# Probabilistic Estimation and Prediction of the Dynamic Response of the Demand at Bulk Supply Points

A thesis submitted to The University of Manchester for the Degree of

**Doctor of Philosophy**

In the Faculty of Engineering and Physical Sciences

**2015**

**Ms Yizheng Xu**

School of Electrical and Electronic Engineering

# TABLE OF CONTENTS

*Word Count: 56,767±5,000*

# LIST OF FIGURES

# LIST OF TABLES

# NOMENCLATURE

## *List of Symbols*

| | |
|---|---|
| $\| \|$ | Absolute value operator |
| $a$ | Output of the artificial neural network |
| $b$ | Bias of the artificial neural network |
| $c$ | Cost of the electricity |
| $C$ | Controllable |
| $d$ | Step change of the voltage |
| $d_p$ | Scaling factor for the real power |
| $d_q$ | Scaling factor for the reactive power |
| $DT$ | Day type |
| $e_{overall}$ | overall error of the disaggregation approach |
| $e_P$ | Load model error for the real power |
| $ep$ | Voltage exponent of the exponential load model for the real power |
| $eq$ | Voltage exponent of the exponential load model for the reactive power |
| $e_Q$ | Load model error for the reactive power |
| $f$ | Frequency |
| $f(), F()$ | A function |
| $H$ | Humidity |
| $J$ | Inertia |
| $k_{pf}$ | Frequency exponent of the exponential load model for the real power |
| $k_{pu}$ | Voltage exponent of the exponential load model for the real power |

| | |
|---|---|
| $K_{pu}$ | Gain |
| $L$ | Inductance |
| $LS$ | Load signature |
| $M$ | Torque |
| $Obj$ | Objective function |
| $P$ | Real power |
| $p$ | Input of the artificial neural network |
| $P_{rated}$ | Rated demand |
| $pf$ | Power factor |
| $Q$ | Reactive power |
| $R$ | Resistance |
| $s$ | Slip |
| $T$ | Temperature |
| $t$ | Time |
| $T_p$ | Recovery time constant for the real power |
| $T_q$ | Recovery time constant for the reactive power |
| $UC$ | Uncontrollable |
| $V$ | Voltage |
| $w$ | Weighting factor |
| $W$ | Weights of the artificial neural network |
| $WS$ | Wind speed |
| $X$ | Reactance |
| $\alpha$ | Voltage exponent of the exponential load model for the real power |
| $\beta$ | Voltage exponent of the exponential load model for the reactive power |

| | |
|---|---|
| $\Delta$ | Operator representing the variation |
| $\lambda$ | Magnetic flux |
| $\mu$ | Mean value |
| $\sigma$ | Standard deviation |
| $\omega$ | Angular frequency |

## *Subscripts*

| | |
|---|---|
| *dsigmf* | Difference-sigmoid shaped membership function |
| *gbellmf* | Bell-shaped membership function |
| *logsig* | Log-sigmoid transfer function |
| *PTRN* | Input matrix of the training process of the artificial neural network |
| *PTST* | Input matrix of the testing process of the artificial neural network |
| *purelin* | Linear transfer function |
| *tansig* | Tan-sigmoid transfer function |
| *trainbr* | Bayesian Regulation backpropagation |
| *traingd* | Gradient descent backpropagation |
| *trainlm* | Levenberg-Marquardt backpropagation |
| *trimf* | Triangle-shaped membership function |
| *TTRN* | Target matrix of the training process of the artificial neural network |
| *TTST* | Target matrix of the testing process of the artificial neural network |

## *Acronyms*

| | |
|---|---|
| A/C | Air conditioning |
| AI | Artificial intelligence |

| | |
|---|---|
| ALME | Absolute load model error |
| ANFIS | Adaptive neuro-based fuzzy inference system |
| ANN | Artificial neural network |
| APE | Absolute percentage error |
| AR | Autoregression |
| ARIMA | Autoregression integrated moving average |
| ARMA | Autoregression moving average |
| ASD | Adjustable speed drive |
| AWFE | Absolute weighting factor error |
| BLF | *Base load* forecasting |
| BLFer | *Base load* forecaster |
| BP | Backpropagation |
| BRBP | Bayesian Regulation backpropagation |
| BSP | Bulk supply point |
| CAPE | Cumulative absolute percentage error |
| CBA | Component-based approach |
| CDF | Cumulative distribution function |
| CdL | Cold load |
| CFANN | Cascade-forward artificial neural network |
| CFL | Compact fluorescent lamps |
| CI | Confidence interval |
| CIGRE | Conseil International des Grands Réseaux Électriques (International Council on Large Electric Systems) |
| CL | Confidence level |
| CLF | *Change load* forecasting |
| CLFer | *Change load* forecaster |

| | |
|---|---|
| CT | Constant torque |
| CTIM1 | Single-phase constant torque induction motor |
| CTIM3 | Three-phase constant torque induction motor |
| CW | Current waveform |
| DADRS | Daily aggregate demand response surface |
| DDLC | Decomposed daily loading curve |
| DFR | Digital fault recorder |
| DFT | Discrete Fourier Transform |
| DG | Distribution generation |
| DRD | Dynamic response of demand |
| DSM | Demand side management |
| EEL | Energy efficient lighting |
| EIG | Eigenvalues |
| EMTP | Electromagnetic Transients Program |
| EPRI | Electric Power Research Institute |
| EV | Electric vehicle |
| FFANN | Feed-forward artificial neural network |
| FIS | Fuzzy inference system |
| GDBP | Gradient descent backpropagation |
| GIL | General incandescent lights |
| HAR | Harmonics |
| HID | High-intensity discharge lighting |
| HV | High Voltage |
| IAW | Instantaneous admittance waveform |
| IEEE | Institute of Electrical and Electronics Engineers |

| | |
|---|---|
| IM | Induction motor |
| IMF | Input membership function |
| IPW | Instantaneous power waveform |
| LED | Light-emitting diode |
| LMBP | Levenberg-Marquardt backpropagation |
| MA | Moving average |
| MAPE | Mean absolute percentage error |
| MBA | Measurement-based approach |
| MCS | Monte Carlo simulation |
| MES | Meat emulsion system |
| MF | Membership function |
| MP | Most probable |
| MPV | Most probable value |
| MV | Medium voltage |
| OMF | Output membership function |
| PDF | Probability density function |
| PQ | Real and reactive power |
| QT | Quadratic torque |
| QTIM1 | Single-phase quadratic torque induction motor |
| QTIM3 | Three-phase quadratic torque induction motor |
| RD | Rated demand |
| REC | Rectifier |
| RLR | Robust local regression |
| RMS | Root mean squared |
| RMSE | Root mean squared error |

| | |
|---|---|
| SEA | Smart energy appliances |
| SES | Smart energy services |
| SG | Savitzky-Golay |
| SMPS | Switch-mode power supply |
| SNRF | Signal-to-noise-ratio figure |
| STLF | Short-term load forecasting |
| STW | Switching transient waveform |
| TF | Transfer function |
| TV | Television |
| WECC | Western Electricity Coordinating Council |
| WF | Weighting factor |
| WFE | Weighting factor error |
| WL | Wet load |
| WSCC | Western Systems Coordinating Council |
| ZIP | Constant impedance, constant current and constant power |

# ABSTRACT

*Probabilistic Estimation and Prediction of the Dynamic Response of the Demand at Bulk Supply Points*

***Ms Yizheng Xu, The University of Manchester, March 2015***

The dynamic response of the demand is defined as the time-domain real and reactive power response to a voltage disturbance, and it represents the dynamic load characteristics. This thesis develops a methodology for probabilistic estimation and prediction of dynamic responses of the demand at bulk supply points. The main outcome of the research is being able to predict the contribution of different categories of loads to the total demand mix and their controllability without conducting detailed customer surveys or collecting smart meter data, and to predict the dynamic response of the demand without performing field tests.

The prediction of the contributions of different load categories and their controllability and load characteristics in the near future (e.g., day ahead) plays an important role in system analysis and planning, especially in the short-term dispatch and control. However, the research related to this topic is missing in the publically available literature, and an approach needs to be developed to enable the prediction of the participation of different loads in total load mix, their controllability and the dynamic response of the demand.

This research contributes to a number of areas, such as load forecasting, load disaggregation and load modelling. First, two load forecasting methodologies which have not been compared before are compared; and based on the results of comparison and considering the actual requirements in this research, a methodology is selected and used to predict both the real and reactive power. Second, a unique methodology for load disaggregation is developed. This methodology enables the estimation of the contributions of different load categories to the total demand mix and their controllability based on RMS measured voltage and real and reactive power. The confidence level of the estimation is also assessed. The methodology for disaggregation is integrated with the load forecasting tool to enable prediction of load compositions and dynamic responses of the demand. The prediction is validated with data collected from real UK power network. Finally, based on the prediction, an example of load shifting is used to demonstrate that different dynamic responses can be obtained based on the availability and redistribution of controllable devices and that load shifting decisions, i.e., demand side management actions, should be made based not only on the amount of demand to be shifted, but also on predicted responses before and after load shifting.

# DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or institute of learning.

# COPYRIGHT STATEMENT

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyrights or related rights in it (the "Copyright") and she has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property Rights") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosures, publications and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy[1], in any relevant Thesis restriction declarations deposited in

[1] see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual- property.pdf

the University Library, The University Library's regulations [2] and in The University's policy on presentation of Theses.

[2] see http://www.manchester.ac.uk/library/aboutus /regulations

# ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor, Prof. Jovica Milanović, for his endless support, guidance and inspiration throughout the period of the research. His commitment to excellence and the patience with which he has reviewed this thesis and the research contained within it has been inspirational.

A special acknowledgement must also go to Engineering and Physical Sciences Research Council (EPSRC) Project *Autonomic Power System* who sponsored the research. I would also like to thank the experts and colleagues working on EPSRC Project *Autonomic Power Systems* for helping strengthening my research background in relevant areas and those working on EU FP7 Project *SuSTAINABLE* and Project *CLASS* for helping strengthening my background in relevant industrial context.

My great appreciation also goes to colleagues from the *Power Quality and Power Systems Dynamics Team* in the *Electrical Energy and Power System Group* of the *School of Electrical and Electronic Engineering* at The University of Manchester. The opportunities to discuss ideas have been significant and have contributed to ensuring the success of the research.

I would like to thank my parents for their unconditional support and provision of resources, information and solutions on both the subject and daily life. I would also like to thank my friends, Dr. Xuecheng Zhang, Ms. Wen Yang, Mr. Shuxuan Chen and Ms. Yaoying Wu, for their spiritual support whenever I need it.

Last but not least, my appreciation will also go to all anonymous reviewers who reviewed our publications and helped us improve the quality of the work.

*To my family*

# 1 Introduction

Load characteristics, if not appropriately accounted for in system studies, can result in an inaccurate representation of system behaviour and lead to suboptimal system performance or, in the worst case scenario, to system instability and collapse. Being able to predict the load characteristics (especially the dynamic characteristics) in advance will not only reduce the probability of the potential system stability issues, but also facilitate the development of advanced control algorithms for effective demand side management.

## 1.1 Load Characteristics

Load characteristics, in the context of this thesis, indicate the relationship between the real or reactive power and power system parameters such as voltage and frequency. They can be broadly classified into steady-state and dynamic load characteristics.

### 1.1.1 Representation of Load Characteristics

Load characteristics are represented by appropriate mathematical load models. Load models can correctly represent the change of the real and reactive power as a function of the variation of power system parameters such as time, voltage and frequency. Generally at present time, and considering tightly coupled power systems

dominated by conventional synchronous generators, the frequency dependence of loads can be neglected [1]. Load models can be further classified into static load models and dynamic load models. The former can represent the steady-state load characteristic, which is usually a relationship between the real or reactive power and the voltage. The latter can represent the dynamic load characteristics, which is the time-domain real and reactive power response to a voltage disturbance (and frequency disturbance if it is concerned). The dynamic load characteristic is also referred to as *dynamic response of the demand* (DRD). There are a variety of load models used in industrial practice, and more details will be given in Chapter 2. Mathematical representation or description of load characteristics is commonly referred to as load modelling.

### 1.1.2 Significance of Load Modelling

Load characteristics influence both steady-state and dynamic performance of power systems significantly. Thus, in order to carry out appropriate power system steady-state and dynamic analysis, appropriate representation of load characteristics (i.e. load models) is required. Computer simulations of power system using, among the others, appropriate load models, is one of the most important (if not the only) approach to understand the system dynamic behaviour and provide suggestions on planning and operation.

The importance of accurate load modelling for understanding power system behaviour was highlighted in a number of historical events which lead to component outages, voltage collapses, and system instability. Some of these are described below.

#### 1.1.2.1 Outage at Western Systems Coordinating Council

An outage event took place at Western Systems Coordinating Council (WSCC) on 10 August, 1996 and resulted in break-up of the integrated system into four islands [2]. The event was initially simulated using the standard WSCC dynamic database (i.e. the load model parameters which represent the aggregate load characteristics). The

simulation results however did not match the recorded response; they were too optimistic and failed to predict the system stability status. To match the simulation results with the recorded response, it was necessary to make modifications to several load models.

### 1.1.2.2 Low Frequency Oscillation in Taiwan, China

A low frequency oscillation event during the trunk line outage was observed in Taiwan Power System, China. Different load models such as the composite load model, the dynamic model and the exponential load model were used to analyse the dynamic process during the outage [3]. The oscillation frequencies and damping constants when one line was tripped out and two lines were tripped out are given in Table 1.1 [3]. The result shows that different load models can have different effects on system stability analysis.

Table 1.1. Effect of Load Models on Oscillation Frequency and Damping Constants (adopted from [3])

| Load type | One line tripped out | | Two trunk lines tripped out | |
|---|---|---|---|---|
| | Oscillation frequency(Hz) | Damping constant | Oscillation frequency(Hz) | Damping constant |
| Field-measurements | 0.815 | 0.0126 | 0.80 | -0.01 |
| Composite model | 0.823 | 0.0135 | 0.813 | -0.00973 |
| Dynamic model | 0.785 | 0.00975 | 0.75 | -0.0165 |
| Exponential model | 0.91 | 0.0216 | 0.903 | 0.0125 |

### 1.1.2.3 Voltage Collapse in Argentina

To analyse the voltage collapse in Argentina [1], a static model and a model with 50% induction motor and 50% static load are used to simulate the voltage dynamics. By analysing the PV curves [4] which indicate the relationship between the voltage and the real power, it was found that the static load model was too optimistic to reflect the voltage collapse dynamics, and that more detailed load models were needed.

### 1.1.2.4 Summary of Past Event Simulations

From the results of past event simulations using different load models, it is clear

that the way of representing load characteristics (i.e. selecting load models) can have a significant effect on the analysis result. Therefore, appropriate representation of load characteristics plays an important role in power system studies and controls.

## 1.2 Load Modelling Approaches

As mentioned in Section 1.1.1, the process of establishing a load model (i.e. representing the load characteristic in a mathematical form) is referred to as "load modelling". Load modelling is not a trivial task. On the one hand, a significant number of factors (e.g. the diversity in types and characteristics of the loads, the lack of information on the load structure, and the difficulties in accessing data during the validation process) need to be considered. On the other hand, temporal and spatial load variation also needs to be considered in system behaviour assessment at different times of the day as well as in different geographic locations. Furthermore, aggregate load models at the medium-voltage (MV) and high-voltage (HV) bulk supply buses adopted for power system studies include implicitly the distribution transformers, shunt compensation and the distribution network feeders, often without accounting appropriately for dynamics of operation of tap changing transformers and other voltage regulators that may be deployed at lower voltage levels.

The two predominantly used load modelling approaches by the majority of researchers and utilities are the measurement-based approach and the component-based approach. The directions of these two approaches are totally different; the former is top-down and the latter is bottom-up. Both of the approaches have been illustrated in [1, 5]. The selection of the approach depends on the actual load characteristics and load modelling requirement in specific cases.

### 1.2.1 Measurement-Based Approach

The measurement-based approach (MBA) is a top-down approach that uses

disturbance data collected at substations or feeders to derive load models. In this approach, a load model structure is assumed, and an appropriate post-processing technique, to screen events suitable for load model parameter identification, is selected.

#### 1.2.1.1 Approach Overview

Measurement-based approach is diagrammatically shown as Figure 1.1[1].

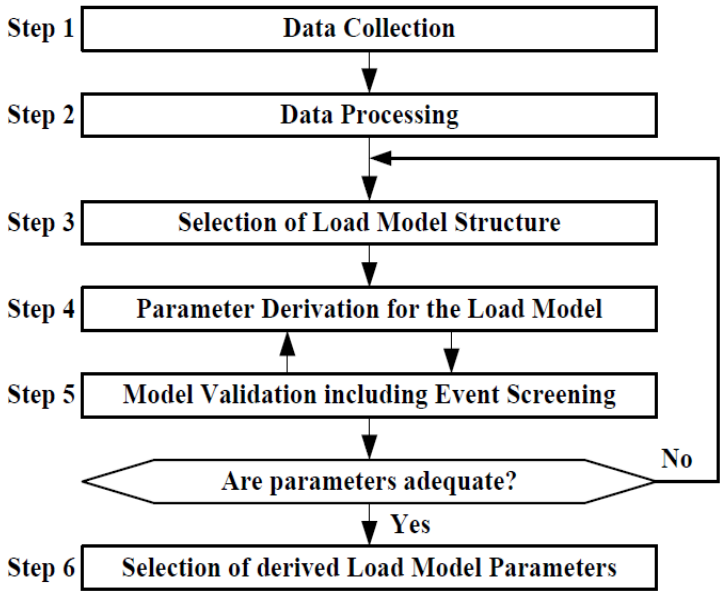


Figure 1.1. Block diagram for the measurement-based approach (adopted from [1])

At the beginning, the disturbance data are collected by data acquisition devices such as power quality monitors or digital fault recorders (DFR). The sampling rate for data acquisition typically ranges from 1Hz to 1kHz, and the selection of the sampling rate depends on the purpose of specific power system studies. For example, for steady-state power flow study, a 1Hz sampling rate is adequate, while for transient stability study, the sampling rate should be at least 100Hz. More details are provided in [1].

If the collected data contains noise, a filter is needed to remove the noise. However, the filtering techniques have to be carefully selected to guarantee that the useful transients in the collected data are retained. The most widely used filtering techniques include moving average (MA), Savitzky-Golay (SG) and robust local

regression (RLR) [6]. In moving average filtering, individual samples are replaced by the average of the neighbouring data points. Savitzky-Golay filter is designed based on the least-square polynomial approximation, and it can achieve a high level of smoothing without significantly distorting the feature of the data. Robust local regression calculates the regression weights for each data point in the selected window. As the noise in the measured data is Gaussian noise in most cases, moving average is usually used for filtering. The mechanism of moving average filtering is provided in Appendix A as an example of the filtering technique. More details about filtering can be found in [6].

After the collected data are filtered, a discrete Fourier Transform (DFT)-based signal processing algorithm (sliding window) is used to convert the three-phase voltage and current into the positive-sequence, negative-sequence and zero-sequence components of the voltage, current, active power and reactive power. After the positive-sequence components are obtained, according to the shape of the load response, a load model structure is selected to postulate the measured data, and an optimisation routine is executed, shown as Equation (1.1) [1]

$$F(\theta) = \sum_{k} (y_m(t_k) - y_c(t_k))^2 \tag{1.1}$$

where $F(\theta)$ is the objective function to be minimised, $\theta$ is the parameter vector to be estimated, $y_m$ is the vector of the measured values, $y_c$ is the vector of the computed values, and $t_k$ is the time step of the simulation. Optimisation techniques, such as least-squares method [5, 7], genetic algorithm [8, 9], support vector machine [10], and simulated annealing [11], can be used for the optimisation process.

With selected optimisation techniques, the value of the parameter that minimises the error between the measured values and the computed values of the function can be obtained by solving Equation (1.1). Then, the derived model is validated using commercially available time-domain simulation tools. The computed real and reactive

power responses are compared with the actual recorded ones. If they do not match reasonably, or the optimised values are not acceptable (i.e. acceptable here means "within typically expected ranges based on past experience"), the process will be repeated until the computed values match the recorded ones and the optimised values are acceptable. After a reasonable set of model parameters is obtained, they are set as the load model parameters and can be used in the above mentioned simulation tools.

The main advantages of the measurement-based approach are as follows:

- ✓ It directly uses recorded load data (after filtering) from the actual system.
- ✓ It can be applied to any type of loads.
- ✓ It can detect temporal changes in connected loads.

The main disadvantages are as follows:

- ✓ Significant data filtering and data conditioning are required
- ✓ The derived model will be inappropriate if data with large disturbances are not available (i.e. large disturbance in the system is generally rare.).
- ✓ It requires detection and removal of natural load changes that are not related to the voltage change.
- ✓ Selected optimisation techniques may result in either no solution or more than one solution for the estimated parameters.
- ✓ The model obtained at one substation based on particular measurements may not be applicable to other substations (unless the load compositions of these substations are very similar) or at different times (as the load compositions might change at different times).

#### 1.2.1.2 Existing Developed Load Models using the Measurement-based Approach

The measurement-based approach has been widely used for load modelling in the past. EPRI's report [12] has provided a comprehensive overview of the measurement-

based approach, including existing monitoring system applied to load modelling, load model structure determination, optimisation techniques for model parameter derivation, and simulation tools during the validation process.

Additionally, in [8], this approach together with genetic algorithm and least square estimation techniques was applied to build a static load model for IEEE 10-machine 39-bus standard test system and Hushitai Power Substation. It was found that the genetic algorithm and the least square estimation techniques could provide different parameters, and that the genetic algorithm was more independent on the initial estimation.

In [13-17], the measurement-based approach was used to identify the steady-state and transient characteristics in the medium-voltage (MV) distribution voltage network. Both of the characteristics have been described mathematically by polynomial and exponential load models. Parameters of a composite model for residential load sector in different day types (weekdays and weekends) of different seasons are given in [18]. In this work, with the uncertainty considered, parameters are presented in the form of Gaussian-fitted probabilistic distribution function. It was concluded that the parameters for different day types and different seasons could be different. Similar to [18], [19] (carried out by the same working group) also used the measurement-based approach to obtain the parameters for different load components in residential load sector. Additionally, [19] extended the work presented in [18] to higher voltage levels based on the responses of 41 transmission system operators and utilities in Europe, which participated in a survey initiated by CIGRE Working Group C4.605. In [20], dynamic load model of some low voltage devices are developed. The results presented in [18-20] are summarised in [1].

The developed load model can be directly applied to power system stability analysis. In [21] , the measurement-based approach is used to develop the composite load model for a Henan Province in Middle China. The developed models are then used

for stability analysis of the area. Similarly, in [9, 14, 22, 23], developed models are validated with real data and used in power system stability analysis.

### 1.2.2 Component-based Approach

Apart from the measurement-based approach, the component-based approach (CBA) has also been widely used to develop composite load model. The component-based approach is a bottom-up methodology that aggregates individual loads according to load mix/classification contribution and load sector contribution. In the component-based approach, the model is developed from: (i) the percentage of different load categories participating in the total demand (or in a specific load sector/class under some circumstances), and (ii) typical characteristics of each load category (usually ZIP + induction motor model).

#### 1.2.2.1 Definitions of Terms used in Load Modelling and Load Classification

Terms such as load category, load controllability, load class and load sector related to component-based load modelling in particular, and to general field of load modelling, need to be defined before describing the approach itself.

##### 1.2.2.1.1 Load Categories

The majority of the currently available electricity consumption statistics and publications usually divide loads into load types based on the specific end-use of electricity (e.g. lighting, heating, cooking etc.). However, it is generally not a suitable categorisation of loads from the point of view of load modelling [1]. For load modelling purposes (especially component-based load modelling purpose), loads are classified into different groups based on the similarity of their electrical characteristics rather than on specific end-use functionality, and such groups are referred to as load categories.

Loads that belong to the same category usually have similar load characteristics

(i.e. *P-V*, *P-t*, *Q-V*, and *Q-t* characteristics), even if the functionalities of the loads are different (i.e. different end-users). For example, in the residential load sector, the residential air conditioner and the refrigerator are different end-users providing different functionalities, but both of them belong to the category "single-phase directly-connected induction motor" and they have similar load characteristics. All types of electrical equipment and devices can be divided into the following general load categories:

- directly-connected motor loads
- drive-controlled motors, or adjustable speed drive (ASD) loads
- resistive loads
- DC power supplies, or switch-mode power supply (SMPS) loads
- energy efficient lighting
- rechargeable/renewable loads (including distribution generators, electric vehicles, solar panel etc.)

It is worth noting that in this context, the distributed generator connected at lower voltage levels is considered as a load at bulk supply points rather than a generator due to assumed small size of individual units and largely distributed nature of LV connected generators [1].

#### 1.2.2.1.2 Controllability of Different Load Categories

In demand side management (DSM), the loads are categorised based not only on load characteristics, but also on their controllability. Based on controllability, loads can be categorised as controllable loads and uncontrollable loads. Controllable load, also referred to as DSM potential load or demand-manageable load, is defined as the load suitable for deferral or shifting [24, 25]. Uncontrollable load therefore refers to the load unsuitable for deferral or shifting. Demand side management potential of the load depends on the load characteristics and the consumer routine. Generally, the

controllable load consists of following types of loads:

- cold loads (e.g. fridge/freezers, air conditioners)
- wet loads (e.g. wash machines and dryers, dish washers)
- heating loads (e.g. heat pump, night storage heating)
- rechargeable/renewable loads (e.g. electric vehicles, solar panel, battery).

Table 1.2. Load Categories and Typical End-users in Each Category

| Main Categories | Sub-categories (if needed in some cases [26, 27]) | Typical End-users | Controllability |
|---|---|---|---|
| Directly-connected induction motor | Single-phase Quadratic Torque (QTIM1) | • Residential/commercial air conditioner<br>• Residential/commercial refrigerator and freezer | Yes |
| | Single-phase Constant Torque (CTIM1) | • Wash machine and dryer<br>• Dish washer | Yes |
| | Three-phase Quadratic Torque (QTIM3) | • Commercial and Industrial pumps or fans | No |
| | Three-phase Constant Torque (CTIM3) | • Heat pump space heating<br>• Heat pump air conditioner<br>• Industrial air conditioner, refrigerator and freezer<br>• Lifts/elevators | Partly Yes |
| Drive-controlled induction motor | Single phase<br>Three phase | • Motor controller<br>• Vector control | Partly Yes |
| Resistive loads | N/A | • Heating (heater /water heater)<br>• Cooking (oven etc.) | Partly Yes |
| Electronic loads | Switch-mode power supply (SMPS) | • PCs/TVs<br>• CD/DVD | No |
| | Rectifier | • Semi-conductor diodes<br>• Thyristor | No |
| Lighting loads | Interior lighting | • General incandescent lights (GIL)<br>• Compact fluorescent lamps (CFL)<br>• Light-emitting diode (LED) | No |
| | Exterior lighting | • High-intensity discharge lighting (HID) | No |
| Rechargeable/Renewable loads | N/A | • Electric vehicles (EVs)<br>• Solar panel<br>• Battery<br>• Micro-CHP | Yes |

According to the categorisation in [1, 26-31], the categorisation of the loads and the controllability of different load categories can be summarised as shown in Table 1.2. For the column that indicates controllability, “Yes” means that the category is controllable, “No” means that the category is uncontrollable, “Partly Yes” means that some of the end-users within the category are controllable while some are not.

If needed, in some cases at higher voltage level, network components (e.g. shunt capacitance/reactance, transformer, underground cables or transmission lines) can also be regarded as a category of “load”, named “network component”.

#### 1.2.2.1.3 Load Sectors

Load sector, also referred to as load class or class of customers, is generally defined as an aggregation or collection of loads from different load categories, representing the typical structure and composition of electrical devices and equipment found in a specific end-use application, where similar activities and tasks are performed. This usually results in inherent similarities in characteristics and patterns of active and reactive power demands of end-users from the same load sector, allowing the use of the same or similar aggregate load models for the representation of their aggregate demands. It should be noted that one load sector may be further divided into several sub-sectors, and further details are in [1].

### 1.2.2.2 Approach Overview

The overall approach to the component-based load modelling is described by Figure 1.2 [1]. It should start from “load characteristics” in Figure 1.2, where the load models for individual load categories are selected and derived according to the measurement-based approach from either field tests or laboratory measurements. After the models of individual categories are developed, the information of the load mix percentage should be collected for specific load sectors or for the overall demand in order to aggregate the individual load categories with appropriate weights. This

approach can be applied to either different load classes (i.e. residential, commercial, industrial etc.) or directly to the aggregate demand at bulk supply point (BSP).

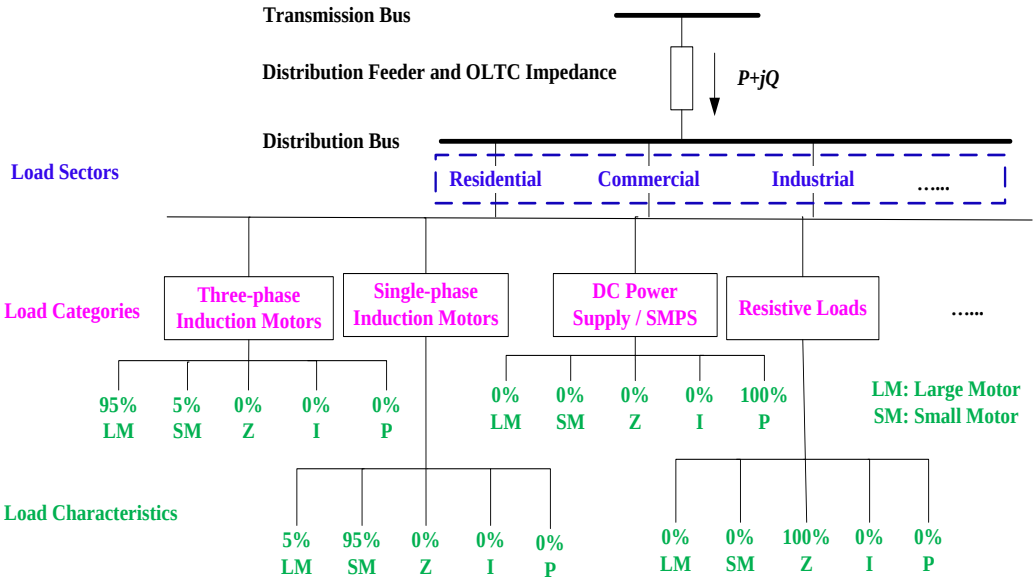


Figure 1.2. Component-based approach (adopted from [1])

The aggregate load can be represented by weighed average of individual load, shown as

$$P_{Agg} = \sum_{i=1}^{N} w_i P_i \text{ and } Q_{Agg} = \sum_{i=1}^{N} w_i Q_i \tag{1.2}$$

where $P_{Agg}$ and $Q_{Agg}$ are aggregate load real and reactive power consumption in per unit, $P_i$ and $Q_i$ are actual real and reactive power consumption of load category $i$ in per unit, $w_i$ is the weighting factor for load category $i$ which satisfies $\sum_{i=1}^{N} w_i = 1$, $N$ is the total number of load categories. The weighting factors are the percentage of different load categories participating in a given load sector or the BSP. In the studies where high accuracy is required, the effect of the network components (e.g. the impedance of the feeders and the transformers, etc.) on proposed load models should be considered as well.

If the percentage of different load categories and the load characteristics of individual load categories are known, the component-based approach can be easily used

to develop composite load model.

The main advantages of the component-based approach are as follows:

- ✓ It can be easily used to develop composite load models with different load mix.
- ✓ It does not require field measurements if the percentage of each load category in the load mix and the load characteristics of each load category are known or could be estimated in advance.
- ✓ It is adaptable to different systems and operating conditions.
- ✓ It allows flexibility for potential demand control and facilitates system studies related to establishment of system performance sensitivity to load composition changes.

The main disadvantages are as follows:

- ✓ The percentage of different load categories can change over time. It is a challenge to acquire the information on the percentage of different load categories at different times of the day as well as different geographical locations.
- ✓ Even if the percentage of each load category in load mix is known, the parameters of the categories (or the types of end-users in the categories) could be different under different operating conditions
- ✓ If a new type of load (which does not belong to any previously defined load categories) is connected, errors in load model parameter identification could be induced.

#### 1.2.2.3 Existing Developed Load Models using the Component-based Approach

In [1, 32], the component-based approach is illustrated to develop a static aggregated load model for residential load sector in a general UK urban area. Both the

exponential model and the ZIP model (more details provided in Chapter 2 and [1]) are developed for power and reactive power. The voltage exponents in the exponential model and the participation coefficients of constant impedance, constant current and constant power load in the ZIP model vary with time, and they are presented by a daily parameter curve which indicates the variation of the parameters against time. This means that the steady-state load characteristics can vary with time.

Apart from the residential sector, [33] provides an aggregated model for the industrial facilities. It classifies different industrial processes into groups, and then represents the industrial processes modelled by different induction motors in each group using an equivalent aggregate induction motor, and finally connects each equivalent aggregate induction motor to the network in parallel with a static load model that can appropriately represent the static part of the aggregate load.

An aggregated model for mixed residential and commercial load sector under severe disturbances (i.e. up to 50% voltage variations) is developed in [34]. Using EMTP (Electromagnetic Transients Program) simulations, the resulting composite model has been validated with field measurements which were recorded during a single-line-to-ground fault.

Western Electricity Coordinating Council (WECC) has also made considerable contribution to the component-based approach. In [35], daily demand data for different load types on different days during different seasons in Californian commercial sector are measured and recorded. An updated composite load model is proposed by WECC in [31] for more accurate load modelling, and the corresponding detailed data of the classification and the percentage of different types of loads used for modelling are provided in [35]. The construction graph is shown in Figure 1.3 [1, 31]. It divides the model into parts shown as below:

- Electronic part

- Static part
- Motor A: three-phase motors driving constant torque loads, such as commercial/industrial A/C compressors
- Motor B: three-phase motors driving speed-squared loads with high inertia, such as fans
- Motor C: three-phase motors driving speed-squared loads with low inertia, such as pumps
- Motor D: single-phase motors driving constant torque loads, such as residential A/C compressors, refrigerators, and heat pumps)

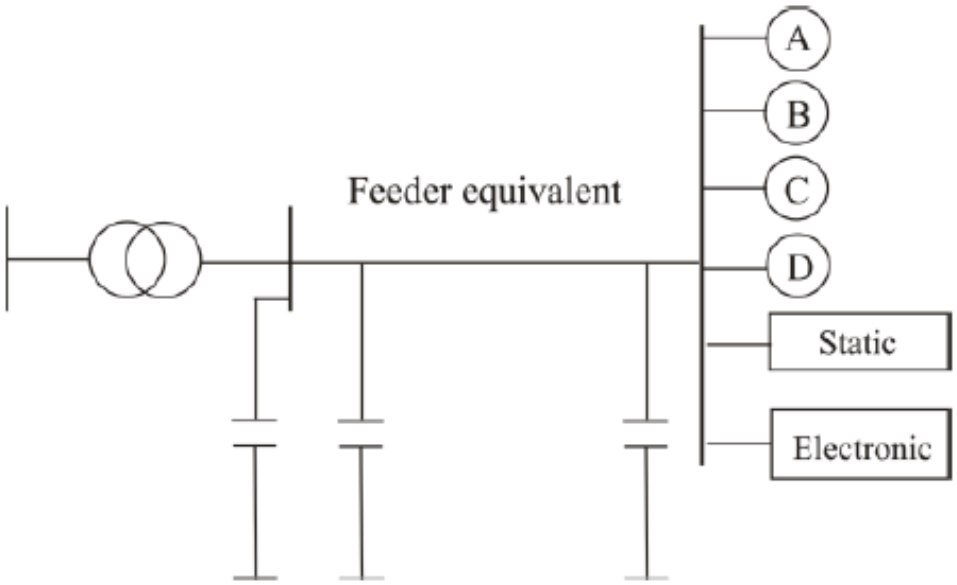


Figure 1.3. WECC composite load model (adopted from [1, 31])

For higher voltage level (over 33kV), apart from the parts mentioned above, following components should also be included:

- Feeder (underground cables or overhead lines)
- Substation shunt
- Transformers

(Note: the bulk supply point voltage level in this study is considered to be up to 11kV.)

