

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

**4,800**

Open access books available

**122,000**

International authors and editors

**135M**

Downloads

Our authors are among the

**154**

Countries delivered to

**TOP 1%**

most cited scientists

**12.2%**

Contributors from top 500 universities



**WEB OF SCIENCE™**

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.

For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Electricity Consumption and Generation Forecasting with Artificial Neural Networks

---

Adela Bâra and Simona Vasilica Oprea

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.71239>

---

## Abstract

Nowadays, smart meters, sensors and advanced electricity tariff mechanisms such as time-of-use tariff (ToUT), critical peak pricing tariff and real time tariff enable the electricity consumption optimization for residential consumers. Therefore, consumers will play an active role by shifting their peak consumption and change dynamically their behavior by scheduling home appliances, invest in small generation or storage devices (such as small wind turbines, photovoltaic (PV) panels and electrical vehicles). Thus, the current load profile curves for household consumers will become obsolete and electricity suppliers will require dynamical load profiles calculation and new advanced methods for consumption forecast. In this chapter, we aim to present some developments of artificial neural networks for energy demand side management system that determines consumers' profiles and patterns, consumption forecasting and also small generation estimations.

**Keywords:** forecast, renewable energy, smart metering, demand side management, consumers' profiles

---

## 1. Introduction

Recently, many national and international communities and authorities developed energy efficiency strategies and programs in order to reduce energy poverty. In Ref. [1], European Economic and Social Committee (EESC) stated that more than 50 million Europeans are affected by energy poverty in 2009. EESC also recommends to establish a European poverty observatory that will bring together all stakeholders to take correct measures to reduce the gaps between different countries and regions in terms of energy poverty and propose a set of statistics indicators to monitor the evolution of energy efficiency. EESC draw attention that energy prices are constantly increasing with more than 10% annually, while most of the Europeans spending an increasing share of their income on energy.

In 2012, European Commission adopted Energy Efficiency Directive that proposed measures to increase with 20% energy efficiency target by 2020 [2]. On November 30, 2016, the Commission updated the Directive, by targeting 30% energy efficiency for 2030. The proposed measures for energy are oriented toward increasing consumers' awareness regarding their consumption management through electronic bills and information and communications technology (ICT) solutions, encourage them to become prosumers by investing in their own generation sources such as photovoltaic (PV) panels, wind turbines and storage devices.

The main objective of the chapter is to present an implementation of artificial neural networks (ANNs) for the electricity consumption management based on smart metering (SM) data. This objective will be reached by following topics:

- determining consumers' profiles and patterns with clustering and self-organizing maps (SOM);
- forecasting aggregated electricity consumption for short-term period on a typical week day with autoregressive (AR) neural networks;
- forecasting energy generation for small wind turbines and photovoltaic panels installed at consumers side (prosumers) with feed-forward neural networks;
- presenting the main components of an informatics prototype that allows the prosumers to configure and schedule their appliances in an interactive manner to optimize the electricity consumption.

The ANN performance will be compared with stochastic methods (classification, ARMA and ARIMA models) and the best solution is adopted for ICT prototype.

## 2. Current problems in electricity consumption and generation

Regarding ICT solutions, the most important measures to reduce the energy poverty and to increase consumers' awareness toward energy efficiency concern both electricity suppliers and consumers. For electricity suppliers, market segmentation can be used to determine dynamic consumers' profiles to better understand consumption behavior and also to set up strategies and plans for different consumers groups. Another important measure for electricity suppliers consists in consumption (load) forecasting for short and medium term, used for planning the grid resources and wholesale electricity markets bids. For consumers, with the introduction of smart metering (SM) systems, their awareness increased and new methods must be taken into consideration such as consumption optimization of household appliances through user-friendly interfaces, micro-generation (through photovoltaic panels, small wind turbines and storage devices), mobile applications for real time billing with detailed information regarding appliances' consumption or own generation sources.

### 2.1. Determine consumers' profiles

Determining dynamic profiles for consumers represents a challenge for energy suppliers due to the widespread implementation of smart metering (SM). Comparing with previous period

before SM implementation, the consumers can play an active role, having the opportunity to control and schedule their consumption by programming some devices such as washing machine, electric heating, ventilation systems or car batteries. Another aspect that can be considered is to use micro-generation sources (photovoltaic panels installed on the roofs or buildings' facades or small wind turbines) to unload back into the grid the generated electricity according to the tariff systems. Based on activities that are carried out by consuming electricity, final consumers are categorized into household and non-household consumers. Final non-household consumers are characterized by specific consumption, defined by profiles (consumption curves) defined in Romania by procedure [3] and are split into several categories such as industry, gas stations, civil works, hospitals, public utilities, hotels, retailers, etc. For household consumers in Romania, there is no official procedure or study used by energy suppliers, although SM is targeted for 80% implementation until 2020. Based on international studies such as [4] conducted in UK during 2010–2011 by the Department of Energy and Climate Change (DECC), the Department for the Environment Food and Rural Affairs (DEFRA) and the Energy Saving Trust (EST), demonstrated clearly the influence of time-of-use tariff (ToUT) on domestic demand response. The consumers shifted their evening peak consumption due to the various ToUT prices without significantly affecting their comfort and lifestyle. In Ref. [5], it is considered a multivariate statistical lifestyle analysis of household consumers in US. The study identified five factors reflecting social and behavioral profiles (patterns) determined by air conditioning, washing machines, climate conditions, PC and TV use. A detailed analyses for market segmentation is presented in Ref. [6] based on effect of lifestyles, socio-demographic factors, smart appliances, electricity and heat supply. Authors also identified the most influenced factors that determined the segmentation: socio-demographic factors (household size, net income and employment status), types of electric appliances and the use of new (smart) technologies. They also correlated the link between socio-demographic factors and the use of new technologies and smart appliances. The household profiles are determined in Ref. [7] by applying a time series auto-regression model—Periodic Autoregression with eXogenous variables (PARX) algorithm, taking into consideration temperature and occupants' daily habits. Thus, the consumption is correlated with life style influenced by temperature that leads to air conditioning and heating electricity consumption variations.

In Ref. [8], we analyzed several methods for profile calculation including fuzzy C-means clustering, autoregression with exogenous variables and multi-linear regression. For National Grid UK, the load profiles are determined with multiple regression taking into consideration seven variables such as noon temperature, two variables regarding sunset moments compare to 6 o'clock in the afternoon and four variables related to the week days [9]. Thus eight profiles are obtained that can be used further for electricity consumption forecasts and simulation.

In Ref. [10], authors applied autoregression on hourly consumption data measured for 1000 household consumers in Canada (Ontario). The household consumption represents 30% of the total consumption and its contribution to the peak load is important due to the ventilation or AC devices. For autoregression, the variables also include hourly temperature data measured at local stations and occupation degree of each house.

In Ref. [11], clustering methods are used for determining load profiles for 300 electricity consumers from Malaysia. Authors proposed C-means fuzzy clustering algorithm by using

average consumption values, thus obtaining four clusters, each of them being split into sub-clusters for a better and detailed typical consumption profiles.

