Carl von Ossietzky Universität Oldenburg

Institute of Physics

European Master in Renewable Energy

Master's Thesis

Estimation and Simulation of a Gas Demand Time Series for the European NUTS 3 Regions

Submitted by:

Javier Enrique Sandoval

EUREC 2019 – 2020 Student

Matriculation Number: 5367077

First Examiner: Prof. Dr. Carsten Agert

Second Examiner: Dr. Herena Torio

Supervisors: Dr. Wided Medjroubi (DLR-VE)

Oldenburg, Germany, 2021

I. Declaration

I hereby state and declare that this thesis was prepared by me and that no other means or sources have been used, except those which I cited and listed in the References section. This thesis is in compliance with the rules of good practice in scientific research of the Carl von Ossietzky Universität Oldenburg.

Javier Enrique Sandoval

Oldenburg, June 28th, 2021

II. Acknowledgments

Working on this project was, for me, a vast part of a new beginning. It gave me focus, a goal, and a new perspective about what interests me and about what I want my professional future to be, and I could not be more grateful about it.

I want to thank my supervisors and colleagues at the DLR Institute of Networked Energy Systems, especially from the SciGRID_gas project: Dr. Wided Medjroubi, Dr. Jan Diettrich, and Dr. Adam Pluta, for giving me the opportunity of developing such an exciting project. Without their support, their advice, and their encouragement, my experience would not have as enriching as it was. Their work has inspired me to pursue a new professional career in the scientific field.

I would like to thank my supervisors at the Carl von Ossietzky Universität Oldenburg, Prof. Dr. Carsten Agert and Dr. Herena Torio, for their expertise in evaluating this work and their understanding disposition.

Finally, I would like to thank my family, especially my mother, and my friends from the PPRE/EUREC Master in Renewable Energy, for their encouragement and presence.

III. Abstract

With the goal of becoming climate-neutral by 2050, in past years Europe has increased its share of electricity generated by renewables. However, the demand for renewable-generated electricity should be decoupled from the time and place of supply. Technologies such as Power-to-gas technology allow to convert the electricity surplus to gas and then transport and store it using existing infrastructure. Hence, an appropriate model of the gas infrastructure is paramount, with gas demand as one of its important aspects.

While figures regarding gas demand at national levels are available, there is lack of data about European regional gas consumption. The aim of this thesis is to address that lack by the estimation of a gas demand time series in Europe with a high spatial resolution. For this, three research questions must be answered: (1) what are the variables that drive the gas consumption in Europe? (2) how is gas demand distributed regionally, and (3) how is gas demand distributed temporally?

In this thesis, a research on demand drivers is performed to answer these questions. Next, taking a top-down approach, the annual gas demand at national level of three economy sectors is disaggregated to regional level, using the corresponding statistical demand driver datasets. Finally, and also with a top-down approach, temporal demand driver datasets such as ambient temperature and type of day are used to estimate daily gas consumption at a regional level.

The results obtained in this thesis are validated and optimized with the help of a different study. They show that, in the residential sector, gas consumption is allocated not only in heavily populated areas, but that regional differences in ambient temperature or gas usage also play a role, with high residential gas consumption observed in Northern Italy and areas of the United Kingdom. Gas demand is in the commercial sector is found to be highly concentrated in major urban centers regardless of their geographical location, with large consumption in cities of Italy, Spain, Germany, and the Netherlands. The industry sector presents important gas consumption figures in regions of Spain and France with a combination of energy-intensive and non-intensive activities. Finally, the seasonal effect is observed in the residential and commercial sector load profiles of the European countries analyzed, being more perceivable in Northern Europe.

IV. Table of contents

I. Declaration i

II. Acknowledgments ii

III. Abstract iii

V. Abbreviations vi

VI. Figures viii

VII. Tables x

1. Introduction 1

2. Theory 3

- 2.1 Definitions and concepts 3
  - 2.1.1 Energy use sectors 3
  - 2.1.2 Energy demand determining variables 4
  - 2.1.3 Spatial disaggregation 5
  - 2.1.4 Temporal disaggregation 8
- 2.2 Statistical data review 8
  - 2.2.1 Energy consumption datasets 9
  - 2.2.2 Energy demand determining variables datasets 9
- 2.3 Literature overview 13
  - 2.3.1 DemandRegio 13
  - 2.3.2 Other publications 15

3. Methodology 18

- 3.1 Residential 21
  - 3.1.1 Definition of household consumers 21
  - 3.1.2 Description of gas demand in households 21
  - 3.1.3 Spatial disaggregation of the residential sector 26
  - 3.1.4 Temporal disaggregation of the residential sector 31
- 3.2 Industry and Commercial, Trade and Services 38
  - 3.2.1 Definition of the Industry and CTS consumers 38
  - 3.2.2 Description of gas demand in the industrial and commercial sectors . 40
  - 3.2.3 Spatial disaggregation of the industry and commercial sectors 43
  - 3.2.4 Temporal disaggregation of the industry and commercial sectors 49

4. Results 53

- 4.1 Presentation of the results 53
- 4.2 Processing of reference datasets 58

4.3 Model optimization 61
4.4 Discussion.. 62
4.4.1 Verification of the temporal disaggregation 62
4.4.2 Verification of the spatial disaggregation 69
4.4.3 Assessment of gas demand in Portugal.. 77
5. Conclusions 81
6. References xi

V. Abbreviations

AGEB	Energy Balance Working Group
API	Application programming interface
ARIMA	Autoregressive integrated moving average
ARMA	Autoregressive moving average
ARMAX	Autoregressive moving average with exogenous inputs
BCM	Billion cubic meters
BGW	Federal Association of German Gas and Water Management
BMWi	German Federal Ministry for Economic Affairs and Energy
C3S	Copernicus Climate Change Service
CTS	Commercial, trade, and services
CSV	Comma-separated values
DWD	German Weather Service
EIA	U.S. Energy Information Administration
EDDV	Energy demand determining variable
EFTA	European Free Trade Association
ERA5	European Centre for Medium-Range Weather Forecasts Reanalysis version 5
EU	European Union
Eurostat	European Statistical Office
FfE	Research Institute for Energy e.V. Munich
GDP	Gross domestic product
GWh	Gigawatt-hour
HDI	Human development index
HDD	Heating degree days
IN	Intermediate
JEVI	Annual survey of energy use in manufacturing
LKD-EU	Long-term planning and short-term optimization of the German electricity system within the European framework
ML	Machine learning

NACE	Statistical Classification of Economic Activities in the European Community
N.E.C.	Not elsewhere classified
NUTS	Nomenclature of territorial units for statistics
OLS	Ordinary least squares
P2G	Power-to-gas
PR	Primarily rural
PU	Primarily urban
SARIMA	Seasonal autoregressive integrated moving average
SBS	Statistical business structure
TJ	Terajoule
TSA	Time series analysis
UTC	Coordinated universal time

VI. Figures

Figure 1: Graphical representation of NUTS regions. Taken from [15]. 7
Figure 2: Methodology flowchart (Industrial and commercial sectors). 19
Figure 3: Methodology flowchart (Residential sector).. 20
Figure 4: Final energy gas consumption in the European Union by sector, 2010-2019 [42]. 22
Figure 5: Final energy consumption in households by use in the European Union - Gas, 2018 [21]... 22
Figure 6: Monthly natural gas demand in Europe (bcm). Taken from [47]. 31
Figure 7: 2-meter temperature (°C) in NUTS 3 regions, 2019.01.01, 00:00........ 33
Figure 8: Zoom-in on Belgium's NUTS 1 regions, illustrating Case 1 (Regions BE2 and BE3), and Case 2 (Region BE1)... 34
Figure 9: Energy mix in European countries in 2019, industry sector [13]. 41
Figure 10: Energy mix in European countries in 2019, commercial sector [13]. . 41
Figure 11: Final energy consumption in industry in Europe, gas, 2010-2019 [13]. 42
Figure 12: Final energy consumption in CTS in Europe, gas, 2010 - 2019 [13].. 43
Figure 13: Example of disaggregation of employees in a NUTS 2 region into NUTS 3 regions (industrial sector). 47
Figure 14: Example of disaggregation of employees in a NUTS 2 region into NUTS 3 regions using GDP (commercial sector). 48
Figure 15: Sample of results for NUTS 3 region AT111. 53
Figure 16: Gas demand in the residential sector, NUTS 3 regions, 2015. 54
Figure 17: Distribution of NUTS 3 regions by gas demand, residential sector, 2015. 55
Figure 18: Gas demand in the industrial sector, NUTS 3 regions, 2015. 55
Figure 19: Distribution of NUTS 3 regions by gas demand, industrial sector, 2015. 56
Figure 20: Gas demand in the commercial sector, NUTS 3 regions, 2015. 57
Figure 21: Distribution of NUTS 3 regions by gas demand, commercial sector, 2015. 58
Figure 22: Sample of DemandRegio's CTS result time series, hourly resolution.59
Figure 23: Sample of DemandRegio's CTS result time series, daily resolution. . 59
Figure 24: Comparison of household gas demand reference and results time series.. 62
Figure 25: Daily percent difference between reference and results datasets, residential sector.. 63
Figure 26: Distribution of days by percent error, residential sector. 64
Figure 27: Comparison of CTS gas demand reference and results time series.. 65
Figure 28: Daily percent difference between reference and results datasets, CTS sector.. 66
Figure 29: Distribution of days by percent error, CTS sector. 66
Figure 30: Comparison of industrial gas demand reference and results time series.. 67
Figure 31: Daily percent difference between reference and results datasets, industry sector. 68
Figure 32: Reference/results ratio in German NUTS 3 regions, residential sector 69

Figure 33: Number of NUTS 3 regions per ratio, residential sector. 70
Figure 34: Reference/results ratio in German NUTS 3 regions, industrial sector 72
Figure 35: Number of NUTS 3 regions per ratio, industrial sector. 73
Figure 36: Reference/results ratio in German NUTS 3 regions, commercial sector 75
Figure 37: Number of NUTS 3 regions per ratio, commercial sector. 76
Figure 38: Total gas demand in households, Portugal, 2015. 78
Figure 39: Total gas demand in the CTS sector, Portugal, 2015. 79
Figure 40: Total gas demand in the industrial sector, Portugal, 2015 80

VII. Tables

Table 1: Energy demand determining variables used in models [15, p. 4]. 5
Table 2: Definition of NUTS levels [17]. 6
Table 3: NUTS regions equivalents of selected countries [18]. 7
Table 4: Example of NUTS codes at different levels. 8
Table 5: NUTS regions count. 8
Table 6: List of datasets used for spatial disaggregation of the residential sector. 10
Table 7: List of datasets used for spatial disaggregation of the Industry and CTS sectors. 11
Table 8: Share of fuels in the final energy consumption in the residential sector for space heating, 2018 [46]. 24
Table 9: Share of fuels in the final energy consumption in the residential sector for water heating, 2018. 25
Table 10: Share of fuels in the final energy consumption in the residential sector for cooking, 2018 [46]. 26
Table 11: Disaggregated energy consumption in households, 2018 [22]. 27
Table 12: Sample of temperatures and HDDs in NUTS 3 regions. 36
Table 13: Final energy consumption by sectors, gas, 2010-2019 [14]. 40
Table 14: Datasets used for the spatial disaggregation of the industrial and commercial sectors. 44
Table 15: Correspondence between gas consumption and number of employees datasets for NACE industrial activities. 45
Table 16: Initial arbitrary values for the industrial sector. 50
Table 17: Initial arbitrary values for the commercial sector. 50
Table 18: DemandRegio's input vs. output yearly gas demand per sector, 2015. 60
Table 19: Project's input vs. output yearly gas demand per sector, 2015. 60
Table 20: Optimal values for temporal disaggregation variables. 61
Table 21: Regions with highest residential gas consumption in Germany, 2015 (DemandRegio project). 70
Table 22: Regions with highest residential gas consumption in Germany, 2015 (Results). 71
Table 23: Regions with highest industrial gas consumption in Germany, 2015 (DemandRegio project) 73
Table 24: Regions with highest industrial gas consumption in Germany, 2015 (Results). 74
Table 25: Regions with highest commercial gas consumption in Germany, 2015 (DemandRegio project). 76
Table 26: Regions with highest commercial gas consumption in Germany, 2015 (Results). 77

1. Introduction

The reduction of carbon dioxide emissions constitutes a global challenge to achieve the goal of reducing global warming [1]. Europe, in particular, intends to be climate-neutral by 2050. With this target, the European Council has set intermediate objectives in this regard, aiming, for example, to reduce greenhouse gas levels by the year 2030 to at least 55% below the figures existing in 1990 [2].

Electricity-producing renewable energies such as wind and solar play a determinant role in the reduction of CO_2 atmospheric concentration levels [3]. However, as the electricity generated by renewables may not be needed at the times and locations of its production, it must be stored or transported.

Renewable energy has largely been used by the power producing sector. In 2020, 38% of the electricity of Europe was generated using renewables, overtaking fossil-fueled generation for the first time in history [4]. However, other areas on the demand side, such as the transport sector, building heating/cooling, and industry, are still largely dependent on fossil fuels [5]. The use of renewables-generated electricity by those consumers would constitute a revolution and would further the decarbonization of European economies. Nevertheless, a society independent of fossil fuels requires the integration of the aforementioned actors in the generation and demand sides. The idea of interconnecting the energy-consuming and power-producing sectors has been named sector coupling [6].

One of today's most promising sector coupling technologies is the Power-to-gas (P2G) technology. It allows to convert a surplus of electricity to gas by means of electrolysis and then transport or store it using existing infrastructure [7]. The development of an appropriate model of such infrastructure is, then, fundamental.

Energy systems modelling is the process of building mathematical and computational models of energy systems in order to analyzed them [8]. It is highly important as it allows to understand and represent, as reliably as possible, the current challenges faced by technologies such as P2G, as well as provide tools for the design and planning of future infrastructure and energy policies.

While the information necessary for the modelling of the European electricity grid is openly available, it is not equally accessible for the gas sector. For this reason, the project SciGRID_gas was started in 2018, with the objective of generating a freely available, open dataset of the European gas grid transmission network [9]. As further details on gas supply and demand are needed, such a dataset should not only include information regarding infrastructure such as pipelines, compressors, LNG terminals, and storages, but also related to gas consumers.

One of the elements considered by the SciGRID_gas project is gas demand. From the supply side, information regarding the amount of gas entering European countries on a daily basis is available [10]. However, this does not apply to the demand side, where there is a lack of gas consumption data. To model the European gas grid adequately, it is necessary to improve this knowledge, especially on a regional scale.

This thesis addresses the creation of the required European gas demand data. The creation of such a dataset is performed using either a top-down or a bottom-up approach. In a top-down approach, gas demand data at a country level is disaggregated, first regionally and then temporally, while in a bottom-up approach, gas demand at a small scale (i.e., industries or households) is analyzed and summed up to estimate gas demand at regional levels or higher.

In this thesis a top-down approach is used, as the necessary statistics for a bottom-up approach are not readily available at a European, country-by-country level.

Hence, this project aims to fill the gas demand data gap by developing a model for the estimation of daily time series of gas demand from 2010 to 2019 in the statistical NUTS 3 regions of 27 selected European countries. Three consumer sectors will be considered (residential, industrial, and commercial) considering gas demand drivers such as population, gross domestic product, and number of households, among others.

The spatial and temporal disaggregation of gas consumption at high spatial resolution has not been approached for the entirety of Europe but for particular cases such as Germany, where an estimation of its gas demand can be found in the DemandRegio project [11], which studied the consumption of gas and electricity at district level taking a similar approach. The results of this thesis are validated against their findings.

2. Theory

In this chapter, a theoretical overview of spatial and temporal disaggregation will be presented. First, basic definitions and concepts will be introduced, followed by a review of the statistical data and literature used to develop this project.

2.1 Definitions and concepts

In this section, some important definitions regarding the topic of energy demand modeling, both general and specific to this project, are presented.

2.1.1 Energy use sectors

For purposes of energy consumption analysis, four sectors are generally considered: **Residential**, **Industrial**, **Commercial**, and **Transportation** [12]. The following definitions are taken from the U.S. Energy Information Administration (EIA) [13]:

Residential (Households): Energy-consuming sector that consists of living quarters for private households. Common uses of energy associated with this sector include space heating, water heating, air conditioning, lighting, refrigeration, cooking, and running a variety of other appliances.

- **Industrial:** Energy-consuming sector that consists of all facilities and equipment used for producing, processing, or assembling goods. The industrial sector encompasses the following types of activity: manufacturing; agriculture, forestry, fishing, and hunting; mining, including oil and gas extraction; electricity, gas, steam, and air conditioning supply; and construction. Overall energy use in this sector is largely for process heat and cooling and powering machinery, with lesser amounts used for facility heating, air conditioning, and lighting. Fossil fuels are also used as raw material inputs to manufactured products.

- **Commercial, Trade and Services (CTS):** Energy-consuming sector that consists of service-providing facilities and equipment of businesses; Federal, State, and local governments; and other private and public organizations, such as religious, social, or fraternal groups. The commercial sector includes institutional living quarters. It also includes sewage treatment facilities. Common uses of energy associated with this sector include space heating, water heating, air conditioning, lighting, refrigeration, cooking, and running a wide variety of other equipment.

- **Transportation:** Energy-consuming sector that consists of all vehicles whose primary purpose is transporting people and/or goods from one physical location to another. Included are automobiles, trucks, buses, motorcycles, trains, subways, rail vehicles, aircraft, ships, barges, and other waterborne vehicles.

In Europe, the amount of gas used for transportation purposes is negligible when compared to the other sectors (1.5% of the total gas consumption in the European Union in 2019 [14]). For this reason, only the **residential, industrial,** and **commercial** sectors were included. A complete definition of each sector is presented in their respective methodology section.

2.1.2 Energy demand determining variables

The modeling of energy demand requires data on the drivers of energy consumption, either at a spatial or a temporal level. Energy demand determining variables (EDDVs) can be either *external*, as in the case of weather or demographic data, or *descriptive*, taking into consideration technical aspects of the energy-consuming system such as thermal insulation or construction material of buildings. Table 1 presents examples of different energy demand determining variables used in the modeling of energy consumption [15, p. 4].

Model inputs	Energy demand determining variables
Historical energy demand	Historical electricity, heat, cooling or natural gas demand
Weather data	Outdoor temperature, cooling/heating degree days, solar irradiance, humidity, wind speed, air pressure
Calendar data	Workdays, holidays, working hours, seasons
Demographic or socioeconomic data	Population, GDP, number of households, number of employees, human development index (HDI)
Technical system information	Devices/Machinery: Efficiency, number of units, nominal power, flow rate. Buildings: Living space, building age, room temperature and humidity, window area, building material, efficiency class
Usage and behavioral data	Data obtained from time use surveys, building usage and occupancy, equipment usage time
Energy economics	Electricity and gas prices; tariffs

Table 1: Energy demand determining variables used in models [15, p. 4].

Some of the variables mentioned above were used for the spatial and temporal distribution of gas and will be listed in Section 2.2.2.

2.1.3 Spatial disaggregation

In the context of this project, the **spatial disaggregation** or **regionalization** of gas is the distribution of the gas supplied to individual European countries among smaller geographical divisions. Since European countries have a variety of administrative subdivisions, it is necessary to adopt a common standard. This is provided by the NUTS system, which will be applied in this project and explained next.

2.1.3.1 Nomenclature of Territorial Units for Statistics (NUTS)

The Nomenclature of Territorial Units for Statistics (NUTS) is a geographical system for the division of the European Union territory into hierarchical levels [16].

The three NUTS levels, respectively called 1, 2, and 3 from lowest to the highest resolution, are defined by minimum and maximum population thresholds, as shown in Table 2:

Level	Minimum inhabitants	Maximum inhabitants
NUTS 1	3 000 000	7 000 000
NUTS 2	800 000	3 000 000
NUTS 3	150 000	800 000

Table 2: Definition of NUTS levels [17].

An unofficial fourth level (NUTS 0) is used by some sources which refers to the national level.

The purpose of the NUTS classification system is:

- The collection, development, and harmonization of European regional statistics.
- Help with socio-economic analysis of regions.
- Framing of European Union policies.

A graphical representation of the NUTS levels can be seen in Figure 1:

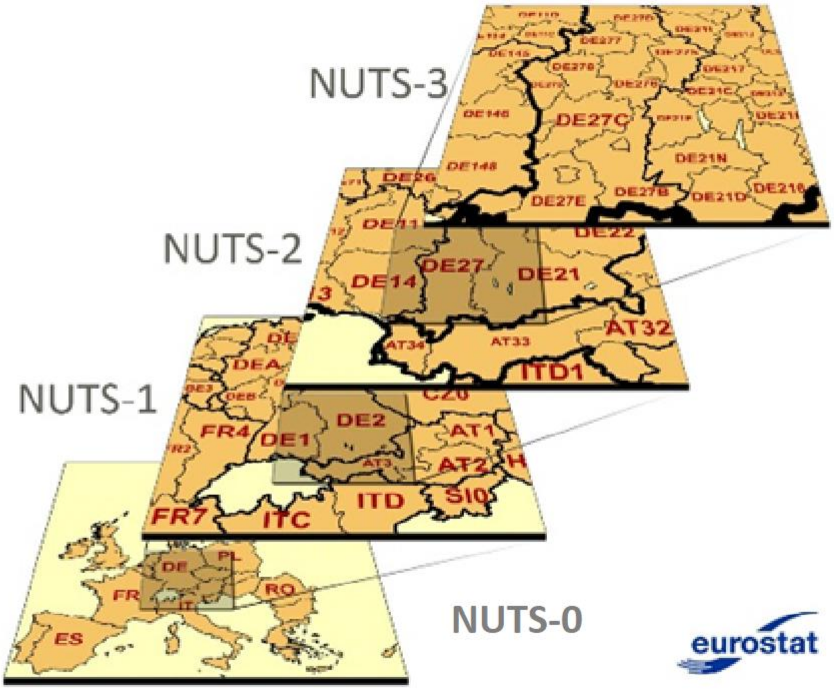


Figure 1: Graphical representation of NUTS regions. Taken from [16].

In most cases, NUTS levels correspond to the administrative division of member countries of the EU. Examples of this are given in Table 3:

Level	Germany	Belgium	Italy
NUTS 0	Country	Country	Country
NUTS 1	States *(Bundesland)*	Regions	Groups of regions
NUTS 2	Government regions *(Regierungbezirk)*	Provinces (+Brussels)	Regions
NUTS 3	Districts *(Kreis)*	Arrondissements	Provinces

Table 3: NUTS regions equivalents of selected countries [18].

While the NUTS classification aims to create comparable statistical areas, on some occasions regions at the same level may differ significantly in regard to their area or their importance from an administrative or an economic viewpoint. In some cases, one region can act as different NUTS levels simultaneously, as in the cases of Berlin and Hamburg (both acting as NUTS 1 and NUTS 2 regions), or Luxembourg, where all three NUTS levels correspond to the same region [19].

NUTS regions are referenced by unique codes. Their length depends on the hierarchical level, with NUTS 1, 2 and 3 regions having 3-, 4- and 5-characters codes, respectively. An example is given in Table 4:

Region	NUTS level	NUTS code
Germany	NUTS 0	DE
Niedersachsen	NUTS 1	DE9
Weser-Ems	NUTS 2	DE94
Oldenburg	NUTS 3	DE943

Table 4: Example of NUTS codes at different levels.

At the beginning of this project, the current and valid NUTS classification was the 2016 version, and hence it was used as reference for the naming and numbering of NUTS regions [20]. The NUTS 2016 Classification listed 104 regions at NUTS 1, 281 regions at NUTS 2, and 1348 regions at NUTS 3. However, as this project includes Norway, a member of the European Free Trade Association (EFTA) with a similar classification code, and excludes two European Union countries (Malta and Cyprus), the number of regions studied here is:

	NUTS 1	NUTS 2	NUTS 3
Count	103	286	1364

Table 5: NUTS regions count.

NUTS regions were modified with the 2021 Classification, which came into effect in January 2021.

2.1.4 Temporal disaggregation

In energy modeling, the **temporal disaggregation** of energy demand is the distribution of energy consumption over a smaller time interval. Temporal resolutions of energy system models vary from low (annual) to high (daily, hourly) and very high resolution (every minute/second) [15]. In the specific case of this project's gas demand modeling, a **daily** resolution has been chosen for the estimation of a time series between the years 2010 to 2019.

2.2 Statistical data review

In this section, relevant statistical datasets regarding European energy consumption and energy demand determining variables will be reviewed.

2.2.1 Energy consumption datasets

Energy consumption statistics are usually presented as annual values for entire countries and are used in different ways in demand modeling. In a top-down approach, they constitute the main input value to be disaggregated, while in a bottom-up methodology they provide a value to reach after summing up specific consumptions [15, p. 26].

2.2.1.1 European Statistical Office Energy Balances

The European Statistical Office (Eurostat) collects and publishes monthly data on energy supply, transformation, and consumption of different fossil fuels such as gas, coal, oil, as well as renewables (wind, hydro, solar, etc.) [21]. Of particular importance is the **Complete energy balances** dataset [14], which contains information regarding annual energy consumption (with gas identified with the code G3000) for the sectors mentioned in Section 2.1.1. The Complete energy balances also provide energy consumption figures for different types of energy-intensive industries (chemical, food, iron, etc.) that are helpful in determining the regional energy distribution in combination with other sources of information.

Eurostat also provides annual figures of energy consumption for different residential applications with the dataset **Disaggregated final energy consumption in households – quantities** [22]. As it will be shown in Section 3.1, this is particularly important given that different energy uses in households are associated with different energy demand determining variables.

2.2.2 Energy demand determining variables datasets

As mentioned in Section 2.1.2, energy demand determining variables (EDDV) can be classified in different types. This section lists the datasets used for the spatial and temporal distribution of gas consumption.

2.2.2.1 EDDV datasets used for spatial distribution

For the spatial distribution of gas demand, all the energy demand determining variables used belonged to the demographic/socioeconomic group. Their corresponding datasets are listed in Table 6 and Table 7.

Sector	EDDV	Dataset	Spatial/ Temporal resolution	Years available	Unit
Residential	**Population**	Population on 1 January by age group, sex, and NUTS 3 region [23]	NUTS 3 / Annual	2014-2020	[Number of persons]
	Number of households	Number of households by degree of urbanization and NUTS 2 regions [24]	NUTS 2 / Annual	1999-2020	[Thousands of households]
	Dwelling area	Average size of dwelling by household type and degree of urbanization [25]	NUTS 0 / Annual	2012	[m^2]
	NUTS typology	Urban-Rural Typology [26]	NUTS 3 / Annual	2016	[-]

Table 6: List of datasets used for spatial disaggregation of the residential sector.

Sector		EDDV	Dataset	Spatial/ Temporal resolution	Years available	Unit
Industrial	Commercial	Number of employees	Employment by age, economic activity, and NUTS 2 regions (NACE Rev. 2) [27][1]	NUTS 2 / Annual	2008-2020	[Thousands of employees]
			Statistical business structure data by NUTS 2 regions and NACE Rev. 2 [28][2]	NUTS 2 / Annual	2008-2018	[Persons employed – Number]
			Employment (thousand persons) by NUTS 3 regions [29][3]	NUTS 3 / Annual	1995-2019	[Thousands of employees]
		Gross Domestic Product (GDP)	Gross domestic product (GDP) at current market prices by NUTS 3 regions [30]	NUTS 3 / Annual	2000-2019	[Millions of Euros]

Table 7: List of datasets used for spatial disaggregation of the Industry and CTS sectors.

For consistency, the source for all datasets utilized for gas spatial distribution is the European Statistical Office (Eurostat), which gathers information supplied by EU members' statistical agencies [31]. However, as the amount and quality of data differs from country to country, missing values in the datasets are not uncommon. Therefore, missing values have been interpolated or extrapolates depending on

[1] Used for the CTS sector
[2] Used for the industry sector
[3] Also used for the industry sector

the completeness of the dataset. The following methods have been used for this purpose: mean value, median value, linear, quadratic, or cubic spline extrapolation.

