

Article

The Impact of the Selection of Exogenous Variables in the ANFIS Model on the Results of the Daily Load Forecast in the Power Company

Janusz Sowinski 

Faculty of Electrical Engineering, Czestochowa University of Technology, Al. Armii Krajowej 17, 42-200 Czestochowa, Poland; janusz.sowinski@pcz.pl

Abstract: Forecasting of daily loads is crucial for the Distribution System Operators (DSO). Contemporary short-term load forecasting models (STLF) are very well recognized and described in numerous articles. One of such models is the Adaptive Neuro-Fuzzy Inference System (ANFIS), which requires a large set of historical data. A well-recognized issue both for the ANFIS and other daily load forecasting models is the selection of exogenous variables. This article attempts to verify the statement that an appropriate selection of exogenous variables of the ANFIS model affects the accuracy of the forecasts obtained ex post. This proposal seems to be a return to the roots of the Polish econometrics school and the use of the Hellwig method to select exogenous variables of the ANFIS model. In this context, it is also worth asking whether the use of the Hellwig method in conjunction with the ANFIS model makes it possible to investigate the significance of weather variables on the profile of the daily load in an energy company. The functioning of the ANFIS model was tested for some consumers exhibiting high load randomness located within the area under supervision of the examined power company. The load curves featuring seasonal variability and weekly similarity are suitable for forecasting with the ANFIS model. The Hellwig method has been used to select exogenous variables in the ANFIS model. The optimal set of variables has been determined on the basis of integral indicators of information capacity H . Including an additional variable, i.e., air temperature, has also been taken into consideration. Some results of ex post daily load forecast are presented.

Keywords: computational intelligence; ANFIS model; short-term load forecasting; input variables selection



Citation: Sowinski, J. The Impact of the Selection of Exogenous Variables in the ANFIS Model on the Results of the Daily Load Forecast in the Power Company. *Energies* **2021**, *14*, 345. <https://doi.org/10.3390/en14020345>

Received: 26 October 2020

Accepted: 6 January 2021

Published: 9 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Forecasting of load in the power system and subsystems is a process of determining future values of time-series based on current and historical values. The forecast horizon may apply either to a single step or to many steps. In the case of long-term forecasts, the length of the step is usually a year, medium-term forecasts a month or day, short-term forecast an hour, and for very short-term forecasts a second (this is the case of real time forecasting).

Short-term load forecasting (STLF) is used for operation and planning the power system and subsystems. Forecasts are necessary for a deregulated energy market in order to plan and implement management operations. These types of forecasts are also necessary at the level of the distribution system operators (DSO) and their sections, i.e., distribution companies. Load profiles at the distribution companies' level are characterized by significant discrepancies related to unstable behavior of the consumers with significant power demand, which places major challenges before the forecasting models for the regions within the distribution company area.

The paper is an attempt to use the forecasting model that minimizes mean absolute forecast error, but at the same time it addresses the problem of developing an algorithm for

selection of variables. Moreover, an analysis of the impact of exogenous variables set on the model accuracy is carried out. Factors that affect the load can be divided into a group specifying the behavior patterns of consumers during a day and a week (time of day, day of a week, Sunday and holiday), another group specifying the type of consumer (residential, municipal, industrial, etc.) and a group specifying weather conditions (temperature, humidity, cloudiness, wind, etc.) [1,2]. Considering the above, one may suspect that the selection of the forecasting model form is strongly affected by daily, weekly and annual variability of a cyclic nature [3]. This feature should be present in short-term forecasting. In the Polish climatic zone, only ambient temperature shows a significant impact on the load profile in the power system and subsystems [2], even though weather conditions are used in many generally applied models [4,5]. The main issue related to the short-term forecasting is the selection of the input quantities in the model [6–8]. The paper focuses on this issue and the impact of the variables on the forecasting process using a model utilizing cyclic changes of the input variables.

This paper puts forth an original approach to forecasting electricity consumption, combining the Hellwig method of selecting a set of exogenous variables with the ANFIS model. It also demonstrates that introducing climate-related variables does not necessarily improve the accuracy of forecasting. The usefulness of ANFIS in modeling periodical processes with random deviations has already been proven in a number of publications. This paper additionally proves that the accuracy of the ANFIS model largely depends on the character of input data. In selected applications, such as the example presented in the software Matlab R2020b, the time series was generated on the basis of the Mackey-Glass time-delay differential equation. The MAPE of the ex post is about 3.5%. The research presented here aims to demonstrate that an increase in random deviations within random exogenous variables of the ANFIS model causes a significant increase in the MAPE of ex post forecasts.

2. Classification of STLF Methods and Forecasting Experience

There are many short-term forecasting methods used to forecast power system and subsystems loads, including but not limited to:

- Linear regression models (auto-regression model of a row p AR(p), moving average model of a row q MA(q), auto-regression and moving average model (p, q) ARMA(p, q), model GARCH),
- Artificial neural networks ANN model,
- Fuzzy logic models,
- Adaptive Neuro-Fuzzy Inference System (ANFIS),
- Data Mining techniques and Support Vector Machines (SVM),
- Wavelet pre-filtering based on ANN and fuzzy models.

In short-term forecasting, ref. [9] various methods based on linear regression are commonly applied due to their simplicity, ease of implementation and reliable results. The regression techniques estimate relationships between exogenous variables and the endogenous variable. They include various variables such as historical loads, weather variables or calendar variables. The review of forecasting method given in Ref. [10] shows that usually the regression factors are estimated based on the least square method.

If chronological series of loads in the power system is characterized with periodic cycles, it is justified to use the following methods for time-series analysis: auto-regression (AR) models and their variants, exponential smoothing models and kernel density estimation (KDE) algorithms. In the AR row model p , identified as AR(p), the load at the instant of time t is the sum of linear combination of historical loads at time instants p and stochastic error. The AR(p) model parameters can be estimated using the Burg method. Some modifications of the AR models with the aggregated load profiles are presented by e.g., Weron [3] and Taylor and McSharry [11].

In 1970, the Box and Jenkins book (former edition of Box et al. [12]) was published for the first time with the description of the autoregressive integrated moving average (ARIMA)

time series models that are used to forecast the loads, including such components as trend, season changes and some irregularities. The moving average model (MA) forecasts the endogenous variable as a sum of the expected value of this variable, linear combination of previous errors of the row q and random noise term. The paper of Singh et al. [9] is one of most interesting comprehensive papers on ARMA model application in forecasting the loads. In the aforementioned paper, the (S)ARMA model has been developed by additional exogenous variables such as temperature, wind, etc., giving rise to the models (S)ARIMAX or (S)ARMAX.

A model that takes into account the annual as well as weekly and daily variables is the model of exponential equalizing, also known as Holt-Winters model [13]. Good results are also obtained using the extension of this model, known as Holt-Winters-Taylor model [14,15], which includes the impact of temperature on the load.

A popular forecasting technique, also used for forecasting the load, is the kernel density estimation (KDE), e.g., in [16] a forecasting model is provided of the load profile of a single consumer of low voltage. A similar topic is discussed in the paper by Dudek [17].

The regression models include also the k -nearest neighbor (kNN) model. The algorithm searches for similar profiles among historical data, considering the Euclidean measure of distance. The forecast is an arithmetic mean of k neighbors. Substituting the Euclidean measure with error estimation p , and the arithmetic mean with the permutation consolidation of the corrected mean, the adjusted feature aware k -nearest neighbor (AFkNN) is derived.

When the load trend is non-linear, acceptable results can be achieved using the following forecasting techniques: fuzzy logic [18] and fuzzy neural networks [19].

A novel method is the improvement of STLF effectiveness through the use of the long short-term memory (LSTM) and GA algorithm to forecast hourly load of the following day [20].

A method that enhances accuracy of the forecasting model using the artificial neural networks is the selection of similar days using the reinforcement learning algorithm based on the Deep Q-Network technique [21].

