Data Mining Techniques for Smart Grid Load Forecasting

Mohamed H. Toukhy

Master Thesis of Computing and Information Science

at

Masdar Institute of Science and Technology

March, 2012

©Masdar Institute of Science and Technology

All Rights Reserved
Abstract

Smart grids, or intelligent electricity grids that utilize modern IT/communication/control technologies, become a global trend nowadays. Forecasting of future grid load (electricity usage) is an important task to provide intelligence to the smart grid. Accurate forecasting will enable a utility provider to plan the resources and also to take control actions to balance the supply and the demand of electricity. In this paper, our contribution is the proposal of a new data mining scheme to forecast the peak load of a particular consumer entity in the smart grid for a future time unit. We utilize least-squares version of support vector regression with online learning strategy in our approach. Experimental results using two datasets, each containing two sub-datasets, show that our method is able to provide more accurate results than an existing forecasting method which is reported to be one of the best. On Germany dataset, our method can provide 98.4–98.7% of average accuracy whilst the state-of-the-art method by Lv et al. is able to provide only 96.7% of average accuracy. On Abu Dhabi dataset, which is less predictable, our method still can provide 95.7–96.7% of average accuracy whilst the method by Lv et al. provides lightly less average accuracy of 95.3%–96.4%. Our method is also computationally efficient and can potentially be used for large scale load forecasting applications.

Keywords—smart grids; data mining; load forecasting; regression analysis; support vector machines.
This research was supported by the Government of Abu Dhabi to help fulfill the vision of the late President Sheikh Zayed Bin Sultan Al Nayhan for sustainable development and empowerment of the UAE and humankind.
Acknowledgments

To Almighty ALLAH, for granting me the wisdom, ability, perseverance, and clarity to do this work.
This research was sponsored by the Government of Abu Dhabi, United Arab Emirates through its funding of the MIT-Masdar Institute Collaborative Research Project on Data Mining for Smart Grids (award number 10CAM1).
First of all I would like to show my special thanks to my supervisor Prof. Zeyar Aung for all his guidance and support. He supervised the whole process of my research, and had provided many technical suggestions. I benefited a lot from his knowledge and experience in research.
I owe my gratitude to the MASDAR management for giving me the chance to study in UAE. Many thanks are due to TRANSCO-NMD, Abu Dhabi, specially Dr. Roshdy Hegazy, Network Management Division Manager for providing the Power Load data for Emirate of Abu Dhabi used in this work.
I wish to thank every individual at Masdar Institute for science and Technology, especially fellow colleagues in CIS department. Without their help and encouragement during the two years in United Arab Emirates, I wouldn’t have finished the Master research work.
I would also like to thank my parents, my brother Ahmed and my sisters Amal and Naema for their support, patience, and understanding of my research in Abu Dhabi. Many thanks are to my ever encouraging lovely wife, Mariam, and my kids Sara and Ahmed for being patient, and supportive throughout the study period.

Mohamed Toukhy,
Masdar City, March 22 2012
## Contents

1 Introduction ........................................................................................................................ 10  
1.1 Summary ....................................................................................................................... 10  
1.2 Scope of Work ............................................................................................................... 13  
1.3 Awareness of the problem .............................................................................................. 14  
1.4 Problem statement ........................................................................................................ 15  
1.5 Proposed Solution Outline ............................................................................................ 17  
1.6 Thesis organization outline ......................................................................................... 18  
1.7 Conclusion ..................................................................................................................... 19  

2 Literature Review ............................................................................................................... 21  
2.1 Overview of load forecasting techniques ....................................................................... 21  
2.2 Classification of forecasting methods ............................................................................. 23  
2.2.1 The Regression Technique ...................................................................................... 23  
2.2.2 The Time Series Technique .................................................................................... 24  
2.2.3 Artificial Intelligence and Neural Networks Techniques .......................................... 25  
2.2.4 Expert System Techniques ..................................................................................... 27  
2.2.5 Fuzzy Logic Techniques ......................................................................................... 28  
2.2.6 Support Vector Machines ....................................................................................... 29  

3 Power System Load Forecasting ........................................................................................ 30  
3.1 Overview ....................................................................................................................... 30  
3.2 Characteristics of a power system load .......................................................................... 31  
3.3 Factors influencing the load .......................................................................................... 31  
3.4 Load features ............................................................................................................... 38  
3.5 Prerequisites of a good STLF System .......................................................................... 41  
3.6 Conclusion ..................................................................................................................... 44  

4 Smart Grids ......................................................................................................................... 45  
4.1 Introduction .................................................................................................................... 45  
4.2 What is Smart Grid? ...................................................................................................... 46  
4.3 Main Goals of smart Grid ............................................................................................. 50  
4.4 Smart Grid vision .......................................................................................................... 50  
4.5 Summary ....................................................................................................................... 55  

5 Proposed Solution ............................................................................................................... 57  
5.1 Summary ....................................................................................................................... 57
5.2 Support Vector Machine ................................................................. 60
5.3 General SVM Formulation ............................................................ 60
5.4 Least Squares SVM Formulation .................................................... 63
5.5 Kernel Function ........................................................................... 66
5.6 Online Learning .......................................................................... 67
5.7 Feature Vector Construction .......................................................... 67
6 Experimental Results ................................................................. 69
  6.1 Experiment on Germany Data ....................................................... 69
  6.2 Experiment on Abu Dhabi Data ..................................................... 75
7 Conclusion ...................................................................................... 79
Bibliography ..................................................................................... 83
Table 1: A side-by-side comparison of the traditional and smart grids......................... 50
Table 2: Feature vector used in forecasting the peak load $P_d$ of a given day $d$. ................. 68
List of Figures

Figure 1: Types of load forecasting and work focus ......................................................... 14
Figure 2: Overview of the proposed load forecasting scheme ........................................... 18
Figure 3: A simple AC circuit ........................................................................................... 31
Figure 4: Possible infusing factors – for STLF. ................................................................. 32
Figure 5: Typical influencing factors – for M-LTLF ......................................................... 36
Figure 6: Architecture and components of the smart grid ................................................. 47
Figure 7: Overview of the proposed load forecasting scheme ........................................... 59
Figure 8: The Daily peak load profile in Lower Saxony (LS), Germany in 2009. The spikes indicate the increases of peak loads on holidays. Electricity usage is higher in winter and lower in summer ................................................................. 70
Figure 9: Example of forecasted results. The actual peak loads and the forecasted values by Lv et al. [11] and our method for the period of December 1 to 31, 2009 on LS Dataset ........ 71
Figure 10: Relative error performances of Lv et al. [11] and our method for the period of June 1, 2009 to December 31, 2009 on LS Dataset ................................................................. 73
Figure 11: Relative error performance of Lv et al. [11] and our method for the period of June 1, 2009 to December 31, 2009 on NRW Dataset ................................................................. 73
Figure 12: Relative error performances of scaling vs. non-scaling approaches in our method. (The results are the averages of those for LS and NRW Datasets for the period of June 1, 2009 to December 31, 2009.) ................................................................. 74
Figure 13: The Daily peak load profile in Abu Dhabi Substation #1 (AD1 Dataset) from August 2009 to June 2011. Electricity usage is higher in summer and lower in winter .......... 76
Figure 14: Example of forecasted results for Abu Dhabi. The actual peak loads and the forecasted values by Lv et al. [11] and our method for the period of June 1 to 30, 2011 on AD1 Dataset ................................................................. 77
Figure 15: Relative error performances of Lv et al. and our method for the period of March 01, 2010 to June 30, 2011 on AD1 Dataset ................................................................. 78
Figure 16: Relative error performances of Lv et al. and our method for the period of March 01, 2010 to June 30, 2011 on AD2 Dataset ................................................................. 78
CHAPTER 1

Introduction

This chapter encapsulates a brief overview of power system load forecasting, awareness of the problem, problem statement, research aim and objectives, hypothesis, assumptions, delimitations and motivation of the research. It also partially covers research design and sampling methodologies. The last part of this chapter provides the organization and outline of the thesis.

1.1 Summary

With the skyrocketing growth of power system networks and the increase in their complexity, many factors have become influential in electric power generation,
demand or load management. Load forecasting is one of the critical factors for economic operation of power systems.

Forecasting of future loads is also important for network planning, infrastructure development and so on. However, power system load forecasting is a two dimensional concept: consumer based forecasting and utility based forecasting. Thus the significance of each forecast could be handled disjointedly. Consumer based forecasts are used to provide some guidelines to optimize network planning and investments, better manage risk and reduce operational costs.

In basic operations for a power generation plant, forecasts are needed to assist planners in making strategic decisions with regards to unit commitment, hydro-thermal co-ordination, interchange evaluation, and security assessments [31] and so on. This type of forecast deals with the total power system loads at a given time, and is normally performed by utility companies.

Nonetheless, power system load forecasting can be classified in three categories, namely short-term, medium term and long term forecasting. The periods for these categories are often not explicitly defined in a number of literature papers. Thus different authors use different time horizons to define these categories. But roughly, short-term load forecasting covers hourly to weekly forecasts. These forecasts are often needed for day by day economic operations of power generation plants.

Medium-term load forecasting deals with predictions ranging from weeks to a year. Outage scheduling and maintenance of plants and networks are often roofed in these types of forecasts.

Long term forecasting on the other hand deals with forecasts longer than a year. It is primarily intended for capacity expansion plans, capital investments, and corporate budgeting. These types of forecasts are often complex in nature due to future
uncertainties such as political factors, economic situation, per capital growth etc. Planning of new and extensions to existing power system networks for both the utility and consumer require long-term forecasts.

The accuracy of the forecast is a critical feature in power system load forecasting. A poor load forecast misleads planners and often results in wrong and expensive expansion plans. From the consumer forecast view, accurate load forecasting is important for distribution system investments, electric load management strategies. This subject also forms part of load rationing strategies i.e. load shedding, DSM (Demand Side Management) initiatives and so on. Without replicating, short term load forecasting is an essential function in daily operational activities especially for utility companies. A negative error in the forecast could severely affect consumer’s production levels, particularly for larger power users. Thus accurate forecasts are required for power system security and its overall reliability.

One of the convincing ways to predict loads that are known to be varying continuously on short-term bases is to rather minimize load sampling points to several minutes or hours. This approach is called very short-term load forecasting and will not be discussed in this work.

