https://doi.org/10.3311/PPme.43518

Periodica Polytechnica Mechanical Engineering, 70(1), pp. 96–116, 2026

Development and Parametric Analysis of Hungary's Residential Building Stock Model on the Example of an Archetype Building

Dávid Erdős[1*], Miklós Horváth[1]

[1] Department of Building Service and Process Engineering, Faculty of Mechanical Engineering, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[*] Corresponding author, e-mail: erdosdavid@edu.bme.hu

Received: 03 February 2026, Accepted: 12 February 2026, Published online: 24 February 2026

Abstract

This paper quantifies uncertainty and identifies the key drivers of simulated energy use in a representative archetype of the Hungarian residential building stock. It specifically examines the impact of standardized occupancy inputs compared to those based on surveys and stochastic methods. A DesignBuilder model of a typical detached house is parameterized using data from Energy Performance Certificates (EPCs) that detail the building's envelope and system characteristics. Additionally, occupant-related parameters such as setpoints, ventilation, domestic hot water (DHW), and internal gains are derived from a comprehensive national survey and relevant literature. To analyze uncertainty propagation and conduct a global sensitivity analysis, Latin Hypercube Sampling is applied across multiple scenarios: a typical meteorological year, two actual years, and a future climate-change scenario, with and without space cooling. Furthermore, an alternative scenario using a 2050 primary energy conversion factor is evaluated. The results indicate that the heating setpoint temperature is consistently the most influential factor for EPBD-based primary energy usage. Other significant contributors, depending on weather conditions and cooling assumptions, include ventilation rates, heating system efficiency, and parameters related to domestic hot water. Overheating hours are primarily affected by factors such as night ventilation, shading, and internal gains. The findings reveal that using standardized assumptions for occupancy can skew both heating and cooling outcomes. Additionally, assumptions regarding climate and primary energy factors can alter the relative significance of key parameters. The proposed workflow enhances the robustness of building-stock assessments and underscores the value of improving input data quality.

Keywords

residential building stock, energy performance certificate, dynamic simulation, uncertainty and sensitivity analysis, overheating

1 Introduction

Improving the energy efficiency of buildings and making them carbon neutral is one of the main priorities of the European climate policy, as buildings are currently responsible for around 40% of EU carbon dioxide emissions and energy consumption, exceeding the industrial and transport sectors [1]. Residential buildings play a key role in this context, making the energy renovation strategy for residential buildings a priority issue and a fundamental tool for supporting decision-making [2]. Determining the energy consumption of the residential building stock is not a simple task, but it can be done with detailed dynamic simulation models compiled from archetypes, which also allow the impact of future renovations to be analyzed. However, dynamic simulation models require a large number of input parameters, which in turn require a comprehensive and detailed database of the building stock. Energy Performance Certificates (EPCs), which are commonly used for similar purposes in other countries, are suitable tools for compiling this database, provided with detailed parametric analysis [3, 4]. However, no comprehensive parametric analysis of the energy consumption of the building stock in Hungary has been carried out so far; existing studies typically focus on only a few selected parameters. Our research addresses this gap by developing and testing a comprehensive parametric analysis framework on a Hungarian residential archetype model, while also assessing the impact of model-development assumptions.

Cite this article as: Erdős, D., Horváth, M. "Development and Parametric Analysis of Hungary's Residential Building Stock Model on the Example of an Archetype Building", Periodica Polytechnica Mechanical Engineering, 70(1), pp. 96–116, 2026. https://doi.org/10.3311/PPme.43518

1.1 Utilization of certificate data for modeling the residential building stock

EPCs were introduced in Hungary in 2006, when Government Decree 7/2006 (V. 24.) TNM entered into force [5]. This decree is a consequence of Directive 2002/91/EC [6] on the energy performance of buildings, which required the introduction of national energy certificates. Today, EPBD (Energy Performance of Buildings Directive) provides regulations for EU member states regarding the energy performance of buildings and energy certificates. The 2018 version of the EPBD directive was implemented in Hungary by Decree 9/2023. (V. 25.) ÉKM [7]. Due to the mandatory nature of EPCs, certificate coverage is high in most European countries; in the United Kingdom, for instance, EPCs cover about 60% of the building stock [8]. For Hungary, no precise and reliable coverage estimate is currently available, but the regulatory requirements imply that the national coverage rate is likely to be relatively high as well. Energy certificates are therefore one of the most comprehensive sources of data on the residential building stock and its energy performance, so it is obvious to use this data when modeling the residential building stock.

Residential building stock can be modeled using various methods, but we can essentially distinguish between "bottom-up" and "top-down" approaches. The development of dynamic simulation archetype models is the basis of the bottom-up method, in which the energy consumption of a typical building is determined and then scaled up. The energy consumption of a typical building can be determined using statistical methods or engineering methods. In the latter case, energy consumption is calculated based on the building's physical parameters – hence dynamic simulation archetype models inherently fall into this category. The great advantage of engineering methods over statistical methods is that they model energy consumption based on the laws of physics, which makes them practical for analyzing renovation scenarios and developing refurbishment strategies. Lim and Zhai [9] also conducted a comprehensive analysis of the methodologies used, which is summarized in Fig. 1.

In their research, they placed particular emphasis on engineering-based models and, within this, on stochastic models. Stochastic methods describe uncertain input parameters with probability distributions and, by extension, specify output variables with distributions. In contrast, deterministic methods give a single specific result from fixed input parameters. If a large, comprehensive database of energy certificates is available, it is possible to create stochastic models, which in some cases can provide more accurate results [9].

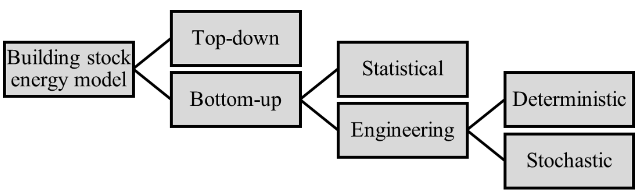


Fig. 1 Methods used in building stock modeling [9]

As mentioned earlier, energy certificates have been used in several countries in research on residential building stock modeling. Pasichnyi et al. provided a comprehensive summary by reviewing 79 articles on the use of energy certificates in this field [10]. They found that most of this research in the European Union and Europe comes from the United Kingdom, with 17 articles. This is followed by Sweden, Italy, and France with 13, 11, and 7 articles, respectively. Since the study was completed, these figures may have increased further, and overall, it can be said that the use of certificates in building stock modeling is indeed a common practice.

One problem with the use of energy certificates was their limited availability in the early stages [11], which determined the number of buildings analyzed in the studies reviewed by Pasichnyi et al. [10]. The number of samples may, in some cases, influence the statistical strength and applicability of the research results. The gradual implementation of the EPBD directives across member states had a major stimulative effect in this regard, and, based on the 79 articles analyzed, an upward trend in the number of buildings examined can be observed. In other words, after 2013, articles in which the number of buildings examined exceeded 100,000 began to appear and became increasingly common, although articles with smaller sample sizes did not disappear from practice [10]. However, data access remains difficult in Hungary, so the order of magnitude of hundreds of thousands is still a question for the near future in our country.

The study also summarized the purposes for which certificates are used in residential building modeling. The most common purpose was to determine building energy consumption, which was reported in 34 articles. The second most common purpose was to analyze the energy renovation of buildings, which was examined in detail in 18 articles. Other common applications included analyzing the impact of certificates on real estate market decision-making, validating and evaluating the quality of certificate data, analyzing performance, and forecasting future energy consumption and CO_2 emissions [10].

However, the positive assessment was undermined in several other articles by conclusions drawn about modeling the residential building stock using building energy certificates. Several problems have come to light; for example, Mangold et al. highlighted in their research analyzing data quality that the determination of heated floor space is inconsistent and often underestimated [12]. In another study, Hardy et al. found that in a significant proportion of certificates, the identification of building structures is also inaccurate, leading to buildings often being incorrectly classified [13]. Another frequently occurring problem is the underrepresentation of certain building types (or residential building types) and distorted models due to the consideration of the most typical system type among building technical systems. Another common problem is that, due to the validity period of energy certificates (5 or 10 years) and renovations carried out in the meantime, the data does not always reflect the current status, resulting in a time lag [14]. The latter problems also apply to Hungarian certificates, and overall, based on the above, it can be said that, with the use of certificates, the main problem in modeling the residential building stock has shifted from data collection to data quality.

Based on the findings above, data quality and data uncertainty are key problems, and it is important to analyze their degree. Van Hove et al. draw attention to the lack of quantification of this in their research and recommend the use of stochastic models to manage uncertainties [15].

Dahlström et al. [14] also emphasized the importance of using stochastic models in their work, which summarized the future research goals identified in articles on energy modeling of the building stock. The second most frequently suggested future research topic in these articles was stochastic occupancy models, and they also suggested the inclusion of other uncertain stochastic parameters, such as infiltration [14].

However, rather than stochastic building usage parameters, energy certificates generally account for standard building usage. This raises the question of how much more accurate the results would be with an average building model based on actual building use, for example, from a survey. Furthermore, as a follow-up, would the aforementioned stochastic (also survey-based) building-use model yield significantly more accurate results? In our research, we seek to answer this question and, using the stochastic model, determine the uncertainty in the energy consumption results arising from input parameters derived from certificates (and other supplementary data sources). Among the input parameters, it is also worth examining which ones are responsible for the largest part of the total uncertainty – that is, which are the most decisive parameters in the model and which ones need to be specified with the greatest emphasis.

1.2 Uncertainty assessment and sensitivity analysis in residential building stock modeling

The uncertainty in energy consumption and other examined building simulation results, as well as the parameters that most significantly determine them, can be identified through uncertainty and sensitivity analyses. Numerous studies have been conducted on this topic, and their findings have been summarized in six comprehensive review articles.

Perhaps the most important lesson and conclusion of these review articles is that greater emphasis is needed on thoroughly determining and quantifying the uncertainty of input parameters (and even conducting further research) to ensure the quality of such analyses. Several review articles clearly identify this as the most difficult task; in many cases, for example, lower and upper limits and distribution curves are missing, or if they are present, they are incomplete [16–18].

In this regard, Tian et al. emphasized that databases created to characterize the uncertainty of input parameters should be structured according to certain indicators, taking into account, for example, the impact of different building types and climatic characteristics [17]. This is confirmed by another study that also attributed a prominent role to weather, which can decisively influence the key parameters of the models. As a result, more and more researchers have been studying the impact of weather in recent times [18].

In addition, building user behavior was also emphasized in the conclusions of the review articles. They highlighted that more attention should be paid to the stochastic modeling of building use, as this parameter is a very complex factor and has a significant impact on the energy consumption of buildings [17].

More recent review articles have compiled a large collection of previously published works, from which interesting conclusions can be drawn. Based on these, it can be said that uncertainty and sensitivity analyses were mainly performed on residential buildings, followed by educational buildings. Most studies were conducted in hot-climate regions, and the vast majority used EnergyPlus or DesignBuilder for modeling. Another important conclusion is that most analyses focused on energy consumption or comfort output parameters (i.e., the uncertainty caused by these input parameters was most often examined).

Furthermore, the heating or cooling setpoint temperature was the most decisive parameter in most studies (in both heating- and cooling-dominant climate zones) [18, 19].

However, in the review articles and the articles referenced therein, numerous other building energy parameters were also examined, with the most frequently occurring input parameters listed in Table 1 [16–21]. Table 1 shows the parameter groups, listed by frequency of occurrence, along with the individual parameters within them. It served as a strong reference point during our research, as we mainly selected the input parameters for our analysis from the frequently collected parameters, taking into account the desired output result, climate, and data availability (since the determining parameters are strongly dependent on these, as highlighted earlier).

