iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open-Source Toolbox

Specifications for data sharing and tools integration

Deliverable title

Deliverable D1.1 - Specifications for data sharing and tools integration

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Document information

Deliverable Number and Name	D1.1: Specifications for data sharing and tools integration
Work Package	WP1: Multi-Physics component models: Implementation and development
Dissemination Level	Public
Author(s)	Teams from SINTEF, TEC, E3M, UD, TUB, IIASA, TUW, EDF, AAU, LE2C, ESO, IPTO, PPC, ACLIMA
Primary Contact and Email	NTNU – Pedro Crespo del Granado, pedro@ntnu.no
Date Submitted	30/05/2025
Status	Final.
Reviewed by (if applicable)	Validated through multiple reviews for different datasets created as part of the deliverable

This document has been prepared in the framework of the European project iDesignRES. This project has received funding from the European Union's Horizon Europe programme under grant agreement no. 101095849.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This deliverable presents the specifications for data sharing and tools integration within the iDesignRES project, establishing a relevant step for open-source energy system modeling at the NUTS Level 2 regional scale across Europe.

Project Objectives: iDesignRES aims to provide public authorities and network operators with modular open-source toolboxes for planning and optimizing renewable energy uptake and enhancing resilient infrastructure at regional, national, and European scales. The project develops optimized long-term planning tools operating at NUTS Level 2 granularity while building on existing European databases to standardize data flows.

Data sharing and tools: The project leverages the IIASA database infrastructure and adopts the IAMC timeseries data format for scenario analysis and model comparison. A centralized Scenario Explorer platform enables transparent data exchange, validation, and collaborative analysis across consortium partners. The standardized approach ensures consistency across models and supports structured reporting of energy, emissions, technologies, and socioeconomic drivers.

Key Deliverable results: Comprehensive datasets have been collected and validated across multiple energy sectors including electricity generation, gas networks, renewable energy sources, industry, transport, and buildings. Specific high-resolution datasets for Spain and the North Sea region have been developed, covering transmission grids, generation assets, demand patterns, and gas infrastructure. All data has been aggregated to NUTS Level 2 level and made publicly available through Zenodo under open licenses. Seven datasets are currently published on Zenodo and eight in iDesignRES scenario explorer, with ongoing efforts to enhance regional granularity.

1. Introduction

1.1 The iDesingRES concept and coordination of data collection

iDesignRES aims to provide public authorities and network operators with modular open-source toolboxes for planning and optimizing the uptake of low and zero emission energy sources and enhancing resilient infrastructure including energy storage at regional, national, and European scales.

The project's core objectives include:

1. **Develop an optimized open-source long-term planning tool** for energy systems at pan-EU level operating at NUTS Level 2 granularity
2. **Build on existing European scale open databases** to standardize data flow and optimize database updates
3. **Develop open-source multi-physics component models** at NUTS Level 2 covering various energy system elements
4. **Develop tools to assemble component models** and improve energy systems modeling
5. **Provide network operators and public authorities with open-source models and visualization tools** compatible with JRC visualization tools
6. **Validate the models and tools** through certification methodologies and 5 European real-life use cases

Data Collection Importance in Task 1.1

Data collection, coordination, and harmonization are crucial to the iDesignRES project as they form the foundation for all modeling activities. The project builds upon the IIASA ixmp database infrastructure, which has been used in IPCC reports and numerous EU research projects (e.g., ECEMF, openENTRANCE, and others).

Key aspects of data collection importance include:

- Creating a **standardized exchange of knowledge and data** among consortium partners and models.
- Ensuring **open accessibility of data** and transparency on methods or source.
- Supporting the **NUTS Level 2 granularity** requirements of the models.
- Enabling **interaction between models** across different scales and discuss assumptions-limitations on data availability
- Mapping a **comprehensive infrastructure bridging sectors and actors**
- Supporting the **modularity and open-source nature** of the project's tools and models

1.2 Scope and report structure

This deliverable focuses on data collection, coordination, and harmonization at NUTS Level 2 granularity to support energy system modeling activities in the iDesignRES project. The scope encompasses comprehensive data mapping across multiple energy sectors including electricity generation and transmission, gas and hydrogen infrastructure, district heating, and end-use sectors (industry, transport, buildings). The work builds upon the IIASA database infrastructure and adopts the IAMC timeseries data format to ensure standardized, transparent data exchange among consortium partners.

Specific focus areas include detailed data collection for two primary use cases: Spain's power and gas systems, and the North Sea region's electricity infrastructure. The deliverable also addresses data aggregation and disaggregation challenges when converting between different spatial resolutions, particularly from plant-level or national data to NUTS Level 2 regional granularity.

The report is organized into eight main chapters, the following chapter covers a detailed description of the Scenario Explorer database infrastructure and IAMC format implementation.

Chapter 3 provides a comprehensive overview of NUTS Level 2 data mapping across various energy technologies and sectors. Chapters 4 and 5 present detailed datasets for the Spanish and North Sea use cases respectively, covering electricity grids, gas networks, and aggregation methodologies. Chapter 6 addresses end-use sectors with particular focus on industry, transport, and building stock data. The report concludes with outreach activities, including participation in the European Open Energy Modelling Workshop, and summarizes the established data infrastructure that supports the project's modeling and tool development activities.

2. iDesignRES and the “Scenario Explorer” app infrastructure

The objective of this chapter is to introduce and describe the database developed for the iDesignRES project. This database serves as the Scenario Explorer—a central platform where diverse datasets are stored, validated, and utilized in various demonstration cases. The chapter provides a comprehensive overview of the key features and primary functions of the iDesignRES Scenario Explorer. Additionally, it summarizes the collaborative efforts undertaken to harmonize definitions and establish consistent nomenclature for naming conventions within energy system modelling.

Figure 1: Landing page of the iDesignRES scenario explorer

2.1 General Introduction

For numerous Horizon 2020 and Horizon Europe projects, IIASA serves as a central data repository during the model comparison phase, supporting the validation and analysis of scenario results across modeling teams from different research institutions.

The database infrastructure facilitates the harmonization of input data and model results across different model families. In addition, it provides automated post-processing workflows, such as translating emissions pathways from energy system models into long-term temperature outcomes using simple climate models.

The iDesignRES Scenario Explorer builds on this infrastructure to enable transparent, efficient, and collaborative scenario analysis. It allows users to browse, analyze, and compare model results through an intuitive web interface, directly connected to a centralized scenario database. It is available through https://idesignres-internal.apps.ece.iiasa.ac.at .

2.2 The IAMC Data Format and Project Codelists

The iDesignRES project uses the IAMC timeseries data format for scenario analysis and model comparison. This standardized format ensures consistency across models and supports structured reporting of inputs, assumptions, and results related to energy, emissions, technologies, and socio-economic drivers.

For a detailed description of the IAMC data format, please refer to:

https://docs.ece.iiasa.ac.at/iamc.html#iamc-format

The project also defines a set of codelists-standardized lists of regions, variables, units, and other key elements-to ensure harmonized reporting. These codelists are typically maintained in a version-controlled GitHub repository.

An example of an IAMC-formatted variable is: Capacity|Electricity|Biomass. Each of these variables has a unique definition within the respository.

This leads to data in the shape:

model	**scenario**	**region**	**variable**	**unit**
EMPIRE	1.5degree	Austria	Capacity|Electricity|Biomass	GW

2.3 The Database Ecosystem

The iDesignRES Scenario Explorer is built on the IIASA database infrastructure, which leverages the following open-source Python packages:

ixmp4 - a backend framework for scenario data storage and access

Documentation: https://docs.ece.iiasa.ac.at/ixmp4

nomenclature - a utility package for validating and processing regional data

Documentation: https://nomenclature-iamc.readthedocs.io/

These tools support consistent validation, region aggregation, and formatting of scenario data. The ixmp4 package also provides a REST API, allowing external tools and web applications to programmatically query and interact with scenario data stored in the central database.

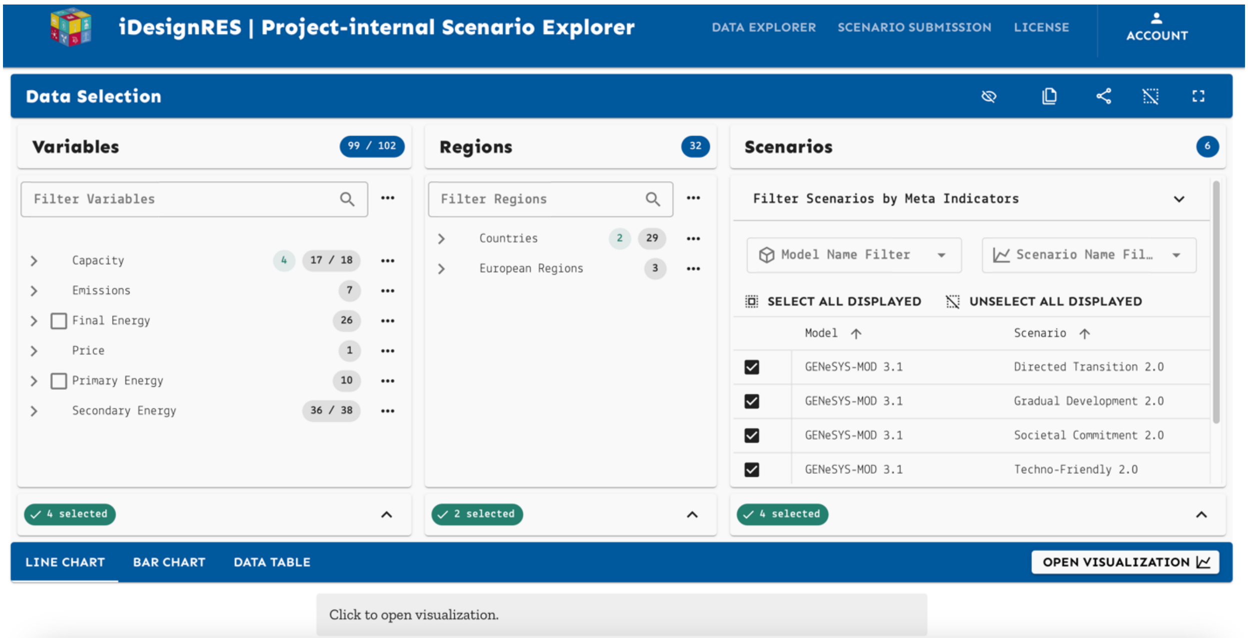


Figure 2: Data selection within the scenario explorer

This ecosystem ensures that scenario data is not only stored securely but also easily accessible and interoperable with other platforms and tools used in energy and climate research.

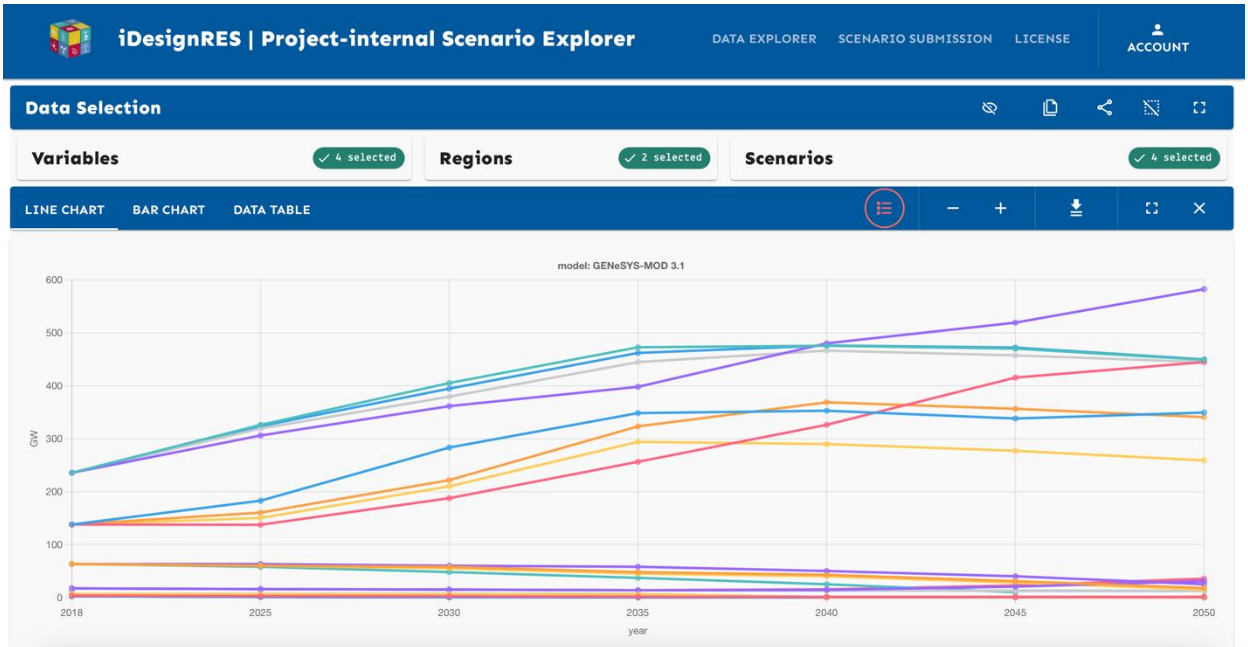
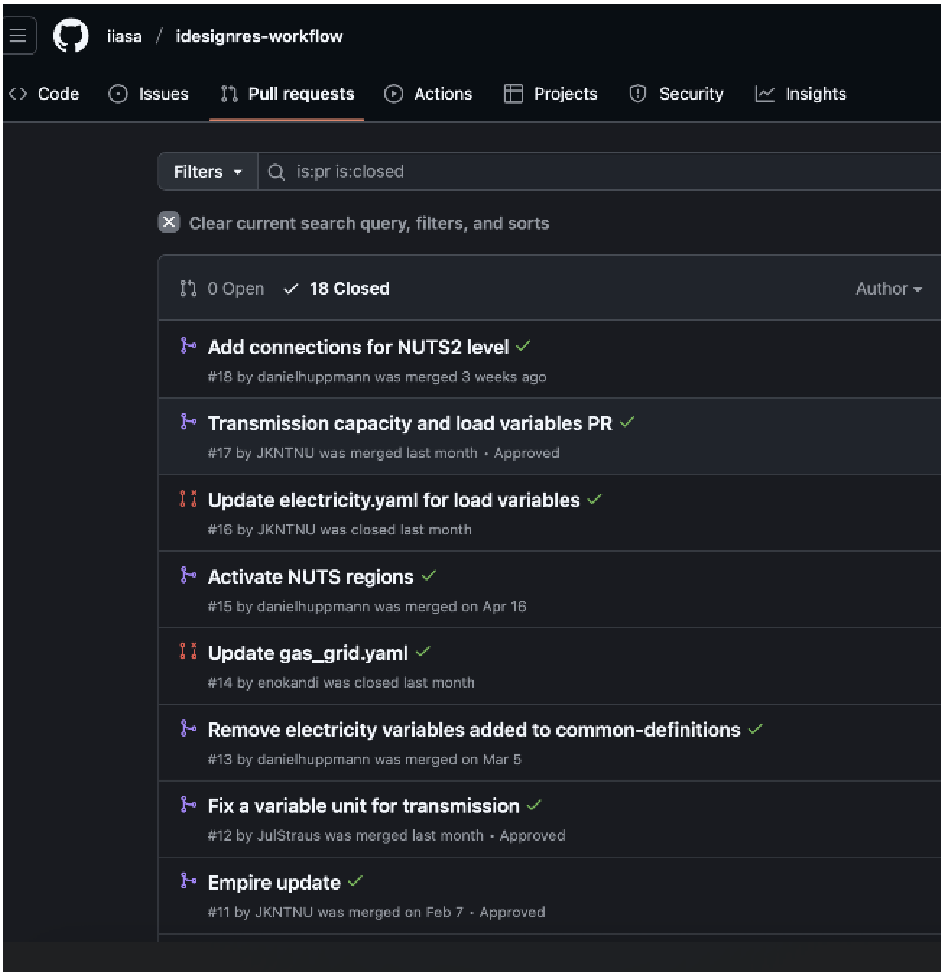


Figure 3: Data visualization within the scenario explorer

Figure 3 illustrates the scenario explorer. It enables users to filter for specific models, scenarios and variables to be shown, including a visualization of the respective time series. The scenario explorer also allows to download the given data in different formats (CSV, Excel, …).

2.4 iDesignRES workflow for nomenclature definition



The iDesignRES-project has set up a specific repository: https://github.com/iiasa/idesignres-workflow where the relevant variables and regional definitions are stored (see Figure above). In this repository, tools and instructions allow the adoption of the IAMC Common-Definition, with the following features and actions:

- **Data Management Plan:** The project's scenario processing and validation workflows are built to use variables from the IAMC common-definitions repository as the primary reference for variable names, units, and definitions (Agreed and outlined on deliverables D2.1 and D7.2).
- **Project-Specific Extensions:** Where higher technological detail or additional variables are required (e.g., specific energy infrastructure or technology types), iDesignRES is extending the IAMC list. New variables are proposed via pull requests to the iDesignRES workflow repository, which are then reviewed and, if appropriate, contributed upstream to the IAMC repository.
- **Templates and Codelists:** iDesignRES in WP2 coordinates an Excel-based template (available for download) that lists allowed variables and regions, drawing from the IAMC repository and project-specific extensions.
- **Validation Tools:** The project uses the nomenclature package and validation scripts to check model data against the common-definitions, ensuring consistency before data is uploaded to the project's Scenario Explorer or shared externally.

Model Registration and Data Submission

- **Model Registration:** Each model participating in iDesignRES registers its variable and region mappings according to the common-definitions template. This process includes:
  - Forking the workflow repository
  - Completing model and region mapping files as per the nomenclature documentation
  - Submitting a pull request for review and integration
- **Data Conversion Scripts:** Custom Python scripts (e.g., iamcconversioninput.py) are used to convert raw model outputs into IAMC-compliant formats, applying mapping rules for technology names, fuels, and regions. These scripts ensure that all data adheres to the shared definitions before integration.

In short, data are stored within the IIASA open database infrastructure and made accessible through the iDesignRES Scenario Explorer, Zenodo, and other open platforms, ensuring traceability via persistent identifiers. All variable definitions, data templates, and scenario results are published under open licenses such as CC BY 4.0 to maximize reuse and facilitate open science practices.

3. Mapping Nuts-Level2 data in iDesingRES

This chapter outlines the purpose and the overview data mapping at the NUTS Level 2 scale, illustrating the diversity of datasets across specific energy sectors, sources, and technologies. It reviews key data sources and discusses aggregation and disaggregation approaches required for NUTS Level 2 data, providing general examples of data mapping and collection for selected energy sectors or components. Subsequent chapters (Chapters 4 and 5) offer a detailed description of the features and availability of the data collected in the project, available at: https://zenodo.org/communities/idesignres/ .

3.1 Role of use cases and demonstrators

Use cases and demonstrators are central to the iDesignRES project, serving as real-world testbeds for validating and refining the project's energy system modelling tools. These use cases—spanning macro-regions like the North Sea and South-Eastern Europe, as well as major industrial clusters such as Lombardy and the Greater Basque Country— represent diverse geographical, technical, and socio-economic contexts across Europe (see Figure 4 for an overview). By implementing and testing the models in these varied environments, iDesignRES ensures that its solutions are robust, adaptable, and aligned with the practical needs of public authorities, network operators, and industrial stakeholders. These use cases provide a regional and more granular understanding of renewable energy deployment and the transformation of the energy sector, hence all are analyzed and implemented at NUTS level2 or finer.

Demonstrators function as pilots, enabling the project team to co-design, test, and validate scenarios with local partners. This iterative process not only verifies the technical soundness and operational feasibility of the models but also ensures that the tools meet end-user requirements. Feedback from these real-world applications is used to refine the models, enhance their modularity, and facilitate their adaptation to different regional or sectoral challenges.

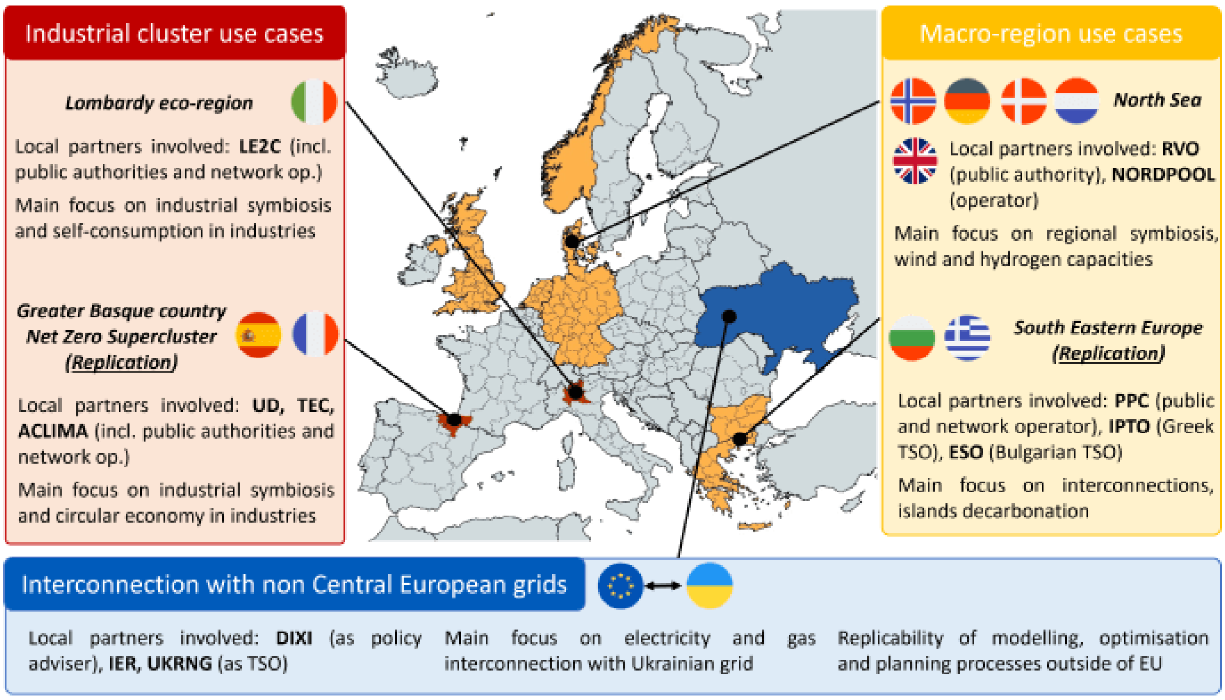


Figure 4: Overview and description of use cases

Importance of Data Mapping Availability and Use at NUTS Level 2

Data mapping, collection and validation at NUTS Level 2 is central to iDesignRES models and analyses planned for the demos. NUTS Level 2 provides a high-resolution spatial framework that enables detailed modelling of energy systems, infrastructure, and demand at the regional level—crucial for effective planning and optimisation of multi-carrier energy networks (electricity, gas, heat, hydrogen). By operating at this granularity, iDesignRES models capture local climate, socio-economic conditions, and sectoral specifics, which are often lost in more aggregated national models.

