INTELLIGENT ENERGY MANAGEMENT SYSTEM – TECHNIQUES AND METHODS

A thesis submitted to
CITY UNIVERSITY LONDON
for the Degree of
DOCTOR OF PHILOSOPHY

by

YINGNAN MA

School of Engineering and Mathematical Sciences

City University London
London EC1V 0HB
United Kingdom

February 2011

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I wish to offer my sincere gratitude to my Ph.D supervisor, Professor Loi Lei Lai, Head of the Energy System Group, City University London, for his great assistance and guidance throughout my Ph.D studies. The completion of this thesis would not have been possible without his unfailing advice and encouragement.

I must also thank my colleagues Dr Qingping Zhang, Dr Harold Braun, Dr Tsu-Yuan Nieh, Mr Long Zhou and Mr Fangyuan Xu for their support and friendship over these challenging years. It has been an honour to work with them.

The financial and psychological support of my parents, Mr Longshuang Ma and Mrs Fengzhi Liu, is greatly appreciated. They have been with me through a long and difficult journey, always demonstrating their unconditional patience, love, inspiration and understanding not only in relation to my Ph.D study but throughout all of my ten years in the UK, even though they are seven-thousand miles away. I am very grateful to my parents.

LIST OF PUBLICATIONS

[1] Yingnan Ma, Long Zhou, Norman Tse, Abdi Osman, Loi Lei Lai, “An Initial Study On Computational Intelligence For Smart Grid”, Proceedings of 2009 International Conference on Machine Learning and Cybernetics IEEE (ICMLC 2009), 2009.07, pp3425-3429, ISBN: 978-1-4244-3702-3

[2] Long Zhou, Fangyuan Xu, Yingnan Ma, “Impact of Smart Metering on Energy Efficiency”, Proceedings of 2010 International Conference on Machine Learning and Cybernetics IEEE (ICMLC 2010), 2010.07 (in print)

[3] Nidul Sinha, Loi Lei Lai, Palash Kumar Ghosh, Yingnan Ma, “Wavelet-GA-ANN Based Hybrid Model for Accurate Prediction Of Short-Term Load Forecast”, Proceedings of 2007 International Conference on Intelligent Systems Applications to Power Systems (ISAP 2007), 2007.11, pp 1 – 8, ISBN: 978-986-01-2607-5

[4] Fangyuan Xu, Long Zhou, Yingnan Ma, Loi Lei Lai, “Extending Version Of Graphical User Interface In Neural Network Toolbox Of Matlab And Engineering Applications”, Proceedings of 2009 International Conference on Machine Learning and Cybernetics IEEE (ICMLC 2009), 2009.07, pp187-192, ISBN: 978-1-4244-3702-3

[5] Qingping Zhang, Yingnan Ma, Loi Lei Lai, “Adaptive RBFN Model for 2D Spatial Interpolation”, Proceedings of 2009 International Conference on Machine Learning and Cybernetics IEEE (ICMLC 2009), 2009.07, pp3418 - 3424, ISBN: 978-1-4244-3702-3

[6] Loi Lei Lai, T. Y. Nieh, D. Vujatovic, Yingnan Ma, Y. P. Lu, Y. W. Yang, H. Braun, “Particle Swarm Optimization for Economic Dispatch of units with Non-smooth Input-output Characteristic Functions”, Proceedings of 2005 International

Conference on Intelligent Systems Applications to Power Systems (ISAP 2005), 2005.11, pp 5, ISBN: 1-59975-174-7

[7] Loi Lei Lai, T. Y. Nieh, D. Vujatovic, Yingnan Ma, Y. P. Lu, Y. W. Yang, H. Braun, “Swarm Intelligence For Optimal Reactive Power Dispatch”, 2005 Power & Energy Society on Transmission and Distribution Conference and Exhibition: Asia and Pacific, (2005 IEEE/PES), 2005, pp1-5, ISSN: 0-7803-9114-4

[8] Loi Lei Lai, H. Braun, Qingping Zhang, Q. Wu, Yingnan Ma, W.C. Sun, L. Yang, “Intelligent Weather Forecast”, Proceedings of 2004 International Conference on Machine Learning and Cybernetics IEEE (ICMLC 2004), 2004.08, pp 4216 - 4221, ISBN: 0-7803-8403-2

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACD	Adaptive Control Design
AI	Artificial Intelligence
AMI	Automatic Metering Infrastructure
AMM	Automated Meter Management
ANN	Artificial Neural Network
APE	Absolute Percentage Error
CI	Computational Intelligence
CO_2	Carbon Dioxide
CPUC	California Public Utilities Commission
Discos	Distribution Companies
DP	Dew Point
DWT	Discrete Wavelet Transform
DWTDNN	Dynamic Weighted Time-Delay Neural Networks
ED	Economic Dispatch
EP	Evolutionary Programming
ERDF	Electricité Réseau Distribution France
EM	Energy Management
FACTS	Flexible Alternating Current Transmission Systems
FIR	Finite Impulse Response
FPGA	Floating Point Genetic Algorithm
F-RBFN	Factorized RBFN
FT	Fourier Transforms
GA	Genetic Algorithm

GAs	Genetic Algorithms
Gencos	Generating Companies
GIS	Geographic Information System
IGA	Improved Genetic Algorithm
IR-RBFN	Incremental Responsivity based RBFN
LADWP	Los Angeles Department of Water and Power
LF	Load Forecasting
MCE	Ministerial Council on Energy
MLP	Multi-layer Perceptron
MSFTC	Missing Sensor Fault Tolerant Control
OLS	Orthogonal Least Squares
ORPD	Optimal Reactive Power Dispatch
PMU	Phasor Measurement Unit
PQ	Power Quality
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
RBFN	Radial Basis Function Network
RF	Radio Frequency
TCQ	Total Cloud Quantity
TDNN	Time-Delay Neural Networks
SA	Simulated Annealing
S.D.	Standard Deviation
SCADA	Supervisory Control and Data Acquisition
SERS	Sensor Evaluation Restoration Scheme
SI	Swarm Intelligence

SSSC	Static Synchronous Series Compensator
STLF	Short-term Load Forecast
UHV	Ultra High Voltage

COPYRIGHT DECLARATION

ABSTRACT

Our environment is an asset to be managed carefully and is not an expendable resource to be taken for granted. The main original contribution of this thesis is in formulating intelligent techniques and simulating case studies to demonstrate the significance of the present approach for achieving a low carbon economy.

Energy boosts crop production, drives industry and increases employment. Wise energy use is the first step to ensuring sustainable energy for present and future generations. Energy services are essential for meeting internationally agreed development goals. Energy management system lies at the heart of all infrastructures from communications, economy, and society's transportation to the society. This has made the system more complex and more interdependent. The increasing number of disturbances occurring in the system has raised the priority of energy management system infrastructure which has been improved with the aid of technology and investment; suitable methods have been presented to optimize the system in this thesis.

Since the current system is facing various problems from increasing disturbances, the system is operating on the limit, aging equipments, load change etc, therefore an improvement is essential to minimize these problems. To enhance the current system and resolve the issues that it is facing, smart grid has been proposed as a solution to resolve power problems and to prevent future failures. This thesis argues that smart grid consists of computational intelligence and smart meters to improve the reliability, stability and security of power. In comparison with the current system, it is more intelligent, reliable, stable and secure, and will reduce the number of blackouts and other failures that occur on the power grid system. Also, the thesis has reported that smart metering is technically feasible to improve energy efficiency.

In the thesis, a new technique using wavelet transforms, floating point genetic algorithm and artificial neural network based hybrid model for gaining accurate prediction of short-term load forecast has been developed. Adopting the new model is more accuracy than radial basis function network. Actual data has been used to test the proposed new method and it has been demonstrated that this integrated intelligent technique is very effective for the load forecast.

Choosing the appropriate algorithm is important to implement the optimization during the daily task in the power system. The potential for application of swarm intelligence to Optimal Reactive Power Dispatch (ORPD) has been shown in this thesis. After making the comparison of the results derived from swarm intelligence, improved genetic algorithm and a conventional gradient-based optimization method, it was concluded that swam intelligence is better in terms of performance and precision in solving optimal reactive power dispatch problems.

Chapter 1

INTRODUCTION

Deregulation of the power utility industry, being a reality today, has resulted in competition in every aspect of power systems; be it in power generation, or in transmission or in energy consumption, professional management of electric energy is of utmost importance [1-10].

Sensing and communication technologies, smart meters for example, are essential to support the development, integration and deployment of flexible, safe, reliable and efficient power distribution management systems. Smart metering is a subject that is attracting much more attention. Smart metering is delivering many benefits in a lot of respects. Many smart metering projects are going on many countries, such as the UK, Italy, the USA and elsewhere. The design, control, management and optimization of these new distributed energy resources and technologies, and their integration into existing energy distribution networks, pose significant technological challenges to ensure their reliability and safety, and to improve and maximize their cost competitiveness.

Accurate prediction of load consumption pattern is becoming a very important function of a utility company, as it is needed to support wiser management decisions. A forecast that exceeds the actual load may lead to extra power being generated and therefore may result in excessive investment in a power plant that is not fully utilized. On the other hand, a forecast that is too low may lead to some revenue loss due to loss of opportunity of selling power to neighboring utilities. Hence, accurate electricity load forecasting (LF), including very short-term, short-term, mid-term, and long-term, plays a vital role in ensuring adequate electricity generation to meet the customer's

demands in the future. LF also helps to build up cost effective risk management plans for companies in the electricity market. Consequently, good operational, planning and intelligent management decision making, such as, economic scheduling of generation capacity, scheduling of fuel purchase, ability to avoid unnecessary start-ups of generating units, planning the scheduling of peaking power, buying or selling electricity at best price, and scheduling of ancillary services can all be carried out based on accurate LF, which forecasts the load of a few minutes, hours, days, weeks, months ahead. The aim of LF is to predict future electricity demand based on historical load data, and currently available data.

In recent years, many solutions to intelligent weather forecast have been proposed, especially on temperature and rainfall; however, it is difficult to simulate the meteorological phenomena and the corresponding characters of weather when some complex differential equations and computational algorithms are already piled up.

As known, Radial Basis Function (RBF) network is considered as an effective methodology to make prediction in spatial space, with spatial information fusion at different layers of RBF; the hidden layers fusion is able to give a better result.

To implement daily optimization task in the power system operation, choosing the appropriate algorithms are also important for the intelligent energy management in terms of performance and precision to get the best results.

1.1 Organization of the Thesis

Chapter 2 reports a framework as to how computational intelligence may be applied in developing smart grid to improve reliability and security of the power system. Some examples are used to demonstrate the feasibility.

Chapter 3 gives a comprehensive review as to the benefits of smart metering in power networks, such as energy efficiency improvement and reduction in greenhouse gas emissions. The benefit of having smart meters to enhance assets management has been highlighted. Numerous case studies worldwide have been referred to. International engineering practices and policy have been discussed.

Chapter 4 demonstrates the use of wavelet techniques and its integration with neural networks for power quality analysis and assessment and load forecasting. This chapter demonstrates some of the techniques that have a high potential to be used in the near future for real-life applications.

Chapter 5 sets out a methodology in respect of short-term temperature and rainfall forecasting over the east coast of China based on the necessary data preprocessing technique and the Dynamic Weighted Time-Delay Neural Networks (DWTDNN), in which each neuron in the input layer is scaled by a weighting function that captures the temporal dynamics of the biological task.

Chapter 6 proposes an extending version of Graphical User Interface in Neural Network Toolbox of MATLAB 7.1, which can release the limit of setting more layers in the feedforward network creating. Users can set up a feedforward network with any architecture.

Chapter 7 describes swarm intelligence to solve the Optimal Reactive Power Dispatch (ORPD) problem. Results derived from swarm intelligence, improved genetic algorithm and a conventional gradient-based optimization method were compared.

Chapter 8 introduces an adaptive RBF network construction method which combines the traditional incremental algorithm and real-time responsivity analysis.

Chapter 9 outlines an application of Particle Swarm Optimization (PSO) to solve economic dispatch (ED) of units with non-smooth input-output characteristic functions.

Chapter 10 summarizes the work done in this research study. Areas for future work and developments are suggested.

1.2 Original Contributions

A summary of the original contributions made in this thesis are given below:

1. Proposal of a solution using computational intelligence on developing Smart Grid to ensure power distribution management systems' reliability and safety, and to improve and maximize their cost competitiveness (Chapter 2).

2. Improvement of energy efficiency by adopting Smart Metering. Listing case studies in this thesis has shown that smart metering is technically feasible (Chapter 3).

3. Development of the proposed hybrid model through the integration of Wavelet transform, Floating Point Genetic Algorithm (FPGA) and Artificial Neural Networks (ANN) for the prediction of short-term load forecast. More accurate predictions are obtained when the proposed model is tested with actual data provided by an electricity market (Chapter 4).

4. Proposal of the solution using data preprocessing technique and the Dynamic Weighted Time-Delay Neural Networks (DWTDNN) to temperature and rainfall estimation. The results confirm that the proposed solution has the potential for successful application (Chapter 5).

5. Development of a new extending version of Graphical User Interface by applying two simple applications which are voting system design and 3–phase generator output

detector design. The contribution is in developing a platform for neural network modelling applied to various problem domains (Chapter 6).

6. Determination of the application of Swarm Intelligence to Optimal Reactive Power Dispatch (ORPD). Comparing the results has shown that Swarm Intelligence gets better results than the conventional methods (Chapter 7).

7. Proposal of an Adaptive Radial Basis Function Network (RBFN) method which combines the traditional incremental algorithm and real-time responsivity analysis for 2D spatial interpolation. Testing the proposed method with the practical data has shown that this intelligent technique is very effective for real-time metrological information processing. The developed algorithm has been implemented in the Shanghai Meteorology Center and is used as one of the tools for weather forecast (Chapter 8).

8. Adoption of Particle Swarm Optimization (PSO) solving economic dispatch of units with non-smooth input–output characteristic functions. The IEEE 30-bus system with 6 generating units has been used as the simulation system to show the effectiveness of the algorithm. Results are compared to those details by Evolutionary Programming (EP). It shows that PSO could be a good technique to solve real life power system problems (Chapter 9).

Chapter 2

AN INITIAL STUDY ON COMPUTATIONAL INTELLIGENCE FOR SMART GRID

2.1 Introduction

The electricity network is facing many problems, from deregulation and load increase to plant problems; which is caused by the fact that the system was not designed to operate at that level. Such problems lead to major events and paralyse the electricity system as has been witnessed in the past. The increasing demand and aging infrastructure made it difficult to keep the system up to date and has this made the system complex; to overcome such complexity it is necessary to come up with a major solution to such problems; in other words the grid has to be made more modern and smarter-that is, generating the electricity and transmitting and distributing in a smarter way [1-10].

The increasing complexity of Electricity Infrastructure is the driving force for smart grid, such complexities arise from various directions such as non-linearity of the system, dynamic behaviour, uncertainties etc. External factors also increase the complexity even further with the increasing development of new technology which requires electricity, such as electric cars, and the development of renewable energy is increasing the electricity challenges particularly in its complexity. Computational Intelligence (CI) holds the key to the development of smart grid to overcome the challenges of managing data and communications, planning and optimization, control and protection of power plants.

Intelligence is required at all levels of the system; it should start from the component level such as plants and substations to grid management. By making each component

computationally intelligent it will make it easy for the grid to be smarter and more alert to any kind of disturbances that might affect the power flow. Improvement is inevitable if the system is to be smart. This can be achieved by upgrading the current system especially in the areas of modeling and optimization, as the current process is slow and not very accurate. To alleviate such problems, there must be some sort of intelligence in the system that is handling these processes.

Smart grid will not only involve power systems but also affect other aspects of technology since a lot of technological developments are taking place towards the electric side and reduce the need for other sources of energy, like oil and gas; as these technologies include electric cars, they will increase the demand and the complexity of the power system, therefore reliable power is a must.

Since the electricity infrastructure is a dynamic system, computational intelligence (CI) must help to adopt changes in all levels, it must be dynamic and fault-tolerant. Furthermore, it must be able to learn from these changes that the system experiences from all directions.

2.2 Smart Grid

The electricity industry was conceived over 50 years ago when the load and generation was less; now in the information age, with a digital society where demand is very high, the electricity infrastructure has been forced to its limits which had not been anticipated; in addition, the electricity demand continues to grow. So the focus of the question has become what needs to be done from a technology prospective to meet that growing demand for electricity, and to do it in a way whereby a greater carbon footprint is not created. Smart grid is the next generation of the electricity infrastructure based on the optimization of the current system at all levels.

Firstly smart grid is equipped with a self-healing ability to response to the threats, material failure and other unexpected problems. Real-time monitoring and reaction can be achieved by using phasor measurement units (PMU) and other sensors to monitor the parameters such as voltage and current of a particular area. Secondly it monitors the performance of generation and searches for any problems that could trigger a disturbance; this will increase system operator awareness.

A self-healing ability involves an intelligent communication and monitoring system. Such technology will enable the system to be a self-healing one. It has the capability of fault-tolerance that will enable the system to operate properly when faced with a fault i.e. it resists attacks. It accommodates all generation and storage options including plug-in vehicles [7]. It will provide a plug-and-play interconnection to any source of power including renewable energy sources such as wind and solar sources and storages. This will increase integration and flexibility in the power resource mix.

Electricity infrastructures are highly interconnected which makes them suitable for intelligent and self-healing technology applications. This can be achieved by using intelligent agents which involve computer science dealing with the simulation of imitating intelligent human beings to sense and measure the parameters for each generator and communication with the central network; such agents need to be adaptable to the environment in order to improve their performance. Different agents are used in different levels of system; some are used for self-healing and control and others are for the operation and management of the system.

It is essential to understand what the drivers that push smart grids are. Until recently the industry was considering smart grids mainly as a concept. This has changed significantly recently. The physical transition of the power system to smart grids is already underway in some countries.

Players tend to adopt state-of-the-art systems, but there is also a risk of insufficient investment at the production level in some countries which will force regulators to consider new incentives to support grid operation in minimum reserve scenarios as too high a risk of blackout is reached.

Without the corresponding development of technical, market and regulatory frameworks in the next few years, the current decentralized system will become unstable and unable to accept further deployment of distributed generation.

The amount of “regulating energy” provision is increasing in value with the increase in stress on the system while the governments continue to strive for distributed resource penetration and launch new energy efficiency ideas.

System management costs are increasing, and the threat to system security is an increasing concern as installed distributed generating capacity in some areas exceeds local demand.

Environmental issues such as conservation, “zero emission buildings”, and grid complexity due to distributed generation, renewable production penetration, back up generation, interconnection, demand side management, network congestions and ancillary services will make future grid operation different from the current one.

The majority of the large transmission and distribution utilities have decided to launch new smart grid prototyping initiatives which, once stabilized, will be scaled to millions of connection points. Most of this initial roll-out is expected to happen by 2012.

From a functional stand-point, the majority of these early adoptions of smart grids concepts are in the domain of system intelligence and the associated challenges which are related to IT system scalability; consisting of applying existing algorithms and processes but in a larger and more distributed manner. It is worth noting that the

majority of the main software building blocks that exist today are used at scales quite different from the future distribution of sensors in grids. Besides, the software and modeling technologies are improving a lot because of other technological trends in other industries like airplanes, mobile phone, and military applications.

2.3 Technologies of the Smart Grid

Smart Grid is a form of electricity network using digital technology. A smart grid delivers electricity from suppliers to consumers using two-way digital communications to control appliances at consumers' homes; this could save energy, reduce costs and increase reliability and transparency if the risks inherent in executing massive information technology projects are avoided.

In order to complete the transition that has already started the following new system intelligence functions will in the future be required to approach the Smart Grid infrastructure targets. However, it should be recognized that these are only some system-level functions and a full implementation of Smart Grids will also require smart equipment such as Fault-Current Limiters, new protection schemes/algorithms to facilitate increased distributed generation connection, some of which are discussed in this chapter, and a massive growth in the use of power electronics; for example, Flexible AC Transmission System (FACTS) etc. and embedded energy storage, even at lower voltage/power levels than those for which are used today.

General technological development for Smart Grid equipment and solutions are now widely shared; for example at the European level through the Smart Grid technological platform for the electricity of the future or in the US through Gridwise. Several research and development programmes have been formed in recent years with

the goal of improving the intelligence of the electric power infrastructure. Several major programmes can be identified for example:

- GROW-DERS: Grid reliability and operability with distributed generation using transportable storage.
- INTEGRAL: achievement of an integrated ICT-platform based on distributed control of decentralised energy resources.

These programmes, along with other established R&D organizations, are addressing the technical, economic, and policy barriers to creating a smarter grid. This abundance of research has bolstered the confidence of industry stakeholders, who now recognize that transformation of the grid indeed can be accomplished within the next few years.

There are strong drivers to smart grid progress and deployment. Many political bodies, in the US and in Europe, at different levels of government regard Smart Grids as a tool that can be leveraged to promote several major policies and tackle several major issues related to the environmental impact of the electricity industry, security of energy supply and economic efficiency of this market. In Europe there is now incredible pressure towards renewable use within the energy mix; there is a target of 20% by 2020. Moreover, as long as countries do not combine energy savings and very strong heat promotion into buildings, all the pressure will be put on electricity.

Technologies that have been developed for smart grid are mainly concentrated in the following three areas.

The first one is integrated communications. This is necessary for Smart Grid since it is required for all levels of the system to make the system smarter and more intelligent. It connects components of the system together; this is to make the components aware of the others' performance.

The second one is sensing and measurements. These technologies support the system to be faster with more accurate responses, such as remote monitoring, data management etc.

The third one is advanced control. It monitors essential components and system areas and enables a rapid diagnosis, in the event of disturbances, and provides an accurate solution for each event. The components are made of superconductor. Improved interfaces and decision support amplify human decision-making and promote grid operators and managers into knowledge workers.

To develop such a Smart Grid system that can provide advanced control and optimization, prediction, monitoring etc. requires fast and dynamic algorithms that can learn the system's behavior; this is where computational intelligence comes in.

2.4 Computational Intelligence (CI) for Smart Grid

To achieve these intelligent technologies for Smart Grid which will operate in an extremely complex and dynamic system, it is necessary to use Computational Intelligence as it provides various solutions and algorithms for such problems. Computational Intelligence is the study of adaptive mechanism to enable or facilitate intelligent behaviour in a complex system. It is based on algorithms and intelligent systems such as neural network, fuzzy systems, evolutionary computation, swarm intelligence, bacteria foraging, ant colony optimization and immune systems.

The challenges that are facing Computational Intelligence include neural network architecture, learning procedure, range of training and long convergence time etc.

To achieve fully a Smart Grid system it is necessary to overcome these very complex and nonlinear systems and in doing so, it will lead to a disturbance free system, and a smarter system that can maximize the utilization of the system.

Since intelligence in a power grid is required in all levels, it would be useful for each substation and power plant to have a processor that monitors and communicates with other processors via smart sensors. These sensors will provide the necessary real-time data from wide area monitoring and provide control to the system controller; such information is vital because it allows the operator to be aware of the systems situation that can be the asset conditions, the operating parameters of the system and it will also help them to make the right decisions. Furthermore it is much faster and more accurate than the traditional SCADA/EM control systems. These sensors also provide the dynamic rating of the lines, and help monitor the line performance, particularly at times of outages and disturbances.

Another challenging issue for the Smart Grid is the extraction of the required information from the large amount of data received from sensors; using advanced control methods, the required features can be extracted using Computational Intelligence and can then be classified accordingly. This is important in power systems because lots of information is redundant; the controller does not require all this data, so only the important information needs to be extracted. On the other hand, if a sensor is missing or faulty it will provide unreliable result, such a problem will lead to the failure of the power controller and might also lead to disturbance in the power system.

To reduce contingencies, a sensor evaluation and (missing sensor) restoration scheme (SERS) using auto-associative neural networks (auto-encoders) and particle swarm optimization was developed for controlling a static synchronous series compensator (SSSC) connected to a power network. This MSFTC improved the reliability, maintainability and survivability of the SSSC and the power network. Such fault-tolerant technologies will be needed in a Smart Grid to improve its reliability and

security.

Computational Intelligence has increased in the field of advanced control, especially in adaptive control design (ACD) and optimization [2]; this is to achieve an intelligent system that has the ability to learn how to coordinate all parts of the system for optimal performance.

Advanced control methods provide real-time prediction via the use of Computational Intelligence methods. It provides applications such as monitoring and collecting data from sensors and analysis of data both online and offline to diagnose and provide solutions.

Advanced components are used in Smart Grid. These include: next generation FACTS, power quality PQ devices, advanced distributed generation and energy storage, fault current limiters, superconducting transmission cable and rotating machines, power electronic advanced switches and conductors. These components will help the system at different levels to ease the complex power system and make it more efficient.

2.5 Conclusion

Streamlining and simplification of existing permission procedures and standardisation of the grid codes for the connection of distributed generation are required to encourage greater distributed resources integration. The standardization in the connection requirements of distributed generation, particularly of the protection equipment and settings criteria, will be very positive for the development of the distributed generation, especially in highly interconnected networks to avoid nuisance tripping and obtain more generation availability and network stability. Also, more intelligent protection is required to overcome some of the protection co-ordination

and sensitivity problems with lower fault current and system stability and network capacity concerns with higher integration of distributed generation, especially if controlled islands are allowed and for full converter wind turbine generators which only provide a small contribution to the fault level. Standardised communications infrastructure will be important to encourage some of these intelligent adaptive and wide area protection schemes.

Sensing and communication technologies, smart meters for example, are essential to support the development, integration and deployment of flexible, safe, reliable and efficient power distribution management systems. The design, control, management and optimization of these new distributed energy resources and technologies, and their integration into existing energy distribution networks, pose significant technological challenges to ensure their reliability and safety, and to improve and maximize their cost competitiveness.

Chapter 3

IMPACT OF SMART METERING ON ENERGY EFFICIENCY

3.1 Introduction

As a part of Smart Grid technology, smart metering is the innovating of electricity and gas meters. Instead of the meter readings and bill-estimation, data from the amount of electricity and gas could be collected real-time to ensure accuracy for customers and source suppliers. Customers will benefit from this application by learning how much energy they use and can then choose the optimal energy consumption. In the application, a communication system will be installed so that the suppliers and the customers could both have the real-time information [11]. A simple relationship in smart metering is shown in Figure 3.1.

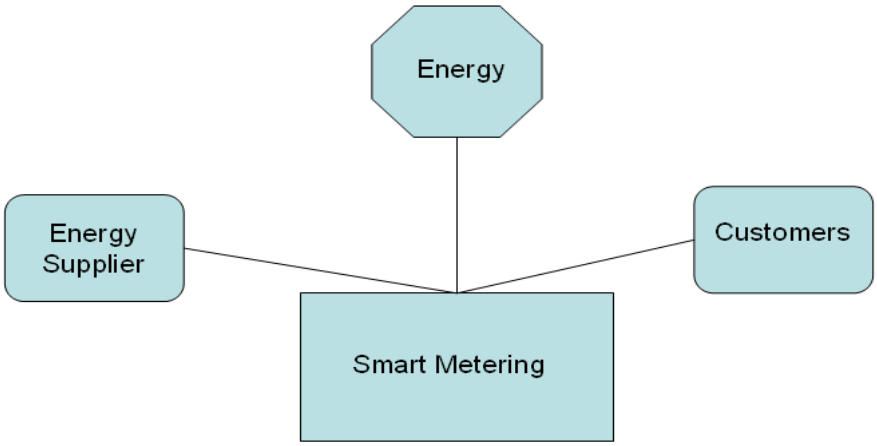


Figure 3.1 Structure of smart metering

The most important application of a smart metering system is smart meter. A Smart meter is fundamentally different from ordinary meters. It can provide a real-time and

accurate record of the gas and electricity customers are using at different times and different costs.

