ECOsystems EMPOWERing at regional and local scale supporting energy communities

February 2025

D2.2 ECOEMPOWER ENERGY-ICT PLATFORM DEVELOPMENT

The project ECOEMPOWER - ECOsystems EMPOWERing at regional and local scale supporting energy communities receives funding from the European Climate, Infrastructure and Environment Executive Agency (CINEA) under Grant Agreement n°101120775.

Co-funded by the European Union

TECHNICAL REFERENCES

Project Acronym	ECOEMPOWER
Project Title	ECOsystems EMPOWERing at regional and local scale supporting energy communities
Funding Programme	LIFE 2027
Call	LIFE-2022-CET
Topic	LIFE-2022-CET-ENERCOM
Project Coordinator	Cinzia Morisco (FBK), cmorisco@fbk.eu
Project Start Date	September 1st, 2023
Project End Date	August 31st, 2026
Project Duration	36 months
Project ID	101120775

Deliverable No.	D2.2 ECOEMPOWER Energy-ICT platform development
Dissemination Level	PU – Public
Work Package	WP2 – Energy and ICT system Analysis
Task	Task 2.2 – Platform development for system planning and decision making
Lead Beneficiary	UBITECH ENERGY (UBE)
Author(s)	Eleni Kotali (UBE)
Contributing Beneficiaries	1 FBK, 2 PAT, 4 AURA-EE, 5 ACV, 6 eza!, 7 BAUM, 8 EAZK, 9 ROCG
Reviewer(s)	PAT, UNITN
Due Date of Deliverable	November 30th 2024
Actual Submission Date	February 28th 2025

REVISION AND HISTORY CHART

Version	Date	Editors	Comment
0.1	22/01/2025	UBE	ToC defined
0.2	19/02/2025	UBE	Initial version of the document
0.3	26/02/2025	UNITN, FBK	Document reviewed

ABBREVIATION LIST

ABBREVIATION	DEFINITION
ACV	Association des Centrales Villageoises
API	Application Programming Interface
BEV	Battery Electric Vehicle
DoA	Description of Action
DSO	Distribution System Operator
EC	Energy Community
EV	Electric Vehicle
EWR	Elektrizitätswerke Reutte
GDPR	General Data Protection Regulation
GSE	Gestore dei Servizi Energetici
HEDNO	Hellenic Distribution Network Operator
HMI	Human Machine Interface
ICT	Information and Communication Technology
IoT	Internet of Things
IT	Information Technology
KPI	Key Performance Indicator
LEC	Local Energy Community
LR	Linear Regression
LS	Least-Squares
ML	Machine Learning
NZEB	Nearly zero-emission building
OSS	One Stop Shop

PAT	Provincia Autonoma di Trento
PV	Photovoltaic
RAE	Regulatory Authority of Energy
REC	Renewable Energy Community
RF	Random Forest
ROCG	Region of Central Greece
ROI	Return Of Investment
SCADA	Supervisory Control and Data Acquisition
SSO	Single Sign-on
UI	User Interface
UC	Use Case
WP	Work Package

DISCLAIMER

The opinion stated in this report reflects the opinion of the authors and not the opinion of the European Commission. The European Union is not liable for any use that may be made of the information contained in this document.

This document will be made available for use and download on the ECOEMPOWER website under a Creative Commons license. It will be used the CC BY 4.0 DEED | Attribution 4.0 (https://creativecommons.org/licenses/by/4.0/). This license enables reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator.

All ECOEMPOWER consortium members are committed to publish accurate and up to date information and take the greatest care to do so. However, the ECOEMPOWER consortium members cannot accept liability for any inaccuracies or omissions, nor do they accept liability for any direct, indirect, special, consequential, or other losses or damages of any kind arising out of the use of this information.

1 EXECUTIVE SUMMARY

This deliverable documents the work conducted in Task 2.2 (T2.2) of Work Package 2 (WP2), offering a detailed overview of the data collected and the algorithmic processes designed for the creation of the ICT Energy Tools and Platform. A key aspect of the development process was the application of a co-creation methodology, which ensured that the tools were designed in alignment with the needs of relevant stakeholders, including energy community managers, aggregators, and prosumers.

Additionally, this deliverable outlines the datasets collected and analyzed from participating energy communities, integrating historical consumption records, renewable generation profiles, financial data, and external sources such as weather and market tariff datasets.

The Energy Forecasting Tool was developed to predict solar PV generation, enabling both long-term strategic planning and short-term operational adjustments. Given the data-driven nature of the tools, significant focus was placed on the available datasets collected from Renewable Energy Communities and various external sources, ensuring reliable model performance. The Energy Modelling and Scheduling Tool was designed to optimize energy consumption, resource allocation, and scheduling by maximizing self-consumption through a centralized decision-making entity. Unlike traditional approaches, this tool enables optimal energy management at the community level without relying on grid operator incentives. The Cost Benefit Analysis and Decision-Making Tool supports energy community leaders, aggregators, and stakeholders in evaluating the financial viability of renewable energy investments by assessing costs, return on investment, and long-term sustainability while also incorporating sensitivity analysis to economic variables.

This deliverable also provides details on the technology stack used, including Python-based data processing, optimization, and financial modeling libraries, as well as the React-based user interface for visualization and interaction.

Finally, this document concludes with insights into the next steps, particularly focusing on the validation of the tools in Task 2.3.

Contents

2 Introduction

2.1 Scope of the deliverable

The aim of WP2 is to provide a user-friendly platform for energy communities (ECs) supporting them in the coordination of their energy resources and future activities planning. To achieve this, WP2 adopts a three-step approach, comprising (i) defining the platform requirements and identifying ECs' Use Cases (UCs) (T2.1), (ii) developing the platform to meet the needs of ECs (T2.2), and (iii) validating the platform and energy analyses for the ECOEMPOWER ECs (T2.3).

This deliverable results from the work carried out in T2.2 – Platform development for system planning and decision making. Its scope contains a detailed description of the co-creation methodology we have followed for the development of the ICT Platform and the three distinct energy tools; the Energy Forecasting Tool, the Energy Modelling & Scheduling Tool and the Cost-Benefit Analysis (CBA) Tool. Additionally, we also describe the relevant data we have received from the 5 Regional Ecosystems (REs) as well as details regarding the algorithmic methodologies and technical implementation of the different UCs described in D2.1.

The aim of the document is also meant to pave the way for the future deliverables of this WP, that will in turn validate, refine and finalize the ICT Platform and its services by the end of the project. Overall, D2.2 aims to provide a first look of the progress of the technical implementation of the Energy Tools and give a more concrete look of the value for each RE.

2.2 Deliverable Structure

D.2.2 is structured in seven sections as follows:

- **Section 1: Executive summary** – This section provides a concise overview of the document's contents.
- **Section 2: Introduction** – This section provides an overview of this document including the description of its purpose, structure, and its interdependencies to the other ECOEMPOWER tasks and deliverables.
- **Section 3: ICT Tool Development Methodologies** – This Section provides an overview of the methodology followed to co-create the tools and collect data.
- **Section 4: Data Collection and Analysis** – This section presents the overview of the data collected by the 5 Regional Ecosystems, along with the external data that were utilized for the development process.
- **Section 5: Energy Tools Overview** – This section of the document outlines the three Energy Tools by covering their methodologies, algorithmic processes and evaluation metrics.
- **Section 6: Technology Stack** – This section provides basic information regarding the technology stacks used for the development of the Platform.
- **Section 7: Conclusions and Next Steps** – In this Section, key deliverable outcomes are summarized and the subsequent steps or actions to be taken within the project are outlined.

The document concludes with the Reference section and the lists of figures and tables.

2.3 Interdependencies with other Tasks and Deliverables

The interdependencies between T2.1, T2.2, T2.3, and T6.2 are fundamental to ensuring the seamless development and implementation of the ECOEMPOWER project. T2.1 lays the groundwork by identifying and characterizing the requirements of ECs, providing crucial input for the development of the software platform in T2.2. The platform developed in T 2.2, tailored to meet the requirements presented in Task 2.1, will then be utilized in T2.3 for platform validation and conducting energy studies in collaboration with the ECOEMPOWER pilot sites. Meanwhile, the data acquired during the initial phase of the project for T2.1, including assessments performed in collaboration with T6.2, will serve as a baseline for benchmarking purposes, providing essential insights into the existing infrastructure and planned initiatives within the targeted regions. The data collection continued in T2.2 as well, where more specialized data from the individual pilot sites were gathered and analyzed as part of the co-creation methodology. Also, a synergy has been identified with T4.4 and the ECOEMPOWER Community Platform as it could potentially be used to host the various Energy Tools presented in this deliverable.

3 ICT Tool Development Methodologies

This chapter outlines the methodologies that support the development of energy tools designed for energy communities. The first subchapter focuses on the broad data-driven decision-making process, emphasizing the framework of co-creation with Pilot Sites and stakeholder collaboration. This ensures that tools are tailored to real-world needs and offer practical value. The second subchapter identifies the primary stakeholders, outlining their roles and how the tools address their unique requirements. Finally, the data collection methodologies are detailed, covering the timeline, sources, and methods used to gather and process energy consumption, generation, and tariff data. Together, these elements form the methodological backbone of the project.

3.1 Data Driven Decision Making Process

The development of the energy tools followed a structured, data-driven decision-making framework centered on collaboration and co-creation with Pilot Sites. This process ensured that the tools address the specific challenges and requirements of the energy communities involved in the project. By engaging community managers, aggregators, and other stakeholders early in the process, the tools were designed to align with the real-world conditions, priorities, and regulatory frameworks specific to each region. Regular workshops, interviews, and feedback loops were conducted to refine the tools based on stakeholder input, fostering a participatory and iterative development approach.

The co-creation process leveraged real historical data, such as energy consumption patterns, generation profiles, and financial data, to create tailored use cases. These use cases, such as energy profile generation, community energy flow modelling, and cost-benefit analysis, served as the foundation for tool design. A modular approach was adopted to ensure scalability and adaptability across diverse energy community contexts. This methodology allowed the decision-making framework to remain flexible, enabling the addition of new data inputs or functionalities as the tools evolve and the needs of the Pilot Sites grow.

By anchoring decision-making in robust data analysis and collaboration, the tools not only provide actionable insights but also empower stakeholders to make informed decisions. The framework emphasizes transparency and inclusivity, ensuring that all stakeholders—from community managers to end-users—can understand and trust the recommendations provided by the tools. This approach builds confidence in the outputs, fosters stakeholder engagement, and ultimately enhances the adoption and success of the energy management solutions developed in the project.

3.2 Target Stakeholders

The energy tools developed as part of the project are designed to serve a wide range of stakeholders, reflecting the diversity of roles and interests within energy communities. These stakeholders include energy community managers, aggregators, policymakers, prosumers, and other relevant actors who play a role in the planning, operation, and optimization of energy systems. This section addresses the key stakeholders involved in the energy tools' development and deployment, emphasizing how their unique needs and challenges are addressed through the methodological framework. A significant element of the methodology is ensuring that the tools are designed to bridge existing gaps in energy management and decision-making processes, making them practical and impactful for diverse users.

By merging fragmented data into actionable insights, the tools directly address challenges in understanding energy consumption patterns at both individual and community levels. They also provide much-needed optimization capabilities for energy communities and managers who may lack the resources for advanced modeling or simulation. Early-stage energy communities, often limited by expertise or infrastructure, benefit from user-friendly tools that guide investment planning and energy flow optimization. Finally, by offering scenario-based analytics, the tools align with policymakers' needs for data-driven insights to craft more effective policies.

This stakeholder-driven approach ensures that the tools are developed not only to address pain points but also to empower users with the knowledge and strategies necessary to enhance energy efficiency, optimize resources, and support long-term sustainability.

Stakeholder	Pain Points	Benefits Provided by Tools
Energy Community Managers and Aggregators	Lack of tools to optimize intra-community energy flows and self-consumption. Limited insights for day-ahead planning and load balancing.	Detailed energy profiles and community-level simulations. Enhanced self-consumption optimization. Data-driven day-ahead scheduling and load management.
Prosumers and End-Users	Limited understanding of personal energy usage patterns. Missed opportunities for reducing energy costs through better self-consumption.	Personalized recommendations for energy efficiency and load-shifting. Insights into usage patterns to optimize energy behavior.
Policymakers and Regulators	Lack of data-driven insights for shaping effective policies. Difficulty in evaluating community-wide energy investment impacts.	Scenario-based insights on economic and social benefits of energy policies. Support for designing policies that promote sustainable energy practices.
Technology Developers and Researchers	Difficulty in testing and validating new energy solutions in realistic settings. Lack of datasets to explore community-level optimization.	Access to data-driven energy profiles, simulated energy flows, and scenario-based tools for validation and experimentation.
Energy Communities at Early Stages	Lack of guidance for investment planning and resource allocation. Difficulty in evaluating long-term benefits of community participation.	Investment planning scenarios tailored to energy communities. Tools to model operational plans without complex real-time data requirements.

