ISSN 2348 – 7968

Single Home Electricity Power Consumption Forecast Using Neural Networks Model

Naser Farag Abed[1] and Milan M.Milosavljevic [1,2]

[1] Singidunum University
Belgrade, 11000, Serbia

[2] School of Electrical Engineering,
Belgrade University, Belgrade, 11000, Serbia

Abstract

This work analyzes an electricity power consumption forecast for a single home using a multi-layer perceptron (MLP) artificial neural network (ANN). The predictor composes of parallel banks of MLP (PBMLP) for each hour within a day. Training PBMLP is performed separately for each day of the week using appropriate training sets of past electricity power consumption. The performance of the predictor was evaluated using real data which represents power consumption per minute measured over almost 4 years for a single home near Paris, France (approximately 2 million data points). Experiments show that proposed system for predicting power consumption one day ahead gave mean absolute value of relative percentage error (MARPE) lower more than 25% comparing to persistent prediction, which is to our knowledge the best reported result up to now.

Keywords: *Energy consumption prediction; Artificial neural network*, *Back propagation algorithm, Mean absolute value of relative percentage error, Persistent prediction.*

1. Introduction

Electricity load forecasting has gained substantial importance nowadays in the modern electrical power management systems with elements of smart greed technology. A reliable forecast of electrical power consumption represents a starting point in policy development and improvement of energy production and distribution. At the level of individual households, the ability to accurately predict consumption of electricity power significantly reduces prices by appropriate systems for energy storage. Therefore, the energy efficient power networks of the future will require entirely new ways of forecasting demand on the scale of individual households.

It is interesting to note that aggregation of consumers at higher levels, for example one block, part of the city or entire regions, make forecasting problem easier due to the effects of the well-known central limit theorem. Basic probabilistic mechanisms governing consumption phenomena become more predictable rendering efficient distribution less problematic.

In [3], the authors claimed the following: "For short forecasting horizons and high granularities of consumption data, persistence forecasts are known as hard to beat by the other methods." In this paper we show that this claim is not generally true. Carefully designed Neural Networks prediction system can be far efficient than persistence method even in the case of very irregular patterns of daily power consumptions.

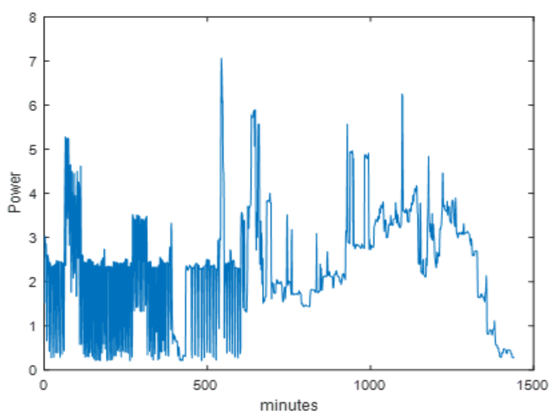
Our paper is organized as follows. In section 2 we describe two publicly available reference data sets, so called REDD and TUM data, along with the data chosen in our investigation, which we named PARIS data. The main reason for such approach is that there are published results regarding accuracy of forecasting energy power consumption based on REED and TUM data sets. In section 3 we present forecasting strategies and important notion of persistent prediction which can serve as the baseline forecast method, not only as a reference for accuracy but also as a measure of complexity of data at hand. Subsequently, in section 4 we describe our novel short term predictor based on neural networks, capable to predict electricity power consumption within forecasting horizon from one day to one week ahead, based on data of 7 to 30 past days. Section 5 contains experimental results obtained for PARIS data set, as well as comparison with REDD and TUMM data sets. In the conclusion we summarize the main results and give some short remarks about possible further extensions.

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 1, January 2016.
www.ijiset.com
ISSN 2348 – 7968

2. Data sets

There is not a lot of previous works on the subject of single-home electricity power consumption prediction. The Smart project at the University of Massachusetts, Amherst, has studied power consumption in several homes equipped with sensors for a period of 2 to 3 months [1].

