An ANN-based Model for Learning Individual Customer Behavior in Response to Electricity Prices

Taha Abdelhalim Nakabi and Pekka Toivanen
University of Eastern Finland, School of Computing, Kuopio Campus, P.O. Box 1627, 70211 Kuopio, Finland
tahanak@uef.fi, Pekka.Toivanen@uef.fi

ABSTRACT

In this paper, we consider the problem of learning the electricity consumption patterns of an individual residential electricity customer, in response to electricity price signals in a demand response program. Two new methods are presented for predicting the hourly loads using the outdoor temperatures, electricity prices and previous loads. The proposed models are based respectively on a fully connected neural network and a Long-Short-term memory network. Both models deal with the uncertainty of household devices and its indoor temperature. Numerical results show the high performance of the proposed methods in terms of accuracy of the predictions. Both models can learn the consumption patterns and are able to give a good approximation of the load profile given a set of prices and temperatures. The proposed architecture can be used to investigate the price elasticity of demand, which can be used in several applications such as optimal pricing, demand flexibility or carbon emission reduction.

Keywords: *Artificial intelligence, Artificial neural networks, Customer behavior learning, Demand response programs, LSTM, Price elasticity of demand, Smart grid.*

NOMENCLATURE

Abbreviations

AMI	Advanced metering infrastructure
ANFIS	Adaptive neural fuzzy inference system
ANN	Artificial neural network
DA	Day-ahead
DR	Demand response
ELT	earliest launching time
HEMS	Home energy management system
LST	latest stopping time
LSTM	Long-short term memory
MAPE	Mean absolute percentage error
MPC	Model predictive control
NN	Neural network
NSGA	Non-dominated Sorting Genetic Algorithm
P2P	Peer to peer
PDDR	Price driven demand response
PLT	preferred launching time
PV	Photovoltaic
RBF	Radial basis function
RMSE	Root mean square error
RT	Real time
TCL	Thermostatically controlled load

Indices

t	Index for time step, t = 1, 2, …, 24

Parameters

α	Penalty factor
D	Device's running period
f	Sigmoid activation function
g	Tanh activation function
L	Load threshold

Variables

b_f	Bias vector for the forget gate
b_i	Bias vector for the input gate
b_o	Bias vector for the output gate
C_t	Cell state vector
$\underline{C_t}$	Candidate state vector
C_{extra}	Extra cost for overconsumption [€ cent]
f_t	Forget gate vector
h_t	Hidden state vector
i_t	Input gate vector (output of the input gate)
o_t	Output gate vector (output of the output gate)
P	Electricity Price [€ cent/kW]
x_t	Input vector
T	Temperature [°C]
T^i	Indoor temperature [°C]
ΔT_t	Gap between the outdoor and indoor temperatures [°C]
U_t	Electricity usage (Load) kW
W_c	Candidate gate weights matrix
W_f	Forget gate weights matrix
W_i	Input gate weights matrix
W_o	Output gate weights matrix

1. INTRODUCTION

The implementation of demand response (DR) programs that can control the consumption of electricity in residential sector and give insight about customers is a primary key for smart grids. The emerging developments of distributed energy resources (including renewable energy resources), will require shifting the flexibility of the power system to the demand side. Therefore, utility companies will have a big interest in developing DR programs that can envolve the users into this flexibility by allowing them to shift their consumption in response to incentives or price signals. Such programs should be based either on an in-depth understanding of the user's behavior and their ability to collaborate and participate in the proposed DR program, using a data-driven methodology or based on smart home technologies that can interact directly with market information and home appliances through control and monitoring of HEMS. In both cases, the residential DR programs should be based on several types of information about the users such as their consumption history, their socio-economic situation and their sensitivity for prices and incentives. This information is extremely valuable for residential price-driven DR programs (PDDR). The user's reaction to electricity prices can enable the retailer/aggregator to extract valuable insights about their customers' behavior by recognizing their consumption patterns. In the case of data-driven methodologies, electricity retailers can use this information for profit maximization through optimal pricing. The retailer's business relies on purchasing electricity by placing orders in the DA electricity markets, then selling it to their customers. This operation is ought to generate maximum profit to the retailer while maintaining his competitivity in the deregulated electricity market [1]. Therefore, having the information about electricity demand and its correlation with prices is required to maintain optimal retail operations. Additionally, the price sensitivity information can help increasing the demand's flexibility through DR programs, which can reduce carbon emission either due to a broad integration of renewables or a less usage of extra traditional power generators. Whereas from smart home technologies' point of view, the aggregators can use the functionalities provided by the smart home technologies to define products with high economic value by optimally managing HEMS resources. It can also enable the implementation of complex DR programs such as P2P trading. A review on price-driven residential DR was presented in [2]. The review has discussed 3 types of PDDR as well as P2P electricity trading. It also shows that research in residential PDDR have taken various paths regarding the parameters influencing the responsiveness of residential users to PDDR such as social contexts, psychology, individual behavior, education, and income level.

Smart home technologies such as HEMS and P2P trading systems can offer a wide range of flexibility and market opportunities both for consumers and aggregators through their functionalities. However, there are several discussions about how much focus should be given to the integration of these technologies compared to the focus given to users in a smart grid. For instance, in [3], several smart grid projects were analyzed, and the conclusions show that more attention should be given to the domestication of these technologies and their adaptation with the users' experience considering their social dimensions. Although these user-centric projects are becoming more popular, the focus is still on technological issues. The authors advise that individual users' daily routines and their social context should be taken more seriously in smart grid projects. The same opinion is shared by authors in [2,4]. Moreover, the financial requirements to implement smart homes technologies in a large scale are too expensive and require large investments. These findings can serve as a benchmark in the discussion of smart homes technologies versus data-driven approaches. Data-driven approaches are focusing more on the analysis of data related to consumption, and daily routines of consumers. They have a big potential in exploring the available data to understand the users' behavior taking into consideration his specific daily routines and societal aspects. However, they can have a big disadvantage compared to smart home technologies as they rely on the users' feedback to learn about their flexibility. This feedback can only be collected after implementing a certain DR program. In this regard, smart home technologies are more powerful as they allow the users to automatically exploit their full range of flexibility in their households. On the other hand, smart meters infrastructure is already deployed in a large scale around the world and still expanding. In 2015, U.S. electric utilities had about 64.7 million advanced (smart) metering infrastructure (AMI) installations [5]. By 2020, it is expected that almost 72% of European consumers will have a smart meter for electricity [6]. Therefore, it is important to develop data-driven methods and techniques for DR programs that can take advantage of this source of information by using the high-resolution data provided by the two-way communication channels in smart grid.

