser15CHP" name="DrivenByDemandUser15CHP"/>
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser17CHP" name="DrivenByDemandUser17CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser19CHP" name="DrivenByDemandUser19CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser52CHP" name="DrivenByDemandUser52CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser55CHP" name="DrivenByDemandUser55CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser59CHP" name="DrivenByDemandUser59CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser61CHP" name="DrivenByDemandUser61CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser67CHP" name="DrivenByDemandUser67CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser68CHP" name="DrivenByDemandUser68CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser74CHP" name="DrivenByDemandUser74CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser78CHP" name="DrivenByDemandUser78CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser79CHP" name="DrivenByDemandUser79CHP"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser84CHP" name="DrivenByDemandUser84CHP"/>
    <service xsi:type="esdl:DrivenByDemand" id="DrivenByDemandUser5GH"
name="DrivenByDemandUser5GH"/>
    <service xsi:type="esdl:DrivenByDemand" id="DrivenByDemandUser9GH"
name="DrivenByDemandUser9GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser13GH" name="DrivenByDemandUser13GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser15GH" name="DrivenByDemandUser15GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser17GH" name="DrivenByDemandUser17GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser19GH" name="DrivenByDemandUser19GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser52GH" name="DrivenByDemandUser52GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser55GH" name="DrivenByDemandUser55GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser59GH" name="DrivenByDemandUser59GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser61GH" name="DrivenByDemandUser61GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser67GH" name="DrivenByDemandUser67GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser68GH" name="DrivenByDemandUser68GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser74GH" name="DrivenByDemandUser74GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser78GH" name="DrivenByDemandUser78GH"/>
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser79GH" name="DrivenByDemandUser79GH"/>
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831

```
    <service xsi:type="esdl:DrivenByDemand"
id="DrivenByDemandUser84GH" name="DrivenByDemandUser84GH"/>
  </services>
  <energySystemInformation xsi:type="esdl:EnergySystemInformation"
id="ESI">
    <carriers xsi:type="esdl:Carriers" id="Carriers">
      <carrier xsi:type="esdl:ElectricityCommodity" voltage="240.0"
name="Electricity" id="Electricity"/>
      <carrier xsi:type="esdl:GasCommodity" name="Gas" id="Gas"/>
      <carrier xsi:type="esdl:HeatCommodity" name="Heat" id="Heat"/>
    </carriers>
  </energySystemInformation>
</esdl:EnergySystem>
```


ENERSHARE has received funding from European Union's Horizon Europe Research and Innovation programme under the Grant Agreement No 101069831