In order to compare our results we chose two different data sets. The Reference Energy Disaggregation Data Set (REDD) is a public data set for energy disaggregation research [2]. This data set is provided by the Massachusetts Institute of Technology and contains power consumption measurements of 6 US households recorded for 18 days between April 2011 and June 2011 and sampled at intervals of 3 seconds. In this paper we will refer to this data set as REDD data set.



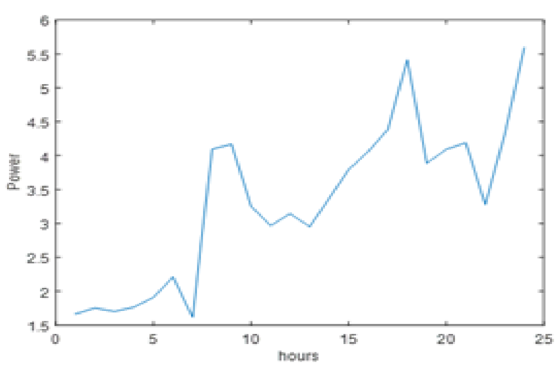


Fig.1 Electricity power consumption in the granularity of minute (upper figure) and hour (bottom figure). In the granularity of one minute and one hour, one day is represented by 1440 data points, and 24 data points, respectively.

In the TUM Home Experiment, a single household in Germany, in the state of Bavaria was equipped with a sensors measuring power from several appliances [3]. The data was collected from February 4th 2013 to October 31st 2013. Henceforth, we will refer to this data set as the TUM data set.

The data set which we named PARIS data was obtained from the UCI data set repository [4]. It represents electricity power consumption per minute measured between December 16th 2006 and November 26th 2010 (47 months) for a single home near Paris, France, containing over 2 million data points. The data set contains nearly 1.25% of missing values. We replace missing values by simple linear interpolation filtering.

All data can be analyzed at different level of time granularity, i.e., sampling frequencies, from one minute to one day. In order to compare our results with the related work in the area, we chose time granularity of one hour. That means that one day of electricity power consumption is represented by 24 data points. Thus, the whole data set is represented by 34440 sampling points. In Fig.1 electricity power consumption is shown in the granularity of one minute (a) and one hour (b) during one day. Change in granularity is implemented by appropriate moving average filtering.

3. Forecasting strategies

In our experiments we use two different strategies to sample the training, validation and test data. In the following we explain these individual strategies.

3.1 Day Type Strategy

The day type approach uses cross-sectional data, joining each day of the week of consecutive weeks into separate data sets. The approach is illustrated in Figure 2. The training, validation and the test data set are then sampled from the individual data sets. An example of such an approach is to join for instance the Saturdays of consecutive weeks. Fig.3 shows tree dimensional representation of whole PARIS data set, while Fig.4 shows cross-section data at 12:00 pm. It is obvious from Fig.4 that data exhibits very clear seasonal behavior, but also strong random component, making forecast across the days very hard task.

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 1, January 2016.
www.ijiset.com
ISSN 2348 – 7968

