ambience

Active managed Buildings with
Energy performaNce Contracting

Deliverable Number (D4.1)

Database of grey-box model parameter values for EU building typologies

The AmBIENCe Consortium

December | 2021

DOCUMENT CONTROL PAGE

PROJECT ACRONYM	AmBIENCe
DOCUMENT	D4.1 Database of grey-box model parameter values for EU building typologies
TYPE (DISTRIBUTION LEVEL)	☒ Public ☐ Confidential ☐ Restricted
DUE DELIVERY DATE	31/08/2021
DATE OF DELIVERY	24/12/2021
STATUS AND VERSION	V02
DELIVERABLE RESPONSIBLE	BPIE
AUTHOR (S)	Ivan Jankovic (BPIE), Xerome Fernandez (BPIE), Jan Diriken (VITO)
OFFICIAL REVIEWER/s	Jessica Glicker (BPIE), Judit Kockat (BPIE), Sheikh Zuhaib (BPIE), Chris Caerts (VITO), Annick Vastiau (VITO), Nuno Mateus (EDP CNET), Joao Cravinho (EDP CNET), Tiago André Soares (INESC TEC).

EXECUTIVE SUMMARY

Deliverable D4.1 "Database of grey-box model parameter values for EU building typologies" describes the development of a European building stock database, collecting information and data on parameters needed for both the assessment of the Energy System Impact and the creation of the performance models. These grey-box model parameters characterize the dynamic thermal behaviour of the respective buildings which is needed to assess the impact of building flexibility.

This deliverable explains how such a building stock database that also captures the dynamic thermal behaviour of the listed buildings, was generated starting from available building stock databases that only contain information about a building's static Energy Efficiency. As the information in available building stock databases is limited and not meant to be the basis for dynamic thermal model creation, some assumptions had to be made that are described in a transparent manner. Similarly, only approximate shoe-box models, i.e., building models with one or just few zones, have been created. But this work lays the foundation to effectively and (semi)automatically create more detailed dynamic thermal models in the future, based on building stock databases containing more detailed information.

In order to answer its requirements, for different building stock segments of EU 27 Member States, the database provides information about the building stock segments' reference building, as well as information on the building stock segments.

Building stock segments present in the database are organised by building use, construction period, and, where applicable, construction methodology and/or materials, while each building stock segment is represented by a single reference building.

The database was fully developed from two publicly available data sources: TABULA and Hotmaps that provide relevant data for residential and non-residential buildings and define building stock segments based on country, building use, construction period, and additional parameters, where applicable.

Following the series of data transformation steps and assumptions used to fill the data gaps, the database provides data on different building stock segments and reference buildings, by focusing on:

- Reference building classification data,
- Reference building geometry data,
- Characteristics of the reference building envelope elements
- Building stock segment data.

Although the database was developed to be used for the purpose of the AmBIENCe project, the database will be publicly available for different purposes. For this reason, one should notice that the database comes with different strengths including its standardized structure, as well as weaknesses such as coarse definition of the material composition. However, once its development and assumptions are well understood, the database should be able to satisfy different research needs in the field of building energy performance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION AND BACKGROUND

1.1 THE CONTEXT

A new concept for performance guarantees of Active Buildings, combining savings from energy efficiency measures and the active control of assets, leveraging the use of flexibility is developed within AmBIENCe. The combination of Demand Response (DR) with current Energy Performance Contract (EPC) schemes establishes the Active Building EPC (AEPC) concept. The AEPC concept uses intelligent and real-time information to offer new combined services, established comfort and safety performance criteria and new levels of flexibility activation and use. The project aims to extend the concept of Energy Performance Contracting to Active Buildings, which are buildings equipped with active control options that can actively participate in demand response and energy efficiency programmes and make it available and attractive to a wider range of buildings.

New data on buildings are needed to introduce DR and flexibility into EPC contracts. Data on buildings have been generated in different EU projects such as EUCalc[1] and ENTRANZE[2], and also, the Building Stock Observatory contains a database. However, these data are often lacking those specific aspects linked to demand response and flexibility. Typically, the details on stereotype buildings representing the European buildings stock are limited to their building geometry and insulation level and associated heat and energy related parameters. The focus is on energy consumption and energy efficiency level, but parameters that allow to develop a dynamic thermal model of a building need to be included, allowing more easily calculations of flexibility measures for a building.

1.2 PURPOSE AND SCOPE OF THE DOCUMENT

Deliverable D4.1 falls within the scope of Work Package 4: "Energy System Impact Assessment Calculation Methodology and Tool," with the main goal to allow calculations of Energy System Impact KPIs in transparent manner.

The deliverable 4.1 describes the data base of grey-box model parameter values for EU building typologies. It aims at creating a building stock information database needed for performance modelling and calculations. This building stock information database is provided on the website of the AmBIENCe project and includes data on the EU27 building stock and its reference buildings. The building stock information database developed in this task, used building stock related databases provided by different EU projects, but adds parameters that allow to develop a dynamic thermal model of a building, facilitating calculations of flexibility measures for a building.

The resulting building stock information database will support Energy System Impact calculation, , as well as creation of performance models that can be added to the Active Building Energy Performance Modelling

---

[1] http://tool.european-calculator.eu/intro
[2] https://www.entranze.eu/

(ABEPeM) tool developed under WP2, e.g., to enable speedier pre-contracting assessments.

As a brief reminder, the ABEPeM tool's purpose is to analyse integrated energy performance of a building where special attention is paid to how value added by energy efficiency and demand response can be combined. Input from the database developed and its dependent models are integrated into the ABEPeM tool, allowing to quantify energy consumption shifts and energy system impact.

From this point onwards, the database developed described in this deliverable will be referred to as database.

The purpose of this document is to transparently explain how the database was prepared and how it could be used. Although developed for the purpose of the AmBIENCe project, the database use should not be limited to the AmBIENCe project and will be publicly available for download at the project's website, which adds another dimension to this report.

In more details, in a simple and transparent manner the report should explain:

- Grey-box model parameter values for EU building stock,

- Main database elements, such as reference building and building stock segment, as well as their meanings and roles in the database,

- Methodology of the database development,

- Data sources used,

- Data collection approach,

- Basic transformations of the source data,

- Basic assumptions needed for filling the data gaps, and

- Fields included in the database.

2. GREY-BOX MODEL PARAMETER VALUE FOR EU BUILDING TYPOLOGY

2.1 PROCESS DESCRIPTION

2.1.1 OVERVIEW AND INTRODUCTION INTO DYNAMIC THERMAL MODELS

By simulating the dynamic thermal behaviour of a building, one can gain insight in the flexibility embedded in the thermal mass of the building under consideration. When it comes to dynamic building simulations three models can be used: white-box, grey-box, and black-box models.

Firstly, white-box models can be constructed. This type of models aims to construct a detailed description of a building. The thermophysical properties of all materials used in the building envelope are specified and each construction element like walls, windows, or roof is individually defined considering the area and orientation. Often the heat distribution system is explicitly included, and a distinction is made between radiators, convectors, or floor heating. This type of model provides a great amount of detail but has the downside that it is computational expensive and specific for a given building.

An alternative approach can be followed using grey-box models, which represent the thermal dynamics in a building in a simplified way. A grey-box model has a fixed structure and includes one or more phenomena that can play a role in the thermal dynamic behaviour of a building. Each model structure is described by a set of model parameters. The values of these parameters will be different from building to building and are identified based on (experimental) data. Although the absolute values of these parameters do not have a physical significance, they can be connected to certain properties of the building. Hence the term "grey-box". Due to their simple structure, grey-box models are fast to calculate and thus suitable for optimization problems and long-term simulations.

The last method relies on the use of black-box models. Using statistical analysis and historical data a relation between inputs and outputs is identified. There is no underlying physical foundation. The results are only valid for the building from which the historical data was used. Common examples are Artificial Neural Networks and Support Vector Machines.

From the properties of the different modelling approaches, grey-box modelling has the most potential to be included in a building stock database. For a given reference building in a database, it is sufficient to specify the model structure and relevant parameters that can be used to describe the thermal behaviour of the building.

The main goal of this task is to develop a procedure to identify the most suitable grey-box model and parameters for reference buildings using the information that is provided by building stock databases. The process is schematically shown in Figure 1. To achieve this ambition on a large scale an intermediate unifying building stock database was created that can be fed from different sources (step 1). The structure of this unifying database, which is a MS Excel table, and sources that were used to populate it are discussed in detail in Section 3 and later. In a first step, the quantitative information present in the unifying building stock database is used to create white-box models for every building present in the database (step 2). Next, all white-box models are simulated to generate data on the thermal behaviour of each building (step 3). In

the last step, this data is used to train the grey-box models, select the most suitable grey-box model structure, and identify the relevant grey-box model parameters (step 4) that are added to the unifying database.
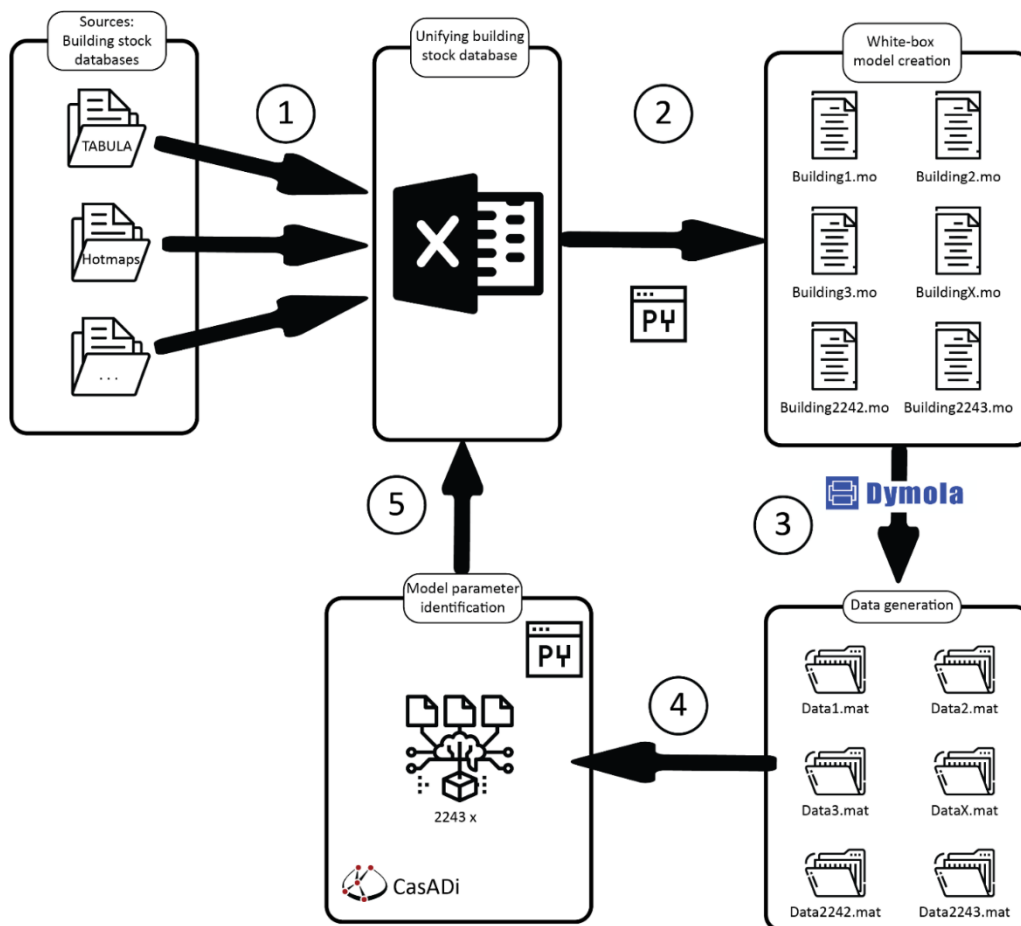
FIGURE 1 - SCHEMATIC REPRESENTATION OF THE AMBIENCE WORKFLOW

In the sections below each step in the process, as well as the assumptions that were made, is described in more detail.

2.1.2 NOMENCLATURE

The following list provides a definition of different terms that are used in the process description below.

− **Database:** Reference to the intermediate unifying building stock database containing information from different sources. This unifying database is the source of data used to create the white-box models.

− **Model code:** Structured text file (extension.mo) which defines a white-box model of a building. This includes a description of the building geometry and materials used. Also, general information needed to perform a simulation is included like reference to climate files, indoor temperature set points, user behaviour and description of the heating system.

− **Building Information Summary File:** Contains all the information calculated during the entire process as well as the results from the grey-box parameter identification step. Its main purpose is to quickly access information needed in the different steps of the process and to identify issues and errors.

2.1.3 DIFFERENT STEPS TO IDENTIFY GREY-BOX PARAMETERS

Database processing and white-box model creation

A tool based on Python programming language[3] was developed to interpret the data included in the unifying database which will be referred to below as the database. This is step 1 in Figure 1. It should be noted that a fixed structured, this means the order and definition of each column, for the MS Excel table is required and that a value for each field must be provided for every reference building. The exact definition of each field or column can be found in Chapter 8. Each line in the database corresponds to a single reference building. Finally, all values within a column must be of the same pre-determined unit (e.g., m² or W/mK).

Once the information from the database is read out, the information is used to construct a simple white-box model for each building defined in the database (Figure 1, step 2). This white-box model is created using the validated open-source IDEAS library within the object-oriented Modelica[4] language. This library contains all components required to build and simulate the thermal behaviour of a building.

