


ClairCity: Citizen-led air pollution reduction in cities

# D3.4 Good practice guidelines for collecting practice-activity data

April 2019

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689289

# Document Details

| | |
|---|---|
| ***Authors*** | Wiet Koren, Olav ten Bosch (CBS Statistics Netherlands), Angreine Kewo (DTU) |
| ***Contact*** | lgh.koren@cbs.nl |
| ***Creation Date*** | 28/06/2018 (Version 1.0) |
| ***Date of Last Revision*** | 17/4/2019 (Version 2.0) |
| ***Description*** | This report is one of six reports in WP3 on behaviour, activity and practices. Its aim is to be a source of information on the ClairCity approach and to provide practical instructions to data managers of cities that may wish to replicate the ClairCity method and examples of where the practice-activity data can be found and how to obtain and process them. |

# Version History

| Version | Updated By | Date | Changes / Comments |
|---|---|---|---|
| V1.0 | Wiet Koren, Olav ten Bosch | June 2018 | First draft |
| V1.1 | Wiet Koren, Angreine Kewo | August 2018 | Addition of modeling approaches and data selection |
| V1.2 | Wiet Koren, Olav ten Bosch | December 2018 | Major rewrite. Addition of description of data sources and good practices (chapter 3) and conclusions (chapter 4) |
| V1.3 | Enda Hayes | December 2018 | Minor amendments and comments |
| V1.4 | Wiet Koren, Olav ten Bosch | December 2018 | Editing for draft for input from WP3 partners |
| V1.5 | Wiet Koren, Olav ten Bosch | February 2019 | Incorporation of input from WP3 partners |

| | | | |
|---|---|---|---|
| V1.6 | Wiet Koren, Angreine Kewo, Olav ten Bosch | March 2019 | Completing and editing all sections |
| V1.7 | Enda Hayes | March 2019 | Minor amendments, points of clarification and comments |
| V1.8 | Wiet Koren and Olav ten Bosch | March 2019 | Version for Executive Management Group (EMG) |
| V2.0 | Wiet Koren and Olav ten Bosch | April 2019 | Comments from annual ClairCity meeting in Aveiro |

# Contributions and Acknowledgements

The authors would like to thank the following people for their important contributions used in the preparation of this final document.

| | |
|---|---|
| **Quality Assurance** | Enda Hayes (UWE) |
| **Native Language Check** | Enda Hayes (UWE) |
| **Project internal comments** | Angreine Kewo (DTU)<br>Carlo Trozzi (TECHNE)<br>Kris Vanherle (TML)<br>Vera Rodrigues (UAVR)<br>Evert Bouman (NILU)<br>Iason Diafas (PBL)<br>Trond Husby (PBL)<br>Ana Patrícia Fernandes (UAVR)<br>Svein Knudsen (NILU)<br>Enda Hayes (UWE)<br>Chris Dunford (EMG, We The Curious)<br>Gert Spaargaren (EMG, Wageningen University)<br>Péter Pápics (TML)<br>Imke van Moorselaar (GGD Amsterdam)<br>Joana Soares (NILU)<br>Irati Artola (Trinomics)<br>Corra Boushel (UWE) |

# Table of Content


Document Details ..... 2

Version History ..... 2

Contributions and Acknowledgements ..... 4

Executive Summary ..... 7

1 Introduction ..... 8

1.1 Objective of this report ..... 8

1.2 What are practice-activity data? ..... 9

1.3 One size fits all? ..... 11

2 Can we identify practice-activity data that are suitable for our purposes? ..... 12

2.1 Modeling ..... 12

2.2 Data selection and availability ..... 23

2.3 Alternative air pollution data sources ..... 28

2.4 Visualizing practice-activity data ..... 33

3 What practice-activity data are collected and what good practices can we identify? ..... 35

3.1 Integrated urban module ..... 35

3.2 Transport module ..... 37

3.3 Industrial, residential, commercial and institutional module ..... 40

3.4 Energy / power generation module ..... 43

3.5 Agriculture module ..... 43

3.6 Natural module ..... 44

3.7 Emission database module ..... 44

3.8 Air quality module ..... 46

3.9 Health module ..... 48
3.10 Carbon footprint module ..... 50
3.11 Identifying good practices ..... 54
4 Conclusion ..... 57
5 Appendix ..... 58

# Executive Summary

Practice-activity data, the data that relate to the *behaviour* and *practices* of people or to any activity that - directly or indirectly - leads to air pollution and/or carbon emissions, is an essential ingredient of the ClairCity approach. This deliverable is about good practices for identifying, collecting, analyzing, processing and disseminating such data, in a broad sense. It is intended to be a valuable source of information for those that want to apply the ClairCity approach – or a similar one – in their own urban and/or regional context.

This document starts by defining the essence of practice-activity data in the scope of air quality improvements. It touches upon some of the more general practices and choices people make in everyday life, having an impact on quality of life. One of the conclusions is that in order to improve air quality, directly or indirectly caused by peoples behaviours, we must model the practices and activities and their impacts on emissions in such detail that we can use it to develop realistic scenarios for a city future. To do this it becomes clear that there is no “one size fits all” approach. As local circumstances and behavioural patterns might differ considerably, the ideal mix of practice-activity data to be used must be carefully considered in every individual case.

To identify suitable data this documents presents various data and activity models, the data modules used to organise the modeling process executed by different ClairCity partners and a number of flowcharts and approaches that can be used to model energy use and the resulting emissions. It elaborates on the common tasks for pre-processing data and lists some alternative data sources and visualization methods. The document enumerates the data sets that were actually used for the modeling work in the six ClairCity pilot cities and regions organised into the various modules: (1) the integrated urban module, (2) the transport module, (3) the industrial, residential, commercial and institutional module, (4) the energy / power generation module, (4) the natural module, (5) the emission database module and (6) the air quality, health and carbon footprint modules.

The document concludes with a lists of good practices identified by the sixteen ClairCity partners working together in this challenging research project. We hope and believe the result can be of use for many future innovation projects that wish to mimic or build upon the ClairCity approach.

# 1 Introduction

## 1.1 Objective of this report

Data is one of the cornerstones of the ClairCity project. Without access to basic demographic urban statistics, data on the presence and use of city transportation, estimates on energy production and consumption of households and organisations, specific emission data for main sources and data on other topics such as industrial, agriculture and natural resources, it is not possible to apportion air pollution and carbon emissions by citizens behaviour, practices and activities and develop a clear view on future scenarios for a city with the aim to improve air quality, reduce carbon emissions and protect public health.

More generally stated, data on the *behaviours* and *activities* of people in a city are needed in order to build a model and derive future scenarios on air quality and carbon footprint impacts. This data is not always readily available, and if available, it is not always of sufficient quality. Ideal data requirements are almost never fully met and therefore in many cases data being used in practice is the result of a compromise and/or using surrogate data or modeled output. One of the results of the ClairCity project is that the ClairCity partners searched for the right data in six very practical case studies driven by the data reality of six diverse European cities / regions – Amsterdam, The Netherlands; Bristol, United Kingdom; Ljubljana, Slovenia; Sosnowiec, Poland; Aveiro, Portugal; and Liguria, Italy. This way ClairCity gained a lot of experience in identifying, analysing and processing the available sources allowing it the ClairCity approach to be both scientifically robust yet flexible and adaptable depending on the city / regions needs.

One of the findings is that each city / region differs not only in a physical and demographic sense, but also from a data perspective. The same holds for countries. From a data perspective each country is organised differently in terms of centralised or specialised organisations measuring and maintaining environmental aspects. The only common denominator, the European Union (EU) and its environmental organisations, do provide lots of data. However these do not always have the granularity needed for detailed city / region modeling as needed in ClairCity. Subsequently, choices had to be made between quantity, granularity, adaptability and quality, and these experiences are worth sharing for future use of the ClairCity approach in other cities or future projects which may adopt similar analysis.

This report is intended to be a source of information for those that want to make the same data journey as we did in ClairCity. It looks back on the pros and cons of the relatively new concept of practice-activity data as explained in the first deliverables from WP3 (see Section 1.2) and used in this collection phase of the ClairCity project. Taking a broad view on the data actually being used in the project and stored in the ClairCity Data Portal, it derives a set of *good practices* (and also mentions less succesful practices) for collecting and processing such data. From there the report aims at giving practical instructions and examples of where practice-activity data can be found and how to obtain and process them.

The audience of this deliverable can be very diverse, ranging from international projects trying to find data for their models to data managers of cities /regions that want to apply the ClairCity approach or a related project approach in their own city.

## 1.2 What are practice-activity data?

The concept of practice-activity data in general is about linking human behaviour and resulting activities. The interactions between activities, in our case activities that result in emission of air pollutants[1], and practices of governments, groups and individuals are more elaborately described in ClairCity Deliverable "D3.1 Review of social science in air quality and carbon management"[2]. Thus, practice-activity data in the context of ClairCity should result, on its own or combined with other data, in clues on how to reduce air pollution and carbon emissions. We take the notion of practice-activity data in a broad sense, including the whole spectrum from choices and actions of citizens to the emission of, and exposure to, polluting substances.

From this point of view, practice-activity data are data that:

- relate to the behaviour and practices of people, or
- relate to the emission of air pollution, for example any activity that leads to the emission of air pollution, or
  relate to any link between these two types of data

Polluting emissions arise from various industrial and household activities and vary in time and space. In a simple scheme (Figure 1), the practices of citizens and their choices and behaviour in those practices are linked to activities resulting in less or more air pollution. In ClairCity, emphasis has been placed upon household activities and transport. Leisure activities and activities at work may be important in specific areas but are sometimes difficult to measure and only taken into account in ClairCity when available.

With the collected data it should be possible to distinguish differences in practices and choices made in daily life. Practices are often habitual, while choices may have a large effect upon the production of air pollution and carbon footprint. Going to work by car may be convenient and quick, alternatively you could take a bike, improve your health and reduce carbon emissions. At longer distances, an e-bike might be a choice, or public transport, giving the opportunity to read.

[1] For the purposes of this report, when we refer to 'air pollution' or 'air pollutants' or 'polluting emissions' etc we are including both traditional pollutants such as Nitrogen Dioxide, Particulate Matter etc and climate related pollutants such as Carbon Dioxide.
[2] This report and other ClairCity reports are available at: http://www.claircity.eu/reports/

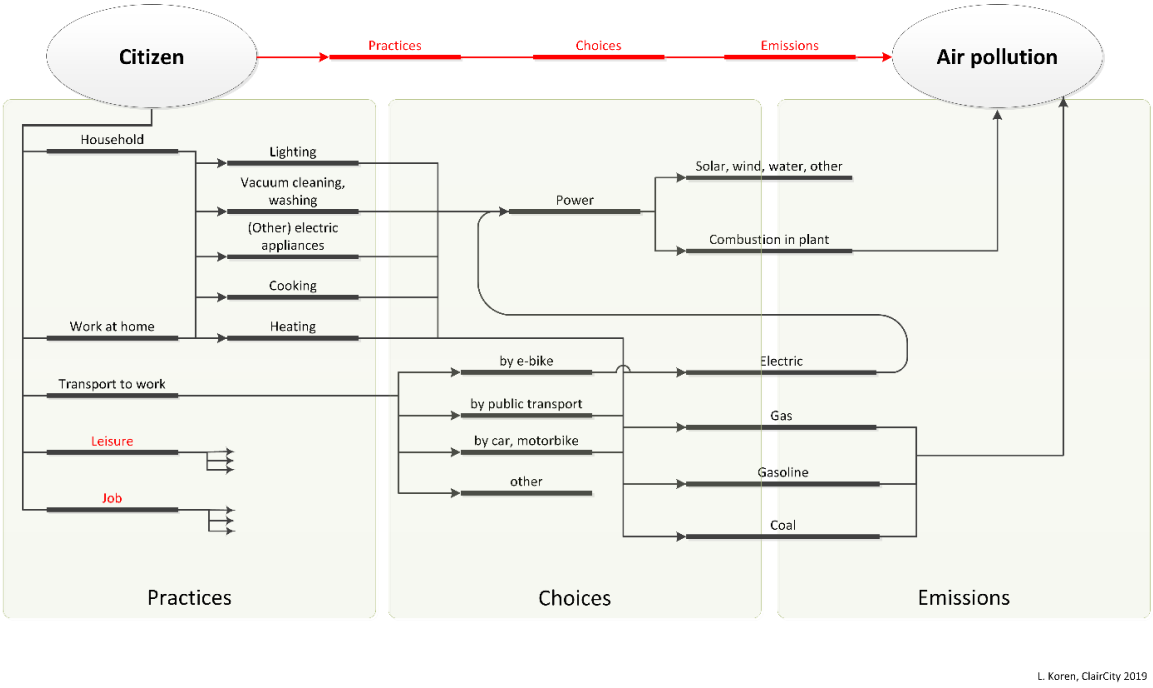



**Figure 1. Practice-activity data of citizens in relation to air pollution**

Bigger behaviour choices may include opting for a new cooking or heating system. Going from coal or wood-fired stoves or gas heating systems to electric heating and cooking systems powered by heat pumps and solar panels is a choice with large initial financial implications but could be finanically beneficial in the long term.

Other choices to increase or reduce the production of air pollution and carbon footprint are not financially restricted: an additional open fireplace in the house beside the closed heating system, barbecueing and camp fires, patio heaters. This type of decision is usually not driven by economic motives. In fact, in many cases environmental aspects do not play a role and the choice for these more or less luxury extras is driven – knowingly or unknowingly – by the attitude of influencers around such as family, friends or colleagues. This makes the practice-activity pattern even more challenging.

Social and economical factors must also be taken into account: city dwellers with low income often do not own any motorized means of transport, and if they have, they do not have the financial means to replace their car, scooter or other transportation to a modern, less polluting one.

In order to improve air quality, directly or indirectly caused by peoples behaviours, we must model the practices and activities of people across and within the city / region and their impacts on emissions. Only then we can design a city-specific long term approach to air quality improvements.

## 1.3 One size fits all?

All cities and regions participating in ClairCity have different focus points depending on local geography, demographics, practices and culture. Local pollution sources are also different and local organisations have different means of measuring and collecting data. This document aims to support and sustain the implementation of the ClairCity practice-activity approach in all cities. Choices of citizens that contribute to better air quality may be positively or negatively influenced by city policies.

In the ClairCity Policy Baseline Reports, examples of city / region diversity are described.

- In Bristol, air quality actions are predominantly road transport related and need to ensure the reduction of social inequality. The role of the citizen is project-based, funding of continuation of projects appeared troublesome (D6.2 Policy Baseline Report Bristol, p. 90-95).
- In Amsterdam, city policies include restricting heavy traffic and stimulating renewable energy (D6.1 Policy Baseline Report, Amsterdam, pp. 14, 61-66).
- Ljubljana, due to its situation in a basin and corresponding negative influence on air pollution, has an active policy towards climate improvement. Air pollution from transport is a major problem. Citizen's engagement is promoted (D6.2 Ljubljana – Policy Baseline Report).
- Economic restructuring in Sosnowiec from coal and heavy industry influences air pollution. Public awareness is limited but necessary as winter smog is frequent, originating from domestic coal and wood boilers and stoves (D6.2 Policy Baseline Report Sosnowiec, p. 67, 72).
- In Liguria, city planning is restricted by the hilly and compact territory and air pollution and climate is as yet not an issue for the general public (D6.2 Policy Baseline Report Liguria, p. 80).
- Aveiro has relatively low levels of air pollution, and climate actions above those prescribed by EU Directives have not been issued, due to lack of finance and political will (D6.2 Policy Baseline Report Aveiro, p. 80).

So although we look for general good practices for collection of practice-activity data to be applied in other cities or projects, this document can not and will not use a "one size fits all" approach. Where possible we will opt for a general approach, where necessary it will be more specific, but the adoption of these guidelines will require some localisation.

# 2 Can we identify practice-activity data that are suitable for our purposes?

In the ClairCity project, various data are selected, collected and modeled. In this chapter, we try to provide guidance to where data may be found, which data are useful, and what parameters must be included to model that data succesfully. Appropriate collecting and modeling techniques and reliability of sources will be discussed.

We start with the modeling approaches and show how these are elaborated into the selection and availability of data.

## 2.1 Modeling

At first, a simple model was put forward to model the emission from citizens' behaviour in time and space from emission data, air pollution producing activities, and practice-activity data (Figure 2).

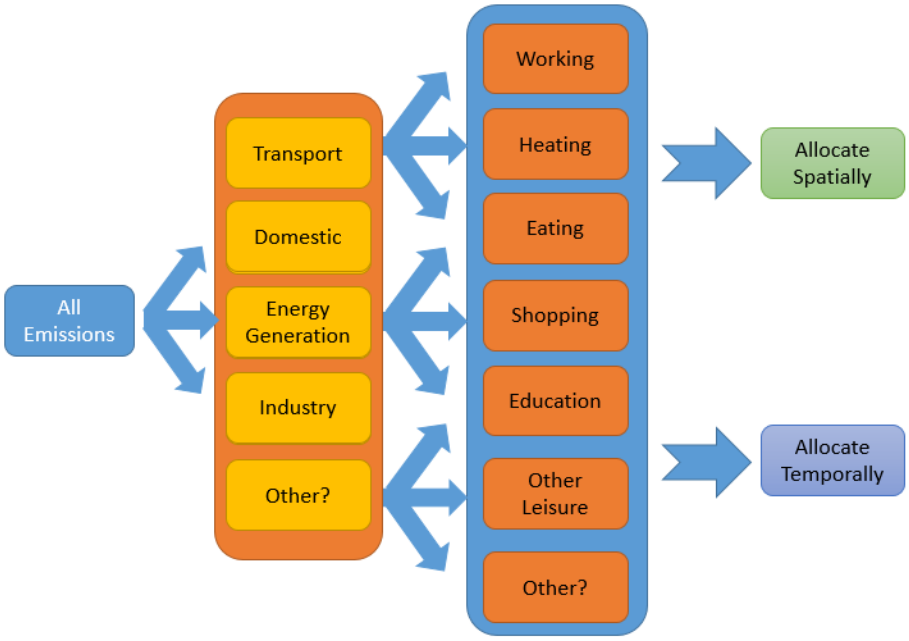



**Figure 2. A simple representation of ClairCity Behaviour/Activity/Practice modeling**

This model was transformed into a datamodel in which tasks were distributed to the partners (Figure 3).