The use of the different EDDV datasets mentioned above will be explained in Section 3.

2.2.2.2 EDDV datasets used for temporal distribution

For the temporal distribution of gas demand, two datasets were used: the **surface air temperature** [32], and the **calendar day type.**

The **surface air temperature**, also called **2 m temperature,** is defined as the temperature of air at 2 meters above the surface of land, sea, or inland waters [33]. It is used not only to help determine the current temperature in a region, but also for temperature forecasting [34].

The source for the surface air temperature dataset is the European Centre for Medium-Range Weather Forecasts Reanalysis version 5 (ERA5), produced by the Copernicus Climate Change Service (C3S). ERA5 provides hourly data from 1950 to the present on atmospheric, land surface, and sea-state parameters, and it is available on a regular 0.25°x0.25° grid.

The 2-m-temperature dataset of ERA5 was used to assign mean daily temperatures to NUTS regions from 2010 to 2019. A thorough explanation will be presented in Section 3.1.3.

The **calendar day type** dataset was created using the Python package **Holidays** [35] to retrieve the public holidays in the 27 European countries from 2010 to 2019. The dataset differentiates between these days, *workdays*, and *weekend days*. Such distinction is useful to determine gas consumption in the commercial and industrial sectors, as it will be seen in Section 3.2.4.

2.3 Literature overview

As mentioned in Section 1, the increased use of P2G systems has made necessary a more accurate model of the European gas infrastructure, which is currently under development by the SciGRID_gas project. An important aspect of such models is energy demand, which is not only relevant as a tool for the design and planning of energy systems, but as a research topic itself.

The importance of energy demand can be appreciated in the increasing amount of research papers on the topic. In this section, a short selection of publications relevant for the development of this thesis are reviewed.

2.3.1 DemandRegio

The DemandRegio project [15], developed between 2017 and 2020 by the Jülich Research Center, the Technical University of Berlin, and the Research Center for Energy Economics (FfE), with funding from the German Federal Ministry for Economic Affairs and Energy (BMWi), is one of the closest precedents to the topic of this thesis. Its goal was to determine the regional distribution of gas and electricity consumption in Germany with a NUTS 3/district resolution.

The DemandRegio team performed an extensive search of energy demand modelling literature; of the search results, the project selected 72 articles, which they analyzed.

According to their research, modeling can be performed by a variety of methods, depending on its application or purpose. These methods can be classified in five categories:

- **Statistical:** Statistical modeling applies statistical analysis to a dataset, identifying relationships between variables, make predictions, and present information in an easier-to-visualize way [36]. These include regression models (using algorithms such as the ordinary least squares (OLS) method to optimize regression parameters) and time series analysis (TSA) models. Ghalehkhondabi et al. [37, pp. 12,18] give examples of TSA models such as ARMA (autoregressive moving average) and its extensions ARIMA, SARIMA, and ARMAX, in which data is forecasted using data from a historical time series of energy consumption.

- **Machine learning:** Machine learning (ML) algorithms are part of artificial intelligence and improve the accuracy of models automatically by finding and applying patterns in data [38]. Two categories are observed here: algorithms with supervised and without supervised learning approach.
- **Metaheuristic:** Debnath et al. [39, p. 306] mention metaheuristic methods that use mechanisms inspired by natural biological evolution in the context of hybrid models. These mechanisms include reproduction, mutation, recombination and selection and the metaheuristic algorithm may be implemented in ML.
- **Uncertainty methods:** Uncertainty methods are used for forecasting energy consumption with the aid of fuzzy logic. They are effective with incomplete or limited datasets [39, p. 306].
- **Engineering methods:** Engineering methods are used to represent energy-consuming units from a bottom-up approach using input-output relationships based in the laws of physics [15, p. 6]. These methods allow for high flexibility in modeling and high level of detail and can be combined with other approaches.

DemandRegio's research also shows some general features:

- While different countries are the object of energy demand studies, some of them, like Australia, have a larger representation, due to more open policies regarding data restrictions.
- Most of the articles are not limited to single consumers but to consumer groups.
- Most articles are limited to a single energy source.
- Most of the studies use historic energy demand as EDDVs, along with weather, calendar and demographic information.
- In terms of temporal resolution, there is a tendency of using hourly time steps or shorter. This depends on the energy source used; in the case of gas, studies do not use time steps shorter than hourly, reflecting the infrastructure used for measuring consumption.
- Spatial resolution is modelled in a relatively small geographical scale, either on the level of individual consumers, buildings or devices, or at regional level within countries.

The DemandRegio project analyzed the energy demand of three sectors: residential, industrial, and commercial. In the case of gas, the methodology used for regionalization is described next:

- Residential sector: The spatial disaggregation of the residential sector was performed using both a top-down and a bottom-up approach. In the top-down approach, three uses of gas were considered, with their respective EDDVs in parentheses: space heating (living space to be heated), water heating (number of persons), and cooking (number of households). The bottom-up approach was used only for space heating considering aspects as the age of buildings, heating degree days, and efficiency of the boilers. A comparison of the results obtained from both methods shows that the difference is negligibly small.
- Industrial and commercial sector: For these sectors, the selected EDDV was the number of employees liable to social insurance. This information is supplied by the German Federal Employment Agency. Other datasets used include the Annual survey on the use of energy in industry (JEVI), the Energy balance of the Federal Republic of Germany (AGEB), and the Final energy consumption by production area (UGR). The input datasets consider the consumer groups (e.g., the type of industry).

For the temporal disaggregation, DemandRegio used standard load profiles. In the residential sector, these load profiles made a distinction between single- and multiple-family homes and were developed using the procedures indicated by the Federal Association of German Gas and Water Management (BGW) with an hourly resolution. In the industrial and commercial sector, two approaches are used. The industrial sector uses generic load profiles with quarterly-hour resolution based on working shifts and typical working hours. The commercial sector uses standard load profiles in hourly resolution, where weekday and outdoor temperature are taking into account.

2.3.2 Other publications

As mentioned in Section 2.1.1, energy modeling usually considers four consumer sectors: residential, industrial, commercial, and transport, the first three being the most commonly analyzed. In the specific case of gas, the LKD-EU study, by Kunz et al. neglects the study of the transport sector due to its low share of gas demand [40]. This can also be observed in an overview of selected studies by

DemandRegio. In the industry sector, Kunz et al. divide industry into eight energy-intensive activities: steel, aluminum, ceramics, glass, paper, food, tobacco, and chlorine. The spatial disaggregation is performed using the number of companies as demand driver. In the residential sector, gas is used for heating purposes. Apart from the number of households, Kunz et al. also take into consideration building structures and heat technologies.

In his analysis of the LKD-EU study, Neuland [41] mentions several limitations of the gas demand model, among them, the lack of representation of public buildings (commercial sector), the neglection of industrial processes using gas as an input, and the use of gross value added (GVA) for the distribution of gas demand of non-energy-intensive industry, which can cause an overrepresentation of urban areas.

In the residential sector, Heitkoetter et al. [42] characterize heat demand in Germany by taking into consideration six building features with three types each: construction year, flat area, number of flats, number of residents, building type, and heating type. Using data from the 2011 census, these building features were used to define 729 building categories. Census data was then also used to obtain the number of buildings per building category at district level. Using data regarding the final energy consumption for space and water heating of more than 500 000 buildings, they analyzed the influence of year of construction and number of households per building.

Having calculated the area-specific heat demand, and by multiplying it by the average floor area, Heitkoetter et al. obtained the absolute annual heat demand of the aforementioned building categories. As this information included both water and space heating, demand of the former was obtained by dividing its final energy demand by the number of inhabitants in Germany. This figure was then multiplied by the average number of residents per household and the number of households per building. The heat demand per building obtained was then subtracted from the total heat demand, to obtain the demand for space heating.

In a separate publication [43], Heitkoetter et al. mentions the use of standard load profiles derived by the German Association of Energy and Water Industries (BDEW) as a commonly used reference for both the residential and commercial sectors. These profiles, in particular the *H0 profile*, were also used by the DemandRegio project.

While the work of Schmidt [44] does not allude directly to gas, it uses four methods for the estimation of NUTS 3 annual energy (electricity) demand in the industry sector in Germany. The first method, plant specific characterization, is done by gathering information regarding specific site production (by identifying individual industrial plants, their locations, and their production capacity) using public sources, such as the records from companies members of national industry associations; this specificity and high-detail data makes it the most reliable method, with the drawback of being time consuming. The other three methods, allocation of NUTS 0 production, allocation of NUTS 0 demand, and NUTS 3 total energy demand method, rely on statistical data.

In the allocation of NUTS 0 production method, Schmidt uses production data obtained from national industrial associations and industry branch employment data to estimate the branch specific production per employee. Then, using NUTS 3 figures for branch employment, she finds the annual production at NUTS 3 level. By combining this annual production with the specific energy demand for production, the annual energy demand for specific applications at NUTS 3 level is found.

The allocation of NUTS 0 energy demand takes a similar approach, using NUTS 0 self-generated electricity, branch consumption, and branch employment figures. With these datasets, the energy demand per employee at NUTS 0 level is found. As on the NUTS 0 production method, branch employment at NUTS 3 level data is then used to find the branch energy demand. Combined with the application specific annual energy demand determined from specific characterization methods at NUTS 3 level, the percentage of the application share of relevant branches annual energy demand is found, finally obtaining the application specific annual energy demand at NUTS 3 level.

The last method, total industrial energy demand, uses a similar approach to the previous ones, using employment data. As this method deals with the specific topic of her work, industrial load shifting, it is not considered relevant for the development of this thesis.

3. Methodology

As previously mentioned in Sections 1 and 2, European gas consumption data is currently available only on a national scale. However, the aim of this thesis is to generate a time series on a higher resolution, e.g., NUTS 3 level. This section will present the methodology used for the generation of such gas demand time series from 2010 to 2019 through means of spatial and temporal disaggregation of the input data and other assumptions.

This thesis is outlined by the following items:

- **Temporal extent:** 2010 – 2019.
- **Temporal resolution:** Daily.
- **Spatial extent:** 27 European countries named *EUR_27* in the context of this project (Austria [AT], Belgium [BE], Bulgaria [BG], Croatia [HR], Czechia [CZ], Denmark [DK], Estonia [EE], Finland [FI], France [FR], Germany [DE], Greece [EL], Hungary [HU], Ireland [IE], Italy [IT], Latvia [LV], Lithuania [LT], Luxembourg [LU], Netherlands [NL], Norway [NO], Poland [PL], Portugal [PT], Romania [RO], Slovenia [SI], Slovakia [SK], Spain [ES], Sweden [SE], United Kingdom [UK]).
- **Spatial resolution:** NUTS 3.
- **Consumer groups:**
  - Residential (Households)
  - Industrial
  - Commercial, trade, and services (CTS)

A flowchart of the methodology is shown in Figure 2 and Figure 3.

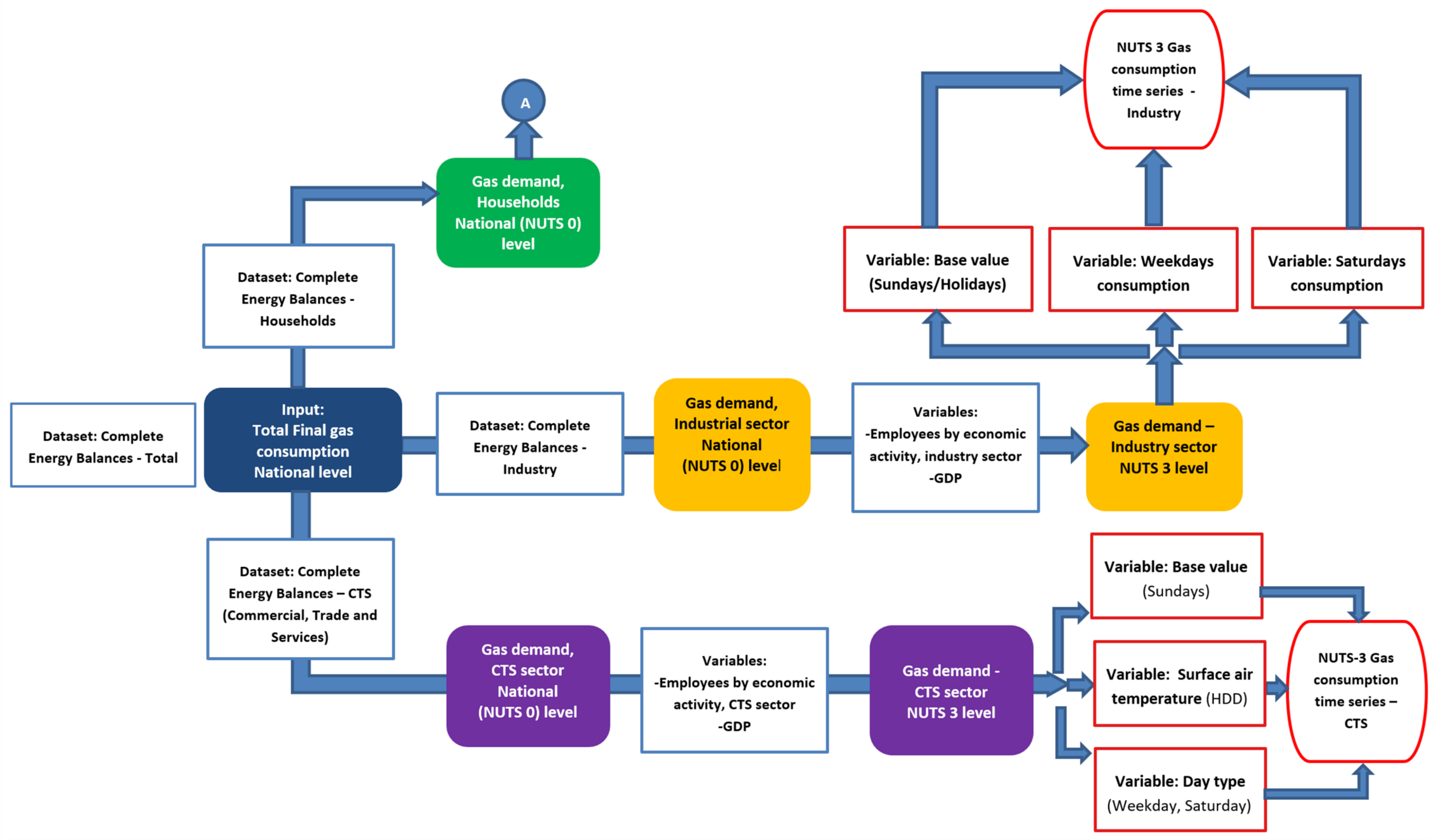


Figure 2: Methodology flowchart (Industrial and commercial sectors).

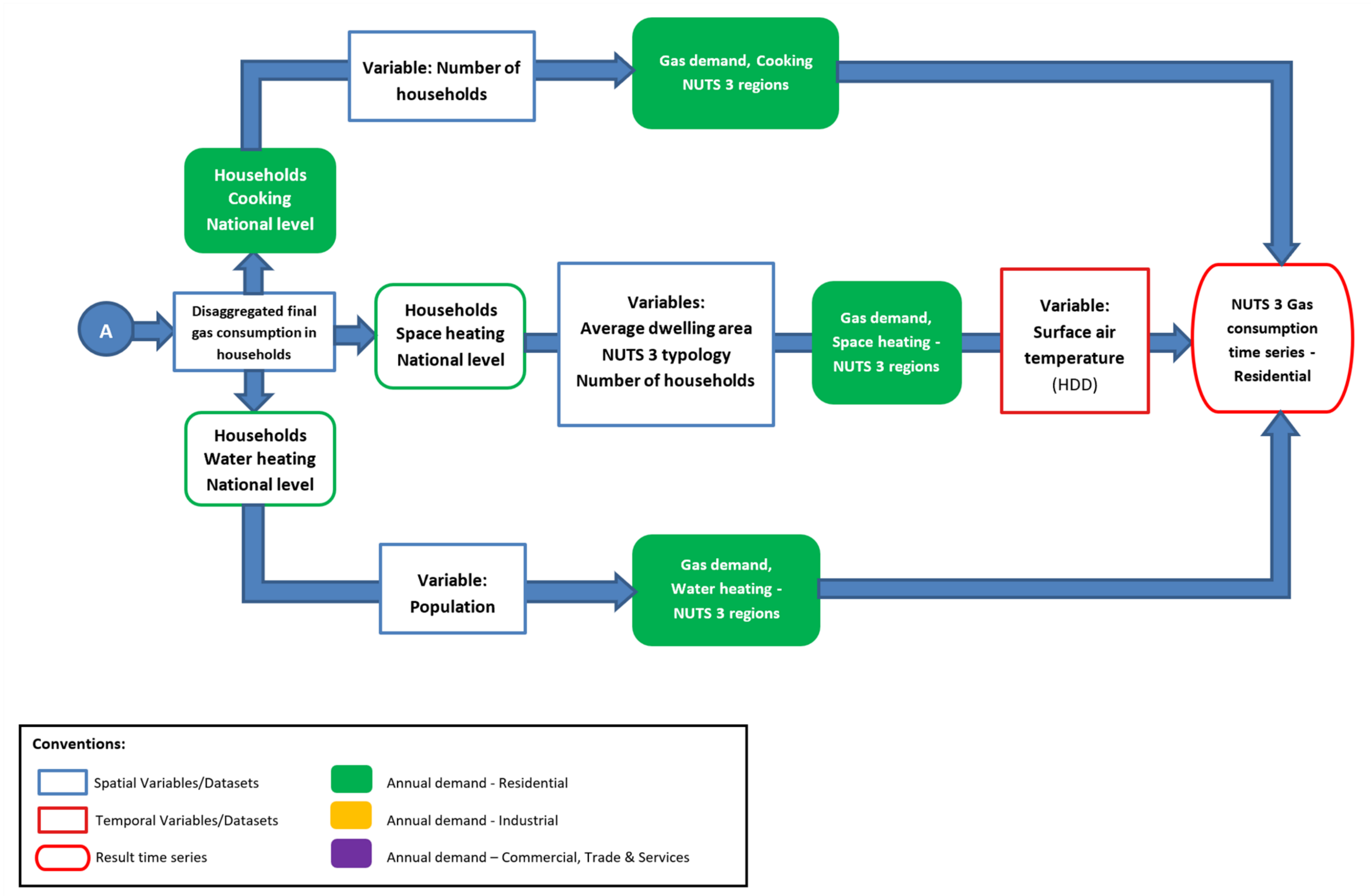


Figure 3: Methodology flowchart (Residential sector)

First, the different consumer groups will be defined, followed by a description of the demand drivers and methods used to spatially disaggregate the gas demand in NUTS 3 regions. Finally, the methodology used for the temporal disaggregation leading to the estimation of a daily household gas demand time series will be presented.

3.1 Residential

3.1.1 Definition of household consumers

In the context of this research, a *household* is defined by the European Statistical Office as a housekeeping unit having common arrangements and sharing expenses as well as a common residence. Households include one or more people, not necessarily related, living at the same address. [45]

Eurostat's household definition includes a reference to collective or institutional households (e.g., hospitals, prisons, boarding houses, etc.) These facilities will be included in the commercial sector. Hence, whenever used in this project, the word *households* will refer only to ***private households***.

3.1.2 Description of gas demand in households

Among the three sectors analyzed in this project, private households consistently present the highest gas consumption during the 10-year scope [46]. Figure 4 shows the final energy consumption of gas by sector in European countries between 2010 and 2019. It can be observed that residential gas consumption ranges between 1080 and 1422 TWh (approximately 43 to 48 percent of the total gas consumption in Europe).

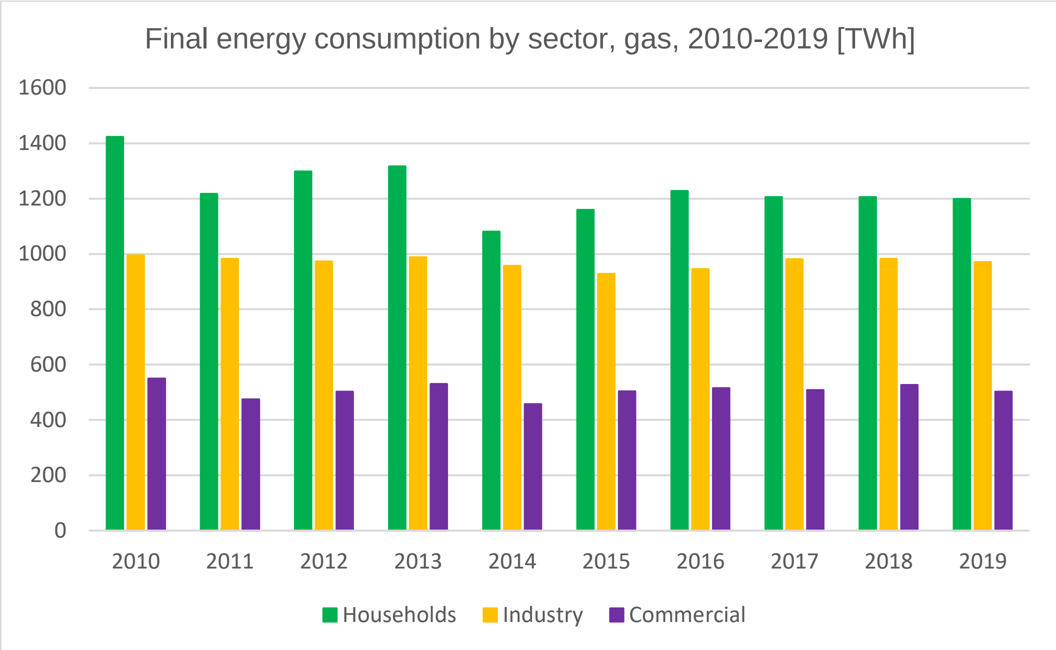


Figure 4: Final energy gas consumption in the European Union by sector, 2010-2019 [47].

Gas consumption in households is based on three main uses, all of them related to heating: **space heating**, **water heating,** and **cooking** [46]. In the European Union, the consumption of gas in households in 2018 amounted to 4.3 million TJ. Figure 5 shows the percentage distribution for these uses.

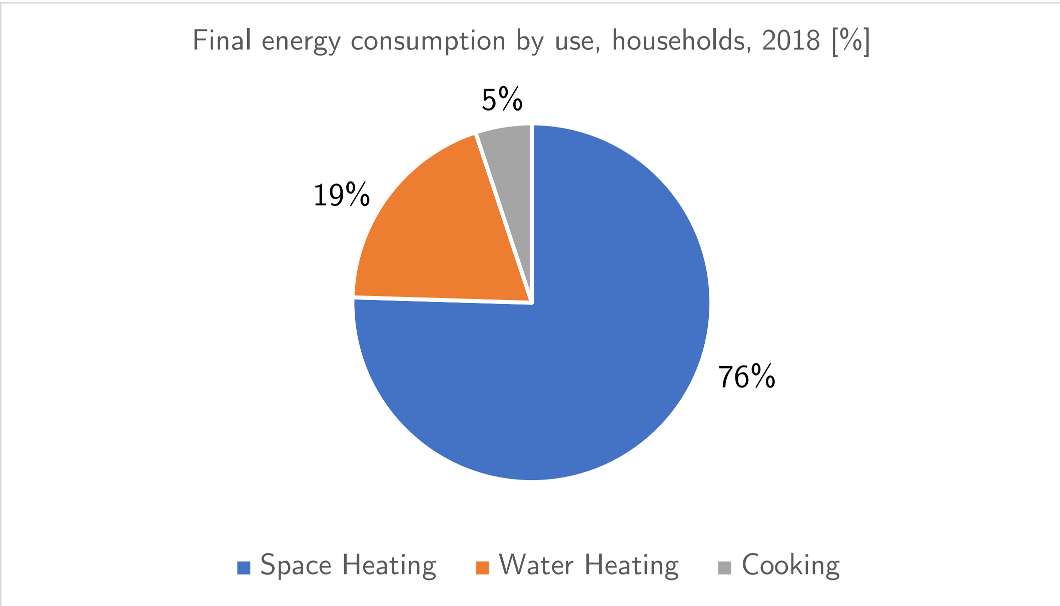


Figure 5: Final energy consumption in households by use in the European Union - Gas, 2018 [22].

According to Eurostat, "In the EU, the main use of energy by households is for home heating (63.6% of final energy consumption in the residential sector)" [46]. Figure 5 shows that in 2018 the highest use of gas in households in Europe was for **space heating**, with 76%, followed by 19% and 5% for water heating and cooking, respectively. This behavior was also observed over all other years of the period analyzed.

Gas is widely utilized in Europe for the three household applications mentioned above and it is the main energy source for two of them (space and water heating). However, when comparing country by county, there are regional differences in the use of gas in households. Table 8 shows that in 2018 gas was predominantly used for space heating in the following ten countries: Germany, France, Croatia, Italy, Luxembourg, Hungary, Netherlands, Romania, Slovakia, and the United Kingdom. On the other side, Portugal, Sweden, Finland, Latvia, Estonia, and Bulgaria, where renewable energies and derived heat play a significant role in space heating, had the lowest share in gas use for this purpose.

Share of fuels in the final energy consumption in the residential sector for space heating, 2018 [%]						
	Gas	Electricity	Derived Heat	Solid fuels[4]	Oil & petroleum products	Renewables and wastes[5]
Germany	48.1	1.8	9.1	1.3	23.6	16.1
France	34.8	12.7	4.0	0.1	14.7	33.8
Croatia	21.9	1.8	6.5	0.1	4.7	65.0
Italy	58.4	0.4	5.1	0.0	7.8	28.3
Luxembourg	54.6	4.2	0.0	0.1	34.0	7.2
Hungary	56.3	0.8	8.3	2.3	0.2	32.0
Netherlands	86.2	2.1	3.5	0.0	0.6	7.6
Romania	30.2	0.2	16.0	0.5	0.0	53.1
Slovakia	65.4	5.7	25.0	1.2	0.3	2.4
United Kingdom	74.7	6.4	1.1	2.0	7.8	7.9
Belgium	47.1	3.2	0.3	1.2	37.9	10.4
Czechia	24.8	4.7	13.5	18.0	0.8	38.1
Denmark	15.5	2.8	36.6	0.0	4.1	41.0
Ireland	25.6	4.0	0.0	17.6	50.1	2.6
Greece	13.6	4.8	2.2	0.2	43.0	36.2
Spain	19.5	7.3	0.0	0.9	31.9	40.5
Lithuania	10.3	1.3	38.4	5.2	1.8	42.9
Austria	26.2	4.8	14.5	0.4	18.6	35.5
Poland	14.9	0.9	20.1	44.9	0.8	18.5
Slovenia	12.5	3.9	9.4	0.0	15.0	59.2
Norway	0.0	63.4	2.9	0.0	1.8	31.9
Portugal	1.0	12.5	0.1	0.0	5.5	80.9
Sweden	0.6	29.3	48.7	0.0	4.0	17.4
Finland	0.5	24.2	29.9	0.1	5.9	39.4
Latvia	7.7	0.7	33.3	0.7	3.0	54.6
Estonia	5.9	5.1	37.4	0.2	0.3	51.2
Bulgaria	5.3	8.8	16.7	9.8	0.1	59.3

Table 8: Share of fuels in the final energy consumption in the residential sector for space heating, 2018 [46].

[4] Hard coal, brown coal, and coal products.
[5] Hydro power, geothermal, solar, wind, biofuels, etc.

The main sources for water heating in Europe are electricity and gas. As Table 9 shows for 2018, gas prevailed over other energy sources in Germany, Spain, Italy, Luxembourg, Netherlands, Romania, Slovakia, and the United Kingdom.