In Section 3.1, we proposed a new method based on self-organizing maps (SOM) that allows us to determine six profiles clearly delimited for consumers having the following types of consumption: heating, cooling, ventilation, interior lighting, exterior lighting, water heating, usual devices (washing machine, refrigerator) and other smaller devices (TV, audio and computer). These profiles were compared with other profiles obtained by stochastic methods such as clustering and classification.

## **2.2. Consumption and micro-generation short-term forecasting**

Regression is seen as part of the first generation of consumption (load) forecasting methods. It is one of most widely used statistical methods due to its undoubtable advantages such as simplicity and transparency. For electricity load forecasting, regression methods are usually applied to effectively model the relationship of consumption level and other factors such as weather (i.e. temperature, humidity, etc.), day type (workdays and holidays) and consumers profiles.

Several methods based on regression have been used for short-term load forecasting with different levels of success such as ARMAX models [12], multiple regression [13, 14] and regression with neural networks [15–17].

In Ref. [18], authors describe several regression models for the next day peak forecasting. Their models incorporate deterministic influences such as weekend days, stochastic influences such as historical loads, and exogenous factors influences such as temperature. In papers [19–22], authors described other applications of regression models to load forecasting.

According to [23, 24], ARIMA models have proven appropriate for forecasting electricity consumption.

In Section 3.2, we proposed a method based on autoregressive neural networks for short-term forecasting the electricity consumption aggregated at supplier's level for a typical day of the week. The forecasting method is applied on each profile previously determined by SOM. Also, we considered ARIMA method for load forecasting and at the end of Section 3.2, we compared the results for both methods.

Regarding the micro-generation forecasting (small wind turbines and photovoltaic panels installed at consumers' side), the methods depend on the time interval. For example, stochastic methods (persistence and autoregressive patterns) are recommended in Ref. [25] for very short-term prediction (up to 4–6 hours). In addition, other authors [26] proposed Kalman integrated support vector machine (SVM) method to achieve a 10% accuracy improvement by comparing with artificial neural networks or autoregressive (AR) methods. Also a consistent approach is given by the use of ANNs for short-term generation forecast in case of wind turbines and photovoltaic (PV) panels. Various ANN-based algorithms are described in [27, 28], it is proposed Bayesian Regularization algorithms for forecasting. Also in [29, 30], authors proposed back propagation neural networks based on the optimization of Swarm particles.

Stochastic methods can be successfully used in order to determine the PV generation. The authors of the paper [31] analyzed and compared different models for forecasting and concluded that the accuracy of the ARMA model is better than other models.

In Section 4, we analyzed two methods for PV and small wind turbines generation: stochastic method based on ARIMA and feed-forward ANN. The results are compared and conclusions are drawn at the end of the section.

### **2.3. Consumption optimization**

Residential consumers usually have certain types of appliances: washing machine, dryer, dish machine, water heater, refrigerator, electric oven/grill, blender, iron machine, electric centralized heating system, coffee maker, vacuum cleaner, AC/ventilation systems, TV and other multimedia appliances. Out of these appliances only some of them can be automatically controlled and used at certain time intervals when electricity price is lower (e.g. washing machine, dish machine, electric oven, car batteries can be charged at night). Electricity suppliers may use several methods for consumption optimization using different optimization functions. In Ref. [32], it is described as stochastic optimization based on Monte Carlo simulation for minimization of estimated payment for entire day. Authors proposed a mixed integer linear programming (MILP) algorithm for optimization of the electricity residential consumption taking into account real time tariffs. In [33, 34], authors proposed genetic algorithms for consumption optimization. We proposed in Ref. [35] a method based on MILP with two optimization functions: cost-based function that minimizes the electricity costs depending on the time-of-use tariffs (ToUT) and a peak-shaving optimization function that minimizes the peak consumption. Both functions provide savings to electricity bills for consumers, but the second method also brings benefits for electricity suppliers and grid operators. In Section 5, we presented informatics solution that integrates the optimization methods presented in Ref. [35] and allows consumers to schedule their appliances in order to minimize the electricity costs.

### **2.4. Smart applications for real time billing**

Until the expansion of smart meters, consumers were charged based on meters' reads made by an electricity supplier employee on different time intervals (usually 2–3 months) and most of the time the bills are based on estimations determined by the energy supplier on historical data regarding each household's consumption. Therefore, the electricity consumer was unable to customize and adjust his consumption based on ToUT or schedule his appliances to avoid peak consumption because he cannot benefit from real time information and his behavior is not reflected in the demand management system in real time. Since the widespread implementation of smart meters in most European countries, various informatics solutions were developed by software companies or by energy suppliers in order to provide consumers accurate electricity bills, near real time. A review of the top utility billing software products is available in Ref. [36]. These solutions are user-friendly, accessible online through mobile devices, intuitive and ease to use even for ICT novices.

Besides information regarding the total consumption, the billing systems have to provide consumptions data for different type of appliances measured by SM or by other smart measurement

devices. Thus, consumers can analyze their consumption for heating, cooling, washing, lighting and other home appliances and they can schedule it based on ToUT. In Section 5, we proposed an informatics solution that provides friendly user interfaces and integrates methods for consumption optimization and micro-generation forecasting for electricity consumers.

### 3. Forecasting the electricity consumption

#### 3.1. Determining the consumption profiles

In order to dynamically determine consumer profiles, first we considered a series of algorithms based on classification and clustering techniques. In order to implement and test the model, we used a data set with hourly electricity consumption recorded in different US cities between January 1, 2014 and December 31, 2014. Each record contains values for the following types of consumption: heating, cooling, ventilators, indoor lighting, outdoor lighting, water heating, household equipment (washing machine and refrigerator) and other interior devices (TV and computer). Data were imported into Oracle Database 11g R2 in the *LOAD\_PROFILE\_T* table with approximately 1,900,000 hourly records for 212 consumers. We analyzed the distribution of electricity consumption at different value ranges, consumption types and time periods as shown in **Figure 1**.

The analyses shows that the consumption curve has the same aspect as the consumption for heating and interior equipment, which makes these types of consumption significant attributes for the total consumption value.

Data being imported into Oracle Database, we consider data mining algorithms developed in Oracle SQL Developer. So, for the first method, we approached support vector machines (SVMs) classification method and we build six profiles (classes) and the profiles with the most cases (over 30,000) have the highest degree of accuracy (about 90%), which can be considered a

