METIS 2 - Technical Note 4

Distribution grid modelling: Data collection, asset disaggregation and methodology to fill gaps

METIS 2 Technical notes
February 2022

Prepared by

Carlos Mateo (Comillas)
Fernando Postigo (Comillas)
Rafael Cossent (Comillas)
Tomás Gómez (Comillas)

Contact: metis.studies@artelys.com

EUROPEAN COMMISSION

Directorate-General for Energy

Directorate A – Energy policy: Strategy and Coordination
Unit A4 – Economic Analysis and Foresight, Recovery

Directorate B – Just Transition, Consumers, Energy Efficiency and Innovation
Unit B5 – Innovation, Research, Digitalisation, Competitiveness

E-mail: ENER-METIS@ec.europa.eu

European Commission
B-1049 Brussels

Content

1. INTRODUCTION

The METIS 2 project involves new developments in the METIS model, that will enable to analyze the interaction of the European Electricity Market with the transmission and distribution networks. A set of archetypes are going to be defined to represent the distribution networks of every country in the METIS distribution network module. These archetypes will be connected to the METIS market model, so that the impact of specific market dispatch solutions on the distribution network archetypes may be evaluated. Every country will have a subset of archetypes. A set of parameters describing them has to be collected in order to model adequately the main characteristics of the distribution networks in each country.

This document summarizes the data collection process, and the methodology to fill the data gaps required to model the distribution networks in the METIS 2 project. It has for purpose to explain the gathering of real-world network data and their processing to describe the representative network archetypes. In Section 2, the data collection is introduced, covering, among others, voltage levels, density of consumers, number of substations, etc. Apart from this data, the consumption and generation profiles, available at country level in the market model, have to be disaggregated for each distribution voltage level, in order to be able to properly analyze the impact on distribution networks. To address this issue, Section 3 describes how demand and generation profiles are modelled in the METIS platform, and propose a methodology to break down the demand and generation profiles per voltage level. It is convenient to have several networks within a country in order to avoid an all-or-nothing effect in the subsequent analyses that will be carried out using the METIS distribution grid module. For this purpose, the division of the countries into several archetypes and zones is addressed in Section 4. In section 5 the methodology to disaggregate demand and generation assets is illustrated using the case of Spain as an example. In a second stage, the data collection process has been improved and complemented with data from JRC DSO Observatory [1, 2], which is currently the most updated source of information from the Distribution System Operators in the European Union. In particular, Section 6 specifies the data collected from JRC.

2. DATA COLLECTION

Required input data to build the archetypes was derived from the technical specifications of the METIS distribution network module. This section briefly summarizes the parameters that have been gathered. It is structured into the following categories:

- General parameters: This section includes general parameters like the voltage levels.
- Decision variables: These parameters reflect topology characteristics of the networks. Some of these variables may be adapted for the purpose of model calibration in order to ensure that the archetypes obtained are representative of the characteristics of the networks in each country.
- Network: These parameters include physical characteristics and costs of the power lines.
- Transformers: These parameters include physical characteristics and investment costs of the transformers.
- Other parameters: These parameters include other data, such as operation costs and equipment life duration.

This section is a summary of the information and the sources of the data collection process, being the full data available in the attached Excel spreadsheet "METIS 2_TN 4_ Distribution grid data collection.xlsx". The following references have been used in the data collection process [1–26].

2.1 General parameters

The general parameters include the voltage levels. The voltage levels are available for most of the countries by Eurelectric [19]. The considered voltage levels per country are included in Annex A. These voltage levels will be represented as nodes in the archetypes, with a corresponding nominal voltage. The division in three voltage levels (low (LV), medium (MV) and high (HV)) is considered in the following sections. In general, data is broken down per voltage level, to be able to model their differential characteristics. Some data is also broken down in urban, semi-urban and rural, to be able to build several archetypes within a country, in order to provide more diversity to the results obtained in each of them.

2.2 Decision variables

The decision variables include, among others, densities of consumers, substations and network length. The following variables have been collected to obtain the densities:

- Surface of the countries
- Length between LV, MV & HV consumers (calculated)
- Number of LV, MV & HV consumers
- Number of MV/LV, TSO substations
- Number of HV/MV substations
- Length of LV, MV & HV network (broken down in underground and overhead)
  This information is disaggregated per country. The main source for this information is Eurelectric [19]. A few countries only provide aggregated values (e.g. total number of consumers). In this case, we inferred the disaggregation from countries with similar characteristics. The number of HV/MV substations is missing in Eurelectric report, so this information is obtained from the JRC DSO Observatory database [1, 2]. The following decision variables have further been collected, where the nominal power is disaggregated per country, and a cable section has been selected.
- Cable sections and nominal power of LV, MV & HV cables
  These decision variables have also been disaggregated per type of area (urban, semi-urban and rural). In this case, the following parameters have been collected.
- Surface of the distribution areas
- Length between LV & MV consumers
- Number of LV & MV consumers
- Number of MV/LV, HV/MV substations
- Length of LV & MV network (broken down in underground and overhead)
  The main source to obtain these parameters disaggregated per type of area have been the representative networks of the DSO Observatory[2]. The previous values for the countries will be used to identify country averages. The information disaggregated per type of area will be applied as weights in order to have a further detail (urban, semi-urban, rural) in every country.

2.3 Network

The following parameters about network physical characteristics have been collected.

- Power factor (defined as the ratio of active power to apparent power)

- LV, MV & HV conductor resistance
- LV, MV & HV conductor reactance
- LV, MV & HV conductor ampacity
- LV, MV & HV conductor nominal voltages
- LV, MV & HV admissible voltage up and down
- LV, MV & HV underground ratios
  The main sources have been publications from transmission system operators, manufacturers, and CEER voltage limits [4, 5, 22, 24–26]. A catalogue of equipment parameters is initially disaggregated in overhead and underground power lines, because this categorization implies using differential constructive components, having different technical characteristics and requiring different costs. The underground ratio of every country is collected, disaggregated in low, medium, and high voltage. This enables to particularize the parameters at country level and per voltage level, by considering the proportion of underground and overhead components, and their differential characteristics. Parameters that relate section and power in cables have also been collected.
- LV, MV & HV Cables, Multiplicative and Exponential Factors, Section vs Power
  The relation between section and power in cables, follows an exponential function as shown by the following equation.

$$p_c = c_{mps} e^{c_{eps} s_c} \qquad (1)$$

Where p_c is the power of the cable, s_c is the section, c_{mps} is the multiplicative factor that relates the section with the power and c_{eps} is the exponential factor that relates the section with the power.

In order to obtain the required multiplicative and exponential factors, first data about a set of equipment has been collected from manufacturer catalogues. These components are then modelled using specific line tendencies. For example, for this specific parameter, the tendency is exponential, requiring a multiplicative and an exponential factor. In order to obtain them, all the collected data components are plotted and a curve regression (with the required shape) is obtained. The coefficients of the curve correspond to the multiplicative and exponential factors needed to build the archetypes. Figure 1 illustrates the power vs section function for MV overhead power lines. Graphs have been obtained for LV, MV & HV, in overhead and underground equipment, to obtain all the corresponding parameters.

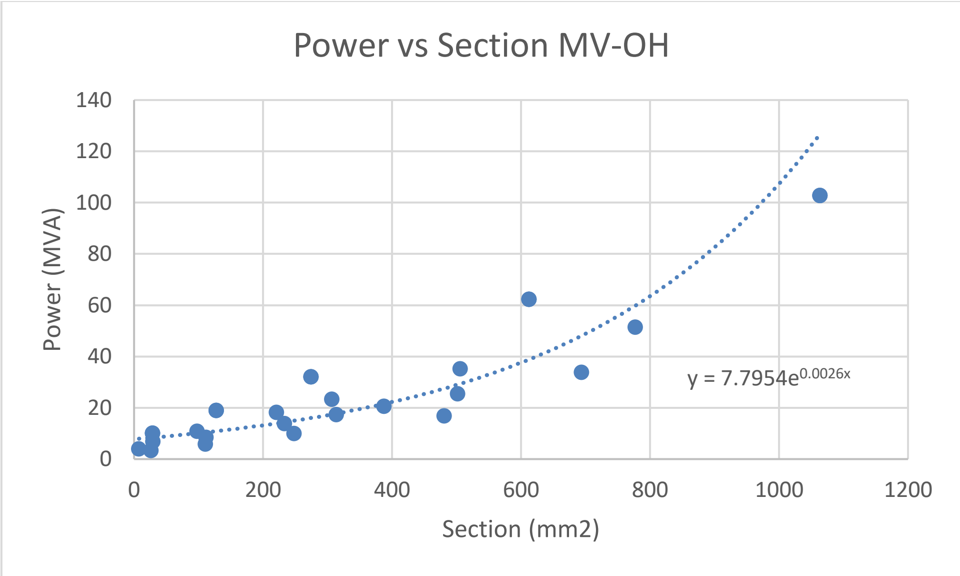

Figure 1 MV power vs section: set of overhead equipment and regression

Again, the parameters have been obtained separately for underground and overhead equipment, which has later allowed us to differentiate per country by taking into account the proportion of underground and overhead installations in each country.

- LV, MV & HV Fixed and Proportional coefficient for cable or power line cost
The relation between section and power in cables, follows a linear function as shown by the following equation.

$$i_c = c_{pis} s_c + c_{fis} \qquad (2)$$

Where i_c is the investment of the cable, s_c is the section of the cable, c_{pis} is the proportional factor that relates the section with the investment and c_{fis} is the fixed factor that relates the section with the power. Figure 2 illustrates the section vs investment cost function for MV overhead power lines. Graphs have been obtained for LV, MV & HV, in overhead and underground equipment, to obtain all the corresponding parameters.

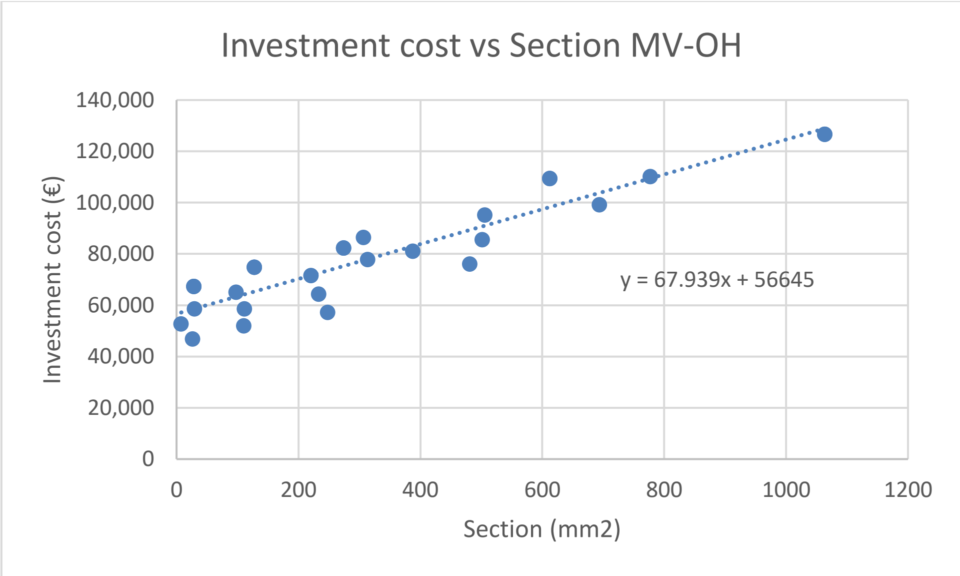

Figure 2 MV cable or power line cost: set of overhead equipment and regression

The main source of the costs have been public unitary reference values from a regulator in Europe [16]. In this case, the fixed cost corresponds to the intersection with the vertical axis, while the proportional coefficient is the slope of the line. As explained for other catalogue components, the parameters are obtained separately for underground and overhead equipment, and per voltage level, because this type of disaggregation has implications in terms of technical characteristics and costs. This enables to finally obtain parameters differentiated per country, by taking into account the proportion of overhead and underground installations in each country and voltage level.

2.4 Transformers

For the HV/MV and the MV/LV transformers the following data has been collected:

- Nominal voltages
- Capacity
- No load losses
- Equivalent resistivity (for copper loss calculations)
- Fixed and proportional factor for iron loss calculations
- Transformer fixed and proportional coefficient used in the cost power relationship
- Cell costs

The relation between iron losses and power in transformers, follows a function as shown by the following equation.

$$l_t = c_{plp} p_t + c_{flp} \qquad (3)$$

Where l_t reflects the iron losses of the transformer, c_{plp} is the proportional factor that relates the power with the iron losses, c_{flp} is the fixed factor that relates the power of the transformer with the iron losses and p_t is the power of the transformer.

The relation between investment cost and power in transformers, follows a function as shown by the following equation.

$$i_t = c_{pip} p_t + c_{fip} \qquad (4)$$

Where i_t is the investment cost of the transformer, p_t is the power of the transformer, c_{pip} is the proportial factor that relates the power with the investment cost of the transformer and c_{fip} is the fixed factor that relates the power of the transformer with investment cost.

The main source has been information from catalogues and manufacturers [5, 21]. The transformer size is available in HV/MV substations and in MV/LV transformers. In the MV/LV transformers, it is broken down in urban, semi-urban and rural. The source for the transformers is public data of the DSO Observatory [2]. For the costs, public unitary costs from European regulators have been applied [16]. The cell costs depend on the type of cell (underground or overhead), enabling to differentiate per country by taking into account the proportion of underground and overhead installations of each country.

The transformer fixed and proportional factors for iron loss calculations have been obtained modelling a set of transformers, obtaining exponential regression curves and identifying their coefficients.