In this regard Task 1.1, looks at the availability, mapping and standardization of data at NUTS Level 2 to allow WP2 perform accurate geolocation of investments, optimise infrastructure allocation, and assess operational resilience under various scenarios. As described in the previous chapter, the adoption of open data standards (such as the IAMC format) and integration with the IIASA database infrastructure ensures consistency, interoperability, and transparency across all models and use cases.

3.2 Overview of Data sources at NUTS Level 2 or finer

The energy systems data at the NUTS Level 2 relies heavily on Eurostat, the EU's central statistical office, which provides harmonized datasets on electricity production, consumption, and renewable energy shares[1]. Eurostat's energy databases, such as the "Energy Statistics - An Overview" and "Electricity and Heat Statistics," offer granular data at regional levels, accessible via APIs. For instance, the Hotmaps Project utilized Eurostat's structural business statistics and climate data to model heating and cooling demand at NUTS-2, combining these with high-resolution spatial data (e.g., 100x100m grids) to estimate building stock characteristics. For the iDesignRES, these sources will be important as they are able to:

- Eurostat's energy databases, such as **nrg_quant (**energy quantities) and **nrg_bal** (energy balances), provide structured annual data on production, consumption, and renewable energy shares across NUTS regions[2]. For example, renewable energy statistics at NUTS Level 2 are accessible through disaggregated datasets that track solar, wind, and hydroelectric contributions[1][4]. Eurostat's regional metadata and GISCO geospatial tools further enable spatial analysis of energy infrastructure[3].
- ENTSO-E Transparency Platform: This platform offers granular data on electricity generation, transmission, and consumption at hourly intervals[4]. While primarily national in scope, ENTSO-E's API allows partial disaggregation to NUTS Level 2 by cross-referencing grid nodes and regional demand patterns

The examples outlined above establish a foundation for understanding the key datasets currently available at the NUTS 2 level or with even finer spatial granularity. Nevertheless, when analyzing specific energy technologies, sectors, or carrier infrastructures, it becomes evident that the spatial resolution and data availability can vary significantly. In many instances, the administrative boundaries defined by NUTS Level 2 do not align precisely with the physical layout or operational regions of energy infrastructures—such as electricity transmission networks, district heating systems, or gas pipelines. This misalignment

[1] https://ec.europa.eu/eurostat/statistics-explained/index.php/Energy_statistics_-_an_overview
[2] https://ec.europa.eu/eurostat/cache/metadata/en/nrg_bal_esms.htm
[3] https://ec.europa.eu/eurostat/web/gisco/geodata/statistical-units/territorial-units-statistics
[4] https://www.entsoe.eu/data/transparency-platform/

can present challenges in accurately mapping and analyzing energy flows, infrastructure capacities, and technology deployments at the regional scale.

For example, large-scale power plants or wind farms may serve multiple NUTS 2 regions, while certain grid assets—such as substations or interconnectors—might be situated near regional borders, complicating the attribution of capacity and output data. Similarly, energy carrier networks (e.g., electricity, gas, hydrogen) are often planned and operated according to technical and economic criteria that do not coincide with administrative divisions. As a result, the granularity and relevance of available data can differ substantially depending on the sector or technology under consideration.

To provide further clarity, the following table present a summary of selected energy technologies, sectors, and network types, highlighting their respective data sources and the typical spatial resolution at which information is available. This overview is intended to illustrate the diversity of data landscapes across the energy system and to underscore the importance of sector-specific approaches when collecting, harmonizing, and interpreting regional energy data.

Table 1: Examples of Energy Technologies, Sectors, and Networks with Potential Data Sources and Typical Spatial Resolution

Category	Example Technologies/Sectors	Typical Data Sources	Typical Spatial Resolution	Notes on Alignment with NUTS 2
Electricity Generation	Wind, Solar, Hydro, Nuclear	Eurostat, ENTSO-E, National TSOs	NUTS 2, plant-level, grid node	Plant locations may cross NUTS borders
Gas Infrastructure	Pipelines, Storage	ENTSOG, National Gas Operators	National, subnational, network	Pipelines often span multiple regions
District Heating	Urban Heat Networks	Local authorities, Euroheat & Power	City, municipality	Often finer than NUTS 2
Renewable Heating/Cooling	Biomass, Heat Pumps	Eurostat, National Energy Agencies	NUTS 2, NUTS 3	Data may be aggregated for privacy
Transport Fuels	EV Charging, Biofuels	European Alternative Fuels Observatory, Eurostat	NUTS 2, national, station-level	Infrastructure may not match NUTS 2

A key challenge addressed in this Deliverable (D1.1) and in Task 1.1 has been the identification and coordination of relevant datasets across each energy technology and sector. This process involves not only cataloguing existing data sources but also assessing their suitability for mapping, analysis, and integration into broader regional studies. The diversity of available datasets is particularly valuable for supporting specific use cases—such as scenario development for regional decarbonization pathways—and for underpinning robust, evidence-based policy recommendations.

Moreover, these additional datasets can play a critical role in validating key assumptions, as well as in performing aggregation or disaggregation procedures necessary to convert data between different NUTS levels (e.g., from NUTS Level 0 to NUTS Level 2, or vice versa). Such conversions are essential for ensuring data consistency and comparability across regions, and for enabling meaningful cross-sectoral and multi-scale analyses.

The **following tables (below) present** further examples of the mapping exercise, highlighting relevant datasets for selected technologies and sectors. This mapping has served as a foundation for the systematic collection and consolidation of data, which is further detailed in subsequent chapters. Additionally, the insights gained from this exercise inform and support ongoing activities in other work packages (WP1, WP2, and WP3), ensuring a coherent and integrated approach to regional energy system analysis throughout the project.

Solar Power Time Series

Data name	**Source**	**Comments, Explanation or specifics**	**pros and cons for Nuts Level 2**
Solar power time series	PVGIS	Provides information on solar radiation and photovoltaic system performance for any location in the world, except the North and South Poles	Data for single PV panel
Solar power time series	Renewables Ninja	API to MERRA-2 reanalysis data with hourly solar radiation data for any location in the world. Useful for investment and operational analyses. Can be converted to power output via assumptions on panel efficiency, orientation and tracking.	Data for single PV panel

Wind Power time Series

Data name	**Source**	**Comments, Explanation or specifics**	**pros and cons for Nuts Level2**
Wind power time series	EMHIRES project	Publically available European wind power generation dataset derived from meteorological sources that is available up to NUTS-2 level – provides meteorologically derived wind power time series at high temporal and spatial resolution to improve the assessment of the wind resource at the precise location of wind farms.	Created specifically for NUTS2 detail level.
Wind power time series	Renewables Ninja	API to MERRA-2 reanalysis data with hourly wind speed and wind power data for any location in the world. Useful for investment and operational analyses.	Allows fast and easy download of data. Conversion from wind speed to wind power within a NUTS2 region can be done with various levels of detail

Wind power time series	windatlas.xyz	API to ERA-5 reanalysis data with hourly wind speed and wind power (and other) data. Useful for investment and operational analyses.	Covers only 200 km offshore, such that the middle of the North Sea is not included. Allows fast and easy download of data. Conversion from wind speed to wind power within a NUTS2 region can be done with various levels of detail

Natural gas and network

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
Power plants burning natural gas	GEM	GEM (Global Energy Monitor) provides a comprehensive list of power plants around the world, including their exact locations, capacities, fuel types, technologies, and more.	Although it is one of the most comprehensive datasets of power plants, it lacks any information about the generation costs of the units. Moreover, it is not entirely complete.
Power plants burning natural gas	OSM	OSM (OpenStreetMap) provide a comprehensive data of electrical infrustructures including power plants around the world. It contains exact location, capacity, using fuels, technologies, etc.	It is one of the main sources for any other dataset of electrical infrastructure. However, the data has not been reviewed or cleaned, and it is also incomplete.
Power plants burning natural gas	PyPSA-Eur	The Prebuilt Electricity Network for PyPSA-Eur, based on OpenStreetMap data, provides a highly detailed dataset of the power systems of European countries, including comprehensive information on individual power plants across the continent. To create the dataset, various sources such as GEM and OSM have been used.	Although it is one of the most comprehensive datasets of power plants, it is not entirely complete.
Total installed capacity	ENTSO-E	The ENTSO-E Transparency Platform provides the total installed capacity of different technologies for European countries.	It is a reliable source but only provides country-level aggregated data.
SciGrid Gas	SciGRID Gas: Gas Infrastructure Data, [Online]. Available:	Contains data on production, power plant capacities, and demand (min, max, average, median hourly values); also includes pipelines, storage, and	Pros: Broad dataset, includes hourly resolution, good for modeling. Cons: Some data imputed via heuristics; pipeline

	https://www.gas.scigrid.de/	LNG terminals. Estimated missing data via AI.	detail limited for multi-region lines.
ENTSOG	ENTSOG: European Network of Transmission System Operators for Gas, [Online]. Available: https://www.entsog.eu/	Not directly elaborated in detail, but implied to provide gas infrastructure and possibly system operations data.	Pros: Likely reliable, standardized data. Cons: Specific granularity and NUTS2 alignment unclear.
GSE Storage Data	Storage Database - Gas Infrastructure Europe (GIE), [Online]. Available: https://www.gie.eu/transparency/databases/storage-database/	Contains working gas volumes, withdrawal capacities, storage sites (operational, under construction, and planned).	Pros: Rich storage-specific data. Cons: May need spatial matching to NUTS2 level.

Coal power plants

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
Power plants burning coal	GEM	GEM (Global Energy Monitor) provides a comprehensive list of power plants around the world, including their exact locations, capacities, fuel types, technologies, and more.	Although it is one of the most comprehensive datasets of power plants, it lacks any information about the generation costs of the units. Moreover, it is not entirely complete.
Power plants burning coal	OSM	OSM (OpenStreetMap) provide a comprehensive data of electrical infrustructures including power plants around the world. It contains exact location, capacity, using fuels, technologies, etc.	It is one of the main sources for any other dataset of electrical infrastructure. However, the data has not been reviewed or cleaned, and it is also incomplete.

Power plants burning coal	PyPSA-Eur	The Prebuilt Electricity Network for PyPSA-Eur, based on OpenStreetMap data, provides a highly detailed dataset of the power systems of European countries, including comprehensive information on individual power plants across the continent. To create the dataset, various sources such as GEM and OSM have been used.	Although it is one of the most comprehensive datasets of power plants, it is not entirely complete.
Total installed capacity	ENTSO-E	The ENTSO-E Transparency Platform provides the total installed capacity of different technologies for European countries.	It is a reliable source but only provides country-level aggregated data.

Oil power plants

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
Power plants burning oil	GEM	GEM (Global Energy Monitor) provides a comprehensive list of power plants around the world, including their exact locations, capacities, fuel types, technologies, and more.	Although it is one of the most comprehensive datasets of power plants, it lacks any information about the generation costs of the units. Moreover, it is not entirely complete.
Power plants burning oil	OSM	OSM (OpenStreetMap) provide a comprehensive data of electrical infrustructures including power plants around the world. It contains exact location, capacity, using fuels, technologies, etc.	It is one of the main sources for any other dataset of electrical infrastructure. However, the data has not been reviewed or cleaned, and it is also incomplete.
Power plants burning oil	PyPSA-Eur	The Prebuilt Electricity Network for PyPSA-Eur, based on OpenStreetMap data, provides a highly detailed dataset of the power systems of European countries, including comprehensive information on individual power plants across the continent. To create the dataset, various sources such as GEM and OSM have been used.	Although it is one of the most comprehensive datasets of power plants, it is not entirely complete.
Total installed capacity	ENTSO-E	The ENTSO-E Transparency Platform provides the total installed capacity of different technologies for European countries.	It is a reliable source but only provides country-level aggregated data.

Biomass

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
Power plants burning biomass	GEM	GEM (Global Energy Monitor) provides a comprehensive list of power plants around the world, including their exact locations, capacities, fuel types, technologies, and more.	Although it is one of the most comprehensive datasets of power plants, it lacks any information about the generation costs of the units. Moreover, it is not entirely complete.
Power plants burning biomass	OSM	OSM (OpenStreetMap) provide a comprehensive data of electrical infrustructures including power plants around the world. It contains exact location, capacity, using fuels, technologies, etc.	It is one of the main sources for any other dataset of electrical infrastructure. However, the data has not been reviewed or cleaned, and it is also incomplete.
Power plants burning biomass	PyPSA-Eur	The Prebuilt Electricity Network for PyPSA-Eur, based on OpenStreetMap data, provides a highly detailed dataset of the power systems of European countries, including comprehensive information on individual power plants across the continent. To create the dataset, various sources such as GEM and OSM have been used.	Although it is one of the most comprehensive datasets of power plants, it is not entirely complete.
Total installed capacity	ENTSO-E	The ENTSO-E Transparency Platform provides the total installed capacity of different technologies for European countries.	It is a reliable source but only provides country-level aggregated data.

Uranium and nuclear power plants

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
Nuclear Power plants	GEM	GEM (Global Energy Monitor) provides a comprehensive list of power plants around the world, including their exact locations, capacities, fuel types, technologies, and more.	Although it is one of the most comprehensive datasets of power plants, it lacks any information about the generation costs of the units. Moreover, it is not entirely complete.
Nuclear Power plants	OSM	OSM (OpenStreetMap) provide a comprehensive data of electrical infrustructures including power plants around the world. It contains exact	It is one of the main sources for any other dataset of electrical infrastructure. However, the data

		location, capacity, using fuels, technologies, etc.	has not been reviewed or cleaned, and it is also incomplete.
Nuclear Power plants	PyPSA-Eur	The Prebuilt Electricity Network for PyPSA-Eur, based on OpenStreetMap data, provides a highly detailed dataset of the power systems of European countries, including comprehensive information on individual power plants across the continent. To create the dataset, various sources such as GEM and OSM have been used.	Although it is one of the most comprehensive datasets of power plants, it is not entirely complete.
Total installed capacity	ENTSO-E	The ENTSO-E Transparency Platform provides the total installed capacity of different technologies for European countries.	It is a reliable source but only provides country-level aggregated data.

Transmission grid for the power system

Data name	**Source**	**Comments, Explanation or specifics**	**pros and cons for Nuts Level 2**
OSM grid data	OSM	"OSM (OpenStreetMap) provides comprehensive data on electrical infrastructure, including power transmission networks across the world. It contains almost all high-voltage transmission lines in Europe, the exact locations of substations, geographical paths of lines, and more.	The data is very comprehensive and highly detailed, but the data is raw, unprocessed, and cannot be used directly.
PyPSA-Eur OSM based grid	PyPSA-Eur	The Prebuilt Electricity Network for PyPSA-Eur, based on OpenStreetMap data, provides a highly detailed dataset of the power systems of European countries, including comprehensive information on the transmission grid across the continent.	The data is cleaned, processed, and ready to use. However, line parameters are estimated and may not be accurate. Additionally, there is some level of aggregation in the data.
ENTSO-E	ENTSO-E	The dataset provides detailed information on the transmission system network operated by members of the European Network of Transmission System Operators.	The data is processed, validated, and considered reliable. However, the network elements are not positioned at their exact geographic locations. Moreover, the data is not easily downloadable. The last officially

			available downloadable version dates back to 2016.

Hydrogen network

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
IEA Hydrogen	IEA, International Energy Agency, Hydrogen Production and Infrastructure Projects Database, [Online]. Available: https://www.iea.org/data-and-statistics/data-product/hydrogen-productionand-infrastructure-projects-database#hydrogen-production-projects	Includes 2023 Hydrogen Production and Infrastructure Projects (corrected/updated). Some overlap between the two files.	Pros: Recent and detailed hydrogen project data. Cons: Overlapping files may cause redundancy or confusion; NUTS2 mapping may require geospatial interpolation.

3.3 Examples of data collected

As part of the activities in WP1, there were various efforts with concrete data collection outputs for specific energy system components. These efforts focused on validating and cross complementing the datasets as the ones with other data mentioned in the previous section, Here there is a summary of three efforts developed and carried out by iDesignRES partners.

3.3.1 Hydropower potentials

iDesginRES has develop a database of hydro potentials for the European countries at a NUTS0 and NUTS2 level. The work included some countries that are not part of the EU, but belong to the Energy Community Contracting Parties, or they are included by EUROSTAT as Statistical Regions (SR) in the NUTS2024 and NUTS2021 database. As a result, an excel file was completed, containing inflow time series for each country, using a daily timestep, consisting of 365 values for dry, average, and wet years. This data is based on a selection of 36 climate years. The inflow time series has been normalized according to the annual inflow, meaning the sum of all values for each country equals. This allows any modeler to use these patterns and, by multiplying them with the assumed annual inflow (whether at the level of a plant, a country, or a region), they can derive the daily inflows in terms of energy.

There were also data discussed with other EU projects for high spatial data on hydropower inflows: https://www.hydrosheds.org/products/hydrobasins

3.3.2 Solar data

iDesignRES has developed a solar power model providing time-series of annual available thermal energy generation capacity in Concentrating Solar Power (CSP) plants and electrical energy generation in large Photovoltaic (PV) plants on hourly basis, for a given NUTS2 region and a specific amount of investment for each of these technologies. iDesignRES has also developed a module for the preprocessing of solar data for any NUTS2 and NUTS3 level region of Europe. The resulting energy profiles are provided aggregated at NUTS2 level and disaggregated at NUTS3 level, representing solar resource variability, not only along the time, but also along the geography of each NUTS3 region. This has demonstrated the accuracy of this approach comparing this solar power model results with PV generation profile in Spanish Peninsular region for 2020 provided by ESIOS and obtaining an estimation error lower than 0.1% for the cumulative annual generation.

3.3.3 Heat demands: potentials for district heating and waste heat

Data collected under a high-resolution spatial dataset quantifying heat demand, district heating (DH) potentials, and waste heat sources across Europe at the NUTS2 regional level. The work addresses both data collection and advanced GIS-based modeling, supporting the planning and integration of renewable energy sources into the European heating sector. By quantifying both the demand and the supply-side potentials (including untapped waste heat), the work supports informed decisions on where district heating expansion and waste heat integration are most viable and cost-effective. The open-source nature of the data facilitates its use by researchers, planners, and policymakers across Europe.

Scope and Coverage: The dataset encompasses all EU27 countries and the United Kingdom, providing aggregated estimates for:

- Heat demand in residential and service buildings.
- District heating potential, covering residential, service, and industrial sectors, with associated investment costs.
- Waste heat potential, disaggregated by source type and temperature level.

A key assumption in the analysis is a 55% market share for district heating across the EU27, with national shares varying according to local potential

The methodology for this analysis builds on previous projects, such as sEEnergies and Heat Roadmap Europe, and incorporates updates from the Pan-European Thermal Atlas (PETA). Heat demand estimation is achieved by spatially disaggregating demand according to the distribution of built areas and construction periods, resulting in refined regional estimates. Zonal statistics are then computed using Python GIS tools to sum heat demand within each NUTS2 boundary. For district heating (DH) potential, the study utilizes PETA6 data at a 100-meter resolution to identify potential DH areas and estimate investment costs, applying standardized cost factors derived from German data. Industrial heat demand is spatially allocated using CORINE land use data and national energy statistics. Areas are prioritized for DH deployment based on the ratio of specific energy costs to heat demand, with the aim of achieving a cumulative 55% market share in the EU27. In the UK, a cost threshold approach defines feasible DH areas, aligning with the highest observed investment costs in the EU27. Waste heat potential is mapped using a top-down approach that complements existing point data, with waste-to-energy (WtE) potentials downscaled to cities based on population and industrial area. Industrial surplus heat is categorized by temperature and spatially allocated to DH areas, while additional sources such as wastewater treatment, supermarkets, and metro systems are mapped using project-specific datasets. All waste heat potentials are then aggregated at the NUTS2 level, facilitating regional analysis and planning.

4. Data Gathered for Use cases: Spain

The objective of this chapter is to describe the data collected and adapted to represent the Spanish power system and the gas-h2 infrastructure. The Spanish Test Case in iDesignRES serves as an early, high-resolution demonstration to validate, implement, and test some multi-physics component models and data workflows for Spain's energy system. It focuses on power, gas, hydrogen, transport, and buildings sectors, enabling robust model certification and scenario analysis at regional and city levels.

4.1 *Electricity Grid*

This dataset focuses of the Spain's electric power system, comprises on the transmission grid at voltage levels of 220 kV and above, provided in its actual form without aggregation or loss of spatial resolution. Additionally, it contains comprehensive generation data for all conventional units with capacities exceeding 1 MW, along with the most accurate possible estimation of renewable generation distribution across the network. While electricity demand time series for the entire country were sourced from the ENTSO-E Transparency Platform, the spatial distribution of consumption across the grid was estimated using population distribution data.

4.1.1.1 Transmission Grid

To obtain transmission network data we used OSM [1] since it provides accurate geographic details of the network. The downloaded data required cleaning due to inconsistencies and missing connections among lines, particularly at substations' busbars. Additionally, there were isolated or missing sections of the network that needed to be corrected.

The first step in cleaning the data involved using an open-source tool called GridTool [2] programmed in MATLAB and specifically developed for processing and cleaning transmission grid data obtained from OSM. However, the GridTool output still required manual fixes due to the absence of some lines and buses in the original data. The manual fixes were done by comparing the resulted grid of GridTool, with the published transmission grid data by ENTSO-E.

4.1.1.2 Generation

To obtain the most accurate and comprehensive data on power plants in the country, we used (Global Energy Monitor) GEM [3] data. To validate the data obtained, we compared the total generation capacity of each type of generation unit with the official data published by the Spanish government (REE). Although we used GEM data as the primary source, we supplemented and improved the data for certain generation technologies using OSM data.

4.1.1.3 Demand

The demand data consists of both a static and a time-dependent component. The static component addresses how the load is distributed across the grid, while the time-dependent component represents the actual hourly load values over a specific period.The time series of total electricity consumption for Spain has been gathered from ENTSO-E Transparency Platform for 2024.

The primary data source used for the static component of demand data is the EU Energy Atlas. This dataset consists of a series of maps of the EU, representing the demand for major energy product groups across various economic activities. The data is based on the 2019 energy balances and the 1 km^2 reference grid from EUROSTAT and the JRC's Energy and Industry Geography Lab. The EU Energy Atlas estimates electricity consumption in EU countries with a spatial resolution of 1 square kilometer. It categorizes consumption by various sectors, such as industry and transportation, while also offering aggregated electricity consumption across all sectors.

Although this dataset is based on 2019 data, here, it is assumed that, while total electricity consumption may have changed since then, the relative spatial distribution of electricity consumption across Europe has remained approximately constant. To estimate the relative load distribution within the grid, the electricity demand from each 1 km² pixel in the Atlas data is aggregated and assigned to the nearest transmission grid bus. Figure 5 illustrates the load distribution estimation across the transmission grid of Spain based on EU Energy Atlas data. The shown values in the figure are scaled proportionally so that the total matches Spain's electricity consumption in 2024, as reported by ENTSO-E Transparency Platform [4].

Regarding the load time series, electricity consumption data for Spain in 15-minute intervals throughout the year 2023. The data was obtained from the ENTSO-E.