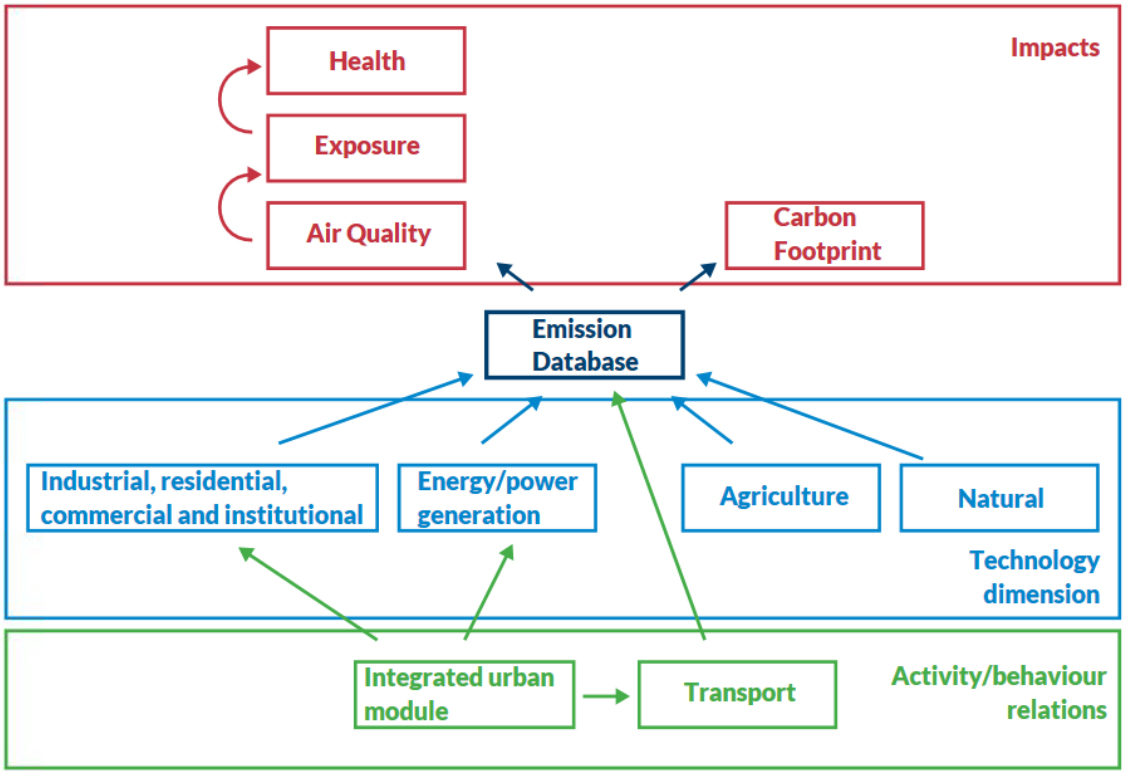



**Figure 3. ClairCity data model and modules**

The blocks were further elaborated into tasks with a commonly agreed set of characteristics. For example, in the flow chart for area sources (Figure 4), the hourly data for mapping air quality are derived from annual emission data for sub-areas, frequent energy consumption data, hourly characteristics of general housing patterns and geographical gridding over the area.

Communication is a key factor in such a complex multi-party scientific modeling project. It is crucial to align the inputs and outputs of each module among project partners. For this reason a working document “ClairCity WP5 Specifications of Modeling Toolset” was designed and maintained during the design phase of the project. In this document, the characteristics of each module were described by those responsible. The description includes the connecting flows to and from each module. The process flows as visualized in the data portal home page, Figure 3, is the result. Thus, everyone in the project has knowledge of what steps have to be taken, what to produce and what to expect from the other modules. Data formats were chosen based on the modeling work to be done in the module of interest.

The modeling approach basically depends on the purpose of the modeling, the availability of input data and the outcome(s). It also takes into account, the data pre-processing techniques, applicable modeling tool (if exists), mathematical model, certain algorithm and

robustness test or validation model to assess the accuracy of the model. All of these considerations in modeling must be in line with the project management. Therefore, an efficient modeling approach is required in the Claircity project.

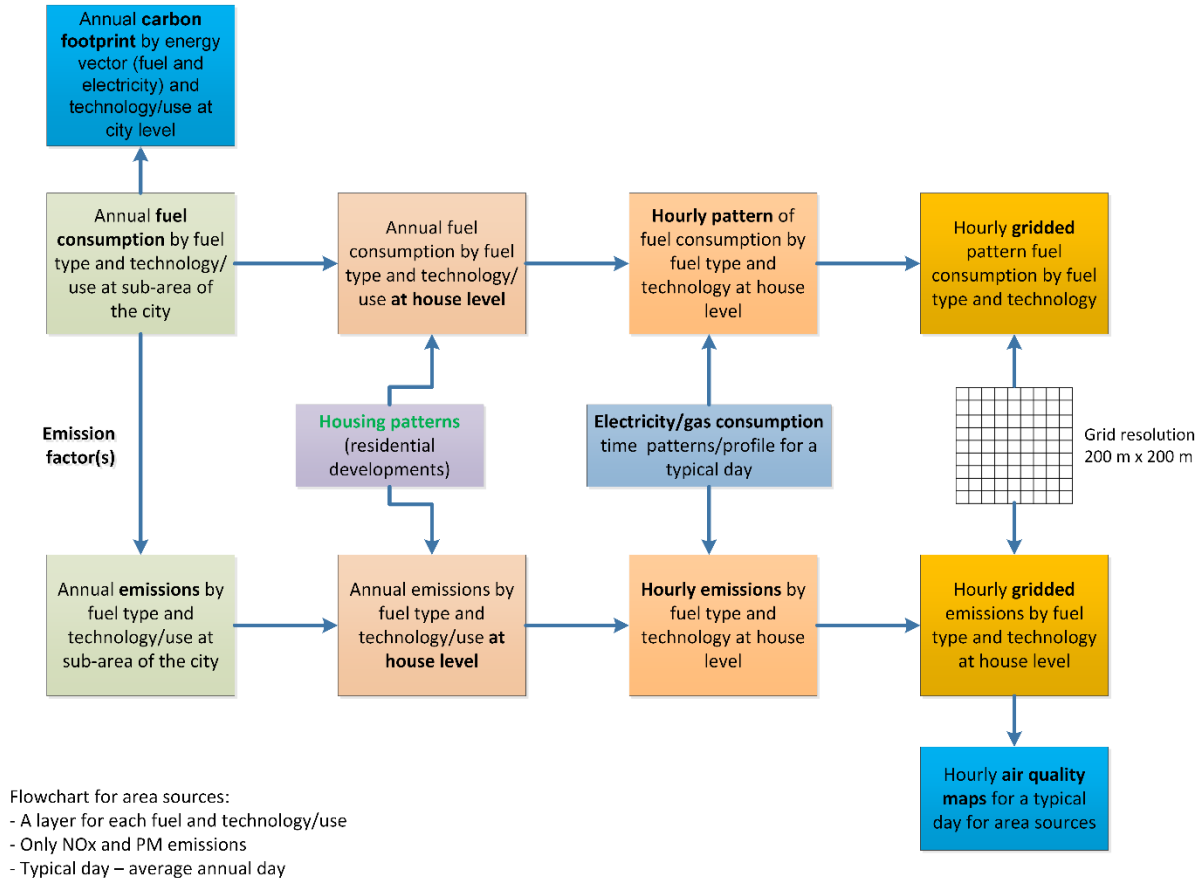



**Figure 4. Flowchart for processing ClairCity data**

### *2.1.1 Case Study 1: Modeling process of the temporal emission load for residential and commerical energy use*

We describe here, as an example, the modeling process of the temporal emission load for residential and commercial energy use. In general, the regional emissions temporal load profile method is the same for all case cities. Figure 5 shows the generic block diagram, where the data inputs are:

a. Monthly gas pattern
b. Hourly local temperature
c. Hourly national electricity load
d. Share of fuel resources (%), especially wood heater emissions.

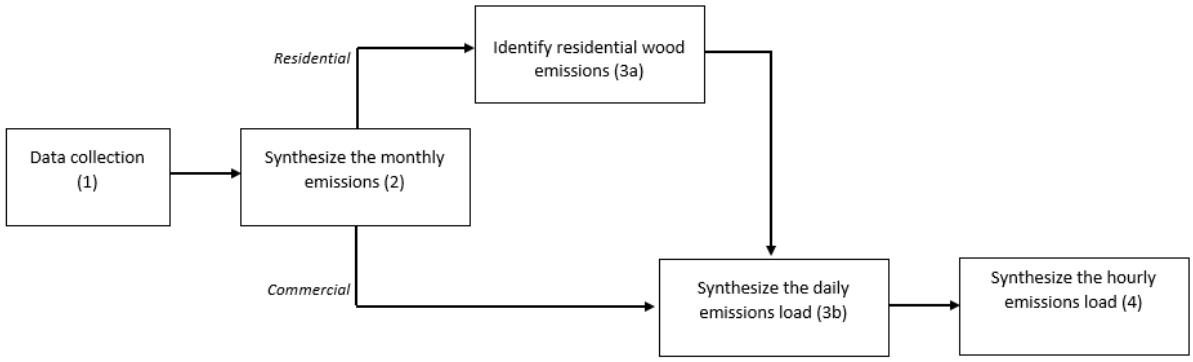



**Figure 5. Diagram of temporal resolution load profile framework for residential and commercial energy**

Explanation of the diagram (Figure 5):

1. Data collection of all required inputs
   *Objective: To identify all related real national or regional patterns.*

2. Synthesize the monthly emissions load according to monthly gas pattern.
   *Objective: To produce the monthly synthetic emissions load that has similar monthly variation pattern with the real monthly gas pattern.*

   The synthetic daily emissions are being distributed from the synthetic total monthly emissions, where the wood emissions are only applied during the winter period.

3a Residential: Identify the residential wood emissions
   *Objective: To identify the residential wood emissions share.*

3b Commercial and Residential: Synthesize the daily emissions load according to the daily average temperature pattern.
   *Objective: To produce the daily synthetic emissions load.*

4. Synthesize the hourly emissions load according to the hourly electricity load.
   The synthetic hourly emissions are being distributed from the synthetic total daily emissions.
   *Objective: To produce the hourly synthetic emissions load.*

In modeling the regional emission temporal load profile, the primary input is listed in Table 1. In case it is difficult to obtain primary data, most of the secondary data are easier to be retrieved.

Due to the different data availability per case study, we have applied the following approaches to solve the matter:

Approach 1: All Primary data are available.

Approach 2: No Hourly Temperature: This approach is employed when there is no hourly local temperature dataset. This is usually the fundamental climate set to model the energy load profile. If the local dataset do not exist, then the hourly temperature of the nearest neighbour city or region or country may be used. Furthermore, if the local or nearby hourly temperature is unavailable and if the residential smart meter data exist, the hourly pattern

may employ the regional smart meter pattern. The rest of the modeling work is similar with Approach 1.

**Table 1 - Primary and secondary data collection of regional emission temporal load profile**

| Primary (P) | Secondary (S) |
|---|---|
| Local temperature dataset<br>Resolution: Hourly<br>*National meteorology office or commercial weather companies* | Nearby or neighbour city/region temperature dataset<br>*National meteorology office or commercial weather companies* |
| National gas pattern<br>Resolution: Monthly<br>*Eurostat database https://ec.europa.eu/eurostat/web/energy/database* | National gas pattern<br>Resolution: Monthly<br>*JODI dataset or related studies*<br>*https://www.jodidata.org* |
| National electricity load pattern<br>Resolution: Hourly<br>*European open data platform ENTSO-E; (https://transparency.entsoe.eu)* | National electricity load pattern<br>Resolution: Hourly<br>*Smart-meter dataset from utilities company or related scientific works.* |
| Emissions dataset<br>Level: City or region<br>*ClairCity partner* | Emissions dataset<br>Level: National<br>*National accounts* |

Approach 3: No monthly gas pattern. In this case we can benchmark to the related studies that provide the monthly gas pattern. If the dataset is unavailable, then we synthesize the monthly gas pattern based on the monthly chart or graph or any other visualisation provide in the related studies. The rest of the modeling work is similar with Approach 1.

Approach 4: No hourly electricity load pattern. If there is no hourly electricity load pattern available publicy, then we can use the smart-meter dataset from the utilities. In this case, we have to make sure that the dataset is reflected the commercial and residential sector, because most of the smart-meter datasets are captured specifically only the households data. Otherwise, we can benchmark to the related studies that provide related dataset.

Approach 5: No **emission** data: In practice, this approach is very unlikely to happen because we typically obtain the emissions area dataset from ClairCity partner. In case there is a delay,

as a backup, we search the emissions data from the national accounts or related studies, although it is not as detailed as the dataset provide by our ClairCity partner.

Approach 6: No Primary Data: When most of the primary data are not available (temperature, national load of electricity and gas, and emissions datasets), the regional emissions temporal load profile may use the input from secondary resources.

In practice, the data preparation or data pre-processing is the success key in modeling. In some cases, data pre-processing may take more than half of the total time spent in the modeling or data mining phase[3]. Before conducting the emissions load profile modeling phase, data pre-processing is necessary. First, the necessary data comes in various data sources and formats[4]. Many common and typical problems in data modeling are caused by missing values, data corruption and outliers[6]. The raw data usually require to be pre-processed due to noise, ambiguity, redundancy, incompleteness and inconsistency[5,6]. Data pre-processing will convert the data into a clean and tidy dataset for the later mining step or statistical analysis[7,8,9].

### *2.1.2 Case Study 2: data pre-processing techniques*

In statistical analysis, the four common tasks in data pre-processing are cleaning, transformation, integration and reduction[8,10] with the aim to assess and improve the data quality for reliable statistical analysis[11].

Table 2 has summarised the data pre-processing techniques applied in each dataset per case city.

[3] Miksovsky P, Matousek K, Kouba Z. Data pre-processing support for data mining. Syst Man Cybern 2002 IEEE Int Conf 2002;5:4 pp. vol.5. doi:10.1109/ICSMC.2002.1176327.
[4] Kewo A, Manembu P, Liu X, Nielsen PS. Data Pre-processing : A Case on Regional Energy Load Profile. ICEECS 2018, 2018.
[5] Bilalli B, Abelló A, Aluja-Banet T, Wrembel R. Intelligent assistance for data pre-processing. Comput Stand Interfaces 2018;57:101–9. doi:10.1016/j.csi.2017.05.004.
[6] Guo Z, Zhou K, Zhang X, Yang S, Shao Z. Data mining based framework for exploring household electricity consumption patterns: A case study in China context. J Clean Prod 2018;195:773–85. doi:https://doi.org/10.1016/j.jclepro.2018.05.254.
[7] Tsai CF, Chou JS. Data pre-processing by genetic algorithms for bankruptcy prediction. IEEE Int Conf Ind Eng Eng Manag 2011:1780–3. doi:10.1109/IEEM.2011.6118222.
[8] MIT Critical Data. Secondary Analysis of Electronic Health Records. 2016. doi:10.1007/978-3-319-43742-2.
[9] Aubrecht P, Mikšovský P, Král L. SumatraTT: A generic data pre-processing system. Proc - Int Work Database Expert Syst Appl DEXA 2003;2003–Janua:120–4. doi:10.1109/DEXA.2003.1232010.
[10] L'Huillier G, Velásquez JD. Advanced Techniques in Web Intelligence-2. vol. 452. 2013. doi:10.1007/978-3-642-33326-2.
[11] Kewo A, Manembu P, Nielsen PS. Data Pre-processing Techniques in the Regional Emissions Load Profile Case. Int. Conf. Control. Decis. Inf. Technol., 2019.

**Table 2. Summary of data pre-processing techniques applied in each dataset per case city**

| Data collection | Property of | Nature and type of data collected | Pre-processing techniques | **Tool or computer application** |
|---|---|---|---|---|
| *1. Temperature* | | | | |
| Bristol, United Kingdom | Met Office UK<br>https://www.metoffice.gov.uk/ | Hourly resolution | Reduction<br>Transformation | Spreadsheet<br>Programming language: Python |
| Amsterdam, The Netherlands | Meteorology office: KNMI<br>https://data.knmi.nl/datasets | Daily (average) resolution | Transformation | Spreadsheet |
| Ljubljana, Slovenia<br><br>Sosnowiec, Poland | Commercial weather services<br><br>https://www.wundergroud.com/ by IBM | 30-minute resolution | Reduction<br>Integration<br>Transformation<br>Cleaning | Spreadsheet<br>Command prompt<br>Programming language: Python<br>Programming language: Python |
| Aveiro, Portugal | ClairCity partner – shared by UAVR, Portugal | 10-minute resolution | Reduction<br>Filtering<br>Selection<br>Transformation | Spreadsheet<br>Programming language: Python<br>Programming language: Python<br>Spreadsheet and Programming language: Python |
| Liguria, Italy | Regional Meteorology Meteorology Regione Ligiuria | Hourly resolution | Reduction<br>Transformation | Spreadsheet |

| Data collection | Property of | Nature and type of data collected | Pre-processing techniques | **Tool or computer application** |
|---|---|---|---|---|
| | | | | Spreadsheet and Programming language: Python |
| *2. Gas national pattern* | | | | |
| Bristol, United Kingdom<br>Amsterdam, The Netherlands<br>Ljubljana, Slovenia<br>Sosnowiec, Poland<br>Aveiro, Portugal<br>Liguria, Italy | Eurostat database<br>Eurostat, Energy | Time series dataset<br>Monthly resolution | Extraction<br>Filtering<br>Selection<br>Transformation | Spreadsheet<br>Spreadsheet<br>Spreadsheet<br>Spreadsheet |
| *3. Electricity national load* | | | | |
| Bristol, United Kingdom<br>Amsterdam, The Netherlands<br>Ljubljana, Slovenia<br>Sosnowiec, Poland<br>Aveiro, Portugal<br>Liguria, Italy | European open data platform ENTSO-E; (https://transparency.entsoe.eu) | Time series dataset<br>Hourly resolution | Reduction<br>Transformation | Spreadsheet<br>Spreadsheet |
| *4. Emissions area* | | | | |
| Bristol, United Kingdom<br>Amsterdam, The Netherlands<br>Ljubljana, Slovenia<br>Sosnowiec, Poland<br>Aveiro, Portugal<br>Liguria, Italy | Owner: ClairCity partner – Techne Consulting | Area dataset | Reduction<br>Filtering<br>Selection<br>Transformation | Programming language: Python<br>Programming language: Python<br>Programming language: Python<br>Spreadsheet and Programming language: Python |

### 2.1.3 *Case Study 3: utilising the Load Profile Generator to determine household energy load*

Specifically for the residential load, we have developed a framework that focuses on the local residential characteristics (Figure 6). Since, the household load profile is modeled using

Noah Plufgradt's model[12,13] through the Load Profile Generator (LPG) tool, before selecting or creating own household profiles, the regional weighting composition of age groups and employment rate is applied. It is the result of households aggregation at the city or regional level. Futhermore, beside the Plufgradt's model, it is also required to enclose the local temperature and holiday parameters on the inputs.

[12] Pflugradt N, Muntwyler U. Synthesizing residential load profiles using behavior simulation. Energy Procedia 2017;122:655–60. doi:10.1016/j.egypro.2017.07.365.
[13] Pflugradt ND. Modellierung von Wasser- und Energieverbräuchen in Haushalten 2016.