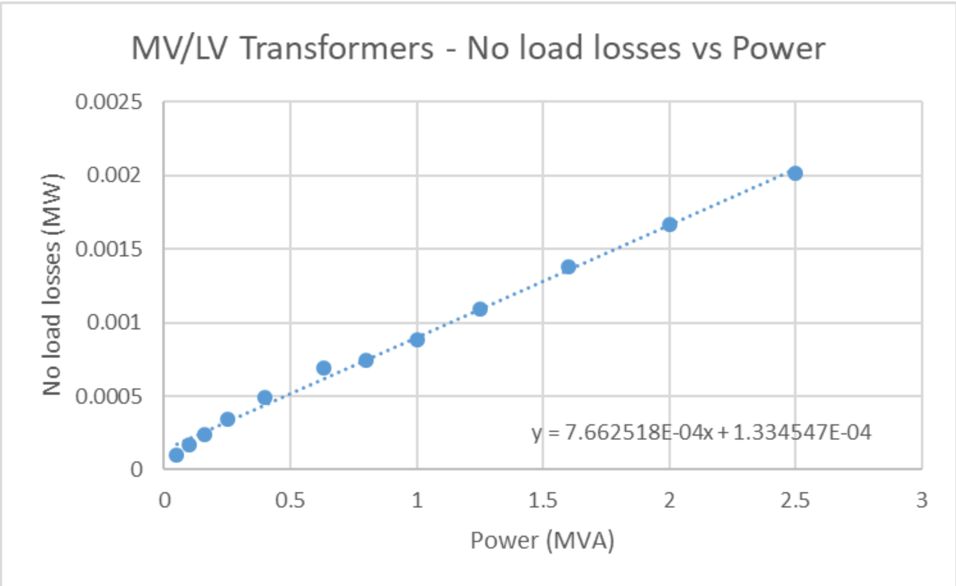

Figure 3 MV/LV transformers. No load losses versus capacity of the transformers. Set of transformers and regression.

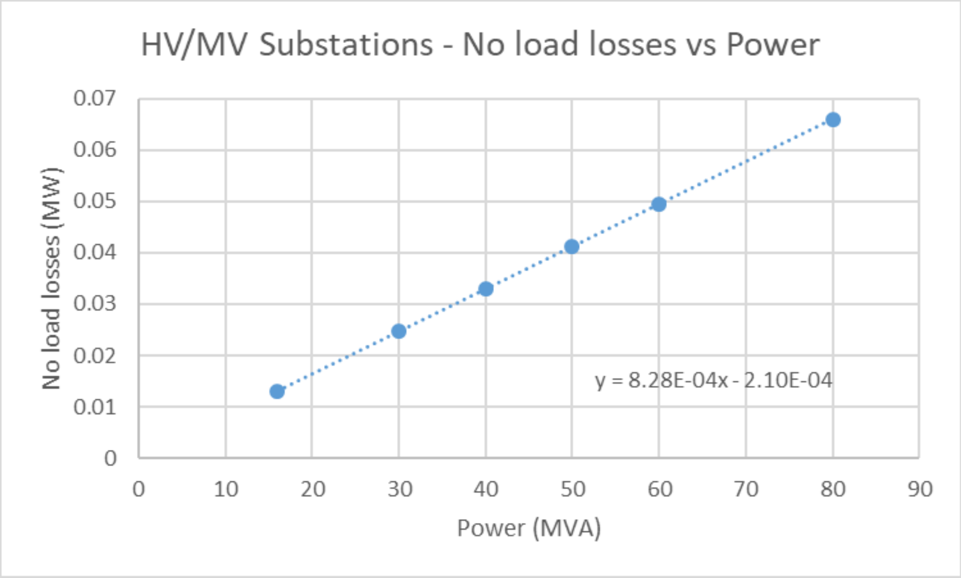

Figure 4 HV/MV transformers. No load losses versus capacity of the transformers. Set of transformers and regression.

For the transformer costs, a set of regression curves have been obtained modelling transformers of different capacities. In this case, the selected representative voltage levels have been taken into account because the costs-power relation depends on the nominal voltage. Figure 5 illustrates the power vs investment relation for 20/0.4kV transformers. The corresponding graphs have also been obtained for HV/MV and for MV/LV transformers of other voltage levels.

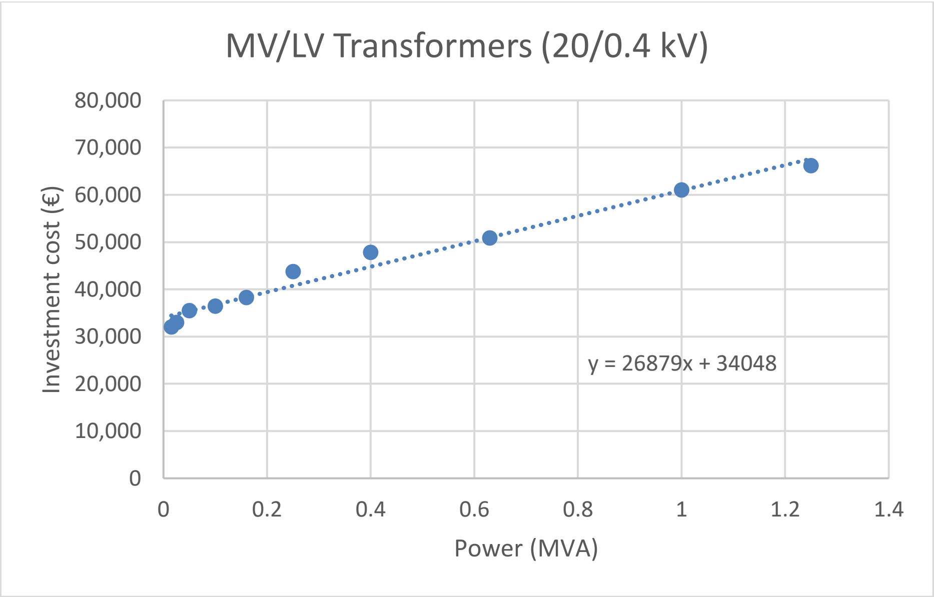

Figure 5 MV/LV transformer costs (20/0.4kV): set of transformers and regression.

2.5 Other parameters

Other parameters involve the return rate, the cost of energy losses and the equipment life duration. The return rate of each country has been calculated following the methodology described by CEER [15], and applied by CNMC to calculate the WACC in Spain [14]. The cost of energy losses is estimated using average prices per country. The equipment life duration includes a range of years for each type of equipment and a median value. The equipment life duration is based on information from IEC [12].

3. PROFILES

The METIS distribution grid module requires demand and generation profiles. Section 3.1 describes hourly profiles obtained from the METIS platform, Section 3.2 describes how to process the demand profiles to disaggregate them per voltage level and assets, and Section 3.3 describes the process to disaggregate generation assets per voltage level.

3.1. Hourly profiles in the METIS platform

This subsection describes the data used from the METIS platform about hourly profiles. In particular, the METIS platform contains national hourly profiles for the following categories:

- Flexible demand
  - Heat pumps
  - Domestic hot water
- Non-Flexible demand
  - Air conditioning
  - Thermosensitive remainder
  - Non-thermosensitive remainder
- Generation
  - Wind Onshore availability
  - Solar availability
  - Hydro RoR availability
  - Biomass availability
  - Waste availability
- Electric vehicles
  - PHEV home charge
  - PHEV work charge
  - BEV home charge
  - BEV work charge
- Storage
  - Batteries

A full year of data has been extracted from the METIS platform about the profiles of these assets, featuring national granularity. This data corresponds to simulations in a scenario in year 2030[1], and covers all countries in Europe. For modelling the distribution networks, the disaggregation approach to the level of network archetypes will be top-down, starting with the national profiles, and obtaining their disaggregation per voltage level. Therefore, the methodology has to take into account the available data. The methodology has to be general, so that it can be applied to other scenarios (e.g. year 2020).

An example of a cumulated demand profile in two consecutive days is shown in Figure 6. The demand profile has two peaks during the day, with significant differences among two consecutive days in this particular case.

[1] We refer here to the METIS EUCO3232.5 scenario, which builds upon the data from the respective scenario of the European Commission.

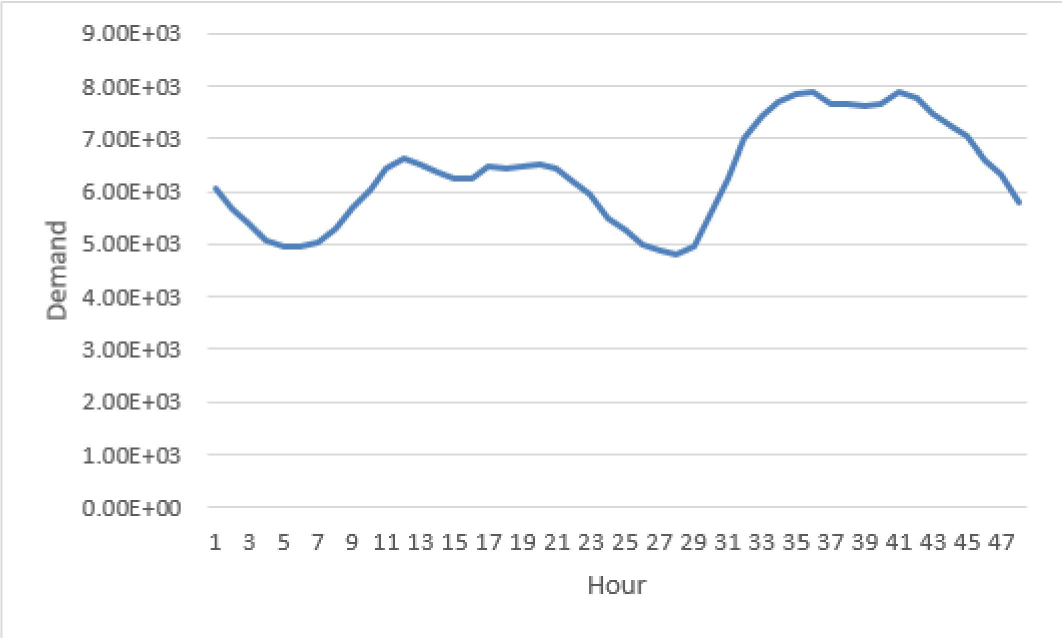

Figure 6 Example of demand profile in the METIS platform (two consecutive days in a country).

Example of photovoltaic and wind profiles are shown Figure 7 and Figure 8, showing also two consecutive days. While in the demand and photovoltaic profiles there is a periodic behavior (despite also recognizing some differences among days), in the wind profile the dynamics do not necessarily have a correspondence with natural days.

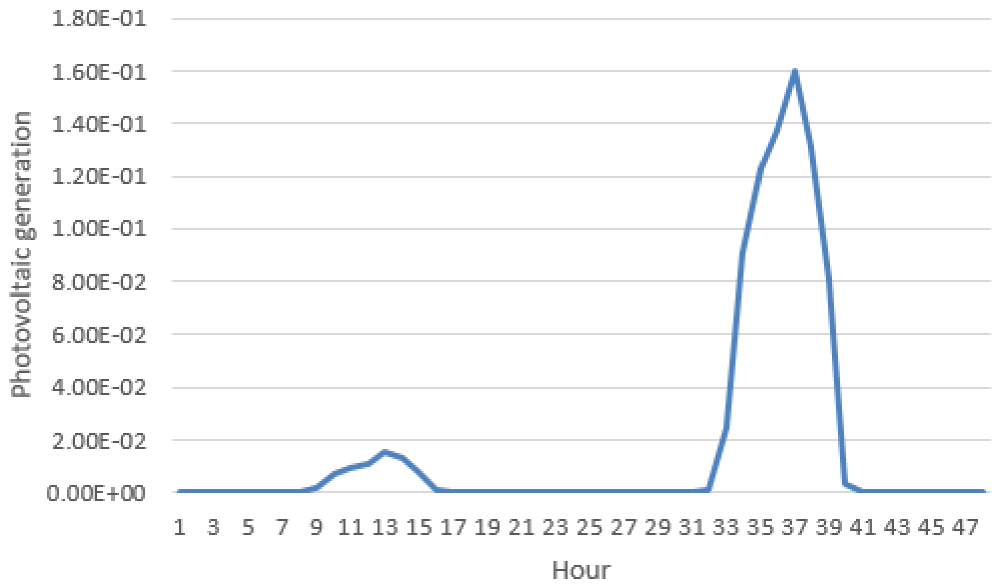

Figure 7 Example of photovoltaic profile in the METIS platform (two consecutive days in a country).

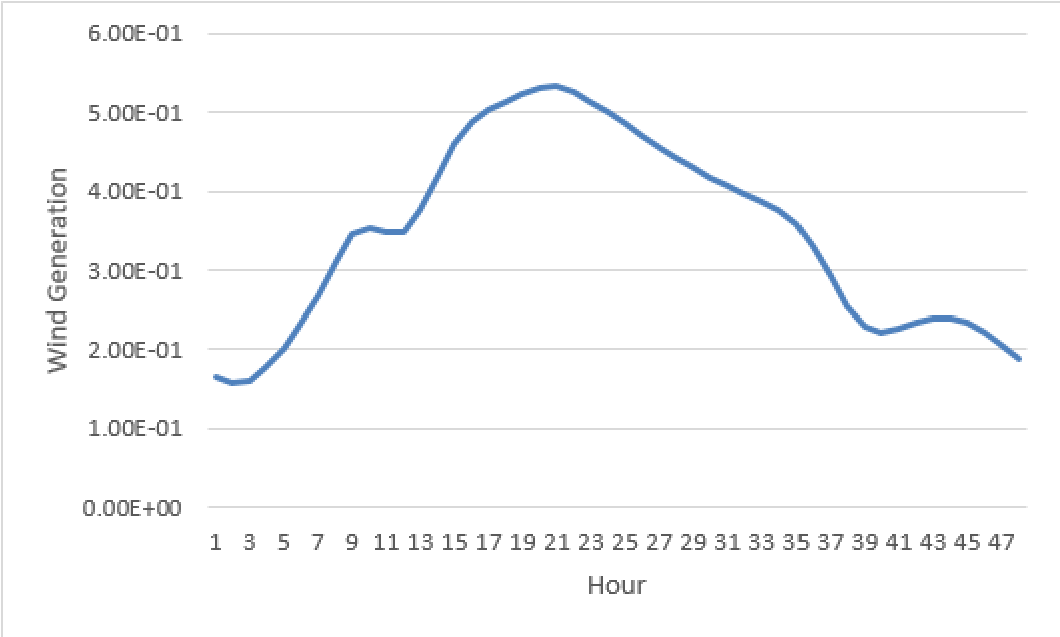

Figure 8 Example of wind profile in the METIS platform (two consecutive days in a country).