Smart meters, which have a visual display, allow people to see clearly how much electricity and gas they are using and send the data to energy firms automatically. Smart metering has a huge potential benefits. In the UK, there is no longer self-sufficient energy, and North Sea oil and gas are depleting.

Smart meters need to play a very important role in reducing the energy consumption [12].

3.2 Standard of Smart Metering

Smart metering is the combination of power system, telecommunication and several technologies. System becomes complicated and the effectiveness are generated. Facing these problems, standards and limitations should be applied.

The smart metering system should be capable of two-way communication. Also, it should be able to contain flexible tariff structures and display consumption details. Other than these minimum requirements, different systems can be flexible based on applications.

In 2007, the Ministerial Council on Energy's (MCE) in Europe agreed to establish a minimum functionality for Smart Meters so that it could set up a platform for smart meter deployment for future government. Based on a cost-benefit analysis of each function, and considering views from subsequent consultation, an initial set of functions are included in the national minimum functionality [13].

This list includes:

- Remotely read interval metering, with the meter capable of daily reads;
- Quality of supply and outage detection to improve consumer supply services;

- Ability to control connection and disconnection remotely and apply supply capacity limits to manage emergency situations;
- Ability to manage load through a dedicated circuit, to support existing off-peak arrangements;
- Supporting management functions such as data security, tamper detections, remote configuration, remote upgrade and plug-and play installation [13].

Other than the MCE standard, ZigBee Alliance has also developed an open standard with IEEE 802.15.4 for implementing secure and convenient home wireless area networks. The profile establishes standards for manufacturers to produce interoperable products, as well as allowing utility companies to set up energy management networks [14].

3.3 Smart Metering Impact on Energy Efficiency

Nowadays, more and more people are encouraged to adopt Renewable Energy in order to reduce in greenhouse gas emission, and to also address the very important issue of climate change. Smart Meters provide knowledge, increase awareness and change customers' behaviour and attitudes in using renewable energy. In the future, renewable energy is the means to reduce carbon emissions and gas emissions during power generation. Although it may not be cheap in terms of the generation cost at the moment, customers can at least reduce the impact on climate change. Hence, if every household can use electricity and gas efficiently, there is a resulting need to minimize power generation as well [15].

3.4 Smart Metering Case Study Worldwide

Most countries of the world are paying more and more attention to smart metering

projects. This section gives a survey of many countries' plans for smart metering and smart grid. Some implementation examples are given such as in the UK, China, Italy, the USA and so on.

3.4.1 Imminent Implementation of Smart Metering in UK

Recently, UK energy smart metering has been rolled out. The government has announced every household will install a smart meter by 2020. The first provider of business and consumer smart meters in the UK as standard was First Utility.

"Smart meters will put the power in people's hands, enabling us all to control how much energy we use, cut emissions and cut bills," said Energy and Climate Change Minister Lord Hunt.

Energy supplier will be in charge of the roll-out of smart meters rather than distribution networks. It will cost £340 per household. However, the money will be recouped through lower bills.

This plan shows that smart meters enable customers to reach a clear understanding of how much electricity and gas they use. This technology will help people to reduce waste, and pay more attention on usage to cut their energy bills. And it can help people realize how to reuse wasted resources. It will achieve maximum savings and rational utilization of human resources, and increase the use of renewable energy.

Currently, the average annual bill in UK is more than £800 for gas and £445 for electricity. It will cost £8bn for the plan to be implemented, and it will help people save at least £28 per year. The Department for Energy and Climate Change hopes that 47 million meters in 26 million properties will be installed by 2020.

"This will be the single biggest revolution in energy use since British Gas converted all the nation's homes to natural gas in the 1970s," said Mark Daeche, of energy

company First Utility.

Energy providers will be responsible for fitting the meters and there will be a huge amount of the work for that as British Gas changed appliances in 17 million homes to natural gas back in the 1970s.

Industry sources in the UK show that the £7bn cost amounts to around £15 per household per year between 2010 and 2020. However, £10 of that will be accounted for in cost savings by the suppliers. For the average consumer it will be possible to save 2% to 3% of energy use per year, and will cut £25 to £35 off their bills.

In total, customers could save more than £20 a year. The government says that they could save around 2% of the total energy use, which would cut £100m from bills by 2020. That means it could reduce CO_2 emissions by 2.6m tones.

In the future the smart meter will need to process a variety of emergency situations, recycling of various resources and effective reuse of resources [16].

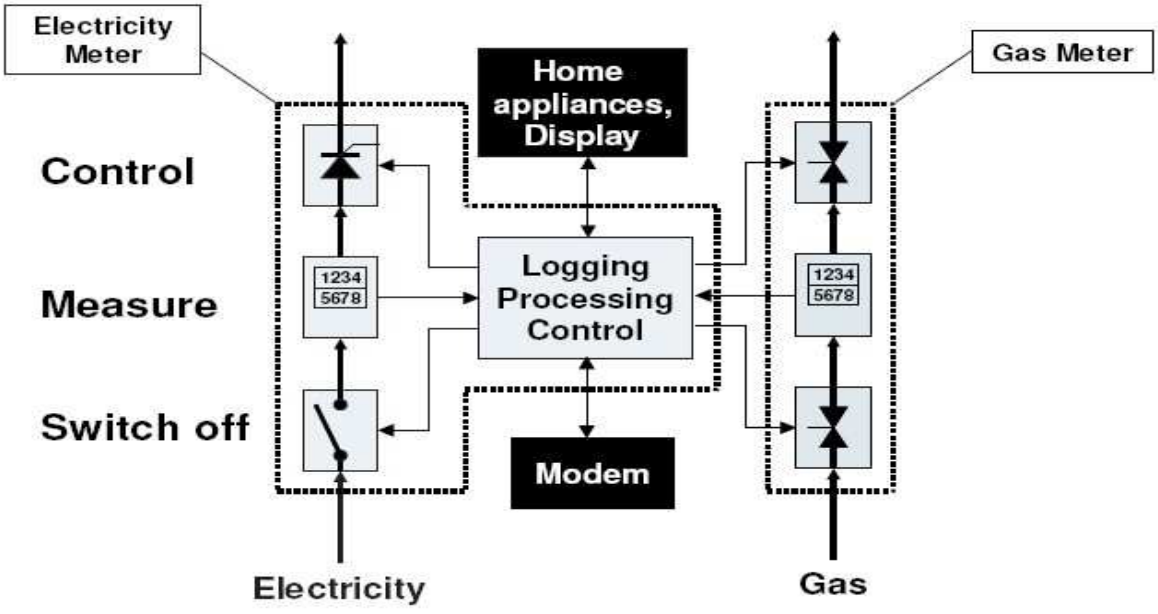


Figure 3.2 Overview of a typical smart meter [17]

3.4.2 China Smart Meter and Smart Grid Project

China has started many projects related to Smart Grid, which will change how the whole country generates and uses energy. All the projects will improve the efficiency and security of the whole power grid.

"A Smart Grid is an inevitable choice for China to address issues in its power industry and develop a lower-carbon economy," said Jiao Jian, an analyst at SYWG Research and Consulting. SYWG is the research arm of the Chinese brokerage Shenyin & Wanguo Securities.

All the proposals for the Smart Grid will increase Chinese investment in the power industry. For years the sector has suffered from lacklustre funding that was resulting in blackouts and the infamous infrastructure collapse during the snow storms in 2008 [17].

There are 1.18 million kilometres of old transmission lines that carry 3 million gigawatts of electricity in the entire Chinese power grid. Around 7% of the 3 million gigawatts is lost through power transmission. But it is expected that demand for power will double by 2020.

Nowadays, most of China's power is generated by unclean coal plants. They generate around 70% of China's energy. The Chinese government has begun to clean up the energy resources. They want to achieve 15% of the total power supply by using renewable power generation by 2020.

Most of the proposals for the smart grid need the integration of the renewable power resources such as wind and solar. And GE will collaborate with State Grid on their smart grid development.

China uses UHV (ultra-high voltage) lines which can afford efficient power flow without big losses. This kind of transmission system is very useful in such big country where resources are rich but unevenly distributed. China has started operating 640-

kilometre UHV lines across the central part of China in January 2009. State Grid is satisfied with the line performance and now they are installing two further lines, which will carry power across 2000 kilometres. State Grid expected investment in the UHV lines can be effective and will therefore spend 300 billion Yuan ($43.94 billion) on the effort by 2012 [17].

3.4.3 Smart Metering Projects in Italy

Italy has the largest number of Smart Metering projects in the world, undertaken by ENEL SpA. As a result, the Italian has become one of the first beneficiaries of the new technologies on energy economy.

Before this project started to deploy in 2001, Italy suffered from a variety factors bringing about high energy cost. The prices of electricity were averagely higher in Italy than in other European countries, partly because around 70% of electricity was generated from hydrocarbons, while a similar percentage was generated from nuclear power and other sources elsewhere in Europe. The amount of fraud and bad debt is another reason worth considering. Also, improving data to achieve better management generation and prevention blackouts is an important business need driving this deployment as well [18].

In 2001, ENEL drew up a 5-year plan on Smart Metering for its entire customer base, which covers nearly 40 million homes and businesses. New meters are own designed with multi abilities integrated such as advanced power measurement and management capabilities and software-controllable disconnect switches. These meters are based on standards-based power line technology from Echelon Corporation, which will automatically reveal the time-of-day pricing to customers and send readings to a central office [19].

Table 3.1 Abilities integrated in Smart Metering

Abilities integrated in New Meters
Bi-directional communications
Advanced power measurement and management capabilities
Software-controllable disconnect switch
An all solid-state design

By 2006 ENEL had invested $3 billion in the whole smart metering deployment, comparing to its 'harvest' of cost savings up to $750 million from new technology annually. These savings include automatically collecting customer data and managing its energy network remotely, instead of paying for a costly technician. To Italian users, electricity pricing of the new meter could form logical energy consuming habits and concretely reduce the family cost of energy consumption.

Smart metering in Italy has been a good experiment for the world. The ENEL's Gallo recommended that corporations need to deploy their technology as quickly as possible. A whirlwind program will help achieve a fast return on investment. It also suggested that corporations make sure their consumers understand the whole system and the advantages of the programs [19].

3.4.4 Smart Metering Implementation in United States

The main driver for introducing a Smart Metering system in the USA, especially in California, is to improve power quality and electricity supply. California has a summer peak demand for power covering about 50-100 hours per year. Hence, more and more air conditioners are used. The main energy supplier company of California

has seen an increasing demand but expects this to decrease. All three main California utility companies developed their own projects to implement the smart metering system or automatic metering infrastructure (AMI) systems for all customers. Deployment plans envisage installing all smart meters and communications infrastructure by 2012 or 2013 [20].

On July 20, 2006, California's energy agency approved a project to roll out conventional meters with communications co-processor electronics, i.e. Smart Metering, to 9 million household customers of Pacific Gas and Electric Company (PG&E) in Northern California. Those meters can record and report electricity and gas consumption on an hourly basis. This action can help PG&E to set pricing, which changes by season and time of the day. And the users can also shift energy use to off-peak time. The peak pricing program will fully roll out over the next five years [21].

PG&E's $1.7 billion Smart Meter proposal received unanimous approval by the California Public Utilities Commission, which allows the utility to move forward with a major investment in new Smart Meters designed to provide a wide range of benefits to household customers while increasing operational efficiencies and energy saving at the utility [22].

PG&E's installation of 9.3 million Smart Meter devices for its 5.1 million electricity and 4.2 million gas customers began in Bakersfield in 2009 and will be completed in 2011 [22].

"PG&E's Smart Meter program is one of the cornerstones of a sweeping effort to take a dramatic leap forward in the way we deliver service to our customers," said Tom King, president and CEO of the Pacific Gas and Electric Company. Customers can obtain more and better information from the Smart Meter and make cost-savings in using energy. The PG&E also can receive message from their customers and give

rapid responses to restoring services, otherwise customers would have to phone the company call centers. When the Smart Metering system is installed, customers will not need to make a special appointment with meter readers to have their meters read. Customer can also check for daily information about energy use online and then they can make better decisions. Because the smart meter can record energy usage every hour, customers can voluntarily adjust their usage depending on differing energy prices that vary by season and the time of the day, by shifting their energy use from peak hours to off-peak hours [22].

A key feature of Smart Meter technology is the ability to reduce peak load on the very hottest days by providing financial incentives to customers who voluntarily shift electricity usage away from critical peaks, which in the case of California will reduce PG&Es need to purchase power to meet demand at the most critical times, help avoid strain on the power grid, and help reduce reliance on fossil-fuel generation. To achieve those benefits, the CPUC (California Public Utilities Commission) approved PG&E's proposal to provide customers with a Critical Peak Pricing choice. The Smart Meter device is almost identical in size and appearance to the existing electric meter. The gas smart meter module is a tiny part which can be installed in the existing gas meter. Many companies in the USA are ready to employ smart meters. However, PG&E's program is the largest one in the United States. PG&E's investment to develop the new smart meter technology is estimated to be $1.74 billion. PG&E projects that these investments will be offset through the energy saving achieved by using the new smart meters [22].

The largest municipal utility in the USA, the Los Angeles Department of Water and Power (LADWP), has chosen to develop its AMI serving their own customers. LADWP has bought 9000 already. All the customers in utilities can then gain as

awareness of their daily energy use from the smart meters, thus creating potential for decreasing their energy use, and contributing to global energy conservation. Austin Energy is the USA's ninth largest community-owned electric utility. There are around 400,000 electricity customers in Austin, Texas, where a two-way radio frequency (RF) mesh network and around 260,000 residential smart meters in 2008 have begun to be deployed. More than 165,000 smart meters had been installed by spring 2009 [11].

3.4.5 Smart Metering Implementation in Other Countries

Including the countries referred to above, smart metering is still fascinating other countries by its pre-eminent role in energy affairs. In Victoria State, Australia, a program named Advanced Metering Infrastructure is being deployed to help Victoria to manage their energy consumption and reduce carbon emissions. Nearly 2.2 million homes and 300,000 businesses have benefitted from this largest energy infrastructure reformation in the state's history. Features are included in AMI, such as a two-way communication between power corporations and the electricity meter at home; permission on accessing accurate electricity pricing reads every 30 minutes, and so on. The decision of the Victoria Government on rolling out smart metering followed an extensive cost-benefit study and a National Cost Benefit Analysis. These studies reveal that the rolling out of smart metering in Victoria could generate net benefits of up to $700 million over the coming two decades. Also, by replacing the existing electro-mechanical accumulation meters, Victoria's electricity users are able to use less than 160 MWh every year [23].

In addition, the Department of Primary Industries has held a series of forums to collect a wide range of ideas and views on AMI. Table 3.2 is part of these forums.

Table 3.2 Part of schedule of series of forums [23]

	Time	Description
Stakeholder Forum 1:	May 2006	Outlined the Victoria Government's decision to commence the AMI project and approach
Stakeholder Forum 2:	Aug 2006	Updated stakeholders on progress, and initial broad-based consultation on draft of AMI functionality
Stakeholder Forum 3:	Apr 2007	Provided a progress update and overview of proposed legislative and regulatory framework
Stakeholder Forum 4:	Dec 2007	Marked the end of the establishment phase and hand-over to industry for the implementation phase

As the largest neighbour of Australia in Oceania, New Zealand started its smart metering program in late 2006. Companies are rolling out their smart metering plan to their customers one after another. But in June 2009, the New Zealand parliament was presented with a report, criticizing the deficiency in smartness of the 150,000 smart meters installed. Commentators said that the questionable meters were insufficient for basic real-time monitoring functions. Also, the meters lacked a microchip at the initial installation stage, which helps meters to communicate with other devices. Other than the problems above, as there are no standard rules, vary differently between companies. It seems that additional resources should be used for retrofitting.

Back to Europe, Electricité Réseau Distribution France (ERDF) has announced a replacement program for smart meters covering 35 million electricity consumers in France, starting with a pilot trial of 300,000 meters. Atos Origin has been chosen as a

consortium manager and the architect for the information system named Automated Meter Management (AMM) [24].

3.5 Conclusion

In the view of the current state of Smart Metering technology, new meters are only developed in providing real-time information without giving any constructive suggestion to consumers. Considering this situation, a new concept, the so-called “smarter meter”, is being introduced. The smarter meter not only provides on real-time information services, but also helps in establishing an optimal energy-consuming plan for a specific user by the AI technologies integrated into the meter. The introduction of AI technology offers opportunities to recognise energy consuming patterns automatically and accurately other than by consumers themselves.

Chapter 4

WAVELET-GA-ANN BASED HYBRID MODEL FOR ACCURATE PREDICTION OF SHORT-TERM LOAD FORECAST

4.1 Introduction

Many power systems are not only being pushed to their limits to meet their customers' demands, but also use a lot of resources on their operation scheduling. Furthermore, power systems need to operate at an even higher efficiency in a deregulated electricity market whereby the generating companies (Gencos) and distribution companies (Discos) have to compete in order to maximize their profits.

To facilitate accurate load-forecasting analysis, a robust noise filtering and trend analysis algorithm must be used to enable effective eventual automation of the analysis of large volumes of data generated by the monitoring and recording of load consumption readings in any particular system. Currently, several forecasting schemes utilize Artificial Intelligence (AI) methods like ANN and GA to perform load-forecasting tasks. The common problem with such a method is that an AI scheme is only as intelligent as the program that trains it. This in turns depends heavily on the reliability of the training data collected. If such training data were in the first place corrupted by noise, it would mean that pre-processing of such data would be necessary. All these add to the implementation cost and set-up time. A good trend analysis scheme should be able to de-noise the electrical noise inherent in the data, and disregard portions of data where monitoring devices might have failed, giving lower resolution readings as a result, and be able to take a macro view of the trend while preserving temporal information. The analysis of non stationary signals like load consumption data often involves a compromise between how well important

transients can be located and how finely evolutionary behaviors can be detected. Extremely noisy data poses a problem to the operator as to how to ascertain the amount of noise in the retained high frequency transient data [25].

Interest in applying neural networks to electric load forecasting began more than a decade ago. Artificial neural networks based methods for forecasting have been able to give better approach in dealing with the nonlinearity and other difficulties in modeling of the time series data. ANNs have been applied recently in the area of time-series forecasting due to their flexibilities in data modelling [26-27]. Most of the approaches reported since are based on the use of an MLP network as an approximator of an unknown nonlinear relation. There have been some pioneering works on applying wavelet techniques together with ANN to time series forecasting, [28-32]. Among ANN based forecasting methods, Radial Basis Function (RBF) networks have been widely used primarily because of their simple construction and easier training as compared to Multi-layer Perceptrons (MLPs) in addition to their capability in inferring the hidden relationship between input and desired target patterns. This capability is attributed to its ability to approximate any continuous function to any degree of accuracy by constructing localized radial basis functions. From the standpoint of preserving characteristics of different classes, this local approximation approach has the advantage over the global approximation approach of multi-layer perception networks.

As large amounts of historical load patterns are needed in a typical load-forecasting algorithm, even low sampling rates of 1 sample per minute generate a huge amount of data. Hence, the effective compression of large data and faithful reconstruction of original signals from compressed data are major challenges for time series data. Also, when an ANN, especially RBF network, is trained with huge data (with noise), it may

result in not only a big network model, and very time consuming training, but also the network may fail to capture the true features in the data. With the development of wavelet transforms, the difficulty of effective data compression and faithful retrieval of original data can be well tackled. This pilot research was intended to test RBF networks model combine with wavelet transformed data for capturing useful information on various time scales. These strategies approximate a time-series at different levels of resolution using multi-resolution decomposition. Recent works [33] stress the use of shift invariant wavelet transforms, which is an auto correlation shell representation technique for making the analysis of time series data easier. This technique is employed to reconstruct singles after wavelet decomposition. With the help of this technique, a time series can be expressed as an additive combination of the wavelet coefficients at different resolution levels. RBF networks are optimized by FPGA. These data are then applied to build Neural-Wavelet based forecasting models to predict electricity demand as from the data obtained from a real electricity market.

In view of the above, the main objectives of the present work are: to develop a wavelet based RBF network model for accurate prediction of short-term load forecast (STLF). The use of neural network will enable online prediction of STLF in an effective and efficient way.

4.2 Wavelet Transforms in Load Forecast

Wavelet transforms, [34, 35] though known previously, have gained much attention only recently. It has been exploited in many fields like seismic studies, image compression, signal processing processes and mechanical vibrations. The flexible time-scale representations of wavelet transform has found its place in many applications that traditionally used modified forms of Fourier Transforms (FT) like

Short Time FT (STFT) and the Gabor Transforms. Its impressive temporal content and frequency isolation features have tempted researchers to use them in the area of power systems analysis.

Wavelet transforms provide a useful decomposition of a signal, or time series, so that faint temporal structures can be revealed and handled by nonparametric models. They have been used effectively for image compression, noise removal, object detection, and large-scale structure analysis, among other applications.

4.2.1 Time Series and Wavelet Decomposition in Load Forecasting

À Trous Wavelet Decomposition

The continuous wavelet transform of a continuous function produces a continuum of scales as output. On the other hand, input data is usually discretely sampled, and furthermore a dyadic or two-fold relationship between resolution scales is both practical and adequate. The latter two issues lead to the discrete transform. Figure 4.1 shows the wavelet decomposition.

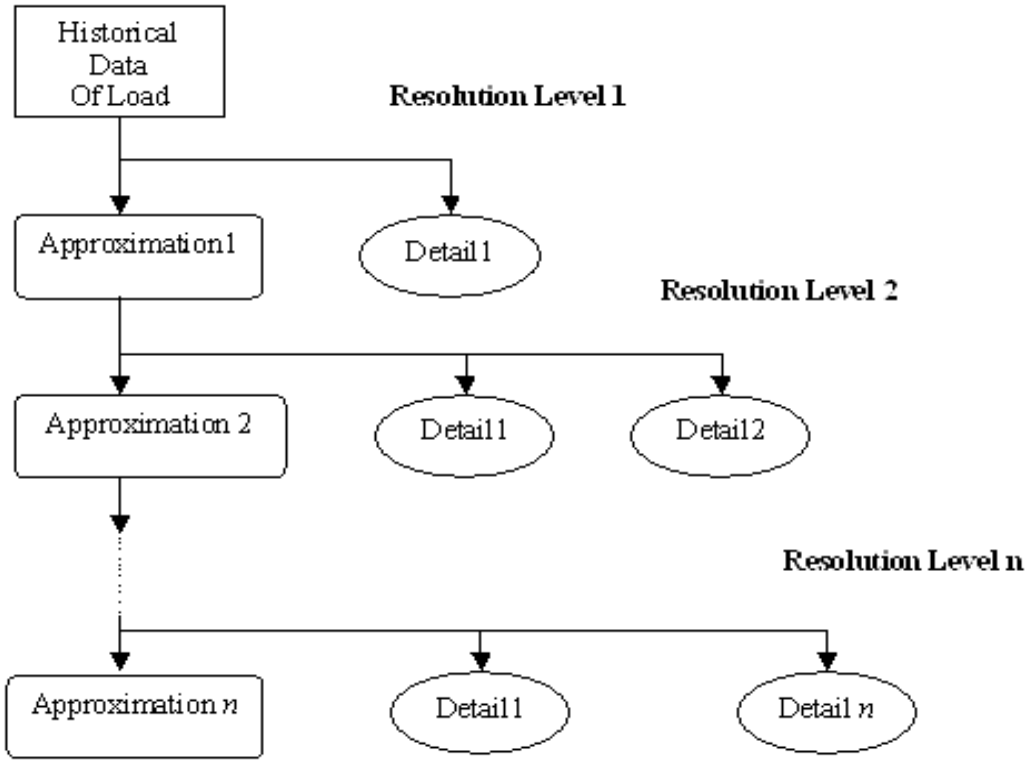


Figure 4.1 Wavelet decomposition process

Wavelet decomposition provides a way of analyzing a signal in both time and frequency domains. For a suitably chosen mother wavelet function ψ a function f can be expanded as:

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} w_{jk} 2^{i/2} \varphi(2^j t - k) \qquad (4.1)$$

where the functions $\varphi(2^j t - k)$ are all orthogonal to each other. The coefficients w_{jk} give information about the behaviour of the function f concentrating on the effects of scale around 2-j near time t×2-j. This wavelet decomposition of a function is closely related to a similar decomposition (the discrete wavelet transform, DWT) of a signal observed in discrete time.

It is well known that DWT has many advantages in compressing a wide range of signals observed in the real world. However, in time series analysis, DWT often

suffers from a lack of translation invariance. This means that DWT based statistical estimators are sensitive to the choice of origin.

The output of a discrete wavelet transform can take various forms [36]. Traditionally, a triangle (or pyramid in the case of 2-dimensional images) is often used to represent all that is worth considering in the sequence of resolution scales. Such a triangle comes about as a result of decimation or the retaining of one sample out of every two. The major advantage of decimation is that just enough information is retained to allow exact reconstruction of the input data. Therefore, decimation is ideal for effective compression. However, it can be easily shown that the storage required for the wavelet-transformed data is exactly the same as is required by the input data. The computation time for many wavelet transform methods is also linear in the size of the input data, i.e. $O(n)$ or *n*-length input time series. Also, with the decimated form of output it is less easy to visually or graphically relate information at a given time point at different scales. More problematic is their lack of shift invariance. This means that, if the last few values of the input time series are deleted, then the wavelet transformed, decimated output data will be quite different from heretofore. One way to solve this problem, at the expense of greater storage requirements, is by means of a redundant or non-decimated wavelet transform.

A non-decimated wavelet transform based on an *n*-length input time series, then, has an *n*-length resolution scale for each of the resolution levels of interest. Therefore, information at each resolution scale is directly related to each time point. This results in shift invariance. The extra storage requirement is by no means excessive.

An à *trous* algorithm is used to realize the shift-invariant wavelet transforms. Such transforms are based on auto-correlation shell representation [33] by dilations and translations of the auto-correlation functions of compactly supported wavelets.

By definition, the auto-correlation functions of a compactly supported scaling function ϕ(x) and the corresponding wavelet Ψ(x) are as follows:

$$\phi(x) = \int_{-\infty}^{\infty} \phi(y)\phi(y-x)dy$$
$$\psi(x) = \int_{-\infty}^{\infty} \psi(y)\psi(y-x)dy \tag{4.2}$$

The set of functions $\{\overline{\psi}_{j,k}(x)\}_{1\le j\le n_0, 0\le k\le N-1}$ and $\{\overline{\phi}_{n_0,k}(x)\}_{0\le k\le N-1}$ is called an auto correlation shell, where:

$$\overline{\psi}_{j,k}(x) = 2^{-j/2}\psi(2^{-j}(x-k))$$
$$\overline{\phi}_{n_0,k}(x) = 2^{-n_0/2}\phi(2^{-n_0/2}(x-k)) \tag{4.3}$$

A set of filters $P = \{p_k\}_{-L+1\le k\le L-1}$ and $Q = \{q_k\}_{-L+1\le k\le L-1}$ can be defined as:

$$\frac{1}{\sqrt{2}}\phi\left(\frac{x}{2}\right) = \sum_{k=-L+1}^{L-1} p_k\phi(x-k)$$
$$\frac{1}{\sqrt{2}}\psi\left(\frac{x}{2}\right) = \sum_{k=-L+1}^{L-1} q_k\phi(x-k) \tag{4.4}$$

Using the filters P and Q, the pyramid algorithm for expanding into the auto-correlation shell can be shown as:

$$c_j(k) = \sum_{l=-L+1}^{L-1} p_l c_{j-1}(k+2^{j-1}l) \tag{4.5}$$

$$w_j(k) = \sum_{l=-L+1}^{l=L-1} q_l c_{j-1}(k+2^{j-1}l) \tag{4.6}$$

These shell coefficients obtained from Equation 4.5 and Equation 4.6 can then be used to directly reconstruct the signals. Given smoothed signal at two consecutive resolution levels, the detailed signal can be derived as:

$$w_j(k) = \sqrt{2}c_{j-1}(k) - c_j(k) \tag{4.7}$$

The process of generating wavelet coefficient series is further illustrated with the block diagram as shown in Figure 4.2.