In this paper, we propose a data-driven approach to study the responsiveness of residential electricity users to PDDRs from the standpoint of individual behavior. We present two models for learning the daily consumption patterns of an individual electricity user participating in a residential PDDR program. The proposed models will learn the daily load curves from two types of loads namely shiftable/flexible and curtailable loads (namely air conditioners and TCLs). The two models are independent. They can be used separately or combined to forecast the overall daily consumption of an individual household depending on its devices. The first model is based on an evolutionary neural network that takes the 24-hours prices and predicts the corresponding shiftable loads. This model can be used alone for households without TCLs or air-conditioning systems. The second model is based on Long-Short-term memory (LSTM) network with a structure tuned by genetic algorithms. This model uses the time series of electricity consumption and corresponding temperatures and electricity prices to predict the power consumption in the next hour given the electricity price. It can be combined with the first model in the case of households with any type of domestic thermostatically controlled loads (TCLs) such as water heaters, heat pumps or air-conditioners.

The rest of this paper is organized as follows. In section 2 an overview of the related literature is provided, and the contributions of this work are explained. The problem statement is presented in section 3. Section 4 presents a theoretical framework for learning models and LSTM networks. Data generation method for price sensitive customers are presented in section 5. Numerical results are presented and discussed in

This work was supported by The Jenny and Antti Wihuri Foundation, FINLAND.

section 6. Section 7 is a conclusion.

2. RELATED WORK AND CONTRIBUTION

The literature contains several works discussing the dependencies between electricity price and demand in a price-responsive environment. Some works have approached this problem as a short-term forecasting problem of the daily load curves of a pool of price-responsive consumers. In one approach, the forecasting problem is divided in two stages; a first stage forecaster makes a normal load forecast with no account of electricity prices followed by a second stage forecaster to take into consideration the price influence. This approach was early described in [7] where an artificial neural network short term load forecaster was used for the first stage, followed by a fuzzy logic system in the second stage enhanced by a genetic algorithm to optimize the number of rules and parameters. This model was applied on simulation data generated by another fuzzy logic system. A similar approach was used in [8] but with a radial basis function RBF network in the first stage forecaster and an adaptive neural fuzzy inference system (ANFIS) in the second stage. Another approach to forecast the price responsive loads is presented as a stochastic regression, where the daily load curve is represented by a set of periodic smoothing-spline basis functions. This approach was described in [9] and applied to data from the OlyPen project [10]; the model parameters were estimated from observational time series using maximum-likelihood methods. Another approach consists of modeling the response to the electricity price as an inverse optimization problem with a set of marginal utility curves and consumption limits as in [11] where authors have introduced a solution for this nonconvex mathematical program. The model considers time and weather variables. The same authors have presented earlier in [12] a data-driven bidding model to determine the optimal market bid based on inverse optimization and bi-level programming. This model was applied to the same data from the OlyPen project. Other works have approached this subject differently. For example, authors in [13], have proposed a hybrid forecasting for electricity price and demand. This approach focuses on the bidirectional price-demand relationships when forecasting electricity market price and demand. The proposed model is composed of three main blocks. The first block uses a multi-input multi-output forecasting engine to generate initial demand and price forecasts. Then using historical market data, interdependencies between price and demand are captured in the second block and presented in the form of IF-THEN rules. In the third block, these rules are applied to the initially generated forecasts and modified accordingly. Another approach is to formulate the problem as a linear regression problem and to consider the aggregated changes in consumption over the distribution network as a weighted sum of all individual changes in consumption. This approach was described in [14] where the authors proposed a hierarchical dynamic linear model and proposed an algorithm for learning the future price elasticity of consumers based on their responses to previous pricing updates. In [15], authors studied the customer price elasticity of demand using an agent-based model. The model was used to demonstrate and quantify the economic impact of price elasticity of demand in electricity markets. Another agent-based model was proposed in [16] to study the behavior of a DA retail electricity market with PDDR. The model focused on air conditioning loads represented by agents reacting to electricity prices given by the retailer agent. The proposed multi-agent framework is based on machine learning dealing with the uncertainty of private data and incomplete information of the aggregate behavior of price sensitive loads. Most of these works have used a data-driven methodology to identify the flexibility of electric loads in response to varying prices. Other works have approached the problem from smart home technology's point of view. One approach is to consider the smart homes as energy prosumers and build optimization frameworks to support the aggregators in the definition of supply and demand bids. The smart home appliances would provide a certain flexibility considering the electricity prices and respond to fast market variations. This approach was introduced in [17]. Authors have proposed two optimization models both for DA energy market and real time management phase. The first optimization model relies on a two-stage stochastic optimization to determine the supply and demand in the DA energy market, while the second model relies on a deterministic optimization incorporated in a MPC to minimize the net cost of buying and selling energy in RT. Another approach to integrate several aspects of HEMS such as load storage, battery degradation costs and uncertainties regarding PV generation, load and electricity prices in the market and analyze them from standpoint of an aggregator trying to optimize these HEMS resources. This approach was discussed in [18] where authors have proposed a framework to optimally manage HEMS resources using an Adjustable robust optimization while considering all the aspects mentioned above. The proposed robust optimization considers different sources of uncertainty such as electrical and thermal demand, PV production and electricity prices. The latter is the main focus of this paper.

The above-mentioned models were built to approach the problem of electricity price sensitivity of a pool of price-sensitive customers in a residential area. In this paper we are interested in the individual consumption patterns as it can give valuable information about each customer's behavior and their potential response to price signals. It can also help in offering targeted electricity demand-response programs and optimal differential pricing for smart grid retail. However, the literature in this context is not very extensive. Authors in [19] presented a model for individual customers. This model can identify valuable information about different behaviors and usage patterns between different customers in response to the price and temperature signals. The proposed model is based on probabilistic Bayesian behavior model to learn the energy usage patterns of shiftable appliances and a price-demand model to predict the hourly energy consumption of air-conditioning systems. The authors also proposed a distributed pricing optimization based on genetic algorithm for the utility company to maximize its profit based on the learning results. Some other works have proposed to consider the individual customer behavior as an implicit part of a pricing optimization problem. For example, authors in [20] developed an online learning approach that can determine electricity prices that can encourage a desirable electricity consumption. This approach is based on an online convex optimization with small assumptions on the dynamics of load levels and consumer elasticity. The latter is learned implicitly using two feedback structures based respectively on full and partial information setups. The full information setup includes individual price elasticity parameters whereas the partial information setup includes only the aggregate load levels.