Undoubtedly, both utility companies and consumers are challenged to accurately predict their respective loads. This challenge has been in existence for decades, thus a variety of load forecasting techniques ranging from classical to intelligent systems have been developed to date, and highlighted in a number of studies. The ultimate distinction of these methods can be drawn on the bases of forecast accuracy.

The most popular techniques used for load forecasting are time series based models, similar-day approach and intelligent system based models. Some of the conventional forecasting methods have major drawbacks especially their inability to map the non-
linear characteristic of the load, thus a substitute of classical methods with intelligent system based models is to a great extent essential.

Most forecasting models use statistical techniques or artificial intelligence algorithms such as regression, neural networks, fuzzy logic, and expert systems.

Amongst all other intelligent techniques, the use of ANN in STLF is very predominant. Most recent load forecasting works are based on Artificial Neural Networks, and a majority of these papers presented good estimates. Because ANNs are capable of generalization and learning non-linear relationships between variables, ANN-based approaches are often favored for STLF problems. The other important feature of ANNs is their capability to iteratively adjust the synoptic weights between layers. Conventional methods on the other hand require static complex mathematical equations and still perform poorly in comparison to intelligent-based approaches.

Another leading load forecasting method is Fuzzy logic. Its application in load forecasting is based on periodical similarity of electric load, where the input variables, output variables and the governing rules are the key points. The scope of this work does not cover this technology despite a brief introduction in following chapter.

1.2 Scope of work

This research work focuses on a specific area of load forecasting, short-term load forecasting in smart grids. The forecasts are achieved by using On Line Support Vector Regression based model using Least Square, The model are applied to the actual load data of two different buildings in two different areas in Germany to forecast what is often referred to as consumer own forecast.
The below attempts to clarify the different types of the power load forecasting and the focus area of this research.

**Figure 1:** Types of load forecasting and work focus

1.3 Awareness of the problem

Short-term load forecasting plays a major role in the real-time control and security functions of an energy management system. Accurate forecasts greatly benefit power system planners to accomplishing a variety of tasks such as economic scheduling of generating capacity, scheduling of fuel purchases etc. However, load forecasting is a difficult task because the consumption is influenced by many factors, such as weather conditions, vacations, economy status, and idiosyncratic habits of individual customers.

Inaccurate load forecasts may increase operating costs. It was observed that reported that a one percent increase in forecasting error of electricity demand resulted in a £10 million increase in operating costs in the British power system. This is purely an error in the utility-type of load forecast.
Evidently, a poor load forecast misleads planners and often results in wrong and expensive expansion plans. While Accurate forecasting will enable a utility provider to plan the resources such as fuel in advance and also to take control actions like switching on/off demand response appliances and revising electricity tariffs, etc. Equally, overestimating future electric loads may result in a redundant reserve of electric power. On the contrary, underestimation of loads causes failure in providing sufficient electric power.

For a planner to neither underestimate nor overestimate the load, convenient forecasting techniques with reasonable degree of accuracy need to be developed. Although different models entertain some superiority in dynamic systems, possibilities to improve associated drawbacks cannot be ruled out.

Therefore there is a need for development of optimal and accurate based load forecasting models to improve (minimize) the forecast error. The main research question of the project can be formulated as: How can Support Vector Regression-based load forecasting models be rationally trained algorithms and possibly obtain optimal improved results for short-term load forecasting?

1.4 Problem statement

Generally, electric load forecasting is a complex exercise. An electric load is a non-linear function; traditional forecasting methods are simply not suitable for the application due to the lack of nonlinear mapping ability. Intelligent techniques on the other hand, require unified training algorithms in order to improve the accuracy of the forecast.
Problem statement: To develop optimized Support Vector Regression-based model for Short-Term Load Forecasting and apply the resulted model to a real life cases to evaluate the performance of the proposed approach and provide one month ahead forecast.

In the research, we will focus on a specific problem of forecasting the peak load (i.e., the maximum electricity usage) of a particular consumer entity for a future time unit. The consumer entity in question can be of various granularity levels. For example, it can be a smart meter (for a household), a cluster of smart meters (for a neighborhood), a power substation (for a town or city), or a power station (for an entire grid covering a large geographical area). Similarly, the time unit in question can be of different lengths. It can be 5 minutes, 15 minutes, 1 hour, 1 days, 1 week, etc.

In this work, we will study a system to forecast the daily peak loads of individual smart meters. However, it should be noted that the same principles and techniques used in our studies are generally applicable to any load forecasting problems with any combinations of consumer entities and time granularities.

Researchers have been trying to solve the problem of electricity load forecasting since 1990’s [1]. A number of methods based on different techniques such as time series analyses (like autoregressive integrated moving average (ARIMA) method [5]), fuzzy logic [14], neuro-fuzzy method [7], artificial neural network (ANN) [3], and support vector regression (SVR) [11] have been proposed.

Among these various techniques, support vector regression (SVR) is one of the latest developments. In [11], it was demonstrated that SVR could provide better results than the older methods like artificial neural network (ANN) [3] could.
1.5 Proposed solution outline

As in the previous work [3] and [11], we try to approach the problem of smart grid’s load forecasting from the data mining perspective. In particular, we propose a peak load forecasting model based on the data mining technique of support vector regression (SVR) using least squares [15]. More specifically, we use the online least-squares SVR proposed by Engel et al. [7].

In order to predict the peak load \( P_d \) of a particular day \( d \), we can roughly consider \( P_d \) as a non-linear combination of a number of attributes from different sources: peak loads of previous \( N \) days, average temperatures of previous \( N \) days, holiday records of previous \( N \) days, forecasted temperature of day \( d \), and whether day \( d \) is a holiday (weekend or public holiday).

So, for each target peak load value in the historical record, we construct a feature vector covering the abovementioned attributes associated with the target. Then, we train our least-square SVR system using a set of <feature vector, target> pairs for a large enough number of days. The result of this training process is a least-squares regressor model.

We can use the resultant regressor model to forecast the peak load value \( P_d \) of a given day \( d \). For that, we have to construct a feature vector for the day \( d \) in the same manner as in the training step. In constructing the feature vector, we need to know the forecasted temperature of the day \( d \) (if it is in the future) and whether it is a holiday (which can be easily known in advance). Then, the feature vector of day \( d \) is supplied to the regressor model to generate the forecasted peak load value of that day.

After the day \( d \) is already passed and its actual peak load value (the target) already known, the regressor model is updated with the <feature vector, actual target> pair for the day \( d \), thus resulting in a fresh model which best reflects the latest trend of events.
A schematic representation of our proposed load forecasting scheme is given in

![Diagram of load forecasting scheme]

**Figure 2**: Overview of the proposed load forecasting scheme.

We tested the proposed method using the smart metering data from the two regions in Germany for the year 2009. Experimental results demonstrate that our approach is both accurate and computationally efficient.

The preliminary results of this thesis work were presented as conference paper [51].

### 1.6 Thesis organization outline

This thesis document comprises seven chapters, arranged in a systematic manner: the current Chapter 1 of the document mainly discusses the background, purpose of the work and breakdown structure of the work.

Chapter 2 covers literature review, i.e., methods used for load forecasting, comparisons of various papers, findings and remarks. In this chapter, drawbacks of
different forecasting methods are also highlighted. This chapter exclusively also discusses the application of genetic algorithm to STLF.

Chapter 3 of this document broadly talks about power system load forecasting: i.e. the factors affecting the load, characteristics and features of the power system load. This chapter also roofs basic requirements of a good STLF system.

An overview on Smart grids and its characteristics is given in Chapter 4. This section covers the different definitions for the smart grid, the main goals of smart grids and the benefits by its usage.

The content of Chapter 5 entails description of the proposed load forecasting method using online least-squares support vector regressions in details.

Chapter 6 presents the experimental results of the proposed method in comparison with a state-of-the-art method.

The last chapter (Chapter 7) of the thesis encapsulates conclusion and future outlook of the project.

1.7 Conclusion

This chapter highlights the basic introduction of the proposed research project. Thus it generally covers the scope of work, significance and objectives of the proposed research project, some brief research methodologies. The chapter also discusses some well thought-out assumptions, delimitations of the research as well as the structure of the proposed work and final organization of the thesis.

In order to establish possible shortcomings of some commonly used forecasting methods, one needs to extensively conduct a literature review on the existing forecasting techniques. This subject is covered in the following chapter and some of
the commonly used forecasting techniques are discussed by explanatory means therein.

Moreover, the following chapter also covers a brief introduction of genetic-based search in STLF with a hope to minimize the error function.

The next chapter also highlights a comprehensive comparison of some previous similar works. The findings, remarks, and conclusions of the literature comparison are also discussed in the next chapter.
The subject of load forecasting has been in existence for decades, and a number of techniques have been developed. These methods are based on either classical or modern approach. This part of the research work is necessary to establish the statistical relevance of the proposed research project, establish a generalized research question, analyze existing methods, and explore areas of possible improvements. This chapter also covers the analysis of various existing load forecasting techniques, a comparative study of reviewed papers, findings and remarks.

2.1 Overview of load forecasting techniques

Load forecasting has become one of the most significant aspects of electric utility planning. The economic consequences of improved load forecasting approaches have kept development of alternate, more accurate algorithms at the forefront of electric power research. Thus the significance of the subject in power systems has drawn alarming interests of many researchers, and to date a number of load forecasting approaches have been developed. Some of the most popular techniques are discussed in this chapter.
Generally, load forecasting models can be classified into two categories: time-of-day models and dynamic models. Time-of-day model is a non-dynamic approach and expresses the load at once as discrete time series consisting of predicted values for each hour of the forecasting period. The second classification involves the dynamic model that recognizes the fact that the load is not only a function of the time of the day, but also of the load most recent behavior.

The load can be represented mathematically as a function of different factors such as time, weather, customer class etc. In a typical additive model, the estimated load is given by:

\[ L^* = L_n + L_w + L_s + L_r \]  

(2.1)

Where \( L^* \) is the total estimated load, \( L_n \) represents the normal part of the load, \( L_w \) represents the weather sensitive part of the load, \( L_s \) is a special event component that creates a substantial deviation from the usual load pattern, and \( L_r \) is a complete random term, the noise.