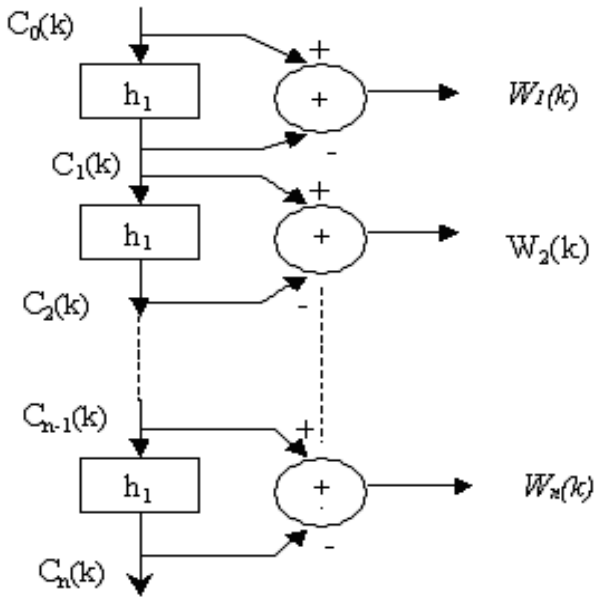


Figure 4.2 À Trous Wavelet transform of a time-series signal

Then the original signal $c_0(k)$ can be reconstructed from the coefficients $\{w_j(k)\}_{1\le j\le n_0, 0\le k\le N-1}$ and residual $\{c_{n_0}(k)\}_{0\le k\le N-1}$:

$$c_0(k) = 2^{-n_0/2} c_{n_0}(k) + \sum_{j=1}^{n_0} 2^{-j/2} w_j(k) \tag{4.8}$$

for k=0,…, N−1, where $c_{n_0}(k)$ is the final smoothed signal.

To make more precise predictions the most recent data shall be used. In case of adaptive learning, the previous data is penalized with forgetting factors. The time-based à trous filters similar to that are used to deal with the boundary condition. Figure 4.3 shows the wavelet recombination process.

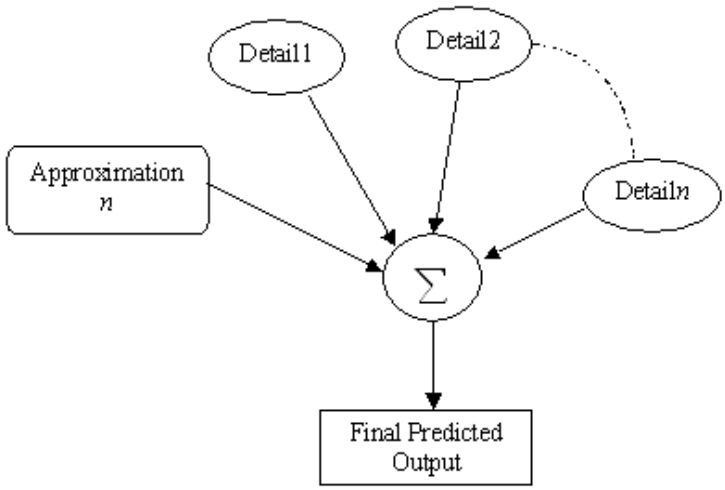


Figure 4.3 Wavelet recombination process

4.3 Radial Basis Networks

An RBF is a function has in-built distance criterion with respect to a center [37]. A typical RBF neural network consists of three layers (input, hidden, output). The activation of a hidden neuron is determined in two steps: the first is to compute the distance (usually the Euclidean norm) between the input vector and a center c_i that represents its hidden neuron; second, a function, that is usually bell shaped, is applied,

using the obtained distance to get the final activation of the hidden neuron. In the present case the well known Gaussian function G(x) is used.

$$G(x) = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma^2}\right) \tag{4.9}$$

The parameter σ is called unit width (spread factor) and is determined using the GA. All the widths in the network are fixed to the same value and the result in a simpler training strategy. The activation of a neuron in the output layer is determined by a linear combination of the fixed nonlinear basis functions, i.e.

$$F(x) = w_0 + \sum_{i=1}^{M} w_i \phi_i(x) \tag{4.10}$$

where $\phi_i(x) = G(\|x - c_i\|)$ and w_i are the adjustable weights that link the output nodes with the appropriate hidden neurons and w_0 is the bias weight. These weights in the output layer can then be learnt using the least-squares method.

The present work adopts a systematic approach to the problem of centre selection. Because a fixed centre corresponds to a given regressed in a linear regression model, the selection of RBF centres can be regarded as a problem of subset selection. The orthogonal least squares (OLS) method [38] can be employed as a forward selection procedure that constructs RBF networks in a rational way. The algorithm chooses appropriate RBF centres one by one from training data points until a satisfactory network is obtained. Each selected centre minimizes the increment to the explained variance of the desired output, and so ill-conditioned problems occurring frequently in the random selection of centres can automatically be avoided. In contrast to most

learning algorithms, which can only work if a fixed network structure has first been specified, the OLS algorithm is a structural identification technique, where the centers and estimates of the corresponding weights can be simultaneously determined in a very efficient manner during learning. OLS learning procedure generally produces an RBF network smaller than a randomly selected RBF network. Due to its linear computational procedure at the output layer, the RBF is shorter in training time compared to its back propagation counterpart.

A major drawback of this method is associated with the input space dimensionality. For large numbers of inputs units, the number of radial basis functions required can become excessive. If too many centres are used, the large number of parameters available in the regression procedure will cause the network to be over sensitive to the details of the particular training set and result in poor generalization performance (overfit).

The present work uses a floating point GA based algorithm for optimizing the centres and spread factors.

4.3.1 A Hybrid Neural-Wavelet Model for Short-Term Load Prediction

The proposed hybrid neural-wavelet model for short-term load prediction is shown in Figure 4.4. Given the time series f(n), n=1,…, N, the aim is to predict the l-th sample ahead, f(N+l), of the series. As a special case, l=1 stands for single step prediction. For each value of l separate prediction architecture is trained accordingly. The hybrid scheme basically involves three stages [28]. At the first stage, the time series is decomposed into different scales by auto correlation shell decomposition; at the second stage, each scale is predicted by a separate RBF network; and at the third

stage, the next sample of the original time series is predicted by another RBF network using the different scale's prediction.

For time series prediction, correctly handling the temporal aspect of data is one of the primary concerns. The time-based à trous transform as described above provides a simple but robust approach. Here we introduce an à trous wavelet transform based on the auto correlation shell representation for the prediction model usage. This approach is realized by applying Equation 4.7 and Equation 4.8 to successive values of t. As an example, given an electricity demand series of 1008 values, it is hoped to extrapolate into the future with 1 or more than 1 subsequent value. By the time-based à trous transform, it is simply necessary to carry out a wavelet transform on values x_1–x_{1008}. The last values of the wavelet coefficients at time-point t=1008 are kept because they are the most critical values for prediction.

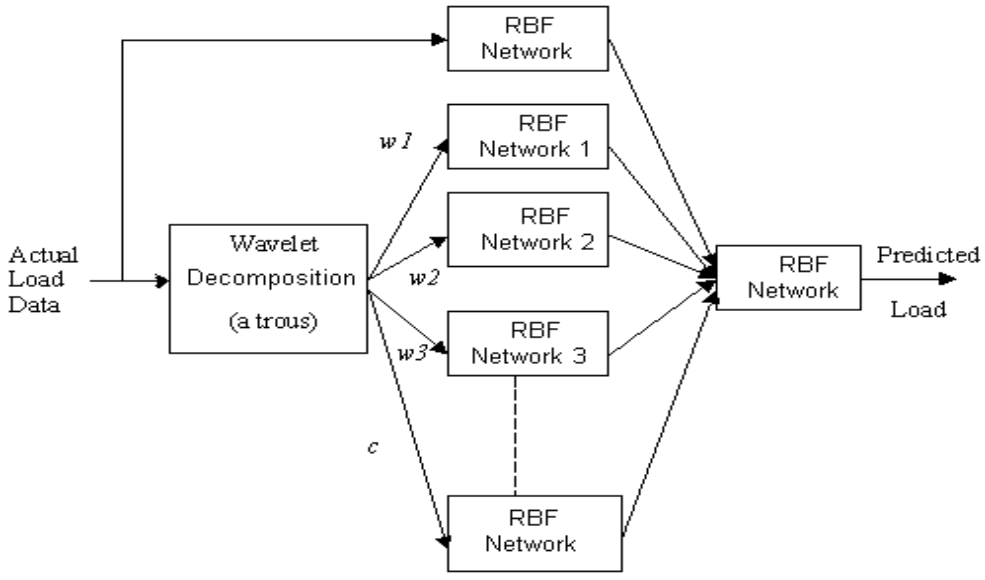


Figure 4.4 Overview of the neural-wavelet multi-resolution forecasting system (w_1, …, w_k are wavelet coefficients, c is the residual coefficient series)

Then the same procedure can be repeated, at time point t=1009, 1010, and so on. It empirically determines the number of resolution levels J, mainly depending on the inspection of smoothness of the residual series for a given J. Many of the high-resolution coefficients are noisy. Prior to forecasting, an over complete, transformed dataset.

Figure 4.5 shows the behaviour of the four-wavelet coefficients over 1008 points for a load series. Note that the data have been normalized for wavelet analysis. Normalization of data is an important stage, for training the neural network. The normalization of data not only facilitates the training process but also helps in shaping the activation function. It should be done in such a way that the higher values should not suppress the influence of lower values and the symmetry of the activation function is retained. The input load data is normalized between the minimum value, -1 and the maximum value, +1 by using the formula.

$$\left(\frac{Actualvalue - Minimum}{Maximum - Minimum}\right) \times (Maximum - Minimum) + Minimum \tag{4.11}$$

The load data should be normalized to the same range of values. The original time series and residual are plotted at the top and bottom in the same figure, respectively. As the wavelet level increases, the corresponding coefficients become smoother. The ability of the network to capture dynamical behaviour varies with the resolution level; this will be discussed in the next section.

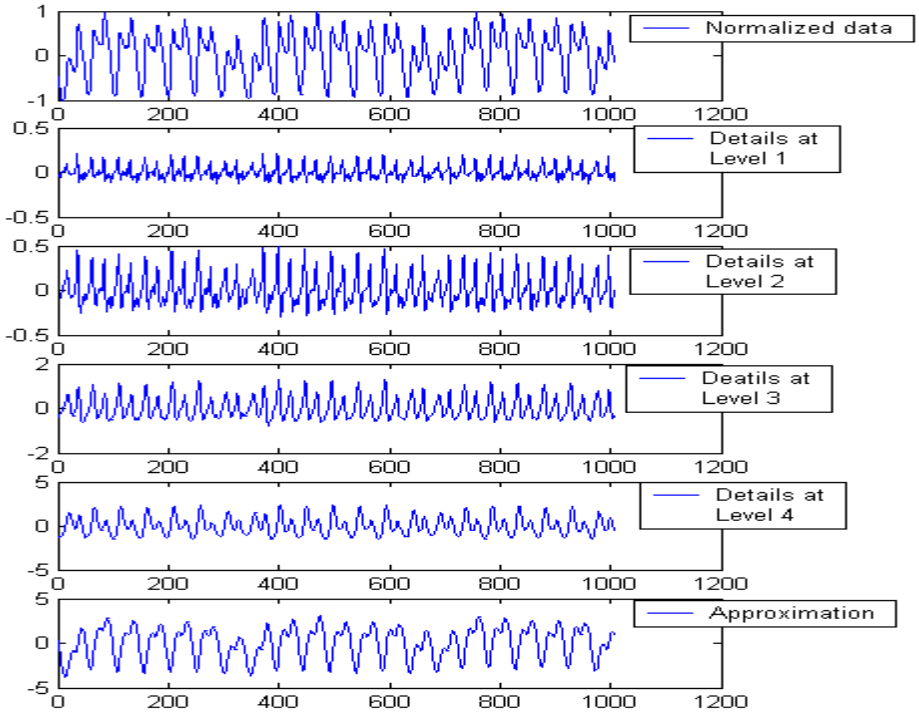


Figure 4.5 Illustrations of the à trous wavelet decomposition of a series of electricity demand

At the second stage, a predictor is allocated for each resolution level and the following wavelet's coefficients $w_i^j(t)$; j=0,…, J; i=1,…, N are used to train the predictor. All networks used to predict the wavelets' coefficients of each scale are of similar feed forward RBF perceptrons with D input units, one hidden layer with radial basis function as an activation function, and one linear output neuron. Each unit in the networks has an adjustable bias. The D inputs to the j-th network are the previous samples of the wavelets' coefficients of the j-th scale. In the proposed model implementation, each network is trained by the orthogonal least squares (OLS) method, which can be employed as a forward selection procedure that constructs RBF networks in a rational way. The procedure for designing neural network structure essentially involves selecting the input, hidden and output layers. At the third stage, the predicted results of all the different scales $\hat{w}_{N+i}^j(t)$, j=0,..., J are appropriately

combined. Three methods of combination have been discussed and compared here. In the first method, was simply applied the linear additive reconstruction property of the à trous, see Equation 4.8. The fact that the reconstruction is additive allows the predictions to be combined in an additive manner. For comparison purposes, a plain RBF was also trained and tested for original time series, denoted as RBF, without any wavelet preprocessing involved.

The target selection is an important issue in applying neural networks to time series forecasting. A neural network, whose output neurons are reduced from two to one, will have half the number of network weights required. It also carries important consequences for the generalization capability of the network. A single output neuron is the ideal case, because the network is focused on one task and there is no danger of conflicting outputs causing credit assignment problems in the output layer. Accordingly, it is preferred to have a forecasting strategy, which proceeds separately for each horizon in the second stage.

4.4 Simulation Results

The proposed model is tested with two sets of historical data containing the electricity load for the month of July 2005 and the month of July 2006, on a half-hourly basis; both are sets of electricity load data from Queensland. The sets of electricity load data are downloaded from the NEMMCO website [39].

The simulation results are obtained through the use of four different programs. These programs were written in MATLAB command line in association with MATLAB toolboxes on wavelet, and neural network. Programs are run on a PC of Pentium IV, 256 MB RAM, 3.2 GHz.

Before the wavelet decomposition technique (à trous) is applied, the sets of historical load data are first normalized.

The model is evaluated based on it prediction errors. A successful model would yield an accurate time-series forecast. The performance of the model is hence measured using the absolute percentage error (APE), which is defined as

$$APE = \left(\frac{|x_i - y_i|}{x_i} \right) \times 100 \qquad (4.12)$$

where x_i is the actual values and y_i is the predicted values at time instance i. This error measure is more meaningfully represented as an average and standard deviation (S.D.) over the forecasting range of interests. Additional measure of the error is defined from the cumulative distribution function as the 90% of the absolute percentage error, which provides an indication of the behaviour of the tail of the distribution of errors and indicates that only 10% of the errors exceed this value.

The forecasting results from the different forecasting schemes are presented in Table 4.1. The RBF network is optimized using FPGA in terms of number of inputs, centres, and spread factor. The number of neurons in the hidden layer is auto-configured by the OLS algorithm. Table 4.1 shows that the à trous wavelet transform system with adaptive combination coefficients for summing up the wavelet coefficients forecasting is the best in seven step ahead forecasting for the testing data, with regards to the mean, variance and percentile over the absolute percentage error (APE).

Parameters for FPGA algorithm:

Population Size = 40

Maximum Iterations = 30

Operators for FPGA:

1. Heuristic crossover

2. Uniform mutation
3. Normalized geometric select function

Table 4.1 Load forecast performance on testing data on APE measure for FPGA optimized spread factor and input

Scheme Number	1	2	3	4	5	6	7
μ_R	5.00	5.11	5.66	6.14	6.73	7.20	7.85
σ^2_R	0.7	1.1	1.6	2.3	2.9	3.6	3.4
η_R	0.086	0.093	0.107	0.123	0.138	0.150	0.156
μ_w	1.43	1.13	1.19	1.53	2.22	1.55	1.14
σ^2_w	0.057	0.0662	0.052	0.065	0.083	0.183	0.175
η_w	0.024	0.0216	0.022	0.025	0.033	0.036	0.034

$\mu_{R,}$ μ_w are mark-up coefficients; η_R, η_w are learning rates; $\sigma^2_R \times 10^{-3}$, $\sigma^2_w \times 10^{-3}$ are the spread factors and subscript R refers to the results with only RBF networks while subscript W refers to the results with hybrid wavelet-RBF model.

4.5 Conclusion

This chapter describes a hybrid model developed through wiser integration of wavelet transforms, floating point GA and artificial neural networks for prediction of short-term load. The use of wavelet transforms has added the capability of capturing both global trend and hidden templates in loads, which is otherwise very difficult to incorporate into the prediction model of ANN. Auto-configuring RBF networks are used for predicting the wavelet coefficients of the future loads. Floating point GA (FPGA) is used for optimizing the RBF networks. The use of GA optimized RBF networks has added to the model the online prediction capability of short-term loads

accurately. The results demonstrate that the proposed model is more accurate when compared to the RBF only model.

Chapter 5

INTELLIGENT WEATHER FORECAST

5.1 Introduction

Nowadays, weather forecasting has become one of the most challenging problems around the world, due to not only its practical value in meteorology, but also because it is a typically unbiased time-series forecasting problem in scientific research. Every sign points to the facts that there is a recognized need for accurate estimates of temperature and rainfall on a variety of temporal and spatial scales. Applications include climate monitoring, drought detection, severe weather and flash flood warnings, river monitoring and control as well as use in numerical weather prediction model initialization and verification. Also many areas of agriculture including crop growth and production, are extremely reliant the weather conditions, especially under some exceptionally high or low temperatures and in particular rainfall [40, 41].

In addition, because of the strong relationship between outside temperature and electric power demand (load), most short term load forecasting techniques use, at least in part, hourly temperature forecasts in generating a load forecast. Often the performance of short term load forecasters as reported in the literature is evaluated using actual temperatures. However, when the short term load forecasting is actually used at a utility, these future temperatures are not known and forecasts must be used instead [42, 43].

In recent years, more and more intelligent weather forecasting based on artificial neural networks (ANNs) have been developed and their performance evaluated. Due to their potential to represent complex nonlinear behavior, such as the relationship

between future temperatures and available data that contribute to certain weather conditions, ANNs are becoming increasingly prominent in many areas of weather forecasting. Also, the considerable success of ANNs in short term load forecasting and agriculture risk management encourages the application of these methods to other forecasting problems.

As reported in the literature [44], the existing multilayer neural networks are not ideally suited to functional approximation of biological activities as they do not capture the time-related dynamics of biological events. This thesis describes the methodology of short-term temperature and rainfall forecasting using dynamic weighted time-delay neural networks (DWTDNN), which are more suited for computations on temporal patterns from biological events. The methodology adapted in this work extends previous studies relying on observational data from a single point station to multiple point stations with time-series weather records on the east coast of China. The corresponding data preprocessing procedures were important for this kind of neural network modeling which was based on the back propagation architecture. This involved variable transformation and classification algorithm for input selection.

5.2 Modeling structure of DWTDNN

Traditionally, neural networks can be broadly classified into static neural networks and dynamic neural networks. Static neural networks like multi-layer perception (MLP) and (radial basis function) RBF are additive networks in which the activations are a function of a set of learned weights. A dynamic neural network, on the other hand, can be thought of as an extension to static networks with a short-term memory mechanism. These networks are also referred to as temporal pattern processors

because they effectively capture patterns in time. The popular dynamic neural networks are the time-delay neural networks (TDNN), the concentration-in-time neural networks and the gamma networks [44].

5.2.1 Structure of TDNN

TDNN, sometimes also called Finite Impulse Response (FIR) networks, are feed-forward networks which are used for a translation invariant recognition of pattern sequences (e.g. time series forecasting, speech processing or wear monitoring).

In a specific application, a set of possible inputs (features) and a set of outputs of the network are given, as well as any tasks performed by humans involving decision making and behavioral responses to spatiotemporal stimulation. Therefore, the recognition and processing of time-varying signals are fundamental to a wide range of cognitive processes. The central problem in processing temporal signals is consideration of representation of past inputs and deciding how the past inputs affect the current activations of the network. The past few years have witnessed an explosion of research on neural networks for temporal processing, most of which is centered on the TDNN, which utilizes an extension of back-propagation algorithm. The short-term memory in the case of TDNN generally consists of delay elements. A sample structure [45] of TDNN is shown in Figure 5.1.

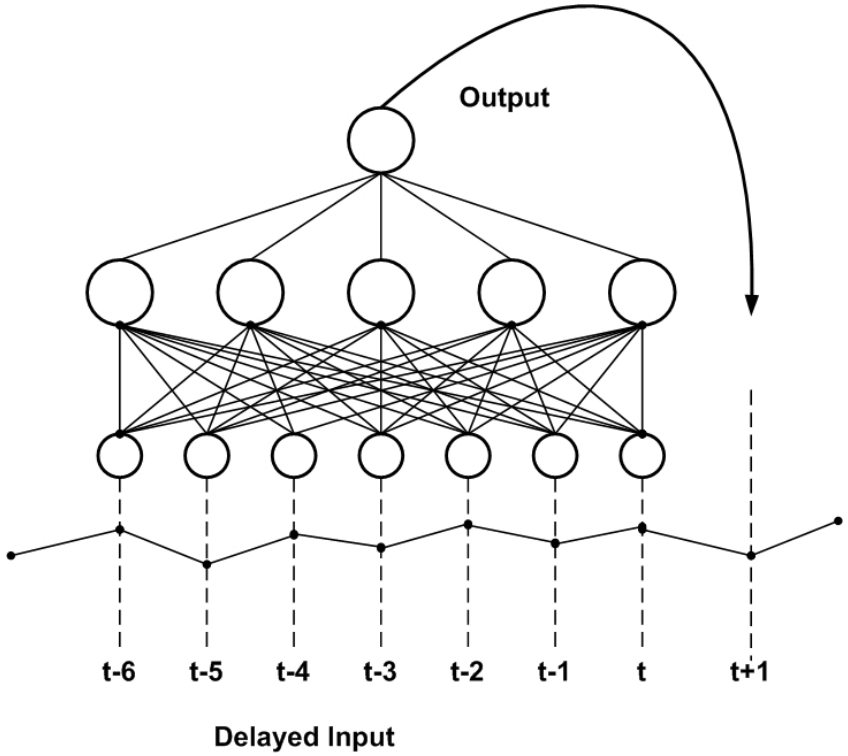


Figure 5.1 A time-delay feed-forward network with one hidden layer, where the input is a sequence of recent values of the time series

5.2.2 Neuron of DWTDNN

V. Venk and J.C. Raj [44] presented a dynamic weighted time-delay neural network (DWTDNN) in which each neuron in the input layer is scaled by a weighting function that captures the temporal dynamics of the biological task. This network is a simplified version of the focused gamma network and an extension of TDNN as it incorporates *a priori* knowledge available about the task into the network architecture [44]. In addition to learning the static weights, the network attempts to identify the parameters of the weighting function that are best suited for the application. Proper choice of this weighting function and its parameters improves the convergence of the network in addition to providing a better approximation.

According to the literature, the architecture of the DWTDNN is shown in Figure 5.2 and the corresponding output of the DWTDNN is given by

$$y(t) = f(\sum_{i=-T}^{T} x(t-i)w_i\Phi_i(\theta)) \tag{5.1}$$

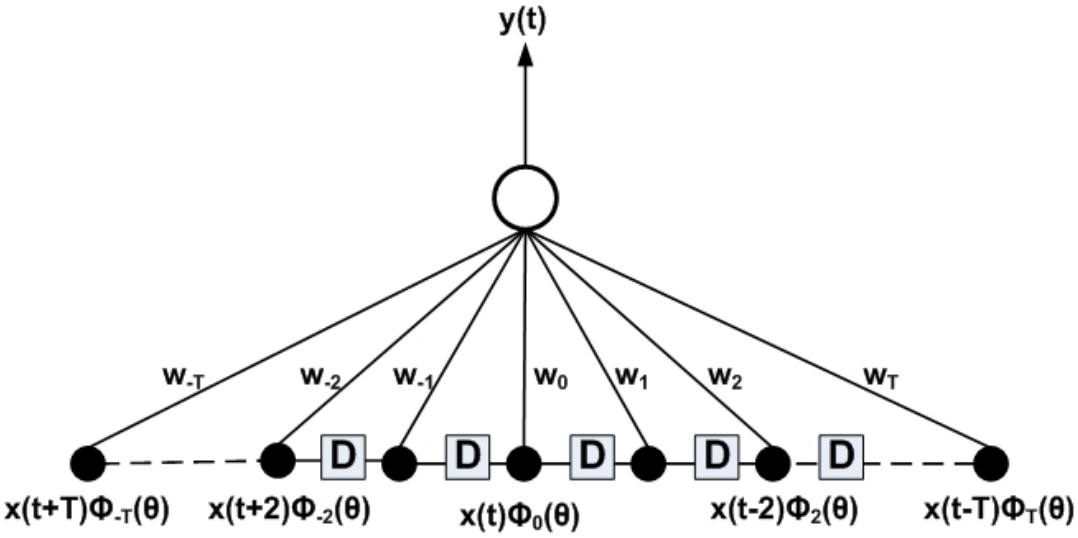


Figure 5.2 Architecture of a dynamic weighted time-delay neuron

where $x(t)$ denotes the time-delayed input signal, w_i denotes the static weights connecting the ith input node to the output neuron and $\Phi_i(\theta)$ denotes the dynamic weighting that scales the input to the ith node.

In order to choose an appropriate weighting function, the prior knowledge available about the task will be used, whilst the parameters θ={ θ_1, θ_2,……., θ_n} will be learnt during the procedure of training. Generally, the network attempts to minimize the given mean square error E: (assume that the input-output pair is (x_p,d_p))

$$E = \frac{1}{2}\sum_{p}(d_p - y_p)^2 \tag{5.2}$$

where y_p is the output of the network according to the input x_p and the weights are adapted using gradient descent as:

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i} = -\eta \sum_p (d_p - y_p) \frac{\partial y_p}{\partial w_i} \tag{5.3}$$

Use y_p given by Equation 5.1 and Equation 5.3 can be extended to Equation 5.4:

$$\Delta w_i = \eta \sum_p (d_p - y_p) x_p(t-i) \Phi_i(\theta) f'\left(\sum_{i=-T}^{T} x_p(t-i) w_i \Phi_i(\theta)\right) \tag{5.4}$$

And there is also:

$$\Delta \theta_l = \eta \sum_p (d_p - y_p) \left(\sum_{i=-T}^{T} x_p(t-i) w_i \Phi'_{il}(\theta)\right) \cdot f'\left(\sum_{i=-T}^{T} x_p(t-i) w_i \Phi_i(\theta)\right) \tag{5.5}$$

Where $\Phi'_{il}(\theta)$ denotes the derivative of $\Phi_i(\theta)$ with respect to θ_l, which is adapted according to Equation 5.5. If sigmoid functions are used, a non-linear mapping is obtained. It should be noted that for learning the parameters of the weighting function, the function has to be continuous and differentiable [44].

The dynamic weighted time-delay neural network (DWTDNN) is a multi-layer feed forward network that consists of a set of dynamic weighted time-delay neurons that constitute the input layer, a set of one or more hidden layers and an output layer both consisting of computation nodes.