The white-box model consists of different thermal zones: one for each floor. Each zone is enclosed by four outer walls, a floor, and a roof. There is a window in every wall. In the case of the ground floor the thermal properties of the floor are those as defined in the database. For the highest floor, the provided definition of the roof is used to describe the roof. In case there are other floors in between a standard composition (concrete and tiles) is used. Windows are divided over each outer wall pro rata the wall area. Also, internal walls, and hence their thermal mass, are included in a lumped manner. On the ground floor a door is included with a fixed area of 2 m². The roof is considered to be flat. Infiltration losses are included and

---

[3] https://www.python.org/
[4] https://modelica.org/

increase with the age of a building. A graphical presentation of the model structure is given in Figure 2. Not all components are shown for simplicity.
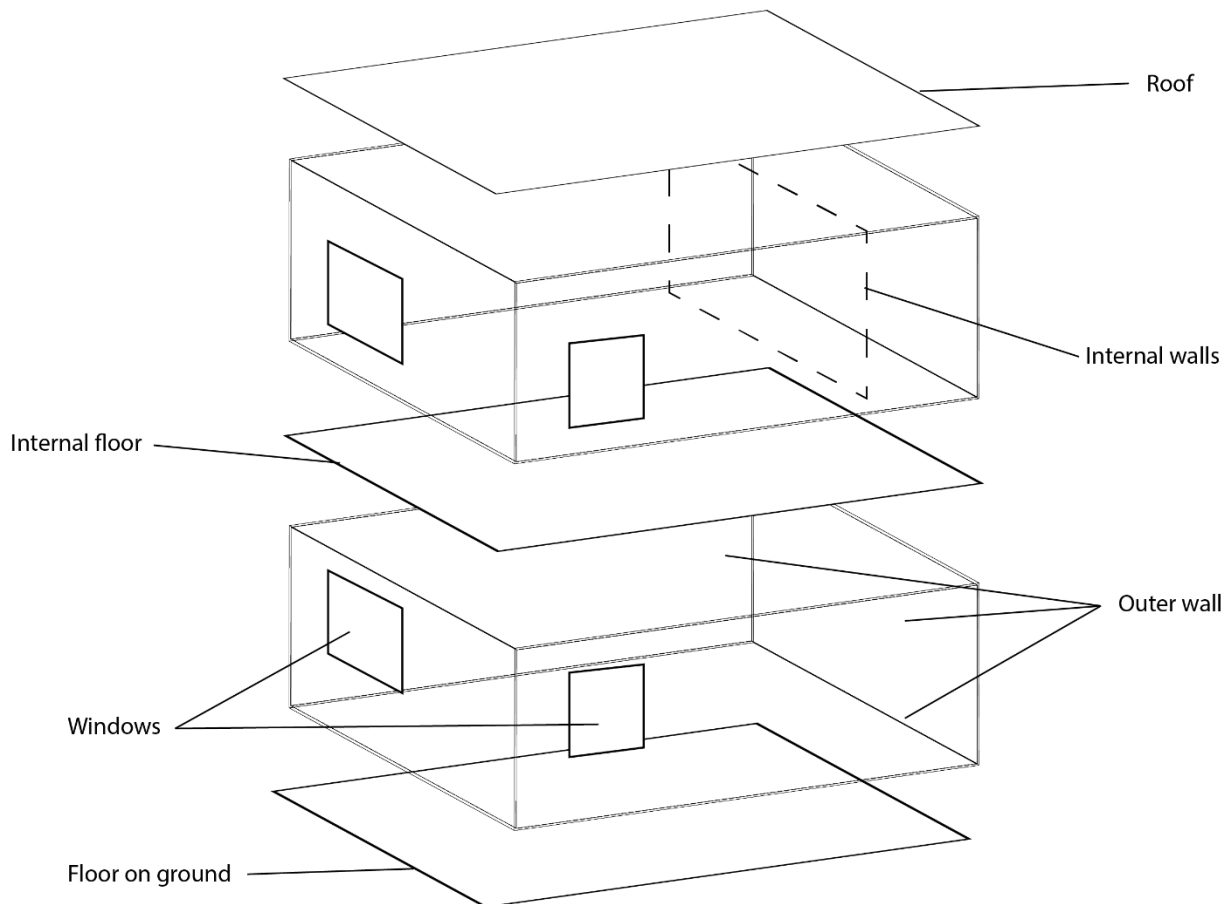
FIGURE 2 – STRUCTURE OF THE WHITE-BOX MODELS (2 FLOOR BUILDING)

Based on the geometrical information in the database, which includes the gross area of the outer walls, windows, floor and roof, the dimensions of the building are determined as follows:

1. The number of floors is determined by dividing the total useful floor space by the building footprint. The value is rounded up to be an integer.

2. The building is assumed to be a cuboid.

3. It is assumed that each floor has a ceiling height of 2.6 m.

4. The aspect ratio of the footprint of the cuboid is the ratio between the longer (l_1) and shorter (l_2) length of the base of cuboid ($AR = \frac{l_1}{l_2}$) and its value is calculated using the information on the floor, wall and window area using the following equation:

$$l_{1,2} = \frac{\frac{A_{facade}}{2\,n_{floors}h_{ceiling}} \pm \sqrt{\left(\frac{A_{facade}}{2\,n_{floors}h_{ceiling}}\right)^2 - 4A_{floor}}}{2}$$

In this equation A_{facade} is the total area of walls, windows, and doors, n_{floors} is the number of floors, $h_{ceiling}$ is the ceiling height and A_{floor} the ground floor area. The two solutions give the longer and shorter length of the walls.

Remark: on some occasions the discriminant is negative and no solutions for $l_{1,2}$ can be found using this expression. In that case, the aspect ratio is fixed to be 1.5 and the ceiling height is determined as to match the resulting area of the walls to those specified in the database. It should be noted that this can lead to unrealistic values of the ceiling height in particular cases, like 1.2 m or lower. Model parameters are also determined for these cases but should be considered indicative.

5. The materials that are described in the database are created within the white-box model framework. Their properties are defined based on the information provided in the database and given a unique name. These properties are the density, thermal capacity, and thermal conductivity of the material.

6. The general building components are defined in the framework. This encompasses combining different materials created in step 4 to form e.g., the outer wall or floor. Each material has a certain thickness which is defined in the database.

7. The individual components of the building envelope are defined. In this step the general definition of e.g., and outer wall is taken and assigned an area and orientation. The dimensions are based on the geometry determined in step 3. In every case the longer wall is oriented along the east-west axis and thus facing north and south. In the case of internal walls, the total area of the included indoor walls equals that of the facade of the floor. This corresponds to dividing the floor into four individual rooms[5]. The area of each window is proportional to the area of the wall it is located in.

8. All components are connected to the relevant zones. Also, the connections with the heating system are defined.

Applying the procedure above to the database will result in a collection of white-box models, namely one for each reference building present in the database. As a final step, the common part of the model code is added. This includes the heating system, control settings, user behaviour and the general simulation manager.

For the heating system a simplified model is used. This model makes the following assumptions:

−   Heat is provided to each zone individually and proportionally to the difference between the actual and set point temperature in the zone,

---

[5] It should be noted that the floor is treated as a single thermal zone that includes the thermal mass as if there are four rooms on that floor.

- The heat is delivered both convectively and radiatively in fixed ratio,

- Cooling is not included.

Control settings include a random set-point regime, both in time and temperature (Figure 3). This non-realistic regime is used to be sensitive to the thermal dynamics of the building, which are less obvious during realistic operation, where the aim is to keep the building at the desired temperature. Realistic operation gives insight in the thermal losses of the building but is less sensitive to the thermal capacity. This justifies the use of more dynamic heating regimes. The same random regime is used in every simulation.

User behaviour represents internal gains and is based on a stochastic profile for a single-family of two adults and two children.
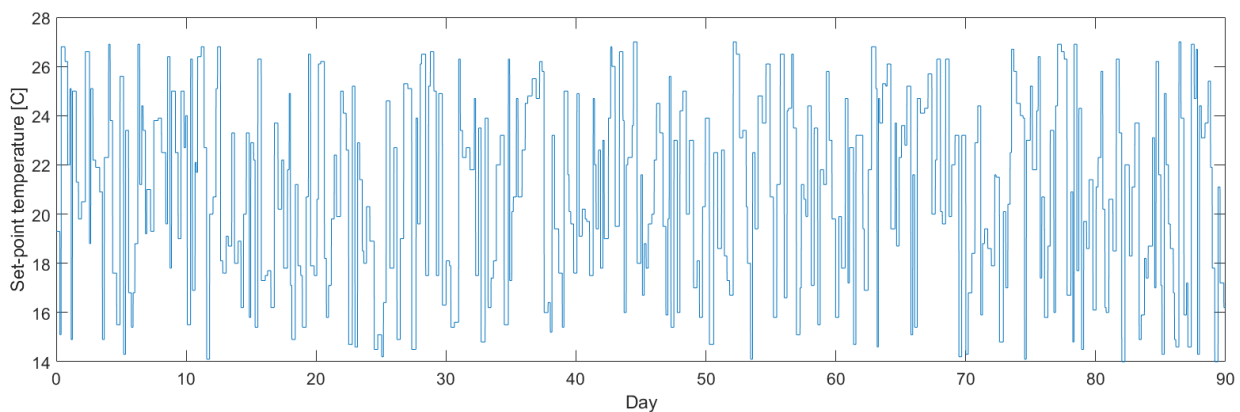
FIGURE 3 - SET-POINT REGIME DURING VIRTUAL EXPERIMENTS

Finally, the simulation manager includes information on the climatic conditions. These are adjusted based on the country in which the reference building is located, see Table 1. Typical Meteorological Year data (TMY) taken from EnergyPlus was used, if available. A single city was used for each country (see table below). The resolution of each TMY file is 1 hour.

Additionally, a building information summary file (csv-file) is created which collects all values calculated during the model creation such as, n_{floors}, $h_{ceiling}$ and $l_{1,2}$ for each reference building. Also, the volume of a single thermal zone is saved. The main purpose of this file is twofold: the information stored in this file can be used during later steps of the process and it can be used to identify potential issues and errors, like unrealistic ceiling heights mentioned earlier.

TABLE 1 - REFERENCE CITIES FOR EU COUNTRIES ACCORDING TO ENERGYPLUS

COUNTRY	CITY	COUNTRY	CITY
Austria	Vienna	Italy	Rome
Belgium	Brussels	Latvia	Kaunas (LTU)
Bulgaria	Sofia	Lithuania (LTU)	Kaunas
Croatia	Banja Luka (BIH)	Luxembourg	Nancy (FRA)
Cyprus (CYP)	Larnaca	Malta	Larnaca (CYP)
Czech Republic	Prague	Netherlands	Amsterdam
Denmark	Copenhagen	Poland	Warsaw
Estonia	Kaunas (LTU)	Portugal	Lisbon
Finland	Helsinki	Romania	Bucharest
France (FRA)	Paris (Orly)	Slovakia	Bratislava
Greece	Athens	Slovenia	Ljubljana
Hungary	Debrecen	Sweden	Stockholm
Ireland	Dublin		

White-box simulation loop

In this step each (white-box) model code, representing a reference building, is compiled and simulated in its representative climatic conditions. A python script is used to loop over all model codes created in the previous step. During every iteration of this loop, the IDEAS library and model code under consideration are loaded and the model code is compiled. The simulation starts on January first at midnight and runs for 3 months. Data is saved every 900 seconds. The LSodar (Livermore Solver for Ordinary Differential equations) is used to solve the differential equations within the model. This is a highly efficient method for both stiff and non-stiff problems by utilizing a variable time-step. The tolerance is set to 1E-5. Dymola is used as software environment.

The time required to simulate a given reference building scales linearly with the number of thermal zones, and thus floors, present in the building. As a rule of thumb, it takes two to three minutes of simulation time per thermal zone. In the case of apartment blocks which contain in some cases more than 20 floors, the simulation time can thus increase to more than an hour. The task was performed on an Intel® Core™ i5-10310U CPU @1.70 GHz, 8 GB RAM, 64-bit MS Windows 10 laptop.

After the simulation is finished the relevant data are exported from the results file. These data are ambient temperature, solar irradiance, total heat input and the indoor temperature in each zone. These data are saved in a csv file and are used in the next step to create the grey-box models and associated parameters.

Grey-box model parameters identification

A second python program runs in parallel with the simulation loop and checks at regular time intervals whether new data csv files have been created. If so, the data is shaped in the right form to apply the grey-box model parameter fitting procedure. The main step is calculating the average indoor temperature. Since each thermal zone in a building has by design the same volume it is not necessary to apply weighting.

The input data is split into two parts: the first two months are used for training, while the third month is used for validation. The total volume of the building[6] is used to initialize the parameter related to the thermal capacity of the building. The procedure will provide the training data to all defined grey-box model structures and apply the fitting routine. After convergence, the parameters are used in the validation phase along with the ambient temperature, heat input and solar irradiance to predict the evolution of the indoor temperature. The resulting temperature evolution is compared to the simulated one. Finally, RMSE (root-mean square error), MAPE (mean absolute percentage error) and variance values for both the training and validation phase are calculated.

In the end the most suitable grey-box model structure and associated model parameters are selected based on the lowest RMSE value. The grey-box model structure identifier and parameters are added to the Building Information File as well as the RMSE, MAPE and variance values for training and validation. Plots depicting the grey-box model results for all structures, as well as their parameters and input data, are saved in a dedicated folder.

Since the entire building is reduced to a single zone, the time required to apply the fitting procedure is independent of the original building structure. The time required to perform a fit of the selected grey-box model structures is two to three minutes.

Data merging

The last step in the process is merging the results of the grey-box model fitting procedure (i.e., selecting the most suitable grey-box model structure and associated model parameters) with the database. This can be done using the unique building identifier present in the database, which was used in every step of the process described above

2.2 GREY-BOX MODEL STRUCTURES

Using grey-box models, the evolution of the internal state of a system can be simulated, given the inputs to the building. The factors/inputs that affect the behaviour of the building in this context are outdoor temperature, heat/cooling input to the building and solar irradiation incident on the building. Grey-box models have the advantage of being flexible and scalable, while still capturing the physical characteristics of the heat flows between various building components. These properties make them an ideal choice for applications involving e.g., model-predictive control or long-term simulations.

---

[6] Obtained from the Building Information File

In this work a family of grey-box models, called RC models, is used. An RC network is built by connecting thermal resistances and capacitances as lumped parameters. There are analogues of electrical circuits and model thermal behaviour instead. The analogues for voltage, current, resistance and capacitance are given by temperature, heat flux, insulation, and the thermal mass respectively (Figure 4).
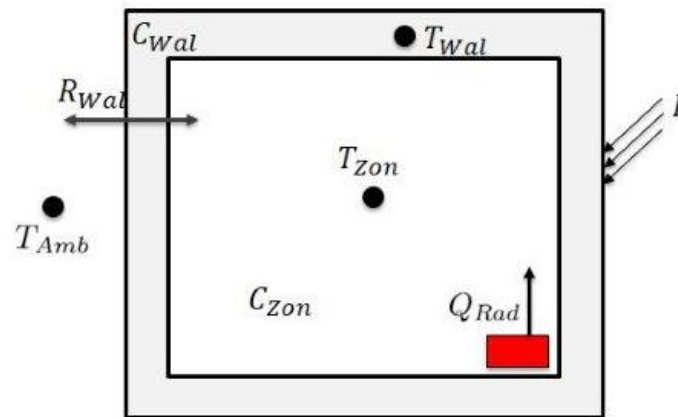
FIGURE 4 – ILLUSTRATION OF A BUILDING AS A THERMAL ZONE

The Grey-Box Model creation toolchain selects the most suitable RC model topology from the available RC model family, as well as the associated parameter values for the selected RC model. The different members of the RC model family that are available and used are Zone D, Zone A, Zone Int A, Zone wall B; Zone Int wall B, and are detailed below.

2.2.1 STRUCTURE 1: ZONE D

This is the simplest structure of the RC model family and described by a single ordinary differential equation (ODE) containing two model parameters to be fitted:

$$C_i \frac{dT_i}{dt} = \frac{T_e - T_i}{R_{vent}} + \dot{Q}_h$$

Where:

- t represents time,
- C_i is the capacitance of the thermal mass,
- R_{vent} is the resistance offered by the wall of the house,
- T_e is the external or ambient temperature,
- T_i is the indoor temperature,
- $\dot{Q}_h$ denotes the heat input into the building (negative when used in cooling mode).