The previous works presented that, better forecasting results could be realized by considering the electricity pricing in mathematical models. The representation of multiplicative approach can be formulated as:

\[ L^* = L_n + F_w + F_s + F_r \]  

(2.2)

Where \( L_n \) is the normal (base) load, and the correction factors \( F_w \), \( F_s \) and \( F_r \) are positive numbers that can increase or decrease the overall load. These corrections are based on current weather (\( F_w \)), special events (\( F_s \)) and random fluctuation (\( F_r \)) which include factors such as electricity pricing or load growth factor.
Forecasting the load using the proposed method involves weighted inputs which are passed through nonlinear transfers. Thus the proposed technique somehow uses a combination of additive and multiplicative methodologies to predict required load values.

2.2 Classification of forecasting methods

The following methods are widely used for load forecasting: similar day approach, regression models, time series, neural networks, expert systems, fuzzy logic, statistical learning algorithms, etc. These methods can be classified in terms of their degrees of mathematical analysis used in the forecasting model and are presented into two basic types, namely: quantitative and qualitative methods.

Large variety of statistical and artificial intelligence techniques have been developed and implemented by researchers to solve the load forecasting task. However, different techniques are widely used, namely; regression, time series, Artificial Intelligence, Neural networks, Expert systems, Fuzzy logic and Support vector machines.

2.2.1 The Regression Technique

Regression is the one of most widely used statistical techniques. For electric load forecasting regression methods are usually used to model the functional relationship of load consumption and other factors such as weather, day type, and customer class. A disadvantage of this technique is that the relationship between the weather components and the load demand is not stationary but rather depends on spatial-temporal components and the regression technique is unable to address this temporal variation [32].
Engle et al. [33] presented several regression models for the next day peak forecasting. Their models incorporate deterministic influences such as holidays, stochastic influences such as average loads, and exogenous influences such as weather.

### 2.2.2 The Time Series Technique

This technique takes a load pattern as a signal in a time series and forecasts the future load. In other words, the future load is only a function of the previous loads. The absence of weather components which strongly effect the energy consumption result in the forecasting being inaccurate and unstable especially when there is a drastic change in the environment (sociological variables) [25].

The ARMA (Auto Regressive and Moving Average models) models are the best example of this technique which assumes that the future load at any particular time can be estimated by a linear combination of a few previous times.

Time series methods are based on the assumption that the data have an internal structure, such as autocorrelation, trend, or seasonal variation. Time series forecasting methods detect and explore such a structure. Time series have been used for decades in such fields as economics, digital signal processing, as well as electric load forecasting. In particular, ARMA (autoregressive moving average), ARIMA (autoregressive integrated moving average), ARMAX (autoregressive moving average with exogenous variables), and ARIMAX (autoregressive integrated moving average with exogenous variables) are the most often used classical time series methods. ARMA models are usually used for stationary processes while ARIMA is an extension of ARMA to non-stationary processes. ARMA and ARIMA use the time
and load as the only input parameters. Since load generally depends on the weather and time of the day, ARIMAX is the most natural tool for load forecasting among the classical time series models.

Fan and McDonald [34] and Cho et al. [29] describe implementations of ARIMAX models for load forecasting.

Yang et al. [46] used evolutionary programming (EP is a method for simulating evolution and constitutes a stochastic optimization algorithm) approach to identify the ARMAX model parameters for one day to one week ahead hourly load demand forecast.

Yang and Huang [47] proposed a fuzzy autoregressive moving average with exogenous input variables (FARMAX) for one day ahead hourly load forecasts.

### 2.2.3 Artificial Intelligence and Neural Networks Techniques

The use of artificial neural networks (ANN) has been a widely studied electric load forecasting technique since 1990’s [43]. Neural networks are essentially non-linear circuits that have the demonstrated capability to do non-linear curve fitting. The outputs of an artificial neural network are some linear or non-linear mathematical function of its inputs. The inputs may be the outputs of other network elements as well as actual network inputs. In practice network elements are arranged in a relatively small number of connected layers of elements between network inputs and outputs. Feedback paths are sometimes used.

Recent progress in the applications of Artificial Neural Networks (ANN) technology to power systems in the areas of forecasting has made it possible to use this
technology to overcome the limitations of the other methods used for electrical load forecasting [35].

This is due to the fact that instead of relying on explicit rules or mathematical functions between past load and temperature variations, neural networks draw a link between input and output data. Thus the neural networks that deviate from relying on statistical models and large historical databases hold a good promise for the purpose of load forecasting. [27]

A neural Network approaches were applied to that area using a new Radial Basis Function Neural Network (RBFN) called Generalized Regression Neural Network (GRNN). This function has been used for short-term load forecasting using hour indicator (some indicators from 1 to 24 points to the significant difference in load magnitudes at different time of a day), day indicator (some indicators from 1 to 7 points to the significant difference in average load magnitudes at different days of the week), weather temperature data and electric price signal as inputs. Results show that the proposed Neural Network method is able to forecast accurate future loads using price and temperature as inputs [34].

Bakirtzis et al. [25] developed an ANN based short-term load forecasting model for the Energy Control Center of the Greek Public Power Corporation. In the development they used a fully connected three-layer feed-forward ANN and back propagation algorithm was used for training. Input variables include historical hourly load data, temperature, and the day of the week. The model can forecast load profiles from one to seven days.

Papalexopoulos et al. [42] developed and implemented a multi-layered feed-forward ANN for short-term system load forecasting. In the model three types of variables are
used as inputs to the neural network: season related inputs, weather related inputs, and historical loads.

Khotanzad et al. [37] described a load forecasting system known as ANNSTLF (Artificial Neural-Network-Based Electric Load Forecasting) which is based on multiple ANN strategies that capture various trends in the data. In the development they used a multilayer perceptron trained with the error back propagation algorithm. ANNSTLF can consider the effect of temperature and relative humidity on the load. It also contains forecasters that can generate the hourly temperature and relative humidity forecasts needed by the system. In the new generation, ANNSTLF includes two ANN forecasters, one predicts the base load and the other forecasts the change in load. The final forecast is computed by an adaptive combination of these forecasts.

Chen et al. [28] developed a three layer fully connected feed-forward neural network and the back propagation algorithm was used as the training method. Their ANN though considers the electricity price as one of the main characteristics of the system load.

2.2.4 Expert System Techniques

Rule based forecasting makes use of rules, which are often heuristic in nature, to do accurate forecasting. Expert systems, incorporates rules and procedures used by human experts in the field of interest into software that is then able to automatically make forecasts without human assistance. Expert systems work best when a human expert is available to work with software developers for a considerable amount of time in imparting the expert’s knowledge to the expert system software. Also, an
expert’s knowledge must be appropriate for codification into software rules (i.e. the expert must be able to explain his/her decision process to programmers).

Ho et al. [36] proposed a knowledge-based expert system for the short-term load forecasting of the Taiwan power system. Operator’s knowledge and the hourly observations of system load over the past five years were employed to establish eleven day types. Weather parameters were also considered. The developed algorithm performed better compared to the conventional Box-Jenkins method.

Rahman and Hazim [48] developed a site-independent technique for short-term load forecasting. Knowledge about the load and the factors affecting it are extracted and represented in a parameterized rule base. This rule base is complemented by a parameter database that varies from site to site. The technique was tested in several sites in the United States with low forecasting errors.

### 2.2.5 Fuzzy Logic Techniques

Fuzzy logic is a generalization of the usual Boolean logic used for digital circuit design. An input under Boolean logic takes on a truth value of 0 or 1. Under fuzzy logic an input has associated with it a certain qualitative ranges. For instance a transformer load may be low, medium and high. Fuzzy logic allows one to (logically) deduce outputs from fuzzy inputs. In this sense fuzzy logic is one of a number of techniques for mapping inputs to outputs (i.e. curve fitting).

According to the described applications of fuzzy logic to electric load forecasting [38], [44], [40]. It is noticed that , Among the advantages of fuzzy logic are the absence of a need for a mathematical model mapping inputs to outputs and the
absence of a need for precise (or even noise free) inputs. With such generic conditioning rules, properly designed fuzzy logic systems can be very robust when used for forecasting. Of course in many situations an exact output (e.g. the precise 12PM load) is needed. After the logical processing of fuzzy inputs, a defuzzification process can be used to produce such precise outputs.

### 2.2.6 Support Vector Machines

Support Vector Machines (SVMs) are a more recent powerful technique for solving classification and regression problems. This approach was originated from [45] statistical learning theory. Unlike neural networks, which try to define complex functions of the input feature space, support vector machines perform a nonlinear mapping (by using so-called kernel functions) of the data into a high dimensional (feature) space [26]. Then support vector machines use simple linear functions to create linear decision boundaries in the new space. The problem of choosing architecture for a neural network is replaced here by the problem of choosing a suitable kernel for the support vector machine [30].

Mohandes [41] applied the method of support vector machines for short-term electrical load forecasting. The author compares its performance with the autoregressive method. The results indicate that SVMs compare favorably against the autoregressive method. Chen et al. [26] proposed a SVM model to predict daily load demand of a month. Their program was the winning entry of the competition organized by the EU-NITE network. Li and Fang [39] also used a SVM model for short-term load forecasting. Lv et al. [11] used the least-square version of SVM regression to forecasting the daily peak load of Hefei region in China.
This chapter discusses the characteristics of power system loads and the effect of endogenous and exogenous factors on an electric load. These factors determine the amount of energy usage at a given time. In fact, the accuracy of a forecast depends predominantly on these factors. In addition, the latter part of this chapter covers general requirements of a good forecasting system.

### 3.1 Overview

An electric load refers to the power consumed by an electric circuit at its output terminal. Fig.3 shows a simple electric circuit presented by a voltage source $V_s$ in series with an internal resistance $R_s$, and then the load at the circuit terminal.
The set-up illustrated in Fig 3 can readily be generalized to a complete AC circuit by introducing line parameters such as reactive and capacitive impedances.

### 3.2 Characteristics of a power system load

The behavior of a power system load depends on a number of factors. Thus the load does not satisfy the superposition principle i.e. the load is not necessary the sum of linear independent variables, but it is rather a nonlinear system. As commonly known, nonlinear problems are often difficult to solve and much less understood than linear problems. Hence, power system load forecasting is a great computational problem for many researchers and network planners.