## 1.3 New Challenges and Motivation

The past event simulations using different load models have proved that accurate load modelling plays an important role in system analysis. Load models, which are used to represent load characteristics, can be built from either the measured response to a disturbance, or the derived or estimated load composition data from customer surveys conducted in the past. However, the role of establishing load models based on historical measured data or historical derived load composition data is limited, as the data only represent situations in the past. It may provide valuable information for long-term planning, but it is not able to appropriately address issues related to short-term prediction, dispatch and control.

One of the most popular activities related to short-term prediction, dispatch and control, is load shifting. This is one of the most widely used short-term demand side management techniques. Load shifting shifts the demand at peak times to off-peak times to reduce the capacity of the traditional generators as well as the expenditure, and it can be achieved by regulating the electricity price, incentive system and policies. During the process, it is necessary to fulfil the following criteria:

- ✓ The demand at peak times are shifted to off-peak times;
- ✓ The total cost (of energy consumption) after load shifting is lower than or equal to that before load shifting;
- ✓ The dynamic load characteristics remain the same (or are satisfactory) before and after load shifting so that the whole power system does not suffer from stability issues.

The dynamic load characteristics in the future are indispensable, but they cannot be obtained from measurements in the field test. Therefore, for load shifting (and the majority of other demand side management schemes), it is necessary to predict not only the total demand, but also the possible compositions of the load as well as their potential

for demand side management (DSM) and dynamic responses of demand before and after DSM, in order to determine the amount of different types of customers to be shifted to meet all the criteria mentioned above.

## 1.4 Review of Past Work

Although the topic about prediction of load compositions and the dynamic response of the demand based only on substation RMS measurements (with 30min resolution) has not been discussed straightforwardly in past work, significant contributions have been made to load disaggregation/decomposition, load forecasting, load shifting and load modelling approaches (i.e. load modelling approaches have been reviewed in Section 1.2, and load models will be described in detail in Chapter 2), and these contributions significantly facilitate the development of the ideas in this research.

### 1.4.1 Load disaggregation/decomposition

Load disaggregation, alternatively referred to as load decomposition, works out the contribution of different end-users or load categories to the total demand. In Chapter 2 of [1], an example of typical load component in Japan is provided from customer surveys. It provides the percentage of motors, lamps, heaters, and other devices in residential, commercial and industrial load sectors and the composition rate under the overall demand.

Decomposed daily loading curves (DDLC), defined as "detailed daily loading curves indicating participation of different types of loads at different times of day for different load sectors during different seasons" [1], is usually used to represent the contribution of different end-users or load categories at different times during the day. An example of a decomposed daily loading curve which represents the percentage of different end-users on a winter's day in the UK residential sector is given as Figure 1.4. It can be clearly seen from the figure that at different times of the day, the contributions

of different end-user types are different. For example, the consumer electronics/ICT load contributes about 1/3 to the total demand at peak time (i.e. 20:30) but only contributes less than 1/10 at 12:00.

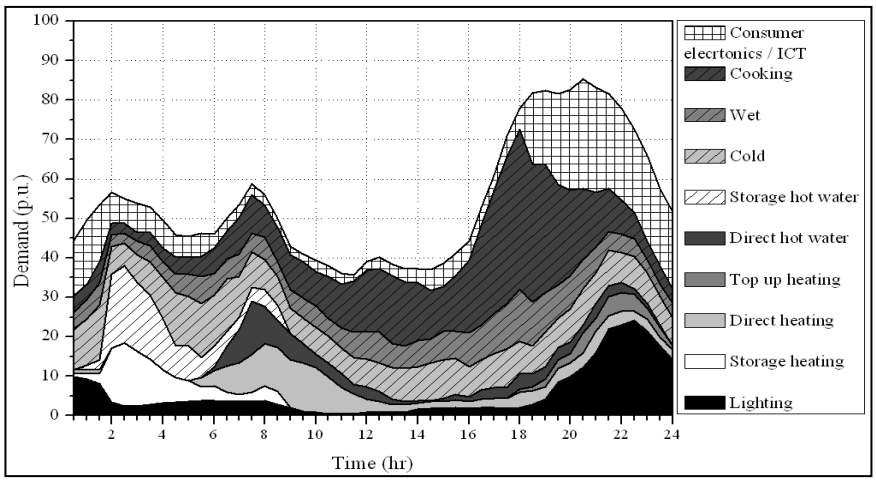


Figure 1.4. An example of decomposed daily loading curves for UK residential load sector in winter (adopted from [1, 24, 32])

As mentioned in Section 1.2.2, the derivation of the contribution of load compositions to the total demand (or decomposed daily loading curve as shown in Figure 1.4) is a challenging task, and great attention has been paid to load disaggregation over past decades. There are two approaches commonly used for load disaggregation: intrusive load monitoring and non-intrusive load monitoring.

#### 1.4.1.1 Intrusive Load Monitoring

Currently, intrusive load monitoring is a widely used approach for load disaggregation/decomposition. It relies on the customer record about appliance operation [36], energy data and billing data [37], and more recently, smart meter data [38]. In intrusive load monitoring, an intermediate monitoring device between the socket and the appliance is installed to record its operation. The customer is additionally required to record the consumption of different appliances at regular time intervals (i.e. every 30 min or 60 min).

Apart from decomposed daily loading curves obtained from intrusive load

monitoring for different load types during the winter in the UK given in Figure 1.4 [1, 24, 32], decomposed daily loading curves of different load sectors (i.e. defined in Section 1.2.2.1) and overall demand for UK in 2010 are also investigated [37]. In addition, a household electricity consumption survey [36] was conducted from 2011 to 2012 by Intertek, which provides the daily consumption of different end-users in different types of households. Simultaneously, the controllability of different load categories in [1, 24, 32, 36] has also been investigated, although the exact percentage of the controllable and uncontrollable parts in load categories is not given. This especially applies to load categories that may contain both controllable and uncontrollable parts (i.e. resistive load includes cooking load which is uncontrollable and heating load which is controllable.).

Furthermore, decomposed daily loading curves are derived for the residential sector in Germany during the winter [39], for all different sectors of California during the summer [40], for different types of households in the Swedish residential sector [41], and for the residential sector in different parts of Europe [38, 42]. Report [42] provides the decomposed daily loading curves of residential load sectors in terms of household appliances for some European countries located in different areas of Europe, such as Denmark, Greece, Italy and Portugal. Report [38] provides decomposed daily loading curves in terms of household appliances during the summer for five areas in Europe including:

- Region A: Southern Europe (Spain, Italy etc.)
- Region B: Scandinavia
- Region C: New Member States
- Region D: Germany/Austria
- Region E: United Kingdom

Both [42] and [38] prove that for different locations, the load subdivision is probably

different due to the variation in climate and lifestyle. Reports [38] and [39] illustrate that even for the same load sector of the same country, decomposed daily loading curves can be different due to the change of seasons.

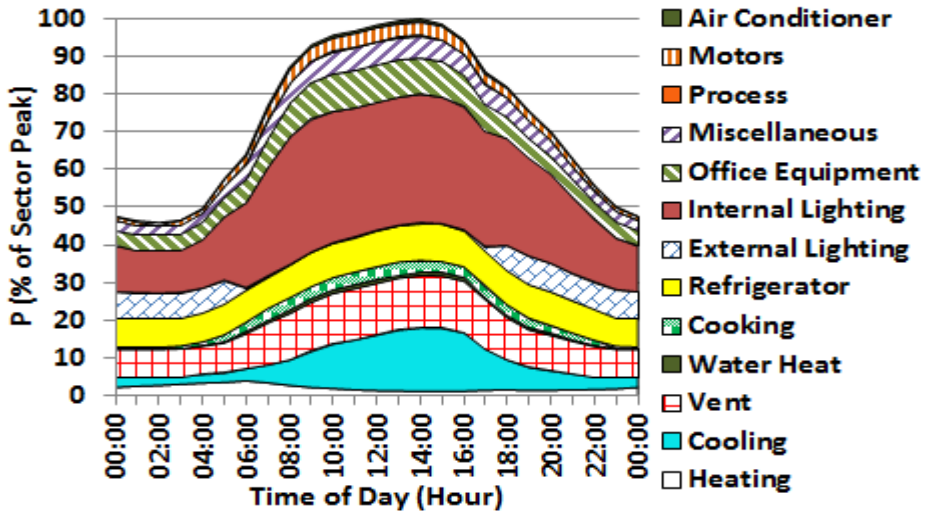


Figure 1.5. DDLC for commercial load sector in California (plotted using the consumption data adopted from [35])

Alternatively, if the measured consumption data of different end-user types are available, decomposed daily loading curves can also be derived directly from those data. Figure 1.5 shows the decomposed daily loading curves for the commercial load sector in California derived using intrusive load monitoring from the measured consumption data [35]. The peak demand is normalised to 100% for ease of comparison. More examples of decomposed daily loading curves derived by the intrusive load monitoring approach for different regions in the world are provided in Appendix B.

The intrusive load monitoring approach can work out the consumption of different end-user types at different times of a normal day. It may provide valuable information for long-term power system planning, but it is inconvenient and expensive for widespread deployment to short-term prediction and control due to a variety of factors, such as the limitation in accessing all appliances in all types of buildings, the vagueness of the customer diary and customer awareness of privacy. And even if the above drawbacks are overcome, the load composition can change as customer behaviour

changes and the composition derived in the past may not be the same as the composition at present or in the future. In other words, to know the load composition at any given time, the meter should keep tracking the households at all times.

#### 1.4.1.2 Non-intrusive Load Monitoring

Intrusive load monitoring (or customer survey) is straightforward, but it is time-consuming and expensive to carry out surveys or collect consumption data of different end-user types. As a result, an approach named "non-intrusive load monitoring" is applied to identify the contribution of different load categories, using high-resolution load signature measurement data such as current waveform, dynamic load response to a voltage step, instantaneous power or real power change [43].

##### 1.4.1.2.1 Load Signature used for Load Disaggregation

A load signature is defined as the electrical behaviour of an individual appliance or a group of appliances when it is under operation [43]. The most typically used load signature includes current waveform (CW), real and reactive power (PQ), harmonics (HAR), instantaneous admittance waveform (IAW), instantaneous power waveform (IPW), eigenvalues (EIG) and switching transient waveform (STW) [43, 44]. Additionally, if the disturbance data are available, a dynamic response of the demand (DRD) can also be used as a load signature. There are two types of load signatures: snapshot form and delta form [43], and both of them can be used for load disaggregation.

###### 1.4.1.2.1.1 Snapshot Form Load Signature

Snapshot form load signatures are the instantaneous snapshots of the load behaviours taken at any fixed time intervals [43]. They are usually micro-level signatures (i.e. load signatures collected with sampling rate equal to or faster than 1 sample per cycle [43]) and can be load signatures of either individual loads or the aggregate load. They can be applied to disaggregation not only for households, but also

for substations, as long as the high-resolution (faster than 1Hz) total load signatures such as current waveforms, harmonics or dynamic responses are available.

Paper [45] derives load compositions for end-users in the residential sector by using integer programming as the optimisation technique. Paper [46] derives load compositions for a utility side by using the conjugate gradient method and Kalman filter as optimisation techniques. Both [45] and [46] adopted current waveforms as snapshot form load signatures. In [47-49], a pattern recognition method is used to identify load compositions, with harmonics of the current waveform adopted as load signatures. Paper [50] uses a quadratic program which has a similar mechanism to the optimisation and pattern recognition method, with daily load curves adopted as load signatures. In [51], power consumptions at different nodes are used as load signatures, and an iterative least square algorithm is used as the optimisation technique to derive the load and generation composition on each node of a distribution grid.

Dynamic responses of the demand have also been used as load signatures for load disaggregation in past work. In [52], the dynamic response of the individual loads and the aggregate loads are used as the snapshot form load signatures to estimate the composition of the loads for the residential and industrial load sector, and fuzzy logic is adopted as the optimisation tool. To validate the algorithm in the reality where the load composition data are usually not available, the computed compositions are used with the dynamic response of the individual loads to calculate the dynamic response of the overall demand (i.e. the component-based load modelling approach), and the calculated response is compared with the measured one to assess the accuracy of the disaggregation.

#### 1.4.1.2.1.2 Delta Form Load Signature

A delta form load signature makes use of the difference between two consecutive snapshot form signatures [43]. If the difference is small enough, it can be assumed that

only one appliance is switched on or switched off within the difference. In other words, it behaves more like a single appliance switching on or off rather than a composite load variation. Delta form load signatures are usually macro-level signatures (i.e. load signatures collected with sampling rate slower than 1 sample per cycle [43]), and they are widely used in non-intrusive load monitoring for a single house. It uses the change of the power consumption to detect what devices are switched on or switched off (i.e. usually the change is in the order of W to kW).

The research on load disaggregation by non-intrusive load monitoring using delta form load signature dates back to Hart's work [53] published in the early 1990s, where the approach was tested on a single household. Similar works have been presented in [54-58], where a more advanced computational algorithm is used to optimise the model used in the approach and therefore increase the accuracy of the disaggregation.

#### 1.4.1.2.2 Derivation of Load Composition

The derivation of load composition is generally an optimisation problem formulated as

$$Minimising \ \ f = \left( LS_{Agg} - \sum_{i=1}^{N} w_i LS_i \right)^2 \tag{1.3}$$

where $f$ denotes the objective function, $LS_{Agg}$ is the load signature of the total demand, $LS_i$ is the load signature of individual load categories or end-users with Index $i$, $N$ is the number of load categories, and $w_i$ is the contribution of individual load categories or end-users with Index $i$ to the total demand. The percentages of main load categories available are worked out by using either optimisation (least square method, genetic algorithm, fuzzy logics etc.) or pattern recognition (artificial neural networks etc.) [44]. The essential mechanism of optimisation and pattern recognition is the same: to minimise the difference between the derived quantities and actual (or measured) quantities. The derived contributions of different load categories to the total demand are

then compared with the corresponding measured consumption data if they are available; if not, the computed load characteristic is compared with the measured one (as mentioned in Section 1.4.1.2.1.1).

#### 1.4.1.2.3 Strengths and Drawbacks of Past Approaches

There is no doubt that non-intrusive load monitoring is more convenient and faster than intrusive load monitoring, as long as the measurements of load signatures with the required sampling rate are available. However, its application to most cases in reality is not wider than intrusive load monitoring.

One reason is that it is difficult to have access to the measurements with the requested sampling rate. Although the non-intrusive load monitoring approach for load disaggregation using snapshot form load signature is applicable to both households and substations, measurements with a sampling rate higher than 1Hz are generally not available under most circumstances, especially in substations. This impedes the frequent and widespread application of the approach.

Another reason is that the approach may only be applicable under limited circumstances. For example, the non-intrusive load monitoring approach for load disaggregation using delta form load signature is only suitable over a small scale area with kW consumption where the W to kW change is detectable. For a large customer or a bulk supply point with MW to GW consumption, such a change is unlikely to be detected by the substation devices. Even if the power change is detected, it is difficult to distinguish the switching of different appliances of the same rating, or to distinguish between simultaneous switching of multiple devices and the switching of a single device with the equivalent rating.

### 1.4.2 Load Forecasting

Accurate prediction of future loads can provide a great saving potential for power system planning and investment activities. Since the 1970s, significant effort has been

made in load forecasting. In [59], which is a comprehensive literature review and summary of load forecasting methods, all methods for load forecasting reviewed are classified into nine approaches, including: (i) multiple regression [60, 61], (ii) exponential smoothing [62, 63], (iii) iterative reweighted least squares [64], (iv) adaptive load forecasting [65, 66], (v) stochastic time series [67-70], (vi) genetic algorithm [71], (vii) fuzzy logic [72-75], (viii) artificial neural network [76-80], and (ix) knowledge-based expert systems [81-83]. The strengths and drawbacks of the nine approaches are also summarised. The multiple regression approach is the basic mathematical model for load forecasting, and the most widely used approaches in recently published research include stochastic time series, fuzzy logic and the artificial intelligence based approach. Due to the superiority of the fuzzy logic and artificial intelligence based approach over other approaches, they have attracted the most attention in recent research.

#### 1.4.2.1 Multiple Regression

The multiple regression approach uses least square estimation or iterative reweighted least square to calculate the statistical relationship between the total demand and the weather as well as the stochastic components. At the beginning, an appropriate model should be selected, and then large quantities of historical data are needed to obtain the statistical relationship. The mathematical theory of multiple regression is introduced in [84].

A multiple regression model for short-term load forecasting is proposed in [60]. Using this model, the forecast can be calculated directly from historical data as a local average of observed past loads and the specific weights on the loads defined by a multivariate product kernel, and the mean absolute percentage error in the case presented in [60] is 2.78%. Additionally, a regression based load forecasting method for Eastern Saudi Arabia is presented in [61], and in this case, the mean absolute percentage

error is about 10%.

Although multiple regression approach is the basic load forecasting approach, it is rarely used in current research due to its time-costing characteristics, complexity in implementation, weak robustness and relatively low accuracy in most cases [59].

#### 1.4.2.2 Stochastic Time Series

The stochastic time series approach is one of the most popular approaches for short-term load forecasting, and it has been used for a long time. It is very likely that this approach will still be used in future research as it can be applied to all types of load forecasting, including load forecasting with wind generators, long-term load forecasting which provides information for long-term power system planning, etc. There are three main types of stochastic time series model: autoregressive (AR) model, autoregressive moving-average (ARMA) model and autoregressive integrated moving-average (ARIMA) model.

##### 1.4.2.2.1 Autoregressive Model

An autoregressive model is used if the load is assumed to be a linear combination of the previous loads [59]. A short-term load forecasting approach using an autoregressive model with optimal threshold stratification is presented in [69]. This model can derive the minimum number of parameters required for the representation of the stochastic components, which removes the subjective judgement and therefore improves the accuracy of the prediction.

##### 1.4.2.2.2 Autoregressive Moving-average Model

An autoregressive moving-average (ARMA) model is used if the current value of the load time series is expressed linearly in terms of its values at previous periods and the previous values of a white noise. The non-linear regression approach such as the maximum-likelihood approach can be used to identify the parameters of an ARMA

model. In [67], a new time-temperature methodology for load forecasting is presented. In this methodology, the time series are decomposed into the deterministic component and the stochastic component, and the latter is determined by an ARMA model. Additionally, an adaptive ARMA model for short-term load forecasting is presented in [70], where the available forecasting errors are used to update the model and the maximum percentage error is 6.98% in contrast with the traditional ARMA model which, in the same operating environment, produces a forecasting error of 13.58%.

#### 1.4.2.2.3 Autoregressive Integrated Moving-average Model

An autoregressive integrated moving-average (ARIMA) model is used if the load variation process is considered as a non-stationary process. In this model, before the forecasting starts, the time series should be transformed into the stationary form [59]. An ARIMA short term load forecasting approach is proposed in [85], and the mean absolute percentage error ranges from 6% to 9% depending on the season.

### 1.4.2.3 Fuzzy Logic

The fuzzy logic approach is also widely applied to load forecasting, and it works in two stages: the training stage and the forecasting stage. In the training stage, large quantities of historical load data are used to train a fuzzy logic based forecaster to generate the pattern database and the fuzzy rule. After training and once validated, the trained forecaster is used for on-line prediction. If the most probable matching pattern with the highest possibility can be found, then an output pattern will be generated through a centroid defuzzifier. More details about fuzzy logic can be found in [86].

A fuzzy logic based model for short-term load forecasting is presented in [74], and Tabu search is used to optimise the fuzzy model structure. The mean square error in this case ranges from 4% to 7%. Similarly, Paper [75] proposes a fuzzy logic based methodology, with a root mean square error of about 4%.

#### 1.4.2.4 Artificial Intelligence based Approach

Due to the complexity and relatively low prediction accuracy of regression models and time-series stochastic models, and subjectivity about the selection of the membership function of fuzzy logic based models, approaches based on artificial intelligence (AI) techniques are gaining more and more attention. They are now rapidly developing because of their high accuracy. Some of the AI based methodologies for load forecasting have already been adopted and widely used by the industry, and the two most widely used AI techniques are the artificial neural network (ANN) and adaptive neuro-based fuzzy inference system (ANFIS) [87].

##### 1.4.2.4.1 Artificial Neural Network

An artificial neural network (ANN) is an interconnected assembly of simple processing elements, units or nodes, whose functionality is inspired by *animal central nervous systems*. The processing ability of the network is stored in the inter-unit connection weights, obtained by a process of adaptation to a set of training patterns. It is now widely applied to several areas such as prediction, curve fitting, optimisation and clustering etc., because of its capability of learning. Similar to fuzzy logics, it also works in the training stage and the on-line prediction stage.

The practicality of Artificial-Neural-Network-based (ANN-based) methodologies for load forecasting has been verified by the fact that some of them have been adopted by industry; for example, ANN-based methodology for load forecasting developed in EPRI project ANNSTLF (Artificial Neural Network Short-term Load Forecasting, [88-90]) is one of the most widely used methodologies for load forecasting. The mean absolute square error (MAPE) of the prediction ranges from 2% to 5%, depending on different locations and utilities. In [79], an ANN approach trained with generalised delta rules (DR) is adopted to forecast demand, and the mean absolute percentage errors are from 1% to 4%. In [80], a wavelet neural network is adopted for short-term load

forecasting of a commercial load, in which different training algorithms are adopted and compared by processing time. The mean absolute percentage errors in this case range from 0% to 5%. A neural network load forecasting approach with weather ensemble predictions for one day to several days ahead of load forecasting is proposed in [91]. The distribution of the load scenarios is used as an input for the estimation of the uncertainty in the forecasting. The mean absolute percentage error in this case ranges from 1.5% to 3%, depending on the number of days in advance that is required to forecast the load.

Apart from short-term load forecasting for real power prediction, the ANN-based approach can also be used for the prediction of reactive power (although it has been applied to only a few cases). A similar ANN-based approach for reactive power prediction is presented in [92, 93], although the reactive power prediction does not acquire the same interest as real power prediction. The mean absolute percentage error ranges from 8% to 17%, depending on the utilities and how long ahead the prediction is (e.g. hours ahead, one day ahead etc.).

#### 1.4.2.4.2 Adaptive Neuro-based Fuzzy Inference System

The adaptive neuro-based fuzzy inference system (ANFIS) is a Sugeno fuzzy inference system (FIS) [94], whose input membership functions are adjusted by either a backpropagation algorithm or hybrid algorithm (a combination of backpropagation and least squares) and output membership functions are either the constant or linear combination of inputs. It is a hybrid approach that incorporates the artificial neural network and fuzzy logic, and is also one of the latest methodologies applied to load forecasting.

Load prediction based on ANFIS with Gaussian-shaped input membership function, constant output function and hybrid training algorithm is presented and compared with the linear regression approach in [95], and a multi ANFIS for short-term

load forecasting for different seasons and different types of days (working days, holidays) is presented in [96]. The mean absolute percentage errors in both cases are up to 3%, in contrast with those for the linear regression approach which, in the same operating environment, are up to 6%. Additionally, in [97], an ANFIS predictor is developed for medium term load forecasting, and the mean absolute percentage error in this case is lower than 2%.

#### 1.4.2.5 Comparison of Different Approaches for Load Forecasting

The comparison of different approaches is as equally important as classifying load forecasting approaches. With comparison, appropriate approaches can be selected for different load forecasting case studies. In most of the past work, when an approach was developed, it was usually compared with other approaches which have been developed before.

One of the most comprehensive and the earliest comparisons is made in [98], where fourteen load forecasting methods were uniformly compared. This work was upgraded in [99], where five extra methods were added. All nineteen methods belong to either the category of regression or the category of time series and have different levels of superiority. The comparisons were tested on urban, rural and agrarian areas. The result shows that in general, the more advanced techniques are justified in urban areas but are usually costly for rural and agrarian areas. Apart from accuracy and cost, growth rates, planning difficulty, requirement to detail and other criteria vary from utility to utility. Therefore, the selection of the methodologies should be made depending on a large number of criteria.

Additionally, different approaches among the nine mentioned at the beginning of Section 1.4.2 were also compared in past work. In [100], three techniques including fuzzy logic, artificial neural network and autoregressive models are compared, and it concludes that fuzzy logic and artificial neural network perform much better than

autoregressive models. In [60], a regression model is compared with the artificial neural network approach in the same operating environment, and the mean absolute percentage error is about 2.78% for the regression model and 2.64% for the artificial neural network, which means that at least in this case, artificial neural network performs slightly better than the regression model.

Based on the comparative result in many publications, paper [59] summarises that some traditional approaches such as regression, iterative reweighted least squares and adaptive load forecasting, seem to be out of favour due to their limited accuracy, and that large amount of research, effort and interest are focussed on fuzzy logic and particularly, artificial neural network (ANN). A clear trend can also be seen towards hybrid methods, which combine two or more of these techniques (e.g. combination of fuzzy logic with ANN [94, 101-105] or a combination of ARIMA with ANFIS [106]) to take advantage of their individual best features. A comparison between one of the typical approaches and one of the hybrid approaches (e.g. ANN and ANFIS) or between different hybrid approaches would be of interest in future studies as it is rare (if not missing) in presently available body of publically available literature.

### 1.4.3 Load Shifting and Shaping of Load Curves

Load shifting, as a potential routine for the shaping of dynamic responses of the demand as well as an effective way for shaping of load curves, has been studied in the past and has become one of the most commonly used techniques for demand side management (DSM) [107]. Load shifting brings benefits to both the environment and economy. On the one hand, as the demand at peak times is shifted to off-peak times, it helps balance the generation and the demand and helps reduce the capacity of the traditional generators and the emission of greenhouse gas; on the other hand, as the price of the electricity at peak times is generally higher than that at off-peak times, load shifting also helps customers and utilities to reduce expenditure on electricity bills.

Great contributions have been made to load shifting since the early 1980s. In [108], a model predictive control strategy with binary integer programing for load shifting is proposed for a water pumping scheme. The results of [108] show that 50%-60% of the energy cost and 50%-90% of the maximum demand cost can be saved depending on different control models. In [107], a day-ahead DSM strategy based on load shifting is proposed to reduce the peak load in a smart grid. Heuristic optimisation is used to minimise the difference between the shaped load curve and the already defined one. The result shows that it achieves up to 10% of cost savings while reducing the peak demand by up to 20%. In [24] where the influence of load shifting on overall performance is investigated, based on the controllability of different load categories, about 40% of the wet loads during the peak time are controlled, and it is concluded that even small changes in load mix can have an impact on the steady-state aggregate load characteristics, system performance and quality of supply.

Similar research that considers smart energy appliances, storage, fuel-cell hybrid vehicles or flow batteries is presented in [25, 109-113]. For example, in load shifting presented in [109], smart energy appliances (SEA) are shifted from peak times to off-peak times and intelligent smart energy services (SES) are provided to achieve economic power regulation in an area with a high penetration of SEA in the Netherlands. It shows that a combination of SEA and SES creates an efficient and flexible power regulation with a capacity of 700MW and an equivalent of 5GWh storage, and the investment on SES is expected to recover within 7 years.

### 1.4.4 Summary of Past Work

The review of past work has identified several areas that need to be addressed, and they are summarised as follows.

- Although intrusive load monitoring and non-intrusive load monitoring can provide a valuable analysis of load compositions, they both have drawbacks

which significantly impede the estimation of load compositions in an effective and practical way. A new disaggregation approach therefore needs to be developed to overcome the drawbacks but simultaneously maintain the strengths.

- Identification of load controllability is usually separated from load disaggregation, and the contribution of the controllable and uncontrollable parts has not been investigated for categories which contain both. Therefore, the developed disaggregation approach should simultaneously address this issue to facilitate more effective demand side management.
- Although comparisons among series of load forecasting algorithms have been made, comparative results between the latest adaptive neuro-based fuzzy inference system (ANFIS) and artificial neural network (ANN, one of the most widely used and the most efficient approach) in load forecasting have not been investigated and are of interest.
- As concluded in [24], even small changes in the load mix can have an impact on the aggregate load characteristics; the effect of load shifting on the dynamic response of the demand to a voltage disturbance requires investigation. Previous research in this area mainly focuses on minimising the cost and the difference between peak and off-peak demand, but the effect on dynamic load characteristics has not been studied thoroughly (if at all).
- To enable demand side management planning in advance, an approach for prediction of load compositions and their controllability in future (at least 24h in advance) needs to be developed. Previous research mainly used the consumption data obtained from either the surveys conducted in the past or the data measured in the past, which cannot represent the cases in the future.
- To exclude potential stability issues in power systems, an approach for

prediction of dynamic responses of the demand without deriving it based on the field test should be developed. Previous research derived load characteristics based on either measured disturbance data, which is difficult to obtain, or the load composition consumption data obtained in the past, which is not able to properly represent the load participations and load characteristics at present or in the future.

## 1.5 Aims and Objectives of the Thesis

### 1.5.1 Aims of the Research

The contribution of different load categories to the total demand as well as the dynamic load characteristics of the aggregate loads are highly important for power system stability studies and demand side management but are also difficult to obtain, estimate or predict. Therefore, the main aim of the thesis is to develop a framework for day-ahead prediction of the percentage of different load categories in total demand mix and the dynamic response of the demand (DRD). (Note: there is no doubt that the same approach also applies to prediction with time frame longer than 24h, e.g., 48h or 72h. However, with longer prediction horizon, higher prediction errors could be induced. This thesis focusses on the one-day-ahead prediction, and the term “day-ahead” refers to one-day-ahead.)

### 1.5.2 The Overview of the Research

Based on the aim, an overview of the research presented in this thesis is developed and illustrated as a block diagram shown as Figure 1.6. The block diagram includes five procedures: (i) total demand forecasting; (ii) actual value to per-unit value transformation; (iii) load disaggregation; (iv) prediction and validation of the dynamic response of the demand (DRD); (v) demand side management (DSM) planning and shaping of the dynamic response of the demand (DRD).

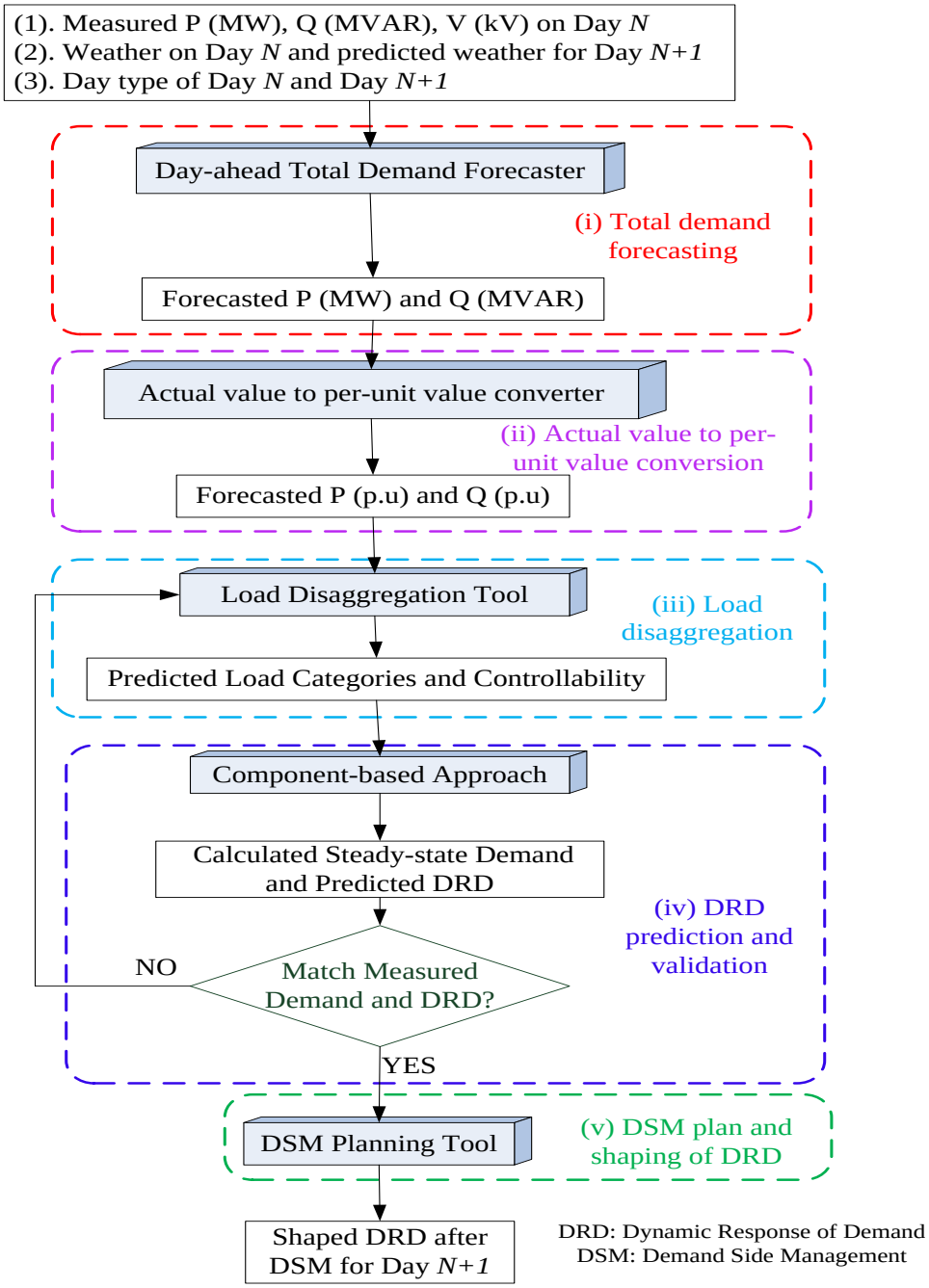


Figure 1.6. Overview block diagram of the research

Procedure (i) forecasts day-ahead total *P* and *Q* demand. Procedure (ii) transforms measured or predicted demand in MW or MVAR into per-unit value so that they can be properly used in the disaggregation approach. Procedure (iii) disaggregates the total demand into different load categories and predicts the contribution and controllability of different load categories participating in the total demand. Procedure (iv) predicts DRD

via the component-based load modelling approach [1] and validates the prediction of both load compositions and DRDs using field measurement data collected at substations in the real distribution network in the UK. Procedure (v) makes plans for load shifting according to the output from (iii) and (iv) and changes the shape of DRD to a preferable one.