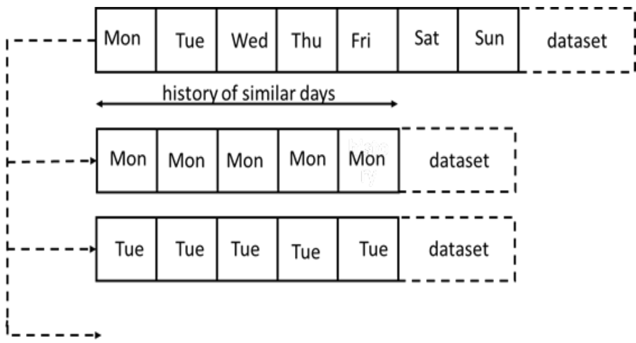


Fig.2 Day Type Strategy

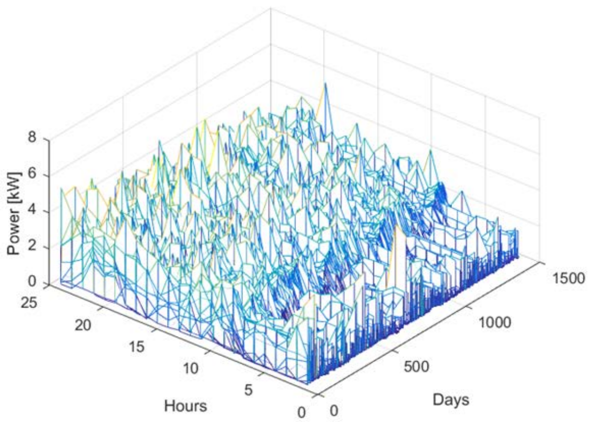


Fig.3 Tree dimensional (Days,Hours,Power) representation of PARIS data

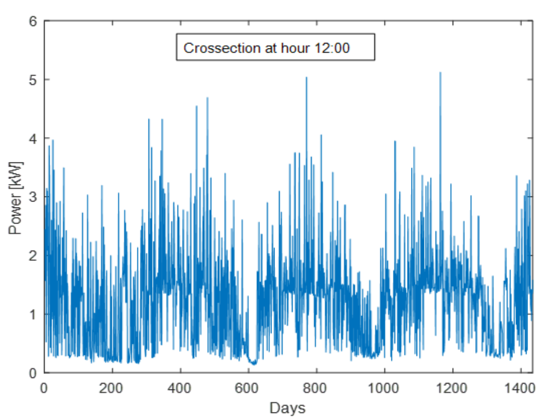


Fig.4 Cross-section data at 12:00 pm

As an illustration of day type strategy, Fig.5 shows tree dimensional (Days,Hours,Power) representation of PARIS data obtained joining Saturdays of consecutive weeks. Fig.6 shows an example of cross-sectioning this data at 12:00 pm.

Day type strategy is convenient for predicting electricity power consumption of chosen day of the week, one or more weeks ahead. Underlying assumption is that every weekday has its own pattern of power consumption and time dynamics. Moreover, it enables synthesis of bank of predictors for each hour cross-section, from 1 to 24 hours.

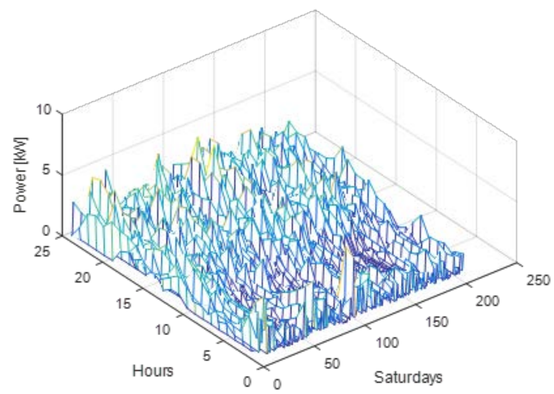


Fig.5 Tree dimensional (Days,Hours,Power) representation of PARIS data obtained joining Saturdays of consecutive weeks.

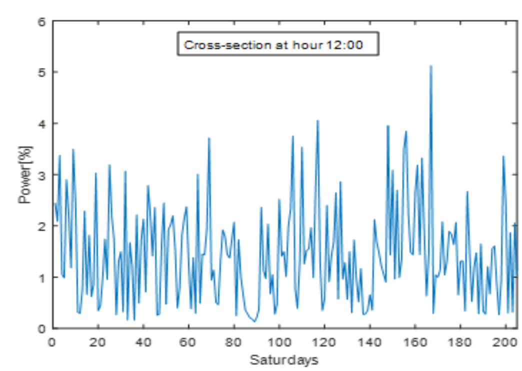


Fig.6 Cross-section data obtained joining consecutive Saturdays at 12:00 pm.