Utility Companies	Difficulty in understanding distributed energy resources within communities. Lack of tools to manage peak loads or grid imbalances.	Support for better demand-side management at the community level. Insights into peak load trends and opportunities for redistribution.

3.3 Data Collection Methods

The data collection process followed a structured approach designed to gather the necessary inputs for the development and implementation of the energy tools outlined in this study. This process was guided by the co-creation framework and iterative discussions with key stakeholders, ensuring that the data collected was both relevant and aligned with the objectives of the use cases. The collection period spanned from M7 to M15 and was carried out as part of the activities within T2.2.

Data from pilot sites formed the backbone of this effort, and at this stage of the project, they could only be provided in the form of static datasets provided in Excel format. These datasets included critical information on energy consumption, generation patterns, and key characteristics of the buildings and assets within the energy communities. The pilot sites shared their data through bilateral discussions, during which specific requirements for the tools were outlined, and gaps were identified and addressed collaboratively. Additionally, workshops were organized with community managers and other stakeholders to refine the scope and applicability of the data.

To complement the static data, open-source datasets were retrieved via Application Programming Interfaces (APIs) from external sources, ensuring the inclusion of market tariffs, weather conditions, and other contextual factors critical for modeling and analysis. These external data sources were selected based on their reliability, relevance, and ability to integrate seamlessly with the tools under development.

The collected data served two primary purposes. First, it was utilized to train the algorithms underlying the energy tools, ensuring that their outputs accurately reflect the patterns and variability present within the real-world energy systems of the pilot sites. Second, the datasets provided crucial insights into the types of inputs that can be reasonably expected from the REs, which informed the design and functionality of the tools. Understanding these input characteristics was essential for aligning the tools' capabilities with the operational context and limitations of the energy communities.

The methodology adopted for data collection mainly prioritized flexibility, and practical applicability. While the process is not exhaustive or entirely automated, as we had originally envisioned in D2.1, this approach ensured that the collected data adequately represented the operational context of the energy communities, and thus kickstarting the development of the Energy Tools.

4 Data Collection and Analysis

This chapter provides an overview of the data gathered to support the development and implementation of the energy tools designed for the project. It details the types and characteristics of data collected from the five REs, encompassing both static and dynamic datasets provided by the pilot sites. These datasets include key information on energy consumption, generation, and asset characteristics, which form the foundation for creating energy profiles and models.

In addition to the pilot-specific data, the chapter also highlights the external datasets sourced from open APIs and publicly available resources. These external data types include market energy tariffs, weather information, and other contextual variables that are essential for enhancing the modeling and analysis capabilities of the tools. Together, these datasets enable the comprehensive analysis of energy behaviors within the pilot communities, providing the insights necessary to achieve the project's objectives.

4.1 Data Collected and Analyzed from the REs

4.1.1 RE1: Autonomous Province of Trento (Italy)

In this section we will describe the basic characteristics of the data provided by PAT, representing the three Energy Communities (Val di Fassa, Levico Terme, Valle dei Laghi) of the RE of Province of Trento.

The datasets received are focused on Pilot Site Levico Terme as it is the one that has the highest technical maturity of the three. They contain detailed records of energy consumption and financial data for the energy community of Levico Terme. They cover energy use for public infrastructure, residential buildings, and general community facilities. These datasets serve as a foundation for energy management, scheduling, and optimization efforts and will be used as the basis for developing the applicable UCs created in D2.1 into the ICT Tools, more specifically the Energy Modelling and Scheduling Tool. The two tables below provide the basic characteristics of the two datasets provided by PAT.

Table 1 RE1 Dataset A (Consumi 23 24)

Characteristic	Details
Location	Levico Terme, Trento, Italy
Date Range	January 2023 - March 2024
Number of Records	2073
Number of Variables	101
Type of Use	Public Lighting, Other Uses, Domestic Tariffs
Energy Data	Detailed energy consumption in time bands (F0-F3).
Financial Data	Extensive data on tariffs, taxes, and billed amounts.

Table 2 RE1 Dataset B (Consumi ante 2023)

Characteristic	Details
Location	Levico Terme, Trento, Italy
Date Range	Before January 2023
Number of Records	6885
Number of Variables	103
Type of Use	Public Lighting, Other Uses, Domestic Tariffs
Energy Data	Less detailed energy data; segmented time bands.
Financial Data	Basic data on tariffs and billed amounts.

The datasets provide extensive information on energy usage and financial metrics for the municipality of Levico Terme, Italy. The first dataset contains records spanning from January 2023 to March 2024, offering granular data for recent energy consumption and billing cycles. The second dataset has records, focusing on historical energy and financial data before 2023, making it valuable for long-term trend analysis.

Both datasets feature over 100 variables, including identifiers for network operators, business partners, and billing accounts. The datasets capture multiple facets of energy operations, including:

- Consumption data categorized by time bands.
- Financial details such as tariffs, taxes, and billing breakdowns.

Both datasets include energy consumption segmented by time bands (F0: total, F1: peak, F2: intermediate, F3: off-peak). Data includes total energy consumed (in kWh) as well as hourly or daily breakdowns for specific periods. Additional metrics such as total energy sold and specialized categories (e.g., PUN hourly and green energy) are recorded. Dataset A provides more recent and complete details, while Dataset B offers historical trends.

The financial data in both datasets is extensive, covering components such as:

- Energy tariffs (fixed and variable).
- Taxes, including VAT and excise duties.
- Adjustments for renewable energy, system charges, and other billing elements.

Notably, Dataset A has better granularity for recent billing trends, while Dataset B captures broader historical aggregates.

Both datasets classify energy consumption into standardized time bands:

- **F0 (Total):** Represents the aggregate energy consumption for all periods.
- **F1 (Peak):** Captures usage during high-demand periods, typically reflecting the highest tariffs.

- **F2 (Intermediate):** Covers mid-level demand periods with moderate tariffs.
- **F3 (Off-Peak):** Reflects low-demand periods, often with the lowest tariffs.

This segmentation allows for detailed profiling of energy consumption patterns. For example, a high proportion of consumption in F3 suggests cost-saving opportunities, while increased F1 usage highlights areas for optimization through load shifting. Additional metrics such as "Off Peak" and "PUN hourly" (Italian energy market hourly price) in both datasets provide further granularity. These enable more sophisticated analyses, such as identifying opportunities for integrating renewable energy during low-cost periods or understanding the impacts of dynamic tariffs.

The financial data in both datasets is extensive and structured, providing valuable insights into the economic aspects of energy use. Shared metrics include fixed and variable tariffs, taxes (e.g., VAT and excise duties), and adjustments for renewable energy and system charges. These financial details offer a comprehensive view of billing dynamics.

Dataset A stands out for its inclusion of recent adjustments and charges, such as:

- **Green Energy Fees:** Reflecting community investments in renewable energy.
- **New Billing Adjustments:** Likely due to regulatory or market changes in recent years.

These recent additions suggest a growing emphasis on sustainability and the incorporation of renewable energy initiatives. On the other hand, Dataset B provides a historical perspective, highlighting simpler or more traditional billing structures used in earlier years. The combination of these datasets enables a comparative analysis of how energy pricing and policy have evolved over time, which is essential for understanding long-term trends and planning future strategies.

4.1.2 RE2: Auvergne-Rhône-Alpes and Grand Est (France)

In this section, we describe the basic characteristics of the data provided for the French pilot sites, representing the three Energy Communities: Centrales Villageoises Vercorsoleil, Centrales Villageoises Vézouze-en-Piémont, and Centrales Villageoises Eau et Soleil du Lac, as have been provided by ACV.

These datasets are focused on energy production metrics and are crucial for developing advanced energy forecasting models as part of the implementation of UC1 through the Energy Forecasting Tool. The datasets include detailed monthly records of solar energy production, segmented by individual installations and their corresponding characteristics (e.g. capacity, building type, orientation etc). This provides a foundation for predicting future energy production, assessing seasonal trends, and optimizing renewable energy utilization across communities. Below, we summarize the datasets for each pilot site and their relevance to energy forecasting.

Table 3 RE2 Centrales Villageoises Vercorsoleil Dataset

Characteristic	Details
Location	Vercors region, France (multiple villages including La Chapelle, Saint Agnan, Saint Julien)
Date Range	January 2017 – April 2024

Number of Installations	29
Building Types	Public buildings (schools, village halls, churches, presbyteries)
Energy Metrics	Monthly energy production (kWh), cumulative production (kWh), and efficiency (kWh/kWc)

Table 4 RE2 Centrales Villageoises Vézouze-en-Piémont Dataset

Characteristic	Details
Location	Vézouze-en-Piémont region, France (villages include Domjevin, Amenoncourt, Repaix)
Date Range	January 2021 – April 2024
Number of Installations	10
Building Types	Public facilities (schools, gyms, town halls, village halls)
Energy Metrics	Monthly energy production (kWh), cumulative production (kWh), and efficiency (kWh/kWc)

Table 5 RE2 Centrales Villageoises Eau et Soleil du Lac Dataset

Characteristic	Details
Location	Aix-les-Bains region, France
Date Range	March 2023 – April 2024
Number of Installations	1
Building Types	School facilities
Energy Metrics	Monthly energy production (kWh), cumulative production (kWh), and efficiency (kWh/kWc)

Each dataset captures key attributes relevant to energy production operations, including:

- **Monthly energy production (kWh):** Total energy generated for each installation.
- **Cumulative energy production (kWh):** Long-term tracking of production output at the community level.

- **Efficiency metrics (kWh/kWc):** A standardized measure of the energy output generated by a solar installation relative to its installed capacity (in kilowatts-peak, kWc). It quantifies how effectively a solar system converts sunlight into electricity, expressed as the number of kilowatt-hours (kWh) produced per kilowatt-peak (kWc) of installed capacity over a specific period (e.g., monthly or annually). This metric allows for performance comparisons across systems of different sizes and configurations.

The datasets highlight key differences in production patterns and efficiencies across the three energy communities. Vercorsoleil provides the most extensive dataset, covering multiple installations over a long timeframe, making it well-suited for broad trend analysis. In contrast, Eau et Soleil du Lac focuses on a single, high-efficiency installation, offering a more concentrated but detailed view of performance. While the datasets include monthly records, they lack finer granularity, such as daily or hourly readings, limiting their usefulness for short-term analyses or peak demand assessments. Additionally, occasional gaps appear in certain months for specific installations, requiring careful handling to avoid distortions in long-term forecasting. Though these gaps are not widespread, they emphasize the need for interpolation or assumptions to maintain analytical consistency.

The datasets provide valuable insights into seasonal variations in energy production, allowing for a better understanding of fluctuations in output across different times of the year. By analyzing data over multiple months:

- Peak generation periods can be identified in summer, with reduced production observed in winter.
- These seasonal insights support planning for energy storage and supplementary generation during low-production months.
- Understanding production cycles helps optimize resource allocation and ensure energy availability throughout the year.

Beyond energy production, the datasets include critical installation attributes that influence efficiency. Each record captures:

- **Building orientation and tilt:** Affects exposure to sunlight.
- **System size (kWc capacity)**: Allows for comparisons across installations.
- **Performance trends**: Helps to identify best practices and areas for improvement.

To ensure comparability across installations and communities, the datasets use standardized energy production metrics, such as:

- **Total output (kWh)**: Insights into absolute production levels.
- **Efficiency ratios (kWh/kWc)**: Performance benchmarking.

These standardized measures facilitate integration into larger analytical models, supporting energy planning at both local and regional levels. Another key strength of the datasets is their chronological structure, which allows for time-series analysis of long-term production trends. This format supports:

- **Evaluation of intervention impacts**: Tracking improvements over time.
- **Forecasting of future energy output**: Aiding in strategic planning.
- **Identification of performance anomalies**: Though short-term fluctuations may be harder to detect due

to the monthly data resolution.

While the datasets provide a solid foundation for macro-level energy forecasting, their limitations—such as occasional data gaps and a lack of high-frequency readings—underscore the need for robust monitoring systems. Enhancing data granularity and ensuring continuous reporting would further improve their value for both operational and strategic energy assessments.