As can be seen from Figure 1.6, the inputs to the whole developed approach are: (1) measured RMS real and reactive powers and voltages on Day *N* with a resolution of 30min; (2) weather data (i.e. temperature, humidity and wind speed) on Day *N* and corresponding forecasted weather data for Day *N+1* with a resolution of 30min; (3) day type (i.e. working days, holidays) of Day *N* and Day *N+1*.

The outputs are the predicted dynamic response of the demand (DRD) at any given time of the day to a given voltage drop and shaped DRD after load shifting. The actual measurements of DRD are used to validate the prediction. Artificial neural network is used to develop the total demand forecasting approach and the load disaggregation approach, and Monte Carlo simulation is used to generate the random voltages and the weighting factor of different load categories in Procedures (ii) and (iii). All simulations are performed in MATLAB 2013a using an Intel(R) Core (TM) i5-2400 CPU @ 3.10GHz computer equipped with 64-bit Windows 7 system.

### 1.5.3 The Scope of the Research

In Procedure (v), this research only focusses on the amount of different loads that should be shifted from the peak period to the off-peak period and how the shifted loads are allocated in the off-peak period. It does not provide solutions or justifications on what price signals or other DSM planning techniques (i.e. selections of objective function or optimisation method) should be used to achieve such load shifting. It is worth noting that Procedure (v) in this research is an illustrated process based on the results of prediction on load compositions which indicates the possibility of DRD

shaping via load shifting; it cannot be validated with field measurements because the load shifting action as illustrated here has not been used in real networks yet.

### 1.5.4 Objectives of the Research

The aim and the conceived overview of the research lead to the following objectives:

1. To investigate load controllability and the general contribution of different load categories to the total demand mix at different times of the day
2. To review and summarise load models and load modelling approach typically used for different load categories and aggregate loads at bulk supply points
3. To illustrate how dynamic responses of the demand can be estimated or predicted without having to perform the field test
4. To select or develop an appropriate approach for day-ahead real and reactive power forecasting
5. To develop an approach for load disaggregation that works out the percentage of different load categories at any given time of the day based only on RMS *P, Q, V* measurements from the bulk supply point
6. To enable day-ahead prediction of load compositions and their controllability and DRD based only on RMS *P, Q, V* measurements and measured or forecasted weather data
7. To validate the load disaggregation, the prediction of load compositions and the prediction of DRD with real power system demand data
8. To illustrate load shifting action based on the prediction of load compositions and their controllability, and assess the DRD after load shifting

## 1.6 Main Contributions of the Research

The research has contributed to several areas in the field of load modelling and load forecasting. These contributions are summarised as follows.

(Note: Paper numbers given in the parentheses indicate that the related results are published in international journals, in proceedings of international conferences or in technical reports. A full list of thesis-based publications is given in Appendix E.)

1. Investigations and summaries of hourly consumption patterns and controllable load penetration for different load categories at different locations in different seasons and years are performed. Based on the investigations, a framework for development of probabilistic decomposed daily loading curves (DDLC) is established. Different from the traditional DDLC, it takes the uncertainty of the demand into consideration in the presentation of the participation of different load categories at different times of the day. It is then extended to a framework for estimation of the dynamic response of the demand (DRD) at different times of the day for different load sectors. The process of DRD estimation based on the customer survey data is illustrated and the typical shapes of DRD for different load sectors are given. Afterwards, the framework is further extended to include probabilistic estimation/prediction of DRD at different times of the day for the total demand by taking into account the uncertainty of both load participation and individual load signatures of different load categories. Development of the probabilistic framework for estimation/prediction of DRD is the first original contribution of the thesis. [E1, E6, E7, E8]
2. Comparative analyses of two widely used artificial-intelligence (AI) based load forecasting approaches, artificial neural network (ANN) and adaptive neuro-based fuzzy inference system (ANFIS), are performed. This work

compares a conventional AI-based approach (i.e. ANN) with a hybrid AI-based approach (i.e. ANFIS) for load forecasting for the first time and concludes that both approaches can achieve comparatively high accuracy with optimal parameter configuration of prediction algorithms. The selected approach is also applied to reactive power prediction. This is the second original contribution of the thesis. [E4, E5, E9]

3. A disaggregation approach that enables the identification of different load compositions at the bulk supply point using only RMS measurements of the voltage, real and reactive powers is developed, and the confidence levels of the disaggregation errors are analysed. This directly facilitates the prediction of DRD without having to perform field tests. The approach is then extended to identify the controllability of different load categories and the effect of missing inputs on the disaggregation errors. This is the third original contribution of the thesis. [E2, E3]
4. Based on the assessment of the effect of missing inputs on the disaggregation errors, an approach that enables prediction of load compositions in the future is developed. This approach integrates the total load forecasting approach and the disaggregation approach, and the confidence levels of the errors on load composition forecasting are analysed. It is the fourth original contribution of the thesis. [E5]
5. An approach that derives the rated demands of the bulk supply point under different voltages is developed. It can provide an appropriate "base value" to transform the predicted or measured demand in MW or MVAR into per-unit values so that they could be used in the disaggregation algorithm. It also enables more accurate load modelling, especially for load modelling with a relatively longer timeframe. This is the fifth original contribution of the thesis. [E10]

6. The whole approach developed for the prediction of load compositions and the dynamic response of the demand (DRD) is validated using real demand data collected at substations of the real distribution network. It is shown that the predicted DRD matches the DRD derived from the field test well. In addition, a load shifting action that shapes the load curves and DRD is also illustrated based on the predicted load compositions. This is the sixth original contribution of the thesis. [E9]

## 1.7 Thesis Overview

The thesis is organised into seven chapters. This chapter (Chapter 1) is an introductory chapter. An overview of each of the remaining chapters is given below.

***Chapter 2-Overview of Power System Load Models***

This chapter reviews typically used load models in research and industrial practice. The voltage-dependent parameters of the most widely used load models for different load categories are described and analysed.

***Chapter 3-Framework for Estimation and Prediction of Dynamic Responses of Demand***

In this chapter, the overall framework for estimation or prediction of the dynamic response of the demand at bulk supply buses is demonstrated considering uncertainties in both load compositions and dynamic responses of different types of load at different times of the day. Load compositions derived based on customer surveys conducted in the past and daily loading curves for different load sectors provided in the past literature are used to illustrate the framework, and the uncertainties are modelled probabilistically. Dynamic responses of the demand of different load categories obtained from field measurements or computer simulations are used to build the probabilistic models. Monte Carlo simulation and the component-based load modelling approach are adopted

for the estimation or prediction of the probabilistic real and reactive power responses as well as the ranges of variation in these responses for any given time of the day.

***Chapter 4-Artificial-Intelligence-based Load Forecasting***

In this chapter, two most widely used artificial-intelligence-based load forecasting methodologies are compared in the same operating environment for the first time. According to the results of the comparison as well as the specific requirements in total demand forecasting in this research, a specific forecasting approach is selected, implemented and validated using real data for forecasting both, the real and reactive power. In the cases involving probabilistic studies or uncertainty analysis, the probabilistic characteristics of forecasting errors are analysed.

***Chapter 5-Load Disaggregation at Bulk Supply Points***

In this chapter, an effective (probabilistic) load disaggregation approach for the load at the bulk supply point is developed. It is based only on the substation RMS measurement. Smart meter data, customer surveys or high-resolution load signatures are not needed. Artificial neural network, Monte Carlo simulation and the component-based load modelling approach are used, and the approach is described in detail in this chapter. It is developed in per-unit as is the case in the vast majority of past studies. Thus, a definition of rated demand needed to transform the predicted or measured demand in MW or MVAR to per-unit values is also given in this chapter. Finally, the effect of the availability of different sets of input data on the accuracy of load disaggregation is analysed and discussed. The developed load disaggregation approach can be connected to the total demand forecasting approach in order to predict load compositions as well as the dynamic response of the demand in the future.

***Chapter 6-Validation of Developed Methodologies and Illustration of Shaping of Dynamic Responses of Demand***

This chapter uses real power system demand data to validate the approach developed for day-ahead prediction of the load composition and the dynamic response of the demand at bulk supply points without having access to the customer consumption data or performing field tests. Based on the results of the prediction of load compositions, load shifting and shaping of dynamic responses of demand are also demonstrated.

***Chapter 7-Conclusions and Future Work***

In this chapter, major conclusions of the research are presented and discussed, and suggestions for further work and development in the area of load modelling and forecasting are provided.

# 2 Overview of Power System Load Models

Load characteristics have significant influence on both steady-state and dynamic performance of power systems, and they are represented by load models. Therefore, to analyse power system behaviour correctly, accurate load models together with appropriate representations of generation, transmission and distribution parts of the system are required. There are a variety of load models used by utilities all over the world, such as the static load model, the induction motor model, and the composite load model. Although the majority of current power system research and industry acknowledges the importance of accurate representation of load characteristics for power system studies, they still use typical static load models (i.e. with the simplest mathematical form) in most cases. For dynamic load models, if used, they are usually presented by induction motor models. This chapter gives an overview of the most commonly used load models.

## 2.1 Load Models

The term ‘load model’ denotes an analytical, mathematical, equivalent-circuit based, physical-component based, or otherwise established or formulated

representation of a load, which correctly represents the changes in real and reactive power demands of the modelled load as a function of certain power system parameter (i.e. voltage, frequency) variations. Load models can be classified into two groups: static and dynamic. A classification diagram of load models is given as Figure 2.1.

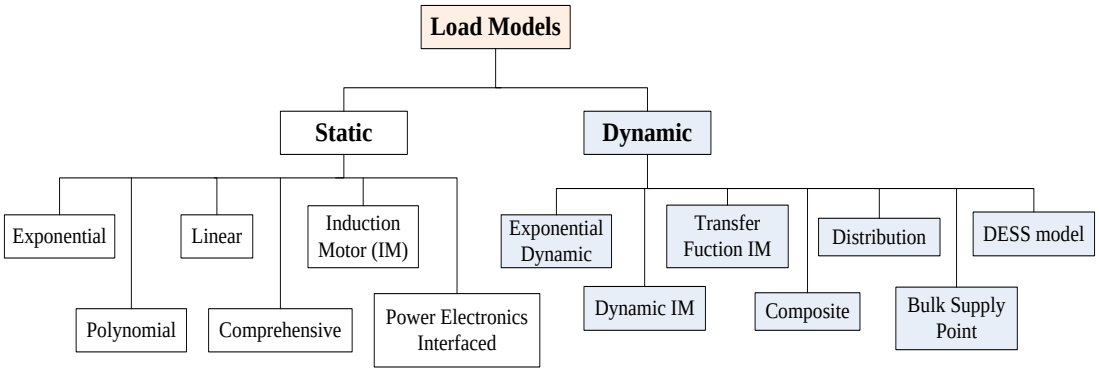


Figure 2.1. Load model classification diagram (adopted from [1])

A static load model is a time-independent load model that provides information on load characteristics as a function of known or specified system parameters. The most widely used static load models include the exponential model, the polynomial model and the static induction motor model. A dynamic load model is a time-dependent load model that provides information on load characteristics as a function of known or specified system parameters and time. The most widely used dynamic load models include the dynamic induction motor model, transfer function model and composite load models.

This section gives an overview of the most frequently used load models. More details on load models are given in CIGRE report [1], one of the most recent and most comprehensive references on load modelling.

## 2.1.1 Static Load Models

### 2.1.1.1 Exponential Load Models

An exponential load model describing the dependence of real power ($P$) on voltage ($V$) and frequency ($f$) is given by (2.1) [1]

$$P = P_0 \left( \frac{V}{V_0} \right)^{k_{pu}} \left( \frac{f}{f_0} \right)^{k_{pf}} \quad (2.1)$$

where $k_{pu}$ and $k_{pf}$ are the voltage and the frequency exponent, respectively. A similar relationship applies to the reactive power ($Q$).

Generally, during normal operation of power system, the voltage may vary within ±10% and the frequency usually vary within ±0.2 Hz (i.e. ±0.4% for a 50Hz system, [114]). In some special cases with extremely light load, the tolerance of the frequency change might go up to ±1% [114]. According to [1], which is one of the most recent and most comprehensive reference on the subject, the values of frequency exponents (for both $P$ and $Q$) for different types of individual loads generally lie between -4.5 and 5 (e.g., agriculture pump and TV sets are some of devices with extreme values of frequency exponents), and the values of voltage exponents for different types of individual loads generally lie within the range [0, 3] for $P$ and within [0, 7] for $Q$ [1]. Further details are available in [1], and typical values of the voltage and frequency exponents for different end-users are given in Appendix D. Considering the weights of different load types in the load mix at the bulk supply point, the average value of the frequency exponent of the aggregate load will be closer to 0 than the frequency exponents of some of the individual devices.

Since the frequency variations in the system are much smaller (almost order of magnitude) and less frequent than voltage variations, load dependence on frequency is generally neglected, except in specific frequency regulation and frequency stability studies [1]. The inclusion of load dependence on frequency in load models would require more variables to be calculated (i.e. frequency exponents for both real and reactive power) which significantly increases memory and processing time requirements without bringing sufficient (if any) improvement in accuracy of the results in most of the power system studies, such as power system stability studies [1].

The recently conducted international survey by CIGRE WG C4.605 on industrial practice on load modelling [115] found that 70% of 95 surveyed utilities and system operators around the world use only static load model for power system stability studies. One of the most commonly used types of static load models is voltage dependent exponential load model, shown as (2.2) [1]

$$P = P_0 \left( \frac{V}{V_0} \right)^{\alpha} \tag{2.2}$$

for the real power, where $P$ is the actual real power, $V$ is the actual load supply voltage, $P_0$ is the initial real power, $V_0$ is the initial load supply voltage, $\alpha$ is the voltage exponent for real power. A similar relationship applies to reactive power $Q$ and its voltage exponent $\beta$. The voltage exponent for constant impedance loads, constant current loads and constant power loads is 2, 1 and 0, respectively.

Although most utilities and system operators currently use static load models for power system stability studies, proliferation of new types of load devices and small distributed generators connected at low-voltage (LV) levels require more accurate load models for system stability studies and fault analysis. In particular, more accurate load models can not only more accurately represent the load characteristics at bulk supply points, but also lead to more accurate simulation results and help design more robust power system, as explained in Chapter 1.

#### 2.1.1.2 Polynomial Load Model

Another commonly used static load model is a polynomial load model. The general form of a polynomial load model is shown as (2.3) [1]

$$P = f_P(V, Para) \text{ and } Q = f_Q(V, Para) \tag{2.3}$$

where $f_P$ and $f_Q$ represent polynomial functions of $V$ for $P$ and $Q$, $Para$ represents parameters other than voltages, such as capacitance or inductance for electronic

devices or energy efficient lighting [28, 30].

A typically used special form of a polynomial load model is a combination of constant impedance, constant current and constant power load models with different weights. It is given as

$$P = P_0\left[a_P\left(\frac{V}{V_0}\right)^2 + b_P\left(\frac{V}{V_0}\right) + c_P\right], \text{ with } a_P + b_P + c_P = 1 \qquad (2.4)$$

where ($a_P$, $b_P$, $c_P$) are polynomial coefficients for real power. A similar relationship applies to reactive power $Q$ and the corresponding polynomial coefficients ($a_Q$, $b_Q$, $c_Q$). It is also called ZIP model, because it consists of the constant impedance ($Z$), constant current ($I$) and constant power ($P$) components.

### 2.1.1.3 Selection on Exponential and Polynomial Model

It is illustrated in [1] that the voltage exponents are much more sensitive to voltage variation than ZIP model or polynomial model coefficients. The load modelling result in [30] shows that under most circumstances, an exponential load model with constant voltage exponents performs less accurately than a ZIP model or polynomial model with constant polynomial coefficients. However, for most case studies, exponential load model is much more convenient than the polynomial model and the ZIP model because it contains only one term and only one parameter for either $P$ or $Q$.

### 2.1.1.4 Static Model of Induction Motors

Figure 2.2 shows the equivalent circuit of an induction motor. In Figure 2.2, $R_s$ is the stator resistance, $R_r$ is the rotor resistance, $X_{\gamma s}$ is the stator leakage reactance, $X_{\gamma r}$ is the rotor leakage reactance, $X_s$=$X_m$+$X_{\gamma s}$ is the shunt reactance, $X_m$ is the magnetising reactance. Typically used parameter values for different induction motor loads are given in Appendix D.

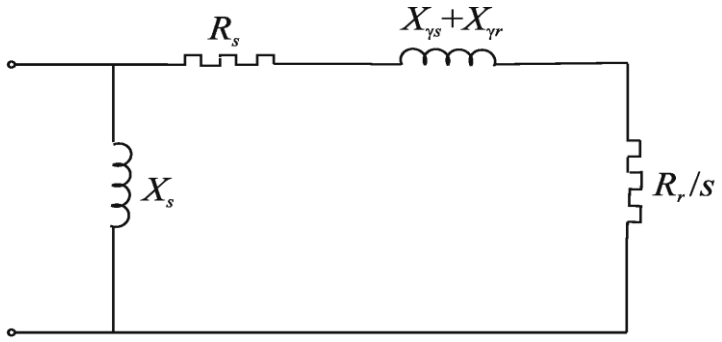


Figure 2.2. Equivalent circuit of an induction motor

The slip, $s$, can be defined as

$$s = \frac{\omega_s - \omega_r}{\omega_s} = 1 - \omega \tag{2.5}$$

where $\omega_s$ is the system angular frequency, $\omega_r$ is the rotor angular speed, $\omega$ is the per-unit rotor angular speed. The static induction motor model is given by (2.6) and (2.7) [1]

$$P = \frac{V^2}{\left(R_s + \frac{R_r}{s}\right)^2 + \left(X_{\gamma s} + X_{\gamma r}\right)^2}\left(R_s + \frac{R_r}{s}\right) \tag{2.6}$$

$$Q = \frac{V^2}{\left(R_s + \frac{R_r}{s}\right)^2 + \left(X_{\gamma s} + X_{\gamma r}\right)^2}\left(X_{\gamma s} + X_{\gamma r}\right) + \frac{V^2}{X_s} \tag{2.7}$$

Even if the model contains the term $V^2$ which makes the real and reactive power look proportional to $V^2$, in reality, induction motor is generally considered as constant power load. In addition, the simulation results also indicate that in the new steady-state achieved after an up to 20% voltage drop (i.e. a voltage drop over 20% may stop the induction motor from working properly), the real power consumption remains fairly the same as it was before the voltage drop occurred. This is because the slip also changes with voltage; when the voltage drops, the slip increases, and $R_r/s$ decreases. To clarify the issues, the swing equation of the rotor of the induction motor (provided in [116-118]) needs to be investigated in more detail. Further details on steady-state load characteristics of the induction motor are given in Appendix C.

## 2.1.2 Dynamic Load Models

For more accurate analysis of power system stability and control issues, dynamic load models are needed. The most frequently used dynamic load models include first-order exponential recovery model, dynamic induction motor model, transfer function model and composite load model.

### 2.1.2.1 First-order Exponential Recovery Model

Based on the general shape of the recorded responses of both individual loads and aggregate loads at bulk supply point, the first-order exponential recovery model can be used to model the loads with significant portion of different types of induction motors or loads composed of static loads and induction motors, or for representation of longer-term load dynamics influenced by operation of tap-changers. It can be represented by (2.8) and (2.9) [1]

$$T_p \frac{dP_r}{dt} + P_r = P_s(V) - P_t(V) = P_0 \left( \frac{V}{V_0} \right)^{\alpha_s} - P_0 \left( \frac{V}{V_0} \right)^{\alpha_t} \tag{2.8}$$

$$P_l = P_r + P_0 \left( \frac{V}{V_0} \right)^{\alpha_t} \tag{2.9}$$

where $P_r$ is the real power recovery after the voltage change, $P_0$ is the initial real power before the voltage change, $P_l$ is the real power consumption, $V_0$ is the RMS value of the initial voltage. Coefficient $\alpha_s$ and $\alpha_t$ are steady-state and transient real power voltage exponent, respectively, and $T_p$ is the real power recovery time constant. A similar relationship applies to the reactive power.

### 2.1.2.2 Dynamic Induction Motor Model

Dynamic induction motor model is used when participation of induction motor is predominant and the short-term dynamics of loads are needed to be modelled. The induction motors considered in system studies include both the three-phase induction

motor and the single-phase induction motor.

#### 2.1.2.2.1 Three-phase Induction Motor

Dynamic induction motor model of a symmetrical three-phase induction motor can be developed based on the equivalent circuit shown as Figure 2.3 [1, 27, 119].

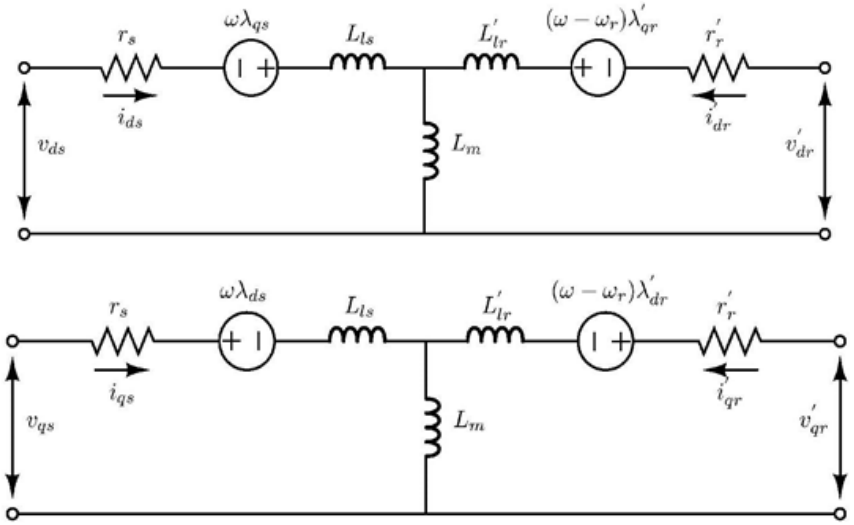


Figure 2.3. Equivalent circuit of a symmetrical three-phase induction motor with *d-q* references (adopted from [1])

In Figure 2.3, $r_s$ is the stator resistance, $r'_r$ is the rotor resistance, $L_{ls}$ is the stator leakage inductance, $L'_{lr}$ is the rotor leakage inductance, $L_m$ is the magnetizing reactance, $\omega$ is the synchronous angular speed, $\omega_r$ is the rotor angular speed. According to Figure 2.3, a 5$^{th}$-order dynamic induction motor model for symmetrical 3-phase induction motor can be represented by (2.10)-(2.18) [1, 119]

$$v_{ds} = r_s i_{ds} - \omega\lambda_{qs} + \frac{d\lambda_{ds}}{dt} \tag{2.10}$$

$$v_{qs} = r_s i_{qs} + \omega\lambda_{ds} + \frac{d\lambda_{qs}}{dt} \tag{2.11}$$

$$v'_{dr} = r'_r i'_{dr} - (\omega - \omega_r)\lambda'_{qr} + \frac{d\lambda'_{dr}}{dt} \tag{2.12}$$

$$v'_{qr} = r'_r i'_{qr} + (\omega - \omega_r)\lambda'_{dr} + \frac{d\lambda'_{qr}}{dt} \tag{2.13}$$

$$\frac{d\omega_r}{dt} = \frac{(M_e - M)}{\omega_b T_m} \tag{2.14}$$

$$\lambda_{ds} = (L_{ls} + L_m) i_{ds} + L_m i'_{dr} \tag{2.15}$$

$$\lambda_{qs} = (L_{ls} + L_m) i_{qs} + L_m i'_{qr} \tag{2.16}$$

$$\lambda'_{dr} = L_m i_{ds} + (L'_{lr} + L_m) i'_{dr} \tag{2.17}$$

$$\lambda'_{qr} = L_m i_{qs} + (L'_{lr} + L_m) i'_{qr} \tag{2.18}$$

Apart from the parameters used in Figure 2.3, in above equations: $v_{ds}$ and $v_{qs}$ are *d*-axis and *q*-axis stator voltage components, $v'_{dr}$ and $v'_{qr}$ are *d*-axis and *q*-axis rotor voltage components, $i_{ds}$ and $i_{qs}$ are *d*-axis and *q*-axis stator current components, $i'_{dr}$ and $i'_{qr}$ are *d*-axis and *q*-axis rotor current components, $\lambda_{ds}$ and $\lambda_{qs}$ are *d*-axis and *q*-axis stator flux linkages, $\lambda'_{dr}$ and $\lambda'_{qr}$ are *d*-axis and *q*-axis rotor flux linkages, $\omega_b$ is the base angular frequency, $M$ is the mechanical load torque, $T_m$ is the mechanical time constant of the motor, $M_e$ is the electromagnetic torque and $M_e = L_m (i_{qs} i'_{dr} - i_{ds} i'_{qr})$. Further details can be found in [119].

Since there are five differential equations forming the dynamic model of symmetrical three-phase induction motor, the dynamic model is fifth-order. If stator transients are neglected (i.e. $\lambda_{ds}$ and $\lambda_{qs}$ become constants), then two differential equations, (2.10) and (2.11), become algebraic equations and the model thus becomes third-order. The fifth-order model should be used for large induction motors [120] as their stator resistances are relatively small and result in large stator transients.

#### 2.1.2.2.2 Single-phase Induction Motor

Directly connected single-phase induction motors are widely used in low-voltage appliances, including refrigerators and freezers, washing machines and dryers,

dishwashers, fans etc. Single-phase induction motors are often referred to as asymmetrical two-phase induction motors, as an auxiliary winding is included and used during the start-up process. An equivalent circuit of the single-phase induction motor is shown as Figure 2.4 [26, 119]. All variables in the *d* axis are referred to the auxiliary winding with $N_S$ effective turns [119], and all variables in the *q* axis are referred to the main winding with $N_s$ effective turns.

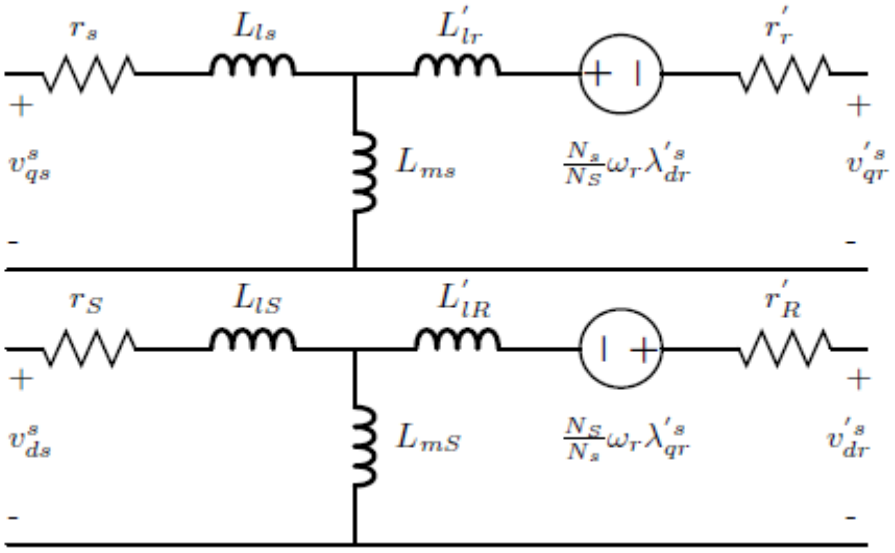


Figure 2.4. Single-phase induction motor equivalent circuit with *d-q* reference (adopted from [1])

According to the equivalent circuit, if the *d* and *q* components of the stator and rotor currents are taken as the variables, the dynamic induction motor model of a single-phase induction motor can be represented using (2.19) [119]:

$$\begin{bmatrix} v^s_{qs} \\ v^s_{ds} \\ v'^s_{qr} \\ v'^s_{dr} \end{bmatrix} = \begin{bmatrix} r_s + (L_{ls} + L_{ms}) \cdot \frac{d}{dt} & 0 & L_{ms} \cdot \frac{d}{dt} & 0 \\ 0 & r_S + (L_{lS} + L_{mS}) \cdot \frac{d}{dt} & 0 & L_{mS} \cdot \frac{d}{dt} \\ L_{ms} \cdot \frac{d}{dt} & -\frac{1}{n}\omega_r L_{mS} & r'_r + (L_{lr}' + L_{ms}) \cdot \frac{d}{dt} & -\frac{1}{n}\omega_r (L_{lR}' + L_{mS}) \\ n\omega_r L_{ms} & L_{mS} \cdot \frac{d}{dt} & n\omega_r (L_{lr}' + L_{ms}) & r'_R + (L_{lR}' + L_{mS}) \cdot \frac{d}{dt} \end{bmatrix} \begin{bmatrix} i^s_{qs} \\ i^s_{ds} \\ i'^s_{qr} \\ i'^s_{dr} \end{bmatrix} \quad (2.19)$$

where: $\omega_r$ denotes the rotor reference frequency, $\frac{d}{dt}$ represents the derivative operator, $n = \frac{N_S}{N_s}$ is the equivalent turn ratio of the auxiliary winding and the main

winding, and $L_{ls}$, $r_s$ are the main winding stator leakage inductance and resistance, $L'_{lr}$, $r'_r$ are the main winding rotor leakage inductance and resistance, $L_{lS}$, $r_S$ are the auxiliary winding stator leakage inductance and resistance, $L'_{lR}$, $r'_R$ are the auxiliary winding rotor leakage inductance and resistance, and $L_{ms}$ and $L_{mS}$ are the magnetising inductance of the main winding and auxiliary winding respectively.

#### 2.1.2.3 Transfer Function Load Model

Considering the complexity of the dynamic induction motor model, in some of the power system studies, the transfer function (TF) load model is used instead of the dynamic induction motor model to simplify the modelling while preserving required accuracy.

##### 2.1.2.3.1 First-order Transfer Function load model

The first-order transfer function (TF) load model can be represented by

$$\Delta P(s) = \frac{k_{pf} + T_{pf}s}{1+T_1 s}\Delta f(s) + \frac{k_{pu} + T_{pu}s}{1+T_1 s}\Delta V(s) \tag{2.20}$$

where $k_{pf}$ is the gain related to frequency change, $T_{pf}$ is the time constant related to frequency change, $k_{pu}$ is the gain related to voltage change, $T_{pu}$ is the time constant related to frequency change, $T_1$ is the overall time constant, $\Delta$ represents the change of different quantities. A similar relationship applies to the reactive power. The effect of frequency sometimes could be neglected when compared with the effect of voltage. The first-order TF load model is usually used to model single-phase induction motors and heat pumps in the residential and small commercial load sectors.

##### 2.1.2.3.2 Second-order Transfer Function load model

The second-order transfer function (TF) load model can be represented as

$$\frac{\Delta P(s)}{\Delta V(s)} = \frac{K_{pu}\left(1+T'_{1p}s\right)}{\left(1+T_{1p}s\right)\left(1+T_{2p}s\right)} \tag{2.21}$$

where $K_{pu}$ is the gain, and $T_{1p}$, $T_{2p}$, $T'_{1p}$ are time constants. A similar relationship applies to the reactive power change, $\Delta Q$. It is generally used to model three-phase induction motors in commercial and industrial load sectors under most circumstances.

#### 2.1.2.3.3 Higher-order Transfer Function Load Model

Sometimes, for higher accuracy of representation of the transient process, transfer function (TF) load model with higher order might be needed. It can be represented in a general form as

$$\frac{\Delta P(s)}{\Delta V(s)} = K_{pu} \frac{\prod_{n=1}^{N-1}\left(1+T'_{np}s\right)}{\prod_{d=1}^{N}\left(1+T_{dp}s\right)} \tag{2.22}$$

where $K_{pu}$ is the gain, and $T_{dp}$, $T'_{np}$ are time constants, $N$ is the order number of the TF load model. It is generally used to model three-phase induction motors in commercial and industrial load sectors when accurate transient processes are concerned.

In load modelling approach, measured responses to system disturbances will be compared with these models. Generally, second-order and higher-order load models are better in capturing the load behaviour during transients (i.e. during the first few hundreds of milliseconds) than the first-order load model, especially for large three-phase induction motors.

### 2.1.2.4 Composite Load Model

The composite load model is also widely-used in modelling the aggregate loads with a mix of static loads and induction motors. The equivalent circuit of a composite load model is given as Figure 2.5. It incorporates both the static part, which is represented by a ZIP model, and the dynamic part, which is represented by an induction motor model. The static load part and the induction motor part are connected in parallel.

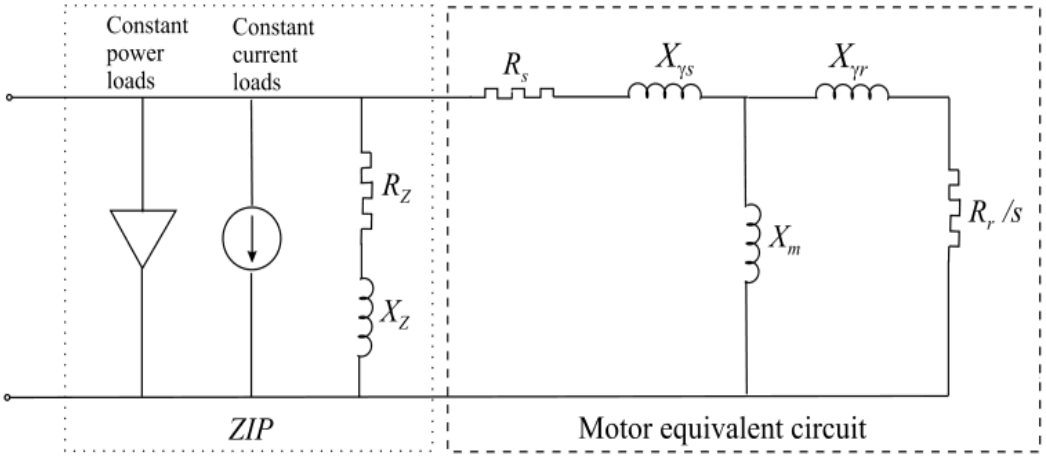


Figure 2.5. Equivalent circuit of composite load model (adopted from [1])

## 2.2 Load Models in Industry Practice

In February 2010, CIGRE established a working group C4.605 to conduct a survey [121] on load modelling and identify current international industry practice on load modelling for steady-state and dynamic power system studies. A questionnaire (shown in Table 2.1 [115]) has been developed and distributed to more than 160 utilities and system operators in over 50 countries in the five continents (shown in Table 2.2 [115]).