The methodology proposed is the basis for breaking down the profiles per voltage level, and into archetypes and climatic zones. The methodology is top-down, because the profiles are already estimated at national level in the METIS platform. The disaggregation per voltage level enables to model scenarios in the distribution networks which are consistent with the scenarios at the national market level. In particular, Section 3.2 shows how to break down the demand profiles per voltage level. In cases in which there is no relevant information about the zones to be used for the profiles, a percentage matrix defined in Section 4 can be directly used to split the profiles among archetypes and zones. In cases like photovoltaics, in which there can be relevant information (like the irradiation levels) in each zone, that information can be used to drive the division into smaller areas.

3.2. Methodology to break down the demand profiles per voltage level

This section introduces a methodology to disaggregate the demand profiles connected to low, medium and high voltage levels. The objective is to be able to develop scenarios for the distribution networks, which are consistent with the national demand profiles in the market model. The methodology takes as input the demand profile of a country of reference [20]. The variability of the hourly information along the year per voltage level is combined and extrapolated, to be applied as percentage of the total demand of other countries taking into account how low, medium and high voltage demand differ depending on the particular weights of residential, commercial and industrial demand in each country, as explained below [11].

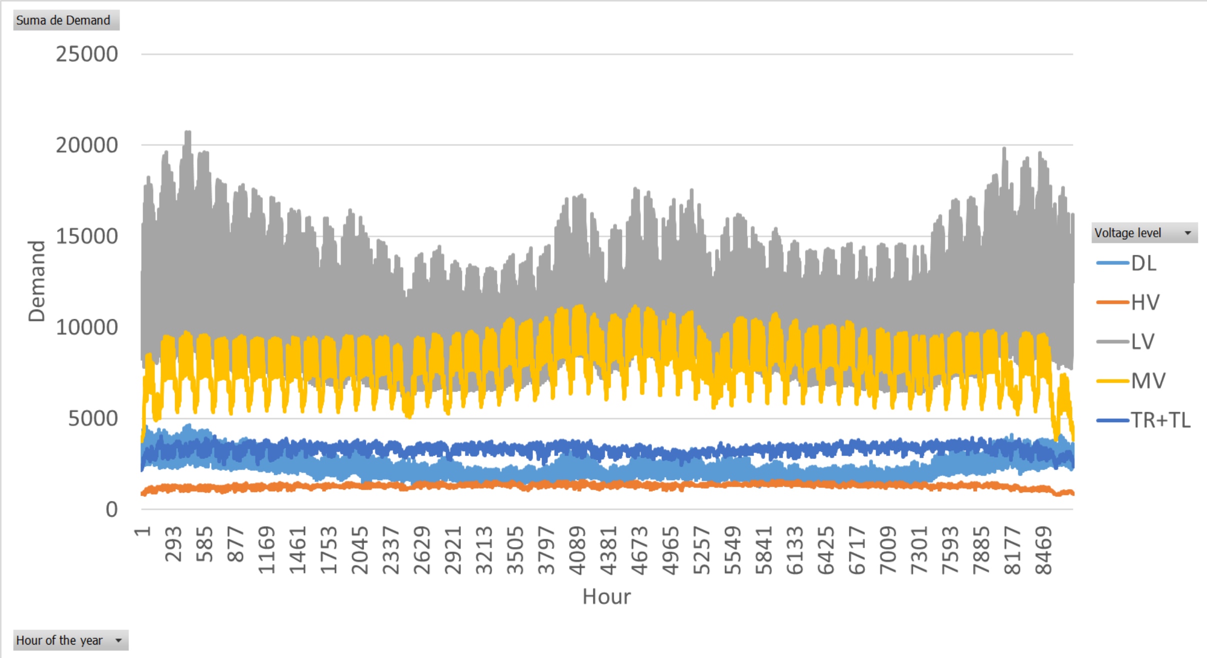

Figure 9 Historical values of demand in low (LV) , medium (MV) , high voltage (HV), distribution losses (DL) and transmission demand and losses (TR+TL)

The historical demand broken down per voltage level is available in a specific country that we will denominate the selected country (SC) [27]. In order to obtain percentage values applicable to different years, these historical values are categorized per hour of the day, day of the week and season. The information of the full year is reformulated with this structure to facilitate applying it to other years. This transformation is carried out because, for demand modeling, it is more relevant the day of the week and the season than the natural day within the year. This processed information is stored into the following variable.

- $PC_{VL,h,d,s}^{SC}$: Percentage of consumption of the voltage level VL in the selected country SC, at hour h, day of the week d and season s.

Figure 10 illustrates how the historical demand in Figure 9 is broken down per season, day of the week and hour of the day to be able to apply it to other years.

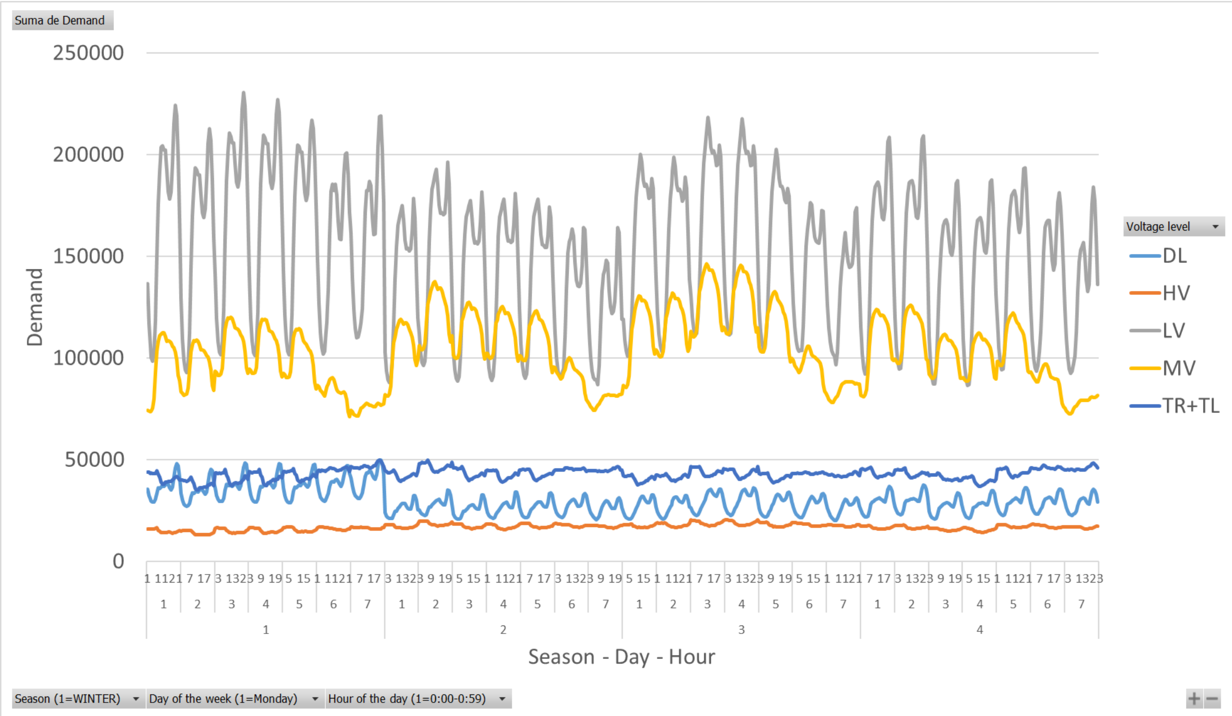

Figure 10 Historical demand in low (LV) , medium (MV) , high voltage (HV), distribution losses (DL) and transmission demand and losses (TR+TL), broken down per season, day of the week and hour of the day.

The percentages of consumption of the residential, commercial, industrial and other sectors are made available for all the European countries by the IEA [13]. This information is included in Annex B.

- PC_{RES}^{C}: Percentage of consumption of the residential sector in country C.
- PC_{COM}^{C}: Percentage of consumption of the commercial sector in country C.
- PC_{IND}^{C}: Percentage of consumption of the industrial sector in country C.
- PC_{OTH}^{C}: Percentage of consumption of others sector in country C.

We can refer generically to any of the sectors with the following notation.

- PC_{SEC}^{C}: Percentage of consumption of the sector Sec in country C.

In order to obtain the percentage that is connected to each voltage level, a table with a structure like the one shown in Table 1 has to be applied.

Table 1 Percentage of each sector that is connected to each voltage level

	Residential	Commercial	Industrial	Others
LV	100	50	0	30
MV	0	50	45	60
HV	0	0	15	5
EHV	0	0	40	5

This table represents a matrix $M_{VL,SEC}$, which indicates for each sector, the assumptions on the percentages that are connected to each voltage level, and calibrated as explained below. It should be noted that the sum of each column is 100%.

The percentage that is connected to each voltage level and sector ($PC^{C}_{VL,SEC}$), is obtained evaluating the product of the previous variables.

$$PC^{C}_{VL,SEC} = M_{VL,SEC} \times PC^{C}_{SEC}$$

The percentage that is connected to each voltage level (PC^{C}_{VL}), is obtained by adding the demand of the different sectors corresponding to that voltage level. We are assuming that residential, commercial and industrial loads are typically connected to certain voltage levels, and depending on the amount of each of them in a country, the demand per voltage level will vary accordingly.

$$PC^{C}_{VL} = \sum_{SEC} PC^{C}_{VL,SEC}$$

The only exception are distribution losses, which are not scaled depending on the residential, commercial and industrial sectors, but scaled proportionally to the low voltage demand, using a correlation with the low voltage demand.

Values in Table 1 have been calibrated so that variable PC^{SC}_{VL} in the selected country SC matches the available historical records of the demand. Once the percentage that is connected to each voltage level (PC^{C}_{VL}) for each country (C) is known, the historical record of demand ($PPC^{C}_{VL,h,d,s}$) is particularized for each country (C), applying the calculated percentages for each voltage level (PC^{C}_{VL}), using the following formula.

$$PPC^{C}_{VL,h,d,s} = PC^{SC}_{VL,h,d,s} \times \frac{PC^{C}_{VL}}{PC^{SC}_{VL}}$$

The result is normalized (PC^{C}_{VL}) to ensure that the sum of the percentages is exactly 100%, this is, that the sum of the disaggregated profiles match the total demand.

$$PC^{C}_{VL,h,d,s} = \frac{PPC^{C}_{VL,h,d,s}}{\sum_{VL} PPC^{C}_{VL,h,d,s}}$$

As explained above the main sources are :

- Annual hourly profile in a selected country [27].
- Percentages of residential, commercial and industrial consumption in each country.
- Assumptions on the amount of load of each sector (residential, commercial and industrial) that is connected to each voltage level.

Using this input data and the above formulas, the annual hourly profiles in the selected country per voltage level are combined into percentage profiles of demand connected to each voltage level in every country, which can be applied to the total hourly consumption in each country in order to disaggregate it.

The result is shown for example in Austria in Figure 11. One of the advantages of this methodology is that the disaggregation is obtained as a percentage, which enables to apply it to different countries with different national demand curves, or even to different years. The disaggregation into different voltage levels will be constant through time in relative terms, but different demand levels may be considered. In this case, the national demand profile of each country and year may be different, and each country will have specific coefficients for the disaggregation into LV, MV, HV & EHV depending on its characteristics.

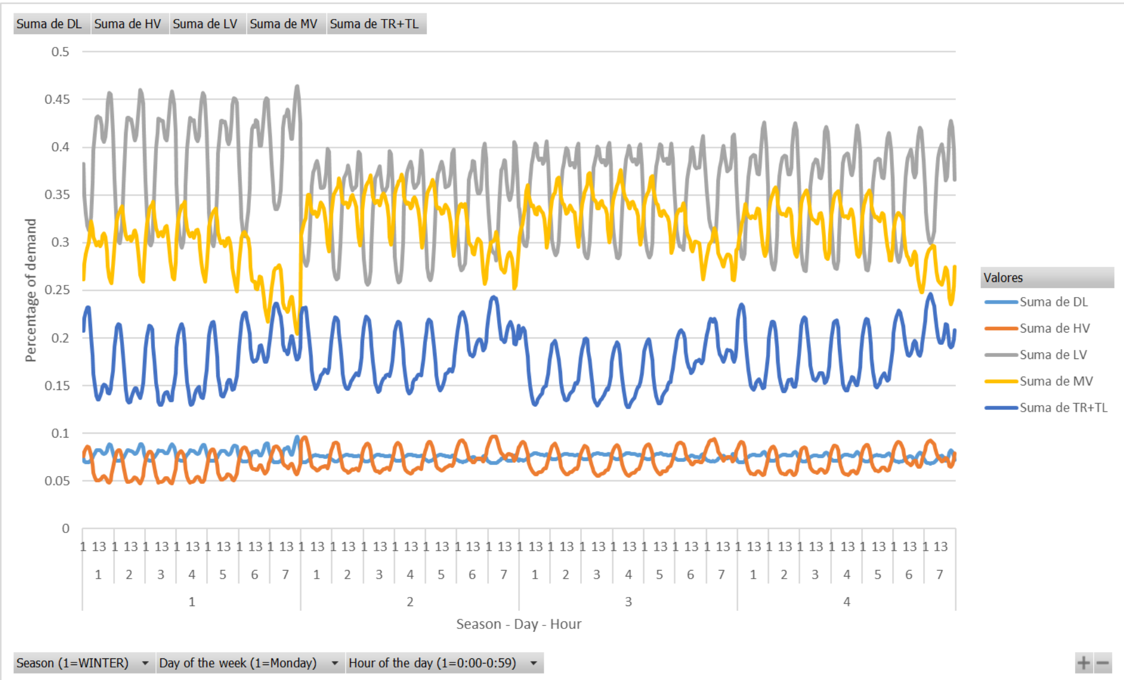

Figure 11 Disaggregation in in low (LV) , medium (MV) , high voltage (HV), distribution losses (DL) and transmission demand and losses (TR+TL), broken down per season, day of the week and hour of the day, in Austria.