5.3 Intelligent Weather Forecast Based on DWTDNN

5.3.1 Original Data Description

Meteorological data during the period between 1999 and 2001 were collected from 14 regional weather stations over the east coast of China as shown in Figure 5.3, and the

corresponding geographical details, such as longitude, latitude, and altitude are listed in Table 5.1. As shown, a finite number of stations will be used in proximity to an area for forecasting, which is a traditional methodology adopted by

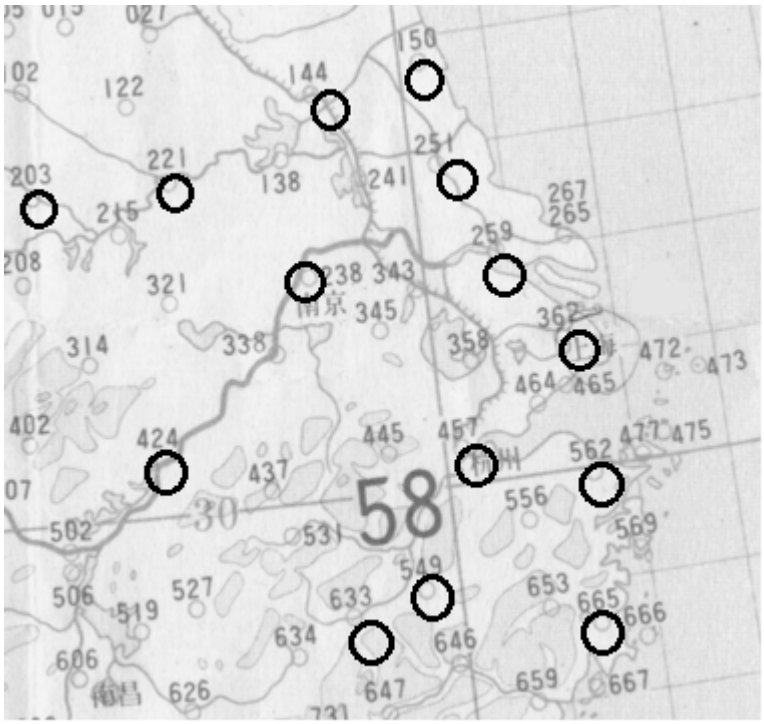


Figure 5.3 Geography information image of 14 regional weather forecast stations

the Bureau of Meteorology. After some extensive data mining and analysis the following 8 meteorological variables (as shown in Table 5.2) have been found to be suitable and reliable in terms of application, data consistency and validation. All the units of the data had been adjusted to international standard units to keep the consistency of the training data and eliminate the complicated and non-linear relationships between individual variables. About 5000 data were collected and interval time was 15 minutes.

Table 5.1 Weather forecast stations information

SNO	City	(longitude, latitude)	Altitude (m)
58-150	Sheyang	(120.25E, 33.77N)	6.7
58-144	Huaiyin	(119.03E, 32.60N)	19.0
58-251	Yancheng	(120.32E, 32.87N)	5.0
58-259	Nantong	(120.85E, 32.01N)	5.8
58-221	Benbu	(117.37E, 32.95N)	22.0
58-203	Fuyang	(115.83E, 32.93N)	38.6
58-238	Nanjing	(118.80E, 32.00N)	12.5
58-362	Shanghai	(121.46E, 31.41N)	8.2
58-457	Hangzhou	(120.17E, 30.23N)	43.2
58-424	Anqing	(117.05E, 30.53N)	19.6
58-562	Yinxian	(121.56E, 29.86N)	5.2
58-549	Jinhua	(119.65E, 29.11N)	64.7
58-633	Quzhou	(118.87E, 28.97N)	67.1
58-665	Hongjia	(122.25E, 28.63N)	2.2

All the values above are from Shanghai Meteorology Center

Table 5.2 Meteorology variable

Meteorological Variable	Unit	Range
Total quantity of cloud	1	0-10
Wind direction	°	0-360
Wind speed	m/s	0-30
Air pressure at sea level or at the station	hpa	0-300
Quantity of 6 hour rainfall	mm	0-50
Dew point	˚C	0-50
Discernable distance	0.1km	0-50
Temperature	˚C	0-50

All the values above are from Shanghai Meteorology Center.

The weather forecast data can be obtained through several approaches; most portable solutions are flat plain ASCII files and XML, data records from databases such as Oracle, DB2, Sybase, or SQL-Server, as well as data instruction by the aid of web services for sharing data in XML across the internet. The data format of the flat plain ASCII file is a-station serial number (long integer), b-longitude, c-latitude, d-altitude (floating point number), e-level of station (integer), f-total quantity of cloud, g-wind direction, h-wind speed, i-air pressure over sea level (or air pressure of self-station), m-quantity of 6 hour rainfall, q-dew point, r-discernable degree, t-temperature. Note that the default value of the missing record is 9999. A rainstorm exists when there is a period of continuous rainfall, spanning 24 hours, that amounts to at least 50mm over each of the four six-hour periods of that total period. For the air pressure over sea level, if the value is less than 70 hpa, e.g. 23, the real value is 1002.3, while if the value is greater than 70, e.g. 96, the value is 999.6.

5.3.2 Data Preprocessing

In order to maintain the integrity of the training data, some incomplete data should be approximated using some linear function interpolating with the nearby values of the same element within the region. If too many data records are missing, e.g. 5 in a block, neural networks or other statistical methods should be used for interpolation of several missing data records. In fact, some obvious training data outliners can be identified by plotting directly as shown in Figure 5.4, where the exceptional value of the air pressure over sea level at the Shanghai weather forecast station at the 16^{th} observation point or so can be recognized automatically on the basis of the probability logic and broken off in time.

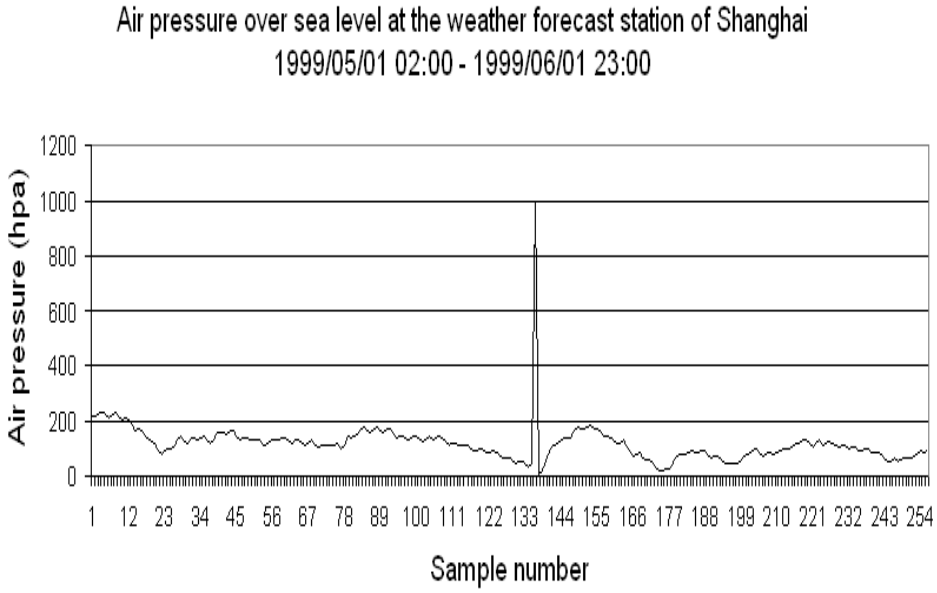


Figure 5.4 Air pressure over sea level at the weather forecast station of Shanghai

For the training architecture of back-propagation neural networks, the procedure of reducing the chaotic behaviors is an important and necessary step, especially for the meteorological variables which change tempestuously, such as wind speed, wind direction, discernable distance and cloud quantity. The filtered curve is much smoother than the unfiltered curve but the general tendency is still contained within the data. That is what the neural networks should understand and learn. Of course it would be better if the neural networks can learn the outliners too. In the end it depends on the filter curve. If the filter curve is weaker in the high frequencies, more details will remain after filtering. Figure 5.5 and Figure 5.6 show the filtering of wind speed on temporal domain and frequency domain respectively.

Low-pass has been adjusted so that general tendency is presented in data. If low pass is too strong (low high frequency) important details may be removed. If, on the other hand, the filter is too weak (more high frequencies) too many outliners may reduce network training success.

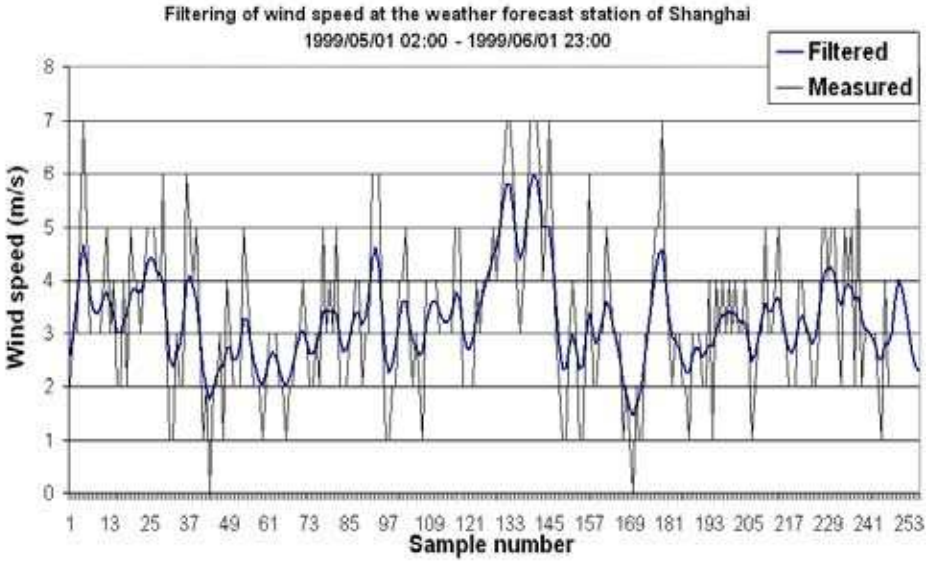


Figure 5.5 Filtering results of wind speed at the weather forecast station of Shanghai

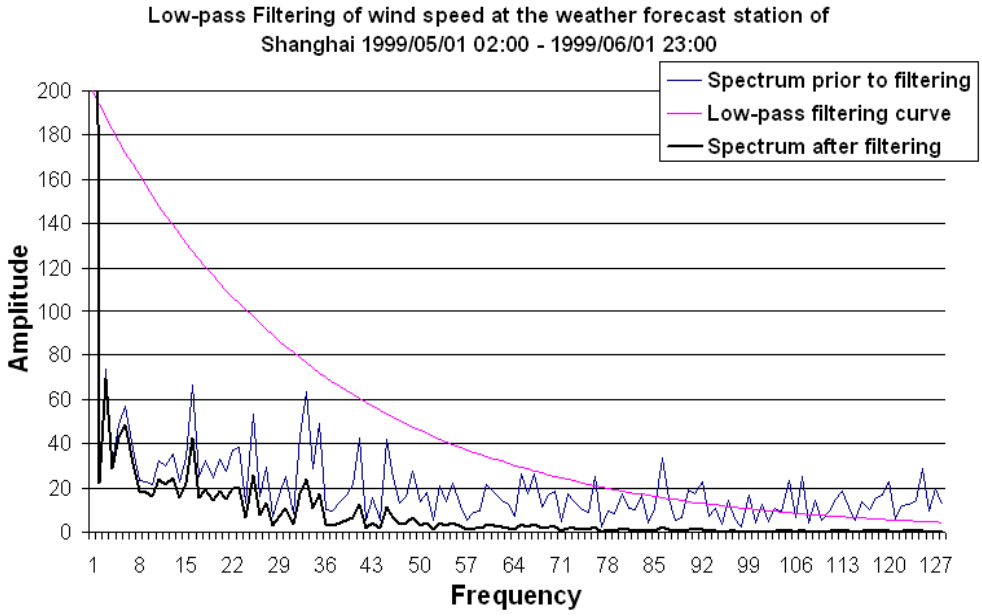


Figure 5.6 Low-pass filtering procedure of wind speed at the weather forecast station of Shanghai

After the generation of training patterns with filtered data to fit data models has been undertaken to determine two different training pattern structures, they need to be normalized depending on the chosen neuron activation function. One weather station can make time series forecasting using its own historical data; or it can use multiple stations' data to predict the one station's weather forecast. If the activation function has a range of [-1, +1], the normalization of input and output will be in the range of [-1, +1]. If the activation function has a range of [0, 1], normalization should be in the range of [0, 1]. It is not absolutely necessary to normalize the input vectors but it is common practice if pattern based normalization is applied.

There are two ways for the normalization of patterns: vector and column based. Vector based normalization is generally chosen if data columns have a similar range and units. Column based normalization is generally chosen if data for each column is unrelated in terms of units and range as it is the case with weather data.

$$\left(\frac{x-min}{max-min}\cdot(high-low)\right)+low \tag{5.6}$$

$$\left(\frac{\log(x)-\log(\min)}{\log(\max)-\log(\min)}\cdot(high-low)\right)+low \tag{5.7}$$

Equation 5.6 is for linear normalization and Equation 5.7 is for logarithmic normalization. Equation 5.6 could be optimized but, for direct comparison with linear equation, log has been added. Any formula or mathematical procedure that is reversible can be used for normalization. In addition, normalization is dependent on neuron activation function especially for output neuron where error difference is calculated.

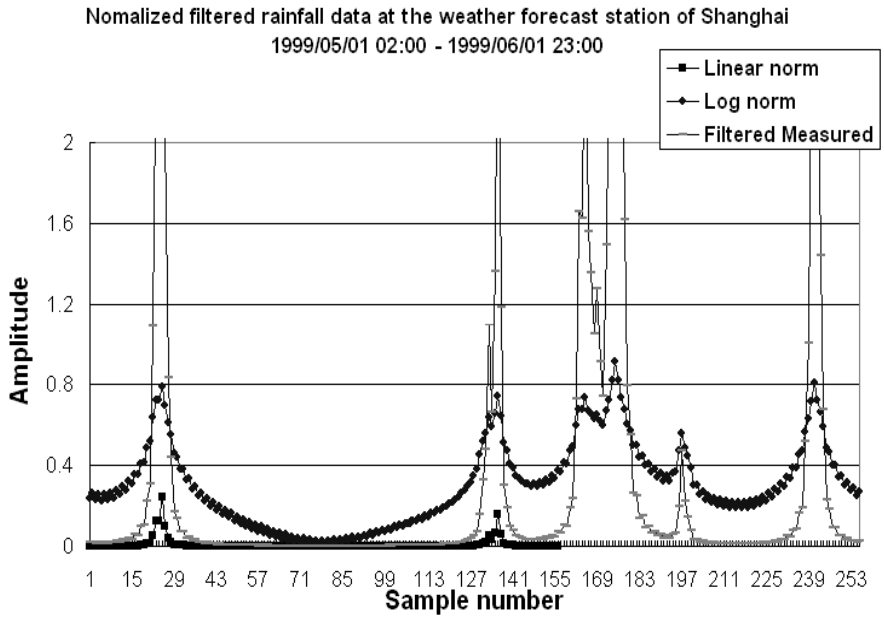


Figure 5.7 Normalized filtered results of rainfall at the weather forecast station of Shanghai using linear and logarithmic functions respectively

Outliners can be kept if logarithmic normalization is used. Outliners can be important e.g. for electricity demand forecasting where there are peaks at certain times during the days and months. Lower values are amplified since log (1000) =3 and log (100) =2, a data range between 100 and 1000 can be transformed into the range of [2, 3]. Furthermore, logarithmic functions are available as standard in almost any maths package and has been optimized in the mathematical coprocessor that is part of the main processor. Figure 5.7 shows the results of normalization for rainfall data after the procedure of filtering.

5.4 Results

There were 14 stations nearby, including the Shanghai weather forecast station

(58362), whilst different distances of surrounding stations will have impact on forecasting accuracy. In this thesis, results for up to 8 stations will be given. Figure 5.8, Figure 5.9 and Figure 5.10 show the results of using 2, 4, and 8 stations during the construction of the input neuron of DWTDNN. Training results will be unrelated if network sizes differ significantly; the same network configuration will be used to have means of comparison. The equation of the generalization error is:

$$SSER = \sqrt{\sum_{c=1}^{C} (T_C - O_C)^2} \tag{5.8}$$

Where T_C, O_C denote the target value and output value respectively. The target value is historical data.

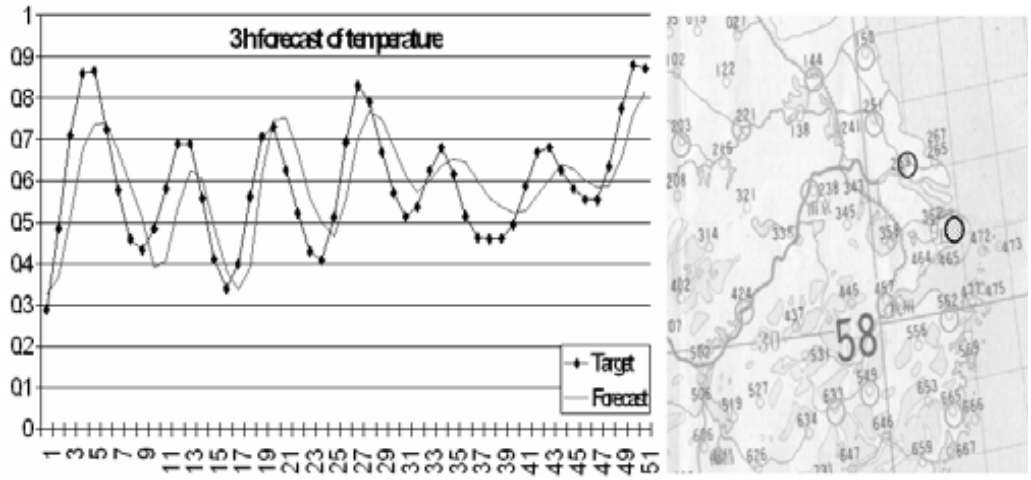


Figure 5.8 Temperature forecasting using 2 stations (58362-58457) with generalisation error: 0.7092

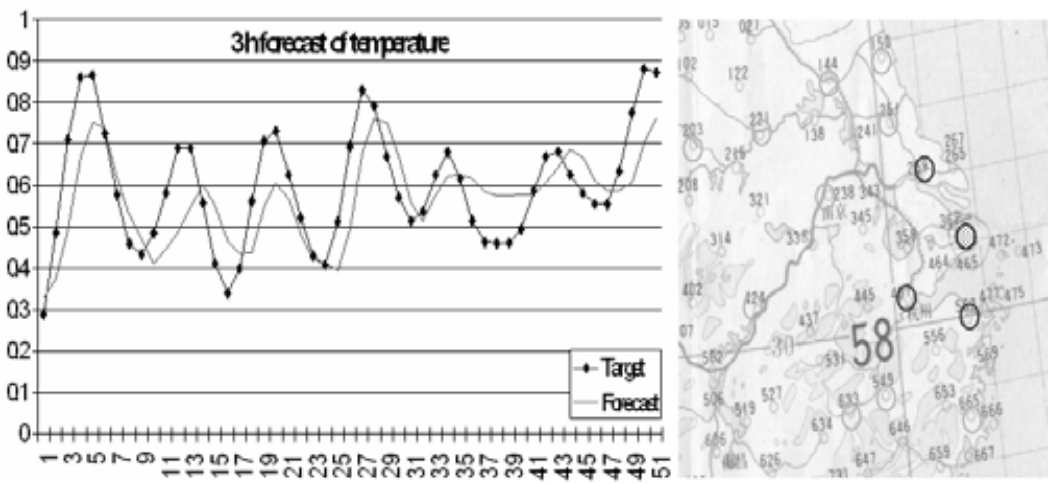


Figure 5.9 Temperature forecasting using 4 stations (58362-58457-58259-58562) with generalization error: 0.7485

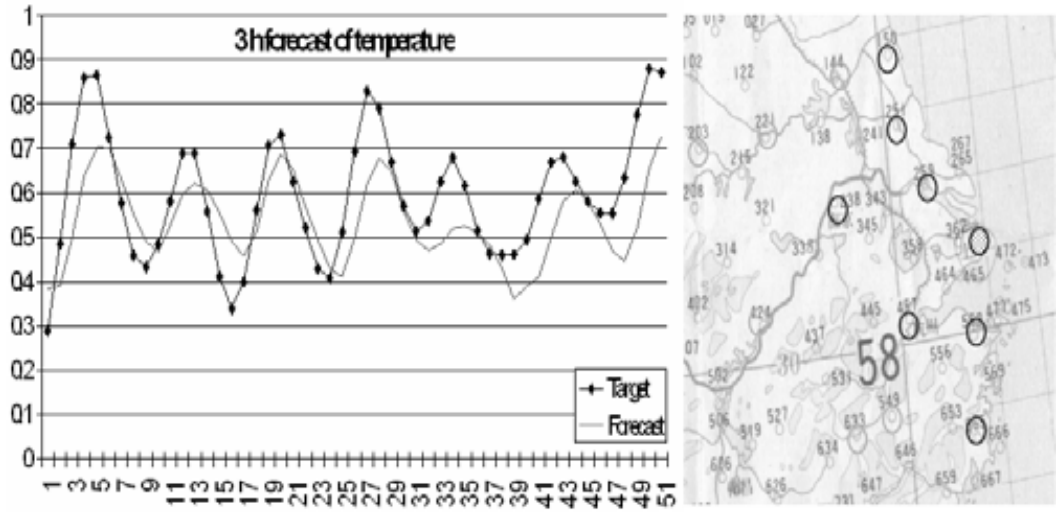


Figure 5.10 Temperature forecasting using 8 stations (58362-58457-58259-58562-58251-58150-58238-58665) with generalization error: 0.8223

5.5 Conclusion

Results have shown that neural networks with as few as a single hidden layer and an arbitrary bounded and non-constant activation function can approximate weather

forecast prediction for rainfall and temperature. The next step aims to show that it can also predict wind direction.

Chapter 6

EXTENDING VERSION OF GRAPHICAL USER INTERFACE IN NEURAL NETWORK TOOLBOX OF MATLAB AND ENGINEERING APPLICATIONS

6.1 Introduction

Artificial neural network is a model abstracted from biological neural networks. In biological neural network, when a neuron receives an excitatory input which is large enough compared with its inhibitory input, it will sends a spike of electrical activity down its axon. Learning occurs by changing the effectiveness of the synapses so that the influence of one neuron on another changes [46].

Compared with other models, the main advantage of artificial neural network is its ability to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. Also, artificial neural network has the ability of self-organization, meaning that the system is intelligent.

Due to the capabilities of artificial neural network, the famous calculating software, MATLAB, made a toolbox especially for artificial neural network. This toolbox gathers lots of functions for basic artificial neural networks. Also MATLAB has a Graphical User Interface of neural network for the introduction of neural network, which can set at most 2 layers for feed-forward network.

In the light of the above, the main objective of this chapter is to introduce a new Graphical User Interface for feed-forward network which can set as many layers as the user wants. Based on this program, 2 applications are applied to introduce the way for designing artificial neural network.

6.2 Extending Graphical User Interface

The graphical user interface is designed to be simple and user friendly. Users can type the data in the relative boxes with easy-to-understand titles instead of a series of codes. In the Neural Network Toolbox of MATLAB 7.1, there is a graphical user interface, 'nntool'. But this interface can only design a feedforward network with 2 layers at most. Here a new interface, 'platform', is introduced in Figure 6.1:

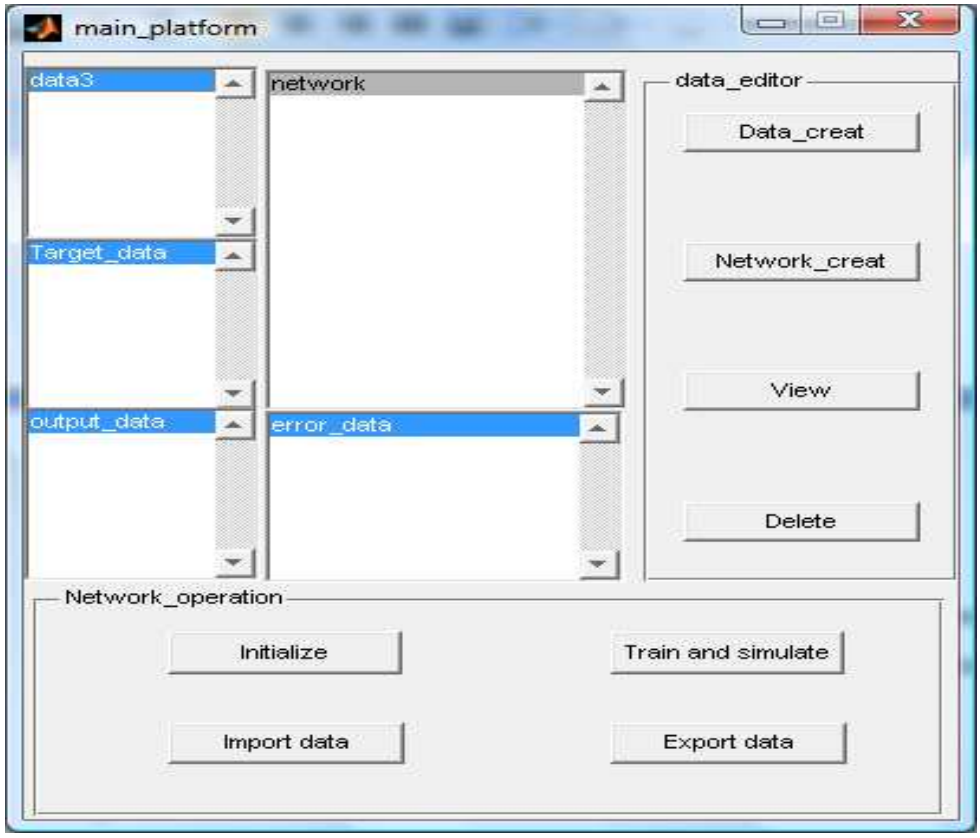


Figure 6.1 The main menu of platform

On this main platform, the user can operate by selecting the buttons:

- 'Data_creat' or 'Network_creat': Create data or neural network in this platform. The created data or network will appear in the corresponding list box.

- 'Initialize': Button only for neural network. Initialize the weight and bias of neural network randomly.
- 'Train and simulate': Button only for neural network. Excite the submenu of network training and simulation.

When pressing 'Network_creat' for setting up a neural network, the submenu is shown in Figure 6.2:

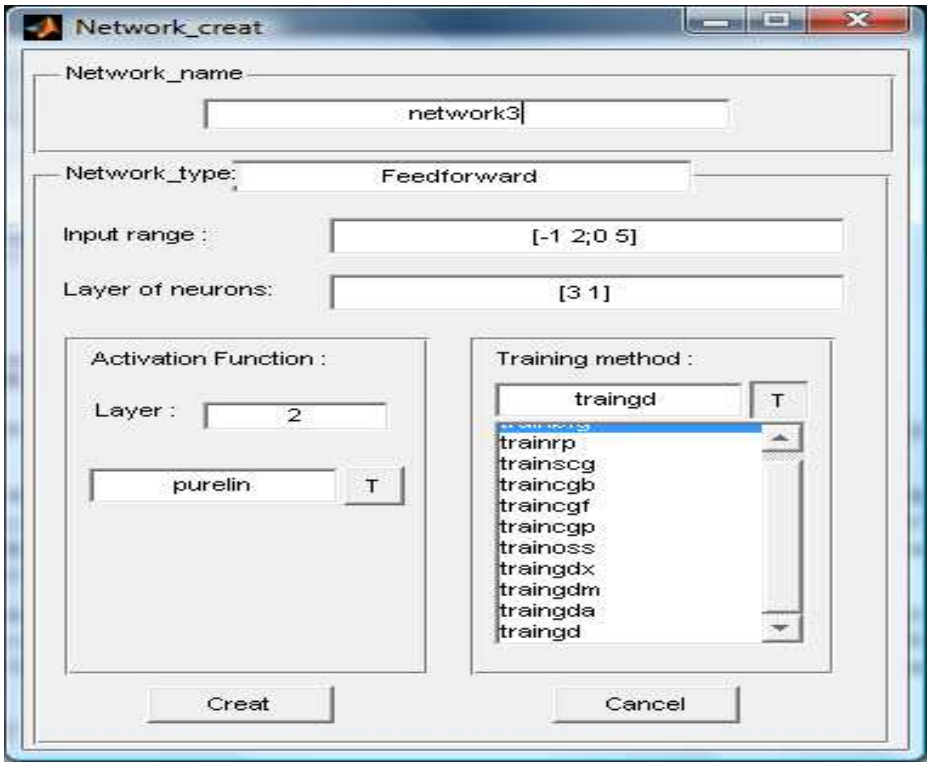


Figure 6.2 Submenu for network building

On this submenu in Figure 6.2, each edit box has its own function:

- 'Layer of neurons': Accepts a 1-by-m matrix in which 'm' is the number of layers. Each of the elements in this matrix stands for the number of neurons in the corresponding column.
- 'Activation function': listbox for selecting activation function in neurons.

