Forecasting Hot Water Consumption in Dwellings Using Artifitial Neural Networks

Linas Gelažanskas
Department of Engineering
Lancaster University
Lancaster, UK
linas@gelazanskas.lt

Kelum A.A. Gamage
Department of Engineering
Lancaster University
Lancaster, UK
k.gamage@lancaster.ac.uk

Abstract— **The electricity grid is currently transforming and becoming more and more decentralised. Green energy generation has many incentives throughout the world thus small renewable generation units become popular. Intermittent generation units pose threat to system stability so new balancing techniques like Demand Side Management must be researched. Residential hot water heaters are perfect candidates to be used for shifting electricity consumption in time. This paper investigates the ability on Artificial Neural Networks to predict individual hot water heater energy demand profile. Data from about a hundred dwellings are analysed using autocorrelation technique. The most appropriate lags were chosen and different Neural Network model topologies were tested and compared. The results are positive and show that water heaters have could potentially shift electric energy.**

Keywords—Hot Water Consumption; Forecasting; Artifitial Neural Networks; Smart Grid, Demand Side Management

I. Introduction

Electric power grid is the largest device ever made by a human being, where it plays an enormous role in every person's life. Traditional power system was designed to be centralised and consists of large generation units generating electricity and many consumers using it. This concept has been gradually changing over the years, particularly with distributed energy resource systems.

The grid is currently transforming into so called the Smart Grid. Consumers are becoming prosumers, meaning they not only consume, but also generate electricity. Electricity now flows both ways (from the grid to the consumer or from consumer to the grid) and the grid could also contain smaller generating units. In Europe, there have been many incentives for wind and solar power pants to be built [1]. But aside of all the environmental advantages of green energy, there are considerable drawbacks of renewable energy generators. The biggest one being the fact that renewable energy generation is intermittent (or forecast) or very hard to control compared to conventional power generation.

This intermittent nature greatly increases the complexity of the supply demand balancing problem. Increasing number of renewables poses real threat to system resilience and affect system stability [2]. To counterweight this issue either more spinning reserve power plants should be built or other alternatives researched and found. This has invoked great attention from researchers around the world.

The authors of this paper are looking into new ways to balance electric energy while keeping current infrastructure and thus minimising upfront cost [3]. In fact the balance should be reached not only by increasing or decreasing generation, but also by changing how people consume electricity. This modification of consumer demand profile through various incentives or education is generally called Demand Side Management (DSM). This term dates back to early 1980s [4], but the attention on it is increased recently due to several reasons. Electricity market deregulation was the first step towards enabling DSM, but the recent attention is mainly due to the development of Smart Grid. The addition of additional communication layer on top of existing power grid allows more precise management and control of electricity. It is expected that the future grid will be able to monitor and control appliances in residential houses. Thus residential customers are allowed to contribute in system balancing act.

There are several kinds of DSM programs. The most popular ones can be grouped into Price Based programs (PBP) and Incentive Based Programs (IBP) [5]. Programs can differ in many aspects. As the group names suggest, some programs are based on Real Time Price (RTP) [6], some give incentives like bill discounts. Some require Direct Load Control (DLC) while others leave customer to make the final decision when to curtail. There are also differences in how many appliances can participate in every residence [7]. All in all, every incentive in DSM program boils down to decreased energy bill.

There are devices in every house that are flexible in terms of when the electricity could be used, i.e. energy use can be shifted in time. One of the biggest electricity users in dwelling is an electric hot water heater [8]. Also due to its large inertia, it can be turned on at different times without a notable change in temperature [9]. This makes it a perfect device candidate that could be used for helping system in reaching perfect balance while optimising existing generation resources.

For electric water heaters to be able to fully participate in demand side management, the first step is to be able to understand how individuals consume hot water. To be more specific, it is required to be able to predict or forecast how every single dwelling or a group of similar dwellings consume hot water. In this particular paper authors are looking into the ability of Artificial Neural Networks (ANN) to learn and predict the hot water consumption patterns in dwellings [10-11]. Having accurately forecast hot water consumption patterns in dwellings, it will then be able to optimally control electricity consumption to maximise supply demand balance by efficiently using existing generation capacity [12-13].