The equation describes that the main contributing factors to the change of the internal temperature is the temperature differential with the outdoor temperature (when $T_e > T_i$, the building is heated due to the difference, and when $T_e < T_i$, heat is lost to the ambience). The label of this model structure D refers to direct heating of the indoor climate. A graphical representation of the zone D model structure is shown in Figure 5.
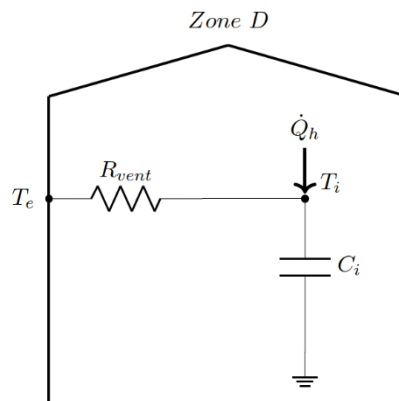
FIGURE 5 - GRAPHICAL REPRESENTATION OF THE ZONE D MODEL STRUCTURE

This model is **included** in the grey-box model identification routine.

2.2.2 STRUCTURE 2: ZONE A

This structure is nearly the same as the model above, but in this case, heat gain due to solar irradiance incident on the windows and surfaces of the building is factored in. The ODE described above is modified as follows:

$$C_i \frac{dT_i}{dt} = \frac{T_e - T_i}{R_{vent}} + \dot{Q}_h + g_A \dot{Q}_{Sun}$$

Where in addition to the notation of Zone D, we have two other quantities

– $\dot{Q}_{Sun}$ denotes the total solar irradiance,

– The parameter g_A is the gain associated with the solar irradiation.

The model structure name "Zone A" originates from the additional parameter g_A.

A graphical representation of the zone A model structure is shown in Figure 6.
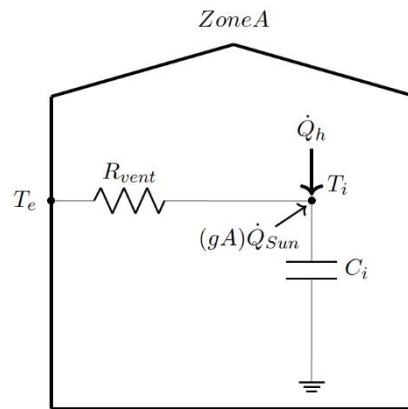
FIGURE 6 - GRAPHICAL REPRESENTATION OF THE ZONE A MODEL STRUCTURE

This model is **included** in the grey-box model identification routine.

2.2.3 STRUCTURE 3: ZONE INT A

Zone A and Zone D were single state systems (i.e., the system is described by the evolution of just one temperature – the indoor temperature Ti). Many times, not all the dynamics of a building can be captured by such single state networks. Hence, a two-state system is introduced which captures the behaviour of the various contents of the house (furniture, indoor walls, etc.) as additional internal mass. The two states are the indoor temperature as before (Ti), and internal temperature (Tint).A graphical representation of the zone int A model structure is shown in Figure 7.
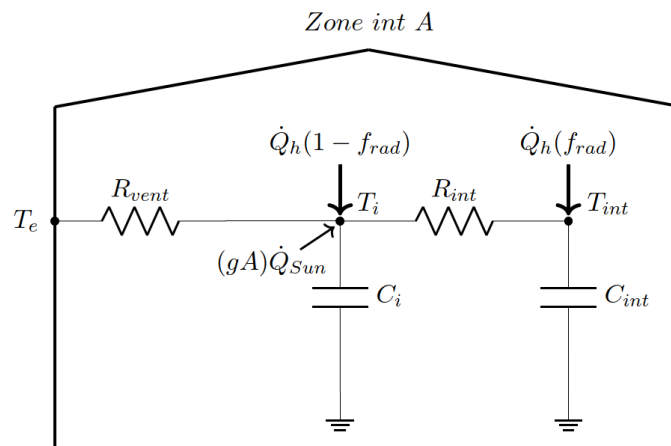
FIGURE 7 – GRAPHICAL REPRESENTATION OF THE ZONE INT A MODEL STRUCTURE

Since there are two states, a system of two ODE's that describe the evolution of these two temperatures based on the inputs can be formulated.

$$C_{int} \frac{dT_{int}}{dt} = \frac{T_i - T_{int}}{R_{int}} + f_{rad}\dot{Q}_h$$

$$C_i \frac{dT_i}{dt} = \frac{T_e - T_i}{R_{vent}} + \frac{T_{int} - T_i}{R_{int}} (1 - f_{rad})\dot{Q}_h + g_A \dot{Q}_{Sun}$$

Where the new parameters/quantities are as follows:

- T_{int} is the internal temperature,
- f_{rad} is the fraction of the input heat that is absorbed by the internal objects. Hence $(1 - f_{rad})$ percent of the heat will go towards heating the air,
- C_{int} is the internal capacitance, representing internal thermal mass,
- R_{int} is the resistance between the internal thermal mass and the air.

Note that T_{int} is a hidden state and cannot be measured. Additionally, it should be noted that C_i has a different meaning here compared to Zone D and Zone A.

This model structure was **not included** during the training procedure (see below for more details).

2.2.4 STRUCTURE 4: ZONE WALL B

This is also a two-state system, where instead of distinguishing between walls/floors and furniture, we distinguish between the building envelope and indoor thermal mass. The system of ODEs that describe the dynamics of this model are given by

$$C_i \frac{dT_i}{dt} = \frac{T_w - T_i}{R_{vent}/2} + \dot{Q}_h + g_A \dot{Q}_{Sun}$$

$$C_w \frac{dT_w}{dt} = \frac{T_i - T_w}{R_{vent}/2} + \frac{T_e - T_w}{R_{vent}/2}$$

Where the new parameters/quantities are as follows

- T_w is the outer wall temperature (as measured indoor),
- f_{rad} is the fraction of the input heat that is absorbed by the internal objects. Hence $(1 - f_{rad})$ percent of the heat will go towards heating the thermal mass (air),
- C_w is the thermal mass of the building envelope.

A graphical representation of the zone wall B model structure is shown in **Error! Reference source not found.**.
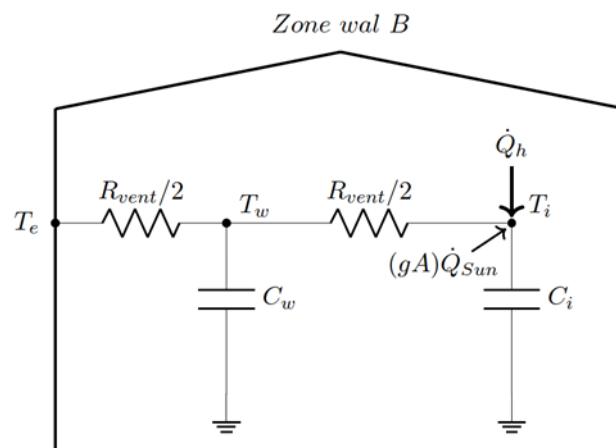
FIGURE 8 – GRAPHICAL REPRESENTATION OF THE ZONE WALL B MODE STRUCTURE

This model structure was **not included** during the training procedure.

2.2.5 STRUCTURE 5: ZONE INT WALL B

This model is a combination of the above two models, where the internal air temperature, temperature of outer walls and the temperature of internal walls and furniture (i.e., 3 states) are included. A graphical representation of the zone int wall B model structure is shown in Figure 9.
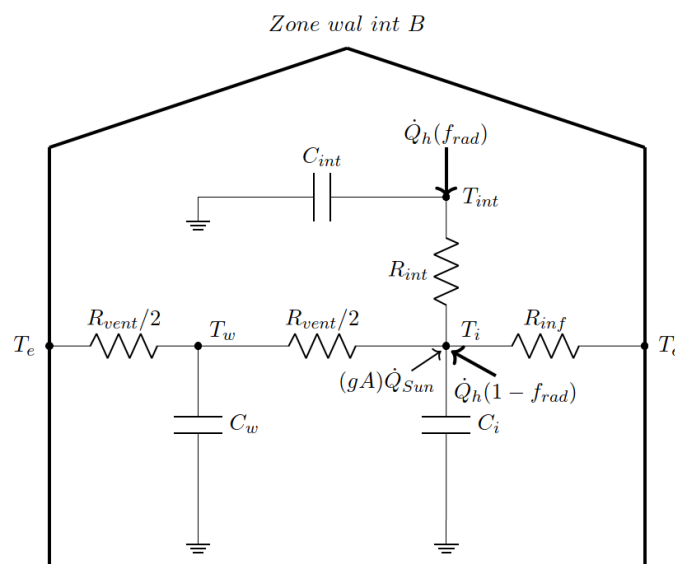
FIGURE 9 – GRAPHICAL REPRESENTATION OF THE ZONE WALL B MODEL STRUCTURE

The following set of ODEs describe the system:

$$C_{int} \frac{dT_{int}}{dt} = \frac{T_i - T_{int}}{R_{int}} + f_{rad}\dot{Q}_h$$

$$C_i \frac{dT_i}{dt} = \frac{T_w - T_i}{R_{vent}/2} + \frac{T_{int} - T_i}{R_{int}} + (1 - f_{rad})\dot{Q}_h + g_A \dot{Q}_{Sun} + \frac{T_e - T_i}{R_{inf}}$$

$$C_w \frac{dT_w}{dt} = \frac{T_i - T_w}{R_{vent}/2} + \frac{T_e - T_w}{R_{vent}/2}$$

Where all the parameters are as described as before and R_{inf} is the resistance between the external and the indoor.

This model structure was **not included** during the training procedure.

Matrix representations

The three more complex models described above have not been included in the training procedure due to time constraints. For the given training time frame, the fitting of the more complex models can take up to 20 minutes to converge, compared to seconds for the simple model structures. For the number of entries in the database this would mean several months (which is not possible in the timespan of the work) instead of several days.

Therefore, two alternative grey-box model structures have been introduced which rely on the same inputs, observables and internal states defined in structures 3, 4 and 5 but in an equation form that can be written as a matrix formulation:

$$X_{t+1} = A X_t + B U$$

Considering grey-box model 4 as an example, X_t is a state vector containing T_i and T_w. U is the input vector with elements T_e, $\dot{Q}_{Sun}$ and $\dot{Q}_h$. It should be noted that this model formulation is intrinsically slightly different than the approach defined above. Here the model equations are given in an explicit Euler formulation since the states at the next steps are written as a function of the current state and input variables. There are 10 model parameters in this case: 4 within the 2×2 A-matrix and 6 in the 3×2 B-matrix. One can also rewrite the equations for grey-box-structure 4 explicitly as:

$$\begin{bmatrix} T_{i,t+1} \\ T_{W,t+1} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} T_{i,t} \\ T_{w,t} \end{bmatrix} + \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \end{bmatrix} \begin{bmatrix} T_e \\ \dot{Q}_{sun} \\ \dot{Q}_h \end{bmatrix}$$

$$T_{i,t+1} = a_1 T_{i,t} + a_2 T_{w,t} + b_1 T_e + b_2 \dot{Q}_{sun} + b_3 \dot{Q}_h$$
$$T_{w,t+1} = a_3 T_{i,t} + a_4 T_{w,t} + b_4 T_e + b_5 \dot{Q}_{sun} + b_6 \dot{Q}_h$$

This particular model structure is referred to as Zone Wall B Matrix. It should be noted that the matrix representation for Zone Int A (structure 3) is the same, albeit with slightly different nomenclature since T_W

has to be replaced by T_{int}. Hence, in the matrix representation these 2 models are the same.

Despite the fact that this model contains more parameters, this representation is linear in its parameters and converges very rapidly (within a couple of seconds), while including the same effects introduced before. From a time-efficiency perspective this model structure is fitted instead of the regular one introduced before. The downside of this approach is the lack of an insight in the fit results since the matrix elements do not have a connection to a physical entity or process.

Also, a matrix representation for the Zone Int Wall B model is included in the procedure. In this case X also contains the hidden state T_{int} and both matrices A and B now have dimension 3×3, leading to 18 parameters to be fitted. Also, this model structure has been included during the fitting routine:

$$\begin{bmatrix} T_{i,t+1} \\ T_{W,t+1} \\ T_{int,t+1} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} T_{i,t} \\ T_{w,t} \\ T_{int,t} \end{bmatrix} + \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} T_e \\ \dot{Q}_{sun} \\ \dot{Q}_h \end{bmatrix}$$

$$T_{i,t+1} = a_1 T_{i,t} + a_2 T_{w,t} + a_3 T_{int,t} + b_1 T_e + b_2 \dot{Q}_{sun} + b_3 \dot{Q}_h$$
$$T_{w,t+1} = a_4 T_{i,t} + a_5 T_{w,t} + a_6 T_{int,t} + b_4 T_e + b_5 \dot{Q}_{sun} + b_6 \dot{Q}_h$$
$$T_{int,t+1} = a_7 T_{i,t} + a_8 T_{w,t} + a_9 T_{int,t} + b_7 T_e + b_8 \dot{Q}_{sun} + b_9 \dot{Q}_h$$