4.1.3 RE3: Allgäu (Germany)

In this section we will describe the basic characteristics of the data provided by eza! and BAUM, representing the three Energy Communities (Elektrizitätswerke Hindelang, Dorfenergie Eppishausen, Elektrizitätswerke Reutte) of the RE of Allgäu.

The datasets received are focused on Pilot Site Elektrizitätswerke Hindelang and contain a detailed report of timestamped generation and consumption energy data from various locations in the community. These datasets serve as the foundation for developing an Energy Modelling and Scheduling Tool, which will enable energy management, scheduling, and optimization.

The datasets cover detailed energy consumption across different sectors, including households, commercial businesses, agricultural enterprises, and specialized operations, recorded at 15-minute intervals. Additionally, a separate dataset contains generation data from multiple locations.

The datasets are categorized into energy consumption and generation datasets, enabling a comprehensive analysis of energy demand and supply. The table below contains the basic characteristics of the consumption datasets:

Table 6 Energy Consumption Datasets for RE3

Category	Description
Households	Energy consumption for residential buildings
Commercial Businesses (Type 1)	General commercial energy consumption
Commercial Businesses (Type 2)	Businesses operating between 8 AM - 6 PM
Commercial Businesses (Type 3)	Businesses with peak energy consumption in the evening hours
Shops & Hair Salons	Businesses with continuous energy consumption throughout the day
Bakeries	Energy consumption for stores and salons
Weekend Businesses	Energy consumption for bakeries with in-house baking facilities
Agricultural Enterprises	General agricultural energy consumption

Dairy Farming	Energy consumption for farms with dairy operations
Other Agricultural Enterprises	Energy use for miscellaneous agricultural businesses

The basic characteristics of the set are the following:

- Time of energy consumption record (15-minute intervals)
- Energy consumption for Winter on Saturdays, Sundays, and weekdays
- Energy consumption for Summer on Saturdays, Sundays, and weekdays
- Energy consumption for transitional seasons on Saturdays, Sundays, and weekdays
- Aggregated energy consumption for a given category (e.g., per 1000 households)

The generation dataset contains measured energy production from various renewable energy installations, recorded at 15-minute intervals. The dataset provides data from multiple locations, including:

Table 7 Generation Dataset Overview

Location	Generation Type
LandWirt + Hotel Eggensberger	PV system
Hotel Eggensberger	Combined Heat & Power system (CHP)
Hoteldorf Hartung	CHP and PV system
Hotel Hirsch	CHP and PV system
Klinik Enzensberg	CHP and PV system

Some initial patterns that can be detected from the generation datasets include:

- **Early Morning (00:00 - 07:00):** No solar PV generation; BHKW (CHP) systems provide steady output.
- **Morning to Midday (07:00 - 12:00):** Gradual increase in PV generation, reaching a peak around midday.
- **Afternoon to Evening (12:00 - 18:00):** Peak PV generation occurs in this timeframe.
- **Nighttime (18:00 - 00:00):** PV generation ceases, leaving BHKW systems as the primary source of generation.

4.1.4 RE4: Zlín Region (Czech Republic)

The following section outlines the fundamental characteristics of the datasets provided by EAZK for the Zlín Region RE. These datasets contain detailed energy consumption data, tariff structures, and operational characteristics related to the municipality's energy infrastructure. They cover public and municipal buildings, energy consumption patterns, distribution tariffs, and self-consumption aspects. These datasets form the foundation for the development of the Energy System Modelling and Scheduling Tool, which will support energy optimization and planning for energy communities.

The datasets focus on the Pilot Site of Vlčnov, which is the one with the highest technical maturity in terms of monitoring infrastructure. They cover different aspects of the energy system like:

- Electricity infrastructure and tariffs for municipal buildings.
- High-frequency energy consumption data at 15-minute intervals for multiple buildings.
- Renewable energy generation and self-consumption, particularly photovoltaic (PV) integration for municipal operations.

A summary of the three can be found in the tables below:

Table 8 Municipality Vlčnov - Electricity Dataset

Characteristic	Details
Location	Vlčnov, Zlín Region, Czech Republic
Date Range	February 2024 – September 2024
Number of Records	33
Number of Variables	17
Type of Use	Infrastructure & Energy Tariffs; Public buildings and municipal energy infrastructure
Energy Data	Tariff structures, high and low tariff energy consumption (kWh), reserved power
Operational Data	Electricity supplier, voltage levels, fuse/circuit breaker capacities, type of metering
Financial Data	Distribution tariffs, billing periods, invoicing details

Table 9 Vlčnov Consumption Data from 3 buildings

Characteristic	Details
Location	Vlčnov, Zlín Region, Czech Republic
Date Range	February 2023 – December 2023

Number of Records	35,043 per building (3 buildings/meters)
Number of Variables	7
Type of Use	Time-series energy consumption data; Municipal buildings, public facilities
Energy Data	Active power consumption (kW), reactive power (kVAr), power metering at 15-minute intervals

Table 10 Municipal House Vlčnov Dataset

Characteristic	**Details**
Location	Vlčnov, Zlín Region, Czech Republic
Date Range	February 2024 – August 2024
Number of Records	20,349
Number of Variables	6
Type of Use	Municipal House energy use and self-consumption
Energy Data	Active consumption (kW), photovoltaic (PV) energy supply (kW)
Operational Data	Metering Status

Together, these datasets provide a comprehensive view of the energy landscape in Vlčnov, offering insights into consumption trends, tariff structures, and renewable energy integration.

The Municipality Vlčnov - Electricity dataset provides high-level information about the electricity infrastructure, metering points, supplier details, tariff structures, and network limitations. This dataset is essential for understanding the billing framework and operational constraints that affect municipal energy management.

The Vlčnov Consumption Data dataset includes detailed time-series energy consumption records for multiple municipal buildings, captured every 15 minutes. These granular readings allow for profiling energy demand, identifying peak consumption patterns, and optimizing scheduling. The dataset also includes reactive power measurements, which are crucial for assessing power quality and network efficiency.

The Municipal House Vlčnov dataset focuses on energy self-consumption and renewable energy generation, specifically photovoltaic (PV) energy production and its impact on energy demand. This dataset allows for evaluating the self-sufficiency potential of municipal buildings and provides insights into the efficiency of renewable integration.

As a holistic overview of the Czech Data, some key insights can be extracted. A key aspect of the data is the energy consumption trends in relation to tariff structures. Consumption is measured at high and low tariff periods, which allows for cost analysis and energy savings through strategic load shifting. Understanding the billing structure and tariff categories enables the municipality to optimize operational costs and select cost-efficient energy agreements.

The time-series datasets offer 15-minute interval recordings of energy usage, making it possible to identify anomalies, inefficiencies, and peak demand periods. This level of granularity supports the development of demand-side energy management strategies that can reduce unnecessary load surges and improve energy efficiency.

In addition to consumption data, the datasets include renewable energy and self-consumption metrics, with a particular focus on photovoltaic (PV) generation at the Municipal House. These records provide insights into the extent to which solar energy contributes to the municipality's energy needs, as well as opportunities for optimizing surplus PV energy distribution and exploring energy storage solutions.

From a broader community perspective, these datasets support energy balancing and resource optimization across multiple municipal buildings. By analyzing these datasets collectively, it is possible to model and simulate various energy flow scenarios, helping to reduce grid dependence and enhance self-sufficiency. More specifically, the datasets enable the following key analyses:

- Energy profiling and demand analysis: Identifying usage patterns and forecasting energy needs.
- Cost and tariff optimization: Aligning consumption with off-peak tariff periods to minimize expenses.
- Renewable energy integration: Assessing the impact of PV generation on municipal energy self-sufficiency.
- Community-wide energy flow simulations: Exploring how energy can be shared between buildings to enhance efficiency.
- Investment planning for energy storage and grid upgrades: Providing data-driven insights into future infrastructure improvements.

4.1.5 RE5: Region of Central Greece (Greece)

In this section we will describe the basic characteristics of the data provided by ROCG, representing the three Energy Communities (Domokos, Amfikleia, Kamena Vourla) of the RE of the Region of Central Greece.

The datasets received are focused on Pilot Site Domokos as it is the one that has the highest technical maturity of the three. They contain detailed records of energy consumption and financial data for different residential, commercial and municipal buildings. The data covers different levels of granularity, offering a comprehensive view of energy usage patterns, cost structures, and potential areas for optimization. Each dataset supports financial and operational decision-making by enabling stakeholders to analyze consumption trends, cost fluctuations, and potential areas for efficiency improvements. Below is a structured breakdown of each dataset, highlighting its key features and relevance.

Table 11 Domokos Dataset A - Residential Consumption Data

Characteristic	Details

Location	Domokos, Greece
Date Range	January 2024 – Dec 2024
Number of Records	160
Number of Variables	6
Type of Use	Household/residential electricity consumption
Financial Metrics	Billed energy costs per kWh, final amounts, and billing type
Energy Metrics	Energy consumption per household across multiple billing periods

Table 12 Domokos Dataset B - Commercial Consumption Data

Characteristic	**Details**
Location	Domokos, Greece
Date Range	January 2024 – Dec 2024
Number of Records	36
Number of Variables	6
Type of Use	Commercial sector energy consumption
Financial Metrics	Detailed billing data including per kWh cost, final billed amounts, and type of billing
Energy Metrics	Aggregated energy consumption per billing period

Table 13 Domokos Dataset C- Municipal Consumption Data

Characteristic	**Details**
Location	Domokos, Greece
Date Range	November 2024
Number of Records	2880

Number of Variables	3
Type of Use	Public buildings' electricity consumption
Energy Metrics	High-frequency time-series energy consumption in kWh, recorded at 15-min granularity

The three datasets together form a comprehensive picture of energy usage within the Domokos Energy Community, covering municipal, commercial, and residential sectors. While **Dataset C (Municipal)** offers a detailed time-series analysis of energy consumption, **Dataset B (Commercial)** and **Dataset A (Residential)** provide structured billing period data with financial insights.

The key differentiators among the datasets include:

- **Granularity:** Municipal dataset provides high-frequency readings, whereas commercial and residential datasets aggregate data at billing periods.
- **Financial Coverage:** Commercial and residential datasets contain detailed cost structures, while municipal data focuses solely on energy usage.
- **Sector Focus:** Municipal dataset helps optimize public infrastructure energy use, commercial dataset aids businesses in cost control, and residential dataset offers insights into household consumption trends.
- **Tariff Diversity:** The commercial and residential datasets include distinct tariff structures, affecting energy pricing and final billed amounts.

The first dataset represents residential electricity consumption patterns, structured similarly to the commercial dataset but focusing on individual households. It provides energy usage statistics alongside financial metrics such as unit costs and total billed amounts. The dataset covers 15 different residential consumers, representing various energy consumption patterns. Additionally it includes two distinct billing formats:

- **Green (Special) Tariff:** Linked to subsidized pricing models. Default for customers who don't choose a plan. Price is announced monthly by each supplier.
- **Yellow (Variable) Tariff:** A pricing model with fluctuating rates, based on wholesale market prices and their fluctuations.

The second dataset captures commercial energy consumption and associated financial costs, segmented by business entities. It includes both energy consumption data (in kWh) and financial breakdowns such as cost per kWh and total billing amounts. The data is aggregated at a broader level on a monthly basis, covering three different small-medium sized businesses. Similarly to the first dataset, it also includes two distinct tariff categories:

- **Green (Special) Tariff**: Linked to subsidized pricing models. Default for customers who don't choose a plan. Price is announced monthly by each supplier.
- **Blue (Fixed) Tariff**: A standard fixed-rate pricing model.

This dataset is particularly useful for tracking business-sector energy costs, assessing tariff impacts, and optimizing electricity expenses for commercial operations.

The third dataset provides detailed municipal energy consumption records, structured with timestamps at the date and time level. The granularity allows for an in-depth analysis of electricity demand patterns, peak usage hours, and overall energy efficiency within municipal or public sector buildings in Domokos. The dataset is valuable for tracking energy demand fluctuations, optimizing operational energy efficiency, and supporting energy-saving initiatives.