3.2 Jumping Window Strategy

In jumping window strategy we obtain training, validation and testing data sets moving window forward on the data set with steps equal to jumping length. In such a way we prepare data for synthesis of predictors for specific day, based on past data within given window. For example, if we are interested in predicting consumption on Sunday, based on the past three days up to Saturday, then we collect training, validation and testing data by jumping window strategy with window length equal to 3+1 days and jumping length equal to 7 days, on condition that the front of starting window is positioned at Saturday.

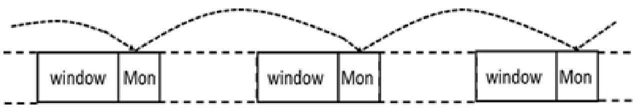


Fig.7 Jumping window strategy. Data sets for training, validation and testing are obtained moving window forward on the data set with steps equal to jumping length.

The approach is illustrated in Fig. 7. The main reason for using the jumping window approach is that we assume different and distinct behavior of daily consumption patterns for each day-to-day transition during a week. In our experiments we use this strategy for obtaining a set of predictors for one day ahead, for each day in the week, based on the past data specified by the window length.

3.3 Model Quality Measure

We need a statistical quality measure that can be used for comparison of the different forecasting methods and strategies. In this work we choose the Mean Absolute Relative Percentage Error (MARPE) as the standard accuracy error measure. Since it's a relative measure, it is very convenient to compare performance of different forecasting techniques on different data sets. MARPE is defined as the mean over the ratio of the absolute difference between the residual and the actual value in percent:

$$MARPE = \frac{1}{n}\sum_{t=1}^{t=n}\left|\frac{x_t-\widehat{x_t}}{x_t}\right| \cdot 100\ [\%]\,, \qquad (1)$$

where x_t is the actual value and $\widehat{x_t}$ is the forecast value. For example, with an actual electricity power consumption of 100 Watt and a corresponding forecasted consumption of 125 Watt, the MARPE would be 25%, because the difference between actual and predicted consumption is 25% of the actual one.

3.4 Persistence forecast

As a benchmark for the analyzed forecasting methods we adopt the so called persistence method. The persistence method assumes that the conditions at the time of the forecast will not change, i.e., all forecasts are equal to the last observation. We will refer to this method as PERSIST. Corresponding MARPE has a form

$$MARPE_{per} = \frac{1}{n-1}\sum_{t=2}^{t=n}\left|\frac{x_t-x_{t-1}}{x_t}\right| \cdot 100\ [\%]. \qquad (2)$$

We extend the notion of persistence forecast to the broader class of simple predictors, which gives forecast according to formula

$$\hat{x}_{t+1} = \frac{1}{N}\sum_{n=t-N+1}^{t} x_n, \qquad (3)$$

i.e., by simple averaging past observations within window of specified length N. We refer to this method as MA_PERSIST (moving average persistence).

4. Short term neural network forecasting model

The artificial neural network (ANN) is a structure of computing elements, which allows the easy integration of a priori knowledge, as well as the ability to learn from data to represent complex dynamic system [6]. In this respect neural networks are very convenient method for solving forecasting problems in complex domains, such as single home electricity power consumption. The functionality of ANN varies considerably depending on its architecture and the criterion function used in learning procedure. In order to avoid over-fitting and poor behavior on unseen data, we adjust complexity of mapping embodied in a neural network with the complexity of training sets.

Analyzing correlation between hourly consumption during a day (see Fig.8) we can conclude that dominant influence on a value at given hour, comes from one hour before and after given hour.