The Figure 10 is an example of a randomly selected reference building trained for both the structure 5 (Zone Int Wall B) and its matrix representation (18 parameters) for the reference building with code BE-HOR-1970-1979-01. Although the predicted evolution of the indoor climate is different between these two grey-box models, the regular (structure 5) Zone Int Wall B model would also still be suitable for practical use. The RMSE in training is 1.46, while during validation it is 1.56, compared to 0.28 and 0.78 for the matrix representation.
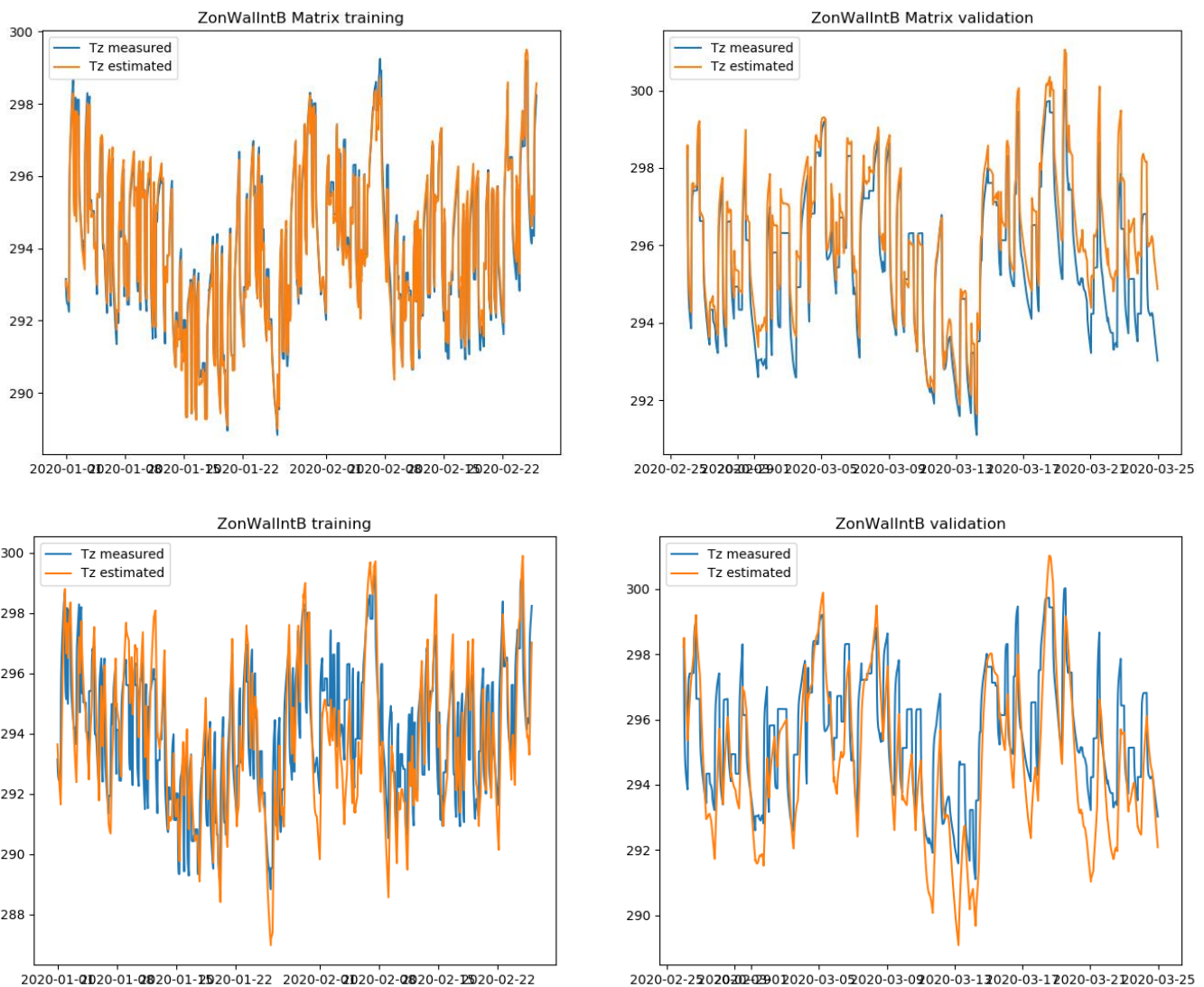
FIGURE 10 - COMPARISON BETWEEN ZONWALINT B RC (TOP) AND MATRIX (BOTTOM) REPRESENTATION

The fact that both models do not generate the same result stems mainly from the fact that the matrix representation allows additional dynamics not included in the RC version. As an example, $\dot{T}_w$ does not depend on T_{int}, Q_{sun} nor Q_h according to the model equations defined in structure 5. There are however matrix elements associated to these variables in the matrix representation, namely a_6, b_5 and b_6. Since these matrix elements are not forced to be zero, they can influence the evolution of the model states. A statistical analysis can be performed on the relative magnitude of these matrix elements to investigate their importance. In an ideal world, where Zone Int Wall B would be a correct description of the different elements that influence the dynamic behaviour of a building, a_6, b_5 and b_6 should have a value of 0. An alternative approach is to force matrix elements that should be zero according to the model equations in structures 3 to 5 to be effectively zero.

It is also possible to compose a set of relations between the elements of matrices A and B and the parameters in the model structures 3 to 5 (like values of the different Rs, Cs, g_A and f_{rad}). Since this is set is overdetermined and non-linear, a dedicated optimization problem should be defined to estimate the values of the Rs, Cs and f_{rad}. When applying this kind of method, one should take care to account correctly for the different model equation formulations (explicit Euler vs. time derivative).

2.3 TRAINING RESULTS

2.3.1 DISTRIBUTION OF SELECTED MODELS

For each reference building the matrix representation of the most complex structure Zone Int Wall B, consists of 18 parameters.

2.3.2 RMSE ANALYSIS

Figure 11 shows the distribution of RMSE values from the training (x-axis) and validation (y-axis) phase. Except for a few outliers, most training RMSE values are smaller than 0.5 °C. During validation RMSE values are in general higher, although exceptions occur. 44% of the trained models has a RMSE value for the validation phase of less than 1 °C.
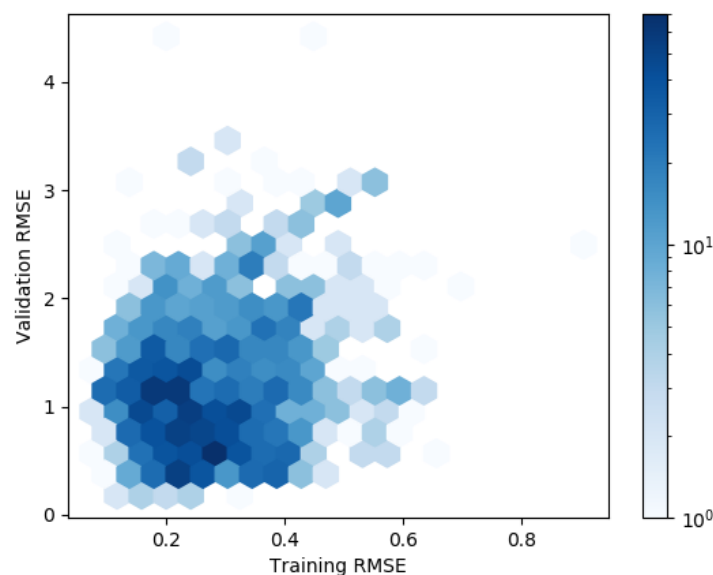
FIGURE 11 - DISTRIBUTION OF RMSE VALUES FOR BOTH TRAINING AND VALIDATION PHASE

2.3.3 MATRIX ELEMENT DISTRIBUTION

Earlier it was mentioned that the matrix representation with 18 free parameters allows more interaction between different states and inputs as compared to the structure 5 RC model. As an example, it was shown that the values of a_6, b_5 and b_6 should be zero since the dynamics in the structure 5 model do not allow them. The distribution of the parameter values of these three matrix elements is shown in Figure . It is clear that in all cases the values of these matrix elements is close to zero thus minimizing the impact of T_{int}, Q_{sun}

and Q_h on the evolution of T_w. As a comparison, also the values of a_5 is shown, which shows a clearly different distribution and thus impact on the dynamics.
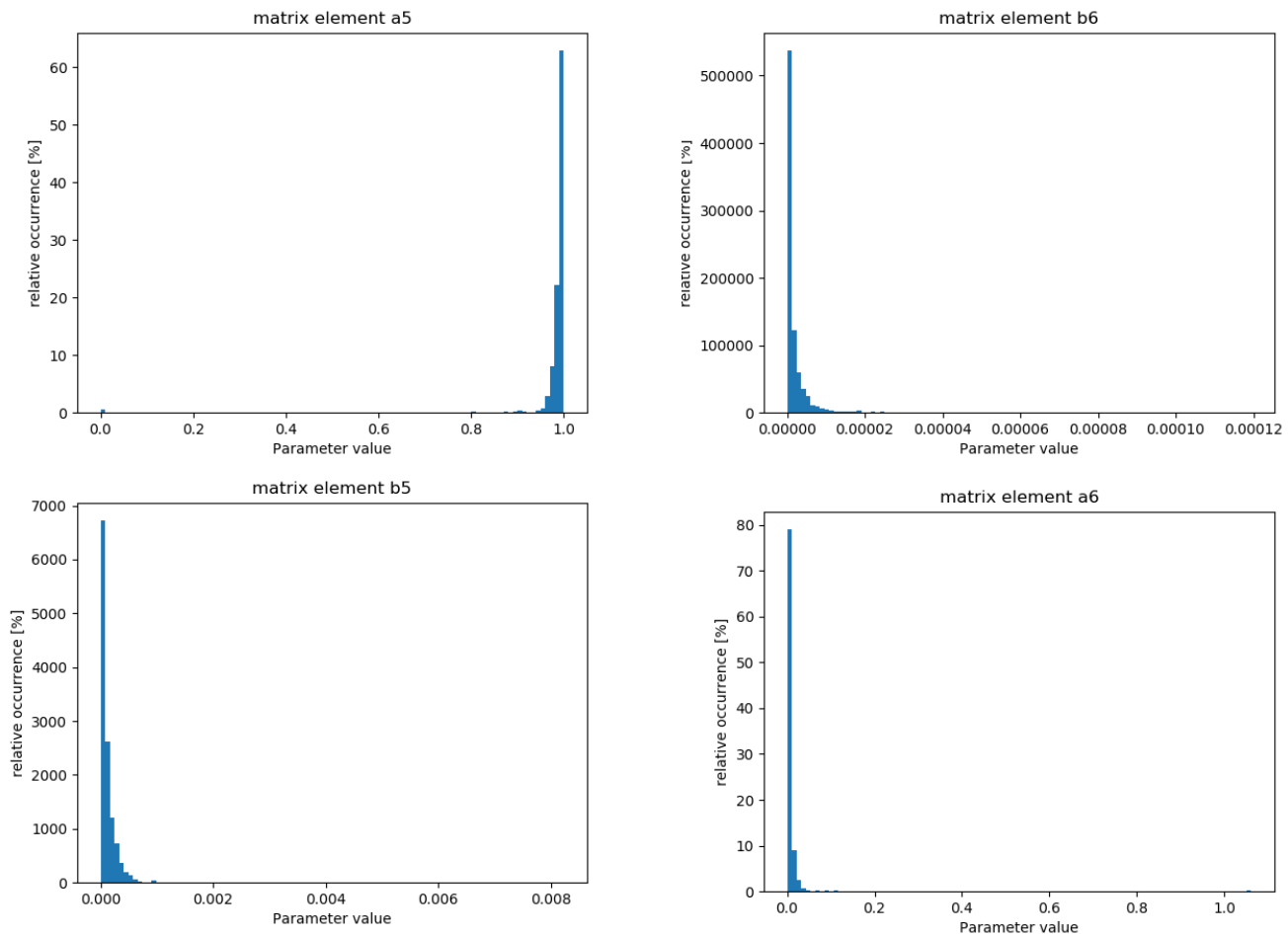
FIGURE 12 - THE DISTRIBUTION OF THE PARAMETER VALUES OF MATRIX ELEMENTS

2.3.4 SELECTED RESULTS

The three results below are randomly selected based on their RMSE values for training and validation. A "very good" model fit has the condition that the RMSE in training is less than 0.2 °C and less than 0.5 °C in validation. A "good" result has a training RMSE between 0.2 and 0.4°C and validation RMSE between 1 and 2 °C. Also, a "bad" result is shown, which has a training RMSE of more than 0.5 °C and validation RMSE of more than 3 °C.

Figure 13 shows the indoor temperature evolution that was obtained from the white-box model (blue) as well as the prediction by the trained grey-box model (blue) for training (left) and validation (right). The randomly selected case has the reference building code "BE-SFH-1991-2005-00", which is a single-family house in Belgium constructed between 1991 and 2005. Both phases show excellent agreement between the estimated and observed temperature. In total 44 reference building qualify as "very good".
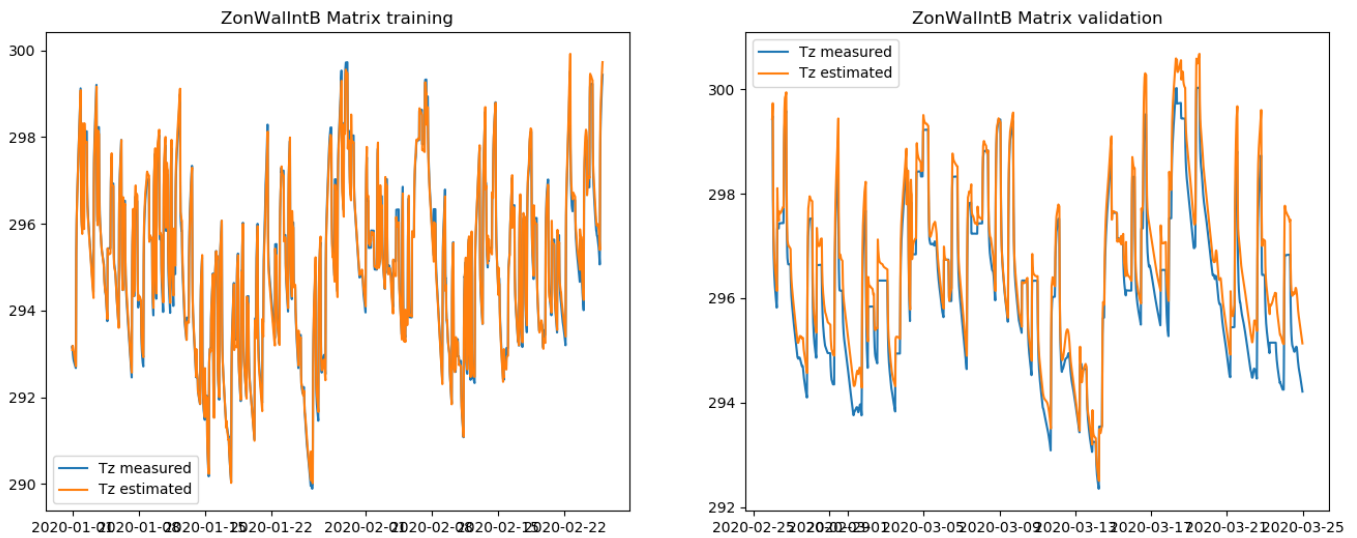
FIGURE 13 - RESULTS FOR BE-SFH-1991-2005-00, CLASSIFIED AS A VERY GOOD MODEL FIT

An example of a building classified as a good model fit is shown in Figure . In total, 687 reference models fall within the bounds defined above. The randomly selected building is CY-EDU-1945-1969-02, an educational building in Cyprus, constructed between 1945 and 1969. The agreement is still very good although some specific spikes cannot be accurately reproduced.
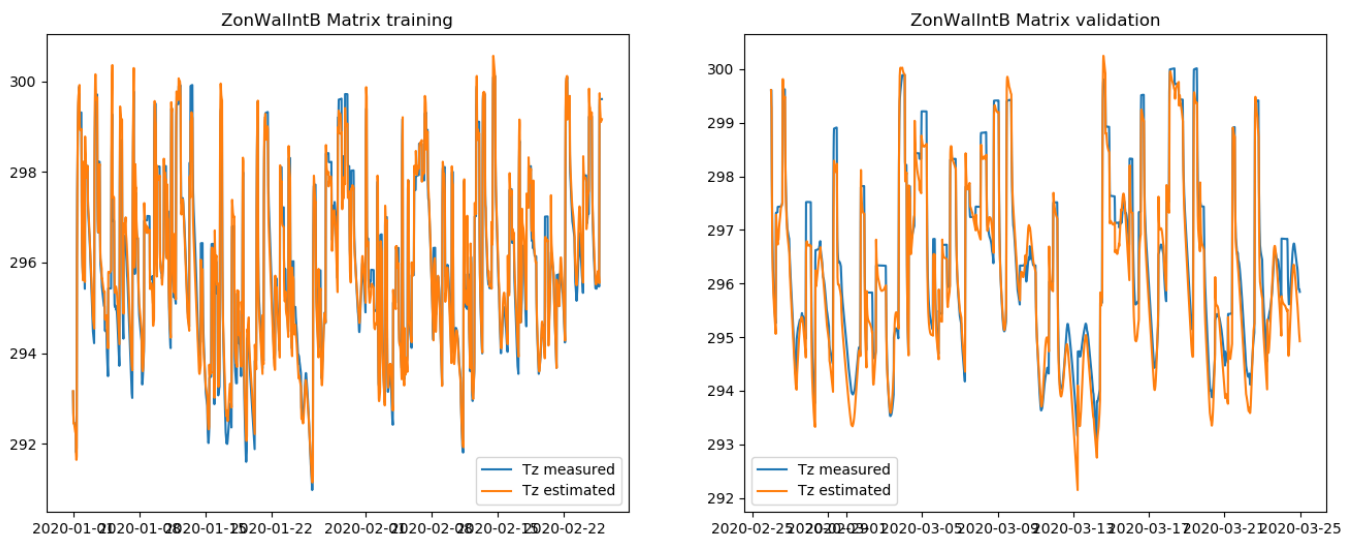
FIGURE 14 - RESULTS FOR CY-EDU-1945-1969-02, CLASSIFIED AS A GOOD MODEL FIT