### 3.3 Factors influencing the load

The system load, in power operation context, is the sum of consumers’ load at the time. A consumer load trend is as different as ‘chalk and cheese’ and influenced by a thousand of factors. There is no any engineering rule that guides the selecting of these factors. Thus this process is mainly based on experience gained from the correlation analysis between the load and potential influencing factors.
The only existing criterion is that load forecasting in power systems can generally be divided into three different time horizons: short, medium, and long term load forecasting. However, factors affecting the load at different time horizons are not necessary the same. The variation in the short-term load depends heavily on time factors i.e. hour of the day, day of the week etc. Medium to long term load is determined by factors such as population growth, per capital income, demographic factors, gross domestic product (GDP) and so on. Most utility companies use this distinctive dependence in selecting the input variables.

Possible factors that may influence the load in various time horizons are discussed the following section.

3.4.1 Short-term load forecasting (STLF)

Generally, it is impractical to determine the major influencing factors for a certain forecast time horizon. However, some of the factors highlighted in Fig 4 could be considered for STLF.

![Diagram](attachment:image.png)

**Figure 4: Possible infusing factors – for STLF.**
• **Consumer category:** residential, commercial, and Industrial – the electrical usage trend is unique for customers that belong to different categories. Normally, industrial type customers consume a large part of electricity. Residential load are normally the smallest and often complex to determine due to different consumer electricity usage routines.

• **Meteorological related-factors:** temperature, humidity, cloud rate, and wind speed. Weather conditions influence the short term load greatly. Literature shows that, weather conditions have a strong correlation with the load, especially the temperature. In fact, load forecasting can be concluded using only weather related inputs.

  A general guideline on the effect of weather related conditions on the system load has been discussed in the work of (Say, 1976). Typically, the impact are as follow: a 1°C temperature change will produce a change in load of about 1 per cent, an overcast sky as compared to a clear sky will increase the load by 3-4 per cent, while thick fog may increase it by 10-12 per cent, and an increase of 1 per cent in load for every 4km/h wind velocity. Authors (Satish et al., 2004), report that STLF can be concluded by only using weather related conditions such as temperature.

• **Recent historical load pattern:** Despite the fact that load forecasting can be concluded by using limited input variables, the recent and corresponding past load trend is mainly the backbone of the forecast. The selected factors will then define the dimension of the training data set. Ideally, the training data set should cover the whole problem space as this enables the network to identify and map input-
output patterns adequately. The biggest problem here is to determine the optimal number of consumption instances, prior to the value to be predicted. Usually, the data points set is arbitrary determined, but taking into consideration the result obtained by using correlation analysis. However, the block entropy analysis performed by (Santos et al., 2006), reports that the use of long chains of contiguous load values does not offer any sort of benefit in the design of the IV of the ANN. For STLF, 1 week past contiguous load values could be adequate to design an IV structure and subsequently forecast the load with reasonable accuracy (Santos et al., 2007).

- **Time of the day:** The load has a cyclic characteristic. Unlike industrial load that is slightly stable, depending on production levels, the residential one changes greatly during the day. Peaks for a residential load are frequently observed in the mornings, mid days, and in the afternoons.

- **Day of the week:** i.e. weekdays or weekends. Generally, each day of the week has its unique electricity demand trend. Sometimes the days of the week could be selected as inputs to a forecasting model. (Adepoju et al., 2007) report that two load patterns namely: weekends and week days load patterns. Their result illustrates that the greatest peak loads are often recorded on Fridays, otherwise the load is slightly constant from Monday through Thursday.

- **Seasonal effects:** The seasonal changes have direct influence on system loads. For supply security and reliability, most utility companies consider worst case
scenarios (seasonal variation effects). The heating loads (in winter) and cooling loads (in summer) are normally viewed as maximum and minimum peak load thresholds respectively. The seasonal variation determines whether a utility company is summer or winter peaking (Gross et al., 1987).

- **Special events:** Depending on the magnitude of an episode in terms of power requirements, the associated demand could significantly influence the planned system load. For instance, the FIFA 2010 World Soccer Cup related activities may greatly influence the load in the South African power system. Often such events are thorny to consider in a load forecasting model.

- **Holidays:** Other factors that add uncertainties in the forecasts are religious or national holidays. Load variations due to these factors are minimal and often negligible, thus often weekend load curves are used instead.

- **Random disturbances:** Unlike residential consumers with smooth load curves, a slight change in industrial loads causes a significant load variation. For instance, if one considers the impact on the load as a result of starting-up or shutting down large loads such as smelters, wind tunnels, or steel mill, indisputably the change is not negligible.

- **DSM (Demand Side Management) Initiatives:** This approach entails actions that influence the quantity or pattern of use of electric energy consumed by end users. A good example would be the prevailing power shortage especially in the South African region which has certainly forced power utility companies in the region to implement ongoing and aggressive DSM initiatives such as compact
fluorescent lamps (CFL) lighting, the use of energy efficient motors, implementation of time-of-use (TOU) tariffs, residential load management.

Other demand management strategies include emergency generation, independent power production, and industrial cogeneration. Although DSM is a short term strategic approach, it helps the utilities to defer the expenditure associated with additional generation capacity by reducing total end-user demand (Boake, 2003). These measures are required immediately to minimize the risk of load shedding of non-essential load until the additional generation capacity can be built, but the reality is that they will be required for the foreseeable future (Ross et al., 2008)

The forecast time horizon determines which factors could be selected in the design of input vector of a forecasting model. For medium-to-long term load forecasting the factors discussed below may need to be considered.

### 3.4.2 Medium-to-long term load forecasting (M-LTLF)

Contrary to the factors affecting STLF, medium-to-long term load forecasting involves numerous uncertainties that are often tricky to predict. Some of these factors are indicated in Fig. 5

![Medium-to-long term load forecasting diagram](image)

**Figure 5: Typical influencing factors – for M-LTLF**
a) **Economic growth:** Depending on the economic growth indicator such as GDP (Gross Domestic Product) of a particular state, an increase in the economic growth would ideally raise the demand (NIEIR, 2007). In fact, economic related factors influence long term load trend greatly, thus the consideration requires a special attention in a long term load modeling.

b) **Energy and environmental policies:** Energy policies (often corresponding with international laws) have direct influences on the system load. This involves policies aimed at promoting energy efficiency, and development of off-grid power generations. Among others, most of the utility companies are fully conversant of the environmental policies i.e. greenhouse polices and/or emission trading schemes. For a utility to operate within these policy bounds, it implies that additional electricity supply resources are required and thus it influences the electricity demand growth.

c) **Technological change:** The change in technology may positively or negatively influence the load. For instance, if one considers an aggressive use of gas air conditioning technology which will as result offset the electric conditioning systems, thereby reducing the system load.

d) **Per capital income:** Often income rate per household determines the corresponding space comforts. Consumer preferences (heating or cooling loads) rely heavily on real income ranges, thus the demand is either positively or negatively influenced.
e) **Demographic factors:** A raise in the population of a certain locality would obviously result in an increase in the demand. For a single locality, the impact seems to be relatively small but the lump-sum effect is surely not negligible.

f) **Customer category:** The impact of a large industrial load growth could affect the forecast in the long run, thus this should be considered.

g) **Political factors:** Politically motivated wars in a particular state can influence inhabitant’s decisions to immigrate to other countries. The movements directly affect the population, and subsequently affect the load negatively and vice versa. Other stakeholders such as independent power producers (IPP) may also find it senseless to invest and operate in a peace challenged country.

**Others:** For example, upstream developments, such as a construction of a new generation plant or regional inter-connectors, will certainly influence electricity prices and demand through final retail prices. This implies that the electricity tariffs are highly likely to be increased, hence affordability determines usability.

### 3.4 Load features

#### 3.5.1 Types of load banks

An electric load is formed by load banks. There are three common types of load banks: resistive load bank, reactive load bank, and capacitive load bank. Load banks are used for different purposes. A general example for a resistive load would be: the conversion of electric energy to heat by power resistor in a circuit. For a resistive
load, the angle between the current and voltage is zero or in-phase, thus the circuit always operates at a unity power factor.

Reactive load bank includes inductive apparatus such as induction motors, transformers, lighting etc, and has a lagging power factor. Capacitive load bank is somehow similar to the reactive load bank in rating and purpose, except that a leading power factor load is introduced.

### 3.5.2 Load categories

Power system loads can be divided into three categories, namely: linear loads, nonlinear loads, and special loads. For linear loads, the load impedance is always constant regardless of the applied voltage and the load current increases proportionally as the voltage increases and vice versa. Typically, linear loads include motors, incandescent lighting, heating loads etc.

For instance, starting a large motor has a great impact on power system load, especially for direct on line (DOL) motors due high starting currents. For a DOL, the starting current is normally 7 times higher than the rated full load current. Thus this type of loads may contribute significantly to cost intensive high peak demands. Other soft-starting techniques (i.e. star-delta, rotor resistance, auto transformer starting etc.) are widely used for economic operations of electric machines.

Load currents in electronic loads such as computers, uninterrupted power supply (UPS), variable speed motor drive, thyristor controlled equipment etc. are neither proportional to the instantaneous voltage nor continuous. These loads are typical
examples of nonlinear loads. Other loads, such as motors used in lift applications, have different nature of operations i.e. here motors are frequently started. Therefore, general design standards specify lower starting currents. Typically, starting current of a lift motor should be less than 75% of the rated current. The contribution of special loads and nonlinear loads to the overall power system load often falls within the average of the load base.

3.5.3 Consumer classes

Electric power is used by different types of consumers: residential, commercial and industrial. Load profile for each consumer class is unique and it depends mainly on types of appliances used, daily activity patterns. Residential consumers generally use electric energy for heating, cooling, and illumination and consume about half of the electric power. Commercial businesses purchase electricity for a variety of commercial functions i.e. office machinery, store display lighting, parking bay lighting, escalators etc. Whilst industrial facilities and plants need electricity to power processes such as compressor, conveyor motor, air conditioning, and other manufacturing applications. This distinction of various consumer classes is essential for electric utilities in determining electricity prices, because rates charged for each class are different.

