UNIVERSITY OF AGDER

Domestic Electricity Demand and Peak Predictions Considering Influence of Weather Parameters

Songpu Ai

Supervisor
Mohan Lal Kolhe
Lei Jiao

This master's thesis is carried out as a part of the education at the University of Agder and is therefore approved as a part of this education. However, this does not imply that the University answers for the methods that are used or the conclusions that are drawn.

University of Agder, 2015
Faculty of Engineering and Science
Department of Information and Communication Technology

Abstract

Predictions for domestic load demand and peak are of great importance for the establishment of home energy management system in order to smooth demand peak by managing/scheduling loads. In this thesis, artificial neural network (ANN) based demand (active and reactive) prediction is built and analysed for a domestic household. Firstly, the prediction is carried out based on historical load consumption data only, and the measurement errors and the missing segments in the available data are processed in order to eliminate their impacts on the predication model. Then, since the environmental parameters have impact on the load consumption in many ways, temperature and solar radiation are employed as external weather parameters in the prediction. Finally, demand prediction of the incoming 1st-24th hour is built for forecasting active and reactive load to predict the highest peak in the next 24 hours with peak load and occurrence timing. Around one and half year' historical domestic load and weather data of a typical Southern Norwegian household is utilized for training and testing purpose. From the results, the performance of predictions is evaluated and compared. It can be observed that data processing improves the prediction efficiency. In addition, active load prediction using load data together with weather data performs better than the case where only load data is adopted. And for reactive load prediction, no significant improvement achieves by introducing weather parameters into the prediction. Furthermore, the performance of peak prediction upgrades when data with longer duration is adopted.

Preface

This thesis is the result of project IKT-590 in the Master program of Information and Communication Technology Department, University of Agder, Grimstad. The work on this project started from 5 Jan. 2012 and ended on 25 May 2012, with 30 ECTS. The thesis is associated with European Union project SEMIAH "Scalable Energy Management Infrastructure for Aggregation of Households". The main goals of project "Domestic Electricity Demand and Peak Predictions Considering Influence of Weather Parameters" have been completed holding the motivation to build load and peak prediction for domestic household. MATLAB® is utilized in the thesis for simulation. The thesis is accomplished by using LaTeX.

Here I would like to show my deepest gratitude to my supervisors, Professor Dr. Mohan Lal Kolhe and Associated Professor Dr. Lei Jiao. They offered this interesting project to me and provided me valuable instructions in every stage of the project. Their impressively patient and professional guidance have helped me overcome the challenges and achieve the targets favourably.

I would like to extend my gratitude to other professors and teachers in the Information and Communication Technology Department. They not only taught me the knowledge in their specific fields, but also taught me more on the way of research and the philosophy of life.

I also gratefully acknowledge my friends. Their encouragement and help gave me faith to persevere in the work and accomplish this thesis.

Last but not least, the appreciation of me should be given to my dear parents and girlfriend. Their everlasting support and love are the foremost source of my endeavour.

Ai, Songpu

Grimstad

May 24, 2015

Acknowledgement

The University of Agder gratefully acknowledge the support for this work of the European Commission under the FP7 program with contract number FP7-ICT-2013-11-619560 for project SEMIAH 'Scalable Energy Management Infrastructure for Aggregation of Households'. http://www.semiah.eu.

Abbreviations

ANN	Artificial Neural Network
BP	Back-Propagation
GSI	Global Solar Irradiance
HEMS	Home Energy Management System
MAPE	Mean Absolutely Percentage Error

Contents

List of Publications

[1] S. Ai, M. L. Kolhe, L. Jiao, and Q. Zhang, "Domestic load forecasting using neural network and its use for missing data analysis," in *Proc. IEEE - 9th International Symposium on Advanced Topics in Electrical Engineering*, May 7-9, 2015. 63

[2] S. Ai, M. L. Kolhe, L. Jiao, N. Ulltveit-Moe, and Q. Zhang, "Domestic demand predictions considering influence of external environmental parameters," in *Proc. IEEE 13th International Conference on Industrial Informatics (INDIN)*, Jul. 22-24, 2015. 67

[3] S. Ai, M. L. Kolhe, L. Jiao, "External Parameter Contribution in a Domestic Load Forecasting Neural Network," in *Proc. 4th IET Renewable Power Generation Conference (RPG 2015)*, provisionally accepted.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Electrical power has been wildly utilized in modern society. Along with plenty kinds of electric equipments entered into people's live, the requirement of electricity has been increasing for many years. To meet the great electrical energy demand from companies and residents, power plants have to increase their generating capacity constantly to fulfil the daily needs. However, due to the limitation of the generation, transmission, distribution, and storage techniques, electricity cannot always be utilized by inhabitants or factories in time. The portion of unexploited energy is wasted. Thereby, additional greenhouse gas and power generation waste are produced. Researchers have struggled on the topic about reduce wasting for years. As one of the effective methods, increase or decrease electricity generation based on the prediction of future load demand is utilized to reduce the waste by reducing the gap between power load supply and actual load consumption. This kind of load forecast-management is usually done by power companies to conserve resources and money.

However, with the arising and popularizing of smart house in recent years, the management of residential electricity load, demand-side management, becomes possible. As a kind of energy management system facing domestic consumer which is also been seen as a portion of smart grid system, home energy management system (HEMS) is introduced by [10]. HEMS is going to be employed on managing/scheduling the operations of controllable power-intensive non-critical domestic loads in a household (demand-side management) to have better effective on using energy serviced by the local power distribution infrastructure. A little more details about HEMS are given in Chapter 2.

In order to manage/schedule the load usage of a household, it is necessary to have a prediction about the future load demand of the household. For this, load demand prediction model for both active and reactive power forecast is required. The active power prediction is going to be utilized on managing the active power supply and demand. While the reactive power prediction will be in charge of managing local voltage profile and reducing the network loss and improving power factor.

It should be noticed that not only load forecast is required by HEMS, the information about peaks occurring in near future is also very important for the system [10]. In the local power distribution network, the overall power consumption fluctuates with the change of simultaneous load usage from the customers. An aggregated or collective behaviour of a certain number of customers may lead to a power consumption peak. For example, the usage of induction cooker and electric oven at dinner time usually leads a local power consumption peak. Due to the regularity of human behaviour, it is hard to avoid the occurrence of daily demand peaks usually. In order to avoid the damage caused from electricity scarcity, such as a power outage, electrical service providers should design their grid to support the maximum potential load peak [11, 18]. Nevertheless, since the continuous increasing of power consumption, the power distribution infrastructures should be updated constantly. The investment is expensive and ineffective. However, HEMS offers a new approach to smooth the local demand peak from each household, which is shifting the operations of controllable domestic loads from a potential power demand peak to an off-peak period. Therefore, an accuracy forecast of the demand peak of a household with the peak value and occurrence timing is important for HEMS when it is used for smoothing potential peak.

1.2 Background

There are numerous of existing electricity demand and peak prediction methods in the research community. For demand prediction, several modelling methods are introduced by [35]. Including which, artificial neural network (ANN) is selected on developing the domestic load prediction model in this thesis. Artificial neural network is a kind of statistical learning algorithm which follows the structure and function of biological neural network. Since ANN is able to estimate non-linear systems, it is widely utilized for modelling and forecasting the complex relationship between input and output [15, 22, 31]. Various studies, for example [21, 29, 30], have been carried out to model or forecast electricity load using ANN. In general, these researches do modelling and forecasting based on regional power consumption data. However, with the development of technology, more data is able to be obtained and analysed. The prediction aiming at domestic electricity power demand has become possible, which is investigated in the thesis. Besides, for the demand peak prediction, several methods are proposed in [11, 16, 28]. However, the peak occurrence timing prediction based on ANN load forecast as a feasible approach has not yet been excavated for the moment. In the thesis, this possibility will be discussed.

The surrounding environment of a household also has an important impact on the power demand of a household. For example, heaters require more electric power to keep the house temperature within the comfortable range when temperature decreases. Also, in cloudy or rainy days (when the solar radiation is low), the dryer is more frequently used which leads more power consumption. Therefore, additional environment meteorology parameters in domestic power prediction should be useful for increasing the prediction accuracy. It has been validated in [3, 12, 30] that weather factors have effect on forecast electric load of a power system. However, the study of using weather data to forecast household power demand is still limited. Therefore, it is interesting to study the load prediction with/without considering external weather parameters and to further build peak prediction based on the load forecast results.

1.3 Thesis Objectives and Limitations

The purpose for this master thesis work is to get prediction of active and reactive power, in order to realize demand-side management in domestic household by using HEMS. Moreover, the impact of weather on the prediction should be studied. Therefore, the tasks that should be accomplished in this project can be listed as follows:

- Establish a demand prediction model for single domestic household;
- Investigate the influence of the weather parameters on the demand prediction;
- Establish a forecast for domestic household to find the demand peak (with load and occurrence timing) in near future.

The main limitations on the way of achieving thesis objectives are listed as follows:

- The available load data has errors and missing segments, in order to avoid the impact of the missing and erroneous data, the selection of training and testing set should be careful;
- The available category of weather data is limited, only temperature and solar radiation data are able to be used;

- The available weather data is the regional data, which is hard to completely reflect the real weather situation around the household;

1.4 Thesis Organization

The introduction, background and objectives of the project is introduced in this chapter. The remainder of this thesis is organized as follows.

Chapter 2 presents a brief introduction of the HEMS firstly. Then the researches on load demand prediction using ANN, on influence of weather parameter for load forecasting, and on peak forecast are described. Besides, the basic knowledge about ANN is also presented.

Chapter 3 describes the specialization of the prediction models as well as the process method for the raw data of power consumption. The errors existing in the raw data are considered and the weather data is processed to fit the format of load data.

Chapter 4 provides the structure of the predictions. Mathematical expressions of load and peak prediction are given. And assessment methods are introduced.

Chapter 5 analyses the performance of one hour demand prediction using only load data firstly. Then the influence of the weather parameters on the one hour demand prediction is investigated. In addition, the performance of load prediction for the incoming 1st - 24th hour is studied and the prediction results are used for finding the highest peak. Finally, discussions about the current result are introduced.

Chapter 6 concludes the thesis and the summarizes the main contributions of the project. Finally, the future scope of the study is discussed.

Chapter 2

Domestic Electric Load Forecasting Using ANN and Peak Prediction

2.1 Introduction

Smart grid is the new generation electric grid which is able to utilize ICT together with power system engineering technology to obtain and manage the behaviours of power supplier and consumers [9, 33]. It is considered as a solution of balancing the demand and response by accessing information from both sides and managing the energy consumption depending on supply condition. As a terminal located in households, HEMS is considered in the smart grid to increase the electricity efficiency by realizing demand-side management in houses.

Briefly, an important function of HEMS utilized in a smart grid is to smooth the load demand peak before occurrence by shifting the energy consumption [10]. A HEMS system should be able to monitors and adjusts energy consumption with the help of sensors, meters, intelligent devices, etc. to provide effective load and peak management [6, 18]. Meanwhile, HEMS is also considered as the core of green buildings and the system is able to benefit the customers under the demand response and time-of-use electricity pricing strategies [6, 20]

In order to manage and schedule household loads, an accurate prediction about the load demand of the household and a forecast about potential demand peak in near future are required by HEMS. In the thesis, above prediction models are built and tested for a typical southern Norwegian household.

2.2 Load Forecasting Using Neural Network

Numerous of load usage/prediction methods based on ANN have been presented in existing articles. At the macro view, power consumption prediction models using ANN are illustrated by [8, 35, 36] for different size of regions in a country. Among these articles, [8, 35] develop the predictions for hourly power usage (24 samples per day) and [36] develops the prediction on the typical hour of days (1 sample per day). For a smaller region, [7] addresses a load prediction for commercial buildings using neural network. Although the forecast region and forecast type of the models are different, it is observed by the power consumption profiles of the regions (states, a city or a building) that fluctuate regularly with time. At the micro view, ANN is utilized for establishing the load usage model for the power components such as heater, cooker in relatively stable use environments [21]. The power consumption profiles of the considered equipment also follow a relative stable curve. Furthermore, for the articles that consider load prediction of household, for example [17] and [39], a processed more regular power usage profile is preferred by the authors. It is because that comparing with power usage of a region, load consumption of a household is more easily influenced by personal behaviours. Techniques such as using aggregated household power usage data and using wavelet transformation are utilized in the papers to reduce impact of the random human activities for the load prediction.

However, in order to forecast the load demand of a household precisely, the impact of human "random" activities should be considered in the prediction model. Moreover, though human actions can be considered as a kind of random in short term, in a longer term observation it may be able to find that there is an action pattern behind the seemingly random events. For example in a family, the father sometimes forget to shut down the oven and the mother is always able to find the problem and shut down the oven in time. This kind of family "hobby" is easily be covered or erased after a widely data processing. Perhaps, these long term "hobbies" are difficult to be clearly described by mathematical equations. However, by training the neural network with long term historical data, it might be possible to forecast the occurrence of those family "hobbies" by using the non-linear algorithm which follows the structure of biological neural networks and as a result, achieving a better load demand prediction of each domestic household.

2.2.1 Artificial Neural Network

Artificial Neural Network (ANN) is a kind of statistical learning model following the operating process of biological neural network. ANN is widely used on classification and function fitting for multivariate non-linear problems [15, 22, 26, 31].

In an ANN, neurons are the basic unit of a neural network. Same as in biological neural network neurons have multiple incentive inputs. An artificial neuron should have several input components. Then, after processing the input components by the neuron, an output should be given. Furthermore, in order to differentiate the relationship between input components and the output, weights are configured for the components. Denote that a neuron named as j has m input components. Each component is denoted by x_k, where $k = 1, 2, ..., m$. And in the neuron j, the weight for each input component is denoted as w_{kj}. The output of neuron j is denoted as o_j. The basic structure of the neuron j is illustrated in Fig. 2.1 [14]. As illustrated in the figure, there are two portions of calculation processed in each neuron. The first portion is to calculate the transfer function, which is to obtain the weighted summation of the input components. The second portion is to calculate the output of the neuron by the activation function f of the neuron.

Therefore, if we denote $x_k(t)$ as the input components at time t, and ignore the synaptic delay, the output of neuron j at time t can be expressed by (2.1) [35], where T_j is the bias value of the neuron, which is considered as the threshold of the neuron.

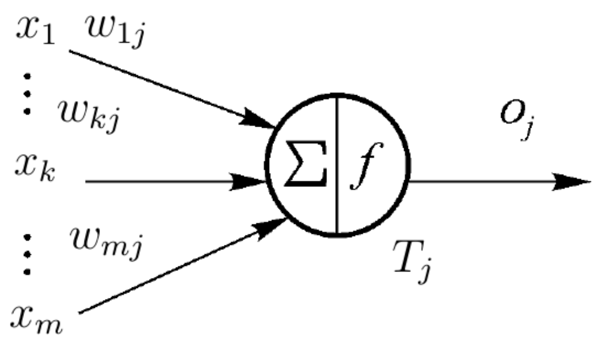


Figure 2.1: The basic structure of a neuron j.

$$o_j(t) = f\left\{\sum_{k=1}^{m} w_{kj}x_k(t) - T_j\right\}. \tag{2.1}$$

The activation function of a neuron should be non-decreasing and differentiable [14, 15, 24]. It is various in different models. A common classification of activation functions is given as threshold type (for example (2.2)), non-linear type (for example (2.3)), piece-wise linear type (for example (2.4)) and probabilistic type [14, 15, 23, 24].

$$f(x) = \begin{cases} 1 \; x \geq 0, \\ 0 \; x < 0. \end{cases} \tag{2.2}$$

$$f(x) = \frac{1}{1+e^{-x}}. \tag{2.3}$$

$$f(x) = \begin{cases} 0 \; x \leq 0, \\ cx \; 0 < x \leq x_c, \\ 1 \; x > x_c. \end{cases} \tag{2.4}$$

Through connections between neurons, a neural network is constructed. The existing topology of ANN is various. Based on the connection between neurons, it can be classified as hierarchical structure and interconnection structure [14]. Based on the information transmission direction in the network, ANN can be classified as feed-forward networks and feed-back networks. The most commonly utilized network structure is feed-forward hierarchical networks [5, 15, 19, 29, 32]. A simple feed-forward hierarchical ANN structure is illustrated in Fig. 2.2, which includes two layer of neurons. Since the network is a feed-forward hierarchical network, connection exists only between neurons in different layers, and neurons in the same layer share the same input at each time. It is denoted that the layer which receives inputs as input layer and the layer giving outputs is named as output layer. The inputs of input layer neurons are the combination of the input components and the inputs of the output layer neuron are the combination of the outputs of the input layer neurons. Denote the input combination of the ANN at time i as $\boldsymbol{X}_i$, the output of the ANN at time i as y_i, and the relative input components as $x_{i,k}$ ($k = 1, 2, ..., m$), where $\boldsymbol{X}_i = [x_{i,1}, x_{i,2}, ..., x_{i,m}]$ holds. Denote the weights of input layer neuron j ($j = 1, 2, ..., n$) as w_{kj}. Denote the weights of the output layer neuron as v_j. And denote the bias value of input layer neurons and the output layer neuron as T_j and T_o respectively. The output of the ANN at time i can be calculated by (2.5) [14], where f_{input} and f_{output} are the activation function of input layer neurons and the output layer neuron respectively.

$$y_i = f_{output}\left(\sum_{j=1}^{n} v_j \; f_{input}\left(\sum_{k=1}^{m} w_{kj} x_{i,k} - T_j\right) - T_o\right). \tag{2.5}$$

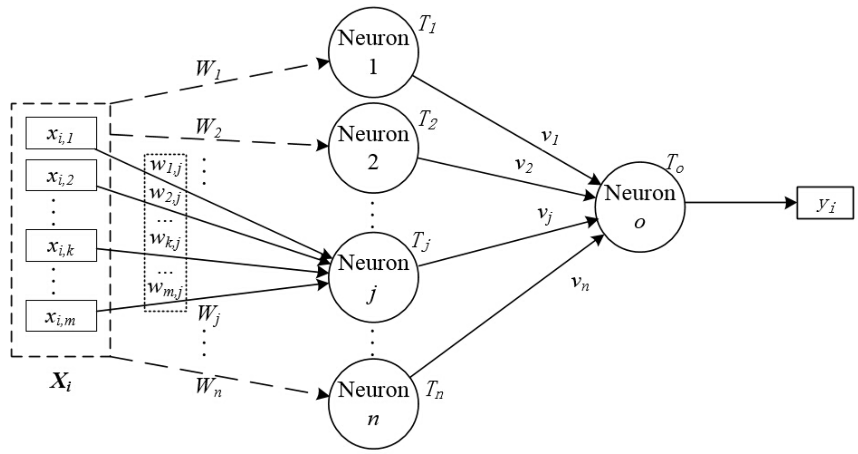


Figure 2.2: A simple example of the network structure of feed-forward hierarchical ANN.