From the submenu of network building, it can be seen that this platform can set up a feedforward network with as many layers as a user wants.

6.2.1 Application 1: Simple Voting System Design

This is a pattern recognition example. The voting system is for making a decision by a group of people on a majority-decision basis.

This system accepts the votes from people where 0 represents not agree and 1 represents agree. So the input is a vector whose members are all 0 and 1. The number of element is the number of people taking part in the voting. The output of this system is a number, 1 for pass and 0 for failure; as shown in Figure 6.3. The winning law is decided by users. A voting system for 5 people deciding the law: a majority win is selected in this example.

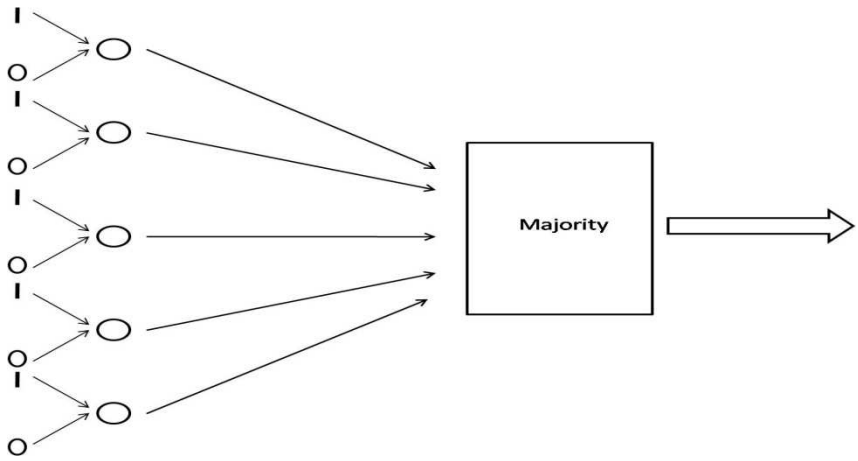


Figure 6.3 Introduction for voting system

In this example feedforward network is chosen. Feedforward neural network is one of the basic neural network models. Usually this kind of networks is combined by one input layer, one output layer and several hidden layers which are all constructed by

neurons. Layers are linked by weights which can be revised. Including input, each neuron also has a bias. Each neuron works as a biological neuron.

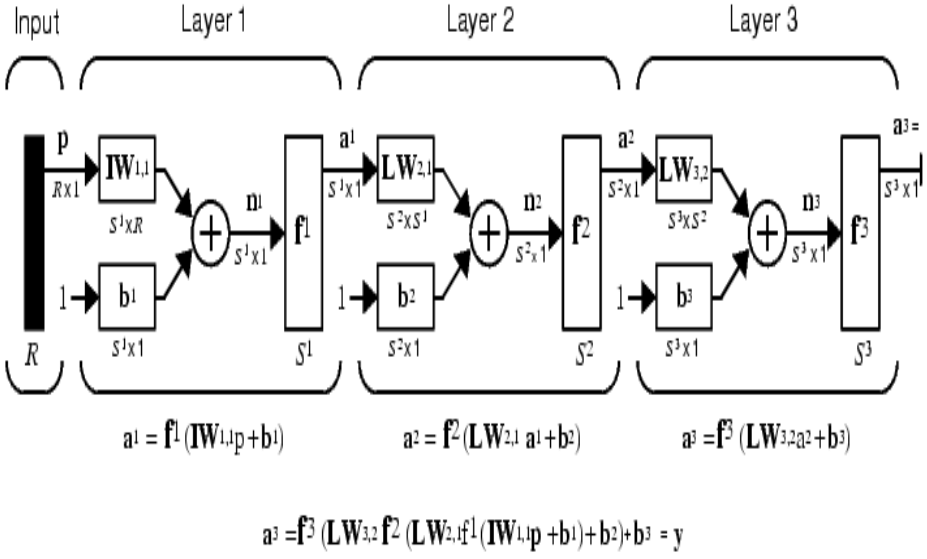


Figure 6.4.Mathematic model of feed-forward network [47]

For training the network, the training input should be a matrix containing all the possible situations, which is shown in Figure 6.5

```
Input =  0 1 0 0 0 0 1 1 1 1 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 1 1 1 1 0 1;
         0 0 1 0 0 0 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 1 1 1 0 1 1;
         0 0 0 1 0 0 0 1 0 0 1 0 0 1 1 0 1 0 0 1 1 0 1 1 0 1 1 1 0 1 1 1;
         0 0 0 0 1 0 0 0 1 0 0 1 0 1 0 1 0 1 0 1 0 1 1 0 1 1 1 0 1 1 1 1;
         0 0 0 0 0 1 0 0 0 1 0 0 1 0 1 1 0 0 1 0 1 1 0 1 1 1 0 1 1 1 1 1

Target = 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1;
```

Figure 6.5 Input and target for training voting system

This input matrix is a 5–by–32 matrix where each column stands for a kind of voting situation. The corresponding target is the result of the system.

For training this network, the first step is to set up a simple network (3 –1), which means there are 2 layers in this network and 3 neurons in the first layer and 1 in the second layer, to select the best training function. All the training functions should be trial and parameters should be changed for each training function so that they all perform best with the network (3 - 1). Table 6.1 shows the result.

Table 6.1 Training function selection

	3-1 ; 'tansig' for hidden layer and 'purelin' for output layer, 300 epochs, goal: 0.0001	
traingda	Best Learning rate: 0.3	Average Performance: MSE = 0.0095
	Momentum: 0.9	
traincgp	Best Learning rate: 2	Average Performance: MSE = goal met: 46 epochs
	Momentum: 0.9	
traincgb	Best Learning rate: 0.2	Average Performance: MSE = goal met: 66 epochs
	Momentum: 0.9	
trainlm	Best Learning rate: 0.15	Average Performance: MSE = goal met: 18 epochs
	Momentum: 0.9	

From Table 6.1, it is obvious that 'trainlm' is the best training method as the learning rate is the smallest.

After training function selection, the complexity of architecture should be increased so it can obtain a more accurate level. In this example there has been selected a new network (3 – 10 - 1) for training and the goal rises to 1 e-006. The training convergence result is shown in Table 6.2.

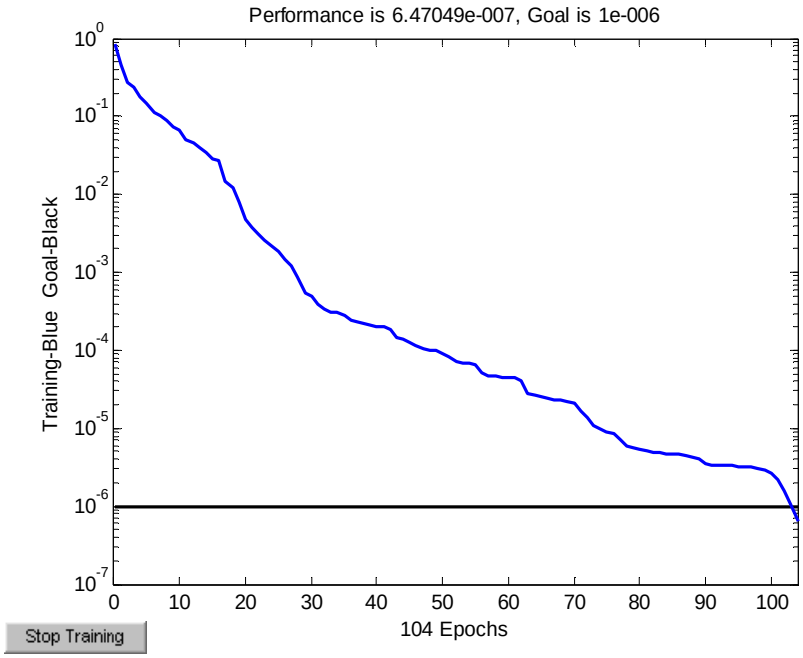


Figure 6.6 Training process of (3-10-1) network

Table 6.2 Result comparing between the two networks

Training Goal	0.0001	1 e-006
Network	(3 - 1)	(3 – 10 - 1)
MSE	9.5278 e-005	6.4705 e-007
Largest error	0.0217	0.0024
CPU time	0.0156	0.0156
Epochs for train	18	104

After training, the network should be tested. In this example the ability of noise resistance is being tested. Select an input + noise, as shown in Table 6.3.

Table 6.3 Code for producing input + noise

```
ip                % input for training
u = rand(size(ip));
u1 = rand(size(ip));
y = ip + n*(u - u1);    % y is the input + noise
```

In the code above, n is a variable for adjusting the amplitude of noise. As n increases, the noise increases and the system begins to have more and more errors.

Table 6.4 Performance of 2 networks when supplying noise

	Network(3-1) trained in goal:0.0001			Network(3-10-1) trained in goal: 1 e-006		
n	MSE	Largest error	Rate of mistake	Mse	Largest error	Rate of mistake
0.1	0.0539	0.6863	9.38%	0.0018	0.0626	0%
0.2	0.23	1.5357	15.63%	0.0174	0.3995	0%
0.3	0.3571	1.6780	25%	0.440	0.8963	3.12%
0.4	0.3320	1.4832	31.25%	0.0663	0.9821	6.25%
0.5	0.2864	1.2057	34.38%	0.0915	0.9984	12.5%

Result from Table 6.4 reveals that the network trained with a better performance has a better ability of noise resistance.

6.2.2 Application 2: 3–phase generator output detector

A 3-phase generator is a source of electricity. Its output is supposed to be 3 sine waves that the phase difference between each 2 sine waves is 120^{o}. In this example a generator is provided without knowing its output. The detector receives the output and it should then be able to tell the users whether the generator is working properly or not.

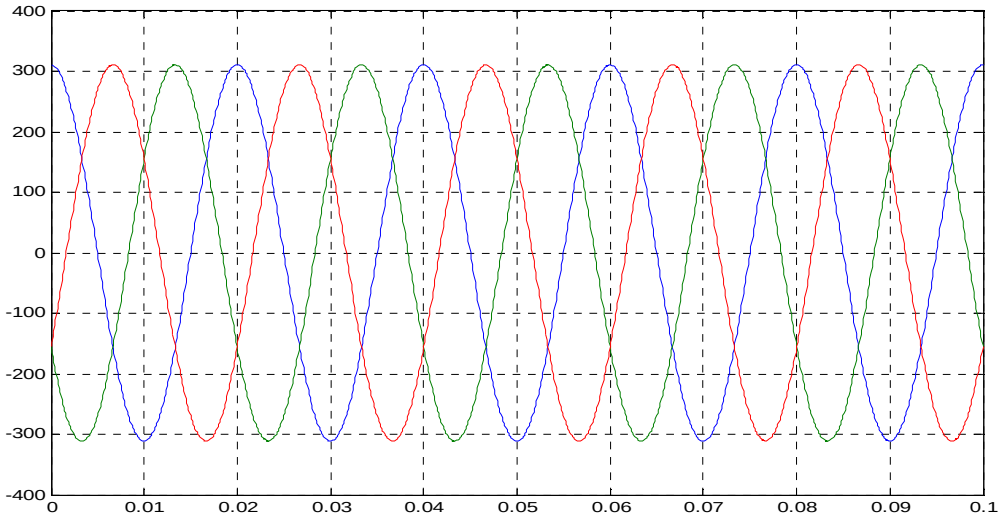


Figure 6.7 The input for the generator detector

Received the signal shown in Figure 6.7, detector should determine:

- What the working point of peak value is and whether the peak values are reasonable
- What the working point of frequency is and whether the frequency is reasonable now
- What the working point of phase difference is and whether the phase difference is reasonable now
- Calculate the RMS of the signal.

Considering the structure of data, each element, like peak value, frequency and so on, has its own value and varying law. So if feedforward neural network is used, each element should be set up with a network of their own. So it increases the complexity of system structure. Due to the disadvantage mentioned above, Radial Basis Network (RBF) is used.

A neuron of RBF network used in this example is shown in Figure 6.8.

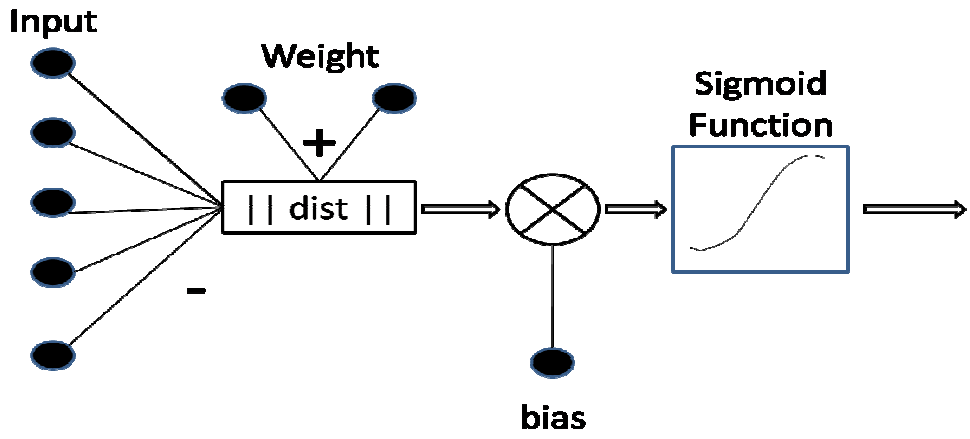


Figure 6.8 A neuron of RBF network

As shown in Figure 6.8, the || dist || box accepts input vector and the single row weight matrix, and then the output of this box is the dot product of the two. The weight is the target that needs to be trained.

If input vector is p, weight vector is w, bias is b and output of this neuron is a, the following equation could reveal the relation between these variables.

$$a = sigmoid(\| w - p \| \times b) \tag{6.1}$$

Based on the data structure, this RBF Network should have a path for peak value, a path for frequency, and a path for phase differences. The main structure is shown in Figure 6.9.

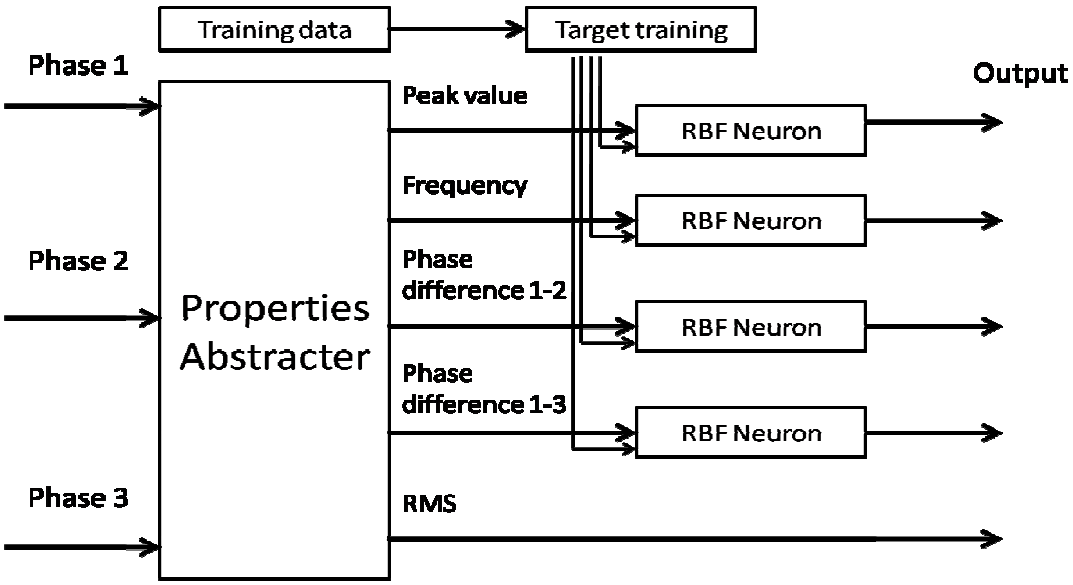


Figure 6.9 Main structure of the generator detector

The output of this detector has 5 elements. Each of them is a vector with 3 elements representing the corresponding phase signal whose value ranges are between -1 to 1, the degree of suitability, with 0 standing for no difference between target and input while 1 stands for much larger than target value; -1 means much smaller than target. Phase difference 1 – 2 means the recognized signal in phase 1 is the standard signal and the phase difference between signal in phase 2 is calculated. A similar reason is applied for phase difference 1 – 3.

When a detector first receives a signal from a generator, the operating point of the generator is unknown. So the first task is to find out the operating point and target of the generator.

The detector collects data from 3 – phase signal, then, from these data, the detector learns the target by Kohonem Learning Rule.

$$IW(q) = IW(q-1) + \alpha(p(q) - IW(q-1)) \qquad (6.2)$$

Equation 6.2 is the basic theory of Kohonem Learning Rule. In this equation, p is input, α is learning rate, IW is the weight or target.

Suppose the data collected from previous input is as in Table 6.5.

Table 6.5 Data collected from previous input

	Peak value (V)	Frequency (Hz)	Phase difference 1-2	Phase difference 1-3
1	309.0718	49.3306	119.4143	239.5816
2	314.4448	51.1216	120.9814	240.7010
3	312.9698	50.6299	120.5512	240.3937
4	420.0000	50.5044	120.4414	240.3153
5	306.7898	48.5699	118.7487	239.1062
6	308.9417	49.2872	119.3763	239.5545
7	313.9540	50.9580	120.8383	240.5988
8	316.4704	51.7968	121.5722	241.1230
9	308.6132	49.1777	119.2805	239.4861
10	313.7737	50.8979	120.7857	240.5612

Figure 6.10 is the training process of peak value with frequency shown in Figure 6.11; Phase differences 1 – 3 in Figure 6.12.

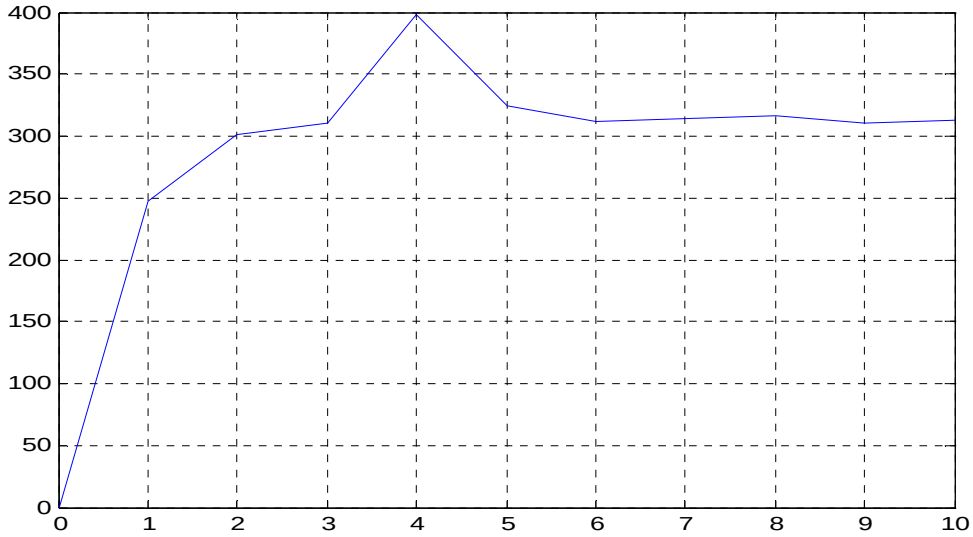


Figure 6.10 Training process of peak value

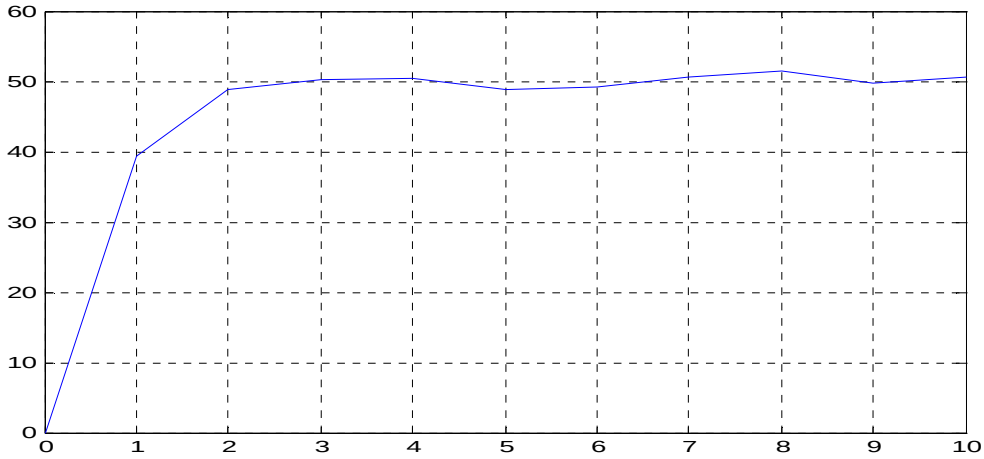


Figure 6.11 Training process of frequency

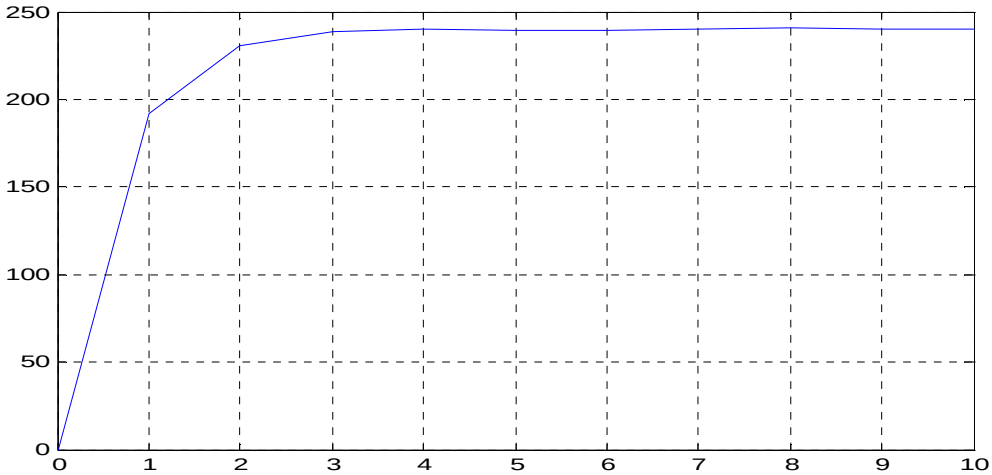


Figure 6.12 Training process of phase difference 1-3

In the peak value training as in Figure 6.10, the operating point should be near 310V. But one data at 4 is up to 420V which cause the target raise to 396V. Due to the intelligence of the network, the system decreases the target by learning more data so the target falls back to near 310V. From this example, it can be seen that RBF network has intelligent capability. The training result is shown in Table 6.6.

Table 6.6 Training result of detector

Learning rate: 0.8	Peak value (V)	Frequency (Hz)	Phase difference 1-2	Phase difference 1-3
Temporary target	313.03	50.65	120.57	240.41

After target selection to avoid saturation, data should be normalized. Suppose the usual varying range for peak value is -450V~+450V, frequency -1000Hz~+1000Hz, phase difference 0~360. The normalizing value is shown in Table 6.7.

Table 6.7 Normalizing value for detector

Name	Normalizing value
Peak value	500V
Frequency	1000Hz
Phase difference 1 – 2	380
Phase difference 1 – 3	380

Until now the detector design is nearly finished. One remaining work is to test and fix the measurement. The signal in Table 6.8 is chosen:

Table 6.8 Code for testing signal

Code	ip1 = A*cos(2*pi*f*t);
	ip2 = A*cos(2*pi*f *t + co1*pi/180);
	ip3 = A*cos(2*pi*f *t + co2*pi/180);

In Table 6.8, ‘A’ is the magnitude of signal; ‘f’ is the frequency; ‘co1’ and ‘co2’ are the phase differences.

Because the detector output is the degree of differences between target and input, the results are shown based on difference. Table 6.9 is the testing result of peak value, while Table 6.10 is for frequency, Table 6.11 is for phase difference.

Table 6.9 Testing result for peak value

Diff (V)	Output on Phase 1	Output on Phase 2	Output on Phase 3	RMS (V)
0.01	1.0e-004 *0.2000	1.0e-004 *0.1697	1.0e-004 *0.1679	221.5021
0.1	1.0e-003 *0.2000	1.0e-003 *0.1630	1.0e-003 *0.1632	221.5658
500	0.7616	0.7616	0.7616	574.9378V
1000	0.9640	0.9640	0.9640	928.0789V

Table 6.10 Testing result for frequency

Diff (Hz)	Output on Phase 1	Output on Phase 2	Output on Phase 3
0.01	1.0e-003* -0.1449	1.0e-003* 0.1114	1.0e-003* 0.1114
0.1	1.0e-003 * 0.1114	1.0e-003* 0.1114	1.0e-003 * 0.1114
1000	0.7858	0.7858	0.7858
1500	0.9051	0.9051	0.9051

Table 6.11 Testing result for phase differences

Diff	Output of phase 1 and 2	Diff	Output of phase 1 and 3
0.01	0.0026	0.01	0.0026
100	0.2597	100	0.2596
180	0.4432	180	0.4391

Because -0.95~+0.95 is the usual operating range for function ‘tansig’, it can be seen from Table 6.9 that the largest difference of peak value between target and input can accurately be tested to near 1000V. Suppose that the smallest value can be shown to users is 1.0 e-003, from Table 6.9 the smallest difference of peak value which can be detected is 0.1V. After deciding the limit on peak value, the measurement fixing is waiting to be done. It can be seen that the smallest difference can be detected but it depends on how small the system it is in terms of giving value to users. The largest difference not only depends on how large the system can show but also the limit of

0.95 in this case. Designers should fix the limit on their own, as well as on frequency and phase differences.

6.3 Conclusion

The new graphical user interface extends the ability of the interface from Matlab and decreases the difficulties of setting up neural networks. Based on this interface, 2 applications are used. Simple voting system is based on feedforward network and the 3–phase generator output detector is designed on Radial Basis Network (RBF). These ones have demonstrated how to design the neural network in applications. Thus, the platform for neural network modelling applied to various problem domains has been developed.

Chapter 7

SWARM INTELLIGENCE FOR OPTIMAL REACTIVE POWER DISPATCH

7.1 Introduction

The purpose of optimal reactive power dispatch (ORPD) is to minimize the network real power loss and improve voltage profiles by regulating generator bus voltage, switching on/off static var compensators and changing transformer tap-settings. The reactive power dispatch has complex problems in a large-scale power system. Many methods based on linear programming and non-linear programming have been proposed to solve this problem [48-52]. These approaches are based on successive linearizations and the first and second differentiations of an objective function and its constraint equations. Such approaches quite often lead to a local minimum point and sometimes result in divergence. Some new methods based on artificial intelligence have recently been used in ORPD planning to solve local minimum problems and uncertainties [53-55]. An improved genetic algorithm (GA) method has been proposed [56] to solve ORPD problem. This thesis proposes an application of Swarm Intelligence (SI) [57-62] to solve ORPD. A standard IEEE 30-bus system has been employed to carry out the simulation study.