An interesting approach is represented by algorithms using machine-type learning, i.e., an automatic system that is able to learn based on available data and gain new knowledge [22,23]. Usually, machine-type learning leads to the selection of the most appropriate forecasting model, as proven in, e.g., [23]. The machine-type learning also uses an adaptive neural fuzzy inference system (ANFIS) to very short-term wind forecasting [24].

Hybrid models are becoming more important and combine various forecasting techniques, e.g., [25,26]. A hybrid algorithm improving the forecasting accuracy is presented in [26]. In order to select the input vector, the non-dominated sorting genetic algorithm II (NSGA II) was used that includes a multi-layer perception neural network (MLPNN). The result of NSGA II, i.e., the best combination of input data is used in the adaptive neuro-fuzzy inference system (ANFIS).

Modern forecasting methods (artificial neural network, evolution algorithms, fuzzy logic) require numerous sets, usually to perform the learning process and then to verify the adopted model. For analyses of time curves of the load in the power subsystems, the paper uses the description of load variability in the power system. The methodology presented in Section 5 is used as a tool for constructing the ANFIS model. The key issue in short-term forecasting is the selection of model input variables. The paper deals with this issue and the influence of variables, and more precisely an additional variable, i.e., ambient temperature, on the accuracy of the forecasting process with a model using cyclic changes of input variables.

3. Description of the Load Variability

The main reason for the phenomenon called load variability is the behavior of the customers, conditioned by the season, tradition and habits as well as the nature of the work that shape the schedule of a working day and a holiday. The system load is affected

by both non-accidental and random factors [27]. The main non-accidental factors are: geographical location of the area under consideration and its climate, change of the sun radiation incidence angle, changes of the time of sunrise and sunset as well as the features of the power system such as reception structure, dynamics of the economy development, statutory length of a work day and the system of working shifts applicable within a given country [1,28,29]. The random factors include: changes of temperature, rainfall, cloudiness, wind rate, failures in the power system, changes in the consumers structure, etc.

Consumer load time series can be represented as a random function [1,2]. The power system load is a result of a superposition of individual load profiles of the consumers and represents a sum of random functions. This is to underline that they are subject to increase within a year that is also of random character.

Behavior patterns of citizens, relationship of such patterns with changing seasons, traditions and habits, as well as the nature of the work performed by persons that control the schedule of a working day and holiday affect the loads in the power system and are the main cause of the load variability phenomenon.

The daily load is affected by work schedule and household activities (Figure 1). The nightly loads drop is noticeable from 11 p.m. to 6 a.m. Most of the working population starts their professional activities between 6 a.m. and 10 a.m. and within this period there is a significant increase in power consumption. Reduction of the load takes place at the turn of the working shifts, i.e., 2 p.m. This is caused by shutting down some of the electricity equipment at industrial plants. In the Polish power system, there is a clear evening peak after people return home from work. So far, it has been attributed to an increase in the use of lighting. Along with technological changes, there are evolutionary changes in the daily load profile, e.g., related to more popular use of air-conditioning.

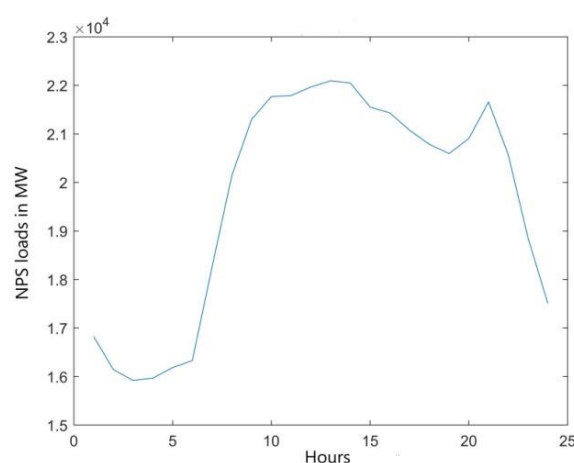


Figure 1. Daily load variability for National Power System(NPS) in Poland on 21 August 2019 (Source: own work based on PSE-Operator S.A. <http://www.pse-operator.pl>).

A different pattern of load reduction on Saturday is presented in Figure 2, which is related to the fact that the working week in Poland is five days long. On Sunday, only the industry of continuous production mode and some of the institutions and public utility companies, hospitals, and transport are working, thus the power demand is much lower when compared to the remaining week days.

The forecasting process capitalizes on the regularities of the daily and weekly load reduction. Stable changes in the remaining weeks of the year cause modern load forecasting methods in the Polish power system to achieve average percentage absolute errors below 2%. However, completely different and burdened with major problems is the forecasting at the level of distribution companies covering small regions. This load is characterized by much more instability of consumers' behavior and the load profiles may significantly differ from the shapes of the national load curves. The paper uses the data of the power company load, which is difficult to forecast load due to significant random deviations.

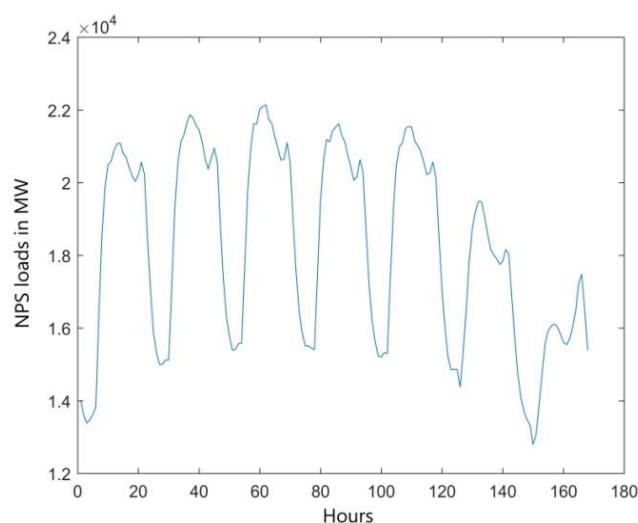


Figure 2. Weekly variability of loads at NPS in Poland from Monday to Sunday 5–11 August 2019 (Source: own work based on PSE-Operator S. A. <http://www.pse-operator.pl>).

4. Statistical Data

Statistical data used in the paper concerns hourly loads from a roughly two-year-long period at the selected power company. The power company loads are characterized by noticeable random fluctuation (Figure 3).

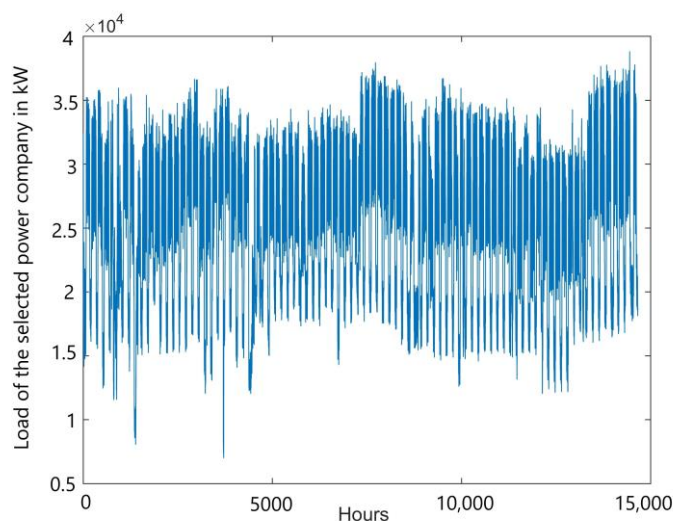


Figure 3. Load of the selected power company-chronological time series, data from 22 months, i.e., ca. 2 years (Source: own work).

The studies have been performed both for the full chronological series and one limited to a selected weekday. When analyzing weekly variability (Figure 2), lower demand for electrical power on Sundays and holidays can be clearly seen. Working days of a week show a significant similarity. Forecasting models were constructed for each day of the week. Wednesday is a typical working day in Poland, therefore the analysis presented in Section 7 refers to that day, but Tables 1 and 2 present the results also for the other days of the week. For each day of the week, a set was created with a number seven times smaller than the input chronological series. The new set is made of only 24-h series of subsequent days of the week and presented in Figure 4. The analysis of the load curves in Figure 4 indicates the special nature of the load profiles of the selected power company, because one can observe random disturbances of significant values in the time series.