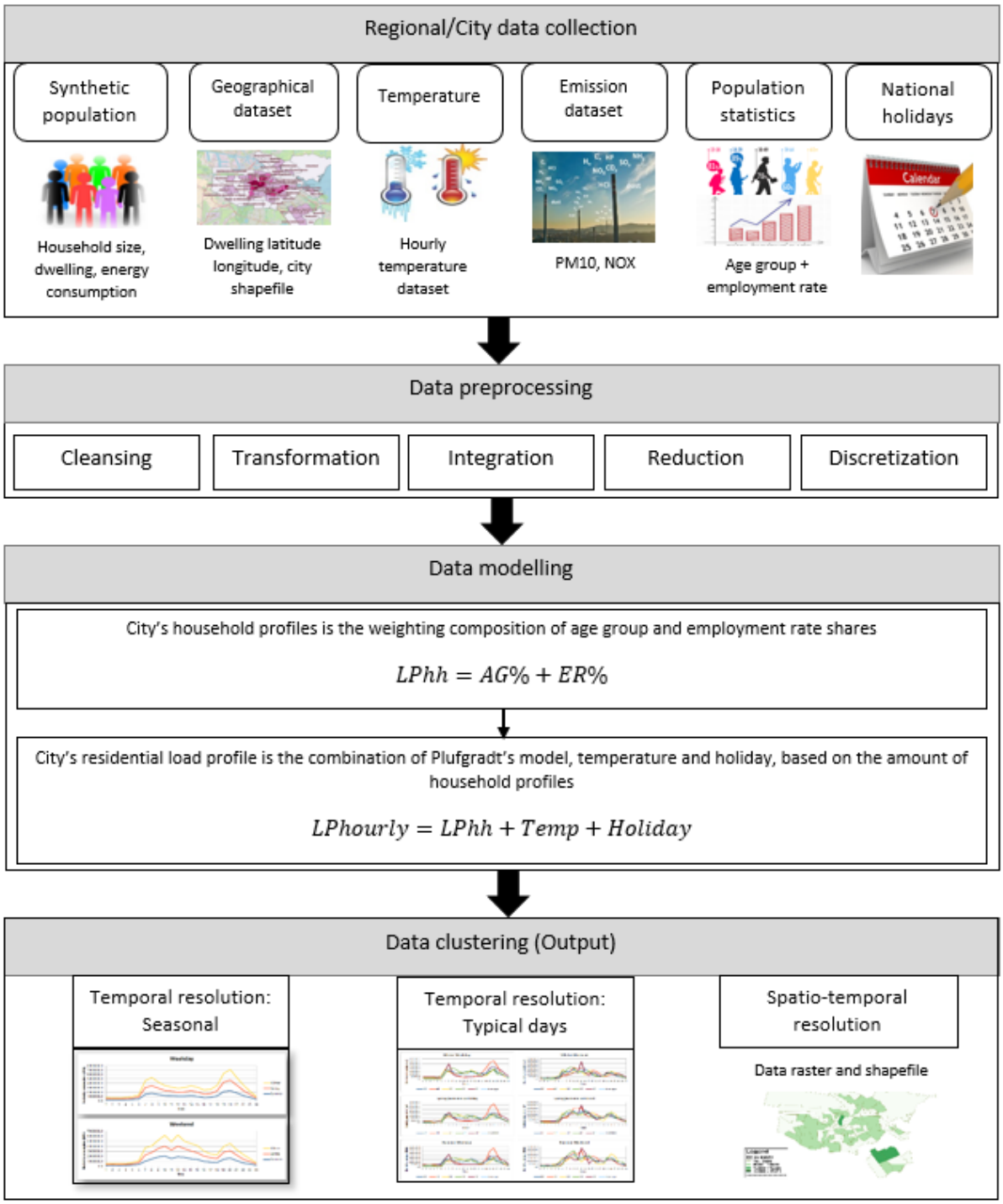



**Figure 6. The residential spatio-temporal resolution load profile method[14]**

[14] Kewo A, et al. Data Pre-processing : A Case on Regional Energy Load Profile. ICEECS. IEEE 2018

In this case, there are three kinds of required input to be enclosed in the LPG tool: temperature, population profile, national public holiday. The rest are following the generic block diagram as listed in table 1: National load of electricity and gas, and emissions datasets.

The primary input  of modeling the residential load profile are in Table 3.

**Table 3 - Primary and secondary data collection of residential load profile modeling used in LPG tool**

| Primary (P) | Secondary (S) |
|---|---|
| Local temperature dataset<br>Resolution: Hourly<br>*National meteorology office or commercial weather companies* | Nearby or neighbour city/region temperature dataset<br>*National meteorology office or commercial weather companies* |
| Population statistics<br>Amount of population, age group and employment rate<br>Level: City or region<br>*National statistics office* | Population statistics<br>Amount of population, age group and employment rate<br>Level: City or region<br>*Synthesized population dataset from ClairCity partner or related studies* |
| National public holiday<br>*National government office* | National  public holiday<br>*Online sources* |

In general the following approaches are possible:

Approach 1: All Primary data are available.

Approach 2: No Hourly Temperature: We employ the same approach as described in the integrated regional emissions load profile of residential and commercial sectors (Table 1), where if the local dataset is unavailable, then the hourly temperature of the nearest neighbour city or region or country may be used. Furthermore, if the local or nearby hourly temperature is unavailable and if the residential smart meter data exist, the hourly pattern may employ the smart meter pattern. The rest of the modeling work is similar with Approach 1.

Approach 3: No Primary Data: When most of the primary data are not available (temperature, population, and national public holiday), the residential load profile modeling may use the input from secondary resources.

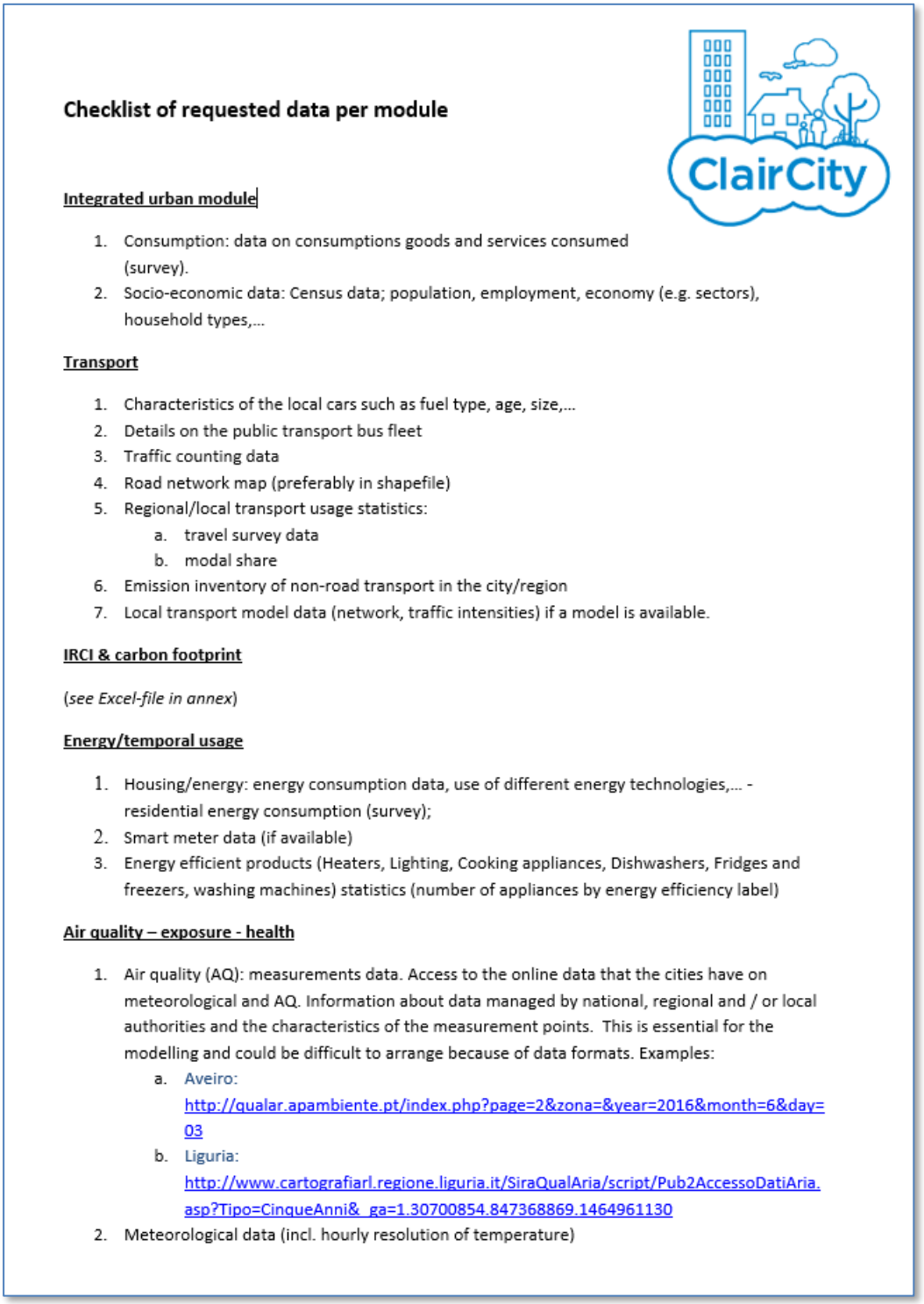
We do not specifically address the approach if there is no population statistics and/or national public holiday information, because these kinds of information are publicy available and easily to be obtained.

## 2.2 Data selection and availability

In the case of ClairCity, data collection plays an essential role to produce the outcome(s) of each module. While elaborating this model, it became obvious that it would not be possible to obtain the same fixed set of data in terms of data structure and format from each city partner. For each module block, the most detailed and highest quality data would be preferred, but less detailed or alternative data could be used too. Dependable data sources and a correct data collection method will allow high quality data, improve reliability and efficiency, resulting in high quality output. For instance, population data may be obtained from local central bureau statistics or temperature data from local meteorology office, to ensure the quality of the data. Since the ClairCity project has a comprehensive governance structure that also involves the city partners as local or domain experts, it makes the data collection process easier. Therefore, it is important to scope, categorize and identify what kind of quantitative data we need.

For this purpose, it was decided to make a checklist of data necessary to fill in the model. A first generic checklist was sent to the different city groups to evaluate and fill in, and amend to the local situation (see Figure 7). In the course of the project, taking one or two cities at a time, the checklist was revised and adapted to the newly acquired knowledge.

As mentioned, it is important to categorize and identify which data are mandatory or optional in each task of the modules. Furthermore, it is required to define the approach(es) in order to produce relevant outcomes based on the data in collection. Basically, there will be one generic or standardized approach as the main solution, and there might be also some alternative approaches, depending on data collection, to be applied in each task of the flowchart (Figure 4).



**Checklist of requested data per module**

**Integrated urban module**

1. Consumption: data on consumptions goods and services consumed (survey).
2. Socio-economic data: Census data; population, employment, economy (e.g. sectors), household types,…

**Transport**

1. Characteristics of the local cars such as fuel type, age, size,…
2. Details on the public transport bus fleet
3. Traffic counting data
4. Road network map (preferably in shapefile)
5. Regional/local transport usage statistics:
    a. travel survey data
    b. modal share
6. Emission inventory of non-road transport in the city/region
7. Local transport model data (network, traffic intensities) if a model is available.

**IRCI & carbon footprint**

(*see Excel-file in annex*)

**Energy/temporal usage**

1. Housing/energy: energy consumption data, use of different energy technologies,… - residential energy consumption (survey);
2. Smart meter data (if available)
3. Energy efficient products (Heaters, Lighting, Cooking appliances, Dishwashers, Fridges and freezers, washing machines) statistics (number of appliances by energy efficiency label)

**Air quality – exposure - health**

1. Air quality (AQ): measurements data. Access to the online data that the cities have on meteorological and AQ. Information about data managed by national, regional and / or local authorities and the characteristics of the measurement points. This is essential for the modelling and could be difficult to arrange because of data formats. Examples:
    a. Aveiro: http://qualar.apambiente.pt/index.php?page=2&zona=&year=2016&month=6&day=03
    b. Liguria: http://www.cartografiarl.regione.liguria.it/SiraQualAria/script/Pub2AccessoDatiAria.asp?Tipo=CinqueAnni&_ga=1.30700854.847368869.1464961130
2. Meteorological data (incl. hourly resolution of temperature)

**Figure 7. City checklist**

Within a block, more detailed data were requested from the city groups. Suggestions were made towards data sources from International or European organisations or EC-induced data regulations.

The list for data requirements to evaluate emissions and carbon footprint from industrial, residential, commercial and institutional (IRCI) sector (Figure 8) indicates the kind of details necessary. There are, however, a number of European-wide sources, like CORINE[15] and E-PRTR[16] and European guidelines providing useful resources. At European level EEA[17] collects and distributes european air quality information reported by EEA member countries, including all EU Member States, as well as EEA cooperating and other reporting countries.

The EEA's air quality database consists of a multi-annual time series of air quality measurement data and calculated statistics for a number of air pollutants. It also contains meta-information on the monitoring networks involved, their stations and measurements, as well as air quality zones, assessment regimes and compliance attainments reported by the EU Member States and European Economic Area countries.

EEA also prepares and publishes annual reports[18] presenting an updated overview and analysis of air quality in Europe from 2000 to last available year data. It reviews the progress made towards meeting the air quality standards established in the two EU Ambient Air Quality Directives[19,20] and towards the World Health Organization (WHO) air quality guidelines (AQGs). It also presents the latest findings and estimates on population and ecosystem exposure to the air pollutants with the greatest impacts and effects. The evaluation of the status of air quality is based mainly on reported ambient air measurements, in conjunction with modeling data and data on anthropogenic emissions and their evolution over time.

In Figure 9 for $NO_2$ and Figure 10 for $PM_{10}$, annual average concentrations across Europe are mapped.

[15] The CORINE programme was initiatived in 1985. The databases, managed by the EEA, provided an inventory of land cover in 44 classes. https://www.eea.europa.eu/publications/COR0-landcover
[16] The E-PRTR is the European Pollutant Release and Transfer Register, managed by the EEA, providing easily accessible environmental data from industrial facilities. https://prtr.eea.europa.eu/#/home
[17] EEA, Air Quality e-Reporting (AQ e-Reporting)
[18] EEA, Air quality in Europe — 2018 report, EEA Report No 12/2018
[19] Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe https://eur-lex.europa.eu/eli/dir/2008/50/oj
[20] Directive (EU) 2015/2193 of the European Parliament and of the Council of 25 November 2015 on the limitation of emissions of certain pollutants into the air from medium combustion plants https://eur-lex.europa.eu/eli/dir/2015/2193/oj

**Data requirements to evaluate emissions and carbon footprint from industrial, residential and commercial (IRC) sector**

1. **Industrial sector**

a) We need the input for main industrial sources as extracted from E-PRTR and regional/local scale emissions inventories.
* E-PRTR ==> The European Pollutant Release and Transfer Register (E-PRTR) is the Europe-wide register that provides easily accessible key environmental data from industrial facilities in European Union Member States and in Iceland, Liechtenstein, Norway, Serbia and Switzerland. It replaced and improved upon the previous European Pollutant Emission Register (EPER). http://prtr.ec.europa.eu/#/home

b)The city/region partners (or buddies) are requested to indicate if there are main industrial sources relevant from air quality point of view.
c) For input to air quality model city/region partners (or buddies) must retrieve information about emissions and plant stacks characteristics (height, diameter or area, velocity and temperature of flue gases) of main sources (as selected by Techne Consulting using results from point a and b).

2. **Residential sector energy data**

a) Energy consumptions in domestic sector by energy vector at the most detailed geoghraphical detail (street/district level)
or, where not available data street/district:
- Energy consumptions in domestic sector by energy vector at city level (for example from Covenant of Mayors BEI). * Covenant of Mayors BEI: http://eumayorsarabic.eu/about/covenant-step-by-step-developments_en.html
- number of dwellings classified for energy vector used at street/district level where available.
- population data at street/district level

b) Share of penetration of technologies (for example traditional vs innovative, stoves vs fireplaces, single vs centralized boilers, boilers vs burners, ecc.) in domestic sector by energy vector at city, regional or national level (gross estimates if detailed data are not available);
c) Land use data (we will use CORINE Land Cover if no more detailed maps are available). * Corine Land Cover monitoring https://land.copernicus.eu/pan-european/corine-land-cover

3. **Commercial & institutional sector energy data**

Data requirements for commercial & institutional sector are:
a) Energy consumptions in commercial & institutional sector by energy vector at street/district level
or, where not available data street/district:
- Energy consumptions in commercial & institutional sector by energy vector at city level (for example from Covenant of Mayors BEI)
- Employees data at street/district level where available
- population data at street/district level

b) Share of penetration of technologies and land use data as in point 2.

4. **Other mandatory and non-mandatory data for Carbon footprint evaluation**

Data required at municipal level for carbon footprint evaluation are:
- Energy consumptions by energy vector (including electricity) for transport sectors (road, rail)
- Electricity consumption for public lighting services
- Final consumptions in other sectors (agriculture, forestry, fishing, small industries)

These data must be derived from transport module, Covenant of Mayors baseline or other sources
Finally the following municipal data, usually available in Covenant of Mayors baseline, are useful (but not mandatory) to compute the carbon footprint:
- Waste filled in landfills
- Waste water managed

These data must be derived from Convenant of Major baseline or other sources.

**Figure 8. Detailed data requirements for IRCI module**

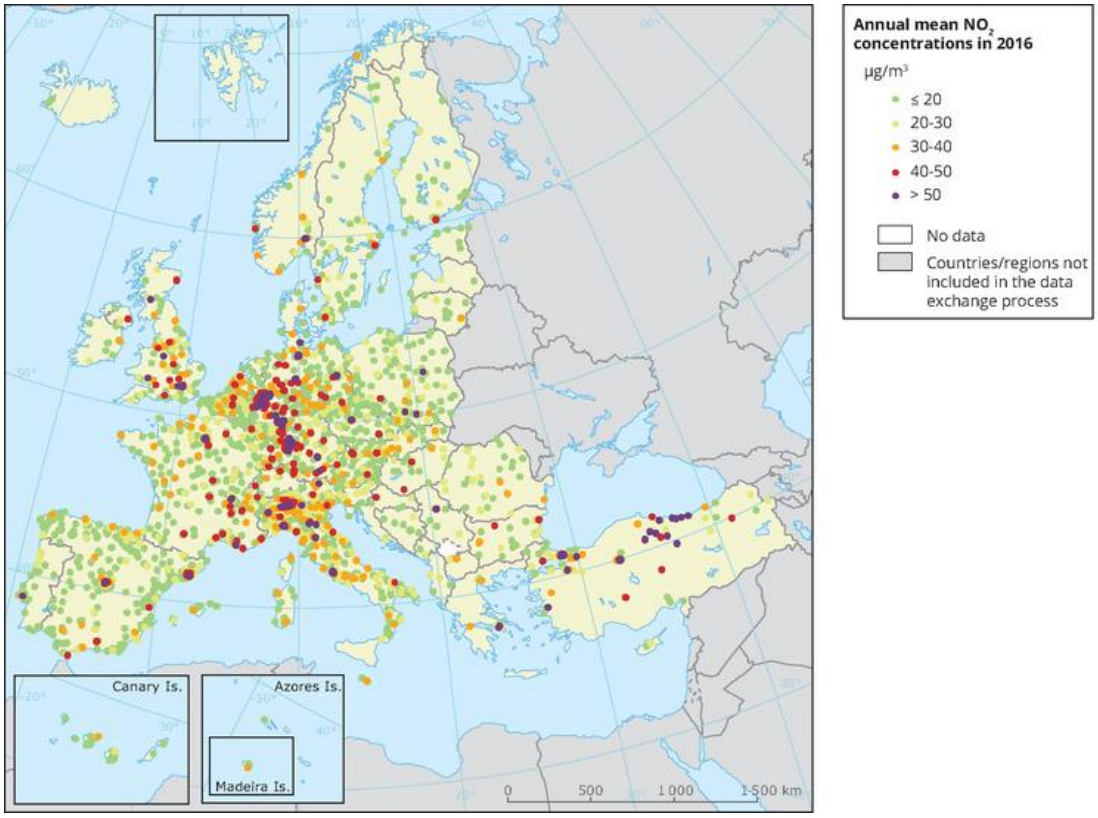



**Figure 9. EU $NO_2$ measurement network annual average concentrations**

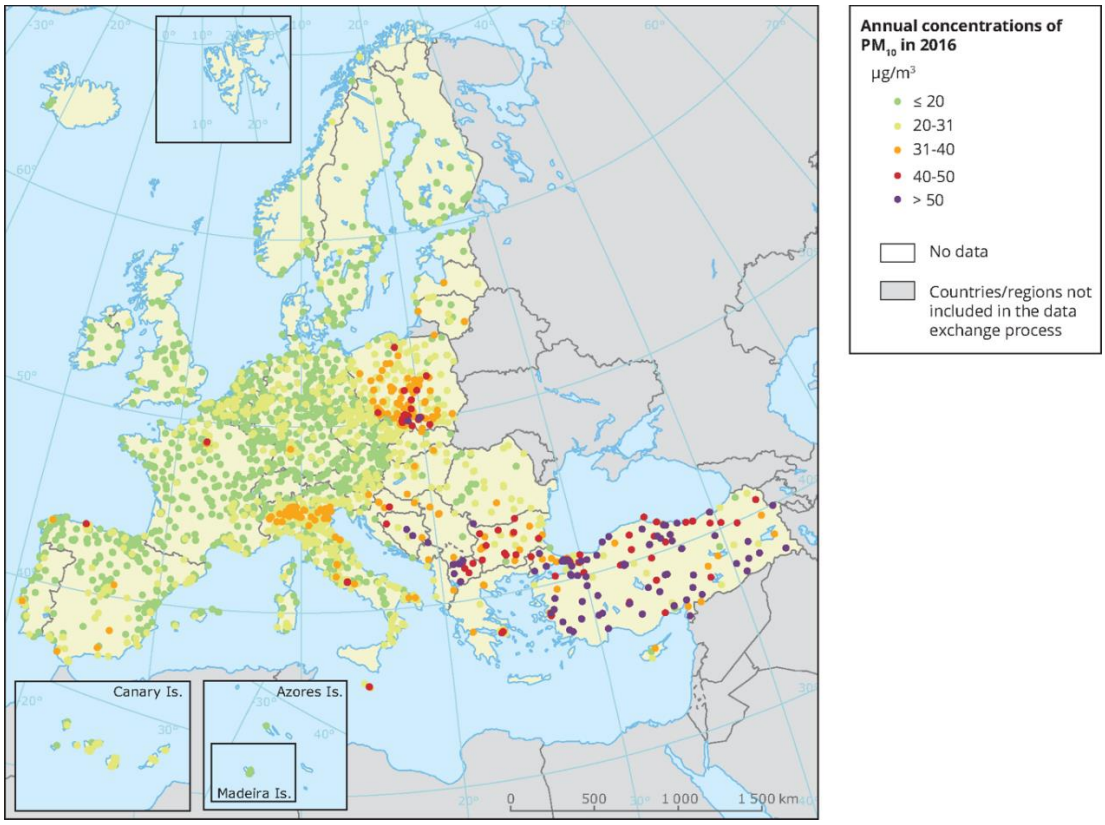



**Figure 10. EU $PM_{10}$ measurement network annual average concentrations**

The Luisa territorial modelling platform (https://ec.europa.eu/jrc/en/luisa) combines data on land-use with an activity-based model based upon dynamic allocation of population, services and activities. It produces a set of territorial indicators, that can be grouped according to the land function (e.g. societal, economic, environmental). Indicators are published in the Urban Data Platform (https://urban.jrc.ec.europa.eu, Figure 11).