Share of fuels in the final energy consumption in the residential sector for water heating, 2018 [%]						
	Gas	**Electricity**	**Derived Heat**	**Solid fuels**	**Oil & petroleum products**	**Renewables and wastes**
Germany	49.3	14.3	3.5	0.0	16.2	16.5
Spain	45.6	18.7	0.0	0.2	23.9	11.6
Italy	65	13.4	4.0	0.0	8	9.5
Luxembourg	56.3	4	0.0	0.0	30.6	9.2
Netherlands	89.2	4	5.4	0.0	0	1.4
Romania	54.5	2	0.0	0.7	5.9	36.8
Slovakia	47.9	20.8	25.5	2.5	1	2.4
United Kingdom	84.4	6.4	0.0	0.3	7.1	1.8
Belgium	48.9	30.4	0.0	0.0	17.3	3.5
Bulgaria	2.7	58.7	31.5	0.5	0.1	6.5
Czechia	36	23	27.3	2.2	0	11.5
Denmark	17.2	4.6	63.2	0.0	8.4	6.5
Estonia	6.4	5.5	59.7	0.1	0.2	28.0
Ireland	29.7	15.5	0.0	5.5	44.5	4.7
Greece	6.8	39.4	0.8	0.0	7.1	45.9
France	27.1	50	6.5	0.0	12.3	4.1
Croatia	37.6	44.5	3.8	0.1	3.4	10.7
Latvia	9.9	12.6	44.5	0.4	3.7	29.0
Lithuania	13.5	9.7	51.8	2.6	4.2	18.2
Hungary	38.4	39.5	16.1	0.0	1.4	4.5
Austria	17.6	28.5	11.7	0.1	10.3	31.8
Poland	29.4	5.7	37.1	17.9	1.2	8.8
Portugal	33.2	5.4	0.0	0.0	42.2	19.2
Slovenia	12.5	29.5	7.5	0.0	9.1	41.5
Finland	0.6	25	55.0	0.1	6.5	12.8
Sweden	0.6	32.8	54.7	0.0	3.2	8.6
Norway	0	96.3	3.7	0.0	0	0.0

Table 9: Share of fuels in the final energy consumption in the residential sector for water heating, 2018.

For cooking purposes, electricity is the main energy source across Europe, followed by gas, and far above other options. According to Table 10, the most prominent use of gas for cooking in 2018 was observed in Italy, with other countries, like Lithuania, Hungary, Netherlands, Poland, Romania, and Slovakia also using gas as their main fuel of choice.

Share of fuels in the final energy consumption in the residential sector for cooking, 2018 [%]						
	Gas	**Electricity**	**Derived Heat**	**Solid fuels**	**Oil & petroleum products**	**Renewables and wastes**
Italy	70.3	15.2	0.0	0.0	10.0	4.5
Lithuania	40.1	15.8	0.0	0.0	34.3	9.1
Hungary	67.9	12.4	0.0	0.0	19.5	0.2
Netherlands	66.2	33.8	0.0	0.0	0.0	0.0
Poland	44.2	16.9	0.0	4.9	30.7	3.3
Romania	62.0	0.1	0.0	0.4	31.3	6.3
Slovakia	69.6	24.4	0.0	3.2	0.5	2.3
Belgium	32.8	61.0	0.0	0.0	6.2	0.0
Bulgaria	2.2	74.2	0.0	0.8	10.9	11.9
Czechia	49.0	45.9	0.0	0.2	2.4	2.5
Denmark	3.1	96.9	0.0	0.0	0.0	0.0
Germany	2.6	92.6	0.0	0.0	0.0	4.9
Estonia	16.7	41.8	0.0	0.0	13.8	27.8
Ireland	22.4	76.1	0.0	0.0	1.5	0.0
Greece	0.9	70.0	0.0	0.0	24.9	4.3
Spain	26.7	53.6	0.0	0.7	16.6	2.4
France	30.3	49.5	0.0	0.0	20.2	0.0
Croatia	24.8	36.4	0.0	0.0	25.2	13.5
Latvia	39.6	14.4	0.0	0.7	15.3	30.6
Luxembourg	53.6	44.7	0.0	0.0	1.1	0.0
Austria	4.8	91.0	0.0	0.0	0.3	3.9
Portugal	9.2	44.4	0.0	0.0	15.6	30.8
Slovenia	15.9	49.9	0.0	0.0	22.9	11.2
Finland	0.8	91.8	0.0	0.0	7.4	0.0
Sweden	1.5	98.5	0.0	0.0	0.0	0.0
United Kingdom	53.8	46.2	0.0	0.0	0.0	0.0
Norway	0.0	100.0	0.0	0.0	0.0	0.0

Table 10: Share of fuels in the final energy consumption in the residential sector for cooking, 2018 [46].

3.1.3 Spatial disaggregation of the residential sector

For all three of the considered sectors in the project, the spatial disaggregation of gas consumption was performed using a **top-down approach.**

For this approach, gas demand was distributed starting at a country level, using data obtained from Eurostat's energy balance [14], and ending at the highest spatial resolution level for statistical purposes in the European Union, NUTS 3.

As mentioned in the Section 3.1.2, the residential sector uses gas for **space heating**, **water heating**, and **cooking.** For the disaggregation of gas consumption in households, each use was separately analyzed as they present different demand drivers.

A sample of the input datasets showing the final energy consumption in European countries in the year 2018, disaggregated by use in households, can be observed in Table 11:

Disaggregated final energy consumption in households, 2018 [TJ]			
	Space heating	**Water heating**	**Cooking**
Belgium	117 494	19 837	1 611
Bulgaria	2 608	450	176
Czechia	51 038	18 432	9 192
Denmark	18 615	7040	96
Germany	735 319	192 979	3 891
Estonia	1 686	297	318
Ireland	18 042	6 595	602
Greece	12 094	1 689	88
Spain	53 137	49 087	12 505
France	399 610	54 214	29 682
Croatia	14 374	3 637	1 549
Italy	521 182	107 680	61 934
Latvia	2 601	941	1 446
Lithuania	4 545	745	1 608
Luxembourg	9 043	901	289
Hungary	98 405	12 010	8 094
Netherlands	220 616	60 135	5 633
Austria	49 305	6 626	349
Poland	79 678	39 502	29 929
Portugal	357	7 063	4 013
Romania	61 611	23 936	19 700
Slovenia	3 409	1 044	193
Slovakia	37 670	5 755	3 419
Finland	795	212	18
Sweden	1 035	259	74
Norway	31	0	0
United Kingdom	756 542	221 203	23 966

Table 11: Disaggregated energy consumption in households, 2018 [22].

The demand drivers, the datasets, and the procedure used for the spatial disaggregation of household gas demand applications are described next.

3.1.3.1 Water heating

Hot water in households is mainly used for showering/bathing and washing activities, making it highly dependent on the number of inhabitants per region. Thus, the amount of gas used for water heating correlates with the **population.** The average daily consumption of water per person in European households amounts to 144 liters [48], with Italy, Bulgaria, and Greece presenting the highest values [49]. The dataset used for spatial disaggregation of gas used in water heating was *Population on 1 January by age group, sex, and NUTS 3 region* [23].

The procedure used for the spatial disaggregation was similar for the three gas applications. First, distribution keys (i.e., ratios) were calculated using the datasets corresponding to each demand driver. Then, these distribution keys were multiplied by the gas consumption at a national level to estimate regional gas consumption.

Given the population at NUTS 3 regions, the **spatial distribution keys for water heating** $D_{\mathrm{pop}}^{r\_R}$ were obtained through:

$$D_{\mathrm{pop}}^{r\_R} = \frac{\mathrm{pop}_r}{\sum \mathrm{pop}_r} \qquad \textit{(Eq. 1)}$$

where pop_r is the population in a NUTS 3 region r belonging to a country R in the set of analyzed countries EUR_27.

The distribution keys obtained were then multiplied by the final gas consumption in households for water heating at a national level, C_{wh}^{R}, allowing an estimation of the final gas consumption in NUTS 3 regions:

$$C_{\mathrm{wh}}^{r} = C_{\mathrm{wh}}^{R} \cdot D_{\mathrm{pop}}^{r\_R} \qquad \textit{(Eq. 2)}$$

3.1.3.2 Cooking:

Cooking is an activity associated with an entire household, and the amount of gas used for this purpose is marginally dependent on its number of residents [48]. Because of this, the demand driver for cooking was only the **number of households**.

The dataset used for the spatial disaggregation was Eurostat's *Number of households by degree of urbanization and NUTS 2 regions* [24]. As this dataset contained only the number of households up to NUTS 2 regions, it was necessary to use the **population** dataset [23] to estimate the number of households at NUTS 3 level. To achieve this, the ratios of inhabitants of the NUTS 3 regions belonging to each NUTS 2 region were calculated. Then, each of these ratios was multiplied by the number of households in NUTS 2 regions, obtaining an approximate number of households at NUTS 3.

Having estimated the number of households in European NUTS 3 regions (2010-2019), the **spatial distribution keys for cooking** $D_{\mathrm{h}}^{r\_R}$ for each NUTS 3 region were calculated by:

$$D_{\mathrm{h}}^{r\_R} = \frac{\mathrm{h}_r}{\sum \mathrm{h}_r} \qquad \textit{(Eq. 3)}$$

where h_r is the number of households in a NUTS 3 region r belonging to a country R in the set of analyzed countries EUR_27.

Gas demand was then estimated in NUTS 3 regions multiplying C_{ck}^{R}, the final gas consumption for cooking in a country R, by the regional distribution keys:

$$C_{\mathrm{ck}}^{r} = C_{\mathrm{ck}}^{R} \cdot D_{\mathrm{h}}^{r\_R} \qquad \textit{(Eq. 4)}$$

3.1.3.3 Space heating

The amount of gas needed for space heating presents a dependency on aspects such as the **household area,** the **thermal insulation of buildings,** and the **efficiency of boilers** [15]. For this project's scope, and due to an uneven availability of European statistics regarding the two latter items, the **household area** was the only demand variable considered for the spatial disaggregation of space heating. To estimate the approximate dwelling area in NUTS 3 regions, three datasets were used:

- **Territorial typology of NUTS 3 regions** from the NUTS 2016 classification [26]: Classifies NUTS 3 regions depending on their typology into Primarily Urban (PU), Primarily rural (PR), or Intermediate (IN).
- **Average dwelling area at NUTS 0 level by typology in 2012** [25]**:** Presents the average dwelling area of Primarily Urban, Primarily Rural, and Intermediate regions in a country.
- **Number of households at NUTS 3 level:** Dataset previously used for the spatial disaggregation of cooking.

The average dwelling surface of the different countries was assigned to their corresponding NUTS 3 regions depending on their topology (PU, PR, or IN), and it was then multiplied by their number of households to obtain an approximate household area of the different NUTS 3 regions.

Having estimated this value, the calculation of the spatial distribution keys for space heating $D_{\mathrm{a}}^{r\_R}$ for the individual NUTS 3 regions *r* in a country *R* was then performed in a similar way than for cooking and water heating, and given by:

$$D_{\mathrm{a}}^{r\_R} = \frac{\mathrm{a}_r}{\sum \mathrm{a}_r} \qquad \textit{(Eq. 5)}$$

Where a_r is the average dwelling area in a region r belonging to a country *R* of the 27 European countries studied.

The allocation of gas demand for space heating at NUTS 3 regions was then estimated similarly as for the previous applications:

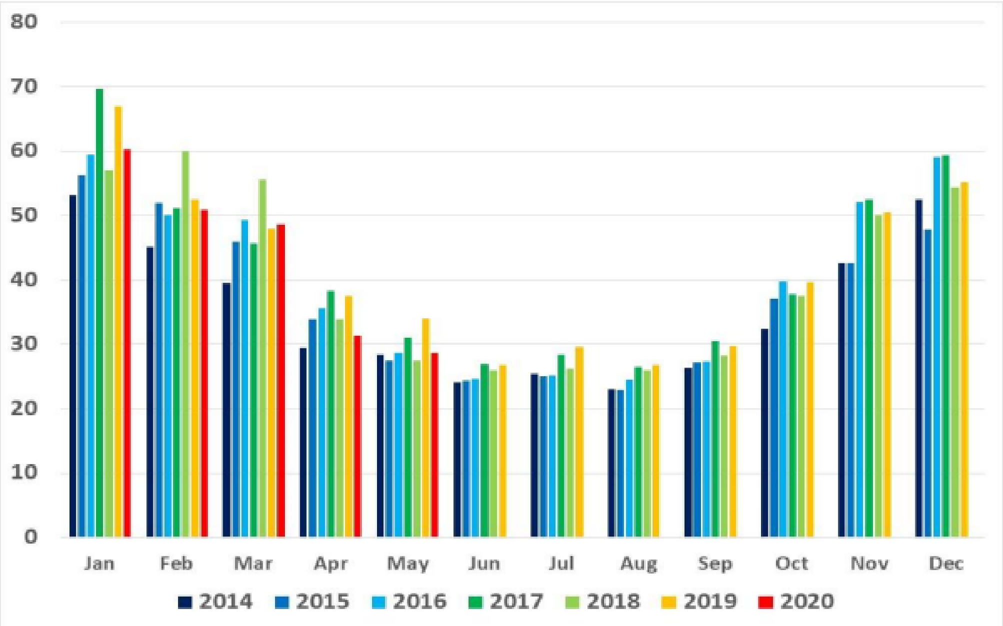
$$C_{\mathrm{sh}}^{r} = C_{\mathrm{sh}}^{R} \cdot D_{\mathrm{a}}^{r\_R} \qquad \textit{(Eq. 6)}$$

3.1.4 Temporal disaggregation of the residential sector

The temporal disaggregation of gas demand in the residential sector was performed by considering the three aforementioned applications separately, as described in the following sections.

3.1.4.1 Space heating

The amount of gas used for space heating in buildings presents an inverse dependency on the **surface air temperature** of the region considered [50] [51]**.** In Europe, this leads to a seasonal effect, with the highest gas consumption occurring during the winter months and the lowest during summer, as it can be observed in Figure 6:



Sources: IEA, Eurostat, Entsog, GRTgaz, Terega, NCG, Gaspool, SNAM, Enagas, NationalGrid and author's calculations

Figure 6: Monthly natural gas demand in Europe (bcm). Taken from [52].

In order to consider the effect of temperature in gas demand for space heating purposes, it is necessary to establish a base temperature below which it is assumed a building will need to be heated. This is associated with the concept of **heating degree day (HDD)**. A *heating degree day* is a measure used to quantify the energy demand needed for heating a building. It is defined with respect to a *base temperature* T_{base} and is derived from measurements of the *surface air temperature* T_{sur} [53]:

$$T_{\mathrm{sur}} < T_{\mathrm{base}} \ \Rightarrow \mathrm{HDD} = T_{\mathrm{base}} - T_{\mathrm{sur}} \qquad \textit{(Eq. 7)}$$

$$T_{\mathrm{sur}} \geq T_{\mathrm{base}} \ \Rightarrow \mathrm{HDD} = 0$$

The base temperature T_{base} may vary depending on the geographical location of the region to study. In the case of Europe, T_{base} is assumed to be **15.5°C** [53]. The amount of energy required for heating a building is considered proportional to the number of heating degree days [53].

To calculate the heating degree days in European NUTS 3 regions for the 10-year scope of this project, it was necessary to obtain the surface air temperatures T_{sur} in respect to the location. The temperature data was provided by the European Center for Medium-Range Weather Forecasts (ECMWF) and its 5th generation meteorological analysis project, ERA5. ERA5 provides hourly estimates for different atmospheric quantities with a 0.25° latitude-longitude step. Among these is the **2-meter temperature**, equivalent to the surface air temperature T_{sur}, and defined as the temperature of the air at two meters above the surface of land, sea, or inland waters [54].

The dataset was obtained in the NetCFD (network common data form) format, and it contained the hourly 2-meter temperature for every day between the years 2010 and 2019 in a rectangular geographical area covering the European landmass.

A visual representation the data contained in the dataset is given in Figure 7. It shows the surface air temperature on January 1st, 2019, at 00:00 UTC in the different points of the 0.25° x 0.25° grid located inside the rectangular area defined by the coordinates above. The European NUTS 3 regions can also be observed as well.

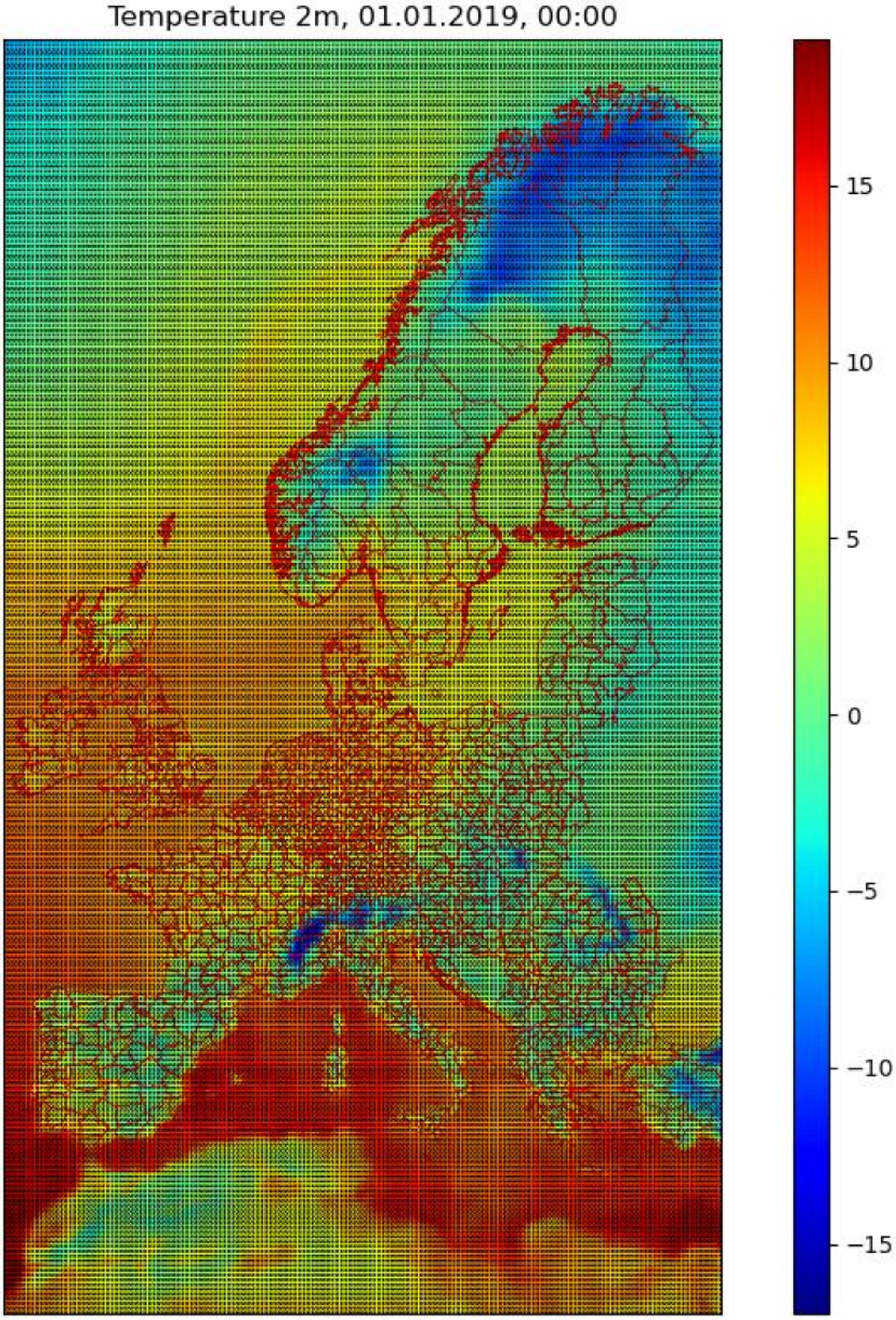


Figure 7: 2-meter temperature (°C) in NUTS 3 regions, 2019.01.01, 00:00.

For each set of coordinates, the mean value of the hourly temperatures of each day was used to calculate the daily 2-meter temperature. These daily temperature values were then assigned to their corresponding NUTS 3 regions according to their coordinates, and considering two possible cases:

- **Case 1:** One or more sets of coordinates inside of a NUTS 3 region.
- **Case 2:** No sets of coordinates inside of a NUTS 3 region.

To illustrate the two cases, an example is given in Figure 8, on which the blue dots represent sets of coordinates. For purposes of clarity regarding the resolution of the images, NUTS 1 regions will be used instead of NUTS 3 regions, but the concept applies to all NUTS levels.

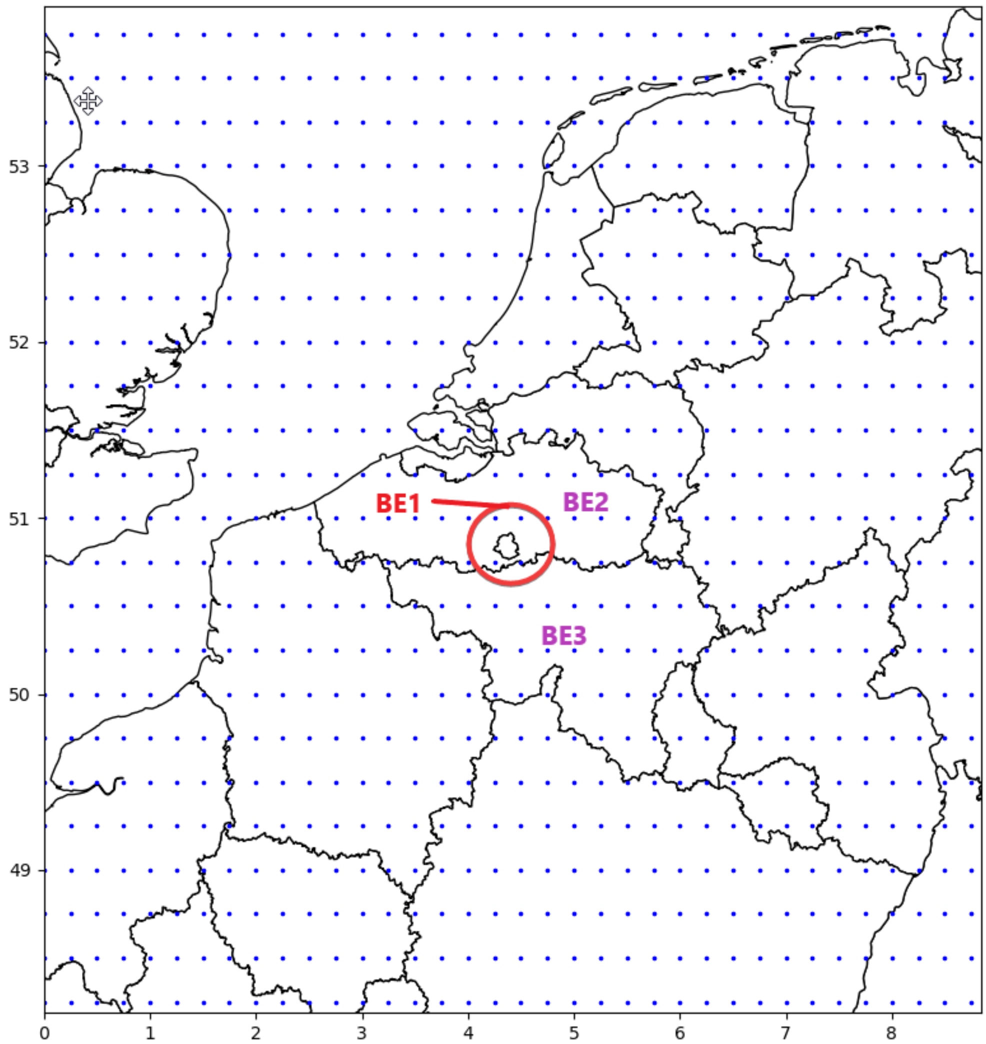


Figure 8: Zoom-in on Belgium's NUTS 1 regions, illustrating Case 1 (Regions BE2 and BE3), and Case 2 (Region BE1).

Figure 8 shows the NUTS 1 regions of Belgium, and the two cases mentioned above can be clearly observed. While two of the Belgian NUTS 1 regions (BE2 – Flanders and BE3 – Wallonia) contain one or more sets of coordinates for which temperature data is available, region BE1 (Brussels) does not have any interior coordinates with temperature information.

To assign temperatures to NUTS regions, it is necessary to analyze the temperature coordinates and assign them to their corresponding NUTS regions. This has been realized with the help of Geopandas[6] and shapefiles[7] Python packages.

- For **Case 1 regions**, the **mean value of all temperatures inside of a region** was calculated and assigned as T_{sur}, the daily surface air temperature for the entire region.
- For **Case 2 regions**, the minimum distance between different sets of coordinates and the points in the perimeter of the region was calculated. Then, the temperature value of this set of coordinates was assigned as T_{sur} to the region in question.

After assigning daily temperature values to all NUTS 3 regions, heating degree days (HDDs) were calculated using Eq. 7 [55]:

Once the HDDs of each day in a region were calculated, the following step consisted in the calculation of the HDDs ratios of a region in a year.

$$D_{\mathrm{HDD}}^{r\_y} = \frac{\mathrm{HDD}_r^d}{\sum \mathrm{HDD}_r^d} \qquad \textit{(Eq. 8)}$$

where r are the NUTS 3 regions and d is a day of the year y.

A sample of the dataset with the 2-m temperatures obtained from ERA5, the heating degree days obtained, and the heating degree days ratios for several NUTS 3 regions can be observed in Table 12:

[6] Geopandas: Open-source project for working with geospatial data in Python.
[7] Shapefile: A vector data file format used to spatially describe and store vector features.

Date	NUTS_ID [-]	Tsur [°C]	HDD [-]	DK_HDD [-]
1/01/2019	AT111	2.8128254	12.6871746	0.00623142
1/01/2019	AT112	4.24085066	11.2591493	0.00613768
1/01/2019	AT113	1.98943667	13.5105633	0.00666016
1/01/2019	AT121	3.03129689	12.4687031	0.00493826
1/01/2019	AT122	3.03274665	12.4672534	0.00523613
1/01/2019	AT123	4.1742678	11.3257322	0.00495074
1/01/2019	AT124	3.69931486	11.8006851	0.00482124
1/01/2019	AT125	4.92157544	10.5784246	0.00518641
1/01/2019	AT126	5.13548183	10.3645182	0.00518791
1/01/2019	AT127	5.0311285	10.4688715	0.00533336
1/01/2019	AT130	5.168505	10.331495	0.00532409
...	...	...	...	...

Table 12: Sample of temperatures and HDDs in NUTS 3 regions.

Finally, the daily gas demand for space heating is calculated by multiplying the annual gas demand of each NUTS 3 region obtained during the spatial disaggregation (see Section 3.1.3.3) by its corresponding heating degree day ratios:

$$C_{\mathrm{sh}}^{r\_d} = C_{\mathrm{sh}}^{r\_y} \cdot D_{\mathrm{HDD}}^{r\_y} \qquad \textit{(Eq. 9)}$$

where r is a NUTS 3 region and d is a day of the year y (2010-2019).

3.1.4.2 Water heating and cooking

As seen in Section 3.1.4.1, a very clear connection between outdoor air temperature and gas used for space heating has been established. This relation is not as clear for the two remaining usages of gas in households. While some publications show sensitivity to temperature conditions for these two applications (see [56], [57]), this phenomenon has not yet been studied in depth. Hence, for purposes of this thesis, the seasonal factor was not considered for the temporal disaggregation of gas demand for water heating and cooking.

For these two applications the gas demand was evenly distributed over the number of days per year as follows:

$$C_{\mathrm{wh}}^{r\_d} = \frac{C_{\mathrm{wh}}^{r\_y}}{\mathrm{d\_y}} \quad \textit{(Eq. 10)} \qquad\qquad C_{\mathrm{ck}}^{r\_d} = \frac{C_{\mathrm{ck}}^{r\_y}}{\mathrm{d\_y}} \quad \textit{(Eq. 11)}$$

where r is a NUTS 3 region, d is a day of a year y, and $\mathrm{d\_y}$ is the number of days per year, taking into consideration leap years.