4.2 Data from External Sources

In our exploration of available datasets to inform the development of our energy management tools, we have identified several external sources that offer valuable data across various domains. While not all have been directly utilized, each provides insights pertinent to energy forecasting, modeling, and financial analysis. Below is a categorized overview of these datasets:

Table 14 Collection of external datasets explored in the development of the ICT Tools

Dataset Name	Data Features	Description
OpenWeatherMap	Weather Data including Temperature, Solar Irradiance and more	Provides comprehensive weather data, including current conditions, forecasts, and historical data. This information is crucial for predicting energy demand and renewable energy generation. Utilized for integrating weather data into energy models.
Copernicus Atmosphere Monitoring Service (CAMS) Solar Radiation Time-Series	Solar Radiation Parameters (GHI, DNI)	Offers historical data on solar radiation, essential for modeling solar energy generation and aligning consumption with generation.
European Climate Assessment & Dataset (ECA&D)	Temperature, Precipitation, Cloud Cover	Provides daily weather observations from across Europe. Valuable for capturing seasonal and historical trends, aiding in long-term photovoltaic (PV) production forecasts and validating forecasting models.
ERA5 Reanalysis Data	Temperature, Solar Irradiance, Wind Speed, Cloud Cover	Offers hourly estimates of atmospheric parameters, supporting short-term PV production forecasting and detailed energy scheduling.

ENTSO-E Transparency Platform	Electricity Generation, Transmission, Consumption, Day-Ahead Prices	Provides comprehensive data on electricity generation, transmission, and market prices across Europe. Instrumental for energy consumption modeling and market analysis.
RAAEY - European Day-Ahead Electricity Prices Map	Day-Ahead Electricity Prices	Offers insights into electricity prices across European markets. Vital for financial modeling and cost-benefit analyses of energy projects.
PRIMES Model	Energy Consumption, Supply, EU Carbon Price Trajectories	Simulates EU energy systems, providing projections on market dynamics and aiding in economic analyses related to energy policies.
Enerdata	Energy Efficiency, CO2 Emissions, Energy Trends	Supports financial analysis by providing insights into energy efficiency, carbon emissions, and consumption trends, aiding in evaluating the cost-effectiveness of energy projects.

The datasets outlined above represent a diverse range of data sources that support the foundational methodologies behind our energy management tools. By leveraging these datasets, we ensure that our tools are informed by robust and reliable data, covering key areas such as weather patterns, energy generation and consumption, and market trends. While not every dataset has been directly integrated into the tools, they collectively provide a valuable repository for future enhancements and refinements, enabling continuous improvement in energy forecasting, modeling, and financial decision-making processes.

5 Energy Tools Overview

5.1 Forecasting Tool

The Energy Forecasting Tool is designed to provide accurate predictions for solar photovoltaic (PV) energy generation across multiple energy communities in the EU. The tool supports both long-term (monthly) forecasting (UC1.a) for strategic planning and short-term (hourly/daily) forecasting (UC1.b) for real-time operational adjustments.

The forecasting framework integrates historical solar energy production data, weather datasets, and machine learning models to predict future energy output with high accuracy. This hybrid approach ensures that both seasonal trends and real-time variations are accounted for, enabling energy community operators to optimize resource allocation and grid management [1].

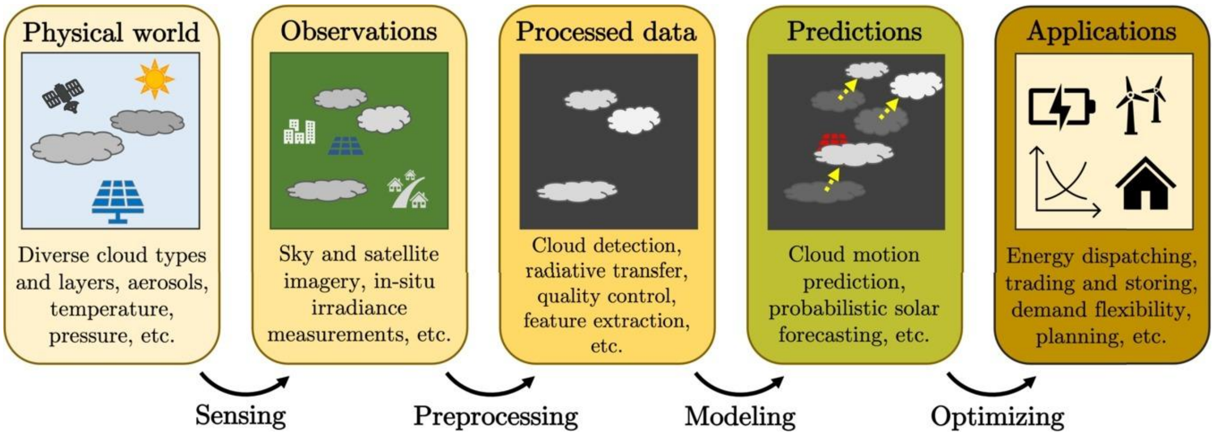

Figure 1 High level approach to Energy Forecasting

The next sections will delve deeper into each phase of the methodology.

5.1.1 Data Processing & Feature Engineering

The Energy Forecasting Tool integrates multiple datasets, including historical PV production records, weather data, and external climate sources, to improve prediction accuracy. This section outlines the data processing pipeline and feature engineering techniques used to optimize forecasting performance.

The forecasting model is built upon structured datasets from multiple pilot sites, external weather sources, and historical records. The primary data sources include:

Pilot Site Datasets	External Datasets
Monthly PV production data (**kWh**)	OpenWeatherMap (Real-time & forecasted weather data**: temperature, irradiance, cloud cover**)
Cumulative energy production (**kWh**)	CAMS Solar Radiation Time-Series (**Solar radiation estimates** for PV efficiency modeling)
Efficiency metrics **(kWh/kWc**)	ERA5 Reanalysis Data (**Historical climate conditions** for long-term trend analysis)

Installation characteristics (**location, orientation, system capacity**)	

Figure 2 Data Processing for the Forecasting Tool

Raw data undergoes multiple preprocessing steps to enhance model accuracy and eliminate inconsistencies.

1. **Data Cleaning**

- Handling missing values using interpolation for time-series continuity.
- Removing outliers with statistical anomaly detection (e.g., Z-score analysis).

2. **Data Normalization & Scaling**

- Normalizing production data using Min-Max Scaling to improve ML model convergence.
- Standardizing weather features (e.g., temperature, irradiance) using Z-score transformation.

3. **Time-Series Structuring**

- Aligning all datasets to a uniform timestamp format (e.g., daily/monthly).
- Resampling high-frequency weather data to match the production data resolution.

To maximize forecasting accuracy, the tool employs a structured approach to feature extraction from raw data, leveraging key techniques to enhance predictive performance.

One critical aspect of this methodology is the incorporation of temporal features. By analyzing the month, day, and hour of energy production, the model captures seasonal variations and daily fluctuations. Additionally, lag-based features, such as the energy output from previous months, are integrated to account for time-series memory effects, ensuring that past trends inform future predictions.

Another important dimension is the use of weather and solar-related features, which play a crucial role in short-term forecasting. The tool applies moving averages and rolling statistical techniques to smooth out noise in weather patterns, improving data stability and reliability. Furthermore, key meteorological factors such as solar radiation intensity and cloud cover are incorporated, as they directly impact photovoltaic (PV) energy generation, helping to refine short-term predictions.

Beyond these core features, the model also derives efficiency metrics to enable more precise performance assessments. Metrics such as energy yield per kilowatt capacity (kWh/kWc) allow for comparisons across different installations, while the capacity utilization factor (CUF) provides insight into seasonal variations in PV performance.

Through a combination of rigorous data cleaning, normalization, and advanced feature engineering, the forecasting model gains a deeper understanding of both long-term seasonal trends and short-term fluctuations. This comprehensive approach enhances predictive accuracy, enabling more reliable energy forecasting.

5.1.2 Forecasting Models

The Energy Forecasting Tool is designed to leverage machine learning techniques to generate accurate predictions for both long-term (monthly) and short-term (hourly/daily) solar energy production. Given the inherent variability in solar energy generation, the forecasting models must effectively capture both seasonal patterns and real-time fluctuations, ensuring optimal decision-making for energy community managers and grid operators.

To address the different forecasting horizons required for this tool, we employ a combination of statistical, machine learning, and deep learning models. The selection of these models is based on their demonstrated effectiveness in handling time-series data, particularly in energy forecasting applications.

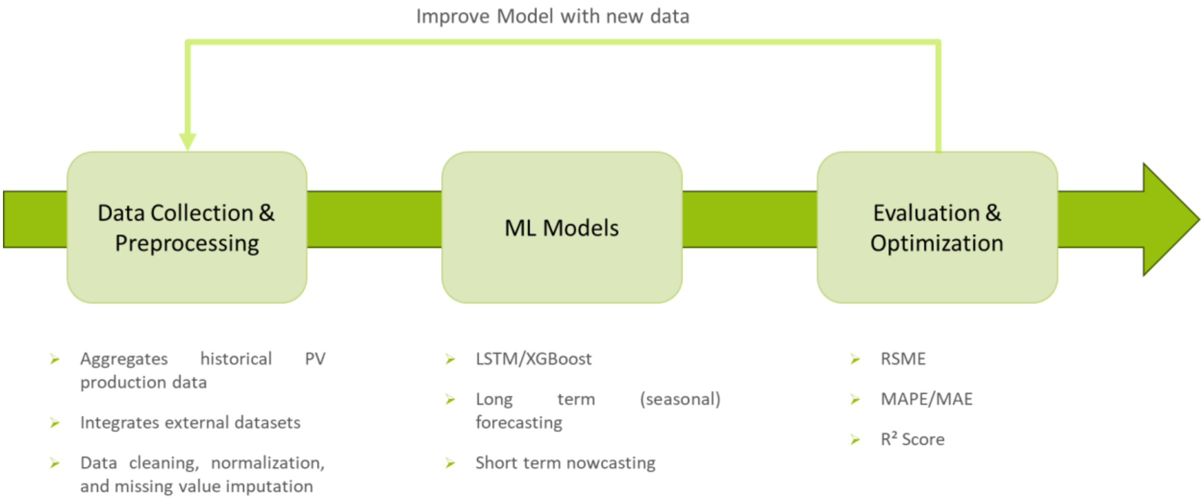

Figure 3 High Level Architecture of the Forecasting Tool

For long-term forecasting (UC1.a), where the objective is to predict monthly PV production, models that excel at identifying seasonal trends and long-term dependencies are prioritized. The Seasonal AutoRegressive Integrated Moving Average (SARIMA) model is particularly suited for this task as it effectively captures cyclical patterns in energy generation by incorporating seasonal differencing into its forecasting framework. Additionally, Long Short-Term Memory (LSTM) neural networks, a type of recurrent neural network (RNN), are used to capture complex temporal dependencies, making them highly effective for time-series forecasting. Finally, Random Forest Regression, an ensemble-based machine learning model, is employed to identify non-linear relationships and interactions among the various input features.

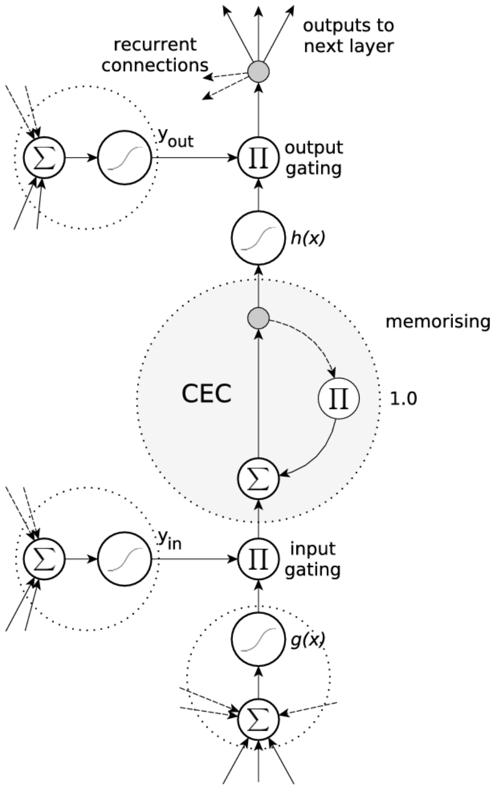

Figure 4 LSTM building blocks as depicted in [2]

For short-term forecasting (UC1.b), where the goal is to generate accurate hourly or daily predictions, the focus is on models that can quickly adapt to rapid fluctuations in solar production due to weather variability. XGBoost (Extreme Gradient Boosting), a high-performance tree-based machine learning algorithm, is employed due to its ability to handle structured time-series data with high accuracy and efficiency. Additionally, LSTM networks with an attention mechanism are incorporated to ensure that recent data points are given greater importance in making short-term predictions. Finally, LightGBM, an optimized gradient boosting framework known for its computational efficiency, is used to improve real-time forecasting performance.

By utilizing this diverse set of models, the Energy Forecasting Tool ensures robust forecasting accuracy across different time horizons, catering to both strategic planning and real-time operational adjustments.