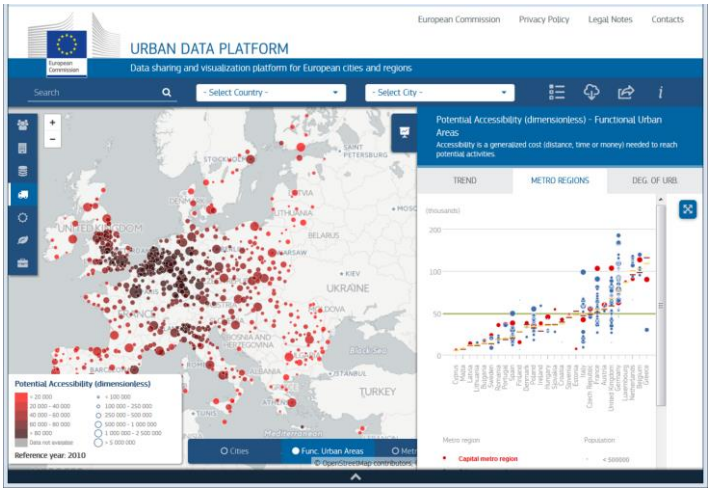



**Figure 11. The Urban Data Platform**

## 2.3 Alternative air pollution data sources

In the countries with ClairCity case studies (i.e. Netherlands, United Kingdom, Portugal, Italy, Poland and Slovenia), numerous monitoring projects have been performed, both with high-quality sensors, and with low-cost sensors. High-quality measurements give the necessary reference points of pollution levels, whereas low-cost sensors are useful for citizen engagement to understand temporal, and to a less extent spatial, variability but are not yet robust enough to accurately undertand concentration dependence or long terms trends. The high quality sensor data are usually official data from governmental organisations and are obviously key input to air quality modeling. The low-cost sensor data however are also interesting because these may be used to measure on more spots at the same time, and citizens may be involved to measure air pollution in their immediate environment. This section highlights a number of such projects.

In the project **Curieuzeneuzen 'Curious Noses'** 20,000 citizens in Flandres, Belgium participate by measuring with low-cost sensors the $NO_2$ concentration at their facades (https://curieuzeneuzen.be/). A pre-project in Antwerp in 2016 showed that low-cost sensors are not only useful for geodistribution modeling but it also showed that the involvement of citizens to measure their own immediate environment induced in these citizens more appreciation for planned measures and also a change in behaviour; more citizens were willing to take a less polluting transport mode to go to work (https://ringland.be/about/curieuzeneuzen/). The data will be used furthermore to examine

the effect of interventions in urban planning and to evaluate and improve modeling of the $NO_2$ distribution model.

Another example is the German measuring project **Luftdaten** (https://luftdaten.info/) with detailed explanations for citizens on how to assemble a laser $PM_{2.5}$/$PM_{10}$- device and how to connect it to the Luftdaten measurement network. This project gained some popularity in other EU countries as well, the network is growing in size and importance. There are even measurement stations in other parts of the world. Figure 12 shows the distribution of Luftdaten measurement stations across Europe and their output on December 20th. Sometimes local groups build on the Luftdaten knowledge, infrastructure and measurement philosophy. One example of such a group is the 'Luchtwachters (air quality guards) Delft' in the Netherlands; https://luchtwachtersdelft.nl (in Dutch).

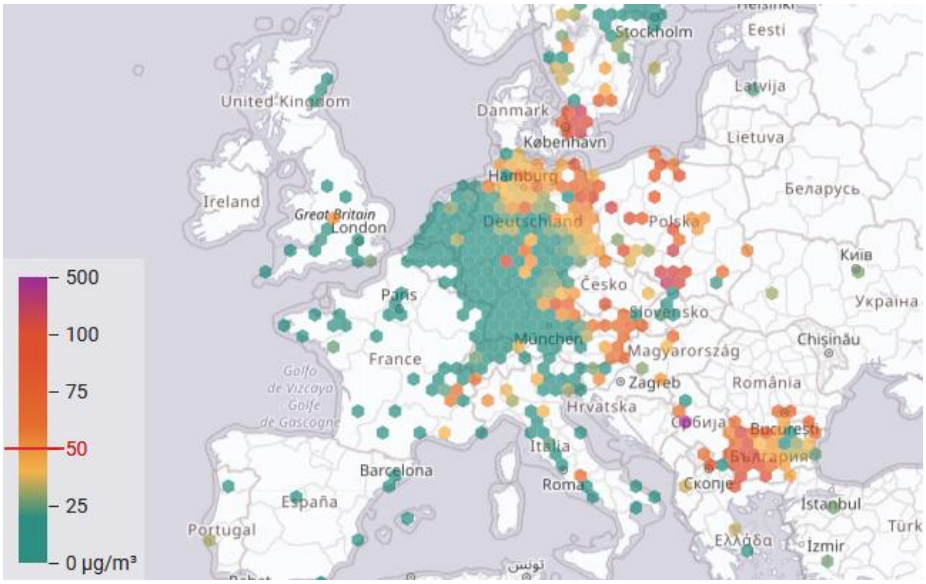



**Figure 12. Luftdaten measurement network.**

In the Netherlands, a project **'How healthy is our air?**' with engaged citizen groups was held in the years 2012-2014, measuring $NO_2$ near city roads. In Amsterdam, 17 locations were sampled. From the final report https://milieudefensie.nl/actueel/eindrapport-hoe-gezond-is-onze-lucht.pdf (in Dutch) a graph of Amsterdam shows a number of hotspots in the city (Figure 13). The actual levels on the measuring points were derived from the Palme's tubes values validated with measuring points of the governmental measuring network. The importance of this project is not the absolute levels measured, but the indication of local high levels –above threshold - and the engagement of citizens. In the Netherlands, the governmental network is maintained by RIVM, the Health and Environment Agency. 'Hollandse Luchten' (Dutch Skies, https://hollandseluchten.waag.org) is another Dutch citizens science project, starting 2019, involving citizens in measuring air quality in their environment. One of the goals of this project is to examine how citizens can gain more insight into air quality through open source and low-cost measuring sensors.

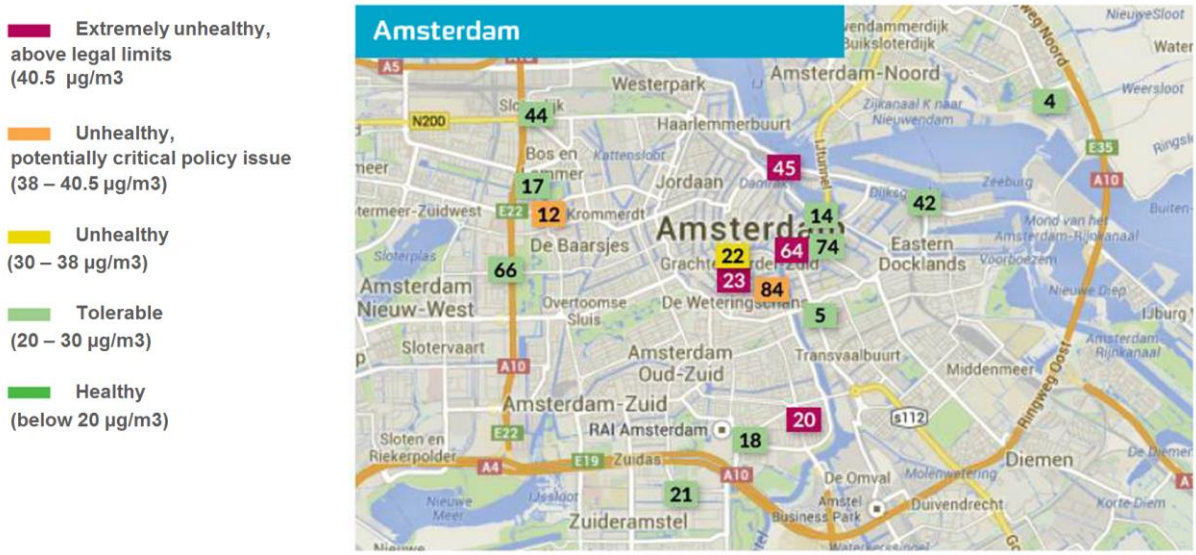



**Figure 13. Pollution levels in Amsterdam, October 2012 – March 2014 (Milieudefensie/RIVM, p. 11)**

In Norway during the Oslo **Green Capital campaign** citizens are called to help researchers measure air quality (https://www.greencapital2019.com/event/help-researchers-measure-oslos-air-quality). The campaign takes place in March-April when air quality deteriorates due to the resuspension of road dust that has accumulated during the winter. Measurement stations are put together from cardboard cartons and Vaseline. Data are collected and compared with the official air quality data.

Another interesting citizen science project is the **Telraam** project (https://telraam.net) from TML in Leuven, Belgium. Instead of measuring air quality, this project measures mobility. Citizens attach a low resolution camera to their window, and Python software running on a raspberry pi calculates hourly traffic counts of pedestrians, bicycles, cars, and trucks. Figure 14 shows the measured traffic intensity based on these data for a part of Leuven.

Yet another project that one might take note of in the Netherlands is the project **City Rhythm**, a cooperation of cities and scientists and the Statistics Netherlands Urban Data Center (UDC). They use hidden-Markov modeling of open and linked data. Repetitive patterns of citizens are identified and modeled like rhythms in music, thus creating new policy making spaces. For instance, identifying and matching the different rhythms of working people, students and mothers with children. Reference: https://books.bk.tudelft.nl/index.php/press/catalog/view/630/737/575-1

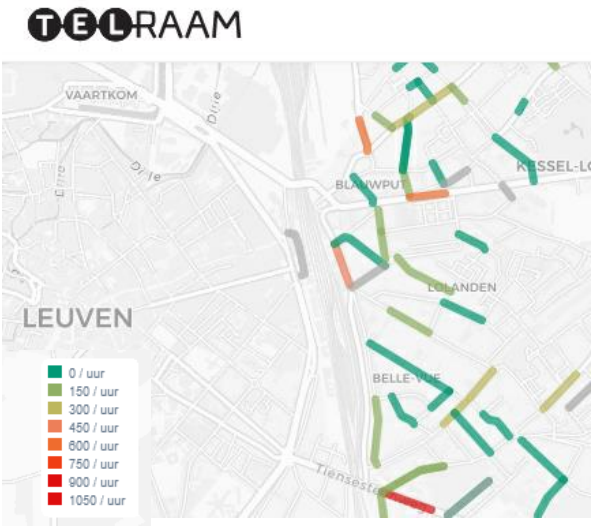



**Figure 14. Telraam traffic count project**

In Spain, Italy and Austria there was a network of citizen-science measurement stations called **captorAIR**, see https://captorair.org/about/. This network is part of the Horizon 2020 CAPTOR project with regard to ozone pollution in Europe. Its main objective was to inform the population about tropospheric ozone concentrations in real time (Figure 15). The captorAIR is free software but activities on the website e.q. data capturing ended on 31 August 2018. The CAPTOR project itself ended 28 February 2015.


**Figure 15. Captorair map**

In Portugal, researchers work together with a school community in an area with a lot of chemical industries on the implementation of a low-cost ambient air quality monitoring micro

sensors for population awareness and to support civil protection: http://ohm-estarreja.in2p3.fr/pt/projetos/details/3/1861. The main objective of this project is to evaluate the use of reliable air quality micro sensors and to contribute towards decision-making and civil protection support, based on big data approach.

**HackAIR** ( https://www.hackair.eu) is another EU funded air quality project which provides an open technology platform that you can use to access, collect and improve air quality information in Europe. Apart from open source software the project also provides some open access datasets. There are $PM_{10}$ and $PM_{2.5}$ measurements collected by hackAIR users with stationary or mobile hackAIR sensors and PM estimations from photos depicting sky from Flickr photos and photos from webcams.hackAIR is currently (March 2019) active.

At the hackair map (Figure 16, https://platform.hackAIR.eu), both hackAIR and Luftdaten sensor data may be viewed

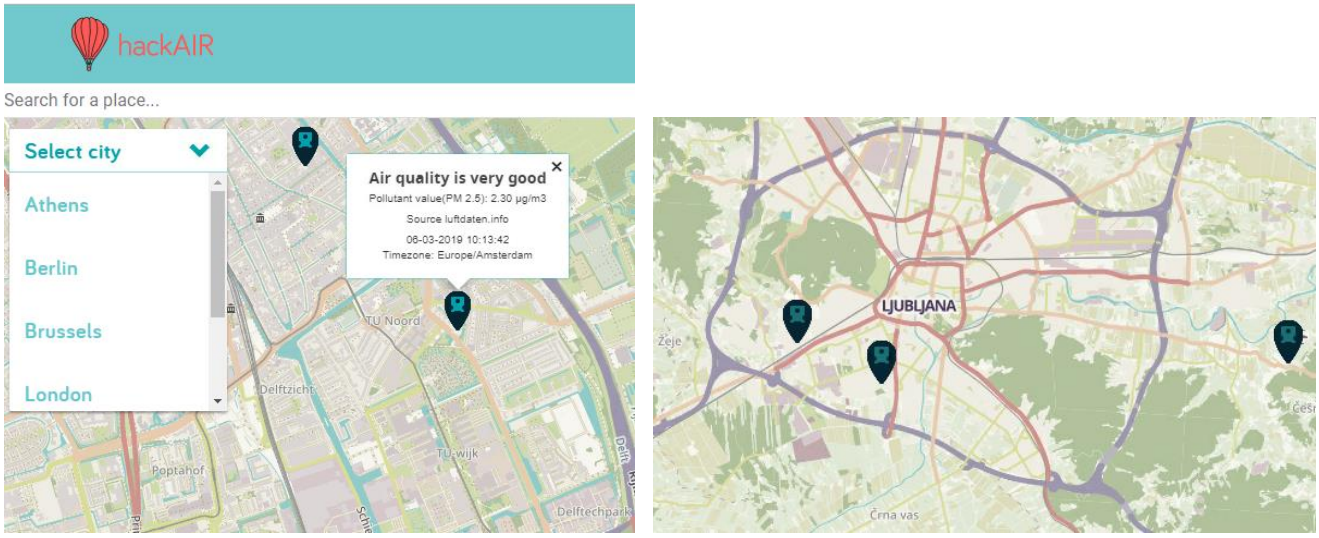



**Figure 16. HackAIR maps**

Two other European funded projects are **CITI-SENSE** (2012-2016, http://co.citi-sense.eu/) and **ACTION** (2019-2021, https://drift.eur.nl/projects/action-participatory-science-toolkit-against-pollution/). The CITI-SENSE project offers a central web portal, which provides interested parties access to products and services, e.g., mobile apps and air quality perception surveys, and serves as a forum for discussion and debate. The ACTION project wants to transform citizen science from a mostly scientist-led process to a more participatory, inclusive, citizen-led one.

The Joint environmental protection agency (DCMR) in the Rijnmond area, the Netherlands, has launched in 2019 a jubilee project to involve citizens in various municipalities in measuring air quality ($NO_2$) with “palmes buisjes” (Palmes diffusion tubes). Website in Dutch: https://www.dcmr.nl/onderwerpen/zelf-meten.html.

The **Air Quality Egg** learning system (https://airqualityegg.com/home) is an education system based on air quality measurements with an egg-like device. It contains 7 air contaminant sensors and can be used to raise awareness by children on air quality.

## 2.4 Visualizing practice-activity data

Practice-activity data and the effect of changing activity patterns on air quality may be complex to explain to the public. Some scenarios may show quick wins but only for a limited period while others may show positive effects for a longer period but need longer time to become visible. ClairCity explored a few ways to communicate practice-activity data to the public. One of these concepts is the concept of stastically safe dotmaps.

Statistically safe dotmaps aim to visualize demographic variables on a detailed regional level while preserving privacy aspects. Dotmaps have been used in earlier research to display population density in combination with ethnic origin in the Netherlands[21]. The same technique appeared to be very useful to visualize variables that are relevant for air quality scenario development. Figure 17 shows the dotmap for the Amsterdam city. The interactive version is available from the ClairCity data portal via https://claircitydata.cbs.nl/pages/dotmaps.

