Application of the Takagi-Sugeno Fuzzy Modeling to Forecast Energy Efficiency in Real Buildings Undergoing Thermal Improvement

Tomasz Szul * , **Krzysztof Nęcka** and **Stanisław Lis**

Faculty of Production and Power Engineering, University of Agriculture, Krakow, 30-149 Kraków, Poland; k.necka@urk.edu.pl (K.N.); s.lis@urk.edu.pl (S.L.)
* Correspondence: t.szul@urk.edu.pl; Tel.: +48-12-662-4647

Abstract: Energy efficiency in the building industry is related to the amount of energy that can be saved through thermal improvement. Therefore, it is important to determine the energy saving potential of the buildings to be thermally upgraded in order to check whether the set targets for the amount of energy saved will be reached after the implementation of corrective measures. In real residential buildings, when starting to make energy calculations, one can often encounter the problem of incomplete architectural documentation and inaccurate data characterizing the object in terms of thermal (thermal resistance of partitions) and usable (number of inhabitants). Therefore, there is a need to search for methods that will be suitable for quick technical analysis of measures taken to improve energy efficiency in existing buildings. The aim of this work was to test the usefulness of the type Takagi-Sugeno fuzzy models of inference model for predicting the energy efficiency of actual residential buildings that have undergone thermal improvement. For the group of 109 buildings a specific set of important variables characterizing the examined objects was identified. The quality of the prediction models developed for various combinations of input variables has been evaluated using, among other things, statistical calibration standards developed by the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE). The obtained results were compared with other prediction models (based on the same input data sets) using artificial neural networks and rough sets theory.

Keywords: energy efficiency; energy saving; energy consumption; forecasting of energy consumption; thermal improved of buildings; Takagi-Sugeno fuzzy models

1. Introduction

Energy efficiency means the amount of energy saved determined by measuring or estimating consumption before and after the implementation of an energy efficiency improvement measure, while ensuring the normalization of external conditions affecting energy consumption [1]. This definition applies to all branches of the European economy, but has a particular impact on the building sector, as energy consumption in the European building sector accounts for about 40% of the total energy demand of the European Union [2]. According to [3], the building sector consumes 35.3% of the final energy demand. This activity is responsible for the emission of about 36% of CO_2 pollution in the European Union [2]. The European Union's (EU) climate-energy policy, including its long-term vision to strive for EU climate neutrality by 2050 and the regulatory mechanisms to stimulate the achievement of effects in the coming decades, has a significant impact on shaping the energy strategy. Achieving the 2020 and 2030 climate and energy targets in the EU is key to a low-carbon energy transition. In line with the EU's decarbonization ambition, in December 2020. The European Council approved in December 2020 a binding EU target to reduce net greenhouse gas emissions by at least 55% by 2030 compared to 1990 levels, thereby increasing the existing 40% reduction target. Additionally, there is a provision on

increasing the share of renewable sources in the final energy consumption to at least 32% and increasing the energy efficiency by 32.5% [4]. Reducing energy consumption and using renewable energy sources to a greater extent is intended to increase the security of energy supply and support technical development in individual countries. Reducing energy consumption in the residential sector is a priority action in the EU member states. These actions allow efficient and sustainable use of fossil fuel potential and reduce gas and dust emissions. In order to improve the efficiency of energy use for heating and air conditioning of buildings and reduce greenhouse gas emissions, in 2018 the European Parliament and the Council amended the Directive 2018/844 on the energy performance of buildings [5]. The provisions set out in the Directive should contribute to the improvement of energy performance of buildings also in Poland, where the covering of thermal needs is done at the local level, therefore it is extremely important to ensure energy planning at the level of municipalities and regions—it is crucial for rational energy management, improvement of air quality and exploitation of local potential. A useful tool will also be the launch of a nationwide heat map (developed in connection with Article 14 of the revised Energy Efficiency Directive 2012/27/EU [6]), which will facilitate planning for covering heat needs. National analyses show that current technical conditions [7] do not guarantee energy consumption at the optimal, economically justified level. The analysis presented in the studies [8–10] shows that the average value of the final energy demand indicator in Poland varies from 85 to 343 kWh/(m^2 year), depending on the type of building, with single and multi-family residential buildings obtaining values much higher than European averages. In residential buildings, the most important area of final energy consumption in a building is heating, which represents 76% of its total consumption [7]. By minimizing heat losses and modernizing the heating system, energy efficiency in residential buildings can be improved. In order to achieve the objectives of Directive 2018/844 [5], the construction sector in Europe is waiting for the transition to energy efficient and passive buildings. No less important is to improve the energy standard of existing buildings through their thermomodernisation. In the case of buildings existing prior to energy saving measures, an assessment of the energy performance of the buildings is necessary to obtain adequate knowledge of the existing energy consumption profile of the building or group of buildings, to determine how and in what quantity cost-effective energy savings can be achieved, and to inform about the results. Therefore, it is necessary to search for tools that will allow estimating the potential for savings in energy consumption for heating buildings. These tools will help policy makers to implement the intentions of the European Union's energy policy, which aims to assess the energy saving potential and thus reduce pollutant emissions from the building sector.

2. Literature Review of Estimation Methods for Building's Energy Demand

The assessment of energy efficiency in buildings can be done in different ways, which can be classified as: engineering calculations, statistical and data-based models (machine learning), and hybrid models [10–36]. Figure 1 shows the techniques of modelling and forecasting energy consumption in buildings with a division into individual calculation methods. Additionally, the diagram contains quotations of works in which the authors used a given method to forecast energy consumption for heating or to determine energy efficiency (energy performance) in various types of buildings.

Literature analysis has shown that statistical and artificial neural network-based models are most commonly used, while fuzzy logic-based models are used much less frequently.
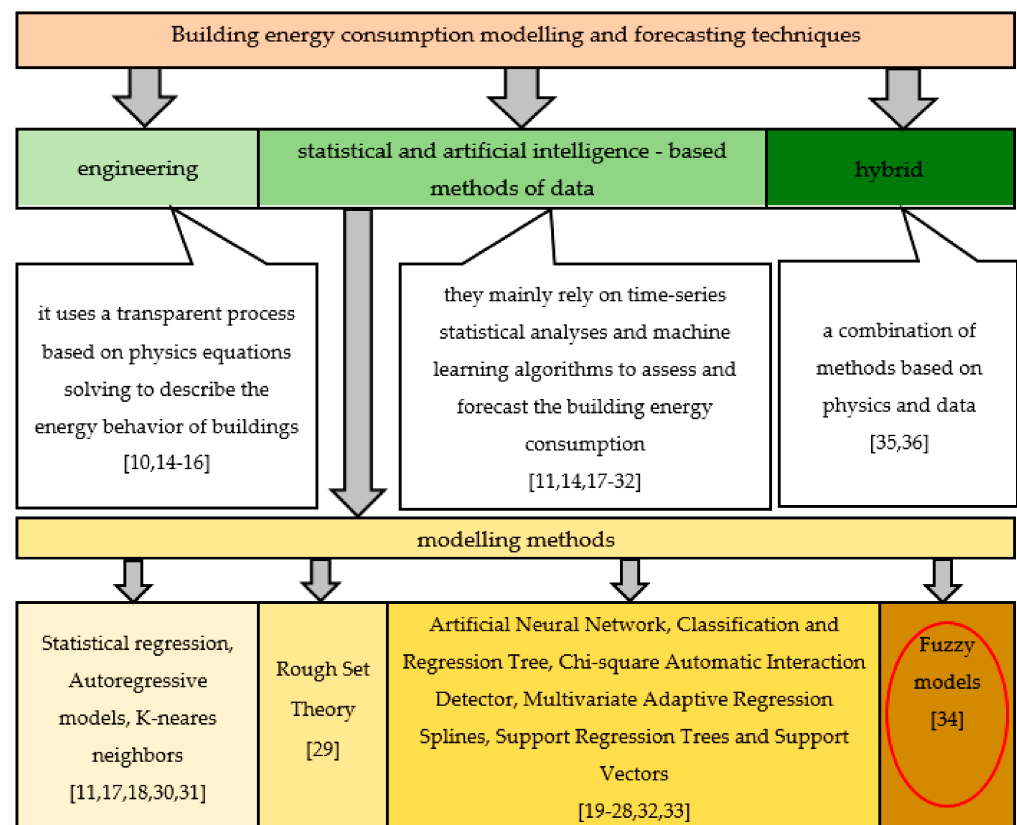
Figure 1. Classification of methods for forecasting energy consumption in buildings.