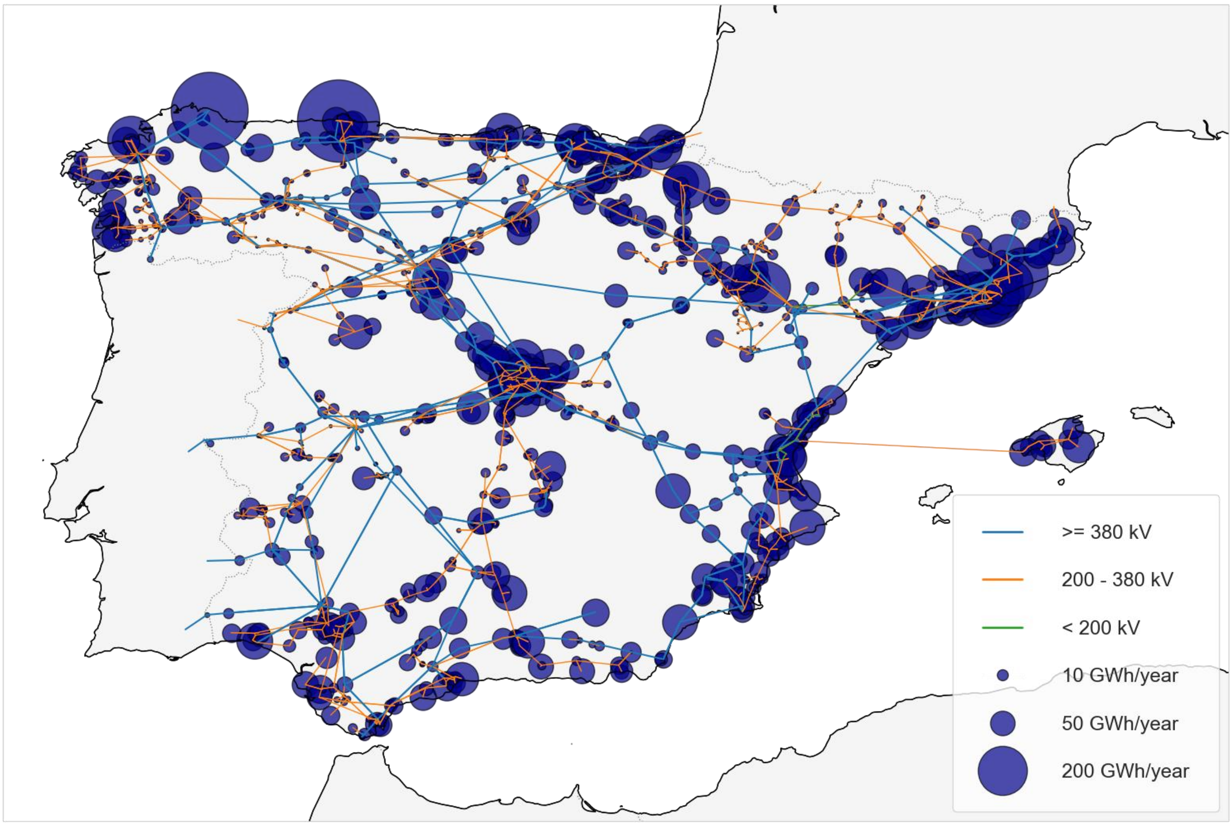


Figure 5: Load distribution through the power transmission network of Spain

4.2 Gas Grid

This dataset provides detailed information on Spain's gas network and has been compiled as part of the iDesignRES project for use within the Multi-Gas Energy Transition (MGET) model. It includes data on natural gas and hydrogen infrastructure, covering Spain, Portugal, and selected neighboring regions to account for cross-border connections. The dataset serves as a basis for modeling energy flows, storage, and supply capacity, particularly in relation to future developments in the gas sector.

4.2.1.1 Geographical Scope and Data Sources

The dataset is structured at the NUTS3 regional level, incorporating information from various sources, including SciGrid Gas, ENTSO-G, the IEA Hydrogen database, GSE Storage data, and Eurostat [5, 6, 7, 8]. The data has been processed and validated to align with the most recent geographical classifications. It includes additional regions outside of Spain to reflect pipeline connections, import routes, and export scenarios, particularly with North Africa and Western Europe.

4.2.1.2 Gas Network Infrastructure

The dataset provides an overview of the existing and planned pipeline network, detailing transport capacities, lengths, and bidirectionality. Additionally, it includes information on gas storage facilities and LNG regasification terminals, highlighting their capacity and operational status. The data also considers efficiency losses in transport and provides estimates for compression losses over long distances.

4.2.1.3 Natural Gas and Hydrogen Production

The dataset includes data on natural gas production, imports, and infrastructure projects, detailing supply sources and regional distribution. For hydrogen, it includes planned production projects and estimated future capacities, based on available industry data. The information is aligned with expected developments leading up to 2030, reflecting projections for increasing hydrogen production and integration into the energy system.

4.2.1.4 Gas and Hydrogen Demand

Gas demand data covers power generation, industrial use, and other consumption sectors, with values provided at the regional level. The dataset includes historical consumption patterns and estimates for future demand, considering the transition from natural gas to hydrogen. Hourly consumption data has been incorporated from multiple sources to provide a detailed view of demand fluctuations over time.

4.2.1.5 Pipeline and Storage Transition for Hydrogen

The dataset includes information on potential repurposing of existing gas infrastructure for hydrogen transport and storage. It considers storage capacities, technical compatibility, and expected future utilization, providing a basis for analyzing hydrogen transport and storage feasibility. The data also reflects estimates for hydrogen injection into the gas grid and the implications for overall network efficiency.

5. Data gathered for Use Cases: The North Sea region

The objective of this chapter is to present the data collected for the North Sea region. This data will support activities in WP2 and WP3, highlighting the importance of mapping, collecting, and validating datasets at the NUTS Level 2 scale. The chapter discusses aggregation and disaggregation challenges, as well as data availability and interoperability, with a focus on converting datasets into the IAMC format.

5.1 *Electricity Grid*

The presented North Sea electric power system dataset covers detailed information on power generation assets, both offshore and onshore, such as wind farms, thermal plants, and other renewable energy sources. Additionally, it contains the high-voltage transmission grid and interconnections between countries. The data on load distribution through the grid, as well as load time series data for 2024 with hourly temporal resolution, has been embedded in the dataset. The dataset spans Norway, Germany, the Netherlands, Denmark, the United Kingdom, and Belgium, with France included for completeness.

5.1.1.1 Transmission Grid

The network data for the North Sea power grid is scraped from PyPSA-Eur [9] , which contains the European High voltage grid (220 kV to 750 kV) and is based on OSM Data. The dataset comprises 3,048 buses, 4,221 AC lines, and 24 DC links, along with data on transformers and converters.

5.1.1.2 Generation

The power-generation dataset contains data on active power plants in the considered North Sea region and the included countries, with France added. It contains information on the type of powerplant and its geographical location.

The generation data is gathered from PyPSA-Europe, specifically the most recent powerplantmatching-publication. Due to varying data across countries for solar and wind installations, their values are proportionally scaled using available, aggregated ENTSO-E data. Additionally, for French nuclear power, the more accurate and detailed Global Energy Monitor data [3] is used instead of PyPSA data. A comparison of the total installed capacity per energy source for each country in our dataset and ENTSO-E is illustrated in Figure 6.

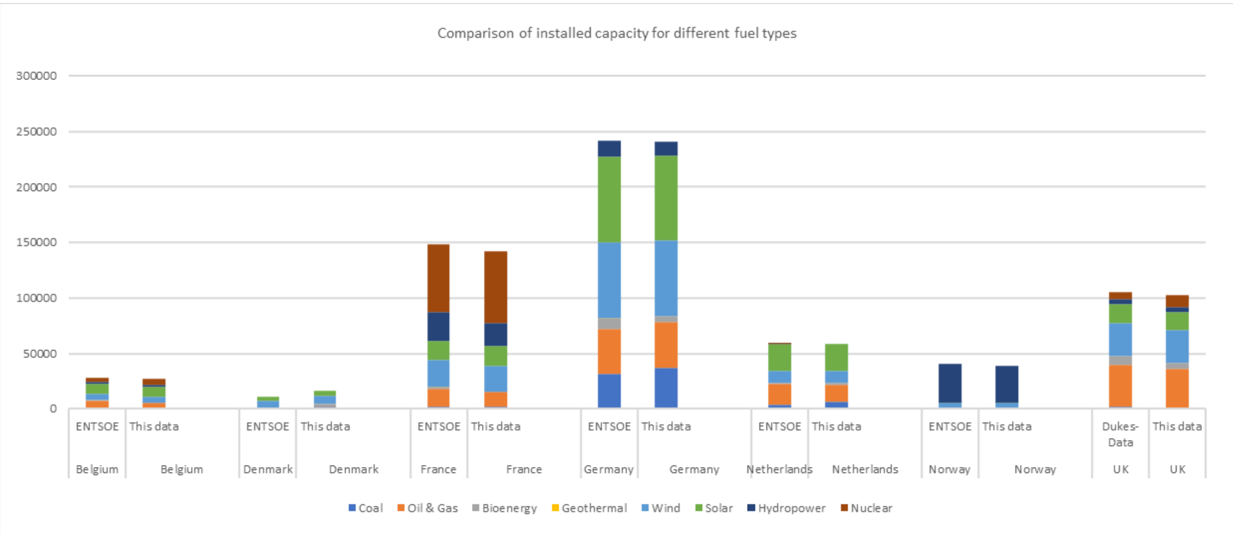


Figure 6: Comparison of installed capacity per fuel type between ENTSO-E and this dataset

5.1.1.3 Demand

The demand data consists of both a static and a time-dependent component. The static component addresses how load is distributed throughout the grid, while the time-dependent component represents the actual hourly load values over a specific period, in our case, the year 2024.

Regarding the load distribution through the grid, the primary data source used is the EU Energy Atlas [10]. The EU Energy Atlas provides an estimate of electricity consumption in EU countries with a spatial resolution of 1 square kilometer. It categorizes consumption by various sectors, such as industry and transportation, while also offering aggregated electricity consumption across all sectors. Although this data is from 2019, it is assumed that while total electricity consumption may have changed between 2019 and 2024, the spatial distribution of electricity consumption across Europe has remained approximately constant.

However, the Atlas data does not cover Norway or the United Kingdom. For the United Kingdom, the current version of our data relies solely on population distribution to estimate the load distribution within the network. For Norway, a combination of geographical locations and annual electricity consumption of major industrial and oil sites has been used to refine the load distribution estimate, improving upon what would be derived solely from population distribution.

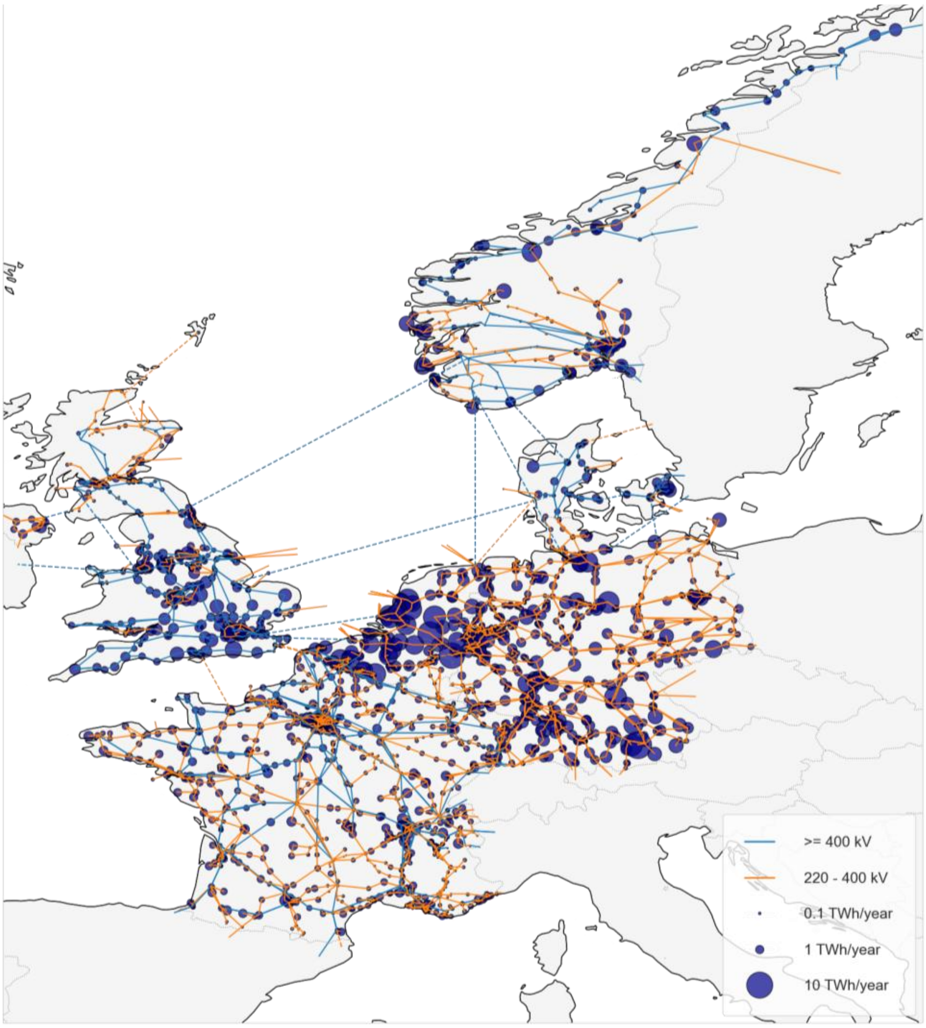


Figure 7: The estimated load distribution through the transmission grid of the North Sea countries. Load values are based on 2024

Finally, Figure 7 illustrates the final estimated load distribution through the transmission grid of the North Sea countries.
Regarding the load time series, electricity consumption data for the North Sea countries and France in hourly intervals throughout the year 2024. The data was obtained from the ENTSO-E Transparency Platform and from the National Energy System Operator (NESO) of the United Kingdom.

5.2 Aggregation for NUTS Level 2

The aggregation was carried out for the electrical power systems of the North Sea countries. After collecting data with high geographical resolution, it was aggregated to the NUTS2 level to ensure compatibility with the models used in the project (WP2 models). The latest version of the NUTS classification, published by Eurostat in 2024, was used for this purpose. However, this update does not include data for the United Kingdom. Therefore, for the UK, the 2021 version of the NUTS classification was used.

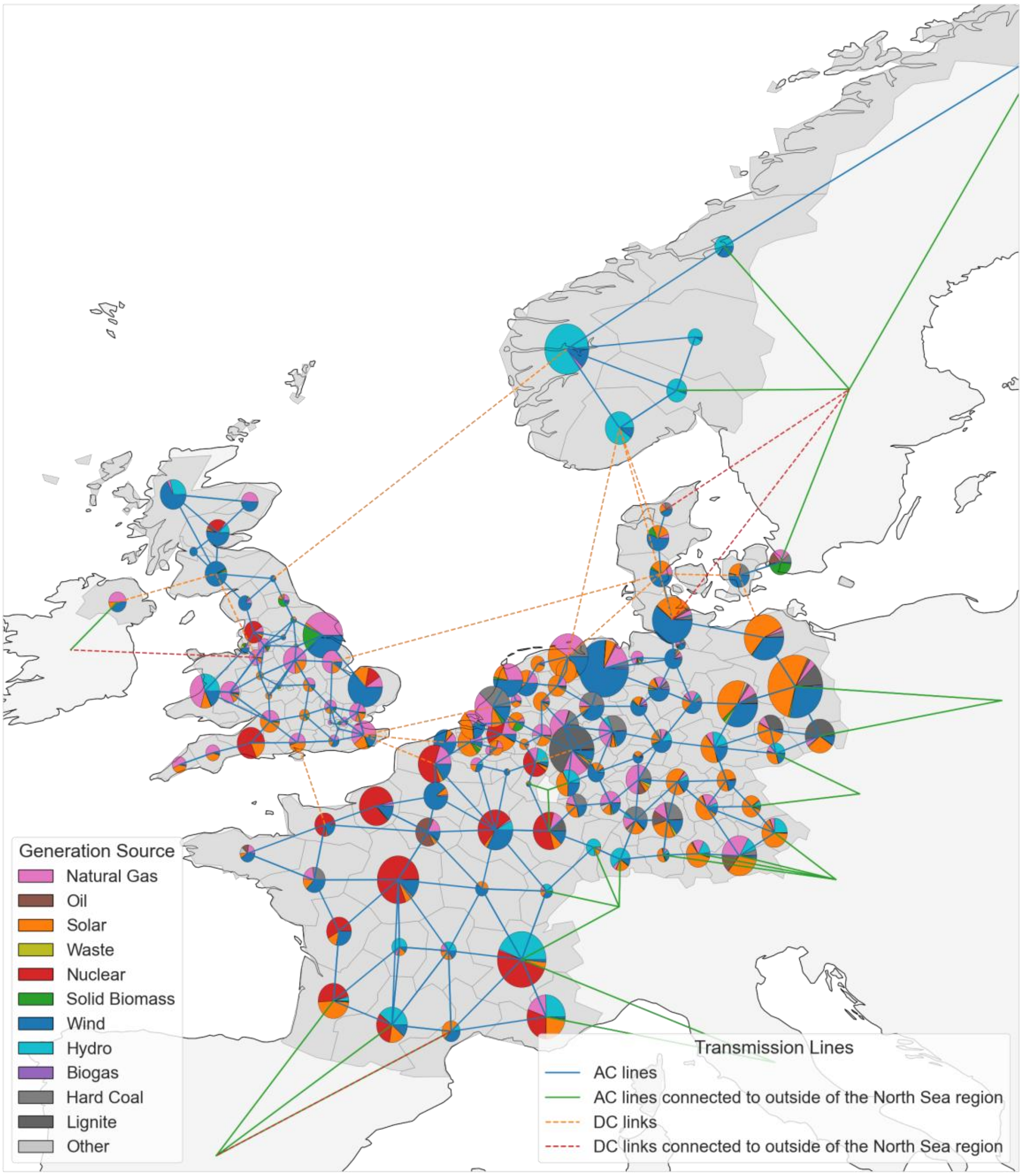


Figure 8: The aggregated installed generation capacity and transmission grid to NUTS2 level.

In the first step of the aggregation process, all buses located within the same NUTS2 region were merged into a single bus, ensuring that each NUTS2 region is represented by only one bus. The total generation capacity and electricity demand for each region were calculated as the sum of the generation capacities and demands of the individual buses located within it. It should be noted that each neighboring country with a directly connected transmission grid has been represented by a single node.

For the aggregation of transmission lines, the process relied solely on the thermal capacities of the lines, rather than a detailed calculation of the equivalent line impedances. The thermal capacity of each equivalent line was calculated by summing the thermal capacities of all transmission lines connecting any buses in one NUTS2 region to another. The aggregated grid and generation installed capacities for different sources has been depicted in Figure 8. **This data has been uploaded to the iDesignRES scenario explorer and converted to the IAMC data format for its wider usage.**

6. End-use sectors: Industry, Transport and Buildings

The objective of this chapter is to detail data that might be used for the energy consumer sectors. For the industry sector detailed data has been collected through interactions with other EU projects. This data has been validated and has a NUTS3 resolution. For the transport and building sectors various data mapping and sources have been identified for collection and validation.

6.1 Industry

The dataset provides data of various industrial sectors at the regional level (NUTS1 and NUTS3). The dataset focuses on technological aspects of industrial processes and contributes to a broader analysis of industrial transformation and decarbonization efforts in Europe. The primary data source is the European AIDRES database (https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/EIGL-Data/AIDRES). However, additional data sources have been incorporated, including those available in the GitHub repository (https://github.com/MOPO-industry/technology-catalogue-and-gap-analysis), to enhance the coverage of industrial processes and regional data. Note that the dataset is subject to changes and might be updated regularly.

- Dataset link: https://github.com/MOPO-industry/technology-catalogue-and-gap-analysis/tree/main/outputs/technology_catalogue

The given relevant dataset (*technology_catalogue_ind_nuts1_nuts3.xlsx)* is structured into several sheets, as detailed below:

The *ind_production_2018_nuts3* and *ind_production_2018_nuts1* sheets contain production volumes for various industrial sectors in 2018. The datasets include key attributes such as country code, NUTS3 or NUTS1 region, sector ID, to_node, industry, unit, and the total production amount in 2018. In these datasets, the column industry maps onto to_node, meaning that the respective industry category corresponds to the sector defined in to_node. Additionally, the ind_production_30_50_nuts3 and ind_production_30_50_nuts1 datasets project production volumes for 2030 and 2050.

The sheet *ind_process_routes_sec* provides insights into energy consumption across different industrial production processes. It details the specific energy inputs required per ton of product, distinguishing between different energy carriers such as electricity, coal, biomass, and other sources. Each row represents an industrial process with a corresponding energy source and its consumption per ton of final product. The dataset presents energy consumption data for multiple years up to 2050. In addition, the *ind_process_route_life* dataset specifies the expected lifespan of various industrial production processes.

Emissions data are provided in the *ind_process_routes_co2_emission* sheet, which quantifies CO_2 emissions per ton of product for different industrial production routes. The dataset includes emission values for future years. Complementing this, the *ind_process_routes_co2_capture* dataset outlines the potential for carbon capture across various industrial processes. It provides estimates of the feasibility and extent of CO_2 capture, indicating the amount of emissions that can be mitigated under different conditions and capture technologies.

The *ind_process_routes_capex* sheet contains information on capital expenditures (CapEx) for different industrial production processes. It specifies the investment required per ton of product and per year for various production technologies and is projected for future years as well.

The *ind_process_routes_fom* sheet contains fixed operating and maintenance costs (FoM) associated with various industrial production processes. It provides recurring cost estimates for maintaining industrial operations per ton of product.

Other important related data for industry has been provided in the appendix.

6.2 Transport

There are various open-source resources for NUTS Level 2 or finer analysis of the transport sector and its links to energy consumption and related dynamics. For the iDesignRES project, combining Eurostat's regional statistics with synthetic freight flow datasets, and real-time mobility data from national sources provides a robust foundation for specifics data for the demonstrators. Each source has unique strengths and limitations; integrating multiple datasets will maximize analytical depth and spatial resolution for the project's objectives. Below Table summarizes some relevant data sources mapped as part of this task.

Source	NUTS Level	Modes Covered	Open Source	Main Advantages	Main Disadvantages	Use potential in iDesignRES
Eurostat Regional Transport Stats	NUTS 2	Road, Rail, Others	Yes	Official, harmonized, updated	Limited NUTS 3, some aggregation	Baseline analysis, cross-country comparison
Synthetic Road Freight Flow Data	NUTS 3	Road Freight	Yes	High spatial detail, projections	Road freight only, modelled	Detailed freight energy modelling
Hotmaps Toolbox	NUTS 2/3	Multiple	Yes	Integrated energy datasets	Some estimated data, focus on heating/cooling	Spatial energy-transport planning
NAPs/Mobility Data Spaces	NUTS 2/3+	Multimodal	Mostly	Real-time, harmonized (improving)	Varying coverage/formats	Urban mobility, real-time applications
EU Open Data Portal	All	All	Yes	Centralized, diverse	Variable quality/granularity	Supplementary, gap-filling

Eurostat Regional Transport Statistics: Eurostat provides the most comprehensive repository of regional transport data for the EU, covering infrastructure, vehicle stocks, energy consumption, and transport flows. Key datasets include:

- **TRAN_R_VEHST**: Stock of vehicles (passenger cars, electric vehicles, road utility vehicles) at NUTS 2 [11] [12].
- **TRAN_R_NET**: Motorway, railway, and inland waterway infrastructure density [13].
- **TRAN_R_ELVEHST**: Electric passenger car registrations by region [14].
- **ROAD_GO_TA_RL/RU**: Road freight performance by loading/unloading region [15].