To the best of the authors' knowledge, there was no previous works that tried to apply neural networks and machine learning methods to detect the dependencies between electricity price and demand in a price-responsive environment at the individual user level. Given the importance of individual residential users in the low voltage network control and customer energy management, we intend to investigate the potential of Artificial neural networks (ANN) and machine learning methods for learning the price elasticity for the individual residential users. The main contributions of this paper are as follows:

- A new modeling approach for individual residential users rather than aggregate loads forecasting. The model considers the shiftable and curtailable loads in an single household. The learning is based on historical data of interactions between the user and the given prices.
- An fully-connected ANN to learn the daily electricity consumption of shiftable devices or households without TCLs systems, given a set of daily prices. The ANN takes the 24h daily prices as input and gives the 24 hours shiftable loads as output.
- An LSTM model to learn the consumption patterns of the household's TCLs and overcome the uncertainty of indoor temperature and thermal insulation.
- A genetic algorithm optimizer to fine tune the structures and parameters of the neural network and the LSTM network.

3. Problem Statement

We consider a smart grid equipped with two-ways communication system (smart meters) that announces the electricity prices corresponding to the next 24 hours. Based on these prices, residential users can schedule their devices either manually or automatically using energy management systems. We want to learn these users' responses to a set of prices proposed by the retailer. The information about the users' preferences, devices, indoor temperatures and thermal insulation is unknown. The idea of this paper is to overcome this uncertainty by extracting an abstract representation of the user's behavior in response to a set of prices and temperatures using historical data of prices, outdoor temperatures and corresponding power consumption. The main goal is to find a mapping between electricity prices, weather conditions and power consumption in a household using its historical data.

To collect the historical data related to the user's responses to electricity prices and weather conditions, we need to implement this DR program for different users and register their behavior for a long period.

We assume that the users have three types of devices:

- Shiftable devices: all devices that have a tolerance period throughout the day. For example, a washing machine or a drying machine.
- Non shiftable devices: all devices that can never be turned off or need to be run in a specific time. For example, a refrigerator.
- Curtailable devices: all devices with adaptable energy consumption levels; mostly TCLs which include water heaters, heat pumps or air-conditioners. The consumption of these devices is usually dependent on the temperature and weather conditions.

For shiftable devices, depending on the user, the scheduling may prioritize minimizing their electricity bill or maximizing their own comfort. We assume that the users are given the DA electricity prices' timetables $(P_1, P_2, \dots, P_{24})$, on which their hourly loads of the whole day $(U_1, U_2, \dots, U_{24})$ will depend. For these devices, we implement a neural network model that will take the whole set of electricity prices during a day $(P_1, P_2, \dots, P_{24})$, and outputs the set of corresponding hourly loads $(U_1, U_2, \dots, U_{24})$. In the case of households with curtailable devices (TCLs and air conditioning systems), the power consumption depends on both electricity prices and temperatures. If the temperature is too hot or too cold, the electricity consumption is expected to be high and even higher if the electricity price is low. On the other side, consumption should be low if the temperature is normal and prices are high. Therefore, we only need the price and temperature at a given hour to predict the consumption of these systems. However, we only have the information about the outdoor temperatures. The TCLs are reacting to the indoor temperature of the house. This information is therefore necessary to predict how much power is going to be used by these systems. We can have an approximate estimation of this temperature based on the outdoor temperatures in previous timesteps but only if we have information about the house's thermal insulation. Since this feature can vary from one house to another, it is not possible to have a generic model that can estimate the indoor temperature based on outdoor temperatures. Another way to approach this problem is to use abstract representations of the hidden features. We will use an LSTM based recurrent neural network to learn the consumption patterns of TCLs using the past outdoor temperatures, power consumption and current electricity price. In the next section, we explain the use of neural networks and LSTM networks and we present details of the two models.

4. Learning Models Based on Artificial Neural Networks

The use of artificial neural networks for this work is justified by their ability to learn the mapping function between the input and output without any prior information about the problem. ANNs are based on the combination of multiple processing layers to learn representations of data with multiple levels of abstraction. They are obtained by composing simple, but non-linear modules. Each of these modules transform the representation at one level into a representation at a higher, and more abstract level. The combination of such transformations in a model can enable it to learn very complex functions [21]. A multilayer NN is composed of an input layer having the same dimension as the input vector, an output layer with dimension of the output vector and the hidden layers composed of several neurons. Whereas an LSTM network is a recurrent neural network with memory blocks capable of learning long-term dependencies. It consists of a set of recurrently connected subnets playing the role of memory chips. These memory cells give the network an ability to learn the contextual information needed to predict the next sequence in a time series [22].

In the first learning model (NN1), we used a multi-layer perceptron architecture with hidden layers, input and output layers both have 24 dimensions to represent respectively the 24-hours prices and the 24-hours loads. The number of hidden layers, hidden activation functions and dropout will be tuned using a genetic algorithm. This choice is justified by the nature of shiftable appliances' loads. The running patterns of shiftable devices are 24 hours periodic and their total amount of energy is supposed to be nearly static.

In the second model, we use an LSTM network that takes as input the outdoor temperatures and power consumption at the previous timesteps and the electricity price at previous and next timesteps and outputs the power consumption at the next timestep. This process can be repeated to recurrently predict a large sequence of loads using temperatures forecasts and projected electricity prices. The choice of this architecture is justified by the following assumptions:

- The TCLs are reacting to indoor temperatures in the house and to electricity prices.
- The indoor temperature in a house depends on the outdoor temperatures in the previous timesteps, the building insulation and the amount of energy spent in heating/cooling in previous timesteps.

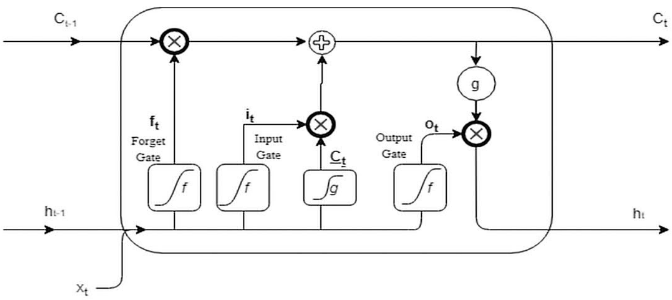

Fig.1: LSTM memory block with one cell. An LSTM cell consists of a cell state vector **C**, a hidden state vector **h** and three gates: input gate, output gate and forget

- The amount of power needed for TCLs can be predicted using the information of current price and previous values of loads and outdoor temperatures.

An illustration of an LSTM memory block with a single cell is provided in Fig.1.