3.2.1. Hourly profiles Methodology to break down the Flexible and Non-flexible demand assets per voltage level

The aim of this sub-section is to define the methodology used to allocate flexible and non-flexible demand assets to the different voltage levels, and specify the associated distribution network losses. This methodology is hourly based, so it can be applied to any time-series profiles. As a starting point, a preliminary voltage-level identification is made for each asset based on expert knowledge, as shown in Table 2. A cross in the table indicates in which levels each asset can be found. Thus, heat pumps can only be found at low voltage.

Table 2 Voltage level (LV, MV, HV, and EHV) to which each asset is connected

Assets	Aggregated assets	LV	MV	HV	EHV
Flexible demand	Heat pumps	X			
	Domestic hot water	X			
Non-flexible demand	Air conditioning	X	X		
	Thermosensitive remainder	X	X		
	Non-thermosensitive remainder	X	X	X	X

As observed in Table 7, *Non-thermosensitive remainder* is the only asset that exists in every voltage level, *Air conditioning* and *Thermosensitive remainder* are connected to LV and MV, and the flexible demand (*Heat pumps* and *Domestic hot water*) to LV only. Following the previous assumptions, an hourly disaggregation of the demand for each asset and voltage level ($D_{VL}^{Asset}(h)$) is made:

- Heat Pumps (HP) disaggregation:

$$D_{Total}^{HP}(h) = D_{LV}^{HP}(h) + D_{Losses}^{HP}(h)$$

- Domestic Hot Water (DHW) disaggregation:

$$D_{Total}^{DHW}(h) = D_{LV}^{DHW}(h) + D_{Losses}^{DHW}(h)$$

- Air Conditioning (AC) disaggregation:

$$D_{Total}^{AC}(h) = D_{LV}^{AC}(h) + D_{MV}^{AC}(h) + D_{Losses}^{AC}(h)$$

- Thermosensitive Remainder (TR) disaggregation:

$$D_{Total}^{TR}(h) = D_{LV}^{TR}(h) + D_{MV}^{TR}(h) + D_{Losses}^{TR}(h)$$

- Non-Thermosensitive Remainder (NTR) disaggregation:

$$D_{Total}^{NTR}(h) = D_{LV}^{NTR}(h) + D_{MV}^{NTR}(h) + D_{HV}^{NTR}(h) + D_{EHV}^{NTR}(h) + D_{Losses}^{NTR}(h)$$

The total demand profile ($D_{Total}(h)$) is obtained by adding the total demand of the different assets.

$$D_{Total}(h) = \sum_{Assets} D_{Total}^{Asset}\,(h)$$

Next, a voltage level based profile break down is performed for the total demand profile ($D_{Total}(h)$). This disaggregation follows the methodology explained in Section 4. Thus, $D_{LV}(h), D_{MV}(h), D_{HV}(h), D_{EHV}(h)\ and\ D_{Losses}(h)$ can be obtained from $D_{Total}(h)$.

$$D_{Total}(h) = D_{LV}(h) + D_{MV}(h) + D_{HV}(h) + D_{EHV}(h) + D_{Losses}(h)$$

The percentage demand $PC_{VL,h,d,s}^{C}$per voltage level (VL) and period (h,d,s) for each country (C) from chapter 4, is used here as input. In this chapter this is named as $D_{LV}(h)$, $D_{MV}(h)$ and $D_{HV}(h)$, $D_{EHV}(h)$ and $D_{Losses}(h)$ referring to the percentage demand in each voltage level (LV, MV and HV and EHV/TR[2]) and to the distribution losses (Losses).

Once demand is disaggregated in LV, MV, HV, EHV and losses, an equation system is proposed to disaggregate the specific assets (HP, DHW, AC, TR and NTR).

- Voltage and losses disaggregation by asset:

$$D_{LV}(h) = D_{LV}^{HP}(h) + D_{LV}^{DHW}(h) + D_{LV}^{AC}(h) + D_{LV}^{TR}(h) + D_{LV}^{NTR}(h)$$

$$D_{MV}(h) = D_{MV}^{AC}(h) + D_{MV}^{TR}(h) + D_{MV}^{NTR}(h)$$

$$D_{HV}(h) = D_{HV}^{NTR}(h)$$

$$D_{EHV}(h) = D_{EHV}^{NTR}(h)$$

$$D_{Losses}(h) = D_{Losses}^{HP}(h) + D_{Losses}^{DHW}(h) + D_{Losses}^{AC}(h) + D_{Losses}^{TR}(h) + D_{Losses}^{NTR}(h)$$

- Asset disaggregation by voltage level:

$$D_{Total}^{HP}(h) = D_{LV}^{HP}(h) + D_{Losses}^{HP}(h)$$

$$D_{Total}^{DHW}(h) = D_{LV}^{DHW}(h) + D_{Losses}^{DHW}(h)$$

$$D_{Total}^{AC}(h) = D_{LV}^{AC}(h) + D_{MV}^{AC}(h) + D_{Losses}^{AC}(h)$$

$$D_{Total}^{TR}(h) = D_{LV}^{TR}(h) + D_{MV}^{TR}(h) + D_{Losses}^{TR}(h)$$

$$D_{Total}^{NTR}(h) = D_{LV}^{NTR}(h) + D_{MV}^{NTR}(h) + D_{HV}^{NTR}(h) + D_{EHV}^{NTR}(h) + D_{Losses}^{NTR}(h)$$

$$D_{Losses}^{HP}(h) = \frac{D_{Losses}(h)}{D_{Total}(h)} * D_{Total}^{HP}(h)$$

$$D_{Losses}^{DHW}(h) = \frac{D_{Losses}(h)}{D_{Total}(h)} * D_{Total}^{DHW}(h)$$

[2] Extra high voltage (EHV) or Transmission (TR)

$$D_{Losses}^{AC}(h) = \frac{D_{Losses}(h)}{D_{Total}(h)} * D_{Total}^{AC}(h)$$
$$D_{Losses}^{TR}(h) = \frac{D_{Losses}(h)}{D_{Total}(h)} * D_{Total}^{TR}(h)$$

$$D_{Losses}^{NTR}(h) = \frac{D_{Losses}(h)}{D_{Total}(h)} * D_{Total}^{NTR}(h)$$

This system has two more unknowns than independent equations, so two additional equations need to be defined. In order to set two extra relationships, the Stratego report about quantifying the heating and cooling demand in Europe [28] is the base to obtain this information. Next, two additional equations are defined:

- **AC disaggregation factor for LV** ($Factor_{LV}^{AC}$): According to [28], AC only represents 2% of the total demand and 33% of the LV profile.

$$Factor_{LV}^{AC} = 33\%$$

$$D_{LV}^{AC}(h) = Factor_{LV}^{AC} * D_{Total}^{AC}(h)$$

- **TR and HP disaggregation factor for LV** ($Factor_{LV}^{TR+HP}$)**:** As [28] explains, the LV fraction that corresponds to the addition of TR and HP is strongly dependent on the country. The factor is shown for all the countries in Table 3.

Table 3 TR and HP disaggregation factor for LV

Zone	Country	$Factor_{LV}^{TR+HP}$
AT	Austria	6%
BE	Belgium	3%
BG	Bulgaria	14%
CY	Republic of Cyprus	27%
CZ	Czech Republic	8%
DE	Germany	7%
DK	Denmark	3%
EE	Estonia	3%
GR	Greece	7%
ES	Spain	18%
FI	Finland	17%
FR	France	13%
HR	Croatia	6%
HU	Hungary	4%
IE	Ireland	5%
IT	Italy	6%
LT	Lithuania	0%
LU	Luxembourg	5%
LV	Latvia	1%
MT	Malta	77%
NL	Netherlands	2%

PL	Poland	1%
PT	Portugal	19%
RO	Romania	1%
SE	Sweden	26%
SI	Slovenia	1%
SK	Slovakia	3%
UK	United Kingdom	9%
BA	Bosnia and Herzegovina	6%
CH	Switzerland	6%
ME	Montenegro	6%
MK	North Macedonia	6%
NO	Norway	6%
RS	Serbia and Kosovo	6%

This factor is included in the equation system through the following equation:

$$D_{LV}(h) * Factor_{LV}^{TR+HP} = D_{LV}^{HP}(h) + D_{LV}^{TR}(h)$$

Considering these additional factors, the equation system can be solved for every hour, obtaining the asset disaggregation by voltage level.

The current methodology allows obtaining unique solutions that comply with the proposed equations. However, in some occasions it is possible to find hours within the Non-thermosensitive remainder category with a negative value for MV assets. In these hours, it has been established that this demand is transferred from MV to LV, thus obtaining consistent results in which there is no negative demand.

In order to exemplify the results of this process, the second season of the year (spring) is shown in Figure 12 Asset disaggregation of the Spanish demand in LV, MV, HV, and EHV.

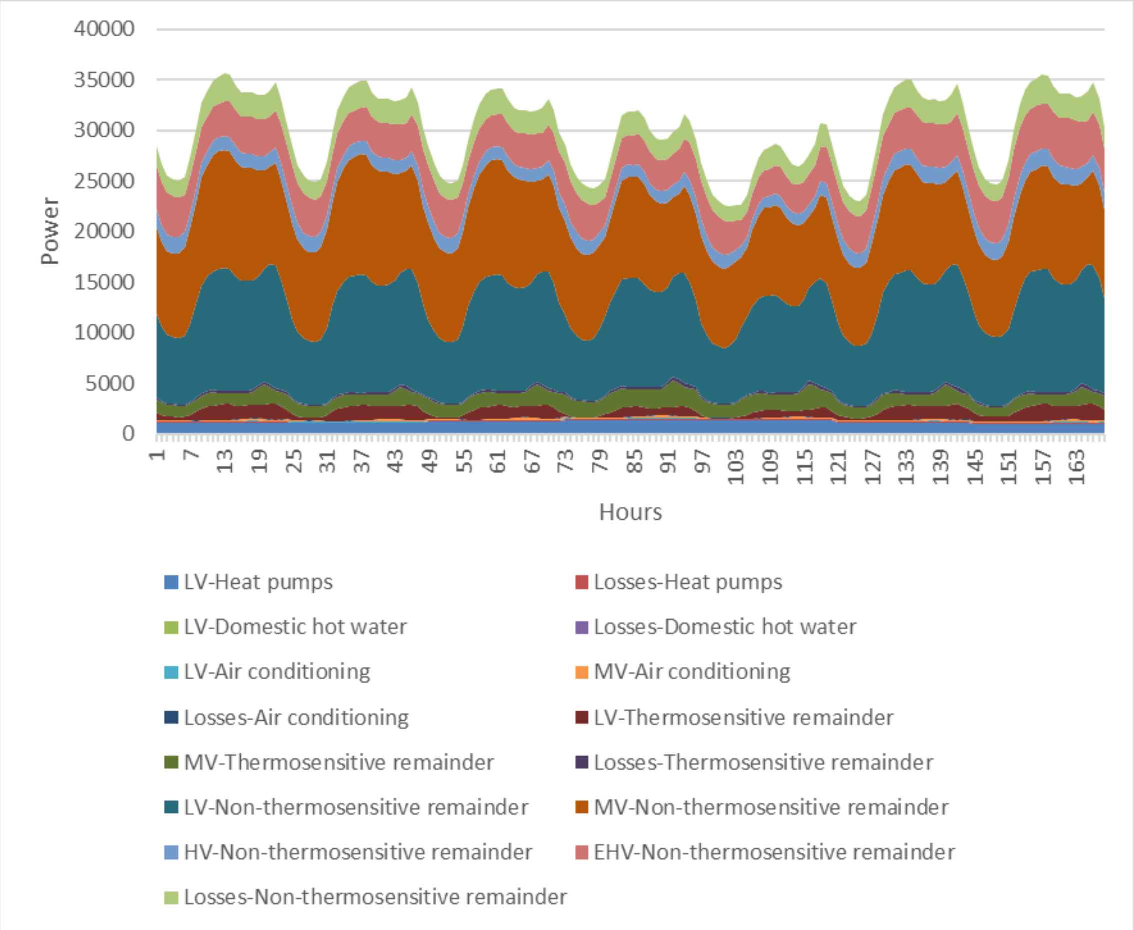

Figure 12 Asset disaggregation of the Spanish demand in LV, MV, HV, and EHV.

3.3. Methodology to break down the generation profiles per voltage level

The aim of this sub-section is to define the methodology used to break down the profiles of different generation assets in voltage levels. This methodology is hourly based, so it can be applied to any time-series profiles. The following assets are analyzed.