A fuzzy approach for determining the energy performance of buildings was applied by Nebot and Mugica [34], where two methods were used for prediction: fuzzy inductive inference (FIR) and adaptive neuro-fuzzy inference system (ANFIS). The study was performed on a set of simulated residential buildings developed by Tsanas and Xifara [30]. The computational results obtained by the authors [34] have been compared with the results obtained for the same input data (for simulated buildings) using static and artificial neural network-based methods [25,28,30,31]. The results showed that the use of methods based on fuzzy logic gives very good results in predicting energy consumption [34]. The presented forecasting models are mainly focused on estimating energy consumption, energy characteristics in simulated objects, for which reliable data on building envelope insulation, ventilation air flows and number of inhabitants can be obtained [28,30–34]. The presented research results are promising and indicate high accuracy in estimating energy consumption in buildings, however, there is a lack of studies on real buildings [14,25–27,29], for which it is difficult to obtain unambiguous data. Individual energy demand in residential buildings is more difficult to estimate due to the lack of data on building occupancy and complexity of occupant behavior. In particular, there is a lack of studies on prefabricated buildings. Buildings erected in the 1960s and 1970s of the last century in the large panel technology, by definition, are not able to meet the thermal and energy requirements of today [29]. Conducted research on the analysis of thermal insulation of external walls of buildings built in large panel technology, confirmed the deterioration of their thermal insulation, mainly due to the use of concrete with increased density. Defects and errors resulting from the execution or damage of thermal insulation layers and the location of thermal bridges also had an impact. The occurrence of these places was caused by insufficient thickness of thermal insulation or its absence [26]. An additional factor making it difficult to estimate energy consumption for heating is the way the apartments are used. It results from inappropriate approach to ventilation of rooms, e.g., covering ventilation grilles, which directly influences the size of ventilation air stream, which also leads to dampness of partitions. A frequent problem in thermal calculations is the lack of complete architectural

and building documentation, which hinders correct estimation of energy consumption for heating [26,29]. Therefore, in our research we want to use fuzzy set theory supported by neural network to predict the energy efficiency of real multifamily residential buildings made with prefabricated technology that have undergone thermal improvement process. We believe that the fuzzy logic-based method provides a convenient tool to describe the uncertainty and imprecision of the input data inherent in the energy demand prediction process. In this research, we propose to apply models with Takagi-Sugeno type inference for residential energy efficiency forecasting. It should be emphasized that, as shown in the literature review, it has not been applied in energy assessment of existing buildings before, and thus is new to this type of research. Due to the different availability and accuracy of data describing buildings, we want to investigate and test different configurations of input variables in order to obtain a trade-off between the auditor's effort to obtain them and the accuracy of the forecast. The results obtained will be compared with other forecasting methods (Neural Methods and Rough Set Theory) [26,29], which were based on the same input data characterizing a set of real residential buildings undergoing thermal improvements.

3. Materials and Methods

The research was carried out on a group of multi-family residential buildings built in the 1960s to 1980s, which were constructed in the technology of large plates (prefabricated elements). The energy characteristics of the analyzed group of buildings have been discussed in the paper [29]. Before proceeding to the analyses, the collected database was subjected to preliminary processing. It consisted among others in elimination of those buildings, which did not have complete information. Thanks to that, it was possible to create a group including 109 buildings. These buildings were characterized by high variability of parameters influencing energy consumption, which is presented in Table 1.

In the next step, qualitative variables such as, information about which partition was undergoing thermal improvement were converted into a quantitative variable. Each qualitative variable could be in two response options "Yes" or "No". These words were replaced by numbers: "1" and "0". The value "1" was put when the event occurred (the partition was thermally improved). Otherwise, when the event did not occur the variable took the value "0". Since zero-one variables sum to unity the resulting matrix is singular. In order to eliminate the linear relationship between the variables in further studies, the variable that most often took the value of "1" was omitted. Next, the independent variables that were not statistically significantly correlated with the explanatory quantity were removed from the collected database. In the last step of pre-processing the data was randomly divided into a learning set and a test set. The learning set was formed from 75% of observations and the rest was the test set. In order to ensure comparability of results obtained for particular methods during the research no objects were moved between particular sets. From the data base prepared in this way, independent variables were selected for four purposefully chosen sets describing the process of energy demand change in buildings after thermomodernization. The individual sets of variables were created based on the statistical analyses performed, in which the requirement was that the independent variables be statistically significantly correlated with the dependent variable, while the degree of correlation between the independent variables could not be greater than 0.3. On this basis, four sets of variables that met these requirements were identified. The required effort to collect all the necessary information was also taken into account when creating each group of variables. The buildings subjected to analysis were characterized by different availability of data, some of the buildings in use had full documentation describing their technical condition and equipment enabling monitoring of energy needs. For older buildings, there has often been a problem in obtaining reliable and up-to-date data on their energy demand.

Table 1. Characteristics of selected parameters (and their variability) affecting heat consumption in buildings.