To ensure the models generalize well to unseen data, a structured data splitting strategy is implemented. The dataset is divided into training and testing subsets, typically using an 80-20 split, where 80% of the historical data is used for model training and the remaining 20% is reserved for validation. Given the time-dependent nature of the dataset, rolling forecasting origin cross-validation is employed to test the models on progressively newer data while preserving the chronological order.

Hyperparameter tuning is a critical step in optimizing the models for accuracy and efficiency. For deep learning models like LSTMs, Bayesian Optimization is used to fine-tune parameters such as the number of layers, dropout rates, and learning rates. For boosting models like XGBoost and LightGBM, grid search techniques are applied to optimize key parameters, including tree depth, learning rate, and the number of estimators. Similarly, the

SARIMA model is fine-tuned by selecting the optimal values for seasonal differencing order, autoregressive, and moving average components. [3]

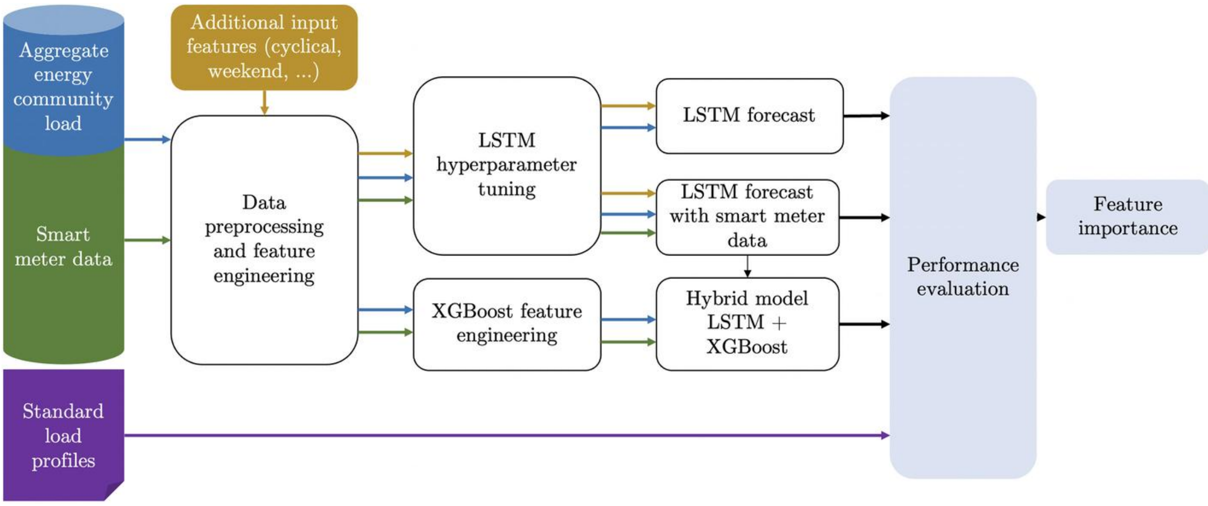

Figure 5 Forecasting Architecture when smart meter data is incorporated as proposed in [3]

To further enhance forecasting accuracy and tool synergy in future versions of the ECOEMPOWER ICT Platform, an ensemble learning approach is adopted, where multiple models are combined to create a more robust prediction system. In particular, a hybrid model that integrates LSTM and XGBoost can be implemented, leveraging the deep learning model's ability to capture complex temporal dependencies while utilizing the gradient boosting model's strength in handling structured data efficiently. The inputs that can be incorporated are the aggregated profiles created in the Energy Modelling Tool and potentially smart meter data to the Pilot Sites with enough technical maturity. The outputs from these models are then combined through weighted averaging, allowing the system to extract load forecasts in addition to RES Forecasts.

5.1.3 Evaluation Metrics

To ensure the reliability and accuracy of the Energy Forecasting Tool, we have selected some evaluation metrics to assess the performance of the forecasting models. Given the importance of both long-term (monthly) and short-term (hourly/daily) solar energy forecasting, multiple error metrics are used to evaluate different aspects of prediction accuracy.

The following **error metrics** are used to evaluate the effectiveness of the forecasting models:

1. **Root Mean Square Error (RMSE):** Measures the overall magnitude of prediction errors by penalizing larger deviations more heavily.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$$

   RMSE is particularly useful for evaluating short-term forecasts (UC1.b), where sudden weather fluctuations may cause large deviations.

2. **Mean Absolute Error (MAE):** Provides an absolute measure of average prediction error.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|$$

MAE is more interpretable than RMSE and is useful for assessing the general accuracy of long-term forecasts (UC1.a).

3. **Mean Absolute Percentage Error (MAPE)**: Evaluates the relative error as a percentage of actual values, allowing for comparisons across different solar installations.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\widehat{y_i} - y_i}{y_i}\right| \times 100$$

4. **Coefficient of Determination (R^2 Score):** Measures how well the model explains variance in the data, with values ranging from **0 to 1** (where 1 indicates a perfect fit).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

R^2 provides insight into the overall predictive power of the model and helps compare different forecasting approaches.

By employing these evaluation metrics and methodologies, the Energy Forecasting Tool ensures high accuracy, reliability, and adaptability, making it a robust solution for energy communities in managing solar PV generation.

5.2 Energy Modelling and Scheduling Tool

The Energy Modelling and Scheduling Tool is designed to provide energy communities with advanced capabilities for optimizing energy consumption, resource allocation, and scheduling, based on the maximization of self-consumption potential (without relying on grid operator's (DSOs/TSOs) incentives). Unlike traditional models that focus on peer-to-peer (P2P) trading or individual prosumers, this tool is structured around a central decision-making entity (at aggregated level) that ensures optimal energy management at the community level, that is usually operated by an aggregator.

The tool integrates three key functionalities based on the defined **Use Cases (UCs)**:

1. **UC2.a – Energy Profiling & Day-Ahead Planning (building level):** Generates detailed energy consumption profiles for buildings and schedules flexible loads to align with renewable energy generation and available electricity tariff structures for cost efficiency.
2. **UC2.b – Community Energy Flow Simulation:** Models energy distribution within the community, identifying surpluses, deficits, and optimization strategies for better resource utilization.
3. **UC2.c – Load Management & Scenario Planning:** Provides tools to assess, optimize, and model community self-consumption while allowing users to explore different renewable energy scenarios to improve sustainability.

The development of this tool is guided by five core principles, ensuring it effectively serves energy communities while remaining adaptable, efficient, and scalable.

First and foremost, the tool follows an Energy Community-Centric Approach, prioritizing decisions that align with the collective needs of the energy community (EC) rather than focusing on individual buildings or prosumers. This ensures that energy management strategies are optimized at a broader level, fostering shared benefits and resource efficiency.

To enhance usability and precision, the tool is designed with Flexibility and Adaptability in mind. Users have the ability to manually input missing data through an intuitive interface, allowing them to refine building consumption profiles and improve the accuracy of scheduling recommendations. This adaptability ensures that the system remains effective even in cases where complete data sets are unavailable.

A key objective of the tool is to maximize the utilization of renewable energy sources, particularly self-consumption of locally generated photovoltaic (PV) power and energy storage assets. By intelligently managing energy flows, the system reduces reliance on external energy sources, promoting sustainability and energy independence within the community.

In addition to sustainability, cost-effectiveness optimization is a fundamental aspect of the tool's design. It strategically aligns energy consumption with periods of low electricity tariffs and peak renewable production, minimizing operational costs while maximizing the efficiency of energy use.

Finally, the tool is built with scalability and future expansion in mind. It is designed to support multiple energy communities with varying data sources, ensuring interoperability and seamless integration across different energy systems. This forward-looking approach guarantees that the tool remains relevant as energy networks evolve and expand.

Overall, the core computational framework of the Energy Modelling Tool consists of the following components:

- **Optimization Algorithms for Load Scheduling** – Applies **Mixed Integer Linear Programming (MILP)** for optimal scheduling of flexible loads.
- **Models of Community-Wide Energy Flows** – Utilizes **agent-based modeling** to analyze energy sharing opportunities within the community.
- **Scenario-Based RES Planning** – Incorporates **what-if analysis** for different renewable energy configurations, allowing users to explore optimal self-consumption strategies.

As a potential improvement we are also considering adding a time-series Energy Profiling Component, that will utilize historical consumption data to create pridictive profiles for the Energy Demand of the buildings or other assets. Since this module will require heavier computational resources and data, we will consider adding this on a later version of the tool.

5.2.1 Data Processing & Feature Engineering

The accuracy and efficiency of the Energy Modelling and Scheduling Tool are heavily dependent on the quality of the energy consumption, financial, and renewable generation datasets available from various Energy Communities (ECs). The tool is designed to integrate both historical and real-time data, ensuring that energy optimization decisions are based on robust and up-to-date information. Given that data gaps and inconsistencies may exist in real-world implementations, the tool also allows users to manually input missing details through the user interface (UI) to ensure accurate modelling and optimization.

This section outlines the data collection process, the preprocessing pipeline, and the feature engineering techniques used to convert raw data into structured inputs suitable for optimization and scenario analysis.

The tool aggregates data from multiple sources, each providing crucial insights into energy consumption patterns, renewable energy availability, and financial implications of energy usage. The datasets utilized in the tool can be classified into three main categories:

Table 15 Data needed for the development of the Energy Modelling Tool

Pilot Site Datasets (Historical & Real-Time Energy Consumption)	Renewable Generation & Self-Consumption Data	User-Provided Inputs via UI
Hourly and 15-minute interval energy usage data for different types of buildings	PV and CHP generation data recorded at 15-minute intervals	Building characteristics (e.g appliances, operational constraints, energy flexibility)
Energy consumption segmentation into time bands (e.g F0-F3) aligned with tariff structures	Battery storage usage metrics, including charge and discharge cycles	Scenario inputs (RES configurations, investment constraints, self-consumption targets)
Financial records detailing tariffs, taxes, and billed amounts for cost optimization		

Similarly to the Forecasting Tool, to maintain data integrity and consistency, a structured preprocessing pipeline is applied before the data is used in optimization models. This ensures that the tool can handle incomplete, noisy, or inconsistent data effectively, improving the reliability of the generated schedules and scenario outcomes.

Real-world datasets often contain gaps or inconsistencies due to sensor errors, communication failures, or incomplete records. To address these issues:

- Interpolation techniques are applied to fill small gaps in time-series energy consumption records, ensuring continuous data for optimization.
- Proxy profiles are generated using clustering-based estimations, which allow the system to infer missing consumption data based on similar buildings with comparable energy usage patterns.

This approach ensures that missing data does not introduce inaccuracies in scheduling and decision-making. Energy consumption data varies significantly across different buildings and operational contexts. To ensure comparability:

- Min-Max Scaling is used to standardize energy consumption values, ensuring that all buildings are analyzed on a uniform scale.
- Z-score normalization is applied to standardize energy demand and renewable generation patterns, ensuring consistency in load scheduling models.

By normalizing the data, the tool enhances the accuracy and efficiency of machine learning models used in scenario-based decision-making.

Energy consumption follows predictable time-based patterns, influenced by factors such as daily routines, weather conditions, and operational schedules. To capture these patterns:

- Time-of-day, weekday/weekend, and seasonal indicators are extracted to analyze variations in demand across different periods.

- Peak vs. off-peak demand segmentation is performed to align energy consumption with tariff structures, enabling cost-efficient scheduling.

These features allow the tool to generate load profiles that align with both operational needs and economic considerations.

Given that maximizing self-consumption is a key objective for Energy Communities, the tool incorporates several renewable energy performance indicators:

- PV generation ratios, which quantify the percentage of total electricity demand met by on-site solar generation.
- Battery charge/discharge cycles, allowing the tool to model how stored energy is utilized to offset peak grid consumption.

These metrics enable accurate renewable energy integration, ensuring that locally generated energy is used efficiently before drawing power from the grid.

Load flexibility is a crucial aspect of energy optimization, as deferrable loads (e.g., HVAC systems, EV chargers) can be rescheduled to reduce costs and increase efficiency. To support this:

- The tool identifies and classifies flexible loads, distinguishing them from non-deferrable energy consumption.
- Scenario-based RES capacity factors are introduced, allowing for investment analysis and long-term energy planning.

By extracting these features, the tool enables intelligent load shifting and scenario-based optimization, ensuring that ECs can adapt to changing energy conditions and future investment opportunities.

5.2.2 Integration of External Datasets

The Energy Modelling and Scheduling Tool leverages external datasets to enhance the accuracy of energy optimization, load scheduling, and scenario planning within Energy Communities (ECs). While historical and real-time consumption data from pilot sites form the core of the system, integrating additional external datasets ensures that the tool is dynamic, adaptable, and capable of making informed decisions based on broader market and environmental conditions.