3.1.4.3 Total final gas demand time series for the residential sector

Having estimated the daily time series for each residential gas application, the **total final gas demand time series in households** in a NUTS 3 region r on a day d was obtained by summing these individual components:

$$C_{\mathrm{total\_hh}}^{r\_d} = C_{\mathrm{sh}}^{r\_d} + C_{\mathrm{wh}}^{r\_d} + C_{\mathrm{ck}}^{r\_d} \qquad \textit{(Eq. 12)}$$

3.2 Industry and Commercial, Trade and Services

In this section the two remaining consumer groups (industrial and commercial) will be discussed. As it will be seen in Section 3.2.2, these two groups have a common energy demand determining variable, e.g., the number of employees, and therefore they will be analyzed jointly.

First, the industrial and commercial sectors will be defined, followed by a discussion on the situation of gas in Europe for both sectors. Then, the spatial and temporal disaggregation methods and their demand drivers will be presented.

3.2.1 Definition of the Industry and CTS consumers

Economic activities can be defined and classified in different ways. A common way of doing this uses the **three-sector theory**, where the economy is divided in three sectors [58]:

- **Primary sector**: Extraction of raw materials and the agricultural (farming, fishing) activities.
- **Secondary sector**: Manufacturing of finished goods, utilities, and construction.
- **Tertiary sector**: Service sector, such as retail, financial activities, hospitality and leisure, communications, and information technology, etc.

In the European Union, the standard classification system of economic activities is based on the three-sector theory, and it is usually known as NACE (***N**omenclature Statistique des **Ac**tivités Économiques dans la **C**ommonauté **E**uropéenne).* The current version, revision 2, was established in the year 2006 and presents four levels, from general to specific economic activities:

- **21 Sections:** Identified by letters from A to U.
- **88 Divisions:** Identified by two-digit numerical codes.
- **272 Groups:** Identified by three-digit numerical codes.
- **629 Classes:** Identified by four-digit numerical codes.

The complete NACE Rev. 2 classification can be found at [59].

In this work, both the **Industrial** sector and the **Commercial, Trade and Services (CTS)** sector were defined with the help of the NACE Rev. 2 sections, as follows:

Industry:

Defined by the following economic activities of NACE Rev. 2:

- Section B: Mining and quarrying.
- Section C: Manufacturing.
- Section F: Construction.

Commercial, Trade, and Services (CTS):

Defined by the following economic activities of NACE Rev.2:

- Section G: Wholesale and retail trade; repair of motor vehicles and motorcycles.
- Section H: Transportation and storage.
- Section I: Accommodation and food service activities.
- Section J: Information and communication.
- Section K: Financial and insurance activities.
- Section L: Real estate activities.
- Section M: Professional, scientific and technical activities.
- Section N: Administrative and support service activities.
- Section O: Public administration and defense; compulsory social security.
- Section P: Education.
- Section Q: Human health and social work activities.
- Section R: Arts, entertainment and recreation.
- Section S: Other service activities.
- Section T: Activities of households as employers; undifferentiated goods and services producing activities of households for own use.
- Section U: Activities of extraterritorial organizations and bodies.

In the industry sector**,** three sections were excluded for reasons related to a lack of relevant statistical data, either demographic (Section A and Section E) or energy-related (Section D). Regarding the latter, the Eurostat Complete Energy Balance dataset lacks the input information necessary for the spatial disaggregation, especially for NACE Rev.2 Class 35.11, “Production of electricity” (see [14], NRG_BAL code NRG_EHG_E, “Energy sector - electricity and heat generation - energy use”).

The separation of the industrial and commercial sector by these codes will be used in the processing of the input datasets used for the spatial disaggregation of gas consumption (see Section 3.2.3).

3.2.2 Description of gas demand in the industrial and commercial sectors

As presented in Figure 4, for the years 2010 and 2019 the industry and the commercial consumers presented respectively the second and third position in terms of gas consumption in the European Union out of the three sectors analyzed in this thesis.

Table 13 shows that for the selected time period the percentages of gas consumption on all three sectors did not present significant changes. In the case of industry, the lowest share of gas consumption took place in 2010 with 34% and peaked in 2014 with 38%. After that year, gas consumption in the industry sector decreased and remained steady around 36%.

Final energy consumption by sector, gas, 2010-2019 [%]			
Year	Households	Industry	CTS
2010	48%	34%	19%
2011	46%	37%	18%
2012	47%	35%	18%
2013	46%	35%	19%
2014	43%	38%	18%
2015	45%	36%	19%
2016	46%	35%	19%
2017	45%	36%	19%
2018	44%	36%	19%
2019	45%	36%	19%

Table 13: Final energy consumption by sectors, gas, 2010-2019 [14].

Table 13 also shows that the percentage of gas demand in the commercial sector has been very stable in the studied decade. The CTS sector consistently presents the lowest consumption of all three sectors.

Along with electricity, gas is one of the two most used fuels in both the industry and the commercial sector, as shown in Figure 9 and Figure 10.

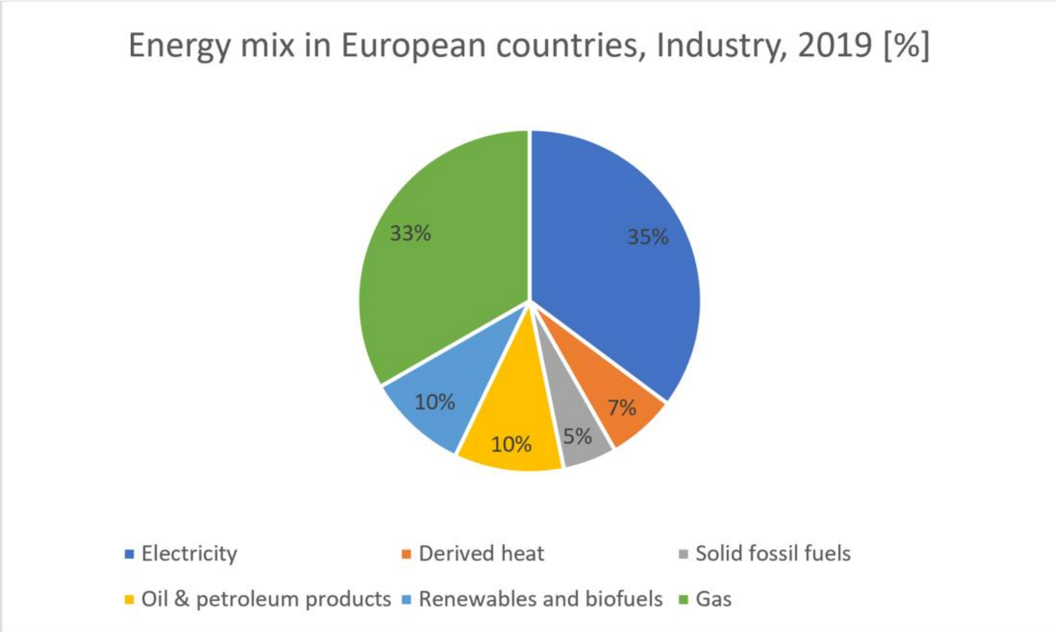


Figure 9: Energy mix in European countries in 2019, industry sector [14].

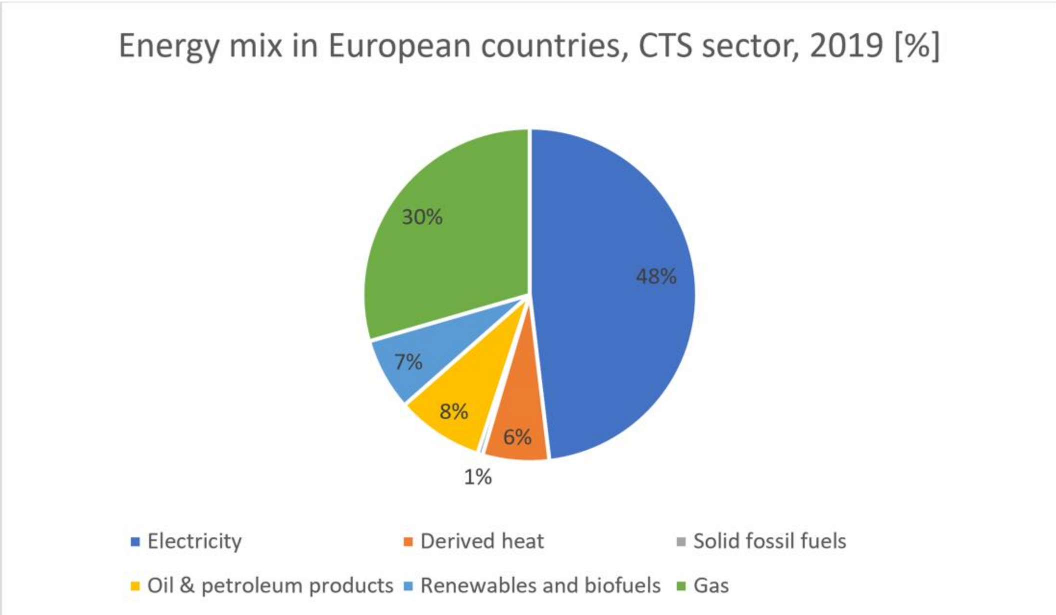


Figure 10: Energy mix in European countries in 2019, commercial sector [14].

As observed in Figure 9, gas had a significant role as a fuel in the industry sector, with a 33% of the energy mix in 2019, almost as important as electricity, the most consumed energy source with 35%. In the CTS sector (Figure 10), gas also is the second most-consumed fuel after electricity. Still, in this case, the difference between both is higher, with electricity taking approximately half of the energy mix of the commercial sector (48%), and gas with almost a third of it with 30%.

At country level, the highest consumption of gas in both the industry and the CTS sector occurred in the same countries: Germany, the United Kingdom, France, Italy, Spain, and the Netherlands.

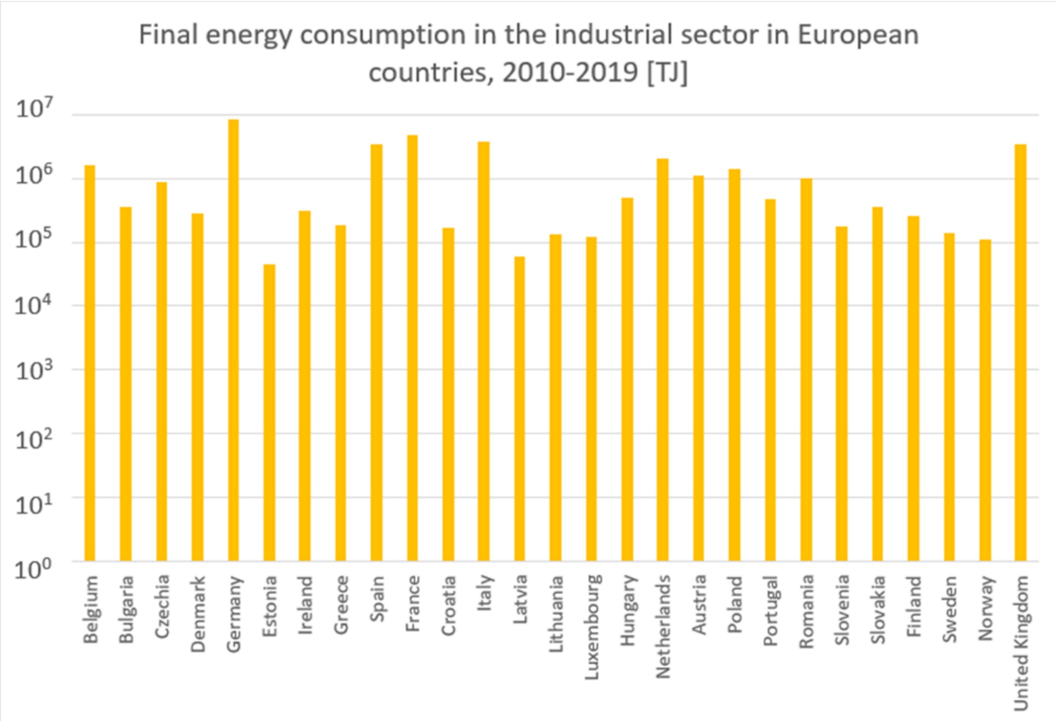


Figure 11: Final energy consumption in industry in Europe, gas, 2010-2019 [14].

Figure 11 shows that he highest consumption of gas in the European industry sector between 2010 and 2019 occurred in Germany (23.5% of the total). Comparatively, it was also significantly higher than other countries, like France (13.4%), Italy (10.6%), United Kingdom (9.5%), and Spain (9.4%).

In the commercial sector (see Figure 12), Germany also had the highest gas demand, consuming 21.4% of the total, and followed by Italy (16.1%), United Kingdom (14.7%), France (13.7%), and Netherlands (7.5%).

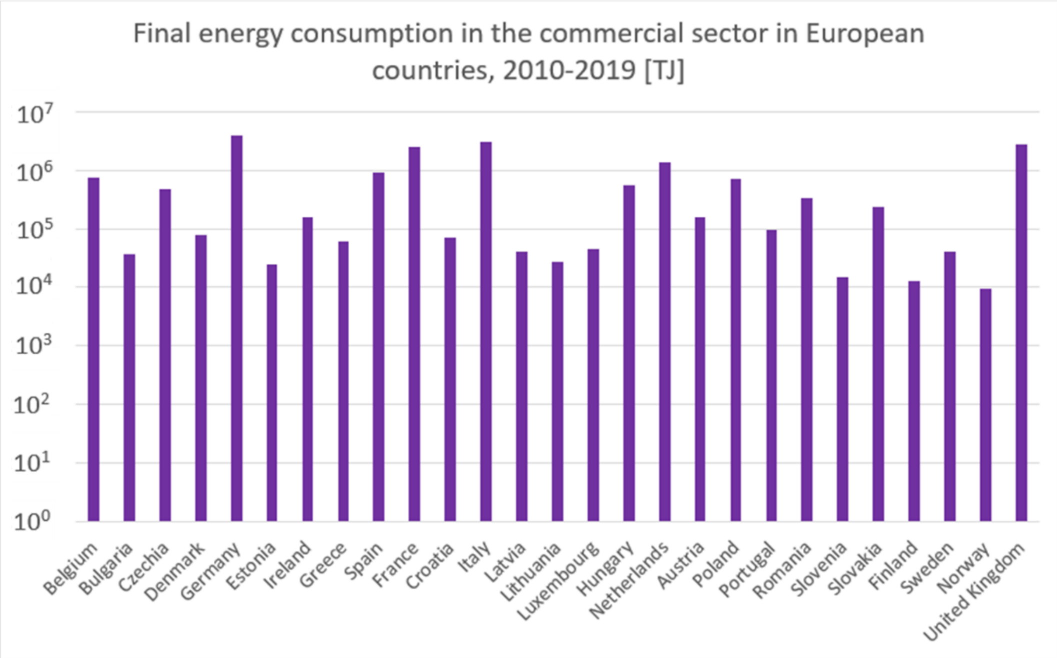


Figure 12: Final energy consumption in CTS in Europe, gas, 2010 - 2019 [14].

3.2.3 Spatial disaggregation of the industry and commercial sectors

As described in Section 3.1.3, the regionalization of residential sector was performed using a **top-down** approach. The spatial disaggregation of the industry and CTS sectors also used a **top-down** methodology, going from gas consumption at a country level to gas consumption at a regional (NUTS 3) level.

As presented in Section 3.2.1, the industrial and commercial sector were defined by a shared source, the Statistical Classification of Economic Activities in the European Community (NACE). Similarly, both sectors had common gas demand drivers: the **number of employees by NACE section**, and the **gross domestic product (GDP)**. Both indicators combined provide an image of the economic activity of a region and were used to estimate the gas consumption for each sector.

The datasets used for the spatial disaggregation of the industrial and commercial sectors are given by Table 14:

Sector	Dataset	Resolution	Notes
Industrial	SBS data by NUTS 2 regions and NACE Rev. 2 [28]	NUTS 2	Number of employees are divided by industrial activity.
	Employment (thousand persons) by NUTS 3 regions [29]	NUTS 3	Number of employees for the industry sector at broader categories.
Commercial	Employment by sex, age, economic activity, and NUTS 2 regions (NACE Rev. 2) [27]	NUTS 2	Number of employees in the commercial sector.
Industrial and Commercial	Gross domestic product (GDP) at current market prices by NUTS 3 regions [30]	NUTS 3	Used in the industrial sector for Norway and Sweden, and for the disaggregation of employees in the CTS sector.

Table 14: Datasets used for the spatial disaggregation of the industrial and commercial sectors.

As mentioned in sector 2.2.1.1, the Complete Energy Balance dataset provides data about gas consumption by different types of industry. This separation allows to avoid agglutinating energy-intensive industries (chemical, food, paper, metallurgic, etc.) with non-energy intensive industries (pharmaceutical, machinery, computer, and electronic products, etc.) [60, p. 113].

The datasets containing national gas consumption by type of industry were associated with the corresponding statistics of number of employees, which were obtained from the SBS (*statistical business structure) by NUTS 2 regions and NACE Rev. 2* dataset, as shown in Table 15. For clarity, this dataset will be referred to as the **SBS dataset.**

National gas consumption dataset	Statistical business structure employees dataset
Iron and steel	Basic metals
Non-ferrous metals	
Mining and quarrying	Mining and quarrying
Chemical and petrochemical	Chemicals and chemical products
Non-metallic minerals	Other non-metallic mineral products
Transport equipment	Motor vehicles, trailers and semi-trailers; other transport equipment
Machinery	Machinery and equipment n.e.c.
Food, beverages, and tobacco	Manufacture of food products; beverages; tobacco products
Paper, pulp, and printing	Paper and paper products; printing and reproduction of recorded media
Wood and wood products	Wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
Construction	Construction
Textile and leather	Textiles; wearing apparel; leather and related products
Not elsewhere specified	Coke and refined petroleum products; basic pharmaceutical products and pharmaceutical preparations; rubber and plastic products; fabricated metal products, except machinery and equipment; computer, electronic and optical products; electrical equipment; furniture; other manufacturing; repair and installation of machinery and equipment

Table 15: Correspondence between gas consumption and number of employees datasets for NACE industrial activities.

For the commercial sector, the dataset used was *Employment by age, economic activity, and NUTS 2 regions.* This dataset will be referred to as the **EN2 dataset.**

As their names indicate, the number of employees datasets used for the CTS and industry sector have a maximum resolution of NUTS 2 level. Hence, it was necessary to use other datasets to estimate the number of employees at NUTS 3 level.

For the industry sector, this estimation was done using the *Employment (thousand persons) by NUTS 3 regions* dataset. For clarity, this dataset will be referred to as the **EN3 dataset.** This dataset has the significant disadvantage of joining different economical activities in broad categories (i.e., Industry, Manufacturing, Construction, etc.). Furthermore, the dataset contains no information of the number of employees in Norway and Sweden NUTS 3 regions.

To estimate the number of employees at NUTS 3 level in the industrial sector, the ratio of employees in manufacturing activities of the NUTS 3 regions belonging to a NUTS 2 region were calculated using the **EN3 dataset**, and then multiplied by the number of employees at NUTS 2 regions from the **SBS dataset.**

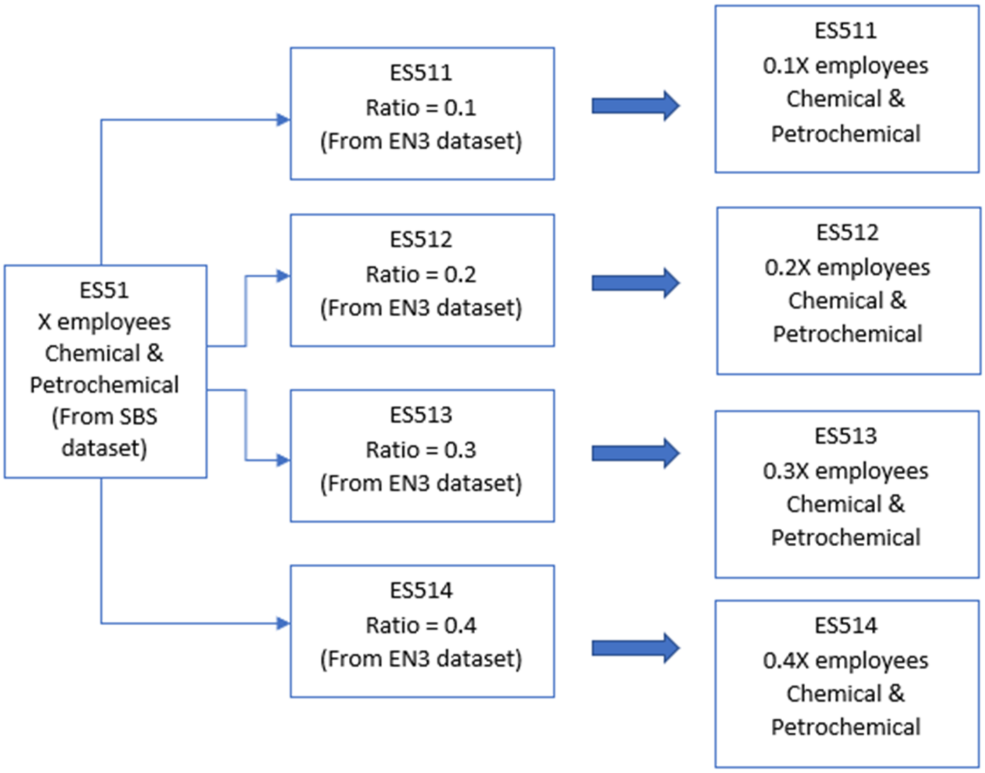


Figure 13: Example of disaggregation of employees in a NUTS 2 region into NUTS 3 regions (industrial sector).

Figure 13 provides an example. Region ES51 (NUTS 2) has X employees working for the chemical industry, according to the SBS dataset. The ratios at the NUTS 3 regions ES11, ES12, ES13 and ES14 have been calculated from the information provided by the EN3 dataset. By multiplying the X employees in the chemical industry at NUTS 2 level by the ratios of each individual NUTS 3 region, the number of employees in the chemical industry at NUTS 3 level is estimated.

As the GDP can act as a proxy of the economy of a region, the *Gross domestic product (GDP) at current market prices by NUTS 3 regions* dataset [30] is used to estimate the number of employees in the CTS sector at NUTS 3 level.

To estimate the number of employees at NUTS 3 level, the GDP ratios of NUTS 3 regions part of a NUTS 2 region were calculated and then multiplied by the number of employees in a NUTS 2 region. An example is given in Figure 14:

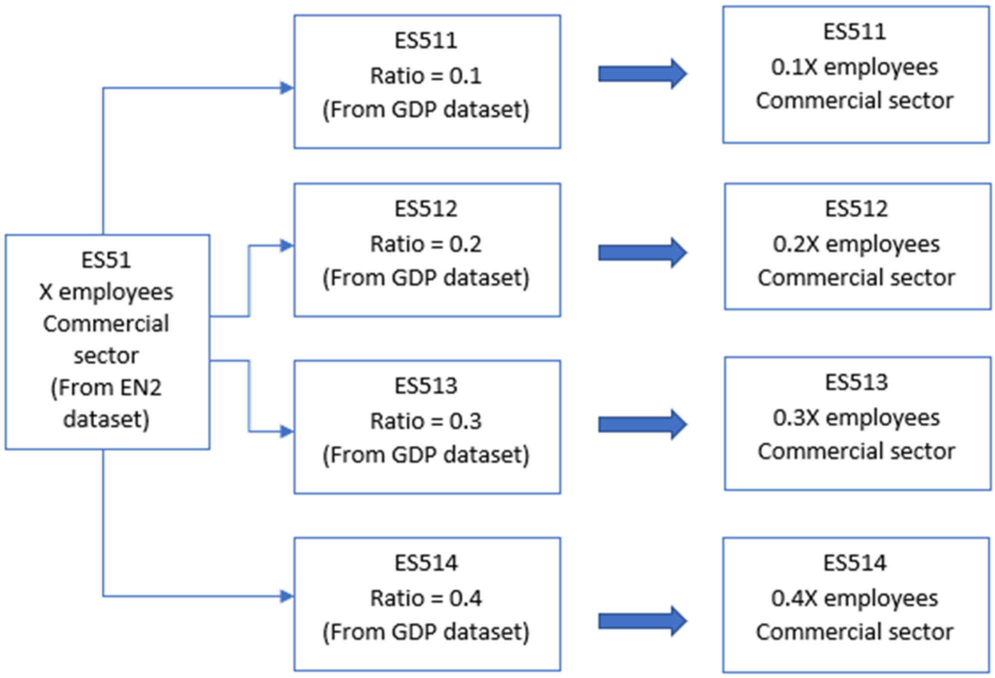


Figure 14: Example of disaggregation of employees in a NUTS 2 region into NUTS 3 regions using GDP (commercial sector).

In a similar way to the previous example, the GDP ratios of the four NUTS 3 regions are calculated and multiplied by the number of employees at NUTS 2 level, obtaining an estimate of employees in the commercial sector of NUTS 3 regions.

Having estimated the number of employees at NUTS 3 level, the spatial disaggregation for both industry and CTS was performed in a similar way as in the previous applications in the residential sector:

- First, NUTS 3 level distribution keys for the demand driver (number of employees) were calculated, as follows:

$$D\_\mathrm{emp}_{\mathrm{sector}}^{r\_R} = \frac{\mathrm{emp}_r^{\mathrm{sector}}}{\sum \mathrm{emp}_r^{\mathrm{sector}}} \qquad \textit{(Eq. 13)}$$

  where emp is the number of employees in the different industrial activities or in the CTS sector in a NUTS 3 region r belonging to a country R.

- Then, the gas demand at country level (for each industrial activity or for the CTS sector) was multiplied by the regional distribution keys:

$$C^{r}_{\text{sector}} = C^{R}_{\text{sector}} \cdot D\_\text{emp}^{r\_R}_{\text{sector}} \qquad (Eq.\ 14)$$

Where C^{r}_{sector} is the final gas consumption of the different industrial activities or the commercial sector in a NUTS 3 region r of an *EUR_27* country R.

- Finally, the gas demand of each industrial activity in a NUTS 3 region was summed up to obtain the total annual gas demand in the industry sector.

3.2.4 Temporal disaggregation of the industry and commercial sectors

For the temporal disaggregation of the industry and CTS sectors, different assumptions for the use of gas were made:

- **Space heating:** As in the residential sector, a part of the gas consumed in the commercial sector is used for space heating of buildings such as offices, stores, and restaurants. The **heating degree days** were then used as a driver for the temporal disaggregation of this aspect. Space heating is not considered a relevant factor of gas consumption in the industrial sector.

- **Production:** Another part of the gas consumption goes to the production of goods and services. For instance, in the industry sector gas is consumed in the production of chemicals, metal, food, and paper. In the CTS sector, gas is also used in the hospitality and leisure areas (for example, in restaurants). Hence, a distinction between *workdays* and *holidays* was relevant in this case. It is assumed that in the industry sector most of the production occurs on weekdays, with lower production on Saturdays and a minimum on Sundays and holidays. In the commercial sector, the workdays are considered to occur from Monday to Saturday, without a special distinction regarding the production on specific days.

- **Base:** It is assumed that on both sectors there is a base gas consumption that takes place **on all days of the year**, as there are downsides to, for example, a complete stop of specific machinery such as ovens in the metal industry, and that commercial establishments such as hotels are functioning even on Sundays and holidays.

Having estimated the annual gas demand in NUTS 3 regions during the spatial disaggregation, initial arbitrary ratios were assigned to the three uses of gas mentioned above.