It is observed that the output of an ANN is determined by the weights and biases of all the neurons. After the network given the output at each time, the output should be compared with the target value and the weights and biases in the neurons should be adjusted according to the comparison result. The back-propagation (BP) algorithm is preferred in articles as the learning rule [8, 15, 22, 23, 24, 28, 35] to adjust the parameters. BP algorithm is a generalization of delta rule for multi-layer feedforward ANNs. As mentioned in [34], [37], a multi-layer feedforward ANN is able to fit non-linear function well using BP algorithm, in the case of offering correct input data and enough hidden neurons. In MATLAB® Neural Network Toolbox™, Levenberg-Marquardt training function is used to achieve the BP algorithm to train the feed-forward ANNs [4], since the use of Levenberg-Marquardt algorithm can make the feed-forward neural network be trained faster [13].

2.3 Influence of Weather Parameters in Load Forecasting

Weather parameters have been considered as a significant effect for the load consumption in a power system [3, 12, 27, 30, 36]. The usage of high power electric equipment, such as air condition and heater, is great influenced by the weather parameters for instance temperature, humidity, solar radiation, etc. In [12, 35], the temperature is considered as the most important weather variable which impacts the load consumption. And the correlation between regional temperature, solar radiation, humidity, pressure, wind speed and precipitation and electric load is investigated in [35] respectively. Furthermore, the impact of temperature, wind speed and cloud cover parameters on the regional short-term load prediction ANNs is discussed in [36]. Also, this paper points out that the mean of data checked in different places in the region can more accurately reflect the weather condition of that region, and using the mean data can obtain better load forecast results.

Although numerous existing papers have studied the correlation between weather parameter and load consumption and further utilized weather data such as temperature and humidity in load prediction models, these researches focus commonly on the prediction of regional power consumption. The study about domestic household is still limited, which is investigated in the thesis.

2.4 Peak Predictions

There are numerous of articles about peak prediction existing in the research community. A neural network based regional long term peak load forecasting is presented in [28]. In that paper a multi-layer feed-forward ANN trained by BP learning algorithm is proposed to forecast the monthly peak load. A short term regional peak prediction using multi-layer feed-forward ANN is given in [16].

Furthermore, an hourly peak prediction method is mentioned in [38] using evolutionary algorithms and fuzzy logic approach. In [11], a possible peak forecasting method is given, in which the regional peak prediction problem is mapped into a classical pattern recognition problem. The prediction results of these articles are impressive and worth remembering.

However, the prediction of the occurrence timing of the peaks, which is required by HEMS to manage and schedule loads, is not covered in the researches above. In addition, the peak predictions above are all made for regional power systems, while our focus is on forecasting peaks in domestic households.

Chapter 3

Domestic Electric Load Data Processing and Utilizing for Demand Forecasting

3.1 Load Demand of Domestic Household

Although numerous of ANN models have been developed for regional load consumption, as presented in Chapter 2, the prediction of load consumption of a domestic household is a more difficult case. Compared with regional consumption, the load usage of a household usually fluctuates more largely and rapidly.

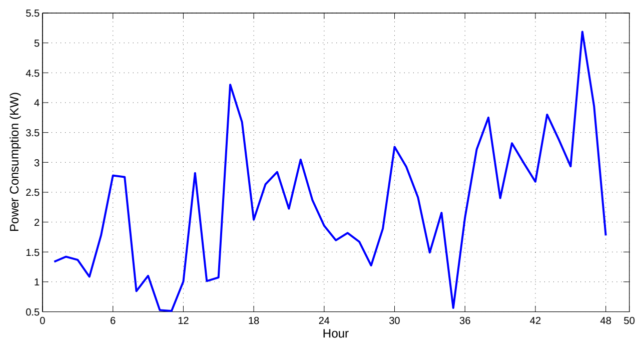
A comparison of hourly domestic household consumption and hourly regional consumption is illustrated in Fig. 3.1, in which the subfigure (a) presents the hourly power consumption line chart of a southern Norwegian domestic household in Dec. 2011. The data obtains from the database used for load and peak prediction of this thesis, which will be introduced later in this chapter. The subfigure (b) shows the power consumption of the entire New England area, America in Jan. 2015 [25]. Both of the subfigures in Fig. 3.1 use hourly data for two continuous workdays. Although due to data limitation, the comparison data are not in the same country at the same days, it is still able to illustrate that comparing with regional consumption, the power usage of one domestic household fluctuates more rapidly.

A straightforward reason is that the activities of a family are easily influenced by random factors. For example, as a dinner dessert, a freshly baked pie requires much more load than an ice cream. However, the impact of household random factors are offset by numerous of family activities when considers a regional consumption, only the common power usage characteristic is shown in the data, e.g., power consumption increases before/at dinner time.

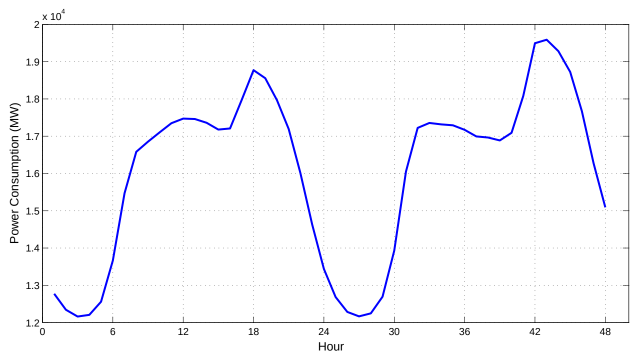
3.2 Power Consumption Data

3.2.1 Raw Load Data

In this thesis, the active and reactive power consumption data of a southern Norwegian domestic household are utilized on developing the load and peak prediction model. The raw load data is the hourly accumulative power consumption value of a southern Norwegian household for about two years. The span of active power accumulation data is from 15:00 5th, Dec. 2011 to 14:00 11th, Nov. 2013, which is illustrated in Fig. 3.2(a). And the span of reactive power accumulation data is from 15:00 6th, Dec. 2011 to 00:00 11th, Dec. 2013, which is presented in Fig. 3.2(b). To simplify expression in the thesis, the ith data in an identified database is called the data at hour No. i. For example, in the database of accumulative active power raw data, the data of hour No. 1 corresponds the accumulative active power at 15:00 5th, Dec. 2011 . It should be noticed that the data of hour No. i in different databases usually represents data for different time points.



(a)



(b)

Figure 3.1: A comparison of hourly domestic household consumption and hourly regional consumption: (a) domestic household consumption and (b) regional consumption.

3.2.2 Load Data Processing

Reading Error

In Fig. 3.2, it is able to be observed that two types of data errors exist in the active and reactive accumulative power consumption raw data. The first type of error is named as reading error, which can be detected if the accumulative power consumption reading of one hour is less than its previous hour. For example, in the

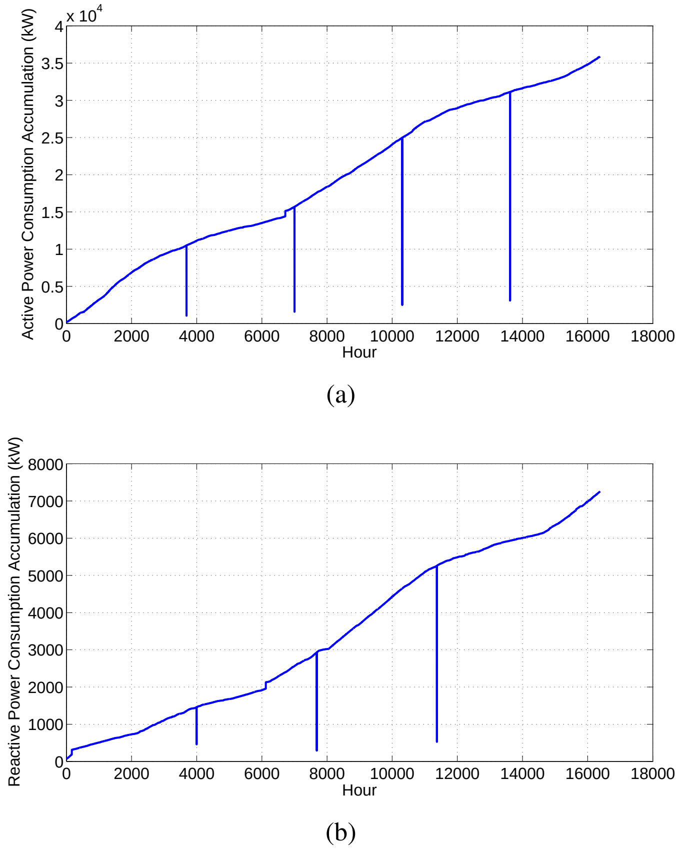


Figure 3.2: (a) Raw accumulative active power consumption data; (b) Raw accumulative reactive power consumption data.

active power raw data, the reading of hour "07/05/2012 04:00:00" (at Fig. 3.2(a) hour No. 3686) is 1049.096 kW. However, the readings of hour "07/05/2012 03:00:00" and "07/05/2012 05:00:00" are 10491.581 kW and 10496.662 kW respectively. Thus the reading of hour "07/05/2012 04:00:00" is detected as a reading error. It is because the slope of an accumulative function cannot less than 0. As illustrated in Fig. 3.2, there are 4 and 3 reading errors existing in the raw active and reactive accumulative power consumption data respectively. To remove reading errors, it is selected to replace the erroneous reading by the mean of its adjacent two readings. In th previous example, the erroneous reading is replaced by

10494.1215 kW ($\frac{1}{2}(10491.581 + 10496.662)$). A necessary assumption of using this approach is that there is no adjacent errors in the raw data.

Gap-data Error

The other type of error is named as gap-data error. This type of error can be detected if the date readings of adjacent data units are not continuous. For example, in the active power raw data, the date reading of hour No. 6719 is "10/09/2012 13:00:00" and the date reading of hour No. 6720 is "27/09/2012 13:00:00", which means that 17 days' load consumption data is missing. This gap-data error is also can be observed in Fig. 3.2(a), in which a data stair exists between hour No.6719 and hour No. 6720. Moreover, similar to the reading errors, the gap-data errors in the data of two power types arise randomly. The gap-data errors in the accumulative active and reactive raw data are illustrated in Appendix A.

The gap-data error is hard to repair since the hourly power usage profile in the gap period is a black-box. Any artificial processes may cause negative impact for the correct training of the prediction model. Nevertheless, the influence of the gap-data error can be avoided by using hourly power consumption data to develop forecast model. Then the short term missing data in the training set does not have considerable influence to the prediction model output. By using this strategy the raw accumulative power consumption data requires to be reformed as hourly power consumption data, in which the missing portions in gap periods are ignored.

Special Holidays

In special holidays like Easter and Constitution Day, the activities of people are usually quite different from common workdays and weekends. Traditional customs and special leisure ways may lead a total different power consumption pattern of a household. For example, if a family goes abroad to travel in the Christmas holiday, their household power consumption of that period should keep in a relative low level. However, if a family stays in their household during the holiday, the power consumption should be much larger than ordinary days. There are two approaches to deal with the load samples in special holidays. One approach is adding a mark column in the input vector to distinguish the samples for holidays with samples for ordinary days. The other approach is removing those special holiday samples from the available database. Both two approaches are utilized in the thesis by different models, based on the different requirements of the models.

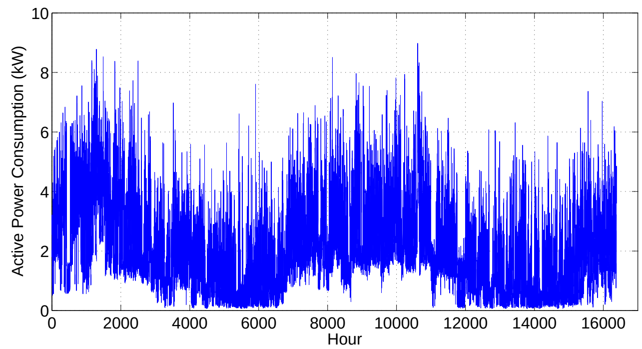
The hourly active and reactive power consumption data of the southern Norwegian domestic household is illustrated in Fig. 3.3(a) and (b) respectively, in which the reading errors are processed and gap parts are ignored. Moreover, active and reactive power consumption of the hour No. i in the database are denoted as $C_{act,i}$ and $C_{re,i}$ respectively. Since the same model (structure and configuration) are utilized to forecast active and reactive power consumption, in unambiguous cases, we utilize C_i to express the power usage of the hour No. i.

3.3 Weather Parameter Data

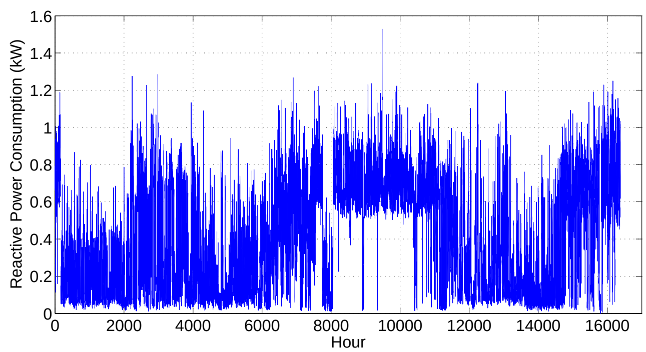
Weather parameter data is utilized in the thesis to investigate the impact of it on the load prediction model for a domestic household.

A house exchanges heat with its surrounding space all the time until reaching balance. In order to keep the inside temperature in a comfortable range, a lot of power should be consumed by air condition/heater. Therefore, the air temperature is considered as an important weather parameter which can influence the power usage. Moreover, house is able to obtain energy from solar radiation. Thus, solar radiation is considered as another weather parameter in the thesis.

The data we utilized on training and testing the forecast model are actual weather values. However, it is impossible to obtain future weather situation in



(a)



(b)

Figure 3.3: (a) Hourly active power consumption data; (b) Hourly reactive power consumption data.

practice, only weather forecast could be used. Therefore, the performance of prediction models using weather parameter data in the thesis could be seen as an upper bound of the practice situation.