No.	Parameter	Average	Coefficient of Variation (%)
1	Φ_h—calculated power consumption of heating systems, (kW)	189.9	63.3
2	FE_0—index of final energy demand for heating before modernization, (kWh·m^{-2}·year^{-1})	287.2	72.4
3	V_e—calculated from exterior measurements the heated volume of building, (m^3)	6392.2	62.9
4	S/V_e—shape coefficient of buildings (the ratio surface to volume), (m^2·m^{-3}), (m^{-1})	0.46	43.7
5	A_f—calculated surface of heated floors from interior measurements, (m^2)	1567.8	53.4
6	A_w—calculated from exterior measurements total walls' surface (net) area, (m^2)	1096.8	49.7
7	A_r—calculated from exterior measurements surface of roof projection area (net), (m^2)	467.0	55.5
8	A_{tw}—calculated from exterior measurements total windows area, (m^2)	290.7	54.1
9	A_{in}—calculated from interior measurements total (net internal area) (m^2)	1764	54.4
10	NO_{pb}—number of living persons per building (N$_b$)	73.9	67.6
11	NOp—number of residential flats, premises (p$_c$.)	32.4	62.8
12	U_w—calculated thermal transmittance of walls components, (W·m^{-2}·K^{-1})	1.1	31.9
13	U_r—calculated thermal transmittance of roof projections components, (W·m^{-2}·K^{-1})	1.25	85.1
14	U_f—calculated thermal transmittance of floors components (floor over basement), (W·m^{-2}·K^{-1})	1.17	37.8
15	U_{win}—thermal transmittance of windows (commercial data), (W·m^{-2}·K^{-1})	1.88	30.7
16	U_g—calculated thermal transmittance of floor components on the ground, (W·m^{-2}·K^{-1})	1.61	51.7
17	FE_1—index of final energy demand for heating after modernization, (kWh·m^{-2}·year^{-1})	143.6	34.7

In this paper, an attempt was made to estimate the energy efficiency of buildings undergoing thermal improvement on the basis of different sets of diagnostic variables. These sets differ not only in the number of variables, but also in the specificity of the information provided. Some of the variables are of typically energetic character, such as heating power demand or seasonal energy demand for heating, while others describe constructional parameters such as: surface areas of partitions through which heat losses occur, heated volumes, heat transfer coefficients through partitions. Utility parameters were also taken into account, such as the number of people using the building. Detailed characteristics of the individual sets of variables are presented in Table 2. The first set contained the

fewest indices, but they were the most strongly correlated with the dependent variable. They were the calculated heat power of the heating system before modernization and the annual final energy demand indicator (before modernization). This set, just like the other three sets, also contained information on which partition was subjected to thermal upgrading. Practical use of this set will be possible only in the objects in which the calculations of design heat load for heating have been carried out according to EN-ISO 12,831 [35]. From the next set of variables, information concerning the demand for heating power has been eliminated and replaced by information concerning characteristic dimensions of particular parts of the building such as: surface area of particular partitions, surface and volume of the building and also indicators characterizing the given building (compactness coefficient of the building, number of people using the building, number of flats). The use of the above-mentioned sets (I and II) will be possible only in those facilities, where the measurements of final energy consumption for heating are performed and archived. The third set of variables, from which information concerning energy consumption was eliminated, was supplemented with heat transfer coefficients for individual rooms. Gathering such an extensive range of information allows for an accurate characterization of the facility but requires a great deal of effort to prepare it reliably. Due to the fact that the third set of variables was extensive and therefore collecting such an extensive range of data is time-consuming, in the last set of input data only variables with direct influence on heat losses in the object were left, such as heat transfer coefficients through partitions and surface areas of partitions through which heat losses occur. Table 2 lists the individual parameters included in the sets of input variables, with the parameter's membership in a given set denoted by 1.

Table 2. Sets of input variables for analyzed predictive models.

Parameter	Set			
	I	**II**	**III**	**IV**
Φ_h—calculated heating consumed power, (kW)	1			
FE_0—index of final energy demand for heating before modernization, (kWh·m^{-2}·year^{-1})	1	1		
V_e—calculated from exterior measurements the heated volume of building, (m^3)		1	1	
S/V_e—shape coefficient of buildings (the ratio surface to volume), (m^2·m^{-3}), (m^{-1})		1	1	
A_f—calculated surface of heated floors from interior measurements, (m^2)		1	1	1
A_w—calculated from exterior measurements total walls' surface (net) area, (m^2)		1	1	1
A_r—calculated from exterior measurements surface of roof projection area (net), (m^2)		1	1	1
A_{tw}—calculated from exterior measurements total windows area, (m^2)		1	1	1
A_{in}—calculated from interior measurements total (net internal area) (m^2)		1	1	
NO_{pb}—number of living persons per building (N$_b$)		1	1	
NOp—number of residential flats, premises (p$_c$.)		1	1	
U_w—calculated thermal transmittance of walls components, (W·m^{-2}·K^{-1})			1	
U_r—calculated thermal transmittance of roof projections components, (W·m^{-2}·K^{-1})			1	1
U_f—calculated thermal transmittance of floors components (floor over basement), (W·m^{-2}·K^{-1})			1	1
U_{win}—thermal transmittance of windows (commercial data) (W·m^{-2}·K^{-1})			1	1
U_g—calculated thermal transmittance of floor components on the ground, (W·m^{-2}·K^{-1})			1	1

The developed sets of variables were used to build a forecasting model. The study tested the usefulness of Takagi-Sugeno type neuro-fuzzy sets for predicting the energy demand index in residential buildings undergoing thermal upgrading.

The prediction of the variable based on the neuro-fuzzy model followed an iterative procedure. Its structure is schematically illustrated in Figure 2.
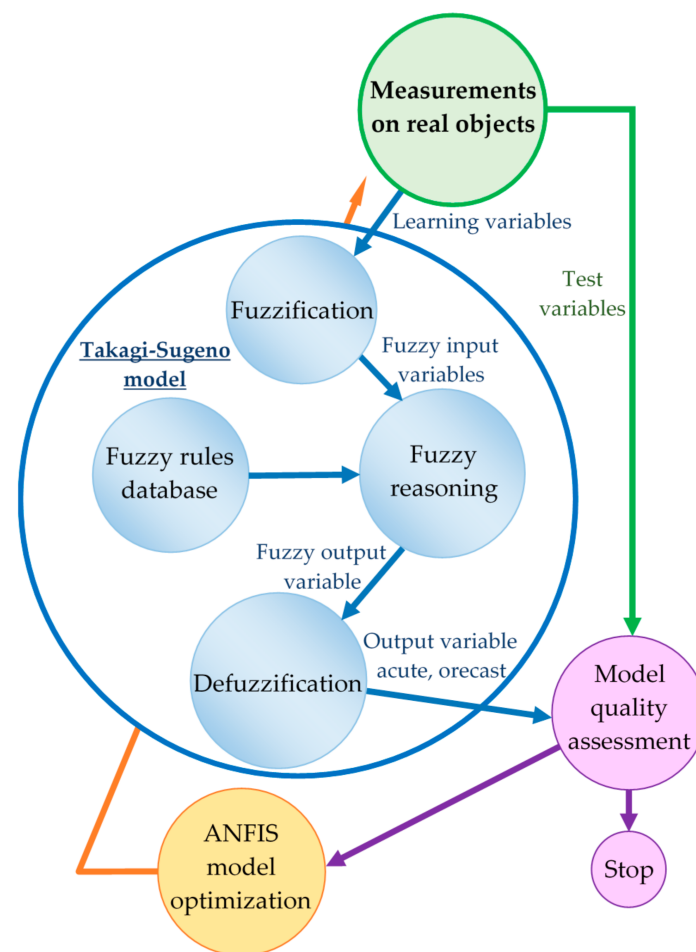
Figure 2. Iterative procedure of variable prediction using neuro-fuzzy model.