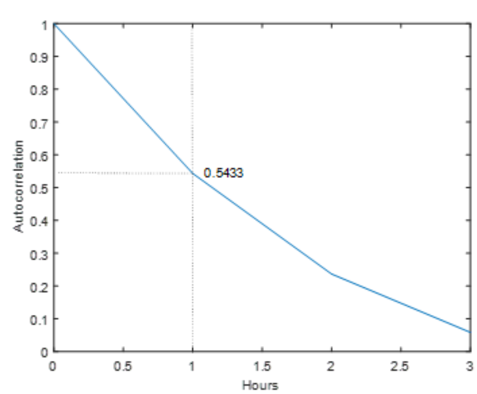


Fig.8 Mean value of autocorrelation function between hourly consumption during a day for data set PARIS.

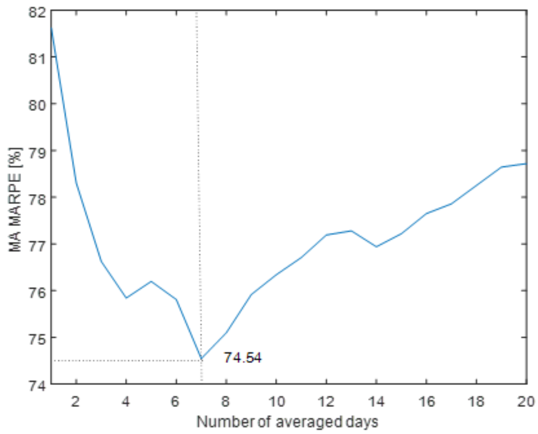


Fig.9 MARPE for persistence forecasting for different value of N, see equation (3). It is clear that on PARIS data, there is no need to go past beyond 7 days from the current day if our forecasting method is based on moving average persistence.

We further investigate accuracy of the persistence forecasting calculated on PARIS data. In Table 1 we summarize obtained results. These results clearly suggest that moving average strategy has significantly lower forecast error. Therefore, it can be incorporated in ANN architecture as useful building block.

Table.1 MARPE for various type of persistence strategy for PARIS data

Persistence forecast		
Type of strategy	Horizon	MARPE [%]
Day type strategy	1 week	83.96
Jumping window strategy	24 hours	81.68
Moving average	24 hours	74.54

Summarizing all obtained experimental evidence, we propose very specific ANN architecture, shown in Fig.10. Inputs to ANN are described as follows:

$$X(h,t) = \begin{bmatrix} x(h,t-1) \\ x(h,t-2) \\ x(h,t-3) \end{bmatrix}, \qquad (4)$$

$$MA_i(h,t) = \begin{bmatrix} ma_i(h,t-1) \\ ma_i(h,t-2) \\ ma_i(h,t-3) \end{bmatrix}, \qquad (5)$$

$$ma_i(h,t) = \frac{1}{N_i}\sum_{n=1}^{N_i} x(h,t-n) \qquad (6)$$

$$\Psi(h,t) = \begin{bmatrix} X(h,t) \\ X(h,t) - MA_1(h,t) \\ X(h,t) - MA_2(h,t) \end{bmatrix}, \qquad (7)$$

where argument h is hour and t denotes day index, i.e., $h \in \{1,2,\cdots,24\}$, $t=1,2,\ldots$.The output is given by

$$\Delta x(t) = x(t) - x(t-1). \qquad (8)$$

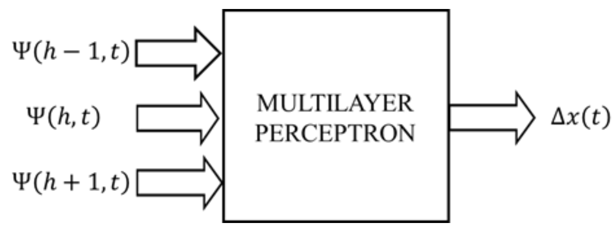


Fig10. Proposed MLP ANN predictor, as one building block for forecasting electricity consumption for a given hour h during the day. 24 such MLPs, each for one hour, make complete parallel bank of MLP ANN (PBMLP) predictors for the entire day.