Sector	Base [-]	Weekday production [-]	Saturday production [-]
Industry	0.4	0.5	0.1

Table 16: Initial arbitrary values for the industrial sector.

Sector	Base [-]	Production [-]	Space heating [-]
CTS	0.2	0.3	0.5

Table 17: Initial arbitrary values for the commercial sector.

The arbitrary values from Table 16 and Table 17 were improved with a validation dataset by an optimization process, as it will be seen in Section 4.2.

The annual final gas consumption in NUTS 3 regions estimated in Section 3.2.3, C^{r}_{sector}, was multiplied by these ratios and separate daily time series for the three uses were estimated as follows:

- **Base:** As the base use assumed presents no dependency on any other factors apart from the number of days of the year, the daily gas consumption for the base usage was estimated by:

$$C^{r\_d}_{\mathrm{base\_sector}} = \frac{C^{r\_y}_{\mathrm{base}}}{\mathrm{d\_y}} \qquad \textit{(Eq. 15)}$$

  Where $C^{r\_d}_{\mathrm{base}}$ is the final gas consumption in a NUTS 3 region r on a day d corresponding to a year y between 2010 and 2019, $\mathrm{d\_y}$ is the number of days of the year, and sector corresponds to Industry or CTS.

- **Space heating:** The temporal disaggregation of gas used for space heating purposes in the commercial sector was performed in the same manner it was done for the residential sector.

Table 12 showed a sample of the dataset containing the daily outside temperatures, heating degree days, and HDD ratios associated to NUTS 3 regions. The HDD ratios were also used here, multiplying them by the yearly gas consumption associated to space heating in the CTS sector estimated with the arbitrary ratios from Table 17:

$$C_{\mathrm{sh\_cts}}^{r\_d} = C_{\mathrm{sh\_cts}}^{r\_y} \cdot D_{\mathrm{HDD}}^{r\_d} \qquad \textit{(Eq. 16)}$$

where $D_{\mathrm{HDD}}^{r\_d}$ are the heating degree day ratios on a day d of a year y in a NUTS 3 region r.

- **Production:** As mentioned above, the production aspect is related to a **day classification** between *workdays* and *holidays***.**
  In this project, for the **industry** sector, the workdays (also called business days) are considered all **weekdays (Monday to Friday)** and **Saturdays**, excluding those when a public holiday occurs. For the **commercial** sector, this project assumed that **Monday to Saturday** are working days, excluding those coinciding with a public holiday.
  As mentioned above, in the industrial sector it was assumed that the production is higher on weekdays, and lower on Saturdays. Production was not assumed to exist on Sundays and holidays.
  In the commercial sector, it was also assumed that the production is evenly distributed during the year (that is, the production is the same across different workdays without distinction).
  The daily time series for the production aspect of the industry and CTS sectors was estimated as:

$$C_{\mathrm{prod\_sector}}^{r\_d} = \frac{C_{\mathrm{prod\_sector}}^{r\_y}}{\mathrm{w}_{\mathrm{sector}}^{r} - \mathrm{h}_{\mathrm{sector}}^{r}} \qquad \textit{(Eq. 17)}$$

where $\mathrm{w}_{\mathrm{sector}}^{r}$ is the number of workdays of the sector and $\mathrm{h}_{\mathrm{sector}}^{r}$ is the number of holidays happening on a workday, both in a NUTS 3 region r.

Finally, the total daily time series for each year was obtained by adding the individual gas consumption daily time series for both sectors:

$$C_{\mathrm{total\_ind}}^{r\_d} = C_{\mathrm{base\_ind}}^{r\_d} + C_{\mathrm{prod\_weekdays}}^{r\_d} + C_{\mathrm{prod\_saturday}}^{r\_d} \qquad \textit{(Eq. 18)}$$

$$C_{\mathrm{total\_cts}}^{r\_d} = C_{\mathrm{base\_cts}}^{r\_d} + C_{\mathrm{sh\_cts}}^{r\_d} + C_{\mathrm{prod\_cts}}^{r\_d} \qquad \textit{(Eq. 19)}$$

4. Results

In this section, the results obtained in the project will be introduced. Then, the reference datasets used for validation will be presented, followed by a description of the model optimization and a presentation and discussion of the project's outcome.

4.1 Presentation of the results

The outcome of the thesis consists in **daily time series for each individual NUTS 3 region from 2010 to 2019**. The time series are presented in csv files and can be found as the CONS dataset of the SciGRID_gas project [61].

A sample of the results for the NUTS 3 region AT111 (Mittelburgenland) is shown in Figure 15:

Date	Households gas consumption	Industry gas consumption	Commercial gas consumption
YYYY-mm-dd	[Mm^3/d]	[Mm^3/d]	[Mm^3/d]
2010-01-01	0.114	0.018	0.007
2010-01-02	0.132	0.019	0.008
2010-01-03	0.166	0.018	0.011
2010-01-04	0.182	0.023	0.012
2010-01-05	0.179	0.023	0.012
2010-01-06	0.172	0.018	0.011
2010-01-07	0.167	0.023	0.011
2010-01-08	0.161	0.023	0.01
2010-01-09	0.148	0.019	0.009
2010-01-10	0.15	0.018	0.009
2010-01-11	0.16	0.023	0.01
2010-01-12	0.162	0.023	0.01
2010-01-13	0.164	0.023	0.01
2010-01-14	0.161	0.023	0.01

Figure 15: Sample of results for NUTS 3 region AT111.

Figure 16 depicts the gas demand in NUTS 3 regions for the residential sector:

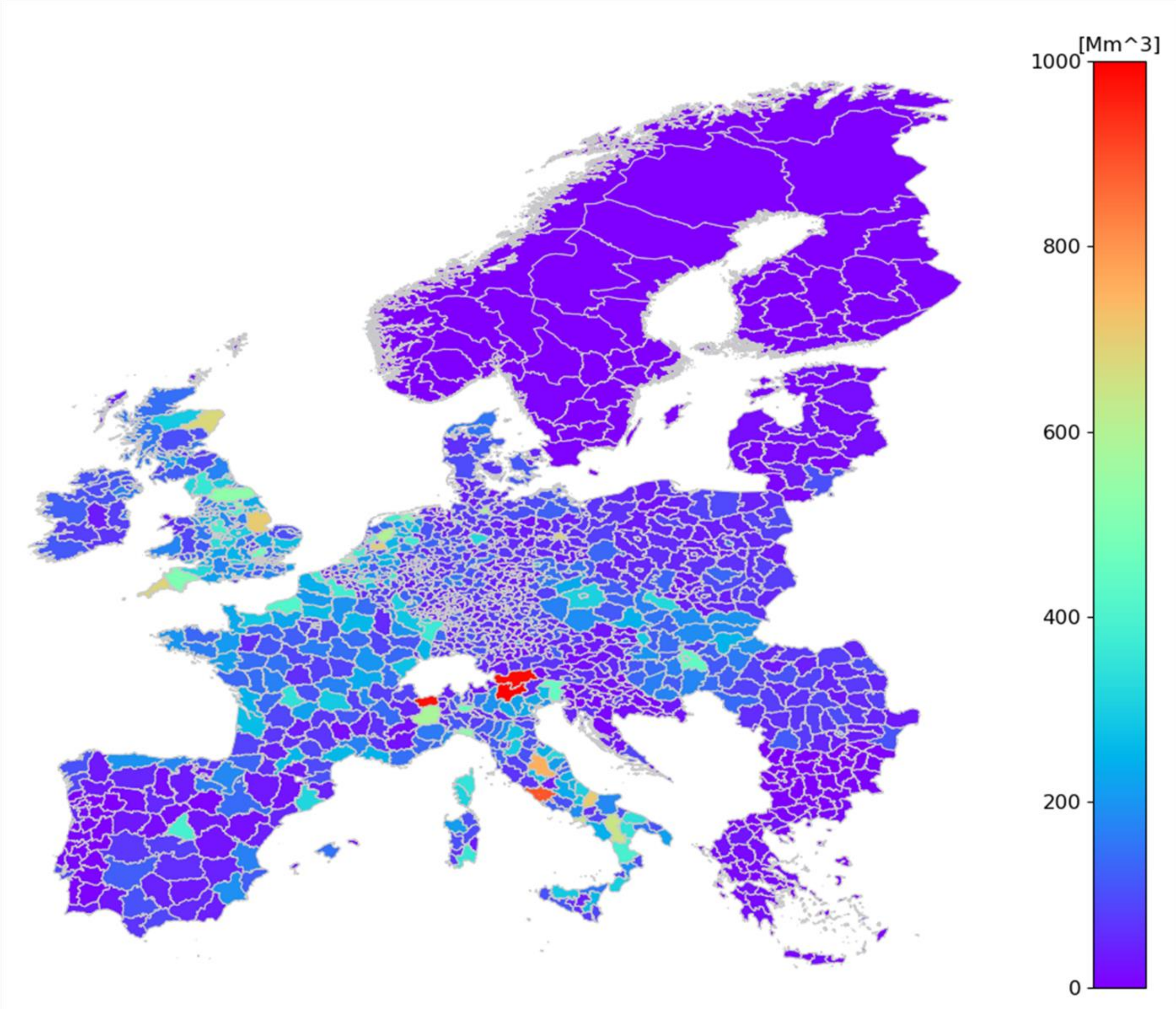


Figure 16: Gas demand in the residential sector, NUTS 3 regions, 2015.

Figure 16 shows that the highest gas consumption occurred in NUTS 3 regions located in the north of Italy (ITH10, ITH20, ITC20), as well as regions in the United Kingdom (UKF30, UKK30, UKM50) and in major cities from several countries (Madrid, Berlin, Rome, Budapest, etc.). In the case of Italy, the country has the highest gas consumption for cooking purposes of all Europe, and the regions near the Alps, close to the border with Switzerland, present a high consumption of gas for space heating due to low average temperatures. The British regions mentioned above also present high consumption for space heating.

Figure 17 shows the distribution of NUTS 3 regions by their residential gas demand:

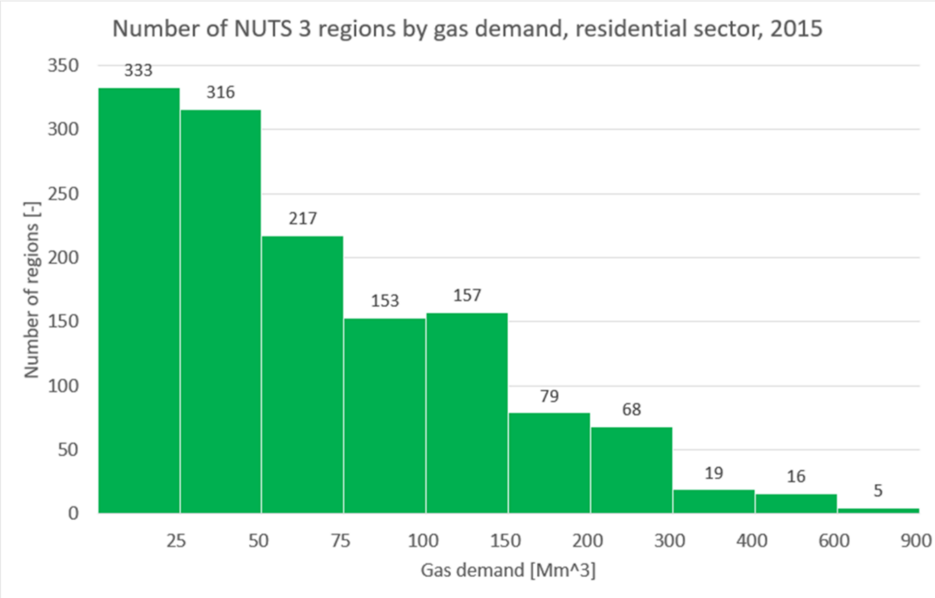


Figure 17: Distribution of NUTS 3 regions by gas demand, residential sector, 2015.

The choropleth map in Figure 18 shows the industrial gas demand in NUTS 3 regions:

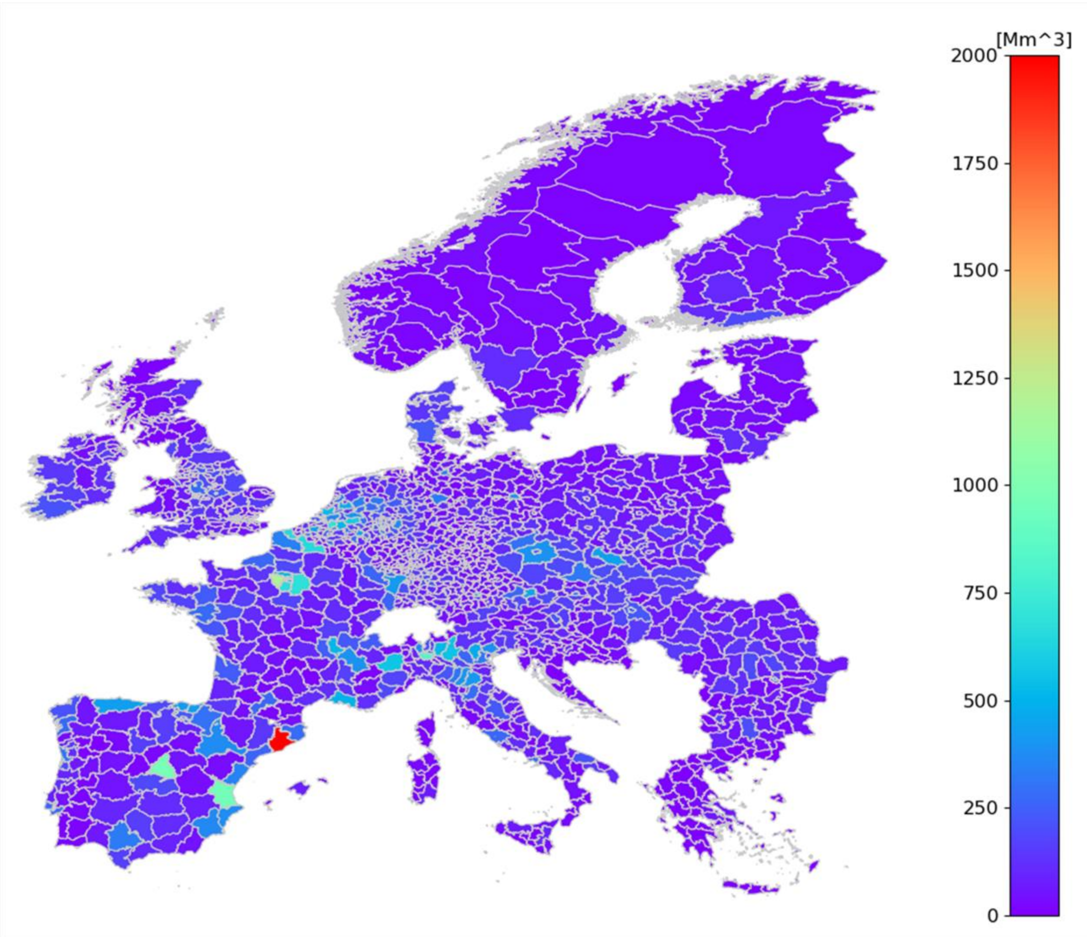


Figure 18: Gas demand in the industrial sector, NUTS 3 regions, 2015.

In the industrial sector, Figure 18 shows that region ES511 (Barcelona) had the highest gas demand of all Europe. Barcelona and its metropolitan area are the most important industrial region of Spain, followed by Madrid (ES300) and Valencia (ES523), with presence of chemical, metallurgy, and automotive companies, among others [62]. The data also shows high industrial gas consumption in the surroundings of the Île-de-France region (FR103 and FR105), one of the most important centers of the world economy and the most important in France. Key industries include defense, automobile, aerospace, and pharmaceuticals.

The distribution of NUTS 3 regions by their gas demand in the industrial sector is depicted in Figure 19.

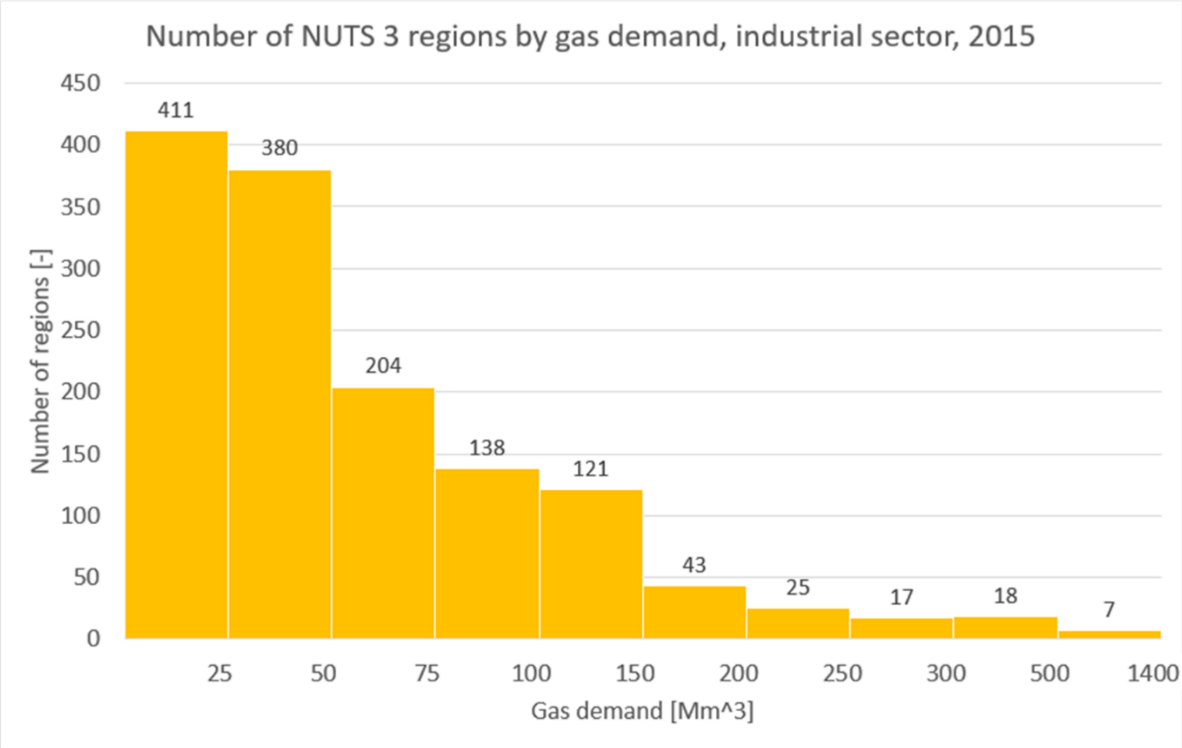


Figure 19: Distribution of NUTS 3 regions by gas demand, industrial sector, 2015.

Finally, the regional distribution of gas in 2015 for the commercial sector is shown in Figure 20:

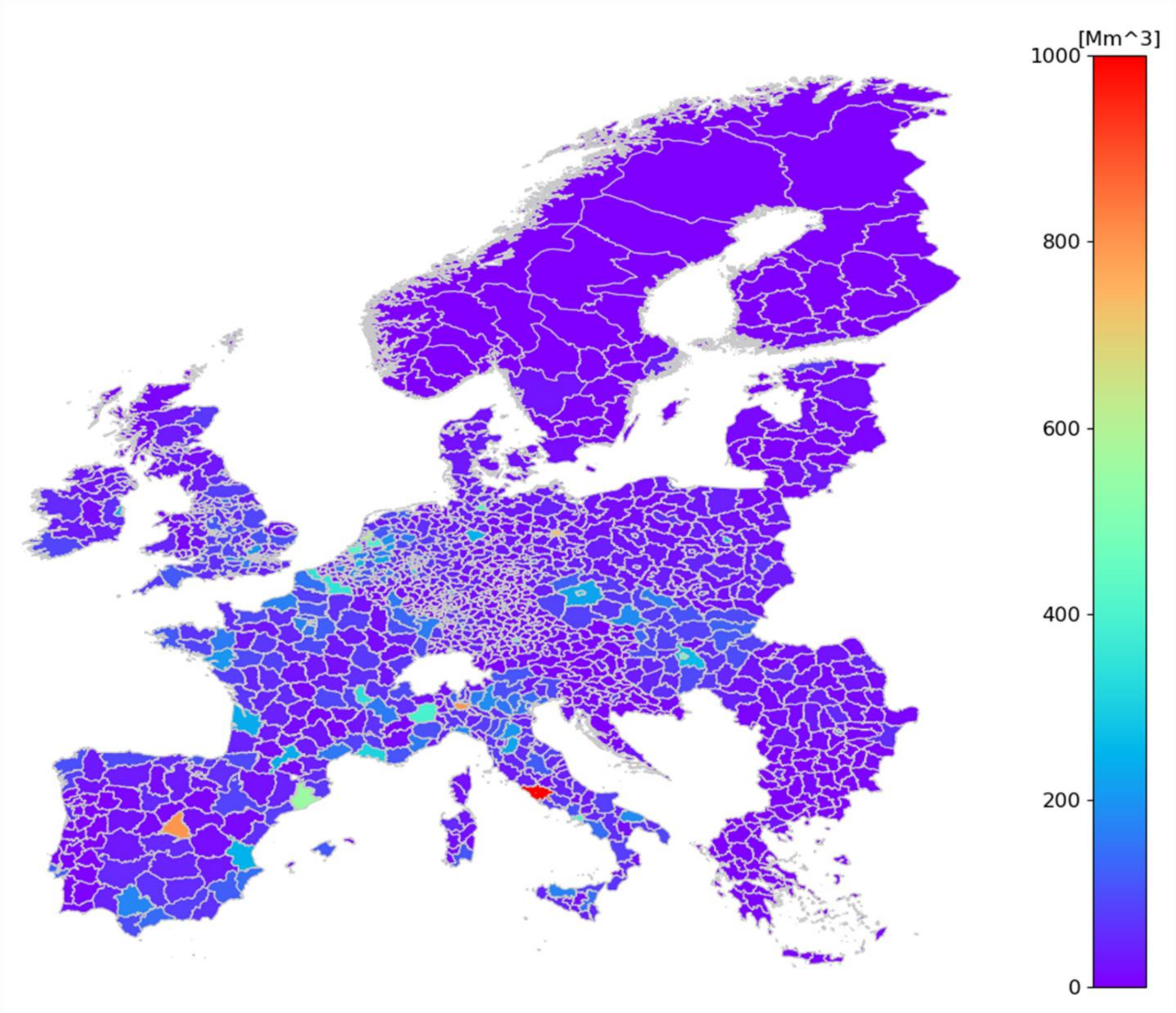


Figure 20: Gas demand in the commercial sector, NUTS 3 regions, 2015.

According to Figure 20, gas in the commercial sector is highly consumed in main urban centers: Rome (ITI43), Madrid (ES300), Berlin (DE300), Hamburg (DE600) and Amsterdam (NL329). Such cities concentrate a large proportion of the tertiary sector of the economy (commerce, financial, services, among others).

The histogram in

Figure 21 presents the distribution of NUTS 3 regions by gas consumption in the commercial sector.

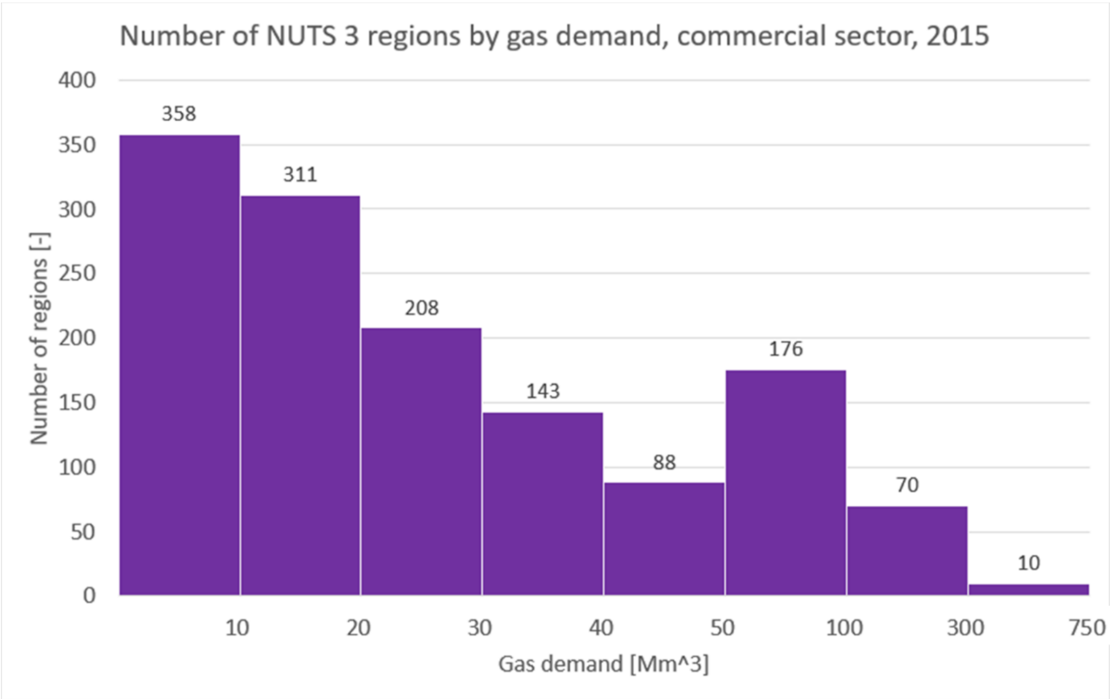


Figure 21: Distribution of NUTS 3 regions by gas demand, commercial sector, 2015

4.2 Processing of reference datasets

For verification of the plausibility of this project's results, the results from the DemandRegio project [15] will be used as reference datasets. These results were obtained using DemandRegio's API [63] and consist of three time series of gas demand in Germany in 2015, one for each analyzed sector (residential, industrial, and commercial).

The DemandRegio project used as its input dataset the Arbeitsgemeinschaft Energiebilanzen (Energy Balance Working Group, AGEB) for Germany [64]. The annual sum of gas used by the different sectors can be observed below:

- **Residential:** 861306 TJ ≙ 22556.01 Mm^3
- **Commercial, trade, and services (CTS):** 408812 TJ ≙ 10706.03 Mm^3
- **Industrial:** 779014 TJ ≙ 20400.94 Mm^3

According to their metadata, the results of DemandRegio used MWh/15 min as units for the industry sector and MWh/h for the household and commercial sector, as shown in Figure 22:

Date	DE111	DE112	DE113	DE114	DE115	DE116	DE117	DE118
2015-01-01 00:00:00	351.921	109.054	134.787	57.2274	140.304	100.608	55.8506	76.0666
2015-01-01 01:00:00	352.013	108.331	134.187	57.2045	140.086	100.454	55.7327	75.8197
2015-01-01 02:00:00	347.428	107.391	132.794	56.4641	138.355	99.1632	55.0334	75.0946
2015-01-01 03:00:00	371.559	115.276	143.239	61.1674	149.834	107.482	59.749	81.1663
2015-01-01 04:00:00	417.256	127.306	158.792	67.8905	166.884	119.652	66.9564	89.1567
2015-01-01 05:00:00	473.519	143.909	180.355	77.212	190.431	136.684	76.8149	100.646
2015-01-01 06:00:00	419.644	132.887	167.741	71.618	175.206	125.918	69.9698	94.4475
2015-01-01 07:00:00	420.963	134.442	170.32	72.8072	178.484	128.244	71.2695	96.4221
2015-01-01 08:00:00	420.675	133.695	169.166	72.2421	176.715	126.948	70.6141	95.4354
2015-01-01 09:00:00	406.875	129.946	164.807	70.252	171.637	123.979	69.1113	91.6481

Figure 22: Sample of DemandRegio's CTS result time series, hourly resolution.

As the results of this project have a **daily resolution**, it became necessary to resample the results of DemandRegio by summing the gas consumption of every day of the year, as shown in Figure 23.