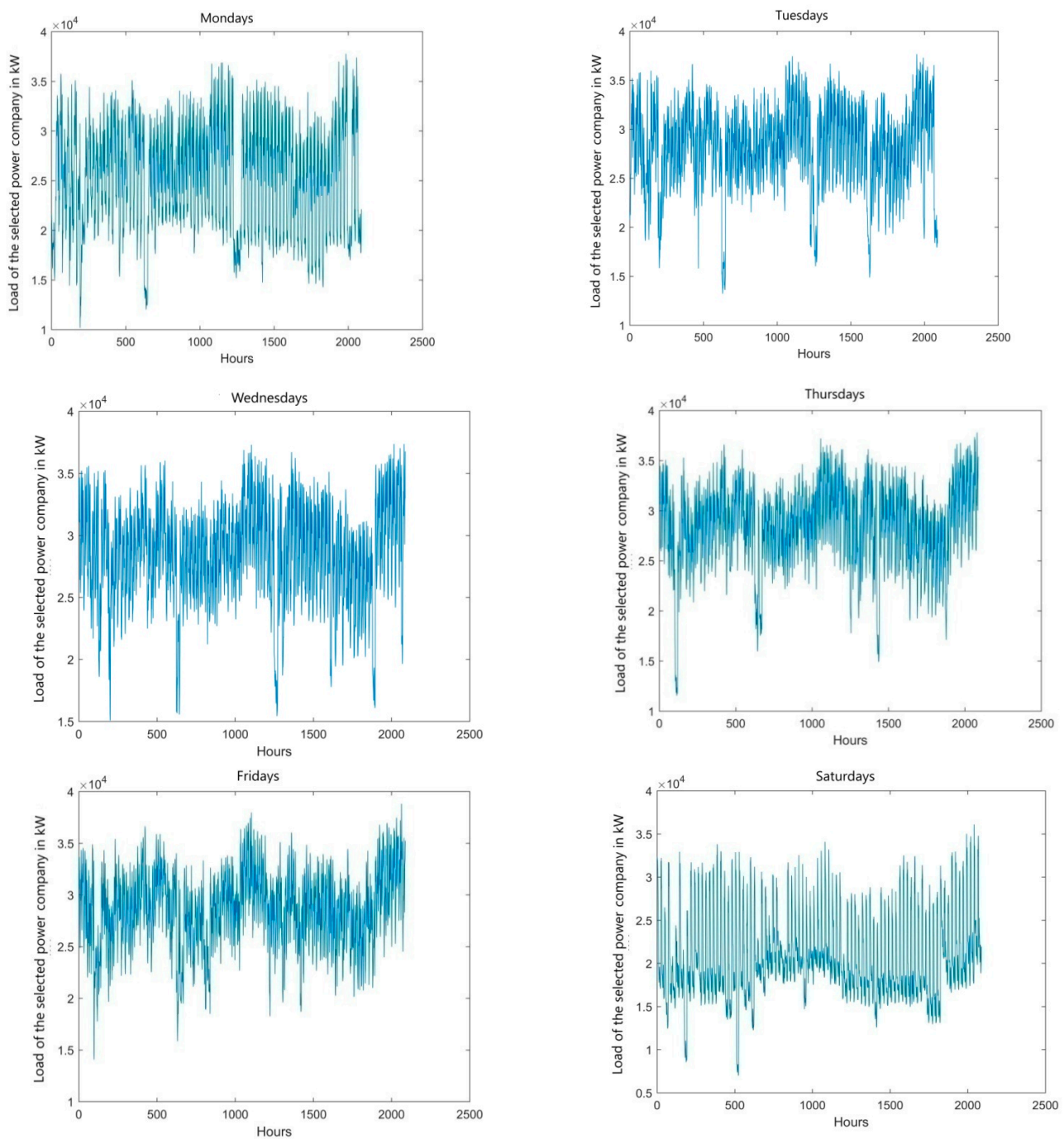


Figure 4. Load of the selected power company within subsequent days of the week—data from ca. 2 years (Source: own work).

Table 1. Integral capacities of information carriers for the selected sets of exogenous variables.

No. of Set	Endogenous Variable and Exogenous Variables	Integral Capacity of Information Carriers H_m						
		Mon	Tues	Wed	Thurs	Fri	Sat	Sun
1	$X(t + 24) X(t) X(t - 24) X(t - 48) \theta(t)$	0.535	0.249	0.371	0.358	0.396	0.716	0.349
2	$X(t + 24) X(t) X(t - 24) \theta(t)$	0.466	0.216	0.364	0.347	0.352	0.666	0.334
3	$X(t + 24) X(t) \theta(t)$	0.413	0.198	0.299	0.334	0.341	0.587	0.307
4	$X(t + 24) X(t) X(t - 24) X(t - 48) X(t - 72)$	0.543	0.255	0.372	0.363	0.399	0.746	0.336
5	$X(t + 24) X(t) X(t - 24) X(t - 48)$	0.545	0.261	0.379	0.355	0.404	0.743	0.331
6	$X(t + 24) X(t) X(t - 24)$	0.489	0.234	0.378	0.349	0.367	0.710	0.326

Table 2. MAPEs for ex post forecasts of loads.

No. of Set	Endogenous Variable and Exogenous Variables	Mean Absolute Percentage Error MAPE [%]						
		Mon	Tues	Wed	Thurs	Fri	Sat	Sun
1	$X(t + 24) X(t) X(t - 24) X(t - 48) \theta(t)$	IM	IM	IM	11.32	IM	IM	IM
2	$X(t + 24) X(t) X(t - 24) \theta(t)$	11.08	10.60	8.95	9.46	8.48	10.14	8.23
3	$X(t + 24) X(t) \theta(t)$	11.29	10.91	9.12	9.68	8.81	10.11	8.70
4	$X(t + 24) X(t) X(t - 24) X(t - 48) X(t - 72)$	IM	IM	IM	IM	IM	9.93	9.01
5	$X(t + 24) X(t) X(t - 24) X(t - 48)$	10.57	10.46	8.42	8.78	IM	9.39	7.86
6	$X(t + 24) X(t) X(t - 24)$	10.70	10.42	8.30	8.95	8.45	9.74	8.40
	<i>Naïve forecast</i>	11.41	11.59	9.49	9.55	9.11	10.59	9.07

IM-Inadequate Model (means that the construction of a model that gives an ex post forecast better than the naive model has failed).

The load curve is slightly affected by weather conditions, the most important parameter of which is ambient temperature. For the load curves presented in Figure 4, curves of mean daily temperatures corresponding to subsequent days of the week were developed (Figure 5).