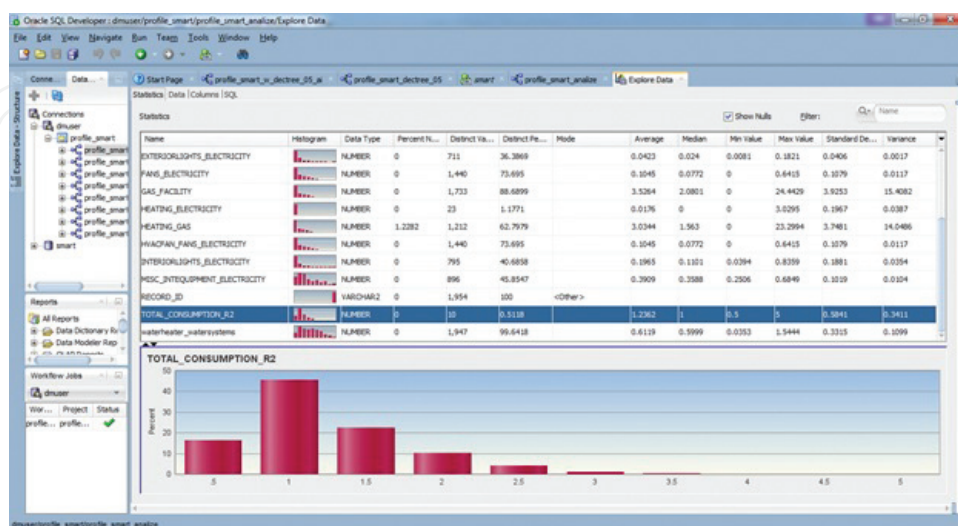
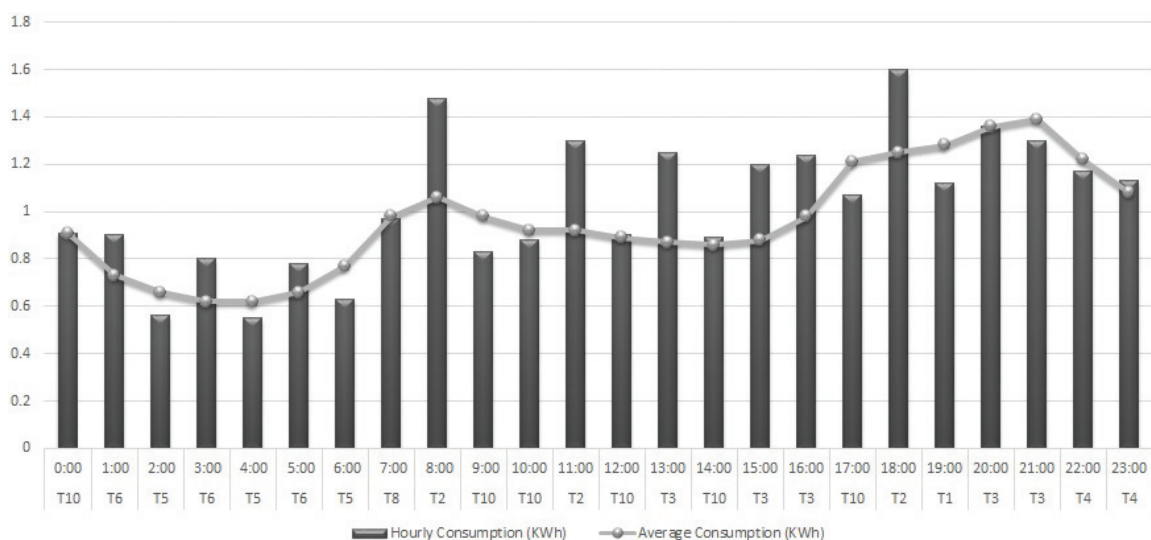


Figure 1. Data set statistics.

good result for classification. Performing classes' analyses, we observed that the profiles are very sensitive to changes in consumer behavior, due to the fact that classes with a small number of items recorded the highest prediction errors.

To eliminate these shortcomings, we considered it useful to apply a second solution for determining profiles dynamically by using clustering methods. For building the profiles, we applied the K-means method and for measuring similarity within a cluster, the variance (the sum of the squares of the differences between the main element and each element) is used, being the best clusters in which the variance is small. We analyzed the confidence level for each cluster and it is noticeable that the confidence is high, in most cases being over 85%. Regarding the clustering rules, from our results we noticed that the grouping rules do not take into account the attributes such as water heating, fans, cooling, household equipment, indoor/outdoor lighting, but only heating and total consumption (the most important attributes). This may be due to the fact that we choose a small number of clusters comparing to the data set population. In conclusion, the lower the number of clusters, the more people in the group and less sensitive to changes in consumer behavior.

In order to divide the obtained profiles into smaller groups, we choose another clustering method in order to establish consumption patterns. So, we refined the K-mean results and we applied O-cluster method (Orthogonal partitioning clustering). This method is owned by Oracle Corporation [37] and uses a recursive data grouping algorithm through orthogonal data partitioning. On top of the previous 6 profiles determined by K-means, we build 10 sub-clusters, representing consumption patterns for each profile on hourly intervals. Analyzing the training rules and the weight of each consumption category in each cluster, we noticed that they have a varied composition, each cluster identifying a primary profile determined by the K-means method and one or more consumption patterns determined by the O-cluster method. For example, we considered the distribution of consumption patterns of a consumer within the P5 profile within 24 h. **Figure 2** shows profile P5 split into 10 patterns (T1, ..., T10) for a detailed perspective on electricity consumption.



**Figure 2.** Profile P5 patterns with O-cluster.



The patterns build with O-cluster refine the clusters and gives a better understanding about consumption behavior regarding smaller groups of consumers and thus adjusting the ToUT for these groups. Also, the consumption patterns shape more accurately the consumer's dynamic behavior within 24 h, the profiles being in fact an approximation of the variation of hourly consumption. The deviations of the actual consumption compared to the average consumption of the profile are small, which again validates the clustering model.

As an option to clustering methods, we approached also a third method based on artificial neural networks (ANN). In Matlab R2015a, we imported data from Oracle Database from *LOAD\_PROFILE\_T* table and we organized input vectors as  $x(t) \in R^n$ , where  $n = 13$  for each consumption type (heating, ventilation, indoor lighting, etc.) and  $t$  represents time interval (hours) between January 1, 2014 and December 31, 2014.

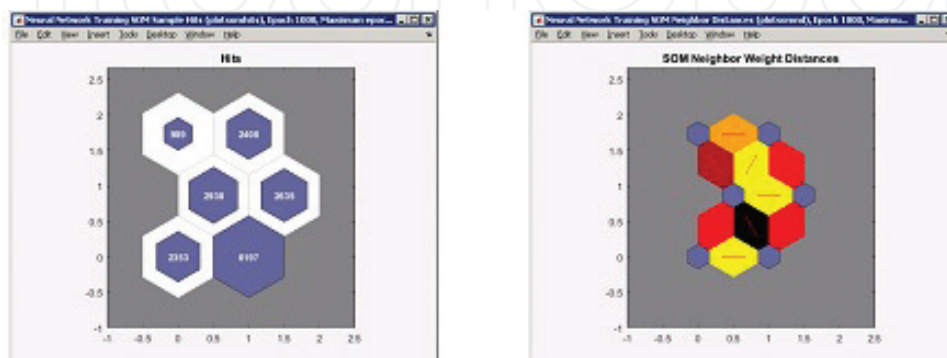
We developed a self-organizing maps (SOM) algorithm, setting the following parameters for the neural network:

- SOM architecture—2D with  $2 \times 3$  neurons/layer (dimensions) = [2 3];
- number of steps for initially processing the input space (coverSteps) = 100;
- initial neighbor (initNeighbor) = 2;
- network topology (topologyFcn) = 'hextop' and
- distance between neurons (distanceFcn) = 'linkdist'.