Among the input parameters listed in Table 1, the most frequently used in the studies were the U-values of different building structures, which can be classified as building envelope parameters, and HVAC system input parameters (e.g., heating setpoint temperature) were also included in many cases. Parameters characterizing user behavior were also included, and the effect of building geometry was analyzed in certain articles, while the effect of weather and climate was also examined as an input parameter [17].

Table 1 Frequently analyzed building energy parameters [16–21]

Number	Parameter group	Parameter
1	Building envelope	Window shading
2		Infiltration rate
3		Roof U-value
4		External wall U-value
5		Ground floor U-value
6		Window U-value
7		Window g-value
8		Surface optical properties
9		Thermal bridges
10		Thermal mass
11	HVAC system	Heating setpoint temperature
12		Cooling setpoint temperature
13		Heating system efficiency
14		Cooling system efficiency
15		DHW system efficiency
16		PV module efficiency
17	Occupant behavior	Lighting power density
18		Ventilation rate
19		Appliances power density
20		Occupant density
21		DHW demand
22	Building geometry	Orientation
23		Window-to-Wall Ratio
24		Conditioned floor area
25	Weather data	Weather file

Most articles on this topic do not refer to building stock, but only to individual buildings. This is pointed out by Tian and colleagues in their work, emphasizing that the main reason for this is the high uncertainty and heterogeneity of the input parameters [17].

Despite the difficulties, several articles have been written specifically on the topic of uncertainty and sensitivity analysis in relation to building stock. For example, Zhao et al. [20] conducted a sensitivity analysis on office buildings, analyzing the impact of a total of 18 input variables on the EUI (Energy Use Intensity) value. For the Chicago climate, the most significant of these variables were lighting and appliances power intensity, floor area, window-to-wall ratio, cooling COP, cooling and heating setpoint temperature, infiltration rate, external wall, and roof U-value [20].

Tian and Choudhary [21] also worked on this topic, creating a stochastic model for educational buildings in London. They found that the heating setpoint temperature, ventilation rate, infiltration rate, and roof U-value were the four most significant input parameters for heating energy consumption. Heating energy consumption accounted for a dominant share of total energy use; therefore, factors that primarily determine electricity consumption (e.g., lighting and equipment power density) were not dominant contributors to overall energy consumption [21].

Booth et al. [22] also relate their work to the topic of uncertainty and sensitivity analysis in the context of building stock, and, in addition, their article specifically focuses on research on residential building stock/housing stock. They conducted a sensitivity analysis for the British temperate climate, identifying the input variables that contribute most to the model's uncertainty. The three most dominant variables were conditioned floor area, heating setpoint temperature, and heating system efficiency, and the infiltration rate was also identified as playing a significant role. The window-to-wall ratio, window, and external wall U-value input parameters were found to be less significant [22].

The literature review presented above highlights that while the use of EPCs in building stock modeling is a common practice, persistent limitations remain regarding input data quality and the explicit quantification of input parameter uncertainty. In the Hungarian context, comprehensive parametric analyses quantifying uncertainty arising from the stochastic variability of input parameters in the residential building stock have not yet been conducted. Therefore, this research proposes and demonstrates a framework supporting stochastic residential building stock modeling through a detailed parametric uncertainty and sensitivity analysis, tested on an archetype building under multiple scenarios. In our research, the development of the uncertain input parameter database for the investigated archetype also represents a novel contribution in itself, particularly as it is constructed with consideration for specific building types and climatic characteristics.

Building use has also been identified as a critical issue in numerous articles; therefore, we compare the baseline model based on standardized EPC building use data with a survey-based average building model and a survey-based stochastic model in terms of primary energy consumption and overheating.

The paper is structured as follows: Section 2 describes the methodology, including the archetype modeling and selection, the definition of input-parameter distributions, and the investigated scenarios. Section 3 presents and discusses the results of the uncertainty and sensitivity analyses, compares the different building use models, and evaluates the impacts of the scenarios. Finally, Section 4 summarizes the key findings and provides recommendations for future work on stochastic modeling of the residential building stock.

2 Methodology

Our research is based on the residential building typology in Hungary, with information from a previous representative survey. Based on the survey, dynamic simulation archetype models were created, and for one of these archetypes, we performed a comprehensive parametric analysis and examined model development using a stochastic model.

A critical step in the parametric analysis was compiling the candidate input parameters. During the literature review, sensitivity tests and uncertainty analyses from individual studies were reviewed to collect typical building physics and technical systems parameters. Candidate parameters are the input parameters expected to be most decisive for the model. In our research, we saw that the key parameters identified by the sensitivity analysis are climate and building type-dependent (and, of course, varied greatly depending on the output variable examined), which we considered when compiling the list of candidate parameters. Based on the survey parameters, the list was adjusted (as detailed in Section 2.3), and we then performed a sensitivity analysis on these parameters during our work and examined the uncertainty they introduced to energy consumption and overheating in a typical building across several scenarios. To do this, it was also necessary to determine the distribution of the selected input parameters, which, in turn, required filtering the datasets in several cases to ensure the quality of the results. We compared the simulation results from the stochastic model obtained in this way with those of a simpler model based on average survey results and a model reflecting standard building use.

2.1 Residential building stock and modeling

The Hungarian residential building stock was represented using a building typology derived from a representative, detailed survey of 2029 residential buildings. Based on the results of the KEOP-7.9.0/12-2013-0019 project completed in 2015, 23 synthetic average buildings were created for the 23 building types in the building typology using certificate data [23]. In this typology, the types were determined by construction period, technology, and size, resulting in 12 single-family house types and 11 multi-family types.

In our research, we used non-existent, artificially created archetype buildings: dynamic simulation models created in the EnergyPlus environment using DesignBuilder software (v2025.1.0.92) [24]. The geometric and building physics characteristics of the archetypes were created according to the certificate database averages, as detailed in the work of Horváth et al. [25], and the methodology used in the present study for the base model follows the approach presented in that paper. However, they also noted in their work that using averages may distort results, a suggestion that partly inspired the current research. Also, certain input parameters could not be obtained from the certificate database, so, for example, for building usage parameters, they used standard values in the model [25].

The specification of the input parameter values is presented in detail later (Sections 2.2 and 2.3); however, describing the schedules and control logic is also necessary, as there are some differences from the model used in the work of Horváth et al. [25]. Ventilation, in particular, must be addressed, as its control logic was defined in great detail in our work and linked to both outdoor and indoor temperatures. Under winter operation, defined by the 24 h rolling average of outdoor temperature, only infiltration is assumed at night (22:00–06:00), while during the day, ventilation air change is also present. During summer operation, ventilation is applied at night (with either an intensive or a normal ventilation air change rate, depending on the indoor temperature). In the transitional seasons, the model assumes a constant ventilation rate when the relevant temperature ranges are met.

A similar logic was applied in our research to define the use of heating or cooling schedules: both systems are activated as a function of the 24 h rolling average of outdoor temperature while maintaining the specified setpoint temperatures.

In our investigations, we analyzed the Type 5 family house, as it represents by far the largest total floor area and the highest number of units within the overall residential building stock. Type 5 of the Hungarian building typology is a single-storey single-family house with a heated floor area of 103.4 m^2. The envelope consists of constructions with poor thermal performance: an uninsulated brick wall and an attic floor slab filled with boiler slag. Based on the above model setup considerations and the envelope's characteristics, the dynamic simulation building model of the archetype is shown in Fig. 2, as implemented in DesignBuilder.

The building is heated by a conventional gas boiler with radiators, and cooling is provided by a split air-conditioning unit. DHW is produced by an electric storage water heater, with no standby losses considered (as they are not typical for this type of system), and storage losses were also neglected in the model. Regarding distribution losses, DesignBuilder does not account for heat loss in the piping network (and it is likely not significant due to the small distribution networks in this building type). A circulation loop is not typical for the examined older building type, so it was not included in the model either.

The HVAC system schematic implemented in DesignBuilder is shown in Fig. 3.

2.2 Input parameters for building use

Specifying building usage parameters was already a major problem in the articles reviewed in the literature review, and in many cases this was the main difficulty encountered in the research. In our case, an important question is how

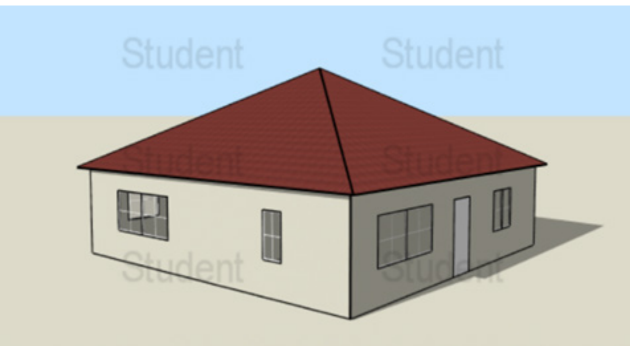


Fig. 2 Building model

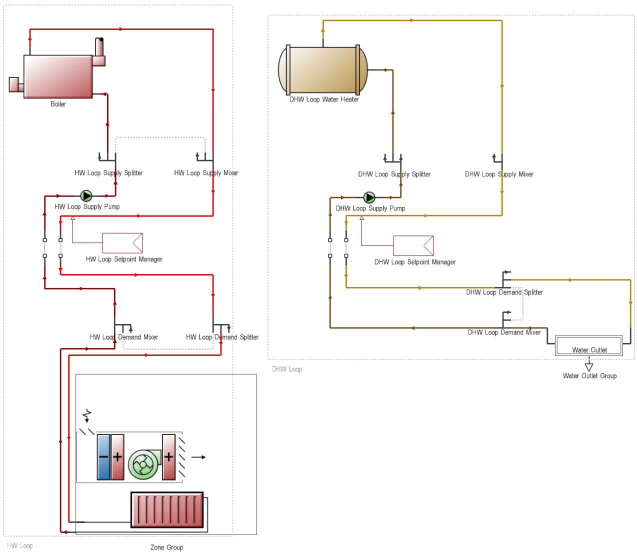


Fig. 3 Detailed HVAC system scheme

large a difference in energy-consumption results arises between a model parameterized under standard-compliant assumptions and a survey-based model (and, subsequently, whether a stochastic model can achieve higher accuracy than the latter).

For standard building use in our research, we took into account the values specified in EN 16798-1:2019 standard [26] Residential – detached house, except for the DHW demand and the air change rate for summer night ventilation. For these, we took into account the standard values according to Hungarian practice (Decree 9/2023. (V. 25.) ÉKM) [7].

When specifying the building usage parameters, we also used the results of REKK survey, a representative survey involving more than a thousand households [27], as the building usage characteristics could not be obtained from the certificates. We included the average results of the REKK building use survey in the model, comparing them with the standard values. The exception to this is the Lighting and equipment power density, for which we adopted literature-based values following the work of Czétány et al. [28] (detailed in Section 2.3). The comparison of the input parameters can be seen in Table 2.

In most cases, there are significant differences between the input parameters, which are likely to be reflected in the simulation energy consumption results.