Table 2.1. Questionnaire (adopted from [115])

| Question No | Question | Category |
|---|---|---|
| Q1 | Types of load models used in static power system studies (e.g. power flow analysis) | 1 |
| Q2 | Types of load models used in dynamic power system studies (e.g. stability analysis) | |
| Q3 | Load models for different load categories/classes | |
| Q4 | Approaches for load model data collection and parameter identification | 2 |
| Q5 | Most recent update of load model parameters | |
| Q6 | Load simulation tools used for system studies | 3 |
| Q7 | Adequacy of available load models for system stability studies | |
| Q8 | Extent of use of user-defined load models | |
| Q9 | Accounting for or inclusion of small distributed generation in load models | 4 |

Table 2.2. Participants and Response Rates (adopted from [115])

| Continent | Number of Utilities/Operators Contacted | Number of Received Responses | Response Rate [%] | Response Rate relative to all 97 responses[%] |
|---|---|---|---|---|
| Africa | 21 | 7 | 33.3 | 7 |
| Americas | 34 | 17 | 50.0 | 18 |
| Asia | 35 | 24 | 68.6 | 25 |
| Europe | 62 | 41 | 66.1 | 42 |
| Oceania | 8 | 8 | 100.0 | 8 |
| **Total** | **160** | **97** | **60.6** | |

The four question categories shown in Table 2.1 are: (1) types of load models used in static and dynamic power system studies; (2) approaches for load model and parameters identification; (3) adequacy of load models and the simulation tools used in the validation process; (4) approaches to incorporate small distributed generation (DG) in load models. Multiple answers were allowed for questions Q1, Q2, Q4, Q5, and Q6, as respondents might use more than one load model, or more than one approach for load model data collection and parameter identification. Multiple answers were not allowed for questions Q3, Q7, Q8, and Q9.Table 2.2 shows the number of utilities/operators contacted in different continents and the number of responses received. The overall response rate is about 60.6% according to the result of the survey.

From the survey, it can be concluded that:

- ✓ For power system stability studies, about 70% of the utilities/operators use only static load models, and around 30% use some form of induction motor (IM) model to represent dynamic loads. In US, a combination of static load model (ZIP typically) and dynamic load model (IM model typically) are in dominant use, which is in consistency with the information provided by WECC.
- ✓ In about 40% of the cases, the parameters of the currently used load models have been updated within the last five years. If equivalent

exponential static load model is used to represent the load characteristics, the value of the exponent is 0 for both real and reactive power in steady-state studies, and is 0.7 and 1.3 for real and reactive power respectively in dynamic studies.

- ✓ Most of the utilities and operators use negative load to represent the characteristics of DG in system studies, without developing DG models explicitly; but some of them have already recognized the importance of appropriate models for DG, especially in future power system studies.

## 2.3 Voltage-dependent Exponential Model Parameter

Sometimes, to make the programming process easier, the exponential load model is selected instead of other models because it has only one parameter. It has been mentioned before in Section 2.1.1.3 that the parameter of the exponential model is more dependent on voltage than that of the ZIP model or the polynomial model. As a result, the exponent-voltage relationship of both the real and reactive power for each typically used load category should be derived before the exponential model is applied.

### 2.3.1 Derivation of Voltage-dependent Exponential Model Parameter

Exponents under different voltages can be calculated using Equation (2.2), ideally with adequate field measurement data (i.e. *P&Q&V* sets). If the field measurement data are not available, the more accurate models such as the polynomial model or the ZIP model can be used alternatively as the 'filed measurement data'. The relationship between the real power exponent $\alpha$ and the supply voltage $V$ ($V \neq V_0$) can be calculated from

$$\alpha(V) = \frac{ln\left(\frac{P_{measurement}}{P_0}\right)}{ln\left(\frac{V}{V_0}\right)} \tag{2.23}$$

when field measurement data are available, where $P_{measurement}$ is measured load power. And when field measurement data are not available, the exponents of $P$ can be

calculated from

$$\alpha(V) = \frac{ln\left[\frac{f_P(V, Para)}{P_0}\right]}{ln\left(\frac{V}{V_0}\right)} \tag{2.24}$$

The definitions of all other variables in (2.23) and (2.24) are the same as those introduced for (2.2) and (2.4). It is worth noting that (2.23) and (2.24) are valid only if the voltage is not equal to the rated voltage $V_0$. When $V=V_0$, the real power is always $P=P_0$=1.0 p.u. regardless of the value of $\alpha$, therefore $\alpha$ in this case is not discussed. A similar relationship applies to the reactive power exponent, $\beta(V)$.

## 2.3.2 Voltage-dependent Exponents of Different Categories of Loads

According to the categories summarised in Table 1.2, the polynomial load models for different load categories used in this research are listed in Table 2.3. In Table 2.3, L represents lighting loads, REC represents rectifier, CdL represents cold loads, WL represents wet loads, and R represents resistive loads.

Table 2.3. Polynomial Load Models of Different Load Categories used in this Research

| Load Type | Mathematical Model | |
|---|---|---|
| L | $P$ | $P = 0.581 + 0.371V - 0.0151P_b + 0.037V^2 + 0.016VP_b - 2.457\times10^{-6}P_b^2$ |
| | $Q$ | $Q = 1.193 + 1.604V - 0.0108P_b + 1.119V^2 - 0.045VP_b - 4.441\times10^{-4}P_b^2$ |
| SMPS | $P$ | $P = P_0$ |
| | $Q$ | $Q = Q_0\left(0.029V^2 + 0.188V + 0.272 + 0.236V^{-0.033} + 0.236V^{-0.033}\right)$ |
| REC | $P$ | $P = 4.6902V^2 - 6.7404V + 2.3222 - 0.85852P_0 + 1.8969P_oV$ |
| | $Q$ | $Q = Q_0\left(0.266V^2 + 0.1641V - 0.042 + 0.243V^{2.17} + 0.243V^{2.17}\right)$ |
| CdL | $P$ | $P = 0.101V + 0.099V^2 + 0.798$ |
| | $Q$ | $Q = -0.905V + 1.402V^2 + 0.503$ |
| WL | $P$ | $P = -0.634V + 0.268V^2 + 1.366$ |
| | $Q$ | $Q = -0.905V + 1.402V^2 + 0.503$ |
| R | $P$ | $P = -0.634V + 0.268V^2 + 1.366$ |
| | $Q$ | $Q = 0$ |
| CTIM3 | $P$ | $P = -0.634V + 0.268V^2 + 1.366$ |
| | $Q$ | $Q = -2.15V + 1.751V^2 + 1.4$ |
| QTIM3 | $P$ | $P = 0.424V - 0.147V^2 + 0.724$ |
| | $Q$ | $Q = -2.15V + 1.751V^2 + 1.4$ |
| $P_b$ is the base power | | |

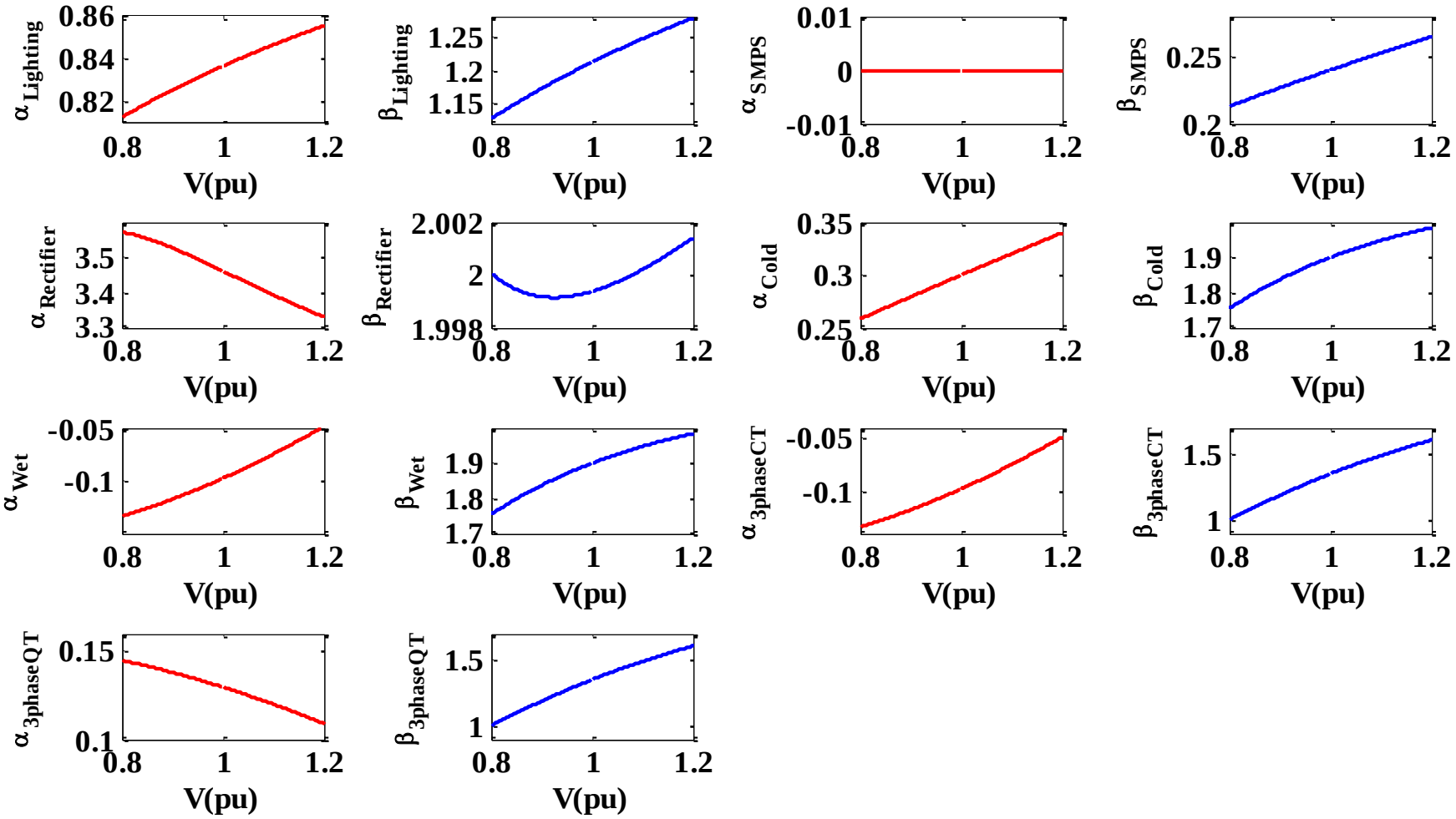


Figure 2.6. *P&Q* exponent-voltage relationship of different load categories

These models were originally developed and validated through laboratory measurements [26-28, 30]. All values of parameters in Table 2.3 are in per-unit. In the model of lighting load, $P_b$ =1 is the base power [28], $P_0$=1 p.u., and $Q_0$ depends on the nominal power factor.

The exponent-voltage relationships for lighting loads, SMPS, rectifier, cold loads, wet loads, CTIM3 and QTIM3 are shown in Figure 2.6. For the resistive load, the real power exponent is always 2; as the reactive power of the resistive load is always 0, its reactive power exponent is not discussed here. It can be seen that apart from the $\beta$-$V$ curve of the rectifier, all other $\alpha$-$V$ and $\beta$-$V$ curves are monotonic or constant. The $\beta$-$V$ curve of the rectifier, however, looks parabolic. This is probably because of the load characteristics of the rectifier, which has the lowest power factor when the voltage is around 0.92p.u [30].

Once obtained, the relationship is saved for further use. With these relationships, actual real and reactive power in p.u. for individual load categories under different load supply voltage can be calculated from Equation (2.2).

## 2.4 Summary

The most frequently used static and dynamic load models are introduced and analysed in this chapter. From the international survey on load modelling, it is found that the majority of the utilities use static load models only. It is also found that for different power system studies, different load models are selected. Generally, in steady-state analysis of power system, constant power loads are usually used; while in power system stability study, static load models (such as exponential load model or ZIP load model) or a combination of static load models and induction motor models are usually preferred than other models.

As the exponents in the exponential load models are more sensitive to the

change of the voltage than the coefficients of ZIP load models and other models, in this chapter, the concept of "voltage-dependent exponential model parameter" is proposed, and the voltage-dependent exponents of some typical load categories are analysed. The result from this analysis will be used in the development of the load disaggregation approach in this thesis.

# 3 Framework for Estimation and Prediction of Dynamic Responses of Demand

The chapter presents the overall framework for estimation/prediction of dynamic response of the demand (DRD) at distribution network buses by taking into account the daily variation in demand compositions and the generic dynamic responses of different types of load. Dynamic load models for different load categories obtained from field or laboratory measurements or through appropriate mathematical modelling are used in combination with hourly load composition at the given bus. For illustration purpose, in this chapter, the load composition is derived based on past demand surveys and daily loading curves for different classes of customers. Uncertainties in both, dynamic load models/responses of individual loads and load compositions at different time of the day, are modelled probabilistically. With established dynamic signatures of different load categories and load compositions at different times of the day, Monte Carlo simulations are used to estimate/predict probabilistic real and reactive power responses, including ranges of variation of these responses, for every hour of the day for a given or anticipated mix of demand.

## 3.1 Dynamic Response of the Demand

As mentioned in Chapter 1, the dynamic response of the demand (DRD) is defined as the time-domain (usually 1 sec to several seconds timeframe) real and reactive power response to a voltage disturbance. Due to inevitable demand variation during the day and season associated with end use customers' operation cycles and habits, the composition of loads forming the total demand at network bulk supply point changes. This results in significant uncertainty in the dynamic response of demand (DRD) following network disturbances (voltage drops, faults, switching operation, load and generation switching, etc.) which could significantly influence the overall network performance.

The proliferation of new types of loads (power electronic interfaced load, efficient lighting, electric vehicles, etc.), small distributed generators and to a certain extent already modified customer behaviour (i.e., compared with the customer behaviour recorded by the surveys conducted in the past) introduces further uncertainty in the demand composition. In addition to uncertainties associated with the demand composition, there is another layer of uncertainty resulting from insufficiently understood responses of these new types of loads to network disturbances, their mathematical models and parameters of those models.

The drive towards reliance on active customer participation in network operation (i.e., willingness to get connected or disconnected at certain time of the day to facilitate appropriate demand profile) further complicates the issue of demand modelling. By shifting certain amount of demand at different time of day (e.g., connecting 10,000 washing machines at 11am instead of 1pm) relying on customers' willingness to participate in demand management, price signals or other market mechanisms would help balance the generation and the demand in the network but

will also change the composition of the demand at those time. (In the example above, over 10,000 single-phase small induction motors would be connected at 11am and disconnected at 1pm). The change in the composition of the demand would affect its dynamic response to network disturbances at those hours, which could significantly influence the overall network voltage and angular stability. Being able to predict the dynamic response of demand (DRD) following network disturbances at any given time of the day would certainly facilitate efficient active demand management and ensure stable and secure operation of the power system as a whole.

In this chapter, the framework based on probabilistic methodology for estimation/prediction of the dynamic response of the aggregate demand with limited information about the demand composition and knowledge of typical dynamic responses of individual load components is introduced. It combines elements of measurement-based and component-based approach to load modelling with probabilistic modelling of the demand composition at bulk supply buses. The framework and underlying methodology does not aim, at this stage, to develop an accurate load model for any particular load but rather to lay foundations for: i) estimation of the dynamic load response at some given time; ii) prediction of the load response at some point in the future based on readily available information without having to perform field tests or measurements. The proposed methodology could be used as an additional input to demand side management (DSM) decision making to ensure desired (or maintain existing) dynamic response of the demand at given hour following the shift of the demand to balance the load and the generation.

## 3.2 The Framework

The framework for estimation and prediction of the dynamic response of the demand (DRD) is shown as Figure 3.1. It can be seen that the information needed

includes demand compositions and individual dynamic load signatures. More details about the framework are provided in the following subsections.

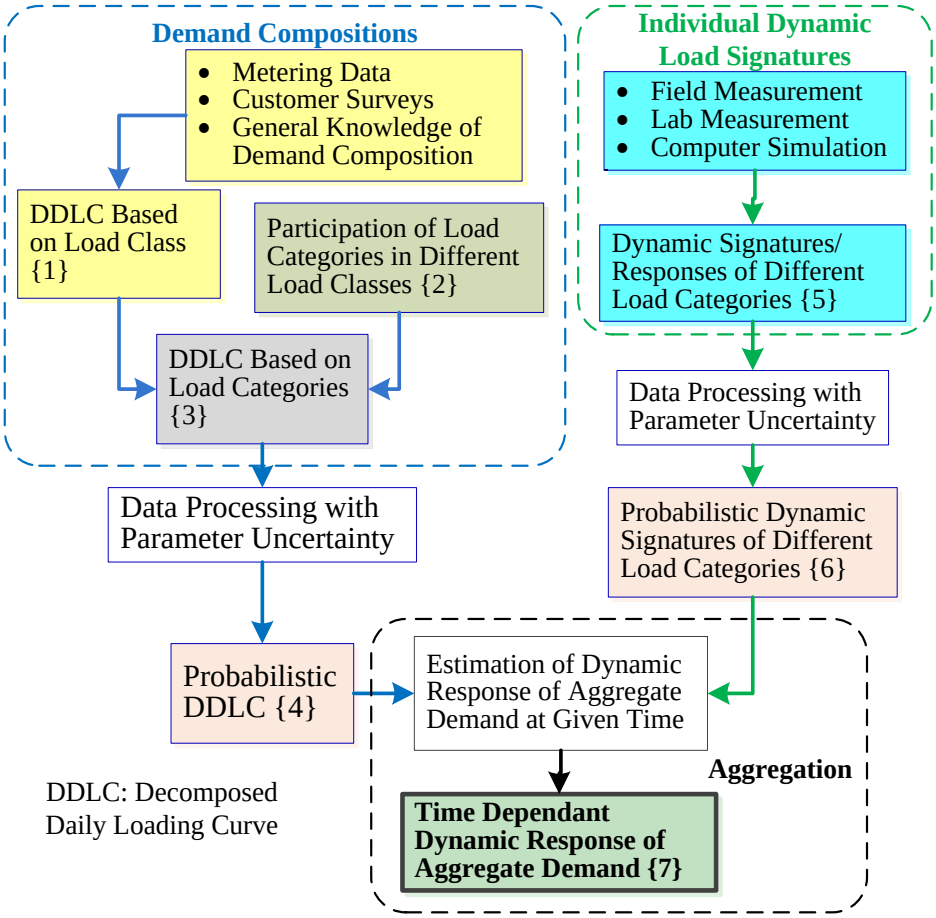


Figure 3.1. Framework for estimation and prediction of DRD

### 3.2.1 Decomposed Daily Loading Curves

Decomposed daily loading curves (DDLCs) are used to show the share of different types or categories of loads at different times of the day. The definition of decomposed daily loading curves (DDLCs) is given in Chapter 1 as "detailed daily loading curves indicating participation of different types of loads at different times of day for different load sectors during different seasons [1]", and DDLCs in different load sectors at different locations during different seasons are reviewed and introduced in Chapter 1.

### 3.2.2 Approaches to Load Decomposition

The participation of different load devices in the overall demand at bulk supply point varies with time and it depends on load classes involved, devices constituting the demand and nature of processes/pattern of device usage in end user facilities. The global participation of different classes of demand (industrial, domestic, commercial, etc.) can be established to a certain extent based on metering data, while it is very difficult to establish participation of different load categories within each class at any given time. In this chapter, whole-house hourly load data and associated survey responses for a sample of households given in [39-42] are used for illustration purpose. (Note: a new approach for load disaggregation or load decomposition which does not need detailed customer surveys is developed in Chapter 5. This is also one of the original contributions of the thesis, as introduced in Chapter 1.)

Once the decomposed daily loading curve (DDLC) based on load classes (box labelled {1} in Figure 3.1) is known, the assessment of participation of different load categories in each class should be determined. Since there are many devices whose individual dynamic signatures (real and reactive power response to voltage change) are very similar, e.g., televisions (TV), personal computers (PCs), DVD, radios etc., they can be classified into the same category, namely power electronics/switch mode power supply (SMPS). Therefore, based on similar reasoning, all devices can be classified into several categories based on the type of expected dynamic response [1, 24, 32]. This classification is discussed in Chapter 1. Participation of each of the main load categories in load class (box labelled {2} in Figure 3.1) can be obtained, ideally, from customer surveys or via computer simulation by applying appropriate clustering and classification algorithms.

Based on these two sets of information, the decomposed daily loading curve

(DDLC) based on load categories can be established (box labelled {3} in Figure 3.1), which will give appropriate mix of load categories at each bulk supply bus. The largely stochastic characteristics of the load (both, the change of load composition at the same time of different days and the variation of parameters of the model of individual load component) is the key aspect that needs to be considered in the modelling process. This however, has been addressed only to a limited extent in the open literature in the past. For examples, the stochastic nature of loads was discussed in [16, 17, 122], and the probabilistic approach was used to derive parameters of load models in [19, 20]. Further processing of DDLC based on load categories therefore is needed to take into account uncertainties involved in establishing exact participation of different load categories in load classes. This yields the probabilistic decomposed daily loading curve (box labelled {4} in Figure 3.1), which is defined as "the decomposed daily loading curve that indicates the customer participation in load mix in terms of mean and standard deviation of probability distribution".

As an aggregate load at bulk supply point consists of different number and type of individual loads, it is necessary to develop a database of appropriate dynamic signatures and corresponding mathematical models of individual load categories. This can be ideally accomplished either by field or laboratory measurements of real and reactive power responses of required load categories or through computer-based simulations (box labelled {3} in Figure 3.1). As it is highly unlikely to have exact dynamic signatures or mathematical models of all load categories connected to a given bulk supply bus in the network, some computer simulations are inevitable to account for numerous uncertainties involved. This post processing (accounting for uncertainties in load dynamic response through computer simulations) of load dynamic responses yields probabilistic dynamic signatures (calculated using probabilistic distribution of parameters and defined in terms of mean and standard

deviation of probability distribution) of different load categories (box labelled {6} in Figure 3.1) that account for uncertainties involved.

Finally, having established probabilistic decomposed daily loading curves (DDLCs) and probabilistic dynamic signatures of different load categories, the dynamic response of aggregate demand at given time (box labelled {7} in Figure 3.1) can be established by weighting individual dynamic signatures of different load categories by participation of corresponding load category in load mix at given time (approach similar to conventional component based approach to load modelling).

## 3.3 Estimation of Demand Composition

The information about the demand composition at bulk supply bus at given time is the essential requirement for estimating the dynamic response of the demand (in addition to individual load component dynamic signatures) in the absence of the measured one. Two approaches that could be used for this purpose are discussed in the following section.

### 3.3.1 Deterministic Decomposition of Daily Loading Curves

Figure 3.2 (reproduced from Chapter 1) shows a decomposed daily loading curve (DDLC) in terms of different load types for commercial load sector for a working day. It is derived from daily demand data for different types of loads provided by [35]. For illustration purpose, in this chapter, the peak demand in the load sector is normalised to 100 for ease of comparison. The key information required for estimating the dynamic response at the bulk supply point however is the type and the percentage of different load categories rather than appliance or end-users involved, i.e., devices for which the dynamic signatures are available. Therefore, a load-category-based daily loading curve is needed instead of that shown in Figure 3.2. Considering demand decomposition shown in Figure 3.2 and according to Table 1.2, it can be seen that one

load category (e.g., the induction motor) includes multiple load types as identified in the figure. Therefore, load types listed in Figure 3.2 are categorised as follows: air conditioning load, motors, process, refrigerator, ventilation and cooking load are categorised as residential-commercial motor load category; office equipment is categorised as power electronics load category; internal and external lightings are categorised as lighting load category; cooking, water heating and heating load are categorised as resistive load. Due to different definitions of miscellaneous load in different areas [39, 123], miscellaneous is classified as a single category. The lines that separate any two load types in Figure 3.2 are called separation lines.

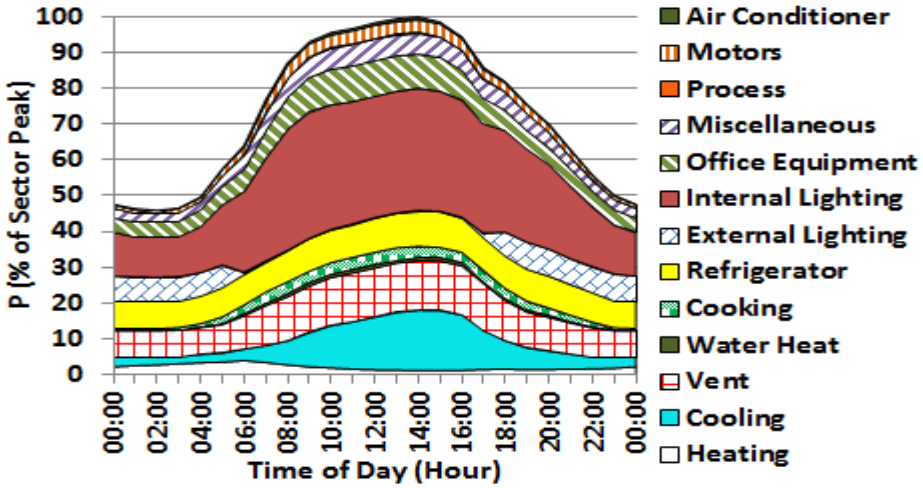


Figure 3.2. DDLC for commercial load sector

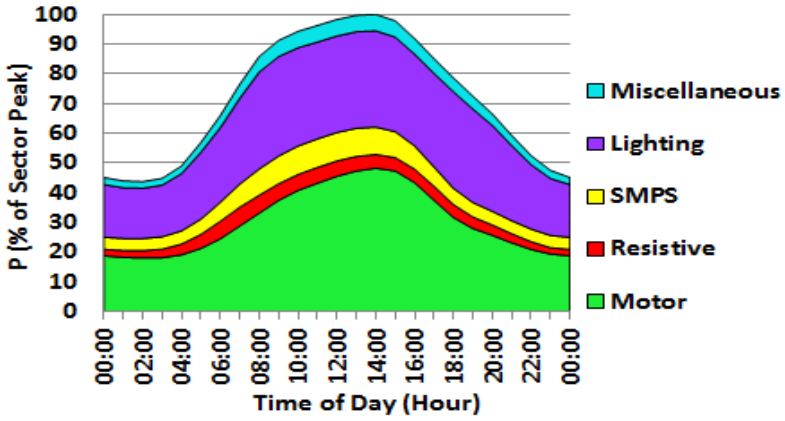


Figure 3.3. DDLC based on load categories for commercial load sector

Following this categorisation, a new load category based decomposed daily loading curve (DDLC) is produced by summing up the consumption of different load types within the same load category, as shown in Figure 3.3, and the separation lines are redistributed. From this figure, it can be easily determined which types of load categories are participating in total demand at the given bus and what shares of total demand it constitutes. Assuming that the dynamic signature of each load category is known (and not affected by the dynamic response of other load categories, as commonly assumed [1]), it will be much easier to estimate dynamic responses of the demand (DRDs) at any chosen time by "summing up" participating load category dynamic signatures after assigning appropriate weighting factors (based on participation in total demand) [1, 5]. Similar to Equation (1.2), this approach is illustrated by

$$P_{Agg}(t) = \sum_{i=1}^{N} w_i P_i(t) \tag{3.1}$$

where $t$ is the time, $P_{Agg}(t)$ is the aggregate dynamic response of the demand (DRD) in the time domain, $N$ is the total number of load categories, $w_i$ is the weighting factor of category $i$, and $P_i(t)$ is the DRD of category $i$ in the time domain. A similar formula can be used to derive DRD for the reactive power, $Q_{Agg}(t)$.

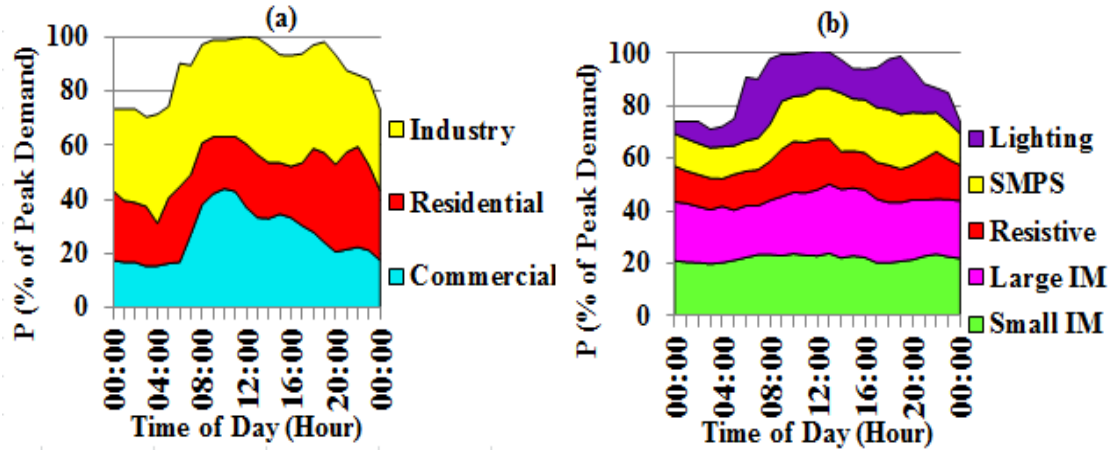


Figure 3.4. DDLC for general network bus based on: (a) load class mix, and (b) load category mix

A similar approach can be used to convert the conventional decomposed daily loading curve (DDLC) based on load class (i.e. residential, commercial, industrial), shown as Figure 3.4(a), to DDLC based on different load categories (i.e. lighting, SMPS etc.) for general network bus, shown as Figure 3.4(b).

### 3.3.2 Probabilistic Decomposition of Daily Loading Curves

Once the decomposed daily loading curve (DDLC) based on load categories, shown in Figure 3.4(b), is established, the uncertainty in participation of different load categories in different load classes has to be taken into account. These uncertainties are the result of not only lack of exact data about participation of particular load category in a given class, but also the variation of end user deployment of different devices at different time caused by factors such as customer behaviours, weather change and special rare events, which is very difficult to establish for a bulk supply bus consisting of very large number of different classes of end users. Therefore, the load category participation in total demand should be modelled probabilistically.

To illustrate this concept, the uncertainty in participation of different categories in Figure 3.4(b) (taken as an example) is modelled in the following way. Firstly, the decomposed daily loading curve (DDLC) of Figure 3.3 is developed for each set of data originally used for producing corresponding DDLC of Figure 3.2 (obtained from customer surveys) and normalised with respect to the peak demand. Then, according to the collected load participation data in different load sectors and the distribution of different load sectors at different time during the day (shown in Figure 3.4(a)), load category-based DDLCs for the total load at the bus are developed (Figure 3.4(b)). Different load category-based DDLCs produced in this way are then overlapped. This results in DDLCs where the separation lines, separating different categories at different time, do not coincide. Therefore, the uncertainty area, i.e., the separation

band (the shaded area in Figure 3.5), replaces the separation line used in deterministic DDLC.

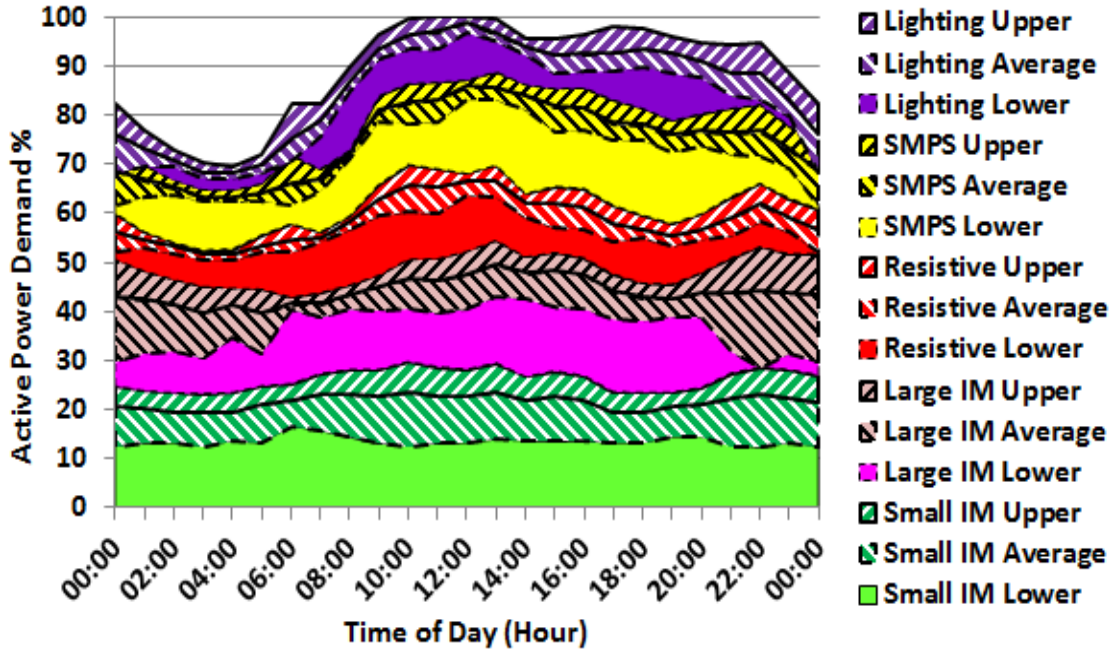


Figure 3.5. Probabilistic DDLC based on load categories

The separation band is produced by connecting the highest and the lowest points of the corresponding separation lines at each given time to produce the upper and lower limit, respectively. From the data collected from the conducted customer surveys, it is found that for each load category at any given moment, its percentage of participation in total demand (weighting factor), can be roughly approximated by a Gaussian distribution with the upper and lower limit at that time set to correspond to the $\mu+3\sigma$ and $\mu-3\sigma$, respectively, where $\mu$ is the mean value of the distribution and $\sigma$ is the standard deviation. In this way, each uncertainty area contains 99.74% of all possible weighting factors. The probabilistic decomposed daily loading curve (DDLC) is then drawn by connecting the average values of weighting factors at different time. Figure 3.6 shows the probabilistic load curve for Small IM (i.e. small induction motor) mentioned in Figure 3.5. In Figure 3.6, terms “upper limit”, “lower limit”, mean value

and the uncertainty area are clearly indicated.