The network is initialized with random values for each neuron. We used the *trainbu* training function that adjusts weights and bias after each iteration. We plotted the results and observed the distribution of the input set in **Figure 3**:

From the representation of the consumption curves corresponding to six clusters, it can be observed a clear delimitation between profiles P2 and P5. Also, a difference of approx. 30% of the evening consumption peak is observed between P6 and P1, P3, P4 (**Figure 4**).

Following the analysis of the obtained results, we noticed a correct and efficient grouping of the consumer profiles using the self-organizing neural networks.



**Figure 3.** Distances among clusters distribution.

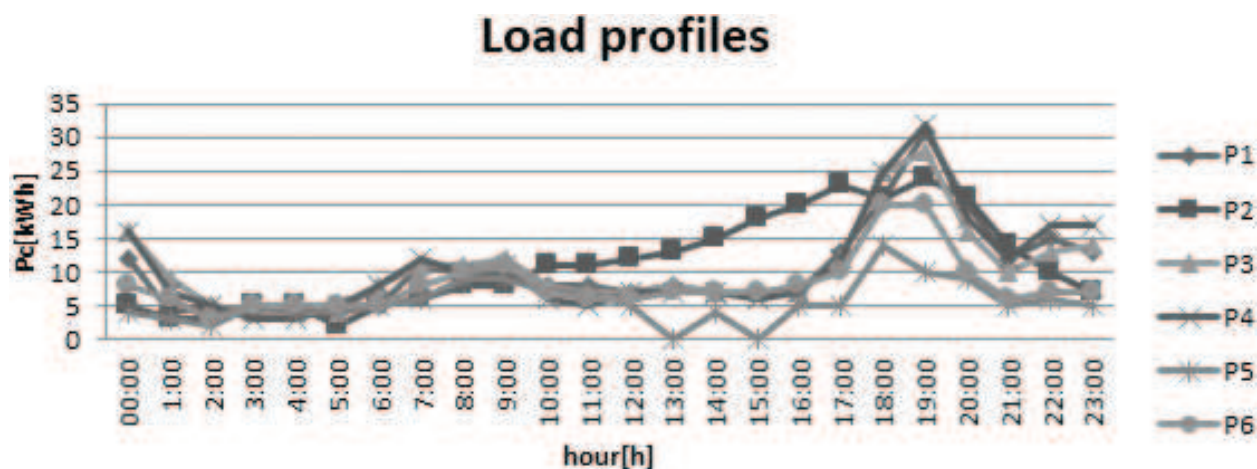


Figure 4. Profiles obtained with SOM.

A short comparison of the results obtained with the three analyzed methods is summarized in Table 1.

From the analysis, we can conclude that for the determination of dynamic consumption profiles, which surprising a series of consumption patterns, the optimal method is the clustering method, and for the determination of clearly delimited profiles the most efficient method is the use of self-organizing maps.

### 3.2. Consumption forecasting solution with ANN

Analyzing the consumption data set for 212 consumers during 4–6 weeks, a regular pattern is observed between working days or weekdays (Monday to Friday) and some differences in weekend or holidays. Therefore, for load forecasting hourly aggregated at grid operator or electricity supplier level for a typical day of the week, we can consider an autoregressive model. In this section, we approach and compare two methods for forecasting electricity consumption: statistics methods based on ARIMA and autoregressive artificial neural networks.

Autoregressive-moving-average (ARMA) models are suitable for stationary series, but most of the series are non-stationary, their mean and variance not being constant over time. The ARMA model was adapted for non-stationary time series that become stationary by differentiation,

Method	SVM	K-means and O-cluster	SOM
Number of profiles	6 profiles	6 profiles with 10 patterns	6 profiles
Sensitivity to consumption variations	High, small classes with lower confidence	Medium, variations are included in patterns	Medium, each group is clearly delimited
Detailed consumption information	High (sub-types of profiles)	High (by patterns of O-cluster)	Low
Overall performance	Medium	High	High

Table 1. Comparison of the profiles obtained by SVM, K-means and O-cluster and SOM.

the resultant models being called autoregressive integrated moving average ARIMA (p, d, q). The ARIMA model (p, d, q) consists of three parts: autoregressive (AR), where p represents the autoregression order, d represents the order of differentiation required for staging the series (I) and the moving average, q being the order of the moving average. Unlike autoregression, the moving average describes phenomena with certain irregularities. Moving average is described by the following equation:

$$Y_t = c + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_p e_{t-p} + e_t \quad (1)$$

where  $Y_t$  is the consumption,  $c$  is a constant coefficient and the  $\theta$  are the parameters of the moving average and  $e_t$  represents the time series error.

To evaluate the results of the analysis, we used the mean squared error (MSE) and also mean absolute percentage error (MAPE) to compare the accuracy of the forecast obtained in various variants of the ARIMA model.

Data from the *LOAD\_PROFILE\_T* table were imported into the SAS Guide Enterprise 7.1. Starting from the input data set, we applied the autoregressive integrated moving average models. In **Table 2**, we presented MAPE for the AR model first order, ARMA(1,1) and ARIMA(1,1,1).

**Table 2** shows that MAPE is the lowest in the autoregressive model, the accuracy of the electricity consumption forecast being the best (about 93%). The accuracy of other forecasts is over 70%. In all analyses, the degree of correlation indicates an average or poor inverse dependence.

In addition to ARIMA models, we approached the autoregressive neural networks in Matlab. We built the *LOAD\_PROFILE\_HOURLY* virtual table based on the *LOAD\_PROFILE\_T* table and the *LOAD\_PROFILE\_SOM\_6* table, which includes six consumption profiles previously determined by the self-organizing maps. For simulation we considered a single profile—P6 with the largest number of consumers (6197).

Due to the structure of the input data and the fact that there is an autoregressive component of electricity consumption during a typical week, we have built a nonlinear autoregressive neural network (*narnet*). We configured ANN parameters as follows:

- `feedbackDelays`—number of delays;
- `hiddenSizes`—number of neurons in the hidden layer;
- `trainFcn`—training function.

We considered 50 neurons in the hidden layer and a single input  $y(t)$ —the total consumption determined according to the formula:

Model	MAPE [%]
AR(1)	7.29
MA(1)	24.45
ARMA(1,1)	29.05
ARIMA(1,1,1)	24.97

**Table 2.** ARIMA models' results.

$$y(t) = f(y(t - 1), \dots, y(t - d)) \quad (2)$$

where  $d$  represents the number of records considered delays. For the first iteration of the model, we considered  $d = 5$  and for the second iteration with better results  $d = 10$ . The architecture of the network is shown in **Figure 5**.

For the hidden layer, we used a bipolar sigmoid activation function and a linear activation function for the output layer. As for the training algorithm, Matlab provides the following algorithms: the Levenberg-Marquardt (LM) algorithm (*trainlm*), the Bayesian Regularization (BR) algorithm (*trainbr*) and the Scaled Conjugate Gradient (SCG) algorithm (*trainscg*). We developed the autoregressive neural network and compared the results obtained with the three training algorithms. The performance of the network is very good, the mean square error (MSE) being 0.0046 attained at epoch 936 for the BR training algorithm and the correlation coefficient  $R$  between the prediction and the actual value is 0.996 (Figure 6).