The waveforms of the variables obtained during the measurements were divided into learning and test matrix (green area in the diagram). The learning matrix, containing the input data and the output variable, was the basis for formulating a predictive model with a Takagi-Sugeno structure. This model consisted of three basic functional modules (blue area in the diagram): blurring, inference, and sharpening. The former module was responsible for the blurring of input variables. It consisted in dividing variables into subsets described by membership functions. The second module—inference—was responsible for processing input data into appropriate value of output variable in the space of fuzzy variables. Inference process was based on rule base describing relations between input variables and output variable. In the last module of sharpening, the resultant value of output variable—forecast—was determined. The processes outlined above, in the blue area of the diagram, occurred automatically. The procedure, however, required the prior definition of a number of parameters related to the optimization of the model, such as: the type of the blurring function, the division into subsets, the type of the output function (linear or constant) and the number of learning epochs. In the next iteration step, a computer simulation was performed, during which the prediction calculated using the model was compared with the actual variables of the test matrix (purple area in the diagram). The result of the analysis was the information about the quality of the model (its measure is the error values described by relations 1–3, which are discussed below). The acceptance of the final version of the model (acceptable error value) meant the end of the iteration procedure—the transition to the stop stage. In the case when the error value exceeded the accepted assumptions, the prediction model was subjected to further optimization—there was a transition to the ANFIS stage (see diagram). At the beginning of this stage, the set of parameters related to the process of fuzzifying the input variables and

to the output function was redefined (in comparison to the previous iterations). ANFIS (Adaptive Neuro-Fuzzy Inference System) using an adaptive five-layer neural network automatically performed Sugeno-type predictive model optimization using two algorithms: back propagation and least squares estimation. The error back propagation algorithm was used to determine the parameters of the fuzzy function. Through the mean square output error minimization algorithm (with respect to the learning data), the weights of the neural network were gradually updated [36].

After optimization using ANFIS, the quality of the model was checked again and on this basis a decision was made to either correct it further or terminate the iterative procedure. For a complete picture of the methodology, apart from its structure (Figure 2), it is necessary to present also the data processing method during the forecast calculation. When formulating neuro-fuzzy models based on the learning matrix, calculations were initially performed, taking into account successively all types of membership functions available for the fuzzification stage (Figure 3) [37].
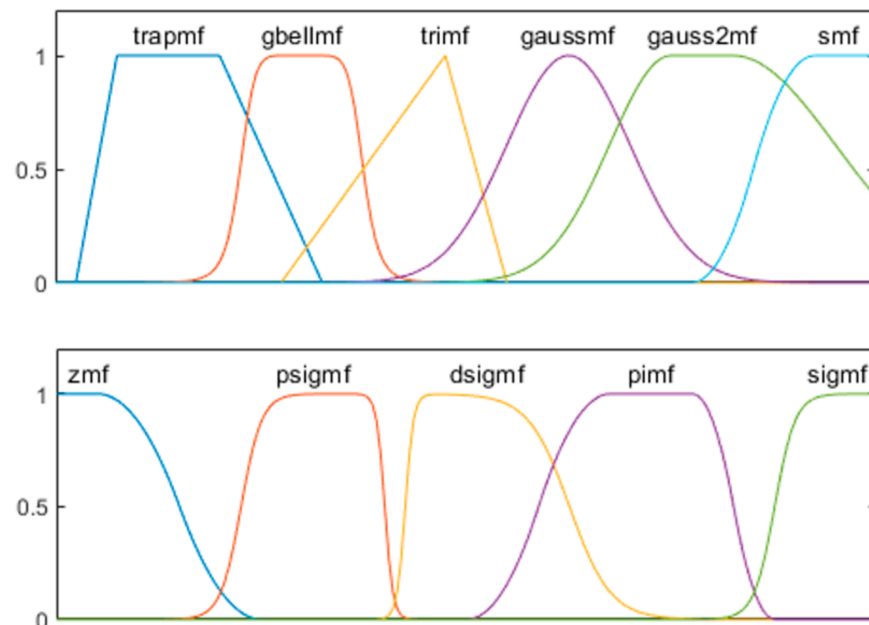

Figure 3. The waveforms of the available membership functions.

Initially, the minimum number of subsets sharing the input variable values and the minimum number of learning epochs (3 subsets and 3 epochs) were used during the fuzzification. Such an example of initial configuration of blurring parameters for one of the input variables is shown in Figure 4.
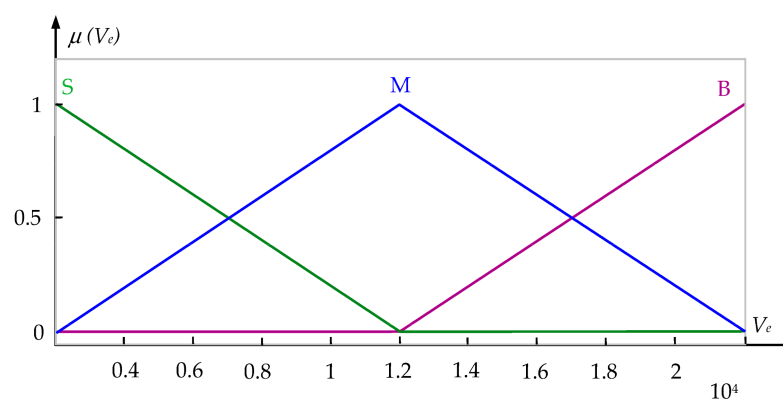

Figure 4. Path of the membership function of the elements of the input value set V_e to the subsets.

As can be observed (Figure 4), the set of values of the input variable V_e (calculated from exterior measurements of the heated volume of building) is divided by the triangular membership function μ (V_e) into 3 subsets: small S, medium M, and big B.

For the prepared models, the most accurate forecasts (with the smallest error) from particular data sets were selected on the basis of error values. For the selected forecasts in subsequent iteration steps, the model optimization parameters were increased by increasing the number of subsets and the number of epochs.

In the course of calculations, the number of learning epochs was maximally increased up to 50, because according to Nebot and Mugica [34], the increase in the number of epochs above this value causes a significant extension (nearly six times) of the computer simulation time and, at the same time, does not affect the reduction of the forecast error value. The quality assessment of the developed models was based on the mean bias error (*MBE*), coefficient of variance of the root mean square error (*CV RMSE*) and coefficient of determination (R^2) which are accepted as statistical calibration standards by ASHRAE Guideline (American Society of Heating, Refrigerating and Air-Conditioning Engineers) 14-2014 [38–40]. Other metrics frequently used in the literature such as *MAE* (mean absolute error) and *MAPE* (mean absolute percentage error) were also used for quality assessment [38].