Date	DE111	DE112	DE113	DE114	DE115	DE116	DE117	DE118
2015-01-01	9410.73	2935.99	3684.64	1568.95	3839.79	2762.85	1542.09	2054.51
2015-01-02	10078.1	3204.41	4057.94	1799.54	4204.53	3009.18	1655.37	2239.67
2015-01-03	8521.83	2763.53	3483.85	1532.27	3597.23	2585.15	1428.62	1925.4
2015-01-04	8750.6	2777.7	3465.19	1536.55	3568.89	2573.04	1425.72	1919.48
2015-01-05	9514.01	3019.83	3752.55	1686.17	3858.97	2781.16	1548.81	2088.86
2015-01-06	9828.96	3074.04	3857.95	1707.52	4018.87	2855.53	1646.28	2207.7
2015-01-07	9828.9	3021.9	3847.18	1680.65	4081.21	2880.97	1684.71	2254.92
2015-01-08	9535.9	2929.53	3718.32	1637.47	3983.27	2798.55	1641.5	2197.75
2015-01-09	8216	2581.11	3219.17	1450.21	3424.85	2423.36	1402.57	1887.96
2015-01-10	5818.83	1901.32	2301.58	1066.76	2391.74	1717.68	986.643	1330.13

Figure 23: Sample of DemandRegio's CTS result time series, daily resolution.

Each sector's time series was converted to Mm³/d of gas for comparison with the results of this project, and the sum of the daily gas consumption on each German NUTS 3 region was calculated. Summing up these figures, the yearly total gas consumption of each sector was obtained. To check for consistency, these output values were compared to the input values mentioned above.

	Residential [Mm³]	Industrial [Mm³]	CTS [Mm³]
Input (AGEB)	22 556.01	20 400.94	10 706.03
Output (DemandRegio)	8 611.35	28 622.02	10 786.36

Table 18: DemandRegio's input vs. output yearly gas demand per sector, 2015.

From Table 18, it can be observed that there is a slight difference between the input and the output gas consumption figure for the commercial sector. However, there is a significant discrepancy between the two numbers in the households and the industry sectors. It is assumed, then, the existence of an error in the DemandRegio industry and households' datasets.

To overcome this issue, scaling factors are calculated:

$$s_h = \frac{d_{\text{in\_h}}}{d_{\text{out\_h}}} = \frac{22556.01\ Mm^3}{8611.35\ Mm^3} = 2.619 \qquad \text{(Eq. 20)}$$

$$s_i = \frac{d_{\text{in\_i}}}{d_{\text{out\_i}}} = \frac{20400.94\ Mm^3}{28622.02\ Mm^3} = 0.712 \qquad \text{(Eq. 21)}$$

By multiplying these scaling factors by the households and industry time series, the new figures are consistent with the input gas demand values from AGEB.

The time series for German NUTS 3 regions obtained in this thesis were also summed up to check for consistency, and as shown in Table 19, the input and the output values match up almost exactly.

	Residential [Mm³]	Industrial [Mm³]	CTS [Mm³]
Input (Eurostat)	21 642.57	21 050.62	10 548.10
Output	21 642.47	21 051.88	10 547.32

Table 19: Project's input vs. output yearly gas demand per sector, 2015.

4.3 Model optimization

As mentioned in Sections 3.1.4 and 3.2.4, initial assumptions were made for the variables determining the temporal disaggregation of the three different sectors.

- **Residential sector:** HDD T_{base} = 15.5 °C
- **Industrial sector:** Base ratio = 0.4
  Weekday production ratio = 0.5
  Saturday production ratio = 0.1
- **Commercial sector:** HDD T_{base} = 15.5 °C
  Base ratio = 0.2
  Production ratio = 0.3
  Space heating ratio = 0.5

These variables were calibrated through iteration processes, first obtaining the results corresponding to different combinations of variables, and then calculating the **absolute difference error** between the reference D_{ref} and the result D_{res} datasets:

$$E_{\text{diff}} = \sum_{d=1}^{365} \left(\sum_{r=1}^{401} |D_{\text{ref}} - D_{\text{res}}| \right) \quad \textit{(Eq. 22)}$$

where *r* are the 401 German NUTS 3 regions and *d* are the 365 days of the reference year 2015.

The minimum error found for each sector determined the optimal values presented in Table 20:

Sector	HDD T_{base} [°C]	Base ratio [-]	Production ratio [-]	Production on weekdays [-]	Production on Saturdays [-]	Space heating ratio [-]
Residential	16.557	N/A	N/A	N/A	N/A	N/A
Industrial	N/A	0.8288	N/A	0.1666	0.004508	N/A
CTS	14.73	0.288	0.69	N/A	N/A	0.022

Table 20: Optimal values for temporal disaggregation variables.

These optimal values were used to estimate the time series final results.

4.4 Discussion

In this section, the results obtained in the project will be reviewed. First, the temporal and spatial disaggregation results in Germany will be discussed, followed by an assessment of the outcome in a different representative country (i.e., Portugal).

4.4.1 Verification of the temporal disaggregation

In order to verify the temporal disaggregation of each sector, the respective gas consumption of every German NUTS 3 region on every day of 2015 was summed up, obtaining a time series for the entirety of the country.

A comparison of the annual gas demand time series for the residential sector is shown in Figure 24.

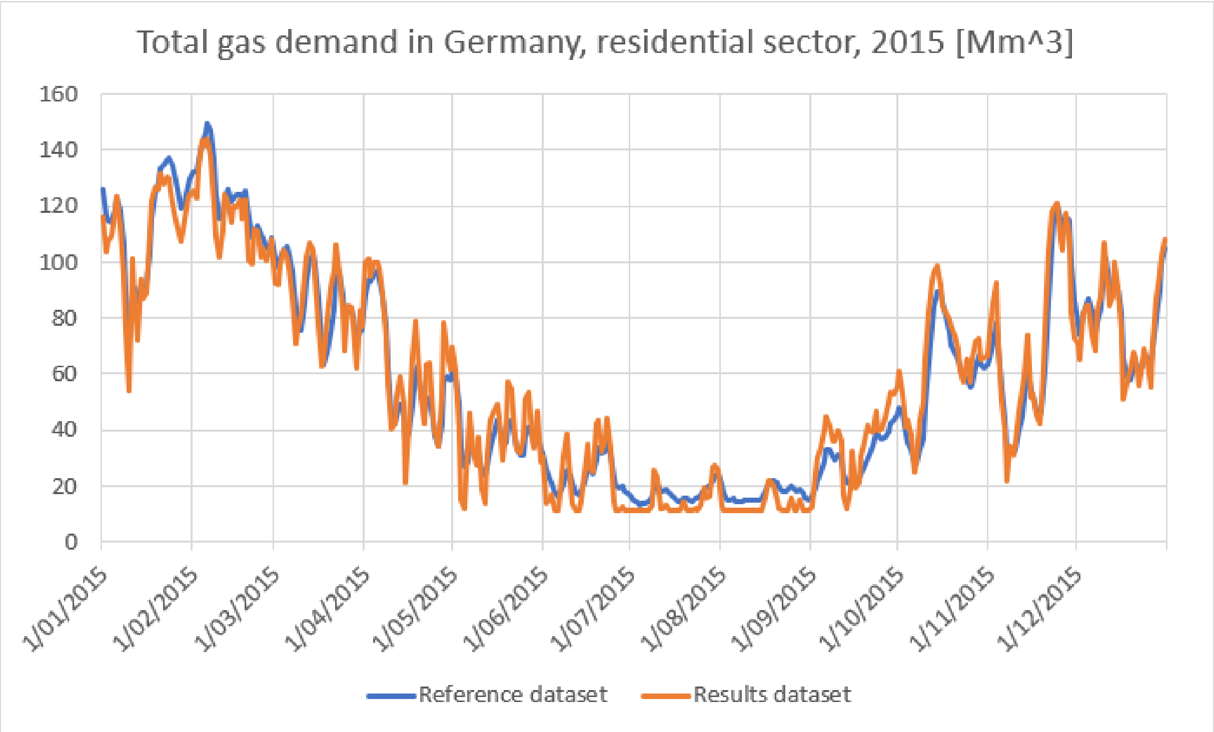


Figure 24: Comparison of household gas demand reference and results time series.

Figure 24 show a strong seasonality in residential gas consumption, with the lowest consumption in the summer months (June to August) and the highest in winter (December to February). This behavior was expected due to the intensive use of gas for space heating in Germany (see Section 3.1.2).

While the results load profile curve closely matches the shape of the validation load profile, there are differences in the daily amount of gas consumed in the German residential sector according to the two datasets. These differences are depicted in Figure 25, which shows that the percent error ranges between -56% and 60%. The distribution of the number of days per range of percent error is represented in Figure 26, which indicates that 105 days (56.4% of the total) present a difference in gas demand ranging between ±15%. The percentage error mean value and the standard deviation have been calculated as μ = -0.0175 and σ = 0.1963, respectively, and the coefficient of variation c_v = 11.22.

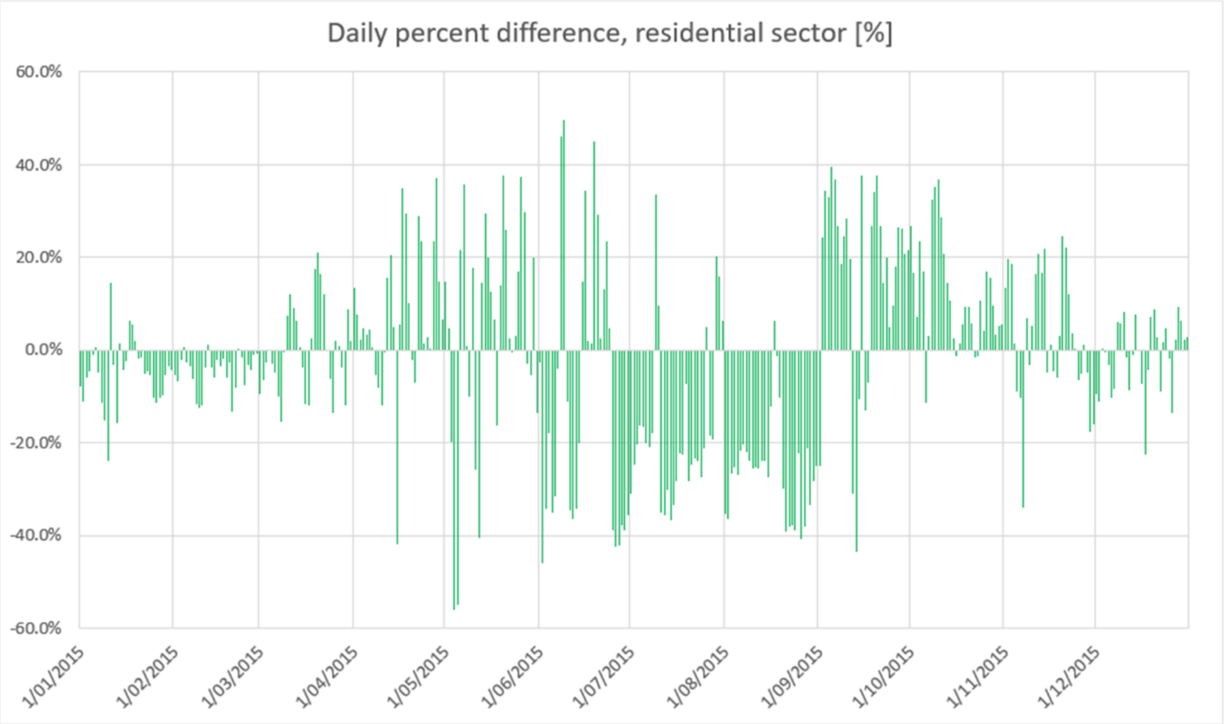


Figure 25: Daily percent difference between reference and results datasets, residential sector.

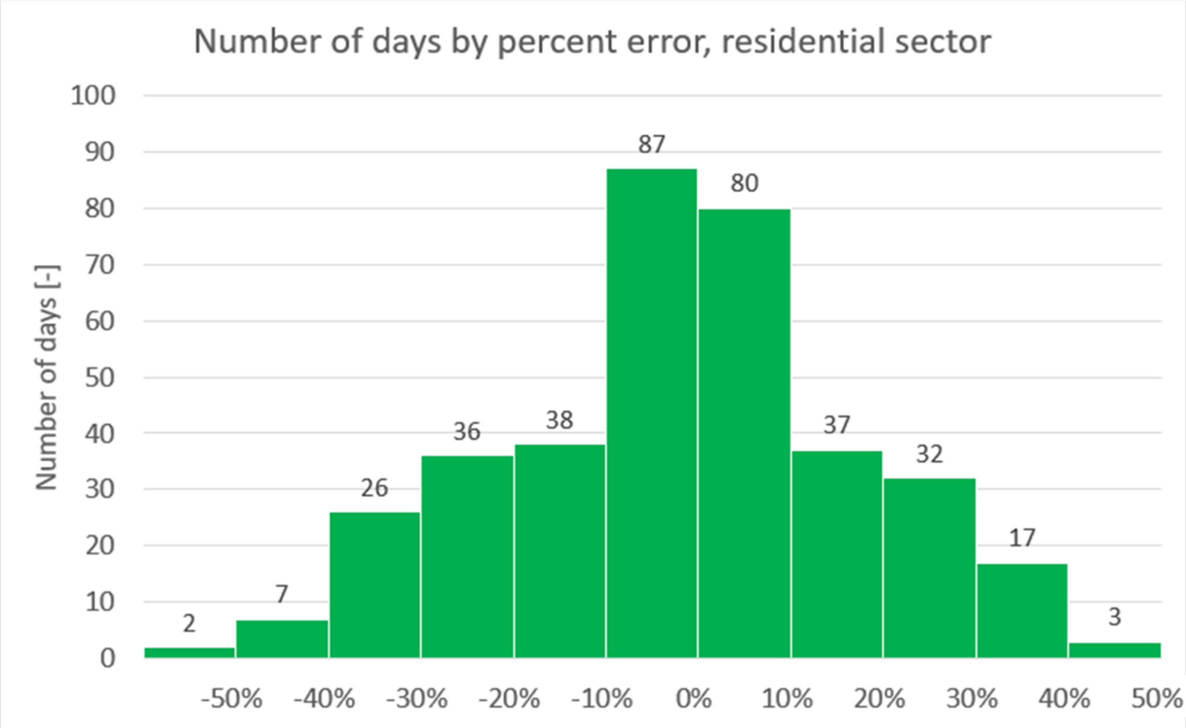


Figure 26: Distribution of days by percent error, residential sector.

The variation in the amount of daily gas consumed in the two projects is possibly explained by the different methodology used for the regional allocation of outside air temperature. As seen in Section 3.1.4.1, in this project, the temperature data obtained from ERA5 was assigned to a NUTS 3 region by averaging the temperature at the coordinates located in its interior. On the other side, the DemandRegio project used the temperature obtained from the German Weather Service (DWD) at either the centroid or the largest settlement of the NUTS 3 region [15, p. 40]. As space heating is the primary use of gas in German households, different assigned temperatures have a decisive effect on regional gas consumption. Figure 25 shows that, in general, DemandRegio's gas demand was higher during the summer months (June to August), while this project estimated a higher gas demand in spring and autumn. The lowest variation in gas consumption is observed in winter.

In a similar way to the residential sector, the seasonal factor plays an essential role in the gas consumption of the commercial sector, as shown in Figure 27, which compares the total gas consumption of the commercial sector in Germany estimated by DemandRegio (blue curve) and this project (in orange). Figure 28 and Figure 29 depict the daily percent difference between the two datasets and the distribution of days by percent difference, respectively.

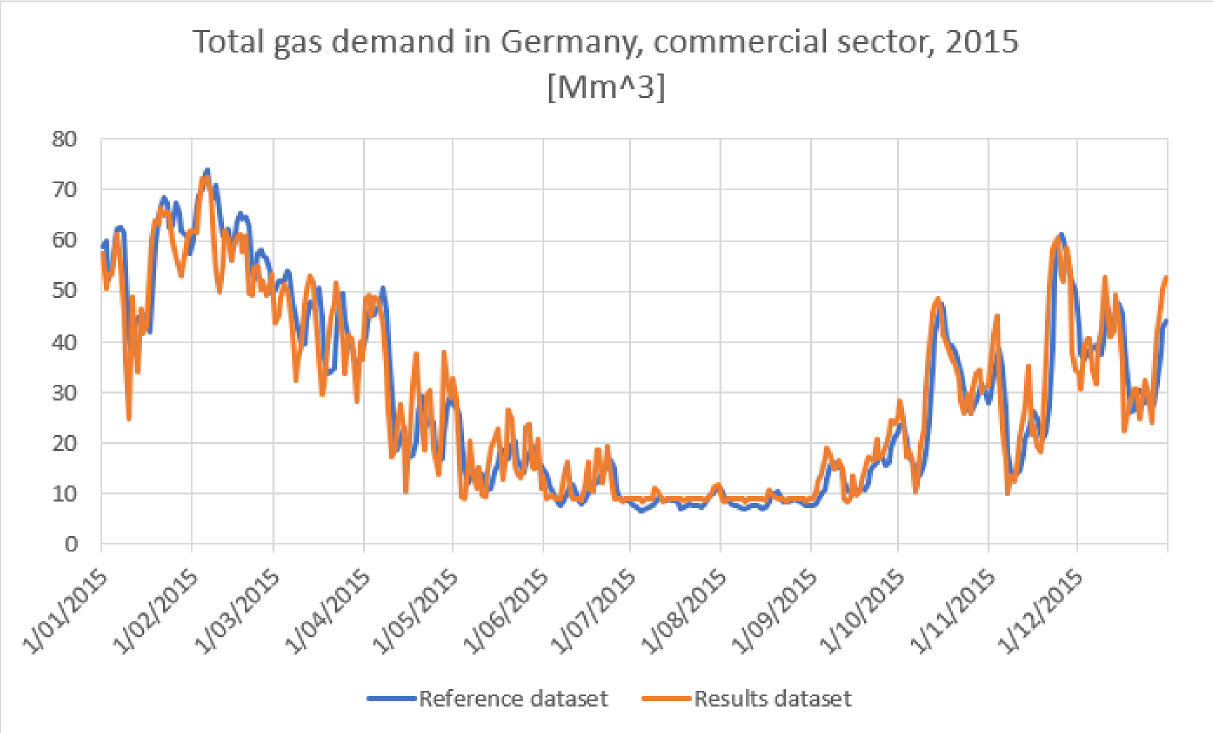


Figure 27: Comparison of CTS gas demand reference and results time series.

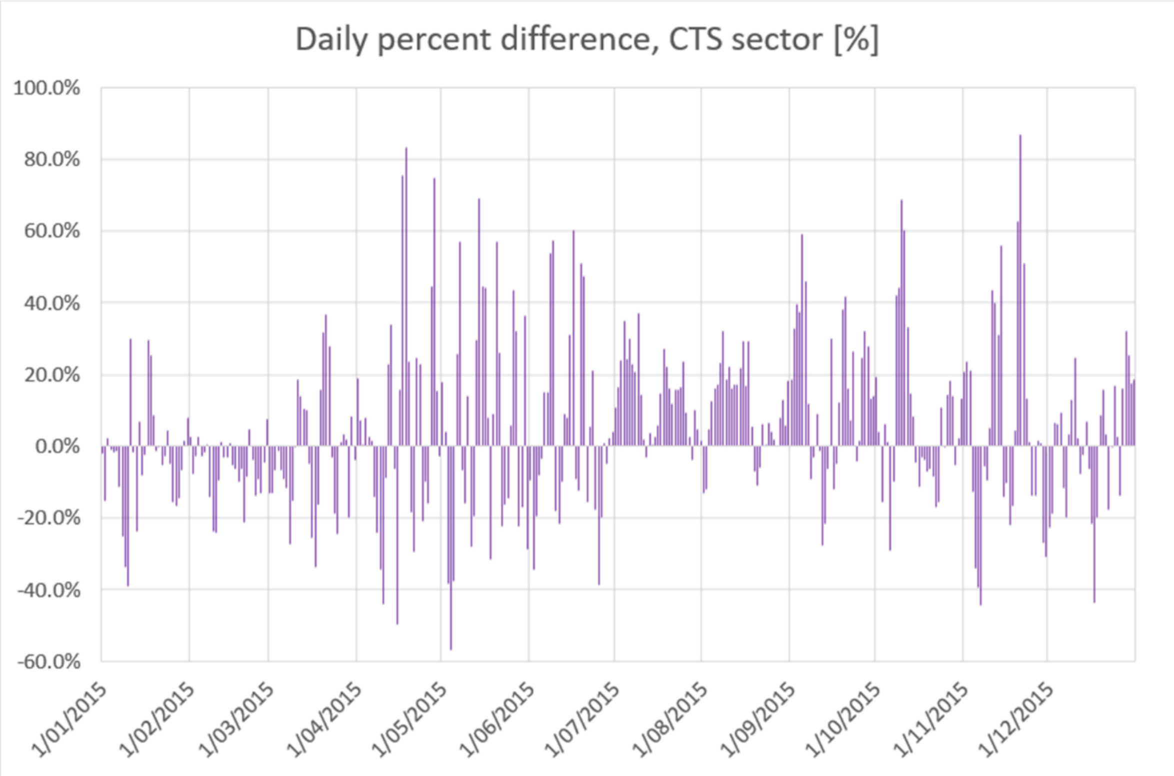


Figure 28: Daily percent difference between reference and results datasets, CTS sector.

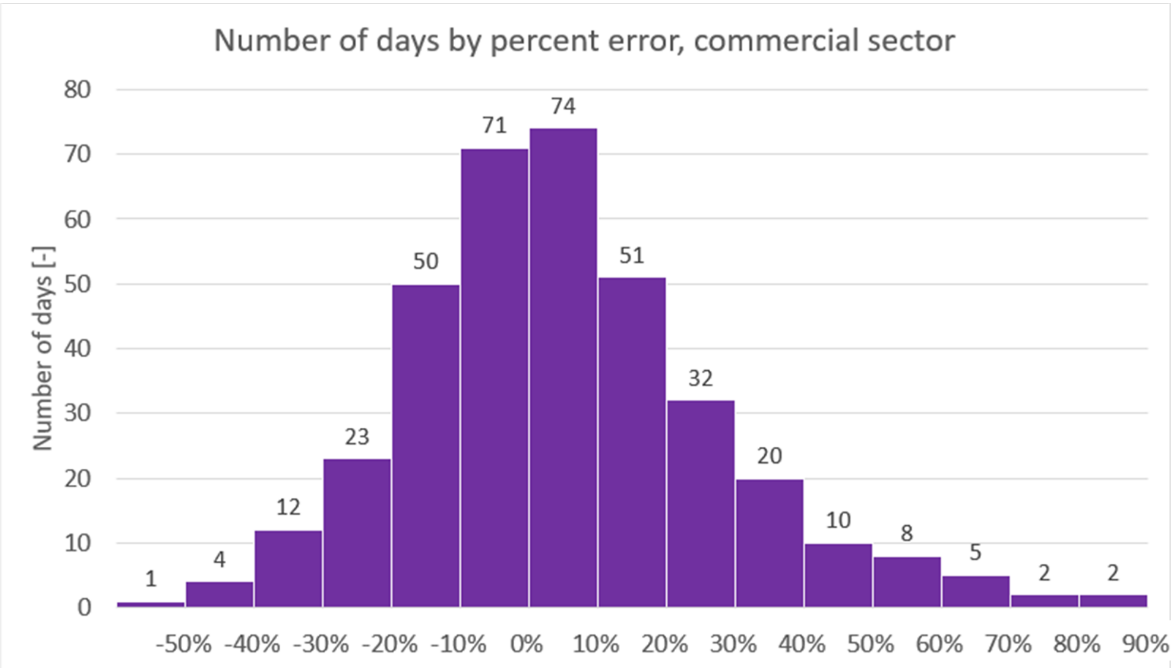


Figure 29: Distribution of days by percent error, CTS sector.

As in the residential sector, Figure 28 shows a similarity in the shape of the curves and differences in the CTS daily gas consumption estimated by the two projects. Again, the different methodologies used for assigning regional temperatures mentioned above explain these differences.

For the commercial sector, the maximum percent difference between the two datasets is 86.9%. The percent error has a mean value μ of 0.051 and a standard deviation σ of 0.2344. With these values, the coefficient of variation $c_v = \sigma/\mu = 4.596$ is calculated, indicating a lower dispersion of the percent difference than in the residential sector.

Finally, Figure 30 illustrates the comparison between the reference dataset for the industrial sector and the results dataset from this project.

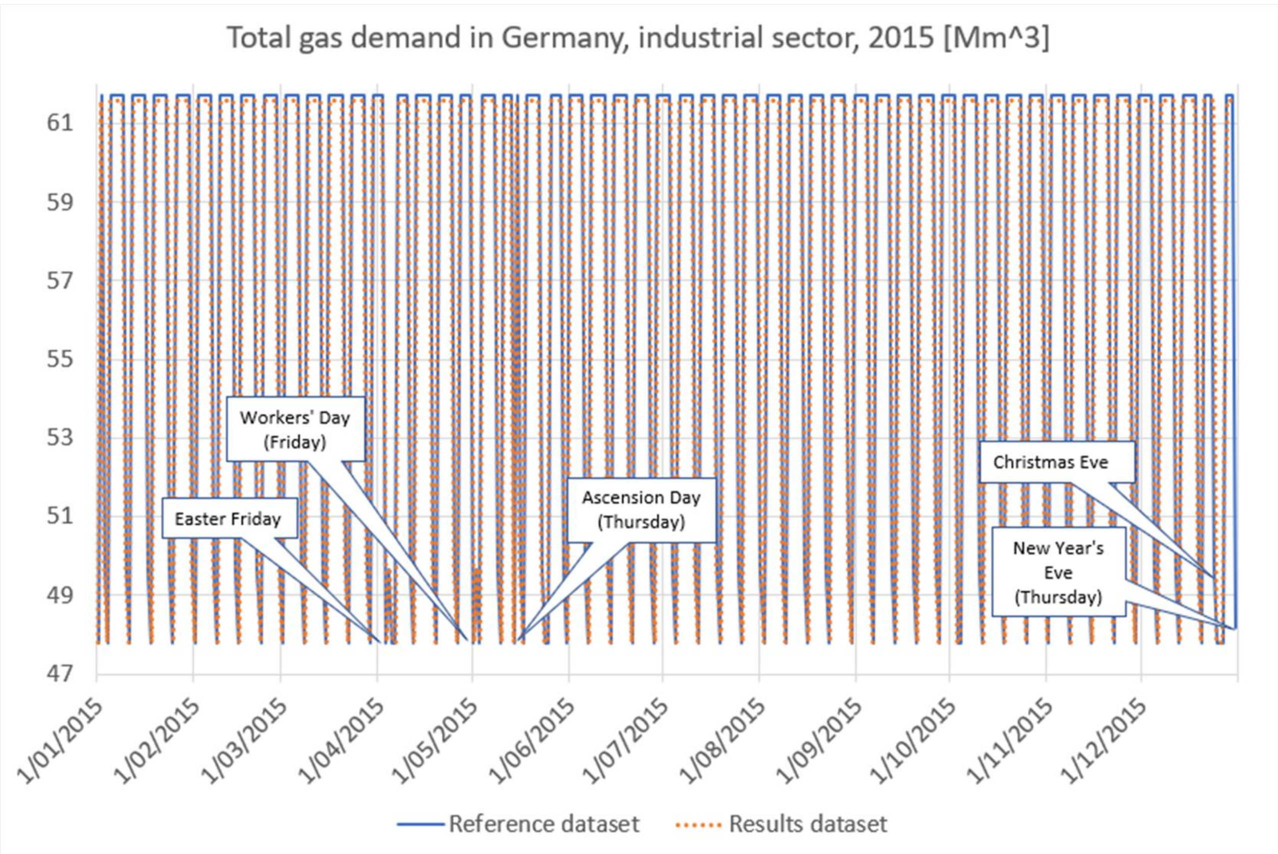


Figure 30: Comparison of industrial gas demand reference and results time series.