To optimize scheduling and decision-making, the tool incorporates data from multiple external sources. These datasets provide critical insights into electricity markets, renewable energy availability, and regional climate conditions, allowing for more precise load optimization and energy cost management. The main external datasets include:

Table 16 External Data used for the Energy Modelling Tool

ENTSO-E Transparency Platform	**European Day-Ahead Electricity Prices Map**	**PRIMES Model & Enerdata Market Forecasts**

Real-time and historical electricity market data	Hourly electricity price forecasts	Long-term energy market projections
Market clearing prices		Demand trends
Day-ahead electricity price variations		Energy mix forecasts

Real time prices are used to signal the potential shift of the flexible loads and to manage storage and are also incorporated as constraints in the load scheduling model. The long-term projections are assisting the decision making process for UC2.c and the long term planning of the ECs by assessing the feasibility of additional RES investments.

5.2.3 Optimization & Scheduling Models

The Energy Modelling and Scheduling Tool is designed to optimize energy consumption, resource allocation, and scheduling within ECs, enabling them to maximize self-consumption, reduce costs, and improve energy efficiency. As mentioned in previous sections, unlike P2P energy trading models, where individual prosumers make autonomous decisions, this tool adopts a centralized optimization approach. A designated aggregator is responsible for making collective scheduling decisions, ensuring that energy resources are distributed efficiently among buildings in the community.

The optimization framework considers multiple factors, including time-of-use tariffs, renewable generation availability, battery storage management, and load flexibility, to determine optimal energy scheduling strategies. This approach allows communities to dynamically balance their energy flows, optimize self-consumption, and plan investments in renewable energy sources (RES).

The tool is built to achieve multiple key objectives that directly benefit Energy Communities:

1. Cost Optimization: By analyzing tariff structures and time-of-use pricing, the tool shifts energy consumption to low-cost periods while avoiding peak demand charges.
2. Self-Consumption Maximization: The tool prioritizes the use of locally generated renewable energy before drawing from external sources, ensuring communities make the most of their existing PV and battery storage assets.
3. Load Balancing Across the Community: The tool identifies surplus and deficit energy periods and redistributes energy across buildings to ensure fair and efficient utilization of resources.
4. Battery Storage Optimization: By scheduling optimal charge and discharge cycles, the tool ensures that batteries are used effectively to store excess PV generation for later use, reducing reliance on the grid.
5. Scenario Planning for Renewable Energy Investments: Users can examine and evaluate different RES configurations, such as adding new solar PV installations or battery storage systems, to assess their impact on energy independence and financial savings.

The optimization framework integrates mathematical optimization, heuristic algorithms, and other models, enabling effective decision-making for energy scheduling and planning.

5.2.3.1 Mixed-Integer Linear Programming (MILP) for Load Scheduling

To optimize day-ahead scheduling of flexible loads, the tool employs Mixed-Integer Linear Programming (MILP), a powerful mathematical optimization technique that ensures optimal and globally efficient scheduling decisions. The model minimizes total energy costs while adhering to operational constraints, such as appliance runtime requirements and user-defined preferences. [4]

The objective function of the MILP model aims to minimize the total cost of energy consumption across all buildings:

$$min \sum_{t} \sum_{b} C_{t,b} \cdot P_{t,b}$$

where $C_{t,b}$ represents the cost of energy for building b at time t, and $P_{t,b}$ represents the power consumption at that time. The optimization process ensures that load shifting strategies are implemented effectively, prioritizing low-tariff periods and high-renewable generation windows.

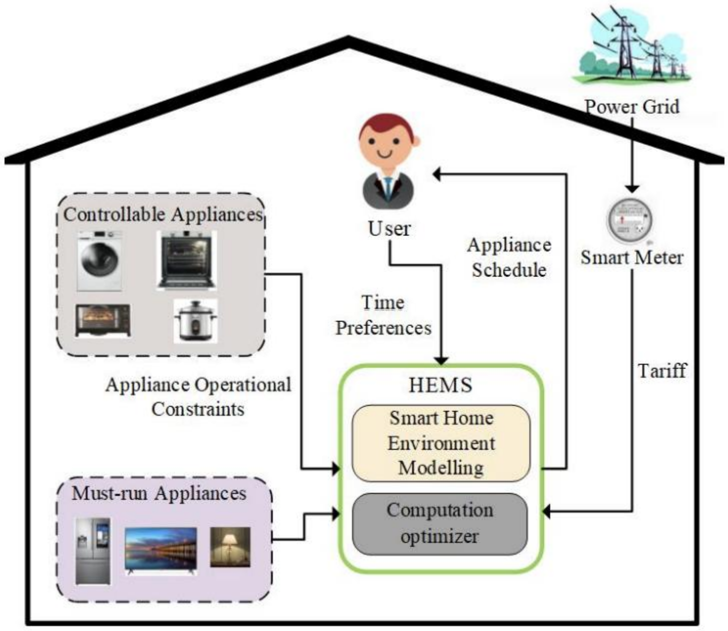

Figure 6 Factors to be taken into account for the MILP Algorithms as mentioned in [4]

5.2.3.2 Quadratic Programming for Battery Charge/Discharge Optimization

To manage battery storage systems efficiently, the tool uses Quadratic Programming (QP) to determine optimal charging and discharging schedules. The model balances minimizing costs with maximizing battery lifespan, ensuring that stored energy is used effectively to reduce peak consumption and reliance on grid energy.

The objective function for battery optimization is defined as:

$$min \sum_{t} (\alpha \cdot SOC_t^2 + \beta \cdot E_t)$$

Where SOC_t represents the battery state of charge at time t and E_t represents the energy exchanged. The quadratic terms are chosen based on a simplified model presented in [5] to provide a smooth and balanced charging pattern, preventing excessive cycling that could degrade batter health.

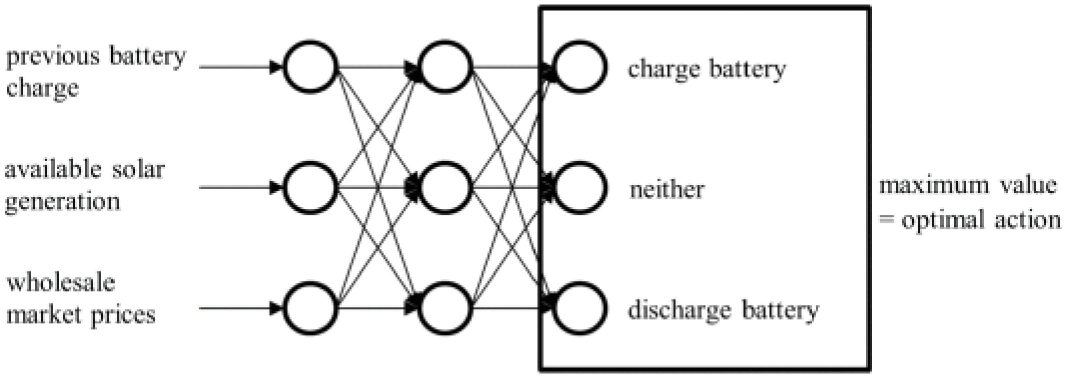

Figure 7 Model approach for battery charge/discharge optimization as described in [5]. We will utilize a simplified format for the first version of the tool.

5.2.3.3 Multi-Agent Modelling for Community-Wide Energy Flow Balancing

To ensure fair and efficient energy distribution across the community, the tool employs multi-agent models. In the energy domain, a multi-agent system (MAS) refers to a network of autonomous agents, which can be software, hardware, or a combination of both, that work together to manage and control various aspects of an electrical system. In this approach, each building acts as an autonomous energy agent, with predefined consumption patterns, possible flexibility levels, and storage capacity [6].

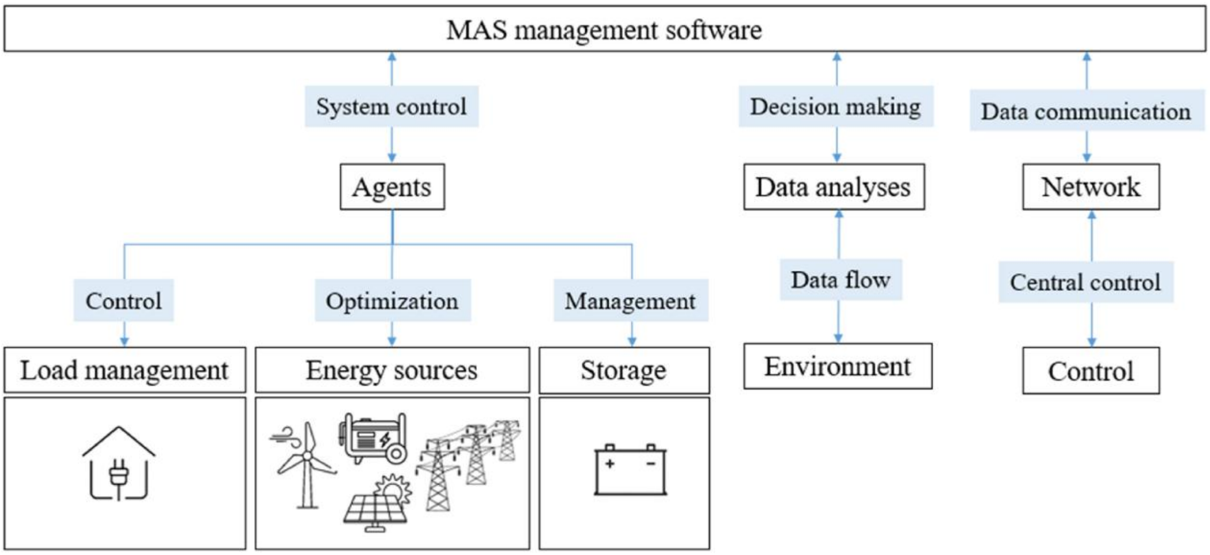

Figure 8 Potential agent types as highlighted in [6], in our approach we focus on the optimization and management side of the communities

Some concessions were made for the ECOEMPOWER Architecture, since there is no grid information or direct DSO involvement for any of the cases we are working on. Namely, there will not be an agent representing and managing the main power grid or an agent that can autonomously make actuation decisions in the ECs. The Building Agents' focus is to provide actionable recommendations about optimized energy scheduling and self-consumption, incorporating available and price signal data.

The aggregator could use the suggestions provided by the agents to dynamically adjust load schedules and thus balance energy surpluses and deficits, ensuring that community-wide energy resources are utilized in a more optimized manner. This model is particularly valuable in scenarios where some buildings generate more renewable energy than they consume, enabling intelligent redistribution within the community.

5.2.3.4 Heuristic Methods and Genetic Algorithms for Scenario-Based RES Planning

For long-term energy planning and investment analysis, the tool uses Heuristic Algorithms to examine the viability of various renewable energy deployment strategies. Users can input different configurations—such as installing new PV panels, increasing battery storage, or modifying load flexibility settings—and the tool evaluates the impact of these changes on overall community performance. Additionally, Genetic Algorithms (GA) have also been examined as a possible solution for this specific optimization problem, as this method does not require a priori knowledge of probability distributions or statistical moments of uncertain parameters, as it is based on their range [7].

The optimization process follows these steps:

- Generate or input an initial population of different RES configurations (e.g., various PV and battery sizes).
- Evaluate each configuration based on predefined criteria, such as self-consumption improvement, cost savings, and return on investment.
- Apply heuristic function or genetic operations (selection, crossover, mutation) to find an optimal solution.
- Output the best RES investment plan that aligns with the community's goals.

5.2.3.5 Decision-Making Workflow

The tool follows a structured decision-making process to ensure accurate and effective scheduling:

Figure 9 High Level Workflow for the Energy Modelling Tool

1. **Data Collection & Preprocessing:**

- Aggregates historical energy consumption, PV generation, tariff structures, and user-provided scenario inputs.

- Cleans and normalizes data to ensure consistency across different sources.

2. **Optimization & Scheduling Execution:**

- MILP-based scheduling generates optimized day-ahead load plans for each building.
- Battery optimization (QP) determines ideal charge/discharge cycles to store excess renewable energy.
- Multi-agent energy flow balancing ensures community-wide optimization.
- Evolutionary scenario analysis explores new RES investment options.

3. **Results & Visualization:**

- The tool presents optimized load schedules, self-consumption metrics, cost savings estimates, and renewable energy recommendations through an interactive dashboard.
- Users can explore different investment scenarios to determine the most effective strategies for improving community energy independence.