The first section of this chapter gives a comprehensive introduction to the power system load. The following section now discusses some requirements for a good and reliable forecasting system.
3.5 **Prerequisites of a good STLF System**

Most of demand or load management programs used by electric utilities comprise STLF units. Every utility intends to have a reliable STLF system for economical operations of power systems. The reliability and robustness of the system primarily depend on the accuracy of the forecasts. Though, there are other important requirements for a good STLF system. These requirements take account of:

- Fast speed
- Accuracy
- Automatic data access
- Friendly interface
- Timely forecast
- Automatic performance evaluation of the obtained forecast
- Automatic bad data detection and forecasting report generation.

### 3.5.1 Fast Speed

Forecasting programs with different techniques, network topologies, training algorithm, desired goals, etc. have diverse convergence periods. This means that the speed of the forecasting depends on the system structure. The idea is to put together an optimal forecasting structure that yields accurate results in the shortest period.

### 3.5.2 Accuracy

Accurate forecasts drastically benefit utilities and other stakeholders in a number of ways. In fact, the main goal of a majority of review papers on STLF is to produce forecasts with a reasonable degree of accuracy. In essence in power generation
sector, just a simple decision with regards to price based unit commitment (PBUC) requires an accurate STLF system.

3.5.3 Automatic data access

Load forecasting models or systems normally require a number of inputs, namely: historical load data, weather related inputs, sinusoidal functions and so on. Thus there exists a need of a database where all relevant model inputs are stored. This implies the system should be able to automatically access and extract required data from the database.

The database is principally for historical data. Some inputs such as weather parameters may need to be captured automatically on-line. For this process, the use of some communication options (internet, telemetry, or external modems) could be enforced. The data base is also used to store the results from the developed forecasting models.

3.5.4 Friendly interface

The forecasting model should be easy to use, a way as possible from complications and convenient to wide range of users with different levels of knowledge.

3.5.5 Timely forecast

If a network operator gets forecasts, even accurate forecasts, from the planning department but outdated, it’s highly likely that the operator will implement unrealistic decisions. Therefore, a good STLF system should yield forecasts which are timely and corresponding to the desired forecast lead periods.
3.5.6 Automatic model performance evaluation

Because accuracy is a key measure in forecasting, the design structure of a good forecasting system should specify some permissible error limits. In this work, the standard mean squared error is employed. The error bound is defined as: %5.2 \leq \text{MSE}. However, this performance measure condition is not applicable when the average forecast error is consistently smaller than these limits.

3.5.7 Bad data detection

Out of range data (outliers) is one of the major challenges in computational problems. In load forecasting for example, historical data are gathered and stored in a database. However, some these data especially estimated data, can fall beyond the reasonable data boundaries, hence a bad data detection strategy should be formulated and implemented to identify and replace the bad data. In the early days, STLF systems relied heavily on the power system operators to interpret the data and subsequently get rid of the bad data (Yang et al., 2006). Detecting bad data manually is a work burden and demanding exercise for operators, thus automated bad data detection can accelerate the forecasting process.

3.5.8 Automatic forecasting report generation

Forecasting results should be presented in a manner that is easily understood and interpreted by the end-user. In simple terms, the results may need to be illustrated in numerical and graphical forms.
3.6 Conclusion
To briefly summarize the preceding chapters, the following concepts have been covered: firstly, a general overview on the subject topic was discussed, and then a review on the existing methods and drawbacks of these techniques were explored. Secondly, here the third chapter covers the basic requirements of a load forecasting model.

The next part of this document now introduces the proposed forecasting method. In other words, (chapter five) now covers the concept of Support vector regression method in general context, training, testing and on line updating paradigms etc. as an attempt to help the reader in understanding the Support vector regression method.
This chapter discusses the definition and characteristics of smart grids and the main goals and benefits of smart grid, and the effect of these features on the power consumption and distribution.

4.1 Introduction

The current electrical grid is seen by some as the greatest engineering achievement of the 20th century, and is considered to be the largest machine on the planet. However, it is increasingly out of date and overburdened, leading to costly blackouts and brownouts. For example, the 2003 blackout in the northern and eastern U.S. and Canada caused a $6 billion loss in economic revenue [1].

Utility providers struggle to monitor the grid’s performance, and in many areas, they depend on customers to report power outages. This is because the grid was designed to meet one main goal — to ensure that the lights kept glowing — before much of the
technology that depends on it existed. It was not originally designed to incorporate other goals such as energy efficiency, reduced environmental impacts, incorporating alternative energy sources, allowing for more consumer choice, and robust cyber security.

The dawn of the 21st century brought us the gift of communication and automation along with ever increasing need for electricity. Centralized electrical supply networks are limited to fossil fuel resources and prone to complete system failures. Decentralized solutions (Micro-Grids) for provision of electricity to every part of the world are gaining popularity especially among the developing countries. These systems can be powered through fossil fuel or renewables and can be linked together to form an electrical network. Integration of these decentralized solutions assisted with high-tech communication systems to build a network brings up the idea of smart grids.

Modernization efforts are underway to make the current electrical grid smarter. The infrastructure that will support the future Smart Grid that, to be capable of informing consumers of their day-to-day energy use, even at the appliance level.

4.2 What is Smart Grid?

In order to fully understand the challenges and issues, it is important to define what a 'Smart Grid' is.

Smart Grid refers to the modernization of the current electrical grid by utilizing information, communication, and control technologies so that to have bi-directional flow of information and electricity to intelligently integrate the actions of all users.
CHAPTER 4 SMART GRID

connected to which - generators, consumers and those that do both – in order to efficiently deliver sustainable, economic and secure electricity supplies [3].

The smart grid is characterized by several new trends and features: smart meters, demand response mechanisms, online customer interactions though PCs/mobile devices, dynamic electricity tariffs, online billing, incorporation of renewable energy generation (such as solar and wind energy) and electric vehicles, more reliable power transmission and distribution, dynamic load balancing, better power quality, better power security, etc. The architecture and components of the smart grid are illustrated in Fig. 6

![Figure 6: Architecture and components of the smart grid.](image)

Communications technology and infrastructure is at the heart of improvements to the electrical grid, which will collate data provided by smart meters, sensors, computer systems, and many other devices into understandable and actionable information for
consumers and utilities [3]. While Information Technology (IT) is one of the major driving forces behind a smart grid as well as various IT systems and techniques such as artificial intelligence, high performance computing, simulation and modeling, data network management, database management, data warehousing, and data mining are to be used to facilitate smooth running of the smart grid [2].

Here are some other definitions of the Smart Grid:

- A power system that contains multiple automated transmission and distribution (T&D) systems, all operating in a coordinated, integrated, efficient, and reliable manner.
- A power system that serves millions of customers and has an intelligent communications infrastructure, enabling the timely, secure, and adaptable information flow, needed to provide power to the evolving digital economy.
- A power system that handles emergency conditions with ‘self-healing’ actions and is responsive to energy-market and utility needs.
- The smart grid is a broad collection of technologies that delivers an electricity network that is flexible, accessible, reliable and economic. Smart grid facilitates the desired actions of its users and these may include distributed generation, the deployment of demand management and energy storage systems or the optimal expansion and management of grid assets [4].

From the Information Technology point of view the Smart Grid technology will significantly increase the amount, quality, and use of data received from various sensors and meters. This will solve two of today's main problems in the grids: environmental concerns and power disturbances. Introduction of Smart Grid will increase both security and efficiency of the supply. New software, implemented in
various microcontrollers will help to avoid grid congestions and enable distributed
generation, making accent on the use of renewable energy resources. Imagine a
network, in which a customer can manage his consumption and take advantage of
pricing schemes, while being able to choose the type of electricity supply (i.e. 100%
renewable energy, conventional energy, mixed mode).

From Energy Industry Point of View the real-time two-way communications available
in a Smart Grid will allow customers to be compensated for their efforts to save
energy and to sell energy back into the grid through Advanced Metering technologies.
After spreading distributed generation concepts such as residential solar panels and
small wind turbines, the Smart Grid will improve the efficiency of energy industry by
providing green energy recourses and reducing peak loads. It will allow small
domestic customers and businesses to sell power to their neighbors or even back into
the distribution grid. The same concept can be applied to larger commercial
organizations that have renewable power systems that can give the excess power back
into the grid during peak demand hours.

The purpose and possibilities of the smart grids could be introduced as a side-by-side
comparison of the traditional and smart grids.

<table>
<thead>
<tr>
<th>Traditional Grids</th>
<th>Smart Grids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centralized power generation</td>
<td>Distributed power generation</td>
</tr>
<tr>
<td>One-way power flow</td>
<td>Two-way power flow</td>
</tr>
<tr>
<td>Empirical-based operation (load indexes)</td>
<td>Renewable power generation</td>
</tr>
<tr>
<td>Limited grid accessibility for new producers</td>
<td>Loads follow generation</td>
</tr>
<tr>
<td></td>
<td>Operation based on real-time data</td>
</tr>
</tbody>
</table>
4.3 Main Goals of smart Grid

The infrastructure platform of Smart Grid enables functioning of different technologies and systems to make intelligent decisions for supplying consistent electricity to consumers) in order to different goals, some of these goals will be listed below:

- Provide consumers with more choices on how, when, and how much electricity they use.
- Self-heal in case of disturbances, physical and cyber-attacks, and natural disasters.
- Link with a wide array of energy sources, in addition to energy produced by power plants, such as renewable energy producers.
- Provide better power quality, and more efficient delivery of electricity.
- Being financially profitable to both generators and distributors.

4.4 Smart Grid vision

The overall vision for the Smart Grid is that it will possess the following qualities: [4]

- **Intelligent** – capable of sensing system overloads and rerouting power to prevent or minimize a potential outage; of working autonomously when
conditions require resolution faster than humans can respond and cooperatively in aligning the goals of utilities, consumers and regulators (Proactive rather than reactive and operate in adaptive and scalable manner).

- **Efficient** – capable of meeting increased consumer demand without adding infrastructure (Optimized for best resource and equipment utilization).

- **Accommodating** – accepting energy from virtually any fuel source including solar and wind as easily and transparently as coal and natural gas; capable of integrating any and all better ideas and technologies – energy storage technologies, for example – as they are market-proven and ready to come online (Open for all types and sized of generation).