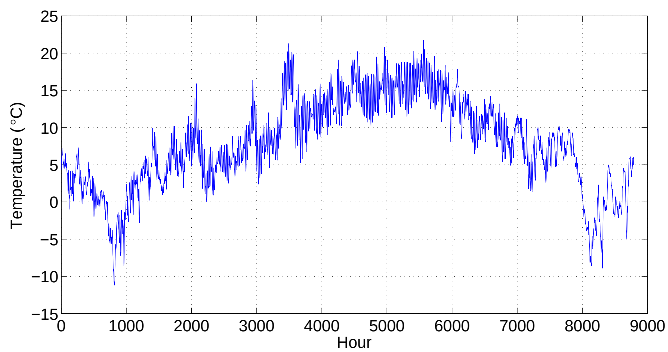
The regional weather data of the southern Norwegian household is offered by website $yr.no$, which is supported by the Norwegian Meteorological Institute. The raw weather data offers values of solar radiation and meteorology temperatures in period from 00:11, 1st, Jan. 2007 to 23:56, 7th, Apr. 2013. For meteorology temperatures, the regional air temperatures at 2 m are given for every 15 min. For solar radiation, values of the global solar irradiance (GSI) are given by the same timings with temperature values. The temperature and solar radiation detection time are the 11th, 26th, 41st, and 56th minute of each hour.

Since the detection interval of hourly power consumption and raw weather data are different, in order to utilize the raw weather data into the prediction model, it is necessary to integrate the time interval of data firstly. Here the weather samples (include temperature and GSI value) detected at 26th minute of each hour are utilized to represent the weather of that hour. In this way, the power consumption data and weather data has the same time interval which is one hour. As an example, the hourly regional temperature and GSI data in 2012 are illustrated in Fig. 3.4(a) and (b) respectively. The x-axis, “Hour”, orders the hour from timing 00:00, 1st, Jan. 2012. Since 2012 is a leap year, the total number of sample in each figure is 8784 ($24 * 366$).

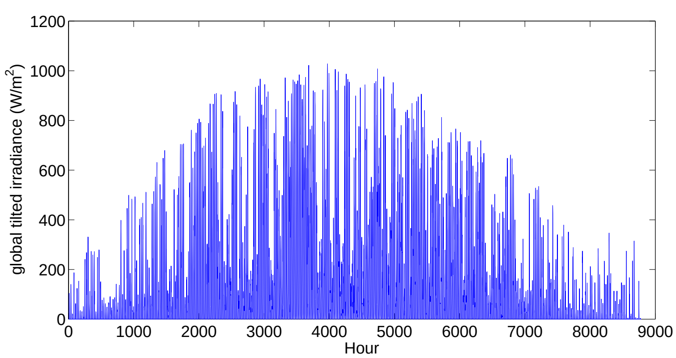
3.4 Peak Prediction of Domestic Household

The results of forecasts for the load demand of the incoming 24 hours are used to predict the peak. Based on different peak determination strategies, the power consumption peak in a short period (24 hours) could be different. In the thesis, the highest peak strategy is utilized on determining peaks.

The highest peak in a period represents the hour with highest power consumption value. The highest power consumption leads the biggest challenge for the local power distribution network. Highest peak forecast is required by HEMS to manage and schedule controllable non-critical load beforehand to smooth the peak.



(a)



(b)

Figure 3.4: (a) Hourly temperature data in 2012; (b) Hourly GSI data in 2012.

Chapter 4

Forecasting Model Building and Synthesis

In this chapter, prediction models are built to forecast the load demand in the incoming 1st - 24th hour. In the thesis, we name the prediction which forecasts the hourly load demand t hours later as “demand prediction model t”, where $t = 1, 2, 3, ...24$. As an example, the demand prediction model 1 represents the one hour demand prediction, which is in charge of forecasting the demand in the next hour.

In order to describe the load prediction mathematically, it is denoted the load demand of t hours after the hour No. i as $L_{i,t}$, where $t = 1, 2, ..., 24$. Since the forecast target of demand prediction model t at hour No. i equals to the actual power consumption of hour No. $i + t$, $L_{i,t} = C_{i+t}$ always holds. And the corresponding demand prediction result is denoted as $\tilde{L}_{i,t}$. For each i, $\tilde{L}_{i,t}$ is predicted by demand prediction model t.

4.1 Input Data for Demand Prediction

In order to compare the demand prediction results between the models with and without considering weather parameters, it is necessary to use the overlapping portion of the load and weather data to train and test the models with and without considering weather parameters. The available period is from 5th, Dec. 2011 to 7th, Apr. 2013.

To make the input vectors using only hourly power usage data (for both active and reactive models), two classes of columns, time columns and historical powr usage columns, are considered. And for prediction in consideration of weather parameters, external weather columns are required.

4.1.1 Time Columns

The time information is required to transform from a string data towards numerical data before imported into the model. For this case, the string time column is transformed into three numerical columns, which are “month”, “hour”, and “day of week”. For models including special holiday samples, one more column, “workday/holiday”, is required.

For hour No. i, the time columns are denoted as $month(i)$, $hour(i)$, $week(i)$, $holidaymark(i)$. The range of function $month(i)$, $hour(i)$, $week(i)$ is $[1, 12]$, $[1, 31]$, and $[1, 7]$ respectively. For $week(i)$, 1 is for Sunday and 7 is for Saturday. The range of function $holidaymark(i)$ is 0 and 1. 0 is for samples of workday, and 1 is for samples of holidays.

4.1.2 Historical Power Consumption Columns (Load Columns)

Through trajectories, the data of "three latest consumption", "the 24-hour lag consumption" and "average consumption of latest 24 hours" are used for the demand prediction. For demand prediction model t and hour No. i, The "three latest consumption" can be expressed as C_i, C_{i-1}, and C_{i-2}. The "24-hour lag consumption" can be presented as C_{i+t-24}. And the "average consumption of latest 24 hours" can be indicated as $\frac{1}{24}\sum_{k=0}^{23} C_{i-k}$.

The combination of time and historical consumption columns forms the input data of the prediction models. $\boldsymbol{X}_i$ is denoted by an arrow vector as the input at hour No. i. Altogether, there are 8 elements in each input vector. As an example, several input vectors around $\boldsymbol{X}_i$ for demand prediction model t are shown in Tab. 4.1.

4.1.3 Weather Columns

For the demand prediction models considering weather parameters, the actual data of temperature and GSI of the hour No. $(i+t)$ are added into each input vectors. Thus each $\boldsymbol{X}_i$ in the prediction considering only temperature/solar radiation has 9 elements, and in the prediction considering both temperature and solar radiation has 10 elements.

4.2 Demand Prediction Using ANN Model

As discussed in Chapter 2, two-layer feed-forward ANN is utilized in the thesis to forecast the power demand of both active and reactive loads. The architecture

Table 4.1: Input vectors around $\boldsymbol{X}_i$ for demand prediction model t

Month	Hour	Day of week	Workday/ Holiday	Three latest			24-hour lag	24-hour average
...								
$month(i+t-1)$	$hour(i+t-1)$	$week(i+t-1)$	$holidaymark(i-1)$	C_{i-1}	C_{i-2}	C_{i-3}	C_{i+t-25}	$\frac{1}{24}\sum_{k=0}^{23} C_{i-k-1}$
$month(i+t)$	$hour(i+t)$	$week(i+t)$	$holidaymark(i)$	C_i	C_{i-1}	C_{i-2}	C_{i+t-24}	$\frac{1}{24}\sum_{k=0}^{23} C_{i-k}$
$month(i+t+1)$	$hour(i+t+1)$	$week(i+t+1)$	$holidaymark(i+1)$	C_{i+1}	C_i	C_{i-1}	C_{i+t-23}	$\frac{1}{24}\sum_{k=0}^{23} C_{i-k+1}$
...								

of the feed-forward ANN is illustrated in Fig. 4.1, which includes an input layer with n neurons and an output layer with a single neuron.

After an input vector $\boldsymbol{X}_i$ is sent into the input layer neurons, each input layer neuron calculates the weighted summation. The number of element in an input vector is denoted as m ($m = 8, 9$ or 10 holds). Denote the input components of $\boldsymbol{X}_i$ as $x_{i,1}$, $x_{i,2}$, ..., $x_{i,m}$. And denote that in the input-layer-neuron j, the weight for $x_{i,k}$ as w_{kj}. Then the transfer function of the jth neuron could be expressed as $t_j = \sum_{k=1}^{m} w_{kj}x_{i,k} - T_j$, where $j = 1, 2, 3...n$ and T_j is the threshold of that neuron. If we denote $x_{i,0} = -1$ and $w_{0j} = T_j$, then the transfer function could be expressed as $t_j = \sum_{k=0}^{m} w_{kj}x_{i,k} = \boldsymbol{X}\boldsymbol{W}_j$, where $\boldsymbol{W}_j$ is the column vector for the weights of the input-layer-neuron j. Hyperbolic tangent sigmoid function, $o_j = 2[1 + e^{-2t_j}]^{-1} - 1$, which is good for training speed, is utilized as the activation function for all input layer neurons [4, 8].

In the output layer, outputs of the input layer neurons are sent to the single output-layer-neuron as its input. Denote the weights of the output-layer-neuron for input-layer-neuron j as v_j, then the weight summation is given as $t_o = \sum_{j=0}^{n} v_j o_j$, where $o_0 = -1$ and $v_0 = T_o$. Since activation function of the output-layer-neuron is configured as linear function ($a_o = t_o$) [8, 35], as mentioned in Fig. 4.1, $a_o = \sum_{j=0}^{n} v_j o_j$ holds.

Therefore, based on (2.5) and network configuration discussed above, the forecast result of input vector $\boldsymbol{X}_i$ can be calculated by (4.1).

$$\tilde{L}_{i,t} = \sum_{j=1}^{n} \frac{(1 - e^{-2(\sum_{k=1}^{m} w_{kj}x_{i,k} - T_j)})v_j}{1 + e^{-2(\sum_{k=1}^{m} w_{kj}x_{i,k} - T_j)}} - T_o. \tag{4.1}$$

The weights and thresholds of neurons are updated in each training using Levenberg-Marquardt algorithm [4, 13].

4.3 Peak Finding Using Load Predictions

Based on the load demand forecast results of the domestic household for the incoming 1-24 hours, it is able to predict the demand peak in the next 24 hours following the highest peak strategy. Before we develop the prediction, the mathematical expression of power demand peak is required.

$P_{i,n}$ is denoted as the peaks occurring in next 24 hours after the hour No. i, where n is the serial number of a peak in the period, i.e., $n \in$ **Z**, $n \in [1, 24]$ holds. Since for the highest peak strategy, there is one and only one peak for each i, which means that $n \equiv 1$, it is able to simplify the mathematical expression of highest peak as P_i in this thesis.

As a power consumption peak, its value and occurrence timing are important. Denote $P_i = (L_{peak}, t)$, where L_{peak} is the power demand of the peak, and $t \in [1, 24]$ is the number of hours that the peak will occur after hour No. i. Based on the 24 demand prediction results, the one with greatest value is selected to be the highest peak, which can be expressed by (4.2).

$$P_i = (max\{L_{i+k}, k \in [1, 24]\}, t_{the\ peak})\,. \tag{4.2}$$

4.4 Assessment of Predictions

4.4.1 Assessment of Load Forecasting

Two evaluation standards are introduced in the thesis to evaluate the load demand prediction result.

Regression value $R \in [-1, 1]$ measures the correlation between the power consumption prediction result and the target power usage, which is able to be calculated by (4.3) [4, 8].

R with a larger positive value means that a closer positive correlation exists between the prediction results $\tilde{L}_{test,t}$ and targets $L_{test,t}$. R with a larger negative value means a more negative correlation relationship. And $R = 0$ means a random relationship. For our prediction models, the larger R the better.

The other evaluation standard is the mean absolutely percentage error (MAPE) [26, 31], which is denoted as M. M of the demand prediction model t can be calculated by (4.4). Literally, MAPE is the mean value of absolute errors in the testing period. That is to say that a smaller M represents a better demand prediction model.

$$R = \frac{cov(L_{test,t}, \tilde{L}_{test,t})}{\sigma_{L_{test,t}} \sigma_{\tilde{L}_{test,t}}}. \tag{4.3}$$

$$M = mean \left\{ \left| \frac{\tilde{L}_{test,t} - L_{test,t}}{L_{test,t}} \right| \right\} * 100. \tag{4.4}$$

4.4.2 Evaluation of Peak Prediction

To evaluate peak prediction, it is defined that if the error of peak occurrence timing is not greater than 2 hours, the forecast is considered "correct" predicting the peak occurrence timing. And in the set of "correct" predicted peaks, MAPE M is calculated to illustrate the performance of peak value.

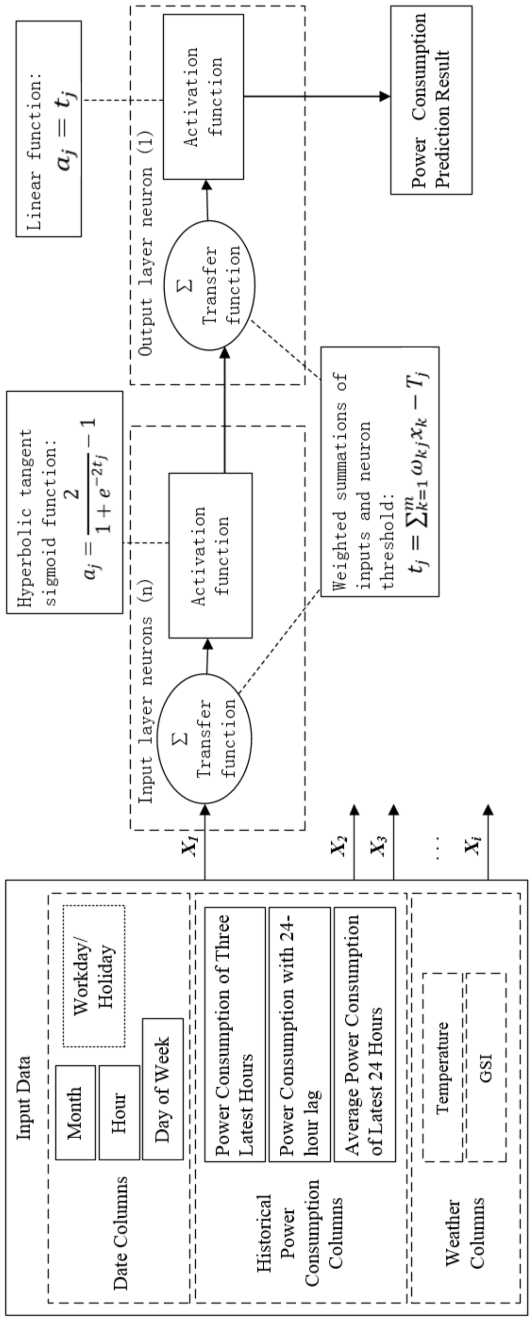


Figure 4.1: The structure of demand prediction ANN model.

Chapter 5

Performance Analysis of Domestic Load Predictions

5.1 One Hour Demand Prediction Using Only Load Data

Based on the ANN structure and input combination given above, in this section, one hour demand prediction (demand prediction model 1) is trained and tested using load data with/without error processing. The performance of prediction results of model using with and without processed data is compared. The results in this section has been published in [2].

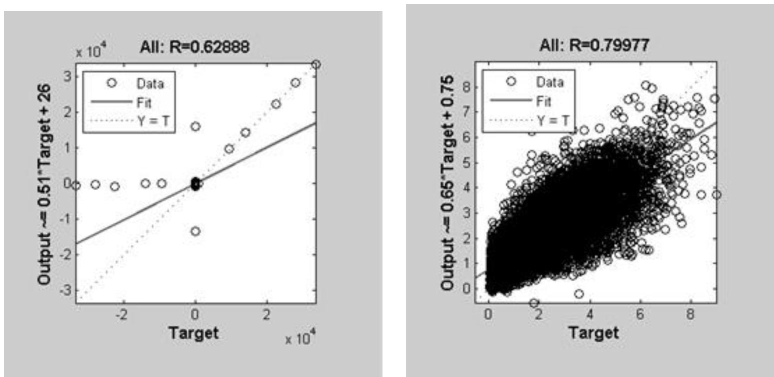
5.1.1 ANN Configuration

The input vectors $\boldsymbol{X}_i$ (with/without error) for the model are built up by time columns and load columns as presented in Tab. 4.1. The setting of the number of neuron in the input layer, n, is a compromise of accuracy and computation amount. In order to achieve better performance, a larger n is required. However, the increase of n leads much more computation as well as operation time. In this section, 100 neurons are used by the input layer of the ANN, since $n = 10$ is not enough to obtain precise predictions and $n = 1000$ is too much to be calculated fast. All the available data are divided using random uniform sampling into three sets, 70% of samples are used for training, 15% for the validation and the last 15% for the testing.