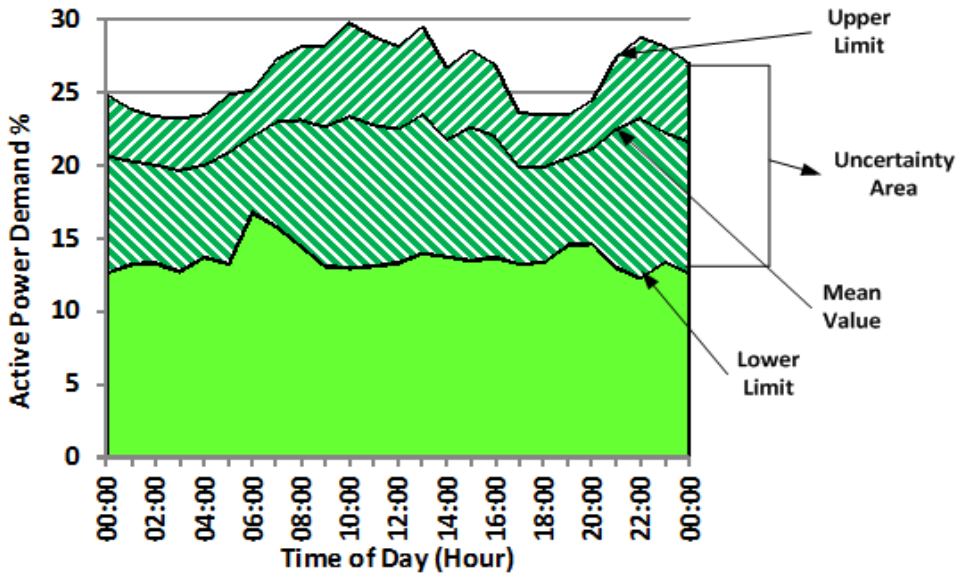


Figure 3.6. Graphical representation of "upper limit", mean value, uncertainty area and "lower limit"

Depending on the skewness of Gaussian distributions at different times, the most probable value of the weighting factor at that time may not coincide with the mean value of the Gaussian distribution, and the resulting line connecting the most probable weighting factors at different time may not run exactly through the middle of the uncertainty area. In the illustrative case in this chapter, the difference between the mean and the most probable value of the distribution at different time is not significant, and the mean values can be used as a first approximation instead of the most probable values for illustration purpose. Furthermore, in general cases, the mean value of the Gaussian distribution may not correspond to the average value of participation at given time depending on the skewness of the distribution. The choice of which value to use in a particular case (the average value, the mean value of the Gaussian distribution or the most probable value) will depend on the actual distribution of values within the area of uncertainty at different time.

An illustrative example of the developed probabilistic decomposed daily loading

curve (DDLC) for a bus is shown in Figure 3.5. In this figure, the shaded areas represent the areas of uncertainty for different load categories participating in the overall demand at different time. The solid lines inside the uncertainty area represent the average participation of particular load categories in total demand at given time. It can be seen that these lines do not run exactly through the middle of the uncertainty area, as discussed above. By comparing calculated hourly average values of participation of particular load categories and mean values and standard deviations of corresponding Gaussian distributions for different load categories, it was found that the average values in this illustrative case can be used as a good approximation of mean values of corresponding Gaussian distributions of participation of different load category in DDLC. These values are then also used as weighting coefficients for weighting and summing up individual load category dynamic signatures to derive the dynamic response of the total demand by using Equation (3.1).

The mean value and the standard deviation of contribution of different load categories to overall demand for the probabilistic DDLC shown in Figure 3.5 are plotted in Figure 3.7, and it will be used in probabilistic estimation of dynamic responses of demand in later sections.

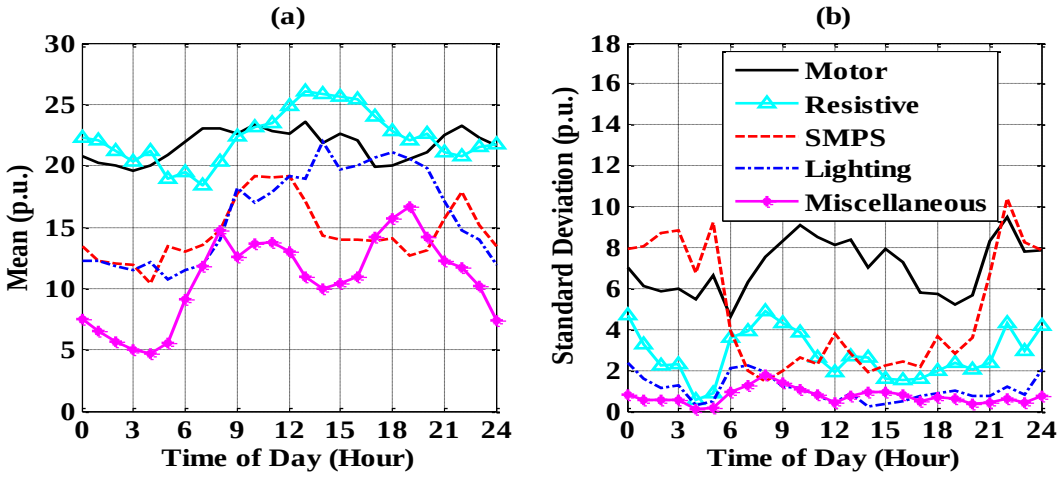


Figure 3.7. (a) Mean, and (b) standard deviation of demand for different load categories for probabilistic DDLC shown in Figure 3.5

## 3.4 Dynamic Load Characteristics of Typical Load Categories

To achieve the objective of estimation and prediction of dynamic responses of demand mentioned in Chapter 1, dynamic load characteristics of typical load categories should be studied according to the measurement-based load modelling approach introduced in Chapter 1. The following sub-sections study the dynamic load characteristics of load categories used in this research. As the concept of probabilistic load model parameters has been proposed in [18, 19] to deal with the uncertainty, for illustration purpose in this chapter, the dynamic load model parameters are also presented with probabilistically distributed function, and the dynamic response of different load categories to a voltage disturbance is also presented probabilistically.

### 3.4.1 Energy Efficient Lighting

The energy efficient lighting (EEL) is one of the load categories with fastest growing rate. It consists of different types of illumination devices including the compact fluorescent lamp (CFL), which is expected to replace the general incandescent light (GIL) to large extent in many parts of the world in the near future [24]. As dynamic processes associated with the operation of the EEL have very short time constants (from tens to hundreds milliseconds), i.e., much faster than a general electro-mechanical dynamic process, the EEL load can be regarded as a static load. Therefore, an exponential load model given by (3.2) and (3.3) [124]

$$P = d_p V^{ep} \tag{3.2}$$

$$Q = d_q V^{eq} \tag{3.3}$$

is used to model EEL, where $d_p$ ( $d_p = \frac{P}{V^{ep}}$ ) and $d_q$ ( $d_q = \frac{Q}{V^{eq}}$ ) are the scaling coefficients for $P$ and $Q$, and $ep$ and $ep$ are corresponding voltage exponents for $P$ and

*Q* [124].

The mean value and standard deviation of model parameters are provided in Table 3.1 based on data reported in [1, 124] and validated via measurement-based approach by [28]. It can be seen that the mean value of *ep* is 1.062. Therefore, to a first approximation, the real power response of the EEL can be modelled as constant current load. The reactive power has a characteristic between constant power and constant current according to the value of *eq*. (Note: the minus sign in front of reactive power coefficient, $d_q$, indicates that the EEL always has leading power factor.)

Table 3.1. Model Parameters for EEL (Adopted from [1, 124])

| | $d_p$ | *ep* | $d_q$ | *eq* |
|---|---|---|---|---|
| Mean | 0.751 | 1.062 | -0.624 | 0.653 |
| Standard Deviation | 0.043 | 0.164 | 0.042 | 0.098 |

Figure 3.8 shows normalised (the value of power before the disturbance set to 1p.u.) real power responses of EEL to a 5% voltage step reduction. Figure 3.8a) illustrates responses obtained in 500 Monte Carlo simulations (MCS) by varying parameters in Table 3.1 based on normally distributed parameter values using provided mean value and standard deviation. Other distributions could be also used though, if available. (Note: Reference [125] specifies how many Monte Carlo simulations are needed to be run to produce statistically relevant samples. According to [125], required number of Monte Carlo simulations for different parameters of different individual load models in case of studies reported in this research varies from 30 to 340. Considering the extreme cases and a reasonable processing time, 500 Monte Carlo simulations are used for deriving the load model responses of each load category.) Figure 3.8b), the probability histogram, illustrates the distribution of the steady-state real power after voltage drop for all possible responses. The most probable steady-state value of the real power response (corresponding to the point where the probability reaches its maximum) can be determined from this plot. Figure

3.8c) establishes the range of possible real power responses with clearly indicated upper and lower limit (derived as described in Section 3.3.2 for the probabilistic decomposed daily loading curve) as well as the most probable real power response of the EEL (obtained from Figure 3.8a)) to a 5% voltage step reduction.

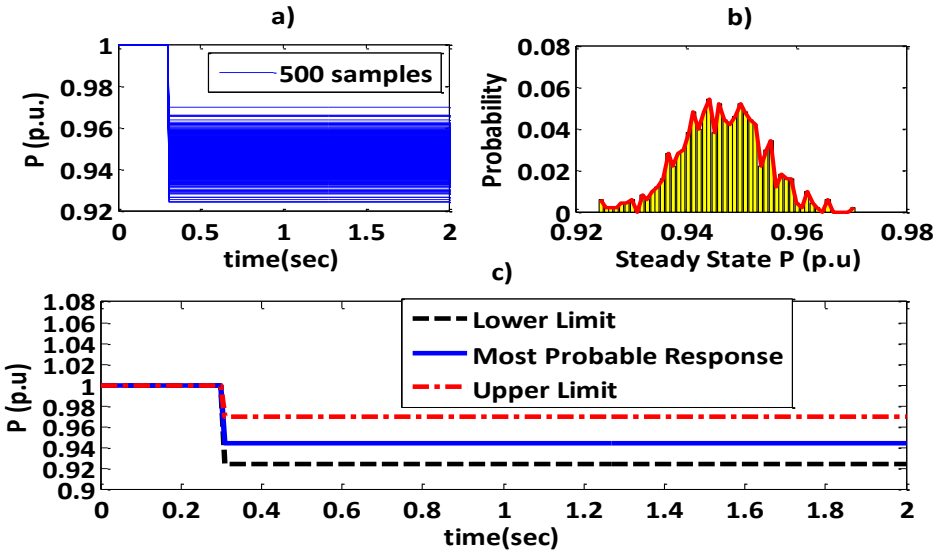


Figure 3.8. Real power responses to a step reduction in voltage for EEL: a) responses obtained from 500 Monte Carlo simulations; b) probability histogram of steady-state power after voltage drop; c) upper and lower limit of the responses and the most probable response

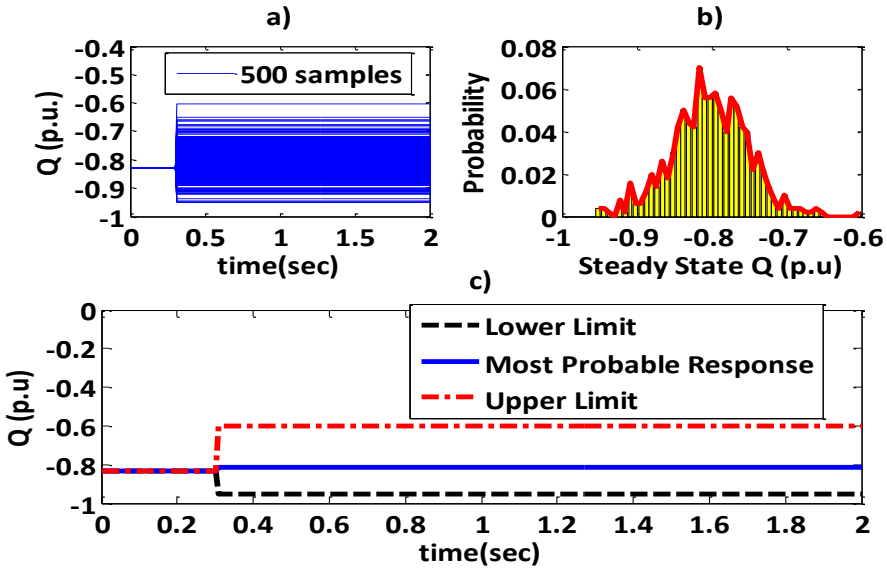


Figure 3.9. Reactive power responses to a step reduction in voltage for EEL: a) responses obtained from 500 Monte Carlo simulations; b) probability histogram of steady-state power after voltage drop; c) upper and lower limit of the responses and the most probable response

A similar approach could be used to obtain the responses of $Q$ and they are plotted in Figure 3.9.

The same approach can be used to assess the responses of EEL to larger voltage drops assuming that there is no disconnection and automatic restart of parts of EEL. The Monte Carlo simulation can be also used to simulate disconnection and reconnection of parts of the load and subsequent derivation of relevant parameters based on obtained responses. This however, is out of the scope of this research.

### 3.4.2 Directly-connected Induction Motor Load

Directly-connected induction motor (IM) load includes residential-commercial induction motors, small industrial induction motors and large industrial induction motors. Due to the noticeable differences of dynamic characteristics among them, their models should be developed separately.

#### 3.4.2.1 Residential and Commercial Induction Motors

Normalised real power responses (to 5% step reduction in voltage) of residential and commercial appliances (dish washer, refrigerator, room air conditioner, washing machine and dryer) are shown in Figure 3.10.

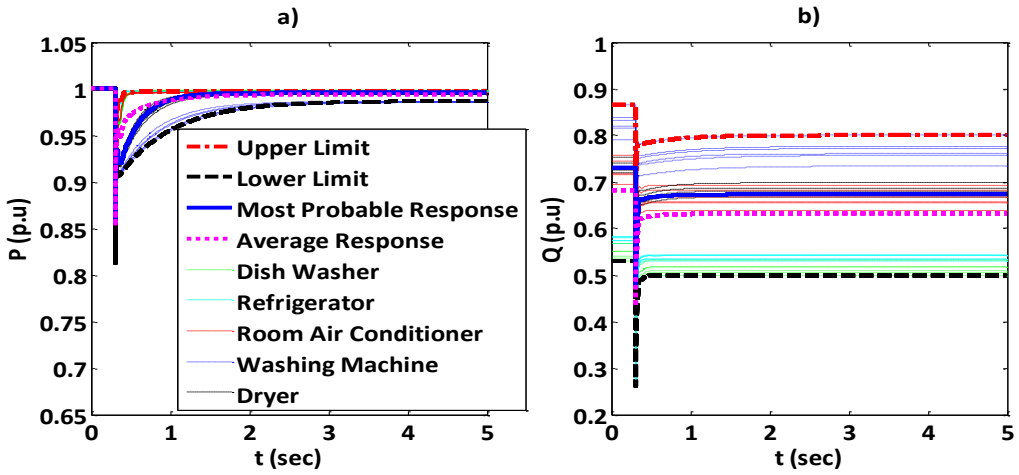


Figure 3.10. a) Real power responses, and b) reactive power response for residential and commercial motors with most probable and average response specified

Responses are generated in DIgSILENT using the 'Asynchronous Machines' option/tool. Six model parameters including rotor resistance, rotor reactance, stator resistance, stator reactance, magnetic reactance and inertia are varied as normally distributed random variables (using 500 Monte Carlo simulations) within a range according to the parameters provided in [1]. Reference [125] provides mathematical deduction for the number of Monte Carlo simulations needed to produce statistically relevant samples. According to [125], required number of Monte Carlo simulations for different parameters of different individual load models for the case studies used in this thesis varies from 30 to 340. Therefore, 500 Monte Carlo simulations are used for deriving the load model responses of each load category to ensure required accuracy of results within reasonable computation time even in most extreme cases.

Based on the results obtained, the upper and lower limits are established as well as the most probable response for this type of load. It can be seen that the responses have relatively similar initial transient power drop and steady-state values, as well as comparable recovery time constants. The real power response of this type of load to small variation in supply voltage can be therefore represented by a first-order exponential recovery model, shown as Equation (2.8) and (2.9). Following the application of Laplace Transform to the equations that made up the model, the analytical solution can be derived as

$$P_l = P_0\left(\frac{V}{V_0}\right)^{\alpha_s} + P_0\left[\left(\frac{V}{V_0}\right)^{\alpha_t} - \left(\frac{V}{V_0}\right)^{\alpha_s}\right]e^{-\frac{t-t_0}{T_p}} \tag{3.4}$$

The first term, known as the steady state component, is the power consumption after the dynamic response stabilizes. The second term, known as the transient component, describes the power drop immediately after the voltage disturbance, which recovers with time constant $T_p$.

In this model, $P_0$ and $P_l$ are initial real power before the voltage change and real power consumption respectively. $V_0$ is the initial RMS value of voltage. Coefficient $\alpha_s$ and $\alpha_t$ are steady-state and transient real power voltage exponent, respectively and $T_p$ is the real power recovery time constant. Time $t_0$ is the moment when the voltage disturbance occurs. The reactive power responses of these loads are shown in Figure 3.10b) and Equation (3.4) is applicable. The mean value and standard deviation of model parameters are shown in Table 3.2, where $\beta_s$ and $\beta_t$ are steady-state and transient reactive power voltage exponent respectively and $T_q$ is the reactive power recovery time constant. The values are obtained by calculating parameters for all sampled $P$ and $Q$ responses first, and then finding corresponding mean and standard deviation values. After the parameters of the model are calculated for each response, the response of the model with estimated parameters is compared with the sampled response. It was found that the two responses match very well. The most probable values (MPV) of parameters, estimated from probability histograms of the parameters (similar to the one shown in Figure 3.8b)), are also included in Table 3.2.

Table 3.2. Model Parameters for General Residential-commercial Motors (Derived using the Data from [1, 4])

| | $\alpha_s$ | $\alpha_t$ | $T_p(sec)$ | $\beta_s$ | $\beta_t$ | $T_q(sec)$ |
|---|---|---|---|---|---|---|
| Mean | 0.1070 | 1.0411 | 0.2047 | 1.4826 | 1.9916 | 0.1557 |
| Standard Deviation | 0.0866 | 0.6064 | 0.2279 | 0.2063 | 0.1189 | 0.2544 |
| MPV | 0.0823 | 1.6287 | 0.2800 | 1.7016 | 1.9352 | 0.2500 |

### 3.4.2.2 Industrial Induction Motors

Real power responses for small and large industrial induction motors (IMs) are more oscillatory than those for residential and commercial motors, as shown in Figure 3.11. The responses are generated in DigSILENT, in a similar way as before, using parameter ranges provided in [1] and in DIgSILENT model library.

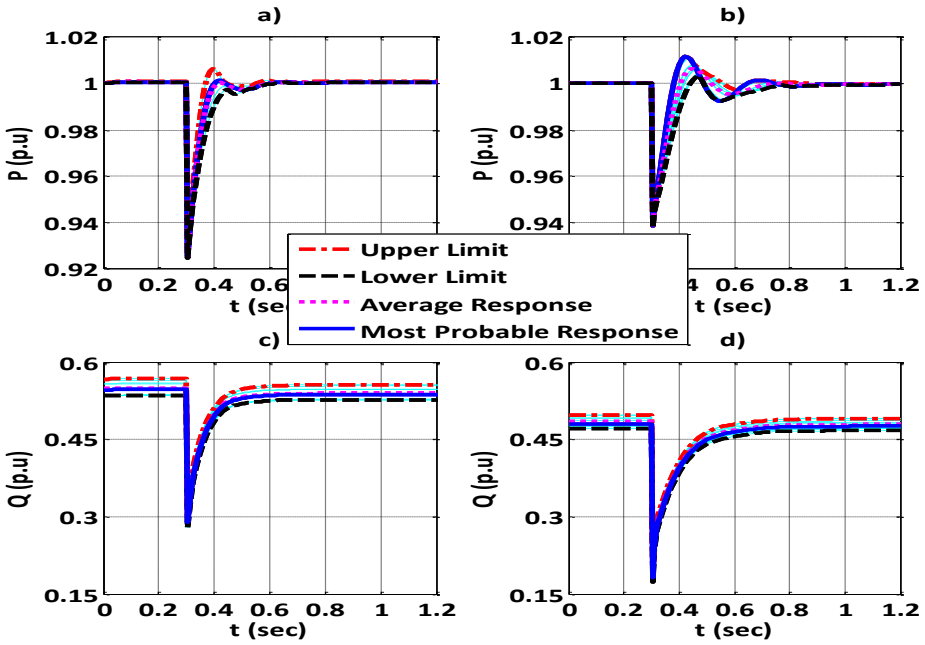


Figure 3.11. a) Small industrial IM P response, and b) large industrial IM P response to step down voltage; c) small industrial IM Q response, and d) large industrial IM Q response to step down voltage. (The upper and lower limit and the average responses are also shown.)

A second-order transfer function model [1] is used to model real power response of this type of load. The model is given by Equation (2.21). It can alternatively be written as

$$\frac{\Delta P(s)}{\Delta V(s)} = \frac{b_1 s + b_0}{s^2 + a_1 s + a_0} \tag{3.5}$$

where $a_0$, $a_1$, $b_0$, $b_1$ are coefficients of the model and $a_1^2 - 4a_0 < 0$. The analytical solution of real power for this type of load depends on the type of voltage change, i.e., $\Delta V(s)$ in (3.5). For a step up or step down voltage change, applying Laplace Transform to (3.5), the analytical solution of real power is derived as

$$P(t) = \left( P_0 + \frac{b_0}{a_0} \Delta V \right) + e^{-\frac{a_1}{2}(t-t_0)} \left[ \frac{2a_0 b_1 - a_1 b_0}{2a_0 \sqrt{a_0 - \frac{a_1^2}{4}}} \sin\left( \sqrt{a_0 - \frac{a_1^2}{4}} (t - t_0) \right) - \frac{b_0}{a_0} \cos\left( \sqrt{a_0 - \frac{a_1^2}{4}} (t - t_0) \right) \right] \Delta V \tag{3.6}$$

$\Delta V$ is the change in voltage magnitude; it is positive for a step-up voltage change and negative for a step-down voltage change. Other parameters have the same definitions as defined in (3.4) and (3.5). All variables are in per unit. Similar to (3.4), (3.6) also has the steady-state component and the transient component. The first term is the final steady-state $P$ consumption. The second term is the transient component, which decays to zero with time. In general, $a_0 >> b_0$ [22]. Therefore, the new steady-state $P$ does not significantly deviate from the initial value. This matches the dynamic responses shown in Figure 3.11a) and Figure 3.11b), as well as the discussion of the static load model of induction motor mentioned in Section 2.1.1.4, Chapter 2 and Appendix C. The range of responses (the upper and lower limit) and the average values are also shown in Figure 3.11.

The average response is calculated in a similar way as before, i.e., based on individually determined model parameters. The most probable responses were calculated as well and also shown in Figure 3.11. They were determined by detecting the most probable value of power at given time and then reconstructing the response. Though there is a difference between most probable and average response, the responses are qualitatively the same. The range of the parameters, as well as the average and the most probable values (MPV) of parameters are given in Table 3.3.

The reactive power responses for this type of induction motors (IMs) are shown in Figure 3.11c) and Figure 3.11d). They are fitted by the first order exponential recovery model, given by Equation (3.4).

The similar ranges of responses (possibly more accurate) for both residential and commercial appliances and industrial IMs could be obtained by considering actual measured responses of number of these devices in field or laboratory measurements, both to small and large voltage disturbances. In spite of available measured responses

though, some simulations would need to be carried out using typical parameters due to large variety of motors, in which case the approach above would need to be used.

Table 3.3. Model Parameters for Small and Large Industrial IM (Derived using the Data from [1, 4])

| Real power | | $a_0$ | $a_1$ | $b_0$ | $b_1$ |
|---|---|---|---|---|---|
| Small | Upper | 3000 | 70 | 2.5 | 150 |
| | Lower | 1600 | 60 | 2.5 | 150 |
| | Average | 2100 | 55 | 2.5 | 150 |
| | MPV | 2100 | 62 | 2.5 | 150 |
| Large | Upper | 1800 | 50 | 0.1 | 110 |
| | Lower | 1200 | 30 | 0.1 | 80 |
| | Average | 1500 | 45 | 0.1 | 100 |
| | MPV | 1500 | 30 | 0.1 | 85 |
| Reactive power | | $\beta_s$ | $\beta_t$ | $T_q(sec)$ | |
| Small | Upper | 0.3691 | 4.9735 | 0.0600 | |
| | Lower | 0.2735 | 4.7946 | 0.0500 | |
| | Average | 0.3174 | 4.9050 | 0.0590 | |
| | MPV | 0.2735 | 4.9735 | 0.0600 | |
| Large | Upper | 0.2562 | 8.1850 | 0.0900 | |
| | Lower | 0.1515 | 7.8799 | 0.0900 | |
| | Average | 0.2082 | 8.0136 | 0.0900 | |
| | MPV | 0.2158 | 7.9832 | 0.0900 | |

### 3.4.3 Resistive Load

The category of resistive load typically includes space heaters, water heaters, ovens and deep fryers [1]. They are static loads and generally have constant impedance characteristic for real power. Therefore, an exponential load model [1] with real power voltage exponent $n_p$=2 is adopted to represent real power response of resistive load. The range of variation in real power responses including the upper and lower limit as well as the most probable values for resistive load is established in similar way as for efficient lighting loads in previous section. The reactive power of resistive load is always zero.

### 3.4.4 Power Electronics / Switch Mode Power Supply Load (SMPS)

The type of load whose participation in the overall demand is increasing at the highest rate is the power electronics based load, categorised as the switch mode power supply (SMPS) load in this thesis. Similar to lighting load, as the dynamic processes

associated with operation of this type of load have time constants much faster than a general electro-mechanical dynamic process, power electronics/SMPS load is modelled by an exponential model given by (3.2) and (3.3). Following a similar procedure as described before, individual responses are generated for randomly chosen values of typical parameters and the most probable response is established. All relevant model parameters adopted from [1, 124] and validated by measurement data provided in [30] are listed in Table 3.4. It was found that this type of load can be modelled, in first approximation, as constant power load for both real and reactive power. (Note: same as the case for efficient energy lighting, the minus sign in front of coefficient $d_q$ indicates that this load category always has leading power factor.)

Table 3.4. Model Parameters for Power Electronics/SMPS (Adopted from [1, 124])

| | $d_p$ | $ep$ | $d_q$ | $eq$ |
|---|---|---|---|---|
| Mean | 0.934 | -0.186 | -0.26 | 0.376 |
| Standard Deviation | 0.002 | 0.124 | 0.062 | 0.146 |

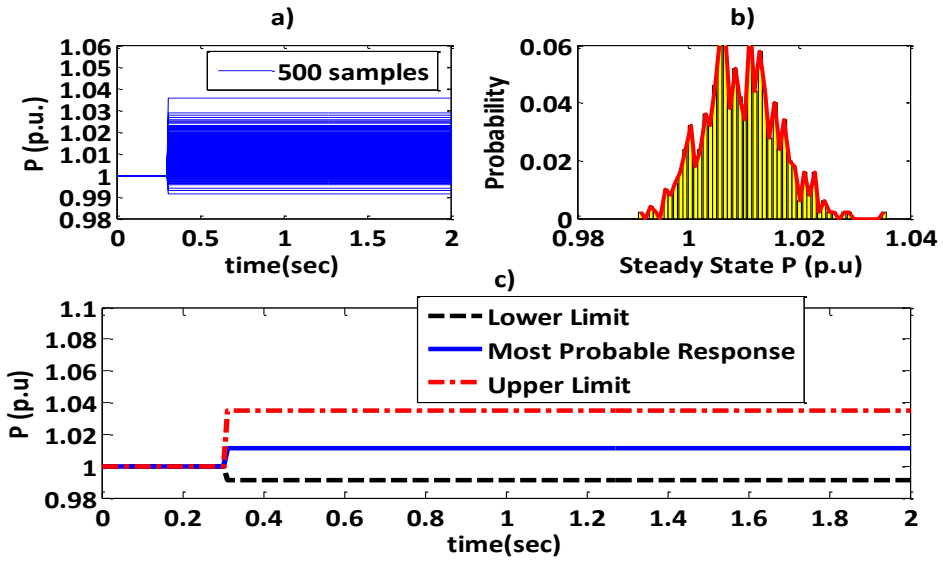


Figure 3.12. Real power responses to a step reduction in voltage for SMPS: a) responses obtained from 500 MC simulations; b) probability histogram of steady-state power after voltage drop; c) upper and lower limit of the responses and the most probable response

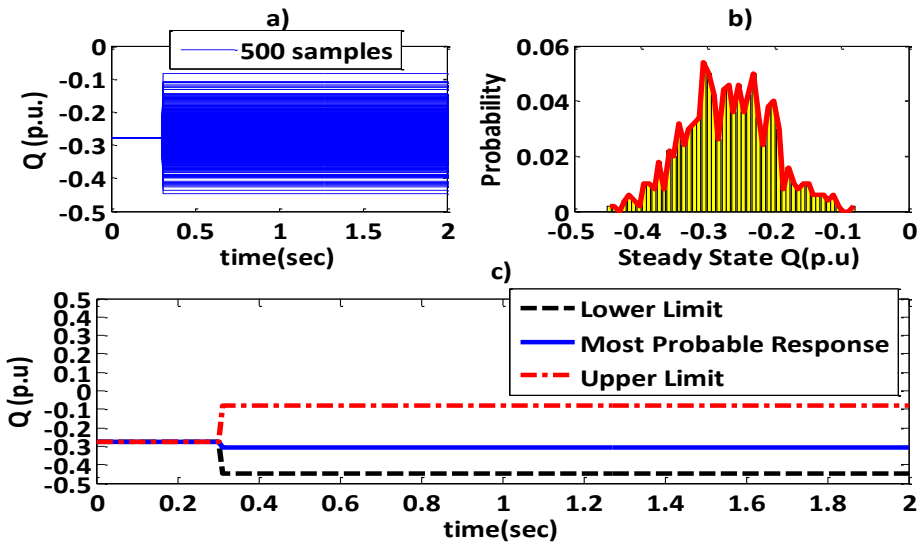


Figure 3.13. Reactive power responses to a step reduction in voltage for SMPS: a) responses obtained from 500 MC simulations; b) probability histogram of steady-state power after voltage drop; c) upper and lower limit of the responses and the most probable response

All possible responses for $P$ and $Q$ are generated using Monte Carlo simulation as before and the most probable and the average responses were found to be almost identical. Similar to EEL, the corresponding responses for $P$ and $Q$ of SMPS are shown in Figure 3.12 and Figure 3.13 respectively.

It is worth noting that the responses of rectifier are similar to those of SMPS, except that the power factor of rectifier is about 0.6 leading [30] whereas the power factor of SMPS is about 0.96 leading [1, 30, 124].

## 3.5 Estimation of Dynamic Response of Demand

Once the individual dynamic load signatures and the load participation in decomposed daily loading curves are available, the aggregate load dynamic response at any given time at bulk supply bus can be established, as indicated by Figure 3.1, by scaling individual load category dynamic responses derived in Section 3.4 by corresponding load contributions at given time derived in Section 3.3 using Equation (3.1). All the simulations related to the dynamic response of the demand at the bulk

supply bus are produced in MATLAB.

### 3.5.1 Most Probable Real and Reactive Power Response during the Day

With the most probable dynamic response of demand (DRD) calculated in Section 3.4 and the most probable decomposed daily loading curve (DDLC) calculated in Section 3.3, using (3.1), the resulting most probable real and reactive power responses to 5% drop in supply voltage throughout the day at bulk supply bus (DDLC shown in Figure 3.5) are shown in Figure 3.14. Figure 3.14 is named "daily aggregate demand response surface (DADRS)". An alternative view of Figure 3.14 is shown in Figure 3.15, where the *t-axis* and the *Hour-axis* are swapped.

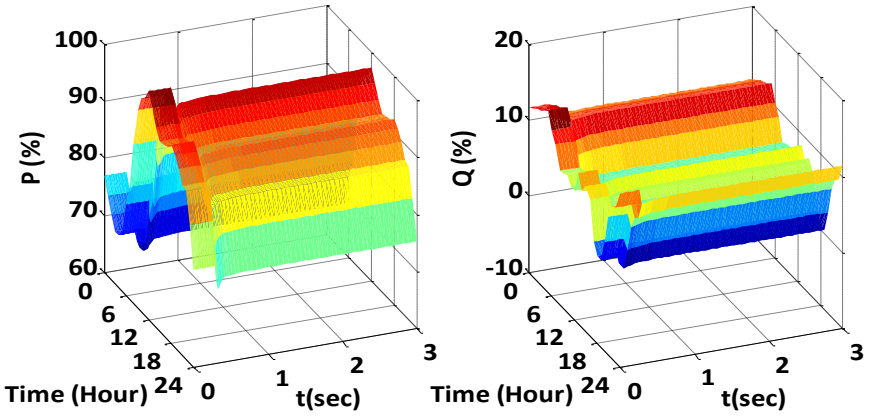


Figure 3.14. The most probable DRD of *P* and *Q* to voltage step during the day

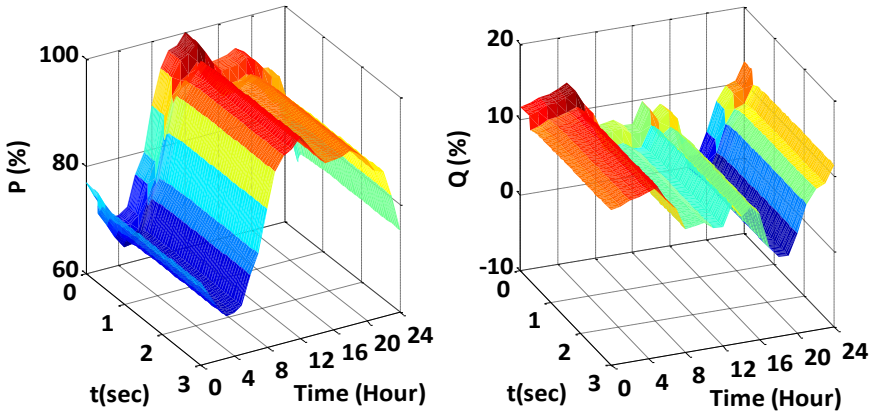


Figure 3.15. The most probable DRD of *P* and *Q* to voltage step during the day

Viewed from *P-t* or *Q-t* Plane, it gives the shape of dynamic real or reactive power responses. While viewed from *P-Hour* or *Q-Hour* Plane, it shows the shape of daily loading curves. It incorporates the information of dynamic response of the demand to a voltage drop and the actual daily demand.

### 3.5.2 Monte Carlo Case Studies and the Most Probable Dynamic Response of Demand

With established mean and standard deviation of hourly load category participation in the decomposed daily loading curve (DDLC), Monte Carlo simulation is applied to generate different load category participations in DDLC at any given time. The percentages of load category participation in load mix at given time are generated randomly within the $\mu \pm 3\sigma$ interval following corresponding normal distributions shown in Figure 3.5 and Figure 3.7 and ensuring that the sum of load category participation at any given time is 100%. Similarly, the dynamic responses of the demand (DRDs) of individual load categories are generated randomly (according to actual probability distributions of parameters of individual load categories discussed in Section 3.4). The number of Monte Carlo simulations for both customer participation and individual load DRD at each hour was 5000 in accordance to [125].