These memory blocks replace the summation units in the hidden layers in a standard recurrent neural network. The input vector is concatenated to the hidden state vector and passed through the forget gate to determine how much of the cell state components can be kept. The same vector is passed through the input gate to determine how much of the new state candidate $\underline{\boldsymbol{C}}$ can pass to the new cell state. Finally, the output gate will decide how much of the transformed state cell vector can be passed to the next hidden state vector $\boldsymbol{h}_t$. The cell state vector $\boldsymbol{C_t}$ is given by:

$$C_t = f_t * C_{t-1} + i_t * \underline{C_t}$$

where the forget vector f_t is given in function of the forget gate's weight matrix W_f and bias vector b_f as follows:

$$f_t = f(W_f.[h_{t-1}, x_t] + b_f)$$

Using the same notations, the input gate vector i_t, the output vector o_t and the cell state candidate $\underline{C_t}$ are given by:

$$i_t = f(W_i.[h_{t-1}, x_t] + b_i)$$

$$o_t = f(W_o.[h_{t-1}, x_t] + b_o)$$

$$\underline{C_t} = g(W_c.[h_{t-1}, x_t] + b_c)$$

Finally, the next hidden state is given by:

$$h_t = o_t * g(C_t)$$

The proposed LSTM network consists of several layers of LSTM cells followed by a fully connected layer. The number of LSTM layers, size of LSTM cells, activation functions and other parameters will be determined by a genetic algorithm to find the optimal structure and hyperparameters.

In the case of our model, the input vector x_t is composed of the current electricity price P_t, the temperature at previous timestep T_{t-1} and the power consumption at previous timestep U_{t-1}. Fig.2 illustrate the architecture of the proposed LSTM network.

5. SIMULATION DATA

Considering the difficulty of obtaining the real data related to customers' response to electricity price signals, we will generate the data needed to test our models using electricity prices and temperatures in Finland. We simulate a customer's response to a set of prices and temperatures using an optimization model from the literature for shiftable devices to train NN1, and a fuzzy logic system for TCL systems to train the LSTM network.

5.1 Shiftable appliances simulation model

For shiftable appliances, we use a device scheduling optimization model from the literature presented in [23]. The idea is to simulate the response of a customer to a set of prices. We suppose that the user is responding to price signals by shifting the usage of certain devices to a certain period of the day to reduce their electricity bill. The user is also supposed to have preferences for running devices at certain times. Therefore, the model should consider the user's comfort function. We assume that a household has 6 shiftable devices, each one has a time window in which it can operate defined by an *earliest launching time (ELT)* and a *latest stopping time (LST)*, a *preferred launching time (PLT)*, a running period (*D)* and a load profile. In addition, the household has a basic load that is running all day and cannot be stopped or shifted. The user prefers to start his devices in the PLT but having information of the whole day electricity prices, he also wants to

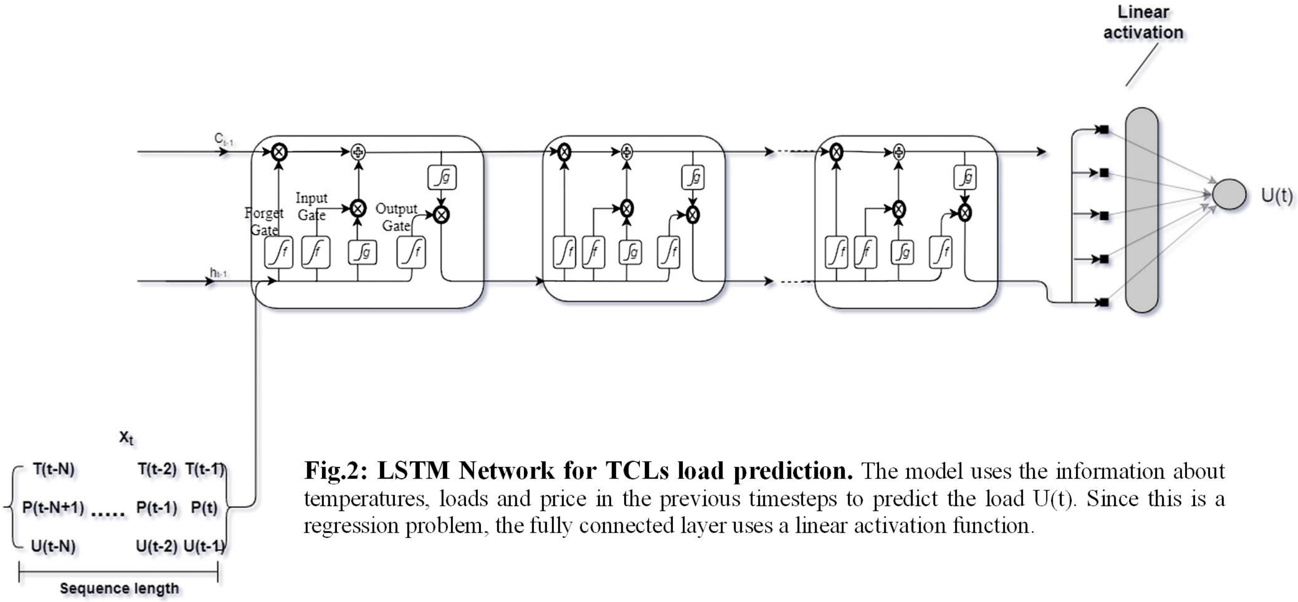

Fig.2: LSTM Network for TCLs load prediction. The model uses the information about temperatures, loads and price in the previous timesteps to predict the load U(t). Since this is a regression problem, the fully connected layer uses a linear activation function.

minimize his bill payment. Therefore, he has interest in shifting their starting time in the time window of each device to minimize the cost function but not far away from the PST to minimize the discomfort function defined by how much the actual starting time is shifted from the PLT. Additionally, the user cannot switch too much load to one time slot because we assume that an extra cost, called penalty cost, would be applied if a load threshold (L) were exceeded in a time slot. This extra cost is calculated as:

$$C_{extra} = \alpha * (U_t - L_t)$$

where α is a constant called penalty factor. This cost (if positive) is added to the cost function for every hour. This is a multi-objective optimization problem that we assume (for data generation purposes) that the users are solving to schedule their appliances either manually or using a device scheduling system. The used values of ELT, LST, PLT and running periods about the 6 shiftable devices are presented in Table 1 and their required loads in (kW) are shown in Table 2. We tried to make the values as realistic as possible for the different kinds of devices. For simplicity, the basic load will be 2kW for the whole day.

The load threshold L for one hour is chosen to be 5 kW and the penalty factor α is 1.0.

To solve this multi-objective optimization problem, the authors in [23] used a Non-dominated Sorting Genetic Algorithm II (NSGA-II). The algorithm gives for each set of 24-hours prices different Pareto optimal solutions [24] to the multi-objective optimization problem in respect of the cost and discomfort functions. It is essential to remind that NSGA-II here is only used to generate a realistic simulation of a user behavior. It doesn't contribute to the main proposed model in any case.