$$MBE = \frac{\sum_{m=1}^{n_g}\left(y_i - y_i^P\right)}{\sum_{m=1}^{n_g} y_i}\cdot 100\% \quad m = 1,2,3\ldots,n_g \tag{1}$$

$$CV\ RMSE = \frac{\sqrt{\sum_{m=1}^{n_g}\frac{\left(y_i - y_i^P\right)^2}{y_i}}}{\frac{1}{n_g}\sum_{m=1}^{n_g} y_i}\cdot 100\% \quad m = 1,2,3\ldots,n_g \tag{2}$$

$$R^2 = \left(\frac{n_g\cdot\sum_{m=1}^{n_g} y_i\cdot y_i^P - \sum_{m=1}^{n_g} y_i\cdot\sum_{m=1}^{n_g} y_i^P}{\sqrt{\left(n_g\cdot\sum_{m=1}^{n_g} y_i^2 - \left(\sum_{m=1}^{n_g} y_i\right)^2\right)\cdot\left(n_g\cdot\sum_{m=1}^{n_g} y_i^{P\,2} - \left(\sum_{m=1}^{n_g} y_i^P\right)^2\right)}}\right)^2 \tag{3}$$

$$MAE = \frac{1}{n_g}\sum_{m=1}^{n_g}\left|y_i - y_i^P\right| \quad m = 1,2,3\ldots,n_g \tag{4}$$

$$MAPE = \frac{1}{n_g}\sum_{m=1}^{n_g}\left|\frac{y_i - y_i^P}{y_i}\right|\cdot 100\% \quad m = 1,2,3\ldots,n_g \tag{5}$$

where: y_i is the actual value (quantity) in the facility i, and $y^p_{\ i}$ is the forecast value (quantity) in the facility i. The difference between y_i and $y^p_{\ i}$ is divided by the actual value y_i and m is the index of number of test object, n_g is the number of objects ($m = 1, 2, 3, \ldots, n_g$).

Two linear ordering methods were used to assess the suitability of each set of variables to determine changes in the annual heat demand index for thermally improved buildings: Simple Additive Weighting (SAW) and a method based on a synthetic measure of development (SMR). *MBE*, *CV RMSE*, R^2, *MAE*, and *MAPE* were used as diagnostic variables to determine the synthetic variable. Initially, the nature of the variables (stimulants, destimulants) and the weights of the variables were determined, taking only the R^2 index as stimulant and equal weights for all diagnostic characteristics.

The conversion of variables of the nature of destimulants to stimulants was done during the normalization itself, using Weitendorf normalization [41]:

Formula normalization of stimulants

$$x_{ij}^* = \frac{x_{ij} - x_j^-}{x_j^+ - x_j^-} \tag{6}$$

Formula normalization of destimulants

$$x_{ij}^* = \frac{x_j^+ - x_{ij}}{x_j^+ - x_j^-} \qquad (7)$$

x_{ij}^*—normalized value of the *i*-th variant according to the *j*-th criterion,

x_{ij}—value of the *i*-th variant according to the *j*-th criterion,

$x_j^+ = \max_i(x_{ij})$, $x_j^- = \min_i(x_{ij})$—maximum or minimum value of the *i*-th variant according to the *j*-th criterion.

The SAW method was first presented in 1954 in a paper by Churchman and Ackoff [42] and over time has become one of the most well-known algorithms for multicriteria analysis, mainly due to the simplicity of determining synthetic variables (p_i) and the intuitiveness of the method.

$$p_i = \sum_{i=1}^{n} w_i \cdot x_{ij}^* \qquad (8)$$

w_i—weight for particular evaluation criteria, $\sum_{i=1}^{n} w_i = 1$.

This paper also analyzed a linear ordering method based on the synthetic SMR measure, which is very popular in Poland. The author of this method is Hellwig, one of the precursors of Multiple Criteria Decision Making (MCDA) methods, and he described the algorithm of the method in 1968 in his paper [43].

$$p_i = 1 - \frac{d_{io}}{di} \qquad i = 1, 2, \ldots, n \qquad (9)$$

where:

$$d_{io} = \sqrt{\sum_{i=1}^{n} \left(x_{ij}^* - x_{oj}^* \right)^2} \qquad (10)$$

$$d_o = d_{o\acute{s}} + 2 \cdot s_d \qquad (11)$$

$$d_{os} = \frac{1}{n} \sum_{i=1}^{n} d_{io} \quad s_d = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (d_{io} - d_{os})^2} \qquad (12)$$

x_{oj}^*—maximum value from the set of features,

n—number of objects.

4. Results

4.1. Analysis of the Prediction Quality of Models for Particular Sets of Variables

For the first set of variables explaining the annual final energy demand after building modernization, only two indices were used: calculated thermal power of the heating system before modernization and the index of annual final energy demand (before modernization). This set, like the other three sets, also included information on which partition was subjected to thermal modernization.

Using such a limited set of variables allowed us to develop neuro-fuzzy models characterized by acceptable quality. According to the adopted methodology, acceptable models were those for which R^2 was above 0.75, *MBE* within ±5% and *CV RMSE* below 15%. These requirements were simultaneously met for only 7 models, the characteristics of which are presented in Table 3.

Table 3. Model quality characteristics for SET I.

Assessment Indicator:	The Name of the Blurring Function:						
	gbellmf	**gaussmf**	**dsigmf**	**gaussmf**	**gauss2mf**	**dsigmf**	**psigmf**
	Number of Blurring Functions:						
	2	**2**	**3**	**4**	**4**	**4**	**4**
R^2	0.77	0.76	0.78	0.84	0.82	0.83	0.85
MBE	−1.86	−2.83	1.24	−3.56	−2.28	−3.49	−4.21
CV RMSE	9.27	9.12	6.75	6.75	6.80	7.05	7.03

Based on the analyses performed, it is difficult to clearly indicate which method should be used for such a simplified set of input variables. The best results were obtained for dsigmf using 3 fuzzy functions and gaussmf and gauss2mf using 4 fuzzy functions. For further comparisons of model quality between sets, it was decided to choose the model with gaussmf and 4 fuzzy functions. It had a high correlation coefficient of 0.84 and the lowest *RMSE CV* error of 6.75. Only in terms of *MBE* error was it one of the worst models. The choice of this model was supported by the fact that it had the lowest *MAE* and *MAPE* error rates of 24.50 and 15.58%, respectively, of all the alternatives evaluated. These indices are not included in the model evaluation by ASHARE [39,40] but they are also used in other works [38]. In the second set of variables, the indicator concerning the demand for heating power was abandoned and replaced with information on characteristic dimensions of particular building components as well as indicators characterizing the way of using the given building. Obtaining the information concerning the calculated thermal power of the heating system before modernization requires detailed and time-consuming calculations of the design heat load for the building on the basis of EN-ISO 12,831 standard [35]. The tested set of variables also has its limitations, as it can only be used in facilities where measurements of final energy consumption for heating are made and archived. For such a set of explanatory variables, only two models met ASHARE requirements (Table 4), that is, trimf and gbellmf with two fuzzy functions.

Table 4. Model quality characteristics for SET II.

Assessment Indicator:	The Name of the Blurring Function:	
	trimf	gbellmf
	Number of Blurring Functions:	
	2	2
R^2	0.82	0.86
MBE	0.38	−3.87
CV RMSE	8.14	8.22

Assessing the quality of the selected models for set II it was decided to choose trimf model for further comparative analysis. It is true that the correlation coefficient was at a lower level than for the gbellmf model, but the other two evaluation indicators were slightly better. In addition, the *MAE* and *MAPE* errors clearly indicated this model. The difference between the models in both evaluation indexes was 0.6% each.