II. DEMAND SIDE MANANEGMENT USING HOT WATER USAGE

There are several reasons why hot water heaters are well suited for the use of demand side management of electricity [14]. Hot water heaters (boilers) are installed in majority of residential houses, lowering the initial installation cost as the infrastructure is already established. Also this makes energy storage distributed and closer to the end user. Secondly, water has high specific heat that allows storing relatively large amount of energy. This enables large power deviations from normal consumption for a reasonably long period. Finally, resistive hot water heaters are virtually 100% efficient as all energy is converted into heat. This fact should be emphasised in cold climate regions.

Sandels et. al in [1] presents a model for forecasting Domestic Hot Water (DHW) and other types of consumers based on non-homogenous Markov chains. The results of the DHW module coincide with the measured consumption, thus confirms that the model is reliable. Another study in [15] focuses on voltage control to reduce domestic hot water loads. In [16], DHW load profiles are simulated using physical models and then Direct Load Control (DLC) switching programs are evaluated for how load-shedding actions change customer comfort level. A peak load reduction is studied in [17] using Time of Use (ToU) and other techniques.

III. DATA USED FOR FORECASTING MODEL

The data used in this paper was taken from a project initiated by the Energy Monitoring Company in conjunction with and on behalf of the Energy Saving Trust, with funding and support of the Sustainable Energy Policy Division of the Department for Environment, Food and Rural Affairs (Defra). The data consists of temperature and volumetric consumption records from 112 different dwellings. Various sensors were fitted in houses that measured hot water volumetric consumption, inlet temperature, outlet temperature and primary circuit temperature (in regular boilers). Some additional devices measuring water temperature were fitted around pipes near kitchen sink, washing machine, bathroom basin, bath, etc (Figure 1). This allowed determining the exact spot in house where energy was used. Also boiler type, geographical region, number of occupants and other parameters were recorded.

For this particular paper it was not desired to consider the location of water use. Instead the interest was focused on the total volumetric and energy consumption of hot water boilers and hence only hot water meter readings and inlet/outlet temperatures were used. Water meter was aggregated for different sampling periods because it was being reset after every reading. Since inlet and outlet temperatures are not constant, the volumetric consumption does not show the exact energy consumption. To calculate the energy consumed it is required to look at both volumetric consumption and difference

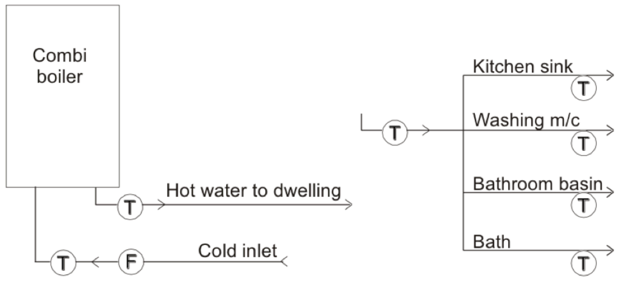

Figure 1 - Various sensor layout inside dwellings [18].

in temperatures. The following formula was used to calculate energy consumed:

$$E_t = (T_{out} - T_{in}) \times V_t \qquad (1)$$

where E_t is the energy stored in water used at time t, T_{out} and T_{in} are the outlet and inlet temperatures respectively, and V_t is the water meter reading at time t.

A. Formating Raw Data

Data of domestic hot water consumption in dwellings were recorded in year 2006. The data was recorded for about one year period at ten minute intervals. When water run-off was detected, the sampling rate increased to five seconds. The data was then resampled at constant intervals of 1, 2 and 3 hours. A range of periods was chosen to test for the best accuracy. The volumetric records where aggregated for every sampling period. Data was then looked through and any outliers or inconsistencies were discarded to improve the quality of the data which is going to be used for the model.

B. Further data analysis

In this paper authors are testing the ability off ANN to learn hot water consumption and predict future consumption. Neural Network Nonlinear Autoregressive (NAR) and Nonlinear Autoregressive Exogenous (NARX) models were tested, where autoregression is the key element in this forecast.