2.3 Setup for sensitivity and uncertainty analysis

When setting up the sensitivity analysis, the first step is to define the output variables, which determine the subsequent steps. In our case, we want to examine the impact of input parameters on the total primary energy consumption of the

Table 2 Standard and survey-based parameters of building use

Parameter	Standard	Survey
DHW demand (m^3/y)	55.74	50.76
Occupant density (people/m^2)	0.0235	0.0301
Lighting & appliances power density (W/m^2)	2.4	3.497
Heating setpoint temperature (°C)	20	21.19
Cooling setpoint temperature (°C)	26	23.59
Ventilation rate, normal (1/h)	0.667	0.334
Ventilation rate, increased (1/h)	5	2.666

building type. In addition, it is important to emphasize that, under the 2018 EPBD provisions as implemented in Hungary, residential energy performance certificates exclude lighting and other household electricity consumption [29]; accordingly, we also use the EPBD primary energy consumption as a target variable. We also aim to assess the buildings' annual overheating hours, defined as the time during which the indoor temperature exceeds 26 °C.

After the output variables, it is necessary to identify the candidate parameters expected to account for the greatest part of the model uncertainty. We did this in Section 1.2, using parameters collected from the relevant literature, and in Section 2.3, we narrow them down to the available and relevant input parameters.

Therefore, we omitted some of the parameters listed earlier in Table 1 from our study for the following reasons. None of the survey databases contain relevant information on the surface optical properties, thermal bridges, and thermal mass input parameters, and they are rarely mentioned in the literature, so there is no reliable information on the distribution of these parameters either. The DHW system efficiency parameter is not relevant in our case, as DHW production is provided by an electric boiler that always operates at 100% efficiency. PV system efficiency is also irrelevant, as solar panels were installed in a negligible number of cases for the building type under investigation [30]. In the case of the building geometry parameter group, we encountered a software-related limitation on the one hand, and on the other hand, it is not possible to extract the necessary data from the certificates for these parameters either – for example, the orientation of the building is only available at the level of individual constructions. Weather data is not treated as a stochastic input parameter, but rather analyzed as scenarios to determine the effect of different weather files on the results (more on this in Section 2.4).

The quality of the data obtained from the surveys can be improved by reviewing and filtering the data sets. We do this in the following, and then we determine the distributions of the filtered data sets obtained from the survey using the XLSTAT plug-in software [31] with the Kolmogorov-Smirnov test [32], which is the next step in compiling the study.

Among the input parameters, we used the certification database for the U-values of building structures. In most cases, we applied literature-based filtering to the U-values found in the certificates to remove implausible or inconsistent entries [33]. For the U-values of external walls, for example, we considered the typical minimum and maximum U-values for structures characteristic of the given building type and period, so we only considered values between 1.2 and 1.6 W/m^2K when adjusting the distribution. This effectively excluded insulated structures and excessively high heat transfer coefficients (presumably incorrectly specified) from the study [32]. The exclusion of thermally insulated structures is due to the fact that, given the original state of the residential building stock, we consider all structures to be uninsulated in the case of the building type under investigation, and also due to the fact that it distorts the data, and no distribution could be fitted on the original data points. We were then able to fit a normal distribution to the truncated set.

We proceeded similarly with the U-values of attic ceilings, taking values between 0.8 and 1.6 W/m^2K into account and thus also fitted a normal distribution. For ground floors, we also removed extreme values, thereby accounting for heat transfer coefficients between 0.57 and 1.22 W/m^2K. We then fitted a log-normal distribution to the dataset; that is, for this construction type, predominantly low thermal transmittance values are observed, alongside a few cases with higher U-values.

For external door U-values, the assessor software [34, 35] typically requires the U-value to be entered directly (not calculated), resulting in a much more quantized dataset. We applied a maximum threshold of 4 W/m^2K, removing higher values considered likely erroneous. A normal distribution was then fitted to the filtered data.

We proceeded similarly with the U-values of windows, taking into account the typical minimum and maximum heat transfer coefficients reported in the literature, i.e., 1.1–4.0 W/m^2K. For windows, we also used the certification database for the g-value, but with corrected values. This is because it contained many incorrect values (e.g., a Window g-value of 0.1 is currently a highly un-common technology), so we used typical g-values assigned to U-values to correct the presumably incorrect data. We then fitted a normal distribution to the data series.

Shading was also taken into account based on the certification database, and the main cases identified in the survey were divided into two large groups, thereby simplifying the model. Thus, we considered only cases with or without external shading, resulting in approximately equal numbers for both, thereby assuming a uniform distribution of the parameter.

In the case of infiltration rate, the certification database also contains many incorrect values, as older certificates – which were still prepared in accordance with Decree 7/2006 (V. 24.) [5] TNM – incorrectly specified an infiltration rate of 0 1/h (at that time, this was also the default value in the certification software [34, 35]). Thus, similar to the g-value, we used the typical infiltration air change rate assigned to the window U-value to correct the presumably incorrect values, and then fitted a normal distribution based on these corrected values.

For the heating system efficiency parameter, we also referred to the certification database; however, because the model considered a traditional gas boiler system, we only used gas boiler efficiency values. We calculated system efficiencies by dividing the net heating demands obtained from the certificates by the final heating energy consumption, and we removed values above 0.8 from the data series (as these cases would have reflected condensing gas boiler systems).

In the case of cooling system efficiency, the certification database was no longer usable due to incomplete data, so we used a database of Varga [36]. We fitted a normal distribution to this dataset as well – similarly to the heating system efficiency – although in this case, the fit was somewhat less accurate.

For the ventilation rate, the certification database was also unusable, as the vast majority of cases used the default value of 0.5 1/h specified in the certification software [34, 35]. Thus, the ventilation air change values were determined based on the specific fresh air demand (per-capita rate) of 30 m^3/h/person [37] and using the REKK building usage survey results for the number of occupants. We fitted a lognormal distribution to the resulting data series and used a truncated lognormal distribution for sampling in the range 0.110–0.860 1/h (reflecting the actual minimum and maximum values obtained with the specific fresh air demand, thus avoiding distortion from unrealistically high or low values).

We also relied on REKK survey data for the air change rate during night ventilation in summer. For the building type examined, intensive night ventilation (accounted for with 5 1/h air change rate) was present in about half of the cases and absent in the other half, so we assumed a uniform distribution for the parameter.

We also used the REKK survey results to determine the distribution of heating setpoint temperatures. However, the survey offered respondents overly quantified response options, resulting in a slightly inaccurate fit. For this parameter, we assumed a normal distribution (also because of its nature). We proceeded similarly for the cooling setpoint temperature, also assuming a normal distribution.

The REKK survey data were also used for occupant density. We fitted a log-normal distribution to the data series and used a truncated log-normal distribution for sampling in the range of 0.00967–0.0774 person/m^2. These values correspond to 1 and 8 people, the actual minimum and maximum values from the survey (again avoiding distortion from unrealistically high or low values).

For lighting and appliances power density, we relied on values from the literature, using the results of Czétány et al. [28]. From the data series in the literature, we filtered out the top 10% of values that resulted in unrealistically high heat loads and electricity consumption (5782 kWh/year, corresponding to an average annual electrical power density of 6.38 W/m^2 when projected onto the floor area). The reason for this is that the exceptionally high consumption values likely reflect consumption outside the building (e.g., for agricultural purposes), which we do not account for (since its effect would not appear as internal heat load). A log-normal distribution could be fitted to the resulting data series.

Finally, DHW demand values were determined based on a specific DHW demand of 50 l/person/day [38] and using the results of the REKK building usage survey for the number of occupants. We fitted a lognormal distribution to the resulting data series and applied a truncated lognormal distribution for sampling in the range of 0.0139–0.111 m^3/s (corrected values for the specified schedule in the model), resulting in a consumption of 18–146 m^3/year. This reflects the actual minimum and maximum values obtained for the specific DHW demand, thereby avoiding distortion from unrealistically high or low values.

The distributions of the input parameters, based on the above considerations, are summarized in Table 3. This database of uncertain input parameters is specific to the investigated archetype.

Most parameters follow a normal distribution, but for building usage parameters, a log-normal distribution is more common. Most parameters are continuous, only in the case

Table 3 Uncertain input parameter database for uncertainty and sensitivity analysis

Number	Parameter group	Parameter	Category	Distribution	Descriptive statistics	
					μ	σ
1	Building envelope	External wall U-value (W/m²K)	Continuous	Normal	1.380	0.086
2		Attic floor U-value (W/m²K)	Continuous	Normal	1.130	0.209
3		External door U-value (W/m²K)	Continuous	Normal	2.69	0.736
4		Ground floor U-value (W/m²K)	Continuous	Log-normal	−0.217	0.222
5		Window U-value (W/m²K)	Continuous	Normal	2.518	0.670
6		Window g-value (–)	Continuous	Normal	0.686	0.049
7		Shading	Discrete	Uniform	–	–
8		Infiltration rate (1/h)	Continuous	Normal	0.280	0.10
9	HVAC system	Heating system efficiency (–)	Continuous	Normal	0.661	0.047
10		Cooling system efficiency (–)	Continuous	Normal	3.843	0.875
11	Occupant behavior	Heating setpoint temperature (°C)	Continuous	Normal	21.186	1.382
12		Cooling setpoint temperature (°C)	Continuous	Normal	23.588	2.190
13		Ventilation rate, normal (1/h)	Continuous	Log-normal	−1.223	0.497
14		Ventilation rate, increased (1/h)	Discrete	Uniform	–	–
15		Occupant density (people/m²)	Continuous	Log-normal	−3.631	0.497
16		Lighting & appliances power density (W/m²)	Continuous	Log-normal	4.131	0.711
17		DHW demand (m³/s)	Continuous	Log-normal	−3.269	0.498

of window shading and air change rate during summer nights are the possible values of the parameters discrete. The two columns on the right-hand side of Table 3 summarize the statistical characteristics of the distributions (for example, in the case of a normal distribution, μ is the mean and σ is the standard deviation), based on which we specified the corresponding distributions in the simulation software [24].

After determining the distributions and limits of the input parameters, the sensitivity analysis method must be selected based on the results of the previous steps. There are two main groups to choose from: local sensitivity analysis (LSA) and global sensitivity analysis (GSA). In our case, we will need to use GSA, which, unlike LSA, allows us to analyze the effects of multiple input parameters. GSA methods include regression-based methods, which introduce the use of correlation coefficients to estimate the relationship between input and output parameters [39]. One such method is the Standard Regression Coefficient (SRC) method, which uses standardized coefficients, making them comparable with each other. In our case, the SRC method is also used in the analysis.

As a final important step, it is necessary to select a sampling method, which can also influence the results. Basically, we can distinguish between random and stratified sampling methods. An important consideration when making this selection may be, for example, the available computing capacity, as different methods have different computational requirements – and, of course, this also significantly impacts the time required for the calculations. The most commonly used sampling method in the literature is Latin Hypercube Sampling (LHS), a stratified method, which we also used in our research. The method ensures thorough mapping of the input parameter space by dividing the input ranges into equal parts and then sampling from each interval. The effectiveness of the method also depends on the sample size, so as a rule of thumb, a minimum of ten times the number of input variables is usually recommended for the sample size [40, 41]. In our case, with 17 input variables, this means a minimum of 170 simulation variants, so we ran 200 simulations per scenario during the runs.

2.4 Scenarios

Section 2.4 presents the factors that greatly influence the results of the models, based on which we have compiled various test scenarios.

2.4.1 Cooling

According to the current state of the Hungarian residential building stock, mechanical cooling is present only to a limited extent; where it does occur, it is most common in newer (nearly zero-energy (nZEB)) building types, which nonetheless represent only a small fraction of the overall stock [30]. However, the warming climate has been a strong driver of the post-installation of mechanical cooling (e.g., split air conditioners) in recent years, raising the question of whether cooling should be included in the model. In the simulations, we therefore examine the results of models without mechanical cooling and also models supplemented with split air conditioning.