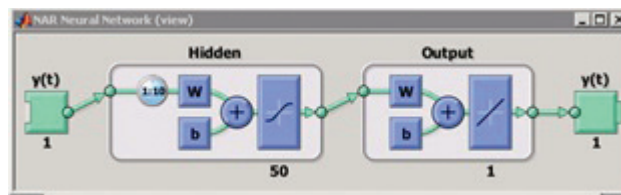


Figure 5. The architecture of the autoregressive neural network.

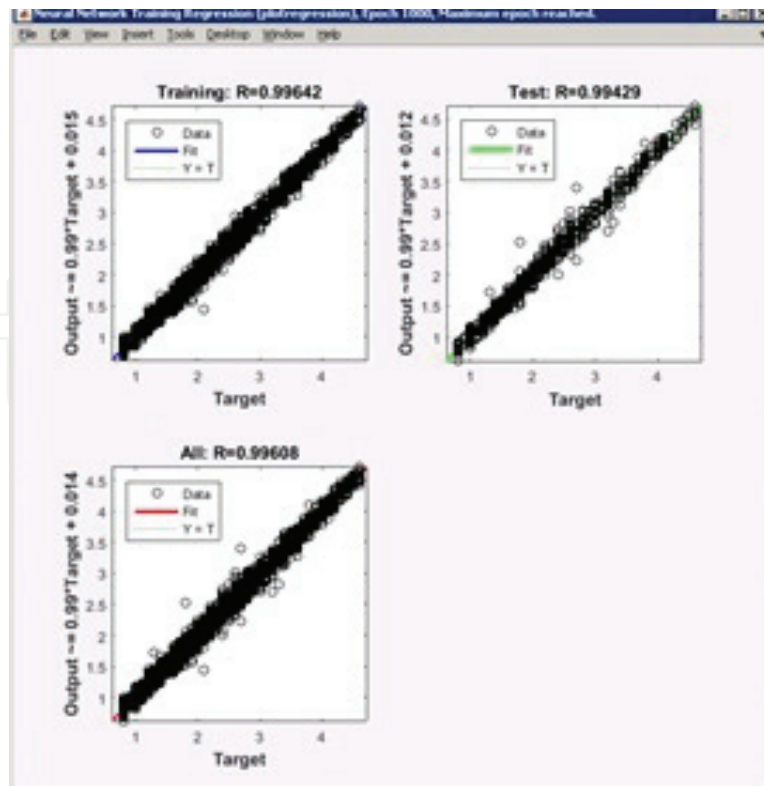


Figure 6. Results for  $R$  coefficient for BR algorithm.

From the error histogram (**Figure 7**), it can be observed that the errors are between  $-0.13$  and  $+0.12$ , which can be considered an acceptable distribution.

We trained the network using the three algorithms (LM, RB and SCG), the best results being recorded using the Bayesian Regularization algorithm, although the Levenberg-Marquardt algorithm recorded good results with an increased performance in training.

In **Table 3**, the results obtained with autoregressive neural networks are compared with stochastic methods (ARMA, ARIMA and AR).

The accuracy of ANN algorithms is better (about 95%) compared to the accuracy of stochastic models. Also, the Levenberg-Marquardt and Bayesian regularization algorithms are also superior regarding the lowest MSE. The R coefficient and error distribution for neural network algorithms are better than AR, MA, ARMA and ARIMA models.

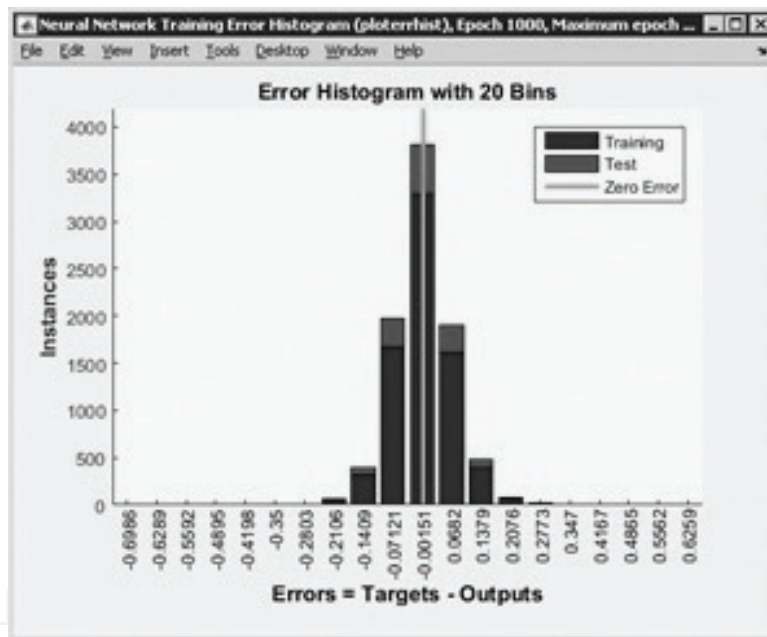


Figure 7. Errors histogram.

Performance/ method	LM	RB	GCS	AR	MA	ARMA	ARIMA
MSE	0.0064	0.0046	0.167	0.0091	0.0275	0.0316	0.0287
MAPE	4.26	4.21	6.21	7.29	24.45	29.05	24.97
Errors distribution	$-0.3$ to $0.12$	$-0.13$ to $0.12$	$-0.18$ to $0.22$	$-1.24$ to $1.16$	$-1.36$ to $1.44$	$-1.11$ to $0.99$	$-1.14$ to $0.66$

Table 3. Autoregressive neural networks versus stochastic methods.

## 4. Forecasting the electricity generation

In this section we will analyze stochastic methods based on ARMA and ARIMA models compared with feed-forward artificial neural networks for small wind turbines and photovoltaic panels generation in case of short-term forecasting.

### 4.1. Photovoltaic panels generation forecasting

In order to forecast the electricity produced by photovoltaic panels, we used input data from one PV power plant of 7.6 kW located in Giurgiu City, Romania, installed at a prosumer side on his building facade. PV generates electricity for the consumption of the prosumer, and when the consumption is lower than the PV output, the electricity is sent to the grid. For our experiments, data was recorded at 10 minutes interval, from January to December 2015 and includes the following attributes: ambient temperature, humidity, solar radiation, wind direction, wind speed and PV output (generated power), having more than 50,000 records.

First, we applied the ARIMA models and we calculated the error distribution, MSE, MAPE and R correlation coefficient (**Table 4**).

From our observations, the correlation coefficient indicates a strong relationship between solar radiation and the PV power forecast. This close dependence showed that regression models are appropriate for this time series. The best results were obtained with ARIMA model where the accuracy is 96.5% that indicates that the model can be used in PV panels generation forecast.

We consider a second method based on feed-forward neural networks. Therefore, we trained and validated a set of ANN in Matlab using Levenberg-Marquardt (LM), Bayesian Regularization (BR) and Scaled Conjugate Gradient (SCG).

For ANN architecture, we analyzed various settings regarding the number of neurons per layer, number of hidden layers and training algorithms. After several tests, we chose the following architecture: the input layer with 5 neurons (ambient temperature, humidity, solar irradiation, wind direction and wind speed) 60 neurons on a hidden layer and a single output (the energy produced).