The next step in analysing the data was to look at auto-correlation of every single dwelling separately. Different patterns were noticed, where Figure 2 shows the first type of auto-correlation when the volumetric hot water usage

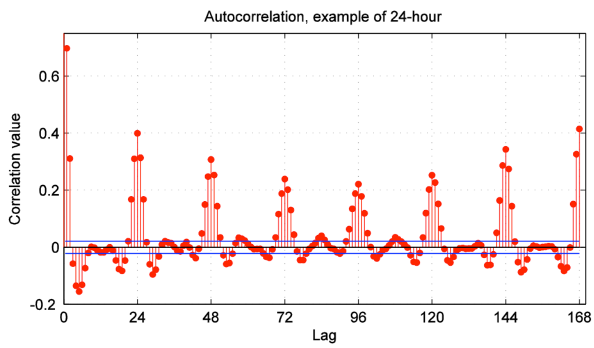

Figure 2 - Autocorrelation example, when data best correlates at 24 hour intervals.

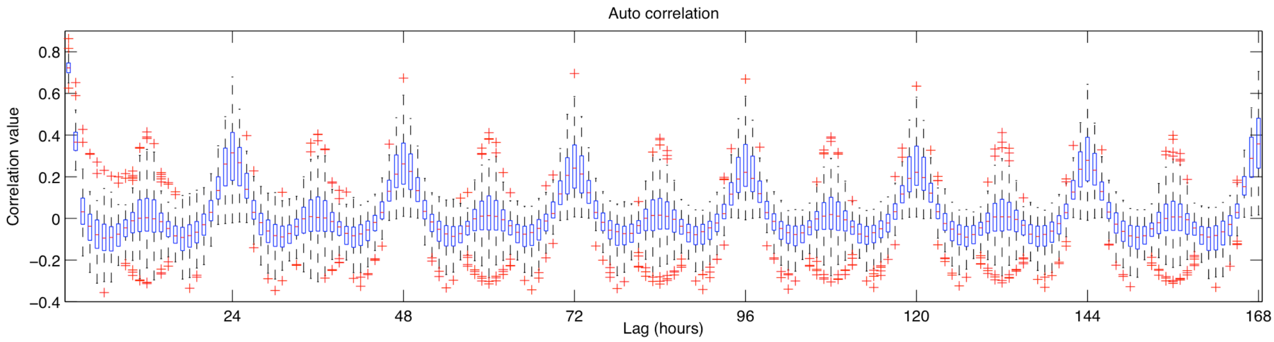

Figure 3 - A box plot of autocorrelation data of all dwellings hourly data.

correlates at every 24-hour interval. This means that occupants of this particular dwelling have strong periodic habits because their water usage follows strong pattern and is repetitive.

Another type of dwelling that can be distinguished is when consumption pattern repeats every 6 hours. In Figure 4 the data autocorrelation has four spikes for each previous day (two positive and two negative). The 12-hour period can be explained that there is a similar consumption during the night and in the middle of the day (consumption is small), and there is a large consumption in the morning and in the evening (12 hours apart). The negative correlation appears every 6-hour because peaks and lows are 6 hours apart.

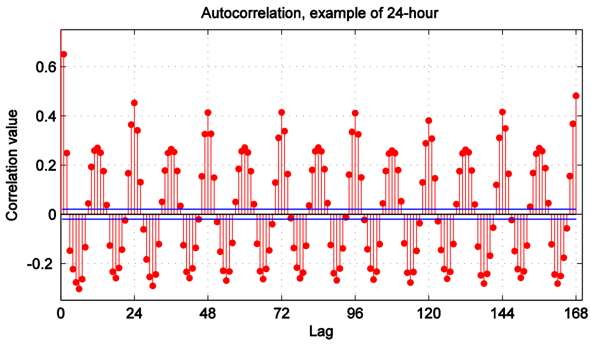

Figure 4 - Autocorrelation example, when data best correlates at 12-hour and 6-hour intervals.

Figure 2 and Figure 4 represents only two dwellings. To represent autocorrelation of all dwellings, a box plot diagram was used. Figure 4 shows a box plot diagram of autocorrelation with a maximum lag of one week. Each box represents how the dataset for particular lag is distributed and the median is represented by the red lines. Red crosses represent the outliers. It can be noticed there is a higher peak at 168th hour (mean correlation value 0.3528), which is the lag of 7th day, comparing with the same hour in every other day of the week (mean correlation value 0.2721), i.e. the data repeats best every week. This shows that not every day in a week looks the same.