Advantages:

- Open access: All datasets are freely available through the Eurostat Data Browser.
- Standardization: Harmonized methodologies ensure cross-country comparability [16] [1].
- Temporal coverage: Annual updates (e.g., vehicle stocks from 2013–2022) [7].

Limitations:

- Incomplete regional granularity: Some indicators (e.g., road freight origin-destination flows) lack NUTS 2 resolution [15] .

- Data gaps: Partial derogations for Finland, Greece, and Spain in energy consumption statistics [18] [19].

ODYSSEE-MURE Energy Efficiency Indicator: The ODYSSEE-MURE platform, managed by ADEME and supported by the EU, offers decomposed energy consumption data for transport, including activity, modal shift, and savings effects [11] [20].

Advantages:

- Analytical tools: The decomposition tool isolates drivers of energy use (e.g., traffic growth vs. efficiency gains) [20] [10].
- Integration with Eurostat: Leverages Eurostat data while adding policy context (e.g., renewable energy targets) [11].

Limitations:

- Limited regional resolution: Primarily national-level data, with sparse NUTS 2 breakdowns [20].
- Complex access: Requires navigation through multiple interfaces (e.g., policy radar tools) [20].

European Environment Agency (EEA) Renewable Energy Statistics: The EEA tracks renewable energy use in transport, including biofuels, renewable electricity, and hydrogen, with preliminary EU-wide estimates [22] [23].

Advantages:

- Policy alignment: Directly links to Renewable Energy Directive (RED II) targets [22] [24].
- Fuel-specific insights: Breaks down renewables by type (e.g., biodiesel vs. biogasoline) [25].

Limitations:

- Regional gaps: No NUTS 2 data for RES shares in transport; national-level only [22] [23].

UNECE E-Road and E-Rail Censuses: UNECE's origin-destination datasets for rail (TRAN_R_RAPA, TRAN_R_RAGO) and inland water freight (IWW_GO_ATYGOFL) provide flow data between NUTS 2 regions [15].

Advantages:

- Flow analysis: Captures cross-border freight movements (e.g., Paris-Vienna rail traffic).
- Multimodal integration: Includes NST2007 goods classification for inland waterways.

Limitations:

- Sparse updates: Rail data collected quinquennially; road freight flows lack comparable detail.

National Statistical Offices and Collaborative Platforms, examples include:

- **Germany (DESTATIS)**: Regional energy consumption in road transport, including heavy-duty vehicles [26].
- **Poland (GUS)**: Local RES adoption rates in transport, aligned with EU cohesion policies [27].
- **OpEnergy Platform**: Aggregates Eurostat data for time-series forecasting of transport energy use [21].

Academic and EU Project Databases, examples:

- **Academic publications**: Analyses of RES adoption in road transport using Eurostat data [7] [15] [13].
- **EU Regional Yearbooks**: Contextualize transport energy trends within economic profiles.

In short, Eurostat seems the main source for NUTS 2 transport energy data, offering regional granularity and open access. However, its limitations in flow analysis and RES specificity necessitate supplementation from UNECE, national databases, and academic studies. For the iDesignRES project, a hybrid approach—combining Eurostat's infrastructure/vehicle data, UNECE's flow statistics, and ODYSSEE-MURE's policy tools—will provide the most robust foundation for modeling sustainable energy transitions in European transport.

6.3 Building stock

For the building stock, key datasets include Eurostat's harmonized statistical repositories, EU-funded initiatives like Hotmaps and TABULA. While these sources provide granular insights into building characteristics, energy performance, and regional disparities, challenges persist in data standardization, temporal coverage, and integration of heterogeneous datasets. The following details the advantages and disadvantages of the datasets usefulness for NUTS Level2 and the iDesignRES project.

Eurostat Energy Consumption Surveys: Eurostat's *Energy Consumption in Households* dataset provides NUTS 2-level data on energy use patterns, including breakdowns by fuel type (natural gas, electricity, renewables) and end-use categories (space heating, water heating, cooling) [29]. The 2022 update reveals that 63.5% of EU household energy is allocated to space heating, with significant regional variations in renewable adoption (e.g., 88.56% in Portugal vs. 28.46% fossil dependence in Poland) [29].

- Advantages: Standardized methodology across EU member states; annual updates; integration with other socioeconomic indicators (e.g., income, urbanization).
- Disadvantages: Limited building-specific metadata (e.g., construction materials, insulation quality); aggregation biases at NUTS 2 level obscure municipal-level variations.

Census Hub and Geostat Gridded Data: Eurostat's 2021 Population and Housing Census offers LAU 2 (municipal-level) data on dwelling characteristics, including construction period, occupancy rates, and heating systems [30]. The 1 km² gridded dataset enables spatial analysis of population density and building distribution, critical for estimating energy demand hotspots [30].

- Advantages: Unprecedented spatial resolution; compatibility with GIS tools for cross-domain analysis.
- Disadvantages: Restricted access to microdata; delayed publication cycles (e.g., 2021 census data fully available by 2024).

Building Permits and Floor Area Statistics

Eurostat's *Building Permits Index* tracks construction activity at NUTS 2, with 2023 data indicating a 14.6% decline in floor area approvals EU-wide, reflecting post-pandemic market adjustments [31]. The dataset differentiates residential and non-residential sectors, enabling demand forecasting for energy retrofits[4].

- Advantages: High-frequency updates (monthly); granular sectoral breakdowns.
- Disadvantages: No linkage to energy performance metrics; incomplete coverage in Eastern European countries.

Hotmaps Toolbox

The Hotmaps project (Horizon 2020) generates 100x100 m raster maps of building stock indicators, including useful energy demand (UED), gross floor area, and heating/cooling requirements [33] [34].

Leveraging Eurostat's CENSUS 2011 and CORINE land cover data, it supports scenario analysis for district heating and renewable integration [35].

- Advantages: Open-source GitLab repository; compatibility with urban energy planning tools.
- Disadvantages: Requires advanced GIS expertise; outdated building age profiles (pre-2011).

TABULA WebTool and Building Typologies

TABULA provides standardized building typologies for 30 European countries, classifying residential structures by construction period, size, and energy performance [36]. For instance, post-2001 buildings in Germany show 40% lower space heating demand compared to pre-1945 stock [36].

- Advantages: User-friendly interface for benchmarking; integration with national energy performance certificate (EPC) databases.
- Disadvantages: Overaggregation at NUTS 1 level; limited coverage of non-residential buildings[7].

In short, Eurostat's focus on household data overlooks a significant portion—about 45%—of commercial building energy use, creating non-residential blind spots. Looking at the data there will be need to preprocess and harmonize heterogeneous data formats, such as merging raster datasets from Hotmaps with vector datasets from TABULA. There is also a misalignment between policy mechanisms and available data: for example, the ETS2 carbon pricing scheme lacks sufficiently granular, building-level emission factors, complicating its application to real-world scenarios. To address these challenges, iDesignRES for some demos could create or adapt a hybrid data sourcing strategy by combining Eurostat's structural data with real-time IoT readings from pilot cities in the demos. The development of open-source Python toolkits to automate aggregation from LAU 2 to NUTS 2 levels would streamline data integration.

7. Outreach and dissemination activities

The objective of this chapter is to provide a concise overview of the main datasets collected in Task 1.1 and to highlight the coordination efforts within WP1 to achieve collection and mapping of high spatial resolution data. It also details the project's commitment to open data availability and summarizes participation in external dissemination events

7.1 Open data sets for Spain and North Sea

So far, several datasets have been successfully compiled and published on the Zenodo platform as part of the project (see https://zenodo.org/communities/idesignres/). The following datasets are currently available:

- **Spanish Power System Data**
  https://zenodo.org/records/14748878
- **Spanish Gas System Data**
  https://zenodo.org/records/14859357
- **North Sea Power System Data**
  https://zenodo.org/records/14918402
- **European Gas Networks and infrastructure data**

  https://zenodo.org/records/15553011
- **Heat demands: GIS-based modelling to generate high-resolution spatial data on heat demand, district heating potentials across Europe, and waste heat sources with their corresponding potential.**

  https://zenodo.org/records/15363552

These datasets were gathered using a wide range of sources, which were carefully selected, harmonized, and processed. Detailed descriptions of the data collection procedures, original data sources, and processing methodologies can be found in the accompanying documentation available with each dataset on Zenodo. The basic description of these data records has been provided in Chapter 3, 4 and 5.

A deeper analysis and aggregation of the data to the NUTS-2 regional level is currently underway and will be included in an upcoming update. This regionalization will enable more targeted analyses and enhance the usability of the data for energy system modelling, planning, and policy support at the sub-national level. This is particularly relevant for WP2 in the iDesignRES project.

7.2 Workshop organization at the 13th European Open Energy Modelling

Location: KTH Royal Institute of Technology, Stockholm

Date: March 25–26, 2025

Introduction and purpose of participation

Our participation in the 13th European Open Energy Modelling Workshop (Openmod) at KTH in Stockholm was part of our ongoing efforts to improved access to high-quality, open-source data for energy system modelling, possibly on the NUTS 2-level. The goal was to identify new data sources, specifically in the context of NUTS2-level energy system data. Further, the goal was to share experiences with the broader modelling community, and exchange best practices related to data handling.

OpenMod 2025 Workshop

OpenMod is one of the largest grassroots initiatives in the field of energy modelling. It serves as a platform for researchers, practitioners, and developers committed to transparency, reproducibility, and open-source tools and data in energy system analysis. The community is known for its collaborative spirit, policy engagement, and its proactive role in addressing the broader challenges of climate change, energy justice, and inclusive capacity building. The 2025 edition of OpenMod[5], hosted by the KTH Division of Energy Systems, placed particular emphasis on inclusiveness in modelling and representation of under-represented. This was especially relevant given our interest in harmonizing open data practices across diverse European sub-national regions, many of which face similar data accessibility challenges.

The event format included lightning talks, poster sessions, and breakout discussions, providing both structured and informal opportunities for engagement. Presentations spanned a wide array of topics: stochastic load profiles, coupling ML models with market simulations, geospatial modelling, metadata annotation, agent-based and dynamic models, model integration, and sector-coupled systems—underscoring the technical depth and diversity of the OpenMod community.

Contributions and Data-Focused Activities

We engaged in three main activities during the workshop, all aligned with our data collection objectives:

- Lightning Talk

As part of the workshop's lightning talk sessions, we presented an introduction to the iDesignRES project, aimed at sparking conversations around collaborative, open-data-driven energy system modelling. Our talk focused on the project's ambition to bridge the Pan-European energy system perspective with a highly resolved, regionalized modelling layer—spanning different NUTS levels and linking national planning with sub-national capacity expansion and operational validation.

A key challenge we highlighted during the talk is the need for robust, interoperable data structures. The iDesignRES framework depends on the seamless flow of various data types—techno-economic, spatial, temporal—across modelling layers. This requires harmonized and standardized datasets that can interface with diverse modelling environments without breaking interoperability or transparency, particularly in an open-source context.

To encourage further exchange and collaboration, we concluded by inviting participants to join our breakout session on the following day, focused on strengthening cooperation and sharing best practices around open energy data. We aimed to engage participants with hands-on experience and fresh perspectives, hoping to co-develop strategies for building a shared data backbone that supports modular, scalable, and resilient modelling approaches across Europe.

[5] Openmod 2025: announcement, concept and agenda 25-26 March 2025 - Workshop - Open Energy Modelling Initiative

Figure 9: iDesignRES project lightning talk at the openmod

Poster Presentation

During the workshop, there was a poster session where we had the opportunity to present two different posters. One of them focused on the recent work done in the frame of the iDesignRES project: “European Energy Vision 2060: Charting Diverse Pathways for Europe’s Energy Transition”. The second one provided an overview of iDesignRES, introducing its objectives, methodologies, and expected outcomes.

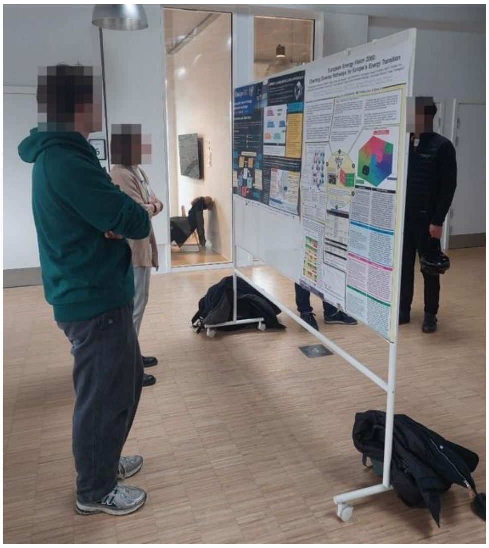


Figure 10: Poster session

Energy Vision Scenarios poster

The poster addresses Europe's urgent need to explore energy transition pathways, given that it is warming at a faster rate than any other continent. In alignment with the European Green Deal’s goal of achieving climate neutrality by 2050, the research develops four long-term scenarios that consider key uncertainties across social, technological, economic, political, and geopolitical dimensions. These scenarios—EU Trinity, REPowerEU++, Go RES, and NECP Essentials—are structured within a three-dimensional space defined by Social Dynamics, Innovation, and Geopolitical Instabilities.

A major contribution of the research is the Qualitative to Quantitative (Q2Q) matrix, which links scenario narratives with quantifiable energy system modeling parameters. The poster facilitated discussions and findings, engaging workshop participants in dialogues about open and reproducible energy system modeling.

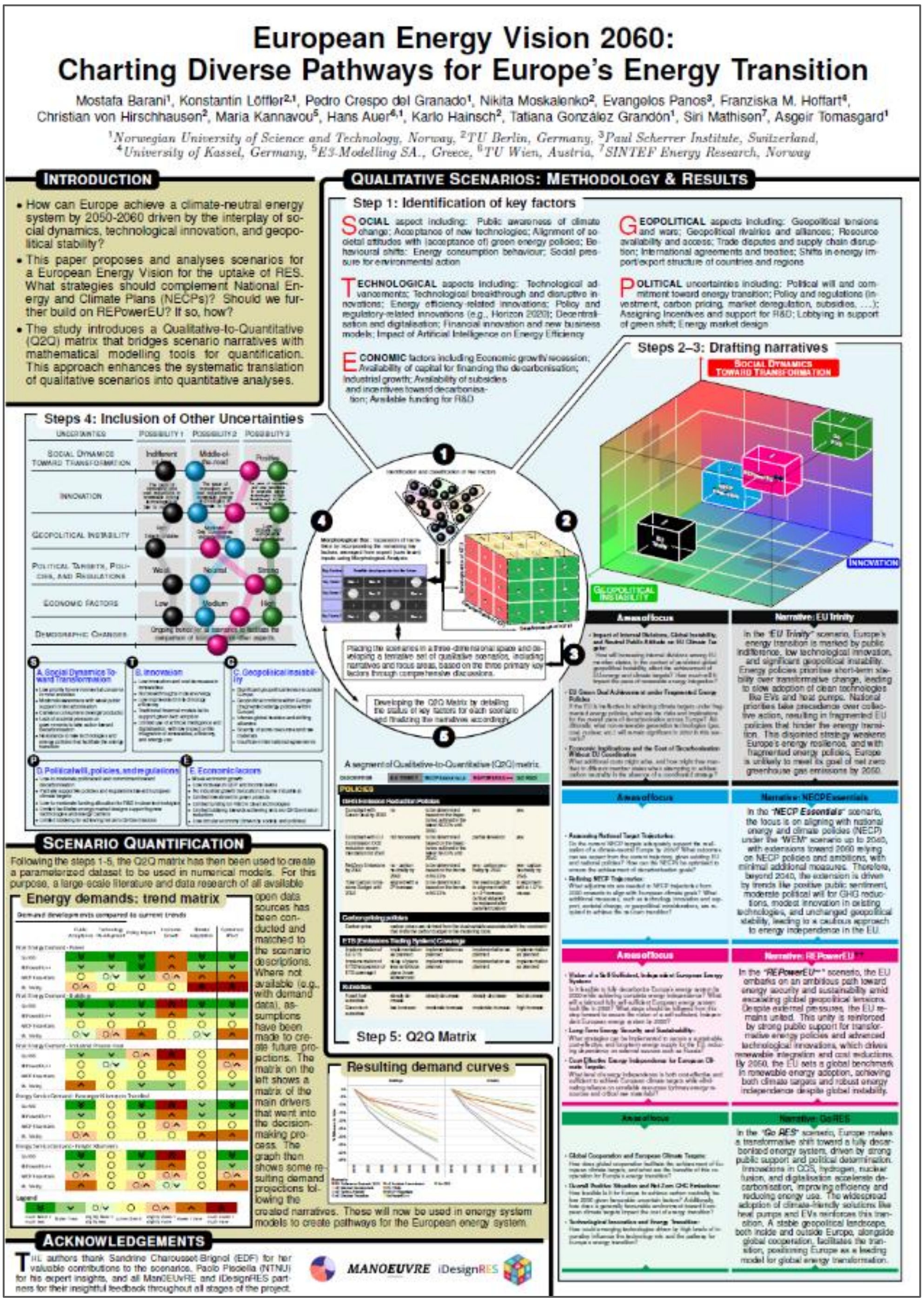


Figure 11: Energy vision scenarios poster

iDesignRES project poster

The poster provided participants with a clear understanding of the project's objectives, methodologies, and key research areas. It served as a starting point for discussions on the development of open-source tools for energy system modeling and decision-making.

In particular, the poster highlights the core components of iDesignRES, showcasing the development of multi-physics energy system models, the integration of open-access data sources, and the use of advanced visualization tools to support scenario-based analysis. Attendees engaged in discussions about how these elements contribute to improving system flexibility, integrating renewable energy sources, and ensuring resilience in the face of uncertainties. A key aspect emphasized was the open-source and collaborative nature of iDesignRES, ensuring transparency and accessibility for the broader research and policymaking community.

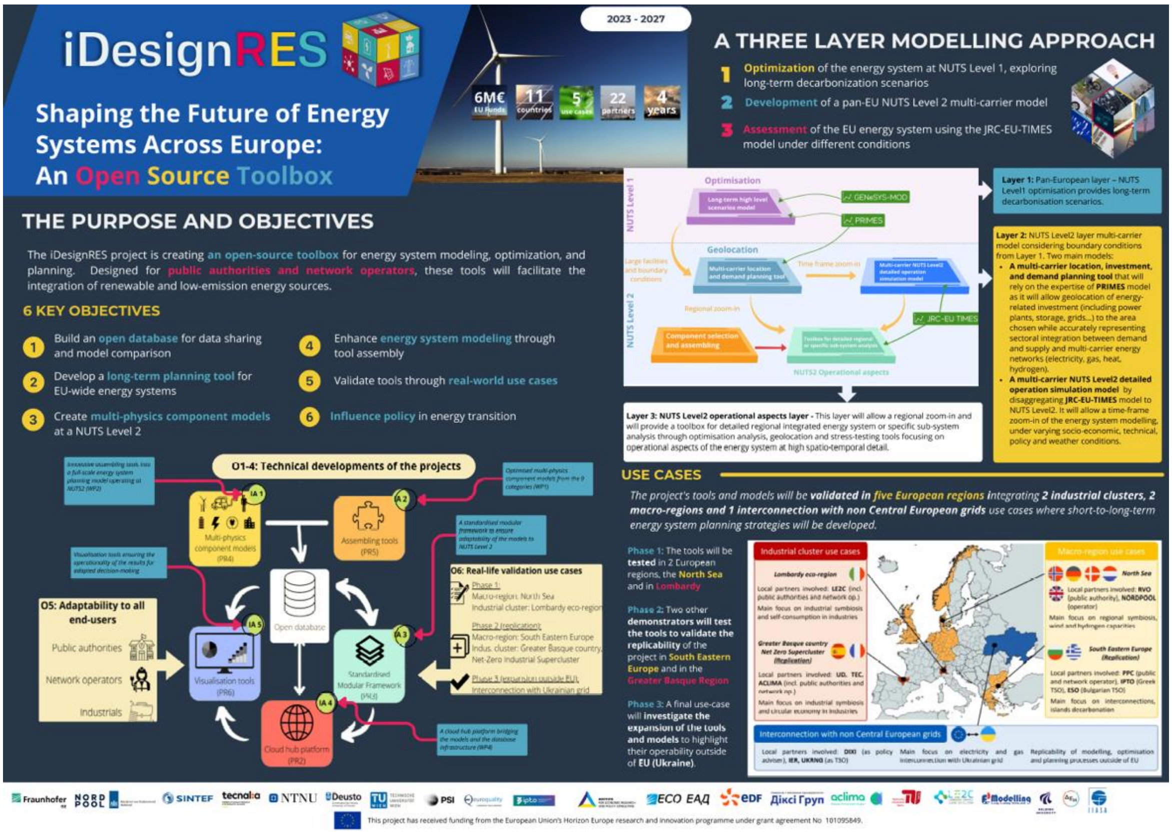


Figure 12: iDesignRES poster and project overview

Breakout Session: High-Resolution Modelling and data at the NUTS2 Level

The goal of our breakout session was to create a participatory space to **share experiences, build connections, and collect insights on useful data sources and best practices** for high-resolution energy system modelling. We focused on the unique challenges and opportunities of working at the regional (NUTS2) level and aimed to foster practical exchange among researchers and practitioners engaged in similar efforts.

To structure the session effectively, we adopted an **interactive, step-by-step approach**. We opened with a short introduction to the **iDesignRES project**, outlining our objectives, current progress, and some of the hurdles we're facing—especially regarding data harmonization across models and regions. This helped set the stage for a deeper, experience-based discussion.

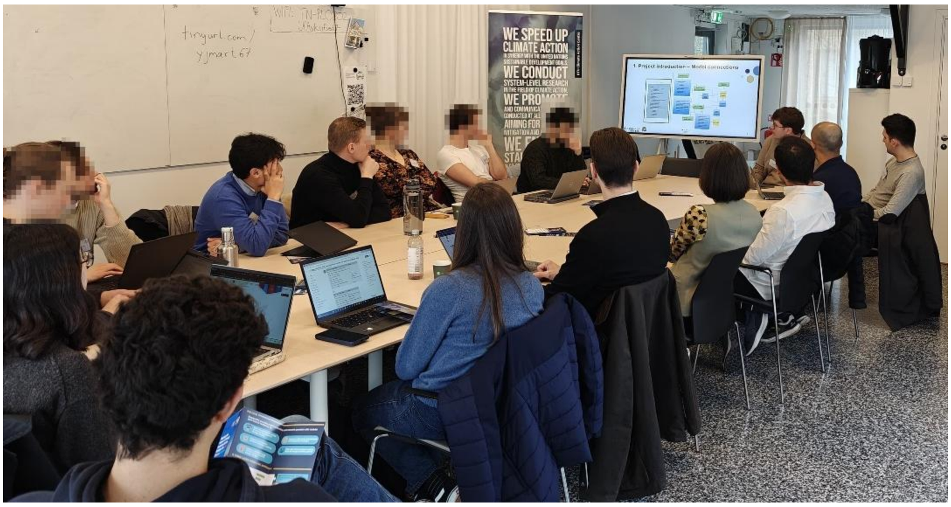
We then turned to **Mentimeter** to actively engage participants. The session was divided into three phases:

1. **Personal Experiences**: We began by asking participants to share their own encounters with high-resolution modelling—both challenges and insights—through open-ended questions and polls. This helped surface the diversity of use cases, sectors, and modelling tools represented in the room.
2. **Data Challenges**: In the second phase, the discussion shifted toward **specific barriers** faced when working with regional-level data. These included data availability, inconsistency across regions or formats, issues with granularity, and difficulties in integrating spatial data into national or continental-scale models.
3. **Solutions and Best Practices**: Finally, we invited participants to contribute **practical strategies**, including successful **aggregation/disaggregation methods**, key **data sources**, and **open-source tools** they've found useful. We also gathered inputs on ways to promote more standardization and collaboration in this space.