- **Motivating** – enabling real-time communication between the consumer and utility so consumers can tailor their energy consumption based on individual preferences, like price and/or environmental concerns (Interactive among customers, retailers, and markets).

- **Opportunistic** – creating new opportunities and markets by means of its ability to capitalize on plug-and-play innovation wherever and whenever appropriate.

- **Quality-focused** – capable of delivering the power quality necessary – free of sags, spikes, disturbances and interruptions – to power our increasingly digital economy and the data centers, computers and electronics necessary to make it run.

- **Resilient** – increasingly resistant to attack and natural disasters as it becomes more decentralized and reinforced with Smart Grid security protocols (Self-healing by its ability to predict/distinguish/bypass abnormal situations).
• **Green** – slowing the advance of global climate change and offering a genuine path toward significant environmental improvement (incorporating renewable energy resources).

Smart Grid technologies are expected to produce valuable cost and energy efficiencies all along the electrical distribution system. One of the first and most important will be to meet peak energy demand more efficiently and with less detriment to the environment. Since storage of electricity is currently very costly, electricity must be consumed the moment it is created. As a result, estimating the correct amount of demand of electricity is very difficult. Without the benefit of this knowledge, electricity providers must use peaker plants when energy demand threatens to exceed supply levels. Peaker plants tend to be older, expensive to bring online, and require fuel to operate, which further creates greenhouse gases.\[5\]

So how does the Smart Grid vision relate to identifiable individuals? [8]

<table>
<thead>
<tr>
<th>With Smart Grid technology people will be able to:</th>
<th>And they will do it by:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understand how their household uses energy, manage energy use better, and reduce their carbon footprint.</td>
<td>Logging into their energy use account and seeing how much energy they are using in real time, and as compared to their neighbors, as reported by smart meters installed at each household.</td>
</tr>
<tr>
<td></td>
<td>Using smart devices, such as a smart thermostat that shows minute-by-minute price of energy. The thermostat could be</td>
</tr>
<tr>
<td>Control expenditure on electricity.</td>
<td>Accessing their account balance, and seeing how many units are being used per day, and which appliances are costing the most money. Taking advantage of energy saver plans offered by the utility to keep energy use in line with a person's budget. For example, if a heat wave hits and the price of electricity peaks, the individual could be notified that they may exceed their budget. The individual would then be in control regarding whether the utility could adjust the temperature of the air conditioning a few degrees when peak energy consumption occurs.</td>
</tr>
<tr>
<td>Experience fewer and shorter power outages, and to be notified when</td>
<td>Having the Smart Grid pinpoint the location of the outage and dispatch workers to the scene immediately. Power will be routed around the outage, so that fewer individuals are</td>
</tr>
<tr>
<td>the power will come back on.</td>
<td>affected by the power outage.</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Signing up to receive alerts when the power goes out via text message to a mobile phone regarding when the power will be back on. Additional messaging services could provide alerts regarding a loved one’s energy restoration time.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control energy devices in the home.</th>
<th>Tying all energy devices that give energy back to the grid, such as a plug-in hybrid vehicle and solar panels, to a central household control which provides up-to-the-minute indication of energy use.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring whether their home is using more energy than it is producing, and adjusting devices so they use less energy. The smart meter tracks this activity, and any surplus in energy shows up as a credit on the person’s utility bill.</td>
<td></td>
</tr>
<tr>
<td>Controlling smart devices and account information over the Internet, allowing individuals to monitor and adjust their house’s energy usage remotely.</td>
<td></td>
</tr>
</tbody>
</table>

Nowadays, it is a priority of many governments worldwide to replace/upgrade their several decades old electricity grids with smart grids. For example, in 2010, the US government spent $7.02B on its smart grid initiative, while the Chinese government used $7.32B for its smart grid program [20].
Significant development towards achieving a vision for a Smart Grid is occurring widely in North America and Europe. For example, the U.S. economic stimulus plan included the passing of the American Recovery and Reinvestment Act of 2009, which is allocating billions of dollars to fund improvements to the electrical grid [9]. The Department of Energy has dispersed 3.4 billion in Smart Grid grants which will lead to the rollout of 18 million smart meters, 1 million in-home management systems, and advanced load management devices. [10]

Homes will be outfitted with smart metering devices and each ratepayer will have access to an in-home energy management website, which will monitor energy use and provide information and recommendations for lowering energy use and costs [12]. Smart Grid technology is also being supported by significant private investment. Companies are pursuing new products in the area of electric vehicles, smart appliances, and energy production technology, such as solar panels for household roofs, as well as new service offerings in the management of energy capacity, location, time, rate of change and quality. Significant private sector investment is occurring with Smart Grid venture capital, valued at over $900 million between the years 2000 and 2008 [19].

Morgan Stanley estimates that the Smart Grid market will be $100 billion in 2030 [20].

### 4.5 Summary

By definition, smart grids are electrical networks that can make intelligent decisions for supplying consistent electricity to consumers while being financially profitable to both generators and distributors. Due to decentralized nature of smart grids, a
consumer can be a generator at the same time. Henceforth, new control techniques for controlling flow of electricity are needed. Electrical devices require either active power, reactive power or both based on their resistive or inductive nature. In conventional grids, generators are designed to provide maximum active power. However, in smart grids new control techniques are needed to balance both active and reactive power needs of the load. This is needed because the generator is satisfying the needs of the local electrical loads while supplying power to the grid and also in case of disconnection from the grid, a condition termed as islanding.

In order for smart grids to be financially profitable, access and availability of system information at generation, distribution and consumption level is crucial. This system information comprises of electricity supply and demand patterns and their effect on electricity price. In conventional grids, this is achieved through metering and by models based on load patterns and weather. However, information access is limited to generators and distributors. In smart grids, consumers also have access to this system information and can also act as a generator or distributor. They can use this information to meet their needs by employing intelligent devices. These devices can either be consumer end devices such as smart washing machines that process laundry at minimal electricity rates or devices that controls the flow of electricity to the grid for making profit.
CHAPTER 5

Proposed Solution

This chapter discusses the main part of this work which is the proposed solution using support vector regression model. In addition reviews the support vector machine classification method to introduce the reader a good background of the method.

5.1 Summary

As in the previous work [3] and [11], we try to approach the problem of smart grid’s load forecasting from the data mining perspective. In particular, we propose a peak load forecasting model based on the data mining technique of support vector regression (SVR) using least squares [15]. More specifically, we use the online least-squares SVR proposed by Engel et al. [7].

In order to predict the peak load $P_d$ of a particular day $d$, we can roughly consider $P_d$ as a non-linear combination of a number of attributes from different sources: peak loads of previous $N$ days, average temperatures of previous $N$ days, holiday records of
previous \( N \) days, forecasted temperature of day \( d \), and whether day \( d \) is a holiday (weekend or public holiday).

So, for each target peak load value in the historical record, we construct a feature vector covering the abovementioned attributes associated with the target. Then, we train our least-square SVR system using a set of <feature vector, target> pairs for a large enough number of days. The result of this training process is a least-squares regressor model.

We can use the resultant regressor model to forecast the peak load value \( P_d \) of a given day \( d \). For that, we have to construct a feature vector for the day \( d \) in the same manner as in the training step. In constructing the feature vector, we need to know the forecasted temperature of the day \( d \) (if it is in the future) and whether it is a holiday (which can be easily known in advance). Then, the feature vector of day \( d \) is supplied to the regressor model to generate the forecasted peak load value of that day.

After the day \( d \) is already passed and its actual peak load value (the target) already known, the regressor model is updated with the <feature vector, actual target> pair for the day \( d \), thus resulting in a fresh model which best reflects the latest trend of events.

A schematic representation of our proposed load forecasting scheme is given in Figure 2: Overview of the proposed load forecasting scheme.
Figure 7: Overview of the proposed load forecasting scheme.

We tested our proposed method using the smart metering data from the two regions in Germany for the year 2009. Experimental results demonstrate that our approach is both accurate and computationally efficient.
The remaining of the sections is organized as follows. This chapter describes our proposed load forecasting method using online least-squares support vector regressions in details.

### 5.2 Support Vector Machine

Support vector machine (SVM) [16] is a kind of maximum margin classifier which was originally proposed to solve the problem of binary classification. Among a large number of training data vectors, only a few are selected as support vectors that define the maximum margin. Only the support vectors are utilized in predicting the classes of the testing data vectors, thus leading to a good generalization.

Later, it was realized that SVM can be adapted to solve the problem of regression [13]. Suykens and Vandewalle proposed a least-squares version of support vector regression (SVR) [15] which is particularly suitable to solve regression problems in time series data. The least-squares SVR tries to find the solution by solving a set of linear equations instead of a convex quadratic programming for classical SVMs.

A brief description of the least-squares SVR is given in the below subsections. This description is adapted and modified from the original ones given in [10], [15], and [19].

### 5.3 General SVM Formulation

Suppose we have a training set of n samples \( \{x_i, y_i\} (i = 1, \ldots, n) \) with input data vector \( x_i \in \mathbb{R}^m \) (where m is the dimensionality of \( x_i \)) and corresponding binary class labels \( y_i \in \{-1, +1\} \). In Vapnik’s original formulation [16], the SVM classifier is defined by the conditions:
\[ \begin{align*}
  w \cdot \phi(x_i) + b & \geq 1, \quad \text{if } y_i = +1 \\
  w \cdot \phi(x_i) + b & \leq -1, \quad \text{if } y_i = -1
\end{align*} \]

These formulations can be rewritten as a single condition:

\[ y_i (w \cdot \phi(x_i) + b) \geq 1, \quad i = 1, \ldots, n \]  \hspace{1cm} (2)

Where \( \phi(x) \) is a nonlinear mapping function of a vector from original space to the high (possibly infinite) dimensional space, \( w \) is a weight vector which defines the separation hyperplane, and \( b \) is an offset of the separation hyperplane from the origin \((0, 0)\).