The mean value of autocorrelation (period of one week) was calculated for every dwelling and arranged in an ascending order as shown in Figure 5. According to the result, no separate regions could be distinguished – the correlation values are distributed evenly.

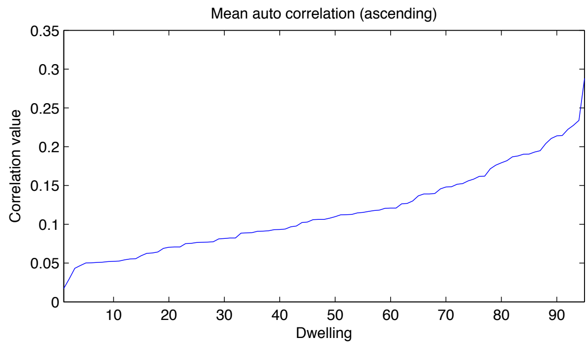

Figure 5 - Mean autocorrelation in an ascending order.

As mentioned above, hot water consumption differs depending on the day of the week. Weekly mean consumption pattern was calculated and is shown in Figure 6. It can be seen that consumption profile of Monday to Friday is different from profile of Saturday and Sunday. Further analysis has been carried to investigate this (Figure 7).

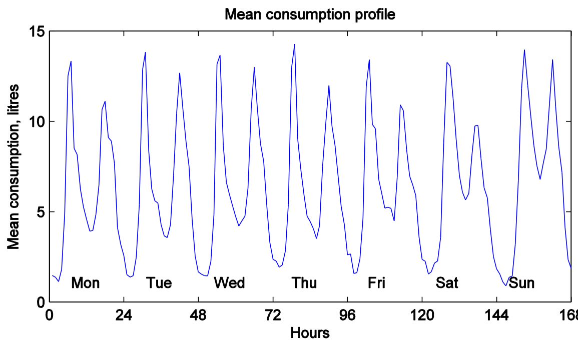

Figure 6 - Weekly mean volumetric hot water consumption pattern.

Figure 7 shows cross-correlation R-values between different days of the week. First of all, it could be divided into two regions using the diagonal. The values above the diagonal were calculated for correlation between weekdays of the same week, whereas the values below the diagonal (inclusive diagonal) are calculated using weekdays of consecutive weeks. Therefore, the matrix is not symmetric.

By looking at the calculated R-values, the matrix can be further divided into 4 regions. The values in region 1 (top left) are between 0.21 and 0.28 – it represents high correlation between consumption profiles of Monday to Friday. Region 2 (top right) and region 3 (bottom left) with low R-values (0.11 to 0.14) correspond to low correlation between working days and weekends. On the other hand region 4 (bottom right – values 0.16 to 0.19) shows that water usage on Saturdays and Sundays are similar. Conclusion could be drawn that ANN needs external input giving information about the day of the week.

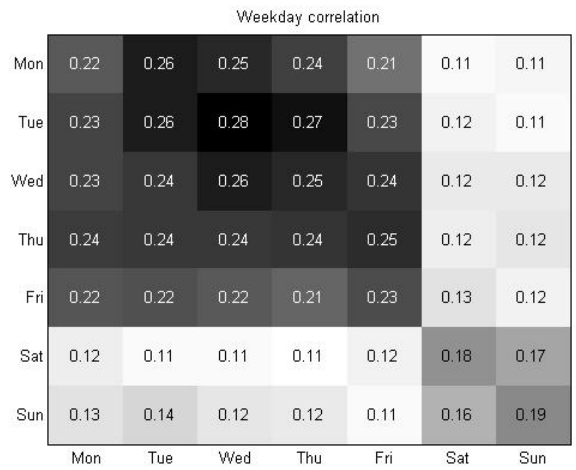

Figure 7 – Weekday correlation.