2.4.2 Weather scenarios

Based on our literature review, it is important to consider the impact of different weather files on simulation results in the sensitivity analysis. Thus, in our research, we analyze four scenarios based on weather files from a colder (2018) and a warmer (2022) actual weather data series (actual meteorological year (AMY)), a typical meteorological year (TMY) based on the period 2007–2016, and a typical weather scenario based on the extreme climate change expected for 2050.

The first three data sets are easy to compile, but weather conditions, assuming future climate change, require a more complex approach. Using the Meteonorm software, we generated a data set based on extreme climate change (RCP 8.5) [42] using the Representative Concentration Pathway, which shows changes in greenhouse gas emissions and atmospheric concentrations (by the end of the 21st century). A comparison of the weather scenarios is shown in Table 4, summarizing the main characteristics.

Table 4 shows that the average temperatures differ by more than 3 °C between the two most extreme cases. The number of heating days decreases by approximately 50 days, while the number of cooling days doubles in the 2050 future scenario compared to the TMY data series. It is also worth examining the annual change in outdoor temperature, shown in Fig. 4 on a monthly basis.

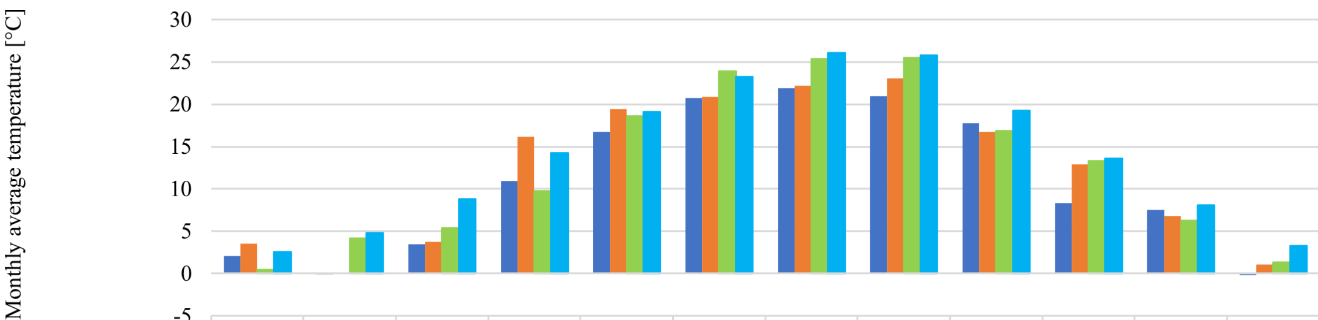
Fig. 4 shows that in the 2022 and 2050 scenarios, the average monthly outdoor temperature rose dramatically during the summer months, while the TMY data series clearly shows the coldest average monthly temperatures.

2.4.3 Primary energy conversion factor

The primary energy consumption of residential buildings examined is strongly influenced by the primary energy conversion factors used in the calculation; for electricity, this means it is important to consider the current non-renewable primary energy conversion factor of 2.3, as specified in Decree 9/2023. (V. 25.) ÉKM [7] decree will change by 2050. The European Union has a well-known goal of decarbonizing the electricity grid energy mix by 2050, and EU measures are also aimed at promoting this; however, it is very difficult to account for the impact of future measures that have not yet been adopted. In our calculations for 2050, we used 0.65 as a scenario assumption for the conversion factor for non-renewable primary energy to

Table 4 Main characteristics of weather scenarios

	TMY (2007–2016)	AMY (2018)	AMY (2022)	RCP 8.5 (2050)
Yearly average temperature [°C]	10.92	12.28	12.62	14.13
Yearly minimum temperature [°C]	−10.30	−18.60	−8.30	−8.40
Yearly maximum temperature [°C]	34.10	34.90	40.10	39.80
Heating days	205	166	174	156
Cooling days	44	60	84	86



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
TMY (2007-2016)	2.08	-0.12	3.50	10.95	16.78	20.78	21.94	20.95	17.74	8.34	7.56	-0.14
AMY (2018)	3.56	0.04	3.74	16.15	19.43	20.88	22.20	23.09	16.74	12.87	6.84	1.06
AMY (2022)	0.42	4.18	5.37	9.73	18.61	23.89	25.35	25.52	16.87	13.32	6.30	1.30
RCP 8.5 (2050)	2.54	4.80	8.82	14.23	19.11	23.27	26.09	25.76	19.27	13.62	8.11	3.27

Fig. 4 Monthly average temperatures

electricity, reflecting a highly decarbonized energy mix. This value was based on an internal, non-citable analysis that was performed at the Budapest University of Technology and Economics to estimate future Hungarian national primary energy factors. In our work, we have designated this scenario 2050*.

3 Results and discussion

To obtain the results, we analyzed a large number of dynamic simulations (2000 in total). In the following, we present the results of the model-uncertainty and sensitivity analysis, as well as the results of the different building-use models.

3.1 Results of uncertainty analysis

During our study, we determined the modeling uncertainty for several target variables, as explained earlier. Since energy certificates were the main source of data, it is advisable to consider the primary energy consumption specified in these certificates in accordance with the 2018 EPBD directive. In models with cooling, this quantity includes the cooling system's energy consumption, whereas in models without cooling, it does not. In addition, total primary energy consumption (including household electricity consumption) is the other target variable in our analysis, and a non-energy variable is also analyzed for models without cooling: the number of overheating hours.

3.1.1 Model uncertainty

The uncertainty of models can be most easily illustrated using box plots. In these boxes, the lower and upper 25% limits are shown, while the median, i.e., the middle value, is also displayed. Thus, the boxes show the middle 50% of the result data, with the other values falling outside this range.

Fig. 5 analyzes the "EPBD-compatible" primary energy consumption for the model without cooling, using values normalized to the floor area of the archetype building.

Fig. 5 shows that the width of the middle 2 quartiles, i.e., the middle 50%, typically ranges from 40 to 65 kWh/m^2 per year in the simulation results for the model without cooling. The box is narrowest for the low primary energy conversion factor, as its low value "squeezes" the range, and widest for the TMY and 2022 weather files. Compared to the coldest TMY scenario, the 2018 and 2022 weather scenarios show a decrease in the building's primary energy consumption for heating and, consequently, in its total primary energy consumption. In the 2050 future scenario, energy consumption decreases further due to even milder winters. Typical primary energy consumption values are

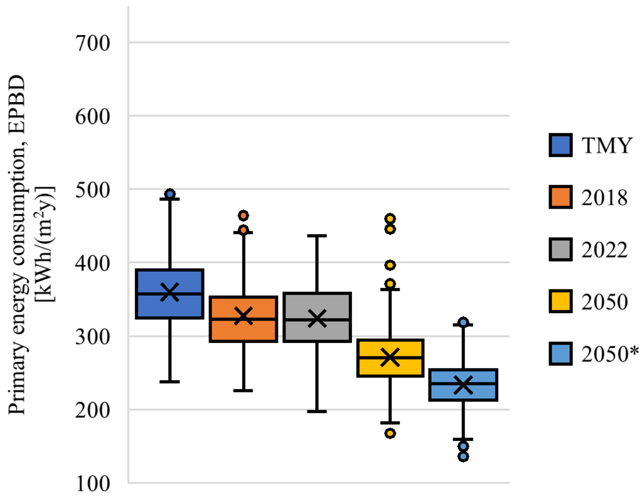


Fig. 5 Uncertainty in primary energy consumption, EPBD – without cooling

325–390 kWh/m^2 per year in the TMY scenario and 215–255 kWh/m^2 per year in the 2050* scenario.

For models without cooling, it is worth examining how the number of overheating hours develops, as shown in Fig. 6.

Based on Fig. 6, it can be clearly concluded that overheating is a much more common problem in the 2022, 2050 (and 2050*) scenarios, as in these cases the middle quartile shifts significantly towards higher values. In the TMY and 2018 scenarios, the annual number of overheating hours often ranges between a few hundred hours, while in the warmer scenarios, 1600–1800 h is more typical, meaning that the importance of cooling comes to the fore with a warming climate. For this reason, it is not advisable to omit cooling from the model (despite it currently appearing in only a relatively small proportion of buildings), as shown in Fig. 7.

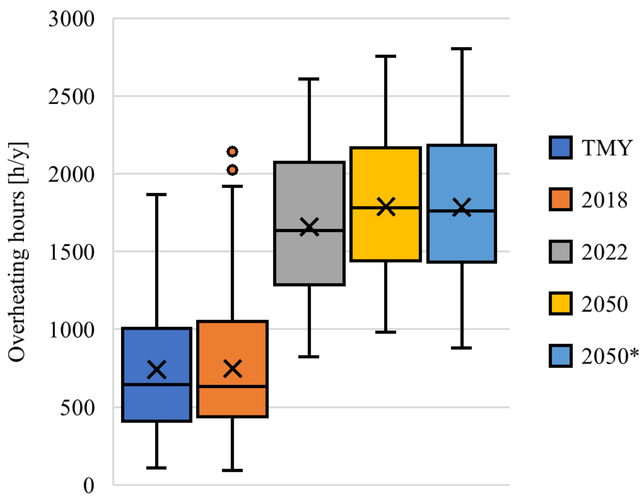


Fig. 6 Uncertainty in overheating hours – without cooling

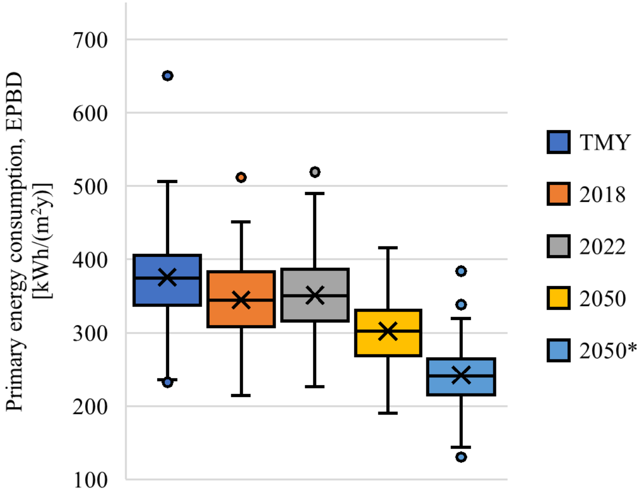


Fig. 7 Uncertainty in primary energy consumption, EPBD – with cooling

Including cooling energy consumption makes the middle two quartiles slightly wider, with values between approximately 50 and 70 kWh/m^2 per year. It can be said that although cooling energy consumption increases in scenarios with a warmer summer (e.g., the 2022 and 2050 scenarios) compared to the TMY scenario, heating energy consumption decreases to a much greater extent due to milder winters, so overall, the warming weather caused by climate change reduces the energy consumption of buildings. Typical primary energy consumption values are 335–405 kWh/m^2 per year in the TMY scenario and 215–265 kWh/m^2 per year in the 2050* scenario.

Subsequently, in addition to cooling energy consumption, we also accounted for lighting and other household electricity consumption, so that the total primary energy consumption (not according to EPBD) is shown in Fig. 8.