Outcomes of the break out room:

The break out room was successful, with a large group of participants and valuable exchange of ideas, challenges and data sources.

Several new datasets and tools were identified during the discussion, including sources for grid infrastructure, generation/demand profiles, and renewable energy potentials. This session proved especially useful for expanding our dataset base and establishing contacts with others working on high-resolution modelling across Europe. The discussions also clarified common challenges, such as inconsistent spatial coverage and limitations in regional statistics. It was a very useful exchange with many other energy system modelers.



In the following section, some highlights of the results of the breakout session are shown:

Challenges

Data sources:

Several new data sources were discovered through the exchange. These include datasets on power generation capacities, weather data and infrastructure data.

Key Outcomes openmod

As a result of the participation in the initiative, several important outcomes were achieved. New open datasets relevant for regional-scale modelling were successfully identified, enhancing the availability of high-quality data for future research and applications. Participants exchanged practical techniques for handling spatial data, contributing to improved workflows and shared best practices across teams. Valuable insights were gained through direct input from other modelling groups, particularly concerning data quality issues, which will support more robust and accurate modelling efforts. Furthermore, the initiative led to the establishment of new professional contacts, laying the groundwork for potential future collaborations and strengthening the network within the modelling community.

8. Summary and conclusion

This report presents the results of the task 1.1 of iDesignRES, focused on data collection and validation. The aim has been to build a consistent, high-quality data foundation to support the development of open-source tools for energy system planning at regional and European levels.

A central component of this work was the preparation of a common data infrastructure to feed the involved model in the project, based on the existing IIASA database ecosystem. This infrastructure supports transparent data exchange, model comparison, and collaborative scenario analysis through the Scenario Explorer platform. The adoption of standardized formats, common nomenclature, and open tools ensures that all project partners can contribute and access data in a harmonized way.

The report documents in detail the data gathered across multiple domains, including the electricity and gas sectors for selected use cases such as Spain and the North Sea region. For these cases, high-resolution spatial and temporal data was collected on generation assets, grid infrastructure, and electricity demand, ensuring consistency with existing European data sources and public datasets. Special care was taken to clean and validate this data for integration into the central database.

In addition to supply-side data, the report present data collected for major end-use sectors, such as industry, transport, and buildings, covering energy use, emissions, technology characteristics, and cost data. These datasets are essential for capturing demand-side dynamics in the modeling framework.

Finally, the collected data was aggregated to the NUTS Level 2 scale to ensure compatibility with the modeling resolution used in the project. Together, these efforts establish a robust data infrastructure that will support the project's modeling and tool development activities in the coming phases.

References

[1] "OpenStreetMap: Interactive Map," [Online]. Available: https://www.openstreetmap.org/. [Accessed 30 09 2024].

[2] R. Gaugl, S. Wogrin, U. Bachhiesl and L. Frauenlob, "GridTool: An open-source tool to convert electricity grid data," *SoftwareX,* vol. 21, p. 101314, 2023.

[3] "Global Energy Monitor," [Online]. Available: https://globalenergymonitor.org/. [Accessed 21 11 2024].

[4] ENTSO-E, "ENTSO-E Transparency Platform," [Online]. Available: https://transparency.entsoe.eu/. [Accessed 10 12 2024].

[5] "ENTSOG: European Network of Transmission System Operators for Gas," [Online]. Available: https://www.entsog.eu/.

[6] "(IEA), International Energy Agency, Hydrogen Production and Infrastructure Projects Database," [Online]. Available: https://www.iea.org/data-and-statistics/data-product/hydrogen-production-and-infrastructure-projects-database#hydrogen-production-projects.

[7] "Storage Database - Gas Infrastructure Europe (GIE)," [Online]. Available: https://www.gie.eu/transparency/databases/storage-database/.

[8] Eurostat, "NUTS Maps - Eurostat," [Online]. Available: https://ec.europa.eu/eurostat/web/nuts/maps.

[9] "PyPSA-Eur: A free, open model of the European power system," 2024. [Online]. Available: https://github.com/PyPSA/pypsa-eur.

[10] I. H. Gonzales and A. Uihlein, "High-resolution energy atlas," *Publications Office of the European Union,* 2023.

[11] Eurostat. (2024, April). **Stock of vehicles at regional level**. https://ec.europa.eu/eurostat/statistics-explained/index.php/Stock_of_vehicles_at_regional_level

[12] ETF Europe. (2025, April). **Road freight transport in the EU: TransFair KU Leuven analysis**. https://www.etf-europe.org/wp-content/uploads/2025/04/Road-Freight-Transport-in-the-EU_TransFair-KU-Leuven-analysis.pdf

[13] Eurostat. (n.d.). **Inland transport infrastructure at regional level**. https://ec.europa.eu/eurostat/statistics-explained/index.php/Inland_transport_infrastructure_at_regional_level

[14] Eurostat. (n.d.). **Transport statistics at regional level**. https://ec.europa.eu/eurostat/statistics-explained/index.php/Transport_statistics_at_regional_level

[15] United Nations Economic Commission for Europe. (n.d.). **UNECE Wiki: Transport statistics**. https://wiki.unece.org/plugins/viewsource/viewpagesrc.action?pageId=166724439

[16] Eurostat. (n.d.). **Regional transport statistics**. https://ec.europa.eu/eurostat/web/transport/information-data/regional-transport

[17] Gnap, J., Konecny, V., & Mrazik, J. (2022). **Energy consumption and emission production of electric vehicles in European cities**. *Energies, 15*(16), 6030. https://www.mdpi.com/1996-1073/15/16/6030

[18] Eurostat. (n.d.). **Regional transport statistics: SEPDF**. https://ec.europa.eu/eurostat/statistics-explained/SEPDF/cache/131944.pdf

[19] Eurostat. (2024). **Final energy consumption in transport – detailed statistics**. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Final_energy_consumption_in_transport_-_detailed_statistics&oldid=652642

[20] Didier, B. (2022, December). **ODYSEE-MURE DRTD2022**. https://transport.danube-region.eu/wp-content/uploads/sites/2/sites/2/2022/12/S2-0-ODYSEE-MURE_DRTD2022_Didier-B.pdf

[21] Zsigraiova, Z., Semiao, V., & Silva, C. M. (2022). **Assessment of the environmental performance of electric buses**. *Energies, 15*(21), 8294. https://www.mdpi.com/1996-1073/15/21/8294

[22] European Environment Agency. (n.d.). **Use of renewable energy for transport**. https://www.eea.europa.eu/en/analysis/indicators/use-of-renewable-energy-for

[23] Sroka, W. (2022). **Sustainable development in transport sector**. *Ekonomia i Środowisko, 82*(3), Article 805. https://www.ekonomiaisrodowisko.pl/journal/article/view/805

[24] Bălăcescu, A., & Bălăcescu, S. (2022). **The impact of electric vehicles on energy consumption**. *Energies, 15*(15), 5576. https://www.mdpi.com/1996-1073/15/15/5576

[25] Dinu, M., & Ilie, A. M. (2021). **Electric mobility in the European Union**. *Energies, 14*(22), 7457. https://www.mdpi.com/1996-1073/14/22/7457

[26] Faria, R., Moura, P., Delgado, J., & de Almeida, A. T. (2023). **Integration of electric vehicles in smart grids**. *Energies, 16*(12), 4569. https://www.mdpi.com/1996-1073/16/12/4569

[27] Papadopoulos, A. (2023). **The impact of green transport policies in the EU**. *European Research Studies Journal, 26*(1), 3225. https://ersj.eu/journal/3225

[28] Sadeghianpourhamami, N., Refa, N., Strobbe, M., & Develder, C. (2023). **Electric vehicle charging: A data-driven review**. *Energies, 16*(18), 6431. https://www.mdpi.com/1996-1073/16/18/6431

[29] Eurostat. (2024). *Energy consumption in households*. https://ec.europa.eu/eurostat/statistics-explained/index.php/Energy_consumption_in_households

[30] Eurostat. (2024). *Population and housing censuses: Information and data*. https://ec.europa.eu/eurostat/web/population-demography/population-housing-censuses/information-data

[31] Eurostat. (2022). *Energy consumption in households*. https://ec.europa.eu/eurostat/statistics-explained/SEPDF/cache/10062.pdf

[32] Statista. (2024). *Building permits for floor area in the European Union (EU) from 2012 to 2022*. https://www.statista.com/statistics/1551590/building-permits-floor-area-eu/

[33] Hotmaps Project. (2018). *D2.3: Hotmaps for upload revised final*. https://www.hotmaps-project.eu/wp-content/uploads/2018/03/D2.3-Hotmaps_for-upload_revised-final_.pdf

[34] Hotmaps. (n.d.). *Building stock*. https://gitlab.com/hotmaps/building-stock

[35] EPISCOPE. (n.d.). *Building typology by country*. https://episcope.eu/building-typology/country/

[36] Concerted Action for the Energy Efficiency Directive (CA EED). (n.d.). *Main features useful for the monitoring of the EED and the building observatory Odyssee-Mure*. https://www.ca-eed.eu/ia-document/main-features-useful-for-the-monitoring-of-the-eed-and-the-building-observatory-odyssee-mure/

Appendix

Energy demand for industrial processes. Some examples

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
AIDRES/MOPO	https://github.com/MOPO-industry/technology-catalogue-and-gap-analysis/tree/main/outputs	NUTS3-level	Loss of granularity through aggregation

Steel

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
AIDRES/MOPO	https://github.com/MOPO-industry/technology-catalogue-and-gap-analysis/tree/main/outputs	NUTS3-level	Loss of granularity through aggregation

Cement

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
AIDRES/MOPO	https://github.com/MOPO-industry/technology-catalogue-and-gap-analysis/tree/main/outputs	NUTS3-level	Loss of granularity through aggregation

Metalworking industry

Data name	Source	Comments, Explanation or specifics	pros and cons for Nuts Level 2
AIDRES/MOPO	https://github.com/MOPO-industry/technology-catalogue-and-gap-analysis/tree/main/outputs	NUTS3-level	Loss of granularity through aggregation

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open-Source Toolbox

More information on iDesignRES project:

https://www.linkedin.com/company/idesignres

https://idesignres.eu/

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open Source Toolbox

Spanish Gas Network data of the MGET-model

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Document information

Deliverable Number and Name	MGET data for Spain
Work Package	WP.1, T1.1 & T1.3
Dissemination Level	Public
Author(s)	Ruud Egging-Bratseth, Jannis Kerl, Ehsan Nokandi
Primary Contact and Email	Ehsan Nokandi, Ehsan.Nokandi@ntnu.no
Date Submitted	tba
File Name	Gas_data_Spain_MGET.xlsx
Status	V1.0.0

This document has been prepared in the framework of the European project iDesignRES. This project has received funding from the European Union's Horizon Europe programme under grant agreement no. 101095849.

TABLE OF CONTENTS

1. Exclusive Summary

The data presented is published as part of the iDesignRES project, specifically under WP 1, Task 1.1 and Task 1.3. These tasks involved gathering detailed energy system data for Spain, covering electricity and gas, to be used as test data for the models employed within the iDesignRES project.

The data which is presented here focuses on the Spain's gas system and is used in the multi-gas model for the Energy Transition (MGET). It comprises all of Spain and Portugal, and several regions in France bordering to Spain with a natural gas pipeline connection. An Algerian and a Moroccan region have been added to allow pipeline imports into Spain. The regions Rotterdam in NLD, and around Marseille, in France, have been added to allow modeling export scenarios.

2. File- and content description

The dataset is published in an Excel file titled *Gas_data_Spain_MGET.xlsx*, which contains several sheets, organized as follows:

2.1 Sheet "N" – Model Nodes and Consumption Shares

Sheet "N" contains various types of information regarding the Spanish gas-nodes:

Model Nodes and Regional Information:

- Contains the set of model nodes and their types, countries, and regions.
- **Column 'N'**: NUTS3 region.
- **Column 'Cn'**: Country information.
- **Column 'NUTS2'**: Region based on NUTS2-defined areas.
- **Column 'Rgn'**: Continent information.

Consumption Shares:

- Contains data on the consumption shares of the NUTS3 region within its NUTS2 region.
- **Columns 'G' and 'H'**: Detail the shared consumption values for each NUTS3 region as a proportion of the total consumption within their NUTS2 region for:
    - Natural gas.
    - Hydrogen.

2.2 Sheet "A" – Arcs and Pipelines

Sheet "A" contains information regarding the arcs between different NUTS3 regions and includes data on the length of pipelines and the energy carrier (Natural Gas-G or Hydrogen-H).

- **Column 'A'**: Arc identifier defined by starting and end node with an underscore in between (also found in 'Start' and 'End' columns, both at NUTS3 resolution).
- **Column 'F'**: Fuel type (Gas-G or Hydrogen-H).
- **Column 'cap'**: Capacity of the arcs. Unit: GWh/h (=GW)
- **Columns 'total length' and 'offshore length'**: Respective length of the pipelines, in total and the offshore part. Unit: km.
- **Column 'Is bidir'**: Binary value indicating if the pipeline is bidirectional.

- **Column 'Reversable?'**: Indicates whether the arc is reversible, i.e., can be made bidirectional.

2.3 Sheet "C" – Consumption Data

Sheet "C" contains the consumption data for NUTS2 regions:

- **First column**: NUTS2-area information.
- **Consumption Data Source -->**The Nodes-sheet draws its share from the values here.
- **Column 'Fuel'**: Fuel type (Gas-G or Hydrogen-H).
- **Column 'POW'**: Consumption value of the NUTS2-region (From SciGrid data).
- **Max/Mean/Med/Min**: hourly consumption statistics (From SciGrid data).

2.4 Sheet "P" – Gas/Hydrogen Production and Projects

Sheet "P" contains data on gas and hydrogen production in each NUTS3 area:

2.4.1 Gas Production and Hydrogen Projects

- **Left table**: Gas production for NUTS3 regions.
- **Right table**: Hydrogen projects.

2.4.2 Gas production Columns:

- **'R'**: Continent/Region of the production site.
- **'C'**: Country.
- **'N'**: NUTS3-node.
- **'F'**: Fuel type (Gas-G or Hydrogen-H).
- **'Y'**: Year.
- **'1'**: Production value.
- **'MC'**: Marginal cost. Assumed unit production cost in €/MWh

2.4.3 Hydrogen Project Data

- **Left column**: NUTS3-region identifier.
- **Columns '2025', '2030', and 'Grand total'**: Respective yearly added and total production values for hydrogen.

2.5 Sheet "W" – Storage Capacities

Sheet "W" contains storage capacities in NUTS3 regions:

- **'NUTS3'**: NUTS3-region.
- **'F'**: Fuel type (Gas-G or Hydrogen-H).
- **'H2-ready'**: Indicates if storage can be used for hydrogen.
- **'W'**: Work capacity, Unit: GW.
- **'I'**: Injection rate, Unit: GW Unit: GW.
- **'X'**: Extraction rate, Unit: GW.
- **'Year'**: First Model year that capacity is available.

2.6 Sheet "R" – Regasifiers

Sheet "R" contains information on regasifiers:

- **'N'**: NUTS3-node information.
- **'F'**: Fuel type (Gas-G or Hydrogen-H).
- **'Y'**: First Model Year that capacity is available.
- **'ub'**: Capacity, Upper bound on regasifier output.

2.7 Sheet "O" – Other Data

Sheet "O" contains definitions and parameter values needed for the model.

3. Data sources:

To compile the dataset the following sources were utilized.

- SciGrid Gas [1]

  Data from SciGrid Gas were used to provide information on production, power plant capacities, and demand from other consumers, including minimum, maximum, average, and median hourly values. It also included data on pipeline segments, storage facilities, and LNG terminals. The dataset was compiled using AI tools and several heuristics to estimate missing values. While this provides a solid starting point for many data categories, there are some unresolved issues that require further data checking and revision.

- ENTSO-G [2]

  Data from the IEA Hydrogen database was used to gather information on hydrogen production and infrastructure projects. This includes the IEA 2023 Hydrogen Production Projects (corrected on 23/01/2024) and the IEA 2023 Hydrogen Infrastructure Projects (updated on 31/10/2023). It is worth noting that there appears to be some overlap between these two files.

- IEA hydrogen [3]

  Data from the IEA Hydrogen database was used to gather information on hydrogen production and infrastructure projects. This includes the IEA 2023 Hydrogen Production Projects (corrected on 23/01/2024) and the IEA 2023 Hydrogen Infrastructure Projects (updated on 31/10/2023). It is worth noting that there appears to be some overlap between these two files.

- GSE Storage data [4]

  The database contains operational data such as working gas volume, and withdrawal capacities and storage facilities as well as the under construction and planned storage sites.

- IEA Global Energy Review 2019 [5]

  The report provides an analysis of gas systems, including information on demand, production, and infrastructure data.

- Global Energy Monitor [6]

  It used for gathering the information if gas system for the North Sea region as for revision process for the Spanish case. It utilizes datasets such as the Global Oil & Gas Extraction Tracker and the Global Gas Infrastructure Tracker. However, these datasets do not include gas storage information. They do provide latitude and longitude coordinates, which can be mapped to NUTS3 region codes. The overall data quality appears to be high, though some minor issues exist, such as North Sea production fields being assigned to the wrong country and pipeline start and end points being misaligned. Additionally, pipelines passing through multiple NUTS3 regions mean that most onshore pipelines lack complete detail. However, pipeline capacity data from this source is generally more reliable than that of SciGrid.

- Various other sources, including company websites and country's central bureaus of statistics.

The main approach to preparing the final dataset involved using elements and data from the SciGrid files as the foundation. These were then validated and supplemented with information from other reliable sources. Where necessary, assumptions and heuristics were applied to fill data gaps and ensure completeness.

4. Geographical scope and regional classification

This data set is compiled at the NUTS3 level. It contains the entirety of Spain and Portugal, and several regions in France bordering to Spain with a natural gas pipeline connection. An Algerian and a Moroccan region have been added to allow pipeline imports into Spain. The regions Rotterdam in the Netherlands, and around Marseille, in France, have been added to allow modelling export scenarios. The needed data extracted from Eurostat-NUTS [7], SciGrid, and Google Maps.

At the time of preparing the data, the most recent NUTS classification was Eurostat NUTS 2024. Several input data sets were compiled for partially outdated NUTS3 regions definitions. As part of the data curation, every attempt was made to update these region definitions to align with the most recent version. This conversion has sometimes been based on finding addresses and/or comparing different geographical maps, so is a "best effort" exercise with a reasonably high level of accuracy, but no guarantees for specific items. It worth noting that regions without supply, demand and pipeline connections have been left out of the data set.

Mohamed Kais Msakni at NTNU has developed a Python script that maps NUTS3 region codes to latitude longitude coordinate pairs. If a region code cannot be assigned, it finds the nearest NUTS3 region and adds a warning flag that this is approximate". These entries are checked manually.

5. Production data

5.1 Natural gas

The main used data source was IGGIELGNC3_Productions.csv file that can be downloaded from SciGrid. Some values for large fields have been adjusted based on evidence indicating whether the fields were still active or decommissioned. Additionally, assumptions were made regarding Algerian gas supply and permitted supply to LNG import terminals where necessary.

5.2 Hydrogen

The main data source used was the IEA Hydrogen published data. Since many projects consume hydrogen at the same location where it is produced, they do not impact the overall network. Therefore, only projects intended for hydrogen transport or injection into the gas grid were selected. Capacities and demand values for natural gas in the MGET model are typically measured in gigawatts, while many projects operate at a megawatt scale. Since such small amounts can generally be blended into existing consumption and transport systems for natural gas, only projects with a capacity of at least 30 MW were included. However, if a project had 'previous stages' with a capacity below 30 MW, those were also included in the dataset.

The European union has set a goal for 2030 aimed for 10 Mtpa of hydrogen production and 10 Mtpa of import. Meeting these targets is estimated to require a total production and import capacity of 38 GW.

In the MGET case study, we have scaled up supply projects data for 2025 to 2030 and later years. To do so, an additional assessment was made based on wind and solar potential maps in conjunction with H2Med supply origins [8] [9] [10].

6. Demand

6.1 Natural gas

The main source files used were SciGrid and data on Gas-Fired Power Generation & "Other Demand." To create a representative hourly NUTS3-level demand as input for the model, gas-fired power generation capacities were combined with the four representative hourly demand values. Additionally, a reduction in natural gas consumption over time was assumed, covering the period from 2025 to 2050.

6.2 Hydrogen

An own assumption was made regarding the phased introduction and gradual increase of hydrogen consumption from 2025 to 2050. The model can impose hydrogen demand at the NUTS2 and/or at the NUTS3 level. NUTS2 level specification allows the model to find long-distance intra-regional "corridors".

7. Pipeline network capacities, existing and planned

7.1 Natural gas

The main source for natural gas data was SciGrid, though some issues were identified regarding directionality and capacity. For example, bidirectionality indicators for segments that meet the criteria for being bidirectional were not always set to 1, and the thresholds for determining bidirectionality appeared somewhat arbitrary. Additionally, two NUTS3 nodes lacked pipeline connections: ES512 (Girona), which receives gas from France, and PT11B (Soutelo de Aguiar), which also lacks consumption data, making its absence of connections reasonable. Some locations are also supplied by LNG tanker trucks, as indicated in the Enagás 2021 Spanish Gas System Report [11] p. 15.

Additional data sources included ENTSOG, as well as the Spanish and Portuguese network operators.

7.2 Hydrogen

The primary data source for the hydrogen pipeline network is the IEA Infrastructure Projects, with updates expected after June 2024. A blending limit of 2.0% hydrogen by volume has been applied at given locations. While this is a conservative estimate, hydrogen can be injected at multiple points along the network. However, tracking the exact hydrogen percentage in the gas flow is computationally intensive and is not done.

7.3 Network losses (efficiencies) due to compression

Network losses have been estimated based on pipeline lengths. For natural gas, losses are assumed to be 0.2% per 100 km, and for hydrogen, the estimated loss is 0.02% per 100 km.

8. Storage injection and extraction

8.1 Natural gas

The data source was SciGrid while capacity revisions and additional storages has been done based on GSE.

8.2 Hydrogen

The ability for repurposing of the existing infrastructures and capacity additions was assessed using GSE as the primary reference.