Note that the proposed architecture has 27 inputs, where each block of inputs $\Psi(h,t)$has dimension 9. The output of the proposed forecasting system gives estimate of $\Delta\hat{x}(t)$, so that according to (8), desired forecast $\hat{x}(t)$ is obtained by

$$\hat{x}(t) = \Delta\hat{x}(t) + x(t-1). \qquad (9)$$

5. Experimental results

The purpose of our experiment is to evaluate proposed forecasting system on PARIS data. MLP ANN consists of two hidden layer of sizes 40 and 4. All neurons except the output one have sigmoidal activation function. Training, validation and test sets are 70%, 15% and 15% of whole available data, respectively. System is trained by conjugate gradient variant of back propagation learning procedure based on 5 fold cross validation scheme. Overfitting is fixed by regularization and early stopping method.

The first group of experiments refers to forecasting electricity power consumption one week ahead by day type strategy. The value of MARPE in terms of day of a week is shown in Fig.11. Notice that the hardest day for weekly prediction is Saturday, and easiest one is Wednesday. In Fig.12 MARPE distribution over 24 hours is shown. It is evident that a period of day between 5 am to 9 am is easier

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 3 Issue 1, January 2016.
www.ijiset.com
ISSN 2348 – 7968

to predict then other parts of day. Mean value of MARPE for all days in week is 71.50%, which is 14.84% lower than the corresponding persistent forecasting MARPE (see, Table.1). In Fig.13 we show one cross section of true and forecasted consumption one week ahead for Fridays at 12:00 pm.

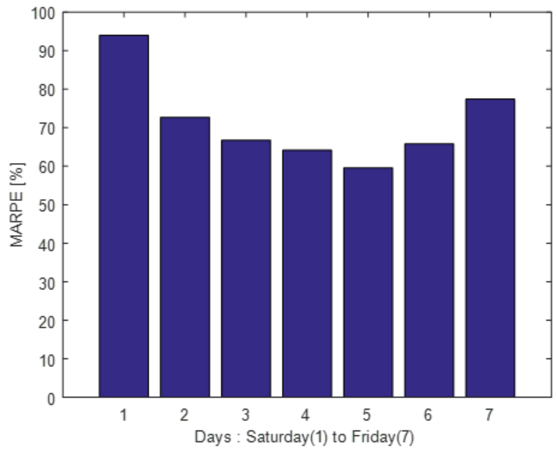


Fig.11 Distribution of MARPE over days of a week in the case of forecasting one week ahead.

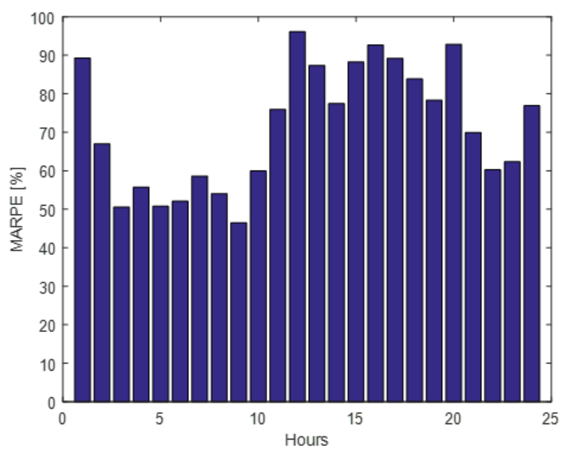


Fig.12 Distribution of MARPE over 24 hours a day in the case of forecasting one week ahead.

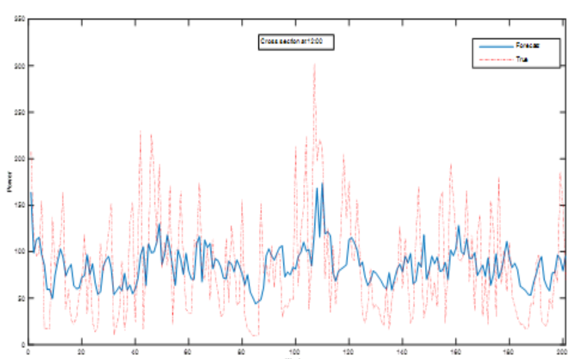


Fig.13 True and forecasted value of electricity power consumption for Fridays at 12:00 pm obtained by day type strategy and proposed BMLP.