To generate the most diversified data for the model, we use electricity prices in Elspot DA electricity prices in Finland in the period between 1st January 2017 and the 7th September 2018 [25]. Fig.3 shows a boxplot representation of electricity prices given in €cents/kW. The simulation algorithm outputs a dataset of the expected loads for each day. A boxplot representation of the generated loads is shown in Fig. 4.

Table 1: Devices usage's data

	ELT	LST	PLT	D (hours)
Electric range	5:00	9:00	7:00	1
Oven	10:00	21:00	19:00	1
Electric water heater	9:00	23:00	10:00	3
Dishwasher	13:00	23:00	19:00	2
Clothes Washer	6:00	22:00	18:00	2
Clothes Dryer	8:00	21:00	19:00	2

Table 2: Devices' loads

	First hour	2nd hour	3rd hour
Electric range	1.6	0.0	0.0
Oven	2.8	0.0	0.0
Electric water heater	2.0	2.0	2.0
Dishwasher	1.3	2.1	0.0
Clothes Washer	1.4	2.0	0.0
Clothes Dryer	4.0	3.5	0.0

The boxplot distribution reflects the consumption constraints defined in Table 1 and Table 2. The hourly load is always higher or equal to 2kW/hour, which is the basic load defined previously. During the night hours (23:00 to 04:00) according to Table 1, no shiftable devices are operating. Consequently, the consumption in this period is equal to the basic load. The boxplot also shows that at 07:00 the consumption is equals to 3.6 most of the time. This can be seen in Table 1 and Table 2 as the load required for the electric range, which has a preferred starting time at 07:00, plus the basic load. The load at 22:00 is equal to 0 in most of the times, which can be explained by the fact that only the electric water heater and the dishwasher can be run at that time and none of them have a preferred running time close to 22:00. The rest of the loads are dependent on the electricity price. The loads did not surpass 6kW/hour, which can be explained by the penalty constraint. This data set will serve our proposed model NN1 for training and testing.

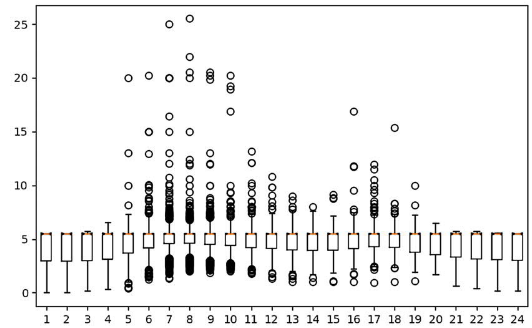

Fig. 3: Boxplot representation of daily prices distribution.

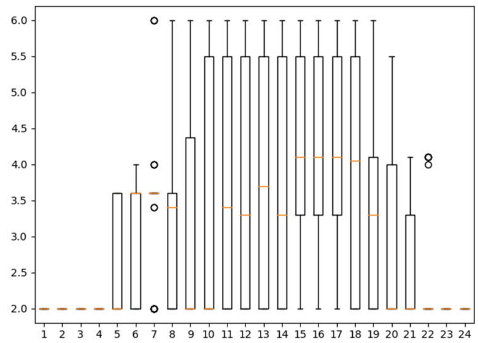

Fig. 4: Boxplot representation of generated daily loads distribution

5.2 Curtailable appliances simulation model

For curtailable devices, we use a fuzzy logic system to model the consumption of TCLs operating during the day to maintain a comfortable temperature of the space while taking into consideration the electricity price in a given hour. Fuzzy logic is the best system in this kind of problems as it can model non-qualitative concepts like *"hot temperature"* or *"low price"*. We use fixed rules for our fuzzy logic system presented as follows:

1) If (P is low) and (T is low) then (U is much-high).
2) If (P is low) and (T is average) then (U is little-high).
3) If (P is low) and (T is high) then (U is much-high).
4) If (P is average) and (T is low) then (U is little-high).
5) If (P is average) and (T is average) then (U is little-low).
6) If (P is average) and (T is high) then (U is little-high).
7) If (P is high) and (T is low) then (U is average).
8) If (P is high) and (T is average) then (U is much-low).
9) If (P is high) and (T is high) then (U is average).

Where P is the electricity price, T is the indoor temperature and U is the expected usage or consumption of the air conditioner.

However, the information about indoor temperature is unknown and need to be simulated. The idea is to consider that the dynamic of indoor temperature depends on the outdoor temperature and the electricity consumption in previous timesteps. The difference between the outdoor and indoor temperature ΔT depends on how much energy is spent during the past timesteps. A second fuzzy-logic system is designed to simulate this dynamic using the following rules:

1) If (U is low) then (ΔT is neutral)
2) If (U is much-low) then (ΔT is neutral)
3) If (T is average) then (ΔT is neutral)
4) If (T is cold) and (U is average) then (ΔT is neutral)
5) If (T is cold) and (U is high) then (ΔT is positive)
6) If (T is cold) and (U is much-high) then (ΔT is positive)
7) If (T is hot) and (U is average) then (ΔT is neutral)
8) If (T is hot) and (U is high) then (ΔT is negative)
9) If (T is hot) and (U is much-high) then (ΔT is negative)

We use a Mamdani type model in this proposed fuzzy-logic system and the centroid method is adopted for defuzzification [26]. Membership functions for price, temperature and usage are presented in Fig.5 and Fig.6

The simulation data is given by the combination of both fuzzy logic systems. The indoor temperature at time t is given by:

$$T_t^i = T_{t-1} + \Delta \mathrm{T_t}$$

It is essential to remind that the indoor temperature is only an intermediate value for the sake of simulation of the house's response to electricity price and temperatures. The second fuzzy

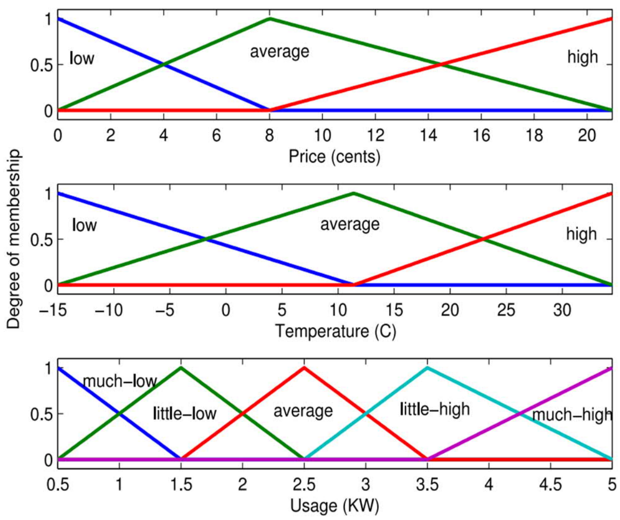

Fig.5. Fuzzy membership functions for air conditioner dynamics (fuzzy system 1).