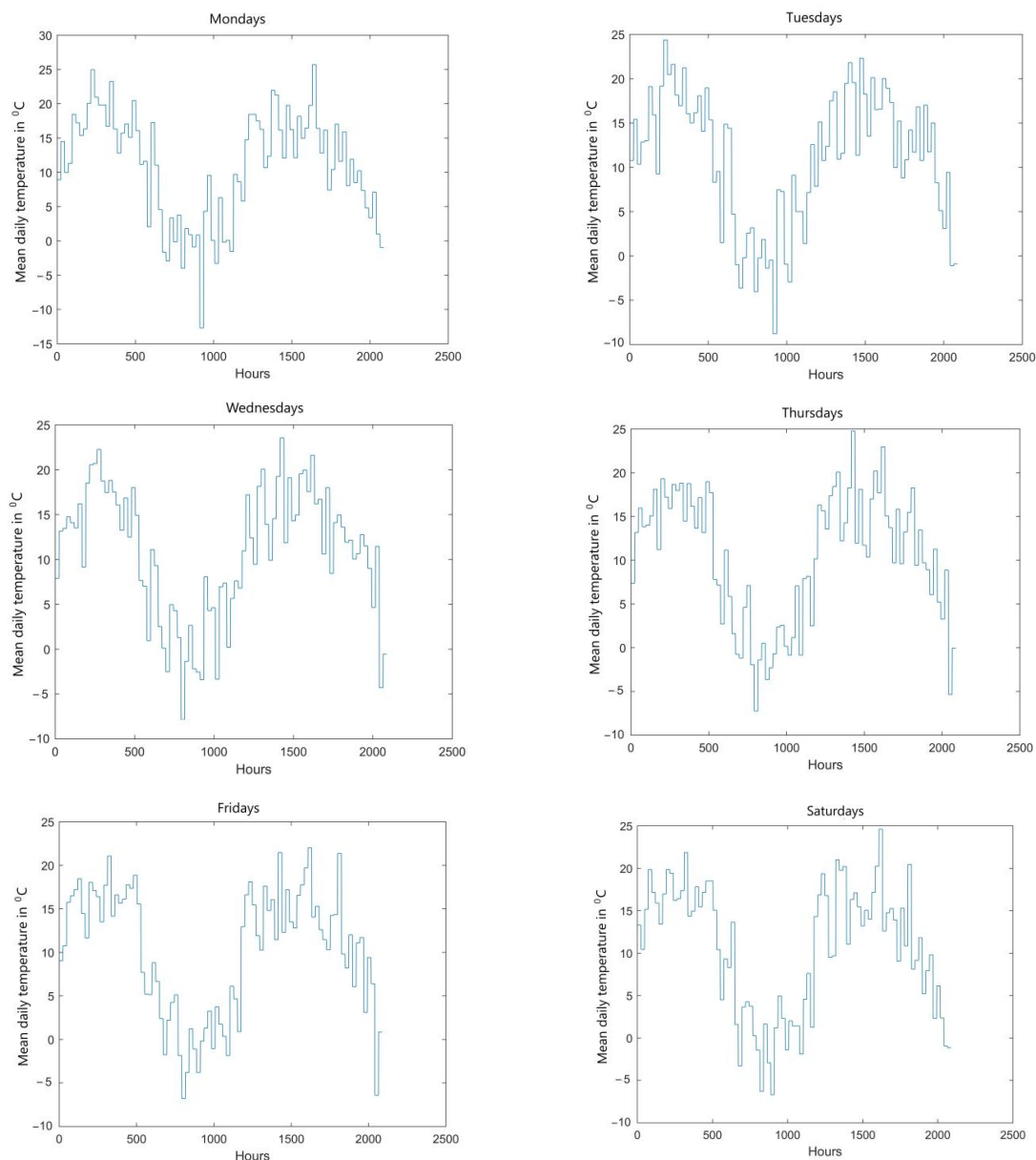


Figure 5. Cont.

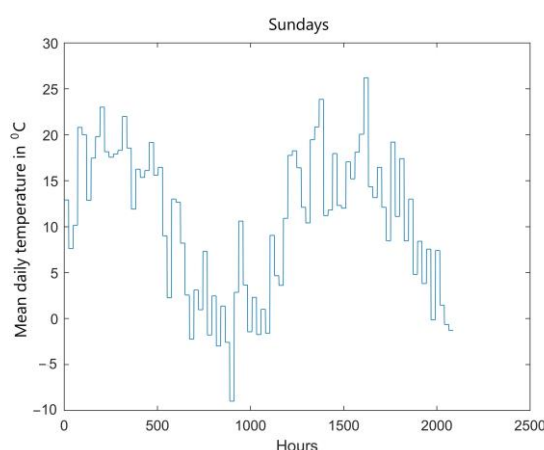


Figure 5. Curves of mean daily temperature (°C) in the area of the power company within subsequent days of the week—data from ca. 2 years (Source: own work).

It can be remarked that there exists a contradictory trend between the load of the power company and the ambient temperature. Within the periods of temperature drop, higher electrical power demand is noticed, however the studies in Poland [1,2] show that temperature is not the only factor affecting the demand increase. The power load changes are also influenced by, for example, religious, sports and cultural events broadcast in real time by the mass media. The aim of the research is to verify the thesis about the significance of the impact of temperature on the load profiles.

5. Method of the Exogenous Variables Selection

5.1. Review of Methods

Many sources [7,8,30–32] provide descriptions of methods for selecting exogenous variables for econometric models. Selection of suitable subsets of variables can be made in order to show the cause and effect relationships, to improve forecast accuracy, to eliminate unimportant variables (noise reduction) or to reduce the modeling time. Current methods of variable selection are based on both multidimensional regression and simultaneous methods. The most common methods include: All Subset Models (ASM), Sequential Search (SS), StepWise regression methods (SW) based on the strategies such as Forward Selection (FS) and Backward Elimination (BE), Genetic Algorithms (GAs), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Least Absolute Shrinkage and Selection Operator (LASSO), Elastic Net as well as Variable Importance on Partial Least Squares (PLS) projections (VIP).

The method of All Subset Models (ASM) [8] consists in analyzing the models of all combinations from among n variables, i.e., $2^n - 1$ models. The method leads to the selection of the best subset of variables however, it is unsuitable for a large number of variables due to calculation time.

Sequential Search (SS) is a simple method of selection of the optimum subset of variables into a form of a model of specified number of variables [31]. Each next variable is selected by substituting it with the remaining variables and selecting the best one according to the adopted criterion.

StepWise regression method (SW) is one of the most commonly applied ones [31,32]. Using the Forward Selection (FS) strategy, the method starts with a model of dimension 0 and the next variables that meet the defined criterion are added. Usually, the variable selected at the next step is to minimize the Residual Sum of Squares (RSS), which enables estimating the F-Snedecor test that compares the models with p and $p + j$ variables.

In the case of the Backward Elimination (BE), the algorithm starts with the p -dimensional model (where p is the maximum number of variables). During the next steps, unnecessary variables are eliminated to minimize the increase of RSS, tested using the F-Snedecor test.

In modern algorithms, the F-test is replaced by minimization of other functions, e.g., the Akaike Information Criterion (AIC) [33].

Within recent years, methods using artificial intelligence become more and more common, e.g., Genetic Algorithms (GAs) [34,35], Particle Swarm Optimization (PSO) [36] or Ant Colony Optimization (ACO) [37]. The aforementioned algorithms were first used as optimization methods, however they were adapted to perform the selection of exogenous variables in the models.

The LASSO method belongs to the regression methods and was developed by R. Tibshirani in 1996 [38]. This method minimizes RSS with an additional condition that checks whether the sum of absolute values of factors is less than the assumed value. This provides the grounds for the selection of a set of variables because some of the regression factors assume values equal to 0.

The extension of the LASSO method when combined with the Ridge regression method results in the Elastic Net method as proposed by Zou and Hastie in 2005 [39]. The additional condition consists of a sum of absolute values of factors and the sum of the components squares. Such design of the criterion provides good selection results of variables strongly correlated with each other.

The Variables Importance Projections (VIP) method is based on the Canonical Powered Partial Least Squares (CPPLS) [40]. For each variable, the VIP factor is defined and the selection consists in removing variables with VIP value less than the defined criterion value (usually equal to 1, because the average VIP mean is 1).

5.2. Hellwig Method Algorithm

Many sources [7,8] provide descriptions of the selection of variables for econometric models. Methods for linear models or those that can be reduced to linear, single-equation methods are best developed. The basic postulate concerning the variables-strong correlation between each explaining and explained variable and at the same time a weak correlation between the explaining variables cause most of variable selection methods to avail of some versions of the linear correlation. A solution found to be an optimal one depends on the lowest number of mutually uncorrelated explaining variables. The explaining variables are as maximally correlated as possible with the explained variable. This general definition translates in the literature to a set of a dozen, and after a few modifications, a few dozen methods of variables selection [6].

The Hellwig method of integral information capacity is one of the most effective and at the same time simple methods according to [6,7]. This method won the greatest recognition among econometricians performing empirical studies.