If the given data set is inseparable (i.e., separating hyperplane does not exist), a slack variable \( \xi_i \) is introduced in such a way that:

\[ \begin{align*}
  y_i (w \cdot \phi(x_i) + b) & \geq 1 - \xi_i, \quad i = 1, \ldots, n \\
  \xi_i & \geq 0, \quad i = 1, \ldots, n
\end{align*} \]

By applying the structural risk minimization principle, the risk bound (i.e., learning error) of the classifier can be minimized by solving the following minimizing problem:
\[
\min \ J_1 (w, \xi) = \frac{1}{2} \| w \|_2^2 + C \sum_{i=1}^{n} \xi_i
\]  

(4)

Subject to the constraints:

\[
y_i [w \cdot \phi(x_i) + b] \geq 1 - \xi_i, \quad i = 1, \cdots, n
\]

\[
\xi_i \geq 0, \quad i = 1, \cdots, n
\]

Where \( C \) is the slack penalty parameter to control the net effect of the slack variables.

In order to remove the complex constraints of the above minimization problem in Equation (4), we introduce Lagrangian multipliers \( \alpha_i \geq 0 \) (\( i = 1, \cdots, n \)) [4]. Thus, the minimization problem becomes:

\[
\min \ L_1 (w, \xi, \alpha) = \frac{1}{2} \| w \|_2^2 + C \sum_{i=1}^{N} \alpha_i (y_i (w \cdot \phi(x) + b) - 1 + \xi_i)
\]  

(5)

Subject to the constraints:

\[
\xi_i \geq 0, \quad i = 1, \cdots, n
\]

The optimal point will in the saddle point of the Lagrangian function. Thus, we have:

\[
\frac{\partial L_1}{\partial w} = 0 \implies w = \sum_{i=1}^{n} \alpha_i \phi(x_i)
\]

\[
\frac{\partial L_1}{\partial b} = 0 \implies \sum_{i=1}^{n} \alpha_i y_i = 0
\]

\[
\frac{\partial L_1}{\partial \xi_i} = 0 \implies 0 \leq \alpha_i \leq C, \quad i = 1, \cdots, n
\]

(6)
By substituting $w$ by its expression, we get the following quadratic programming problem:

$$\max Q_1(\alpha) = \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$  \hspace{2cm} (7)

Subject to the constraints:

$$0 \leq \alpha_i \leq C, \quad i = 1, \cdots, n$$

Here, $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ is called the kernel function (which will be elaborated below in Section II.C). By solving this quadratic programming problem subject to the constraints, we will get the separating hyperplane in the high dimensional space, that is, the classifier in the original space.

### 5.4 Least Squares SVM Formulation

Suykens and Vandewalle derived the least squares version of the SVM classifier by reformulating the minimization problem as below [15]:

$$\min J_2(w, b, e) = \mu \left( \frac{1}{2} \| w \|^2 \right) + \zeta \left( \frac{1}{2} \sum_{i=1}^{n} e_i^2 \right)$$  \hspace{2cm} (8)

Subject to the equality constraints:
\[ y_i(w \cdot \phi(x_i) + b) = 1 - e_i, \quad i = 1, \ldots, n \]

The least-squares SVM classifier formulation above implicitly corresponds to a regression interpretation with binary targets \( y_i = \pm 1 \).

Both \( \mu \) and \( \zeta \) are parameters to tune the amount of regularization versus the sum squared error. The solution does only depend on the ratio \( \gamma = \mu / \zeta \), therefore the original formulation uses only \( \gamma \) as tuning parameter. Therefore, we have:

\[
\min J_2(w, b, e) = \frac{1}{2} \| w \|^2 + \gamma \sum_{i=1}^{n} e_i^2
\]  (9)

The solution of the least-squares regressor is obtained after the Lagrangian function is constructed as follows:

\[
L_2(w, b, e, \alpha) = J_2(w, b, e) - \sum_{i=1}^{n} \alpha_i (y_i(w \cdot \phi(x_i) + b) - 1 + e_i)
\]

\[
= \frac{1}{2} \| w \|^2 + \gamma \frac{1}{2} \sum_{i=1}^{n} e_i^2 - \sum_{i=1}^{n} \alpha_i (y_i(w \cdot \phi(x_i) + b) - 1 + e_i)
\]  (10)

Where \( \alpha_i \in \mathbb{R} \) (\( i = 1, \ldots, n \)) are the Lagrange multipliers. Again, the conditions for optimality are:
\[
\frac{\partial L_2}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{n} \alpha_i y_i \phi(x_i)
\]

\[
\frac{\partial L_2}{\partial b} = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0
\]

\[
\frac{\partial L_2}{\partial e_i} = 0 \Rightarrow \alpha_i = \gamma e_i, \quad i = 1, \cdots, n
\]

\[
\frac{\partial l_2}{\partial \alpha_i} = 0 \Rightarrow y_i (w. \phi(x_i) + b) - 1 + e_i = 0,
\]

\[
\quad i = 1, \cdots, n
\]

By the elimination of \(w\) and \(e\), we will have a linear programming problem instead of a quadratic programming one:

\[
\begin{bmatrix}
0 & y^T \\
y & \Omega + \gamma^{-1} I_n
\end{bmatrix}
\begin{bmatrix}
b \\
\alpha
\end{bmatrix} =
\begin{bmatrix}
0 \\
1_n
\end{bmatrix}
\]   \hspace{1cm} (12)

where \(y = [y_1, \cdots, y_n]\), \(1_n = [1, \cdots, 1]\), \(\alpha = [\alpha_1, \cdots, \alpha_n]\), and \(I_n\) is an \(n \times n\) identity matrix.

Here \(\Omega \in \mathbb{R}^{n \times n}\) is the kernel matrix whose individual element \(\Omega_{ij}\) (\(i, j = 1, \cdots, n\)) is defined as follows [6], [10].

\[
\Omega_{ij} = \phi(x_i, x_j) = K(x_i, x_j)
\]   \hspace{1cm} (13)
5.5 Kernel Function

For the kernel function $K(\bullet, \bullet)$ one typically has the following choices [10], [19]:

Linear kernel:

$$K(x_i, x_j) = x_i \cdot x_j$$  \hspace{1cm} (14)

Polynomial kernel of degree $p$:

$$K(x_i, x_j) = \left( \frac{x_i \cdot x_j}{c} \right)^p$$  \hspace{1cm} (15)

Radial basis function (RBF) kernel:

$$K(x_i, x_j) = \exp \left( - \frac{\| x_i - x_j \|^2}{\sigma^2} \right)$$  \hspace{1cm} (16)

Multi-layer Perceptron (MLP) kernel:

$$K(x_i, x_j) = \tanh(k \cdot x_i \cdot x_j + \theta)$$  \hspace{1cm} (17)

Where $p, c, \sigma, k$ and $\theta$ are constants. Here, Mercer’s condition holds for all $c, \sigma \in \mathbb{R}^+$ and $p \in \mathbb{N}$ values in the polynomial and RBF case, but not for all possible choices of $k$ and $\theta$ in the MLP case. The scale parameters $c, \sigma$ and $k$ determine the scaling of the inputs in the polynomial, RBF and MLP kernel functions. This scaling affects the
kernel’s bandwidth, which is an important factor in generalization of a kernel method [10], [19].

5.6 Online Learning

In our proposed method, we employ a version of least-squares SVR, namely the online SVR proposed by Engel et al. [7]. In this online learning setup, we first train our least-squares SVR system with a large enough set of data in a batch mode. Then, we deploy the system for regressing (forecasting) an unknown future data. When the actual value of the forecasted data is came to know, the SVR system is updated using this actual data. In this way, the SVR system is always up-to-date and can truthfully represent the latest trend of the data. Therefore, online learning enables us to reduce the effect of the concept drift phenomenon [18] which usually occurs in time-series data, and thus improving the accuracy of forecasting.

5.7 Feature Vector Construction

In order to forecast the peak load \( P_d \) of a given day \( d \), we construct (encode) a feature vector with 32 attributes as listed in Table 2. These 32 attributes are empirically chosen.

<table>
<thead>
<tr>
<th>Attribute ID</th>
<th>Feature Description</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 28</td>
<td>Peak load of previous 28 days</td>
<td>( P_{d-1} ) to ( P_{d-28} )</td>
</tr>
<tr>
<td>29</td>
<td>Average peak load of previous 7 days</td>
<td>( 1/7 \ (P_{d-1} + P_{d-2} + \ldots ) )</td>
</tr>
</tbody>
</table>

\( + P_{d-7} \)
Table 2: Feature vector used in forecasting the peak load $P_d$ of a given day $d$.

The individual and the average peak load information can be obtained from the given peak load data set itself. The historical temperature information for different regions of the world can be extracted from the Weather Underground website [24]. Holiday information of countries all over the world is available from the Holidays-Info.com website [22].

We use the scaling facility of the LibSVM software [21] to map the values of each attribute into the range of $-1$ to $+1$. This scaling exercise helps us improve the forecasting accuracy by a considerable extent. The experimental results of scaling vs. without scaling are discussed in the next chapter.
Experimental Results

In this Chapter, we will discuss about the datasets that we use in our experiment, how the datasets are of training and testing subsets, the results that our proposed method achieved in comparison with an existing state-of-the-art method, the effects of scaling vs. non-scaling, and finally the computational efficiency of the method.

6.1 Experiment on Germany Data

We use two datasets in our experiment. The first one is the electricity usage data for the year 2009 logged by a smart meter deployed in a household in Lower Saxony (LS) region of Germany. We will call this dataset as LS Dataset. The second one is the data for 2009 logged by a smart meter installed in a household in North Rhine-Westphalia (NRW) region of Germany. We will name this dataset as NRW Dataset. Each original datasets contains the electricity usage readings of the smart meter at every 15 minutes. From these readings, we extract the peak load (i.e., the maximum reading) for each day. The daily peak load profile of LS Dataset for the whole year of 2009 is illustrated in Fig. 8.
Figure 8: The Daily peak load profile in Lower Saxony (LS), Germany in 2009. The spikes indicate the increases of peak loads on holidays. Electricity usage is higher in winter and lower in summer.

Training and Testing Sets

For each dataset, we use each daily peak load value and its associated feature vector (as discussed before) from February 01, 2009 to June 30, 2009 (150 days) as the training data for our forecasting system. (Note: we simply cannot start from January 01, 2009 because we need the data for the previous 4 weeks, i.e., 28 days, to construct a feature vector.)

The remaining days of the year from July 01, 2009 to December 31, 2009 (184 days) are used for testing (as well as for model updating in our online learning setup).