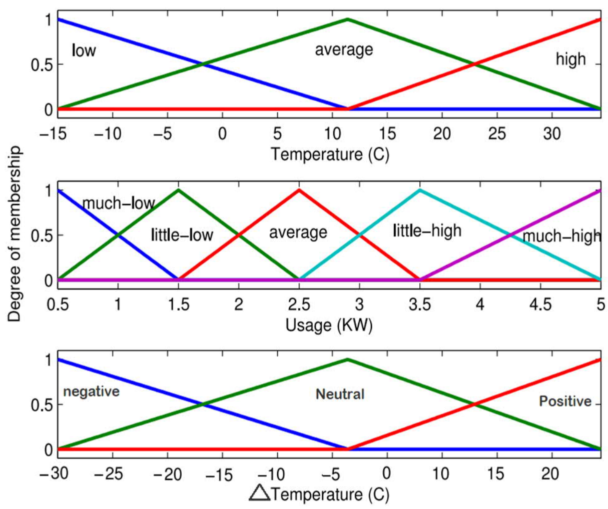

Fig.6. Fuzzy membership functions for indoor temperature dynamics (fuzzy system 2)

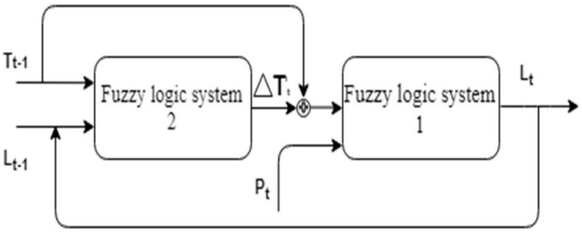

Fig.7 Simulation model for air conditioning devices in response to electricity prices and outdoor temperatures.

logic system is not claimed to give an accurate approximation of the actual indoor temperature. Fig.7 illustrates the simulation mechanism combining the first and second fuzzy systems.

We use the same dataset of prices as for the first model in addition to historical air temperatures data related to the same period. We used the weather data from *Kaisaniemi* observation station in Helsinki available online in [27]. This dataset will serve as a training dataset for the LSTM network.

6. Numerical results

In this section, we evaluate the results of the two learning models and compare them with other methods. After generating the training data for shiftable and curtailable appliances, we split the datasets into training, validation and testing datasets. We use the measures of the root mean square errors (RMSE) and the mean absolute percentage error (MAPE) to evaluate the performance of the models against the testing datasets. All the developments and models in this paper are implemented using Python3. The NN models are implemented using TensorFlow library and Keras for the second model.

6.1 Multi-layer perceptron results

After convergence, the model is tested against the test dataset. Results of MAPE and RMSE at each hour for both training and test data are given in Fig. 8 and Fig. 9. The RMSE and MAPE tend to be null at the night (22-03) and get bigger in the early morning and afternoon (6-7, 9-10 and 15-16). The number of devices supposed to be run in each period, can explain the difference between the daily hours in terms of MAPE and RMSE. In the night period, no devices are supposed to run, which makes the total consumption equal to the basic load for all days. This made it easy for the model to learn this pattern with high accuracy. For the morning and afternoon periods, several devices can be run separately or simultaneously depending on the price. This variety can sometimes confuse the model which makes it give relatively bigger error. The model's RMSE and MAPE are generally very small and can give proof of the efficiency of the model. Additionally, the errors related to training and test data are of the same magnitude (test error seem to be less than training error in RMSE) which confirms that the model is not overfitting to the training data.

We used the same test data to evaluate the results of two alternative models namely linear regression and a simple one-layer neural network. The results of MAPE and RMSE errors are given in Fig. 10 and Fig 11. The results show that both models are radically left behind the NN1. The MAPE error of linear regression reaches over 350% and RMSE reaches over 50. Whereas the simple one-layer NN has given more 200% of