The second group of experiments refers to forecasting electricity power consumption one day ahead by the jumping window strategy. The value of MARPE in terms of day of a week is shown in Fig.14. Notice that the hardest day for weekly prediction is Saturday, and the easiest one is Wednesday. In Fig.15 MARPE distribution over 24 hours is shown. It is evident that the period between 5 am to 9 am is easier to predict than other parts of a day. Mean value of MARPE for all days in a week is 61.16%, which is 25.12% lower than the corresponding persistent forecasting MARPE (see Table.1). In Fig.16 we show one cross section of true and forecasted consumption one day ahead for Fridays at 12:00 am, obtained by jumping window strategy.

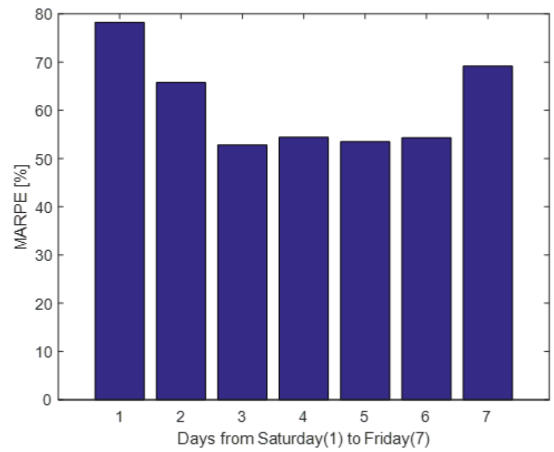


Fig.14 Distribution of MARPE over days of a week in the case of forecasting one day ahead obtained by jumping window strategy.

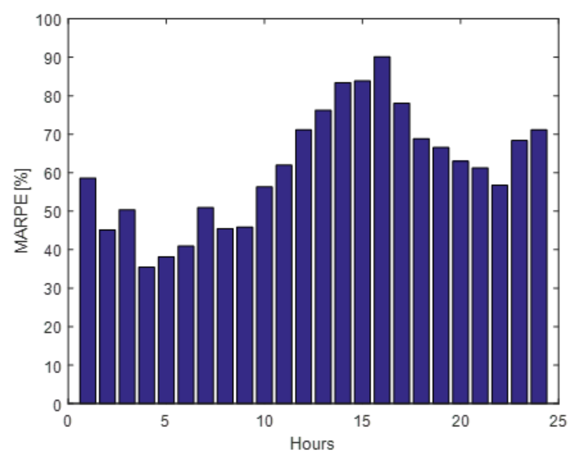


Fig.15 Distribution of MARPE over 24 hours in the case of forecasting one day ahead obtained by jumping window strategy.

Finally we compare our results with the results reported in [3] for TUM and REDD data. We summarize all relevant findings in Table. 2. Notice that in 3 out of 4 data sets, persistent forecast is better than ANN, which supports the claim that for short forecasting horizons and high granularities of

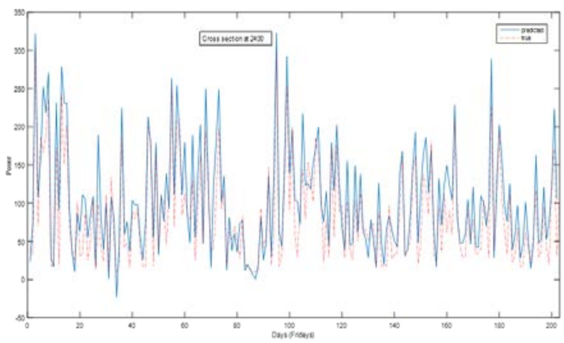


Fig.16 True and forecasted value of electricity power consumption for Fridays at 12:00 am obtained by jumping window strategy and proposed BMLP.

consumption data, persistence forecasts are hard to beat by the other methods [3]. In this context, our results show significant support for the opposite claim.