7.2 Problem Formulation

List of symbols

NE = set of branch numbers

N i = set of numbers of buses adjacent to bus *I* including bus *i*

NPQ = set of PQ- bus numbers

Ng = set of generator bus numbers

N = set of numbers of tap-setting transformer branches

NB = set of numbers of total buses

NB-1 = set of numbers of total buses, excluding slack bus press

Ploss = network real power loss

gk = conductance of branch k

Vi = voltage magnitude at bus *i*

ii θ = voltage angle difference between bus *i* and bus j

Qi = reactive power injected into network at bus *i*

Gij,Bij = mutual conductance and susceptance between bus *i* and bus j

Gii,Bii = self conductance and susceptance of bus *i*

Qgi = reactive power generation at bus *i*

Tk = tap-setting of transformer branch k

NV lim = set of numbers of buses of voltages outside the limits

NQg lim= set of numbers of buses of reactive power generations outside the limits.

The objective function of ORPD can therefore be expressed as follows:

$$\min P_{loss} = \sum_{\substack{k \in N_E \\ k=(i,j)}} g_k(V_i^2 + V_j^2 - 2V_iV_j\cos\theta_{ij})$$

$$s.t. \quad 0 = P_i - V_i \sum_{j \in N_i} V_j(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}) \qquad i \in N_{B-1}$$

$$s.t. \quad 0 = Q_i - V_i \sum_{j \in N_i} V_j(G_{ij}\sin\theta_{ij} + B_{ij}\cos\theta_{ij}) \qquad i \in N_{pQ}$$

$$Q_{gi}^{min} \le Q_{gi} \le Q_{gi}^{max} \qquad i \in N_g$$

$$T_k^{min} \leq T_k \leq T_k^{max} \qquad k \in N_T \tag{7.1}$$

$$V_i^{min} \leq V_i \leq V_i^{max} \qquad i \in N_B$$

where power flow equations are used as equality constraints; reactive power generation restrictions, transformer tap-setting restrictions and bus voltage restrictions are used as inequality constraints. The transformer tapsetting T and generator bus voltages V_g are control variables so they are self-restricted. The load bus voltage V_{load} and reactive power generations Q_i are state variables, which are restricted by adding them to the objective function as the quadratic penalty term to form a penalty function. Equation 7.1 is therefore changed to the following generalized objective function:

$$\min f_q = P_{loss} + \sum_{i \in N_{V\lim}} \lambda_{V_i} (V_i - V_i^{\lim})^2 + \sum_{i \in N_{Q_{gi}\lim}} \lambda_{Q_{gi}} (Q_{gi} - Q_{gi}^{\lim})^2$$

$$s.t. \quad 0 = P_i - V_i \sum_{j \in N_i} V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \qquad i \in N_{B-1}$$

$$0 = Q_i - V_i \sum_{j \in N_i} V_j (G_{ij} \sin\theta_{ij} + B_{ij} \cos\theta_{ij}) \qquad i \in N_{PQ} \tag{7.2}$$

Where λ_{Vi} and λ_{Qgi} are the penalty factors which can be increased in the optimization procedure; $V_i^{\lim}$ and $Q_{gi}^{\lim}$, are defined in the following equations:

$$V_i^{\lim} = \begin{cases} V_i^{\min} & if \quad V_i < V_i^{\min} \\ V_i^{\max} & if \quad V_i > V_i^{\max} \end{cases}$$

$$Q_{gi}^{\lim} = \begin{cases} Q_{gi}^{\min} & if \quad Q_{gi} < Q_{gi}^{\min} \\ Q_{gi}^{\max} & if \quad Q_{gi} > Q_{gi}^{\max} \end{cases} \tag{7.3}$$

It can be seen that the generalized objective function f_q is a non-linear and non-continuous function. Gradient-based conventional methods are not good enough to solve this problem.

7.3 Swarm Intelligence

Swarm Intelligence (SI) was introduced in 1995 [57-62]. The method has been developed through a simulation of simplified social models. The method is based on researches on swarms such as fish schooling and bird flocking.

According to the research results for bird flocking, birds are finding food by flocking (as opposed to individually). Flocking prompted the assumption that information is owned jointly.

The SI algorithm can start with a population of particles with random positions. In swarm intelligence, a single particle is a solution in the search space. All particles have fitness values, which are evaluated by the fitness function to be optimized, and velocities, which direct the flying of the particles.

Each individual knows its best value so far (pbest) and its position. Moreover, each individual knows the best value so far in the group (gbest) among pbests. Modification of an individual position is carried out by the position and velocity information.

Consider a swarm of p particles; with each particle's position representing a possible solution point in the design problem space. For each particle *m*, Kennedy and Eberhart proposed that its position x^m be updated in the following manner:

$$x_{k+1}^m = x_k^m + v_{k+1}^m \tag{7.4}$$

With a velocity v_{k+1}^m calculated as follows:

$$v_{k+1}^m = w_k v_k^m + c_1 r_1 (p_k^m - x_k^m) + c_2 r_2 (p_k^g - x_k^m) \quad (7.5)$$

Here, subscript *k* indicates a time increment, p_k^m represents the best ever position of particle *m* at time *k*, and p_k^g represents the global best position in the swarm at time *k*. r_1 and r_2 represent uniform random numbers between 0 and 1. To allow the product $c_1 r_1$ or $c_2 r_2$ to have a mean of 1, c_1, c_2 are constant values typically in the range of 2 to 4. Kennedy and Eberhart proposed that the cognitive and social scaling parameters c_1 and c_2 be selected such that $c_1 = c_2 = 2$. These two rates control the relative influence of the memory of the neighbourhood the individual. The factor *w* is the inertia weight. If the inertia weight is large, the search becomes more global, while for smaller inertia the search becomes more local. The coefficients c_1 and c_2 are learning factors, which help particles to accelerate towards better areas of the solution space.

The best ever fitness value of a particle at design co-ordinates p_k^m is denoted by f_{best}^m and the best ever fitness value of the overall swarm at co-ordinates p_k^g by f_{best}^g. The algorithm flow can be described as follows:

(a) Set constants $c1$, $c2$, and k.

(b) Initialize dynamic maximum velocity v_k^{max} and weight w_k.

(c) Randomly initialize particle positions.

(d) Randomly initialize particle velocities.

(e) Evaluate fitness values f_0^m using design space co-ordinates x_0^m for *m*=1,…, *p*.

(f) Set $f_{best}^m = f_0^m$, $p^m = x$ for *m*=1,…, *p*.

(g) Update particle velocity vector v_{k+1}^m

(h) Update particle position vector x_{k+1}^m

(i) Update maximum velocity v_k^{max} and weight w_k.

(j) Evaluate fitness values f_k^m

(k) If stopping condition is satisfied then terminate

(l) Increment *m*, increment time, go to (g)

The steps used within the proposed algorithm for ORPD application are as follows:

Step 1 Input data for network configuration, load value and range of variables. Control variables include generator voltage, transformer tap and shunt capacitor. Load-bus voltage and reactive power generation are dependent variable.

Step 2 Initialize randomly the speed and position of each particle.

Step 3 Calculate the load flow and system loss.

Step 4 Calculate particle's coordinates according to the values of its objective function.

Step 5 Calculate the speed and its new position of each particle using Equations 7.4 and Equation 7.5.

Weight is computed according to the following equation:

$$w = w_{max} - \frac{w_{max} - w_{min}}{iter_{max}} \times iter \tag{7.6}$$

w_{max} and w_{min}: Maximum and minimum values of weights respectively.

$iter_{max}$: Maximum iteration.

$iter$: Current iteration number.

7.4 Comparison Methods

Two comparison methods have been considered.

7.4.1 Improved Genetic Algorithm (IGA)

GAs are search algorithms based on the mechanics of natural genetics and natural

selection. The GA used here is the binary-coded one. Each control variable is encoded into a series of binary bits. Each bit simulates a gene. The series of binary bits of all control variables compose a string, which simulates a chromosome. GA is a population search method. A population of strings is kept in each generation. The next generation is produced by the simulation of natural process of reproduction, gene crossover and mutation. To make GA practicable in the real-life large- scale systems, an improved GA (IGA) is needed.

In general, GAs mutation probability is fixed throughout the whole search processing. However, in practical applications, a small fixed mutation probability can only result in premature, while the search with a large fixed mutation probability will not converge. An adaptive mutation probability has been proposed [56].

Regarding crossover, each control variable is coded in a short fixed bit-length sub string. A dynamical hierarchy coding system was developed to code the large number of control variables in a real-life system with a reasonable length string without losing the resolution of the result. The search process will be divided into several stages with different change-steps for control variables.

7.4.2 Broyden's Method

Turning to Broyden's method, it is popularly applied in practice [63]. The only disadvantage of Broyden's method is that the inverse of full Hessian matrix has to be computed. However, in modem computer, this is not a difficult problem. In solving reactive power optimization problems, more precise methods have to be used than those used in real power dispatch problems. Broyden's method is briefly described in the following iterative equations to minimize the unconstrained function *f(x):*

$$x_{k+1} = x_k - B_K^{-1} \nabla f(x_k)$$

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)s_k}{s_k^T s_k}, k = 0, 1, \ldots$$

$$\nabla f(x_k) = \frac{\delta f}{\delta x} x_k$$

$$y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$
$$s_k = x_{k+1} - x_k \qquad (7.7)$$

where B_0 is a unit matrix.

7.5 Numerical Results

In this section, IEEE 30-bus system has been used to show the effectiveness of the algorithm. The network parameters of the system are given in [64]. The network consists of 6 generator-buses, 21 load-buses and 43 branches, of which 4 branches, (6, 9), (6, 10), (4, 12) and (28, 27), are under- load-tap-setting transformer branches. Two cases have been studied. Case 1 is of light loads whose loads and the active power generations of P-V buses are the same as in [64]. Case 2 is of heavy loads whose load and the active power generations of P-V buses are twice as those of Case 1. The reactive power generation limits are listed in Table 7.1. Voltage end tap-setting limits are listed in Table 7.2. All power and voltage quantities are per-unit values.

Initial Condition

All initial generator voltages and transformer taps are set to 1.0. The loads are given as:

Case 1: P_{load}=2.834, Q_{load}=1.262

Case 2: P_{load}=5.668, Q_{load}=2.524

The initial generations and power losses are obtained as in Table 7.3. The variables outside their limits are listed in Tables 7.4 and 7.5. In Case 2, because of heavy loads, almost all bus voltages and all reactive power generations are outside their limits.

Table 7.1 Reactive power generation limits

Bus	1	2	5	8	11	13
Case 1	0.596	0.48	0.6	0.53	0.15	0.155
Case 2	1.192	0.96	1.2	1.06	0.30	0.31
	-0.298	-0.24	-0.3	-0.265	0.075	0.078

Unit: per unit

Table 7.2 Voltage and tap-setting limits

V_g^{max}	V_g^{min}	V_{load}^{max}	V_{load}^{min}	T_k^{max}	T_k^{min}
1.1	0.9	1.05	0.95	1.05	0.95

Unit: per unit

Table 7.3 Generations and power losses

	P_g	Q_g	P_{loss}	Q_{loss}
Case 1	2.89388	0.98020	0.05988	-0.28180
Case 2	5.94588	3.26368	0.27788	0.73968

Unit: per unit

Table 7.4 Voltages outside limits

Case 1							
Bus	26	29	30				
	0.93	0.94	0.93				
Case 2							
Bus	9	10	12	14	15	16	17
	0.94	0.91	0.94	0.90	0.89	0.91	0.90
Bus	18	19	20	21	22	23	24
	0.87	0.86	0.873	0.87	0.88	0.87	0.85
Bus	25	26	27	29	30		
	0.86	0.81	0.88	0.83	0.80		

Unit: per unit

Table 7.5 Reactive power generations outside limits

Bus	1	8	13
Case1		0.596	
Case2	-0.402	1.497	0.439

Unit: per unit

7.6 Optimal Results and Comparison

The number of individuals is set to 50. 100 trials are performed for other simulations using different random numbers. At each trial, the best-evaluated value is stored within 100 searching iteration. w is set to 1.0 and c_1 and c_2 are 2.

The optimal generator bus voltage, transformer tap-settings, generations and power losses are obtained as in Tables 7.6, 7.7 and 7.8. All the state variables are regulated back into their limits.

Table 7.6 Generator bus voltages

Bus		1	2	5	8	11	13
Case 1	IGA	1.07	1.06	1.04	1.04	1.08	1.06
	Brydn	1.07	1.06	1.04	1.02	1.09	1.07
	PSO	1.07	1.06	1.04	1.04	1.08	1.06
Case 2	IGA	1.10	1.09	1.05	1.05	1.10	1.09
	Brydn	1.07	1.06	1.04	1.02	1.09	1.07
	PSO	1.10	1.09	1.06	1.06	1.10	1.10

Unit: per unit

Table 7.7 Transformer Tap-settings

Branch		(6.9)	(6.10)	(4.12)	(28,27)
Case 1	IGA	0.98	1.03	1.02	1.04
	Brydn	0.96	1.00	1.00	1.02
	PSO	0.98	1.03	1.03	1.04
Case 2	IGA	1.00	1.10	1.04	1.10
	Brydn	1.07	1.02	1.07	1.10
	PSO	1.00	1.10	1.05	1.10

Unit: per unit

Table 7.8 Generations and power losses

		P_g	Q_g	P_{loss}	Q_{loss}
Case 1	IGA	2.88357	0.87389	0.04957	-0.38811
	Brydn	2.88798	0.92305	0.05398	-0.33895
	PSO	2.88408	0.87588	0.05001	-0.38618
Case 2	IGA	5.87926	2.86863	0.21126	0.34463
	Brydn	5.89406	2.93719	0.22607	0.41320
	PSO	5.87514	2.85933	0.20945	0.34424

Unit: per unit

The power savings are given as follows:

Case 1

IGA:

$$P_{save}\% = \frac{P_{loss}^{init} - P_{loss}^{opt}}{P_{loss}^{init}} \times 100 = \frac{0.05988 - 0.04957}{0.05988} \times 100 = 17.22\%$$

Broyden:

$$P_{save}\% = \frac{0.05988 - 0.05398}{0.05988} \times 100 = 9.85\%$$

PSO:

$$P_{save}\% = \frac{0.05988 - 0.05001}{0.05988} \times 100 = 16.48\%$$

Case 2

IGA:

$$P_{save}\% = \frac{0.27788 - 0.21126}{0.27788} \times 100 = 23.97\%$$

Broyden:

$$P_{save}\% = \frac{0.27788 - 0.22607}{0.27788} \times 100 = 18.64\%$$

PSO:

$$P_{save}\% = \frac{0.27788 - 0.20945}{0.27788} \times 100 = 24.63\%$$

7.7 Conclusion

The potential for application of Swarm Intelligence to ORPD has been shown in this paper. It can be seen that in the three cases, Swarm Intelligence produced similar results to those from IGA and the results from Broyden's method were the worst. It

seems that SI could be a good technique to solve real life power system problems in the near future.

Chapter 8

ADAPTIVE RBFN MODEL FOR 2D SPATIAL INTERPOLATION

8.1 Introduction

The GIS (geographic information system) is a system for capturing, storing, analyzing and managing data and associated attributes which are spatially referenced to the Earth. According to research on applications and theories of intelligent weather forecast, generally, large volumes of complicated and precise computation are required in many related applications, e.g. meteorological data training for neural networks, tracking process of typhoon eye, nephogram decomposition and details fusion, etc [65-67]. Geographical information systems could be improved by adding procedures for geography statistical spatial analysis to existing facilities.

Spatial interpolation is a major component in many GIS-based analyses. Often data are obtained in one form and need to conduct analyses with surfaces or other point data sets. Creating surfaces from points is a very complex process-both from a concept point-of-view (knowing which method is most appropriate) and from a process point-of-view (the mathematics can be complex). As such, spending some time learning about the intricacies of spatial interpolation will prove beneficial to all GIS analysts. Most traditional methods of interpolation are based on mathematical as distinct from stochastic models of spatial variation. Spatially distributed data behave more like random variables; fortunately regionalized variable theory provides a set of stochastic methods for analyzing them.

Some efforts in this direction already exist, such as Voronoi Diagram - sometimes also known as a dirichlet tessellation [68, 69]; TIN (Triangulated Irregular Network) which is a form of the tessellation model based on triangles [70]; IDW (inverse distance

weighted) which is one of the most commonly used techniques for interpolation of scatter points [71]; Kriging which is a group of geo-statistical techniques to interpolate the value of a random field at an unobserved location from observations of its value at nearby locations [72, 73]; Topogrid, and RST (Regularized Splines with Tension), generally, these two applications are available for interpolating areas: change of the resolution of raster data (re-sampling) and filling up of incomplete data (interpolation). Splines-Interpolation (RST) is a major interpolation module for the second mentioned application [74, 75]. Proposed methodologies will start from radial basis function network, which has high performance advantages, for real-time 2D spatial interpolation and similar 2D information analysis; it helps to increase the processing efficiency to a certain extent.

8.2 Radial Basis Function Networks

8.2.1 Radial Basis Function

A radial basis function (RBF) is a real-valued function whose value depends only on the distance from the origin, so that $\phi(x) = \phi(\|x\|)$; or alternatively on the distance from some other point c, called a center, so that $\phi(x, c) = \phi(\|x - c\|)$. Any function ϕ that satisfies the property $\phi(x)=\phi(\|x\|)$ is a radial function. The norm is usually Euclidean distance. Radial basis functions are typically used to build up function approximations of the form

$$y(x) = \sum_{i=1}^{N} w_i \left(\|x - c_i\|\right) \tag{8.1}$$

where the approximating function *y(x)* is represented as a sum of *N* radial basis functions, each associated with a different center c_i, and weighted by an appropriate

coefficient w_i. Approximation schemes of this kind have been used particularly in time series prediction and control of nonlinear systems exhibiting sufficiently simple chaotic behavior.

The sum can also be interpreted as a rather simple single-layer type of artificial neural network called a radial basis function network, with the radial basis functions taking on the role of the activation functions of the network. It can be shown that any continuous function on a compact interval can in principle be interpolated with arbitrary accuracy by a sum of this form, if a sufficiently large number (N) of radial basis functions are used.

The approximant $y(x)$ is differentiable with respect to the weights w_i. The weights could thus be learned using any of the standard iterative methods for neural networks. But such iterative schemes are not in fact necessary: because the approximating function is linear in the weights w_i, the w_i can simply be estimated directly, using the matrix methods of linear least squares.

Construction of radial basis function neural networks (RBFN) involves selection of radial basis function centroid, radius (width or scale), and number of radial basis function (RBF) units in the hidden layer. The K-means clustering algorithm is frequently used for selection of centroids and radii. However, with the K-means clustering algorithm, the number of RBF units is usually arbitrarily selected, which may lead to suboptimal performance of the neural network model. Besides, class membership and the related probability distribution are not considered.

8.2.2 RBFN Kernel Construction

Radial basis function (RBF) networks have a static Gaussian function as the nonlinearity for the hidden layer processing elements. The Gaussian function responds only to a small region of the input space where the Gaussian is centered. The key to a successful implementation of these networks is to find suitable centers for the Gaussian functions. This can be done with supervised learning, but an unsupervised approach usually produces better results [76, 77].

The simulation starts with the training of an unsupervised layer. Its function is to derive the Gaussian centers and the widths from the input data. These centers are encoded within the weights of the unsupervised layer using competitive learning, which is a form of unsupervised learning in artificial neural networks, in which nodes compete for the right to respond to a subset of the input data. During the unsupervised learning, the widths of the Gaussians are computed based on the centers of their neighbors. The output of this layer is derived from the input data weighted by a Gaussian mixture.

Once the unsupervised layer has completed its training, the supervised segment then sets the centers of Gaussian functions (based on the weights of the unsupervised layer) and determines the width (standard deviation) of each Gaussian. Any supervised topology (such as a MLP) may be used for the classification of the weighted input.

The advantage of the radial basis function network is that it finds the input to output map using local approximators. Usually the supervised segment is simply a linear combination of the approximators. Since linear combiners have few weights, these networks train extremely fast and require fewer training samples.

8.3 Adaptive RBFN Design

8.3.1 Factorized RBFN (F-RBFN)

The specific network architecture that will be investigated in this thesis is a hybrid between the Factorized Radial Basis Function Networks and the fuzzy/neural networks for implementing fuzzy controllers capable of learning from a reinforcement signal [78-83]. The activation function used in an F-RBFN with n input units is defined as the product of n one-dimensional radial functions, each one associated to one of the input features. Therefore an F-RBFN can be described as a network with two hidden layers. The neurons in the first hidden layer are feature detectors; each associated to a single one-dimensional activation function and connected to a single input only. If using Gaussian functions, the neuron r_{ij} (the *i-th* component of the *j-th* activation area) computes the output:

$$\mu_{ij} = e^{-(\frac{I_i - C_{ij}}{\sigma_{ij}})^2} \tag{8.2}$$

The neurons in the second hidden layer simply compute a product and construct multi-dimensional radial functions:

$$r_j = \prod_i \mu_{ij} \tag{8.3}$$

Finally, the output neuron contributes to the composite functions computed in the second hidden layer. In our architecture, a choice among two different activation functions is possible for the output neuron. The output function, normally adopted for RBFNs, is a weighted sum

$$Y = \sum_j w_j r_j \tag{8.4}$$

The usual methods for learning from examples Radial Basis Function Networks are based on a two step learning procedure. Firstly a statistical clustering algorithm, such as K-Means is used to determine the centers and the amplitude of the activation functions. Then the weights on the links to the output neuron are determined by computing the coefficients of the pseudo-inverse matrix. The method is usually faster than gradient descent. On the other hand, the second is preferable, because it is more simple to implement and suitable for real-time learning.

As the adopted activation functions are continuous, which are derivable in the whole domain, it is to apply the classical error gradient descent technique in order to finely tune the weights and the parameters in the activation functions. More specifically, let *E* be the quadratic error evaluated on the learning set and let, moreover, λ_k indicate a generic parameter in the network, such as, a weight *w* on a link, the width σ or the center *c* of a Gauss activation function, or the threshold of the sigmoid in the output neuron. Then all the necessary derivatives can be computed and the learning rule can be described as:

$$\Delta\lambda_i = -\eta \frac{\partial E}{\partial \lambda_k} \tag{8.5}$$

Dynamic competitive learning algorithm is used to assist network construction, which consists of a combination of the gradient descent algorithm and improved growing cell structure algorithm. The main algorithm includes two parts: one for generating adaptive network parameters and another one for creating rules to add new neurons.

To improve the adaptability of center locating and structure setting up, it is necessary to build up a neighborhood topology. In the most known approaches, it is required to

choose both the dimensions and the structure type in advance. However, as data distribute with a high-dimension, to find out a topological structure closely that can reflect the data topology is not a trivial task. An effective way to learn such a structure is by using competitive Hebbian learning [84, 85]. The locating of the multi-dimensional Gaussian centers can be executed by aid of an algorithm similar to the self-organizing feature mapping networks.

Generally, there are two completely different cases from the learning point of view. In first cases, all errors have the same sign and then training the rule will decrease the error to a certain extent where the error is either zero or the conditions fall in the second case. In second cases the errors have different sign and training does not help because changing the variances has just a very limited influence on that problem and changing the output weight decreases the error of one type but increases the other one and vice versa. In order to have a direct measurement of the error difference, for each rule, two variables, the accumulated error E_s and the accumulated absolute error E_s^{abs}, are maintained. Once an input data is received, the variances of the best matching rule are updated according to:

$$\Delta E_S = (X - Z) \tag{8.6}$$

$$\Delta E_S^{abs} = |X - Z| \tag{8.7}$$

Notice that the best matching rule is the most activated one instead of the nearest one; what is more important is that for each activation factor of the network, the variables E_c and E_c^{abs} associated to all rules are decreased with the quantities:

$$\Delta E_C = -\alpha E_C \tag{8.8}$$

$$\Delta E_c{}^{abs} = -\alpha E_c{}^{abs} \tag{8.9}$$

In this way errors which occur far in the past are getting less weighted than errors that occurred recently in the past. This kind of averaging can also be seen as a filter so that noisy inputs are getting averaged and could be avoided to create new rules directly because of disturbances. Now the accumulated error variance σE_c at any time for any rule is computed in the following way.

$$\sigma E_c = \left\| E_c \right| - \left| E_c{}^{abs} \right\| \tag{8.10}$$

It is an indirect measurement for the variance of the error. If the errors share the same sign, i.e. σE_c becomes zero and can also get larger with increasing errors. Once a stable and robust insertion criterion is obtained, focus could be directed to inserting a new rule. That presents a problem: how to set the criterion of stopping the growing process, which is fairly easy in supervised learning tasks; however, for incremental real-time learning tasks, without precise error heuristic information is really a problem.
This dynamic competitive learning method uses a kind of incremental clustering logic, which is similar to that used in Knhonen's maps, in order to distribute the hidden unit over the domain of the target function. The advantage of it is that it does not need to remember the history; additionally, the incremental clustering algorithm is able to provide good coverage since the hidden units can move everywhere in the input space following the data.

8.3.2 Construction of RBF Networks with Responsivity Analysis

As known, an RBF classifier is a three-layer neural network model, in which an N-dimensional input vector $\vec{x} = (x_1, x_2, ...x_N)$ is broadcast to each of *K* neurons in the hidden layer. Each hidden neuron produces an activation function, typically a Gaussian kernel:

$$h_i = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma_i^2} \right), \qquad i = 1, 2, ..., K \tag{8.11}$$

where c_i and σ_i are the center and width of the Gaussian basis function of the *i-th* hidden unit, respectively. The units in the output layer have interconnections with all the hidden units. The *j-th* output neuron has the form:

$$f_j(x) = \vec{w}_j \vec{h} = \sum_{i=1}^{K} w_{ij} \exp\left(-\frac{\|x - c_i\|^2}{2\sigma_i^2} \right), \qquad i = 1, 2, ..., K \tag{8.12}$$

where $\vec{h} = (h_1, h_2, ...h_K)$ is the input vector from the hidden layer, and the w_{ij} is the interconnection weight between the *j-th* output neuron and the *i-th* hidden neuron.

In this section, an RBF classifier's responsivity is defined as the mathematical expectation of the square of output deviations caused by the oscillation of RBF centers. An algorithm as described below will be used to select kernel vectors. The symbols $\hat{c}_i$ and $\hat{\sigma}_i$ are used to denote the values of the center and width of the *i-th* hidden neuron with an oscillation. Then the deviation resulting from this oscillation is:

$$\Delta y_j = \vec{\hat{w}}_j \vec{\hat{h}} - \vec{w}_j \vec{h}$$
$$= \sum_{i=1}^{K} \hat{w}_{ij} \exp\left(-\frac{\left\| x - \hat{c}_i \right\|^2}{2\hat{\sigma}_i^2} \right) - \sum_{i=1}^{K} w_{ij} \exp\left(-\frac{\left\| x - c_i \right\|^2}{2\sigma_i^2} \right) \quad (8.13)$$

Here, $\hat{c}_i = c_i + \Delta c_i$ are the centers deviated from the centers under the oscillations, and the interconnection weights under the oscillations are $\hat{w}_j = w_j + \Delta w_j$, where w_j can be calculated using a pseudo matrix inversion.