- Wind onshore fleet ($G^{W}(h)$)
- Solar fleet ($G^{S}(h)$)
- Hydro RoR fleet ($G^{H}(h)$)
- Biomass fleet ($G^{B}(h)$)
- Waste fleet ($G^{Wa}(h)$)

First, an hourly disaggregation of the generation profiles for each asset and voltage level ($G_{VL}^{Asset}(h)$) is carried out as the following equations show.

$$G_{Total}^{W}(h) = G_{LV}^{W}(h) + G_{MV}^{W}(h) + G_{HV}^{W}(h) + G_{EHV}^{W}(h)$$

$$G_{Total}^{S}(h) = G_{LV}^{S}(h) + G_{MV}^{S}(h) + G_{HV}^{S}(h) + G_{EHV}^{S}(h)$$

$$G_{Total}^{H}(h) = G_{LV}^{H}(h) + G_{MV}^{H}(h) + G_{HV}^{H}(h) + G_{EHV}^{H}(h)$$

$$G_{Total}^{B}(h) = G_{LV}^{B}(h) + G_{MV}^{B}(h) + G_{HV}^{B}(h) + G_{EHV}^{WB}(h)$$

$$G_{Total}^{Wa}(h) = G_{LV}^{Wa}(h) + G_{MV}^{Wa}(h) + G_{HV}^{Wa}(h) + G_{EHV}^{Wa}(h)$$

The following equation shows the breakdown process, assuming that $Factor_{VL}^{Asset}$ represents the fraction of the total generation of a specific $Asset$ in a voltage level VL.

$$G_{VL}^{Asset}(h) = Factor_{VL}^{Asset} * G^{Asset}(h)$$

On the one hand, the disaggregation between distribution and transmission is performed according to the data provided by an expert of a transmission system operator and assumptions. The disaggregation factors are shown in Table 4.

Table 4 Asset disaggregation for voltage profiles in Distribution and Transmission (Source: REE + Assumptions)

	Distribution	**Transmission**
Wind onshore fleet	35%	65%
Solar fleet	95%	5%
Hydro ROR fleet	90%	10%
Biomass fleet	85%	15%
Waste fleet	85%	15%

On the other hand, the distribution disaggregation in LV, MV and HV is performed based on data collected from JRC DSO Observatory. The disaggregation factors are shown in Table 5.

Table 5 Asset disaggregation for voltage profiles in LV, MV and HV (Source: DSO observatory)

	D-LV	**D-MV**	**D-HV**
Wind onshore fleet	5%	20%	9%
Solar fleet	68%	25%	2%
Hydro ROR fleet	14%	32%	4%
Biomass fleet	8%	29%	13%
Waste fleet	1%	31%	18%

In order to combine both tables (Table 4 and Table 5), the following equations should be taken into account.

$$Factor_{LV}^{Asset} = Factor_{Distribution}^{Asset} * Factor_{D-LV}^{Asset}$$

$$Factor_{MV}^{Asset} = Factor_{Distribution}^{Asset} * Factor_{D-MV}^{Asset}$$

$$Factor_{HV}^{Asset} = Factor_{Distribution}^{Asset} * Factor_{D-HV}^{Asset}$$

$$Factor_{EHV}^{Asset} = Factor_{Transmission}^{Asset}$$

Finally, the break down factors ($Factor_{VL}^{Asset}$) used for the generation disaggregation are shown in Table 6.

Table 6 Asset disaggregation for voltage profiles

	LV	MV	HV	EHV
Wind onshore fleet	5.4%	20.4%	9.2%	65.0%
Solar fleet	67.9%	25.5%	1.6%	5.0%
Hydro_ROR_fleet	25.4%	58.0%	6.6%	10.0%
Biomass fleet	13.2%	49.4%	22.3%	15.0%
waste fleet	2.4%	51.9%	30.7%	15.0%

Considering the previous factors, **Figure 13**shows the asset disaggregation for the second season of the year (spring) in the Spanish case. This figure also shows the peak generation produced by solar energy during the central hours of the day, mainly connected in low and medium voltage.

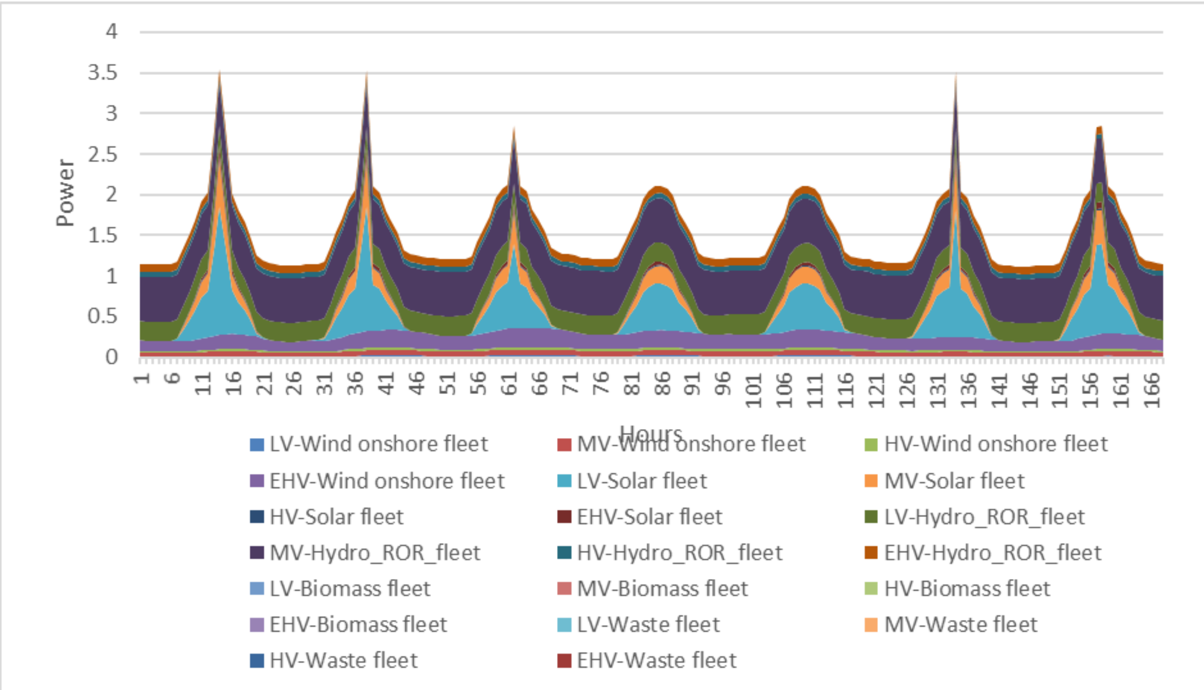

Figure 13 Asset disaggregation of the Spanish distributed generation in LV, MV, HV and EHV

4. DIVISION IN ARCHETYPES AND ZONES

In order to avoid an all-or-nothing effect in the analyses[3], it is necessary to model several network archetypes within each country. A division into three main archetypes for each country is proposed: **Urban, Semi-Urban and Rural**. This division has been chosen, because the type of area impacts in the grid characteristics and design criteria. For example, rural networks are much more dispersed and tend to suffer undervoltages, while urban networks are more concentrated and tend to experience overloads problems. The land and population percentages of each of these categories within every country are available in the JRC human settlement database [11], and included in Annex C. This type of division has already been applied for example in the DSO Observatory [1, 2], which will be one of the sources to obtain the required information that will drive this disaggregation. The generated archetypes have been obtained in such a way that their parameters (length of lines, number of substations, etc.) are calculated per unit area. In this way, the archetypes can be used in diverse distribution areas with different surface area values.

Additionally, the archetypes can be broken down according to **climatic zones** since this will affect the profiles of grid users, i.e. generation profiles of intermittent generation due to wind availability or level of solar irradiation, as well as demand profiles, e.g. in the case of thermo-sensitive loads. The divisions are based on criteria like temperature, irradiation, and wind production. This further division in zones is not expected to modify the per square kilometer models of the distribution networks but can influence the demand and generation profiles, which is expected to impact the results. Besides, a percentage matrix for each country has to be obtained (see example in Figure 14) to disaggregate the archetypes within each country. This will also have an impact on the absolute values of each archetype and zone.

[3] If a single archetype were built for each country, the analysis could detect that there is congestion in the LV power line, but that would imply that all the LV power lines of the country would be congested. To avoid this effect, several archetypes will be built, considering also several demand and generation profiles, that will have a correspondence with each of the zones identified within each country.

Figure 14 illustrates an example of the type of disaggregation matrix that is expected to be obtained, where each climatic zone has a percentage of the assets of the country, and the assets are also distributed among archetypes (urban, semi-urban, rural). For each climatic zone and archetype there is a percentage of the assets of the country, which can be calculated as the product of the associated probabilities.

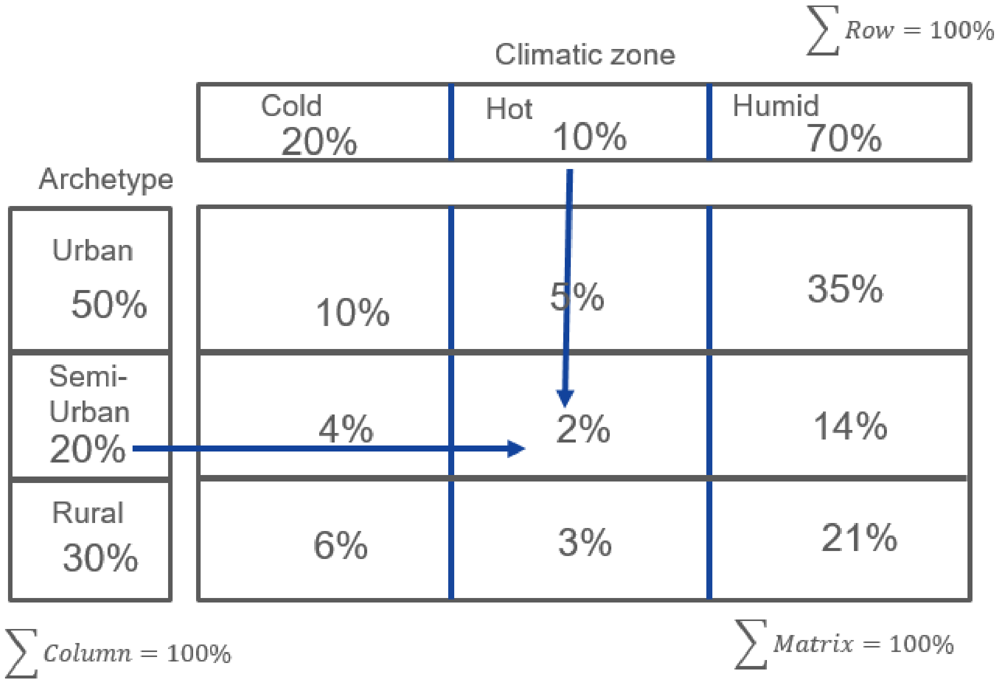

Figure 14 Example of percentage matrix of each zone and archetype within a country.

5. DISAGGREGATION OF THE ASSETS: SPAIN CASE STUDY

The disaggregation of the demand and generation assets per climatic zone, area (urban, semi-urban, rural) and voltage level is illustrated in this section, applying it to the case study of Spain.

5.1 Climatic zones

As example to test the methodology, we have applied it to the case of Spain. In this section it is shown how the division into climatic zones is applied in this country. The first step is the identification of the climatic zones. For the division into climatic zones, we use the following information:

- Settlement identification from the JRC human settlement database [11].
- Irradiation, temperature and wind generation yearly profiles in each of the settlements according to [29, 30].

Using this input, we analyze the number of clusters of the climatic zones, and how representative they are of the situation in the country. This method allows selecting the optimal number of clusters for the existing settlements according to three metrics: (1) sum of the mean square error, (2) average root mean square error of the duration curve, and (3) the average correlation error. Finally, the number of clusters in a country is obtained through a clustering process called Elbow method based on K-means[4].

Following this approach, the selected number of clusters for Spain is equal to 6. These clusters are represented as different colors in the following figure. In addition, it can be observed that, as expected, the clusters obtained include nearby settlements since their characteristics are similar.

[4] k-means is a clustering technique that aims to assign several observations (settlements in this case) into clusters in which each observation belongs to the cluster with the nearest mean.

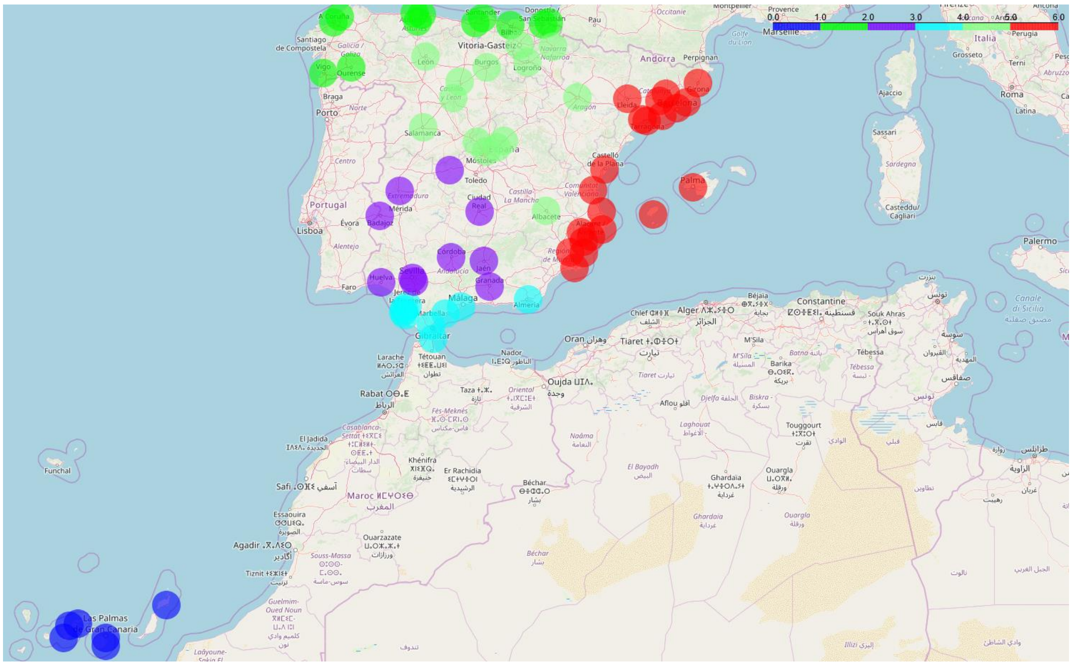

Figure 15 Climatic zone clusters in Spain.

Once the clusters are identified, the variables available to distribute the assets between the clusters of climatic zones are (1) the population, (2) the surface (proxy: number of clusters), (3) temperature, (4) irradiation and (5) wind generation. Each asset can be allocated following different criteria. We propose to use the following variables to drive the assignment of the assets to the climatic zones.