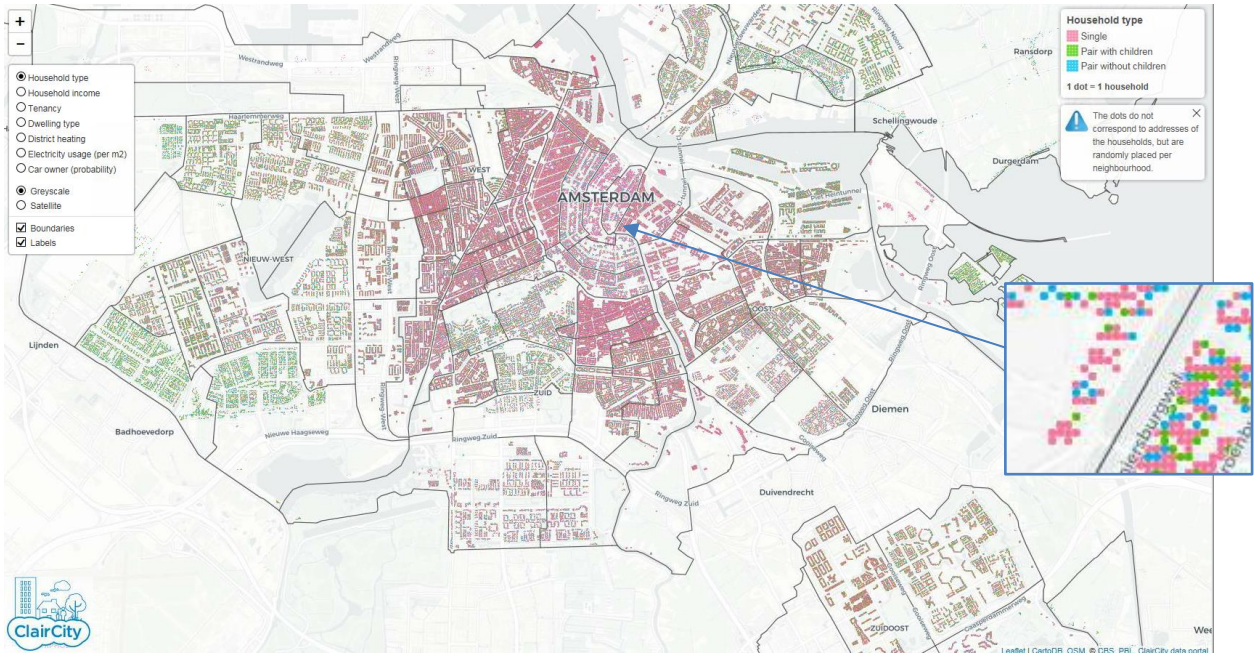



**Figure 17. Dotmap for Amsterdam**

Several variables can be selected such as household type, household income, dwelling type, the use of district heating, number of cars, the existence of a wood stove or fireplace and emissions from gas consumption. The distribution of the indicator chosen is shown in dots

[21] See https://www.cbs.nl/en-gb/our-services/innovation/project/cbs-experimenting-with-dot-maps

with different colors. The dots are distributed over the regions where people actually live. It is planned to extend these maps with air quality layers and a time dimension so that dependencies among practise-activity data, emissions and scenarios can be shown.

The level of data detail needed to make such a map is high. Practise-activity data on a low regional level is needed. In addition, even more detailed land-use data is needed on the actual borders of areas that have a residential function. In a first version of the Dotmap for Amsterdam, this data was taken from the register of addresses and buildings (BAG) of the Netherlands. However, such a register is not available in every European country participating in ClairCity. Moreover it was felt that it would be too strict to require such register for each country where one would apply the ClairCity approach.

Therefore a second, less detailed, version of the dotmap for Amsterdam was created using community data from OpenStreetMap (https://openstreetmap.org; regions with attribute land-use = residential). It was expected that the level of detail of OpenStreetMap was good enough to apply this in every ClairCity city. Unfortunately this was not the case. At the time of writing, a generic solution has not yet been found. For the time being we note that a good practice is to look for detailed data on residential areas in a national or local administration, or to derive it from a community source such as OpenStreetMap or another public map or satellite image.

# 3 What practice-activity data are collected and what good practices can we identify?

As mentioned in the introduction of this document, one of the goals of this report is to put together all the knowledge for those that want to make the same data journey as in ClairCity for modeling air quality scenario's for their own region or city. In this chapter, we take a very practical approach to that. We look at the data stored in the ClairCity Data Portal, their origin as well as their use in the ClairCity modules and we infer good practices.

In the following subsections, we perform an analysis of the data stored in the ClairCity Data Portal and the data sources that were used to derive them. We identify for all modules and pilot cities / regions what data from which organisations and, wherever possible, from which data source were derived to perform the ClairCity modeling for that city / region. This shouldn't be interpreted too strictly, the aim is to identify the main data sources needed to be able to derive the good practices for consecutive projects.

For practical reasons, if the same data source is used in multiple cities / regions it is listed at the first city / region only. In addition, generic data that is in principle usable for any city is listed under a generic section. For completeness, we note that at time of writing, the ClairCity project is progressing considerably but it is not completely finished yet. Therefore, the data landscape sketched here has to be interpreted as a snapshot. However, since data projections have been calculated for a majority of the cities, we feel the picture is complete enough to draw conclusions on the good (and less good) practices collecting practice activity data for ClairCity-like activities.

In the final sub-section of this chapter, we translate and interpret the observed data sources into a set of practical instructions of where practice-activity data can be found and how to obtain and process them. Wherever possible we include references to organisations and data repositories and some hints on how to connect, download and process the data.

## 3.1 Integrated urban module

The integrated urban module is the module that elaborates the household and dwelling characteristics. The integrated urban models households' choices with a direct or indirect effect on emissions, focussing on geographical and temporal demands for transport, housing and residential energy consumption, and consumption of goods. The module of Transport and the module of Industrial, Residential, Commercial and Institutional build on this module.

### *3.1.1 Generic*

For the integrated urban module data, there are two major generic data sources available.

**Danish Time-Use Survey (TUS) Data:** Time-use-survey data collected from randomly drawn families registered in Danish population registration office, from March 2008 to March

2009, including 41 activity types at 10-minute intervals. This data set was used as a generic starting point for time practice activity data for those cities where no more detailed regional time use data was present. Alternative data may be found at https://www.timeuse.org/information/access-data.

**Synthetic Electricity consumption time series based on Danish TUS data**: Synthetic time series of residential electricity consumption. The synthetic data is based on the Danish time-use-survey data sets. This data set was used as a generic profile for those cities where no more detailed regional data was present.

### *3.1.2 Bristol*

**Bristol synthetic population**: A synthetic population dataset for Bristol with data on dwelling type, tenure, vehicle ownership, household composition (size and members by age), central heating fuel and household fuel consumption (for oil, gas and electricity) was created. Data on population projections was generated using population and household projections of Bristol's household population by age group of household head, and household structure (household size and number of children) from the official ONS (Office for National Statistics) statistics office https://www.ons.gov.uk/. Also, Bristol Area census data on the level of LSOAs (Lower Super Output Areas) including electricity and gas usage was used (a single LSOA contains approximately 700 households). The age group and employment rate information used in the weighting composition was obtained from Bristol's open data: https://opendata.bristol.gov.uk/pages/homepage/, https://gov.uk/government/statistics/ and https://www.nomisweb.co.uk/

**Bristol LSOA D3 shapefile**: Shapefile containing the modeling boundary for Bristol (20 x 20 km) and the definition of regions to be used for modeling. For Bristol this is the level of Lower level Super Output Areas (LSOA) as defined by the Office for National Statistics (ONS).

### *3.1.3 Amsterdam*

**Amsterdam population**: A synthetic population dataset was developed for Amsterdam, generated using a combination of the Iterative Proportional Fitting (IPF) algorithm and mixed integer programming. Synthesized population based on Woningonderzoek Nederland (WoON) 2015, Onderzoek Verplaatsingen in Nederland (OViN 2011-2015) and the Buurt- and Wijk-kaarten from CBS (2014): https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data. Households electricity was estimated using two simple linear regression models.

**Amsterdam activity data**: This data set contains the predicted activity of the synthetic population in terms of 9 distinct activities and 13 time slots.

**Amsterdam Buurt D3 Shapefile**: Shapefile containing the modeling boundary for Amsterdam (25 x 20 km) and the definition of regions to be used for modeling. For

Amsterdam this is the level of neighbourhoods as defined by the Statistics Netherlands (CBS).

### *3.1.4 Ljubljana*

**Ljubljana population**: Due to lack of appropriate data with which to generate a **synthetic** population, we followed Approach 2, as described in Section 2.1. Population data is available from the statistics office of the republic of Slovenia – SURS. The online database can be reached via http://pxweb.stat.si/pxweb/Dialog/statfile1.asp and contains national and regional data; GIS data can be obtained via the STAGE app: http://gis.stat.si/.

**Ljubljana Naselja D3 shapefile:** Shapefile containing the modeling boundary for Ljubljana (20 x 20 km) and the definition of regions to be used for modeling. For Ljubljana this is the level of settlements as defined by the Republic of Slovenia Statistical Office (SURS).

### *3.1.5 Sosnowiec*

**Sosnowiec Gminy D3 shapefile:** Shapefile containing the modeling boundary for Sosnowiec (20x20 km) and the definition of regions to be used for modeling. For Sosnowiec this is the level of communities as defined by Statistics Poland.

### *3.1.6 Liguria*

**Liguria SezCens D3 shapefile**: Shapefile containing the modeling boundary for Liguria (25 x 15 km) and the definition of regions to be used for modeling. For Liguria this is the level of census sections as defined by Statistics Italy (ISTAT).

### 3.1.7 Aveiro Region

**Aveiro Region synthetic population**: A synthetic population dataset was obtained from Census 2011 from the National Statistical Institute of Portugal (https://ine.pt/). The dataset include information about number of inhabitants, age groups, gender, education and employment status at municipality level.

**Aveiro Region CIRA shapefile**: Shapefile containing the modeling boundary for CIRA / Aveiro (40 x 55 km) and the definition of regions to be used for modeling.

## 3.2 Transport module

The transport module is the module where emissions from current and future transport of persons and goods are modeled.

### *3.2.1 Generic*

For transport emission estimates, there is a variety of public sources that produce city-independent data. First of all, we rely on a single approach for the emission factors:

**COPERT**: The EU standard vehicle emissions calculator was used, available at: https://www.emisia.com/utilities/copert

Secondly, for simulation of the emission factors in the future, we rely on a fleet model that simulates vehicle turnover (scrappage, sales, retention) and uptake for new technology in the fleet. It combines the following generic car fleet modules: TREMOVE https://www.tmleuven.be/nl/navigation/TREMOVE, MOVEET https://www.tmleuven.be/nl/navigation/MOVEET, TRANSPHORM http://www.transphorm.eu/ and the LUTI model for combining transport and land use.

Data with specific practice-activity characteristics in **bold**:

- Fleet statistics (ACEA, EuroStat, TiF): **age structure**, fuel types, emission standard, **vehicle size**: National statistics are typically sufficient for modeling. We found that for Bristol, the city-specific fleet does not differ that much from the national fleet; furthermore the city-specific fleet (ownership) is not, per se, a good reflection of the actual **active** fleet in the city. National statistics suffice and are publically available, see separate city / region sections below.
- Travel survey data: **mode choice, motive for travel**: These data are essential for the practice-activity relation: What is driving transport demand? Full survey data are available for Bristol (see Section 3.2.2), but not for others. However, limited survey data or national data are valuable as a proxy.
- Traffic counts by individual road links, by **time of day**, by vehicle category: These data may be available for larger cities, but not for smaller cities. The capability for transport modeling is not available in each city; however using proxy data, a good estimate may still be made, using for instance the OpenStreetMap - TRANSPHORM mesh.
- Detailed population data: Detailed data on **population properties** are necessary to link transport demand to population (see Section 3.1 Integrated urban module*).*

Other transport data without an explicit practice-activity link:

- TRANSPHORM data: General transport indicators such as average occupation of transport modes and emission factors: http://www.transphorm.eu
- Road network data: OpenStreetMap (http://www.openstreetmap.org
- EU Emission inventory data of non-transport (EMEP: https://www.emep.int). The European Monitoring and Evaluation Programme (EMEP) is a scientifically based and policy driven programme under the Convention on Long-range Transboundary Air Pollution (CLRTAP) for international co-operation to solve transboundary air pollution problems.

### *3.2.2 Bristol*

Data on car use for travel to work from UK Office for National Statistics. Nomis statistics. https://www.nomisweb.co.uk/census/2011/lc7401ew

UWE has obtained low-level data on vehicle characteristics and calculated mileages linked to LSOA, from UK Department for Transport (https://www.dft.gov.uk/).

- Road network including annual traffic volumes (countings + modeled)

### 3.2.3 Amsterdam

**Dutch Governmental pollutant release and transfer register**: Data from the Dutch emission register was used. Accessible via: http://www.emissieregistratie.nl/erpubliek/bumper.en.aspx

- **Road transport emission maps – Amsterdam:** 2x7 separate layers of transport emission for PM and $NO_x$ for 7 vehicle types.
- (Detailed) travel survey data.
- Road network including annual traffic volumes (countings + modeled)
- Fleet data
- Emission-inventory data for shipping
    - Maritime
    - Inland waterway
    - Pleasure crafts

**Amsterdam passenger trips**: This data set contains activity specific origin-destination data on postal code level (4 digits) per day and hour for Amsterdam.

### 3.2.4 Ljubljana

**Car fleet characteristics**: The OPSI Open data portal Republic of Slovenia has open data on car fleet and characteristics, see: https://podatki.gov.si/dataset/evidenca-registriranih-vozil-presek-stanja

**Public transport network**: Data on the public transport network is avaliable from: http://www.lpp.si/

**Traffic counting**: Data on traffic counts is available from: https://prominfo.projekti.si/web

**Transport usage**: National data on the use of transport is available from the national statistical office: https//www.stat.si/StatWeb/en/Field/Index/24/82. More regional data on transport usage can be obtained from the following document (in Slovenian): http://www.mzi.gov.si/fileadmin/mzi.gov.si/pageuploads/Kabinet_ministra/Prometne_navade_prebivalcev.pdf.

**Transport emissions**: Non-road transport emissions are available at: http://okolje.arso.gov.si/onesnazevanje_zraka/uploads/datoteke/IIR_2016_Slovenia.pdf and (in Slovenian): https://www.ljubljana.si/assets/Uploads/Ocena-emisij-v-zrak-za-leto-2016.pdf.

### 3.2.5 Aveiro

**Transport network:** Three shapefiles were obtained from CIRA, namely:

(i) road network in the projected coordinate system ETRS_1989_TM06-Portugal (*confidential data - data can not be used on other projects or by people outside the ClairCity project*);
(ii) (ii) cycle network in the geographic coordinate system GCS_WGS_1984; and
(iii) rail network in the projected coordinate system ETRS_1989_TM06-Portugal (*confidential data - data can not be used on other projects or by people outside the ClairCity project*).

## 3.3 Industrial, residential, commercial and institutional module

The IRCI module integrates distinct emissions sources from 4 sectors.

### *3.3.1 Industrial sector*

For industrial sources emissions, a specific tool has been developed to include all the emissions from the European Pollutant Release and Transfer Register (E-PRTR)[22] facilities and national and local Registers or emissions inventories (national, regional and local scale). Where specific facilities are individually identified that don't have known emissions data, ad hoc estimates are obtained using available information and emission factors from EMEP/EEA Guidebook[23]. The emissions have been geographically allocated by the coordinates of the emission source. When data on a single facility are not known, emissions are evaluated from statistical sources as area sources and allocated using land cover maps.

Large point sources have been individually identified and reported, source by source, with emissions and characteristics of emissions. Large point sources are defined as sources that emit more than 100 Mg of $NO_x$ or $PM_{10}$.[24]

### *3.3.2 Residential, commercial and institutional sector*

#### *3.3.2.1 Emission modeling*

The IRCI tool:

- evaluates emissions at the most detailed administrative territorial units' level,
- uses emission factors from EMEP/EEA Guidebook,
- calculates emissions as:

$$E = A_{ij} F_{ij}$$

[22] European Pollutant Release and Transfer Register (E-PRTR)
[23] EMEP/EEA(2016), Air pollutant emission inventory guidebook
[24] Trozzi C, Villa S, Vaccaro R, Piscitello E. National and regional emission projections in Europe. Methodology, tool and case studies. Conference paper, United States Environmental Protection Agency 20th Emissions Inventory Conference "Emission Inventories - Meeting the Challenges Posed by Emerging Global, National, and Regional and Local Air Quality Issues". Tampa, Florida, August 13 - 16, 2012

where:

- $A_{ij}$ is the indicator of the activity i in the territorial unit j,
- $F_{ik}$ is the emission factor of pollutant k for activity i (expressed in grams per unit of activity).

For small combustion sources in residential, commercial and institutional sector, emissions have been evaluated using emission factors from EMEP/EEA Guidebook and at specific activity level.

#### *3.3.2.2 Emission factors*

In Table 4 combustion air pollutants emission factors used are reported[25].

**Table 4 – Residential, Commercial & Institutional Air pollutant Emission Factors**

| Fuel | Nitrogen oxides [$gNO_x/GJ$] | Particle Matter with diameter less than 10μ [$gPM_{10}/GJ$] |
|---|---|---|
| Boilers – Natural gas | 42 | 0,2 |
| Boilers – LPG | 40 | 2 |
| Boilers – Gas/Diesel Oil | 69 | 1,5 |
| Boilers – Wood | 80 | 480 |
| Fireplaces – Wood | 50 | 840 |
| Stoves – Wood | 50 | 760 |
| Energy Efficient Fireplaces – Wood | 80 | 380 |
| Energy Efficient Stoves – Wood | 80 | 380 |
| Advanced Fireplaces – Wood | 95 | 95 |
| Advanced Stoves – Wood | 95 | 95 |
| Pellets Stoves – Wood | 80 | 29 |
| Boilers - Hard Coal | 158 | 225 |

#### *3.3.2.3 Data evaluation on domains and subdomains*

Air pollution emissions are evaluated on defined subdomains. When data are available only in aggregate figures (overall domain) or at national level, they are allocated to subdomains using a “proxy” variable available at subdomains level and are evaluated using other proxy variables known at the subdomains. Proxy variables allow obtaining information on a certain spatial resolution assuming that it is known for larger spatial resolutions.

When data are available at city/region domain, data at subdomain level is evaluated using the following equation:

$$A_i = A * P_i / \Sigma_i P_i$$

where: $A_i$ and $P_i$ are the values of variable A and proxy variable P in the subdomain i, and A

[25] EMEP/EEA(2016), Air pollutant emission inventory guidebook,1.A.4 Small combustion

is the total of variable A in the domain.

When data are available at national level, data at subdomain level is evaluated using the following equation:

$$A_i = (A * Q_d / \Sigma_d Q_d) * P_i / \Sigma_i P_i$$

where: $A_i$ and $P_i$ are the values of the variable A and proxy variable P in the subdomain i, $Q_d$ is the values of the proxy variable Q in domain d, and A is the national value of proxy variable A.

### *3.3.3 Data sources*

#### *3.3.3.1 Bristol*

The following tables in the appendix record the methodology and data used for:

- Industrial sources (Table 8);
- Residential and commercial sources (Table 9);
- Wood statistics (Table 10);
- Aggregate fuel consumptions data subdivision (Table 11); and
- *LSOA* regional disaggregation variables (Table 12).

#### *3.3.3.2 Amsterdam*

The following tables in the appendix record the methodology and data used for:

- Industrial sources (Table 13)
- Residential and commercial sources (Table 14)
- Wood statistics (Table 15); and
- *Buurt* regional disaggregation variables (Table 16)

#### *3.3.3.3 Ljubljana*

The following tables in the appendix record the methodology and data used for:

- Industrial sources (Table 17);
- Residential and commercial sources (Table 18);
- Wood statistics (Table 19); and
- *Naselje* regional disaggregation variables (Table 20).

#### *3.3.3.4 Aveiro*

**Point sources:** A dataset obtained from multiple sources includes information about the main industrial point sources, namely: (i) localization (lat, lon); (ii) emissions of $NO_x$ and $PM_{10}$; (iii) stack height and diameter; and (iv) velocity and temperature of flue gases.