Finally, Figure 15 shows the result for a what is considered a bad fit. Although the results during training are comparable to those considered to be good, a clear drift during the validation phase is seen. Only 6 reference buildings are classified as bad models, all of them are constructed before 1969 and located in either Malta or Croatia. A possible method to improve bad fits is to expand the training data period to include more dynamics in the training phase.
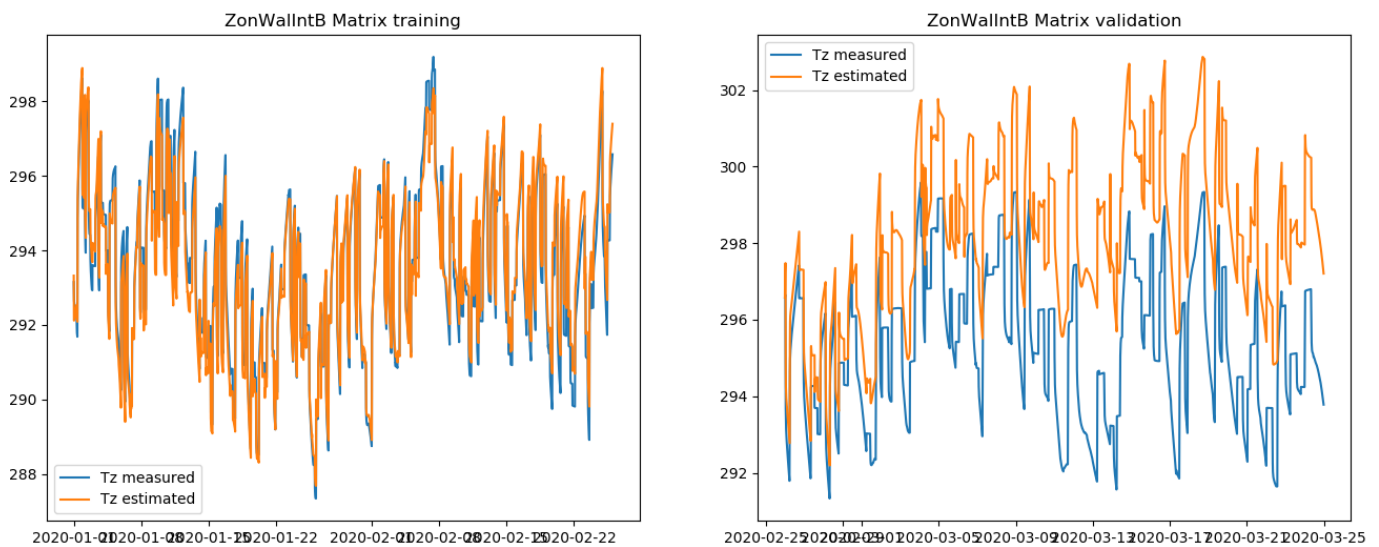
FIGURE 15 - RESULTS FOR MT-OFF-1850-1944-02, CLASSIFIED AS A GOOD MODEL FIT IN THE TRAINING PHASE

2.4 YEARLY ENERGY CONSUMPTION

An interesting metric is the yearly thermal energy consumption per m² of a building. One way to estimate this, is by running a second simulation over the course of an entire year, using a realistic indoor temperature set point regime for each reference building. Since this would take four times as long as the data generation step it was not feasible to do this within the project.

An alternative approach was therefore followed in which for each country a random building was simulated for a full year. By applying the rule of three it is possible to estimate the total yearly consumption of each reference building in that country as:

$$\frac{Yearly\ energy\ consumption_{full\ simulation,random\ buidling}}{Energy\ consumption_{training,random\ building}} = \frac{Yearly\ energy\ consumption_{reference}}{Energy\ consumption_{training,reference}}$$

These approximated yearly energy consumptions are divided by the total useful floor space and added to the database.

3. DATABASE CONTENT AND METHODOLOGY

3.1 DATABASE CONTENT

For different building stock segments of EU 27 Member States, the database provides:

- Information about the building stock segment's reference building
  - o Reference building classification data,
  - o Reference building geometry data,
  - o Characteristics of the reference building's envelope elements
    - Ground floors,
    - Walls,
    - Windows,
    - Roofs,
- Information on the building stock segment itself.

3.1.1 BUILDING STOCK SEGMENT

Building stock segment is a part of a national or regional building stock that can be represented and described by its respective reference building. Building stock segments are defined by the following parameters:

- Country,
- Building use,
- Construction period, and
- Specifics of construction methodology and/or materials used, if applicable.

For example, one of the building stock segments covered in the database consists of German hotels and restaurants built between 1970 and 1979 with walls made of concrete.

Building stock segments are defined by using the parameters listed above in case of both residential and non-residential buildings, which allows merging these into a single database.

3.1.2 REFERENCE BUILDINGS

Reference buildings presented in the database are individual buildings that should be used as representatives of their respective building stock segments. This means that analysis of such buildings will provide sufficient and reliable conclusions to represent their respective building stock segments.

For example, by analysing the energy needs of a reference building, one should (with or without additional analysis of results) be able to understand the energy needs of the building stock segment represented by the reference building.

One of the main features of the database is that residential reference buildings were simply taken from the TABULA[7] database, while non-residential reference buildings primarily came from Hotmaps[8] data source (for instance their size and composition of walls), both EU funded projects. More details about this, as well as about TABULA and Hotmaps, can be found in the remaining sections of the report.

3.2 METHODOLOGY OF THE DATABASE DEVELOPMENT

The database development includes two main tasks: First, defining the database structure and fields, and secondly populating the database. The database was developed through the series of steps:

1. Definition of potential data sources and their comparison with the database requirements

2. Final selection of the data sources

3. Transformation of data provided by data sources, which included other relevant inputs, such as information on properties of building construction materials

4. Defining and filling the data gaps

5. Merging residential and non-residential databases into a final single database

6. Adding the relevant grey-box model parameters.

Steps 1-4 were performed in parallel for residential and non-residential buildings, which resulted in two separate databases available at the end of the step 4. In the step 5 these two databases were merged into one. Step 6 is described in more detail in chapter 2.

Steps 3 and 4 – data transformation and filling the data gaps, were the most important steps of the database development. These aimed at simplifying the raw data available from the data sources, adjusting data to the needs of the database, and enabling a final merge of residential and non-residential databases.

---

[7] https://episcope.eu/building-typology/
[8] https://www.hotmaps-project.eu/

4. DATA SOURCES

All the work done in on the database is based on two publicly available data sources (data sets) delivered by two EU funded projects:

- TABULA[9], and

- Hotmaps[10].

4.1 TABULA DATABASE

4.1.1 BUILDING STOCK SEGMENTS AND DATA

The TABULA project, funded by the Intelligent Energy Europe Programme, developed residential building typologies for residential building stock segments in 17 European countries. TABULA also provides a set of residential reference buildings representing their specific building segment stocks. These reference buildings are imagined as typical representatives of their building segments. These reference buildings were the most important input for the residential part of the database.

TABULA defines building stock segments based on the following parameters:

- Country[11],

- Building use,

- Construction period, and

- Additional parameters, where applicable.[12]

Reference buildings in TABULA are imagined as typical representatives of their building segments. For reference buildings TABULA provides different sets of data describing them in detail, such as:

- Building size
  - o Useful floor area,
  - o Area of walls, windows, roof, and ground floor,
- For each building envelope element (ground floor, walls, windows, roofs)
  - o U-value, or thermal transmittance, as the rate of heat transfer through a structure per temperature difference,
  - o Existing insulation thickness,
  - o Construction components (structural or insulation layer) and materials (e.g., brick or

---

[9] https://episcope.eu/building-typology/
[10] https://www.hotmaps-project.eu/
[11] In some of the countries TABULA provides also regional data. However, in task 4.1 only national data are considered.
[12] https://webtool.building-typology.eu/#bm

concrete).

To enhance information exchange, TABULA established the TABULA WebTool13 where detailed information on the reference buildings is available. On top of the TABULA WebTool, a set of reports on national building typologies, as well as various excel files including raw data, are publicly available for download. More about TABULA sources used for developing the database can be found in chapter 5. [1].

4.1.2 THE SPECIFICS OF TABULA DATA

Covered countries

Data available from TABULA covers the following 17 EU countries: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Denmark, Spain, France, Greece, Hungary, Ireland, Italy, The Netherlands, Poland, Sweden, and Slovenia.

Construction periods

In both TABULA and Hotmaps, construction years are organised into different groups. While these groups are standardised in Hotmaps, in TABULA grouping of construction years differs among countries analysed. As a rule, for buildings built before the mid-1900s, TABULA provides larger groups of construction years, while it provides smaller groups (usually decades) for those built after that.

An example is a set of nine construction periods for Austria: before 1918, 1919-1944, 1945-1960, 1961-1970, 1971-1980, 1981-1990, 1991-2000, 2001-2009, 2010 and after.

Building use

TABULA recognizes the following four types of residential building use14:

- Single family houses,

- Terraced houses,

- Multi-family houses,

- Apartment buildings.

Definitions of these are different among analysed countries and may depend on multiple factors such as useful floor size, number of floors, and similar.

Building size

Regarding the reference buildings' size, TABULA provides two important sets of data:

- Useful floor size, as a sum of areas covered with all building floors, and

- Size of the building envelope, by providing area of walls, windows, doors, and roofs.

---

13 https://episcope.eu/building-typology/webtool/

14 What is here referred to as building use is usually denoted as building type. However, building use is considered as more appropriate for the purpose of the Task 4.1 and the database development.

Building envelope properties

In general, TABULA provides information on more than one type of building envelope element. For example, the building's wall may contain two wall types: bricks and concrete. TABULA's reference buildings allow at most two types of floors, windows, or roofs, and three types of walls. There is a single door type allowed per reference building.

4.2 HOTMAPS DATABASE

4.2.1 BUILDING STOCK SEGMENTS AND DATA

The EU-funded Hotmaps project is a toolbox to support public authorities and other stakeholders in strategic planning on local, regional, and national levels, and in line with EU policies.

As one of its outputs, Hotmaps provides a publicly available building stock dataset on EU28 countries covering both the residential and non-residential segments. However, in its consistency and level of detail, data for the residential building stock significantly differs from data for non-residential buildings. For example, information concerning construction materials and methodology is provided for all individual residential typologies in each country, while this information is provided in very general terms for non-residential typologies.

Hotmaps building stock segments are defined based on the following parameters:

- Country,
- Building use,
- Construction period.

For such building stock segments, Hotmaps provides different sets of data. Data relevant for this database are:

- Number of buildings,
- Constructed floor area,
- Useful energy demand for space heating,
- For each building envelope element (ground floor, walls, windows, roofs):
  o U-values,
  o For residential buildings: construction methodology and materials used,
  o For non-residential buildings: construction materials used.

4.2.2 SPECIFICS OF THE HOTMAPS DATA

Construction periods

Contrary to TABULA, Hotmaps applies a single set of eight construction time-periods, unique for all countries and building uses. These construction periods are before 1944, 1945-1969, 1970-1979, 1980-1989, 1990-1999, 2000-2010, 2001-2009, 2011 and after.

Building use

Building use in Hotmaps is defined at two levels:

  — 1st level: residential vs. non-residential (or service) use,

  — 2nd level: more detailed use within each of these two groups.

As for residential buildings, Hotmaps defines three building purposes:

  — Single family houses,

  — Multi-family houses, and

  — Apartment blocks.

Hotmaps defines apartment blocks as high-rise buildings that contain several dwellings and have more than four floors, while single- and multi-family houses are not defined further.

Within non-residential buildings, Hotmaps provides six different types of building use:

  — Offices,

  — Trade,

  — Education,

  — Health,

  — Hotels and restaurants,

  — Other non-residential buildings including buildings such as warehouses, transportation and garage buildings, military barracks, and agricultural buildings. [2]

Useful energy demand

As explained in the Hotmaps guideline [2], useful energy demand represents the net energy required to cover the different energy needs of a building, such as space heating, space cooling, or domestic hot water preparation.

We should note that useful energy demand differs from the final energy consumption [2].

For developing the database, due to limitations in Hotmaps data, only useful energy demand for space heating is considered. This data is expressed as amount of energy needed to satisfy annual heating needs, with the unit of TWh/year. Useful energy demand should be distinguished from energy demand per area – kWh/m2/year, that is frequently used in building related databases.

4.3 REASON FOR SELECTING TABULA AND HOTMAPS AS DATA SOURCES

Despite the availability of other sources of data on EU building stock, TABULA and Hotmaps were selected mostly due to the following reasons:

- Size of the data sets,
- Level of details provided,
- Possibility for filling the gaps and data merge.

4.3.1 SIZE OF DATA SETS

Both Hotmaps and TABULA cover enough countries and provide enough data for proper database development. Due to such large size of TABULA and Hotmaps it was possible to avoid introducing additional data sources, which would bring more of complexity than of accuracy.

Even in case of TABULA, where some EU countries are missing, the number of available countries is large enough.

4.3.2 LEVEL OF DETAILS PROVIDED

Both sources provide significant and sufficient level of detail, which is especially true given the requirements needed to introduce flexibility by using the grey-box model. This allowed focus on these sources only without the need to analyse and select other sources.

4.3.3 POSSIBILITY FOR FILLING THE DATA GAPS AND DATA MERGE

As extremely important factor, one should notice that finding highly detailed data in few, rather than many sources, allows easy and precise identification and filling of the data gaps. As an illustration, an inconsistency related to number of buildings, for example, would be noticed and fixed much easier when dealing with few sources only, and when this inconsistency is most likely the same within all the building types.