The starting point here is estimation of the matrix R and vector R_0 . The matrix R corresponds to a matrix of correlation factors between individual k explaining variables that may be expressed as follows:

$$R = \begin{bmatrix} 1 & r_{12} & \cdot & \cdot & \cdot & r_{1k} \\ r_{21} & 1 & \cdot & \cdot & \cdot & r_{2k} \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ r_{k1} & r_{k2} & \cdot & \cdot & \cdot & 1 \end{bmatrix}. \quad (1)$$

However, vector R_0 is a vector of correlation factors between the explained variable and the following explaining variables. This vector can be expressed based on the matrix using the following equation:

$$R_0 = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix} \quad (2)$$

Then, the following calculations are performed of the so-called individual capacities of information carriers X_j of endogenous variable Y composed of different combinations of elements of a given explaining set of variables. It is known that the general number of these combinations is 2^{k-1} . The variable X_j is the better carrier of information of the Y variable, the closer to one is the linear correlation factor r_j of the module. Note that the variable X_j is the clear information carrier only when it is not correlated with other explaining variables. However, when the X_j variable is correlated with them, we assume that X_j is a contaminated information carrier of endogenous variable Y .

Contamination of an individual information carrier X_j is the quantity:

$$g_j = \frac{1}{k-1} \sum_{i \neq j} |r_{ij}| \quad (3)$$

It may be easily checked that the inequality $0 \leq g_j \leq 1$ is always fulfilled, wherein $g_j = 0$, when X_j is a clear information carrier of variable Y and $g_j = 1$, when information carrier contamination is complete.

Individual capacity of the information carrier X_j of variable Y is calculated using the following equation:

$$h_j = \frac{r_j^2}{1 + \sum_{i \neq j} |r_{ij}|} = \frac{r_j^2}{1 + (k-1)g_j} \quad (i, j = 1, \dots, k; i \neq j). \quad (4)$$

In turn, the following equation is used to calculate the integral capacity of information carriers:

$$H_m = \sum_{j=1}^k h_j = \sum_{j=1}^k \frac{r_j^2}{1 + \sum_{i \neq j} |r_{ij}|} \quad (m = 1, 2, \dots, 2^k - 1) \quad (5)$$

It is possible to demonstrate that the H_m parameter is a standardized quantity in the range of $0 < H_m < 1$. If this capacity is close to one, it means that the variables constituting a given combination provide almost a complete resource of information of endogenous variable Y . It follows that introduction of any different combination of explaining variables into the model may only worsen our knowledge about the variable Y .

The above-presented procedure enables one to select an optimum combination of explaining variables. The selection criterion of such combination can be expressed as follows:

$$H_{m0} = \max_m H_m, \quad (6)$$

where H_{m0} means the integral capacity of information carriers for the optimum combination of variables.

5.3. Selection of Variables

The forecasting model presented in the paper may use variables shifted in time and weather variables. Due to the daily variability of loads, the analysis uses the endogenous variable $X(t + 24)$ and the exogenous variables from a single week day shifted by one week $X(t)$, by two weeks $X(t - 24)$, by three weeks $X(t - 48)$ and by four weeks $X(t - 72)$, and daily mean air temperature $\theta(t)$ within the week preceding the occurring load. The Table 1

presents a combination of variables and corresponding calculated integral capacities of information carriers H_m .

For example, in the case of the analysis concerning Wednesday, an optimum set of variables providing the maximum integral capacity of information carriers $H_{mo} = 0.379$ is the set number 5 and the second best one that does not differ much with respect to the capacity $H_m = 0.378$ is the set number 6.

When analyzing the integral capacity of information carriers H_m in Table 1, one can risk a statement that temperature in Poland has little influence when explaining the changes in the load profiles in the selected power company for weekdays. Thus, the information value of the variable in the form of daily mean air temperature is of little importance. The situation is quite different in the case of Sundays, because the influence of temperature on the load seems to be significant. Most industrial plants are closed on Sundays, while the residential sector determines the load profile. Further studies are necessary to show whether the inclusion of the daily mean air temperature as an exogenous variable in the ANFIS model improves the mean absolute percentage error of ex post forecasts.

6. The General Description of the Adaptive Neuro-Fuzzy Inference System

The load profiles of the power system and its subsystems are influenced both by deterministic and random factors. Periodic functions can approximate the time series of many quantities explaining the variation of power load within a year [1,2]. Yearly increases or decreases of load are taken into account in the dynamic functions. The utility loads are presented in Figures 3 and 4. The analysis of time series confirms the assumption that the load curves are characterized by periodic changes with dynamic increments.

An adaptive system of inference applying neural networks seems to be an appropriate tool in the prediction process of time series with periodic variation. In the literature [41–47], there are many examples of the use of adaptive inference systems to forecast power and energy demand in the power system.

Lofti A. Zadeh [48,49] was the creator of both fuzzy sets (1965) and fuzzy logic (1973) theories. The theory of fuzzy sets is an extension of the classical set theory. The fuzzy set is a mathematical object for which the membership function has been defined. The membership function takes values from the interval $[0, 1]$ defining the degree of membership of the element in the fuzzy set. Functions of the class S , π , γ , t or L are used as typical membership functions. Due to fuzzy sets and the membership function used in them, it is possible to represent imprecise statements [41]. The classical Mamdani-Zadeh fuzzy system is a linguistic model with a set of rules. The fuzzy system usually consists of several blocks, the most important of which are fuzzification, inferencing, and defuzzification.

Qualitative dependences between variables in the fuzzy model are determined by the “if-then” rules. Development of the fuzzy model requires definition of input variables that take real sharp (or “crisp”) values. Subsequently, the information is processed by the fuzzification block, whose role is to estimate the membership degree of the sharp values in fuzzy sets using membership functions. Various forms of the membership function are used in the fuzzy set theory. The trapezoidal, triangular and so-called S-functions have been assumed in the research. Neural networks (e.g., backpropagation algorithm with the least squares method) have been applied for the precise estimation of membership function coefficients. The output function is estimated from the aggregation of conclusions derived from all the inferences. In order to implement the process, it is necessary to know the membership degrees of the input variables. The output fuzzy set defines the output membership function. A sharp value is obtained by the defuzzification of this set, determination of the sharp output value requires finding the center of weight, center of maximum or center of sums.

The adaptive neuro-fuzzy inference system ANFIS was presented for the first time in 1993 [41]. The ANFIS model applies a well-known technique of data learning that avails itself of both neural networks and fuzzy logic [42]. Determination of the parameters in the fuzzy inference system is the task of neural networks. The software user does not have to

carry out effort-consuming analyses, yet the user is able to implement a fairly complicated model with ease. The procedures of the ANFIS system make it possible to implement an accurate and fast process of learning, the availability of a large number of membership functions allows for multi-variant calculations. Yet another advantage of the system is an adequate description of the properties of the model by means of fuzzy rules.

In the construction of the fuzzy inference system, some functions of the Fuzzy Logic Toolbox, Matlab®, MathWorks, Inc. (Natick, MA, USA) were used.

The ANFIS system utilizes fuzzy logic that has proved its usefulness in engineering applications. It is a state-of-art tool for development of intelligent systems with the ability to generalize knowledge. Modeling in ANFIS is similar to other identification techniques and requires splitting the set of input data into two subsets. The process of learning employs one subset of data, primarily in order to estimate membership functions. The other subset, which is not used in the learning process, is applied to validate the model.

The verified ANFIS model can be used as a forecasting tool. In the load forecasting of power subsystem, the Sugeno-type fuzzy inference was chosen in the ANFIS model. Mamdani and Sugeno methods are similar in many respects, e.g., the fuzzification blocks are the same, the main difference between the two methods relates to the output membership functions. The Sugeno-type inference system allows only singleton output membership functions that take the form of a constant or a linear function of the input values. The final sharp output value is calculated as the weighted average of all rule outputs in the defuzzification process.

7. ANFIS Model of Load Forecast

The values of daily loads $X(t)$ for $t = 1, \dots, T$ and daily mean air temperature $\theta(t)$ for $t = 1, \dots, T$ for the chosen weekday (Wednesday) in the period of almost two years are known. A few-dimensional input vector of learning data was selected $w(t) = [\theta(t) \dots X(t - 48) X(t - 24) X(t)]$. Taking into account the form of the ANFIS model, the learning set with the output data corresponding to the prediction trajectory $s(t) = X(t + 24)$ was defined. The weekly shift in the input data justifies the cycle of variation for each of the variables.