With all generated customer participation data and dynamic responses of the demand (DRDs), using (3.1), 5000 aggregate DRDs are produced. Similar to the approach described in Section 3.3 and Section 3.4, the "range" and the most probable DRD are estimated for every hour of the day to ensure adequate coverage of possible responses. (Similar to estimation of the probabilistic decomposed daily loading curve and probabilistic individual DRD, this range encompasses the vast majority of all possible responses, i.e., the probability of the event that the response will fall outside the range is small enough to be neglected.) Only some characteristic hours are

discussed here for the illustration of the approach, namely 3:00 (lowest demand), 4:00 (largest participation of IMs), 12:00 (peak demand) and 18:00 (lowest participation of IMs).

Figure 3.16 shows, as an example, all possible responses (as percentage of peak demand) generated for 4:00am.

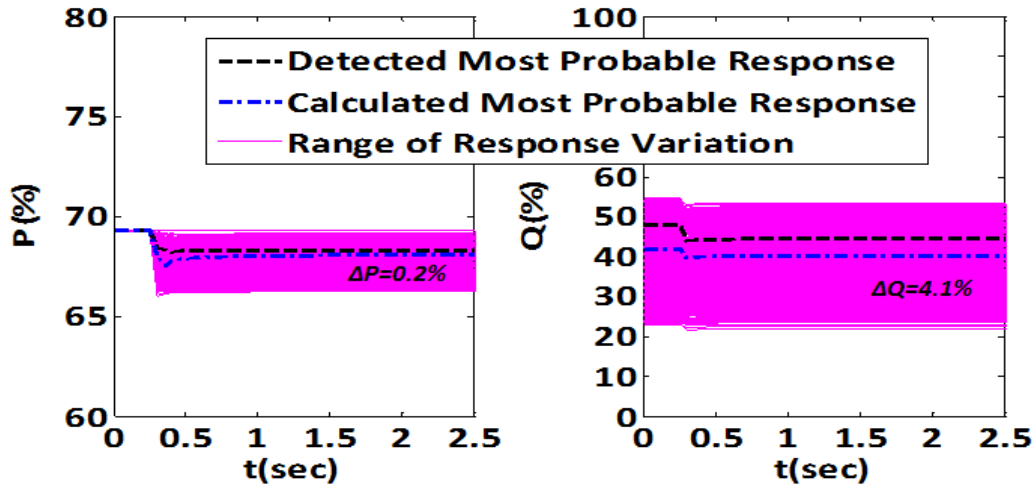


Figure 3.16. Range and most probable DRDs for *P* and *Q* at 4:00am

According to a similar approach of deriving the most probable response for the dynamic response of the demand (DRD) of individual load categories described in Section 3.4, the most probable DRD is established as the most frequent response among all obtained responses. This response is obtained by simultaneous variation of load category participation in generation mix and the individual load category response (i.e., corresponding load model parameters). It is indicated by a dashed line in Figure 3.16 and referred to as detected most probable response. Slightly different responses were obtained when only the most probable response for each individual load category (established in Section 3.4) and the most probably load participation in hourly demand mix (obtained in Section 3.3) were used. This response is indicated by a dash-dot line in Figure 3.16, and referred to as calculated most probable response. The difference between the two responses is generally very small (the selected case is

actually the one where the observed difference is the largest) with deviation in steady state value up to 1% for $P$ and up to 5% for $Q$. It is important to note that in a few cases, qualitative difference in the nature of response was also observed. For example, one curve indicates static response, while the other shows first order exponential response. This is discussed further in the following subsection. In this case, the dynamic response has time constant typically shorter than 0.1 sec, so it could be approximated by static response for most practical power system dynamic studies.

### 3.5.3 Different Dynamic Responses of the Demand at Different Hours

In order to highlight observed differences in dynamic responses of the demand (DRDs) at different time of the day, the most probable $P$ and $Q$ responses at characteristic hours are plotted on the same graph in Figure 3.17. Figure 3.17 shows $P$ and $Q$ responses at four specific hours during the day as well as normalised relative variation in $P$ and $Q$. It can be seen that there is a noticeable difference in transient drops and steady-state values of responses at different hours as well as moderate difference in the recovery time during the transient process. The reason is that the most probable load participations at different hours are different.

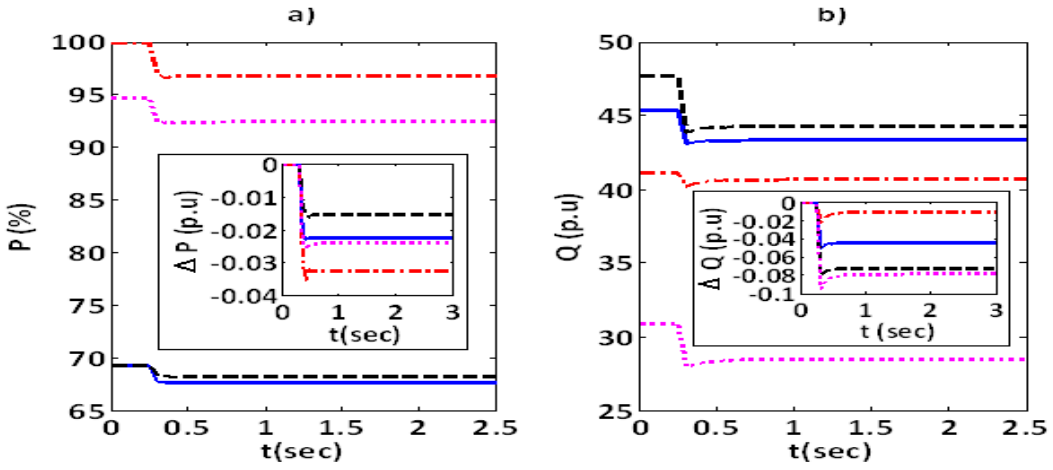


Figure 3.17. Comparison of different most probable a) $P$ and b) $Q$ responses at different times of day (solid line: 3:00; dashed line: 4:00; dash-dot line: 12:00; dotted line: 18:00)

### 3.5.4 Illustration of Possible Dynamic Responses of the Demand

Previous analysis illustrated that dynamic responses of the demand (DRDs) change during the day. The observed differences, for the considered case study, are not significant enough though. Considering all uncertainties involved and identified ranges of possible responses at given time and assuming reasonably small overall variation in composition and responses of individual loads (as previously discussed), two representative DRDs among all possible responses of *P* and *Q* at 12:00 hours are selected and illustrated in Figure 3.18.

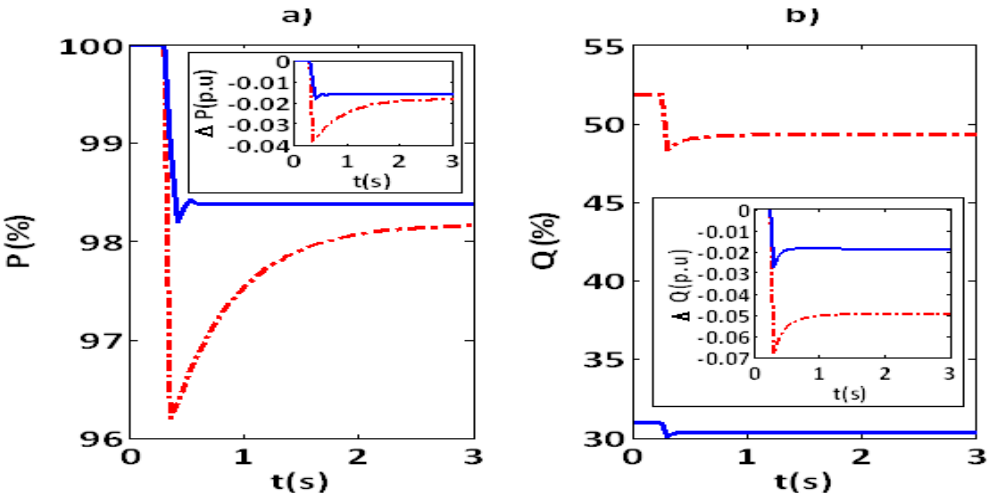


Figure 3.18. Possible difference in a) *P* and b) *Q* response at 12:00 (dashed line and dashdot-line represents two possible different responses among all possible responses)

It can be seen that even at the same time of the day, the dynamic responses of the demand (DRD) could be significantly different. The response of *P* varies from almost static (very fast second order initial transient) to clearly first order exponential recovery while the response of *Q* exhibits different transient drops even though the shape of the response is similar. It is the uncertainty in demand composition that causes the difference in DRD. It is worth noting that the *P* and *Q* responses shown in Figure 3.18 are not correlated (in spite of using the same line type); they are shown here for illustrative purposes only.

## 3.6 Summary

Due to the uncertainty in demand compositions and individual load type static and dynamic characteristics (in particular for new types of device), it is very difficult to establish accurate dynamic responses of the demand (DRDs) without continuous or, at least initially long term measurements at the bus. Both of these might be prohibitively expensive and impractical for real life implementation. Estimation and prediction of DRDs without field measurement, on the other hand, could greatly facilitate demand contribution to static and dynamic operation of future power networks. The methodology presented in this chapter is inspired by conclusions of recently completed work of *CIGRE WG C4.605 Modelling and aggregation of loads in flexible power networks* and enables estimation of DRDs at any bulk supply bus in the network based on reasonably limited information about actual demand composition at the bus and known generic dynamic signatures (i.e. responses) of individual load categories. It combines probabilistic demand compositions at different times of the day at given buses (obtained by processing and classifying large amount of demand data available from previous demand surveys for the bus of interest or for buses with similar demand composition) and individual probabilistic dynamic responses of different load categories (obtained from field/laboratory measurements or computer simulations) to estimate dynamic responses of demand (real and reactive power responses) at different times of the day. It can be seen that the variation in the participation of different load categories results in the variation in dynamic responses of the aggregated demand. Furthermore, the most probable response derived from all responses is very close to the one calculated from the most probable dynamic responses and participation of individual load categories; this finding will be useful for the validation process of the proposed methodology discussed in detail in Chapter 6.

Even though the methodology is illustrated here using only dynamic responses of different load categories to small disturbances in supply voltage, a similar approach can be equally well applied to analyse the dynamic response of the demand (DRD) of the aggregate load to large disturbances if the DRDs of the individual load to large disturbances are available since the models and parameters of individual load categories are input values to the methodology.

# 4 Artificial-Intelligence-based Load Forecasting

Accurate prediction (forecasting) of the load plays an indispensable role in power system planning and electricity market analysis. Load forecasting based on artificial intelligence (AI) techniques received significant attention in the past and it is rapidly developing because of its high accuracy. Some of the AI based methodologies for load forecasting have already been adopted and widely used by the industry. From the literature review, it was found that on the one hand, among the nine most frequently used approaches in the past, artificial neural network (ANN) is the most efficient one with the highest accuracy. The adaptive neuro-based fuzzy inference system (ANFIS), on the other hand, is a new hybrid approach for load forecasting that combines techniques of both ANN and fuzzy logic. However, the performance of these two methods (when applied to load forecasting in the same operation environment) has not been compared in the past to establish whether one of them is superior to the other.

This chapter presents a comparative analysis of the state of the art of ANN and ANFIS load forecasting methodologies in the same operating environment. It demonstrates that, by the appropriate setting of the relevant parameters, a very high

level of accuracy of load forecasting can be achieved by both of them. The mean absolute percentage error (MAPE) is adopted as the first performance indicator in this study. The second performance indicator used is the processing time. Based on the defined performance indicators for different parameter settings, the optimal set of parameters is proposed for both algorithms. Finally, based on the conclusion drawn from the comparison and considering the load composition forecasting approach (i.e. this will be introduced in later chapters), an approach is selected, improved and implemented for both one-day-ahead real and reactive power predictions, and actual demand data (both in MW and MVAR) is used to validate the predictions. With the development of the load disaggregation approach (described in Chapter 5), load forecasting discussed in this chapter can be integrated with load disaggregation to enable prediction of load composition and its controllability, prediction of dynamic responses of demand, load shifting and shaping of dynamic responses of demand (described in Chapter 6).

## 4.1 Importance of Comparing AI-based Load Forecasting Approaches

As the prediction accuracy and processing time are dependent on a variety of factors including the number of inputs, parameter configuration in the artificial intelligence (AI) tools, the software package version and the operating system (i.e. Mac OS X, Linux, Windows) used, it is difficult to rank different AI tools based on just one aspect. In [126], both ANN and ANFIS are applied to predict the effect of different parameters to a meat emulsion system (MES), and the performances of both with different parameter configurations for all MES parameters are compared. The result shows that in MES, sometimes ANN performs better than ANFIS and sometimes vice versa. Therefore, it is important to compare the performance of ANN and ANFIS on load forecasting with multiple factors considered first and then to

choose appropriate parameter settings to enhance the performance accuracy of each of them prior to carrying out further comparisons and making any conclusions about their performance.

Apart from the prediction of total demand, the prediction of time-varying load compositions is gaining the attention of electricity companies and DNOs because of its potential ability to inform demand side management decisions and help with power system planning. The time varying load composition data for training and validation of forecasting tools however, are rare. As a result, it is necessary to use available demand data (much easier to obtain from substation operation data) to train, validate and finely tune developed AI tools for load forecasting first, and then further develop it to load composition prediction.

The artificial neural network (ANN) and the adaptive neuro-based fuzzy inference system (ANFIS) and their applications to load forecasting are introduced in the following sections.

## 4.2 Artificial Neural Network

Two-layer ANNs are adopted in this analysis as they can potentially represent almost all input-output relationships with a finite number of discontinuities as long as an appropriate number of neurons is assigned to the hidden layer [127].

### 4.2.1 ANN Type and Structure

There are mainly two types of ANN, feed-forward ANN (FFANN) and cascade-forward ANN (CFANN).

#### 4.2.1.1 Feed-forward Artificial Neural Network (FFANN)

The structure of a two-layer feed-forward artificial neural network (FFANN) with a hidden layer and an output layer is shown in Figure 4.1.

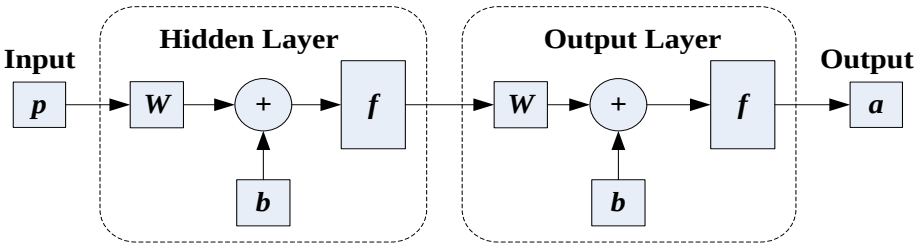


Figure 4.1. Structure of an FFANN

Either layer contains an input vector, a weight matrix $\boldsymbol{W}$, a bias vector $\boldsymbol{b}$, a sum operator, a transfer function $f$ and an output vector $\boldsymbol{a}$. The weighting matrix weighs the input elements, the bias vector biases the weighed inputs via the sum operator, the sum operator gathers the weighed inputs and the biases to produce an intermediate variable for the transfer function, and the transfer function produces the final output of the layer. The output of the hidden layer is the input of the output layer. The relationship between the input and the output in either layer can be represented by

$$\boldsymbol{a} = f\left(\boldsymbol{W}^{T}\boldsymbol{p} + \boldsymbol{b}\right) \tag{4.1}$$

The full description of this type of ANN and its parameter setting rules are given in [78, 127, 128].

#### 4.2.1.2 Cascade-forward Artificial Neural Network (CFANN)

The structure of a two layer cascade-forward artificial neural network (CFANN) is shown in Figure 4.2. Apart from the connections discussed in the case of FFANN, CFANN has connections between the input and each layer and connections between each layer and its successive layers. Such connections are likely to improve the learning speed of ANN [127].

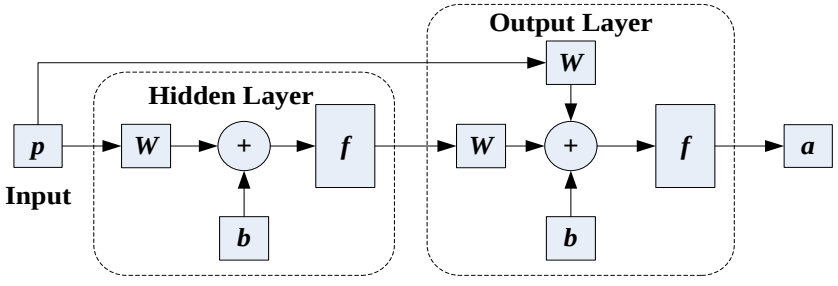


Figure 4.2. Structure of a CFANN

### 4.2.2 Main Parameters

The main parameters of an ANN to be configured include the type of ANN, the hidden layer size, transfer functions and the training algorithm.

#### 4.2.2.1 Type of ANN

Either the feed-forward ANN or the cascade-forward ANN can be selected as the tool for load forecasting. In MATLAB ANN Toolbox, feed-forward ANN and cascade-forward ANN are created by functions *newff()* and *newcf()* respectively.

#### 4.2.2.2 Hidden Layer Size

In [129], a novel approach is proposed to optimise the number of neurons in the hidden layer to avoid over-fitting, and the mathematical approximation for $N/d>>30$ (i.e. $N$ is the number of the training sets, and $d$ is the input dimension) is shown as

$$n = \sqrt{\frac{N}{d \ln N}} \tag{4.2}$$

where $n$ is the number of neurons in the hidden layer. If $N/d$ is smaller than or close to $30$, optimal $n$ most frequently occurs on its maximum value, which has been proved to be $N/d$ [129]. In [130] and [131], signal-to-noise-ratio figure (SNRF) and genetic algorithm is proposed to optimise the size of hidden layers respectively. It is illustrated there that the hidden layer size is always finally determined by trial and error with different numbers of neurons around the value estimated from (4.2). In most cases however, the nearest integer to the estimation of (4.2) works well.

#### 4.2.2.3 Transfer Function

The transfer function can be any differentiable function. The most commonly used transfer functions are log-sigmoid (*logsig*), tan-sigmoid (*tansig*) and linear (*purelin*) transfer function. The output ranges of the three transfer functions are respectively $[0,1]$, $[-1,1]$ and $[-\infty,+\infty]$ [127].

#### 4.2.2.4 Training Algorithm

Before training, the weights are initialised to small random values. The training process needs a set of examples of appropriately selected network behaviours, inputs and targets. During the training process, the input progresses as indicated by Figure 4.1 or/and Figure 4.2, and the outputs are created and compared with the target values. The differences between the created outputs and the target values are recorded and propagated backwards through the network, and the weights and biases of the network are adjusted to minimise the difference. This process is repeated until either the difference is within a predefined range or the maximum epoch is reached. As the error is propagated backwards, this training process is called the backpropagation training algorithm.

Backpropagation (BP) [128] is widely used in ANN training. The basic backpropagation training algorithm, which is also called gradient descent backpropagation (GDBP), adjusts the network weights and biases in the direction of the negative of gradient, in which the performance function decreases most rapidly. The gradient descent backpropagation (GDBP) is illustrated in iteration as Equation (4.3) [127]

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k \boldsymbol{g}_k \tag{4.3}$$

where $\boldsymbol{x}_k$ is the value of the weight or the bias for the $k^{th}$ iteration, $\alpha_k$ is the learning rate for the $k^{th}$ iteration and $\boldsymbol{g}_k$ is the current gradient for the $k^{th}$ iteration. In MATLAB ANN Toolbox, they can be called by setting the training algorithm parameters to *traingd*.

There are also many variations of backpropagation training algorithm, such as resilient backpropagation (RBP), conjugate gradient backpropagation (CGBP), quasi-Newton backpropagation (QNBP), Levenberg-Marquardt backpropagation (LMBP) and Bayesian Regulation backpropagation (BRBP). These variations optimise the

basic network and speed up the training process.

According to the MATLAB Neural Network Toolbox [127], Levenberg-Marquardt backpropagation (LMBP) is usually preferred and selected as the training algorithm in past works due to its high speed and high performance accuracy. The weights and biases update for LMBP is given as Equation (4.4)

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \left[\boldsymbol{J}^T\boldsymbol{J} + \mu\boldsymbol{I}\right]^{-1}\boldsymbol{J}^T\left(\boldsymbol{t}-\boldsymbol{a}\right) \tag{4.4}$$

where $\boldsymbol{J}$ is the Jacobian matrix which contains 1st partial derivatives of the network error with respect to the weights and the biases, and $\boldsymbol{t}$ is the target value vector, and $\boldsymbol{a}$ is the network output vector. When the Levenberg's damping factor, $\mu$, is zero, it becomes Quasi-Newton algorithm using the approximate Hessian matrix $\boldsymbol{J}^T\boldsymbol{J}$, and when $\mu$ is very large, it becomes gradient descent with a small step size ($\boldsymbol{t}$-$\boldsymbol{a}$). The drawback is that the fast speed of LMBP induces relatively large uncertainties, especially when over-fitting occurs.

Therefore, in specific cases, Bayesian Regulation backpropagation (BRBP) is used to make the network stable and increase the robustness of the ANN. Bayesian Regulation Backpropagation (BRBP) updates the weight and bias based on Levenberg-Marquardt optimisation. It minimises a combination of squared errors and weights and determines the correct combinations in order to produce a network that might generalise better [19]. The drawback of this approach is that it is time-consuming. The results of both algorithms are compared in this chapter. In MATLAB ANN Toolbox, Levenberg-Marquardt Backpropagation and Bayesian Regulation Backpropagation can be called by setting the training algorithm parameters to *trainlm* and *trainbr*, respectively.

## 4.3 Adaptive-neuro Fuzzy Inference System

Adaptive-neuro Fuzzy Inference System (ANFIS) is a Sugeno fuzzy inference

system (FIS). The input membership functions (IMF) are generally adjusted by hybrid algorithms and the output membership functions (OMF) are either a constant or a linear combination of the inputs. It allows the fuzzy inference system to decide the rules itself by learning from the input and the target. Hybrid learning algorithms combine backpropagation and least squares and always perform better than the basic backpropagation algorithm on its own [94].

### 4.3.1 Structure

The structure of an ANFIS with two inputs, two membership functions and one output is shown in Figure 4.3 [94].

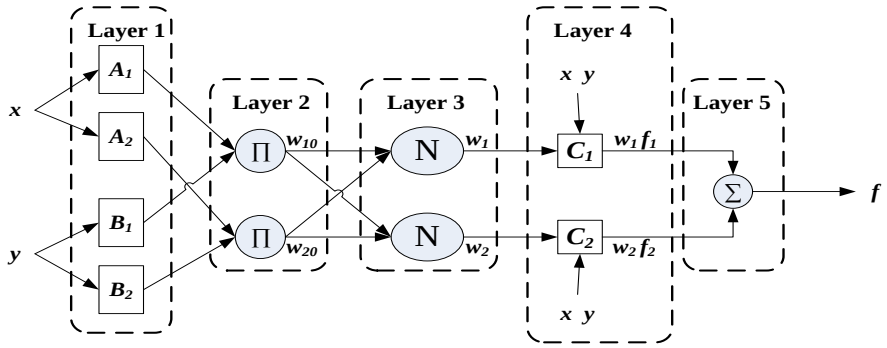


Figure 4.3. Structure of an ANFIS

Layer 1 is the fuzzification layer, where the input data are fuzzified using input membership functions (IMF). The neuron $A_i$, $B_i$ are the linguistic labels which describe the characters of the input data, such as small and large, low and high etc., and are associated with a membership function which specifies the degree to which the given input satisfies the description of the neuron.

Layer 2 executes the fuzzy *AND* Function of the antecedent part of the fuzzy rules. Generally, either soft-min or product [86] is used as the rules in Layer 2. The outputs from Layer 2 are referred to as firing strengths [94].

In Layer 3, each of the firing strengths is normalised and the normalised firing strengths are the weights assigned to corresponding rules.

In Layer 4, the consequent part is executed by the implication method. A consequent is a fuzzy set represented by a membership function [86], referred to as output membership function (OMF). It could be a linear combination of inputs or a constant function. The form of Sugeno rules can be described as '*if x is $A_1$ and y is $B_1$, then f1=f1(x,y,z,c)*'. The input of the implication process is a single number obtained from the antecedent (also a single number). This single number is used to reshape the corresponding consequent. The output of Layer 4 is the reshaped consequent weighted by the corresponding normalised firing rate.

Layer 5 is a sum operator that superimposes all outputs of Layer 4 and produces the final output.

Adaptive neuro-based Fuzzy Inference System (ANFIS) is a robust AI tool, whose output is unique once the number and type of membership functions and the training algorithm are defined.

### 4.3.2 Input and Output Membership Functions

The input membership function (IMF) could be any differentiable function or piecewise function. The most commonly used IMFs include the triangle-shaped transfer function (*trimf*), the bell-shaped transfer function (*gbellmf*) and the difference-sigmoid shaped transfer function (*dsigmf*) [86]. The output membership function (OMF) could be either linear or constant.

## 4.4 General Framework for AI-based Short-term Load Forecasting (STLF)

According to past works, a general framework for day-ahead load forecasting (i.e. predicting demand on Day *N+1* with the data collected on Day *N*) using the artificial intelligence (AI) techniques, is represented by Figure 4.4. The trained AI tool will be used for demand forecasting after it is validated. The inputs (also shown in box

labelled {3}) are: (i) measured demand in MW or/and MVAR on Day *N*; (ii) forecasted weather data (including temperature, humidity and wind speed) for Day *N+1*; (iii) day type (working days, holidays) of Day *N+1*. The outputs are the forecasted demand in MW or/and MVAR. For comparison purposes, only the prediction of real power is implemented. The prediction of reactive power will be implemented after a suitable forecasting approach is selected and upgraded.

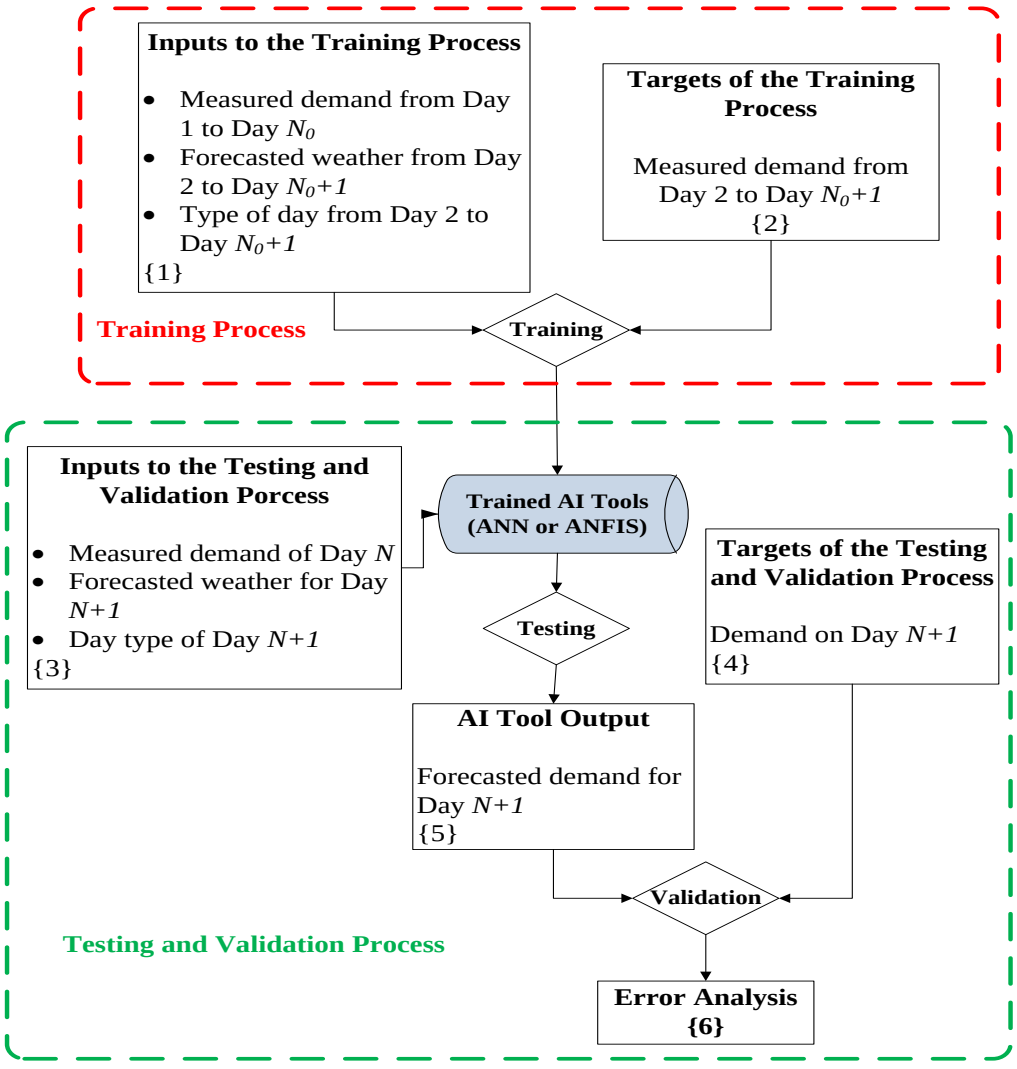


Figure 4.4. Block diagram for demand forecasting

### 4.4.1 Description of Training and Validation Process

In the training process, large quantities of demand data, weather data and day type data are required. The inputs in the training process (box labelled {1}) include

measured demand data from Day 1 to Day $N_0$, forecasted weather data (i.e. temperature, humidity and wind speed) from Day 2 to Day $N_0$+1 and day types (i.e. 1 for working days, 0 for holidays) from Day 2 to Day $N_0$+1. (Note: to distinguish the training process and the general forecasting process, $N_0$ is used instead of $N$ to describe the training process.) It is denoted by ***PTRN*** and if written in the format of the input matrix in MATLAB ANN toolbox, it is in the following form:

$$\boldsymbol{PTRN} = \begin{bmatrix} P_{1,0:00} & P_{1,0:30} & \cdots & P_{k_0,0:00} & \cdots & P_{N_0,23:00} & P_{N_0,23:30} \\ T_{2,0:00} & T_{2,0:30} & \cdots & T_{k_0+1,0:00} & \cdots & T_{N_0+1,23:00} & T_{N_0+1,23:30} \\ H_{2,0:00} & H_{2,0:30} & \cdots & H_{k_0+1,0:00} & \cdots & H_{N_0+1,23:00} & H_{N_0+1,23:30} \\ WS_{2,0:00} & WS_{2,0:30} & \cdots & WS_{k_0+1,0:00} & \cdots & WS_{N_0+1,23:00} & WS_{N_0+1,23:30} \\ DT_{2,0:00} & DT_{2,0:30} & \cdots & DT_{k_0+1,0:00} & \cdots & DT_{N_0+1,23:00} & DT_{N_0+1,23:30} \end{bmatrix} \quad (4.5)$$

where $P$, $T$, $H$, $WS$, $DT$ are demand (real power), temperature, humidity, wind speed and day type, respectively. $N_0$ represents Day $N_0$, and $N_0+1$ represents Day $N_0+1$. $P_{k0,0:00}$ represents the demand at 0:00 on Day $k_0$ ($1 \le k_0 \le N_0$), and similar definitions apply to other elements. The day type can be obtained based on the date when the demand data are collected, represented by 1 for working days and 0 for holidays. The targets are the measured demand data from Day 2 to Day $N_0$+1, correspondingly denoted as ***TTRN*** and written in the form:

$$\boldsymbol{TTRN} = \begin{bmatrix} P_{2,0:00} & P_{2,0:30} & \cdots & P_{k_0+1,0:00} & \cdots & P_{N_0+1,23:00} & P_{N_0+1,23:30} \end{bmatrix} \quad (4.6)$$

The inputs and the targets are used to train and create an AI tool (ANN or ANFIS).

After training, the trained AI tool is validated with another set of data. In the testing and validation process, the inputs to the trained AI tool (box labelled {3}) are the measured demand of Day $N$, the forecasted weather of Day $N$+1 and the day type of Day $N$+1. Based on these inputs, the trained AI tool will create the outputs (box labelled [1]), which are the predicted demands on Day $N$+1. The predicted demands on Day $N$+1 are compared with the measured demand data on Day $N$+1, which are the targets in the testing and validation process (box labelled {4}), and the error between

the predicted demand and the measured demand is assessed (box labelled {6}). Once validated, the trained AI tool can be used for demand forecasting. The details of MATLAB coding for training and validation used in this study can be found in [127] for ANN and [86] for ANFIS.

### 4.4.2 Data Collection

Demand data used in this study are collected from the distribution network of a local utility. Data are collected every 30 minutes, i.e, 48 samples per day. Measured and forecasted weather data for every 30 minutes, including temperature (°C), humidity (p.u) and wind speed (km/h), are obtained from the official website of Weather Underground [132].

### 4.4.3 Parameter Settings for AI Tools

The parameter setting and the training algorithm selection for both ANN-based and ANFIS-based approach are demonstrated in the following subsections.

#### 4.4.3.1 ANN based methodology

Both FFANN and CFANN are implemented and compared. For the case study used in the thesis, the number of training sets, the number of testing sets and the input dimension are 864 (18 days), 96 (2 days) and 5, respectively. According to (4.2), the estimated value of the optimal neuron number is 5.06. Therefore, the neuron number in the hidden layer is set to 5 and 6, and the results for both settings are compared.

Considering the nonlinearity of the prediction [127], a sigmoid function (*logsig* or *tansig*) rather than *purelin* is selected as the transfer function of the output layer. As *logsig* is too insensitive to its input, if selected as the transfer function of the output layer, it will restrict the final ANN output to a quite small range that may not be able to cover the range of the demand. Therefore, *tansig* is selected as the transfer function of the output layer. According to the shape of *tansig* and considering the ANN output

range, *logsig* is selected as the transfer function of the hidden layer. Default settings are assigned to other parameters as they are applicable to most case studies [127].

Three most commonly used training algorithms discussed in Section 4.2.2.4 are used and compared in the study. The results obtained with different numbers of hidden layers and different training algorithm settings are compared and discussed in subsequent sections.

#### 4.4.3.2 ANFIS based methodology

The most commonly used input and output membership functions with different parameter settings are compared in the study. The input membership functions considered include *trimf*, *gbellmf* and *dsigmf*, and the output membership functions include linear and constant functions. According to the number of inputs considered, the number of input membership functions [94] is set to 2 to ensure a reasonable processing time and to prevent MATLAB toolbox from running out of memory.

### 4.4.4 Testing and Validation

After training, the trained AI tool (ANN or ANFIS) is tested with another set of data and the results are validated against the actual measured demand. As the relative error between the predicted load and the actual load is of more concern than the root mean squared error (RMSE) which is usually used in AI computation, in this study, the mean absolute percentage error (MAPE) defined by

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|\boldsymbol{a}_i - \boldsymbol{t}_i|}{\boldsymbol{t}_i}\times 100\% \tag{4.7}$$

is adopted for error analysis, where $\boldsymbol{t_i}$ is the $i^{th}$ measured demand, $\boldsymbol{a_i}$ is the $i^{th}$ predicted demand, and $N$ is the total number of demands predicted.

## 4.5 Results of Comparison of ANN and ANFIS

In the comparison results, only real power prediction is implemented for

illustration. The approach to reactive power prediction is similar.