Gas consumption in the commercial and the industrial sector is influenced by the type of day. While this impact is not clearly seen in the commercial sector load profile plot (Figure 27), it is easily noticed in the industrial sector, with demand plateaus on weekdays followed by low peaks on weekends. National holidays, such as Easter (four non-working days at the beginning of April), International Workers' Day (May the 1st), and Ascension Day (May 14th) are also observed in irregularities of the weekday-weekend pattern and represented by drops in gas consumption.

Unlike the residential and commercial sectors, outside air temperature doesn't influence gas consumption in the industry sector, as shown by the uniform pattern of its load profile. The independence from temperature also allows a closer match in the shape and the daily gas consumption of both datasets, as depicted in Figure 31. Two low peaks are observed on December 24th and December 31st, as DemandRegio considers those days as semi-holidays.

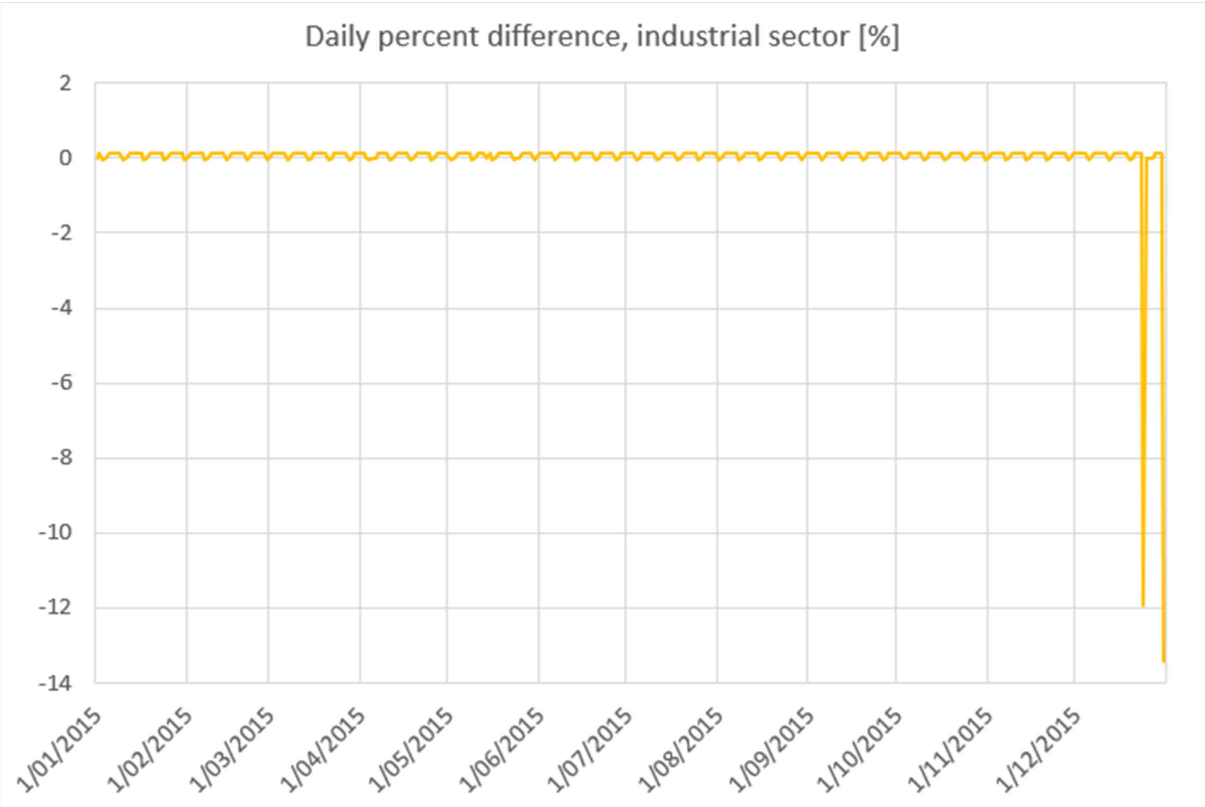


Figure 31: Daily percent difference between reference and results datasets, industrial sector.

4.4.2 Verification of the spatial disaggregation

In order to verify the spatial disaggregation of gas from national to regional level, the gas demand ratio between the reference and the results datasets was calculated in all three sectors.

The choropleth map in Figure 32 shows the ratios in the German NUTS 3 regions for the residential sector:

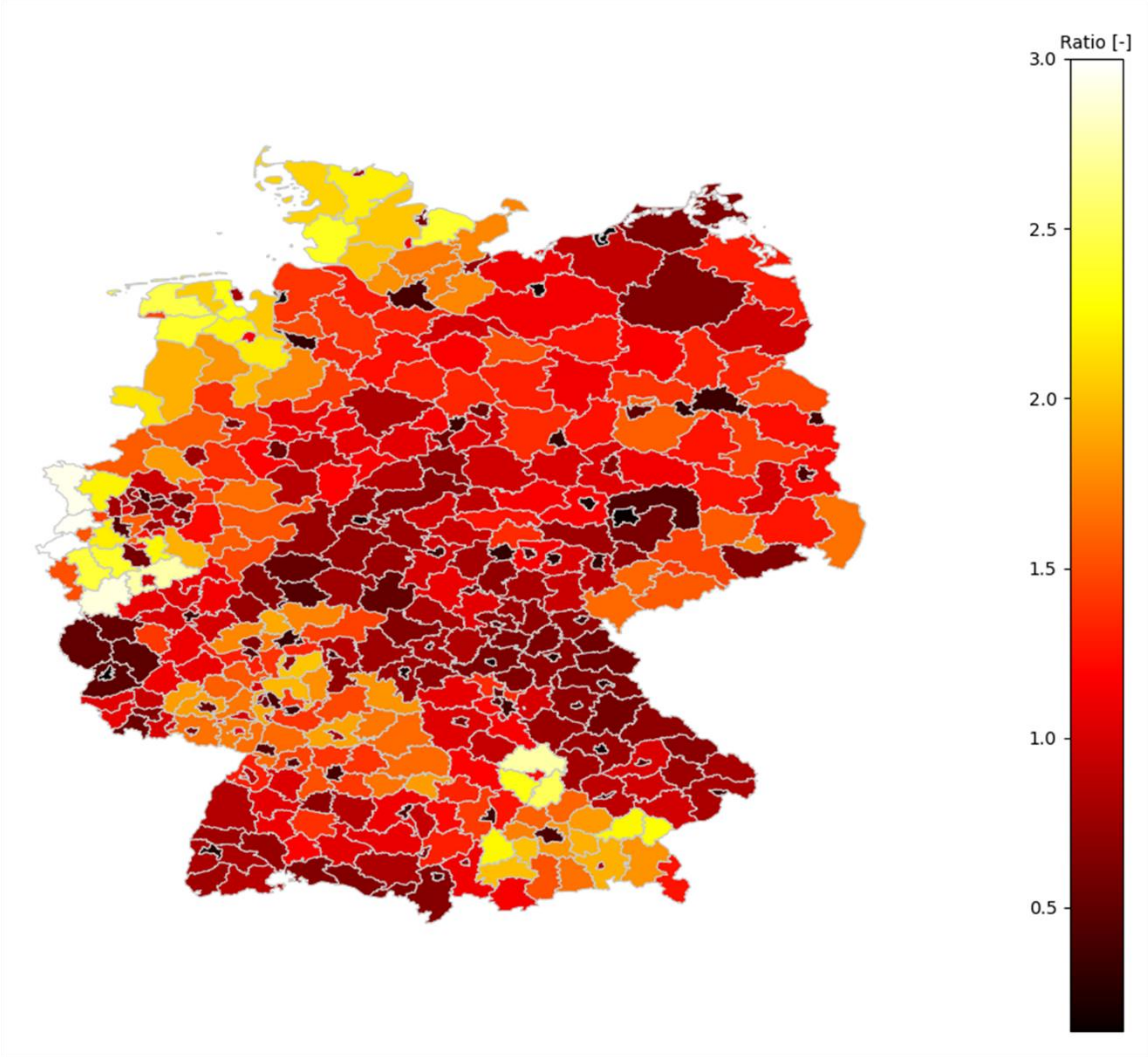


Figure 32: Reference/results ratio in German NUTS 3 regions, residential sector.

The distribution of the NUTS 3 regions by ratios can be observed in Figure 33:

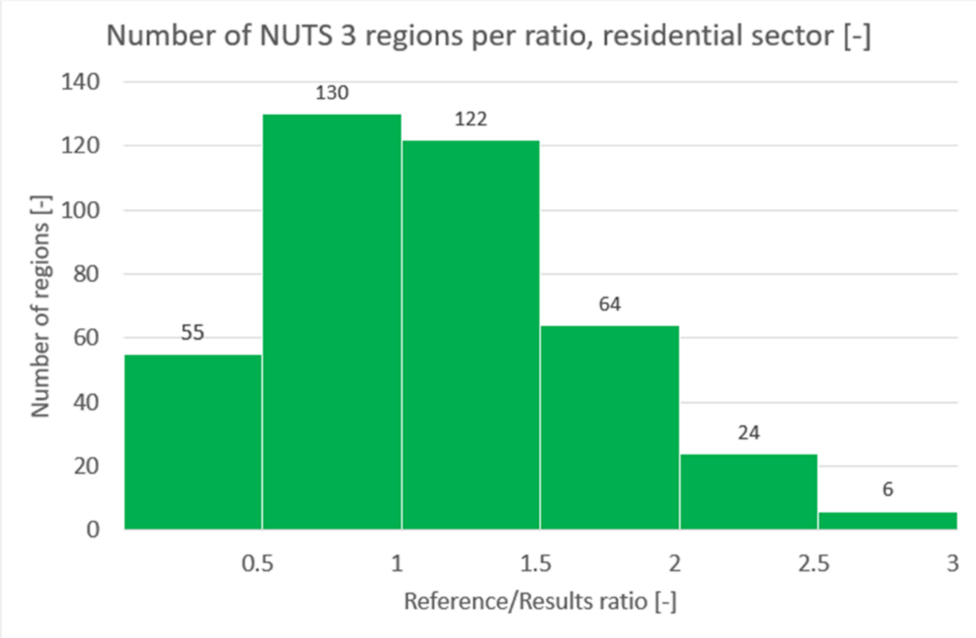


Figure 33: Number of NUTS 3 regions per ratio, residential sector.

As seen in Figure 33, most NUTS 3 regions (252, corresponding to 62.8%) are within a ratio of 0.5 to 1.5. The highest ratios are concentrated close to the borders with Austria (particularly in the surrounding areas around Munich and Ingolstadt) and the Netherlands. On the other side, state capitals (Berlin, Hamburg, Stuttgart, Munich), as well as the area around Trier show a significantly lower gas consumption in the DemandRegio project, compared to the results obtained.

The NUTS 3 regions with the highest gas demand in Germany in 2015 for the DemandRegio project are shown in Table 21:

NUTS 3 code	Region name	DemandRegio gas demand [Mm^3]	Project results gas demand [Mm^3]	Reference/Results ratio [-]
DED21	Dresden	267.22	154.79	1.72
DE929	Region Hannover	217.47	257.69	0.84
DEA2C	Rhein-Sieg-Kreis	215.92	79.52	2.71
DED42	Erzgebirgskreis	207.36	133.77	1.55
DED45	Zwickau	199.77	124.97	1.59

Table 21: Regions with highest residential gas consumption in Germany, 2015 (DemandRegio project).

Table 22 shows the regions with the highest gas demand from the results obtained:

NUTS 3 code	Region name	DemandRegio gas demand [Mm³]	Project results gas demand [Mm³]	Reference/Results ratio [-]
DE300	Berlin	193.28	547.73	0.35
DE600	Hamburg	183.56	463.73	0.39
DE501	Bremen	107.56	341.23	0.31
DE929	Region Hannover	217.47	257.69	0.84
DED51	Leipzig	32.36	235.07	0.13

Table 22: Regions with highest residential gas consumption in Germany, 2015 (Results).

It can be observed that, except from Region Hannover (DE929), the regions with highest demand do not coincide between the two datasets and that the estimated gas demand is significantly different.

As seen in Section 3.1.3, the regionalization of gas demand in households depends on demographical factors (population, number of households) as well as on dwelling area. Therefore, it would be expected higher gas consumption in highly populated areas, as the results obtained show. However, the reference dataset presents a high residential consumption in NUTS 3 regions with considerably fewer inhabitants than main urban areas.

Figure 34 shows the reference to results ratios in NUTS 3 regions for the industrial sector:

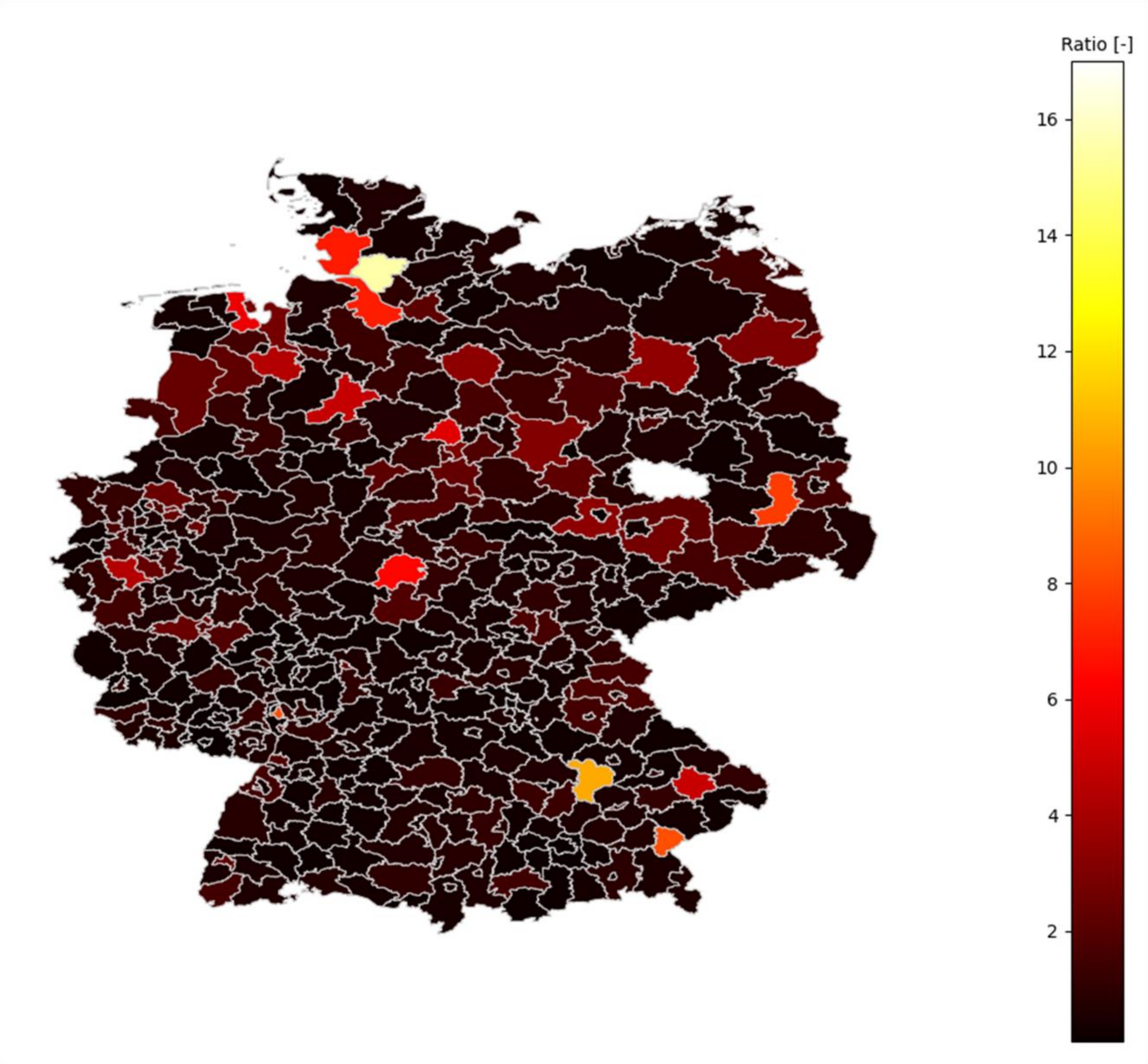


Figure 34: Reference/results ratio in German NUTS 3 regions, industrial sector.

The predominance of dark colors indicates two things. First, according to DemandRegio, several regions (such as DEE0E, in white; DEF0E, in yellow; and DE226, in orange) have a very high gas consumption. Second, that, in general, DemandRegio assigns a lower gas demand to most regions when compared to the results obtained. The distribution of NUTS 3 regions according to their ratios is presented more clearly in Figure 35.

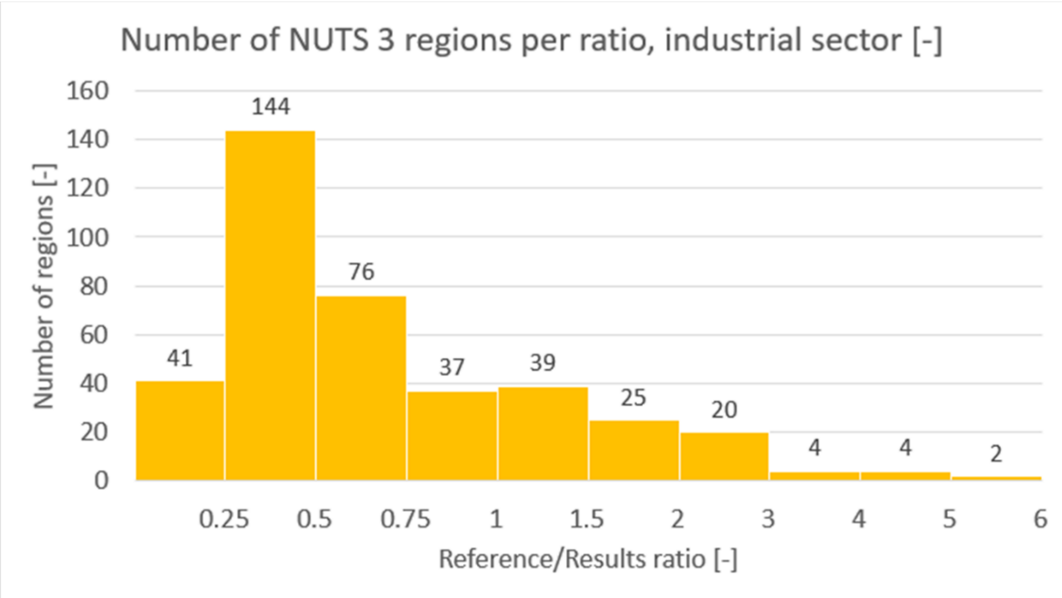


Figure 35: Number of NUTS 3 regions per ratio, industrial sector.

The distribution of NUTS 3 regions according to their ratio is almost evenly divided in two ranges: 152 regions (38%) have ratios between 0.5 and 1.5, while 144 (36%) are between 0.25 and 0.5. Meanwhile, a number of regions present very high ratios, being 16.5 the highest. The distribution indicates that the results obtained are more evenly distributed in NUTS 3 regions, compared to the reference dataset that has values ranging from very low to very high regional gas demand.

The NUTS 3 regions with the highest gas demand in the reference and the results datasets are shown in Table 23 and Table 24, respectively.

NUTS 3 code	Region name	DemandRegio gas demand [Mm^3]	Project results gas demand [Mm^3]	Reference/Results ratio [-]
DEB34	Ludwigshafen am Rhein	2023.16	251.84	8.03
DEE0E	Wittenberg	822.55	49.57	16.59
DEA23	Köln	535.18	227.23	2.35
DE600	Hamburg	476.09	226.14	2.1
DE214	Altötting	425.74	53	8.03

Table 23: Regions with highest industrial gas consumption in Germany, 2015 (DemandRegio project)

NUTS 3 code	Region name	DemandRegio gas demand [Mm^3]	Project results gas demand [Mm^3]	Reference/Results ratio [-]
DE212	München	70.58	305.82	0.23
DEB34	Ludwigshafen am Rhein	2023.16	251.84	8.03
DE300	Berlin	95.24	250.07	0.38
DEA1C	Mettmann	55.65	248.72	0.22
DEA58	Märkischer	120.61	241.93	0.49

Table 24: Regions with highest industrial gas consumption in Germany, 2015 (Results).

Of the five regions with the highest gas consumption, only Ludwigshafen am Rhein (DEB34) appears in both datasets, although with a vast difference in amount. As the headquarters of BASF (the largest German chemical company, an energy-intensive type of industry) are located in Ludwigshafen, a large gas demand is foreseeable. Similarly, DemandRegio assigns a large gas demand to Wittenberg (DEE0E), the region with the largest difference in consumption between the results and the reference figures, and the home of SKW Piesteritz, the main producer of ammonia and urea in Germany.

The other four regions with highest gas consumption in the results dataset present a combination in industrial activities. As an example, regions DEA1C and DEA58 belong to the Rhine-Ruhr area, one of the ten largest economic regions in the world [65], and are located in the proximities of Dusseldorf, where energy-intensive activities such as the production of iron, steel, and chemicals [66] are developed.

Lastly, Figure 36 depicts the reference-to-result ratios in Germany for the commercial sector.

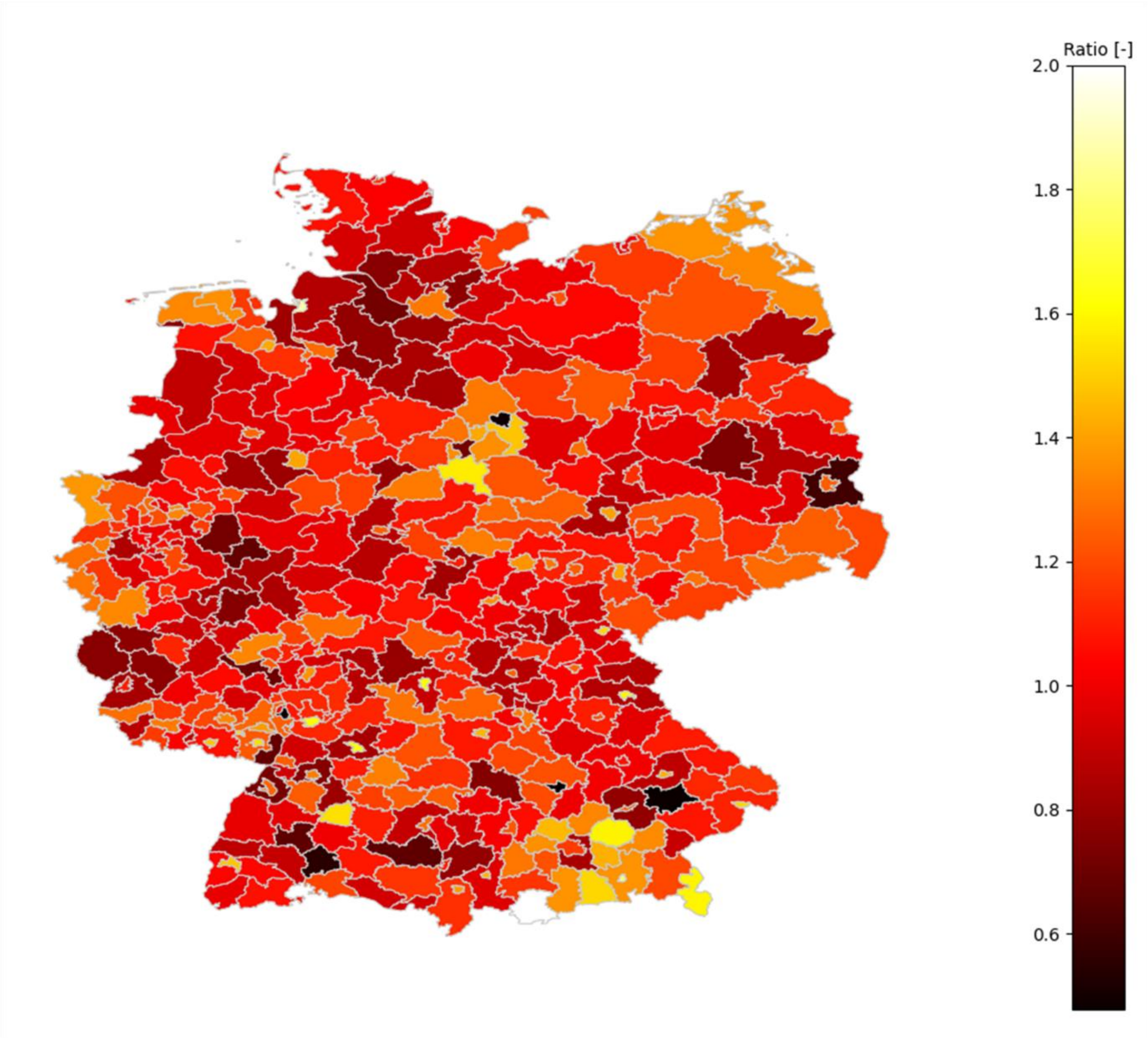


Figure 36: Reference/results ratio in German NUTS 3 regions, commercial sector.

In the commercial sector, the minimum and maximum ratios of all 401 German NUTS 3 regions are 0.47 and 1.79. This indicates that the results from both datasets are more similar than in the residential and the industrial sector.

The distribution of NUTS 3 regions for the commercial sector can be observed in Figure 37.

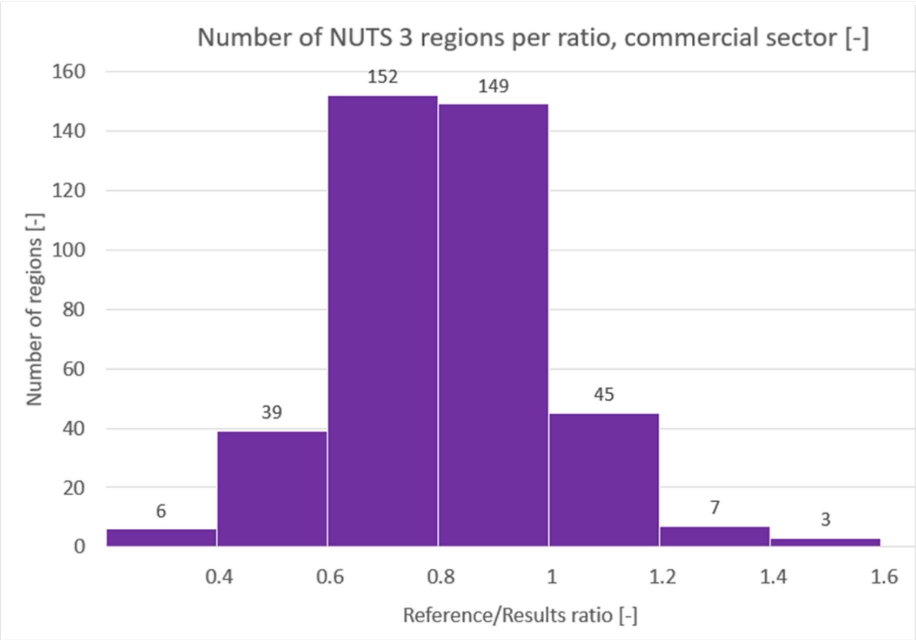


Figure 37: Number of NUTS 3 regions per ratio, commercial sector.

The histogram in Figure 37 shows that 301 NUTS 3 regions, 75% of the total, have ratios among 0.8 and 1.2, indicating a very similar region-to-region gas demand between the two datasets.

Table 25 and Table 26 show the regions with the highest gas consumption in the commercial sector for both projects.