Although there are various ways to specify RBF widths [86], the most common method for selecting RBF widths is to make all of them equal to a constant value depending on the prior knowledge of the given application. With pre-defined RBF widths, the focus is on just the oscillations on the centers and their interconnection weights. The oscillation on the *i-th* RBF center and the weights connected to the *j-th* output, Δc_i and Δw_j, can be generated following a Gaussian distribution with means 0 and variances $\sigma_{c_i}, \sigma_{w_j}$:

$$o(\Delta c_i) = \frac{1}{(\sqrt{2\pi\sigma_{c_i}})^N} \exp(-\frac{\Delta c_i^T \Delta c_i}{2\sigma_{c_i}^2})$$
$$o(\Delta w_j) = \frac{1}{(\sqrt{2\pi\sigma_{w_j}})^K} \exp(-\frac{\Delta w_j^T \Delta w_j}{2\sigma_{w_j}^2}) \quad (8.14)$$

where *N* is the dimension of the input $\vec{x}$, *K* is the number of RBF centers. The RBF centers will be selected recursively in the next subsection. To make the responsivity analysis a pre-requisite for the construction of RBF networks, a recursive definition of responsivity is given below. At the *K-th* time point, suppose there are a number (*K*–1) of RBF centers fixed already, the newcomer c_i is observed. Hence, the *j-th* output neuron's responsivity to the current number *K* of RBF centers is defined as the mathematical expectation of $(\Delta y_j)^2$ (square of output deviations caused by the

oscillations of RBF centers) with respect to all Δc_i and the training example set $D = \{x_l\}_{l=1}^{L}$.

Similar to the comparison between support-vector-based and clustering-based RBF networks [87], the difference between the responsivity-based and the conventional RBF networks can be described as: for data points of two classes responsivity-based RBF, centers are the vectors reactive to the classification, and located relatively close to the newly-introduced input items; while for conventional RBF, centers are usually located at the cluster centroids.

8.4 Experimental Results

As shown in Figure 8.1, a finite number of stations will be used in proximity to an area for forecasting, which is a traditional methodology adopted by the Bureau of Meteorology. After extensive data mining and analysis, 8 meteorological variables have been chosen as being suitable and reliable in terms of application, data consistency and validation. All the units of the data have been adjusted to international standard units, to maintain the consistency of the training data, and eliminate the complication and non-linear relationships between individual variables.

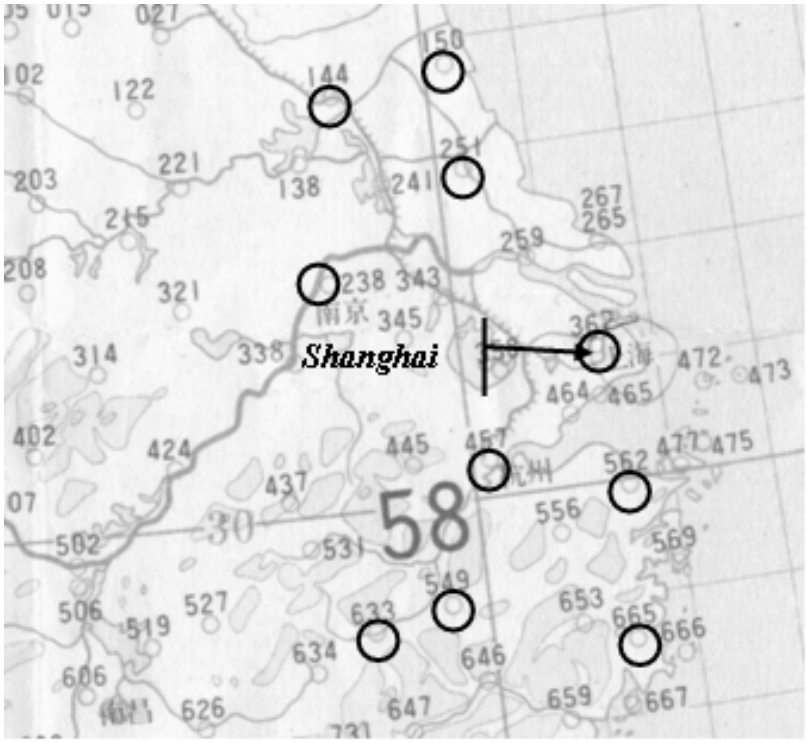


Figure 8.1 Map of stations that will be taken into RBFN training and results verification

Here is a section of filtered meteorological data, which were recorded from the Shanghai station (Site No.: 58-362); the sampling time duration was from 2:00 AM on 01/05/1999 until 14:00 PM on 02/05/1999; [65] the time gap is 3 hours, which is a general sampling frequency for meteorological data analysis. Incremental RBFN were used to evaluate the ratio of TCQ (total cloud quantity) to DP (dew point) at 10 separate stations.

With responsivity real-time evaluation based construction method, the classification centers are very sensitive to newly-introduced sample points (ratio of TCQ vs DP in this case), which is different from other general RBFN. As the weather forecast data is with time-delay attributes, the factorized RBF is specifically effective to this case, and the incremental RBFN is able to behave efficiently to the newly-added points, with a 3-hour time delay. Based on the constructed ratio RBFN, TCQ levels can be estimated with input DP degrees, to evaluate the effectiveness; the error rate will be verified

using the average deviation. Figure 8.2 shows calculated centers of the recorded ratios of 2 stations for 256 sampling points, which means the record duration is 1 month.

Figure 8.3 shows the results comparison between IR-RBFN (incremental responsivity based RBFN), V-Diag, IDW, and Topogrid. The size of training data is increased by 64 per each 192 hours. It should be noted that the initial data size is 64 and practically this represents a continuous meteorological data of 192 hours. Along with the increasing of the size of data set, IR-RBFN behaves better than others, since its responsivity is good at making decision on newly-added testing data items, and the incremental attributes help decrease the training time cost a lot, compared to with traditional RBFN.

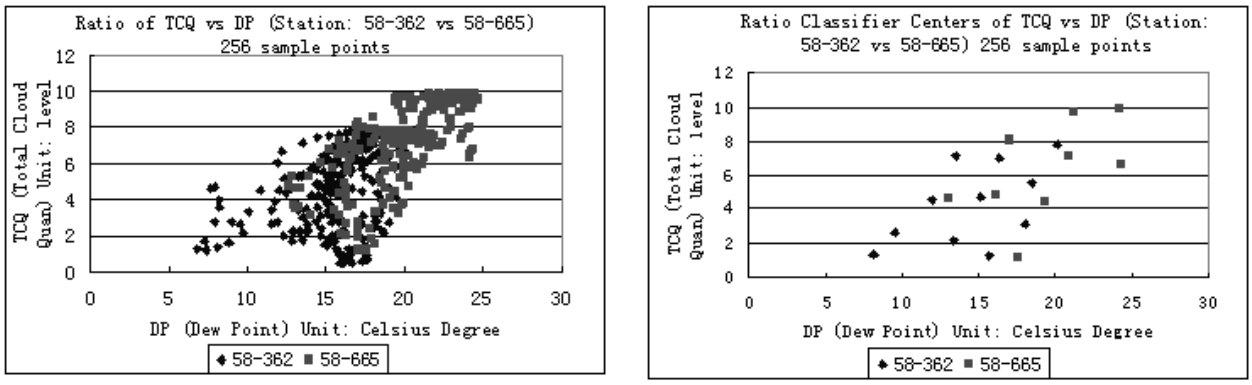
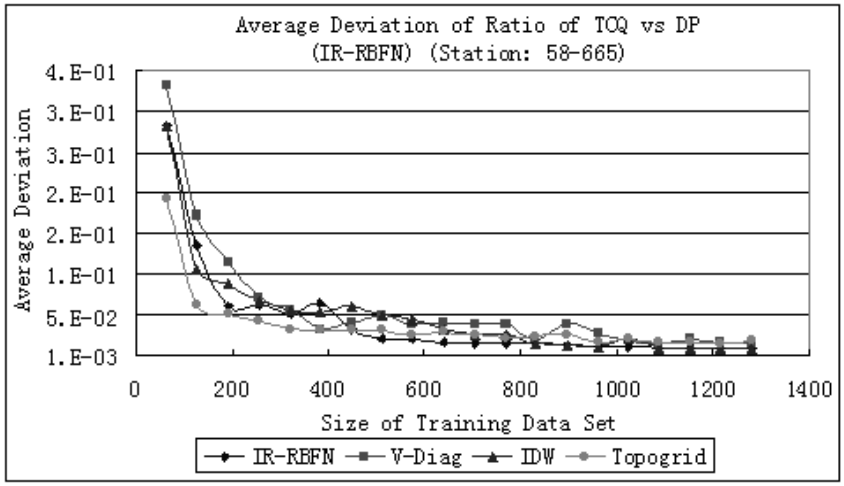


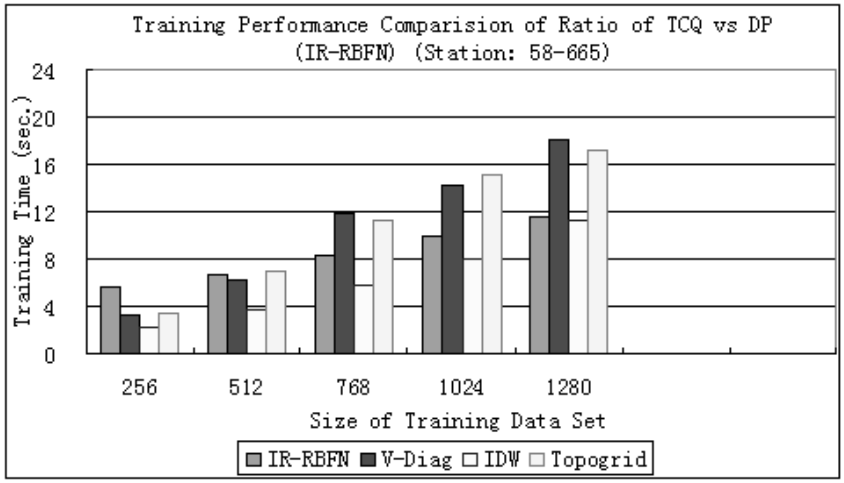
Figure 8.2(a)(b) Ratio centers demonstration of 32 and 256 sample points' cases, the classification centers of ratio are shown as well

According to the training performance comparison between algorithms, shown in Figure 8.3 (a), the incremental properties of IR-RBFN brings benefits in terms of time-shift data input set; although initially there is a relatively small training data set, the performance of the proposed method is not the best; however, with a small increase in the size of training data, it performs better than other methods. Regarding the CPU time usage, it is shown in Figure 8.3 (b) that the increase in CPU time is

much faster as training data size increase for the other three conventional methods; however, the new proposed method has the benefit that this CPU time is converged much more quickly.



(a)



(b)

Figure 8.3 (a) Average deviation comparison between 4 algorithms, to evaluate the precision of estimated ratio of TCQ vs DP; (b) The performance comparison results between 4 algorithms, to evaluate the precision of estimated ratio of TCQ vs DP

8.5 Conclusion

To improve the applicability of traditional RBFN, especially in 2D spatial interpolation, such as the time-delay meteorological data described in this chapter,

optimization in terms of incremental attributes of RBFN and a responsivity based construction method has been proposed. The benefit of utilizing factorized RBFN is its high performance when introducing new training data set, and the responsitivity accelerates the center close to the latest added data item. Practical data has been used to test the proposed new method and it has demonstrated that this integrated intelligent technique is very effective for real-time metrological information processing. The developed algorithm has been implemented in the Shanghai Meteorology Center and used as one of the tools for weather forecasting.

Chapter 9

PARTICLE SWARM OPTIMIZATION FOR ECONOMIC DISPATCH OF UNITS WITH NON-SMOOTH INPUT-OUTPUT CHARACTERISTIC FUNCTIONS

9.1 Introduction

Economic dispatch (ED) is an important daily optimization task in the power system operation. Various mathematical programming methods and optimization techniques have been applied to ED. Most of them are calculus-based optimization algorithms, which are based on successive linearizations and use the first and second differentiations of objective function and its constraint equations as the search directions. They usually require the heat input-electric power output characteristics of generators to be of monotonically increasing nature or of piece-wise linearity. However, large modem generating units with multi-valve steam turbines exhibit a large variation in the input-output characteristic functions. The valve-point effects, due to wire drawing when each steam admission valve is starting to open, will typically produce a ripple-like heat rate curve. The conventional optimization methods are not suitable to solve such a problem. Dynamic programming had been proposed to solve the ED problem with non-smooth cost curves of generators [88]. However, the dimensions of the problem will become extremely large with the increase of the variables. Simulated annealing (SA) [89], genetic algorithms (GAs) [90] and EP [91] based on Artificial Intelligence have been proposed to solve such non-smooth ED problems.

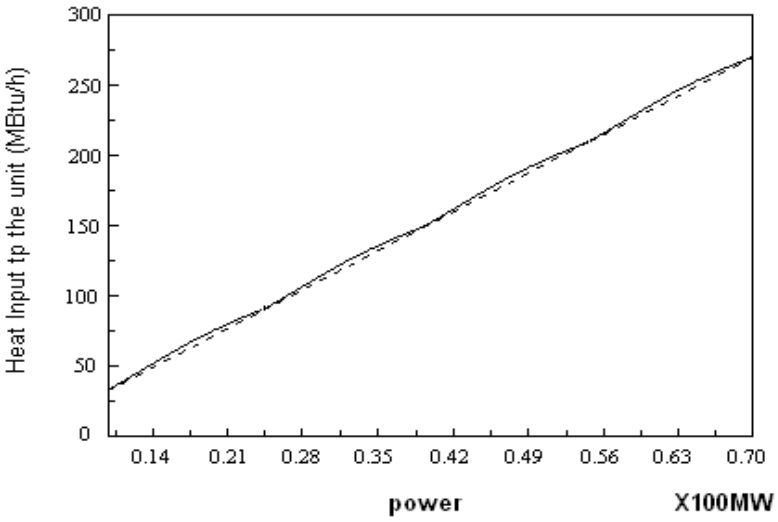
This thesis proposes an application of Particle Swarm Optimization (PSO) to solve ED of units with non-smooth input-output characteristic functions. The IEEE 30-bus

system has been used as the simulation system to show the effectiveness of the algorithm. The network consists of 6 generator units. The input-output curves of the generating units are represented by second order polynomial functions, superimposed with the valve-point effects. The curves are therefore non-monotonically increasing curves with multiple local minimum points.

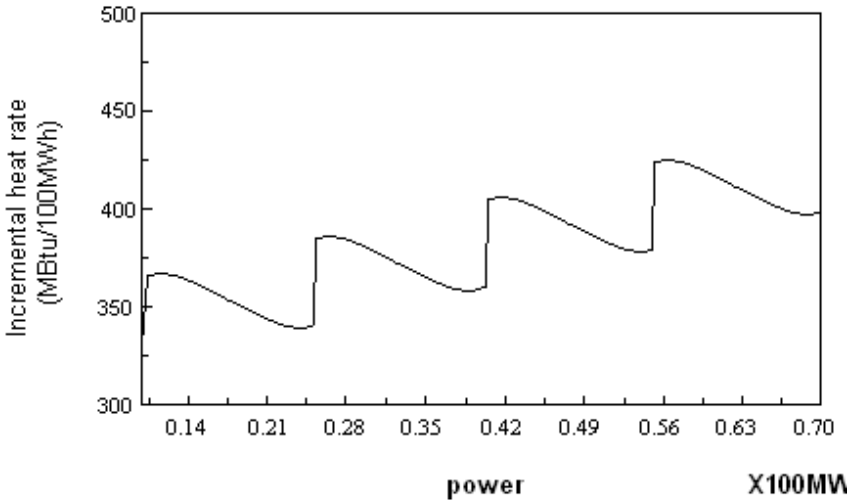
9.2 Problem Formulation

9.2.1 Valve Point Effect

Large steam turbine generators usually have a number of steam admission valves that are opened in sequence to obtain ever-increasing output from the unit. Figure 9.1 shows both input-output and incremental heat rate characteristic for a unit with four valves. As the unit loading increases, the input to the unit increases. However, when a valve is just opened, the throttling losses increase rapidly so that the input heat increases rapidly and the incremental heat rate rises suddenly. This gives rise to the non-smooth type of heat input and discontinuous type of incremental heat rate characteristics. This type of characteristic should be used in order to accurately schedule steam units, but it cannot be used in traditional optimization methods because it does not meet the convex condition.



(a)



(b)

Figure 9.1(a)(b) Characteristics of a steam turbine generator with four steam admission valves

9.2.2 Objective Function

To model the effects of valve-points, a recurring rectified sinusoid contribution is added to the second-order polynomial functions to represent the input-output equation as follows [90]:

$$f_p = \sum_{i \in N_g} (a_i + b_i P_{gi} + c_i P_{gi}^2 + \left| e_i \sin(f_i (P_i - P_i^{\min})) \right|) \tag{9.1}$$

The objective function of ED is to minimize the above non-smooth function by regulating active power outputs of generators, subject to all operational and secure constraints. The objective function of ED is therefore expressed as follows:

$$\min f_p = \sum_{i \in N_g} (a_i + b_i P_{gi} + c_i P_{gi}^2 + \left| e_i \sin(f_i (P_i - P_i^{\min})) \right|)$$

$$s.t. 0 = P_i - V_i \sum_{j \in N_i} V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \qquad i \in N_{B-1}$$

$$0 = Q_i - V_i \sum_{j \in N_i} V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \qquad i \in N_{PQ} \tag{9.2}$$

$$P_{gi}^{\min} \le P_{gi} \le P_{gi}^{\max} \qquad i \in N_g$$

Where power flow equations are used as equality constraints, which represent the network power losses but are more accurate than *B* matrix representation that is used in most ED programs, active power generation restrictions are used as inequality constraints. The active power generation P_{gs} at slack bus, the subscript s represents the slack bus, is the state variable, which is restricted by a quadratic penalty term added to

the objective function to form a penalty function. Equation 9.2 is therefore changed to the following generalized objective function:

$$\min f_p = \sum_{\substack{i \in N_g \\ i \neq s}} (a_i + b_i P_{gi} + c_i P_{gi}^2 + \left| e_i \sin(f_i (P_i - P_i^{\min})) \right|) + \lambda_{pgs} (P_{gs} - P_{gs}^{\lim})^2$$

$$s.t. 0 = P_i - V_i \sum_{j \in N_i} V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \qquad i \in N_{B-1}$$

$$0 = Q_i - V_i \sum_{j \in N_i} V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \qquad i \in N_{PQ}$$

$$P_{gi}^{\min} \leq P_{gi} \leq P_{gi}^{\max} \quad i \in N_g \quad i \neq s \tag{9.3}$$

where $\lambda_{p_{gs}}$ is the penalty factor which can be increased in the optimization procedure. The inequality constraints control variable constraints so they are self-restricted; is defined in the following equation:

$$P_{gs}^{\lim} = \begin{cases} P_{gs}^{\min} & if \quad P_{gs} < P_{gs}^{\min} \\ P_{gs}^{\max} & if \quad P_{gs} > P_{gs}^{\max} \end{cases} \tag{9.4}$$

9.3 Numerical Examples

The number of individuals is set to 40. 40 trials are performed for other simulations using different random numbers. At each trial, the best-evaluated value is stored within 100 searching iteration. w is set to 0.8 and c_1 and c_2 are 2.

In this section, IEEE 30-bus system has been used to show the effectiveness of the algorithm. The one-line diagram is shown in Figure 9.2. The network parameters are given in [92, 93]. Six generators are used in dispatch. The parameters of the characteristics of the steam turbine generators are given in Table 9.1. The powers in the table are in per unit on a power base of

$$s_B=100MVA$$

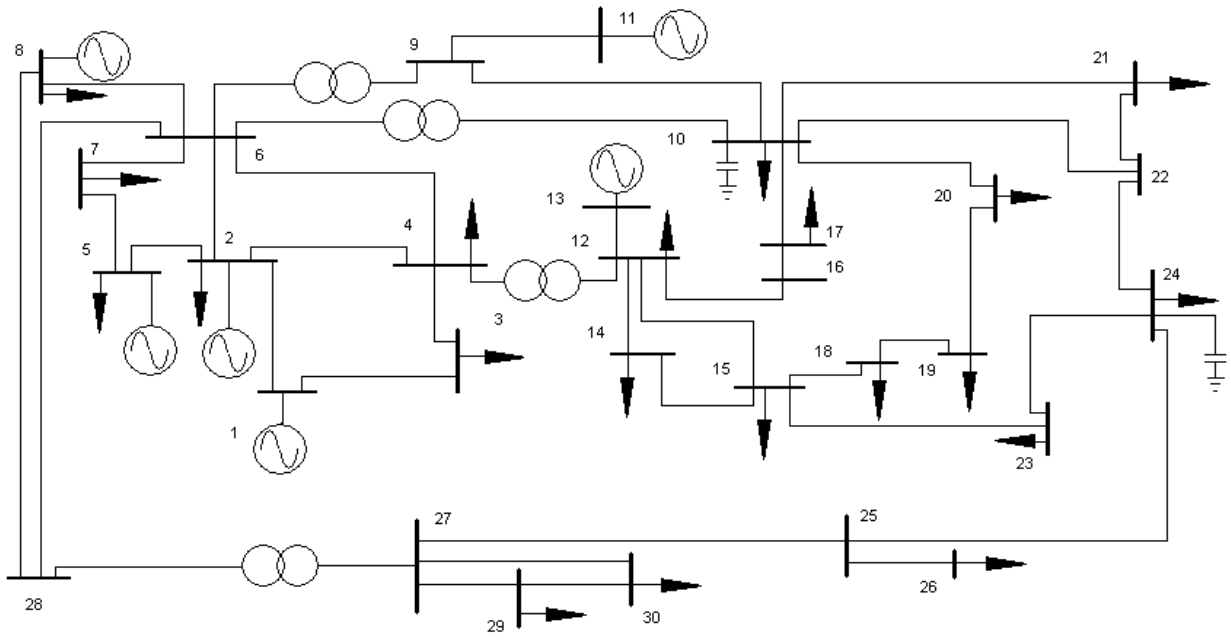


Figure 9.2 One-line diagram of IEEE 30-bus system

Table 9.1(a)(b) Parameters of characteristics of steam turbine generators

Generator	Unit 1	Unit 2	Unit 3
Bus	1	2	5
P_{max}(pu)	2.5	1.6	1
P_{min}(pu)	0.5	0.2	0.15
a(Mbtu)	0	0	0
b(Mbtu/puW)	200	175	100
c(Mbtu/puW^2)	37.5	175	625
e(Mbtu)	15	10	10
f(rad/puW)	6.283	8.976	14.784

(a)

Generator	Unit 4	Unit 5	Unit 6
Bus	8	11	13
P_{max}(pu)	0.7	0.6	0.8
P_{min}(pu)	0.1	0.1	0.12
a(Mbtu)	0	0	0
b(Mbtu/puW)	325	300	300
c(Mbtu/puW^2)	83.4	250	250
e(Mbtu)	5	5	5
f(rad/puW)	20.944	25.133	18.48

(b)

The load periods are simply divided between the valley load duration and the peak load duration. The valley loads and the initial generations are the same as in [93]. The peak loads are twice the valley loads. The results are given in Table 9.2. The heat saved is 9.6% in the valley load duration and 14.8% in the peak load duration.

Table 9.2 Optimization results

Population Size			40			
Valley Loads						
Initial Results						
Generator	Unit 1		Unit 2		Unit 3	
P_g(pu)	1.00		0.8		0.5	
Generator	Unit 4		Unit 5		Unit 6	
P_g(pu)	0.2		0.2		0.2	
Heat Consumed (MBtu)			902.40			
Generator	Unit 1		Unit 2		Unit 3	
P(pu)	EP	PSO	EP	PSO	EP	PSO
	1.99	1.98	0.51	0.52	0.15	0.15
Generator	Unit 4		Unit 5		Unit 6	
P(pu)	EP	PSO	EP	PSO	EP	PSO
	0.1	0.1	0.1	0.1	0.12	0.12
Heat Consumed (MBtu)			EP		PSO	
			815.27		816.11	
Peak Loads						
Initial Results						
Generator	Unit 1		Unit 2		Unit 3	
P_g(pu)	2.15		1.6		1	
Generator	Unit 4		Unit 5		Unit 6	
P_g(pu)	0.4		0.4		0.4	
Heat Consumed (MBtu)			2518.20			
Generator	Unit 1		Unit 2		Unit 3	
P(pu)	EP	PSO	EP	PSO	EP	PSO
	2.50	2.51	1.25	1.26	0.36	0.36
Generator	Unit 4		Unit 5		Unit 6	
P(pu)	EP	PSO	EP	PSO	EP	PSO
	0.7	0.7	0.6	0.59	0.58	0.59
Heat consumed (Mbtu)			EP		PSO	
			2144.45		2143.87	

9.4 Conclusion

The potential for application of PSO to ED has been shown in this chapter. Results from EP and simple PSO method are very similar. It can be foreseen that with parallel hybrid adaptive particle swarm optimization algorithm, the use of larger population sizes may result in improved convergence rates to the global solution. An adaptive PSO algorithm that increases population size incrementally may also improve algorithm convergence characteristics. It seems that PSO could be a good technique to solve real life power system problems in the near future.

Chapter 10

CONCLUSIONS AND FUTURE WORK

10.1 Accomplishments of the Thesis

This thesis has explored the feasibility of smart meters which are essential to the power distribution management systems and how their application can impact on energy efficiency.

A wavelet-GA-ANN based hybrid model is developed for accurate prediction of short-term load forecasting in power systems. The structure of the auto-configuring RBF network is optimized using floating point GA. The short-term load data are transformed into wavelet coefficients using à trous wavelet transform before using them for training the RBF network. Use of wavelet transform ensures extraction of more hidden features (both in time and frequency) in a compressed form. The compression of data enables the RBF network to be more efficient. In the thesis, the results from different practical load data demonstrate that the proposed model is capable of accurately predicting the STLF for seven steps ahead.

It has been shown that Dynamic Weighted Time-Delay Neural Network (DWTDNN) is a simplified version of the focused gamma network and an extension of TDNN as it incorporates *a priori* knowledge available about the task into the network architecture. As an example, the estimations produced by the methodology were applied on 8 different weather forecasting data provided by the Shanghai Meteorology Centre to make the result more practical. The results confirm that the proposed solutions have the potential for successful application to the problem of temperature and rainfall estimation, and the relationships between the factors that contribute to certain weather conditions can be estimated to a certain extent. Also, two applications based on

extending version of Graphical User Interface have been introduced; one is a simple voting system, which is designed on feedforward network; and the other is a 3-phase generator output detector, which is designed by Radial Basis Network (RBF). These examples present how to design the neural network in applications.

In this thesis, the comparison of swarm intelligence, improved genetic algorithm and a conventional gradient-based optimization method show that Swarm Intelligence would produce similar results on the IEEE 30-bus system as those from improved genetic algorithm while the conventional method was not as good as the other two in solving ORPD problem; therefore, Swarm Intelligence (SI) could be a good technique to solve real life power system problems in the near future.

10.2 Trends

The future of power transmission and distribution grids are expected to involve an increasing level of intelligence and integration of new information in every aspect of the electricity system, from demand-side devices to wide-scale distributed generation to a variety of energy markets. Price awareness and sensitivity are possible, and shared between energy suppliers and customers, creating a sophisticated real-time energy market.

The future smart metering system will rely on policies and governmental fiscal stimuli. However the cyber security requirement will also be an issue.

A smart meter that deals with different energy-supplying forms will be required in the future. AI technology to find out an optimal plan for complex usage from different types of energy resources is necessary. This smart meter will form an instrumented part of a micro-grid which further aggregates together to develop a smart grid. It is foreseen that the cost of smart meters could be reduced significantly; therefore it will

be widely used by many countries.

Recently, smart grid has attracted a great deal of attention all over the world; here are some references from PES annual meeting in 2010 for information purpose [94-135].

As to Adaptive RBFN model for 2D spatial interpolation, future works should focus on how to improve the stability of calculation, and the regulation of how to choose more effective initial kernel and deviation. Moreover, the performance of parameters calculation may limit the number of selected kernels.

Further research works also will be conducted to modify the proposed algorithms for optimising economic dispatch with demand side management.