8.3 LNG import terminal capacities, existing and planned

While SciGrid has been used as the main data source, for capacity revisions GIIGNL was used.

9. Other

9.1 Volumetric conversion factors

Given a specific pressure level, per unit of energy hydrogen requires a higher volume than natural gas.

For storages, the amount of energy that can be stored in the form of hydrogen, may be a half to one order of magnitude lower than in the form of natural gas, dependent on various storage characteristics.

For pipelines, there are several sources indicating that at the same pressure difference between inlet and outlet, hydrogen molecules move so much faster, that the amount of energy transported is only about 10%-20% lower compared to natural gas molecules. Other sources, however, indicate that higher pressures are needed. This, but also higher-pressure variations over time, may cause higher wear and tear on pipelines if they are to be repurposed from natural gas to hydrogen.

Values used in June 2024:

- Pipelines: 1.3
- Storages 3.0

These values can be revised based on input from project partners and be infrastructure item specific.

9.2 Costs

There are no reliable cost estimates at the level of detail required. Given the nature of the optimization, it matters more that relative orders of magnitudes of values are correct, than that specific values are very precise. This is especially the case for repurposing of natural gas infrastructure to hydrogen and must be aligned with volumetric conversion factors.

For the analyses, the values that have the right order of magnitude have been used.

10. References

[1] "SciGRID Gas: Gas Infrastructure Data," [Online]. Available: https://www.gas.scigrid.de/.

[2] "ENTSOG: European Network of Transmission System Operators for Gas," [Online]. Available: https://www.entsog.eu/.

[3] "(IEA), International Energy Agency, Hydrogen Production and Infrastructure Projects Database," [Online]. Available: https://www.iea.org/data-and-statistics/data-product/hydrogen-production-and-infrastructure-projects-database#hydrogen-production-projects.

[4] "Storage Database - Gas Infrastructure Europe (GIE)," [Online]. Available: https://www.gie.eu/transparency/databases/storage-database/.

[5] International Energy Agency, "Global Energy Review 2019: The Latest Trends in Energy and Emissions in 2019," OECD Publishing, 2020.

[6] Global Energy Monitor, "Global Energy Monitor," [Online]. Available: https://globalenergymonitor.org/.

[7] Eurostat, "NUTS Maps - Eurostat," [Online]. Available: https://ec.europa.eu/eurostat/web/nuts/maps.

[8] H2med Project, "Europe’s first major green hydrogen corridor," [Online]. Available: https://h2medproject.com/.

[9] "The European Hydrogen Backbone (EHB) initiative," [Online]. Available: https://ehb.eu/.

[10] ATALAYA Generación, "Wind and Solar Sectors in Spain," [Online]. Available: https://www.atalaya.eu/en/renewables/wind-and-solar-sectors.php.

[11] Enagás, "2021 Spanish Gas System Report (Annual Report)," Enagás, 2021.

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open Source Toolbox

More information on iDesignRES project:

https://www.linkedin.com/company/idesignres

https://idesignres.eu/

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open Source Toolbox

North Sea Power System Data

WP1 -- Multi-Physics component models: Implementation and development

Task 1.1 – Data gathering at NUTS level 2: The North Sea.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Document information

Deliverable Number and Name	North Sea Power System Data
Work Package	WP.1, T1.1
Dissemination Level	Public
Author(s)	Ehsan Nokandi, Jannis Kerl
Primary Contact and Email	Ehsan Nokandi, ehsan.nokandi@ntnu.no
Date Submitted	24/2/2025
File Names	North_Sea_Power_System_Data.zip
Status	V.0.1

This document has been prepared in the framework of the European project iDesignRES. This project has received funding from the European Union's Horizon Europe programme under grant agreement no. 101095849.

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

The data presented is published as part of the iDesignRES project, specifically under WP 1, Task 1.1. This task involved collecting detailed energy system data, here for the North Sea region and focusing on electricity, to serve as test data for the models utilized within the iDesignRES project.

The data presented here covers electricity generation, consumption, and the power grid of the North Sea region and is used as a macro-region validation use case for testing and validating the project's models. It includes detailed information on power generation assets—both offshore and onshore—such as wind farms, thermal plants, and other renewable energy sources. Additionally, it contains the high-voltage transmission grid and interconnections between countries.

Data on load distribution through the grid, as well as load time series data for 2024 with hourly temporal resolution, has been embedded in the dataset. The dataset spans Norway, Germany, the Netherlands, Denmark, the United Kingdom, and Belgium, with France included for completeness.

2. Network Data

The network data for the North Sea power grid is scraped from PyPSA [1], which contains the European High voltage grid (220 kV to 750 kV) and is based on OpenStreetMap Data. This dataset is filtered to the needed countries for the iDesignRES North Sea-case - included are lines, links, buses and data for transformers of the said countries. In addition, connections to buses in neighboring countries are also included. The grid has been depicted in Figure 1.

The network data consists of the following files:

- buses.csv
- neighboring_buses.csv
- lines.csv
- links.csv
- transformers.csv
- converters.csv

2.1 buses.csv and neighboring_buses.csv

While buses.csv contains data on transmission buses located within North Sea countries, neighboring_buses.csv provides information on buses outside these countries that are directly connected to their transmission grid.

Both buses.csv and neighboring_buses.csv have the same structure. An excerpt from the bus_data.csv file is shown as an example in Figure 2. The files contain the following columns.

bus_id: Contains a unique identifier for each bus in the system.

voltage: Represents the voltage level of each bus, expressed in kilovolts (kV). The voltage levels range from 220 kV to 520 kV.

dc: Indicates whether the bus is located in the DC part of the grid or not. The values can be f (False) or t (True)

y and x: Contain latitude and longitude of the buses respectively.

country: Contains country code where the bus is located. The country codes are based on ISO 3166-1 Alpha-2 standard.

An excerpt from the buses.csv and neighboring_buses.csv files can be seen in Figure 2.

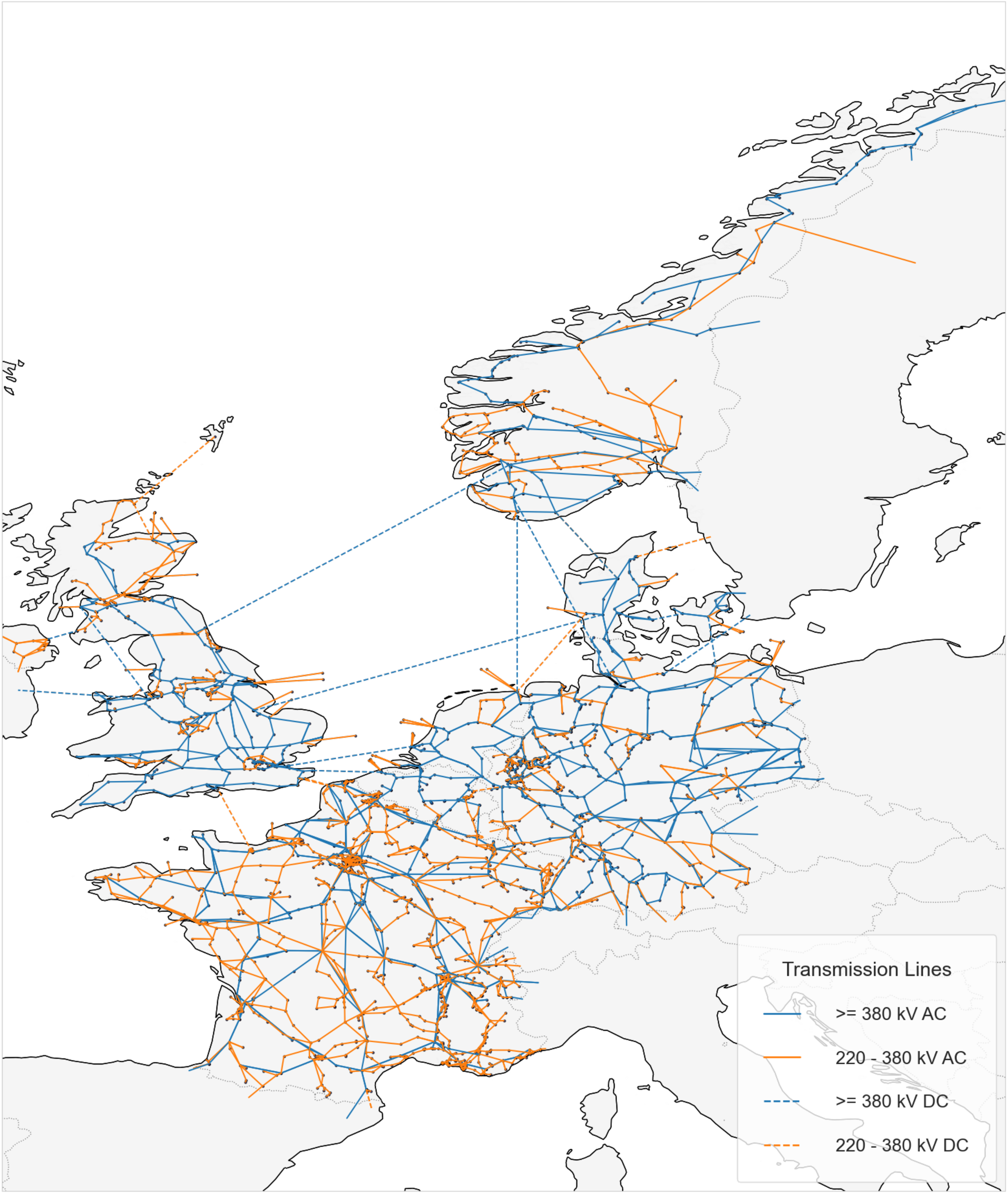


Figure 1: The power transmission grid of the North Sea countries, obtained from PyPSA.

bus_id	voltage	dc	x	y	country
BE1-220	220	f	2.917742	51.6198	BE
DE100-22(	220	f	7.135589	51.41707	DE
DE100-38(	380	f	7.135882	51.41726	DE
FR80-400	400	f	3.489653	48.44471	FR
FR81-225	225	f	-4.18182	48.42753	FR

Figure 2: An excerpt from the buses.csv and neighboring_buses.csv files.

2.2 lines.csv

The lines.csv file contains information regarding the AC transmission lines of the grid. The file includes the following columns.

line_id: A unique identifier for each transmission line in the system.

bus0: Indicates the "from" bus of the transmission line, represented by its unique bus ID.

bus1: Indicates the "to" bus of the transmission line, represented by its unique bus ID.

voltage: Represents the voltage level of the transmission line, expressed in kV. The voltage level of the dataset's lines varies from 220 to 420 kV.

i_nom: Specifies the maximum allowable current for the transmission line circuit, expressed in kiloamperes (kA).

circuits: Specifies the number of circuits that the transmission line includes.

s_nom: Specifies the nominal power transfer capability of the transmission line, expressed in mega voltamperes (MVA).

r: Contains the resistance of the transmission line, measured in ohms (Ω).

x: Contains the reactance of the transmission line, measured in ohms (Ω).

b: Contains the susceptance of the transmission line, measured in siemens (S).

length: Provides the physical length of the transmission line in meters (m).

underground: Indicated whether the line is underground (t) or overhead (f).

Type: Indicates the conductor type of the line.

An excerpt from the lines.csv file can be seen in Figure 3.

line_id	bus0	bus1	voltage	i_nom	circuits	s_nom	r	x	b	length	undergrounc	type
merged_ı	way/88	way/50	220	1.29	1	491.56	0.57	2.87	3.75e-0	9550.2	f	Al/St 240/40 2-bundle 220.
merged_ı	way/66	way/39	220	1.29	1	491.56	4.47	22.4	0.0003	74463	f	Al/St 240/40 2-bundle 220.
merged_ı	way/15	way/66	220	1.29	1	491.56	2.4	12	0.0002	40032	f	Al/St 240/40 2-bundle 220.

Figure 3: An excerpt from lines.csv file.

2.3 links.csv

The links.csv file contains information regarding the DC links of the grid. The file includes the following columns:

link_id: Contains a unique identifier for each link in the dataset.

bus0: Indicates the "from" bus of the link, represented by its unique bus ID.

bus1: Indicates the "to" bus of the link, represented by its unique bus ID.

voltage: Represents the voltage level of the link, expressed in kV.

p_nom: Specifies the nominal active power transfer capability of the link, expressed in megawatts (MW).

length: Provides the physical length of the link, expressed in meters (m).

An excerpt from the links.csv file can be seen in Figure 4.

link_id	bus0	bus1	voltage	p_nom	length
relation/10377412-320-DC	way/753113423	way/636630368	320	1000	230137
relation/13295785-515-DC	relation/12832053	way/642490160	515	1400	723362
relation/14126301-450-DC	way/109189896	way/920127890	450	1000	253393

Figure 4: An excerpt from links.csv file.

2.4 transformers.csv

The transformers.csv file contains information about the transformers in substations of the transmission system. The file includes the following columns:

transformer_id: Contains a unique identifier for each of the transformers in the dataset.

bus0: Indicates the bus on the primary side of the transformers, represented by its unique bus ID.

bus1: Indicates the bus on the secondary side of the transformers, represented by its unique bus ID.

voltage_bus0: Represents the voltage level of the primary side of the transformers, expressed in kV.

voltage_bus1: Represents the voltage level of the secondary side of the transformers, expressed in kV.

s_nom: Specifies the nominal power of the transformers, expressed in mega voltamperes (MVA).

An excerpt from the transformers.csv file can be seen in Figure 5.

transformer_id	bus0	bus1	voltage_bus0	voltage_bus1	s_nom
DE100-220-380	DE100-220	DE100-380	220	380	3397
DE102-220-380	DE102-220	DE102-380	220	380	3397
DE107-220-380	DE107-220	DE107-380	220	380	3397

Figure 5: An excerpt from transformers.csv file

2.5 converters.csv

The converters.csv file contains information about the converters of the system which connects AC and DC parts of the grid. The file includes the following columns.

converter_id: Contains a unique identifier for each of the transformers in the dataset.

bus0: Indicates the bus on the primary side of the transformers, represented by its unique bus ID.

bus1: Indicates the bus on the secondary side of the transformers, represented by its unique bus ID.

voltage: Represents the voltage level of the connected DC link, expressed in kV.

p_nom: Specifies the nominal active power of the converters, expressed in megawatts (MW).

An excerpt from the converters.csv file can be seen in Figure 5.

converter_id	bus0	bus1	voltage	p_nom
conv-relation/10695345	relation/10695345	relation/10695345-400	400	1000
conv-relation/11981383	relation/11981383	relation/8894066-420	450	700
conv-relation/12832053	relation/12832053	relation/12832053-420	515	1400

Figure 6: An excerpt from converters.csv file

3. Generation data

The power-generation dataset contains data on active power plants in the considered North Sea region and the included countries, with France added. It contains information on the type of powerplant and its geographical location.

The Energy Types column includes Nuclear, Hard Coal, Hydro, Lignite, Oil, Natural Gas, Solid Biomass, Wind, Solar, Other, Biogas, and Waste, while the Technologies column contains Steam turbine, Steam Turbine, Run-Of-River, Pumped Storage, Reservoir, CCGT, Offshore, Onshore, PV, Marine, Combustion Engine, PV, and CSP.

The generation data is gathered from PyPSA-Europe, specifically the most recent powerplantmatching-publication [2]. It merges powerplant-data from several sources and databases, making it the most comprehensive and high-quality dataset available for European power plants.

The processing is carried out by filtering out powerplants to include only those with no specified decommissioning date or with dates later than 2024, ensuring all active plants up to 2024 are considered.

Due to varying data across countries for solar and wind installations, their values are proportionally scaled using available, aggregated ENTSO-E data. Additionally, for French nuclear power, the more accurate and detailed Global Energy Monitor data [3] is used instead of PyPSA data. A comparison of the total installed capacity per energy source for each country in our dataset and ENTSO-E is illustrated in Figure 7 .

Note that the granularity of technologies varies between the PyPSA-data and the ENTSO-E source. Therefore, differences might occur. Also note that some of the technologies covered by the PyPSA-data are not covered by ENTSO-E, resulting in differences (most notable for Denmark).

Lastly, powerplants are assigned to their geographically closest buses based on the given coordinates.

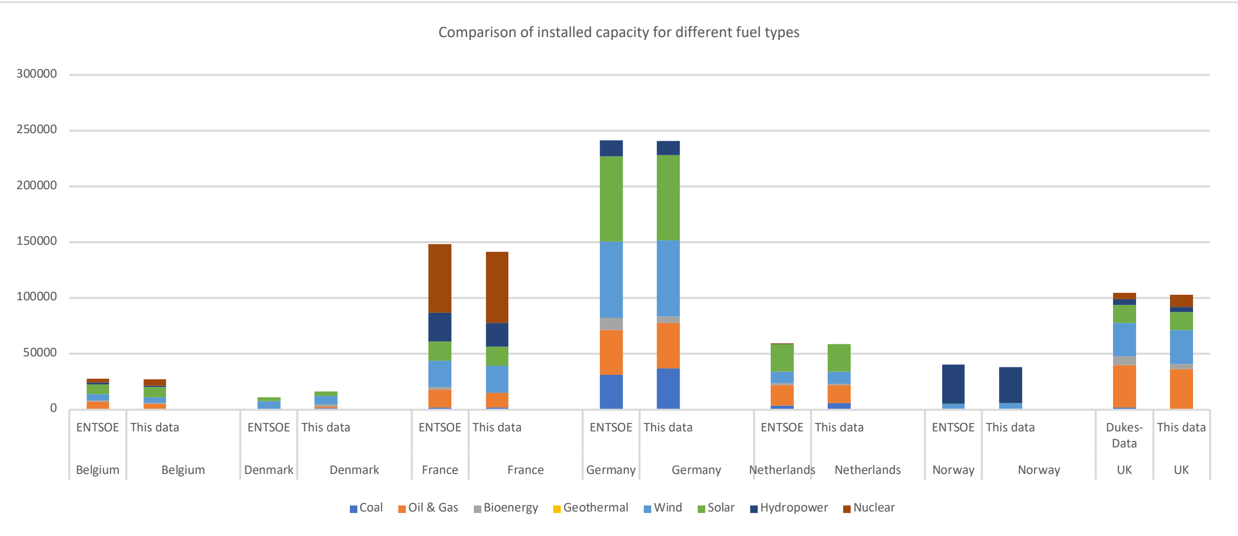


Figure 7: Comparison of installed capacity per fuel type between ENTSO-E and this dataset.

3.1 generation.csv

The file includes the following columns.

unit_id: A unique identifier assigned to each generation unit, following the format "G0001", "G0002", etc.

country: The country where the power plant is located

source: The primary energy source used for power generation, such as "Nuclear" or other fuel types

technology: The specific technology used for power generation, e.g., "Steam turbine"

capacity_mw: The installed capacity of the power plant, measured in megawatts (MW)

y: The geographical latitude coordinate of the power plant

x: The geographical longitude coordinate of the power plant

bus_id: The identifier of the electrical bus to which the power plant is connected, based on the uploaded network-data

An excerpt from the generation.csv file has been depicted in Figure 8.

unit_id	country	source	technolo	capacity_	y	x	bus_id
G0001	FR	Nuclear	Steam tu	1363	47.51	2.87501	way/41829825-400
G0002	FR	Nuclear	Steam tu	1363	47.51	2.87501	way/41829825-400
G0003	FR	Nuclear	Steam tu	951	45.256	-0.6932	way/105515586-225

Figure 8: An excerpt from the generation.csv.

4. Load data

The demand data consists of both a static and a time-dependent component. The static component addresses how load is distributed throughout the grid, while the time-dependent component represents the actual hourly load values over a specific period, in our case, the year 2024.

4.1 load_distribution.csv

This file contains information about the share of each bus in the total electricity consumption of its respective country. Consequently, the sum of the values assigned to the buses in each country equals one.

The primary data source used is the EU Energy Atlas [4]. This dataset consists of a series of maps of the EU in TIFF format, representing the demand for major energy product groups across various economic activities. The data is based on the 2019 energy balances and the 1×1 km reference grid from EUROSTAT and the JRC’s Energy and Industry Geography Lab.

The EU Energy Atlas provides an estimate of electricity consumption in EU countries with a spatial resolution of 1 square kilometer. It categorizes consumption by various sectors, such as industry and transportation, while also offering aggregated electricity consumption across all sectors.

Although this data is from 2019, it is assumed that while total electricity consumption may have changed between 2019 and 2024, the spatial distribution of electricity consumption across Europe has remained approximately constant.

To estimate how the load is distributed through the grid, each pixel is assigned to the closest transmission grid of the respective country. Figure 9 illustrates the estimated of electricity demand of EU countries, as provided by EU Energy Atlas. However, the Atlas data does not cover Norway or the United Kingdom.

For the United Kingdom, the current version of our data relies solely on population distribution to estimate the load distribution within the network.

For Norway, a combination of geographical locations and annual electricity consumption of major industrial and oil sites has been used to refine the load distribution estimate, improving upon what would be derived solely from population distribution. Figure 10 illustrates population density alongside the electricity consumption of industrial and oil sites in Norway.

Finally, Figure 11 illustrates the final estimated load distribution through the transmission grid of the North Sea countries.

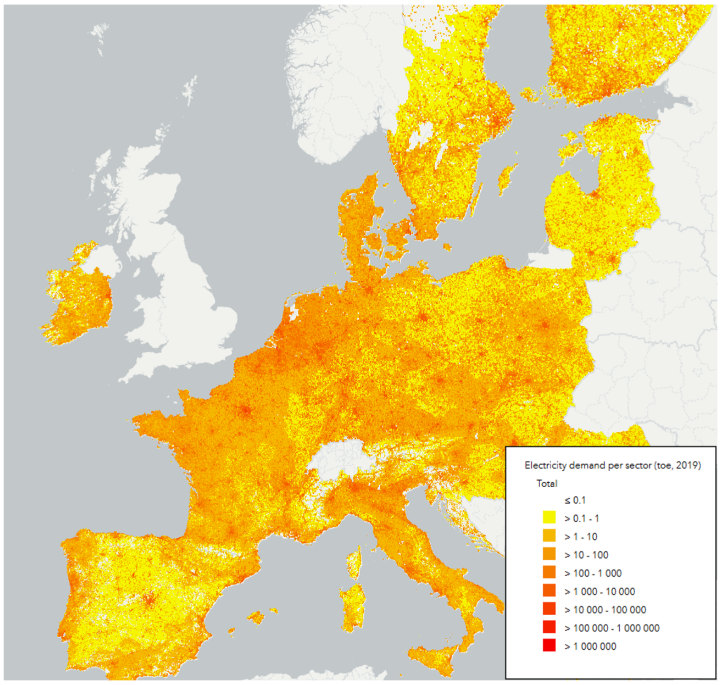


Figure 9: The estimation of electricity demand of Europe in 2019 by EU Energy Atlas (figure from Energy and Industry Geography Lab [5]).