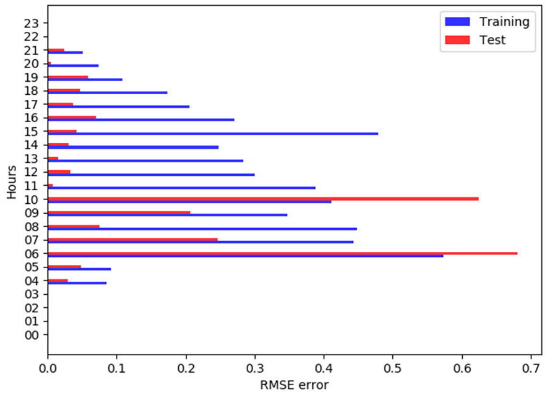

Fig.8: RMSE error for 24 hours

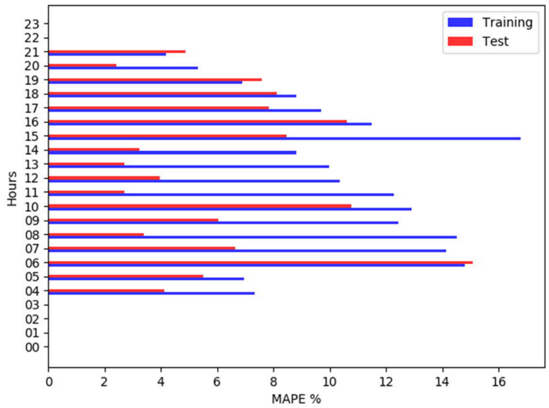

Fig.9: MAPE percentage error for 24 hours

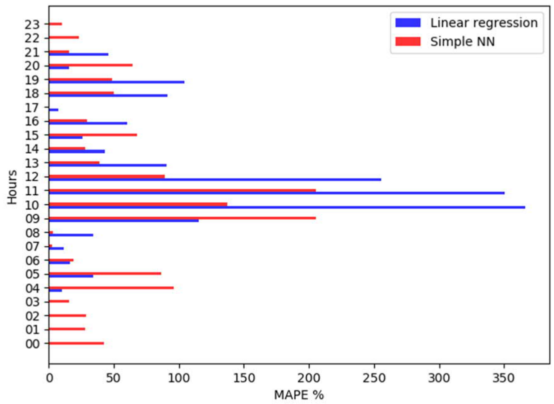

Fig.10: MAPE error for linear regression and simple

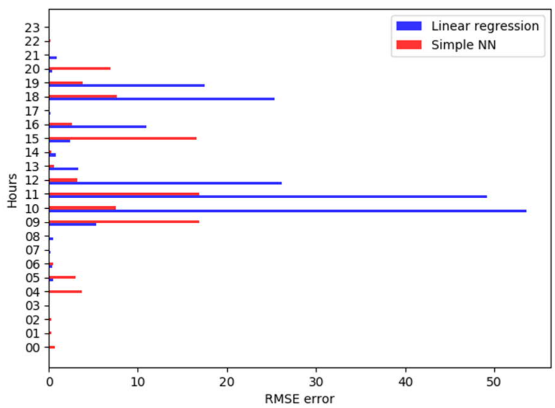

Fig.11: RMSE error for linear regression and simple NN

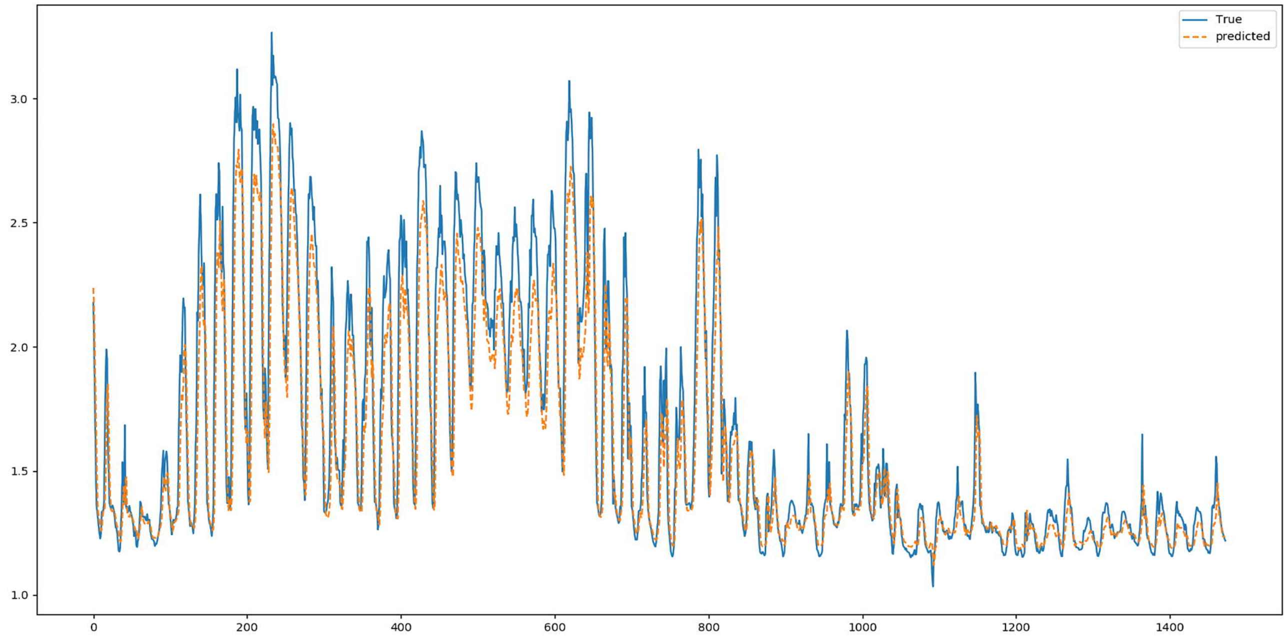

Fig.12: Comparison between the actual and predicted loads.

MAPE and almost 20 of RMSE. These results show that these models cannot be useful for this problem. When evaluating the model against the test data, it has made new predictions based on daily prices.

6.2 LSTM network results

We train the LSTM network on a dataset of 11788 datapoints and test on 2947 datapoints. The hyperparameters and structure of the network are optimized using a genetic algorithm. The parameters that we need to optimize are given in Table 3. The utility function to minimize is the RMSE error between the predicted and actual loads of the testing dataset. The genetic algorithm is run on a population of size 50 over 100 iterations. The results of the best solution are given in Table 4. The comparison between the real and predicted loads is graphically presented in Fig12.

The small values of RMSE and MAPE can confirm the efficiency of the model. LSTM network is giving a good prediction of electricity load given a sequence of past electricity prices and temperatures.

Having the trained models, we will be able to predict the hourly energy consumption for a residential customer, given a DA set of prices and temperatures. This is a very useful tool for testing the responsiveness of residential customers to variable prices of electricity. This tool will give a big opportunity for utility companies to implement DR programs or optimize prices for profit maximization.

7. Conclusion

In the future smart grids, residential customers are supposed to take part of the grid's flexibility by participating in DR programs and reacting to electricity price signals. Consequently, the electricity retailers or the utility companies will have a big interest in getting the information about the responsiveness of the users to certain electricity prices. Hence, we proposed a new approach to learn about the individual user's behavior in a price sensitive environment. Two new machine learning-based models were proposed to detect the consumption patterns for shiftable appliances and TCLs in a household. The first model predicts the 24-hours loads from shiftable appliances and basic loads given a set of 24 hours prices. It can be used for households without any TCLs, whereas the second model predicts a one-hour consumption from TCLs given a sequence of electricity prices and the outdoor temperature. It can be combined with the first model to predict the overall consumption of a household with TCLs.

The first model is based on a multi-layer perceptron, and the second is a recurrent model based on LSTM network. The two models were trained and tested using a simulation data, generated from an optimization model simulating households' device and a fuzzy logic system simulating TCLs. The results were evaluated using MAPE and RMSE errors. Both models gave a good accuracy over the time period of tests, which was confirmed by the small values of the errors. The results showed

Table 3 Hyperparameters of the LSTM network. The values of activation functions and optimizing algorithms are taken from TensorFlow documentation on the possible values of these parameters.

Hyperparameter	**Description**	**Values**
Sequence length	The number of previous datapoints used in the input	range (2, 24)
LSTM cell size	The size of memory cell vector	{10, 20, …, 50}
LSTM cells	Number of consecutive LSTM layers	range (1, 5)
Dropout	Percentage of random neurons ignored in the training process	{0%, 10%, …, 40%}
Activation	'g' function	'relu', 'elu', 'selu', 'tanh', 'sigmoid',
Recurrent activation	'f' function	'hard_sigmoid'
Optimizer	Training algorithm	'adadelta', 'adagrad', 'rmsprop', 'adam', 'adamax', 'nadam', 'sgd'

Table 4: NN2 Results of Hyperparameters optimization

Sequence length	2
LSTM cell size	30
LSTM cells	2
Dropout	0.2
Activation	'tanh'
Recurrent activation	'selu'
Optimizer	'rmsprop'
RMSE	0.184
MAPE	6.16%

that we can learn the consumption patterns of a residential consumer in a price varying environment using their historical data. The LSTM network has shown its good ability to learn accurately the electricity usage's patterns with incomplete information about the indoor temperatures of households. LSTM network is used here for the first time in the context of learning the electricity consumption of TCLs in response to electricity prices and outdoor temperatures. Thanks to its abstract extraction of features, it was able to learn the dynamics of the electricity consumption of TCLs from the past sequence of loads, prices and outdoor temperatures. The combination of these two models can be used in a variety of cases such as price optimization for retailers' profit maximization, reducing energy consumption at peak hours, balancing between the production and demand, or reducing carbon emissions by shifting the flexibility to the demand side and replacing traditional generators by renewable, yet less flexible, energy resources.