IV. ARTIFITIAL NEURAL NETWORK FORECASTING TECHNIQUE

An ANN model was created in MATLAB programing environment. To forecast hot water consumption, ANN technique was chosen for couple of reasons. ANN is a learning algorithm that can be adapted to different consumption profiles. The goal is to learn individual consumption habits of families and maximise the amount of energy that could potentially be shifted in time to reduce the overall peak seen by the generation or to better match the demand with supply. Secondly, ANN learning algorithm is very appealing because it mimics nature. Although, it is a high level algorithm that requires a relatively large amount of processing power, which nowadays becomes easily available [17].

A. Time-series forecasting using NAR model

A NAR model was created using one hidden layer with 10 neurons, where Figure 8 shows a simplified model diagram. The data division for training, testing and validation was chosen to be random for these time series. The performance was measured using Mean Square Error (MSE). ANN was trained using Levenberg-Marquardt training algorithm. An individual dwelling hourly volumetric hot water consumption time series were used to train the network. According to auto-correlation analysis, different sets of lags were tested to find the best performance. The lag configuration in the first 10 cases was in ascending order in difficulty (see TABLE 1).

Case 1 uses only 6 past inputs. Each of them is a past input of exactly the same hour of the day from past six days. Figure 4 suggests that these inputs should have the biggest weight when predicting future consumption. Case 2 has the addition of the 7th day – the same exact hour from the previous week. Figures 1-3 show that there are additional correlation peaks every 12 and 6 hours so cases 3 and 4 have additional inputs of every 12 and 6 previous hours respectively. Case 5 is the same as case 2 with the addition of the most recent consumption reading. Case 6 contains 24 hour consumption profile the same day from previous week. Case 7 and 9 has inputs from the most recent day. Finally, case 8 and 10 contains a combination of 6 with 7 and 6 with 9 respectively.

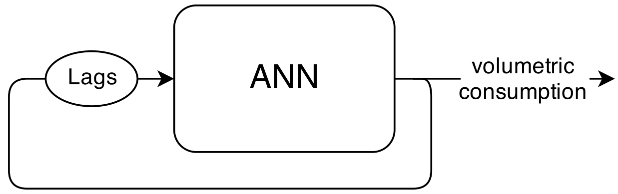

Figure 8 - Simplified NAR model.

After ANN ware trained for all dwellings, the output-target correlation R-values were recorded to judge the performance of the model. The simulations for the same 10 cases were repeated with extended lag configurations – for every lag between 24 and 168, there were adjacent lags added: t-1 and t+1. For example instead of lag 48, now ANN receives lags 47, 48 and 49. TABLE 1 in the results section shows the corresponding results for both NAR and NAR extended configurations.

B. Time-series forecasting using NARX model

The ANN was converted from NAR to NARX by adding external inputs (see Figure 9). As the data analysis above suggested, the ANN should be supplied with information containing the day of the week and whether it is a weekend or not. As a result, 6 dummy variables were constructed to represent weekday and additional dummy variable was used as a Boolean for marking weekends. Also, average hourly consumption profile (average value for the hour that is being predicted) was fed in as an external input.

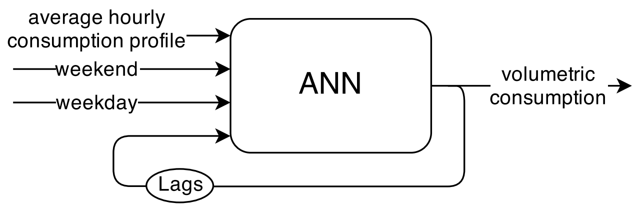

Figure 9 - Simplified NARX model.

TABLE 1 – The R-values of NAR and NARX models for different cases.

Case	Lags (NAR)	Mean R values			
		NAR	NAR ext.	NARX	NARX ext.
1	[24:24:24*6]	0,49	0,49	0,58	0,58
2	[24:24:24*7]	0,52	0,52	0,58	0,58
3	[24:12:24*7]	0,53	0,53	0,58	0,58
4	[24:6:24*7]	0,54	0,53	0,59	0,58
5	[1 24:24:24*7]	0,77	0,77	0,79	0,80
6	[1 24:24:24*7 24*7+1:24*8]	0,79	0,79	0,80	0,80
7	[1:23 24:24:24*7]	0,87	**0,88**	**0,88**	**0,88**
8	[1:23 24:24:24*7 24*7+1:24*8]	**0,88**	**0,88**	**0,88**	**0,88**
9	[2:23 24:24:24*7]	0,74	0,74	0,75	0,76
10	[2:23 24:24:24*7 24*7+1:24*8]	0,73	0,74	0,75	0,74

The simulations were then run again using lags from previous 10 cases and the results can be found in the following section.