The uncertainty obtained from lighting and household electricity consumption is significantly greater than in

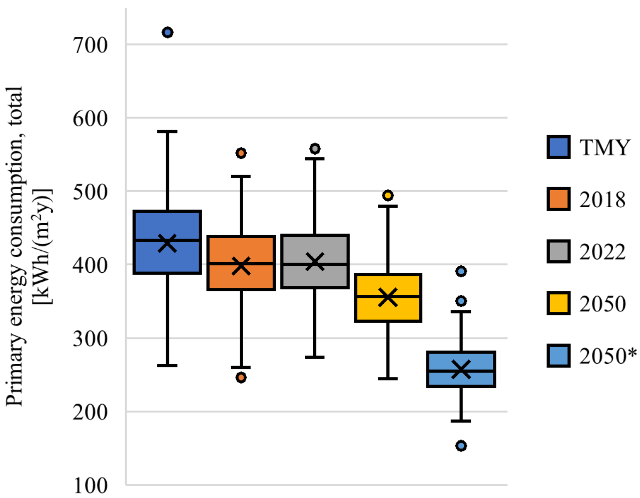


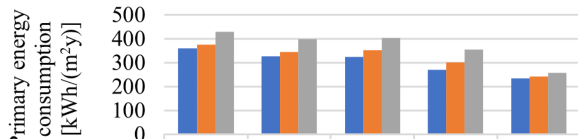
Fig. 8 Uncertainty in primary energy consumption, total – with cooling

previous cases. In other words, the middle two quartiles in this case fall between 50 and 85 kWh/m^2 per year, mainly because the model accounted for the power density of lighting and electrical equipment using a log-normal distribution. As a result, the upper bounds of this parameter can be extremely high, leading to greater uncertainty in energy consumption. The 50 kWh/(m^{2}year) box width observed in the 2050* case is approximately the same as in the previous case, due to the low value of the primary energy conversion factor for electricity. With lighting and household electricity consumption, the typical values for primary energy consumption are 390–475 kWh/(m^{2}year) in the TMY scenario (which is significantly higher than in the previous case), while in the 2050* scenario they are 235–280 kWh/(m^{2}year).

After reviewing the results of the different primary energy use indicators separately, it is worthwhile to analyse how these indicators relate to each other – i.e., by how much cooling, lighting, and other household uses increase the total primary energy use. This is illustrated in Fig. 9.

Based on Fig. 9, cooling increases the total primary energy use by approximately 16–17 kWh/(m^{2}year) in the scenarios with cooler summers, whereas in the warmer-summer scenarios (2022 and 2050) the increase rises to about 27–31 kWh/(m^{2}year). When lighting and other household electricity uses are also included, the increase is larger: total primary energy consumption grows by around 53–54 kWh/(m^{2}year). In the 2050* scenario, the additional primary energy use is lower, since both cooling and lighting/other household energy use are electricity-based.

To summarize the uncertainty analysis, we present the Coefficient of Variation (CV) values for each scenario. CV is defined as the ratio of standard deviation to mean value, expressed as a percentage, and is also referred to as Relative Standard Deviation (RSD) in some literature. Table 5 summarizes the CV values for all scenarios.



	TMY	2018	2022	2050	2050*
Primary energy consumption, EPBD - without cooling	360.07	327.44	324.22	271.26	233.34
Primary energy consumption, EPBD - with cooling	376.05	344.81	351.37	302.21	242.20
Primary energy consumption, total - with lighting and appliances	429.28	398.03	404.45	355.45	257.40

Fig. 9 Primary energy consumption of stochastic models

Table 5 Coefficient of variation for each scenario

	Without cooling		With cooling	
	Primary energy consump-tion, EPBD	Overheating hours	Primary energy consump-tion, EPBD	Primary energy consump-tion, total
TMY	14.60	57.47	15.02	14.84
2018	15.15	64.87	14.47	13.63
2022	14.50	27.81	15.36	14.10
2050	16.44	24.31	14.84	13.70
2050*	15.10	24.34	15.83	14.48

Overall, the CV value for primary energy consumption typically falls between 14 and 16%, i.e., the relative standard deviation is approximately ±15%. However, a different trend is observed in the number of overheating hours, and it is important to note that, while the results for primary energy consumption show a nearly symmetrical distribution, this is less clear for overheating hours. For this reason, the interpretation of mean- and SD-based indicators may be more limited, and it is advisable to treat the results with caution and to also consider the quartile-based indicators briefly analyzed earlier.

For overheating hours, the high CV values (57–65%) obtained for the TMY and 2018 scenarios indicate that, in these cases, differences among individual model variants have a substantial impact on the results. In other words, there are parameter combinations that already cause significant overheating, and there are others that cause almost no overheating at all in buildings without mechanical cooling – the most decisive parameters here are the best passive heat protection measures, which will be determined in a later sensitivity analysis (Section 3.2). In contrast, the CV value is much lower (24–28%) in the 2022 and 2050 (and 2050*) scenarios, where the "average" overheating is very high and remains so across almost all model variants (as we saw this trend earlier in the box plots for overheating hours). In other words, passive heat protection measures are no longer sufficient in these cases, and mechanical cooling is needed to ensure adequate comfort in summer. It is important to note that the decrease in CV is mainly due to the increase in the average number of overheating hours, rather than a significant change in the standard deviation.

3.1.2 Comparison of models

We compare the results of three models regarding building use. A model reflecting standard building use, a model based on survey averages, and a stochastic model based on survey values (in this case, the average and median of the model simulation results).

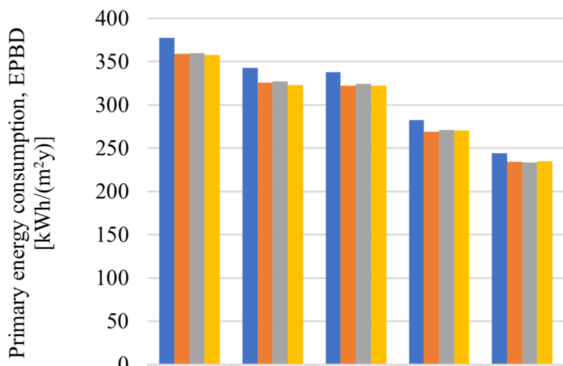
In this analysis, we begin by examining the results of models without cooling, shown in Fig. 10 for the different scenarios.

Based on Fig. 10, without cooling, the model overestimates energy consumption across all scenarios when using standard building usage parameters. Furthermore, a practically highly relevant finding is that there is no significant difference between the model based on the survey's average values and the stochastic model based on the survey, mainly because most of the stochastic input parameters examined follow a normal distribution. Based on this the use of average survey values appears sufficiently accurate in this case.

In the 2050* scenario, primary energy consumption decreases with a lower electricity primary energy conversion factor. Electricity consumption is much lower than natural gas consumption based on the current state of the building stock (or the model representing it), but this is expected to change significantly with the renovation of the building stock. With electrification, the decarbonized network energy mix will play a much greater role than the current state of the building stock suggests.

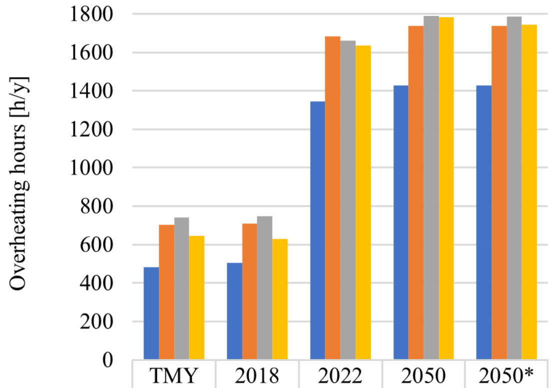
In the model without cooling, this study also discusses overheating in buildings. Fig. 11 compares the results of the different models in terms of overheating hours.

The note accompanying Fig. 11 is that, of course, the number of overheating hours in the 2050 and 2050* scenarios show virtually identical results (since only the primary energy conversion factor has changed). The results show that the model based on standard building use is, in all cases, significantly lower than the survey results for



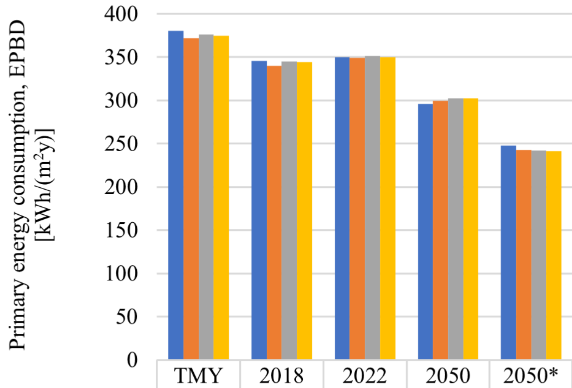
	TMY	2018	2022	2050	2050*
Standard parameters	377.29	342.56	338.17	282.19	244.07
Survey - Averaged parameters	359.03	325.71	322.57	269.19	234.44
Survey - Stohastic mean	360.07	327.44	324.22	271.26	233.34
Survey - Stohastic median	357.45	322.80	321.94	270.73	234.92

Fig. 10 Comparing models' primary energy consumption, EPBD – without cooling



	TMY	2018	2022	2050	2050*
Standard parameters	481	506	1346	1427	1427
Survey - Averaged parameters	704	709	1683	1738	1738
Survey - Stohastic mean	742	748	1659	1789	1785
Survey - Stohastic median	646	631	1636	1783	1744

Fig. 11 Comparing models' overheating hours – without cooling



	TMY	2018	2022	2050	2050*
Standard parameters	380.37	345.82	349.85	296.25	247.91
Survey - Averaged parameters	371.99	339.85	349.27	299.35	243.06
Survey - Stohastic mean	376.05	344.81	351.37	302.21	242.20
Survey - Stohastic median	374.55	344.29	350.17	302.48	241.36

Fig. 12 Comparing models' primary energy consumption, EPBD – with cooling

overheating hours; this difference is even more pronounced in warmer weather scenarios. Thus, standard behavior as described in EN 16798-1:2019 [26] leads to a serious underestimation of cooling comfort, which is also expected to result in an underestimation of cooling energy consumption.

In the stochastic model, the median is typically significantly lower than the average, especially when using TMY and 2018 weather data. The reason is that, in these two scenarios, overheating is still primarily driven by internal heat gains (i.e., lighting and appliance power density), for which we assumed a log-normal distribution (i.e., non-symmetric).

After all this, we also included cooling in the models, and Fig. 12 compares the results.

Taking cooling into account, the model's overestimation under standard parameters is no longer clear. This is because standard building use results in overestimation for heating and DHW production, and underestimation for cooling. This is analyzed in detail in Fig. 13, which compares models parameterized using standard and survey-average values for each end use.

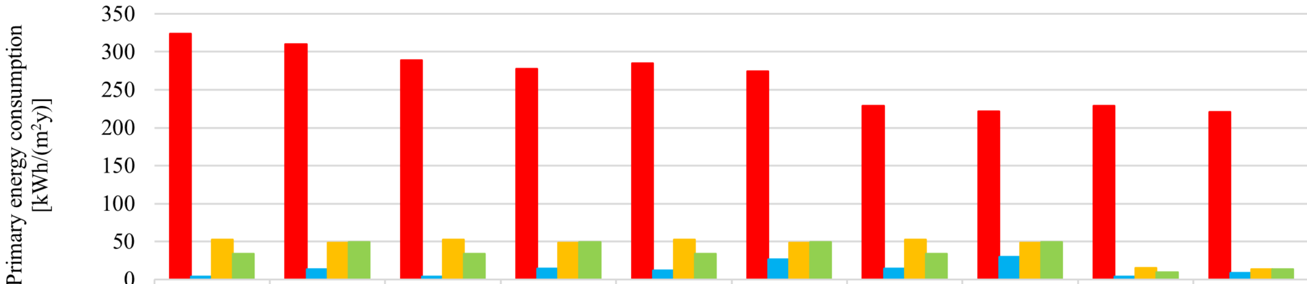
Fig. 13 confirms the above findings and also shows that the standard underestimates lighting and other household electricity consumption.