### 4.5.1 ANN based methodology

Table 4.1 and Table 4.2 show MAPE (%) and processing time (sec) of a feed-forward ANN (FFANN) and a cascade-forward ANN (CFANN) with different combinations of training algorithms and hidden layer size configurations. The result shows that in both cases, the best accuracy is achieved by adjusting the number of neurons in the hidden layer to 5 and training algorithm to *trainbr*, and that the fastest training algorithm is *trainlm*. With the same parameter configuration, it could be seen that FFANN generally operates slightly faster than CFANN, although the processing time in both cases is comparable. This is probably because CFANN has a more complex structure. Under the same parameter setting, the CFANN does not necessarily improve the prediction accuracy. It can also be seen from Table 4.1 that *trainbr* performs better when the hidden layer size is 5 than when it is 6, while *trainlm*, on the contrary, performs better when the hidden layer size is 6. Therefore, it can be concluded that *trainbr* works more effectively than *trainlm*.

Table 4.1. MAPE and Processing Time of FFANN with Different Parameter Configuration

| | **MAPE (%)** | | | **Processing Time (sec)** | | |
|---|---|---|---|---|---|---|
| **Hidden Layer Size** | ***traingd*** | ***trainlm*** | ***trainbr*** | ***traingd*** | ***trainlm*** | ***trainbr*** |
| **5** | **10.47** | **6.28** | **3.49** | **1.2243** | **0.4140** | **1.4127** |
| **6** | **6.56** | **5.02** | **4.80** | **1.2380** | **0.7399** | **1.5526** |

Table 4.2. MAPE and Processing Time of CFANN with Different Parameter Configuration

| | **MAPE (%)** | | | **Processing Time (sec)** | | |
|---|---|---|---|---|---|---|
| **Hidden Layer Size** | ***traingd*** | ***trainlm*** | ***trainbr*** | ***traingd*** | ***trainlm*** | ***trainbr*** |
| **5** | **8.19** | **5.74** | **4.51** | **1.3717** | **0.6334** | **1.5594** |
| **6** | **7.55** | **7.01** | **5.21** | **1.2758** | **0.6520** | **1.5045** |

From the performance of the tested ANN, it can be concluded that CFANN has a relatively larger performance error than FFANN. The fastest training algorithm is *trainlm*. The slowest training algorithm is *trainbr;* however *trainbr* is typically more stable than *trainlm* and the processing time is acceptable though slightly longer.

According to the conclusion drawn from Table 4.1 and Table 4.2, Table 4.3 provides suggestions (marked with ticks) for selection of training algorithms for ANN methodology for load forecasting when factors such as accuracy, processing time, robustness and effectiveness are considered.

Table 4.3. Suggestions of Training Algorithm Selection

| | ***traingd*** | ***trainlm*** | ***trainbr*** |
|---|---|---|---|
| **Accuracy** | | | ✓ |
| **Processing Time** | | ✓ | |
| **Robustness** | | | ✓ |
| **Effectiveness** | | | ✓ |

### 4.5.2 ANFIS based methodology

The performance result of ANFIS are summarised in Table 4.4. It shows that constant output membership functions (OMFs) always result in lower MAPE and a much faster performance than linear OMFs. For constant OMFs, the processing time is about 1-2s; while for linear OMFs, the processing time is about 11s, i.e., significantly longer. Although linear OMFs with 11s processing time are acceptable in situations with only one case or a few cases to be trained, they are not suitable for situations with a large number of cases. Besides, with the same IMF setting, constant OMFs always perform better than linear OMFs.

Table 4.4. MAPE and Processing Time of ANFIS with Different Parameter Configuration

| | **MAPE (%)** | | | **Processing Time (sec)** | | |
|---|---|---|---|---|---|---|
| **Output &Input MF** | ***trimf*** | ***gbellmf*** | ***dsigmf*** | ***trimf*** | ***gbellmf*** | ***dsigmf*** |
| **Constant** | **3.48** | **4.26** | **6.67** | **1.3229** | **1.4542** | **1.4049** |
| **Linear** | **6.68** | **12.07** | **9.05** | **11.2455** | **11.3320** | **11.2714** |

With the same setting of OMF, *trimf* always processes the data faster than *gbellmf* and *dsigmf*, and its MAPE is always the smallest among the three IMFs. Therefore, if ANFIS is adopted as the forecasting tool, simpler IMFs such as *trimf* generally perform faster and more accurately than more complicated ones such as *gbellmf* and *dsigmf*. Furthermore, a simpler OMF (i.e. constant OMF) always performs much faster and more accurately than a more complicated OMF (i.e. linear

OMF).

According to the conclusion drawn from Table 4.4, Table 4.5 provides recommendations for selection of input membership functions for ANFIS based methodology for load forecasting considering key performance indicators, i.e., accuracy, processing time, and effectiveness. The recommended options are marked with ticks. As the output of ANFIS is unique once the parameter configuration is decided, the robustness is not discussed in the table.

Table 4.5. Suggestions of IMFs Selection with Constant OMF

| | ***trimf*** | ***gbellmf*** | ***dsigmf*** |
|---|---|---|---|
| **Accuracy** | ✓ | | |
| **Processing Time** | ✓ | | |
| **Effectiveness** | ✓ | | |

The results of the comparative analysis of ANN and ANFIS methodologies for efficient and accurate load forecasting based on several performance indicators are summarised in Table 4.6. For each performance indicator, better performance is marked with a tick. It can be seen that the performances of both methodologies are comparable if appropriate parameter configurations are provided.

Table 4.6. Suggestions of ANN and ANFIS Selection

| | **ANN** | **ANFIS** |
|---|---|---|
| **Accuracy** | ✓ | ✓ |
| **Processing Time** | ✓ | **constant OMF comparable** |
| **Robustness** | ***trainbr* comparable** | ✓ |

The load prediction results obtained by both methodologies using appropriately selected parameters as discussed above are plotted against the actual measured demand curve in Figure 4.5(a). It can be seen that the predicted demand obtained from both ANN and ANFIS matches the measured curves well. The absolute percentage error (APE) of the predicted demand defined as

$$APE = \frac{|\boldsymbol{a} - \boldsymbol{t}|}{\boldsymbol{t}} \times 100\% \tag{4.8}$$

for every sampling point is plotted Figure 4.5(b) in addition to MAPE, where $\boldsymbol{t}$ is the

measured demand, ***a*** is the predicted demand. Figure 4.5 shows that with appropriate parameter configurations, both ANN and ANFIS can perform well in demand prediction, and their performances are comparable.

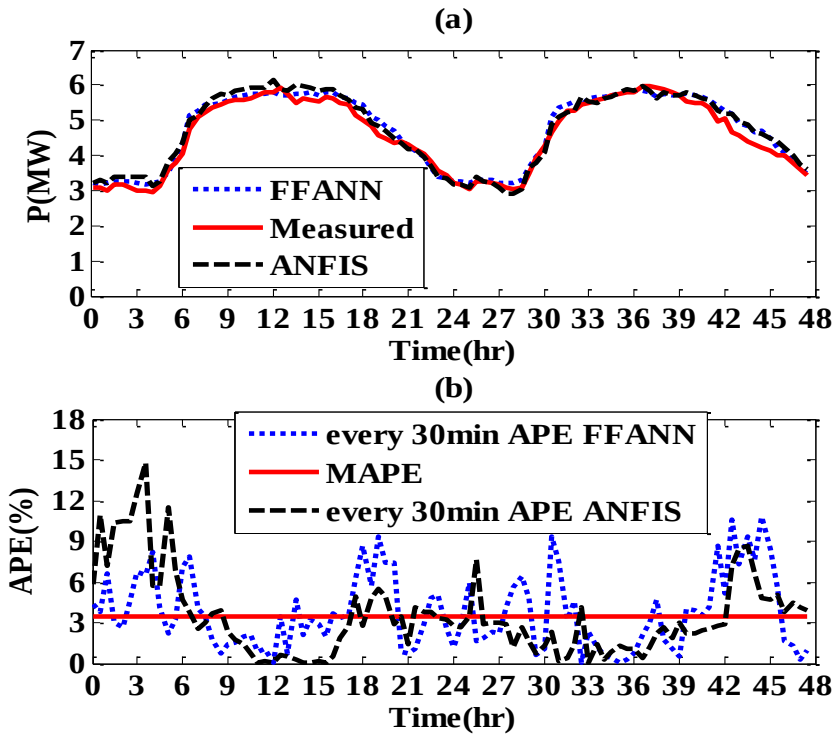


Figure 4.5. (a) Day-ahead forecasted demand curves against actual demand; (b) APE at different sampling times against MAPE

## 4.6 Advantages and Disadvantages of ANN and ANFIS

In spite of the fact that the performances of ANN and ANFIS are comparable, both of them have their own advantages and disadvantages. Therefore, the selection of the methodology in reality depends not only on the prediction accuracy, but also on the specific requirements (e.g. the number of elements in the output and robustness of the algorithm) in different case studies.

The main advantages and disadvantages of both ANN and ANFIS are summarised in Table 4.7. In general, ANN has a stronger adaptation (i.e. more adaptive to the daily variation of the demand) than ANFIS and it can support multi-output modules where the number of the elements in the output is more than one.

However, as the initial weights of ANN are randomly set by MATLAB, the result could be different in different runs, although such a difference is moderate. In ANFIS, once the parameters and membership functions are set, the result is unique and unchanged. Nevertheless, ANFIS can only support single-output modules, where the number of the element in the output is one, and the selection of membership functions in the ANFIS approach is experience-dependent. In other words, ANFIS has a higher robustness and lower uncertainty than ANN.

Table 4.7. Advantages and Disadvantages of ANN and ANFIS

| | ANN | ANFIS |
|---|---|---|
| Advantages | ✓ Strong adaptation<br>✓ Support multi-output module | ✓ High robustness<br>✓ Low uncertainty |
| Disadvantages | • Low robustness<br>• High uncertainty | • Only supports single output<br>• Membership function is experience-dependent |

## 4.7 Upgraded Day-ahead Load Forecasting Approach

As ANN supports the multi-output modules, it is adopted for load disaggregation as well as total demand forecasting in this research.

The total demand forecasting framework as discussed in Section 4.4 is more suitable for forecasting the demand on the target day if there are no significant differences between the demand on two adjacent days (i.e. Day $N$ and Day $N$+1 in Figure 4.4). In the case that the demand on the target day significantly differs from the demand collected on the day before the target day (e.g. the demand pattern could dramatically change when the day transfers from a working day to a holiday or vice versa), the framework given as Figure 4.4 could induce errors. As a result, another approach should be implemented in parallel to cater for the considerable change in demand pattern. Therefore, the upgraded overall methodology for demand forecasting is given as Figure 4.6.

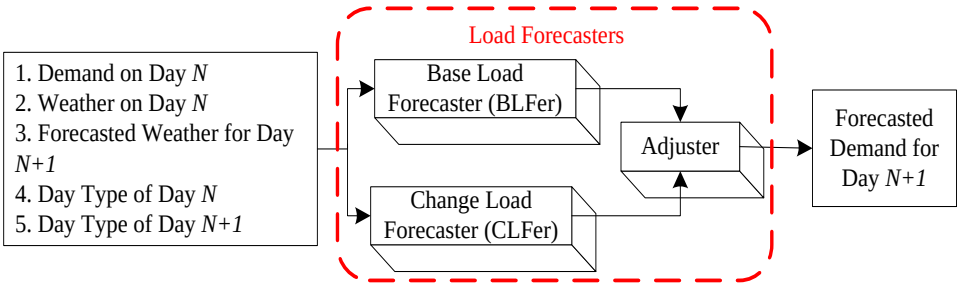


Figure 4.6. Total load forecasting framework

It consists of a *base load forecaster* (BLFer), a *change load forecaster* (CLFer) and an adjuster. *Base load forecaster* (BLFer) is trained to forecast the regular load pattern for the next day. *Change load forecaster* (CLFer) is trained to forecast the change in hourly demand from Day *N* to Day *N*+1 caused by the change of day type or the change of weather. The *base load forecaster* and *the change load forecaster* complement each other. The former emphasises the regular load patterns more and it responds slowly to rapid changes in demand, whereas the latter focuses more on the demand of the day before the target (or predicted) day and it responds more quickly to sudden changes than the former does. Therefore, the combination of the two forecasters leads to improved accuracy, especially in cases when sudden load changes occur [89]. As a result, an adjuster, which can be either a trained ANN or a least square algorithm box, is installed after BLFer and CLFer to take advantage of both. More details about the adjuster are given in later subsections.

This model originated from the 3$^{rd}$ generation of the ANNSTLF tool developed by EPRI [89, 90] and has been upgraded to be applied to the study in this thesis. The difference between the EPRI's ANNSTLF tool and the upgraded tool used in this research is mainly that: in this research, all weather data including temperature, humidity and wind speed are put together in the ANN for training; while in the tool developed by EPRI, an effective temperature which combines all weather data into a cumulative index is obtained and used as an equivalent temperature. The process of derivation for effective temperature is however experience dependent, and it takes

longer to process than directly using all the weather data for ANN training.

As shown in Figure 4.6, the inputs of the upgraded methodology are: (i) measured demand on Day *N*, (ii) weather (temperature, humidity and wind speed) on Day *N*, (iii) forecasted weather for Day *N+1*, (iv) day type of Day *N*, and (v) day type of Day *N+1*. When compared with the inputs given in Figure 4.4, there are two more sets of inputs here, the weather on Day *N* and the day type of Day *N*. They are needed because the demand pattern variations from Day *N* to Day *N*+1 are considered. The output of the upgraded methodology is the forecasted demand for Day *N+1*. The data collection approach is exactly the same as that introduced in Section 4.4. The approach is also applicable to the reactive power prediction.

### 4.7.1 *Base Load* Forecasting

*Base load* forecasting forecasts the regular demand pattern (or the *base load*) of the next day.

#### 4.7.1.1 Training Process of *Based Load* Forecasting

Figure 4.7 shows the training process of *base load* forecasting. It is similar to the training part as shown in Figure 4.4, except that there are more inputs.

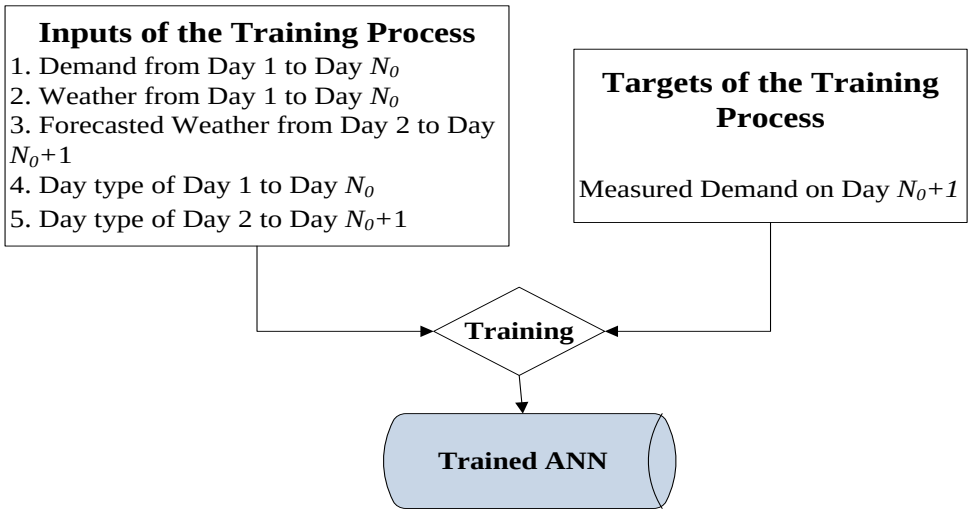


Figure 4.7. Base load forecasting training process

If written in the format of the input matrix and the target matrix in MATLAB programming [127], the input is written as:

$$\boldsymbol{PTRN}_{BLF} = \begin{bmatrix} P_{1,0:00} & P_{1,0:30} & \cdots & P_{k_0,0:00} & \cdots & P_{N_0,23:00} & P_{N_0,23:30} \\ T_{1,0:00} & T_{1,0:30} & \cdots & T_{k_0,23:00} & \cdots & T_{N_0,23:00} & T_{N_0,23:30} \\ H_{1,0:00} & H_{1,0:30} & \cdots & H_{k_0,23:00} & \cdots & H_{N_0,23:00} & H_{N_0,23:30} \\ WS_{1,0:00} & WS_{1,0:30} & \cdots & WS_{k_0,23:00} & \cdots & WS_{N_0,23:00} & WS_{N_0,23:30} \\ T_{2,0:00} & T_{2,0:30} & \cdots & T_{k_0+1,23:00} & \cdots & T_{N_0+1,23:00} & T_{N_0+1,23:30} \\ H_{2,0:00} & H_{2,0:30} & \cdots & H_{k_0+1,23:00} & \cdots & H_{N_0+1,23:00} & H_{N_0+1,23:30} \\ WS_{2,0:00} & WS_{2,0:30} & \cdots & WS_{k_0+1,23:00} & \cdots & WS_{N_0+1,23:00} & WS_{N_0+1,23:30} \\ DT_{1,0:00} & DT_{1,0:30} & \cdots & DT_{k_0,23:00} & \cdots & DT_{N_0,23:00} & DT_{N_0,23:30} \\ DT_{2,0:00} & DT_{2,0:30} & \cdots & DT_{k_0+1,23:00} & \cdots & DT_{N_0+1,23:00} & DT_{N_0+1,23:30} \end{bmatrix} \quad (4.9)$$

where definitions of all variables are the same as those of Equation (4.5). The target in the training process is correspondingly written as:

$$\boldsymbol{TTRN}_{BLF} = \begin{bmatrix} P_{2,0:00} & P_{2,0:30} & \cdots & P_{k_0+1,0:00} & \cdots & P_{N_0+1,23:00} & P_{N_0+1,23:30} \end{bmatrix} \quad (4.10)$$

and the definitions of all variables are exactly the same as those mentioned for Equation (4.5).

#### 4.7.1.2 Testing and Validation Process of *Base Load* Forecasting

The testing process is similar to the testing and validation part shown in Figure 4.4, except that there are more inputs. The format of the inputs for the testing and validation process, denoted by ***PTST***, is the same as that represented by Equation (4.9). The output of the trained ANN is predicted day-ahead demand for the *base load* with a resolution of 30min, and it will be compared with the measured demand on the predicted day.

### 4.7.2 *Change Load* Forecasting

*Change load* forecasting *(CLF)* forecasts the change in hourly demand from Day *N* to Day *N*+1.

#### 4.7.2.1 Training Process of *Change Load* Forecasting

The training process of *change load* forecasting is very similar to that for *base load* forecasting shown in Figure 4.7. The only difference is the target. In *change load* forecasting, the target becomes the demand change (in MW or MVAR) from Day $N_0$ to Day $N_0+1$. The demand change from Day $N_0$ to Day $N_0+1$ at time $T$, which is denoted by $\Delta P_{N0+1,T}$ is represented by:

$$\Delta P_{N_0+1,T} = P_{N_0+1,T} - P_{N_0,T} \tag{4.11}$$

Thus, the target in the training process of the *change load* forecasting, which is denoted by $\boldsymbol{TTRN}_{CLF}$ is represented by:

$$\boldsymbol{TTRN}_{CLF} = \left[ \Delta P_{2,0:00} \quad \Delta P_{2,0:30} \quad \dots \quad \Delta P_{k_0+1,0:00} \quad \dots \Delta P_{N_0+1,23:00} \quad \Delta P_{N_0+1,23:30} \right] \tag{4.12}$$

where $\Delta P_{k0+1,0:00}$ is the demand change at 0:00 from Day $k_0$ to Day $k_0$+1 ($1 \le k_0 \le N_0$), and similar definitions apply to other variables.

#### 4.7.2.2 Testing and Validation Process of *Change Load* Forecasting

In the testing and validation process of *change load* forecasting, the inputs are exactly the same as those for *base load* forecasting. The output of *change load* forecasting is the predicted demand change at different times of the day from Day $N$ to Day $N+1$. Using (4.11), this output is superimposed on the measured demand of Day $N$ to calculate the predicted demand for Day $N+1$. Then, the predicted demand is compared with the measured demand on Day $N+1$.

### 4.7.3 Adjuster

An adjuster is installed after BLFer and CLFer to take advantage of both. The adjuster takes into account the demand change induced by the change of the day-type (e.g. from working days to weekends, or vice versa) [89]. There are generally two types of adjusters, a least square algorithm box and a trained ANN.

#### 4.7.3.1 Least Square Algorithm Box

The problem could be formulated as an optimisation problem which derives the parameters to make the computed value as close to the measured value as possible, shown as:

$$Minimise\ Obj = (P_m - P_{BC} w_{BC})^T (P_{m.} - P_{BC} w_{BC}) \tag{4.13}$$

where *Obj* is the objective function. Vector $P_m$ is the measured demand on the prediction day and

$$P_m = \begin{bmatrix} P_1 & P_2 & \dots & P_{k-1} & P_k \end{bmatrix}^T \tag{4.14}$$

where $P_k$ represents the $k^{th}$ value of measured demand on the prediction day. Matrix $P_{BC}$ is the forecasted demand using *base load* forecasting and *change load* forecasting on the prediction day and

$$P_{BC} = \begin{bmatrix} P_{1,BLF} & P_{2,BLF} & \dots & P_{k-1,BLF} & P_{k,BLF} \\ P_{1,CLF} & P_{2,CLF} & \dots & P_{k-1,CLF} & P_{k,CLF} \end{bmatrix}^T \tag{4.15}$$

where $P_{k,BLF}$ represents the $k^{th}$ value of predicted demand from *base load* forecasting, and $P_{k,CLF}$ represents the $k^{th}$ value of predicted demand from *change load* forecasting. The weighting factor vector, $w_{BC}$, indicates the contribution of *base load* forecasting and *change load* forecasting, and it can be represented by:

$$w_{BC} = \begin{bmatrix} w_B & w_C \end{bmatrix}^T \tag{4.16}$$

where $w_B$ is the weighting factor of *base load* forecasting and $w_C$ is the weighting factor of *change load* forecasting. Taking the derivative of Equation (4.13) and setting it to zero, the weighting factor vector is derived as

$$w_{BC} = \left( {P_{BC}}^T P_{BC} \right)^{-1} {P_{BC}}^T P_m \tag{4.17}$$

After derivation, the derived weighting factors are saved for further use.

#### 4.7.3.2 Trained ANN

Instead of a least square algorithm box, a trained ANN can also be used as an adjuster. In the training process, the inputs of the ANN are forecasted demand using *base load* forecasting and *change load* forecasting for the prediction day, and the target is the measured demand on the prediction day. If written in the format of MATLAB code, the input is the transpose of Vector $P_{BC}$ (shown as (4.15)), and the target is the transpose of Vector $P_m$ (shown as (4.14)).

### 4.7.4 Case Study

The methodology is applied to both $P$ and $Q$ predictions to predict day-ahead demand for "Bus BRW" in a local distribution network. Demand data over a period of 30 days with 30min resolution are used to illustrate the methodology. Among the 30-day data, the first third are used for the training of *base load* forecasting and *change load* forecasting, the second third are used for validation of the two sub-approaches, and the last third are used for training and validation of adjustments for *base load* forecasting and *change load* forecasting.

The prediction result using the proposed approach for "Bus BRW" is shown in Figure 4.8.

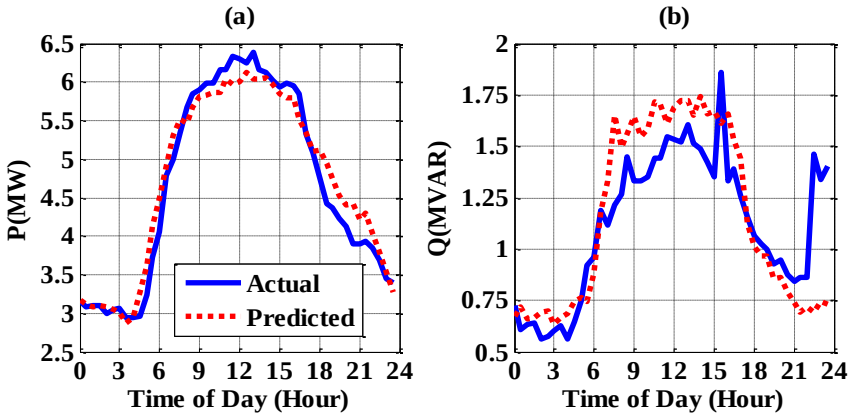


Figure 4.8. Predicted demand versus actual demand: (a) real power, (b) reactive power

It is shown that the predicted load curves for both real and reactive power are close to

the actual curves. The mean absolute percentage errors (MAPEs) of real and reactive power are 4.27% and 15.29%, respectively. MAPE of reactive power is higher than that of real power as the (absolute) value of reactive power (in MVAR) is much smaller than that of real power (in MW). A similar finding regarding MAPE of the reactive power prediction is reported in [92].

## 4.8 Probabilistic Characteristics of Absolute Percentage Error for Total Load Forecasting

In some cases involving probabilistic studies, the probabilistic characteristics of absolute percentage error (APE) for total load forecasting are of interest. The APE of total load forecasting for $P$ or $Q$ at every sampling point is therefore included in a sample set, and then the probability density function (PDF) and cumulative distribution function (CDF) of APEs are calculated.

### 4.8.1 PDF & CDF of APE for either *P* or *Q* Prediction

Following the developed approach and the case study presented in Section 4.7, the PDF and CDF of APEs for total $P$ and $Q$ forecasting are plotted and shown in Figure 4.9, where $|e_P|$ and $|e_Q|$ are APEs of $P$ and $Q$ prediction respectively. From Figure 4.9(a), it can be seen that the most probable APE for total $P$ prediction is about 1.95%. Among all the predictions, there are about 65% of predictions with APEs ≤ 5%, 88% with APEs ≤ 10%, and 100% with APEs ≤ 15%; in 90% of the cases, APEs≤10.5%. (Note: The negative values on the horizontal axis in Figure 4.9 are induced by MATLAB plotting, but the data used for plotting are all positive. This also applies to similar situations in Chapter 5.) The result of Figure 4.9(a) indicates a relatively high accuracy and high confidence level of real power forecasting. From Figure 4.9(b), it is apparent that the probability density of APEs for $Q$ prediction is nearly the same within the interval [8%, 18%], although there are two local peaks at

APE=8.1% and APE=17.22%. Among all the predictions, there are about 19% with APEs ≤ 5%, 37% with APEs ≤ 10%, 58% with APEs ≤15%, 72% with APEs ≤20%, 90% with APEs ≤ 30%. The result of Figure 4.9(b) shows that the accuracy and confidence level of reactive power forecasting are not as high as those for real power forecasting but still reasonable and in line with previously reported results [92, 93]. This is predominantly due to much higher variability of the reactive power induced among the others by voltage regulating devices (e.g. reactive power compensators).

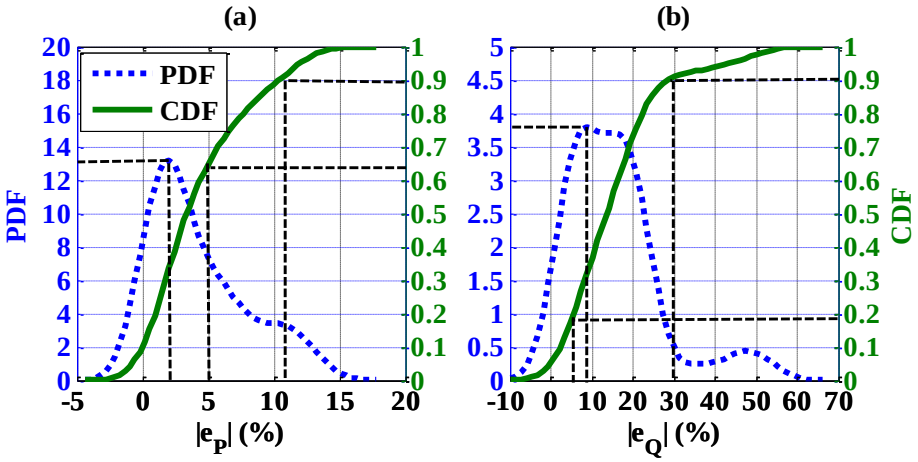


Figure 4.9. PDF and CDF of APEs for (a) total P forecasting, and (b) total Q forecasting

### 4.8.2 PDF & CDF for the Weighted Average of APE for *P* and *Q* Prediction

As real and reactive power are related, $|e_P|$ and $|e_Q|$ are weighted and combined to create a new index, the cumulative absolute percentage error (CAPE), for total demand forecasting to describe the total demand forecasting error. In this work, the cumulative absolute percentage error (CAPE) is denoted by $e_F$ and represented by:

$$e_F = w_{|e_P|}\left|e_P\right| + w_{|e_Q|}\left|e_Q\right| \tag{4.18}$$

where $w_{|ep|}$ and $w_{|ep|}$ are the weights of $|e_P|$ and $|e_Q|$, respectively. The weights of $|e_P|$ and $|e_Q|$ are dependent on the importance of accuracy of $P$ and $Q$ prediction for specific studies. For illustration purposes in this study, the weights of both $|e_P|$ and $|e_Q|$ are set to 50%.

With weights of 50% for both $|e_P|$ and $|e_Q|$, PDF and CDF curves of the cumulative absolute percentage error (CAPE) are obtained and shown in Figure 4.10. It can be seen from the CDF that the most probable CAPE is about 9.2%. Among all the samples forecasted, there are about 19% with CAPEs ≤5%, 58% with CAPEs ≤10%, 88% with CAPEs ≤15% and 93% with CAPEs ≤20%; in 90% of the cases, CAPEs≤16%.

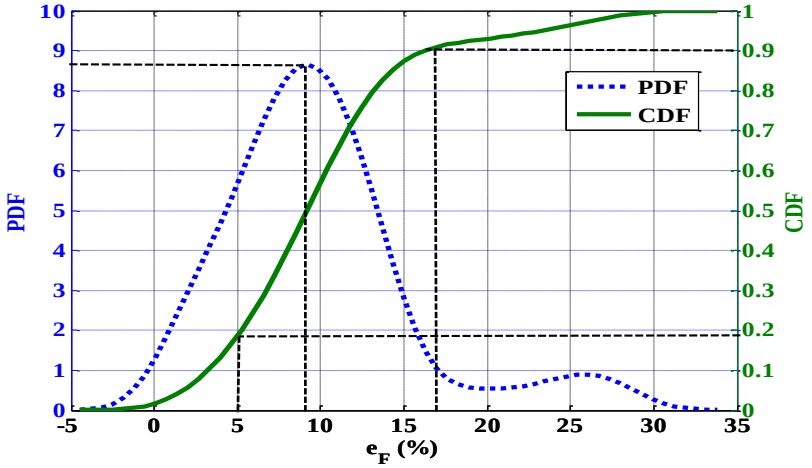


Figure 4.10. PDF and CDF of CAPE for total demand forecasting when the weights of APEs for both *P* and *Q* forecasting are 50%

## 4.9 Summary

The chapter introduced and presented a comparative analysis of two widely used artificial-intelligence-based (AI-based) approaches, the artificial neural network (ANN) and the adaptive neuro-based fuzzy inference system (ANFIS), for load forecasting. It demonstrated that both tools can achieve a high level of accuracy of prediction with optimal parameter configurations of prediction algorithms and that their performances are comparable when all the parameters are configured properly. Recommendations with respect to optimal configuration of parameters for both methods are made.

Additionally, the chapter reported improvements made to the ANN short-term load forecasting (ANNSTLF) tool for load forecasting developed by EPRI. The

improved forecasting methodology is illustrated using case studies from real UK distribution network. Probabilistic characteristics of absolute percentage errors (APEs) of total demand forecasting for real power, reactive power and the weighted average of APE for total $P$ and $Q$ forecasting are also analysed.

# 5 Load Disaggregation at Bulk Supply Points

Load disaggregation is a task that estimates the contribution (as well as the controllability in some cases) of different categories to the total demand mix at any given time of the day. The knowledge of real time load composition will contribute to demand side management (DSM). Past work mainly focussed on disaggregating the load via either intrusive or non-intrusive load monitoring. However, due to the difficulty in accessing all houses via smart meters at all times and the unavailability of frequently measured high-resolution load signatures at bulk supply points, neither of the two approaches is suitable for frequent or widespread application.

This chapter employs artificial neural network (ANN) to develop a fast and effective load disaggregation approach for bulk supply points (BSP) based on the substation RMS measurement without relying on smart meter data, customer surveys or high-resolution load signatures. For convenience, the disaggregation approach is developed in per-unit. As a result, a rated demand (RD) is needed to convert the predicted or measured demand in MW or MVAR to per-unit values. Monte Carlo Simulation (MCS) is used to obtain the most probable rated demand and to generate the training and validation data in the disaggregation approach. Load compositions

obtained using ANN are compared with the validation data and used for load characteristics estimation and validation. Probabilistic distributions and confidence levels of different confidence intervals for errors of load compositions and load characteristics are also derived. In addition, the effect of input absence on load disaggregation has also been assessed. The developed disaggregation approach could be further extended to an approach that enables the prediction of load compositions as well as dynamic responses of the demand (DRD) in the future, if it is integrated with the total demand forecasting approach.

## 5.1 Rated Demand

The demand at bulk supply point (BSP) varies continually at different times during the day because the customer activities vary continually. At the same time under normal operating conditions, system voltage always remains (or is controlled) at a relatively stable level (not necessarily at 1.pu. but close to it) to maintain the secure and stable operation of the power system. In other words, the rated demand (RD) at different times during the day also varies continually as the customer consumption behaviour changes.

### 5.1.1 Definition of the Rated Demand

If the actual voltage at a bulk supply point is the same as its rated voltage, the demand of the bulk supply point at this moment is referred to as rated demand. For example, if the demand at a 6.6kV-level bus is 5MW and the actual voltage ($V$) of the bus is 6.6kV, then the rated demand of the bus at this moment is 5MW. For the same demand but with a 6.5kV voltage, i.e.0.985 p.u. if 6.6kV is taken as the base value, rated demand will assume a different value. The rated demand is generally denoted by $P_{rated}$.

The following case provided by Table 5.1 will be used to describe the rated

demand in more detail using the exponential load model. It gives an illustrative example with voltages and real power measurement at three different times for a bus with voltage level $V_0$.