Table 7 Variables to drive the distribution of assets between climatic zones

Asset category	**Asset names**	**Climatic zone driver**
Flexible demand	Heat pumps (HP) Domestic hot water	Temperature and population Population
Non-flexible demand	Air conditioning (AC) Thermosensitive remainder (TR) Non-thermosensitive remainder	Temperature and population Temperature and population Industrialisation
Generation	Wind onshore fleet Solar fleet Hydro_ROR_fleet Biomass fleet Waste fleet	Wind and surface Irradiation and surface Surface Surface Surface
EV	PHEV home charge PHEV work charge BEV home charge BEV work charge	Population Population Population Population
Storage	Batteries	Population

The non-thermosensitive remainder has been distributed using data about the industrial sector in each region (source: Eurostat data [31]), as this demand depends mainly on the industry. The

Eurostat data specifies the gross value added at basic prices by NUTS 3 region. In order to evaluate the gross value added at each of the locations of the climatic zones, the Voronoi diagram has been applied to identify the nearest location for each NUTS 3 region. This process is illustrated in Figure 16.

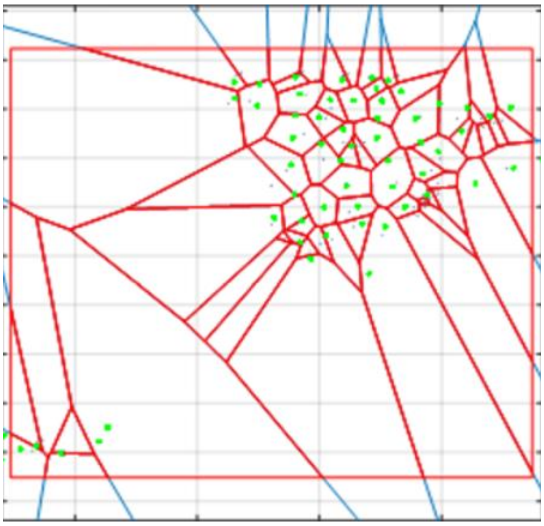

Figure 16 Voronoi diagram to assign NUTS level 3 regions to locations inside the countries.

The result is the gross value added at each location of the climatic zones. This is illustrated in Figure 17.

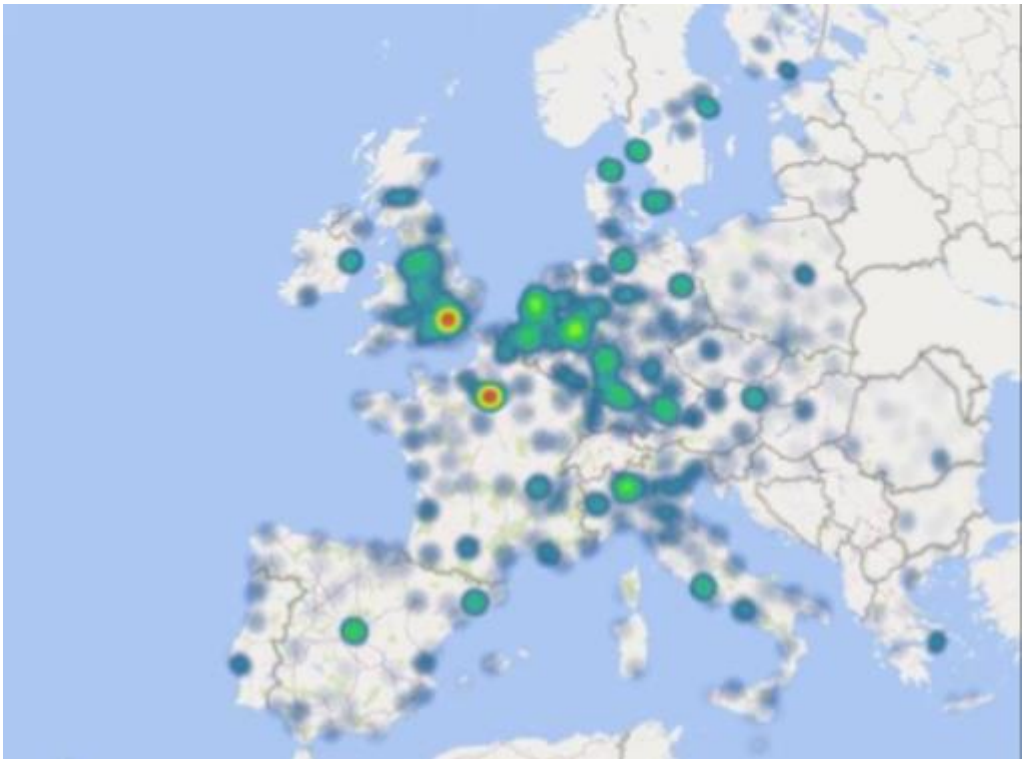

Figure 17 Heat map of the industrialisation level.

For the temperature, we consider the relation between load and temperature identified in [32], which has the following curve and equation.

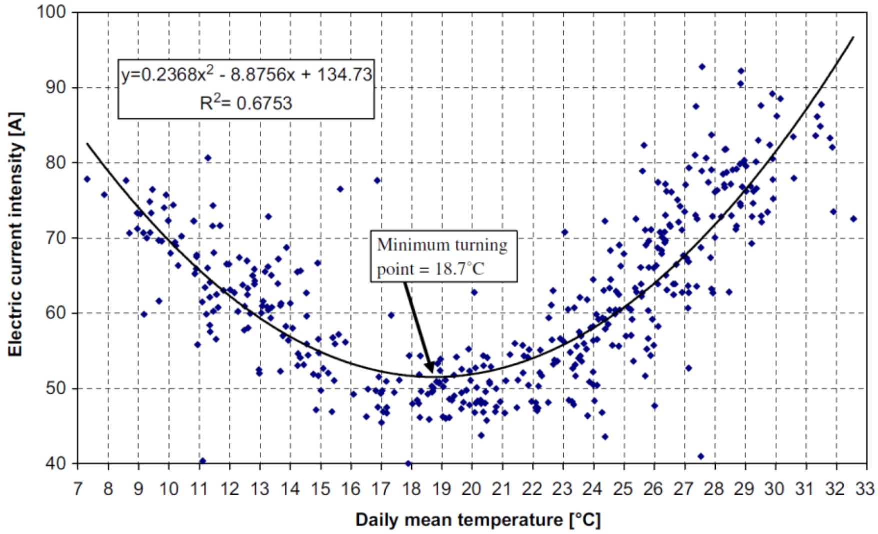

Figure 18 Relation between temperature and electric demand.

We use the shape of the relation between temperature and demand above the minimum in the example provided in Figure 18 to model the thermosensitive remainder, which is the remainder of the demand that depends on temperature.

Figure 19 shows the relation between the heating and cooling demand and temperature, according to [33]. We use the shape of the whole curve for the relation among temperature and demand of heat pumps, and only use the cooling curve (on the right of the figure) for modelling the relation among temperature and demand of air conditioning.

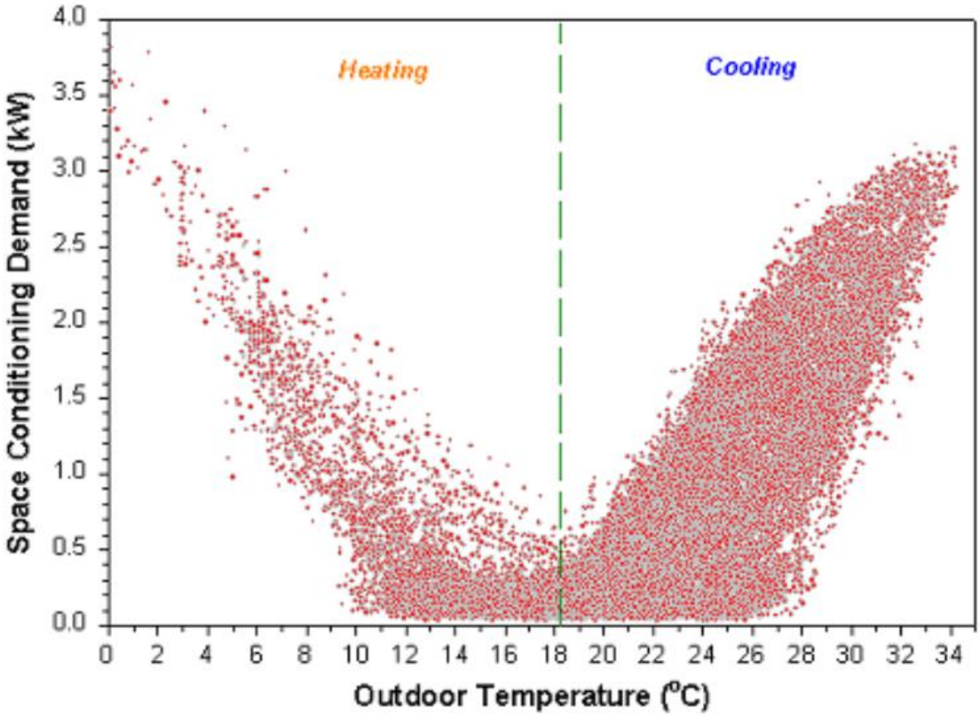

Figure 19 Relation between temperature, heating and cooling demand.

We use the above curves to define the relation between temperature and demand, and according to them, we obtain linear and quadratic regressions that model them, resulting in the following equations for the weights that depend on temperature.

$$weigth_{T,TR} = 0.2368 \times T^2 - 8.8756 * T + 134.73$$

$$weigth_{T,HP} = \begin{cases} 0.0135 \times T^2 - 0.4413 * T + 3.76 & T < 16 \\ 0.1552 + (0.2934 - 0.1552) * (T - 16)/(18 - 2) & 16 \leq T \leq 18 \\ 0.00714 \times T^2 - 0.2 * T + 1.58 & T > 18 \end{cases}$$

$$weigth_{T,AC} = \begin{cases} 0.2934 & T \leq 18 \\ 0.00714 \times T^2 - 0.2 * T + 1.58 & T > 18 \end{cases}$$

Where $weigth_{T,TR}$ is the non-thermosensitive remainder weight associated to temperature, $weigth_{T,HP}$ is the heat pump weigth associated to temperature and $weigth_{T,AC}$ is the air conditioning weight associated to temperature.

Figure 20 shows the curve that we use to model the relation between the heating and cooling demand and temperature (according to the input from Figure 19), using the previous equations.

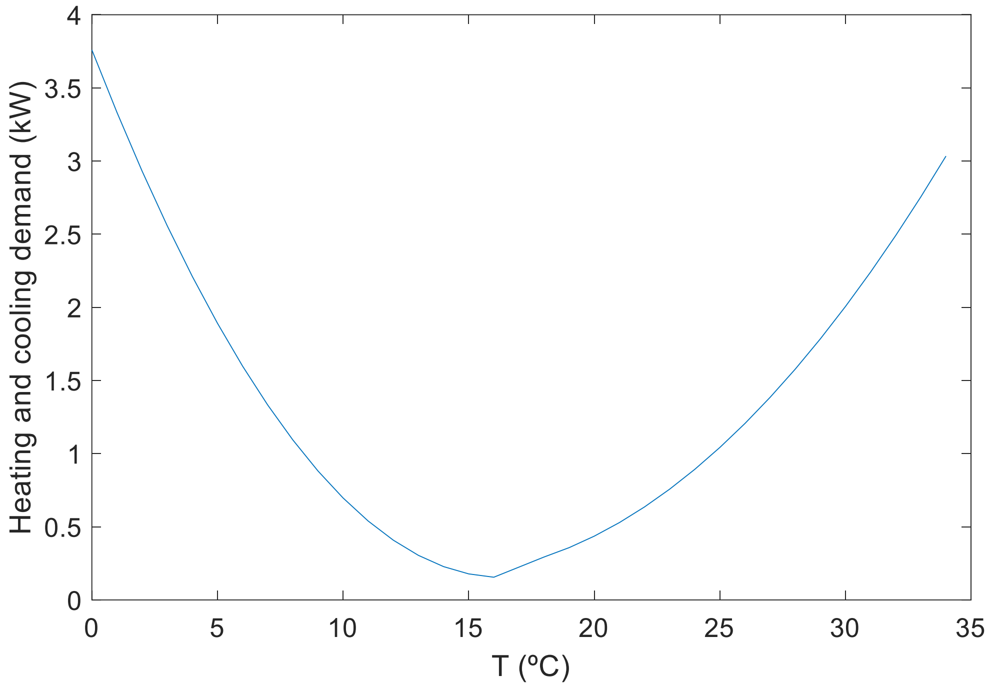

Figure 20 Proxy to model the relation between temperature, heating and cooling demand.

With the given clusters in each country, calculated as explained above, we obtain the number of settlements in each cluster (from the JRC human settlement database [11]) and their population.

Table 8 Values of the number of clusters and population to drive the distribution of assets between climatic zones

Cluster	Number of settlements	Population
1	6	1,056,701
2	12	2,451,744
3	10	1,906,553
4	9	1,840,233
5	14	7,210,966
6	21	7,741,643
Total	72	22,207,841

As the JRC human settlement database does not cover 100% of population and settlements, we normalize the values, obtaining percentages.

Table 9 Percentages of the number of clusters and population to drive the distribution of assets between climatic zones

Cluster	Number of settlements (proxy of size)	Population
1	8.3%	4.8%
2	16.7%	11.0%
3	13.9%	8.6%
4	12.5%	8.3%
5	19.4%	32.5%
6	29.2%	34.9%
Total	100.0%	100.0%

We obtain the following (dimensionless) weights for temperature, solar and wind yearly profiles for each climatic zone, where the weight for solar is derived from the irradiation.