Returning to the analysis of Fig. 12, it can be said that although the two opposing effects (overestimation of heating and DHW production, underestimation of cooling) largely "cancel each other out" in the current model, it may be worthwhile to extend the research in the future to better insulated building types and multi-family houses, where the ratios presented here are expected to be different.

If we further analyze the results in Fig. 12, we can see that, for TMY and the 2018 weather, the model still slightly overestimates the standard building usage parameters. In the 2022 weather scenario, the model based on standard building use yields results that are approximately the same as those of the survey-average and stochastic models. In the extreme climate change scenario (2050), the model based on standard parameters underestimates, as cooling energy consumption has a much greater weight in total primary energy consumption. In the 2050* scenario, however, with the reduced primary energy conversion factor for electricity, cooling energy consumption again accounts for a smaller share of total primary energy consumption, so the model based on standard parameters again overestimates primary energy consumption.

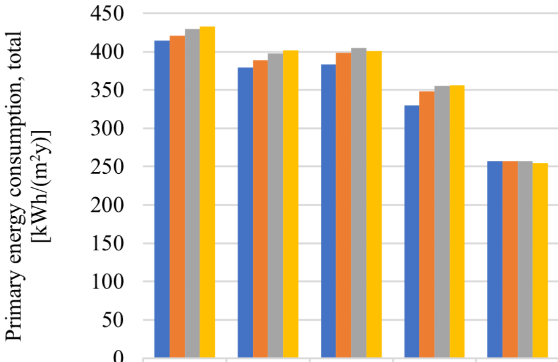
When we add lighting and household electricity consumption, the situation becomes a little more complicated. The results for total primary energy consumption, not according to the EPBD, are shown in Fig. 14.

If we also take into account lighting and household electricity consumption (i.e., the estimated electricity consumption of lighting and electrical equipment in addition to their heat load), it can be seen that the model based on standard parameters also underestimates this end use (since, instead of higher energy consumption than in previous models based on surveys, we obtain lower energy consumption, as shown in Fig. 13). This phenomenon is less noticeable in the TMY weather scenario, but becomes increasingly apparent in warmer weather scenarios. On the



	Standard	Survey	Standard	Survey	Standard	Survey	Standard	Survey	Standard	Survey
	TMY		2018		2022		2050		2050*	
Heating	323.99	310.51	289.31	277.39	285.02	274.38	229.16	221.15	228.87	220.86
Cooling	3.56	13.38	3.69	14.36	12.00	26.79	14.26	30.10	4.05	8.54
DHW	52.82	48.10	52.82	48.10	52.82	48.10	52.82	48.10	14.99	13.66
Light. & appl.	33.66	49.03	33.66	49.03	33.66	49.03	33.66	49.03	9.55	13.92

Fig. 13 Comparison of primary energy consumption by end use



	TMY	2018	2022	2050	2050*
Standard parameters	414.03	379.48	383.50	329.91	257.46
Survey - Averaged parameters	421.03	388.88	398.31	348.39	256.98
Survey - Stohastic mean	429.28	398.03	404.45	355.45	257.40
Survey - Stohastic median	432.91	401.94	400.73	356.47	255.08

Fig. 14 Comparing models' primary energy consumption, total – with cooling

one hand, the underestimation will result in lower lighting and household electricity consumption, but on the other hand, it will also reduce internal heat gains used for heating, thus increasing heating energy consumption. In warmer weather scenarios, however, the weight of heating energy consumption decreases, resulting in a greater difference between the models parameterized based on the standard and the survey averages. In the 2050* scenario, the weight of heating increases again with reduced electricity primary energy conversion factors because natural gas is the energy source used for heating, so the two opposing effects "balance each other out", i.e., we get approximately the same primary energy consumption with standard parameters and those according to the survey.

Taking lighting and household electricity consumption into account, even stochastic models show deviations from those parameterized with the survey's average values. The reason for this is that the uncertainty of the specified power density of lighting and electrical equipment is very high, and moreover, it follows a strongly log-normal distribution (as explained earlier). Taking this into account, the previously symmetrical, normal distribution of the simulation energy consumption results is distorted, so the consumption results from the model parameterized with the survey averages are less consistent with the consumption predicted by the stochastic model average.

3.2 Results of sensitivity analysis

When analyzing the results of sensitivity tests, one key value is the adjusted R-squared. It reflects the overall model fit, showing how much of the output variation is explained by the analyzed input parameters, and it ranges from 0 to 1. In all analyzed cases and across all output variables, the value was between 0.95 and 0.99, indicating that most key parameters were identified in our models.

Another important feature is the p-value, which shows whether the given input parameter has a statistically significant effect on the output variable. If the p-value is greater than 0.05, the reliability of the regression results for the input parameter is low (i.e., the given input parameter is probably not decisive for the given output variable). In our study, p-values greater than 0.05 occurred only in a few cases.

In Sections 3.2.1 to 3.2.3, we provide a detailed analysis of the Standardized Regression Coefficient (SRC) values for the given input parameters in detail. This value shows the relative sensitivity of the input parameters to the given output variable. Based on its absolute value, it is possible to rank the most decisive parameters, but it is important to note that its value can also be negative. For example, the heating system efficiency value has a negative SRC for primary energy consumption, which means that a positive change in heating system efficiency results in a decrease in primary energy consumption. The results of certain key cases are presented in detail below.

3.2.1 Influence of cooling

To illustrate the influence of cooling, we compare results from models without cooling to those with cooling. The results for the primary energy consumption output variable according to the EPBD for the TMY weather scenario are shown in Fig. 15.

If there is no cooling in the model, then naturally the Cooling setpoint temperature and Cooling system efficiency will not appear as input parameters. In this case, Window shading and Summer night ventilation rate will not be statistically significant either, as their *p*-values will be greater than 0.05. The heating setpoint temperature is the most influential input parameter across all cases, as observed in the literature review. In addition, DHW demand, Heating system efficiency, and Ventilation rate are also very decisive parameters. Attic floor U-value, Infiltration rate, and Lighting and appliances power density can also be classified as key parameters.

If cooling is also accounted for in the model, the cooling setpoint temperature appears among the input parameters

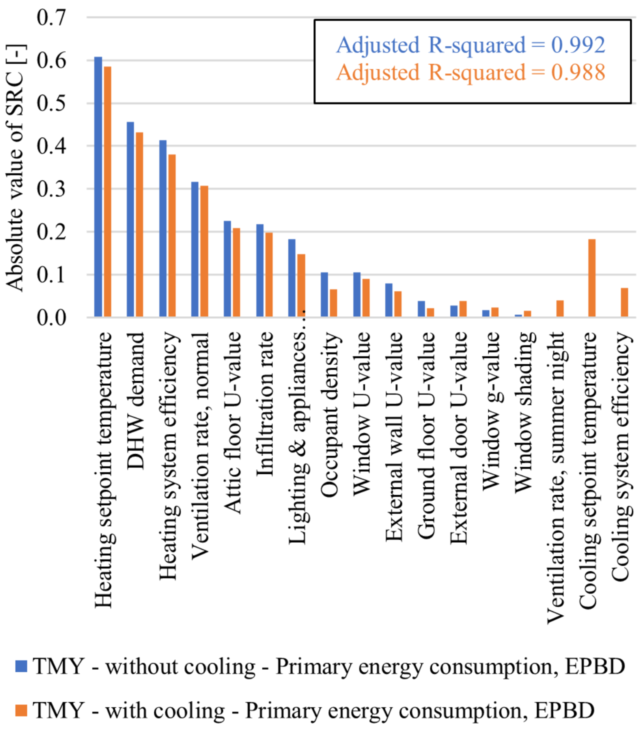


Fig. 15 Results of sensitivity analysis – influence of cooling

and becomes one of the most decisive. The cooling system efficiency will be a slightly less decisive parameter, but its role cannot be neglected either. However, the cooling setpoint temperature lags significantly behind the absolute value of the heating setpoint temperature SRC, meaning that when using TMY weather data series, heating clearly plays a dominant role. The other parameters also show this: in the model with cooling, the role of Window g-value, Window shading, and Summer night ventilation rate increases, but these parameters are not decisive for primary energy consumption in the TMY scenario.

When analyzing the role of cooling, it is also important to examine which parameters are most decisive for the number of overheating hours in the model without cooling. The most decisive parameters in terms of the number of overheating hours are Summer night ventilation rate, Window shading, and Lighting and appliances power density. Moderately decisive parameters include Window g-value, Occupant density, and Infiltration rate. The minimally influential input parameter is the Ground floor U-value, which affects heat transfer to the cooler ground in summer. The most important and decisive parameters were the same for all weather scenarios, with only minimal changes in their order between cases. Thus, with respect to passive thermal protection, the most promising intervention points for this model are the Summer night ventilation rate, Window shading, and Window g-value parameters.

However, for the number of overheating hours, several parameters (window U-value, DHW demand, external door U-value, heating setpoint temperature, external wall U-value, heating system efficiency and attic floor U-value) are not influencing parameters based on their *p*-values.

In the following, we present the results of models that include cooling, as the high number of overheating hours observed in the warmer weather scenarios indicates that cooling will be essential (and, in some years, already critical).

3.2.2 Influence of weather data

When analyzing the impact of different weather data sets, we first examine differences between actual weather data files to see the results of the 2018 and 2022 data sets for primary energy consumption according to the EPBD in Fig. 16.

There are no significant differences between the results of the two weather data sets, but the absolute values of the SRC values for the Heating setpoint temperature and Heating system efficiency parameters decrease, while the absolute values of the SRC values for the cooling system efficiency, attic floor U-value, window shading, and window

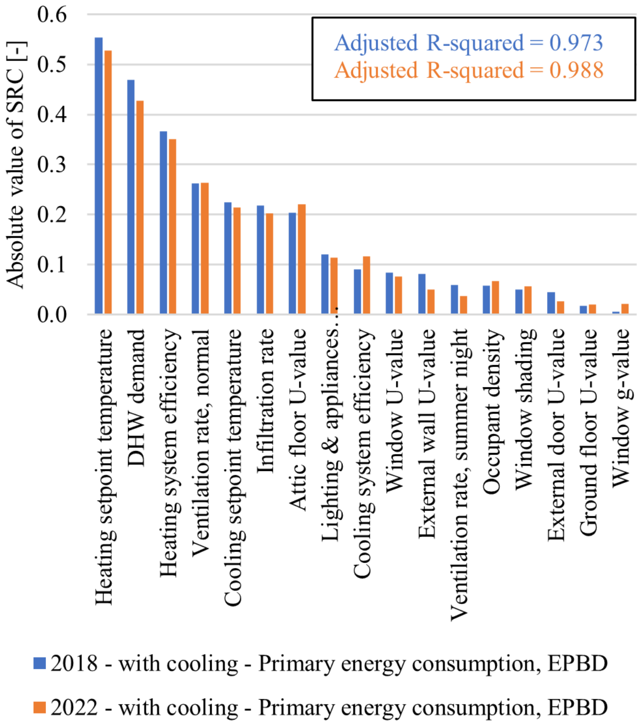


Fig. 16 Results of sensitivity analysis – influence of actual weather data

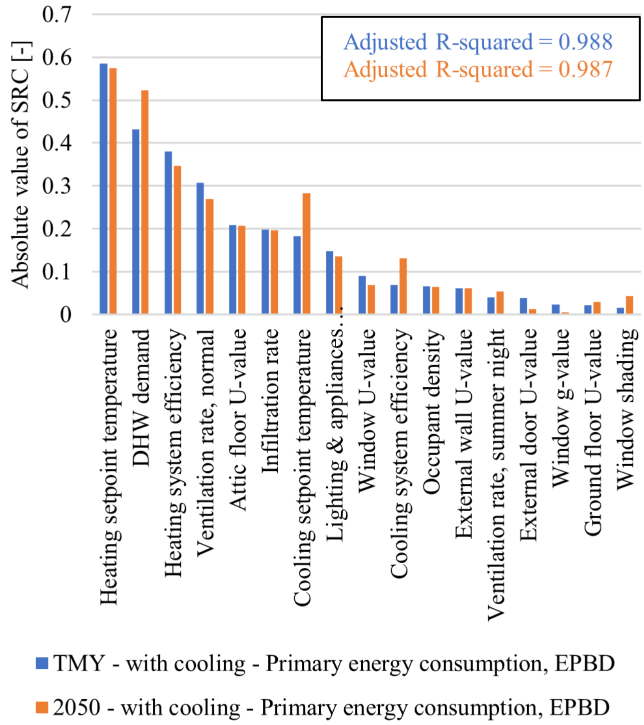


Fig. 17 Results of sensitivity analysis – influence of climate change

g-value parameters increase in the 2022 weather scenario compared to the 2018 scenario. As a result of the change in SRC values, the parameter order has also changed. For example, in the 2022 scenario, cooling system efficiency is placed before the lighting and appliances power density, which has a decreasing absolute SRC value. This shows that in a year with a hotter summer, the internal heat load is relatively less significant in the model, while the cooling system's efficiency becomes more significant.