Models	MSE	R	MAPE [%]
AR(1)	0.006505	0.989558	3.716167
AR(2)	0.013464	0.978219	5.008174
MA(1)	0.094253	0.948153	59.89228
MA(2)	0.102734	0.923199	63.59939
ARMA(1,1)	0.006478	0.98958	3.803265
ARMA(2,2)	0.013297	0.978492	5.439663
ARIMA(1,1,1)	0.006372	0.989939	3.633844
ARIMA(2,1,2)	0.006435	0.989787	3.566530

**Table 4.** Stochastic models results in case of PV generation forecasting.

For training, validating and testing: we have allocated 70% of the records for the training process, 15% for the validation process and 15% for the testing process. For training errors, we used the mean square error (MSE) by applying an error normalization process by configuring the normalization parameter to “standard”. Thus, output parameter values were standardized, ranging from  $[-1, 1]$ .

Taking into account the seasonal variations of the influence factors in Romania, we built artificial neural networks based on the three algorithms for each month and we compared the results in **Table 5**.

Comparing the ARIMA and ANN results, we consider that the most efficient approach is to use ANN on monthly data sets, which leads to excellent accuracy for every analyzed month. We also found that in almost 70% of cases, BR algorithm has a better generalization than LM or SCG algorithms. In 30% of cases, the highest accuracy was obtained with LM algorithm.

#### 4.2. Wind turbines generation forecasting

For forecasting simulations, we used recorded data for one small wind turbines (micro-generators) with a total power of 5 kW and another wind turbine of 10 kW, belonging to two consumers-producers (prosumers) located in two different areas of County of Tulcea, Romania. The two prosumers uses the energy produced by wind turbines for pumping water. For each turbine, the data set contains hourly data recorded from January 2013 to December 2014 with the following attributes: ambient temperature, wind direction, wind speed, atmospheric pressure and humidity. For each wind turbine, the measuring devices provides different values such as average, maximum and minimum wind speeds. Starting from the input data,

Period	MSE			Coefficient R		
	LM	BR	SCG	LM	BR	SCG
January	0.0818	0.0872	0.1132	0.9994	0.9993	0.9992
February	0.0387	0.0317	0.0679	0.9994	0.9993	0.9990
March	0.0617	0.0570	0.1174	0.9985	0.9987	0.9978
April	0.0201	0.0191	0.0319	0.9990	0.9989	0.9984
May	0.0539	0.0505	0.0640	0.9980	0.9981	0.9975
June	0.0705	0.0675	0.0865	0.9991	0.9992	0.9989
July	0.0439	0.0474	0.0577	0.9967	0.9969	0.9962
August	0.0870	0.0658	0.1019	0.9991	0.9993	0.9989
September	0.0348	0.0308	0.0512	0.9997	0.9997	0.9996
October	0.0601	0.0628	0.0877	0.9996	0.9997	0.9996
November	0.0369	0.0295	0.0650	0.9999	0.9999	0.9999
December	0.1002	0.0839	0.1091	0.9997	0.9997	0.9997

**Table 5.** Performance of ANN for solar generation forecasting in case of the PV panel.

we have developed several forecast scenarios by applying ARIMA models. The data recorded at the turbines' locations were analyzed using ARIMA Modeling and Forecasting time series in SAS Guide Enterprise 7.1. Applying the autoregressive model of the first AR (1), we obtained an extremely high average error (MAPE) of 86.5% for the wind turbine of 10 kW, which means that the accuracy of the model is only 13.5%. We tested also with first-order moving average (ARMA(1,1)) for both turbines and obtained the following results:

- for 5 kW turbine: AR(1) MA(1) model with MSE = 0.047 and MAPE = 93.6 and
- for 10 kW turbine: AR(1) MA(1) model with MSE = 0.021 and MAPE = 50.7.

Although the moving average improves the results (especially for the 10 kW wind turbine) as a consequence of the wind's unpredictable nature, because of the very low accuracy, the ARIMA models cannot be used for forecasting wind turbines' generation.

As a consequence, we approached the feed-forward neural networks. We used the three algorithms available in Matlab for feed-forward ANN: Levenberg-Marquardt (LM), Bayesian Regularization (BR) and Scaled Conjugate Gradient (SCG).

Considering seasonal characteristics of wind generation in Romania, where during spring and autumn the weather is windy, we split the data set into four seasons and trained and tested a dedicated ANN for each data set.

The ANN architecture was as follows: 5 neurons for the input layer (wind speed, wind direction, atmospheric pressure, temperature and humidity), 50 neurons for the hidden layer and one output (generated energy). The data set was randomly divided as follows: for training 70% of the records, for testing 15% and the remaining 15% for validation. The results obtained from each network testing and validation are synthesized in **Table 6**, which shows that the LM algorithm obtained the best correlation coefficient R (0.96) for spring and autumn data sets.

The results are good for all algorithms, analyzing the errors distribution we observed that most of them are between -0.1 and +0.1, which can be considered acceptable for the 5 kW turbine.

We experiment the ANN training and validation for the second turbine of 10 kW registering similar results, so we can conclude that the networks are efficient for generation forecasting in case of small wind turbines.

Season	MSE			Coefficient R		
	LM	BR	SCG	LM	BR	SCG
Spring	0.0372	0.0387	0.0489	0.96	0.96	0.95
Summer	0.0578	0.0603	0.0610	0.95	0.95	0.94
Autumn	0.0553	0.0610	0.0845	0.96	0.94	0.93
Winter	0.0604	0.0623	0.0636	0.95	0.95	0.94

**Table 6.** Performance of ANN for wind generation forecasting in case of 5 kW turbine.



## 5. Informatics solutions for electricity consumption and generation

In order to increase the consumer awareness toward energy efficiency, new informatics solutions must be developed and offered by the electricity supplier. An informatics solution for demand management must fulfill the following requirements:

- describing and modeling consumer's electrical appliances;
- real time consumption monitoring;
- monitor and forecast generation in case of prosumers;
- scheduling electrical appliances;
- optimizing the consumption;
- offering advanced analysis for consumption and micro-generation and
- monitoring the costs of electricity consumption according to advanced tariffs systems.

Our proposed informatics solution is part of a research project and it is addressed mainly to household consumers, but it also contains a management consumption module for electricity supplier. The informatics solution contains the following modules: data acquisition from smart metering and smart appliances, models for consumption management and user-friendly interfaces.

### 5.1. Data management

Data acquisition module extract data from heterogeneous sources such as smart meters and appliances in *.csv* or *.raw* format, micro-generation equipment (small photovoltaic panels, wind turbines and electrical vehicles), manual reading done by electricity supplier' employees or via web interfaces. Data are loaded first into local databases (concentrators) via Wi-Fi or RF. From local concentrators, data are synchronized periodically and loaded into a central data stage for proper cleansing and validation. We also designed an extract, transform and load (ETL) patterns for different types of SM and also procedures for extracting data from heterogeneous appliances. After ETL process completes, data are loaded into a central relational database running Oracle Database 12c or MySQL for operational management and then into a data warehouse for advanced analytics.