4. **User Feedback & Adjustments:**

- Energy community managers can modify inputs, adjust constraints, and rerun the algorithms to refine scheduling decisions.
- The tool adapts over time as new data is continuously integrated into the optimization framework.

5.2.4 Evaluation Metrics

To ensure the accuracy, efficiency, and reliability of the Energy Modelling and Scheduling Tool, a structured evaluation framework is implemented. The tool is designed to optimize energy consumption, scheduling, and scenario planning for Energy Communities (ECs), making it essential to assess its performance across multiple dimensions. The evaluation methodology focuses on cost efficiency, energy self-sufficiency, optimization quality, and system adaptability.

The tool is assessed using a range of quantitative performance indicators, each providing insight into a specific aspect of system performance. These include:

Table 17 Energy Modelling Tool Evaluation Metrics

Metric	Definition	Purpose	Formula
Total Cost Savings (%)	Measures the percentage reduction in total electricity costs compared to a baseline (e.g., no optimization).	Evaluates the tool's ability to shift consumption to low-cost tariff periods and reduce overall community energy expenses.	None
Self-Consumption Ratio (%)	The percentage of locally generated renewable energy that is consumed	Higher self-consumption rates indicate better use of locally generated PV	$CR = \frac{E_{self\_consumed}}{E_{total\_generated}} \times 100$

	within the community rather than exported to the grid.	energy, reducing dependency on external suppliers.	
Renewable Utilization Index (RUI)	Assesses the fraction of community energy demand met by renewables, combining self-consumption and optimized storage usage.	Evaluates how efficiently the tool integrates renewables and energy storage to meet community energy needs.	$SCR = \frac{E_{self\_consumed}}{E_{total\_generated}} \times 100$

5.3 Cost-Benefit Analysis and Decision-Making Tool

The Cost Benefit Analysis (CBA) and Decision-Making Tool is designed to support energy community leaders, aggregators, and other stakeholders in making informed financial decisions about energy investments. It provides a structured framework for evaluating the financial viability of various renewable energy projects, allowing users to assess potential interventions based on cost, return on investment (ROI), long-term financial sustainability, and sensitivity to economic variables.

Energy communities face growing challenges in managing their financial resources efficiently, particularly as they seek to increase renewable energy penetration, enhance self-consumption, and optimize energy contracts. The CBA Tool addresses these needs by offering quantitative financial analysis, scenario-based investment planning, and tariff optimization capabilities, ensuring that projects align with both short-term operational goals and long-term economic viability.

The CBA methodology evaluates projects across different time horizons, ensuring that both short-term financial gains and long-term economic impacts are considered. The tool provides a structured approach to financial analysis, enabling users to:

- Identify financially viable energy projects that contribute to long-term sustainability.
- Optimize capital allocation by prioritizing high-impact investments.
- Adapt to changing economic conditions by modeling different financial scenarios.

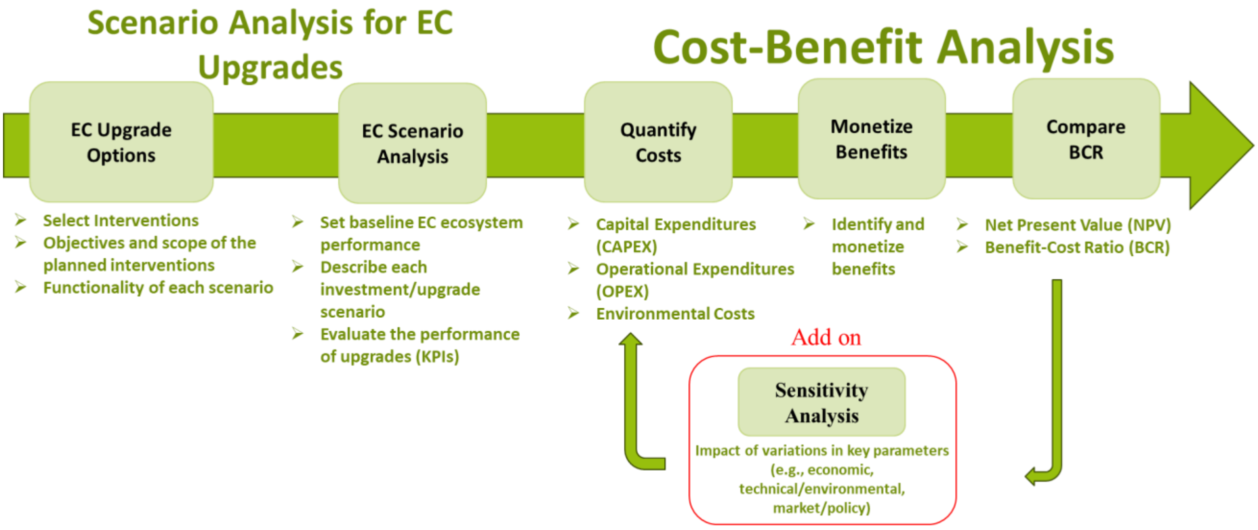

Figure 10 High Level Workflow of the CBA Tool

The computational framework behind the CBA Tool is designed to handle large-scale energy investment scenarios, using mathematical formulations and financial evaluation methods to generate accurate and reliable projections. To provide a comprehensive and structured analysis, the tool follows a five-step methodology, guiding users through:

- **Scenario Selection & Investment Planning** – Users define proposed interventions, such as expanding renewable generation, adding battery storage, or improving energy efficiency. Each scenario includes cost estimates, expected benefits, and key assumptions.
- **Performance Evaluation** – The tool assesses baseline energy consumption and financial performance, providing insights into current inefficiencies and estimating how interventions impact energy self-sufficiency, grid reliance, and revenue generation.
- **Cost Quantification** – Users input financial details to create a full cost breakdown for each investment scenario.
- **Benefit Monetization** – The tool translates energy savings, emission reductions etc. into monetary terms, enabling a direct comparison between different investment options.
- **Cost-Benefit Analysis (CBA) Computation** – The system calculates key financial indicators, to determine whether a project is financially viable. A sensitivity analysis is also performed, evaluating how economic fluctuations affect project feasibility.

5.3.1 Data Processing & Feature Engineering

The accuracy and reliability of the CBA Tool depend on the availability of high-quality financial, energy consumption, and investment-related datasets. Given that energy community investments often involve long-term financial projections, ensuring clean, structured, and well-integrated data is critical for obtaining accurate cost-benefit assessments and investment recommendations.

This section describes the data sources, preprocessing pipeline, and feature engineering techniques used to structure raw data into meaningful inputs for cost-benefit modeling, investment evaluation, and scenario analysis.

Data Category	Variables	Usage
Historical Energy and Financial Data	Energy Demand & Consumption Trends, segmented by time-of-use (peak, off-peak, intermediate hours)	Establish baseline energy consumption profiles and benchmark financial indicators for comparing against proposed investment scenarios.
	Self-Consumption Rates	
	Tariff Structures	
	Investment Records of past renewable energy installations and storage deployments	
Renewable Energy and Storage Data	Solar and wind energy data sourced from pilot sites or standardized regional datasets	Enable scenario-based analysis to assess how new renewable energy or battery storage deployments impact energy costs, revenues, and self-sufficiency.
	Battery Storage Data	
	Energy Market Prices	
User-Provided Investment Scenarios & Market Data	Capital Expenditures (CAPEX)	Allow users to input custom financial assumptions and key market variables to model and compare investment scenarios for more tailored and informed decision-making.
	Operational Expenditures (OPEX)	
	Policy Incentives like subsidies, feed-in tariffs, or carbon pricing mechanisms	
	Discount rates, and energy price trends	

To maintain data integrity, all collected datasets undergo a structured preprocessing pipeline before they are used in financial modeling. This step is crucial for ensuring that raw financial and energy data is transformed into a structured and standardized format suitable for cost-benefit analysis. The data preprocessing pipeline consists of two main steps:

Investment datasets frequently contain incomplete, inconsistent, or missing records, which can introduce biases into financial assessments. To mitigate these issues, the tool applies:

- Interpolation techniques to estimate missing values in energy consumption, financial expenditures, and investment cost datasets.
- Proxy modeling approaches, where missing investment cost data is approximated based on historical benchmarks and similar energy projects.

- Outlier detection methods, which identify and remove abnormal or extreme values in energy pricing, revenue estimates, or cost projections.

To enable a fair comparison between different investment options, financial data must be normalized across different time periods, cost structures, and energy pricing mechanisms. The tool applies:

- Standardized financial metrics, aligning energy cost data with historical and forecasted market trends to create meaningful financial comparisons. These metrics are further explained in Section 5.3.4.
- Unit-based cost indexing, normalizing all investment costs per installed capacity (€/kW) or energy output (€/MWh), ensuring compatibility with industry-standard financial assessments.

To extract meaningful insights from preprocessed datasets, the tool generates a comprehensive set of financial indicators and cost-benefit metrics, equipping decision-makers with the necessary data to evaluate investment feasibility and optimize financial planning. By analyzing these metrics, energy communities can assess the long-term profitability, risk exposure, and overall economic impact of renewable energy and storage investments.

5.3.2 CBA Methodology

The CBA Tool employs a structured methodology to assess the financial viability of energy investments, ensuring that energy community stakeholders can make data-driven decisions. The methodology follows a systematic approach to quantifying costs and benefits, applying financial modeling techniques such as Net Present Value (NPV), Benefit-Cost Ratio (BCR), and sensitivity analysis. This is a widely accepted methodology as described in ENTSO-E's guidelines [8], that we have modified to cover the specific needs of the ECs. These financial indicators allow users to compare multiple investment scenarios, optimize resources, and ensure long-term economic sustainability.

As summarized in the introductory section of this chapter, the CBA framework implemented in this tool consists of five key steps that are detailed in the following sections.

5.3.2.1 Identification of Interventions & Investment Scenarios

The first step in the cost-benefit analysis process is the identification of potential interventions and investment scenarios. Energy communities often face multiple investment opportunities, each with varying financial implications. The tool allows users to define:

- Increased renewable energy integration, such as expanding solar PV capacity, wind energy installations, or hybrid RES projects.
- Energy storage deployment, including battery energy storage systems (BESS) or other demand-side flexibility solutions.
- Demand-side management technologies, such as smart energy management systems, heat pumps, and electric vehicle (EV) charging infrastructure.

Each investment scenario is analyzed based on technical feasibility, expected output (e.g energy savings) and financial parameters, ensuring a comprehensive evaluation of potential projects.

5.3.2.2 Performance Evaluation & Baseline Comparison

Before assessing the financial implications of new investments, it is essential to establish a baseline performance profile of the energy community. The tool analyzes:

- Historical energy consumption trends, identifying peak demand periods and inefficiencies.
- Self-consumption rates, determining the proportion of locally generated energy used within the community.
- Existing financial performance, analyzing historical energy tariffs, operational costs, and revenues from energy projects.

Once a baseline scenario is established, investment interventions are evaluated in terms of their impact on:

- Increased renewable energy penetration – Assessing the reduction in grid dependency.
- Enhanced self-consumption rates – Estimating improvements in on-site energy usage.
- Potential revenue streams – Identifying new income sources from energy trading or grid services.

This comparative analysis ensures that stakeholders can assess the financial and operational improvements offered by different investment options.

5.3.2.3 Cost Quantification

A key component of the cost-benefit analysis is the identification and classification of costs associated with each investment scenario. The tool categorizes costs into three main groups:

1. **Capital Expenditures (CAPEX):** One-time investment costs incurred at the beginning of the project including:
    - Equipment and infrastructure costs (e.g., PV panels, wind turbines, battery storage units).
    - Installation and commissioning costs (e.g., labor, engineering, grid connection fees).
2. **Operational Expenditures (OPEX):** Ongoing costs incurred over the project’s lifespan
    - Maintenance and repair expenses, ensuring long-term asset performance.
    - Insurance and administrative costs, which impact profitability.
    - Energy purchasing costs, if applicable (e.g., for hybrid systems).
3. **Environmental Costs**
    - Carbon emissions and environmental impact (monetized using carbon pricing models).

5.3.2.4 Benefit Monetization

To ensure a fair comparison of investments, all energy savings, revenue streams, and environmental benefits are translated into monetary terms. The tool considers the following benefit categories:

1. **Energy Cost Savings**:
    - Self-consumption rate improvement, reducing grid electricity purchases.
    - Lower energy tariffs, achieved through optimized contracts or demand-side management.
2. **Revenue Streams**:
    - Selling surplus energy to the grid at feed-in tariff rates.
    - Participating in ancillary grid services, such as demand response programs.