In addition, selecting TABULA and Hotmaps made data merge possible, which was another enabling factor for covering the data gaps. As explained in 4.1.1 and 4.2.1, TABULA and Hotmaps both follow the same logic in defining the building stock segments, thanks to which it was possible to use assumptions from one source for filling the data gaps in the other. An example here can be using residential insulation thickness, provided by TABULA, as an assumption for insulation thickness missing in non-residential buildings in Hotmaps.

4.3.4 NATIONAL DATABASES

Throughout the database development, special attention was paid to the national databases as potential buildings data sources. Since for Portugal, as one of the project pilot countries, no data on residential reference buildings was available in TABULA and Hotmaps, a set of information on Portuguese buildings was additionally considered.

Portuguese building data sources provided by the Portuguese national energy agency (ADENE) include databases of EPC available at the national level. Although these national databases showed a respectable

level of detail, a thorough review revealed that:

− Such databases do not include information on reference buildings. Instead, aggregated average information on a large number of individual buildings, subject to energy performance certification, is provided.

− Although with significant reduction in accuracy, the best proxy of a reference building available from these databases would depend on the transformation of individual building data. For example, the floor area of the reference building could be approximated as an average of individual building floor areas. However, even in this case, Portuguese national databases could not provide values for all the fields included in the database.

− It is of utmost importance that reference building proxies, obtained after transforming individual building data, included highly inconsistent building characteristics. For instance, in many cases, provided U-values of walls (but also of other building envelope elements) significantly differed from what could be calculated from provided thicknesses and materials of the corresponding wall layers.

Due to these important reasons, and after a thorough evaluation by the project participants concerned, it was finally agreed that the retrieved Portuguese national data do not meet, at this stage, the minimum requirements to be included in the database.

Moreover, it is reasonable to suspect that similar problems may be found in other national databases as well. This, in addition to the reasons presented in chapters 4.3.1 to 4.3.4, led to the conclusion that TABULA and Hotmaps should be kept as the only sources for the database development.

4.3.5 OTHER DATA SOURCES

Other available data sources dealing with building stock data, such as Building Stock Observatory (BSO), EUCalc, or ENTRANZE projects, were not selected due to their lower compatibility with the database requirements than what can be achieved by using TABULA and Hotmaps. For example, it was noticed that all or some of these sources usually do not provide satisfactory level of detail (lack of information about construction/insulation materials or building size), do not deal with reference buildings, or simply do not provide data in a way that would ensure data merge from different sources. All of these reasons, in addition to the high relevance of TABULA and Hotmaps explained above, was enough to exclude these potential data sources from the database development.

5. DATA COLLECTION

5.1 DOWNLOADABLE DATA FILES

Full set of Hotmaps and TABULA data were downloaded from their respective websites. Data was available through downloadable .xls and .csv files. At the moment of writing this report, the files were available via the following links:

- For TABULA database:
  - o [tabula-calculator.xlsx,](#)
  - o [tabula-values.xlsx.](#)
- For Hotmaps database:
  - o [building_stock.csv](#)

5.2 ADDITIONAL DATA COLLECTION FROM THE TABULA WEBSITE

To accomplish and to have better understanding of the data provided in the files mentioned above, composition of the building envelope elements in residential buildings – floors, walls, windows, and roofs – was read from the TABULA website. Data collected in this way was added to the database built from the files listed under 5.1.

Screenshot of a typical TABULA webpage used for this purpose is presented in Figure 16.
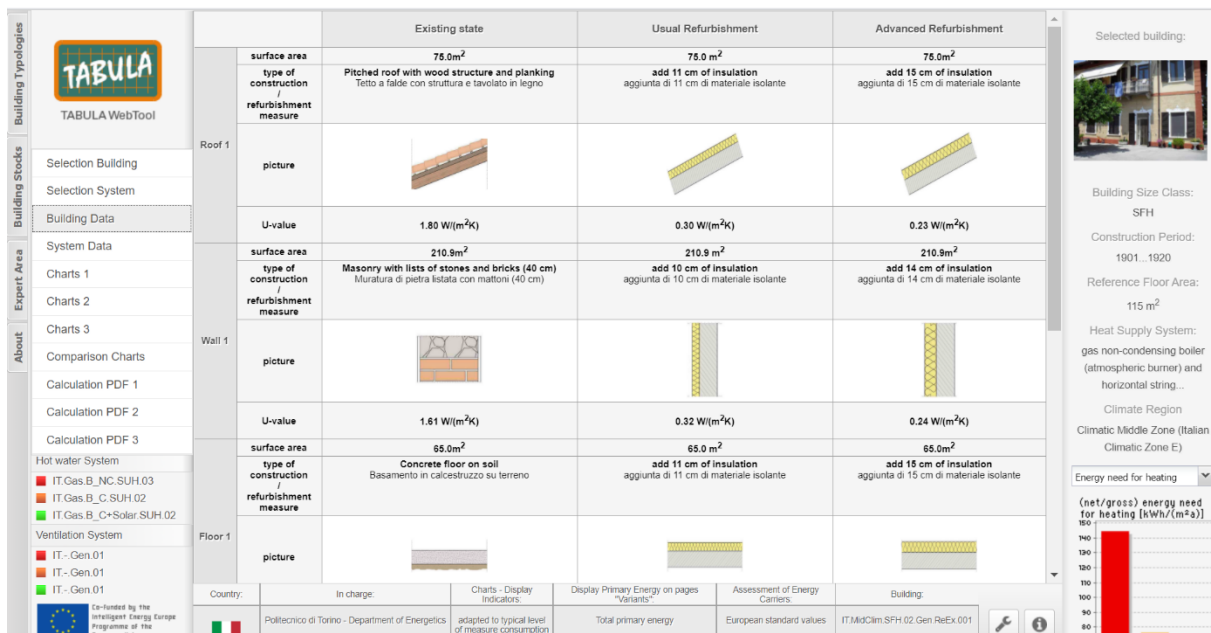
FIGURE 16 - BUILDING DATA FROM TABULA WEBPAGE

6. DATA TRANSFORMATIONS

Building stock segments of the database are organised by building use, construction period, and, where applicable, construction methodology and/or materials. In addition, each building stock segment shown in the database is represented by a single reference building.

Because of this, the first step was to define reference buildings and their code names. A unique reference building code is therefore used to define both the reference building and represent its building stock segment.

Residential reference buildings cover the following:

- 27 EU countries,

- Building use: single-family buildings, multi-family buildings, and apartment blocks,

- Construction periods that differ among covered countries.

Different combinations of these parameters defined 453 residential reference buildings. An example of a reference building is therefore a multifamily building in Germany built in the 1970s.

Non-residential reference buildings are defined for:

- 27 EU countries,

- Use: office, education, health, hotels and restaurants, other non-residential buildings,

- Construction periods: before 1944, 1945-1969, 1970-1979, 1980-1989, 1990-1999, 2000-2010, 2001-2009, 2011 and after,

- Specifics of construction methodology and/or materials used, where applicable.

By combining these parameters 1,792 non-residential reference buildings were defined. An example of such reference building is educational building located in Spain built in the 1970s and with brick walls.

Combining the residential and non-residential reference buildings yielded the 2245 buildings that were included in the database.

Once this was done, the following steps were applied to residential and non-residential buildings.

6.1 RESIDENTIAL BUILDINGS

In a first step all residential reference buildings and their code names were identified and used to define both the reference building and represent its building stock segment. Once this was done, the following steps were applied to residential buildings:

1. Since each of the residential reference buildings referred to more than one TABULA reference buildings, preliminary selection of TABULA reference buildings was required:

    a. TABULA reference buildings described as "Generic" were selected to ensure coverage of the total building stock, since TABULA reference buildings contain certain subtypes

not targeting the whole building stock of their respective countries.[15]

b. Wherever possible, reference buildings defined as "National" were selected to ensure covering the largest possible share of the territory in each country. In case where a country is split into different regions (e.g., Greece), the maximum number of regions was selected.

The filtering explained here was necessary to exclude TABULA reference buildings not relevant for further analysis.

2. Within each TABULA reference building, the following transformations were implemented:

a. If a reference building's envelope element consisted of more than one type, e.g., a wall consisting of brick wall and a concrete wall, these were merged. This had the following implications:

i. The total area of the building element is calculated as a sum of its different parts. In the example above, total wall area is the sum of areas made of bricks and concrete.

ii. The U-value of the building element is an area weighted average of different element types' U-values. As for the example above, wall U-value is the weighted average of U-values of brick wall and concrete wall, where the weights are based on the share of wall's areas in the total wall area.

iii. The insulation thickness follows the same approach applied to U-values: within each building element, insulation thickness is an area weighted average of insulation thicknesses corresponding to different element types.

3. For each reference building, items listed below were defined as simple averages[16] of their respective values found in TABULA reference buildings. For example, useful floor area of a reference building is an average of useful floor areas defined for TABULA reference buildings.

For each reference building, the following data were defined:

a. Useful floor area,

b. Number of reference buildings,

c. Number of floors, and

d. Area of building envelope elements: ground floor, walls, windows, doors, and roof.

4. For each reference building's ground floor, wall, and roof the following were determined:

a. Construction material type: concrete, brick, wood, or stone, read from the relevant

---

[15] https://webtool.building-typology.eu/#bm

[16] Weighted averages were not possible due to very limited information on number of TABULA reference buildings in the building stock segments they represent.

TABULA reference buildings in the TABULA WebTool.

b. U-value, as an average of respective values from TABULA reference buildings,

c. Insulation thickness, as an average of respective values from TABULA reference buildings.

5. Using the results of the step 4, for each reference building's ground floor, wall, and roof, construction material thickness and construction and insulation materials' thermal conductivity was determined.

This was possible due to the following dependence between the building element's U-value, properties of the element's layers, and element's thermal interaction with the surrounding air:

$$U = \frac{1}{R_{si} + \frac{d_{cons}}{\lambda_{cons}} + \frac{d_{ins}}{\lambda_{ins}} + R_{se}},$$

<div align="center">

EQUATION 1 – U-VALUE OF A BUILDING ENVELOPE ELEMENT
</div>

Elements of this equation are:

− U: U-value of the building envelope element,

− R_{si}, R_{se}: Resistance to heat convection from/to internal and external surface of the element, respectively,

− d_{cons}, d_{ins}: Thickness of construction and insulation layer, respectively, and

− λ_{cons}, λ_{ins}: Thermal conductivity of construction and insulation material, respectively.

For given U-value, R_{si}, R_{se}[17], and d_{ins}, the equation shown above can be resolved by implementing following limitations:

− λ_{cons} and λ_{ins} must stay within the range of values expected for the construction material specified under step 4 of this list, and the most common insulation materials, respectively, and

− d_{cons} must stay within the range of construction material thicknesses typical for the building element it refers to.

In other words, by varying λ_{cons} and λ_{ins} within their allowed ranges, it was possible to find the construction material thickness, d_{cons}, that stays within its expected range of values, which resolved the equation and resulted in combination of construction material thickness and thermal conductivities of construction and insulation materials.

Thermal conductivities determined in this way were then used for specifying precise construction and insulation material in ground floor, wall, and roof of each reference building.

---

[17] R_{si} and R_{se} are considered constant for specific building element (not dependent on construction/insulation material).

For example, if in step 4 of this list it was determined that a wall was made of concrete, now it was possible to specify concrete type used, such as cast concrete 2000. This resulted in determining other thermal properties of materials found in ground floors, walls, and roofs, such as density and heat capacity of the construction and insulation layers.

Inputs used for this step are summarized in Table 6 - Properties of materials, Table 7 - Thermal resistances of surfaces, and Table 5 - Allowed thicknesses of construction layers.

6. For each reference building's windows, data listed below was read from the TABULA WebTool:

    a. Glazing type: single, double, or triple glazing,

    b. Frame material type: wood, plastic, steel, or aluminium,

    c. Existence of low-E layer,

    d. Gas used for filling the glazing gap: air, or argon, if applicable,

    e. Window U-value, as an average of TABULA reference building's values.

7. For each reference building, keeping in mind its U-value from the step 6 of this list, window properties were found by varying:

    a. Within each glazing type – glazing gap,

    b. Within the frame material type – more detailed frame material and thicknesses of the frame,

    c. In case of existing low-E layer – its emissivity.

For example, combining the window U-value with window properties read from TABULA: double glazed, air filled, coated window with wooden frame, allowed more precise characteristics of the same window: double glazed window, with 6 mm wide air-filled gap and low-E layer with emissivity of 0.05, and 30 mm thick wooden frame.

Inputs used for this step are summarized in Table 9 - Glazing properties and Table 10 - Window frame properties.

8. Building stock segment represented by each reference building was described by providing the following data from the Hotmaps database:

    a. Useful floor area,

    b. Number of buildings (not equal to the number of reference buildings),

    c. Useful annual energy demand for heating.

6.2 NON-RESIDENTIAL BUILDINGS

In a first step all non-residential reference buildings and their code names were identified and used to define both the reference building and represent its building stock segment. Once this was done, the following steps were applied to non-residential buildings:

1. For each non-residential reference building, by using Hotmaps data, the following information was defined:

   a. Useful floor area, after dividing total useful floor area and number of buildings in the building stock segment represented by the reference building,

   b. Number of reference buildings that is equal to the number of buildings in the building stock segment.

2. For different uses of non-residential buildings, number of floors, floor-to-floor height, and window to floor ratio, was assumed. As a result, for each non-residential reference building it was possible to define areas of

   a. Ground floor,

   b. Walls,

   c. Windows, and

   d. Roof.[18]

   Inputs for this step are summarized in Table 8 - Reference building size.

3. For each reference building's ground floor, wall, and roof the following information was determined:

   a. Construction material type: concrete, brick, wood, or stone, as read from the Hotmaps database,

   b. U-value, read from the Hotmaps database,

   c. Insulation thickness, as an average of insulation thickness found in the residential part of the AmBIENCe database, in the reference buildings of the same country and construction period.

   By following the procedure and inputs applied to residential buildings and described under 5 of the chapter 6.1, construction material thickness, as well as exact construction and insulation material in ground floor, wall, and roof of each reference building was determined. In addition, thermal conductivity, density, and heat capacity of these were resolved.

4. For each non-residential reference building, window properties were established by following the procedure and inputs explained under 6 and 7 of the chapter 6.1.

---

[18] Door area of non-residential buildings is not available.

5. A building stock segment represented by each reference building was described by providing the following data from the Hotmaps database:

    a. Useful floor area,

    b. Number of buildings,

    c. Useful annual energy demand for heating.

The number of buildings in the non-residential building stock (and its segments) is equal to the number of non-residential reference buildings.

7. DATABASE ASSUMPTIONS AND FILLED GAPS

The most important data gaps identified, and the assumptions adopted to resolve them, are explained in this chapter. Wherever possible, the data gaps and assumptions are linked to the steps listed in chapter 6.

1. The most important data gap is the fact that TABULA provides reference buildings only for a limited number or EU countries. To resolve this, the following mapping is proposed for the missing countries, Table 2.