The database includes 611 days, 24 hourly load data for each day, which gives a total of 14,664 power observations. In addition, average air temperatures for 611 days are available. The set was limited to data regarding only Wednesdays, i.e., 87 days with 24 h load data, which gives a total of 2088 observations.

If we consider a model of two input variables (as an example), for each t in the interval $3 \div 87$, we are able to construct the data set containing 84 records. The first 42 records were used in the learning process and the other 42 in the verification process. In the ANFIS model, the number of fuzzy rules depends on the number of input variables. For the model with two input variables, it is necessary to create four fuzzy rules. In turn, this number implies the number of estimated parameters equal to 24. The number of records in the learning set is only twice as big as the number of parameters. Figure 6 presents the architecture of the ANFIS model, described in Section 6, to the prediction of the output variable $X(t + 24)$ taking into consideration two input variables $X(t - 24)$ and $X(t)$. The Figures 7–10 present the results of estimation and validation of the above forecasting model based on ANFIS. The time series of the output variable $X(t + 24)$, loads on Wednesdays, are characterized by the sinusoidal changeability with the trend (Figure 4). The curve of average temperature on Wednesdays is characterized by similar variability (Figure 5). Figure 7 presents estimated membership functions $w(t)$. Consequently, the RMSE for subsequent epochs in the learning process records a certain decrease (Figure 8). The comparison of real data of loads on Wednesdays with the predictions of the ANFIS model (Figure 9) indicates that there are some differences, especially when analyzing values that deviate from the mean profiles, both up and down. Absolute errors of the ex post loads prediction are shown in Figure 10, and MAPE obtained from them is 8.30% (Table 2).

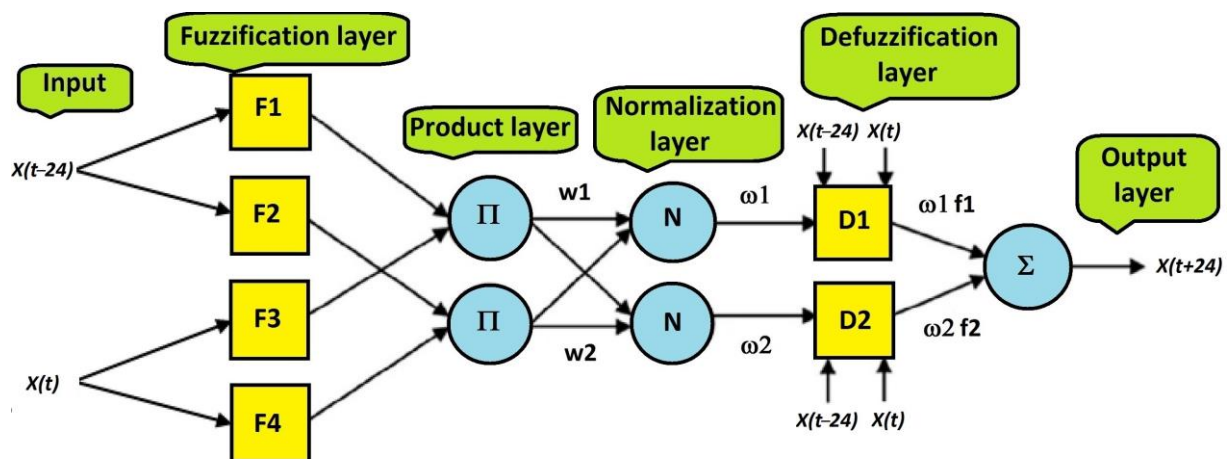


Figure 6. The architecture of the ANFIS model to the prediction of the output variable $X(t + 24)$ taking into consideration two input variables $X(t - 24)$ and $X(t)$.

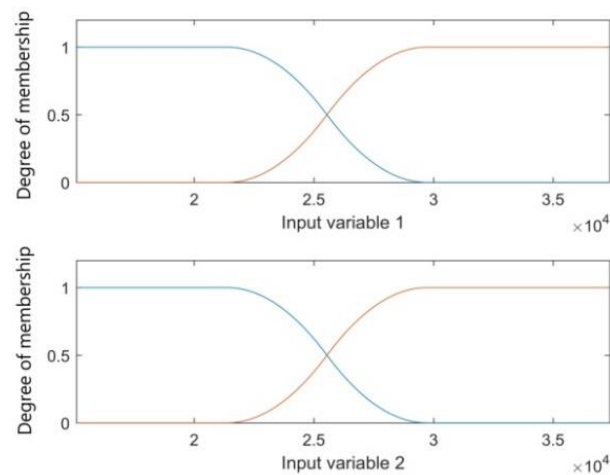


Figure 7. Estimated membership functions for the prediction model.

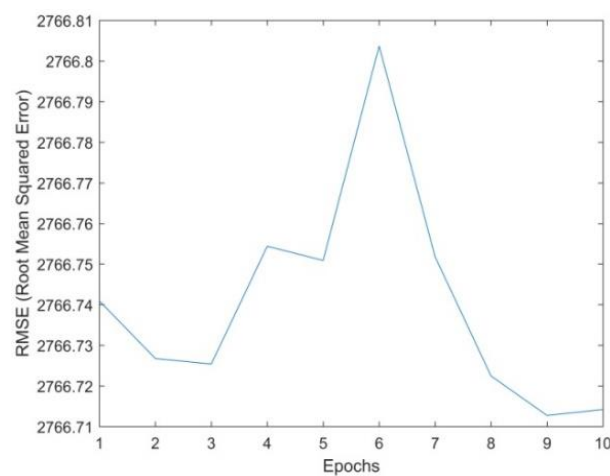


Figure 8. Root Mean Square Error (RMSE) as a function of epochs for the prediction model.

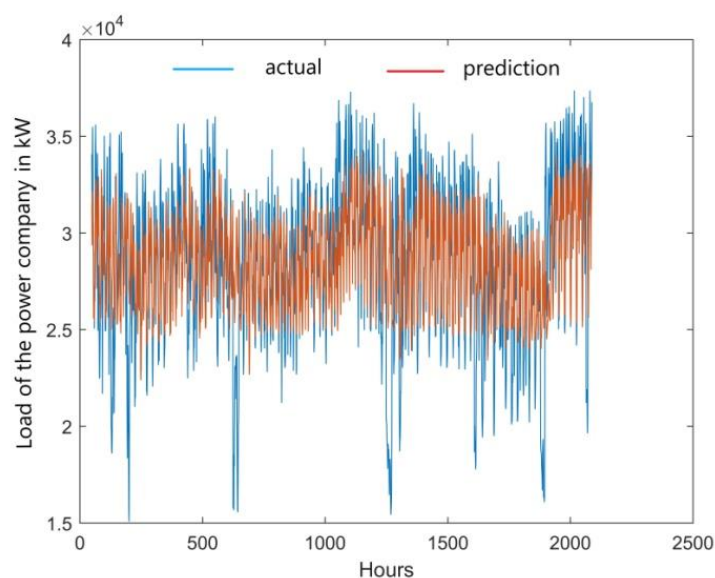


Figure 9. Comparison of Wednesday loads with the prediction.

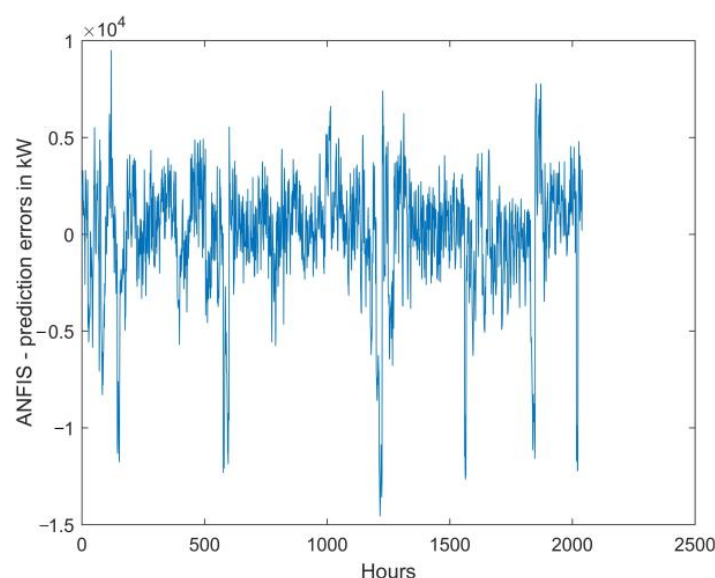


Figure 10. Absolute errors in the prediction model.