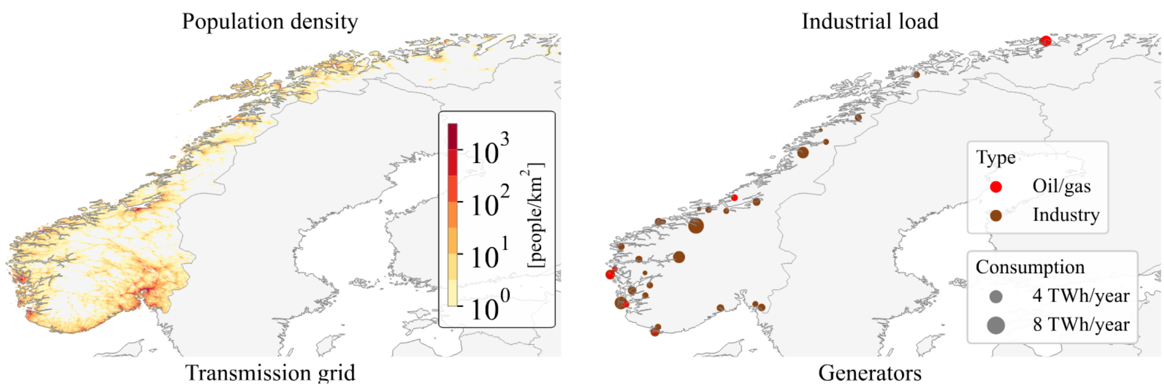


Figure 10: The data of the population density and consumption of the industrial and oil sites of Norway

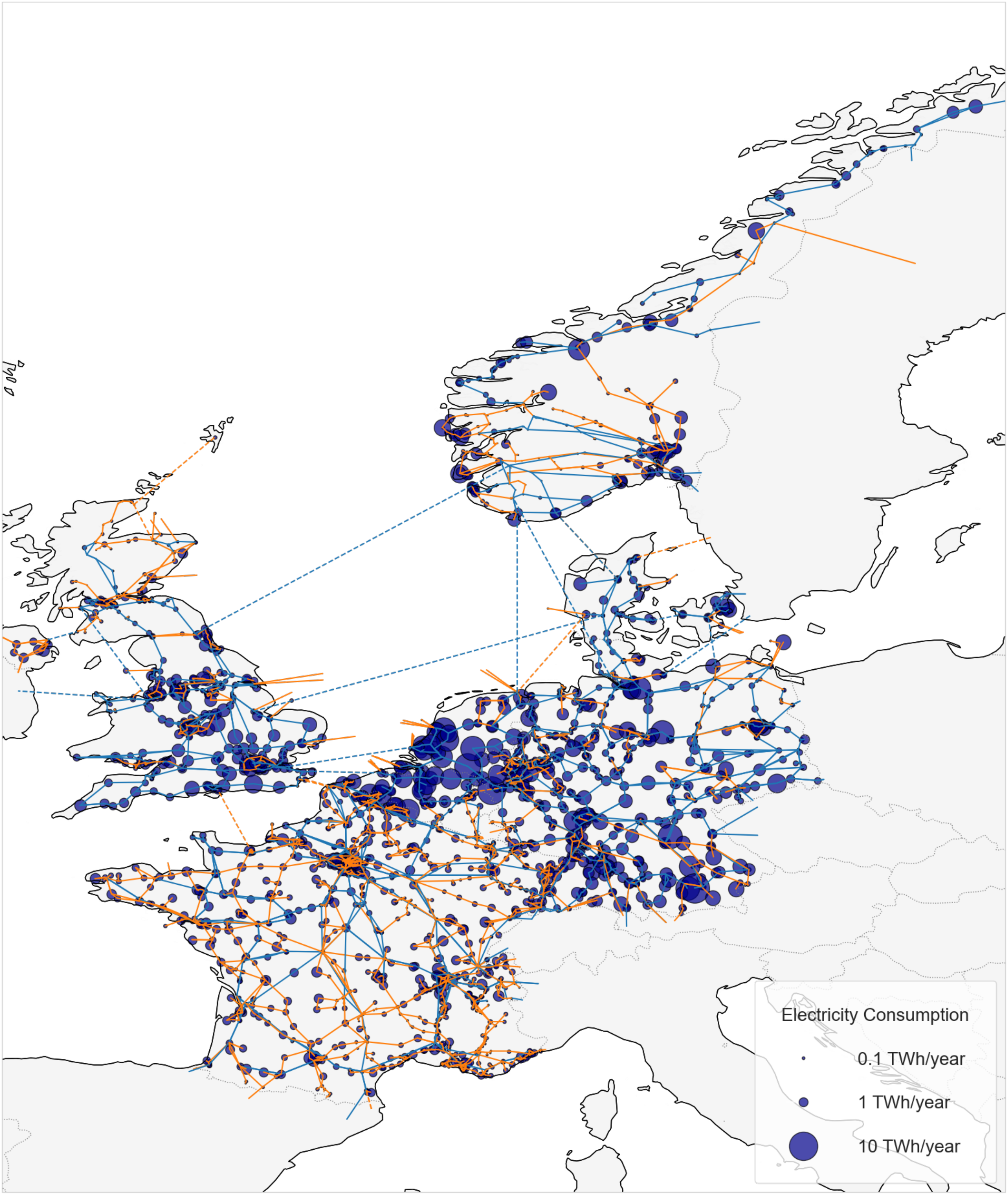


Figure 11: The estimated load distribution through the transmission grid of the North Sea countries.

4.2 load_time_series.csv

The file contains electricity consumption data for the North Sea countries and France in hourly intervals throughout the year 2024. The data was obtained from the ENTSO-E Transparency Platform [6] and from the National Energy System Operator (NESO) of the United Kingdom.

The dataset contains the following columns:

timestamp: This column specifies the timestamp for each recorded data point, indicating the exact date and time at 1-hour intervals throughout the year 2024. The timestamps are in ISO 8601 format, including the time zone offset.

{Country}: This column represents the electricity consumption in the respective country at the corresponding time slot. The values are expressed in megawatts (MW), indicating the total power demand during the 1-hour interval.

An excerpt from the file is shown in Figure 10.

timestamp	Belgium	Denmark	France	Germany	Netherlai	Norway	United Kingdom
1/1/2024 0:00	7335	3704	52846	39337	11312	18364	21783
1/1/2024 1:00	7143	3638	51599	38408	11127	18224	22194
1/1/2024 2:00	6878	3660	50992	37308	10848	18030	20619

Figure 10: An excerpt from the load_time_series_data.csv file.

The load data was gathered through the ENTSO-E transparency platform, which contains time series on loads for different European countries. Since there is no data for the UK from 2020 onwards, datasets provided in [7] by the national system operator (NESO) of the UK are used. There, the entries of ND (National Demand) are used to create the load time series. It is defined as:

"*National Demand is the sum of metered generation, but excludes generation required to meet station load, pump storage pumping and interconnector exports.*" [8] This is used since it is the closest load definition compared to the ENTSO-E load data.

5. References

[1] B. F. D. N. F. R. I. &. B. T. Xiong, „Prebuilt Electricity Network for PyPSA-Eur based on OpenStreetMap Data," 2024. [Online]. Available: https://zenodo.org/records/14144752.

[2] F. H. e. al., „PyPSA/powerplantmatching: v0.7.1," [Online]. Available: https://zenodo.org/records/14785651.

[3] Global Energy Monitor, „Global Energy Monitor," [Online]. Available: https://globalenergymonitor.org/.

[4] I. H. Gonzales und A. Uihlein, „High-resolution energy atlas," *Publications Office of the European Union,* 2023.

[5] European Commission, „Energy and Industry Geography Lab," EU Science Hub, [Online]. Available: https://energy-industry-geolab.jrc.ec.europa.eu/.

[6] ENTSO-E, „ENTSO-E Transparency Platform," [Online]. Available: https://transparency.entsoe.eu/. [Zugriff am 10 12 2024].

[7] N. -. N. E. S. Operator, „Historic Demand Data," 2024. [Online]. Available: https://www.neso.energy/data-portal/historic-demand-data.

[8] N. -. N. E. S. Operator, „Historic Demand Data 2023," 2023. [Online]. Available: https://www.neso.energy/data-portal/historic-demand-data/historic_demand_data_2023.

[9] Eurostat, „NUTS Maps - Eurostat," [Online]. Available: https://ec.europa.eu/eurostat/web/nuts/maps.

[10] „PyPSA-Eur: A free, open model of the European power system," 2024. [Online]. Available: https://github.com/PyPSA/pypsa-eur.

[11] „Overpass Turbo: OpenStreetMap Query Wizard," [Online]. Available: https://overpass-turbo.eu/.

[12] „OpenStreetMap: Interactive Map," [Online]. Available: https://www.openstreetmap.org/. [Zugriff am 30 09 2024].

[13] „Global Energy Monitor," [Online]. Available: https://globalenergymonitor.org/. [Zugriff am 21 11 2024].

[14] „ENTSO-E Transmission System Map," [Online]. Available: https://www.entsoe.eu/data/map/.

[15] J. H. D. S. T. Brown, „PyPSA: Python for Power System Analysis," 2018. [Online]. Available: doi: 10.5334/jors.188.

[16] D. F. F. N. I. R. T. B. Bobby Xiong, „Modelling the High-Voltage Grid Using Open Data for Europe and Beyond," 2024. [Online]. Available: https://arxiv.org/abs/2408.17178.

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open Source Toolbox

More information on iDesignRES project:

https://www.linkedin.com/company/idesignres

https://idesignres.eu/

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open Source Toolbox

Heat demands: potentials for district heating and waste heat at NUTS2 level

WP1 -- multi-physics component models: Implementation and development

Task 1.1 -- Centralized data collection and validation

Task 1.3 -- Energy transmission/distribution and storage

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Document information

Deliverable Number and Name	Data for district heating and waste heat potential
Work Package	WP1
Dissemination Level	Public
Author(s)	Marina Georgati, Steffen Nielsen
Primary Contact and Email	marinag@plan.aau.dk
Date Submitted	17/03/2025
File Name	district_heating_excess_heat.zip
Status	v.0.1

This document has been prepared in the framework of the European project iDesignRES. This project has received funding from the European Union's Horizon Europe programme under grant agreement no. 101095849.

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

The data presented here has been published as part of the iDesignRES project under Work Package 1 (WP1), specifically for Task 1.1 and Task 1.3. These tasks focused on data collection and comprehensive GIS-based modelling to generate high-resolution spatial data on heat demand, district heating potentials across Europe, and waste heat sources with their corresponding potential.

The dataset includes aggregated estimates at the NUTS2 level for heat demand in residential and service buildings; district heating potential—covering residential, service, and industrial sectors—along with associated investment costs based on a 55% market share; and the potential of waste heat, disaggregated by source and temperature level. The data spans all EU27 countries as well as the United Kingdom (UK). The 55% market share refers on an accumulated share across EU27, but varies among the countries, depending on where the potentials are high.

2. Data presentation

The dataset is published in an Excel file entitled dh_eh.xlsx, which contains one sheet with the following attributes:

- **NUTS_ID**: Unique identifier of NUTS2 regions.
- **LEVL_CODE**: Level of country division by NUTS, equal to 2.
- **CNTR_CODE**: Code of country in two-letter abbreviations.
- **Heat demand [GWh]:** Heat demand of residential and service buildings, estimated in GWh.
- **DH heat demand [Total,GWh]:** Heat demand in district heating areas including both industrial, residential and service demands, estimated in GWh.
- **DH heat demand [Res/Ser,GWh]:** Heat demand in district heating areas including residential and service demands, estimated in GWh.
- **DH heat demand [Ind,GWh]:** Heat demand in district heating areas including both industrial, estimated in GWh.
- **Grid cost [MEUR]:** Total investment costs for district heating areas in the corresponding NUTS2 areas, estimated in MEUR.
- **Waste heat [Total,GWh]:** Total waste heat potentials overlapping in the modelled district heating areas aggregated at NUTS2 level, summed of the below categories. Estimated in GWh.
- **Industry High Temp [GWh]:** Waste heat from industrial activities in high temperatures. Estimated in GWh.
- **Industry Medium Temp [GWh]:** Waste heat from industrial activities in medium temperatures. Estimated in GWh
- **Industry Low Temp [GWh]:** Waste heat from industrial activities in low temperatures. Estimated in GWh.
- **WtE [GWh]:** Waste heat from waste to energy in high temperatures. Estimated in GWh.
- **WWT [GWh]:** Waste heat from wastewater treatment plants in medium temperatures. Estimated in GWh.
- **Supermarkets [GWh]:** Waste heat from supermarkets in low temperatures. Estimated in GWh.
- **Food [GWh]:** Waste heat from food production in low temperatures. Estimated in GWh.
- **Metros [GWh]:** Waste heat from metros in low temperatures. Estimated in GWh.

3. Data generation

3.1 Heat demands

Building on the work of the sEEnergies project [1] and incorporating the heat scenarios from the Heat Roadmap Europe project [2], the Pan-European Thermal Atlas has undergone a comprehensive update. This revision significantly enhances the accuracy of heat demand estimates, particularly in previously misrepresented uninhabited areas where earlier versions erroneously allocated demand. The updated methodology involves a detailed disaggregation process that breaks down national-level heat demand estimates by leveraging the spatial distribution of built area, differentiated by construction period. This approach ensures that demand is allocated in a way that reflects the age and type of building stock across regions. A thorough description of this methodology is provided in a working paper titled *Assessing the regional differences of heat market shares with DH potential* by the report's authors. The heat atlas covers all EU27 countries and the UK, excluding the French, Spanish, and Portuguese islands in the Atlantic.

The dataset uploaded here includes a key attribute, 'Heat demand [GWh]', which is derived through spatial analysis from the above heat atlas. Specifically, the zonal statistics function from the rasterstats library [3] is used in Python in combination with the rasterio package. This function calculates the sum of heat demand raster values within each polygon of the NUTS2 administrative boundaries vector layer—finally representing the total heat demand in the NUTS2 areas required by residential and service buildings.

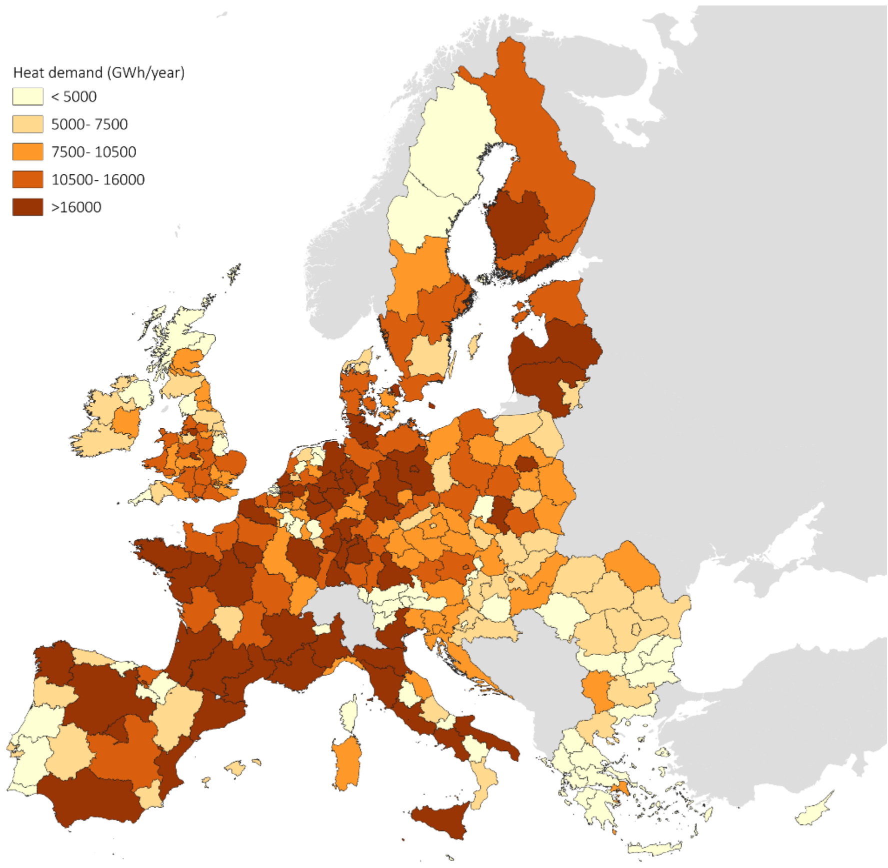


Figure 3.1 Heat demand distribution at NUTS2 level in Europe.

3.2 District heating potentials

Using the updated Pan-European Thermal Atlas (PETA6) as input, representing heat demand at a spatial resolution of 100 meters, the district heating potentials are estimated following the methodological approach outlined by [4]. In addition to delineating the boundaries of potential district heating areas, the model also estimates the corresponding investment costs and heat losses for both 3rd and 4th generation systems.

The model requires district heating pipe costs as input, which are simplified using two cost factors: c1 [EUR/m] and c2 [EUR/m²]. While in practice these costs vary across countries and cities, for consistency with energy system models, the same cost assumptions are applied uniformly across all countries. The cost values used are 664 EUR/m for c1 and 2,610 EUR/m² for c2, based on updated German cost data from [5].

In addition to estimating heat demand for the residential and service sectors, the district heating areas are further refined to incorporate industrial heat demand within the corresponding zones. To achieve this, areas classified as industrial land use according to the CORINE land cover dataset (code 121) are used as guide. The national-level industrial heat demand, obtained from EnergyPLAN, are then spatially distributed across these areas in proportion to their overlap with the identified district heating zones.

Once all potential district heating areas are modelled including heat demand for residential, service and industrial buildings, their suitability is assessed by prioritizing areas according to the ratio of specific energy costs to heat demand. This approach favours larger areas with higher heat demand—even if associated costs are higher—as these tend to be more cost-effective for long-term investments. Then the potential district heating market share across the EU27 is estimated, focusing on areas that collectively support achieving a 55% market share, as this is the proposed target in the Heat Roadmap Europe project [2]. Areas exceeding this threshold are excluded from further analysis.

For the UK, a different approach is applied as it is not included in the Heat Roadmap Europe analysis. Instead of estimating the district heating market share, a cost threshold is defined based on the maximum investment costs observed within the selected EU27 dataset. This threshold serves as a realistic upper limit, under the assumption that investments exceeding this level would likely be economically unfeasible in practice. By aligning the UK assessment with the highest cost observed in the EU27, a conservative yet consistent framework for evaluating district heating potential outside the EU27 scope is maintained.

The dataset provided here includes the attributes of 'DH heat demand [Total,GWh]', 'DH heat demand [Res/Ser,GWh]', 'DH heat demand [Ind,GWh]' and 'Grid cost [MEUR]', deriving from the summation of the corresponding attributes in the DH areas to the NUTS2 level, with each district heating area being assigned to the NUTS2 region where it has the greatest overlap.

While this method generally works well, four problematic cases were identified in which the district heating heat demand exceeds the total heat demand for the NUTS2 region. This issue arises when district heating areas span across multiple NUTS2 boundaries. Potential solutions could involve the dissolvement of the district heating areas to the administrative boundaries of the local authorities that are responsible for the development of district heating systems in each country.

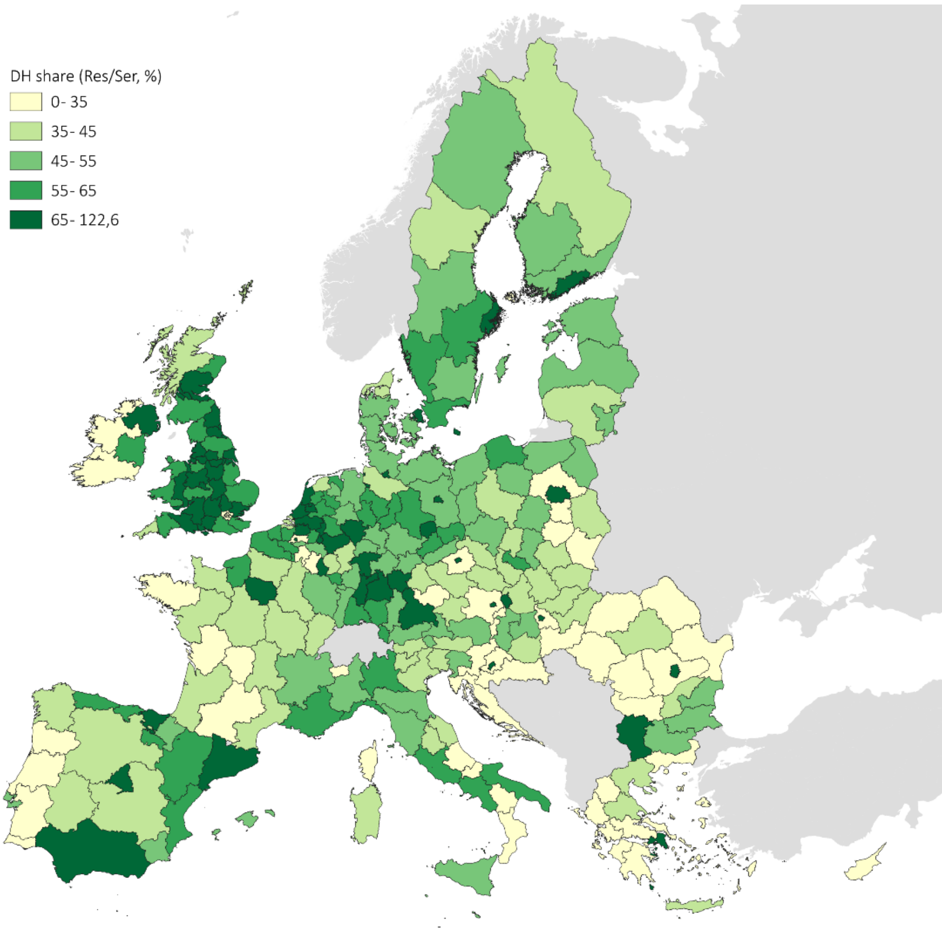


Figure 3.2 Potential district heating share at NUTS2 level in Europe for a EU27 market share of 55% [2].

3.3 Waste heat potentials

In order to assess the potential of waste heat sources in district heating areas, new datasets were developed using a top-down approach to complement existing point data from sEEnergies on current high, medium and low temperature sources.

For Waste to Energy (WtE), cities with a population of around 60,000 were identified as suitable locations based on a sensitivity analysis comparing city size and existing WtE potential. City boundaries were defined using the GHS POP layer and industrial areas within these boundaries were identified using CORINE land cover data. National WtE potentials derived from the Heat Roadmap Europe analysis were then downscaled to city level using a factor that takes into account both population and industrial area size. In cases where a WtE site overlapped several DH areas - such as in Cologne - the excess heat was allocated proportionally based on the local heat demand.

For industrial surplus heat (EH), the national technical potential from the IndustryPLAN model [2] was divided into high (H), medium (M) and low (L) temperature categories. High temperature sources were mapped using the location of existing industrial sites, while medium and low temperature potentials were distributed in proportion to the extent of industrial sites within each EH area.

Additional low and medium temperature sources - such as wastewater treatment plants, supermarkets and metro systems - were mapped using point data from PETA5.2 and the ReUseHeat project.

Their heat potentials were mapped to the modelled DH areas by spatial intersection, and after the filtering process described in the previous section, the waste heat potentials in each DH area were aggregated to the NUTS2 level, resulting in the following attributes of the uploaded dataset: ‘Industry High Temp [GWh]’, ’Industry Medium Temp [GWh]’, ‘Industry Low Temp [GWh]’, ‘WtE [GWh]’, ‘WWT [GWh]’, ‘Supermarkets [GWh]’, ‘Food [GWh]’, ‘Metros [GWh]’, and the total sum of the above attributes at ‘Waste heat [Total,GWh]’

4. REFERENCES

[1] Möller B, Wiechers E, Sánchez-García L, Persson U. Spatial models and spatial analytics results. 2022.