5.1.2 Performance of One Hour Demand Prediction with/without Data Processing

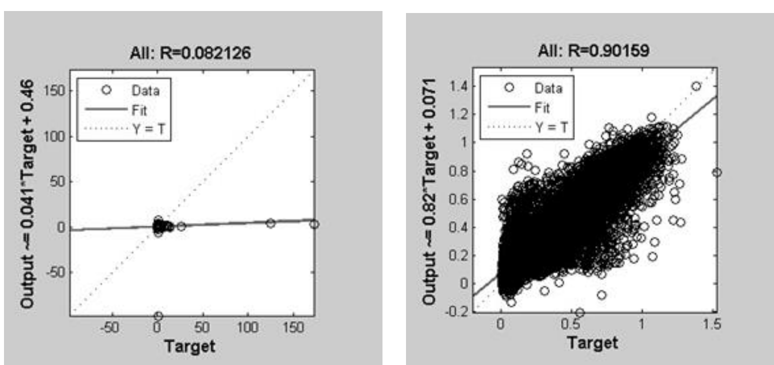
It can be observed that data processing upgrades the prediction performance. For active load prediction, the data processing procedure improves the regression value R from $R = 0.62888$ to $R = 0.79977$, as illustrated in Fig. 5.1. In the figure, each point includes the information of a forecast result (y-axis) and its corresponding actual consumption (x-axis). A point on the line $y = x$ represents that the forecast equals to the target, which means a prefect forecast. And for reactive load prediction, the improvement achieved from data processing is more significant. As presented in Fig. 5.2, the reactive load prediction using unprocessed data has $R = 0.082126$, and the regression value of prediction using processed data achieves $R = 0.90159$.

It is observed from the evaluation results that both active and reactive predictions are not precise enough to be used in HEMS. In order to achieve more accurate prediction, it may be required several influential parameters to be introduced as input components. As a candidate, the impact of weather parameter on the one hour demand prediction is discussed in the next section.



(a) (b)

Figure 5.1: (b) Regression value of all samples for active power consumption prediction and target without data processing; (a) Regression value of all samples for active power consumption prediction and target with data processing.



(a) (b)

Figure 5.2: (b) Regression value of all samples for reactive power consumption prediction and target without data processing; (a) Regression value of all samples for reactive power consumption prediction and target with data processing.

5.2 Influence of Weather Parameters on One Hour Demand Prediction

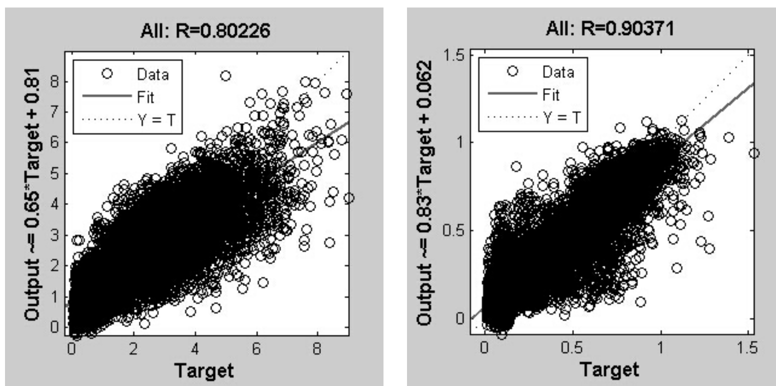
As mentioned in Chapter 2, weather parameters, such as temperature and solar radiation, are considered as important factors which are able to influence the load consumption in regional power system. In this section, the impact of weather parameters on the one hour active and reactive demand prediction for domestic household is investigated. Temperature and solar radiation data are introduced into the one hour demand prediction, and the efficiency of predictions are evaluated and compared. The results in this section has been published in [1].

5.2.1 ANN Configuration

The input vectors $\boldsymbol{X}_i$ used this section consist of time, load and weather columns. In order to distinguish the impact of temperature data and solar radiation (GSI) data, predictions considering temperature are simulated firstly. Then, the predictions with both temperature and solar radiation data are trained and tested. The prediction results are compared with the results of predictions using only load data, which are presented in Section 5.1. In order to make the comparison meaningful, the network configuration of this section should follow the settings in Section 5.1, i.e., input layer neuron is configured as 100. And 70% of the samples are used for training, 15% for the validation and 15% for the testing.

5.2.2 Performance of One Hour Demand Prediction Considering the Influence of Weather Data

For the predictions consider only temperature (time columns + load columns + temperature column), the regression value R for active and reactive load prediction is $R_{act} = 0.80226$ and $R_{react} = 0.90371$ respectively, which is presented in Fig. 5.3. The MAPE for active and reactive prediction is $M_{act} = 49.4703$ and $M_{react} = 48.6966$ respectively. After comparing the evaluation results with the results in Section 5.1, it is observed that for the active prediction, the introduction of temperature data leads a little improvement on the regression but upgrades a lot on reducing MAPE. However, for the reactive prediction, temperature data does not have significant impact for the prediction.

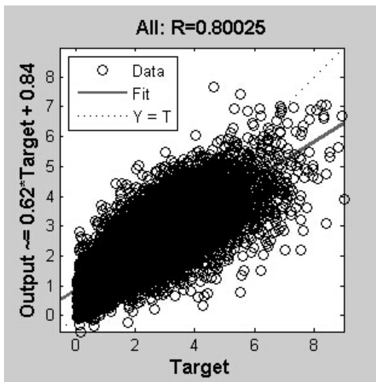


(a) (b)

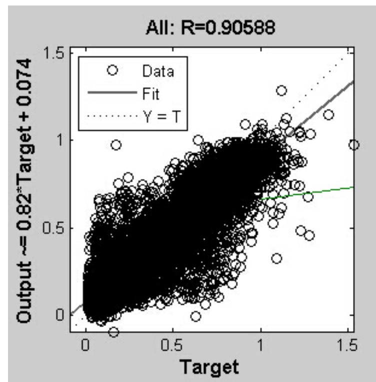
Figure 5.3: (a) Regression of all samples for active prediction considering temperature data; (b) Regression of all samples for reactive prediction considering temperature data.

The different impact of temperature data on active and reactive load prediction might come from the different usage mode of active and reactive loads. Actually, comparing with reactive power, active power is more sensitive to temperature. When the temperature becomes cold/sticky, heaters/air-conditions are commonly used in a household, which directly results the increasing of active power consumption. Thus it is reasonable that the temperature data has benefit for the active load prediction. However, since that there is no direct connection between temperature and reactive power consumption, temperature data is difficult to have significant impact on reactive power prediction.

For the active and reactive predictions considering both temperature and GSI (time columns + load columns + temperature column + GSI column), regression value R is $R_{act} = 0.80025$ and $R_{react} = 0.90588$ respectively, and the MAPE M is $M_{act} = 46.9245$ and $M_{react} = 48.3425$ respectively. The regression results are presented in Fig. 5.4.



(a)



(b)

Figure 5.4: (a) Regression of all samples for active prediction considering both temperature and GSI data; (b) Regression of all samples for reactive prediction considering both temperature and GSI data.

Comparing with temperature data, it can be noticed that there is no significant performance change by introducing solar radiation (GSI) data into the prediction model. One possible reason is that the weather data is not taken from the same location as the load data. The obtained temperature and GSI data is the comprehensive record of the city where the household is located. The temperature variation within a city may not be significant. Thus the temperature data is able to impact the prediction performance in certain level. However, the solar radiation is a different story. According to the locations, the solar radiation of two houses could be widely different with each other. Thus the solar radiation data for a city maybe do not have close relativity with the household considered in the thesis. Therefore, in order to achieve better forecasting performance, more precise weather data for the household surrounding environment might be required.

5.3 Use of Load Prediction for Peak Finding and Demand-side Management

In Section 5.1 and 5.2, one hour demand prediction (demand prediction model 1) is built and the influence of weather parameters is analysed. However, in order to achieve demand-side management, only prediction for the next hour is not enough. Finding the potential peak in next 24 hours is required by HEMS to managed the loads in the incoming day efficiently. Therefore, the prediction of the load demand in each hour of the incoming day (demand prediction model 1-24) is required. In addition, considering the practical application scenario of the peak finding approach in the HEMS, the configuration of load prediction ANN should have some adjustments in this section.

5.3.1 ANN Configuration

As mentioned in Chapter 3, the load demand in special holidays is hard to be predicted. In previous sections, one additional input component, "workday/holiday", is utilized to distinguish rest days from workdays. In order to forecast the peak precisely, in this section, the load consumption data in special holidays are removed from the available database. Correspondingly, the input component "workday/holiday" is removed from the input combination. It should be noticed that the samples for weekend are kept in the available data, and they are able to be distinguished from sample for weekday by input column "Day of week". Through the steps of taking the overlapping portion of load and weather data, processing the raw data into the input form, and removing the data related with gap-data errors and special holidays, the list of input vectors is obtained.

In addition, the training of the load prediction ANN in the HEMS terminal is a challenge. As a terminal located in household, the operation capability of HEMS might not be able to train the network good enough. Therefore, the training and validation process should be done when the HEMS is installed. In order to simulate this scenario, the available input vectors are divided into three sets by blocks. The first 70% of data is used for training the ANN. Next 15% of data is used on validation and the last 15% is used for testing the performance of the ANN.

It should be noticed that in order to utilize the load prediction result in peak prediction, the date of the input vectors in the test set should be continuous. Otherwise the peak forecast results will have some troubles. The available data for active predictions is in a 1.3 year span, and the available data for reactive predictions is in a 1.2 year span, approximately.

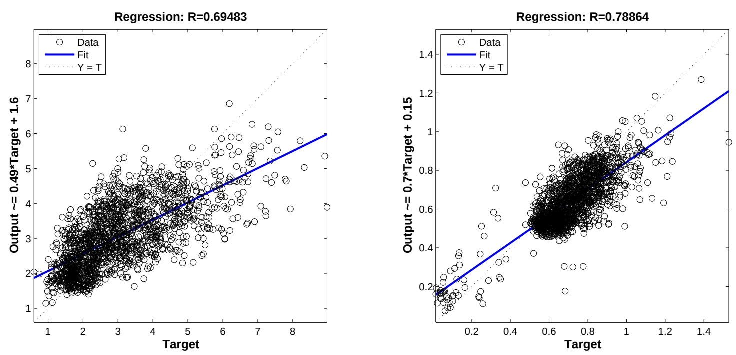
The ANNs for demand prediction model $1 - 24$ are established based on the configurations discussed in Chapter 4. And the number of hidden layer neuron, n, is configured as 10. Although the size of the hidden layer is insufficient to obtain a precise enough prediction, the relatively small number of neuron guarantees the operation time of the training and testing of the 24 prediction models could be controlled within an acceptable range.

In order to find the peak in the incoming 24 hours, the correlation between the prediction result and the target is important. A greater regression value means that the prediction fits the load usage profile better. So the peak forecast using load predictions with greater R should performs better. Thus, in this section, the regression value of the demand prediction models are presented.

5.3.2 Performance of Load Predictions for the Incoming 24 Hours Using Only Load Consumption Data

After setting the models, it is able to investigate the performance of load prediction models using only load consumption data. As mentioned, 8 components (3 for time, 5 for load) are included in each input vector.

The regressions of demand prediction model 1 (train and test for once) for active and reactive power are presented in Fig. 5.5(a)-(b). Although there are other configurations difference, it is still able be observed that the performance difference between the prediction with 100 input neurons (Figs.5.1(b) and 5.2(b)) and the prediction with 10 input neurons (Fig. 5.5).

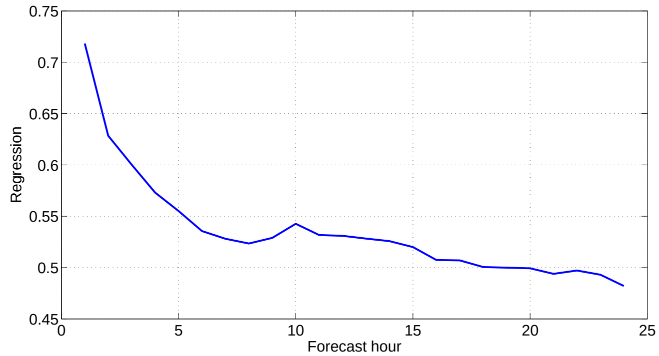


(a) (b)

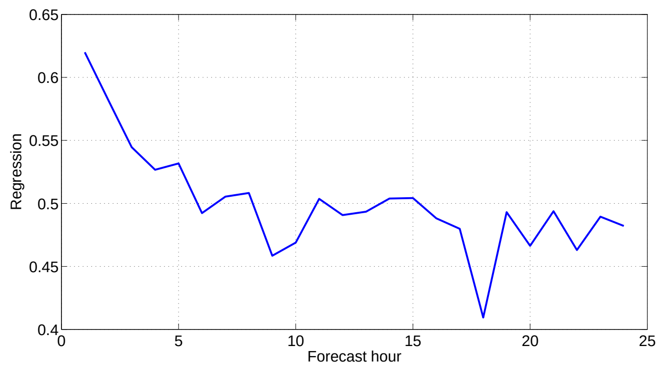
Figure 5.5: (a) Regression of demand prediction model 1 for active power using only load data; (b) Regression of demand prediction model 1 for reactive power using only load data.

The demand prediction model t ($t = 1, 2, ..., 24$) is trained and tested for 15 times. The mean regression value R of the tests for the 24 models are illustrated in Fig. 5.6(a) (for active power) -(b) (for reactive power). It is observed that the value of R decreases when t increases, which means that the forecast result and actual consumption become less correlation in the model which is in charge of forecasting a farther hour.

Furthermore, it is observed a tendency that the regression of load prediction models decreases fast in the predictions of closer hours (forecast hour 1-5 approximately). And the decline rate of regression decreases in the predictions of farther hours (forecast hour 6-24 approximately). However, due to the number of neuron configured in the input layer and the operation time are limited, the curve of mean regression is not stable enough. The relationship between forecast hour and the



(a)



(b)

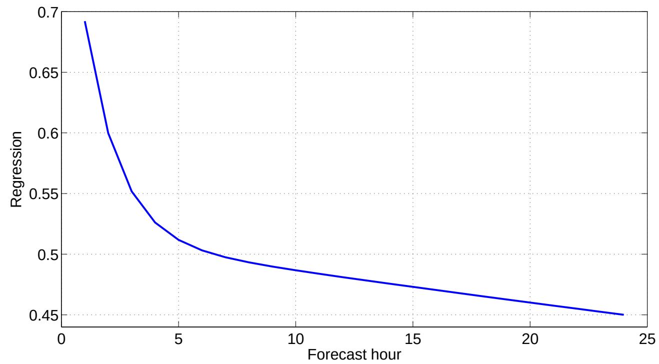
Figure 5.6: (a) Mean regression for active load predictions using only load data; (b) Mean regression for reactive load predictions using only load data.

regression value R is hard to be illustrated clearly. Therefore, in order to reflect the characteristic of the data, MATLAB® Curve Fitting Tool™is utilized here to fit the regression line chart (active and reactive) as a continuous function. There are numerous types of curve fitting method with different critical norms. Which fitting method reflects the characteristic of the relationship between regression values and forecast hour better is still an open question. When the optimal fitting method is found, the fitting result might be more insightful, which is able to be used on improving the algorithm.

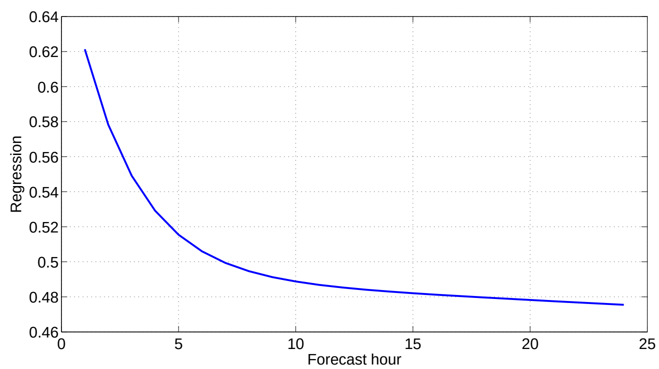
Here, based on the profile of the curves, two terms exponential function ($f(x) = a * e^{bx} + c * e^{dx}$) is selected as the target function format, where b and d should be negative. The curves (active and reactive) after fitting are presented in Fig. 5.7. It is able to illustrate clearly that the decline rate of regression value decreases rapidly in the figure.

Since same ANN structure and same input components are utilized for the predictions, the change of the decline rate might reflect the change of the relativity between the forecast hour and existing data (input components). The shorter interval between existing data and forecast hour, the larger impact the existing data has, the better performance the prediction achieves. However, when forecast the load in farther future, the impact of existing data reduces. For this reason, the decline rate decreases rapidly in the predictions for close hours.