In the third set, information concerning energy consumption in the building was eliminated and the set of independent variables was supplemented with heat transfer coefficients for particular partitions. Gathering such an extensive range of information allows for accurate characterization of the object, but requires a lot of work to prepare it reliably. Such a set of variables, contrary to our expectations, did not allow to elaborate a much bigger number of models fulfilling requirements (Table 5).

Table 5. Model quality characteristics for SET III.

Assessment Indicator:	The Name of the Blurring Function:		
	gaussmf	trimf	gaussmf
	Number of Blurring Functions:		
	2	3	3
R^2	0.83	0.92	0.84
MBE (%)	−6.12	−4.15	2.42
CV RMSE (%)	8.70	8.21	11.42

During the study, it was observed that increasing the number of fuzzy functions did not significantly improve the quality of the models. For tasks in which a set of observations of one hundred objects described by a dozen or so features is available, simple models with two or three fuzzy functions are preferred. Further increase of the number of functions significantly increased computation time and did not improve quality. Additionally, increasing the number of learning epochs beyond the accepted level of 50 did not affect the quality of the model. This observation is consistent with Nebot and Mugica [34].

4.2. Comparison of Model Quality for Different Sets of Variables

In the next stage of the research an attempt was made to indicate which set of variables allows to make the best prediction of the annual final energy demand indicator after building modernization, using fuzzy logic methods. For this purpose, a comparison was made between the most effective models from different sets of variables. The obtained values of evaluation indices proposed by ASHRAE Guideline [40] for the best models are shown in Figure 5.
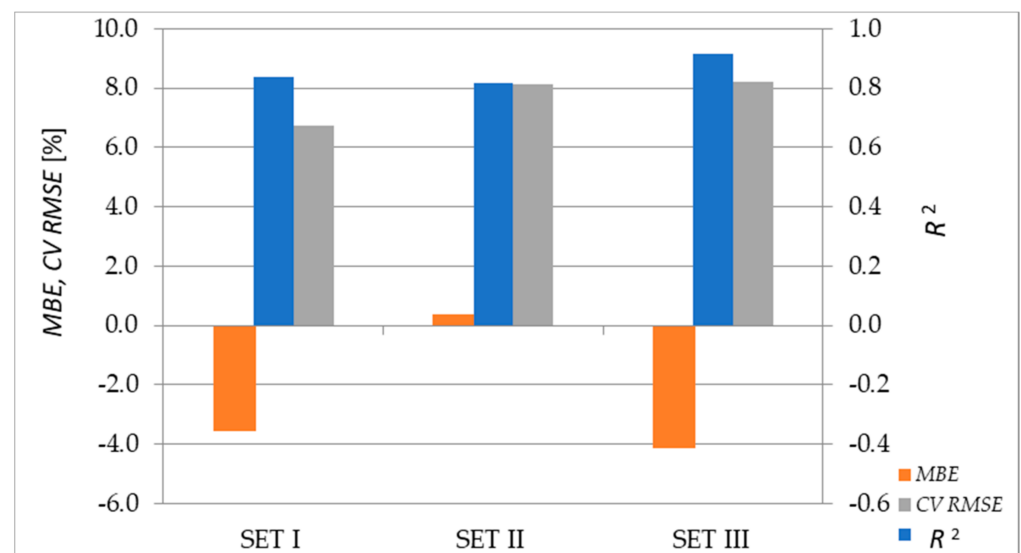

Figure 5. Indicators for assessing the quality of the selected models for each set of variables. SET I—two gaussmf blurring functions; SET II—two trimf blurring functions; SET III—3 trimf blurring functions.

Looking at the data shown in the figure above, it is difficult to indicate which of the selected models is the best. This is because there is no model that is simultaneously the best for at least two indicators of their quality assessment. Set I has the lowest *CV* error of the *RMSE*. The lowest *MBE* error has set II, and the highest correlation coefficient occurs for set III of variables. However, taking into account the effort to determine the actual coefficients of heat transfer through partitions and the fact that their inclusion resulted only in an improvement of the quality assessment indicators in terms of R^2, it was decided that the best results for this type of research, in which the learning sample is of small size and the independent variables are characterized by high variability, is preferred to set II. Models developed on its basis allowed obtaining *MAE* and *MAPE* errors on the level of 24.36 and 14.92%. At the same time, these are the lowest values of the evaluation indices for all the studied samples. This choice was also confirmed by the linear ordering methods used in the study: Simple Additive Weighting (SAW) and the method based on the synthetic measure of development (SMR). The values of aggregated variables for particular sets of variables, which were determined on the basis of analyzed evaluation indicators, are presented in Figure 6.
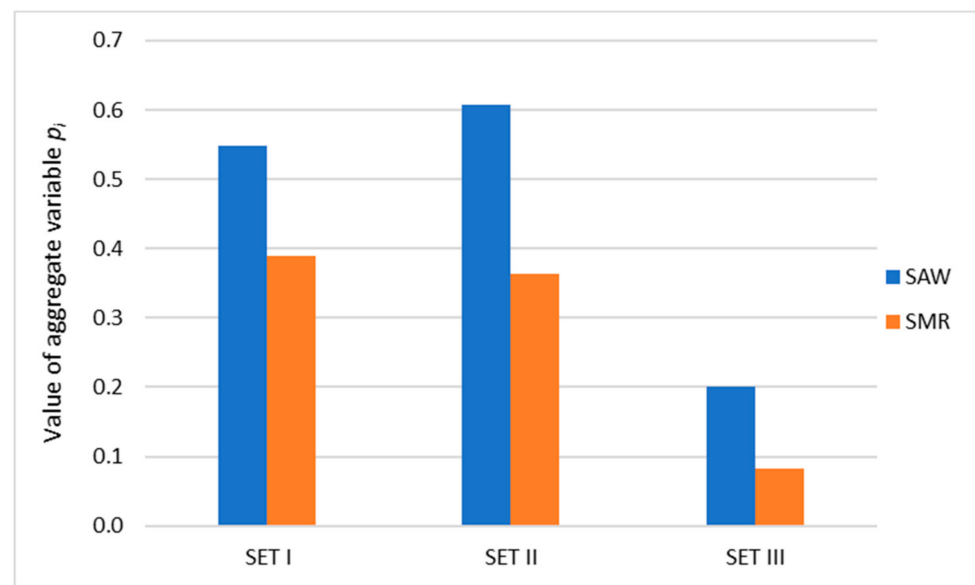
Figure 6. Synthetic indices of model quality assessment for each set of explanatory variables.

The value of the aggregate variable is contained in the range 0–1. The best set of variables describing the studied phenomenon is the one, for which the value of aggregate variable is the highest. In the SAW method the model developed on the basis of the second set of variables can be regarded as the best one. However, its advantage over SET I is small. In addition, the SMR method indicated a slight advantage of the model developed on the basis of SET I over SET II. Both methods unanimously indicated that the third set of variables should not be used.