### 5.2. Models

The models' layer integrates previously described methods for determine the consumers' profiles based on SOM, methods for short-term consumption forecasting based on autoregressive neural networks and in case of prosumers, short-term generation forecasting from small wind turbines and PV panels. Also the layer includes an optimization model based on two optimization functions as described in detailed in Ref. [35].

### 5.3. Interfaces

Our proposed solutions are integrated into a web-based application using business intelligence components that allow both electricity supplier and prosumers to interact with the proposed models. We set up a business intelligence server installed in a cloud computing using Software as a Service (SaaS) that offers access for prosumers/electricity supplier to services via internet connection for advanced analytics. The application includes the following consumers' facilities (**Figure 8**):

- monitoring their own consumption/generation;
- customize appliances and schedule their consumption;
- optimize consumption/generation based on advanced tariffs systems and
- real time electricity bills.

For electricity suppliers, the application will include an advanced analytics interface (**Figure 9**) that allow them to:

- set up advanced tariffs systems;
- analyze aggregate consumption;
- determine consumers' profiles and
- forecast electricity consumption.

The informatics solution is developed on a scalable platform, using Java with Application Development Framework and Oracle Database 12c that enables Cloud management and services. Thus, the solution can be adopted by the energy suppliers without expensive investments in infrastructure. Also, it offers an user-friendly interfaces that can be easily understand and managed by end-users on personal computers and mobile devices.

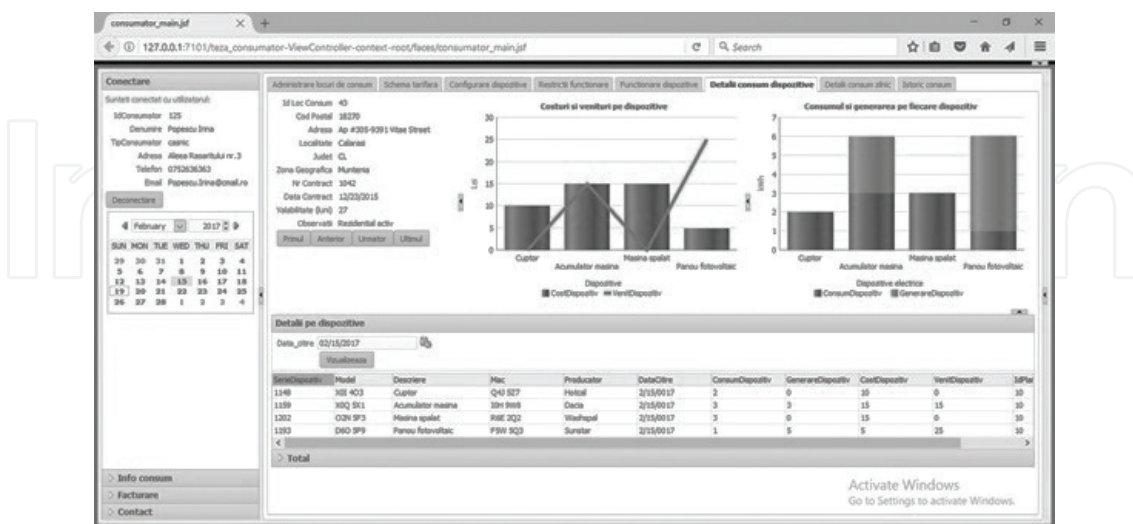


Figure 8. Monitoring consumption and generation in case of prosumers.

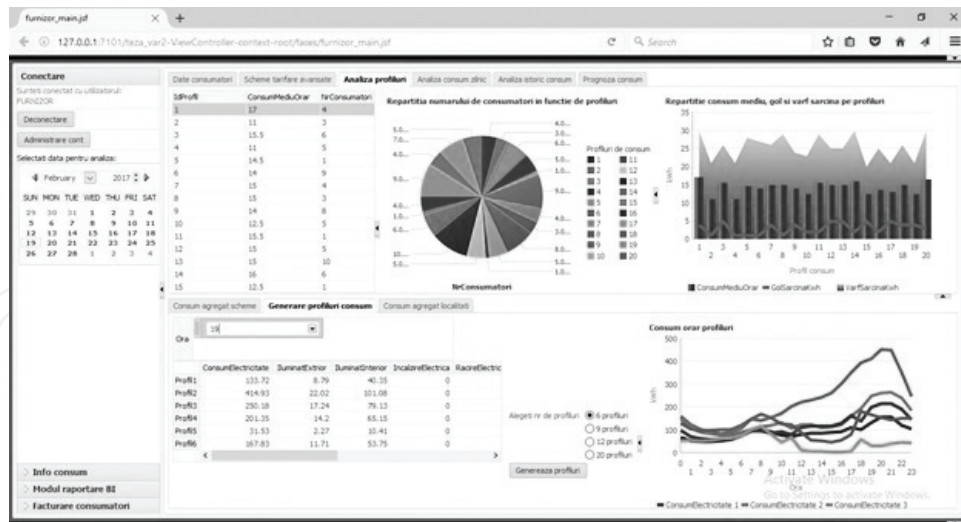


Figure 9. Determine consumers' profiles through web interfaces in case of electricity suppliers.

## 6. Conclusion and future work

Householders' profiles and patterns will allow electricity suppliers to understand consumers' behavior, set up more flexible and customized electricity prices to avoid peak consumption. On the other hand, prosumers will benefit from the forecasting solutions that will estimate wind and PV generation, therefore they will schedule their appliances according to electricity prices and their generation resources.

From our experiments, we consider artificial neural networks a good solution for determining the consumption profiles, for short-term load forecasting on each profile and also for short-term micro-generation forecasting.

A disadvantage of neural networks is that the most appropriate solution in a particular case is found by successive attempts on the number of hidden layers and the number of neurons on each layer, so in the case of another set of data from another geographic area with different characteristics regarding consumption or meteorological conditions that affect the wind or solar generation, it is necessary to re-configure the ANN parameters.

An advantage of artificial neural networks in case of consumption and generation forecasting is that they perform predictions with very good results in a very short time, which makes ANN particularly useful for real time short-term forecasting.

## Acknowledgements

This paper presents some results of the research project: Informatics solutions for optimizing the operation of photovoltaic power plants (OPTIM-PV), project code: PN-III-P2-2.1-PTE-2016-0032, 4PTE/06/10/2016, PNIII - PTE 2016.