TABLE 2 - MAPPING OF COUNTRIES FOR COMPLETING TABULA REFERENCE BUILDINGS

Country missing in TABULA	Existing TABULA country used to replace the missing one
Estonia	Sweden
Finland	Sweden
Croatia	Slovenia
Lithuania	Sweden
Luxembourg	Netherlands
Latvia	Sweden
Malta	Spain
Portugal	Spain
Romania	Bulgaria
Slovakia	Czech Republic

The mapping was established based on different criteria, such as:

– Geographical proximity,

– Similar weather profiles,

– Country size,

– The same or similar building codes applicable in the past[19].

One should note that, since it does not refer to reference buildings, Hotmaps database was not suitable for replacing missing TABULA countries. Using Hotmaps for this purpose would result in inconsistent level of accuracy among residential reference buildings, which was avoided by the mapping presented above.

2. Another crucial data gap is linked to any building envelope element where adequate

---

[19] As in case of Slovakia and Czech Republic, or Croatia and Slovenia.

construction layer thickness could not be found, following the procedure explained under 5 of chapter 6.1.

In this case, the building envelope element, was replaced by the same element, where the construction material thickness was determined, found in a building of the same country and use, and built during the closest construction period. As a rule, if there were two elements to choose from, equally close in terms of construction period[20], the element from the older building was selected.

In most instances, the nearest eligible element was found within the same country, building use, and material combination (where applicable), which helped in creating a consistent database.

In an extreme case, when the nearest element could not be found within the same country, an element from another country was used. This happened only in few cases listed in Table 3.

TABLE 3 - MAPPING OF UNRESOLVED BUILDING ENVELOPE ELEMENTS

Missing building envelope element	Replaced by
Floors and roofs in non-residential buildings in Estonia	Floors and roofs in non-residential buildings in Poland
Floors and roofs in non-residential buildings in Finland	
Floors and roofs in non-residential buildings in Sweden	
Floors and walls in residential buildings in Sweden	Floors and walls in residential buildings in Poland
Roofs in residential buildings in Poland	Roofs in residential buildings in Germany

3. Each building element of the reference building is assumed to consist of two layers only:

    a. Structural element, made of one of the following materials:

        i. Brick,

        ii. Concrete,

        iii. Stone,

        iv. Wood.

---

[20] In practice this implies two available elements: one in the succeeding and the other in the preceding construction period.

    b.   Insulation element.

Such an approach is used since, at the same time, it:

- Provides enough information about the thermal mass of the building envelope element, and

- Allows specification of layers' thicknesses and other parameters.

4. As explained under 4.1.2 and 4.2.2, for defining its residential building stock segments, TABULA includes one building use type more than Hotmaps: terraced houses (TH).

Merging two databases assumed that reference buildings for single-family houses include terraced houses as well. In other words, reference buildings for terraced houses are not provided in the database but covered by the single-family reference buildings.

5. Hotmaps does not provide information about insulation thicknesses in non-residential buildings, which prevents definition of the building envelope layers.

To resolve this, it was assumed that insulation thickness in a building envelope element of a non-residential reference building, is equal to the average insulation thickness found in residential reference buildings built in the same country and during the same or similar construction period.

For EU27 countries not covered by the TABULA database (i.e., not found in the list of residential reference buildings), the mapping shown in Table 2 was used. For instance, this resulted in insulation thickness in non-residential buildings in Estonia being equal to thickness of the insulation layer found in Swedish residential buildings.

6. Due to lack of data, non-residential reference buildings are determined from the data on the non-residential building stock: useful floor area of non-residential reference building is equal to average useful floor area of the building stock segment, while number of reference buildings is equal to the number of buildings in the same segment.

7. In some of the countries covered by TABULA, reference buildings refer to lower number of building use types than what is provided by Hotmaps database, such as in case of Sweden where TABULA provides only single-family (SFH) and multi-family (MFH) reference buildings.

In these cases, quantification of the residential building stock segments required adding the missing building uses to the existing ones. The mapping used for this purpose is shown in Table 4.

TABLE 4 - MAPPING OF MISSING BUILDING USES IN TABULA

Country	Missing residential building stock segment	Included in residential building stock segment
Sweden	ABL	MFH
Cyprus	ABL	MFH
Denmark	MFH	SFH
Greece	ABL	MFH
Ireland	MFH	SFH
Hungary	ABL	MFH

SFH – single-family buildings, MFH – multi-family buildings, ABL – apartment blocks

8. Due to the lack of data and relatively low impact on energy performance of reference buildings, the door is not considered as part of the building envelope. This means that building envelopes of reference buildings consist of ground floor, walls, windows, and roof.

9. The possible range of construction layer thicknesses was a starting point for determining construction and insulation material thickness in reference building's ground floor, wall, and roof. For this purpose, the following list of possible thicknesses of construction layers was assumed, Table 5.

TABLE 5 - ALLOWED THICKNESSES OF CONSTRUCTION LAYERS

LAYER	MINIMUM ALLOWED THICKNESS (m)	MAXIMUM ALLOWED THICKNESS (m)
Floor	0.10	1.50
Wall	0.10	1.00
Roof	0.10	1.00

10. Calculation of a window's U-value depends on the U-values of its glazing and frame, as well as on the share of glazing and frame in the total window area. For this purpose, window frame is assumed to occupy 20%, and the glazing 80% of the window area.

8. DATABASE STRUCTURE AND FIELDS

The database consists of 64 fields that are either taken from the source databases or calculated in the way explained in this report. The fields are organised into five major sections:

1. Reference building classification, used to describe the reference building and the building stock segment it represents,

2. Reference building geometry, that explains geometry and size of each reference building,

3. Reference building envelope elements, that for each building element of a reference building list the element's layers and their main physical properties, and

4. Building stock segment, explaining the size and energy demand of the building stock segments represented by reference buildings.

5. Grey-box model related fields, providing the name of the most suitable grey-box model structure and parameters obtained from the training algorithm. Metrics from training and validation phase include RMSE, MAPE and variance. Yearly energy use for space heating (per m²) is also given.

8.1 REFERENCE BUILDING CLASSIFICATION

The reference building classification describes the reference building and the building stock segment it represents and is composed out of 6 fields.

Field 1.   REFERENCE BUILDING CODE

The reference building code has the following structure – CC-UUU-yyyy-YYYY-MM, where specific characters represent:

- 1st and 2nd – CC: the country code,

- 4th to 6th – UUU: the building use, with possible values,

- 8th to 11th – yyyy: the lower year of the construction period,

- 13th to 16th – YYYY: the upper year of the construction period, and

- 18th to 19th – MM: used to indicate when a building stock segment of specific country, building use, and construction period, consists of buildings built with different construction methodologies/materials. More information about these characters is provided under field 6.

For example, for an Italian single-family building built between 1961 and 1975, the code IT-SFH-1970-1979-00 defines both its building stock segment and the reference building. For the sake of database convenience, elements of the reference building code are presented separately in the fields 2 to 6.

Field 2.   COUNTRY CODE

Each EU 27 country has been assigned a two-letter country code.[21]

Field 3.   REFERENCE BUILDING USE CODE

Use of each reference building is determined by this field that contains the following values:

- SFH – single family houses,
- MFH – multi-family houses,
- ABL – apartment blocks,
- OFF – offices,
- TRA – trade buildings,
- EDU – education buildings,
- HEA – health buildings,
- HOR – hotels and restaurants,
- OTH – other non-residential buildings.

Field 4.   REFERENCE BUILDING CONSTRUCTION YEAR LOW

Starting year of the period when the building was built.

Field 5.   REFERENCE BUILDING CONSTRUCTION YEAR HIGH

Ending year of the period when the building was built. One construction period includes years from the fields 4 and 5.

Field 6.   REFERENCE BUILDING MATERIAL COMBINATION CODE

This code is used to distinguish two or more construction methodologies and/or materials when these exist within a reference building defined by country, use, and construction period. In case of value 00 there is a single construction methodology and/or material used, otherwise numbers 01 and 02 are used[22].

8.2 REFERENCE BUILDING GEOMETRY

The reference building geometry explains the geometry and size of each reference building and if formed by the following fields:

Field 7.   REFERENCE BUILDING USEFUL FLOOR AREA (m2)

Useful area of the reference building expressed in m^2. This can be considered as conditioned (heated,

---

[21] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Country_codes
[22] There is maximum of 2 different construction methodologies/materials used in a reference building defined by country, use, and construction period. These are denoted with 01 and 02.

cooled, ventilated) and used area of the reference building.

Field 8.   NUMBER OF REFERENCE BUILDINGS IN THE BUILDING STOCK SEGMENT

The number of reference buildings representing the building stock segment defined in the fields 2 to 6.

Field 9.   NUMBER OF REFERENCE BUILDING STOREYS

Number of storeys in the reference building.

Field 10.   REFERENCE BUILDING GROUND FLOOR AREA (m2)

Area of the reference building's floor in contact with the ground, expressed in m^2.

Field 11.   REFERENCE BUILDING WALL AREA (m2)

Total wall area of the reference building, expressed in m^2.

Field 12.   REFERENCE BUILDING WINDOW AREA (m2)

Total window area of the reference building (not included in wall area) and expressed in m^2.

Field 13.   REFERENCE BUILDING ROOF AREA (m2)

Total roof area of the reference building, expressed in m^2.

8.3 REFERENCE BUILDING ENVELOPE

Reference building envelope elements indicate for each building element of a reference building list the element's layers and their main physical properties in the following fields:

Field 14.   REFERENCE BUILDING FLOOR MATERIAL

Material used for the structural element of the reference building's floor.

Field 15.   REFERENCE BUILDING FLOOR MATERIAL THICKNESS (m)

Thickness of the structural element included in the reference building's floor, expressed in meters.

Field 16.   REFERENCE BUILDING FLOOR MATERIAL THERMAL CONDUCTIVITY (W/m/K)

Thermal conductivity of the reference building floor's structural material, expressed in W/m/K.

Field 17.   REFERENCE BUILDING FLOOR MATERIAL DENSITY (kg/m3)

Density of the reference building floor's structural material, expressed in kg/m^3.

Field 18.   REFERENCE BUILDING FLOOR MATERIAL SPECIFIC HEAT CAPACITY (J/kg/K)

Specific heat capacity of the reference building floor's structural material, expressed in J/kg/K.

Field 19.   REFERENCE BUILDING FLOOR INSULATION MATERIAL

Material used for the insulation element of the reference building's floor.

Field 20. REFERENCE BUILDING FLOOR INSULATION MATERIAL THICKNESS (m)

Thickness of the insulation element included in the reference building's floor, expressed in meters.

Field 21. REFERENCE BUILDING FLOOR INSULATION MATERIAL THERMAL CONDUCTIVITY (W/m/K)

Thermal conductivity of the reference building floor's insulation material, expressed in W/m/K.

Field 22. REFERENCE BUILDING FLOOR INSULATION MATERIAL DENSITY (kg/m3)

Density of the reference building floor's insulation material, expressed in kg/m^3.

Field 23. REFERENCE BUILDING FLOOR INSULATION MATERIAL SPECIFIC HEAT CAPACITY (J/kg/K)

Specific heat capacity of the reference building floor's insulation material, expressed in J/kg/K.

Field 24. REFERENCE BUILDING FLOOR U-VALUE (W/m2/K)

U-value of the reference building's floor, expressed in $W/m^2/K$.

Field 25. REFERENCE BUILDING WALL MATERIAL

Material used for the structural element of the reference building's wall.

Field 26. REFERENCE BUILDING WALL MATERIAL THICKNESS (m)

Thickness of the structural element included in the reference building's wall, expressed in meters.

Field 27. REFERENCE BUILDING WALL MATERIAL THERMAL CONDUCTIVITY (W/m/K)

Thermal conductivity of the reference building wall's structural material, expressed in W/m/K.

Field 28. REFERENCE BUILDING WALL MATERIAL DENSITY (kg/m3)

Density of the reference building wall's structural material, expressed in kg/m^3.

Field 29. REFERENCE BUILDING WALL MATERIAL SPECIFIC HEAT CAPACITY (J/kg/K)

Specific heat capacity of the reference building wall's structural material, expressed in J/kg/K.

Field 30. REFERENCE BUILDING WALL INSULATION MATERIAL

Material used for the insulation element of the reference building's wall.

Field 31. REFERENCE BUILDING WALL INSULATION MATERIAL THICKNESS (m)

Thickness of the insulation element included in the reference building's wall, expressed in meters.

Field 32. REFERENCE BUILDING WALL INSULATION MATERIAL THERMAL CONDUCTIVITY (W/m/K)

Thermal conductivity of the reference building wall's insulation material, expressed in W/m/K.

Field 33. REFERENCE BUILDING WALL INSULATION MATERIAL – DENSITY (kg/m3)

Density of the reference building wall's insulation material, expressed in kg/m^3.

Field 34. **REFERENCE BUILDING WALL INSULATION MATERIAL SPECIFIC HEAT CAPACITY (J/kg/K)**

Specific heat capacity of the reference building wall's insulation material, expressed in J/kg/K.

Field 35. **REFERENCE BUILDING WALL U-VALUE (W/m2/K)**

U-value of the reference building's wall, expressed in $W/m^2/K$.

Field 36. **REFERENCE BUILDING WINDOW TYPE**

Type of the window used in the reference building. This code contains information about window glazing type, frame material, low-E layer, and gas used for filling the window. For the sake of convenience, information about each of these is presented in the fields 37 till 40.

Field 37. **REFERENCE BUILDING WINDOW GLAZING TYPE**

Glazing of the window used in the reference building.

Field 38. **REFERENCE BUILDING FRAME MATERIAL**

Material of the window frame used in the reference building.

Field 39. **REFERENCE BUILDING WINDOW COATED**

Existence of the low-E layer in the reference building's window, expressed as *Coated/Non-coated*.

Field 40. **REFERENCE BUILDING WINDOW FILLING GAS**

Gas used for filling the window panels, if applicable, expressed as Argon/No gas. This means that in case of gas-filled windows only argon-filled windows are used to model the reference buildings.

Field 41. **REFERENCE BUILDING WINDOW U-VALUE (W/m2/K)**

U-value of the reference building's window, expressed in W/m2/K.

Field 42. **REFERENCE BUILDING ROOF MATERIAL**

Material used for the structural element of the reference building's roof.

Field 43. **REFERENCE BUILDING ROOF MATERIAL THICKNESS (m)**

Thickness of the structural element included in the reference building's roof, expressed in meters.

Field 44. **REFERENCE BUILDING ROOF MATERIAL THERMAL CONDUCTIVITY (W/m/K)**

Thermal conductivity of the reference building roof's structural material, expressed in W/m/K.

Field 45. **REFERENCE BUILDING ROOF MATERIAL DENSITY (kg/m3)**

Density of the reference building roof's structural material, expressed in kg/m^3.

Field 46. **REFERENCE BUILDING ROOF MATERIAL SPECIFIC HEAT CAPACITY (J/kg/K)**

Specific heat capacity of the reference building roof's structural material, expressed in J/kg/K.

Field 47.   REFERENCE BUILDING ROOF INSULATION MATERIAL

Material used for the insulation element of the reference building's roof.