Moreover, in Fig. 5.7, it is observed that the regression value of predictions for farther hours (forecast hour 6-24) presents a kind of stability. The values of R keep in a certain low level and decrease slowly. A possible reason of this phenomenon is that there is a kind of short term feature existing in the power consumption of a household. The load usage in an hour implies the power consumption feature of the coming few hours. Hence, without the short term load consumption feature information offered by existing data (when the forecast hour > 5), the predictions for farther hours provides a kind of generality load forecast. Since the latest three hours' load consumption is included in the input combination of each prediction, the range of short term load consumption feature might be in $3 - 5$ hours. For this reason, the performance declines rapidly for demand prediction model $1 - 5$ and for the predictions for farther hours perform similar.



(a)



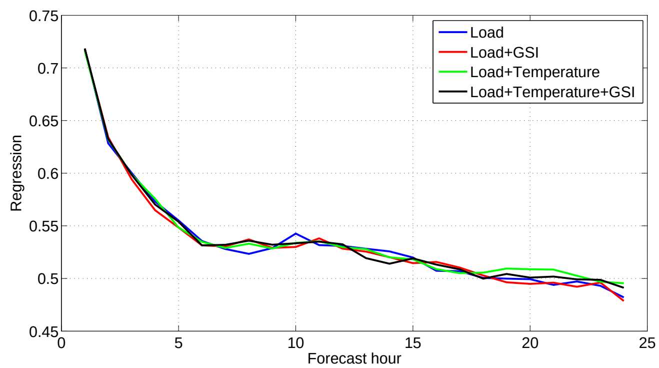
(b)

Figure 5.7: (a) Fitted mean regression for active load predictions using only load data; (b) Fitted mean regression for reactive load predictions using only load data.

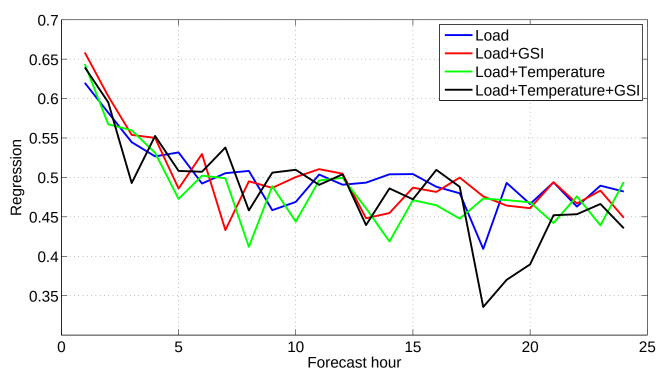
5.3.3 Performance of Load Predictions Considering Weather Data

Based on the same ANN structure and configuration, the load prediction models using load consumption data and weather data (1 or 2 external input components) are built. The model performance of network using temperature and/or solar radiation data is compared with the one using only load consumption data in this subsection.

The 24 load prediction models using load consumption together with only temperature, with only solar radiation, and with both temperature and solar radiation are trained and tested separately for 15 times. The mean regression of test results for active and reactive power are illustrated in Fig. 5.8(a) and (b) respectively.



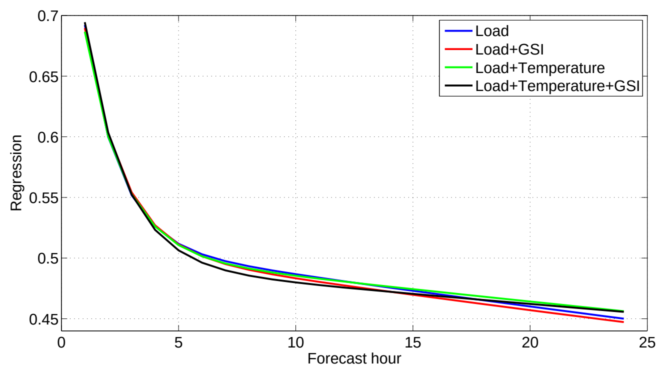
(a)



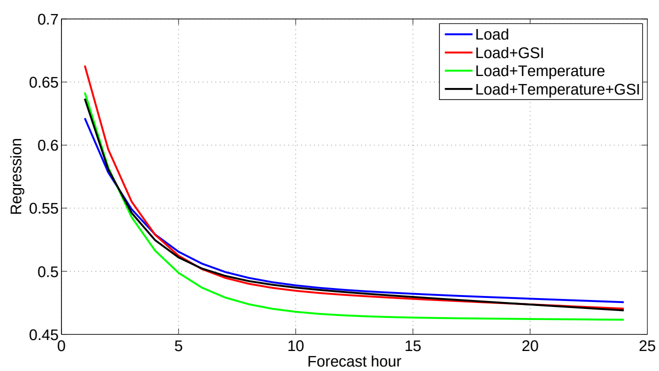
(b)

Figure 5.8: (a) Mean regressions for active predictions; (b) Mean regressions for reactive predictions.

It is observed that the performance of predictions considering weather parameters keeps the characteristics of predictions using only load data which are discussed in the previous subsection. Similarly, to illustrate the relationship between curves, the MATLAB® Curve Fitting Tool™is utilized again to fit the regression values on each curve. Similar to the prediction with only load data, two terms exponential function ($f(x) = a * e^{bx} + c * e^{dx}$) is selected as the form of the fitting function. The curves after function fitting are presented in Fig. 5.9.



(a)



(b)

Figure 5.9: (a) Fitted mean regressions for active predictions; (b) Fitted mean regressions for reactive predictions.

It is observed in Fig. 5.9(a) that the regression of active power predictions for relative close hours (1-3 hour) has no significant difference. However, when the ANN model is utilized on forecasting relative farther hours (> 15 hour), it is noticed that the performance of models using temperature data as a weather parameter (“load + temperature” and “load + temperature + GSI”) becomes better than the others. It means that by introducing temperature data into the prediction, regression value has a little advantage comparing with the predictions without temperature data. In addition, it can be observed that the models using GSI data (“load + GSI” and “load+ temperature + GSI”) as inputs perform worse than the models which do not using GSI data (“load” and “load+ temperature” respectively), meaning that introducing GSI data as an input component does not have positive impact for active power predictions.

However, in reactive power predictions, the performance is quite different. The prediction using only load data performs better on the predictions for farther hours. And predictions using temperature data is worse than the predictions without using, which means that temperature data does not have benefit for reactive power prediction. It can be observed that the prediction result here conforms the analysis about the possible reason of the different impact of temperature data on active and reactive load prediction in Section 5.2. It is that comparing with reactive load, active load is more sensitive to temperature. Furthermore, it is also can be observed that the solar radiation data does not have significant impact for active and reactive predictions too.

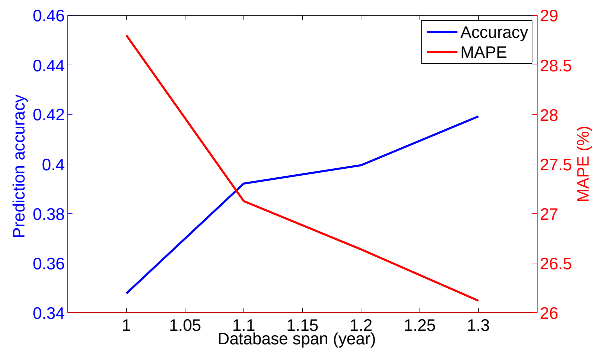
5.3.4 Performance of Peak Finding

The highest peaks in the next 24 hours of the household are predicted based on the demand prediction model $1 - 24$ considering temperature and GSI. In order to avoid the situation that two highest peaks with same load consumption exist in the prediction, two measures are taken.

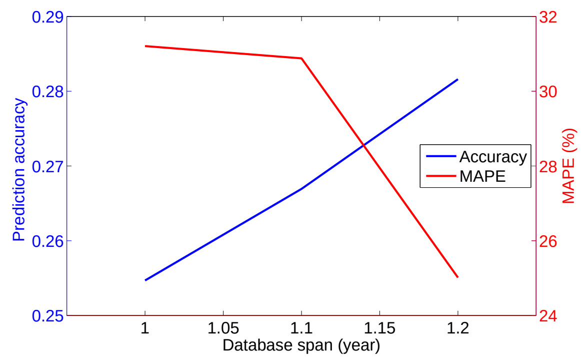
1. The forecast and target load value are set as type “double” in MATLAB® codes;

2. If more than one highest peaks (with equal “double” value) exist, the first one will be claimed as the peak (since in practice, the closer peak is more impor-

tant for the HEMS).



(a)



(b)

Figure 5.10: (a) Accuracy and MAPE for active peak prediction with different database spans; (b) Accuracy and MAPE for reactive peak prediction with different database spans.

It is observed that regression value of load prediction decreases with increase of forecast hour. Based on current knowledge, when the forecasting a farther hour, it is natural that the forecast accuracy decreases. Thereby the peak forecast relying on load prediction results inherits the prediction errors in the load predictions. According to the peak evaluation method introduced in Chapter 4, the forecast “correct” percentage (accuracy) and MAPE of active and reactive peak predictions are calculated. For the active peak prediction, the accuracy is $P_{act} = 41.985\%$, and the MAPE of peak load prediction is $M_{act} = 26.120\%$. For the reactive peak prediction, the accuracy is $P_{react} = 28.160\%$, and the MAPE of peak load prediction is $M_{react} = 33.662\%$.

The peak prediction results based on different data spans are illustrated in Fig. 5.10, in which 1-1.3 year's active data and 1-1.2 year's reactive data are utilized to train and test in the peak prediction model respectively. It can be observed that by using data with longer span, the model performs better (with greater accuracy and less MAPE). It should be mentioned that in order to avoid gap-data error existing in the test set, the available data is limited. Therefore the investigation range is 1-1.3 year for active data and 1-1.2 year for reactive data.

5.4 Discussions

5.4.1 Load Prediction Model

In the thesis, hourly active and reactive load demand in the 1st - 24th incoming hours of a typical domestic household is predicted using load data with/without weather data (temperature and solar radiation). Due to the limitation of time and device, the simulations adopt a configuration with relative less number of hidden layer neuron and testing times. The prediction results, especially the results of predictions for farther hours, are not very impressive compared with other ANN based load consumption predictions (predictions for power system) [21, 24, 29, 30]. Increasing the number of neuron is an approach to increase the forecast accuracy. And according to the result of [39], wavelet transformation of the load data is another way to improve the accuracy. However, this approach is not utilized in this study since this approach may erase the power consumption “hobby” of the family in certain extent. In addition, perhaps through extracting several main waveform of the load consumption using Fourier transformation and using the frequency data to train the model is also a way to process the load data beforehand which can improve accuracy. This possibility might be discussed in the future work.

For regional demand prediction considering weather parameters, it has been presented in [36] that the prediction performs better when using the mean of data monitored by weather stations in the region rather than using the data monitored by a certain station. Nevertheless, for domestic demand prediction, more specific weather data might leads to a more accurate forecast result. A more specific weather data includes more details of the household surrounding which might influence the load consumption. Even, it might be required to include continuous weather data for considering influence of the weather parameters.

For the gap-data errors in the raw data, using the load prediction model to estimate the usage of each missed hour could be a way to refill the missing segments. The estimation result might be used for other researches or statistic processes which are not very sensitive on load usage but sensitive on missing data.

Moreover, about selecting input combination, using different input component scenarios for different demand prediction models (with different t) may improve the prediction performance when forecasting farther hours. Introducing correlation analysis into the component selection might be a good approach.

5.4.2 Peak Prediction Model

Peak prediction is built in the thesis based on the load forecast results. Since regression level in the load predictions cannot be controlled well, accuracy of the peak prediction is also not impressive. A straightforward thought to increase the accuracy is to utilize mean error of each demand prediction model to correct the peak prediction. However, since the errors of load prediction are relative unstable, this approach might be hard to have significant effect. Based on the peak prediction results, training the model using data with longer span should be considered as an approach to increase the forecast performance. Another approach that may improve the peak prediction accuracy is to employ the results of “vertical-direction” load predictions as inputs to predict the peak. “vertical-direction” load prediction, means that using every day the same hour’s load to forecast the load demand of the next day at that hour (for example, load usage data at 0:00 is utilized on forecast the demand of the coming day, 0:00).

In the thesis, the continuous hourly data is utilized to forecast the load consumption in the incoming 1st - 24th hour. This kind of prediction can be called as “horizontal-direction” load prediction. This kind of “horizontal-direction” prediction leads to a problem that the time interval between existing data and the forecast target can be different. For example, for demand prediction model 1, the time interval between existing data and forecast target is 1 hour. However, for demand prediction model 24, the interval enlarges to 24 hours. This kind of interval diverse results in a complexity in the control of the prediction. However, for "vertical-direction" load predictions, the time interval between existing hour and the forecast target is the same (24 hours). Thus, the performance of the load predictions is able to be controlled easier. As a result, by using the results of "vertical-direction" load predictions, the accuracy of peak prediction model might be improved.

Chapter 6

Conclusions and Further Scope of the Study

6.1 Conclusions

In the thesis, one hour active and reactive demand prediction for a typical Southern Norwegian domestic household has been built based on two layer feed-forward ANNs using load data with/without processing or using load data together with weather parameters (temperature and solar radiation) as input. It is observed from the numerical results that data processing upgrades the forecast performance. In addition, active load prediction using load data together with weather data performs better than the case where only load data is adopted. And for reactive load prediction, no significant improvement achieves by introducing weather parameters into the prediction.

In addition, hourly active and reactive loads of the incoming 1st - 24th hour for the household have been predicted based on the ANNs with/without considering the impact of weather parameters (temperature and solar radiation). The prediction result conforms to the results in one hour demand prediction. Furthermore, the regression value decreases with increase of forecast hour. At the same time, the decline rate decreases, so that the regression value of prediction for closer hours (forecast hour 1-5) drops fast and the performance of predictions for farther hours (forecast hour 6-24) converges to a constant. A possible reason is that a 3-5 hours' short term load consumption feature exists in the power consumption of the household.

Furthermore, the demand peak with load and occurrence timing of the household in next 24 hours is found out based on the load predictions. It is observed that the peak prediction performs better when data with longer span is utilized for training.

6.2 Contributions

- Hourly active and reactive load predictions (including one hour demand prediction) have been established for domestic household in active and reactive consumption using ANN with/without considering weather parameters (temperature and solar radiation).

- A near future peak finding approach has been built based on the load predictions.

- The predictions have been validated by a typical Southern Norwegian household load consumption data around one and a half years.

- Forecast results have been evaluated and compared. Possible reasons for comparison results have been discussed.

6.3 Further Scope of the Study

In order to obtain better load prediction results, especially for forecasts of farther hour ahead, correlation analysis is going to be used on selecting input power components. Different input combination scenarios will be adopted in the predictions of different incoming hours. In addition, as discussed in the Section 5.4, "vertical-direction" load predictions will be considered to improve the peak prediction accuracy (load and occurrence timing).

Moreover, the load prediction based on the data after the operation of peak-reduction (demand-side management) algorithm is going to be considered in the future study. So that the prediction can be carried out according to the revised load profile.

Furthermore, predictions for peak determined by other strategies, for example, unexpected peak and demand limit peak, are going to be established.

Unexpected peaks are the time instances which consume more loads than expected. This type of peak determination strategy is required by HEMS since the aggregation of predicted power consumption, additional power consume and the loads moved towards this hour may lead to a new peak. To avoid this problem, information about unexpected peak is necessary.

Demand limit peaks are the time instances which consume more loads than the demand limit which is given by the local distribution network. Since the HEMS is a terminal of smart grid system at domestic households. A domestic demand limitation could be broadcast by the upper node in the distribution network. Therefore, in order to manage and schedule loads beforehand to keep the power demand under the demand limit as possible, the demand limit peak prediction is required.

Bibliography

[1] S. Ai, M. L. Kolhe, L. Jiao, N. Ulltveit-Moe, and Q. Zhang, “Domestic demand predictions considering influence of external environmental parameters,” in *Proc. IEEE 13th International Conference on Industrial Informatics (INDIN)*, Jul. 22-24, 2015.

[2] S. Ai, M. L. Kolhe, L. Jiao, and Q. Zhang, “Domestic load forecasting using neural network and its use for missing data analysis,” in *Proc. IEEE - 9th International Symposium on Advanced Topics in Electrical Engineering*, May 7-9, 2015.

[3] Y. Al-Rashid and L. D. Paarmann, “Short-term electric load forecasting using neural network models,” in *Proc. IEEE 39th Midwest symposium on Circuits and Systems*, Ames, IA, Aug. 18-21, 1996.

[4] M. H. Beale, M. T. Hagan, and H. B. Demuth, “Neural network toolbox 7,” *Users Guide, MathWorks*, 2010.

[5] W. Charytoniuk and M.-S. Chen, “Very short-term load forecasting using artificial neural networks,” *IEEE Trans. Power Syst.*, vol. 15, no. 1, pp. 263–268, Feb. 2000.