Table.2 Accuracy of the best ANN forecasting methods reported in [3]. The last row shows results of the proposed BMLP system tested on PARIS data.

Data	Persistent	ANN	Accuracy gain
REDD1	117	102	12.82
TUM1	29	50	-72.41
REDD2	48	65	-35.42
TUM2	25	31	-24.00
PARIS	**81.68**	**61.16**	**25.12**

6. Conclusions

We have evaluated a new short-term forecasting method for household electricity power consumption based on ANN. Test data represents power consumption of a single home near Paris, France, measured during almost 4 years and containing over 2 million data points. The proposed method is based on the aggregated power consumption and does not need any type of distributed sensors for measuring power consumption of individual appliances. We show that carefully designed ANN system can beat persistent forecast method, opposite to wide spread opinion. Further research can be application of proposed synthesis methodology to disaggregated electricity power consumption data.

References

[1] Mishra, A., Irwin, D., Shenoy, P., Kurose, J., & Zhu, T. (2012, May). SmartCharge: cutting the electricity bill in smart homes with energy storage. In Proceedings of the 3rd International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet (p. 29). ACM.

[2] J.ZicoKolter and Matthew J. Johnson. REDD: A public data set for energy disaggregation research. In proceedings of the SustKDD workshop on Data Mining Applications in Sustainability, 2011.

[3] Veit, A. et al. “Household Electricity Demand Forecasting – Benchmarking State-of-the-Art Methods.” Eprint arXiv:1040.0200 (2014).

[4]http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption, source: Georges HÃ©brail (georges.hebrail @edf.fr), Senior Researcher, EDF R&D, Clamart, France ,Alice BÃ©rard, TELECOM ParisTech Master of Engineering Internship at EDF R&D, Clamart, France

[5] Ljiljana Cander, Milan Milosavljević, Srdjan Stanković, Saša Tomašević, "Ionospheric forecasting technique by artificial neural network", Electronic Letters, 34(16):1573 -- 1574, 1998.

[6] S.S Stankovic, M.M Milosavljevic, Training of multiplayer perceptions by stochastic approximation, In Neural networks: concepts, applications, and implementations 4, 201-239, Prentice Hall, 1991.

Naser Farag Abed was born on 1972. He received the B.S. degree and M.Sc. degree in power Electrical Engineering from Brno University of Technology in 2006. He now is doing PhD study program in Advanced Information Safety System, field of natural and mathematical science, information and computing at Singidunum University, Belgrade, Serbia. His scientific research areas include Information Security, Machine Learning and Neural Networks.

Milan M.Milosavljevic received the B.Sc, M.Sc and Ph.D. degree in electrical engineering from the School of Electrical Engineering, Belgrade University, in 1976, 1979, and 1982, respectively. He has held a variety of positions in academia and scientific community in former Yugoslavia and nowadays Republic of Serbia. From 1982 to 2003 he was the Deputy Director for Research at the Institute for Applied Mathematics and Electronics, Belgrade, responsible for all scientific research projects of the Institute. Under his leadership the Institute for Applied Mathematics and Electronics became the leading scientific institutions in the former Yugoslavia. Since 2003 he has worked as a full professor of Electrical Engineering at the Department of Signals and Systems, Belgrade University, and at the same time at Singidunum University, where he held many important positions. During the school year 1987/1988 and 1997 he was a visiting professor at Cornell University, Ithaca, USA and the University of Paris XIII, Paris, France, respectively. Since 2005, he is external evaluator at the master's and doctoral degree program Information Security, University of Gjovik, Gjovik, Norway. His areas of expertise are pattern recognition, artificial intelligence, digital signal processing, machine learning, discrete optimization, cryptanalysis and information security. He has published over 400 scientific bibliographic references and was a mentor to over 30 doctoral dissertations.