**Energy data:** Six datasets were obtained from Census 2011 (https://censos.ine.pt) and National Statistical Institute of Portugal (https://ine.pt/). The datasets include information

about: (i) annual consumption of natural gas per 1000 inhabitants ($Nm^3$) at municipality level for the years 2011 to 2016; (ii) annual household consumption of electric energy by inhabitant (kWh/ inhab.) at municipality level for the years 2011 to 2016; (iii) number of dwellings, according to type of heating available (electricity, wood, gas, ...), at the municipality and parish (freguesia) levels for the year 2011; (iv) consumption of electric energy (kWh) by type of consumption at municipality level for the years 2011 to 2016; (v) number of enterprises by economic activity at the municipality and parish levels for the years 2011 to 2016; and (vi) number of persons employed of establishments by economic activity, at municipality level for the years 2011 to 2016.

**Fuel consumption**: Two datasets were obtained from Pordata (https://www.pordata.pt/) and National Statistical Institute of Portugal (https://ine.pt/). The datasets include information about: (i) fuel sales for consumptions (tonne) at municipality level for the years 2011 to 2017; and (ii) consumption of motor fuel by inhabitant (toe/ inhab.) at municipality level for the years 2011 to 2016.

## 3.4 Energy / power generation module

The energy / power generation module distributes annual energy consumption and emissions into an hourly resolution.

### *3.4.1 Generic*

The generic methodology for calculating energy / power models is described in Section 2.1 through our case study examples. Four methods of modeling were described depending on the availability of detailed regional energy data. There is an example of the work carried out for Bristol. This method was used for all other cities.

## 3.5 Agriculture module

The agriculture module calculates emission contributions from agricultural activities including livestock.

### *3.5.1 Generic*

**Emission data:** CEIP (the EMEP Centre on Emission Inventories and Projections) collects emissions of air pollutants (gases and particulate matter) from parties to the convention on Long-Range Transboundary Air Pollution (LRTAP) and makes available an European emission inventory over a grid at 0.1 x 0.1 degrees resolution (~10 km) covering the whole Europe - http://www.ceip.at/ms/ceip_home1/ceip_home/new_emep-grid/01_grid_data/ . Gridded emissions are available by pollutant and by activity sector (GNFR classification)

**Land cover data**: The Open Street Map database is a powerful geographic tool that represents physical features on the ground (e.g., roads or buildings) using tags attached to its basic data structures (its nodes, ways, and relations). Each tag describes a geographic attribute of the feature being shown by that specific node, way or relation. Several land use feature types are included. The Open StreetMap is available for free download at http://download.geofabrik.de/

For all case studies, the EMEP European emission inventory for the year 2015 was used to derive the agriculture and livestock emissions for the urban domains, by performing a disaggregation of emissions by farms, meadows, vineyards land uses classified in the Open Street Map database.

## 3.6 Natural module

The Natural module adds the natural background emission contributions.

### *3.6.1 Generic*

**Emission data**: For all case studies, the EMEP European emission inventory (described in Section 3.5.1) for the year 2015 was used to derive the natural emissions for the urban domains, by performing a disaggregation of emissions by forest, pasture, parks and nature reserve areas classified in the Open Street Map database, as mention in the previous section.

### *3.6.2 Bristol*

**Reference air quality data**: Reference air quality data on NOx, NO2, PM10, PM2.5, CO, benzene, and 1,3-butadiene available at https://laqm.defra.gov.uk/review-and-assessment/tools/background-maps.html

## 3.7 Emission database module

The emissions database module is the module that combines the emissions of Transport, IRCI, Energy/power, Agriculture, and Natural background.

### *3.7.1 Generic*

In ClairCity, $NO_x$, $PM_{10}$ and $PM_{2.5}$ were the primary emissions of concern. The emission values were divided by three main emission sources: point, line and area sources. **The same methodology is used for all case studies.**

**Input:**

**Point sources:** The emissions are provided from large industries with detailed source characterization and annual emission value higher than 100 Mg/year.

**Line sources:** Emissions from main traffic lines in Mg/h. These emission data include the daily profile of the weekday and weekend for six traffic categories (car, motor, light commercial vehicle (LCV), medium truck, heavy truck, and bus).

**Area sources:** Emissions from residential, commercial and industries (industrial sources with an annual emission less than 100 Mg/year) were considered. Daily and hourly profiles were applied to the annual emission value. The hourly profiles were provided for four typical days considering the annual seasonality (weekday and weekend for winter and summer seasons). In addition, emission data from large industries with an annual emission rate lesser than 100 Mg/year were considered as area sources.

**Methodologies applied:**

**Point sources:** The point source emission data were processed using the intersect tool embedded into ArcGIS. The main objective was to identify, in the simulation grid, the location of the point sources before this information was adapted to include in the model inputs.

**Line sources:** The line sources were processed using one set of the tools embedded into ArcGIS to gridded the emissions into 200m x 200m cells (horizontal resolution used in urban air quality model). In the first step, the intersect tool was applied between the road network and the model grid. A factor was applied based on the road length. The dissolve tool was then used to aggregate emission values based on the model grid identification. The result from this methodology is the traffic emission data gridded into cells of 200m x 200m spatial resolution. This methodology was applied differentiating only two typical days (weekday and weekend) with hourly temporal resolution.

**Area sources:** The residential and commercial emissions were processed using the total annual emission values for the different neighbourhoods included within the simulation domain. In the first step, the total annual values were processed using the ArcGIS intersect tool to grid the emissions into cells of 200m x 200m. The union tool was used to join information from links and the model grid. The dissolve tool was then used to aggregate emission values based on the model grid identification. Finally, the daily and hourly profiles were used to allocate the emission data in an hourly-basis temporal resolution.

The industrial sources with an annual emission value less than 100 Mg/year were processed applying the same methodology as used for point sources. The intersect tool was applied to allocate the location of the industrial sources into the model grid. Finally, the emisson data were included in the model inputs as area sources.

**Output:**

**Point sources:** The output is one CSV file with point source emission characterization and identification of their corresponding location in the model grid.

**Line sources:** The outputs are twelve CSV files, six files (one for each traffic category) for two typical days with hourly spatial emissions for grid model input.

**Area sources:** The outputs are four CSV files. These files contain the daily emissions profile for each typical days representative of the annual seasonality (weekday and weekend in winter and summer).

Other output is one CSV file with emission characterization from industry sources (with annual emission value less than 100 Mg/year) including the identification and their corresponding location in the model grid.

Below we summarize the auxiliary data sources used in the emission database module.

**Input data:**

- **Point sources:** Location of the emissions point sources with detailed emissions characterization.

**Line sources:** Road map with the emissions for typical days (weekday and weekend) with different hourly profile.

- **Area sources:** Spatial distribution of the residential and commercial emissions (total annual values)**; l**ocation of the industries sources with emission value less than 100Mg/year with detailed emissions characterization.

**Output data**:

- **Point sources:** Position of the point sources over the model grid.

**Line sources:** Example for one hour of the spatial distribution of the traffic emission.

- **Area sources:** Example for one hour of the spatial distritubion of the commercial and residential emissions.

## 3.8 Air quality module

The Air quality module is the module that assesses the air quality based on the emissions from the Emission module.

### *3.8.1 Generic*

Two air quality dispersion modeling tools were used:

**WRF-CAMx**: A combination of the meteorological model WRF, Weather Research and Forecasting model (https://www.mmm.ucar.edu/weather-research-and-forecasting-model), a next-generation mesoscale numerical weather prediction system and CAMx, the Comprehensive Air Quality Model with extensions (http://www.camx.com/), a multi-scale photochemical modeling system for gas and particulate air pollution. CAMx includes a source

apportionment (SA) tool that estimates the contributions from multiple source areas, categories, and pollutant types to the spatial and temporal distribution of PM.

**URBAIR model**: The URBAIR model is a Gaussian dispersion model with continuous varying parameters with the atmospheric stability; it is developed by the research group on Emissions, Modeling and Climate Change of the University of Aveiro, more information at: https://www.ua.pt/gemac/.

The numerical model, URBAIR was setup and run at urban scale for the computational domain defined around the urban area. The baseline scenario simulations were performed using the meteorological vertical profiles and the background concentrations provided by the WRF-CAMx system as input data. The air quality simulations were performed for the full-year based on a full year of meteorological data and the emissions established for the typical days over summer and winter times considering the weekday and the weekend. The concentrations of $NO_2$ and $PM_{10}$ were simulated using the emission rates available in the ClairCity emission database, including the point sources with the large industry emissions, the line sources with the road traffic emissions and the area sources covering the residential, commercial and industrial emissions. The meteorological inflow conditions were used in an hourly basis and the emission data were replicated also in an hourly basis in order to perform hourly air quality simulations. The $NO_2$ and $PM_{10}$ annual averaged concentrations, without and with background concentrations, considering all emissions sectores, are available in the ClairCity dataportal, both the figures and the matrix data (in .mat files).

### *3.8.2 Bristol*

**Online air quality data**: There is online data on air quality for Bristol available from: https://opendata.bristol.gov.uk/pages/aqdashboard/#about-air-quality

UK air quality data and forescasts are available from: https://uk-air.defra.gov.uk/

### *3.8.3 Ljubljana*

**Online air quality data**: There are city and national air measurement stations in Ljubljana generating data, available at: https://www.ljubljana.si/sl/moja-ljubljana/varstvo-okolja/stanje-okolja/kakovost-zraka/

### *3.8.4 Liguria*

**Online air quality data**: There is online data on air quality available from: http://www.cartografiarl.regione.liguria.it/SiraQualAria/script/Pub2AccessoDatiAria.asp?Tipo=CinqueAnni&_ga=1.30700854.847368869.1464961130

### *3.8.5 Aveiro*

**Online air quality data**: There is online data on air quality available from: http://qualar.apambiente.pt/index.php?page=2&zona=&year=2016&month=6&day=03

## 3.9 Health module

The health module assesses the health impacts based upon the air quality data from the Air quality module.

### *3.9.1 Generic*

Gridded annual concentration levels of $PM_{2.5}$, $PM_{10}$, and $NO_2$ from the Air Quality module are used, together with population statistics, to calculate mortality and morbidity indicators related to air pollution. Mortality endpoints include the number of premature deaths and years of life lost (YLL), whilst the morbidity endpoint used is Restricted Activity Days (RAD). According to de Leeuw and Horalek[26] premature deaths are deaths that occur before a person reaches an expected age (typically the age of standard life expectancy for a country and gender), whilst YLL are the years of potential life lost due to premature deaths. Since YLL take into account the age at which deaths occur, a greater weight is given to deaths at a younger age than at an older age. A RAD is defined as a day when individuals reduce their normal activities, including days of missed work, absences from school and other more minor reductions in daily activity. It is thus an indirect measure of all morbidity, and can be converted to socio-economic costs.

Pollution concentration response functions (CRFs) are assumed to be linear in line with WHO recommendations[27,28] and are based on long-term exposure (annual average). Relative risk ratios indicate the slope of the CRF per 10 µg/m$^3$. Relative risks for mortality and morbidity (only $PM_{2.5}$) are given in Table 5.

**Table 5 - Exposure-response coefficients for mortality and morbidity used in this study**

| | Pollutant | Exposure-response coefficient and/or relative risk (RR) for mortality | | |
|---|---|---|---|---|
| | Value [per 10 µg/m$^3$] | Type | Reference |
| $PM_{2.5}$ | 6.2 % Risk ratio 1.062 (95 % CI 1.040-1.083). No threshold | All-cause (natural) mortality in ages above 30 (ICD-10 codes A00-R99). | WHO 2013. Health risks of air pollution in Europe - HRAPIE project. Recommendations for concentration-response functions for cost-benefit analysis of particulate matter, ozone and nitrogen dioxide. Copenhagen, Denmark. |
| $PM_{2.5}$ | 4.7 % Risk ratio 1.047 (95% CI 1.042-1.053) No threshold | Morbidity - Reduced activity days (all ages) | WHO 2013. Health risks of air pollution in Europe - HRAPIE project. Recommendations for concentration-response functions for cost-benefit analysis of particulate matter, ozone and nitrogen dioxide. Copenhagen, Denmark. |
| $PM_{10}$ | 4% Risk ratio 1.04 (95% CI, 1-1.09) | All-cause (natural) mortality in ages above | Beelen et al. 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European |

[26] De Leeuw, F. & Horalek, J. 2016. Quantifying the health impacts of ambient air pollution: methodology and input data. European Topic Centre on Air Pollution and Climate Change Mitigation.
[27] WHO 2013. Health risks of air pollution in Europe - HRAPIE project. Recommendations for concentration-response functions for cost-benefit analysis of particulate matter, ozone and nitrogen dioxide. Copenhagen, Denmark.
[28] WHO 2013. Review of evidence on health aspects of air pollution - REVIHAAP Project Technical Report. Copenhagen: World Health Organization.

| | | | |
|---|---|---|---|
| | No threshold | 30 (ICD-10 codes A00-R99). | cohorts within the multicentre ESCAPE project. The Lancet 383 pp 785-795 |
| $NO_2$ | 5.5 % Risk ratio 1.055 (95 % CI 1.031-1.08%). Threshold 10 µg/$m^3$ | All-cause (natural) mortality in ages above 30 (ICD-10 codes A00-R99). | WHO 2013. Health risks of air pollution in Europe - HRAPIE project. Recommendations for concentration-response functions for cost-benefit analysis of particulate matter, ozone and nitrogen dioxide. Copenhagen, Denmark. |

The **Population Attributable Fraction (PAF)**, defined as the proportional reduction in population disease or mortality that would occur if exposure to a risk factor were reduced to zero, can then be calculated from the linear relative risk CRF, according to:

$$PAF = (RR - 1)/RR$$

Where the population is considered with only one concentration level.

Subsequently, premature deaths and years of life lost can be estimated at grid cell level by multiplying the PAF with crude death rates, broken down by age cohort and sex, and total population within the grid cell.

$$PremDeath = \sum_i CDR_i * PAF * Pop$$

Crude death rates for 5-year age cohorts (all ages above 30), and distinguishing for male and female were calculated from natural all cause mortality in 2015 (ICD codes A00-R99) and total population, both at country level.

Population data were obtained via the population pyramids for 2015 are obtained from the UN.
https://population.un.org/wpp/Download/Standard/Population/
Tables:

- POP/DB/WPP/Rev.2017/POP/F15-3
- POP/DB/WPP/Rev.2017/POP/F15-2

Natural all cause mortality data were obtained from the WHO: http://apps.who.int/healthinfo/statistics/mortality/causeofdeath_query/start.php

Year**s** of **Life L**ost **(YLL)** are calculated at grid cell level by multiplying premature deaths with life expectancy by age cohort and sex.

$$YLL = \sum_i LE_i * PremDeath_i$$

Life expectancy data per age cohort and sex is calculated averaging life expectancy at exact age and the average age at death for separate age cohorts.

Life expectancy at exact age is obtained from the UN.
https://population.un.org/wpp/Download/Standard/Mortality/
Tables:

- POP/DB/WPP/Rev.2017/MORT/F16-2

- POP/DB/WPP/Rev.2017/MORT/F16-3

Data on average age at death is obtained from Annex 3.1 in the following document. https://www.who.int/quantifying_ehimpacts/publications/en/9241546204chap3.pdf

**Morbidity (in reduced activity days)** is calculated by multiplying the PAF for morbidity with the baseline reduced activity days per age cohort and sex.

$$RAD = \sum_i PAF * RAD_i * Pop$$

Using the population pyramid data and a baseline reduced activity of 19 RAD[29] per person per year $RAD_i$ can be calculated by age cohort and sex.

The expected burden of disease attributable to air pollution in one specific area (e.g. Bristol) can finally be found by summing over all grid cells in the area of interest for the indicator of interest (e.g. YLL).

The expected burden of disease attributable to air pollution in one specific area (e.g. Bristol) can finally be found by summing over all grid cells in the area of interest for the indicator of interest (e.g. YLL).

## 3.10 Carbon footprint module

The Carbon footprint module calculates the carbon footprint from the emissions collected in the Emission database module.

### *3.10.1 Clair city approach to carbon footprint*

The carbon footprint evaluation was undertaken following an emission inventory approach similar to the approach utilised by the Covenant of Mayors[30] and using both:

- "Standard" emission factors (emissions were evaluated using methodologies and emission factors from 2006 IPCC Guidelines for National Greenhouse Gas Inventories[31] and specific activity level); the methodology cover all the $CO_2$ emissions that occur due to energy consumption within the territory of the city/region, either directly due to fuel combustion within the city/region or indirectly via fuel combustion associated with electricity and heat/cold usage within their area;
    - the standard emission factors are based on the carbon content of each fuel, like in national greenhouse gas inventories in the context of the UNFCCC and

[29] WHO 2013. Health risks of air pollution in Europe - HRAPIE project. Recommendations for concentration-response functions for cost-benefit analysis of particulate matter, ozone and nitrogen dioxide. Copenhagen, Denmark.
[30] Covenant of Mayors (2010), How to develop a Sustainable Energy Action Plan (SEAP) – Guidebook Part II, Baseline emissions inventory
[31] 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 2 Energy

the Kyoto protocol; in the standard approach, the $CO_2$ emissions from the sustainable use of biomass/biofuels, as well as emissions of certified green electricity, are considered to be zero; emission are reported as
    - $CO_2$ only emissions, the most important greenhouse gas,
    - $CO_2$ equivalent emissions, including calculation of the emissions of $CH_4$ and $N_2O$ with emission factor from 2006 IPCC Guidelines for National Greenhouse Gas Inventories[31] and will be included calculated using the Global Warming Potential (GWP) with 100 year time horizon[32]:

$$1 \text{ Mg } CO_2 = 1 \text{ Mg } CO_2\text{-eq}$$
$$1 \text{ t } CH_4 = 21 \text{ Mg } CO_2\text{-eq}$$
$$1 \text{ t } N_2O = 310 \text{ Mg } CO_2\text{-eq}$$

    - In our approach, $CO_2$ is the most important greenhouse gas, and the emissions of $CH_4$ and $N_2O$ will be not calculated; furthermore, the $CO_2$ emissions from the sustainable use of biomass/biofuels, as well as emissions of certified green electricity, are considered to be zero;
- LCA (Life Cycle Assessment) emission factors, which take into consideration the overall life cycle of the energy carrier; includes not only the emissions of the final combustion, but also all emissions of the supply chain; it includes emissions from exploitation, transport and processing (e.g. refinery) steps in addition to the final combustion; this hence includes also emissions that take place outside the location where the fuel is used. In this approach, the GHG emissions from the use of biomass/biofuels, as well as emissions of certified green electricity, are higher than zero. In the case of this approach, other greenhouse gases than $CO_2$ may play an important role, therefore the LCA approach will report emissions as $CO_2$ equivalent; as a default will be used the LCA emission factors given in Covenant of Mayors guidelines, based on JRC European Reference Life Cycle Database; specific national emission factors will be investigated.

### *3.10.2 Carbon footprint modeling*

The Carbon Footprint tool:

- evaluates emissions at most detailed administrative territorial units level;
- uses emission factors from Covenant of Mayors guidelines;
- calculates emissions as:

$$E_k = A_{ij} F_{ik}$$

where:

- $A_{ij}$ is the indicator of the activity i in the territorial unit j
- $F_{ik}$ is the emission factor for different Carbon footprint indicators k for activity i (expressed in grams per unit of activity);
- k is the carbon footprint indicator used: $CO_2$, $CO_{2eq}$, $CO_{2eq,LCA}$

[32] IPCC, 1995. Contribution of Working Group I to the Second Assessment of the Intergovernmental Panel on Climate Change

### *3.10.3 Emission factors*

In Table 6 the $CO_2$ fuel emission factors used are reported[33].