REFERENCES

[1] P. Valtonen, "Operating and market environment, in Distributed energy resources in an electricity retailer's short-term profit maximization", [dissertation]. PhD thesis, Lappeenranta University of Technology 2015.

[2] X. Yan, Y. Ozturk, Z. Hu, and Y. Song, "A review on price-driven residential demand response," Renewable and Sustainable Energy Reviews, vol. 96, pp. 411–419, 2018.

[3] G. Verbong, S. Beemsterboer, and F. Sengers, "Smart grids or smart users? Involving users in developing a low carbon electricity economy," Energy Policy, vol 52, pp. 117–125, 2013

[4] M. Hansen, M. Borup, "Smart grids and households: how are household consumers represented in experimental projects?", Technology Analysis & Strategic Management, vol. 30, no. 3, pp. 255–267, 2018.

[5] The U.S. Energy Information Administration (EIA), FAQ. "How many smart meters are installed in the United States, and who has them?" https://www.eia.gov/tools/faqs/faq.php?id=108&t=3 (accessed 13th February 2019).

[6] European Commission, Energy, Smart grids and meters, https://ec.europa.eu/energy/en/topics/markets-and-consumers/smart-grids-and-meters. (accessed 13th February 2019)

[7] A. Khotanzad, E. Zhou, and H. Elragal, "A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment," IEEE Transactions on Power Systems, vol. 17, no. 4, pp. 1273–1282, 2002.

[8] Z. Yun, Z. Quan, S. Caixin, L. Shaolan, L. Yuming, and S. Yang, "RBF neural network and anfis-based short-term load forecasting approach in real-time price environment," IEEE Transactions on Power Systems, vol. 23, no. 3, pp. 853–858, 2008.

[9] J. Hosking, R. Natarajan, S. Ghosh, S. Subramanian, and X. Zhang, "Short-term forecasting of the daily load curve for residential electricity usage in the smart grid," Applied Stochastic Models in Business and Industry, vol. 29, no. 6, pp. 604–620, 2013.

[10] D.J. Hammerstrom, et al. Pacific northwest gridwise testbed demonstration projects, Part I. Olympic peninsula Project, Report no PNNL-17167, Pacific Northwest National Laboratory, Richland, Washington, 2007.

[11] J. Saez-Gallego, J. M. Morales, "Short-term forecasting of price-responsive loads using inverse optimization", IEEE Transactions on Smart Grid, vol. 9, no. 5, pp. 4805–4814, 2017.

[12] J. Saez-Gallego, J. M. Morales, M. Zugno, and M. Henrik, "A data-driven bidding model for a cluster of price-responsive consumers of electricity," IEEE Transactions on Power Systems, vol. 31, no. 6, pp. 5001-5011, 2016.

[13] A. Motamedi, H. Zareipour, and W. D. Rosehart, "Electricity price and demand forecasting in smart grids," IEEE Transactions on Smart Grid, vol. 3, no. 2, pp. 664–674, 2012.

[14] V. Gómez, M. Chertkov, S. Backhaus, and H. J. Kappen, "Learning price-elasticity of smart consumers in power distribution systems," In: Proceedings of the IEEE Third International Conference on Smart Grid Communications. pp. 647–652, 2012.

[15] P. R. Thimmapuram and J. Kim, "Consumers' price elasticity of demand modeling with economic effects on electricity markets using an agent-based model," IEEE Transactions on Smart Grid, vol. 4, no. 1, pp. 390–397, 2013.

[16] K. Dehghanpour, M. H. Nehrir, J. W. Sheppard, and N. C. Kelly, "Agent-Based Modeling of Retail Electrical Energy Markets with Demand Response," IEEE Transactions on Smart Grid, vol. 9, no. 4, pp. 3465–3475, 2018.

[17] J. Iria, F. Soares, and M. Matos, "Optimal supply and demand bidding strategy for an aggregator of small prosumers," Applied Energy, vol. 213, pp. 658–669, 2018.

[18] C. A. Correa-Florez, A. Michiorri, and G. Kariniotakis, "Robust optimization for day-ahead market participation of smart-home aggregators," Applied Energy, vol. 229, no. 1, pp. 433–445, 2018.

[19] Fan-Lin Meng, Xiao-Jun Zeng, "A Profit Maximization Approach to Demand Response Management with Customers Behavior Learning in Smart Grid," IEEE Transactions on Smart Grid, vol. 7, no. 3, pp. 1516–1529, 2016.

[20] S.J. Kim, G.B. Giannakis, "An Online Convex Optimization Approach to Real-Time Energy Pricing for Demand Response," IEEE Transactions on Smart Grid, vol. 8, no. 6, pp. 2784 – 2793, 2017.

[21] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[22] F. A. Gers, J. Schmidhuber, F. Cummins, "Learning to Forget: Continual Prediction with LSTM", vol. 12, no. 10, pp. 2451–2471, 2000.

[23] Z. Garroussi, R. Ellaia, E. Talbi, "Appliance scheduling in a smart home using a multiobjective evolutionary algorithm", In: International Renewable and Sustainable Energy Conference (IRSEC), pp. 2380–7393, 2016.

[24] K. Deb, H. Gupta, Searching for Robust Pareto-Optimal Solutions in Multi-objective Optimization. In: C.A. Coello Coello, A. Hernández Aguirre, E. Zitzler, editors. Evolutionary Multi-Criterion Optimization. Lecture Notes in Computer Science, vol 3410, pp 150–164, Springer, Heidelberg, EMO 2005.

[25] Nord Pool, Elspot Day-ahead, Prices, 1st January 2017 to 7th September 2018. https://www.nordpoolgroup.com/Market-data1/Dayahead/Area-Prices/ALL1/Hourly/?view=table (accessed 8th September 2018).

[26] J.L. Castro, "Fuzzy logic controllers are universal approximators," In: IEEE Transactions on Systems, Man, and Cybernetics, vol. 25, no. 4, pp. 629–635, 1995.

[27] Finnish meteorological institute, weather observations, *Kaisaniemi* observation station Helsinki, 1st January 2017 to 7th September 2018, https://en.ilmatieteenlaitos.fi/download-observations (accessed 8th September 2018).