Field 48.   REFERENCE BUILDING ROOF INSULATION MATERIAL THICKNESS (m)

Thickness of the insulation element included in the reference building's roof, expressed in meters

Field 49.   REFERENCE BUILDING ROOF INSULATION MATERIAL THERMAL CONDUCTIVITY (W/m/K)

Thermal conductivity of the reference building roof's insulation material, expressed in W/m/K.

Field 50.   REFERENCE BUILDING ROOF INSULATION MATERIAL DENSITY (kg/m3)

Density of the reference building roof's insulation material, expressed in kg/m^3.

Field 51.   REFERENCE BUILDING ROOF INSULATION MATERIAL SPECIFIC HEAT CAPACITY (J/kg/K)

Specific heat capacity of the reference building roof's insulation material, expressed in J/kg/K.

Field 52.   REFERENCE BUILDING ROOF U-VALUE (W/m2/K)

U-value of the reference building's roof, expressed in $W/m^2/K$.

8.4 BUILDING STOCK SEGMENT

There fields from this database section explain the size and energy demand of the building stock segments represented by reference buildings:

Field 53.   BUILDING STOCK SEGMENT USEFUL FLOOR AREA (m2)

Useful floor area of the building stock segment represented by the reference building, expressed in m^2.

Field 54.   BUILDING STOCK SEGMENT NUMBER OF BUILDINGS

Number of buildings present in the building stock segment represented by the reference building.

Field 55.   BUILDING STOCK SEGMENT USEFUL ENERGY DEMAND HEATING (TWH/YEAR)

Useful energy demand, or energy need of the building stock segment represented by the reference building, expressed in TWh/year. This is the amount of energy required to cover heating energy need of a building.

8.5 GREY-BOX MODEL RELATED FIELDS

Field 56.   GREY-BOX MODEL IDENTIFIER

Provides the name of the most suitable model structure (ZoneWallIntB_matrix for all reference buildings).

Field 57.   MODEL PARAMETERS

List of all estimated model grey-box parameters. Format is a python dictionary: {'parameter name': value, …}. Since the most suitable model structure is the same for each reference building, the dictionary always

contains 18 parameters, namely a1 to a9 and b1 to b9.

Field 58.   FIT MAPE TRAINING PHASE

Mean Absolute percentage error (MAPE), calculated as follows:

$$MAPE = \frac{1}{N_{training}} \sum_{training} \frac{|T_{calculated} - T_{measured}|}{T_{measured}}$$

The absolute relative error is calculated for each time step. Note that temperatures are expressed in Kelvin. The average is taken by summing over all discrete time steps in the training phase and dividing by the number of time steps $N_{training}$.

Field 59.   FIT RMSE TRAINING PHASE

Root-mean square error (RMSE) of the fit during the training phase:

$$RMSE = \sqrt{\sum_{training} \frac{(T_{calculated} - T_{measured})^2}{N_{training}}}$$

The sum spans over all discrete time steps during the training phase and adds all square errors (difference between calculated and measured indoor temperature). By dividing by $N_{training}$, the total number of discrete time steps in the training phase, the average square error is found. Finally, taking the square root gives the RMSE value for the training phase.

Field 60.   FIT VARIANCE TRAINING PHASE

The variance of the absolute difference between calculated and measured temperature is calculated by the NumPy function var. This value is a measure for the spread of the distribution of the values provided.

Field 61.   MAPE VALIDATION PHASE

Identical to field 58 but instead of using data from the training phase, the validation phase is used to calculate the metric

Field 62.   RMSE VALIDATION PHASE

Identical to field 59 but instead of using data from the training phase, the validation phase is used to calculate the metric

Field 63.   VARIANCE VALIDATION PHASE

Identical to field 60 but instead of using data from the training phase, the validation phase is used to calculate the metric

Field 64.   YEARLY ENERGY USE [kWh/m²]

Estimated heating demand per square meter of the reference building.

9. INPUTS

This section explains the main inputs, other from TABULA and Hotmaps data sets, used for the database development. The following inputs are considered:

1. Properties of the materials,

2. Thermal resistances of services,

3. Reference building size,

4. Glazing properties, and

5. Window properties.

9.1 PROPERTIES OF MATERIALS

TABLE 6 - PROPERTIES OF MATERIALS

MATERIAL	THERMAL CONDUCTIVITY (W/m/K)	DENSITY (kg/m3)	SPECIFIC HEAT CAPACITY (J/kg/K)	MATERIAL TYPE
Cement fibre slabs, shredded wood	0.08	350	1 030	Insulation
Asbestos fibre	0.060	640	840	Insulation
Perlite board, expanded	0.052	16	1 260	Insulation
Rock wool	0.034	200	710	Insulation
Urea formaldehyde resin foam	0.054	14	147	Insulation
Mineral wool	0.042	12	1 030	Insulation
Polystyrene, expanded	0.035	23	1 470	Insulation
Polyurethane foam	0.025	30	1 400	Insulation
Polyisocyanurate, aged, with facers	0.020	25	1 470	Insulation
Concrete block (dense) (exposed)	1.87	2 300	1 000	Concrete
Concrete block (dense) (protected)	1.75	2 300	1 000	Concrete
Precast concrete (dense) (exposed)	1.56	2 100	1 000	Concrete
Precast concrete (dense) (protected)	1.46	2 100	1 000	Concrete
Cast concrete 2000	1.33	2 000	1 000	Concrete
Cast concrete 1800	1.13	1 800	1 000	Concrete
Brick, fired clay, 2400	1.34	2 400	800	Brick
Brick, fired clay, 2240	1.19	2 240	800	Brick
Brick, fired clay, 2080	1.02	2 080	800	Brick
Brick, fired clay, 1920	0.90	1 920	800	Brick
Brick, fired clay, 1600	0.68	1 600	800	Brick
Brick, fired clay, 1280	0.48	1 280	800	Brick

Brick, fired clay, 1120	0.41	1 120	800	Brick
Basalt	3.49	2 880	840	Stone
Granite, red	2.49	2 650	900	Stone
Sandstone	1.83	2 200	710	Stone
Limestone	1.50	2 180	720	Stone
Sandstone tiles	1.20	2 000	840	Stone
Oak, beech, ash, walnut	0.23	650	3 050	Wood
Maple, oak, and similar hardwoods	0.16	720	1 260	Wood
Timber	0.14	720	1 680	Wood

Source: This table is provided from [3] and [4].

9.2 THERMAL RESISTANCES OF SURFACES

The thermal resistance of surfaces is shown in Table 7 and expressed in $(W/m2/K)^{-1}$.

TABLE 7 - THERMAL RESISTANCES OF SURFACES

SURFACE RESISTANCES TYPE	SURFACE RESISTANCE $((W/m2/K)^{-1})$	COMMENT
External surface resistance Rse	0.04	The same for all elements
Internal surface resistance Rsi - floor	0.17	Ceilings or floors, downward
Internal surface resistance Rsi - roof	0.10	Ceilings or roofs, upward
Internal surface resistance Rsi - wall	0.13	Walls, horizontal

Source: These inputs are provided from [3].

9.3 REFERENCE BUILDING SIZE

The reference building size is described in Table 8. Information is provided as reference building use code, reference building number of floors, reference building floor to floor height and reference building floor to floor ratio.

TABLE 8 - REFERENCE BUILDING SIZE

REFERENCE BUILDING USE CODE	REFERENCE BUILDING NUMBER OF FLOORS	REFERENCE BUILDING FLOOR TO FLOOR HEIGHT (m)	REFERENCE BUILDING WINDOW TO FLOOR RATIO
OFF	6		20%
TRA	1	3.5	5%
EDU	2		15%

HEA	5		15%
HOR	5		20%
OTH	3		15%

Sources: Number of floors is provided from [5]. Floor to floor height obtained from [6] page 369. Window to floor ratio provided from [7] section 2.3.

9.4 GLAZING PROPERTIES

An overview of the input for the glazing properties is shown in Table 9. Information is provided as glazing type, (non-) coated, emissivity, gas, spacing and the U-value.

TABLE 9 - GLAZING PROPERTIES

GLAZING TYPE	COATED	EMISSIVITY	GAS	SPACING (mm)	U-VALUE (W/m2/K)
Double	Coated	0.05	Argon	6	2.05
Double	Coated	0.05	Argon	9	1.61
Double	Coated	0.05	Argon	12	1.34
Double	Coated	0.05	Argon	16	1.19
Double	Coated	0.05	Argon	20	1.21
Double	Coated	0.1	Argon	6	2.16
Double	Coated	0.1	Argon	9	1.75
Double	Coated	0.1	Argon	12	1.5
Double	Coated	0.1	Argon	16	1.36
Double	Coated	0.1	Argon	20	1.38
Double	Coated	0.2	Argon	6	2.35
Double	Coated	0.2	Argon	9	1.98
Double	Coated	0.2	Argon	12	1.76
Double	Coated	0.2	Argon	16	1.63
Double	Coated	0.2	Argon	20	1.65
Double	Coated	0.05	No gas	6	2.48
Double	Coated	0.05	No gas	9	1.98
Double	Coated	0.05	No gas	12	1.67
Double	Coated	0.05	No gas	16	1.42
Double	Coated	0.05	No gas	20	1.45
Double	Coated	0.1	No gas	6	2.57
Double	Coated	0.1	No gas	9	2.1
Double	Coated	0.1	No gas	12	1.8
Double	Coated	0.1	No gas	16	1.57
Double	Coated	0.1	No gas	20	1.6
Double	Coated	0.2	No gas	6	2.71
Double	Coated	0.2	No gas	9	2.29
Double	Coated	0.2	No gas	12	2.02
Double	Coated	0.2	No gas	16	1.82
Double	Coated	0.2	No gas	20	1.85

Double	Non coated	n/a	No gas	6	3.28
Double	Non coated	n/a	No gas	9	3.01
Double	Non coated	n/a	No gas	12	2.85
Double	Non coated	n/a	No gas	16	2.73
Double	Non coated	n/a	No gas	20	2.74
Double	Non coated	n/a	No gas	25	2.76
Single	Non coated	n/a	No gas	0	5.75
Triple	Non coated	n/a	No gas	6	2.29
Triple	Non coated	n/a	No gas	9	2.04
Triple	Non coated	n/a	No gas	12	1.89
Triple	Non coated	n/a	No gas	16	1.78
Triple	Non coated	n/a	No gas	20	1.71
Triple	Non coated	n/a	No gas	25	1.72

Sources: [3].

9.5 WINDOW FRAME PROPERTIES

The following properties of materials are depicted in Table 10: the frame material type, frame material and U-value.

TABLE 10 - WINDOW FRAME PROPERTIES

FRAME MATERIAL TYPE	FRAME MATERIAL	U-VALUE (W/m2/K)
Wood	Wood 30 mm thick	2.30
Wood	Wood 40 mm thick	2.15
Wood	Wood 50 mm thick	2.02
Wood	Wood 60 mm thick	1.90
Wood	Wood 70 mm thick	1.78
Wood	Wood 80 mm thick	1.67
Wood	Wood 90 mm thick	1.57
Wood	Wood 100 mm thick	1.48
Plastic	Polyurethane	2.80
Plastic	PVC, two hollow chambers	2.20
Plastic	PVC, three hollow chambers	2.00
Aluminium	Aluminium, thermal break 4 mm thick	4.40
Aluminium	Aluminium, thermal break 8 mm thick	3.90
Aluminium	Aluminium, thermal break 12 mm thick	3.50
Aluminium	Aluminium, thermal break 16 mm thick	3.20
Aluminium	Aluminium, thermal break 20 mm thick	3.00
Steel	Steel	6.90

Source: [3].

10. CONCLUSIONS

Within this document a highly efficient and (semi)automatic manner to create grey-box models is described. Supplied with the reference building data, these models allowed dynamic calculations of building stock thermal performance. However, the existing building stock information is fragmented, different databases contain different pieces of information and there is lack of information that support measuring of flexibility. An intermediate database format was identified as a unified source needed to (semi) automatically create white-box models in Modelica, that subsequently serve as a source required for (semi)automatic creation of grey-box models. For adding more building typologies (for instance, typologies that may be available from some future databases), it will be sufficient to add the information about additional typologies to the unified intermediate database, and subsequently create grey-box models in a (semi)automatic manner.

In addition to describing grey-box models, this report explains development of the intermediate building stock database through selection of data sources, collection of relevant reference building data, and development of the database needed for the assessment of the Energy System Impact and the creation of the performance models.

After considering various sources, the database was developed from two publicly available data sources: TABULA and Hotmaps that provide necessary data for residential and non-residential buildings in EU. For different building stock segments of EU 27 Member States, the database provides information about the building stock segments' reference building, as well as information on the building stock segments.

Due to the limited information in existing building stock databases, in addition to the data retrieved from the TABULA and Hotmaps database a number of assumptions had to be made to populate the unified intermediate database. These assumptions are described in a transparent manner and could be changed if needed. This might be done in future work where scenario's development is taken into consideration. But ideally, future work on building stock creation should capture the needed information so that number and scope of such assumptions is reduced as much as possible.

The unified intermediate database was set up to create white-box models in Modelica that subsequently were used as input for the grey-box models. The parameters for these grey-box models were calculated and added to the database.

The results compiled in the database can serve as a starting point to assess the flexibility potential when investigating a particular case for which no data/not enough useful data is available. This is relevant for most buildings in EU where examples from the database can be used as input for the calculating the flexibility potential with the ABEPeM platform.

11. REFERENCES

[1]   TABULA project, "Tabula Final Report," 2012.

[2]   Hotmaps project, "D2.3 WP2 Report – Open Data Set for the EU28," 2019.

[3]   The Chartered Institution of Building Services Engineers London, Environmental design CIBSE Guide A, 2006.

[4]   American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., ASHRAE Handbook, Fundamentals, 2009.

[5]   U.S. Department of Energy, "Commercial Reference Building Models of the National Building Stock," 2011.

[6]   B. Chun and J.-M. Guldman, "Two- and Three-Dimensional Urban Core Determinants of the Urban Heat Island: A Statistical Approach," 2012.

[7]   N. Ibrahim and S. Hayman, "Daylight rules of thumb," 2005.

[8]   I. collaboration, "OpenIDEAS: An open framework for integrated building and district energy simulations," 2021. [Online]. Available: https://github.com/open-ideas.

[9]   EnergyPlus, "EnergyPlus," 2021. [Online]. Available: https://energyplus.net/weather.

ABBREVIATIONS AND ACRONYMS

ABL	Apartment blocks
AEPC	Active Building Energy Performance Contract
BSO	Building Stock Observatory
DR	Demand Response
EDU	Educational buildings
EPC	Energy Performance Contract
HEA	Health buildings
HOR	Hotels and restaurants
MAPE	Mean absolute percentage error
MFH	Multi-family houses
ODE	Ordinary differential equation
OFF	Offices
OTH	Other non-residential buildings
RMSE	Root-mean square error
SFH	Single-family houses
TH	Terraced houses
TMY	Typical Metrologic Year
TRA	Trade buildings