**Table 6 – $CO_2$ Emission Factors**

| Fuel | Standard Emission Factors[34] [Mg $CO_2$/MWh] | Standard Emission Factors[34] [Mg $CO_2$eq/MWh] | LCA Emission Factors[35] [Mg $CO_2$-eq/MWh] |
|---|---|---|---|
| Motor Gasoline | 0.249 | 0.250 | 0.299 |
| Gas oil, diesel | 0.267 | 0.268 | 0.305 |
| Residual Fuel Oil | 0.279 | 0.279 | 0.310 |
| Anthracite | 0.354 | 0.356 | 0.393 |
| Other Bituminous Coal | 0.341 | 0.342 | 0.380 |
| Sub-Bituminous Coal | 0.346 | 0.348 | 0.385 |
| Lignite | 0.364 | 0.365 | 0.375 |
| Natural Gas | 0.202 | 0.202 | 0.237 |
| Municipal Wastes (non-biomass fraction) | 0.330 | 0.337 | 0.330 |
| Municipal Wastes (biomass fraction) | 0 | 0.007 | 0.106 |
| Wood[1] | 0 – 0.403 | 0.007 – 0.410 | 0.017 – 0.416 |
| Plant oil[1] | 0 – 0.287 | 0.001 – 0.302 | 0.182 – 0.484 |
| Biodiesel[1] | 0 – 0.255 | 0.001 – 0.256 | 0.156 – 0.411 |

[1] Lower value if fuel meets carbon neutrality criteria, higher otherwise

In Table 7 electricity consumptions national emission factor are reported to use for $CO_2$ indirect emissions from electricity consumptions.

**Table 7 – $CO_2$ National Electricity Emission Factors**

| Fuel | Standard Emission Factors [Mg $CO_2$/MWh] | Standard Emission Factors [Mg $CO_2$eq/MWh] | LCA Emission Factors [Mg $CO_2$-eq/MWh] |
|---|---|---|---|
| Bristol (UK)[36] | 0.4585 | 0.46219 | 0.5224 |
| Amsterdam (NL)[37] | 0.53 | 0.53 | 0.60 |

[33] JRC (2017), Covenant of Mayors for Climate and Energy: Default emission factors for local emission inventories
[34] 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 2 Energy
[35] ELCD (2015), European Reference Life Cycle Database (ELCD) Release 3.2. LCA data sets of key energy carriers, materials, waste and transport services of European scope
[36] UK Government conversion factors for Company Reporting (LCA EF derived using the same ratio as JRC)
[37] CBS, Rendementen en CO2-emissie elektriciteitsproductie 2015 (LCA EF derived using the same ratio as JRC)

### 3.10.4 Data evaluation at domains and subdomains

Carbon footprint is evaluated at city domains. When data are available only at national level, it is allocated to subdomains using a “proxy” variable available at city levels.

When data are available at national level, data at subdomain level is evaluated using the following equation:

$$A_i = A * P_i / \Sigma_i P_i$$

where: $A_i$ and $P_i$ are the values of variable A and proxy variable P in the city i and A is the national value of proxy variable P.

### 3.10.5 Data sources

#### 3.10.5.1 Bristol

The following tables in the appendix document the methodology and data used for:

- Industrial sources (Table 21);
- Residential and commercial sources (Table 22);
- Aggregate fuel consumtions data subdivision (Table 23); and
- *LSOA* disaggregation variables (Table 24)

#### 3.10.5.2 Amsterdam

The following tables in the appendix document the methodology and data used for:

- Industrial sources (Table 25);
- Residential and commercial sources (Table 26); and
- *Buurt* disaggregation variables (Table 27).

#### 3.10.5.3 Ljubljana

The following tables document the methodology and data used for:

- Industrial sources (Table 28);
- Residential and commercial sources (Table 29); and
- *Naselje* disaggregation variables (Table 30).

## 3.11 Identifying good practices

Based on the review of data sources being used in the ClairCity approach as described in the previous sections, we identify the following good practices:

- In the data modeling and data selection sections (Section 2.1 and Section 2.2) different approaches are described to identify primary data sources or secundary alternatives.

- Data collection may be easier when partners know what to look for. Checklists of preferred data are useful for this purpose.

- Simplify as much as possible. It is better to build your model on a few good data sources than to try to be complete and take all data of all aspects into consideration.

- One of the cornerstones of the ClairCity model is detailed data on the households and population in a city / region. These data can usually be retrieved from the national statistical office, either in the regular population projections or otherwise from the census data. If data offered by the national statistics office is not detailed enough, one might consult the demographic department of the city council, however we did not find such a need in the current ClairCity project to date. Due to the crucial role that data from National Statistical Institutes (NSI's) play in the modeling, we herewith provide a link to a page from the European Statistical Institute (Eurostat) which contains directions to all European statistical institutes https://ec.europa.eu/eurostat/web/links.

- Another important data source for ClairCity is data on time-use. The Danish time use survey (TUS) can be used for any of the cities where more specific data is not available for that region. In addition, national statistics offices might have time use data available based on the Harmonised European Time Use Surveys (HETUS), see https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-08-014. UK, Poland, Italy and Slovenia, as well as 11 other European countries have contributed data.

- Satellite data, especially from Copernicus - Urban Atlas, is another source of information that is used in many of the ClairCity modules. It is worth to be mentioned here as one of the generic data sources, usable for any city that applies the ClairCity model: https://www.copernicus.eu/.

- For transport many generic fleet data is already present on the European level. The EU standard vehicle emissions calculator is a valuable source to be used in this module. When combined with very specific regional data on transport, detailed emissions data can be calculated, as has been shown in Bristol and Amsterdam. In addition local open data portals might offer additional data. For example, in Ljubljana car characteristics can be accessed from the open data portal of the Republic of Slovenia.

- National statistics on fleet data are typically sufficient for modeling. We found that for Bristol, the city-specific fleet does not differ that much from the national fleet; furthermore the city-specific fleet (ownership) is not per se a good reflection of the actual active fleet in the city. National statistics suffice and are publically available.

- Openstreetmap can be a valuable source to create a network for transport modeling.

- Full travel survey data are available for Bristol, but not for others. However, limited survey data or national data are valuable as a proxy.

- Traffic counts data may be available for larger cities, but not for smaller cities. However using proxy data, a good estimate may still be made, using for instance the OSM-TRANSPHORM mesh.

- Another important data source is the European institute where national institutes report to the E-PRTR. Data can be found on their website: https://prtr.eea.europa.eu.

- For visualizing practice-activity data, for example in dotmaps, one needs detailed geo information on residential areas in cities. A good practice is to look for such data in a national or local administration or to derive it from a community source such as openstreetmap or another public map or a satellite image.

- Meteorological data is an additional data source in some of the ClairCity modules. Obviously these can be obtained from national meteorological institutes (listed at http://eumetnet.eu/members-partners/). For city modeling however, it is crucial to find a nearby measurement station for applicable data. In Amsterdam it was taken from measurement stations of the nearby airport. This is only applicable to cities that indeed have a nearby airport. If not, one might consider using citizen science data, such as for example from https://opensensemap.org.

- In the era of big data the possibilities of identifying, quantifying and visualising behaviour of citizens are vast. While in the 20$^{th}$ century behaviour data collection often included diligent hour by hour labour of engaged citizens, currently diverse tools are available to collect similar data automatically from minute to minute. As an example, citizens become active measuring air quality themselves. Alternative data sources, as mentioned in Section 2.3 are useful to take into account for local air quality modeling. This hold especially for cross-border citizen science projects such as the German project with self-assembled laser $PM_{2.5}$/$PM_{10}$-measuring devices, extended to devices in most other European countries: https://luftdaten.info.

- A good practice - not for collecting but for communicating air quality aspects - can be found in the indicator developed by the city of Amsterdam: the number of passive

smoked cigarettes. This method expresses the air quality in practical terms, recognizable for citizens[38].

- A good practice for collecting behaviour data on transport is to draw on existing public opinion data for opinions on commuting/transport. Examples are Bristol satisfaction by mode of transport (https://travelwest.info/businesses/travel-work-survey/2017-survey-results), Amsterdam satisfaction with public transport (https://en.gvb.nl/nieuws/ov-klantenbarometer-2016-reizigerstevredenheid-stabiel-gebleven), Ljubljana (in Slovenian, https://www.ljubljana.si/sl/aktualno/celostna-prometna-strategija-mol-3/). A similar approach could hold for other behaviour / targeted citizen engagement data, for example poorer people who drive, families who cycle, people using solid fuel, people using district heating.

[38] van der Zee, Saskia & Fischer, Paul & Hoek, Gerard. (2016). Air pollution in perspective: Health risks of air pollution expressed in equivalent numbers of passively smoked cigarettes. Environmental Research. 148. 475-483. 10.1016/j.envres.2016.04.001.

# 4 Conclusion

In this deliverable we reviewed the ClairCity data journey from the perspective of re-use. The aim was to deliver a general view on how the ClairCity approach could be applied to other cities / regions, and especially what good practices from the existing project could be re-used and applied in a new context.

First the concept of practice-activity data was explored. This type of data is the foundation of the ClairCity model and one of the first questions to be asked when applying the model in another context is the availability of such data. Detailed data – it could be synthetic but has to be representative enough - on households, energy use, transportation, peoples activities and industry are necessary to build scenarios for clean air cities. Data can come from many sources, European, national, local, official, citizen-generated, from databases, satellites, sensors or big data, the key thing is that it is a valuable input to any of the modules from the basic ClairCity modeling framework (see Figure 3).

Then, the main types of data being used in each of the ClairCity modules were identified. Using the central ClairCity data portal (http://claircitydata.cbs.nl) as a guide, a long list of data sources was put together categorised into generic data sources and those specific to a region. This led to a list of good practices including (deep)links to data sources for re-use. Finally some good practices from other data intense projects were mentioned.

We hope that this deliverable is a valuable source of information for other projects working towards cleaner air or other environmental improvements. Although every project has its specific needs from a data and modeling perspective, there is one commonality: no progress in research without access to good quality data.

# 5 Appendix

**Table 8 – Methodology and source of data for Bristol emissions evaluation - Industrial sources**

| Activity | Data availability | Source | Publication | Reference |
|---|---|---|---|---|
| Industrial sector | Single facility | UK Department for Environment Food & Rural Affairs | Emissions from NAEI large point sources | http://naei.beis.gov.uk/data/map-large-source |

**Table 9 – Methodology and source of data for Bristol fuel consumptions evaluation - Residential and services sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| Residential sector | Natural Gas | Level 3 (LSOA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas gas consumption 2018 update | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676340/LSOA_domestic_gas_2015.xlsx | Consumption (kWh) | None |
| | Wood | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 20% of Column M (Bioenergy & Waste) [see Share of wood on biomass in Table 10 for technology spit] | households not connected to the gas network (Table 12) |
| | LPG | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 10% of Column D (Petroleum; domestic) [Table 11 for percentage] | households not connected to the gas network (Table 12) |
| | Gasoil | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 90% of Column D (Petroleum; domestic) [Table 11 for percentage] | households not connected to the gas network (Table 12) |
| | Coal | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | Columns J+L (Coal+ Manufactured Solid Fuels; domestic) [Table 11 for percentage] | households not connected to the gas network (Table 12) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Service sector | Natural gas | Level 2 (MSOA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas gas consumption 2018 update | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676344/MSOA_non_dom_gas_2015.xlsx | Consumption (kWh) 42% [Table 11 for percentage]; totals at LA level obtained as sum from MSOA data are directly allocated to LSOA (°) | Employees (Table 12) |
| | LPG | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/statistical-data-sets/estimates-of-non-gas-non-electricity-and-non-road-transport-fuels-at-regional-and-local-authority-level | 30% of Column F+G (Petroleum; Public Administration + Commercial) [Table 11 for percentage] | Employees (Table 12) |
| | Gasoil | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 70% of Column F+G (Petroleum; Public Administration +Commercial) [Table 11 for percentage] | Employees (Table 12) |
| | Wood | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | Negligible share of Column M (Bioenergy & Waste) [see Share of wood on biomass in Table 10 for percentage] | Value=0 no disaggregation |
| | Coal | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | Columns I (Coal; Industrial & Commercial) [Commercial share 1,5%; Table 11 for percentage] | Employees (Table 12) |

(°) if MSOA data are used to evaluate LSOA a bias is introduced due to different distribution industry/services in different MSOA

**Table 10 – Methodology and source of data for Bristol fuel consumptions evaluation – Wood statistics**

| Variable | Data availability | Sources | Publication | Reference | Note |
|---|---|---|---|---|---|
| Share of wood on biomass | Level 1,5 (LA) | Ricardo Energy & Environment | Personal comunication | | The following share is evaluated: wood domestic 20%. In commercial sector only, wood wastes and plant biomass |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | assumed included in point sources and globally negligible |
| Technologies split | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Summary results of the domestic wood use survey<br>Table 2.7 Final energy calculation | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/576953/Summary_Tables_Domestic_Wood_Survey.xlsx | On the basis of available data, the following shares are evaluated: conventional stoves 16%, high efficiency stoves 17%, advanced stoves 13%, conventional fireplaces 43%, high efficiency fireplaces 4%, advanced fireplaces 4%, boilers 4%. |

**Table 11 – Methodology and source of data for Bristol fuel consumptions evaluation – Aggregate fuel consumptions data subdivision**

| Energy vector | Data availability | Source | Publication | Reference | Note |
|---|---|---|---|---|---|
| Natural Gas | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Digest of UK Energy Statistics: natural gas: commodity balances (DUKES 4.1) | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632524/DUKES_4.1.xls | On the basis of available data, the following share is evaluated:<br>SERVICES Natural gas 42%, industrial 52%, others (agriculture, miscellaneous) 6% |
| LPG, Gasoil | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Digest of UK Energy Statistics: Petroleum products: commodity balances (DUKES 3.2-3.4) | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632507/DUKES_3.2-3.4.xls | On the basis of available data, the following shares are evaluated:<br>SERVICES LPG 30% Gasoil 70%<br>(in gasoil we include also an 8% of fuel oil)<br>RESIDENTIAL LPG 10% Gasoil 90%<br>(in gasoil kerosene is included). |
| Coal | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Digest of UK Energy Statistics: : solid fuels and derived gases: commodity balances (DUKES ) | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632497/DUKES_2.4.xls | On the basis of available data, the following shares are evaluated for Coal:<br>SERVICES 1,5% INDUSTRIAL 98,5% |

**Table 12 – Methodology and source of data for Bristol fuel consumptions evaluation – LSOA disaggregation variables**

| Variable | Data availability | Sources | Publication | Reference | Fields |
|---|---|---|---|---|---|
| Households not connected to the gas network | Level 3 (LSOA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas gas consumption 2018 update | https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/712424/LSOA_domestic_nongas_2016.xlsx | Estimated number of households not connected to the gas network |
| Employees | Level 3 (LSOA) | UK Office for National Statistics | All people aged 16 to 74 in employment the week before the Census<br>Occupation by industry<br>2011 Occupies | https://www.nomisweb.co.uk/census/2011/ks605uk | Geography<br>All of the following:<br>2001 super output areas - lower layer<br>Cell<br>SERVICE SECTOR, Table CAS039, Occupation by industry<br>select columns<br>• G Wholesale and retail trade; repair of motor vehicles and motor cycles<br>• H Transport and storage<br>• I Accommodation and food service activities<br>• J Information and communication<br>• K Financial and insurance activities<br>• L Real estate activities<br>• M Professional, scientific and technical activities<br>• N Administrative and support service activities<br>• O Public administration and defense; compulsory social security<br>• P Education<br>• Q Human health and social work activities<br>• R, S, T, U Other<br>INDUSTRIAL SECTOR<br>select columns<br>• B Mining and quarrying |

**Table 12 – Methodology and source of data for Bristol fuel consumptions evaluation – LSOA disaggregation variables**

| Variable | Data availability | Sources | Publication | Reference | Fields |
|---|---|---|---|---|---|
| | | | | | • C Manufacturing<br>• D Electricity, gas, steam and air conditioning supply<br>• E Water supply; sewerage, waste management and remediation activities<br>• F Construction<br>The table provides information that classifies usual residents aged 16 to 74 in employment the week before the census by the industry in which they work, for United Kingdom as at census day, 27 March 2011. |
| properties with heating fuel oil | LSOA | The non-gas map | Kiln for Affordable Warmth Solutions, in conjunction with the Department for Business, Energy and Industrial Strategy | https://www.nongasmap.org.uk | Heating_fuel_oil |
| properties with heating fuel oil | LSOA | The non-gas map | Kiln for Affordable Warmth Solutions, in conjunction with the Department for Business, Energy and Industrial Strategy | https://www.nongasmap.org.uk | Heating_fuel_oil |

**Table 13 – Methodology and source of data for Amsterdam emissions evaluation - Industrial sources**

| Activity | Data availability | Source | Publication | Reference |
|---|---|---|---|---|
| Industrial sector | Single facility | Emissieregistratie | | http://emissieregistratie.nl |

**Table 14 – Methodology and source of data for Amsterdam fuel consumptions evaluation - Residential and commercial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| Residential sector | Natural Gas | Level 3 (Buurt) | CBS | Wijk - en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | G_GAS_TOT*WONINGEN where: [G_GAS_TOT]: Average total natural gas consumption [WONINGEN]: Housing stock | None |
| | Wood | Level 2 (Gemeente) | RIVM | Klimaatmonitor | https://klimaatmonitor.databank.nl/dashboard/ | Wood burning stoves dwellings hern. heat [TJ] (see Share of wood on biomass in Table 15 for technology spit) | Population (Table 16) |
| | LPG | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: households Period: 2015 Energy commodities: LPG | Population (Table 16) |
| | Gasoil | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: households Period: 2015 Energy commodities: Heating and other gas oil | Population (Table 16) |
| Service sector | Natural gas | Level 2 (Gemeente) | RIVM | Klimaatmonitor | https://klimaatmonitor.databank.nl/dashboard/ | Gas use commercial Services [m3] + Gas use Public Services [m3] | Services Companies number (Table 16) |
| | Wood | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: services waste and repairs Period: 2015 Energy commodities: Solid and liquid biomass | Services Companies number (Table 16) |
| | LPG | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: services waste and repairs Period: 2015 Energy commodities: LPG | Services Companies number (Table 16) |
| | Gasoil | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: services waste and repairs Period: 2015 Energy commodities: Heating and other gas oil | Services Companies number (Table 16) |

(*) For non-domestic MSOA data, industrial half hourly (HH) consumption is provided for each local authority but it is not disaggregated further as doing so would break the UK Statistics Authority's Code of practice for Official Statistics relating to data disclosure. These LA data has been allocated to MSO using not-half hours MSOA data.