## Author details

Adela Bâra\* and Simona Vasilica Oprea

\*Address all correspondence to: [bara.adela@ie.ase.ro](mailto:bara.adela@ie.ase.ro)

The Bucharest University of Economic Studies, Romania

## References

- [1] European Economic and Social Committee. For coordinated European measures to prevent and combat energy poverty (own-initiative opinion) (2013/C/341/05). Official Journal of the European Union. 2013
- [2] Directive 2012/27/EU of the European Parliament and of the Council (2012/27/EU). Official Journal of the European Union. 2012
- [3] ANRE Decision 143/23.01.2013. The procedure for elaboration and application of specific consumption profiles in the license area of CEZ Distribuție. 2013
- [4] Chan A, Moreno JG, Hughes M. Electricity Price Signals and Demand Response [Internet]. 2014. Available from: <https://www.gov.uk/government/collections/household-electricity-survey>. [Accessed: 24-04-2017]
- [5] Sanquista TF, Orrb H, Shuic B, Bittnerd AC. Lifestyle factors in U.S. residential electricity consumption. *Energy Policy*. 2012;**42**:354-364
- [6] Hayn M, Bertsch V, Fichtner W. Electricity load profiles in Europe: The importance of household segmentation. *Energy Research & Social Science*. 2014;**3**:30-45
- [7] Ardakanian O, Koochakzadeh N, Singh RP. Computing electricity consumption profiles from household smart meter data. In: 3rd Workshop on Energy Data Management (EnDM 2014); 28 March 2014; Athens, Greece; 2014
- [8] Oprea SV, Bâra A. Electricity load profile calculation using self-organizing maps. In: 20th International Conference on System Theory, Control and Computing (ICSTCC), (Joint conference SINTES 20, SACCS 16, SIMSIS 20); 13-15 October 2016; Sinaia, Romania; 2016
- [9] ELEXON. Load Profiles and their use in Electricity Settlement, Guidance [Internet]. 2013. Available from: [https://www.elexon.co.uk/wp-content/uploads/2013/11/load\\_profiles\\_v2.0\\_cgi.pdf](https://www.elexon.co.uk/wp-content/uploads/2013/11/load_profiles_v2.0_cgi.pdf) [Accessed: 2017]
- [10] Ardakanian O, Koochakzadeh N, Singh RP, Golab L, Keshav S. Computing electricity consumption profiles from household smart meter data. In: Workshop Proceedings of the EDBT/ICDT 2014 Joint Conference. Greece: Athens. p. 2014
- [11] Lo KL, Zakaria Z, Sohod MH. Determination of consumers' load profiles based on two-stage fuzzy C-means. In: Proceedings of the 5th WSEAS Int. Conf. on Power Systems and Electromagnetic Compatibility; 2005; Greece. 2005

- [12] Yang HT, Huang CM, Huang CL. Identification of ARMAX model for short term load forecasting: An evolutionary programming approach. *IEEE Transactions on Power Apparatus and Systems*. 1996;**11**(1):403-408
- [13] Haidda T, Muto S. Regression based peak load forecasting using a transformation technique. *IEEE Transactions on Power Apparatus and Systems*. 1994;**9**(4):1788-1794
- [14] Charytoniuk W, Chen MS, Van Olinda P. Nonparametric regression based short-term load forecasting. *IEEE Transactions on Power Apparatus and Systems*. 1998;**13**(3):725-730
- [15] Ko CN, Lee CM. Short-term load forecasting using SVR (support vector regression)-based radial basis function neural network with dual extended Kalman filter. *Energy*. 2013;**49**:413-422
- [16] Sevlian R, Rajagopal R. Short Term Electricity Load Forecasting on Varying Levels of Aggregation [Internet]. 2014. Available from: <https://arxiv.org/abs/1404.0058> [Accessed: 2017]
- [17] Aung Z, Toukhy M, Williams JR, Sanchez A, Herrero S. Towards accurate electricity load forecasting in smart grids. In: *The Fourth International Conference on Advances in Databases, Knowledge and Data Applications*; 2012
- [18] Engle RF, Mustafa C, Rice J. Modeling peak electricity demand. *Journal of Forecasting*. 1992;**11**:241-251
- [19] Hyde O, Hodnett PF. An adaptable automated procedure for short-term electricity load forecasting. *IEEE Transactions on Power Systems*. 1997;**12**:84-93
- [20] Ruzic S, Vuckovic A, Nikolic N. Weather sensitive method for short-term load forecasting in electric power utility of Serbia. *IEEE Transactions on Power Systems*. 2003;**18**:1581-1586
- [21] Haida T, Muto S. Regression based peak load forecasting using a transformation technique. *IEEE Transactions on Power Systems*. 1994;**9**:1788-1794
- [22] Charytoniuk W, Chen MS, Van Olinda P. Nonparametric regression based short-term load forecasting. *IEEE Transactions on Power Systems*. 1998;**13**:725-730
- [23] Mohamed Z, Bodger PS. Forecasting Electricity Consumption: A Comparison of Models for New Zealand [Internet]. Available from: [http://ir.canterbury.ac.nz/bitstream/handle/10092/821/12593635\\_C47.pdf?sequence=1](http://ir.canterbury.ac.nz/bitstream/handle/10092/821/12593635_C47.pdf?sequence=1) [Accessed: 2017]
- [24] Chujai P, Kerdprasop N, Kerdprasop K. Time series analysis of household electric consumption with ARIMA and ARMA models. In: *Proceedings of the International MultiConference of Engineers and Computer Scientists*; 13–15 March 2013; Hong Kong; 2013
- [25] Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitosa E. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews*. 2008;**12**(6):1725-1744

- [26] Yu Chen KJ. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. *Applied Energy*. 2014;**113**:690-705
- [27] Chitsaz H, Amjady N, Zareipour H. Wind power forecast using wavelet neural network trained by improved clonal selection algorithm. *Energy Conversion and Management*. 2015;**89**:588-598
- [28] Blonbou R. Very short-term wind power forecasting with neural networks and adaptive Bayesian learning. *Renewable Energy*. 2011;**36**(3):1118-1124
- [29] Yeh WC, Yeh YM, Chang PC, Ke YC, Chung V. Forecasting wind power in the Mai Liao Wind Farm based on the multi-layer perceptron artificial neural network model with improved simplified swarm optimization. *International Journal of Electrical Power & Energy Systems*. 2014;**55**:741-748
- [30] Ren C, An N, Wang J, Li L, Hu B, Shang D. Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting. *Knowledge-Based Systems*. 2014;**56**:226-239
- [31] Huang R, Huang T, Gadh R, Li N. Solar generation prediction using the ARMA model in a laboratory-level micro-grid. In: *IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*; 5–9 November 2012. Taiwan: Tainan. p. 2012
- [32] Zhi C, Lei W, Yong F. Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization. *IEEE Transaction on Smart Grid*. 2012;**3**:1949-3053
- [33] Yang X, Guo S, Yang HT. The establishment of energy consumption optimization model based on genetic algorithm. In: *IEEE Int. Conf. on Automation and Logistics*; 1–3 September 2008; Chindao, China. New York, NY, USA: IEEE. 1426-1431
- [34] Omari M, Abdelkarim H, Salem B. Optimization of energy consumption based on genetic algorithms optimization and fuzzy classification. In: *2nd World Symposium on Web Applications and Networking*; 21-23 March 2015; Sousse, Tunisia. New York, NY, USA. IEEE. 2015
- [35] Oprea S, Bara A, Cebeci ME, Tör OB. Promoting peak shaving while minimizing electricity consumption payment for residential consumers by using storage devices. *Turkish Journal of Electrical Engineering & Computer Sciences*. 2017;**25**(5):3725-3737. DOI: 10.3906/elk-1606-152
- [36] Captterra. Top Utility Billing Software Products [Internet]. Available from: <http://www.captterra.com/utility-billing-software> [Accessed: 24-04-2017]
- [37] Campos MM, Milenova BL. O-cluster: Scalable Clustering of Large High Dimensional Data Sets. Oracle Data Mining Technologies. Oracle Corporation; 2002