[6] R. B. Chaudhari, D. P. Dhande, and A. P. Chaudhari, “Home energy management system,” *International Journal of Advanced Electronics and Communication Systems*, vol. 3, no. 3, 2014.

[7] V. Cherkassky, S. R. Chowdhury, V. Landenberger, S. Tewari, and P. Bursch, “Prediction of electric power consumption for commercial buildings,” in *Proc. the 2011 International Joint Conference on Neural Networks (IJCNN)*, San Jose, CA, Jul. 31- Aug. 5, 2011.

[8] A. Deoras. (2014, Dec.) Electricity load and price forecasting webinar case study. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/28684-electricity-load-and-price-forecasting-webinar-case-study

[9] H. Farhangi, "The path of the smart grid," *IEEE Power Energy Mag.*, vol. 8, no. 1, pp. 18–28, Jan. 2010.

[10] T. Gjosaeter, N. Ulltveit-Moe, M. L. Kolhe, R. H. Jacobsen, and E. S. M. Ebeid, "Security and privacy in the semiah home energy management system," in *Proc. 17th EUROMICRO conference on Digital System Design (DSD 2014)*, Verona, Italy, Aug. 27-29, 2014.

[11] M. Goodwin and A. Yazidi, "A pattern recognition approach for peak prediction of electrical consumption," in *Artificial Intelligence Applications and Innovations*. Springer, 2014, pp. 265–275.

[12] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proc. of IEEE*, vol. 75, no. 12, pp. 1558–1573, Dec. 1987.

[13] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–993, Nov. 1994.

[14] L. Han, *Artificial Neural Network Guide*. Beijing University of Posts and Telecommunications, 2006.

[15] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, Feb. 2001.

[16] Y.-Y. Hsu and C.-C. Yang, "Design of artificial neural networks for short-term load forecasting. ii. multilayer feedforward networks for peak load and valley load forecasting," *IEE Proceedings of Generation, Transmission and Distribution*, vol. 138, no. 5, pp. 414–418, Sep. 1991.

[17] S. Humeau, T. K. Wijaya, M. Vasirani, and K. Aberer, "Electricity load forecasting for residential customers: Exploiting aggregation and correlation between households," in *Proc. Sustainable Internet and ICT for Sustainability (SustainIT 2013)*, Palermo, Italy, Oct. 30-31, 2013.

[18] R. H. Jacobsen and E. Ebeid, “Semiah: Scalable energy management infrastructure for aggregation of households,” in *Proc. 17th Euromicro Conference on Digital Systems Design (DSD 2014)*, Verona, Italy, Aug. 27-29, 2014.

[19] C. H. Jeon, Y. J. Cheon, S. W. Sung, C. Lee, C. Yoo, and D. R. Yang, “Development of a supervisory training rule for multilayered feedforward neural network using local linearization and analytic optimal solution,” in *Proc. ICCAS-SICE 2009*, Fukuoka, Japan, Aug. 18-21, 2009.

[20] A. Kailas, V. Cecchi, and A. Mukherjee, “A survey of communications and networking technologies for energy management in buildings and home automation,” *Journal of Computer Networks and Communications*, 2012.

[21] J.-P. Lee, P.-S. Ji, J.-Y. Lim, K.-D. Kim, S.-W. Park, and J.-H. Kim, “A load modeling using ann for power system analysis,” in *Proc. TENCON’99*, Cheju Island, Korea, Dec. 15-17, 1999.

[22] K. Y. Lee, Y. T. Cha, and J. H. Park, “Short-term load forecasting using an artificial neural network,” *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 124–132, Feb. 1992.

[23] F. Liu, R. D. Findlay, and Q. Song, “A neural network based short term electric load forecasting in ontario canada,” in *Proc. International Conference on Computational Intelligence for Modelling, Control and Automation*, Sydney, NSW, Nov. 28 - Dec. 1 2006.

[24] T. G. Monohar and V. C. V. Reddy, “Load forecasting by a novel technique using ann,” *ARPN Journal of Engineering and Applied Sciences*, vol. 3, no. 2, 2008.

[25] NEdata. (2015, Mar.) Pricing reports- zonal information. NEdata. [Online]. Available: http://iso-ne.com/isoexpress/web/reports/pricing/-/tree/zone-info

[26] Z. H. Osman, M. L. Awad, and T. K. Mahmoud, “Neural network based approach for short-term load forecasting,” in *Proc. IEEE Power Systems Conference and Exposition (PSCE ’09)*, Seattle, WA, Mar. 15-18, 2009.

[27] D. R. Palacio, “Simple model for load forecast weather sensitive,” in *Proc. 16th International Conference and Exhibition on Electricity Distribution*, Amsterdam, Holland, Jun. 18-21, 2001.

[28] S. Phimphachanh, K. Chammongthai, P. Kumhom, and A. Sangswang, "Using neural network for long term peak load forecasting in vientiane municipality," in *Proc. TENCON 2004*, Chiang Mai, Thailand, Nov. 21-24, 2004.

[29] M. Ramezani, H. Falaghi, M. Haghifam, and G. A. Shahryari, "Short-term electric load forecasting using neural networks," in *Proc. IEEE International Conference on Computer as a Tool (EUROCON 2005)*, Belgrade, Serbian, Nov. 12-24, 2005.

[30] P. J. Santos, A. G. Martins, and A. J. Pires, "Short-term load forecasting based on ann applied to electrical distribution substations," in *Proc. 39th International Universities Power Engineering Conference (UPEC 2004)*, Bristol, UK, Sep. 8, 2004.

[31] T. Senjyu, P. Mandal, K. Uezato, and T. Funabashi, "Next day load curve forecasting using recurrent neural network structure," *IEE Proceedings of Generation, Transmission and Distribution*, vol. 151, no. 3, pp. 388–394, May 2004.

[32] A. P. Singh, S. S. Sodhi, P. Chandra, and C. S. Rai, "Metrics for weight stuck-at-zero fault in sigmoidal ffanns," in *Proc. 2013 European Modelling Symposium (EMS)*, Manchester, UK, Nov. 20-22, 2013.

[33] A. Singhal and R. P. Saxena, "Software models for smart grid," in *Proc. 2012 International Workshop on Software Engineering for the Smart Grid (SE4SG),*, Zurich, Switzerland, Jun. 3-3, 2012.

[34] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," *Chemometrics and intelligent laboratory systems*, vol. 39, no. 1, pp. 43–62, 1997.

[35] E. L. Taylor, "Short-term electrical load forecasting for an institutional/industrial power system using an artificial neural network," Master's thesis, University of Tennessee, Knoxville, 2013.

[36] J. W. Taylor and R. Buizza, "Neural network load forecasting with weather ensemble predictions," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 626–632, Aug. 2002.

[37] L. Xu, J. Chen, D. Huang, J. Lu, and L. Fang, "Analysis of boundedness and convergence of online gradient method for two-layer feedforward neural

networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 8, pp. 1327–1338, Aug. 2013.

[38] H.-T. Yang and C.-M. Huang, "A new short-term load forecasting approach using self-organizing fuzzy armax models," *IEEE Trans. Power Syst.*, vol. 13, no. 1, pp. 217–225, Feb. 1998.

[39] H.-T. Yang, J.-T. Liao, and C.-I. Lin, "A load forecasting method for hem-s applications," in *Proc. IEEE PowerTech*, Grenoble, France, Jun. 16-20, 2013.

Appendix A

Gap-data Errors

In Appendix A, the gap-data errors in the accumulative active and reactive raw data are presented in Tabs. A..1 and A..2 respectively.

Table A..1: Gap-data errors in raw active accumulative power consumption data

Adjacent hour No.	Start timing	End timing	Gap period (day)	Gap value (kW)
6719-6720	10/09/2012 13:00:00	27/09/2012 13:00:00	17	722.261
13313-13314	29/06/2013 06:00:00	29/06/2013 23:00:00	< 1	37.486
13422-13423	04/07/2013 11:00:00	06/07/2013 23:00:00	2	68.069
13814-13815	23/07/2013 06:00:00	23/07/2013 23:00:00	< 1	18.851
13990-13991	31/07/2013 06:00:00	01/08/2013 23:00:00	1	49.340
14454-14455	21/08/2013 06:00:00	21/08/2013 23:00:00	< 1	30.180
15446-15447	02/10/2013 06:00:00	02/10/2013 23:00:00	< 1	42.929
16078-16079	29/10/2013 06:00:00	29/10/2013 23:00:00	< 1	39.208
16134-16135	01/11/2013 06:00:00	01/11/2013 23:00:00	< 1	44.760
16215-16216	05/11/2013 07:00:00	05/11/2013 23:00:00	< 1	42.037
16295-16296	09/11/2013 06:00:00	09/11/2013 23:00:00	< 1	49.597

Table A..2: Gap-data errors in raw reactive accumulative power consumption data

Adjacent hour No.	Start timing	End timing	Gap period (day)	Gap value (kW)
165-166	13/12/2011 11:00:00	06/01/2012 00:00:00	23	125.131
1753-1754	12/03/2012 03:00:00	12/03/2012 14:00:00	< 1	2.775
6121-6122	10/09/2012 13:00:00	27/09/2012 13:00:00	17	172.980
6276-6277	03/10/2012 23:00:00	05/10/2012 10:00:00	1	9.864
12681-12682	29/06/2013 6:00:00	29/06/2013 23:00:00	< 1	6.031
12790-12791	04/07/2013 11:00:00	06/07/2013 23:00:00	2	13.103
13182-13183	23/07/2013 06:00:00	23/07/2013 23:00:00	< 1	3.561
13358-13359	31/07/2013 06:00:00	01/08/2013 23:00:00	1	10.994
13822-13823	21/08/2013 06:00:00	21/08/2013 23:00:00	< 1	7.706
14814-14815	02/10/2013 06:00:00	02/10/2013 23:00:00	< 1	10.138
15446-15447	29/10/2013 06:00:00	29/10/2013 23:00:00	< 1	11.490
15502-15503	01/11/2013 06:00:00	01/11/2013 23:00:00	< 1	13.068
15583-15584	05/11/2013 07:00:00	05/11/2013 23:00:00	< 1	9.301
15639-15640	08/11/2013 06:00:00	09/11/2013 23:00:00	1	26.217

Appendix B

Publications From This Work

The publications based on the work of this Master Thesis project are presented in Appendix B., which are:

B.1: S. Ai, M. L. Kolhe, L. Jiao, and Q. Zhang, "Domestic load forecasting using neural network and its use for missing data analysis," in *IEEE - 9th International Symposium on Advanced Topics in Electrical Engineering*, May 2015.

B.2: S. Ai, M. L. Kolhe, L. Jiao, N. Ulltveit-Moe, and Q. Zhang, "Domestic demand predictions considering influence of external environmental parameters," in *IEEE International Conference on Industrial Informatics (INDIN)*, Jul. 2015.

B.3: There is another paper "External Parameter Contribution in a Domestic Load Forecasting Neural Network" written by S. Ai, M. L. Kolhe and L. Jiao has been provisionally accepted by the 4th IET Renewable Power Generation Conference (RPG 2015).

THE 9th INTERNATIONAL SYMPOSIUM ON ADVANCED TOPICS IN ELECTRICAL ENGINEERING
May 7-9, 2015
Bucharest, Romania

Domestic Load Forecasting Using Neural Network and Its Use for Missing Data Analysis

Ai Songpu[1], Mohan Lal Kolhe[1,§], Lei Jiao[1], Qi Zhang[2]
[1]Faculty of Engineering and Science, University of Agder, PO 422, Kristiansand, NO 4604, Norway
[2]Department of Engineering, Aarhus University, 8200, Aarhus, Denmark
[§]Email(s): songpa13@student.uia.no, mohan.l.kolhe@uia.no, lei.jiao@uia.no, qz@eng.au.dk

***Abstract*- Domestic demand prediction is very important for home energy management system and also for peak reduction in power system network. In this work, active and reactive power consumption prediction model is developed and analysed for a typical Southern Norwegian house for hourly power (active and reactive) consumptions and time information as inputs. In the proposed model, a neural network is adopted as a main technique and historical domestic load data of around 2 years are used as input. The available data has some measurement errors and missing segments. Before using the data for training purpose, missing and inaccurate data are considered and then it is used for testing the model. It is observed that the possible reasons of prediction errors may be due to local external parameters (e.g. ambient temperature, moisture, solar radiation etc.). It may be required to include analysis of these external parameters on domestic demand prediction model with peak prediction and timing and this will be carried out in our further work.**

Keywords: Domestic load, forecasting, home energy management.

I. INTRODUCTION

Domestic demand prediction is very important for home energy management system and for peak reduction in power system network. It is very important to get prediction of active and reactive power for managing / scheduling the operations of controllable power intensive non-critical domestic loads. The forecasting of active power is necessary for managing the active power supply and demand, and reactive power prediction is also important in managing the voltage profile in the network as well as generating reactive power locally (as local reactive power management will help in reducing network losses and improving power factor). Scheduling / operational timing of these loads will help in reducing the peak demand in a distributed network. Therefore, it is required to have information on prediction of domestic peaks with occurring timing and it is required in home energy management system for scheduling the operation of these loads.

There are numerous existing articles on power prediction in electric network based on artificial intelligence. In [1], a power load modelling method using neural network is presented, in which neural network algorithm is used to establish models of different load components of a house (for example, TV, heater, and so on). The focus of that paper is the construction of component load model, rather than the household-level prediction. In [2], a prediction for whole-building power consumption is addressed with focus on power consumption in commercial buildings. Based on regional power data, a prediction method is reported in [3]. A classical pattern recognition problem is discussed in [4] for peak detection based on average and variance. Based on these literature surveys, we have realized that it is interesting to predict short-term domestic load (active and reactive power) for a typical Southern Norwegian house. In this work, a prediction model for active and reactive is developed using neural network and it is analysed for a typical historical domestic load profile of a Southern Norwegian house.

For building a neural network based prediction model, two tasks are carried out. The first task is process of historical data and it is required for training. However, the available data has some measurements errors and missing segments. Before we use the data for training purpose, it is of great importance to process it in order to reduce the influence of the missing and inaccurate data on prediction. After the initial data processing, the second task has been carried out for training purpose in model in order to develop and analyse the prediction model. The paper is organized in following sections: Section II provides data processing and prediction model, Section III is giving prediction results of this model, which are illustrated and analysed and Section IV provides conclusion with future scope of work.

II. FORECASTING MODEL DEVELOPMENT

As mentioned earlier, the proposed model has two tasks: (i) data processing, and (ii) predictions based on neural network. The forecasting model and data processing are discussed in following sub sections.

A. Forecasting Model

There are many algorithms which can be adopted for prediction purpose and some of them are discussed in [5]. First, linear regression is considered for prediction, but the obtained results are not precise enough. It might have been due to nonlinear behaviour of the data. Therefore, a nonlinear algorithm is adopted. A few load forecasting methods are discussed in ref [5-6] and based on them a neural network is selected to establish a domestic power (active and reactive) consumption prediction model. Use of Levenberg-Marquardt

978-1-4799-7514-3/15/$31.00 ©2015 IEEE

algorithm can make the feed-forward neural network faster in training the data set [7]. A two layer feed-forward neural network with sigmoid hidden neurons and linear output neurons is selected to forecast the curve with multiple inputs and it is also reported in our earlier work of wind power forecasting [8].

The overall structure of the model is illustrated in Fig. 1. To establish a prediction model using the hourly power usage data, the input of this model is taken combination of time and power (active and reactive) usage data. In this work, through trajectories, the combinations of "month, hour, day of week, workday/holiday, last three hour's power consumption, yesterday the same hour's power consumption, and the average power consumption of last 24 hours" are taken as input data. By taking these 9 components as inputs, the highest average regression value R could be achieved by comparing with other input combinations. Regression value R is introduced here to evaluate and compare the prediction results. R measures the correlation between the power consumption prediction result and the practical / actual power usage. R=1 means that a very close relationship exists between the prediction results and targets. In contrast, R=0 means random relationship.

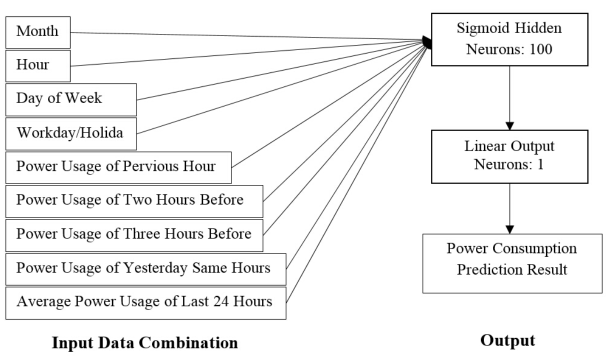


Fig. 1. The two-layer model of neural network

B. Data Processing

In the raw data that we used from a typical Southern Norwegian house (identified as house ID) has accumulative active and reactive power and the detection timing of the accumulative power (active, reactive power respectively) are with some errors. Hence, the raw data is required to be processed in a suitable way as inputs to the neural network. Through analysing the available raw accumulative data, two kinds of data errors are discovered. The first kind of error is named as reading error. For example, at a particular time point, the accumulative power consumption is read as 1582.2 W. However, in the previous hour and the hour after that, the readings are 15680.2 W and 15684.3 W respectively. It is obvious that one digit is missing in 1582.2 W. To remove such type of of errors, the erroneous reading value is replaced by the mean of its two neighbours. In the previous example, 1582.2 W is replaced by 15682.3 W. It is noted that there is an assumption that no adjacent error exists in the neighbouring data by using this method for handling reading errors.