3. **Environmental Benefits**:
    - Emission reductions, calculated using carbon pricing (€ per ton of CO_2 avoided).
    - Energy independence improvements, reducing reliance on external suppliers.

5.3.2.5 Cost-Benefit Analysis Computation

The tool computes several key financial indicators that provide a clear understanding of an investment's expected performance:

- **Payback Period** – This metric estimates the time required to recover the initial investment costs through accumulated operational savings. A shorter payback period indicates a quicker return on investment, making the project financially attractive.
- **Revenue Potential** – The tool projects potential income streams generated from surplus renewable energy exports and participation in grid services markets. These revenue sources can enhance the overall financial viability of the project by offsetting costs and providing additional returns.

To further support financial decision-making, the tool generates a set of **cost-benefit metrics**, allowing stakeholders to compare different investment options and assess their financial sustainability:

- **Net Present Value (NPV)** – This metric evaluates the discounted value of total expected benefits and costs over time. A positive NPV indicates a financially viable project, as the total benefits outweigh the total expenditures. The tool calculates NPV for each investment scenario, allowing stakeholders to compare multiple interventions and select the most financially viable option.

$$NPV = \sum_{t=1}^{T} \frac{(B_t - C_{op,t})}{(1+r)^t} - C_{inv} - C_{env}$$

  where:
    - B_t = Benefits from energy savings, revenue streams, and environmental incentives in year t.
    - $C_{op,t}$ = Operational costs in year t (e.g., maintenance, insurance, administrative costs).
    - C_{inv} = Initial capital expenditure (CAPEX) for infrastructure and equipment.
    - C_{env} = Environmental costs (e.g., land use, emissions pricing).
    - r = Discount rate.
    - T = Total time horizon (e.g., 20-30 years).
- **Benefit-Cost Ratio (BCR)** – By comparing total financial returns against total investment costs, the BCR provides a clear indication of whether an investment will yield a positive financial outcome. A BCR greater than 1 suggests that the benefits outweigh the costs, making the investment economically justifiable. The BCR metric helps decision-makers compare multiple investment options, prioritize projects with higher cost-efficiency, and ensure that resources are allocated to interventions with the highest financial returns. All variables retain their previous meaning.

$$BCR = \frac{\sum_{t=1}^{T} \frac{B_t}{(1+r)^t}}{C_{inv} + \sum_{t=1}^{T} \frac{C_{op,t}}{(1+r)^t} + C_{env}}$$

- **Sensitivity Analysis & Risk Assessment** – The tool incorporates sensitivity parameters to analyze how variations in key financial assumptions (such as energy prices, policy incentives, or maintenance costs) affect overall project performance. This helps decision-makers identify potential financial risks and uncertainties, ensuring more resilient investment strategies. Given that energy investments are subject to economic fluctuations, the tool performs a sensitivity analysis, testing variations in:
    - **Discount rates** (ranging from 5% to 15%).
    - **Energy price trends** (assessing market volatility).
    - **Regulatory and policy changes**, including carbon pricing adjustments.

By integrating these **financial performance indicators and cost-benefit metrics**, the tool enables energy communities to make well-informed, data-driven investment decisions, ensuring economic feasibility and long-term sustainability.

5.3.3 Evaluation Metrics & Performance Indicators

The CBA Tool evaluates investment feasibility by analyzing key financial, environmental, and operational metrics that measure the overall impact and effectiveness of proposed energy interventions. These indicators enable decision-makers to assess the profitability, cost efficiency, risk exposure, and sustainability benefits of energy projects, ensuring that investments align with both financial and strategic objectives.

The evaluation framework consists of three primary metric categories:

1. **Financial Metrics** – Assess the economic viability and return on investment (ROI) of a project.
2. **Operational Metrics** – Measure the increase in renewable energy use, self-consumption, and revenue generation from surplus energy sales.
3. **Environmental** – Quantify CO_2 emission reductions, ensuring alignment with sustainability and climate targets.

By combining these quantitative indicators, the tool provides a holistic investment evaluation, ensuring that decision-makers have transparent and reliable data to support financially and environmentally sustainable energy planning.

Financial indicators are the core assessment tools used to determine whether an energy investment is economically viable. The CBA Tool applies standard financial evaluation methods to calculate profitability, payback time, and cost-effectiveness.

Metric	Definition	Purpose	Interpretation
Net Present Value (NPV)	NPV represents the total expected financial return of an investment, discounted to the present value.	Determines whether the benefits of an investment exceed the total costs when accounting for the time value of money.	NPV > 0 → The investment is profitable. NPV < 0 → The investment is financially unviable.

Benefit-Cost Ratio (BCR)	The BCR compares total discounted benefits to total discounted costs, providing a measure of cost efficiency.	Assesses whether an investment delivers more financial value than it costs.	BCR > 1 → The investment is financially viable. BCR < 1 → The investment is not cost-effective.
Payback Period	The payback period estimates the time required for an investment to recover its initial costs through operational savings.	Helps decision-makers prioritize investments with shorter payback times to reduce financial risk.	Shorter payback periods indicate faster capital recovery and lower financial risk. Longer payback periods require additional sensitivity analysis to assess economic viability.
Internal Rate of Return (IRR)	The IRR is the discount rate at which the NPV of an investment becomes zero.	Measures the expected annual return of an investment, allowing comparisons with alternative projects.	IRR > Discount Rate → The project is financially attractive. IRR < Discount Rate → The project may not be viable.
Renewable Energy Share Increase (%)	The percentage increase in energy demand covered by renewable energy sources (RES) after the investment.	Evaluates the extent to which an investment reduces dependency on fossil fuel-based electricity.	Higher values indicate greater renewable energy penetration and improved sustainability.
Self-Consumption Increase (%)	The improvement in the percentage of renewable energy produced that is consumed directly by the energy community.	Measures the effectiveness of energy storage and demand-side management strategies in reducing grid reliance.	Higher self-consumption rates indicate better energy efficiency and lower energy procurement costs.
Revenue Streams from Surplus Energy Sales (€/kWh Sold)	The additional income generated from selling excess renewable energy to the grid or local energy markets.	Quantifies the financial return from surplus energy exports, supporting tariff and contract negotiations.	Higher revenue streams indicate greater financial sustainability and energy market participation.

CO_2 Reduction (tons avoided per year)	The total carbon dioxide emissions avoided due to increased renewable energy generation and reduced reliance on fossil fuels.	Measures the environmental benefits of an investment, supporting climate policy objectives and carbon credit mechanisms.	Higher CO_2 reductions indicate stronger environmental impact and regulatory compliance.

6 Technology Stack

The implementation of the ECOEMPOWER ICT Platform was conducted through a structured, phased approach to ensure modularity, scalability, and efficiency and will be thoroughly presented in D2.3. In this section we will describe the basic technology stack we have used to develop it.

We have integrated multiple data sources, including the Pilot Datasets discussed in Section 4, and the external weather APIs, market tariff datasets. To successfully feed the datasets into the three Energy Tools, some automated data cleaning and normalization routines have been implemented according to the needs of each tool, as described in Section 5 of this deliverable. To handle missing or faulty data, some mechanisms have been incorporated using Pandas and NumPy libraries in Python.

The ICT Tools that comprise the ICT Platform were created using Python and relevant libraries such as Pandas, SciPy, requests and PuLP and designed to be containerized to be deployed in various Platforms and infrastructures, such as the ECOEMPOWER Community and Engagement Platform developed in T4.7, potentially the different OSSs etc.

Regarding the UI, a React-based user interface has been developed with interactive energy visualization tools, including Recharts and D3.js. Dashboards have been built to present all the key evaluation metrics of the tools, as described in Section 5. The table below contains a summary of the core components and the specific technologies that we explored for the implementation.

Component	Description	Library/Technology
Data Integration & Preprocessing	Handles data ingestion, cleaning, normalization, and fault detection for pilot site datasets and external APIs.	Python: Pandas, NumPy, SciPy, requests
Energy Forecasting Tool	Provides long-term PV production forecasts (monthly) and short-term generation forecasts (hourly/day-ahead).	Python: SARIMA (Statsmodels), LSTM (TensorFlow), SciPy
Energy Modelling & Scheduling Tool	Optimizes load shifting, peer-to-peer energy trading, and renewable storage scheduling using multi-agent optimization.	Python: PuLP (Linear Programming), SciPy Optimization
Cost Benefit Analysis Tool	Evaluates investment feasibility using NPV, BCR, payback period, and sensitivity analysis.	Python: NumPy (Financial Modelling), Monte Carlo Simulations (SciPy), Matplotlib (Financial Visualization)

UI	Web dashboard for data visualization, scenario creation and financial analysis	React.js, Recharts, D3.js (Graph-Based Visualization)

7 Conclusions and Next Steps

In this deliverable we described the overview of the methodology and algorithmic processes developed to create the three ECOEMPOWER ICT Tools. The Energy Forecasting Tool was designed to predict solar PV energy generation, supporting long-term strategic planning and short-term operation adjustments. As data-driven solutions, an important aspect that was discussed was the available data collected from the REs and various external sources. The Energy Modelling and Scheduling Tool optimizes energy consumption, resource allocation, and scheduling for energy communities by maximizing self-consumption potential through a central decision-making entity, ensuring optimal community-level energy management without relying on grid operator incentives. Finally, the Cost Benefit Analysis and Decision-Making Tool was designed to help energy community leaders, aggregators, and stakeholders evaluate the financial viability of renewable energy projects by assessing costs, ROI, long-term sustainability, and sensitivity to economic variables. Specifics regarding the technology stacks that were used were also described, as well as some sample test cases that we developed to assess the different functionalities and evaluation metrics in the following months and deliverables of WP2.

Taking into consideration the project roadmap and the challenges we have faced do far, the next steps can be summarized below:

- The UI of the ICT Tools will be presented in D2.3, which will be officially submitted in M20. The decision to postpone the submission was made to offset the delays caused by the change of Pilot Partners and consequently the rejection and resubmission of D2.1 as well as the mismatch of data received from the different Pilot Sites.
- To have a unified storytelling approach for the whole of ECOEMPOWER, the integration of the technical tools of all WPs should be contained in one place. For this reason, the integration of the ICT Tools created in WP2 with the central hub for ECOEMPOWER, that is the Community Platform developed in WP4 should be prioritized and be explored in depth.
- The validation of the Energy Tools will be thoroughly documented in the next three deliverables outlined in the DoA. To facilitate the validation, dedicated workshops for the REs should be arranged to create detailed test cases and scenarios for the three Tools for all ECs. The results of these case studies will be documented in deliverables D2.4, D2.5 and D2.6.
- The development of the Tools will continue throughout the end of the project to ensure that their results are accurate, and the ECs find meaningful use for them. Each new cycle of development and refinement will also be documented in D2.4, D2.5 and D2.6.

8 References

[1] Q. Paletta, G. Terrén-Serrano, Y. Nie, J. Bieker, W. Zhang, L. Dubus, S. Dev and C. Feng, "Advances in solar forecasting: Computer vision with deep learning," *Advances in Applied Energy,* 2023.

[2] R. C. Staudemeyer and E. Rothstein Morris, "Understanding LSTM a tutorial into Long Short-Term Memory Recurrent Neural Networks," *arXiv,* 2019.

[3] L. Semmelmann, S. Henni and C. Weinhardt, "Load forecasting for energy communities: a novel LSTM-XGBoost hybrid model based on smart meter data," *Energy Informatics,* 2022.

[4] R. Deng, F. Luo, G. Ranzi, Z. Zhao and Y. Xu, "A MILP Based Two-Stage Load Scheduling: Approach for Building Load's Peak-to-Average," *Power and Electrical Engineering (ACPEE), 5th Asia Conference,* 2020.

[5] R. Rana, K. Berg, M. Z. Degefa and M. Löschenbrand, "Modelling and Simulation Approaches for Local Energy Community Integrated Distribution Networks," *IEEE Access,* vol. 10, 2022.

[6] D. E. Hafiane, A. El Magri, H. Eddine Chakir, R. Lajouad and S. Boudoudouh, "A multi-agent system approach for real-time energy management and control in hybrid low-voltage microgrids," *Results in Engineering,* vol. 24, 2024.

[7] B. B. Oliveira, M. A. Carravilla and J. F. Oliveira, "A diversity-based genetic algorithm for scenario generation," *Innovative Applications of O.R.,* 2022.

[8] ENTSO-E, "Implementation Guidelines for TYNDP 2022 based on 3rd ENTSO-E guideline for cost benefit analysis of grid development projects," TYNDP, 2023.

A. List of Tables

B. List of Figures