5. Discussion

The literature review presented in the introduction shows that most of the presented forecasting models focus mainly on estimating energy consumption, energy characteristics in new or simulated buildings, for which precise data can be obtained [28,30–34]. In this area, very little attention is paid to issues related to modeling the impact of thermal improvement measures for existing residential buildings. Unfortunately, it is not possible to compare the presented research results for real buildings made in prefabricated technology (big plate) with simulated or new objects. In the case of a set of simulated residential buildings developed by Tsanas and Xifara [30], the authors using machine learning models such as polynomial regression, support vector machines (SVM), artificial neural networks (ANN), decision trees, fuzzy methods, and hybrid methods obtained very low *MAE* errors of 0.24–2.14 [34].

Due to the fact that the results presented in this paper relate to real buildings, the usefulness of fuzzy logic methods was decided to compare with models relating to real buildings undergoing thermal upgrading [26,29]. The comparison of indicators for assessing the quality of forecasting models is summarized in Table 6.

The performed comparative analysis shows that the use of the neuro-fuzzy method based on the Takagi-Sugeno model for estimating the energy demand index in buildings undergoing thermal improvement is expedient. The use of this method improved the quality of constructed models in most cases. Table 6 shows the ranges of changes in quality assessment indices. Comparing the selected methods, it was observed that the ranges of changes of the evaluation indicators take similar values. However, evaluating the best models for each method, a slight advantage of fuzzy logic was noticed. The selected model (trimf 2 fuzzy functions) was characterized by indicators at: *MAPE* = 12%, *MBE* = 0.4%, *CV RMSE* = 8%, and R^2 = 0.8. Of the remaining group, the MARS and SRT model were noteworthy, whose errors were at: *MAPE* = 17%, *MBE* = 4%, *CV RMSE* = 14%, and R^2 = 0.8.

Table 6. Comparison of the mean absolute error (*MAE*), mean absolute percentage error (*MAPE*), mean bias error (*MBE*), coefficient of variance of the root mean square error (*CV RMSE*), and coefficient of determination (R^2), of selected methods estimating the reduction energy requirement taking into account all accessible input variables for test.

Method Estimating	Index				
	MAE	*MAPE* [%]	*MBE* [%]	*CV RMSE* [%]	R^2
ANN [25]	-	23 ÷ 29	4 ÷ 13	14 ÷ 24	0.6 ÷ 0.8
MARS [25]	-	17 ÷ 35	4 ÷ 14	15 ÷ 37	0.3 ÷ 0.8
SRT [25]	-	16 ÷ 27	5 ÷ 12	14 ÷ 28	0.4 ÷ 0.8
Rough Set Theory [28]	18 ÷ 25	14 ÷ 18	−16 ÷ 2	18 ÷ 32	-
Fuzzy logic	15 ÷ 26	12 ÷ 25	−4 ÷ 12	7 ÷ 29	0.7 ÷ 0.9

6. Conclusions

On the basis of 109 prefabricated residential buildings, for which the authors of the study performed energy audits, specific sets of purposefully selected variables characterizing the buildings in terms of heat demand were extracted. These variables were used to assess the usefulness of neuro-fuzzy models for prediction of heat demand after thermomodernization.

- During the study it was observed that the increase in the number of fuzzy functions did not result in significant improvement of model quality. For tasks in which a set of observations of about a hundred objects described by a dozen or so features is available, simple models with two or three fuzzy functions are preferred. Further increase of the number of functions significantly prolonged computation time and the number of built models fulfilling the assumptions was decreasing.

- Evaluating the usefulness of particular sets of variables according to the indicators proposed by ASHARE, it was observed that it is very difficult to unequivocally indicate the set of variables that allows to develop the best predictive model. Taking into account the quantitative indicators and the effort required to prepare the independent variables, it was concluded that the second set of variables should be preferred. It includes the following data described in Table 2: FE_0, V_e, S/V_e, A_f, A_w, A_r, A_{tw}, A_{in}, NO_{pb}, NO_p. This choice was confirmed by the line order analysis of multivariate objects using SAW and SMR methods. Using fuzzy logic on the test dataset describing real buildings, errors for the test set were obtained at *MAPE* = 12%; *MBE* = 0.4%; *CV RMSE* = 8%; and R^2 = 0.8.

- From the comparison of selected methods for predicting the energy demand of thermally improved real buildings, it is concluded that neuro-fuzzy methods and MARS and SRT models should be recommended.

In the future, the authors of the paper plan to once again use the studied group of objects to test hybrid forecasting methods. The research will cover all possible combinations of methods tested so far. It is also planned to extend the existing database with analogous objects as well as with objects constructed in a different technology or buildings used in different climatic zones. The current database for machine learning methods seems to be insufficient due to a very large number of features describing a limited number of objects.

Author Contributions: Conceptualization, T.S.; software, T.S., K.N. and S.L.; data curation, T.S. and K.N.; investigation, T.S. and K.N.; methodology, T.S. and S.L.; project administration, T.S.; supervision, T.S.; Writing—Original draft, T.S., S.L. and K.N.; Writing—Reviewing and editing, K.N., S.L. and T.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

References

1. Act of 20 May 2016 on Energy Efficiency. Journal of Laws 2016, Item 831. Available online: https://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20160000831/U/D20160831Lj.pdf (accessed on 18 September 2020).
2. Ballarini, I.; Corrado, V. Application of energy rating methods to the existing building stock: Analysis of some residential buildings in Turin. *Energy Build.* **2009**, *41*, 790–800. [CrossRef]
3. Chan, H.-Y.; Riffat Saffa, B.; Zhu, J. Review of passive solar heating and cooling technologies. Renewable and Sustain. *Energy Rev.* **2010**, *14*, 781–789. [CrossRef]
4. Energy Policy of Poland until 2040. Available online: https://www.dziennikustaw.gov.pl/MP/2021/264 (accessed on 21 March 2021).
5. The European Parliament and the Council of the EU Directive 2018/844 of 30 May 2018 Amending Directive 2010/31/EU on the Energy Performance of Buildings and Directive 2012/27/EU on Energy Efficiency. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018L0844&from=EN (accessed on 21 September 2020).
6. Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on Energy Efficiency, Amending Directives 2009/125/EC and 2010/30/EU and Repealing Directives 2004/8/EC and 2006/32/EC. Available online: http://data.europa.eu/eli/dir/2012/27/oj (accessed on 21 March 2021).
7. Kasperkiewicz, K. *Termomodernizacja Budynków. Ocena Efektów Energetycznych*; Wydawnictwo Naukowe PWN: Warszawa, Poland, 2018; ISBN 978-830-119-61-34.
8. Praca Zbiorowa. Stan Energetyczny Budynków w Polsce, Raport, BuildDesk Polska. 2011. Available online: http://www.builddesk.pl/files/BuildDesk/Consultancy/PL%20BD%20Analytics/2010-12-stan-energetyczny-budynkow.pdf (accessed on 24 September 2020).
9. TABULA. Polish Building Typology. In *Scientific Report*; Narodowa Agencja Poszanowania Energii: Warszawa, Poland, 2012; Available online: https://episcope.eu/fileadmin/tabula/public/docs/scientific/PL_TABULA_ScientificReport_NAPE.pdf (accessed on 15 January 2020).
10. Szul, T. *Ocena Efektywności Energetycznej Budynków*; Wydawnictwo Naukowe INTELLECT: Waleńczów, Poland, 2018; ISBN 978-83-950526-3-7.
11. Bourdeau, M.; Zhai, X.-Q.; Nefzaoui, E.; Guo, X.; Chatellier, P. Modelling and forecasting building energy consumption: A review of data-driven techniques. *Sustain. Sustain. Cities Soc.* **2019**, *48*, 101533. [CrossRef]
12. Seyedzadeh, S.; Rahimian, F.; Glesk, I.; Roper, M. Machine learning for estimation of building energy consumption and performance: A review. *Vis. Eng.* **2018**, *6*, 5. [CrossRef]
13. Zhao, H.; Magoulès, F. A review on the prediction of building energy consumption. *Renew. Sustain. Energy Rev.* **2012**, *16*, 3586–3592. [CrossRef]
14. Fumo, N. A review on the basics of building energy estimation. *Renew. Sustain. Energy Rev.* **2014**, *31*, 53–60. [CrossRef]
15. Foucquier, A.; Robert, S.; Suard, F.; Stéphan, L.; Jay, A. State of the art in building modelling and Energy performances prediction: A review. *Renew. Sustain. Energy Rev.* **2013**, *23*, 272–288. [CrossRef]
16. Costanzoa, V.; Fabbrib, K.; Piraccini, S. Stressing the passive behavior of a Passivhaus: An evidence-based scenario analysis for a Mediterranean case study. *Build. Environ.* **2018**, *142*, 265–277. [CrossRef]
17. Mat Daut, M.A.; Hassan, M.Y.; Abdullah, H.; Rahman, H.A.; Abdullah, M.P.; Hussin, F. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. *Renew. Sustain. Energy Rev.* **2017**, *70*, 1108–1118. [CrossRef]
18. Fan, C.; Xiao, F.; Wang, S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Appl. Energy* **2014**, *127*, 1–10. [CrossRef]
19. Ahmad, T.; Chen, H.; Guo, Y.; Wang, J. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. *Energy Build.* **2018**, *165*, 301–320. [CrossRef]
20. Tardioli, G.; Kerrigan, R.; Oates, M.; O'Donnell, J.; Finn, D. Data Driven Approaches for Prediction of Building Energy Consumption at Urban Level. *Energy Procedia* **2015**, *78*, 3378–3383. [CrossRef]
21. Yildiz, B.; Bilbao, J.I.; Sproul, A.B. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renew. Sustain. Energy Rev.* **2017**, *73*, 1104–1122. [CrossRef]