The slight changes observed in the two actual weather data series alone indicate that the strength and order of the input parameters also depend on the yearly variation of the weather. It is therefore worth examining the differences between an extreme climate change scenario and the typical current weather scenario – Fig. 17 shows the primary energy consumption results for these two scenarios according to the EPBD.

The impact of extreme climate change results in considerably stronger changes in the TMY–2050 comparison than in the earlier case, where differences were assessed only between two actual meteorological years. There is a striking difference in terms of DHW demand: in the 2050 scenario, the previously dominant high heating demand (and its role) decreases significantly, while low cooling energy consumption increases. Although the energy used to produce DHW does not change, it will account for a larger share of total primary energy consumption under the EPBD in the 2050 scenario, thus increasing the role of the related input parameter.

A further change in the 2050 case is that, as the absolute SRC value of the Cooling setpoint temperature increases, this parameter becomes the fourth most influential, closely following the Heating system efficiency. The absolute SRC value of the cooling system efficiency also increases markedly, roughly doubling. In addition, the influence of the summer night ventilation rate and window shading parameters increases, while the relative importance of internal heat gains decreases, even in the 2050 scenario (whereas the role of solar gains increases). The absolute SRC values for the heating setpoint temperature, heating system efficiency, and ventilation rate decrease, consistent with the trends observed above. Climate change is expected to further increase the dominance of cooling.

3.2.3 Influence of primary energy conversion factor

In future scenarios, as already highlighted, it is also worth considering changes in the primary energy conversion factor. The effect of its declining value in the case of electricity is shown in Fig. 18, which presents primary energy consumption according to the EPBD.

When comparing the results of the resulting 2050* scenario with those of the 2050 scenario, it can be seen that the absolute SRC values of parameters related to both the cooling system (e.g., cooling setpoint temperature, cooling

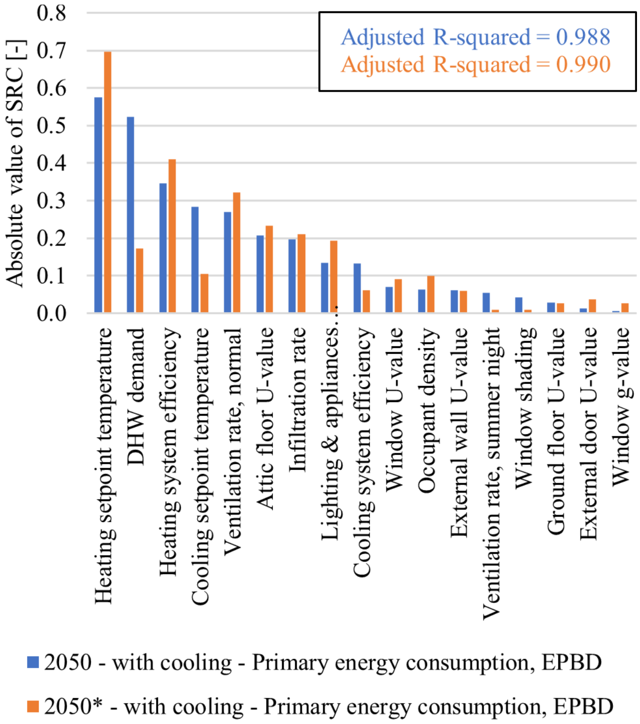


Fig. 18 Results of sensitivity analysis – influence of primary energy conversion factor (EPBD)

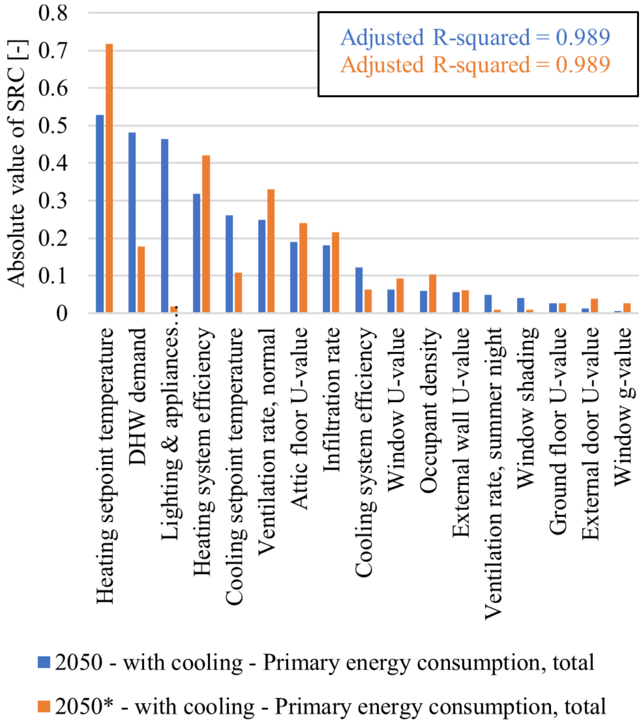


Fig. 19 Results of sensitivity analysis – influence of primary energy conversion factor (total)

system efficiency, summer night ventilation rate, window shading) and the DHW system (e.g., DHW demand) decrease substantially. This is because these end uses rely on electricity, and their contribution is therefore weighted by a primary energy factor of 0.65 instead of the previous 2.3. By contrast, space heating is supplied by natural gas; consequently, its relative weight increases. Accordingly, parameters such as the heating setpoint temperature, heating system efficiency, ventilation rate (normal), envelope U-values, and the infiltration rate become more influential in the 2050* scenario. The power density of lighting and plug loads – i.e., internal heat gains – which reduce heating demand, also gains relative importance in the 2050* case; however, its effect on EPBD primary energy use appears only indirectly.

However, if we also take lighting and other household electricity consumption into account in total primary energy consumption, the above-mentioned parameter (lighting and equipment power density) no longer appears only indirectly in the model as internal heat load, but also as electricity consumption. The results are shown in Fig. 19.

Considering the results for non-EPBD primary energy use, a notable change is that the lighting and appliances power density parameter becomes the third most influential parameter in the 2050 scenario (a substantial increase compared to the EPBD primary energy results). In the 2050* scenario, however, this parameter's influence decreases sharply due to the reduction in the electricity primary energy factor: in addition to the lower conversion factor, its role as an internal heat gain on the (already more significant) cooling energy use in the 2050 scenario also diminishes, since cooling is likewise electricity-based.

4 Conclusions

During the parametric analysis of the residential building stock, we observed that the distributions of the 17 investigated input parameters are generally well approximated by normal distributions, whereas log-normal distributions are more typical for building-use parameters. Specifically, ventilation rate, occupant density, lighting and appliances power density, and DHW demand were best represented by log-normal distributions. The ground-floor U-value also followed a log-normal distribution. In addition, the certificate-derived datasets require systematic quality control and filtering, as they frequently include erroneous or inconsistent entries.

When determining uncertainty, we found that the relative standard deviation in primary energy consumption results for the archetype building examined is approximately 14–16% in all scenarios. We also examined the relative standard deviation of the results for overheating hours and found that differences between model variants are even more significant in cooler summer scenarios. That is, with more favorable parameters, the building without cooling hardly overheats in summer, so passive thermal protection could prove sufficient in such scenarios. In contrast, in warmer summers or in the event of extreme climate change, the building overheats in all variants for a high number of hours per year, so passive heat protection is no longer sufficient and mechanical cooling is required.

Looking at the results of the sensitivity analyses, there are no significant differences between the actual historical weather files, but in a warmer year, for example, the role of the cooling system efficiency and window shading parameters becomes slightly more decisive. However, if we consider a scenario of extreme climate change compared to the current typical meteorological year, there are much more pronounced changes: for example, the cooling setpoint temperature becomes the fourth most important parameter, and its weight is almost equal to that of the heating system parameters. The assumed future decrease in the primary energy conversion factor for electricity also fundamentally changes the order of the determining parameters, as the parameters of subsystems that use electricity are naturally pushed back in determining total primary energy consumption.

With regard to building-use models, we have seen that the model parameterized according to the standard overestimates primary energy consumption for heating and DHW, but underestimates it for cooling, lighting, and other household electricity consumption. In terms of total primary energy consumption, it is highly dependent on weather scenarios, the consideration of cooling, and the non-renewable primary energy conversion factor of electricity. It also significantly underestimated the number of overheating hours, with the survey results showing much higher annual overheating hours. Based on this, it can be said that, as a further development of the Hungarian residential building stock model, it is definitely worthwhile to use survey results when specifying the building-use parameters of the model.

In contrast, stochastic models did not cause any significant difference in most scenarios compared to the model parameterized with survey averages, because the input parameters typically followed a symmetric (normal) distribution, so the average building result is close to the average result of the stochastic models. Only in the case of overheating hours was there a minimal difference with the stochastic models, as in this case the output results show a less symmetrical distribution. It is definitely worthwhile to conduct further research on the application of stochastic models in the further development of the residential building stock, extending the investigations to multi-family houses and thermally insulated buildings.

The methodology itself can be transferred to other building types; however, in the case of multi-family types, completely different HVAC system configurations may occur, and substantial differences can also be found in other parameters – such as the window-to-wall ratio – which significantly affects the building's heat balance. For newly built types, well-insulated envelope assemblies are typical, while their building technical systems can also differ considerably. As a result, these type-specific differences may carry over to the modeling outcomes and could lead to somewhat more significant differences between the survey results for average buildings and stochastic model results.

Accordingly, other conclusions drawn from the examined building type cannot necessarily be extended to the entire residential building stock either, or at least they may not hold in exactly the same way: the ranking and relative importance of parameters may differ by type, and the relative standard deviation of primary energy use and overheating hours may also change to some extent. The main takeaways – such as the increasing share of model variants affected by overheating under hotter summers, or the biasing effect of standardized building-use parameters – will likely evolve in a similar manner across types; nevertheless, identifying the precise differences would require additional analyses for other building types as well.

Acknowledgements
The project supported by the Doctoral Excellence Fellowship Programme (DCEP) is funded by the National Research Development and Innovation Fund of the Ministry of Culture and Innovation and the Budapest University of Technology and Economics.

The work has been carried out within the research project "Integrated development of the residential building stock and the electricity mix models for decarbonized building stock scenarios". The project (no. K 142992) has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the K_22 funding scheme.