The ANFIS models of load forecast were constructed for each day of the week. Table 2 presents the values of MAPE errors for ex post forecasts of loads in the case of the discussed models.

Some kind of estimation of the ANFIS forecasting models quality is their comparison with the naive model results. For example, in the case of the daily load forecast concerning Wednesday, the naive forecast consists in adopting the relationship of loads from the week preceding the forecast as the forecast proper. Error in case of the naive forecast from the whole set of data, i.e., 2064 observations is MAPE = 9.49%.

Models with the set of variables 5 and 6 for Wednesday characterized by the lowest MAPE errors and at the same time the highest values of integral capacities of information carriers H_m point to the proper selection of exogenous variables.

The results presented in Table 2 confirm the conclusion that adding temperature to the exogenous variables does not improve the accuracy of the forecasting models for the weekdays but improves the accuracy for Sundays when the load is determined by the residential sector. The conclusion concerns the specific Polish power company.

8. Comparison of the Results Presented in Other Publications

Artificial intelligence (genetic algorithms, neural networks and neuro-fuzzy systems) is currently used for forecasting electricity demand and forecasting electrical loads. The models use information in the form of electric, climate and calendar data. The results are usually presented for typical weekdays or holidays. In the case of holidays, relatively different forecasting errors are obtained, as the power load patterns of holidays show large differences compared to typical weekdays. Furthermore, the first day of the week (Monday) and the last day (Friday) are characterized by significant deviations from the middle days of the week. This fact justifies the development of forecasting models for a specific day of the week.

The results, presented in many publications [19,26,43,44,46,47,50,51], confirm that the ANFIS model is a reliable forecasting tool. Particularly good results are achieved in the cases of time series with cyclical variability. A mean square error (MSE) below 3% is obtained for different series of power system loads but higher MSE for power company profiles.

The forecast results for the next week of the power system loads, presented in [50], are recognized to be perfect and consistent. A very high accuracy of the ex post forecast was achieved. On the other hand, when using the ANFIS model for medium- or long-term forecasting, the results are usually even better, e.g., the MAPE is equal to 0.4% when monthly electricity demand was forecasted in [51].

It may seem that the results of the ex post forecast accuracy presented in the article are not good because the MAPE is equal to approx. 8–9%. However, the results should be compared with the MAPE equal to approx. 9–11% for the naive forecast, which is always a certain point of reference. The MAPE of the ex post load forecast depends on the nature of the time series of the power company selected for the analysis.

Many publications [4,5,28,50] indicate that the global demand for electricity is influenced by climate variables such as temperature and cloudiness. However, their impact depends on the climate zone. A change in summer temperature in France by one degree causes a change in demand by 500 MW and in winter by about 1450 MW [50]. The power demand changes depend on the use of heating or air-conditioning systems. Climate variables should definitely be applied in medium and long-term forecasting of electricity demand whereas taking them into consideration in the STLF may lead to an imprecise forecast. As shown in the article, the mean daily temperature as the exogenous variable does not improve the accuracy of the ANFIS model for the weekdays. However, the influence of temperature on the load profile on Sundays was observed.

9. Discussion

Recently, a trend to avail of artificial intelligence methods in short-term forecasting has been clearly visible. Using such methods, we lose an opportunity to detect the cause and effect relationships. Most of the machine-type learning algorithms are used as “black-boxes” and the researcher has limited capability to shape the final form of the model. The database and its information capacity, prepared by a researcher, affect usually the final result. Here, the rule garbage in, garbage out (GIGO) is used. Each of the methods adopts one or the whole set of measures of the ex post forecasting model (e.g., Mean Absolute Percentage Error MAPE, Root Mean Square Error RMSE, Mean Absolute Scaled Error MASE, Sum of Squared Errors SSE, Random-Walk Coefficient of Determination R^2 , Root Mean Squared Scaled Error RMSSE, etc.). Note that using different measures generates extremely different forecasts. In general, using the measures of square errors (RMSE, SSE, MSE or R^2) and their optimization leads to an optimum forecast being a conditional expected value, the use of MASE and MAE measures model a conditional median and MAPE measures provide a forecast that is difficult to interpret. Artificial intelligence methods are based on the process of learning from pieces of information given in the database. The verification process is performed on a separated part of the database. This procedure undoubtedly provides very good results in ultra short-term forecasts, usually used in the process of controlling various objects. Additionally, in short-term forecasting or even in medium-term forecasting of

stable behavior of variables, good results can still be achieved. However, in the case of forecasting technological changes, the results may be unsatisfactory.

The paper shows that the proper selection of the exogenous variables set has an impact on the accuracy of the model. The Hellwig method using integral capacities of information carriers for the selection of an optimum set of exogenous variables is used to construct models reflecting relationships between explaining (exogenous) variables and the explained (endogenous) variable.

The article examines the impact of taking into account an additional exogenous variable—the ambient temperature—on the accuracy of ex post load forecasts. It was demonstrated that for the Polish data adapted for the construction of the ANFIS model, the ambient temperature as an additional input variable does not improve the accuracy of forecasts for the weekdays. The influence of temperature on the load forecasts on Sundays is significant, which confirms the sensitivity of the residential sector loads to temperature changes. It should be however emphasized that this is only a case study for the data of the power subsystem with a load structure with a predominance of the industrial sector. It is difficult to draw general conclusions from a case study, so it makes sense to continue the research for a broad spectrum of data.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

- Meier, A. *Electric Power Systems: A Conceptual Introduction*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006.
- Popławski, T.; Lyp, J. Forecasting Methods Supporting the Operator of Transmission System—Chosen Issues. *Przegląd Elektrotechniczny* **2010**, *86*, 192–195.
- Weron, R. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2006.
- Mirasgedis, S.; Sarafidis, Y.; Georgopoulou, E.; Lalas, D.P.; Moschovits, M.; Karagiannis, F.; Papakonstantinou, D. Models for Mid-Term Electricity Demand Forecasting Incorporating Weather Influences. *Energy* **2006**, *31*, 208–227. [\[CrossRef\]](#)
- Taylor, W.J.; Buizza, R. Using weather ensemble predictions in electricity demand forecasting. *Int. J. Forecast.* **2003**, *19*, 57–70. [\[CrossRef\]](#)
- Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*, 2nd ed.; OTexts, 2018. Available online: <https://otexts.org/fpp2/> (accessed on 25 October 2020).
- Hellwig, Z. *Linear Regression and Its Application to Economics*; Pergamon Press: Oxford, UK, 1963.
- Cassotti, M.; Grisoni, F. Variable Selection Methods: An Introduction. University of Milano-Bicocca. Available online: www.moleculardescriptors.eu (accessed on 15 April 2020).
- Singh, R.P.; Gao, P.X.; Lizotte, D.J. On hourly home peak load prediction. In Proceedings of the 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm), Tainan, Taiwan, 5–8 November 2012; pp. 163–168.
- Hong, T.; Fan, S. Probabilistic electric load forecasting: A tutorial review. *Int. J. Forecast.* **2016**, *32*, 914–938. [\[CrossRef\]](#)
- Taylor, J.W.; McSharry, P.E. Short-term load forecasting methods: An evaluation based on European data. *IEEE Trans. Power Syst.* **2007**, *22*, 2213–2219. [\[CrossRef\]](#)
- Box, G.; Jenkins, G.; Reinsel, G. *Time Series Analysis: Forecasting and Control*; Wiley Series in Probability and Statistics; Wiley: Hoboken, NJ, USA, 2008.
- Taylor, J.W. Short-term electricity demand forecasting using double seasonal exponential smoothing. *J. Oper. Res. Soc.* **2003**, *54*, 799–805. [\[CrossRef\]](#)
- Arora, S.; Taylor, J.W. Short-term forecasting of anomalous load using rule-based triple seasonal methods. *IEEE Trans. Power Syst.* **2013**, *28*, 3235–3242. [\[CrossRef\]](#)
- Taylor, J.W. Triple seasonal methods for short-term electricity demand forecasting. *Eur. J. Oper. Res.* **2010**, *204*, 139–152. [\[CrossRef\]](#)
- Arora, S.; Taylor, J.W. Forecasting electricity smart meter data using conditional kernel density estimation. *Omega* **2016**, *59*, 47–59. [\[CrossRef\]](#)
- Dudek, G. Short-Term Load Forecasting Based on Kernel Conditional Density Estimation. *Przegląd Elektrotechniczny* **2010**, *86*, 164–167.