References

[1] M. Amin, B. Wollenberg, "Towards a smart grid", IEEE Power and Energy Magazine, Vol. 3, No 5, September~October 2005, pp.34-38.

[2] S. Ray, G.K. Venayagamoorthy, "Real-time implementation of a measurement based adaptive wide area control system considering communication delays", IET Proceedings on Generation, Transmission and Distribution, Vol.2, No.1, January 2008, pp.62-70.

[3] W. Qiao, G.K. Venayagamoorthy, R.G. Harley, "Missing-sensor-fault-tolerant control for SSSC FACTS device with real-time implementation", IEEE Transactions on Power Delivery, Vol.24, No.2, April 2009, pp.740-750.

[4] T. Hammons, L.L. Lai, K.P. Wong, "International practices in smart grid design, operation, control and management", IEEE Power and Energy Society Annual Meeting, Canada, July 2009.

[5] R. Lee, L.L. Lai, "Smart metering in micro-grid applications", IEEE Power and Energy Society Annual Meeting, Canada, July 2009.

[6] G.K. Venayagamoorthy, "Potential and promises of computational intelligence for the smart grid", IEEE Power and Energy Society Annual Meeting, Canada, July 2009.

[7] L. Wang, P. Kulshrestha, M.Y Chow, "Intelligent energy management system for PHEVs at municipal parking deck in a smart grid environment", IEEE Power and Energy Society Annual Meeting, Canada, July 2009.

[8] C. Sasse, "Electricity Networks of the Future", IEEE PES General Meeting, 2006.

[9] European Commission, "European Electricity Projects 2002 – 2006", 2007

[online] Available:

http://ec.europa.eu/research/rtdinfo

Accessed on November 07 2008.

[10] F. Sissine, "Energy Independence and Security Act of 2007: A Summary of Major Provisions", CRS Report for Congress, Order Code RL34294, 21st December 2007.

[11] Energy retail association, "Smart Meters"

[Online] Available:

http://www.energy-retail.org.uk/smartmeters.html.

Accessed on November 30 2009.

[12] China to announce massive Smart Meter & Grid Project

[Online] Available:

http://www.smartmeters.com/the-news/547-china-to-announce-massive-smart-meter-a-grid project.html.

Accessed on November 30 2009.

[13] MCE Decision paper, "A National Minimum Functionality for Smart Meters"

[Online] Available:

http://www.ret.gov.au/Documents/mce/_documents/MCE_Decision_Paper_A_National_Minimum_Functionality_for_Smart_Meters20071213153847.pdf.

Accessed on December 2 2009.

[14] Smart Energy Overview

[Online] Available:

http://www.zigbee.org/Markets/ZigBeeSmartEnergy/ZigBeeSmartEnergyOverview/tabid/431/Default.aspx.

Accessed on December 3 2009.

[15] P.K. Lee, L.L. Lai, "A practical approach of smart metering in remote monitoring of renewable energy applications", Power & Energy Society General Meeting, 2009, PES, 26-30 July 2009, pp.1 – 4.

[16] What smart meters will do for you

[Online] Available:

http://news.bbc.co.uk/1/hi/business/8023507.stm.

Accessed on December 6 2009.

[17] R.J.F. van Gerwen, S.A. Jaarsma, F.T.C. Koenis, 'Domme meters worden slim? Kosten-baten analyse slimme meetinfrastructuur' (Are dumb meters getting smarter? Cost-benefit analysis for a smart metering infrastructure), KEMA report 40510016-TDC, 30th August 2005.

[18] Italy – 27 Million Smart Meters

[Online] Available:

http://neuralenergy.blogspot.com/2009/07/italy.html.

Accessed on December 6 2009.

[19] Italy's Smart Meter Revolution: Showing the World how it's done

[Online] Available:

http://www.energyboom.com/emerging/italys-smart-meter-revolution-showing-world-how-its-done.

Accessed on 8 December 2009.

[20] R.J.F. van Gerwen, S.A. Jaarsma, R. Wilhite, "Smart Metering", Distributed Generation, July 2006, pp.7.

[21] Smart Meter

[Online] Available:

http://en.wikipedia.org/wiki/Smart_meter.

Accessed on December 9 2009.

[22] Pacific Gas and Electric Company's Smart Meter Proposal Approved by California Public Utilities Commission

[Online] Available:

http://www.pge.com/about/news/mediarelations/newsreleases/q3_2006/060720a.shtml

Accessed on December 11 2009.

[23] Advanced Metering Infrastructure for Victoria

[Online] Available:

http://www.dpi.vic.gov.au/dpi/dpinenergy.nsf/LinkView/4EC2E4EA42B821FCCA2572B10079A930A8BAF6E4E66C900FCA2572B20004C403.

Accessed on December 13 2009.

[24] ERDF and Atos Origin Partner on Smart Meters

[Online] Available:

http://smartelectricnews.blogspot.com/2008/07/erdf-and-atos-origin-partner-on-smart.html

Accessed on December 13 2009.

[25] P.J. Santos, A.G. Martins, A.J. Pries, J.F. Martins, R.V. Mendes, “Short-Term Load Forecast Using Trend Information and Process Reconstruction”, International Journal of Energy Research, 2006, pp.811-822.

[26] D.K. Ranaweera, G.G. Karady, R.G. Farmer, “Effect of probabilistic inputs on neural network-based electric load forecasting”, IEEE Transactions on Neural Works, Vol.7, No.6, 1996, pp.1528–1532.

[27] J. N. Bastian, J. Zhu, V. Banunarayana, R. Mukerji, Forecasting energy prices in a competitive market, IEEE Computer Applications in Power, Vol.13, No.3, 1999, pp. 40–45.

[28] Amir B. Geva, “ScaleNet - Multiscale neural-network architecture for time series prediction”, IEEE Transactions on Neural Networks, Vol.9, No.5, 1998, pp.1471-1482.

[29] A. Oonsivilai, M.E. El-Hawary, “Wavelet Neural Network Based Short Term Load Forecasting of Electric Power System Commercial Load”, IEEE Proceedings, 1999, pp.1223-1228.

[30] O. Renaud, J.L. Starck, F. Murtagh, “Wavelet-based Forecasting of Short and Long Memory Time Series”, Econpapers, May 2002.

[31] Cheng-Jian Lin, “Wavelet Neural Networks with a Hybrid Learning Approach”, Journal of Information Science and Engineering 22, 2006, pp.1367-1387.

[32] C. Kim, I. Yu, Y.H. Song, “Kohonen neural network and wavelet transform based approach to short-term load forecasting”, Electric Power Systems Research 63, 2002, pp.169-176.

[33] N. Satio, G. Beylkin,“Multiresolution representations using the auto-correlation functions of compactly supported wavelets”, IEEE Transactions on Signal Processing 41, No.12, 1993, pp.3584–3590.

[34] I. Daubechies, Ten Lectures on Wavelets, CBMS-NSF, SIAM, 1992.

[35] K.P. Soman, K.I. Ramachandran, “Insight into Wavelets from Theory to Practice”, Prentice-Hall of India, 2004.

[36] D. Benaouda, F. Muratagh, J.L. Starck, O. Renaud, “Wavelet-Based Nonlinear Multiscale Decomposition Model for Electricity Load Forecasting”, Neurocomputing 70 Elsevier, 2006, pp.139-154.

[37] J. Park, I.W. Sandberg, “Universal approximation using radial basis function networks”, Neural Computation, Vol.3 (2), 1991, pp.246-257.

[38] S. Chen, C.F.N. Cowan, P.M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks", IEEE Transactions on Neural Networks, Vol.2, 1991, pp.302–309.

[39] Australian National Electricity Market Management Company Limited (NEMMCO).

[Online] Available:

http://www.nemmco.com.au/.

Accessed on April 12 2008.

[40] James.N.K. Liu, Raymond.S.T. Lee, "Rainfall Forecasting from Multiple Point Sources Using Neural Networks", Proceeding of IEEE International Conference on Systems, Man and Cybernetics, Vol.3, 1999, pp.429-434.

[41] J. McCullagh, K. Bluff, E. Ebert, "A Neural Network Model for Rainfall Estimation", ANNES Proceedings, November 1995, pp.389-392.

[42] A. Khotanzad, M. H. Davis, A. Abaye, D. J. Maratukulam, "An Artificial Neural Network Hourly Temperature Forecaster with Applications in Load Forecasting", IEEE Transactions on Power Systems, Vol.11, No.2, May 1996, pp.870-876.

[43] L.L. Lai, H. Subasinghe, N. Rajkumar, E. Vaseekar, B. J. Gwyn, V. K. Sood, "Object-Oriented Genetic Algorithm Based Artificial Neural Network for Load Forecasting", IEEE SEAL 1998, pp.462-469.

[44] V. Venkatraman, J.C. Rajapakse, "Dynamically Weighted Time-Delay Neural Networks for Functional Approximation", Proceedings of IJCNN 2001, Vol.1 of 4, 2001, pp.2768-2773.

[45] R. Sitte, J. Sitte, "Analysis of the Predictive Ability of Time Delay Neural Networks Applied to the S&P 500 Time Series", IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews, Vol.30, No.4, November 2000,

pp.568-572.

[46] Christos Stergiou, Dimitrios Siganos. “Neural Networks”

[online] Available:

http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html

Accessed on November 25 2009

[47] Matlab help --- Neural Network toolbox -> backpropagation The MATH WORKS Inc.

[48] O. Alsac, J. Bright, M. Prais, B. Stott, “Further developments in LP-based optimal power flow”, IEEE Transactions on Power Systems, 1990, pp. 697-711.

[49] N. Deeb, S.M. Shahidehpour, “Linear reactive power optimization in a large power network using the decomposition approach”, IEEE Transactions on Power Systems, PWRS-5, 1990, pp.428-438.

[50] K. Mamandur, R. Chenoweth, “Optimal control of reactive power flow for improvements in voltage profiles and for real power loss minimization”, IEEE Transactions on Power Apparatus & Systems, PAS100, 1981, pp.3185-3194.

[51] M.O. Mansour, T.M. Abdel-Rahman, “Non-linear VAr optimization using decomposition and coordination”, IEEE Transactions on Power Apparatus & Systems, PAS-103, 1984, pp.246-255.

[52] J. Qiu, S.M. Shahidehpour, “A new approach for minimizing power losses and improving voltage profile”, IEEE Transactions on Power Systems, PWRS-2, 1987, pp.287-295.

[53] K.H. Abdul-Rahman, S.M. Shahidehpour, “Application of fuzzy sets to optimal reactive power planning with security constraints”, IEEE Transactions on Power Systems, PWRS-9, 1994, pp.589- 597.

[54] G. Boone, H. D. Chiang, “Optimal capacitor, Placement in distribution systems

by genetic algorithm", Electrical Power & Energy Systems, Vol.15, 1993, pp.155-162.

[55] Y.T. Hsiao, C.C. Liu, H.D. Chiang, Y.L. Chen, "A new approach for optimal VAR sources planning in large scale electric power systems", IEEE Transactions on Power Systems, PWRS-8, 1993, pp.988-996.

[56] J.T. Ma, L.L. Lai, "Application of genetic algorithms to optimal reactive power dispatch of power systems", IFAC Symposium on Large Scale Systems, City University, London, UK, July 1995, pp.615-620.

[57] H. Yoshida, K. Kawata, Y. Fukuyama, Y. Nakanishi, "A particle swarm optimization for reactive power and voltage control considering voltage security assessment," IEEE Transactions on Power Systems, Vol.15, No.4, 2001, pp.1232-1239.

[58] K.E. Parsopoulos, M.N.Vrahatis, "Recent approaches to global optimization problems through particle swarm optimization", Natural Computing, Kluwer, Netherlands, 2002, pp.235-306.

[59] J.F. Schutte, J.A. Reinbolt, B.J. Fregly, R.T. Haftka, A.D. George, "Parallel global optimization with the particle swarm algorithm", International Journal for Numerical Methods in Engineering, Wiley, 2004, pp.2296-2315.

[60] H.H. Balci, J.F. Valenzuela, "Scheduling electric power generators using particle swarm optimization combined with the Lagrangian Relaxation method," International Journal of Applied Mathematics and Computer Science, Vol.14, No.3, 2004, pp.411-421.

[61] J. Kennedy, R. Eberhart, Swarm Intelligence – San Francisco: Morgan Kaufmann, 2001.

[62] J. Kennedy, R. Eberhart, "Particle swarm optimization," Proceedings of the

International Conference on Neural Networks, IEEE, Vol.4, Perth, Australia, 1995, pp.1942-1948.

[63] G.L. Nemhauser, A.H.G. Rhmooy Kan, M.J. Todd, Optimization, Vol.1, Elsevier Science Publishers B.V., Amsterdam, 1989.

[64] K.Y. Lee, Y.M. Park, J.L. Ortiz, "A united approach to optimal real and reactive power dispatch", IEEE Transactions on Power Apparatus and Systems, PAS-104, 1985, pp.1147-l153.

[65] Q.P. Zhang, L.L. Lai, H. Wei, "Continuous space optimized artificial ant colony for real-time typhoon eye tracking", IEEE International Conference on Systems, Man, and Cybernetics (SMC'2007), 2007.

[66] L.L. Lai, H. Braun, Q.P. Zhang, Q. Wu, Y. Ma, W.C. Sun, L. Yang, "Intelligent weather forecast", Proceedings of the Third International Conference on machine Learning and Cybernetics, IEEE, Shanghai, August 2004.

[67] Q.P. Zhang, L.L. Lai, "Intelligent location of tropical cyclone centre", Proceedings of 4th IEEE International Conference on Machine Learning and Cybernetics, 2005.

[68] N. Amenta, M. Bern, M. Kamvysselis, "A new voronoi based surface reconstruction algorithm", Proceedings of SIGGRAPH '98, ACM Press, 1998, pp.415-421.

[69] L. Mitas, H. Mitasova, Multidimensional Spatial Interpolation, Geographic Modeling Systems Laboratory, University of Illinois at Urbana-Champaign, 1998.

[70] H. Hoppe, "View-Dependent Refinement of Progressive Meshes", Proceedings of SIGGRAPH 1997.

[71] L. Li, P. Revesz, "The relationship among GIS-oriented spatiotemporal databases", Proceedings of the Third National Conference on Digital Government

Research, Boston, 2003.

[72] K.E. Kerry, K.A. Hawic, “Kriging Interpolation on High-Performance Computers”, Technical Report DHPC-03, 1998.

[73] J. Morrison, “Kriging in a parallel environment”, School of Computer Science, KIS5B6, Canada, 1998.

[74] H. Hild Fritsch, “Generation of a consistent spatial database for a land use planning project in sourth eastern China (SILUP)”, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol.34, Part 6/W6, 1998.

[75] C. Furlanello, M. Neteler, Stefano Merler, “GIS and the random forest predictor: integration in R for tick-borne disease risk assessment”, DSC 2003 Working Papers, 2003.

[76] J.S.R. Jang, C.T. Sun, “Functional equivalence between radial basis function networks and fuzzy inference systems”, IEEE Transactions on Neural Networks, Vol.4, 1993, pp.156-159.

[77] Y. Jin, M. Olhofer, B. Sendhoff, “A framework for evolutionary optimization with approximate fitness functions”, IEEE Transactions on Evolutionary Computation, Vol.6, No.5, 2002, pp.481-494.

[78] C. M. Bishop, “Improving the generalization properties of radial basis function neural networks”, Neural Computation, Vol.3, No.4, 1991, pp.579-581.

[79] J. Park, I.W. Sandberg, “Approximation and radial basis function networks”, Neural Computation, 1993, pp.305-316.

[80] L. Xu, A. Krzyzak, A. Yuille, “On radial basis function nets and kernel regression: statistical consistency, convergence rates, and receptive field size”, Neural Networks, Vol.7, No.4, 1994, pp.609-628.

[81] T. Poggio, F. Girosi, “Networks for approximation and learning”, Proceedings of the IEEE 78(9), 1990, pp.1481-1497.

[82] H.R. Berenji, “Fuzzy logic controllers,” An Introduction to Fuzzy Logic Applications in Intelligent Systems (R.R. Yager and L.A. Zadeh, Eds.), Kluver Academic Publishers. 1992

[83] V. Tresp, J. Hollatz, S. Ahumad, “Network structuring and training using rule-based knowledge”, In Advances in Neural Information Processing Systems 5, 1993.

[84] T.M. Martinetz, “Competitive hebbian learning rule forms perfectly topology preserving maps”, International Conf. on Artificial Neural Networks, 1993, pp.427-434.

[85] G. Tesauro, D.S. Touretzky, T.K. Leen, “A growing neural gas network learns topologies,” In Advances in Neural Information Processing Systems, MIT Press, 1995.

[86] T. Kohonen, “Self-organized formation of topological correct feature maps”, Biological Cybernetics, 1982, pp.59-69.

[87] L. Xu, Y. Yang, “Bayesian Learning (I): A unified perspective for statistical modeling”, Intelligent Technologies for Information Analysis, Springer, 2004, pp.615-659.

[88] R.R. Shoults, S.V. Venkatesh, S.D. Helmick, C.L. Ward, M.J. Lollar, “A dynamic programming based method for developing dispatch curves when incremental heat rate curves are non-monotonically increasing”, IEEE Transactions on Power Systems, Vol.PWRS-1, No.1, February 1986, pp.10-16.

[89] K.P. Wong, C.C. Fung, “Simulated annealing based economic dispatch algorithm”, IEE Proceedings of Generation, Transmission and Distribution, Vol.140, No.6, 1993, pp.509 -515.

[90] D.C. Walters, G.B. Sheble, “Genetic algorithm solution of economic dispatch with valve point loading”, IEEE Transactions on Power Systems, PWRS-8, 1993, pp.1325-1332.

[91] L.L. Lai, J.T. Ma, K.P. Wong, “Evolutionary programming for economic dispatch of units with non-smooth input-output characteristic functions”, Proceedings of the 12th Power Systems Computation Conference Dresden, August 19 - 23, 1996, pp.492-498.

[92] S. Naka, T. Genji, T. Yura, Y. Fukuyama, “Practical distribution state estimation using hybrid particle swarm optimization”, Proceedings of IEEE Power Engineering Society inter Meeting, January-Feburary 2001.

[93] O. Alsac, B. Stott, “Optimal load flow with steady-state security”, IEEE Transactions on Power Apparatus and Systems PAS-93, 1974, pp.745-751.

[94] R.R. Mariani, R.K. Rayudu, M.S. Witherden, E. Lai, “Power Quality Indices of Compact Fluorescent Lamps for Residential Use – A New Zealand Study”, IEEE Power and Energy Society General Meeting 2010, paper ID:000024 , July 2010.

[95] R.H. Salim, M. Oleskovicz, R.A. Ramos, “Assessment of Voltage Fluctuations Induced by Electromechanical Oscillations in Distributed Generation Systems”, IEEE Power and Energy Society General Meeting 2010, paper ID:000028 , July 2010.

[96] L. Gao, W. Li, W. Zhang, F. Zhu, Q. Chu, Z. Liu, H. Zhao, “Verification of System Dynamic Performance of the 1000kV Synchronized Grid”, IEEE Power and Energy Society General Meeting 2010, paper ID:000034 , July 2010.

[97] D.Le. Cuong, Y.H. Irene, H.J. Math, “Joint Causal and Anti-Causal Segmentation and Location of Transitions in Power Disturbances”, IEEE Power and Energy Society General Meeting 2010, paper ID:000040 , July 2010.

[98] E.C. Guardia, A.R. Queiroz, J.W. Marangon Lima, “Estimation of Electricity Elasticity for Demand Rates and Load Curve in Brazil”, IEEE Power and Energy Society General Meeting 2010, paper ID:000044 , July 2010.

[99] A. Bracale, G. Carpinelli, P. Caramia, P. Verde, R. Chiumeo, F. Tarsia, “System Voltage Quality Regulation for Continuous Disturbances”, IEEE Power and Energy Society General Meeting 2010, paper ID:000059 , July 2010.

[100] Dr.Tevfik Sezi, “Fast and Accurate Measurement of Power System Frequency”, IEEE Power and Energy Society General Meeting 2010, paper ID:000107 , July 2010.

[101] S. Corsi, F.De. Villiers, R. Vajeth, “Secondary Voltage Regulation Applied to the South Africa Transmission Grid”, IEEE Power and Energy Society General Meeting 2010, paper ID:000154 , July 2010.

[102] B. Ramachandran, S.K. Srivastava, D.A. Cartes, C.S. Edrington, “Distributed Energy Resource Management in a Smart Grid by Risk Based Auction Strategy for Profit Maximization”, IEEE Power and Energy Society General Meeting 2010, paper ID:000192 , July 2010.

[103] A. Motamedi, H. Zareipour, W.D. Rosehart, “Electricity Market Price Forecasting in a Price-responsive Smart Grid Environment”, IEEE Power and Energy Society General Meeting 2010, paper ID:000193 , July 2010.

[104] M.F. Faisal, A. Mohamed, H. Shareef, “A Novel Method for Upstream Asymmetrical Voltage Sag-Source Detection”, IEEE Power and Energy Society General Meeting 2010, paper ID:000240 , July 2010.

[105] A. Tolk, “Architecture Constraints for Interoperability and Composability in a Smart Grid”, IEEE Power and Energy Society General Meeting 2010, paper ID:000247 , July 2010.

[106] A. Nourai, R. Sastry, T. Walker, “A Vision & Strategy for Deployment of Energy Storage in Electric Utilities”, IEEE Power and Energy Society General Meeting 2010, paper ID:000250 , July 2010.

[107] A. Nagarajan, W. Shireen, “Grid Connected Residential Photovoltaic Energy Systems with Plug-In Hybrid Electric Vehicles(PHEV) as Energy Storage”, IEEE Power and Energy Society General Meeting 2010, paper ID:000284 , July 2010.

[108] A. Kowli, M.N. Pincetic, G. Gross, “A Successful Implementation with the Smart Grid: Demand Response Resources”, IEEE Power and Energy Society General Meeting 2010, paper ID:000312 , July 2010.

[109] H. Liu, Z.B. Zabinsky, W. Kohn, “Rule-based Control System Design for Smart Grids”, IEEE Power and Energy Society General Meeting 2010, paper ID:000315 , July 2010.

[110] M. Lauby, “Reliability Considerations for Application of Smart Grid Technologies”, IEEE Power and Energy Society General Meeting 2010, paper ID:000318 , July 2010.

[111] G.W. Chang, C. Chen, “Performance Evaluation of Voltage Sag Detection Methods”, IEEE Power and Energy Society General Meeting 2010, paper ID:000335 , July 2010.

[112] S.T. Mak, “Sensor Date Output Requirements for Smart Grid Applications”, IEEE Power and Energy Society General Meeting 2010, paper ID:000371 , July 2010.

[113] J.A. Kavicky, “Impacts of Smart Grid Data on Parallel Path and Contingency Analysis Efforts”, IEEE Power and Energy Society General Meeting 2010, paper ID:000391 , July 2010.

[114] P.J.M. Heskes, J.M.A. Myrzik, W.L. Kling, "Harmonic Distortion and Oscillatory Voltages and the Role of Negative Impedance", IEEE Power and Energy Society General Meeting 2010, paper ID:000413, July 2010.

[115] P. Wang, J.Y. Huang, Y. Ding, P. Loh, L. Goel, "Demand Side Load Management of Smart Grids Using Intelligent Trading/Metering/ Billing System", IEEE Power and Energy Society General Meeting 2010, paper ID:000500, July 2010.

[116] T.J. Overbye, J.D. Weber, "The Smart Grid and PMUs: Operational Challenges and Opportunities", IEEE Power and Energy Society General Meeting 2010, paper ID:000580, July 2010.

[117] S.E. Widergren, H.Kirkham, "Smart Grid - Transforming Power System Operations", IEEE Power and Energy Society General Meeting 2010, paper ID:000606, July 2010.

[118] D. Haughton, G.T. Heydt, "Smart Distribution System Design: Automatic Reconfiguration for Improved Reliability", IEEE Power and Energy Society General Meeting 2010, paper ID:000662, July 2010.

[119] D. Pudjianto, C.K. Gan, V. Stanojevic, M. Aunedi, P. Djapic, G. Strbac, "Value of Integrating Distributed Energy Resources in the UK Electricity System", IEEE Power and Energy Society General Meeting 2010, paper ID:001510, July 2010.

[120] C. Singh, A. Sprintson, "Reliability Assurance of Cyber-Physical Power Systems", IEEE Power and Energy Society General Meeting 2010, paper ID:001511, July 2010.

[121] S. Suryanarayanan, P.F. Ribeiro, M.G. Simões, "Grid Modernization Efforts in the USA and Brazil – Some Common Lessons Based on the Smart Grid Initiative", IEEE Power and Energy Society General Meeting 2010, paper ID:001512, July 2010.

[122] S. Sridhar, G. Manimaran, “Data Integrity Attacks and their Impacts on SCADA Control System”, IEEE Power and Energy Society General Meeting 2010, paper ID:001522, July 2010.

[123] R. Podmore, “Smart Grid Restoration Concepts”, IEEE Power and Energy Society General Meeting 2010, paper ID:001540, July 2010.

[124] B. Kroposki, G. Martin, “Hybrid Renewable Energy and Microgrid Research Work at NREL”, IEEE Power and Energy Society General Meeting 2010, paper ID:001545, July 2010.

[125] J. McCalley, E. Ibáñez, Y. Gu, K. Gkritza, D. Aliprantis, L. Wang, A. Somani, R. Brown, “National Long-Term Investment Planning for Energy and Transportation Systems”, IEEE Power and Energy Society General Meeting 2010, paper ID:001550, July 2010.

[126] R. Moreno, D. Pudjianto, G. Strbac, “Future Transmission Network Operation and Design Standards to Support a Low Carbon Electricity System”, IEEE Power and Energy Society General Meeting 2010, paper ID:001552, July 2010.

[127] M. Saeedifard, M. Graovac, R.F. Dias, R. Iravani, “DC Power Systems: Challenges and Opportunities”, IEEE Power and Energy Society General Meeting 2010, paper ID:001568, July 2010.

[128] Z.A. Vale, H. Morais, H. Khodr, B. Canizes, J. Soares, “Technical and economic resources management in Smart Grids using Heuristic Optimization methods”, IEEE Power and Energy Society General Meeting 2010, paper ID:001618, July 2010.

[129] A. Bose, “Models and Techniques for Reliability Analysis of the Smart Grid”, IEEE Power and Energy Society General Meeting 2010, paper ID:001646, July 2010.

[130] M.R. Haghifam, A. Hadian, “Simultaneous Placement of Conventional and Renewable DGs in Electric Distribution System”, IEEE Power and Energy Society General Meeting 2010, paper ID:001662, July 2010.

[131] J. Taylor, A. Maitra, M. Alexander, D. Brooks, M. Duvall, “Evaluations of Plug-in Electric Vehicle Distribution System Impacts”, IEEE Power and Energy Society General Meeting 2010, paper ID:001665, July 2010.

[132] A.J. Flueck, C.P. Nguyen, “Integrating Renewable and Distributed Resources - IIT Perfect Power Smart Grid Prototype”, IEEE Power and Energy Society General Meeting 2010, paper ID:001679, July 2010.

[133] J. Ning, W. Gao, “Multi-feature Extraction for Power System Disturbances by Wavelet Transform and Fractal Analysis”, IEEE Power and Energy Society General Meeting 2010, paper ID:001685, July 2010.

[134] A. Baggini, F. Bua, “Overview on PQ development in Europe”, IEEE Power and Energy Society General Meeting 2010, paper ID:001691, July 2010

[135] S.M. Amin, “Electricity Infrastructure Security: Toward Reliable, Resilient and Secure Cyber-Physical Power and Energy Systems”, IEEE Power and Energy Society General Meeting 2010, paper ID:001719, July 2010.