Table 10 Values of the variables to drive the distribution of assets between climatic zones

Cluster	Weight solar	Weight temperature air conditioning	Weight temperature heat pumps	Weight temperature thermosensitive remainder	Weight wind	Industrialisation
1	0.290	0.392	0.398	1.502	0.411	35,626
2	0.181	0.120	0.696	14.740	0.307	131,938
3	0.243	0.394	1.010	19.399	0.237	114,637
4	0.261	0.295	0.515	7.592	0.292	57,020
5	0.213	0.199	1.298	27.503	0.279	312,388
6	0.227	0.348	0.721	12.102	0.239	329,250
Average	0.229	0.285	0.816	15.103	0.279	

We scale the values of the weights to normalize them respect the average value that we set to 1.0[5].

Table 11 Percentage variables to drive the division into climatic zones

Cluster	Normalized weigth solar	Normalized weight temperature air conditioning	Normalized weight temperature heat pumps	Normalized weight temperature thermosensitive remainder	Weight wind	Industrialisation
1	1.27	1.38	0.49	0.10	1.47	0.04
2	0.79	0.42	0.85	0.98	1.10	0.13
3	1.06	1.39	1.24	1.28	0.85	0.12
4	1.14	1.04	0.63	0.50	1.05	0.06
5	0.93	0.70	1.59	1.82	1.00	0.32
6	0.99	1.22	0.88	0.80	0.86	0.34
Average	1.00	1.00	1.00	1.00	1.00	

The values of these variables make sense to the best of our knowledge. For example, Cluster 1 (which is Canary islands) has low population and size (see Table 9), with high irradiation (normalized weight solar), high air conditioning consumption and low heat pumps consumption, thermosensitive remainder and low industrialisation. Cluster 3 (which is in the south, but far from the coastline) also requires a lot of air conditioning. Cluster 5 (which is in the north, but far from

[5] In principle, the normalization does not affect the results and could be omitted.

the coastline) is the one with more heat pump consumption and thermosensitive remainder. It includes Madrid and has high industrialisation. Cluster 6 (which corresponds to “Levante”, in the eastern coastline of Spain), has significant air conditioning demand and high industrialisation.

We then apply the variables that we identified in Table 7 to determine the percentage of each asset in each climatic zone. For example, air conditioning depends on temperature and population. Therefore, for this asset both weights are multiplied and combined, normalizing again the result to obtain percentages that sum 100%. For the demand assets we obtain the disaggregation shown in Table 12.

Table 12 Percentages of the demand assets in each climatic zone

Asset	**Heat pumps**	**Domestic hot water**	**Air conditioning**	**Thermosensitive remainder**	**Non-thermosensitive remainder**	**EV and batteries**
Variables	**Temperature and population**	**Population**	**Temperature and population**	**Temperature and population**	**Industrialisation**	**Population**
Cluster 1	2.1%	4.8%	6.7%	0.4%	3.6%	4.8%
Cluster 2	8.6%	11.0%	4.8%	9.5%	13.5%	11.0%
Cluster 3	9.7%	8.6%	12.3%	9.7%	11.7%	8.6%
Cluster 4	4.7%	8.3%	8.9%	3.7%	5.8%	8.3%
Cluster 5	46.9%	32.5%	23.4%	52.1%	31.8%	32.5%
Cluster 6	28.0%	34.9%	43.9%	24.6%	33.6%	34.9%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

And for the generation assets, we obtain the disaggregation shown in Table 13.

Table 13 Percentages of the generation assets in each climiatic zone

Asset	Wind onshore fleet	Solar fleet	Hydro_ROR_fleet	Biomass fleet	waste fleet
Variables	**Wind and size (proxy of surface)**	**Irradiation and size (proxy of surface)**	**Size (proxy of surface)**	**Size (proxy of surface)**	**Size (proxy of surface)**
Cluster 1	12.3%	10.6%	8.3%	8.3%	8.3%
Cluster 2	18.3%	13.2%	16.7%	16.7%	16.7%
Cluster 3	11.8%	14.8%	13.9%	13.9%	13.9%
Cluster 4	13.1%	14.3%	12.5%	12.5%	12.5%
Cluster 5	19.4%	18.2%	19.4%	19.4%	19.4%
Cluster 6	25.1%	29.0%	29.2%	29.2%	29.2%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

5.2 Urban, semi-urban, rural

We propose the following variables to drive the distribution of the assets between urban, semi-urban and rural.

Table 14 Variables to derive the distribution of assets between urban, semi-urban and rural

Aggregated assets	Individual assets	Urban-semi urban-rural
Flexible demand	Heat pumps	Population
	Domestic hot water	Population
Non-flexible demand	Air conditioning	Population
	Thermosensitive remainder	Population
	Non-thermosensitive remainder	Population[6]
Generation	Wind onshore fleet	Surface
	Solar fleet	Surface
	Hydro_ROR_fleet	Surface
	Biomass fleet	Surface
	Waste fleet	Surface
EV	PHEV home charge	Population
	PHEV work charge	Population
	BEV home charge	Population
	BEV work charge	Population
Storage	Batteries	Population

[6] In the case of the archetypes there is not a matching with locations and therefore, population and surface are the only drivers to distribute the assets.

We use as input the JRC Human Settlement Database [11]. We assign urban centers (with a density of at least 1500 inhabitants per km2 and a minimum total population of 50000) to the urban archetype. We assign urban clusters (with a density of at least 300 inhabitants per km2 and a minimum population of 5000) to semi-urban. Finally, we assign rural grid cells (with a density below 300 inhabitants per km2 outside urban clusters or centers) to rural. Using this classification, we obtain the following values for population and surface in each archetype in Spain.

Table 15 Population and surface in urban, semi-urban and rural

	Population	Population (%)	Area (km2)	Area (%)
Urban	22,469,245	48.1%	4,745	0.9%
Semi-urban	13,688,107	29.3%	11,761	2.3%
Rural	10,575,686	22.6%	488,276	96.7%
Total	46,733,038	100.0%	504,782	100.0%

Using the variables proposed in Table 14 for each of the assets, we obtain the following disaggregation for demand, EVs and batteries.

Table 16 Disaggregation of demand assets.

Asset	Heat pumps	Domestic hot water	Air conditioning	Thermosensitive remainder	Non-thermosensitive remainder	EV and batteries
Variable	**Population**					
Urban	48.1%					
Semi-urban	29.3%					
Rural	22.6%					
Total	100.0%					

Similarly, the disaggregation of the generation assets in urban, semi-urban and rural is shown in Table 17, where generation is mostly installed in rural areas, being the urban and semi-urban areas significant in terms of population and demand, but small in terms of surface and generation.

Table 17 Disaggregation of generation assets.

Asset	Wind onshore fleet	Solar fleet	Hydro_ROR_fleet	Biomass fleet	waste fleet
Variable	**Surface**				
Urban	0.9%				
Semi-urban	2.3%				
Rural	96.7%				
Total	100.0%				

5.3 Disaggregation of network assets

The disaggregation of the network assets is illustrated in this section with the disaggregation of the MV/LV substations in Spain, but the same procedure can be applied to:

- Density of MV/LV transformer substations.
- Density of HV/MV substations.
- Density of TSO-DSO interconnection points.
- Density of LV, MV & HV consumers.
- Density of LV, MV & HV network length.

The input data for urban, semi-urban and rural is the density of MV/LV substation in each type of network (from DSO Observatory networks) and the population and surface of urban, semi-urban and rural (from the JRC Human Settlement Database). To obtain the percentages of MV/LV substations within each area, the density of substations is multiplied by the surface, and the result is normalized to add 100%[7]. The idea behind this calculation is to multiply the density of MV/LV substations by the surface to obtain the number of MV/LV substations in urban-semi-urban and rural areas, as shown in Table 18.

On rare occasions, the values obtained using this methodology were missing or not in a reasonable or expected range. In these cases, it has been decided to perform an inter- or extrapolation of the densities analyzed with countries with a population density similar to the country with the inconsistencies.

Table 18 Urban, semi-urban and rural input

	Density substation/km2 MV-LV	Population %, Spain	Surface %, Spain	Estimation of percentage of substations[8]
Urban	48.65	48.1%	0.9%	28.2%
Semi-urban	13.53	29.3%	2.3%	19.4%
Rural	0.88	22.6%	96.7%	52.4%
Total		100.0%	100.0%	100.0%

[7] For the density of consumers, the population density of the country in urban, semi-urban and rural could be used alternatively to the density of consumers in the urban, semi-urban and rural networks from the DSO Observatory. The advantage of using the consumer density in the urban, semi-urban and rural networks is that the criterion is homogeneous for all the variables and the differentiation between LV, MV & HV.

[8] Density of substations multiplied by surface and normalized.

For the climatic zones, the input data to obtain the percentage of substations is the size / number of clusters (proxy of surface) and the density of population (proxy of the density of MV/LV substations). Its product is evaluated and the result normalized to add 100%[9]).

Table 19 Cluster input

Cluster	Size (proxy of surface)	Population	Density of population	Estimation of percentage of substations[10]
1	8.3%	4.8%	57.1%	4.8%
2	16.7%	11.0%	66.2%	11.0%
3	13.9%	8.6%	61.8%	8.6%
4	12.5%	8.3%	66.3%	8.3%
5	19.4%	32.5%	167.0%	32.5%
6	29.2%	34.9%	119.5%	34.9%
Total	100.0%	100.0%		100.0%

The percentage of substations is obtained in Table 18 for urban, rural and semi-urban areas, and in Table 19 for each cluster / climatic zone. Using this input, the percentage of the number of MV/LV substations is obtained for each archetype and cluster, by multiplying them (assuming that they are independent variables).

Table 20 Percentage of MV/LV substations for each archetype and cluster

			Urban	Semi-urban	Rural	
			Density of substations and surface			
			28.2%	19.4%	52.4%	100.0%
Cluster 1	Density of population (proxy of density) and size (proxy of surface)	4.8%	1.3%	0.9%	2.5%	
Cluster 2		11.0%	3.1%	2.1%	5.8%	
Cluster 3		8.6%	2.4%	1.7%	4.5%	
Cluster 4		8.3%	2.3%	1.6%	4.3%	
Cluster 5		32.5%	9.2%	6.3%	17.0%	
Cluster 6		34.9%	9.8%	6.8%	18.3%	
		100.0%				100.0%

[9] As observed in the table, this is equivalent to using the percentages of population, but conceptually it is the product of a density per the surface.

[10] Density of population (proxy of substation density) multiplied by size (proxy of surface) and normalized

Using these percentages and the total number of MV/LV substations in Spain (317,326), the respective number of MV/LV substations in each archetype and cluster is obtained.

Table 21 Number of MV/LV substations for each archetype and cluster

			Urban	Semi-urban	Rural	
			Density of substations and surface			
			89442	61657	166228	317326
Cluster 1	Density of population (proxy of density) and size (proxy of surface)	15099	4256	2934	7909	
Cluster 2		35033	9874	6807	18352	
Cluster 3		27243	7679	5293	14271	
Cluster 4		26295	7412	5109	13774	
Cluster 5		103037	29042	20020	53975	
Cluster 6		110620	31180	21493	57947	
		317326				317326

For the percentages of surface, the size (proxy of surface) is used for the clusters.

Table 22 Percentage of surface for each archetype and cluster

			Urban	Semi-urban	Rural	
			Surface			
			0.9%	2.3%	96.7%	100.0%
Cluster 1	Size(proxy of surface)	8.3%	0.1%	0.2%	8.1%	
Cluster 2		16.7%	0.2%	0.4%	16.1%	
Cluster 3		13.9%	0.1%	0.3%	13.4%	
Cluster 4		12.5%	0.1%	0.3%	12.1%	
Cluster 5		19.4%	0.2%	0.5%	18.8%	
Cluster 6		29.2%	0.3%	0.7%	28.2%	
		100.0%				100.0%

In order to obtain the surface of each archetype and cluster, the total surface of Spain (504,782 km2) is combined with the percentages of the surface (Table 22).

Table 23 Surface (km2) for each archetype and cluster

			Urban	Semi-urban	Rural	
			Density of substations and surface			
			4745	11761	488276	504782
Cluster 1	Size (proxy of surface)	42065	395	980	40690	
Cluster 2		84130	791	1960	81379	
Cluster 3		70109	659	1634	67816	
Cluster 4		63098	593	1470	61034	
Cluster 5		98152	923	2287	94942	
Cluster 6		147228	1384	3430	142414	
		504782				504782

The density of MV/LV substations in each archetype and cluster, is obtained by dividing the number of MV/LV substations by the respective surface.

Table 24 Density of MV/LV (1/km²) substations for each archetype and cluster

		Urban	Semi-urban	Rural	
		18.85	5.24	0.34	0.63
Cluster 1	0.36	10.76	2.99	0.19	
Cluster 2	0.42	12.49	3.47	0.23	
Cluster 3	0.39	11.65	3.24	0.21	
Cluster 4	0.42	12.50	3.48	0.23	
Cluster 5	1.05	31.48	8.75	0.57	
Cluster 6	0.75	22.53	6.27	0.41	
	0.63				0.63

It can be checked that the total density of MV/LV substations coincides with the average in Spain (0.63 MV/LV substations per km2). The percentages are different for different archetypes and also for different clusters of climatic zones within each archetype.

6. DATA COLLECTED FROM JRC

In the second stage of the data collection process[11] the collected data was refined and contrasted. During November 2019, Comillas' staff has travelled one week to JRC to gather additional data, in order to further improve and complete the data collection process.

The JRC DSO Observatory project, initially carried out with the collaboration of Comillas Pontifical University, is the most updated and complete data source about Distribution System Operators in the European Union [1, 2]. For this reason, the data collected in Section 2 has been improved and completed using this source. The data collected in JRC refers mainly to parameters that are used to calculate the densities of consumers, substations and network length. This has allowed to review and improve the data that has already been collected, and to fill in missing gaps. The parameters collected from the JRC, specifying the granularity are shown in Table 25. Some parameters, like the number of HV/MV substations and the installed capacity of the HV/MV and MV/LV substations are not available from other sources and have only been collected from the DSO Observatory. This data is not available in the public reports of this project [1, 2]. Therefore, in order to collect it, the collaboration of JRC Energy Security, Distribution and Markets Unit has been indispensable.