NUTS 3 code	Region name	DemandRegio gas demand [Mm^3]	Project results gas demand [Mm^3]	Reference/Results ratio [-]
DE300	Berlin	554.21	526.13	1.05
DE600	Hamburg	338.95	283.37	1.19
DE212	München	304.49	283.53	1.07
DEA23	Köln	201.41	207.13	0.97
DE712	Frankfurt	197.10	198.52	0.99

Table 25: Regions with highest commercial gas consumption in Germany, 2015 (DemandRegio project).

NUTS 3 code	Region name	DemandRegio gas demand [Mm^3]	Project results gas demand [Mm^3]	Reference/Results ratio [-]
DE300	Berlin	554.21	526.13	1.05
DE212	München	304.49	283.53	1.07
DE600	Hamburg	338.95	283.37	1.19
DEA23	Köln	201.41	207.13	0.97
DE712	Frankfurt	197.10	198.52	0.99

Table 26: Regions with highest commercial gas consumption in Germany, 2015 (Results).

From Table 25 and Table 26 it can be observed that the five German NUTS regions with the highest gas consumption in the CTS sector are present in both datasets. The percentage error is relatively low, confirming the similarity in regional gas demand observed in Figure 37.

4.4.3 Assessment of gas demand in Portugal

As described in Section 3.1.2, the amount of gas used for different households' applications varies from country to country. For example, while German homes uses gas mostly for space and water heating (78.9% and 20.7% respectively in 2018; see Table 11), Portugal uses most gas for water heating and cooking (61.8% and 35.1% in 2018; see Table 11), leaving only 3.1% of the gas in Portuguese households for space heating. Due to these differences and the fact that the two countries belong to different geographical regions, the gas consumption of Portugal's three sectors will be analyzed in this section.

Figure 38 shows the total gas consumption in the residential sector in Portugal in 2018.

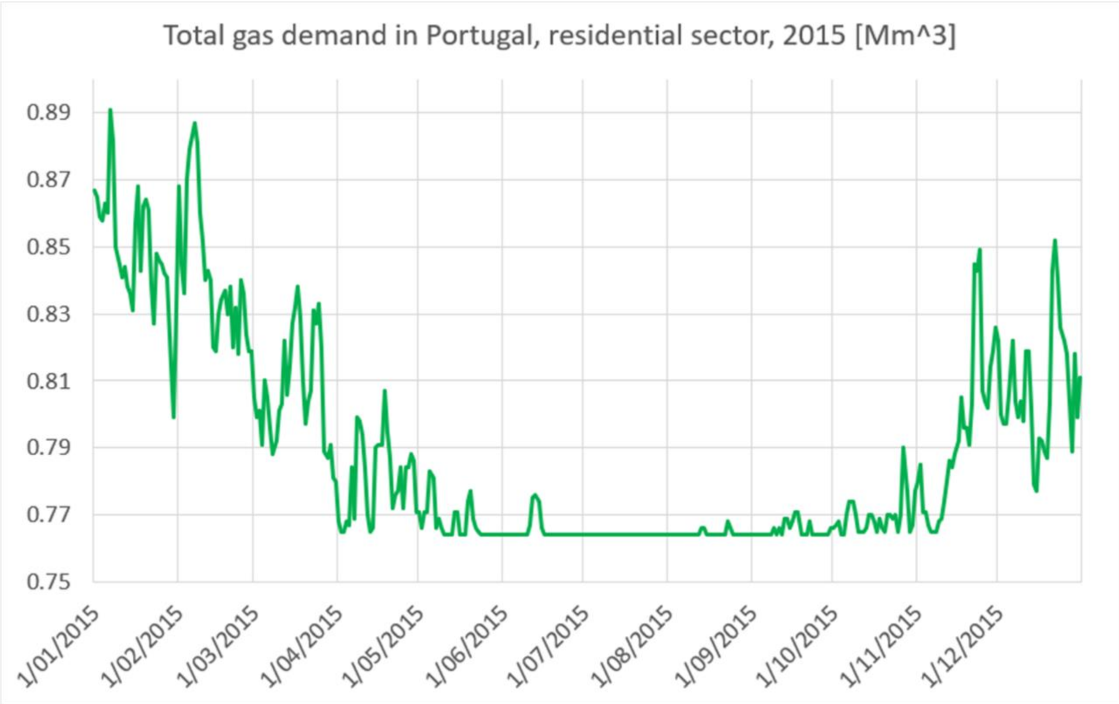


Figure 38: Total gas demand in households, Portugal, 2015.

As less gas is used for space heating, Portugal's residential gas consumption is less influenced by the seasonal effect, as it can be seen when compared to Germany. Figure 24 shows that, for Germany, the peak-to-base ratio is approximately 12.86 (144.176/11.21 = 12.86) and for Portugal, this ratio is only 1.16 (0.891/0.764 = 1.16). The higher average temperature characteristic of Southern European countries like Portugal is also reflected in longer periods of time close to baseload values (July and August in Germany, June to October in Portugal).

In Portugal, the highest consumption of gas in households happens in the regions PT170 (Lisboa), PT11A (Porto), and PT150 (Algarve), some of the most populated urban centers or regions in the country (Algarve in particular acts as a NUTS 2 and NUTS 3 region).

The annual gas demand for the commercial sector in Portugal in 2015 can be observed in Figure 39.

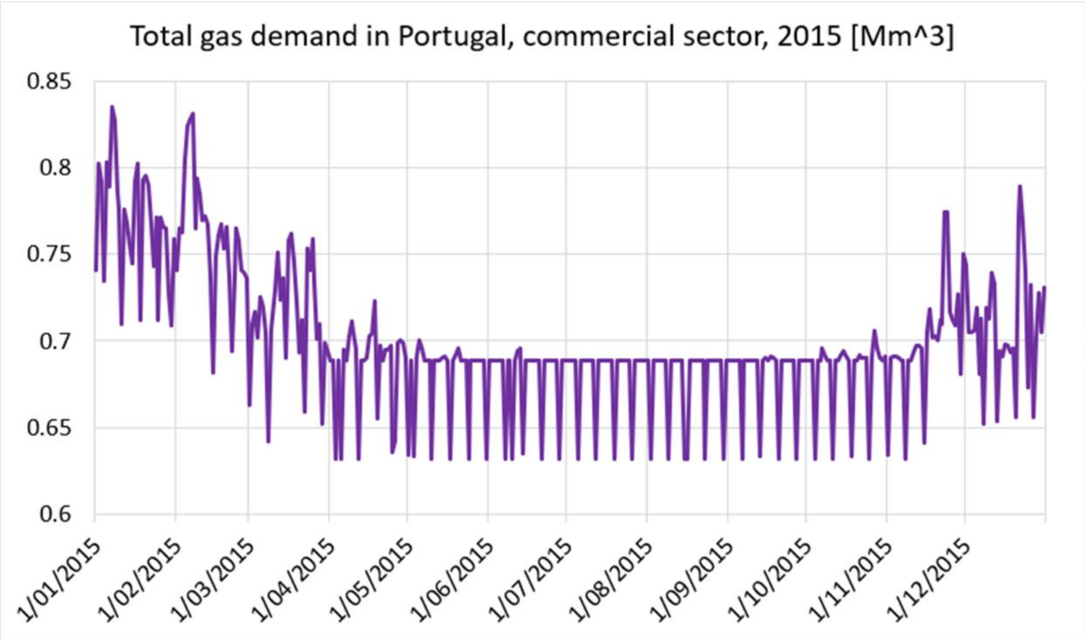


Figure 39: Total gas demand in the CTS sector, Portugal, 2015.

While the effect of the day type on the German CTS sector was not easily noticeable (see Figure 27), the general lack of impact of temperature on gas demand during the summer months in Portugal clearly shows consumption plateaus on weekdays and Saturdays followed by lower gas demand on Sundays.

As in the residential sector, the highest gas demand for the commercial sector in Portuguese NUTS 3 regions occurs in Lisboa, Porto and Algarve, with the latter being a popular touristic destination with a strong infrastructure of hotels, resorts and other hospitality services.

The estimated load profile for the industry sector in Portugal is shown in Figure 40.

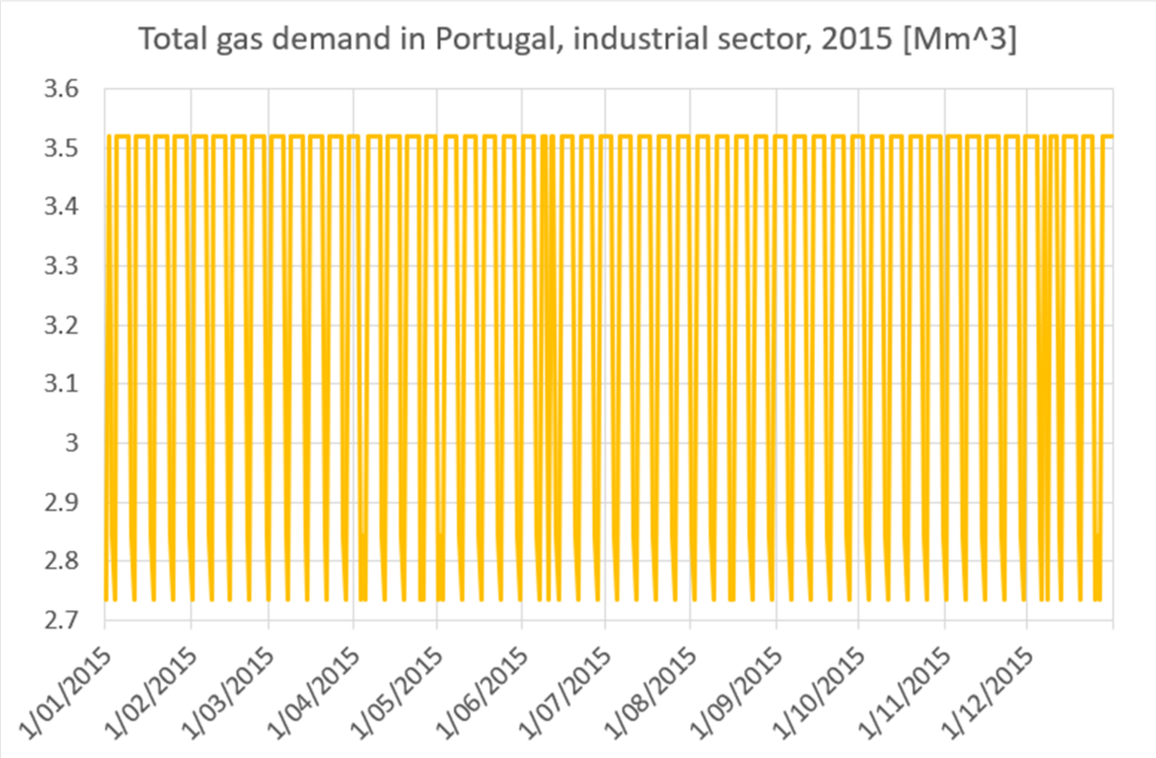


Figure 40: Total gas demand in the industrial sector, Portugal, 2015.

As in the German industry gas consumption annual time series seen in Figure 30, the total gas demand for Portugal's industry sector does not depend on the temperature but only on the day type, with high gas demand plateaus observed on weekdays and lower consumption on Saturdays, Sundays and holidays such as Easter Friday (April 3rd), May Day (May 1st), and Portugal Day (June 10th).

The results show that the highest gas demand in Portugal for the industry sector is located in the NUTS 3 regions PT11A (Porto), PT170 (Lisboa), and PT16D (Aveiro). The two latter regions account for three-fourths of Portugal's net industrial output with chemical, metallurgic, machinery, and paper activities, and the former with textile and food processing industries [67].

5. Conclusions

The objective of this thesis consisted in the estimation of a gas demand time series from 2010 to 2019 for the European NUTS 3 regions. Three sectors were considered: residential, industrial, and commercial.

To achieve this goal, two main tasks were performed: first, the spatial disaggregation of gas demand from national to regional levels, and second, the temporal disaggregation of regional annual gas demand for the generation of a daily time series.

The spatial disaggregation of gas was performed taking a top-down approach, on which a series of demographic statistics were used to estimate gas demand at NUTS 3 level, taking into consideration the individual demand drivers of each analyzed sector.

In the residential sector, the statistics show that gas is mainly used for space heating purposes. Therefore, temperature and geographical location play a direct role in the gas consumption of households. Combined with other demand variables, the results show that a large amount of the gas consumed in the residential sector is concentrated in Italian and British regions located in low-temperature areas, as well as major European urban centers and their surroundings.

In the industry sector, it was found that the highest gas consumption occurred in important manufacturing centers in Spain and the metropolitan area around Paris. Smaller regions with energy-intensive industries such as chemical, paper, and metallurgic, had also a higher share in gas consumption.

The largest figures in the commercial sector were found in capital cities in Italy, Spain, Germany, and Netherlands, as they gather a high share of activities in the tertiary sector of the economy.

The way some countries or regions are divided by the NUTS classification may be misleading regarding their gas demand. For example, while Germany is the main consumer of gas in the industrial sector, its high number of NUTS 3 regions compared to the European total suggests a reduced regional gas consumption. This also happens in the residential and the commercial sector; while cities like Berlin, Madrid, or Rome have their own individual NUTS 3 region, the metropolitan area of cities like Paris or London is split in several NUTS 3 regions, conveying the impression of a lower overall gas consumption.

The temporal disaggregation of gas demand was also performed with a top-down approach, from an annual to a daily resolution, and with different considerations and initial assumptions for each sector.

The load profiles in the residential and commercial sector show a strong seasonal effect in their gas consumption, with the highest figures during winter and the lowest during summer. The seasonal effect is especially noticeable in countries with high gas consumption for space heating and average lower temperatures, such as Germany, Netherlands or United Kingdom. In South European countries, where average temperatures are higher, the seasonality of gas consumption is less pronounced, but still present. As the industrial sector is considered to be independent from temperature, there is no seasonal effect observed, and the load profiles were found to be highly uniform, with variations in gas consumption depending on the type of day.

The validation of the results against the DemandRegio datasets showed strong similarities in the regional gas demand of the commercial sector and in the temporal gas demand of the industry sector. Differences were observed in the spatial disaggregation of the residential and the industrial sector. In the former, the reference dataset assigns lower gas consumption to urban areas compared to the results obtained by this thesis, and in the latter, the outcome of the DemandRegio project depicts regions with extremely high gas consumption, while the results show a more even distribution of gas for industrial purposes in NUTS 3 regions.

In the temporal disaggregation, and while the curves of the load profiles in the residential and commercial sector show a similar shape, there are discrepancies in the daily gas demand values. As mentioned before, these differences are caused by the different approach taken when assigning temperatures to individual NUTS 3 regions.

A future availability of relevant statistics, especially at a high-resolution regional level, may lead to the improvement of the results of this thesis. Datasets regarding gas used for power production, average dwelling area, building age, household numbers, and employment in the industrial sector separated by branch of activity, would be particularly beneficial for gas demand modelling. A new version of the

European Union Census, which will take place during 2021, will likely address some of these statistical gaps and allow an enhanced regionalization of gas demand in Europe.

6. References

[1] World Wildlife Fund for Nature, "International Work on Climate Change," WWF, [Online]. Available: https://www.wwf.org.uk/what-we-do/projects/international-work-climate-change. [Accessed 18 05 2021].

[2] European Comission, "2030 climate & energy framework," European Comission, [Online]. Available: https://ec.europa.eu/clima/policies/strategies/2030_en. [Accessed 18 05 2021].

[3] European Environment Agency, "Renewables successfully driving down carbon emissions in Europe," European Environment Agency, 13 01 2017. [Online]. Available: https://www.eea.europa.eu/highlights/renewables-successfully-driving-down-carbon.

[4] D. Vetter, "It's Official: In 2020, Renewable Energy Beat Fossil Fuels Across Europe," Forbes, 25 01 2021. [Online]. Available: https://www.forbes.com/sites/davidrvetter/2021/01/25/its-official-in-2020-renewable-energy-beat-fossil-fuels-across-europe/.

[5] K. Appunn, "Sector coupling - Shaping an integrated renewable energy system," CleanEnergyWire, 25 04 2018. [Online]. Available: https://www.cleanenergywire.org/factsheets/sector-coupling-shaping-integrated-renewable-power-system.

[6] Next Kraftwerke, "What Is Sector Coupling?," Next Kraftwerke, [Online]. Available: https://www.next-kraftwerke.com/knowledge/sector-coupling. [Accessed 25 06 2021].

[7] ProGeo, "Power-to-gas system enables massive storage of renewable energy," ProGeo, 24 01 2020. [Online]. Available: https://cordis.europa.eu/article/id/413313-power-to-gas-system-enables-massive-storage-of-renewable-energy.

[8] W. contributors, "Energy modeling," Wikipedia, The Free Encyclopedia., 15 06 2021. [Online]. Available: https://en.wikipedia.org/wiki/Energy_modeling.

[9] SciGRID_gas, "Motivation and Approach," SciGRID_gas, [Online]. Available: https://www.gas.scigrid.de/pages/motivation-and-approach.html. [Accessed 18 05 2021].

[10] ENTSOG, "EntsoG Transparency Platform," European Network of Transmission System Operators for Gas, [Online]. Available: https://transparency.entsog.eu/#/map. [Accessed 18 05 2021].

[11] Forschungsstelle für Energiewirtschaft e.V., "Harmonization and development of methods for a spatial and temporal resolution of energy demands (DemandRegio)," FfE, [Online]. Available: https://www.ffe.de/en/topics-and-methods/production-and-market/736-harmonization-and-development-of-methods-for-a-spatial-and-temporal-resolution-of-energy-demands-demandregio. [Accessed 18 05 2021].

[12] Eurostat, "What kind of energy do we consume in the EU?," Eurostat, [Online]. Available: https://ec.europa.eu/eurostat/cache/infographs/energy/bloc-3a.html.

[13] U.S. Energy Information Administration, "Glossary," EIA, [Online]. Available: https://www.eia.gov/tools/glossary/. [Accessed 18 05 2021].

[14] Eurostat, "Complete energy balances," 08 02 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nrg_bal_c&lang=en.

[15] F. Gotzens, B. Gillessen, S. Burges, W. Hennings, J. Müller-Kirchenbauer, S. Seim, P. Verwiebe, T. Schmid, F. Jetter and T. Limmer, "DemandRegio: Harmonisierung und Entwicklung von Verfahren zur regionalen und zeitlichen Auflösung von Energienachfragen," 2020.

[16] Eurostat, "Background," [Online]. Available: https://ec.europa.eu/eurostat/web/nuts/background.

[17] Eurostat, "Principles and characteristics," [Online]. Available: https://ec.europa.eu/eurostat/web/nuts/principles-and-characteristics.

[18] "Nomenclature of Territorial Units for Statistics," Wikimedia Foundation, 22 04 2021. [Online]. Available: https://en.wikipedia.org/wiki/Nomenclature_of_Territorial_Units_for_Statistics.

[19] Destatis, "NUTS classification," Statistisches Bundesamt, [Online]. Available: https://www.destatis.de/Europa/EN/Methods/Classifications/OverviewClassification_NUTS.html.

[20] EUR-Lex, "Document 02003R1059-20180118," European Union Law, 18 01 2018. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:02003R1059-20180118&qid=1519136585935.

[21] Eurostat, "Energy Data," Eurostat, [Online]. Available: https://ec.europa.eu/eurostat/web/energy/data.

[22] Eurostat, "Disaggregated final energy consumption in households - quantities," 06 05 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nrg_d_hhq&lang=en.

[23] Eurostat, "Population on 1 January by age group, sex and NUTS 3 region," 11 03 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_r_pjangrp3&lang=en.

[24] Eurostat, "Number of households by degree of urbanisation and NUTS 2 regions (1 000)[lfst_r_lfsd2hh]," 08 02 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfst_r_lfsd2hh&lang=en.

[25] Eurostat, "Average size of dwelling by household type and degree of urbanisation," 08 02 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_hcmh02&lang=en.

[26] Eurostat, "Urban-Rural Typology," 17 01 2019. [Online]. Available: https://ec.europa.eu/eurostat/documents/35209/35256/Urban-rural-NUTS-2016.xlsx.

[27] Eurostat, " Employment by age, economic activity and NUTS 2 regions (NACE Rev. 2)," Eurostat, 14 04 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfst_r_lfe2en2&lang=en.

[28] Eurostat, "SBS data by NUTS 2 regions and NACE Rev. 2 (from 2008 onwards)," Eurostat, 16 03 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=sbs_r_nuts06_r2&lang=en.

[29] Eurostat, "Employment (thousand persons) by NUTS 3 regions," Eurostat, 30 04 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_3empers&lang=en.

[30] Eurostat, "Gross domestic product (GDP) at current market prices by NUTS 3 regions," Eurostat, 31 03 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_3gdp&lang=en.

[31] European Comission, "Eurostat - European statistics," European Comission, [Online]. Available: https://ec.europa.eu/info/departments/eurostat-european-statistics_en. [Accessed 18 05 2021].

[32] Copernicus Climate Change Service, "ERA5 hourly data on single levels from 1979 to present," C3S, 14 06 2018. [Online]. Available: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form.

[33] European Centre for Medium-Range Weather Forecasts, "ERA5 hourly data on single levels from 1979 to present - Overview," ECMWF, 14 06 2018. [Online]. Available: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview.

[34] Northwest Regional Modeling Consortium, "Surface (2m) temperature," Northwest Regional Modeling Consortium, [Online]. Available: https://a.atmos.washington.edu/wrfrt/descript/pages/tsfc.html. [Accessed 18 05 2018].

[35] R. S and M. Montel, "Holidays 0.11.1," 02 04 2021. [Online]. Available: https://pypi.org/project/holidays/.

[36] T. Stobierski, "What is Statistical Modeling For Data Analysis?," Northeast University Graduate Programs, 29 10 2019. [Online]. Available: https://www.northeastern.edu/graduate/blog/statistical-modeling-for-data-analysis/.

[37] I. Ghalehkhondabi, E. Ardjmand, G. Weckman and W. Young II, "An overview of energy demand forecasting methods," Springer, 2016.

[38] K. Hao, "What is machine learning?," MIT Technology Review, 17 11 2018. [Online]. Available: https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/.

[39] K. Debnath and M. Mourshed, "Forecasting methods in energy planning models," *Renewable and Sustainable Energy Reviews,* no. 88, p. 306, 2018.

[40] F. Kunz, M. Kendziorski, W.-P. Schill, J. Weibezahn, J. Zepter, C. von Hirschhausen, P. Hauser, M. Zech, D. Möst, S. Heidari, B. Felten and C. Weber, "Electricity, heat, and gas sector data for modeling the German system," DIW Berlin, Berlin, 2017.

[41] N. Neuland, "Spatially and temporally resolved potential analysis and allocation optimization of power-to-gas plants in Gemany," Aachen, 2021.

[42] W. Heitkoetter, W. Medjroubi, T. Vogt and C. Agert, "Regionalised heat demand and power-to-heat capacities in Germany - An open dataset for assessing renewable energy integration," Elsevier, Oldenburg, 2019.

[43] W. Heitkoetter, B. Schyska, D. Schmidt, W. Medjroubi, T. Vogt and C. Agert, "Assessment of the regionalised demand response potential in Germany using an open source tool and dataset," DLR Institute of Networked Energy Systems, Oldenburg, 2020.

[44] D. Schmidt, "NUTS-3 regionalization of industrial load shifting potential in Germany using a time-resolved model," Carl von Ossietzky Universit• at Oldenburg, Oldenburg, 2019.

[45] Eurostat, "Glossary:Household - social statistics," Eurostat, 20 03 2017. [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Household_-_social_statistics.

[46] Eurostat, "Energy consumption in households," Eurostat, 02 06 2020. [Online]. Available: ec.europa.eu/eurostat/statistics-explained/index.php/Energy_consumption_in_households.

[47] Eurostat, "Complete energy balances," Eurostat, 15 04 2021. [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nrg_bal_c&lang=en.

[48] European Environment Agency, "Water use in Europe — Quantity and quality face big challenges," 23 November 2020. [Online]. Available: https://www.eea.europa.eu/signals/signals-2018-content-list/articles/water-use-in-europe-2014. [Accessed 26 3 2021].

[49] Statista, "Where Europeans Consume The Most Tap Water," 9 October 2019. [Online]. Available: https://www.statista.com/chart/19591/average-consumption-of-tap-water-per-person-in-the-eu/. [Accessed 26 3 2021].

[50] U. E. I. Administration, "Short-Term Energy Outlook Supplement: Weather Sensitivity in Natural Gas," U.S. Energy Information Administration, Washington, D.C., 2014.

[51] C. Marinosci, G. L. Morini, G. Semprini and M. Garai, "Preliminary energy audit of the historical building of the School of Engineering and Architecture of Bologna," in *69th Conference of the Italian Thermal Engineering Association, ATI 2014*, Bologna, 2014.

[52] A. Honoré, «Natural gas demand in Europe: The impacts of COVID-19 and other influences in 2020,» The Oxford Institute for Energy Studies, Oxford, 2020.

[53] European Environment Agency, "Heating and cooling degree days," 23 02 2021. [Online]. Available: https://www.eea.europa.eu/data-and-maps/indicators/heating-degree-days-2.

[54] "ERA5 hourly data on single levels from 1979 to present," Copernicus Climate Change Service, 14 06 2018. [Online]. Available: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview.

[55] D. S. Pisupat, "Heating Degree Days Examples," Pennsylvania State University, [Online]. Available: https://www.e-education.psu.edu/egee102/node/2076.

[56] M. P. Bouchelle, D. S. Parker and M. T. Anello, "Factors influencing water heating energy use and peak demand in a large scale residential monitoring study," in *Proceedings of the Twelfth*

Symposium on Improving Building Systems in Hot and Humid Climates, San Antonio, TX, 2000.

[57] N. L. Lam, B. Upadhyay, S. Maharjan, K. Jagoe, C. L. Weyant, R. Thompson, S. Uprety, M. A. Johnson and T. C. Bond, "Seasonal fuel consumption, stoves, and end-uses in rural households of the far-western development region of Nepal," Environmental Research Letters, Volume 12, Number 12, 2017.

[58] T. Pettinger, "Sectors of the economy," EconomicsHelp.org, 19 12 2019. [Online]. Available: https://www.economicshelp.org/blog/12436/concepts/sectors-economy/.

[59] Eurostat, "NACE Rev. 2 - Statistical classification of economic activities in the European Community," Eurostat, [Online]. Available: https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF.

[60] U.S. Energy Information Administration, "International Energy Outlook 2016," 2016.

[61] J. E. Sandoval, W. Medjroubi, J. Diettrich and A. Pluta, "SciGRID_gas Cons," SciGRID_gas, 17 06 2021. [Online]. Available: https://zenodo.org/record/4922884.

[62] Countrystudies.us, "Economy of Spain - Regional concentration," Countrystudies.us, [Online]. Available: http://countrystudies.us/spain/63.htm. [Accessed 22 06 2021].

[63] Forschungsstelle für Energiewirtschaft e.V., "The FfE Open Data Portal - How to," FfE, [Online]. Available: http://opendata.ffe.de/how-to/. [Accessed 15 05 2021].

[64] AG Energiebilanzen e.V., "Energy Balance 2015 (update August 2017)," 18 10 2017. [Online]. Available: https://ag-energiebilanzen.de/index.php?article_id=29&fileName=bilanz15d_engl.xlsx.

[65] Prologis Germany, "Rhine-Ruhr, Industrial Capital of Western Germany," Prologis Germany, 14 06 2021. [Online]. Available: https://www.prologisgermany.de/en/industrial-logistics-warehouse-space/europe/germany/rhine-ruhr-industrial-capital-western-germany.

[66] Encyclopaedia Britannica, "Düsseldorf," Encyclopaedia Britannica, 03 05 2006. [Online]. Available: https://www.britannica.com/place/Dusseldorf.

[67] Country-data.com, "Portugal Industrial Regions," Country-data.com, [Online]. Available: http://www.country-data.com/cgi-bin/query/r-10929.html. [Accessed 15 06 2021].