The second kind of error is named as missing data error, which means that part of data is missing between the two time points. For example, one missing data error is observed in available load data, in which 17 days are missing between the time points "10/09/2012 13:00:00" and "27/09/2012 13:00:00". To handle such type of missing data error, we have selected to ignore the missing data part and remove the corresponding forecast results (which are dependent on these missing data), since short term missing data do not have considerable influence to the prediction model output. In the raw data, the starting time point of active and reactive accumulation power are usually not the same. Moreover, errors for the two power types also arise randomly. Therefore, the error removing process is carried out for active and reactive power separately.

III. Prediction Results and Analysis

A. Data Processing

The hourly active and reactive power consumption data of a typical Southern Norwegian house for around two years are obtained and it is observed that there are some missing intervals. The hourly active and reactive power consumption data (without errors) of the house with ID "73500*** *****839" are illustrated in Figs. 2 and 3 respectively.

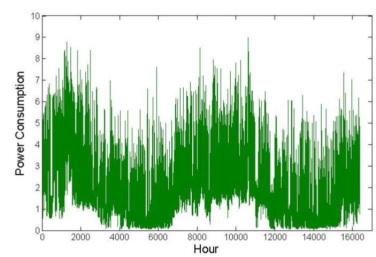


Fig. 2. Active Power Consumption of the House

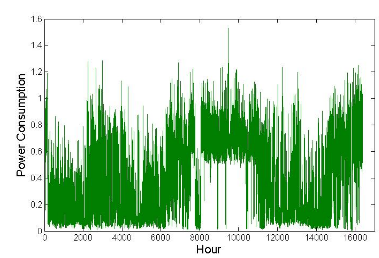


Fig. 3. Reactive Power Consumption of the House

B. Demand Prediction

In this work, the number of hidden neurons for the neural network is configured as 100. Among all the data samples, 70% of samples are used for training, 15% for the validation, and 15% for the testing. Using the input combinations with configuration discussed above, a power consumption prediction can be made to fit the power using pattern of the typical Southern Norwegian house. In the prediction of active power, data processing makes a certain level of contribution for improving the prediction performance. The data processing procedure improves the regression value *R* for active power prediction from 0.62888 to 0.79977 (Figs. 4 and 5). The performances of power consumption prediction without data processing are low. The reactive power prediction without data processing has *R*=0.082126 (Fig. 6), which has poor prediction. However, the reactive prediction with data processing performs much better (Fig. 7), and it has *R*=0.90159. The prediction results of active and reactive power usage with data processing are illustrated in Figs. 8 and 9 respectively. Fig. 10 illustrates the detail of the relationship between the power consumption predictions and targets. From this Fig. 10, it is obvious that the prediction is not precise enough.

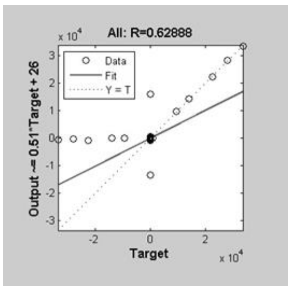


Fig. 4. Regression value of all samples for active power consumption prediction and target without data processing

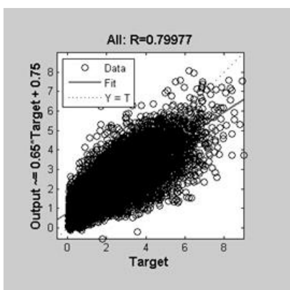


Fig. 5. Regression value of all samples for active power consumption prediction and target with data processing

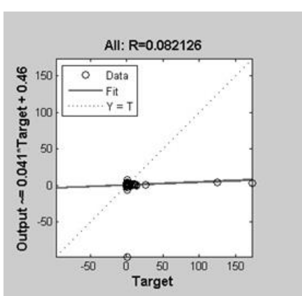


Fig. 6. Regression value of all samples for reactive power consumption prediction and target without data processing

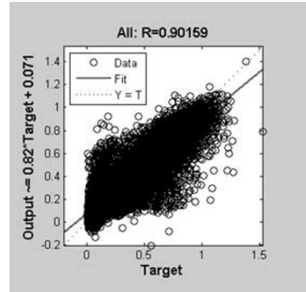


Fig. 7. Regression value of all samples for reactive power consumption prediction and target with data processing

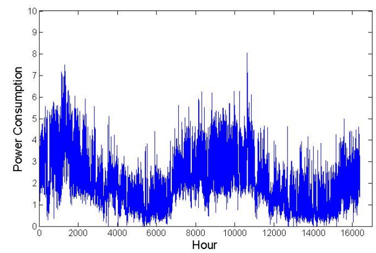


Fig. 8. Result of active power consumption prediction with data processing of a house

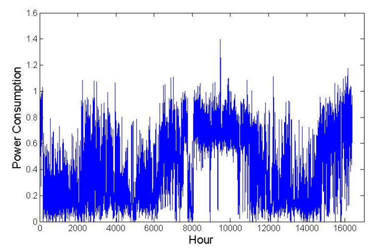


Fig. 9. Result of reactive power consumption prediction with data processing of a house

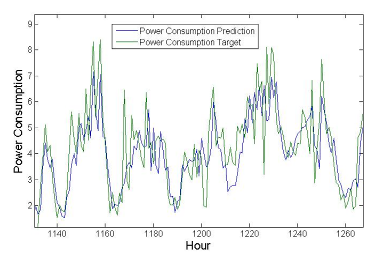


Fig. 10. Detail view of relationship between active power consumption predictions (with data processing) and targets

After running the prediction model with data processing on several typical houses load profiles, the average *R* value for active / reactive power predictions are obtained. The average *R* of active power consumption predictions (with data processing) is 0.79286. And the average *R* value of reactive power consumption predictions (with data processing) is 0.90269. The evaluation results show that both of these power consumption usages are not precise enough to utilize on forecasting peak or making / deciding peak-shift schedule. To make predictions more precise, it may be required to take into account several parameters as inputs. During this analysis only the power usage and time information are used. Therefore, the missing influential data as inputs should be some external parameters. When looking back to the active and reactive power consumption data (Figs. 2 and 3 respectively), a fluctuating phenomenon of the power consumption (especially for the active power usage) with time should be noticed. The power usage in the winter (hour No.0~2000; 8000~10000 approximately) is normally larger than the power usage in summer (hour No.4000~6000; 12000~14000 approximately). It can be observed that there may be some influence of local weather condition (external parameters) on power usage and its predictions. It means that the prediction errors are needed to be improved by including the local external parameters (e.g. temperature, solar radiation, humidity etc.) into the power consumption forecasting model.

IV. CONCLUSION AND SCOPE FOR FUTURE WORK

In this paper, active and reactive power consumption prediction model is developed and analysed for a typical Southern Norwegian house for hourly power (active and reactive) consumptions and time information as inputs. Two kinds of error in raw data are discovered and accordingly they are processed. The forecasting results are presented and evaluated. It is observed that the possible reasons of prediction errors may be due to local external parameters (e.g. ambient temperature, moisture, solar radiation etc.). It may be required to include the analysis of these external parameters on domestic demand prediction model and it will be carried out in our further work.

It is important to predict the demand peak level and its occurring timing. This information will be very useful for home energy management system for scheduling the operation of the non-critical power intensive loads with consideration of user behaviour/comfort. Change of load patterns after scheduling will be useful for forecasting the future domestic load patterns considering demand side management in the housing sector. This work will continue on developing / analysing domestic demand forecasting model and demand peak prediction model using neural network / artificial intelligence and / or with combination of genetic algorithm. The impact of external parameters (e.g. solar radiation, ambient temperature, humidity, etc.) will also be considered in this model for forecasting / analysing the domestic demand and peak predictions.

ACKNOWLEDGMENT

This work is supported by the European Commission under the FP7 programme with contract number FP7-ICT-2013-11-619560 for project SEMIAH 'Scalable Energy Management Infrastructure for Aggregation of Households'. http://www.semiah.eu

REFERENCES

[1] J-P. Lee, P-S. Ji, J-Y. Lim, K-D. Kim, "A load modeling using ANN for power system analysis," *TENCON 99. Proceedings of the IEEE Region 10 Conference,* vol.2, pp. 1475,1478 Dec 1999. doi: 10.1109/TENCON.1999.818712.

[2] V. Cherkassky, S.R. Chowdhury, V. Landenberger, S. Tewari, P. Bursch, "Prediction of electric power consumption for commercial buildings," *IEEE, 2011 International Joint Conference on Neural Networks (IJCNN),* pp. 666-672, Jul. 2011. doi: 10.1109/IJCNN.2011.6033285

[3] A. Deoras, "Electricity load and price forecasting webinar case study". online: http://www.mathworks.com/matlabcentral/fileexchange/28684-electricity-load-and-price-forecasting-webinar-case-study, 29 Oct. 2014.

[4] M. Goodwin, A. Yazidi, "A pattern recognition approach for peak prediction of electrical consumption," *Artificial Intelligence Applications and Innovations, IFIP Advances in Information and Communication Technology,* vol. 436, pp 265-275, 2014. doi: 10.1007/978-3-662-44654-6_26.

[5] E.L. Taylor. "Short-term electrical load forecasting for an institutional/industrial power system using an artificial neural network" M S. thesis, Univ. Tennessee, 2013.

[6] L. Xu, J. Chen, D. Huang, J. Lu, L. Fang, "Analysis of boundedness and convergence of online gradient method for two-layer feedforward neural networks," *IEEE Trans. on Neural Networks and Learning Systems,* vol. 24, no. 8, pp. 1327-1338, Aug. 2013. doi: 10.1109/TNNLS.2013.2257845

[7] M.T. Hagan, M.B. Menhaj, "Training feedforward networks with the Marquardt algorithm." *IEEE Trans. on Neural Networks,* vol. 5, no. 6, pp. 989-993, Nov. 1994. doi: 10.1109/72.329697.

[8] M. Kolhe, T-C. Lin, J. Maunuksela, "GA-ANN for Short-Term Wind Energy Prediction". *IEEE, 2011 Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp. 1-6, Mar. 2011. doi: 10.1109/APPEEC.2011.5749029.

Domestic Demand Predictions Considering Influence of External Environmental Parameters

Ai Songpu, Mohan Lal Kolhe,
Lei Jiao, Nils Ulltveit-Moe
University of Agder,
Faculty of Engineering and Science,
PO Box 422, NO 4604, Kristiansand, Norway.
Email(s): songpa13@student.uia.no, mohan.l.kolhe@uia.no,
lei.jiao@uia.no, nils.ulltveit-moe@uia.no

Qi Zhang
Aarhus University,
Department of Engineering,
Aarhus, 8200, Denmark.
Email: qz@eng.au.dk

***Abstract*— A precise prediction of domestic demand is very important for establishing home energy management system and preventing the damage caused by overloading. In this work, active and reactive power consumption prediction model based on historical power usage data and external environment parameter data (temperature and solar radiation) is presented for a typical Southern Norwegian house. In the presented model, a neural network is adopted as a main prediction technique and historical domestic load data of around 2 years are utilized for training and testing purpose. Temperature and global irradiation (which illustrates the solar radiation level quantitatively) are employed as external parameters. From the results, the efficiency of predictions are evaluated and compared. It can be observed from the numerical results that predictions using historical power data together with external data perform better than the case where only power usage data are adopted.**

Keywords—Domestic load; forecasting; home energy management; environment parameters.

I. INTRODUCTION

Predicting domestic demand in a precise manner is of great importance for reducing the damage caused by power overload and establishing home energy management system (HEMS) and even in using demand side management in virtual power plant. Active power prediction are required by the HEMS to reduce the power demand peaks for shifting the operating times of controllable power intensive non-critical loads from a potential power demand peak to an off-peak period. Active power prediction is required to manage power flow in the network, and forecasting of reactive power is required for managing the voltage profile in the network as well as for power factor improvement [1]. Therefore, domestic demand prediction of both active and reactive power is necessary for HEMS to schedule load operations for reducing peak demand and improving power factor in the distributed power network.

There are numerous articles in literature for electric power prediction based on artificial intelligence. A neural network based power load model is reported in [2] and it has focused on establishing models for power components like heater, incandescent light and cooker, rather than at household-level prediction. In [3], a method of modelling dynamic aggregated power demand response of a group of thermostatically controlled loads is given to calculate the energy savings of these loads. The modelling method has provided an interesting way of building power usage model of a load by using aggregated data. In [4], a possible method of peak prediction is discussed and the forecasting potential peaks are mapped to a classical pattern recognition problem, but the prediction in general is not covered. The power consumption of an entire commercial building is reported in [5]. Several power usage prediction approaches are reviewed in [6], which are based on the regional power usage data in America. A short term power prediction for regional power data based on neural network is provided in [7]. A domestic power demand prediction method based on neural network using only historical power consumption data is proposed in [1], but without external environmental parameters.

The surrounding environment of a household has an important impact on the power demand of a household. For example, heaters require more electric power to keep the room temperature within the comfortable range when temperature decreases, also, in cloudy or rainy days (when the solar radiation is lack), the dryer is more frequently used which leads more power demand. Therefore, considering environment parameters in domestic power prediction should be useful in helping the prediction accuracy. However, existing prediction models are usually based on historical power usage data alone to forecast the future demand. The study of using external data (environment parameters) on neural network based power prediction is still limited. In this paper, a prediction model for active and reactive by using historical power usage data together with environment external data (temperature and solar radiation) is developed using neural network and it is analysed for the load profile of a typical southern Norwegian house with the environment parameter data of the city that house located.

In order to establish a prediction model based on neural network, data processing and model configuration are two main tasks that should be solved. Since the available data have different detection interval, it is necessary to integrate the data interval before training. Also, measurement errors and missing segments in the integrated data also should be processed to reduce the negative impact caused by the inaccurate and

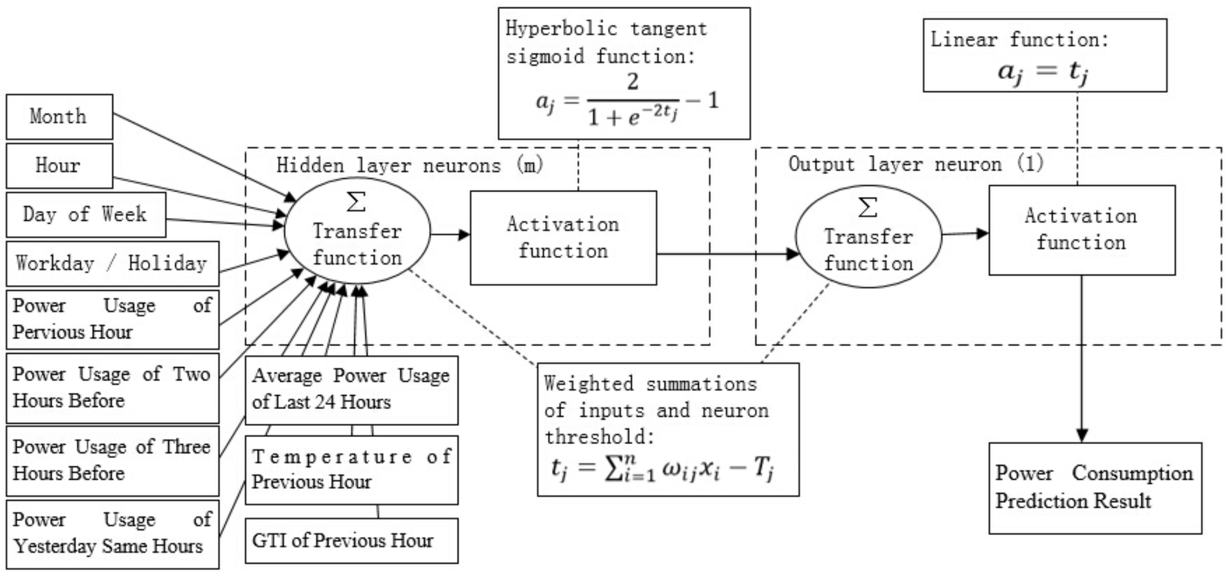


Figure 1 The structure of prediction model.

missing data. Then, the neural network should be configured and trained by historical power usage data combining with external temperature and global solar irradiation (GSI, which reflects the solar radiation level quantitatively) data. The remainder of this paper is organized as follow. Section 2 presents the model configuration and data processing. In Section 3, the results of prediction with external data are given and compared with the prediction without external data. Analysis of the comparison of prediction with external temperature and prediction with both external temperature and GSI is also given in Section 3 before we conclude the work in Section 4.