We use dlib C++ library [17] for the implementation of online least-squares SVR algorithm by Engel et al. [7]. We use a radial basis function (RBF) kernel, which is described in Equation (16), with the kernel scaling parameter $\sigma = 15$, which is empirically determined.
Results

We compare the accuracy performance of our proposed method with another least-square SVR-based method by Lv et al. [11], which uses a different feature vector encoding. A RBF kernel is also used for it with the parameter $\sigma = 18$, which is the optimum for that method. To enable a fair comparison, their regression model is also re-trained after every test instance in order to ensure an up-to-date model.

The forecasted peak load values for 184 test days from July 01, 2009 to December 31, 2009 are computed using both methods and are compared against the actual peak load values.

An example of the forecasted values by the two methods and the actual values for the month of December 2009 for LS Dataset are demonstrated in Fig. 9. It can be observed from the figure that our method can predict daily peak loads more accurately than the method by Lv et al.

![Figure 9: Example of forecasted results. The actual peak loads and the forecasted values by Lv et al. [11] and our method for the period of December 1 to 31, 2009 on LS Dataset.](image)
In order to systematically analyze the performance of the two methods, we use two criteria: relative error and accuracy in our experiment. For each testing day, the relative error and the accuracy of the forecasted peak load are calculated as follows.

\[
\text{relative error} = \frac{|\text{actual peak load} - \text{forecasted peak load}|}{\text{actual peak load}} \times 100\% \quad (18)
\]

\[
\text{accuracy} = 100 - \text{relative error} \quad (19)
\]

For the testing period of 184 days, the comparisons of relative error values of the two methods are given in Fig. 10 for LS Dataset and Fig 11. for NRW Dataset respectively. We can visually observe from the figures that our proposed method provides lower relative errors than the method by Lv et al. in a majority of cases.

For LS Dataset, the average relative error of our method is 1.3\% (i.e., 98.7\% average accuracy) whilst that of Lv et al. is 3.3\% (i.e., 96.7\% average accuracy).

For NRW Dataset, the average relative error of our method is 1.6\% (i.e., 98.4\% average accuracy) whilst that of Lv et al. is still 3.3\% (i.e., 96.7\% average accuracy).

Load forecasting is quite a mature technology in which many methods are able to provide an accuracy level of ~95\% in general [11]. For example, the method by Lv et al. provides 96.7\% of accuracy in our experiment. Here in our research, we are able to further improve the forecasting accuracy to the level of 98.4–98.7\%. Although the 1.7–2.0\% increase in accuracy may be small in absolute terms, this achievement is non-trivial. It is a common phenomenon that improving the performance of a
particular technology is quite difficult when it becomes mature (i.e., when the higher plateau in its technology S-curve is reached) [12].

Figure 10: Relative error performances of Lv et al. [11] and our method for the period of June 1, 2009 to December 31, 2009 on LS Dataset.

Figure 11: Relative error performance of Lv et al. [11] and our method for the period of June 1, 2009 to December 31, 2009 on NRW Dataset.
Figure 12: Relative error performances of scaling vs. non-scaling approaches in our method. (The results are the averages of those for LS and NRW Datasets for the period of June 1, 2009 to December 31, 2009.)

### Scaling vs. Non-scaling

As mentioned in the above chapter, we use the scaling facility of LibSVM [21] to map the values of each of our 32 attributes into the range of −1 to +1. Scaling helps significantly reduce the relative error (i.e., improve the accuracy) of our proposed method over directly using the data without any scaling. We can clearly observe this trend in Fig. 12. The overall average of the relative errors for both LS and NRW Datasets with scaling is 1.4% while that without scaling is 6.2%. The superiority of scaling is because it prevents the dominance of attributes with larger value ranges over those with smaller ranges in calculating the Lagrange multipliers $\alpha$ in Equation 12.

### Computational Efficiency

The proposed method is developed in C++ and tested on a modest laptop PC with Intel Core Duo 1.83 GHz processor and 2GB of main memory running Windows

The method is found to be quite efficient and scalable. The overall running time of the training for 150 days and the testing (and re-training) for 184 days is only 210 milliseconds for LS Dataset and 220 milliseconds for NRW Dataset respectively. Thus, our proposed method can be potentially deployed in a larger scale to forecast the loads of tens of thousands of consumer entities like smart meters on a distributed computing platform.

6.2 Experiment on Abu Dhabi Data

In order to further investigate the general usability of our method, we do an experiment using data from Abu Dhabi. The data contains the daily peak load values of two substations in Abu Dhabi Emirate. Those data were kindly provided to us by TRANSCO, Abu Dhabi.

We name the data for Substation #1 as AD1 Dataset and those for Substation #2 as AD2 Dataset respectively. Each data set contains the daily peak load values of its substation from August 01, 2009 to June 30, 2011 (699 days).

We show the peak load data of AD1 Dataset in Fig. 13 for illustration. We can compare it to Fig. 8 and observe that the Abu Dhabi data is noisier and less periodically regular than the data from Germany.
Training and Testing Data

For each of AD1 and AD2 datasets, we use the daily peak loads from September 01, 2009 to February 28, 2010 (i.e., 181 days) for training and March 01, 2010 to June 30, 2011 (i.e., 487 days) for testing. (Again, we simply cannot start the training from August 01, 2009 because we need the data for the previous 4 weeks, i.e., 28 days, to construct a feature vector.)

Results

In the same way as in the experiment with the Germany data, we compare the accuracy performance of our proposed method with another least-square SVR-based method by Lv et al. [11], which uses a different feature vector encoding. A RBF kernel is also used for it with the parameter $\sigma = 18$, which is the optimum for that method. To enable a fair comparison, their regression model is also re-trained after every test instance in order to ensure an up-to-date model.
The forecasted peak load values for 487 test days from March 01, 2010 to June 30, 2011 are computed using both methods and are compared against the actual peak load values.

An example of the forecasted values by the two methods and the actual values for the month of June 2011 for LS Dataset are demonstrated in Fig.14. It can be observed from the figure that our method can predict daily peak loads more accurately than the method by Lv et al.

![Figure 14: Example of forecasted results for Abu Dhabi. The actual peak loads and the forecasted values by Lv et al. [11] and our method for the period of June 1 to 30, 2011 on AD1 Dataset.](image)

For the testing period of 487 days, the comparisons of relative error values of the two methods are given in Fig. 15 for AD1 Dataset and Fig. 16 for AD2 Dataset respectively. We can visually observe from the figures that our proposed method provides lower relative errors than the method by Lv et al. in a majority of cases.

For AD1 Dataset, the average relative error of our method is 4.3% (i.e., 95.7% average accuracy) whilst that of Lv et al. is 4.7% (i.e., 95.3% average accuracy).

For AD2 Dataset, the average relative error of our method is 3.3% (i.e., 96.7% average accuracy) whilst that of Lv et al. is 3.6% (i.e., 96.4% average accuracy).
Because the Abu Dhabi data (AD1 and AD2) are noisier and less periodically regular than the Germany data (LS and NRW), the accuracies by our method are slightly lower now (i.e., 95.7%– 96.7%). However, our results are still better than those by Lv et al. albeit by a narrower margin now.
In this work, we have presented an accurate load forecasting method which can potentially provide greater intelligence (smartness) to the upcoming smart grids. In our approach, we adopt the least-squares support vector regression technique incorporated with online learning. Experimental results show that our method is able to provide more accurate results than an existing forecasting method by Lv et al. [11], which is reported to be one of the best methods, and is also computationally efficient. As the future work, we intend to explore the idea of automatic feature selection for our regression model in order to further improve its accuracy. In addition, we plan to rigorously test our method with multiple smart grid load datasets from different countries/industries and fine tune the method so as to ensure its general usability. Finally, we hope our proposed method with these future improvements can be potentially useful to utility companies in their large-scale load forecasting applications for consumer entities at any granularity levels (such as individual households,
neighborhoods, towns, cities, and large geographical regions) by providing results with better precisions.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>Artificial Neural Networks</td>
</tr>
<tr>
<td>ANNSTLF</td>
<td>Artificial Neural-Network-Based Electric Load Forecasting</td>
</tr>
<tr>
<td>ARIMA</td>
<td>Auto Regressive Integrated Moving Average</td>
</tr>
<tr>
<td>ARIMAX</td>
<td>Auto Regressive Integrated Moving Average with eXogenous variables</td>
</tr>
<tr>
<td>ARMA</td>
<td>Auto Regressive and Moving Average models</td>
</tr>
<tr>
<td>ARMAX</td>
<td>Auto Regressive Moving Average with eXogenous variables</td>
</tr>
<tr>
<td>CFL</td>
<td>compact fluorescent lamps</td>
</tr>
<tr>
<td>DOL</td>
<td>Direct On Line motors</td>
</tr>
<tr>
<td>DSM</td>
<td>Demand Side Management</td>
</tr>
<tr>
<td>EP</td>
<td>Evolutionary Programming</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross Domestic Product</td>
</tr>
<tr>
<td>GRNN</td>
<td>Generalized Regression Neural Network</td>
</tr>
<tr>
<td>IPP</td>
<td>Independent Power Producers</td>
</tr>
<tr>
<td>LS</td>
<td>Lower Saxony</td>
</tr>
<tr>
<td>MLP</td>
<td>Multi-layer Perceptron</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>M-LTLF</td>
<td>Medium-to-long term load forecasting</td>
</tr>
<tr>
<td>MSE</td>
<td>mean squared error</td>
</tr>
<tr>
<td>NRW</td>
<td>North Rhine-Westphalia</td>
</tr>
<tr>
<td>PSLF</td>
<td>Power System Load Forecasting</td>
</tr>
<tr>
<td>RBF</td>
<td>Radial Basis Function</td>
</tr>
<tr>
<td>RBF</td>
<td>Radial basis function</td>
</tr>
<tr>
<td>RBFN</td>
<td>Radial Basis Function Neural Network</td>
</tr>
<tr>
<td>STLF</td>
<td>Short Term Load Forecasting</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>SVR</td>
<td>Support Vector Regression</td>
</tr>
<tr>
<td>T&amp;D</td>
<td>Transmission and Distribution</td>
</tr>
<tr>
<td>UPS</td>
<td>Uninterrupted Power Supply</td>
</tr>
</tbody>
</table>
Bibliography