[2] Mathiesen BV, Gonzalo EG, Georgati M, Nielsen S, Nikolic J. Heat Roadmap Europe – a long-term path to resilient and competitive heating systems. Working report

[3] rasterstats — rasterstats 0.10.3 documentation. 2025.

[4] Fallahnejad M, Kranzl L, Haas R, Hummel M, Müller A, García LS et al. District heating potential in the EU-27: Evaluating the impacts of heat demand reduction and market share growth. Appl Energy 2024;353:122154.

[5] Sánchez-García L, Averfalk H, Persson U. sEEnergies special report : Construction costs of new district heating networks in Germany. 2022:26.

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open Source Toolbox

More information on iDesignRES project:

https://www.linkedin.com/company/idesignres

https://idesignres.eu/

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open Source Toolbox

Publicly available Gas network datasets & models

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Document information

Deliverable Number and Name	Publicly available Gas network datasets & models. Part of D1.1.
Work Package	WP.1, T1.1 & T1.3
Dissemination Level	Public
Author(s)	Parinaz Aliasghari, Pedro Crespo del Granado, Jannis Kerl
Primary Contact and Email	Parinaz Aliasghari, parinaz.aliasghari@ntnu.no
Date Submitted	May 30th 2025
File Name	
Status	V0.0.1

This document has been prepared in the framework of the European project iDesignRES. This project has received funding from the European Union's Horizon Europe programme under grant agreement no. 101095849.

TABLE OF CONTENTS

1. Executive Summary

Gas infrastructure continues to play a significant role in the European energy system—both as a source of primary energy and as critical infrastructure for energy transport and flexibility. Natural gas remains a key fuel in industry, heating, and electricity generation. Simultaneously, the existing gas network is expected to adapt to accommodate emerging energy carriers such as hydrogen, especially within decarbonization pathways.

Accurate representation of the gas sector in energy system models is therefore essential. This includes the availability and integration of relevant data on infrastructure, demand, and flows. Depending on a study's objective and scope, different spatial resolutions and modeling approaches may be appropriate. Coarse-grained models at national or regional level are well-suited for long-term scenario analysis, while high-resolution network models are necessary for identifying infrastructure bottlenecks, assessing cross-border capacity, and optimizing flow dynamics.

In this context, we compare two complementary gas infrastructure datasets used in the iDesignRES project's MGET energy system model:

- **The Global Gas Model (GGM)** offers a broad global market perspective at national (NUTS0) and major regional (NUTS1) levels. It supports long-term infrastructure and trade scenario analysis through a multi-period equilibrium framework.
- **SciGRID_gas** provides a detailed spatial representation of Europe's gas transmission network at NUTS3 level. Its high spatial resolution allows for flow-based simulations and the identification of regional constraints, which are critical for infrastructure investment decisions and transition modeling.

A Spanish case study validated the importance of high-resolution modeling for detecting localized bottlenecks. Building on these insights, a NUTS2-level North Sea dataset—developed using SciGRID methods and structured for GGM compatibility—is in progress. This dataset aims to bridge the gap between detailed local analysis and global market modeling, enhancing regional planning capabilities in integrated energy system studies.

Together, these datasets enable consistent and transparent modeling of the gas sector at multiple scales, forming the basis for robust analysis within the MGET framework and beyond.

2. Gas modeling at different spatial resolutions

Low-resolution models use aggregated nodes for entire countries or regions, enabling global simulations but hiding intra-national bottlenecks. High-resolution models map individual pipeline segments, allowing for localized simulations, bottleneck detection, and more accurate infrastructure assessments. Combining both scales supports robust energy planning model and dataset for Multi-gas energy transition (MGET) model.

2.1 GGM (NUTS0)

The Global Gas Model (GGM) is a multi-period partial equilibrium model of the world's natural gas market ntnu.edu. It captures the entire value chain from production to final consumption in an optimization framework. The model is implemented in GAMS as a quadratic program and computes a market equilibrium (maximizing overall welfare/profit) subject to infrastructure and supply–demand constraints. The model's core structure is illustrated in Figure 1, which shows how gas flows are modeled between countries via pipelines and LNG trade, supported by storage and processing facilities.

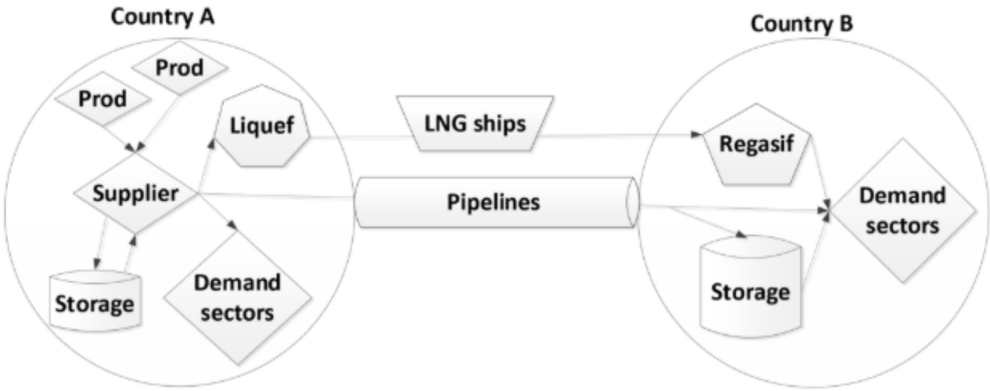


Figure 1: Simplified structure of the Global Gas Model (GGM) Source: GGM-Website

In each country, producers supply gas to a national supplier, which distributes it to domestic demand sectors or exports it via pipelines or as LNG. Liquefaction (Liquef) enables gas to be shipped overseas, while regasification (Regasif) terminals receive LNG in importing countries. Storage facilities provide buffer capacity in both supply and demand regions. This structure allows GGM to simulate international gas flows, accounting for transport modes, infrastructure constraints, and market interactions. Key features and structure of GGM include:

- **Geographic Scope and Spatial Resolution (NUTS0–1):** GGM models 90+ countries, each represented as a single node (NUTS0). Larger countries (e.g., USA, Russia) are split into several sub-national nodes (NUTS1-like) for regional accuracy.
- **Temporal and Seasonal Resolution:** GGM uses 5-year time steps from a base year (e.g., 2020) up to 2050 to reflect investment dynamics and long-term planning. Each year is divided into seasonal segments (e.g., summer/winter) to model demand variation and enable storage-cycle representation.
- **Full Value Chain Modeling & Optimization:** GGM captures the full gas value chain (production, pipelines, LNG, storage) and solves a quadratic program to compute market equilibrium under capacity and cost constraints.
- **Scenario-Based Input–Output Structure:** The model uses input data on demand, production, infrastructure, and costs to simulate future market states. Outputs include regional flows, infrastructure use, prices, and welfare indicators.

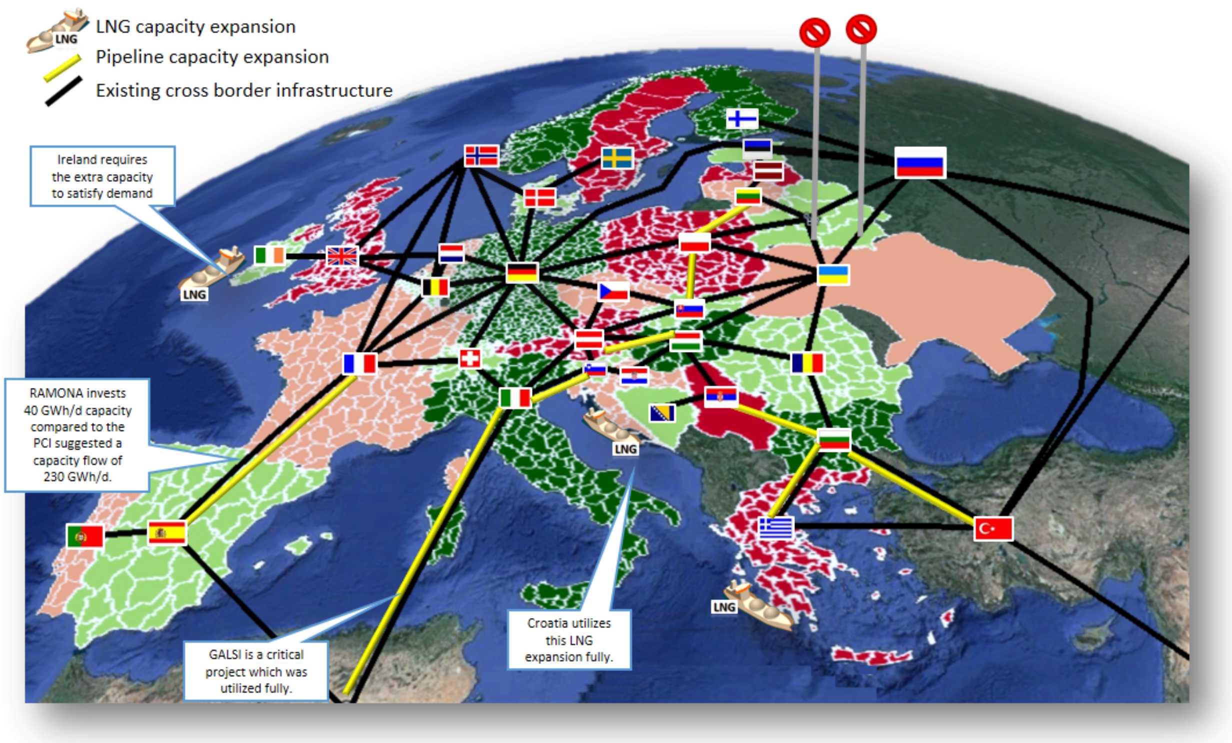


Figure 2: Visualization of the GGM-Model on a European Scale, Source: GGM-Website

2.2 High spatial resolution gas network

To support more detailed analysis in MGET, high-resolution gas network data is essential. Unlike aggregated country-level models, these datasets capture infrastructure and demand at the subnational level, allowing realistic simulation of gas flows, regional constraints, and investment needs. The following dataset, SciGRID_gas, serves as the core input for modeling Europe's gas grid at this level of detail.

2.2.1 SciGRID_gas (NUTS-3)

SciGRID_gas is an open-source geospatial gas network dataset that represents Europe's gas transmission system in high detail. Instead of aggregating entire countries into one node, SciGRID_gas maps out the actual pipeline network and nodes at a roughly NUTS3 level of granularity (district/county level within countries) for EU and neighboring states. The data model is structured as a graph: physical components are nodes, and pipelines are edges connecting them. An overview of this network structure and component types is illustrated in Figure 3.

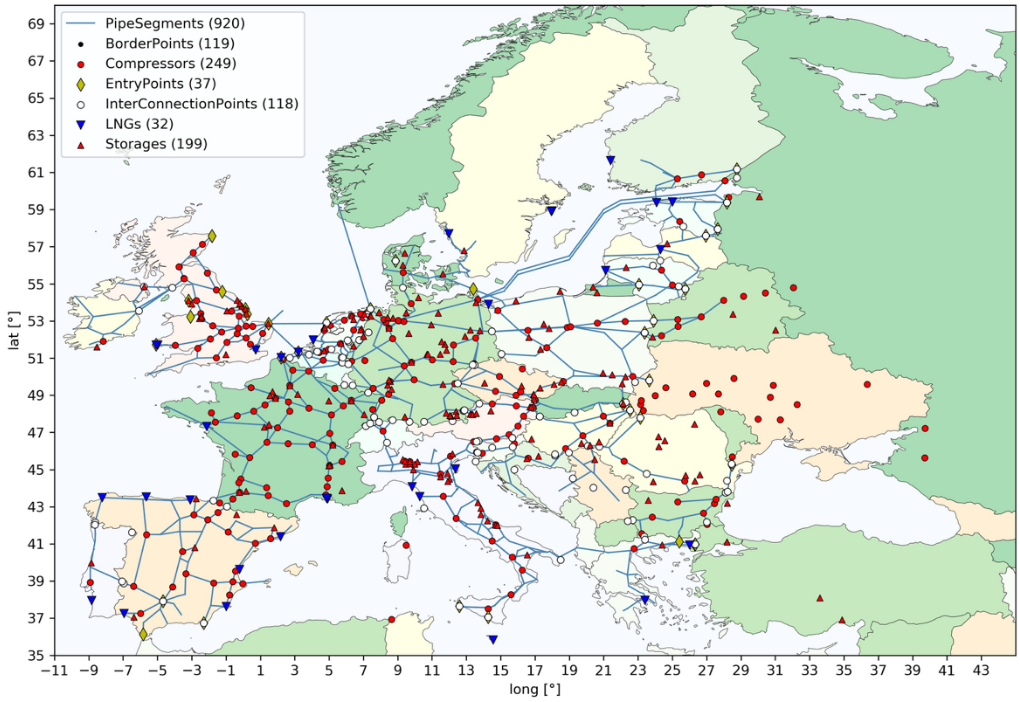


Figure 3: Schematic representation of the SciGRID_gas dataset and its main network component; Source: Scigrid-Gas

Key classes of components in dataset included are:

- **PipeSegments (PS)**: Individual pipeline segments, each defined by a path (with GIS coordinates) between two nodes. These segments carry gas and have attributes like length and diameter (and optionally capacity/pressure if available). The full network includes around 230,000 km of pipelines, which is comparable to the known length of Europe's transmission network .
- **Nodes representing facilities:** Several types of node objects are included to model where pipelines start, end, or intersect:
- **BorderPoints (BP):** Cross-border interconnection points where pipelines cross country borders. These connect the national networks.
- **Compressors (CS):** Compressor stations that boost gas pressure to push gas through pipelines.
- **LNG terminals (LNG):** Both import (regasification) and any export (liquefaction) terminals in the region.
- **Storages (ST):** Underground gas storage facilities (e.g. depleted fields, salt caverns).
- **Productions (PO/PD):** Major gas production sites or entry points from production into the transmission network. This can include onshore fields or entry from offshore pipelines.
- **PowerPlants (PP):** Gas-fired power plants connected to the network (treated as offtake nodes).

- **Consumers (CO):** Aggregated demand nodes representing gas consumers in an area. In SciGRID_gas, artificial consumer nodes are created to represent demand in each region, typically tied to administrative regions (NUTS zones).

The SciGRID_gas data model is built by merging multiple datasets from industry associations and TSOs. Sources include **Gas Infrastructure Europe (GIE)** (https://www.gie.eu/) for pipelines and storage, the **ENTSOG transmission network map** (https://transparency.entsog.eu/) for pipeline routes, the **International Gas Union (IGU)** (https://www.gie.eu/) for global stats, national sources like Great Britain's and Norway's TSO data, and others zenodo.org. These inputs were carefully processed: pipelines from different sources were combined, overlapping entries identified and merged, and missing attributes (e.g. pipeline diameters, capacities) were estimated using statistical and heuristic methods. The result is a comprehensive integrated network model.

Each node and pipeline segment in the dataset is annotated with key attributes—such as diameter, length, capacity (where available), or storage volume—that enable technical modeling of network operation. Pipeline routes are recorded as geospatial polylines, making the network suitable for both simulation and spatial analysis. A notable feature is the demand representation: SciGRID_gas includes daily sector-specific gas consumption time series (2010–2019) for NUTS3 regions across 27 countries. Each artificial consumer node is connected to the closest pipeline segment, enabling granular mapping of demand within the transmission system. The dataset is published in multiple regional aggregation levels (NUTS1, NUTS2, NUTS3) and provided in accessible formats (CSV, GeoJSON) under a CC-BY license. This structured, component-level architecture supports direct integration into the MGET model and has already been validated through its application in the Spanish case study. It also forms the technical foundation for the forthcoming NUTS2-level North Sea dataset being developed for iDesignRES.

3. Datasets available

3.1 NUTS0-1: GGM

The GGM dataset is an open-source aggregated gas market dataset at national (NUTS0) and major regional (NUTS1) resolution, developed by NTNU for analyzing the world natural gas market (ntnu.edu). It represents more than 90 countries as nodes practically the entire global gas production and consumption with links for cross-border pipelines, LNG liquefaction/regasification terminals, and underground storage facilities (all with associated capacity data). Large countries are subdivided into multiple production/consumption nodes to capture internal supply and demand regions (e.g. the USA has ten nodes, Russia three). The dataset provides detailed infrastructure capacities alongside regional supply and demand figures (base year values and future projections), forming the input for GGM's multi-period market equilibrium simulations. This allows scenario-based modeling of gas markets and analysis of large-scale trade flows and infrastructure utilization across regions. The GGM data (and model) are openly available through NTNU's Energy Models Hub, supporting studies on gas network development, market dynamics, and cross-border trade in energy system analyses. The GGM dataset is based on version 3.0 of the Global Gas Model (https://ideas.repec.org/p/diw/diwddc/dd100.html), and its 2023 update incorporates more recent data sources for key infrastructure components (https://doi.org/10.1016/j.esr.2025.101646). These include detailed information on LNG terminals, transmission pipelines, and gas storage facilities. Key data sources used in the update are:

- **GIIGNL Annual Report 2022**– for subcountry-level LNG liquefaction and regasification capacities.https://giignl.org/wp-content/uploads/2023/07/GIIGNL-2023-Annual-Report-July20.pdf

- **ENTSOG Transmission Capacity Map 2021** – for European cross-border pipeline capacities https://www.entsog.eu/maps#transmission-capacity-map-2021

- **Global Gas Infrastructure Tracker** (Global Energy Monitor, 2023) – for global infrastructure status and developments https://globalenergymonitor.org/projects/global-gas-infrastructure-tracker/

- **Gas Infrastructure Europe (GIE), Aggregated Gas Storage Inventory (2023)** – for updated gas storage data across Europe https://agsi.gie.eu

All GGM data and model files are openly available in the updated 2023 version on the GitHub repository linked here.

3.2 NUTS3: Scigrid gas

The SciGRID_gas dataset provides an open, high-resolution model of the European gas transmission system. Built primarily from OpenStreetMap and enriched with additional publicly available sources such as Gas Infrastructure Europe and national regulators, SciGRID_gas includes over 230,000 kilometers of pipeline data. It covers key infrastructure elements such as compressor stations, border entry and exit points, gas-fired power plants, storage facilities, LNG terminals, and production sites. Where necessary, missing technical or spatial attributes are estimated using statistical or engineering-based heuristics.

This dataset enables spatially explicit gas network modeling and has been used in research and policy analyses concerning gas flow optimization, sector coupling, and infrastructure resilience. Its open-source structure facilitates reproducibility and extensions by the wider energy modeling community.

4. Catalogue and Description of NUTS-2 Level Data for Hydrogen in Europe

At the NUTS-2 regional level, hydrogen potentials and related data varies significantly due to differences in renewable energy resources, industrial demand, and infrastructure. Regions with high wind and solar capacity, such as parts of the Nordics, Iberian Peninsula, and Western Europe, present the greatest potential for green hydrogen production, often exceeding 100 MWh/km² in supply density[1]. These regions are positioned as future net exporters, while others with less renewable potential will rely on imports via emerging hydrogen corridors[2]. These are important considerations to ensure the adaptions and extensions of GGM emphasize these hydrogen potentials and developments. Here some sources and inputs to consider in the adaptation of hydrogen data[3]:

[1] https://hydrogeneurope.eu/wp-content/uploads/2024/11/Clean_Hydrogen_Monitor_11-2023_DIGITAL.pdf

[2] https://hydrogeneurope.eu/wp-content/uploads/2024/11/Clean_Hydrogen_Monitor_11-2023_DIGITAL.pdf

[3] https://ehb.eu/files/downloads/1653999355_EHB-Supply-corridors-presentation-Full-compressed-1.pdf

Data on Production: By 2023, Europe's hydrogen production capacity was concentrated in Germany, the Netherlands, Poland, France, and Italy, collectively accounting for over half of total capacity[3]. Utilization rates vary: Portugal and Finland achieved over 90%, while some Eastern European regions fell below 50%. The pipeline for clean hydrogen projects is rapidly expanding, with nearly 25 Mt/year of announced capacity by 2040, though most projects remain in early development stages[4].

Use: Hydrogen's primary end-uses at the regional level are in ammonia, steel, refining, methanol, and e-fuels production, with nearly half of future capacity not yet assigned to a specific sector. Industrial clusters in NUTS-2 regions drive localized demand, particularly in heavy industry and chemical sectors.

Prospects: Europe aims for 10 Mt of domestic hydrogen production and 10 Mt of imports by 2030. Regional disparities in supply and demand underscore the importance of cross-border hydrogen corridors and storage infrastructure, such as salt caverns and repurposed pipelines, to balance the system and enable emissions reductions of up to 25% by 2050[5]

5. Next datasets and ongoing work

Based upon the presented dataset, more spatially detailed datasets are to be set up, spanning up until a NUTS3-level.

The shift from aggregated country-level gas modeling (NUTS0) to detailed subnational network representation (NUTS2/NUTS3) marks an important step forward in energy system analysis. While the GGM dataset provides a global overview suitable for market-based simulations and long-term infrastructure planning, its coarse spatial resolution limits the ability to assess intra-country flows, regional bottlenecks, and localized supply and planning strategies. In contrast, SciGRID_gas introduces a highly detailed, open-access model of Europe's transmission system that includes thousands of components and regional demand nodes, enabling fine-grained flow analysis, contingency studies, and coupling with other sectoral models (such as electricity).

To support the iDesignRES platform—specifically the MGET module the gas model—our approach builds on these two foundations. The MGET model requires spatially explicit gas infrastructure data to realistically simulate flow-based interactions, investment needs, and resilience across interconnected regions. This has already been successfully demonstrated in the Spanish case study, where a NUTS2/NUTS3-level network was created to capture both interprovincial dynamics and cross-border infrastructure links.

Building on that experience, the forthcoming North Sea NUTS2 gas network dataset will extend this approach to one of Europe's most critical gas corridors. It will feature regionalized representations of offshore supply pipelines, coastal entry terminals, LNG facilities, and demand zones across countries including Norway, the UK, the Netherlands, Germany, Belgium, and Denmark. The dataset will be compatible with MGET's architecture and delivered in standardized, open formats, aligning with

---

[3]https://observatory.clean-hydrogen.europa.eu/sites/default/files/2024-11/The%20European%20hydrogen%20market%20landscape_November%202024.pdf

[4]https://hydrogeneurope.eu/wp-content/uploads/2024/11/Clean_Hydrogen_Monitor_11-2023_DIGITAL.pdf

[5]https://ehb.eu/files/downloads/1653999355_EHB-Supply-corridors-presentation-Full-compressed-1.pdf

SciGRID_gas conventions. Ultimately, the NUTS2 gas dataset for the North Sea will serve as a benchmark for further regional modeling within iDesignRES, supporting scalable, transparent, and interoperable energy planning across Europe.

iDesignRES

Integrated Design of the Components of the Energy System to Plan the Uptake of Renewable Energy Sources: An Open Source Toolbox

More information on iDesignRES project:

https://www.linkedin.com/company/idesignres

https://idesignres.eu/

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.