V. Results

The extended data analysis was focused on autocorrelation. The results are as expected and show that daily volumetric hot water consumption in dwellings is similar. It is also clear that water usage habits during workdays and weekends are different (Figure 7).

The goal of this paper was to assess the ability of ANN to predict hot water consumption for separate dwellings. Figure 10 depicts how well NAR model predicts. It can be seen from the boxplot that cases 1 to 4 are quite unreliable as there is a large spread in performance variable (R value) throughout dwellings. On the other hand cases 5 to 10 show that NAR model performs well and the R-values are about 0.8 with a narrow spread.

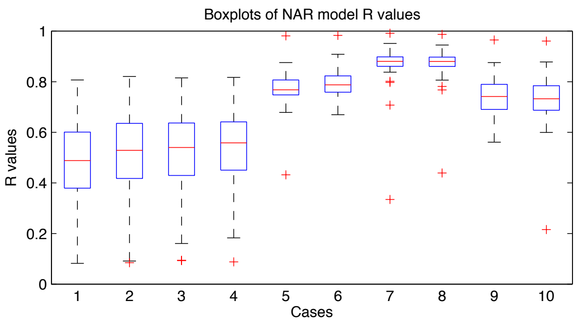

Figure 10 - Boxplot of output-target correlation R-values, NAR model simulation.

Finally, Figure 11 compares results between NAR and NARX models. It can be seen that NARX model predicts better in all cases, though the relative difference is minute in some cases. By looking at Table TABLE 1, it can be seen that cases 7 and 8 perform the best.

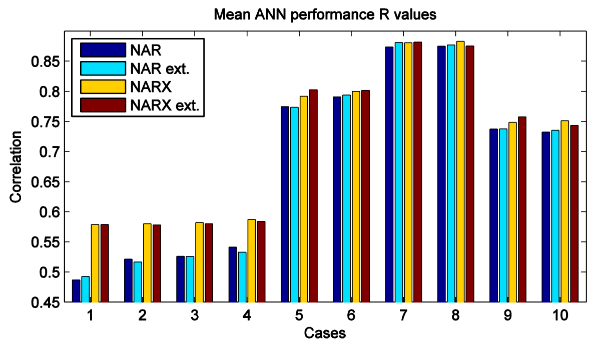

Figure 11 - Mean R values from all simulations. Graphical representation of TABLE 1

Acknowledgment

The authors would like to acknowledge the financial support of Department of Engineering and Faculty of Science and Technology, Lancaster University, UK. The authors would also like to thank the Energy Saving Trust together with Department for Environment, Food and Rural Affairs for providing the necessary data.

References

[1] C. Sandels, J. Widén, L. Nordström, Forecasting household consumer electricity load profiles with a combined physical and behavioral approach, Applied Energy, Volume 131, 15 October 2014, Pages 267-278, ISSN 0306-2619, http://dx.doi.org/10.1016/j.apenergy.2014.06.048.

[2] Rabindra Nepal, Tooraj Jamasb, Security of European electricity systems: Conceptualizing the assessment criteria and core indicators, International Journal of Critical Infrastructure Protection, Volume 6, Issues 3–4, December 2013, Pages 182-196, ISSN 1874-5482, http://dx.doi.org/10.1016/j.ijcip.2013.07.001.

[3] Goran Strbac, Demand side management: Benefits and challenges, Energy Policy, Volume 36, Issue 12, December 2008, Pages 4419-4426, ISSN 0301-4215, http://dx.doi.org/10.1016/j.enpol.2008.09.030.

[4] Gellings, C.W., "The concept of demand-side management for electric utilities," *Proceedings of the IEEE* , Volume 73, Issue 10, Pages 1468-1470, Oct. 1985, http://dx.doi.org/10.1109/PROC.1985.13318.