II. Demand Prediction Model and Data Processing

A. Demand prediction model

Numerous load prediction methods are discussed in [7-8], in which neural network is considered as a fitting method to predict power usage. Moreover, [1] provides a feasible input combination of power usage data, which is utilized by load input portion in this work.

As mentioned in [1, 7-8], multi-layer feed-forward neural network using back-propagation training is able to fit non-linear function well, if the input data is correct and enough hidden neurons are given [7], [9]. In this work, a two-layer feed-forward neural network is selected to forecast the active and reactive power demand. The two layers are denoted as hidden layer and output layer. Neurons in the hidden layer receive the combination of processed input data and then each neuron sends its action to the output layer neuron based on its activation level. There is only one neuron in the output layer. It is in charge of synthesis the actions of all the hidden layer neurons and giving the demand prediction. The structure of this prediction model is illustrated in Fig. 1. In each neuron, transfer function is calculated firstly. Transfer function is the weighted summations of each input combination and the bias values of the neuron and it denotes the number of input data component as n and the number of hidden layer neuron as m. Components of each data are denoted as combination i.e. x_i, where i=1, 2, 3...n. The bias values of each neuron is denoted as T_j, where j=1, 2, 3...m. To denote each neuron, the weight of each input component as ω_{ij} is considered. Then, the transfer function of each neuron could be expressed as $t_j = \sum_{i=1}^{n} \omega_{ij} x_i - T_j$. After obtaining the weighted summation, activation function is utilized by each neuron to decide the action. In hidden layer neurons, hyperbolic tangent sigmoid function is utilized as activation function and in the output layer neuron, linear function is selected [10]. The load usage prediction of next hour (next step) is obtained using (1), where v_j (j=1,2,…m) and T_{out} expresses the weight and bias value of the output layer neuron respectively.

$$\text{Output} = \sum_{j=1}^{m} \frac{2v_j}{1 + e^{-2(\sum_{i=1}^{n} \omega_{ij} x_i) - T_j}} - T_{out}. \tag{1}$$

The weights and bias values of each neuron are adjusted in each training repetition to obtain lower error between prediction and practical power usage. In this work, Levenberg-Marquardt algorithm is utilized on data training, since this algorithm can make training period shorter [11].

In order to train the domestic demand prediction model, a combination of power usage data and external environment (temperature and GSI) data can be selected as input. It is observed that the power usages of a day follow some basic regulars. Generally, in work days, people seldom stay in their house at work time. However, in holidays, people can either stay in their house or go out travelling. Thus, the date information (month, day of week and work day / holiday) is important for prediction. Since the power demand in different periods of a single day is also regulated, the "hour" information is required. In addition, previous power usage data is required to forecast the demand of next hour. "Last

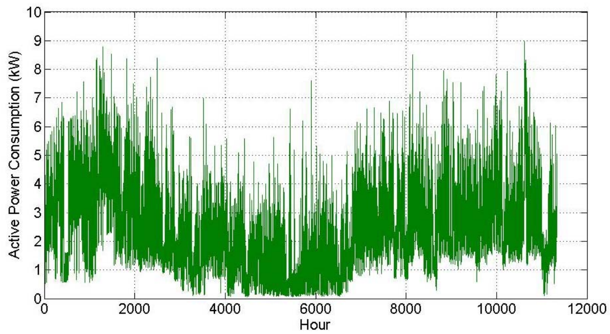


Figure 4 Active Power Consumption of the House.

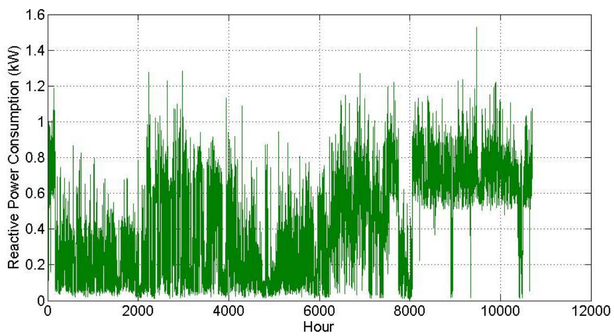


Figure 5 Reactive Power Consumption of the House.

three hour's power consumption, yesterday the same hour's power consumption, and the average power consumption of last 24 hours" are chosen to reflect the previous hourly power usage information. Thus, based on our trajectory work in [1], the combinations of "month, hour, day of week, workday / holiday, last three hour's power consumption, yesterday the same hour's power consumption, and the average power consumption of last 24 hours" are taken as input data of power usage portion. Although plenty of power data combinations are able to be utilized in the prediction, the combination shown above is more proper for our prediction. In addition, the combinations of "temperature of previous hour and GSI of previous hour" are taken as the external portion data.

To analyse and evaluate the prediction model, two evaluation methods are introduced to evaluate prediction result. Regression value R is one evaluation method which is utilized in this work to evaluate and compare the prediction results. $R \in [0, 1]$ measures the correlation between the power consumption prediction result and the target power usage. A larger R illustrates a closer relationship existing between the prediction results and targets. In contrast, a less R means a more randomly relationship. One other evaluation method utilized in this work is mean absolutely percentage error (MAPE), which is denoted as M. M measures the accuracy of a method for calculating the average absolute occurred errors.

By taking the hourly active and reactive power components and external data components described above as inputs, the power demand of the next hour is to be predicted. The forecasting results are evaluated and analysed by R and M.

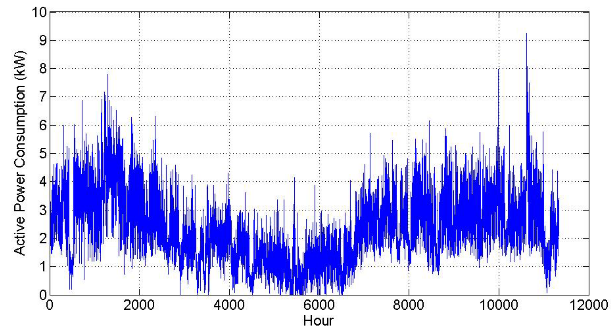


Figure 2 Result of active power demand prediction with both external temperature and GSI data of a house.

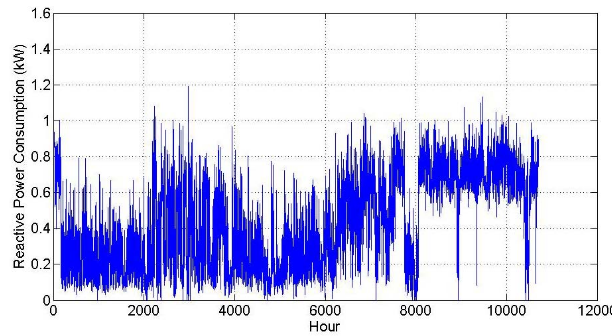


Figure 3 Result of reactive power demand prediction with both external temperature and GSI data of a house.

B. Data processing

In order to utilize the raw data into the prediction model, it is necessary to integrate the time intervals of data firstly.

Temperature and GSI are considered as the external parameters in this work. The data of GSI and air temperature of the city the house located are obtained and used in 15 minutes' time interval. The temperature samples are taken at the 0th, 15th, 30th, and 45th minute of each hour, and the GSI data is taken at the 11th, 26th, 41st, and 56th minute of each hour. Also, the interval of power usage data of a typical Southern Norwegian house and external data are not the same. The interval of power usage data is 1 hour. It means that one power usage sample corresponds with 4 external data samples. In order to integrate the time interval, for temperature data, the sample taken at 0th minute of an hour is used to express the sample of the whole hour. And for GSI data, the sample taken at the 56th minute of an hour is utilized to express the sample of the hour. In this way, the power usage data and external data have same time interval which is one hour.

Missing segments in the integrated data are considered in ref [1]. To handle missing segment errors, it has been considered to ignore the missing data in the training and validation sets. Make sure that the missing periods wouldn't appear in the test set. It is because that short term missing data does not considerable influence to the accuracy of prediction model.

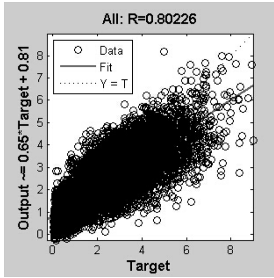


Figure 6 Regression value of all samples for active power consumption prediction with external temperature data and target.

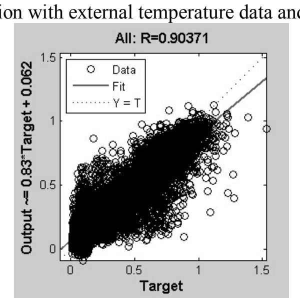


Figure 7 Regression value of all samples for reactive power consumption prediction with external temperature data and target.

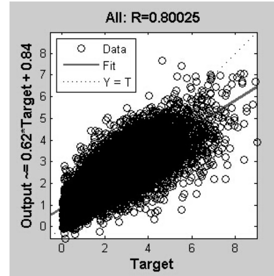


Figure 8 Regression value of all samples for active power consumption prediction with external temperature and GSI data and target.

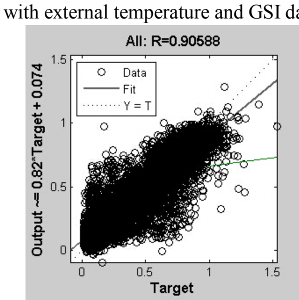


Figure 9 Regression value of all samples for reactive power consumption prediction with external temperature and GSI data and target.

III. PREDICTION RESULTS AND ANALYSIS

A. Data processing

The processed hourly active and reactive power consumption data of a typical Southern Norwegian house about two years are illustrated in Fig. 2 and Fig. 3 respectively. The hour number is ranked at x-axis from the beginning hour (0) to the ending hour (about 11400), in which the missing segments are ignored. Since the missing segments take place in active and reactive data separately, the number of hours in Fig. 2 and Fig. 3 are different.

B. Prediction Results and Analysis

All the data combination samples are divided using random uniform sampling into three sets. 70% of data combination samples are utilized on training the neural network, 15% of samples are utilized on validation and the remaining 15% are utilized on testing the network performance.

The configuration about the number of neuron in the hidden layer, m, is a compromise of accuracy and computation amount. Since m=10 is not enough to obtain a precise forecast, and m=1000 is too much to be calculated fast, the number of hidden neurons for the neural network based prediction model is configured as 100.

Using the combination of the historical power usage data and the external data (temperature and GSI) with configuration discussed above, the power demand prediction, based on power and environmental parameters, is made to fit the power using pattern of the typical Southern Norwegian house. The remaining parts of this section are organized as follow. Firstly, prediction based on historical power usage data and temperature data are trained. The forecasting results are compared with the prediction without using external data. Then, the prediction results based on power data and both temperature and GSI data are presented and analysed. The active and reactive power prediction results using temperature and GSI data are shown in Fig. 4 and Fig. 5 respectively.

In the situation of using external temperature data in active power prediction, it has R=0.80226 (Fig. 6) and M=49.4703. And the reactive power prediction using external temperature data has R=0.90371 (Fig. 7) and M=48.6966. For the predictions without using external data, the active power prediction has R=0.79977, M=87.2701, and reactive power prediction has R=0.90159, M=49.5252. It is able to observe that by introducing temperature data into the prediction, regression value R has a little advantage comparing with the results of without external data prediction. However, MAPE in the power and temperature based prediction decreases a lot. These results illustrate that although the correlation between prediction result and practical usage does not become closer, the absolute value distance between prediction result and practical usage reduces a lot.

In the predictions of using power data and both temperature and solar radiation data in active power prediction, it has R=0.80025 (Fig. 8) and M=46.9245. The reactive power prediction using temperature and solar data has R=0.90588 (Fig. 9) and M=48.3425. It can be observed that there is no significant improvement by introducing solar radiation data into the prediction model. The meteorological data and load data may not be from the same location and due to that significant improvement is not observed in the forecasting. The obtained temperature and GSI data is the comprehensive record of the city, where the household is located. Temperature variation within city may not be significant. It is noticed that ambient temperature has significant impact on active power prediction.

However, the solar radiation data is a different story. Since the different locations, the solar radiation level of two households could be widely different from each other. Thus, the GSI of the city can hardly represent the solar radiation level of a household. Therefore, introducing GSI into the prediction does not significantly improve the forecasting accuracy. To make predictions more precise, it may be required to have more precise external environmental data for the household surrounding area.

IV. Conclusion

In this paper, the impact of external data on active and reactive power demand prediction is discussed. The load (active and reactive) predictions have been done using available load data and considering external environmental parameters i.e. temperature and GSI are done. The accuracy of the predictions are evaluated and compared by regression value R and MAPE M. As a result, the prediction based on power and temperature data reduces the value M compare with the prediction without considering external data. And the prediction based on power and both temperature and solar radiation data does not have significant improvement compare with the prediction based on only power and temperature data.

One possible reason of prediction errors may be non-availability of site specific external environmental data. In further work, combination of neural network / artificial intelligence and / or with genetic algorithm will be considered for load prediction. The work is continuing and also going to consider use of outputs in finding peak predictions with timing, which is required for HEMS.

Acknowledgment

This work is partially supported by the European Commission under the FP7 program with contract number FP7-ICT-2013-11-619560 for project SEMIAH ‘Scalable Energy Management Infrastructure for Aggregation of Households’. http://www.semiah.eu.

The University of Agder gratefully acknowledge the partial support of the Research Council of Norway and partners in the project ‘Electricity Usage in Smart Village Skarpnes’.

References

[1] S. Ai, M. L. Kolhe, L. Jiao, and Q. Zhang, “Domestic load forecasting using neural network and its use for missing data analysis,” in *IEEE - 9th International Symposium on Advanced Topics in Electrical Engineering*, Bucharest, Romania, 7-9, May 2015.

[2] J-P. Lee, P-S. Ji, J-Y. Lim, K-D. Kim, "A load modeling using ANN for power system analysis," in *IEEE Region 10 Conference - TENCON 99,* pp. 1475-1478, Dec. 1999. doi: 10.1109/TENCON.1999.818712.

[3] C. Perfumo, J. Braslavsky, and J. Ward, “Model-based estimation of energy savings in load control events for thermostatically controlled loads,” *IEEE Trans. on Smart Grid*, pp. 1410–1420, May 2014, doi: 10.1109/TSG.2014.2298840.

[4] M. Goodwin, A. Yazidi, "A pattern recognition approach for peak prediction of electrical consumption," in *Artificial Intelligence Applications and Innovations (AIAI)*, pp 265-275, 2014. doi: 10.1007/978-3-662-44654-6_26.

[5] V. Cherkassky, S.R. Chowdhury, V. Landenberger, S. Tewari, P. Bursch, "Prediction of electric power consumption for commercial buildings," in *IEEE - 2011 International Joint Conference on Neural Networks (IJCNN)*, pp. 666-672, Jul. 2011. doi: 10.1109/IJCNN.2011.6033285

[6] A. Deoras, "Electricity load and price forecasting webinar case study". online: http://www.mathworks.com/matlabcentral/fileexchange/28684-electricity-load-and-price-forecasting-webinar-case-study, 29 Oct. 2014.

[7] E. L. Taylor. "Short-term electrical load forecasting for an institutional/industrial power system using an artificial neural network" M S. thesis, Univ. Tennessee, 2013.

[8] L. Xu, J. Chen, D. Huang, J. Lu, L. Fang, "Analysis of boundedness and convergence of online gradient method for two-layer feedforward neural networks," *IEEE Trans. on Neural Networks and Learning Systems*, pp. 1327-1338, Aug. 2013. doi: 10.1109/TNNLS.2013.2257845

[9] D. Svozil, V. Kvasnicka, and J. Pospichal, “Introduction to multilayer feed-forward neural networks,” in *Chemometrics and intelligent laboratory systems*, pp. 43–62, Nov. 1997 doi:10.1016/S0169-7439(97)00061-0.

[10] M. Kolhe, T.C. Lin, J. Maunuksela, "GA-ANN for Short-Term Wind Energy Prediction," in *IEEE - 2011 Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp.1-6, 2011. doi: 10.1109/APPEEC.2011.5749029.

[11] M.T. Hagan, M.B. Menhaj, "Training feedforward networks with the Marquardt algorithm." *IEEE Trans. on Neural Networks*, pp. 989-993, Nov. 1994. doi: 10.1109/72.329697.