References

[1] Cichowicz, R., Jerominko, T. "Comparison of calculation and consumption methods for determining Energy Performance Certificates (EPC) in the case of multi-family residential buildings in Poland (Central-Eastern Europe)", Energy, 282, 128393, 2023. https://doi.org/10.1016/j.energy.2023.128393

[2] Bertoldi, P., Maduta, C., Tsemekidi-Tzeiranaki, S., D'Agostino, D., Paci, D., Clementi, E. L., Castellazzi, L. "Transformative investment strategy for carbon–neutral residential buildings in the European Union", Energy and Buildings, 347, 116232, 2025. https://doi.org/10.1016/j.enbuild.2025.116232

[3] Pongelli, A., Priore, Y. D., Bacher, J.-P., Jusselme, T. "Definition of Building Archetypes Based on the Swiss Energy Performance Certificates Database", Buildings, 13(1), 40, 2023. https://doi.org/10.3390/buildings13010040

[4] Raushan, K., Uidhir, T. M., Salvador, M. L., Norton, B., Ahern, C. "Filtered dataset of Irish energy performance certificates: A data-driven approach for enhanced building stock modelling", Data in Brief, 59, 111281, 2025. https://doi.org/10.1016/j.dib.2025.111281

[5] Government of Hungary "7/2006 (V. 24.) TNM rendelet az épületek energetikai jellemzőinek meghatározásáról" (7/2006 (V. 24.) TNM decree on the energy performance of buildings), Budapest, Hungary, 2006. [online] Available at: https://njt.hu/jogszabaly/2006-7-20-6F [Accessed: 01 February 2026] (in Hungarian)

[6] The European Parliament and the Council of the European Union "Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings", Brussels, Belgium, 2002. [online] Available at: https://eur-lex.europa.eu/eli/dir/2002/91/oj/eng [Accessed: 01 February 2026]

[7] Government of Hungary "9/2023. (V. 25.) ÉKM rendelet az épületek energetikai jellemzőinek meghatározásáról" (9/2023. (V. 25.) ÉKMdecree on the energy performance of buildings), Budapest, Hungary, 2023. [online] Available at: https://njt.hu/jogszabaly/2023-9-20-8X [Accessed: 01 February 2026] (in Hungarian)

[8] Krapf, S., Mayer, K., Fischer, M. "Points for energy renovation (PointER): A point cloud dataset of a million buildings linked to energy features", Scientific Data, 10(1), 639, 2023. https://doi.org/10.1038/s41597-023-02544-x

[9] Lim, H., Zhai, Z. J. "Review on stochastic modeling methods for building stock energy prediction", Building Simulation, 10(5), pp. 607–624, 2017. https://doi.org/10.1007/s12273-017-0383-y

[10] Pasichnyi, O., Wallin, J., Levihn, F., Shahrokni, H., Kordas, O. "Energy performance certificates — New opportunities for data-enabled urban energy policy instruments?", Energy Policy, 127, pp. 486–499, 2019. https://doi.org/10.1016/j.enpol.2018.11.051

[11] Amecke, H. "The impact of energy performance certificates: A survey of German home owners", Energy Policy, 46, pp. 4–14, 2012. https://doi.org/10.1016/j.enpol.2012.01.064

[12] Mangold, M., Österbring, M., Wallbaum, H. "Handling data uncertainties when using Swedish energy performance certificate data to describe energy usage in the building stock", Energy and Buildings, 102, pp. 328–336, 2015. https://doi.org/10.1016/j.enbuild.2015.05.045

[13] Hardy, A., Glew, D. "An analysis of errors in the Energy Performance certificate database", Energy Policy, 129, pp. 1168–1178, 2019. https://doi.org/10.1016/j.enpol.2019.03.022

[14] Dahlström, L., Broström, T., Widén, J. "Advancing urban building energy modelling through new model components and applications: A review", Energy and Buildings, 266, 112099, 2022. https://doi.org/10.1016/j.enbuild.2022.112099

[15] Van Hove, M. Y. C., Delghust, M., Laverge, J. "Quality assurance for building-stock energy models: a performance comparison of eleven uncertainty and sensitivity analysis methods", Journal of Building Performance Simulation, 17(2), pp. 149–175, 2024. https://doi.org/10.1080/19401493.2023.2248063

[16] Hopfe, C. J. "Uncertainty and sensitivity analysis in building performance simulation for decision support and design optimization", PhD Thesis, Eindhoven University of Technology, 2009. https://doi.org/10.6100/IR643321

[17] Tian, W., Heo, Y., de Wilde, P., Li, Z., Yan, D., Park, C. S., Feng, X., Augenbroe, G. "A review of uncertainty analysis in building energy assessment", Renewable and Sustainable Energy Reviews, 93, pp. 285–301, 2018. https://doi.org/10.1016/j.rser.2018.05.029

[18] Pang, Z., O'Neill, Z., Li, Y., Niu, F. "The role of sensitivity analysis in the building performance analysis: A critical review", Energy and Buildings, 209, 109659, 2020. https://doi.org/10.1016/j.enbuild.2019.109659

[19] Roka, R., Figueiredo, A., Vieira, A., Cardoso, C. "A Systematic Review of Sensitivity Analysis in Building Energy Modeling: Key Factors Influencing Building Thermal Energy Performance", Energies, 18(9), 2375, 2025. https://doi.org/10.3390/en18092375

[20] Zhao, F., Lee, S. H., Augenbroe, G. "Reconstructing building stock to replicate energy consumption data", Energy and Buildings, 117, pp. 301–312, 2016. https://doi.org/10.1016/j.enbuild.2015.10.001

[21] Tian, W., Choudhary, R. "A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London", Energy and Buildings, 54, pp. 1–11, 2012. https://doi.org/10.1016/j.enbuild.2012.06.031

[22] Booth, A. T., Choudhary, R., Spiegelhalter, D. J. "Handling uncertainty in housing stock models", Building and Environment, 48, pp. 35–47, 2012. https://doi.org/10.1016/j.buildenv.2011.08.016

[23] Csoknyai, T., Farkas, J., Formanek, L., Horváth, M. "Lakossági épület energetikahatékonysági potenciál felmérése" (Assessment of energy efficiency potential of a residential building), Government of Hungary, Budapest, Hungary, Rep. KEOP-7.9.0/12-2013-0019, 2015. (in Hungarian)

[24] DesignBuilder "DesignBuilder, (v2025.1.0.92)", [computer program] Available at: https://designbuilder.co.uk/download/release-software [Accessed: 01 February 2026]

[25] Horváth, M., Gergely, L. Z., Takács, Z., Szagri, D., Szalay, Z., Csoknyai, T. "High saving potential at low Cost: Decreasing heating demand by operation in the existing building stock", Journal of Building Engineering, 111, 113153, 2025. https://doi.org/10.1016/j.jobe.2025.113153

[26] CEN "EN 16798-1:2019 Energy performance of buildings - Ventilation for buildings - Part 1: Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics - Module M1-6", European Committee for Standardization, Brussels, Belgium, 2019.

[27] Csoknyai, T., Szalay, Z., Gergely, L. Z., Szagri, D., Horváth, M., Takácsné Tóth, B., Kotek, P. "A rezsicsökkentés szabályváltozásának hatása a magyar lakóépületszektor gázfogyasztására" (The impact of the changes in the utility cost reduction system on the gas consumption of the Hungarian residential sector), [pdf] REKK: Regionális Energia- és Infrastruktúra-Politikai Együttműködésért Alapítvány, 2022. Available at: https://rekk.hu/downloads/academic_publications/rekk_policybrief_hu_2022_05.pdf [Accessed: 01 February 2026] (in Hungarian)

[28] Czétány, L., Vámos, V., Horváth, M., Szalay, Z., Mota-Babiloni, A., Deme-Bélafi, Z., Csoknyai, T. "Development of electricity consumption profiles of residential buildings based on smart meter data clustering", Energy and Buildings, 252, 111376, 2021. https://doi.org/10.1016/j.enbuild.2021.111376

[29] The European Parliament and the Council of the European Union "Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency (Text with EEA relevance)", Brussels, Belgium, 2018. [online] Available at: https://eur-lex.europa.eu/eli/dir/2018/844/oj/eng [Accessed: 01 February 2026]

[30] Hungarian Central Statistical Office "Népszámlálás 2022" (Hungarian Census Report 2022), (online) Available at: https://nepszamlalas2022.ksh.hu/ [Accessed: 01 February 2026] (in Hungarian)

[31] Lumivero "XLSTAT, (2025.1.3)", [computer program] Available at: https://www.xlstat.com/download [Accessed: 07 February 2026]

[32] Massey Jr., F. J. "The Kolmogorov-Smirnov Test for Goodness of Fit", Journal of the American Statistical Association, 46(253), pp. 68–78, 1951. https://doi.org/10.1080/01621459.1951.10500769

[33] ÉMI Építésügyi Minőségellenőrző és Innovációs Nonprofit Kft. "Nemzeti Épületenergetikai Stratégia" (National Building Energy Strategy), [pdf] ÉMI Építésügyi Minőségellenőrző és Innovációs Nonprofit Kft., Szentendre, Hungary, 2015. Available at: https://2015-2019.kormany.hu/download/d/85/40000/Nemzeti%20E%CC%81pu%CC%88letenergetikai%20Strate%CC%81gia%20150225.pdf [Accessed: 01 February 2026] (in Hungarian)

[34] Auricon Mérnöki Ltd. "Auricon Energetic, (5.5.14)", [computer program] Available at: https://energetic.auricon.hu/webshop/Auricon-Energetic-EVES-p536079900 [Accessed: 07 February 2026]

[35] Bausoft Ltd. "WinWatt, (9.54)", [computer program] Available at: http://www.bausoft.hu/telepit.htm [Accessed: 07 February 2026]

[36] Varga, D. "Elektromos fűtési megoldások energetikai és gazdaságossági értékelése" (Energy and economical analysis of electric heating systems), BSc Thesis, Budapest University of Technology and Economics, 2020. (in Hungarian)

[37] DIN "DIN 1946-6 Ventilation and air conditioning - Part 6: Ventilation for residential buildings - General requirements, requirements for design, construction, commissioning and handover as well as maintenance", Deutsches Institut für Normung, Berlin, Germany, 2019.

[38] Horváth, M. "Calculation of Solar Yield on Tilted and Oriented Surfaces, Utilization in Buildings", [pdf] Budapest University of Technology and Economics, Budapest, Hungary, 2016. Available at: https://repozitorium.omikk.bme.hu/bitstream/handle/10890/5351/tezis_eng.pdf?sequence=4&isAllowed=y [Accessed: 01 February 2026]

[39] Nouri, A., van Treeck, C., Frisch, J. "Sensitivity Assessment of Building Energy Performance Simulations Using MARS Meta-Modeling in Combination with Sobol' Method", Energies, 17(3), 695, 2024. https://doi.org/10.3390/en17030695

[40] Pannier, M.-L., Recht, T., Robillart, M., Schalbart, P., Peuportier, B., Mora, L. "Identifying optimal renovation schedules for building portfolios: Application in a social housing context under multi-year funding constraints", Energy and Buildings, 250, 111290, 2021. https://doi.org/10.1016/j.enbuild.2021.111290

[41] Hosamo, H., Mazzetto, S. "Performance Evaluation of Machine Learning Models for Predicting Energy Consumption and Occupant Dissatisfaction in Buildings", Buildings, 15(1), 39, 2025. https://doi.org/10.3390/buildings15010039

[42] Meteotest "Meteonorm, (v8.2.0)", [computer program] Available at: https://mn8.meteonorm.com/en/meteonorm-version-8 [Accessed: 01 February 2026]