22. Wang, J.; Srinivasan, R.S. A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renew. Sustain. Energy Rev.* **2017**, *75*, 796–808. [CrossRef]

23. Deb, C.; Lee, S.E. Determining key variables influencing energy consumption in office buildings through cluster analysis of pre-and post-retrofit building data. *Energy Build.* **2018**, *159*, 228–245. [CrossRef]

24. Szul, T. Comparison of methods in the definition of home energy characteristics in the context of the European Union Directives. *Barom. Reg.* **2015**, *13*, 73–80. Available online: http://br.wszia.edu.pl/zeszyty/pdfs/br41_08szul.pdf (accessed on 21 September 2020).

25. Chang, C.; Zhu, N.; Yang, K.; Yang, F. Data and analytics for heating energy consumption of residential buildings: The case of a severe cold climate region of China. *Energy Build.* **2018**, *172*, 104–115. [CrossRef]

26. Szul, T.; Nęcka, K.; Mathia, T.G. Neural Methods Comparison for Prediction of Heating Energy Based on Few Hundreds Enhanced Buildings in Four Season's Climate. *Energies* **2020**, *13*, 5453. [CrossRef]

27. Fayaz, M.; Kim, D. A Prediction Methodology of Energy Consumption Based on Deep Extreme Learning Machine and Comparative Analysis in Residential Buildings. *Electronics* **2018**, *7*, 222. [CrossRef]

28. Castelli, M.; Trujillo, L.; Vanneschi, L.; Popovič, A. Prediction of energy performance of residential buildings: A genetic programming approach. *Energy Build.* **2015**, *102*, 67–74. [CrossRef]

29. Szul, T.; Kokoszka, S. Application of Rough Set Theory (RST) to Forecast Energy Consumption in Buildings Undergoing Thermal Modernization. *Energies* **2020**, *13*, 1309. [CrossRef]

30. Tsanas, A.; Xifara, A. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning to OLS. *Energy Build.* **2012**, *49*, 560–567. [CrossRef]

31. Chou, J.S.; Bui, D.K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy Build.* **2014**, *82*, 437–446. [CrossRef]

32. Cheng, M.Y.; Cao, M.T. Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. *Appl. Soft Comput.* **2014**, *22*, 178–188. [CrossRef]

33. Li, K.; Sua, H.; Chua, J. Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: A comparative study. *Energy Build.* **2011**, *43*, 2893–2899. [CrossRef]

34. Nebot, À.; Mugica, F. Energy Performance Forecasting of Residential Buildings Using Fuzzy Approaches. *Appl. Sci.* **2020**, *10*, 720. [CrossRef]

35. CEN. *European Standard: Heating Systems in Buildings*; ISO 12831-1:2017-08; CEN: Brussels, Belgium, 2017.

36. Nauck, D.; Klawonn, F.; Kruse, R. *Neuro-Fuzzy Systems*; John Wiley & Sons: Hoboken, NJ, USA, 1997.

37. Mathworks. Membership Function Gallery. Available online: https://www.mathworks.com/help/fuzzy/membership-function-gallery.html (accessed on 5 November 2020).

38. Ruiz, G.R.; Bandera, C.R. Validation of Calibrated Energy Models: Common Errors. *Energies* **2017**, *10*, 1587. [CrossRef]

39. ASHRAE. *ASHRAE Guideline 14-2002 for Measurement of Energy and Demand Savings*; American Society of Heating, Refrigeration and Air Conditioning Engineers: Atlanta, GA, USA, 2002; Available online: https://scholar.google.com/scholar_lookup?title=ASHRAE+Guideline+14-2002+for+Measurement+of+Energy+and+Demand+Savings&author=ASHRAE&publication_year=2002 (accessed on 14 January 2021).

40. ASHRAE. *American Society of Heating, Ventilating, and Air Conditioning Engineers (ASHRAE), Guideline 14-2014, Measurement of Energy and Demand Savings; Technical Report*; American Society of Heating, Ventilating, and Air Conditioning Engineers: Atlanta, GA, USA, 2014; Available online: https://scholar.google.com/scholar_lookup?title=American+Society+of+Heating,+Ventilating,+and+Air+Conditioning+Engineers+(ASHRAE).+Guideline+14-2014,+Measurement+of+Energy+and+Demand+Savings&author=ASHRAE&publication_year=2014 (accessed on 14 January 2021).

41. Weitendorf, D. Beitrag zur Optimierung der Räumlichen Struktur Eines Gebäude. Ph.D. Thesis, Hochschule für Architektur und Bauwesen, Weimar, Germany, 1976.

42. Churchman, C.W.; Ackoff, R.L. An Approximate Measure of Value. *J. Oper. Res. Soc. Am.* **1954**, *2*, 172–187. [CrossRef]

43. Hellwig, Z. Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr. *Przegląd Stat.* **1968**, *4*, 307–327.