[5] Albadi, M.H.; El-Saadany, E.F., "Demand Response in Electricity Markets: An Overview," *Power Engineering Society General Meeting, 2007. IEEE* , Pages 1-5, 24-28 June 2007, http://dx.doi.org/10.1109/PES.2007.385728.

[6] Alessandro Di Giorgio, Francesco Liberati, Near real time load shifting control for residential electricity prosumers under designed and market

indexed pricing models, Applied Energy, Volume 128, 1 September 2014, Pages 119-132, ISSN 0306-2619, http://dx.doi.org/10.1016/j.apenergy.2014.04.032.

[7] Mohammad Rastegar, Mahmud Fotuhi-Firuzabad, Farrokh Aminifar, Load commitment in a smart home, Applied Energy, Volume 96, August 2012, Pages 45-54, ISSN 0306-2619, http://dx.doi.org/10.1016/j.apenergy.2012.01.056.

[8] Negnevitsky, M.; Wong, K., "Demand-side Management Evaluation Tool," *Power Systems, IEEE Transactions on* , Volume PP, Issue 99, Pages 1-11, http://dx.doi.org/10.1109/TPWRS.2014.2329323.

[9] Julien Eynard, Stéphane Grieu, Monique Polit, Predictive control and thermal energy storage for optimizing a multi-energy district boiler, Journal of Process Control, Volume 22, Issue 7, August 2012, Pages 1246-1255, ISSN 0959-1524, http://dx.doi.org/10.1016/j.jprocont.2012.05.011.

[10] A.S. Ahmad, M.Y. Hassan, M.P. Abdullah, H.A. Rahman, F. Hussin, H. Abdullah, R. Saidur, A review on applications of ANN and SVM for building electrical energy consumption forecasting, Renewable and Sustainable Energy Reviews, Volume 33, May 2014, Pages 102-109, ISSN 1364-0321, http://dx.doi.org/10.1016/j.rser.2014.01.069.

[11] Julien Eynard, Stéphane Grieu, Monique Polit, Wavelet-based multi-resolution analysis and artificial neural networks for forecasting temperature and thermal power consumption, Engineering Applications of Artificial Intelligence, Volume 24, Issue 3, April 2011, Pages 501-516, ISSN 0952-1976, http://dx.doi.org/10.1016/j.engappai.2010.09.003.

[12] Sossan, F.; Kosek, A.M.; Martinenas, S.; Marinelli, M.; Bindner, H., Scheduling of domestic water heater power demand for maximizing PV self-consumption using model predictive control, *Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES* , Pages 1-5, 6-9 Oct. 2013

[13] Takeyoshi Kato, Yasuo Suzuoki, Autonomous scheduling of heat-pump water heaters for mitigating voltage rise caused by photovoltaic power generation systems, Applied Thermal Engineering, Available online 5 November 2013, ISSN 1359-4311, http://dx.doi.org/10.1016/j.applthermaleng.2013.10.029.

[14] A. Arteconi, N.J. Hewitt, F. Polonara, State of the art of thermal storage for demand-side management, Applied Energy, Volume 93, May 2012, Pages 371-389, ISSN 0306-2619, http://dx.doi.org/10.1016/j.apenergy.2011.12.045.

[15] Nehrir, M.H.; Runmin Jia; Pierre, D.A.; Hammerstrom, D.J., "Power Management of Aggregate Electric Water Heater Loads by Voltage Control," *Power Engineering Society General Meeting, 2007. IEEE* , Pages 1-6, 24-28 June 2007, http://dx.doi.org/10.1109/PES.2007.386024.

[16] A. Gomes, A. Gomes Martins and R. Figueiredo, “Simulation-Based Assessment of Electric Load Management Programs”, International Journal of Enegy Reseach, Volume 23, Issue 2, Pages 169-181 (1999).

[17] van Tonder, J.C.; Lane, I.E., "A load model to support demand management decisions on domestic storage water heater control strategy,"*Power Systems, IEEE Transactions on*, Volume 11, Issue 4, Pages 1844-1849, Nov 1996, http://dx.doi.org/10.1109/59.544652.

[18] EST, Measurement of Domestic Hot Water Consumption in Dwellings, DEFRA, 2008.