18. Ciabattini, L.; Grisostomi, M.; Ippoliti, G.; Longhi, S. A Fuzzy Logic tool for household electrical consumption modeling. In Proceedings of the IECON 2013—39th Annual Conference of the IEEE Industrial Electronics Society, Vienna, Austria, 10–13 November 2013; pp. 8022–8027.
19. Zahedi, G.; Azizi, S.; Bahadori, A.; Elkamel, A.; Wan Alwi, S.R. Electricity demand estimation using an adaptive neuro-fuzzy network: A case study from the Ontario province—Canada. *Energy* **2013**, *49*, 323–328. [\[CrossRef\]](#)
20. Santra, A.S.; Lin, J.-L. Integrating Long Short-Term Memory and Genetic Algorithm for Short-Term Load Forecasting. *Energies* **2019**, *12*, 2040. [\[CrossRef\]](#)
21. Park, R.-J.; Song, K.-B.; Kwon, B.-S. Short-Term Load Forecasting Algorithm Using a Similar Day Selection Method Based on Reinforcement Learning. *Energies* **2020**, *13*, 2640. [\[CrossRef\]](#)
22. Alpaydin, E. *Introduction to Machine Learning*, 4th ed.; MIT: Cambridge, MA, USA, 2020; Volume xix, pp. 1–3, 13–18, ISBN 0262043793.
23. Aurangzeb, K. Short Term Power Load Forecasting using Machine Learning Models for energy management in a smart community. In Proceedings of the 2019 International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, 3–4 April 2019; pp. 1–6.
24. Negnevitsky, M.; Mandal, P.; Srivastava, A.K. Machine Learning Applications for Load, Price and Wind Power Prediction in Power Systems. In Proceedings of the 15th International Conference on Intelligent System Applications to Power Systems, Curitiba, Brazil, 8–12 November 2009; pp. 1–6.
25. Wei, D.; Wang, J.; Ni, K.; Tang, G. Research and Application of a Novel Hybrid Model Based on a Deep Neural Network Combined with Fuzzy Time Series for Energy Forecasting. *Energies* **2019**, *12*, 3588. [\[CrossRef\]](#)
26. Jin, X.B.; Wang, J.F.; Zhang, H.Y.; Cao, L.H. ANFIS Model for Time Series Prediction. *Appl. Mech. Mater.* **2013**, *385–386*, 1411–1414. [\[CrossRef\]](#)
27. Zalewski, W. Statistical Analysis of the Load Variability in Distribution Network. *Econ. Manag.* **2010**, *4*, 203–213.
28. Pardo, A.; Meneu, V.; Valor, E. Temperature and Seasonality Influences on Spanish Electricity Load. *Energy Econ.* **2002**, *24*, 55–70. [\[CrossRef\]](#)
29. Akil, Y.; Miyauchi, H. Seasonal Regression Models for Electricity Consumption Characteristics Analysis. *Engineering* **2013**, *5*, 108–114. [\[CrossRef\]](#)
30. Ullah, A.; Wang, H. Parametric and Nonparametric Frequentist Model Selection and Model Averaging. *Econometrics* **2013**, *1*, 157–179. [\[CrossRef\]](#)
31. Miller, A. *Subset Selection in Regression*, 2nd ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2002.
32. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 2nd ed.; Springer: Berlin, Germany, 2009.
33. Akaike, H. Fitting autoregressive models for prediction. *Ann. Inst. Stat. Math.* **1969**, *21*, 243. [\[CrossRef\]](#)
34. Goldberg, D.E. *Genetic Algorithms in Search, Optimization, and Machine Learning, Reading*; Kluwer Academic: Boston, MA, USA, 1989.
35. Leardi, R.; Boggia, R.; Terrile, M. Genetic algorithms as a strategy for feature selection. *J. Chemom.* **1992**, *6*, 267–281. [\[CrossRef\]](#)
36. Clerc, M. *Particle Swarm Optimization*; ISTE Ltd.: London, UK, 2006.
37. Dorigo, M.; Blum, C. Ant colony optimization theory: A survey. *Theor. Comput. Sci.* **2005**, *344*, 243–278. [\[CrossRef\]](#)
38. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. B* **1996**, *58*, 267–288. [\[CrossRef\]](#)
39. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. B* **2005**, *67*, 301–320. [\[CrossRef\]](#)
40. Indahl, U.; Liland, K.; Næs, T. Canonical partial least squares: A unified PLS approach to classification and regression problems. *J. Chemom.* **2009**, *23*, 495–504. [\[CrossRef\]](#)
41. Jang, J.-S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [\[CrossRef\]](#)
42. Rutkowski, L. *Flexible Neuro-Fuzzy Systems: Structures, Learning and Performance Evaluation*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2004.
43. Sowiński, J.; Szydłowski, M. Forecast of Electricity Supply Using Adaptive Neuro-Fuzzy Inference System. In Proceedings of the 18th International Scientific Conference on Electric Power Engineering, Kouty nad Desnou, Czech Republic, 17–19 May 2017; Rusek, S., Gono, R., Eds.; IEEE: New York, NY, USA, 2017.
44. Sowiński, J. Medium-Term Forecast of Monthly Supply of Electricity and Maximum-Demand-Day Power Using Adaptive Neuro-Fuzzy Inference System. In Proceedings of the 9th International Scientific Symposium on Electrical Power Engineering, Stara Lesna, Slovakia, 12–14 September 2017; Kolcun, M., Kolcunova, I., Kurimsky, J., Eds.; Curran Associates, Inc.: Stara Lesna, Slovakia, 2017.
45. Ying, L.-C.; Pan, M.-C. Using adaptive network based fuzzy inference system to forecast regional electricity loads. *Energy Convers. Manag.* **2008**, *49*, 205. [\[CrossRef\]](#)
46. Prado, F.; Minutolo, M.C.; Kristjanpoller, W. Forecasting based on an ensemble Autoregressive Moving Average—Adaptive neuro—Fuzzy inference system—Neural network—Genetic Algorithm Framework. *Energy* **2020**, *197*. [\[CrossRef\]](#)
47. Acakpovi, A.; Ternor, A.T.; Asabere, N.Y.; Adjei, P.; Iddrisu, A.-S. Time series prediction of electricity demand using Adaptive Neuro-Fuzzy Inference Systems. *Math. Probl. Eng.* **2020**, *2020*, 4181045. [\[CrossRef\]](#)
48. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [\[CrossRef\]](#)

-
49. Zadeh, L.A. Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. *IEEE Trans. Syst. Man Cybern.* **1973**, SMC-3, 28–44. [[CrossRef](#)]
 50. Mordjaoui, M.; Boudjema, B. Forecasting and Modelling Electricity Demand Using Anfis Predictor. *J. Math. Stat.* **2011**, 7, 275–281. [[CrossRef](#)]
 51. Tay, K.G.; Muwafaq, H.; Tiong, W.K.; Choy, Y.Y. Electricity Consumption Forecasting Using Adaptive Neuro-Fuzzy Inference System (ANFIS). *Univers. J. Electr. Electron. Eng.* **2019**, 6, 37–48. [[CrossRef](#)]