[11] This corresponds to the second delivery of the data collection process described in the proposal for the second year.

Table 25 Data collected from the JRC DSO Observatory

Parameters in DSO Observatory	Granularity
Number of HV/MV Substations	Per country
Number of MV/LV Secondary Substations	Per country
Total installed capacity of HV/MV Substations (MVA)	Per country
Total installed capacity of MV/LV Secondary Substations (MVA)	Per country
Area of Distribution Activity (approximately) (km2)	Per country
Number of TSO-DSO interconnection points	Per country
Toral Installed Capacity [MW]	Per country
Total Gross Electricity Generation [GWh]	Per country
DG Connected to HV (>36kV) [%]	Per country
DG Connected to LV (1kV) [%]	Per country
DG Connected to MV (1-36 kV) [%]	Per country
Technologies: Photovoltaic, Wind, Biomass, Waste, Hydro	Per country
DSO Network length per voltage level: HV (> 36 kV): overhead	Per country
DSO Network length per voltage level: HV (> 36 kV): underground	Per country
DSO Network length per voltage level: LV (< 1 kV): overhead	Per country
DSO Network length per voltage level: LV (< 1 kV): underground	Per country
DSO Network length per voltage level: MV (1-36 kV): overhead	Per country
DSO Network length per voltage level: MV (1-36 kV): underground	Per country
Number of HV (> 36 kV) Customers	Per country
Number of LV (< 1 kV) Customers	Per country
Number of MV (1- 36 kV) Customers	Per country

4. Annex A: Archetype voltage levels per country

Table 26 shows the considered voltage levels per country, based on Eurelectric [19].

Table 26 Voltage levels per country

		Voltage LV (kV)	Voltage MV (kV)	Voltage HV (kV)
Austria	**AT**	0.4	20	110
Belgium	**BE**	0.4	15	-
Bulgaria	**BG**	0.4	20	110
Cyprus	**CY**	0.4	22	110
Czech Rep	**CZ**	0.4	22	110
Germany	**DE**	0.4	20	110
Denmark	**DK**	0.4	20	60
Estonia	**EE**	0.4	20	-
Greece	**GR**	0.4	20	110
Spain	**ES**	0.4	20	132
Finland	**FI**	0.4	20	110
France	**FR**	0.4	20	-
Croatia	**HR**	0.4	20	110
Hungary	**HU**	0.4	20	120
Ireland	**IE**	0.4	20	110
Italy	**IT**	0.4	20	150
Lithuania	**LT**	0.4	20	-
Luxembourg	**LU**	0.4	20	-
Latvia	**LV**	0.4	20	-
Malta	**MT**	0.4	11	132
Netherlands	**NL**	0.4	20	110
Poland	**PL**	0.4	20	110
Portugal	**PT**	0.4	15	-
Romania	**RO**	0.4	20	110
Sweden	**SE**	0.4	24	110
Slovenia	**SI**	0.4	20	110
Slovakia	**SK**	0.4	20	110
United Kingdom	**UK**	0.4	33	150
Bosnia and Herzegovina	**BA**	0.4	20	110
Switzerland	**CH**	0.4	20	110
Montenegro	**ME**	0.4	20	110
North Macedonia	**MK**	0.4	20	110
Norway	**NO**	0.4	22	132
Serbia and Kosovo	**RS**	0.4	20	110

5. ANNEX B: PERCENTAGES OF CONSUMPTION OF RESIDENTIAL, COMMERCIAL AND INDUSTRIAL SECTORS

Table 27 shows the considered percentages or consumption of residential, commercial and industrial sectors per country, based on the IEA [13].

Table 27 Percentages or consumption of residential, commercial and industrial sectors per country

Country	ID	%RES	%COM	%IND	%OTHERS
Austria	AT	28%	19%	48%	5%
Belgium	BE	23%	27%	49%	2%
Bulgaria	BG	38%	30%	32%	1%
Cyprus	CY	39%	49%	12%	0%
Czech Rep	CZ	27%	28%	42%	3%
Germany	DE	25%	29%	44%	2%
Denmark	DK	33%	36%	29%	1%
Estonia	EE	28%	39%	32%	1%
Greece	GR	38%	37%	24%	0%
Spain	ES	31%	32%	35%	2%
Finland	FI	28%	22%	49%	1%
France	FR	37%	33%	28%	3%
Croatia	HR	39%	36%	23%	2%
Hungary	HU	30%	22%	45%	3%
Ireland	IE	31%	28%	41%	0%
Italy	IT	23%	33%	40%	4%
Lithuania	LT	29%	34%	37%	1%
Luxembourg	LU	15%	33%	50%	2%
Latvia	LV	26%	44%	28%	2%
Malta	MT	33%	47%	20%	0%
Netherlands	NL	23%	37%	37%	2%
Poland	PL	22%	35%	41%	2%
Portugal	PT	28%	36%	35%	1%
Romania	RO	29%	19%	49%	2%
Sweden	SE	36%	22%	40%	2%
Slovenia	SI	25%	26%	48%	2%
Slovakia	SK	19%	30%	48%	2%
United Kingdom	UK	36%	32%	31%	2%
Bosnia and Herzegovina	BA	42%	19%	38%	1%
Switzerland	CH	33%	29%	31%	7%
Montenegro	ME	45%	28%	26%	1%
North Macedonia	MK	51%	25%	23%	0%
Norway	NO	36%	22%	41%	1%
Serbia and Kosovo	RS	50%	19%	30%	1%

6. ANNEX C: PERCENTAGES OF POPULATION AND LAND IN URBAN, SEMI-URBAN AND RURAL AREAS

Table 28 shows the considered percentages or population and land in urban, semi-urban and rural areas, based on the JRC human settlement database [11].

Table 28 Percentages of population and land in urban, semi-urban and rural areas, based on the JRC human settlement database.

		Population, %Urban	Population, %Semiurban	Population, %Rural	Land, %Urban	Land, %Semiurban	Land, %Rural
Austria	**AT**	30.90%	26.57%	42.53%	0.76%	2.88%	96.36%
Belgium	**BE**	32.43%	45.05%	22.52%	4.55%	20.89%	74.56%
Bulgaria	**BG**	26.56%	29.49%	43.95%	0.42%	1.51%	98.07%
Cyprus	**CY**	38.72%	30.34%	30.94%	1.77%	3.95%	94.28%
Czeh Rep	**CZ**	22.00%	36.70%	41.30%	0.91%	4.30%	94.79%
Germany	**DE**	32.87%	36.68%	30.45%	2.53%	7.89%	89.58%
Denmark	**DK**	29.28%	32.81%	37.91%	1.18%	3.87%	94.95%
Estonia	**EE**	33.13%	29.66%	37.21%	0.28%	0.72%	99.00%
Greece	**GR**	45.12%	22.75%	32.13%	0.53%	1.59%	97.88%
Spain	**ES**	48.08%	29.29%	22.63%	0.94%	2.33%	96.73%
Finland	**FI**	24.68%	33.75%	41.57%	0.15%	0.56%	99.29%
France	**FR**	35.48%	26.29%	38.23%	1.02%	2.87%	96.11%
Croatia	**HR**	28.16%	29.16%	42.68%	0.66%	2.20%	97.14%
Hungary	**HU**	27.04%	35.06%	37.90%	0.81%	3.47%	95.72%
Ireland	**IE**	29.03%	24.82%	46.15%	0.57%	1.20%	98.23%
Italy	**IT**	35.19%	39.08%	25.74%	1.79%	5.32%	92.89%
Lithuania	**LT**	33.35%	34.20%	32.45%	0.50%	1.38%	98.12%
Luxembourg	**LU**	21.08%	43.80%	35.12%	1.63%	9.41%	88.96%
Latvia	**LV**	34.74%	29.98%	35.28%	0.32%	0.80%	98.88%
Malta	**MT**	66.75%	27.78%	5.47%	17.89%	22.56%	59.55%

Netherlands	**NL**	46.58%	35.70%	17.72%	6.89%	15.05%	78.06%
Poland	**PL**	27.55%	32.64%	39.81%	1.03%	3.73%	95.24%
Portugal	**PT**	35.10%	31.88%	33.02%	1.06%	3.45%	95.49%
Romania	**RO**	29.88%	25.10%	45.02%	0.51%	1.70%	97.79%
Sweden	**SE**	29.16%	34.92%	35.92%	0.20%	0.77%	99.03%
Slovenia	**SI**	16.13%	33.87%	50.00%	0.57%	3.42%	96.01%
Slovakia	**SK**	14.96%	36.90%	48.14%	0.51%	3.22%	96.27%
United Kingdom	**UK**	58.40%	26.40%	15.20%	4.27%	4.77%	90.96%
Bosnia and Herzegovina	**BA**	27.30%	34.36%	38.35%	1.15%	4.49%	94.36%
Switzerland	**CH**	27.30%	34.36%	38.35%	1.15%	4.49%	94.36%
Montenegro	**ME**	27.30%	34.36%	38.35%	1.15%	4.49%	94.36%
North Macedonia	**MK**	27.30%	34.36%	38.35%	1.15%	4.49%	94.36%
Norway	**NO**	27.30%	34.36%	38.35%	1.15%	4.49%	94.36%
Serbia and Kosovo	**RS**	27.30%	34.36%	38.35%	1.15%	4.49%	94.36%

7. References

1. Prettico, G., Flammini, M.G., Andreadou, N., Vitiello, S., Fulli, G., Masera, M.: Distribution System Operators observatory 2018. (2018)

2. Prettico, G., Gangale, F., Mengolini, A., Lucas, A., Fulli, G.: Distribution System Operators observatory. (2016)

3. Ministerio: Resolución de 10 de marzo de 2000, de la Secretaría de Estado de Industria y Energía, por la que se aprueba el procedimiento deoperación del sistema (P.O. - 7.4) «Serviciocomplementario de control de tensión de la red de transporte». (2000)

4. CEER: Electricity Voltage Quality. (2016)

5. IIT: Reference Network Models: Catalogs of components. (2019)

6. EC: Country Surface, www.europarl.europa.eu/elections-2014/website/.../datasheet.xls%0A

7. EC: Q4 Quarterly Report on European Markets. (2017)

8. EC: Q3 Quarterly Report on European Markets. (2017)

9. EC: Q2 Quarterly Report on European Markets. (2017)

10. EC: Q1 Quarterly Report on European Markets. (2017)

11. EC/JRC: Global Human Settlement Layer, http://ghsl.jrc.ec.europa.eu

12. IEC: Strategic Asset Management of Power Networks. White Pap. 1–84 (2015)

13. IEA: IEA: Statistics. Global energy data at your fingertips, https://www.iea.org/statistics/?country=AUSTRIA&year=2016&category=Key indicators&indicator=TPESbySource&mode=chart&dataTable=ELECTRICITYANDHEAT

14. CNMC: Acuerdo por el que se aprueba la propuesta de metodología de cálculo de la tasa de retribución financiera de las actividades de transporte y distribución de energí eléctrica para el segundo periodo regulatorio 2020-2025. (2018)

15. CEER Secretariat: CEER Report on Investment Conditions in European Countries. 200 (2017)

16. Ministerio: Orden IET/2660/2015, de 11 de diciembre, por la que se aprueban las instalaciones tipo y los valores unitarios de referencia de inversión, de operación y mantenimiento por elemento de inmovilizado y los valores unitarios de retribución de otras tareas reg. (2015)

17. Schörkhuber, C., Klapuri, A., Sontacchi, A.: Audio pitch shifting using the constant-Q transform. AES J. Audio Eng. Soc. 61, 562–572 (2013)

18. Gargour, C., Gabrea, M., Ramachandran, V., Lina, J.M.: A Short Introduction to Wavelets And Their Applications. IEEE Circuits Syst. Mag. 9, 57–68 (2009). https://doi.org/10.1109/MCAS.2009.932556

19. Eurelectric: Power Distribution in Europe - Facts and Figures. (2013)

20. REE: esios: Sistema de Información del Operador del Sistema,

https://www.esios.ree.es/es/generacion-y-consumo

21. SchneiderElectric: Transformadores de distribución MT con dieléctrico líquido hasta 36 kV. (2019)

22. Nexans: Aluminium cable 132 kV 630 R, https://www.nexans.co.uk/eservice/UK-en_GB/navigateproduct_510056/Aluminium_cable_132_kV_630_R_aluminium_.html#characteristics

23. Estralin: Cables de potencia XLPE Y sistemas de cable 6-220 KV. (2019)

24. SolidAl, Quintas&Quintas: Conductores Eléctricos: Catálogo. (2010)

25. TopCable: Catálogo de Cables para la distribución de energía eléctrica. (2019)

26. CablesRCT: Cables 0,6/1 kV. (2019)

27. Demand measured per voltage level (Spanish), https://www.esios.ree.es/es/generacion-y-consumo

28. Persson, U., Werner, S.: STRATEGO: Quantifying the Heating and Cooling Demand in Europe. (2015)

29. Modern-Era Retrospective analysis for Research and Applications (MERRA), https://disc.gsfc.nasa.gov/datasets?project=MERRA

30. PVGIS Photovoltaic Geographical Information System, https://joint-research-centre.ec.europa.eu/want-make-best-solar-energy-where-you-live_en

31. Eurostat, https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_3gva&lang=en

32. Beccali, M., Cellura, M., Lo Brano, V., Marvuglia, A.: Short-term prediction of household electricity consumption: Assessing weather sensitivity in a Mediterranean area. Renew. Sustain. Energy Rev. 12, 2040–2065 (2008). https://doi.org/10.1016/j.rser.2007.04.010

33. Parker, D.S.: Research highlights from a large scale residential monitoring study in a hot climate. Energy Build. 35, 863–876 (2003). https://doi.org/10.1016/S0378-7788(02)00244-X