**Table 15 – Methodology and source of data for Amsterdam fuel consumptions evaluation – Wood statistics**

| Variable | Data availability | Sources | Publication | Reference | Note |
|---|---|---|---|---|---|
| Technologies split | Level 3 (National) | CBS | Houtverbruik bij huishoudens (Wood consumption in households) | https://www.cbs.nl/-/media/imported/documents/2010/18/2010-houtverbruik-bij-huishoudens-art.pdf?la=nl-nl | On the basis of available data, the following shares are evaluated: stoves 55% fireplaces 45%. Using national EMEP $PM_{10}$ data the following shares are derived: traditional 30% advanced 70%.)<br>Service sector allocated to boilers. |

**Table 16 – Methodology and source of data for Amsterdam fuel consumptions evaluation – Buurt disaggregation variables**

| Variable | Data availability | Sources | Publication | Reference | Fields |
|---|---|---|---|---|---|
| Population | Level 3 (Buurt) | CBS | Wijk-en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | AANT_INW (Number of inhabitants) |
| Services Companies number | Level 3 (Buurt) | CBS | Wijk-en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | [A_BED_GI]+[A_BED_HJ]+[A_BED_KL]+[A_BED_MN]+[A_BED_RU]<br>where:<br>[A_BED_GI]: Number of companies and catering trade<br>[A_BED_HJ]: Number of companies in transport, information, communication<br>[A_BED_KL]: Number of firms financially property<br>[A_BED_MN]: Number of companies in business services<br>[A_BED_RU]: Number of companies in culture, recreation, other |
| Industry Companies number | Level 3 (Buurt) | CBS | Wijk-en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | [A_BED_BF]: Number of companies in industry and energy |

**Table 17 – Methodology and source of data for Ljubljana fuel consumptions/emissions evaluation - Industrial sources**

| Activity | Data availability | Source | Publication | Reference | Disaggregation variable |
|---|---|---|---|---|---|
| Industrial sector | Single facility | EIONET | Reporting Obligations Database (ROD), Deliveries for National Emission Ceiling Directive (NECD) - Large point source (LPS) emissions data by source category (GNFR) Slovenia NECD 2017 Report LPS emissions 2007 2015 | http://cdr.eionet.europa.eu/si/eu/nec_revised/lps/envwox5ng | None (Point sources) |
| Industrial sector | Level 2 (Občine) only for Ljubljana | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Corine Land Cover for industrial plants and direct allocation for not point sources energy transformation plants |

**Table 18 – Methodology and source of data for Ljubljana fuel consumptions evaluation - Residential and commercial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Disaggregation variable |
|---|---|---|---|---|---|---|
| Residential sector | Natural Gas | Level 2 (Občine) only for Ljubljana | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |
| | Wood | Level 2 (Občine) only for Ljubljana | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |

**Table 18 – Methodology and source of data for Ljubljana fuel consumptions evaluation - Residential and commercial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Disaggregation variable |
|---|---|---|---|---|---|---|
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |
| | LPG | Level 2 (Občine) only for Ljubljana | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |
| | Gasoil | Level 2 (Občine) only for Ljubljana | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |
| | Coal | Level 2 (Občine) only for Ljubljana | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |
| Service sector | Natural Gas | Level 2 (Občine) | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |

**Table 18 – Methodology and source of data for Ljubljana fuel consumptions evaluation - Residential and commercial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Disaggregation variable |
|---|---|---|---|---|---|---|
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |
| | Wood | Level 2 (Občine) | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |
| | LPG | Level 2 (Občine) | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |
| | Gasoil | Level 2 (Občine) | EnerGisSolution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | Population |
| | | Level 1 National for all other areas | Republika Slovenija, Ministrstvo za Infrastrukturo | Portal Energetike: Energetska Bilanca Republik Eslovenije - Zaleto 2015 | http://www.energetika-portal.si/fileadmin/dokumenti/publikacije/energetska_bilanca/ebrs_2015.pdf | Population |

**Table 19 – Methodology and source of data for Ljubljana fuel consumptions evaluation – Wood statistics**

| Variable | Data availability | Sources | Publication | Reference | Note |
|---|---|---|---|---|---|
| Technologies split | Level 1 (National) | Slovenian Environment Agency | Slovenia's Informative Inventory Report 2017. Submission under the UNECE Convention on Long-Range Transboundary Air Pollution and Directive (EU) 2016/2284 on the reduction of national emissions of certain atmospheric pollutants, Ljubljana, March 2017 | http://cdr.eionet.europa.eu/si/un/clrtap/iir/envwmaww/Slovenia_IIR__2017.pdf | In the year 2015 there were 67 % conventional boilers burning wood and similar wood waste, 12 % advanced / ecolabelled stoves and boilers burning wood, 5 % pellet stoves and boilers burning wood pellets, 1 % open fireplaces burning wood, 15 % conventional stoves burning wood and similar wood waste |

**Table 20 – Methodology and source of data for Ljubljana fuel consumptions evaluation – Naselje disaggregation variables**

| Variable | Data availability | Sources | Publication | Reference | Fields |
|---|---|---|---|---|---|
| Population | Level 3 (Naselje) | Statistical Office of the Republic of Slovenia | SI-Stat Database | https://pxweb.stat.si/pxweb/Dialog/varval.asp?ma=05C5004E&ti=&path=../Database/Demographics/05_population/10_Number_Population/25_05C50_Population_naselja/&lang=1 | Population by large and 5-year age groups and sex, settlements, Slovenia, annually |
| Industrial areas coverage | Level 3 (Naselje) | Copernicus Land Monitoring Service | CORINE Land Cover | https://land.copernicus.eu/pan-european/corine-land-cover | A GIS query has been used to evaluated the coverage of industrial area on each Naselje and industrial emissions are allocated to the area based on dimension of area itself. |

**Table 21 – Methodology and source of data for Bristol fuel consumptions evaluation - Industrial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| Industrial sector | Natural Gas | Level 2 (MSOA) | UK Department for Business, Energy & | Lower and Middle Super Output Areas gas consumption 2018 update | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676344/MSOA_non_dom_gas_2015.xlsx | 52% of non-domestic Consumption (kWh) [see – for percentage]; totals at LA level obtained as sum from | Employees (Table 24 –) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | Industrial Strategy | | | MSOA data are directly allocated to LSOA (°) | |
| | LPG | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 11% of Column D (Petroleum; Industrial) [see – for percentage] | Employees (Table 24 –) |
| | Gasoil | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 89% of Column D (Petroleum; Industrial) [see – for percentage] | Employees (Table 24 –) |
| | Coal | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 98,5% of Columns I+K (Coal+ Manufactured Solid Fuels; Industrial+Commercial) [see – for percentage] | Employees (Table 24 –) |
| | Electricity | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas electricity consumption - MSOA non domestic electricity 2015 | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676475/MSOA_non_dom_electricity_2015.xlsx | Consumption (kWh) 50% [see – for percentage]; totals at LA level obtained as sum from MSOA data are directly allocated to LSOA (°) | Employees (Table 24 –) |

(°) if MSOA data are used to evaluate LSOA a bias is introduced due to different distribution industry/services in different MSOA

**Table 22 – Methodology and source of data for Bristol fuel consumptions evaluation - Residential and services sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| Residential sector | Natural gas | Level 3 (LSOA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas gas consumption 2018 update | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676340/LSOA_domestic_gas_2015.xlsx | Consumption (kWh) 42% [see – for percentage] | None |
| | Wood | Level 1,5 (LA) | UK Department for Business, | Residual fuel consumption at | https://www.gov.uk/government/uploads/system/uploads/attachment_ | 56% of Column M (Bioenergy & Waste) [see Share of wood | households not connected to the |

**Table 22 – Methodology and source of data for Bristol fuel consumptions evaluation - Residential and services sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| | | | Energy & Industrial Strategy | regional and local authority level | data/file/647698/residual_fuels_2005-2015.xlsx | on biomass in – for percentage] | gas network (Table 24 –) |
| | LPG | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 10% of Column D (Petroleum; domestic) [see – for percentage] | households not connected to the gas network (Table 24 –) |
| | Gasoil | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 90% of Column D (Petroleum; domestic) [see – for percentage] | households not connected to the gas network (Table 24 –) |
| | Coal | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | Columns J+L (Coal+ Manufactured Solid Fuels; domestic) [see – for percentage] | households not connected to the gas network (Table 24 –) |
| | Electricity | Level 3 (LSOA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas electricity consumption - MSOA non domestic electricity 2015 | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676473/LSOA_domestic_electricity_2015.xlsx | Consumption (kWh) | None |
| Service sector | Natural gas | Level 2 (MSOA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas gas consumption 2018 update | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676344/MSOA_non_dom_gas_2015.xlsx | Consumption (kWh) 42% [see – for percentage]; totals at LA level obtained as sum from MSOA data are directly allocated to LSOA (°) | Employees (Table 24 –) |
| | LPG | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/statistical-data-sets/estimates-of-non-gas-non-electricity-and-non-road-transport-fuels-at-regional-and-local-authority-level | 30% of Column F+G (Petroleum; Public Administration + Commercial) [see – for percentage] | Employees (Table 24 –) |

**Table 22 – Methodology and source of data for Bristol fuel consumptions evaluation - Residential and services sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| | Gasoil | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | 70% of Column F+G (Petroleum; Public Administration+Commercial) [see – for percentage] | Employees (Table 24 –) |
| | Wood | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | Negligible share of Column M (Bioenergy & Waste) [see – for percentage) | Value=0 no disaggregation |
| | Coal | Level 1,5 (LA) | UK Department for Business, Energy & Industrial Strategy | Residual fuel consumption at regional and local authority level | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/647698/residual_fuels_2005-2015.xlsx | Columns I (Coal; Industrial & Commercial) [Commercial share 1,5%; see – for percentage] | Employees (Table 24 –) |
| | Electricity | Level 2 (MSOA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas electricity consumption - MSOA non domestic electricity 2015 | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676475/MSOA_non_dom_electricity_2015.xlsx | Consumption (kWh) 50% [see – for percentage]; totals at LA level obtained as sum from MSOA data are directly allocated to LSOA (°) | Employees (Table 24 –) |

(°) if MSOA data are used to evaluate LSOA a bias is introduced due to different distribution industry/services in different MSOA

**Table 23 – Methodology and source of data for Bristol fuel consumptions evaluation – aggregate fuel consumtions data subdivision**

| Energy vector | Data availability | Source | Publication | Reference | Note |
|---|---|---|---|---|---|
| Wood | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Digest of UK Energy Statistics (DUKES): renewable | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632546/DUKES_6.1-6.3.xls | On the basis of available data the following share is evaluated: wood domestic 56%. In commercial sector only wood wastes and plant biomass assumed; consumtions valutated as included in point sources and globally negligible |

| | | | | | |
|---|---|---|---|---|---|
| | | | sources of energy | | |
| Natural Gas | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Digest of UK Energy Statistics: natural gas: commodity balances (DUKES 4.1) | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632524/DUKES_4.1.xls | On the basis of available data the following share is evaluated:<br>SERVICES Natural gas 42%, industrial 52%, others (agriculture, miscellaneous) 6% |
| LPG, Gasoil | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Digest of UK Energy Statistics: Petroleum products: commodity balances (DUKES 3.2-3.4) | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632507/DUKES_3.2-3.4.xls | On the basis of available data the following shares are evaluated:<br>SERVICES LPG 30% Gasoil 70%<br>(in gasoil we include also a 8% of fuel oil)<br>RESIDENTIAL LPG 10% Gasoil 90%<br>(in gasoil kerosene is included).<br>INDUSTRIAL LPG 11% Gasoil 89%<br>(in gasoil we include also a 5% of fuel oil) |
| Coal | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Digest of UK Energy Statistics: : solid fuels and derived gases: commodity balances (DUKES ) | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632497/DUKES_2.4.xls | On the basis of available data the following shares are evaluated for Coal:<br>SERVICES 1,5% INDUSTRIAL 98,5% |
| Electricity | Level 3 (National) | UK Department for Business, Energy & Industrial Strategy | Digest of UK Energy Statistics: Electricity: commodity | https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/632598/DUKES_5.1.xls | On the basis of available data the following shares are evaluated for electricity:<br>SERVICES 50% INDUSTRIAL 50% |

| | | | | | |
|---|---|---|---|---|---|
| | | | balances (DUKES 5.1) | | |

**Table 24 – Methodology and source of data for Bristol fuel consumptions evaluation – LSOA disaggregation variables**

| Variable | Data availability | Sources | Publication | Reference | Fields |
|---|---|---|---|---|---|
| households not connected to the gas network | Level 3 (LSOA) | UK Department for Business, Energy & Industrial Strategy | Lower and Middle Super Output Areas gas consumption 2018 update | https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/767332/LSOA_estimates_of_households_not_connected_to_the_gas_network_2016.xlsx | Estimated number of households not connected to the gas network |
| Employees | Level 3 (LSOA) | UK Office for National Statistics | All people aged 16 to 74 in employment the week before the Census Occupation by industry 2011 Occupies | https://www.nomisweb.co.uk/census/2011/ks605uk | Geography<br>All of the following:<br>2001 super output areas - lower layer<br>Cell<br>SERVICE SECTOR<br>Table CAS039 Occupation by industry<br>select columns<br>• G Wholesale and retail trade; repair of motor vehicles and motor cycles<br>• H Transport and storage<br>• I Accommodation and food service activities<br>• J Information and communication<br>• K Financial and insurance activities<br>• L Real estate activities<br>• M Professional, scientific and technical activities<br>• N Administrative and support service activities<br>• O Public administration and defence; compulsory social security<br>• P Education<br>• Q Human health and social work activities<br>• R, S, T, U Other<br>INDUSTRIAL SECTOR<br>select columns<br>• B Mining and quarrying |

**Table 24 – Methodology and source of data for Bristol fuel consumptions evaluation – LSOA disaggregation variables**

| Variable | Data availability | Sources | Publication | Reference | Fields |
|---|---|---|---|---|---|
| | | | | | • C Manufacturing<br>• D Electricity, gas, steam and air conditioning supply<br>• E Water supply; sewerage, waste management and remediation activities<br>• F Construction<br>The table provides information that classifies usual residents aged 16 to 74 in employment the week before the census by the industry in which they work, for United Kingdom as at census day, 27 March 2011. |

**Table 25 – Methodology and source of data for Amsterdam fuel consumptions evaluation - Industrial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| Industrial sector | Natural Gas | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: Industry (Total_26)<br>Period: 2015<br>Energy commodities: Natural gas | Industry Companies number (Table 27 –) |
| | Gasoil | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: Industry (Total_26)<br>Period: 2015<br>Energy commodities: heating and other gasoil | Industry Companies number (Table 27 –) |
| | Coal | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: Industry (Total_26)<br>Period: 2015<br>Energy commodities: Total Coal Product | Industry Companies number (Table 27 –) |
| | Electricity | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: Industry (Total_26)<br>Period: 2015 | Industry Companies |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | Energy commodities: Electricity | number (Table 27 –) |

**Table 26 – Methodology and source of data for Amsterdam fuel consumptions evaluation - Residential and commercial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| Residential sector | Natural Gas | Level 3 (Buurt) | CBS | Wijk - en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | G_GAS_TOT*WONINGEN<br>where:<br>[G_GAS_TOT ]: Average total natural gas consumption<br>[WONINGEN]: Housing stock | None |
| | Wood | Level 2 (Gemeente) | RIVM | Klimaatmonitor | https://klimaatmonitor.databank.nl/dashboard/ | Wood burning stoves dwellings hern. heat [TJ] | Population (Table 27 –) |
| | LPG | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: households<br>Period: 2015<br>Energy commodities: LPG | Population (Table 27 –) |
| | Gasoil | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Selection:<br>Topic: households<br>Period: 2015<br>Energy commodities: Heating and other gas oil | Population (Table 27 –) |
| | Electricity | Level 3 (Buurt) | CBS | Wijk - en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | G_ELEK_TOT*WONINGEN<br>where:<br>[G_ELEK_TOT ]: Average total Electricity consumption<br>[WONINGEN]: Housing stock | None |
| Service sector | Natural gas | Level 2 (Gemeente) | RIVM | Klimaatmonitor | https://klimaatmonitor.databank.nl/dashboard/ | Gas use commercial Services [m3] + Gas use Public Services [m3] | Services Companies |

**Table 26 – Methodology and source of data for Amsterdam fuel consumptions evaluation - Residential and commercial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| | | | | | | | number (Table 27 –) |
| | Wood | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: services waste and repairs Period: 2015 Energy commodities: Solid and liquid biomass | Services Companies number (Table 27 –) |
| | LPG | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: services waste and repairs Period: 2015 Energy commodities: LPG | Services Companies number (Table 27 –) |
| | Gasoil | Level 1 (National) | CBS | Energy balance sheet supply consumption | https://opendata.cbs.nl/statline/portal.html?_la=en&_catalog=CBS&tableId=83140ENG&_theme=1028 | Topic: services waste and repairs Period: 2015 Energy commodities: Heating and other gas oil | Services Companies number (Table 27 –) |
| | Electricity | Level 2 (Gemeente) | RIVM | Klimaatmonitor | https://klimaatmonitor.databank.nl/dashboard/ | Electricity use commercial Services [kWh] + Electricity use Public Services [kWh] | Services Companies number (Table 27 –) |

(*) For non-domestic MSOA data, industrial half hourly (HH) consumption is provided for each local authority but it is not disaggregated further as doing so would break the UK Statistics Authority's Code of practice for Official Statistics relating to data disclosure. These LA data has been allocated to MSO using not-half hours MSOA data.

**Table 27 – Methodology and source of data for Amsterdam fuel consumptions evaluation – Buurt disaggregation variables**

| Variable | Data availability | Sources | Publication | Reference | Fields |
|---|---|---|---|---|---|
| Population | Level 3 (Buurt) | CBS | Wijk-en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | AANT_INW (Number of inhabitants) |
| Services Companies number | Level 3 (Buurt) | CBS | Wijk-en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | [A_BED_GI]+[A_BED_HJ]+[A_BED_KL]+[A_BED_MN]+[A_BED_RU]<br>where:<br>[A_BED_GI]: Number of companies and catering trade<br>[A_BED_HJ]: Number of companies transport, information, communication<br>[A_BED_KL]: Number of firms financially property<br>[A_BED_MN]: Number of companies business services<br>[A_BED_RU]: Number of companies culture, recreation, other |
| Industry Companies number | Level 3 (Buurt) | CBS | Wijk-en buurtkaart 2015 | https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015 | [A_BED_BF]: Number of companies industry and energy |

**Table 28 – Methodology and source of data for Ljubljana fuel consumptions/emissions evaluation - Industrial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| Industrial sector | Natural Gas | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Wood | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| | | | | (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | | | |
| | Coal | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Biogas | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | LPG | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Gasoil | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Coal | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Electricity | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |

**Table 29 – Methodology and source of data for Ljubljana fuel consumptions evaluation - Residential and commercial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| Residential sector | Natural Gas | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Wood | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | LPG | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Gasoil | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Coal | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Electricity | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| Service sector | Natural Gas | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Wood | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |

**Table 29 – Methodology and source of data for Ljubljana fuel consumptions evaluation - Residential and commercial sources**

| Activity | Energy vector | Data availability | Source | Publication | Reference | Field | Disaggregation variable |
|---|---|---|---|---|---|---|---|
| | LPG | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Gasoil | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |
| | Electricity | Level 2 (Občine) | EnerGis-Solution | Energy Balance And Emission Estimation - City Of Ljubljana (MOL) (Project version: 2016.MOL.1996-2015) Table EB-A: Energy Balance | http://www.energis-solutions.com/en/EB-Ljubljana-MOL/energy-balance | | Population |

**Table 30 – Methodology and source of data for Ljubljana fuel consumptions evaluation – Naselje disaggregation variables**

| Variable | Data availability | Sources | Publication | Reference | Fields |
|---|---|---|---|---|---|
| Population | Level 3 (Naselje) | Statistical Office of the Republic of Slovenia | SI-Stat Database | https://pxweb.stat.si/pxweb/Dialog/varval.asp?ma=05C5004E&ti=&path=../Database/Demographics/05_population/10_Number_Population/25_05C50_Population_naselja/&lang=1 | Population by large and 5-year age groups and sex, settlements, Slovenia, annually |
| Industrial areas coverage | Level 3 (Naselje) | Copernicus Land Monitoring Service | CORINE Land Cover | https://land.copernicus.eu/pan-european/corine-land-cover | A GIS query has been used to evaluated the coverage of industrial area on each Naselje and industrial emissions are allocated to the area based on dimension of area itself. |