Table 5.1. Voltages and Real Powers of a Bus with Voltage Level $V_0$ at Three Different Times

| $t$ | $t_0$ | $t_1$ | $t_2$ |
|---|---|---|---|
| $V$ | $V_0$ | $V_0$ | $0.98V_0$ |
| $P$ | $P_n$ | $1.05P_n$ | $1.05P_n$ |

At time $t_0$, $t_1$, $t_2$, the voltages are $V_0$, $V_0$, $0.98V_0$, and the powers are $P_n$, $1.05P_n$, $1.05P_n$. In the exponential load model shown as Equation (2.2), $P_0$ is the rated demand. Therefore, when $t=t_0$, $t_1$ and $t_2$, the exponential load model can be written in the formats of (5.1), (5.2) and (5.3)

$$P_{t=t_0} = P_n\left(\frac{V}{V_0}\right)^{\alpha_0} \tag{5.1}$$

$$P_{t=t_1} = \left(1.05P_n\right)\left(\frac{V}{V_0}\right)^{\alpha_1} \tag{5.2}$$

$$P_{t=t_2} = \left(1.05P_n\right)\left(\frac{V}{0.98V_0}\right)^{\alpha_2} = \left(1.05\times 0.98^{-\alpha_2}P_n\right)\left(\frac{V}{V_0}\right)^{\alpha_2} \tag{5.3}$$

respectively. The corresponding $P_{rated}$ are $P_n$, $1.05P_n$ and $1.05\times 0.98^{-\alpha_2}P_n$. Thus, it can be seen that the rated demand can change at any time of the day, even if the voltage remains (or is controlled to be) the same.

## 5.1.2 Significance of the Rated Demand

The phenomenon described above has brought difficulties in a variety of research tasks, such as load modelling and load disaggregation.

### 5.1.2.1 Rated Demand in Load Modelling

A typical task that encounters difficulties is measurement-based load modelling, especially for load modelling with a relatively long timeframe or *natural demand change* (definition: natural demand change refers to the demand change induced by the variation of the customer consumption behaviour), self-disconnection or self-

recovery.

Publications studying the measurement-based load modelling approach are listed in Chapter 1. In [1], which is the most recent and most comprehensive reference on load modelling, the parameters of different types of load models for both individual load types and aggregate load at different customer sectors are provided. The result of the recovery time constant of dynamic load models of the aggregate load, which is shown in Table III-A-5 of Appendix 3-A of [1] and also presented in [133], exceeds 100s. This indicates that the demand takes about 7-8 min (i.e. 4-5 times the recovery time constant, which is about 400-500s) or even longer to achieve a new steady-state after the step-down voltage disturbance occurred.

Papers [4, 134] published almost at the same time as [133] provide the parameters of the induction motor model (i.e. resistance/reactance of rotors/stators, inertia) for different types of load. If those values are simulated in power system analysis software such as DIgSILENT on the "Asynchronous Machine" model with the same voltage disturbance (implemented in Chapter 3), it can be found that the recovery time constants for all different motors are less than 1 second, and the result presented in [1, 5, 20] supports this conclusion.

A similar diversity in recovery time constants exists between [20] and [18], which were created by the same working group and published almost at the same time. In [20], the recovery time constant of individual dynamic loads are close to 0s; while in [18], the recovery time constants of aggregate demands exceed 100s. As the demand is composed of static and dynamic loads, and the recovery time of the static load is zero and that of the dynamic load is less than 1s, according to the superposition theory, the recovery time constant is technically supposed to be less than 1s and the time for recovery should technically be less than 5s, which is much shorter than 7-8 min.

Figure 5.1 (adopted from [18, 133]) shows an example of measurement-based load modelling. According to the analysis above, even though the parameters obtained in [18, 133] fit the measured dynamic response well (shown as Figure 5.1(b)), they did not accurately present the actual load characteristics(e.g. the parameter such as the recovery time constants, the steady-state voltage exponents), and the most likely reason is that the gradual load change contributed by customer behaviour change distorts the already achieved steady-state and makes it look like a recovery behaviour with long recovery time (shown as Figure 5.1).

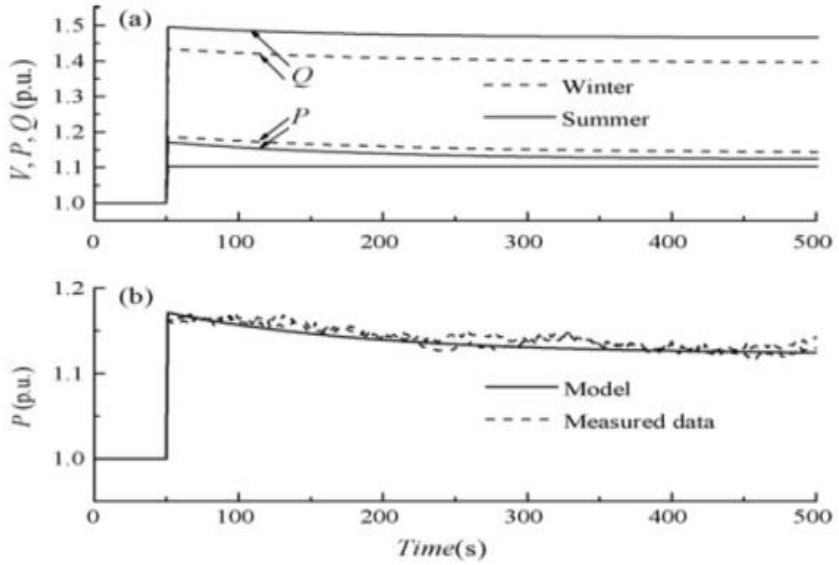


Figure 5.1. Real and reactive power responses to a 10% voltage step-up: (a) simulated for winter and summer and (b) simulated and measured real power responses for summer (adopted from [18, 133]).

Moreover, it is highly likely that the customer behaviour changes within 7-8min, which causes the change of load compositions and eventually the change of the rated demand during the period. For example, for a 10kW load which contains 400 efficient light bulbs (i.e. 25W each [28]) which can be approximated to constant current load [1], when the voltage drops from 1p.u to 0.95p.u, the load immediately drops to 9.5kW and remains unchanged. If from the moment the voltage drop occurs, other bulbs are turned on one after another at one-second intervals, then after 10s, the load will gradually increase to 9.7375kW. The load is still made up of light bulbs with the same load characteristics, but the change of the bulb number makes the whole process

look like a recovery behaviour with a recovery time constant of approximately 2-2.5s, and it will mislead the model structure selection from a static model to a first-order exponential recovery model.

Therefore, for load modelling with relatively long timeframe measurement data where customer behaviour changes may occur, it is important to decide the load changes contributed by both voltage changes and customer behaviour changes quantitatively.

#### 5.1.2.2 Rated Demand in Load Disaggregation

The rated demand (RD) also plays an indispensable role in load disaggregation at the bulk supply point (BSP) with limited data (i.e. RMS measured voltage, real power and reactive power). The ANN-based approach for load disaggregation developed in later sections of this chapter will enable estimation of load categories participating in the total demand at any time. As the approach is developed using load models in per unit systems, when applied to the real measurement data in MW and MVAR, appropriate 'base values' at different times are required to transform the real value into the per-unit value so that the approach can be applied.

In the majority of the literature studying load modelling as those listed above or in Chapter 1, the real power ($P$) and reactive power ($Q$) are presented with per unit value; the process that transforms the real values into per unit values is missing. For example, in EPRI's load modelling report [5], even though the natural load increasing at 10:00am at Bus Oakridge 743 is considered and a value is assigned as the rating, it is not justified why this value is selected and whether it remains the same at different times. Therefore, rated demand plays a significant role as the 'base value' to transform data given in MW or MVAR into appropriate per-unit values.

To fill the gap between the past work and ongoing research, this section proposes a Monte-Carlo based approach for probabilistic estimation of the rated

demand (RD) at the bulk supply point (BSP) at different times during the day. The data are collected from real substations in UK distribution networks and include measured RMS real power ($P$), reactive power ($Q$) and voltage ($V$) at BSP. The voltage level of BSP is selected as the rated voltage $V_0$. Monte-Carlo Simulation (MCS) is applied to generate the voltages and the weighting factors (WF) of different load categories (i.e. contribution of different load categories to total demand) and calculate per-unit total $P$ and $Q$ for all the $V$ and WF. Probability distribution is used to present the characteristics of the rated demand.

### 5.1.3 Derivation of Rated Demand

Figure 5.2 shows how the rated demand is derived in this study and how the actual demand is transformed into per-unit value.

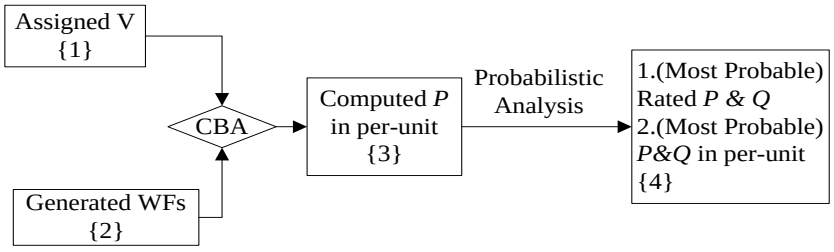


Figure 5.2. Framework for transforming actual demand into per unit value

First of all, operating voltages are assigned (box labelled {1} in Figure 5.2) according to the actual voltage range in the real UK distribution network. Then, weighting factors (WFs) that indicate the contribution of different load categories are generated randomly (box labelled {2}). With the assigned voltages and the generated WFs, using Equation (2.2), (2.23) or (2.24) and (1.2) and Monte Carlo Simulation, the real powers in per-unit for different voltages and different WFs are computed (box labelled {3}). These computed real powers (in per-unit) are classified based on corresponding voltages. Using probabilistic analysis techniques, for any given voltage, the probabilistic characteristics of the computed real powers in per-unit are obtained (box labelled {4}). The rated demands for real powers are then calculated by dividing the

actual real powers by computed real powers in per-unit (box labelled {4}). According to the power factor calculated from the actual real and reactive powers, the rated demands for reactive powers and the probabilistic characteristics of reactive powers in per-unit for any given voltage are computed (box labelled {4}). More details are provided in the following sub-sections.

### 5.1.3.1 Voltage Range Selection

Based on the data obtained from substation in real UK distribution network, the voltages of all buses vary between 0.95p.u. and 1.05p.u. (i.e. the voltage level of the substation is selected as the rated voltage). To consider any extreme operating condition, the voltage range in this study is set to [0.9p.u, 1.1p.u]. Therefore, to generate cases for Monte Carlo simulation, voltages within [0.9p.u, 1.1p.u] are selected and assigned to an array $V_a$

$$V_a = \begin{bmatrix} V_1 & V_2 & \dots & V_n & \dots & V_{N_v} \end{bmatrix} \tag{5.4}$$

with some given step change (e.g., a step change of 0.005p.u) satisfying

$$V_{n+1} - V_n = d \tag{5.5}$$

where $N_V$ is the total number of voltages assigned to $V_a$, $V_n$ is the $n^{th}$ term of $V_a$ ($1 \le n \le N_V$), $d$ is the given step change. The $n^{th}$ term $V_n$ can be further represented by an arithmetic sequence shown as

$$V_n = 0.9 + d(n-1) \tag{5.6}$$

According to the voltage range selected, 0.9 is the $1^{st}$ term and 1.1 is the last (i.e. the $N_V{}^{th}$) term. Therefore, if 0.005p.u. is selected as the step change of the voltage, $N_V$=41.

### 5.1.3.2 Weighting Factor Generation

For the purpose of the study, according to Table 1.2, the loads are classified into ten load categories, shown in Table 5.2. From Table 5.2, it can be seen that the ten load categories are grouped into two parts, the controllable part and the uncontrollable

part; either of the two parts contains five categories. According to recent UK customer surveys from [36, 37], the controllable load varies from 10% to 50%. Therefore, the controllable load range in this study is set as [10%, 50%], and the uncontrollable load is set as [50%, 90%]. To ensure that all possibilities are considered within the range, the controllable load varies gradually from 10% to 50% with a 5% step increase, and the uncontrollable part correspondingly varies from 90% to 50% with a 5% step decrease. Therefore, there are 9 possibilities of the controllable-uncontrollable combinations in total.

Table 5.2. Load Categories and Their Controllability in This Study

| Group | Category |
|---|---|
| Controllable | ✓ CTIM1<br>✓ QTIM1<br>✓ Controllable CTIM3<br>✓ QTIM3<br>✓ Controllable resistive |
| Uncontrollable | ✓ Uncontrollable CTIM3<br>✓ Uncontrollable resistive<br>✓ Rectifier<br>✓ SMPS<br>✓ Lighting Loads |

For each of the 9 possibilities, the weighting factor (WF) which indicates the percentage of each load category participating in the total demand are generated randomly for different load categories within corresponding parts using Monte Carlo simulation. For example, for a case with a "25% controllable part and 75% uncontrollable part" combination, the WFs of different controllable loads are generated from 0% to 25%, and the WFs for all the uncontrollable loads are generated from 0% to 75%. The sum of WFs for all loads in either of the controllable and uncontrollable part remains unchanged, and all generated WFs are ensured positive. The number of WFs generated for each controllable-uncontrollable combination is recorded as $N_W$. Therefore, there are $9 \times N_W$ controllable-uncontrollable combinations

in total.

#### 5.1.3.3 Per-unit Total *P* Calculation

After *V* and *WF* for all possibilities are generated, all possible *V*-*WF* combinations are obtained to calculate the per-unit aggregate *P*, via Equation (2.2) and (1.2). The exponential load models of individual load categories with voltage-dependent exponents have been derived in Chapter 2. There are $N_V \times 9 \times N_W$ *V*-*WF* combinations as well as calculated per-unit *P* in total. Each calculated per-unit aggregate *P* is one possible rated demand (in p.u.) of real power for the corresponding voltage.

#### 5.1.3.4 Probability Distribution of Per Unit Total Demand

After the per-unit *P* is calculated, the probability distribution of the per-unit *P* under each corresponding voltage can be generated, and the value that has the highest probability density is defined as the most probable per-unit real power. The rated demands of the real power are derived by dividing the actual demand by the calculated per-unit real power, and the most probable rated demand of the real power is derived by dividing the actual demand by the calculated most probable per-unit real power.

Measured real power in per-unit can be determined via Monte Carlo simulation in the two following ways: (1) divide the measured real power in MW by the corresponding rated demand for the real power; (2) directly take the per-unit value of the real power calculated for the corresponding voltage at a given time. The two are actually equivalent. Similarly, the most probable value of measured/predicted real power in p.u. could be determined: (1) divide the measured/predicted real power in MW by the most probable rated demand for the real power; (2) directly take the most probable per-unit real power calculated for the corresponding voltage at a given time. As power factor can be obtained from the measurement, the rated demands for the reactive power (in MVAR) and measured reactive power in per-unit can be calculated

easily.

#### 5.1.3.5 Application to Predicted Demand

It is worth noting that in day-ahead total load forecasting, the quantities to be predicted and used for load composition prediction are *P* and *Q*, and the voltage cannot be predicted in advance. However, when predicting *P* and *Q*, voltage is not taken as an index in ANN training, either. In other words, day-ahead total demand prediction is executed considering that the voltage on Day *N*+1 at given time remains the same as it is on Day *N*. Therefore, when determining the rated demand or predicted per-unit *P&Q* demands for Day *N+1*, the voltage of Day *N+1* is assumed to be the same as that of Day *N*.

Using the assumed voltage, the rated demand for *P* can be determined via Monte-Carlo simulation, and the most probable predicted *P* in per-unit could be determined in the similar ways mentioned above. According to the predicted *P* and *Q*, predicted *Q* in p.u. could be calculated using the predicted power factor. The predicted *P* and *Q* in p.u. will be used in the load disaggregation approach to predict day-ahead load composition at different times of the day.

#### 5.1.3.6 Probabilistic Rated Demand Curve

As per-unit *P* is probabilistically distributed, RD is also probabilistically distributed. To obtain the rated demands at different times of the day as well as probabilistic rated demand curve, a general way is using inverse cumulative distribution function (ICDF) generation function (in MATLAB or other computation software). According to the obtained probability distribution function (PDF), a group of per-united *P* can be generated. Dividing the actual demand by them at different times of the day, the probabilistic rated demand curves can be obtained. Sometimes, if the derived PDF looks like one of the widely used distributions such as Gaussian distribution, it can also be approximated by the corresponding distribution, which will

simplify the computational task.

### 5.1.4 Results and Discussion

In this study, 0.005 p.u is selected as the voltage step change, and for each voltage assigned to $V_a$, there are 33,600 calculated per-unit values of the real power in total.

#### 5.1.4.1 Probability Distribution of Per Unit Total Demand

Figure 5.3 shows the probability distribution of per-unit total demand when per-unit voltage is 0.97 and 1.03p.u respectively as illustrative examples. The result indicates that the probability distribution can be very accurately fitted by a Gaussian distribution (i.e. red solid line in Figure 5.3). Probability distributions of per-unit real power at other voltages are not shown, but they have similar shapes as those shown in Figure 5.3.

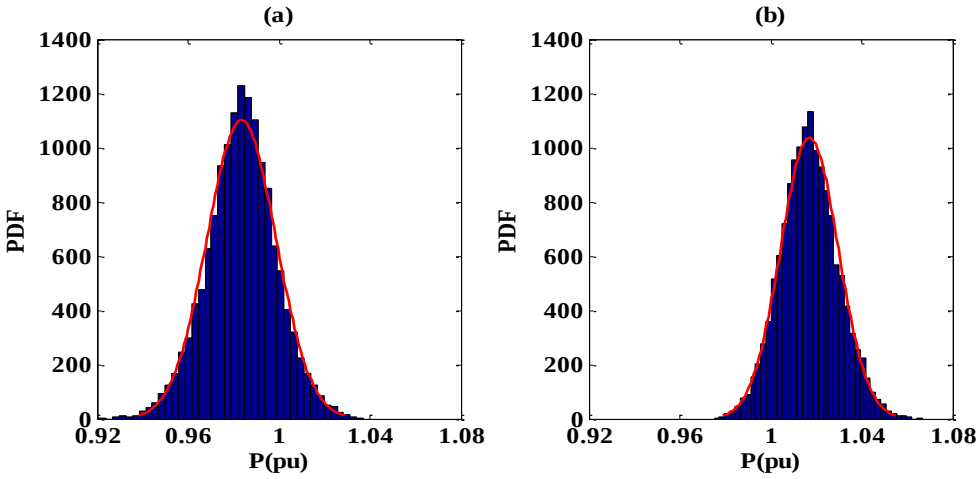


Figure 5.3. Probability distribution of per-unit total demand when voltage is: (a) 0.97 p.u, and (b) 1.03 p.u

Figure 5.4 shows the fitted mean ($\mu$) and standard deviation ($\sigma$) of calculated per-unit $P$ for different voltages from 0.9 to 1.1p.u. It can be seen that the mean value (also the most probable value, if a Gaussian distribution is used) of calculated per-unit total $P$ increases as the voltage increases, while the standard deviation decreases as the voltage approaches 1.00 p.u. The means and the deviations are in per-unit.

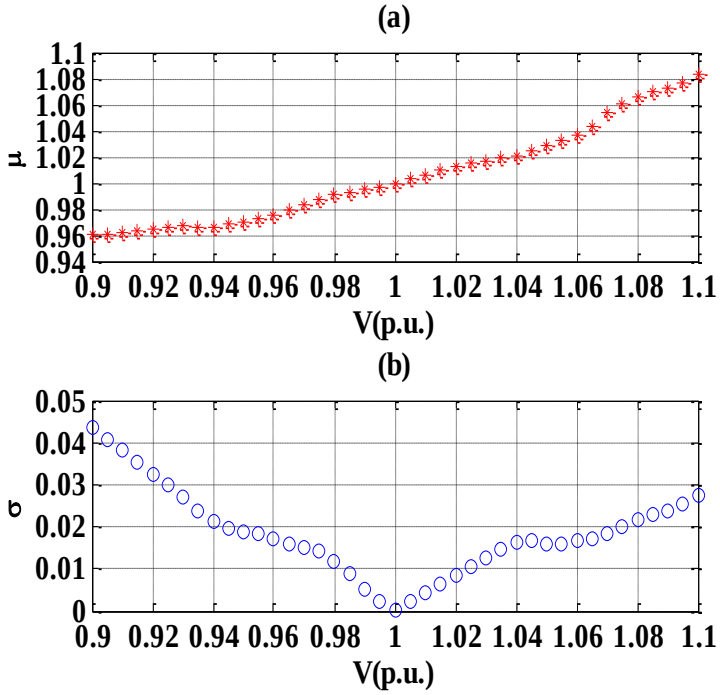


Figure 5.4. Fitted mean and deviation for different per unit voltage

### 5.1.4.2 Daily Probabilistic Rated Demand

Figure 5.5(a) shows the actual voltages at the selected 6.6kV bus "Bus DKST" from 8:00 to 17:00, on 25 June, 2014. The voltages are lower than 6.6kV at all times during the day. Figure 5.5(b) shows the boxplot of the probabilistic rated demand for real power and the actual real power demand at the same bus during the same period. From Figure 5.5(b), it can be seen that the majority of the rated demands for the real power are slightly higher than the actual real power demand.

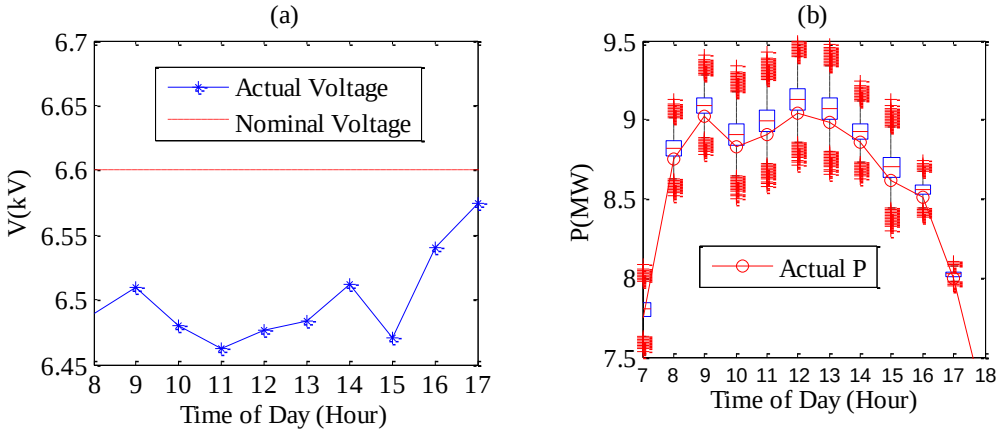


Figure 5.5. (a) The actual voltage verses voltage level, and (b) boxplot of probabilistic rated demand against actual demand at a 6.6kV ENW bus from 8:00 to 17:00.

According to the measured *P* and *Q*, the magnitude and the characteristics of power factors (i.e. lagging or leading) at different times during the day can be calculated. Based on the calculated power factors, the daily probabilistic rated demand for *Q* can be calculated, correspondingly. Figure 5.6 shows the boxplot of the probabilistic rated demand for *Q* and the actual *Q* demand at the same bus as for Figure 5.5 during the same period. From Figure 5.6, it can be seen that the majority of the rated demands for *Q* as well as the most probable values are slightly higher than the actual *Q* demand.

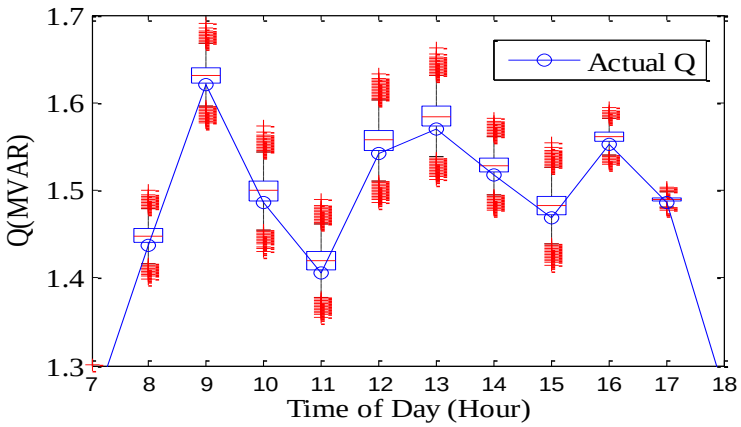


Figure 5.6. Boxplot of probabilistic RD for *Q* against actual *Q* demand at Bus "DKST" from 8:00 to 17:00 on 25 June, 2014.

### 5.1.4.3 Rated Demand Applied to Real and Reactive Power for Disaggregation

It is worth noting that, for convenience in the disaggregation and validation process in later chapters, generally, only the most probable values for rated demand and predicted per-unit real and reactive power are used. If the methodology for rated demand and per-unit real and reactive power derivation is applied to the demand data on 25 June, 2014 of "Bus BRW", the derived most probable rated demand for *P* and most probable per-unit *P* and *Q* are shown in Figure 5.7.

Figure 5.7(a) shows measured voltages in kV at "Bus BRW" verses the rated voltage (i.e. 6.6kV) with a 30min resolution. It can be seen that the system voltage

fluctuates around the rated voltage. Therefore, rated demands at different times of the day should be different; correspondingly, the calculated per-unit *P* and *Q* at different times of the day should also be different. Figure 5.7(b) shows the actual *P* demand verses the most probable rated demand of *P* in MW at different times of the day. Generally, rated demand is different from the actual demand, although the difference is moderate in this case. The most probable rated demand is used to calculate the most probable per-unit *P* and *Q* at different times of the day, shown as Figure 5.6(c) and Figure 5.6(d), respectively. It can be seen from Figure 5.7(c) and Figure 5.7(d) that during the day, the most probable per-unit values of the real power range from 0.985p.u to 1.01p.u and those of Q range from 0.2p.u to 0.4p.u.

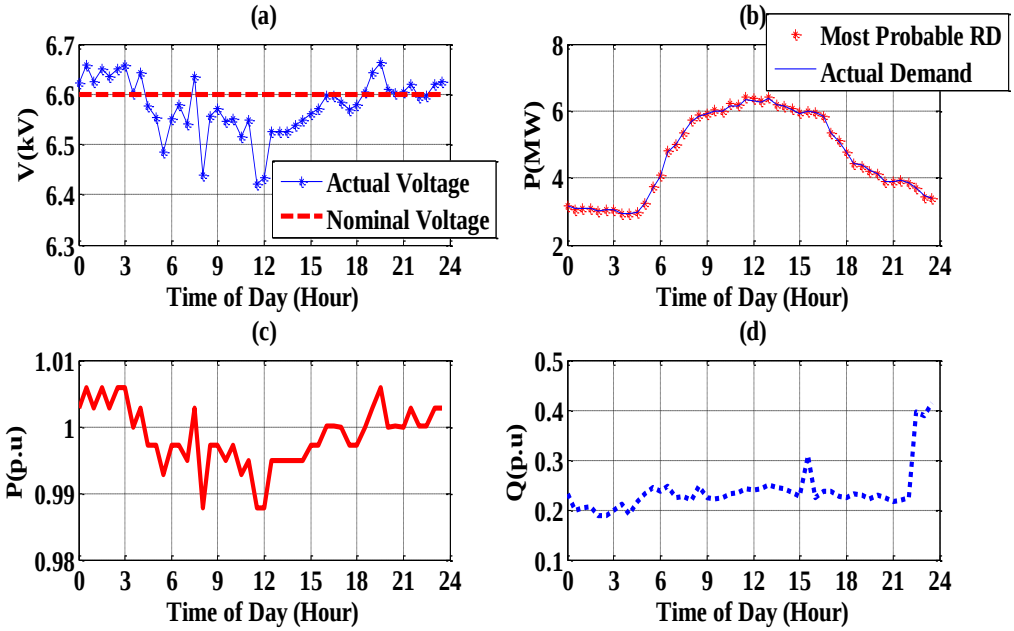


Figure 5.7. (a) Actual voltage measurements verses voltage level; (b) rated *P* demand verses actual *P* demand; (c) Per-unit *P* at different times of the day; (d) Per-unit *Q* at different times of the day

### 5.1.5 Summary about Rated Demand

Section 5.1 described a Monte-Carlo-based approach to obtain the rated demand as well as per-unit *P* and *Q* at different times during the day. The results are used for more accurate load disaggregation (and the prediction of load compositions) as illustrated in the following subsections.

## 5.2 Load Disaggregation Based on Load Categories and Controllability

Significant attention was paid to load disaggregation in past work. Two approaches commonly used for load disaggregation are intrusive load monitoring and non-intrusive load monitoring, which have been introduced in Chapter 1. Considering the deficiency of the past approaches, this section develops a new load disaggregation approach to estimate the load composition at bulk supply busses based only on general substation measurements, i.e., RMS of real power ($P$), reactive power ($Q$) and voltage ($V$). It does not require the involvement of customers and facilitates the real time estimation of dynamic load characteristics at the bulk supply point without carrying out voltage disturbance tests. ANN is used to develop the approach.

By classifying the customers into ten load categories mentioned in Section 5.1 and making use of the appropriate voltage-dependent load models for different load categories mentioned in Chapter 2, the algorithm captures the relationships among the voltage, real and reactive power consumptions, and appropriate weighting factors of the load categories. Monte Carlo simulation is used to generate the weighting factors for load composition within the total load and bus voltages, and to establish $P$ and $Q$ of the total load based on individual contributions of different load categories. The inputs to the ANN in the training process are the voltage and the total load real and reactive power, and the targets are the weighting factors (WF) of different load compositions. The error between the calculated WFs and the WFs obtained from validation data is assessed and its probability density function (PDF) and the confidence level (CL) of different confidence intervals are calculated. The estimated WFs are then used as an input to the estimation of aggregate load characteristics and the results are compared with the generated validation data.

### 5.2.1 Disaggregation Procedure

The ANN based load disaggregation approach considers all possible combinations of load supply voltage and load compositions at the bulk supply point, and captures the relationship among the load compositions, the voltage, the real and reactive power at the bulk supply point under all circumstances via offline training.

The approach can be summarised into the following steps: i) derive voltage dependent $P$ and $Q$ exponential load model coefficients for different individual load categories; ii) generate random voltages and weighting factors for training; iii) derive all combinations of voltages and weighting factors; iv) use the result in i) and iii) to define the input and the target for the ANN; v) train ANN and save trained ANN; vi) repeat ii)-iv) to generate validation data; vii) validation, probability distribution derivation and confidence level analysis. A flowchart of the proposed methodology is shown in Figure 5.8. Among the procedures, the details of Step i) have already been completed in Section 2.3. For other steps, more details will be provided in the following text.

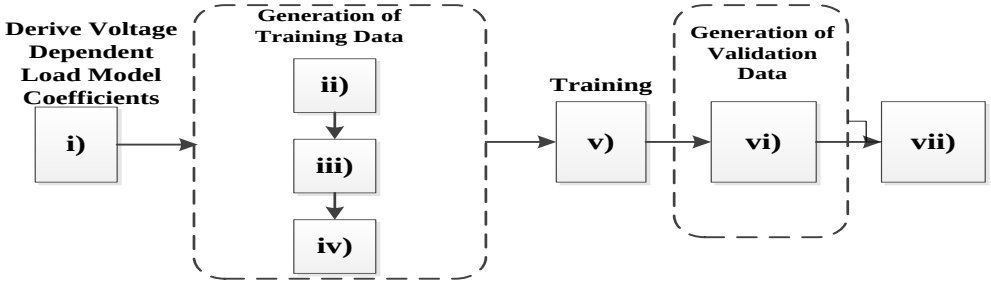


Figure 5.8. Flowchart of the proposed methodology

## 5.2.2 Generation of the Training Data for ANN

### 5.2.2.1 Generation of Voltages

In most cases, the power system is operating in normal conditions, with the bus voltages ranging from 0.95 p.u to 1.05 p.u. Considering the potential margins in extreme cases and ensuring that the developed tool could be widely applied to networks and areas both within and outside the UK, the voltage range used in load disaggregation is [0.9, 1.1] p.u. With Monte Carlo simulation, voltages ranging from 0.9 p.u to 1.1 p.u are generated randomly using uniform distribution, ensuring that the voltage is sampled with equal probability for training. The number of voltage samples is denoted as $N_V$.

### 5.2.2.2 Generation of the Weighting Factor

The procedure of the generation of weighting factors (WFs) is exactly the same

as the circumstance of the derivation of the rated demand, discussed in Section 5.1.3.2. The number of WFs generated for each of the 9 controllable-uncontrollable combination possibilities is denoted as $N_W$.

#### 5.2.2.3 Combinations of Voltage and Weighting Factor

After random voltages and WFs are generated, all possible combinations of voltages and weighting factors are obtained in order to generate the inputs for ANN. There are in total $N_V \times 9 \times N_W$ $V$-WF combinations.

#### 5.2.2.4 ANN Input and Target

For each combination of the voltage ($V$) and the weighting factor (WF), the total real and reactive power at the bulk supply point is calculated by Equation (1.2). They are regarded as the measurement (RMS value) collected from the bulk supply point here. The real and reactive power in Equation (1.2) for individual load categories under different voltages can be calculated with the exponent-voltage relationship given in Figure 2.6, by Equation (2.23) or (2.24), (2.2) and (1.2). The inputs for ANN are total load real power $P_{Agg}$, total load reactive power $Q_{Agg}$, and supply voltage $V$ at the bulk supply point. They are written in an input matrix ***PTRN*** represented by

$$\boldsymbol{PTRN} = \begin{bmatrix} P_{Agg,1} & \cdots\cdots & P_{Agg,N_V \times 9 \times N_W} \\ Q_{Agg,1} & \cdots\cdots & Q_{Agg,N_V \times 9 \times N_W} \\ V_1 & \cdots\cdots & V_{N_V \times 9 \times N_W} \end{bmatrix} \tag{5.7}$$

Sometimes when $V$ or $Q$ measurements are not available, the inputs for ANN become "$P$ and $Q$" or "$P$ and $V$" only. In the former case, the third row of Equation (5.7) is removed; in the latter case, the second row of Equation (5.7) is removed. The targets of ANN are corresponding WFs used to calculate aggregate real and reactive power, written as a target matrix ***TTRN*** represented by

$$\boldsymbol{TTRN} = \begin{bmatrix} w_{1,1} & \cdots\cdots & w_{1,N_V \times 9 \times N_W} \\ \vdots & \vdots\ w_{i,j}\ \vdots & \vdots \\ w_{10,1} & \cdots\cdots & w_{10,N_V \times 9 \times N_W} \end{bmatrix} \tag{5.8}$$

where $w_{i,j}$ is the WF of load category $i$ for $j^{th}$ $V$-WF combination, with $1 \le i \le 10$, $1 \le j \le N_V \times 9 \times N_W$. All data in input matrix and target matrix are in per unit values.

### 5.2.2.5 ANN Training

After the input and the target are defined, an ANN is employed as the training tool. The flow chart of the training process as well as ANN settings are shown in Figure 5.9. The size of the ANN hidden layer is configured as the nearest integer to the estimated value of Equation (4.2). The transfer function for the hidden layer and the output layer are configured as *logsig* and *tansig* respectively to ensure that the output of the ANN (i.e. the weighting factors) is within [0,1]. (Note: As mentioned in Section 4.4.3.1, the output of a *logsig* function is insensitive to its input. If selected as the transfer function of the output layer, it could limit the final output of the ANN to a small range which cannot cover the range of the weighting factors. Therefore, the best way is using a *tansig* transfer function, with positive inputs to the output layer. To ensure the inputs of the output layer (also the output of the hidden layer) are positive, a *logsig* is selected as the transfer function for the hidden layer.).

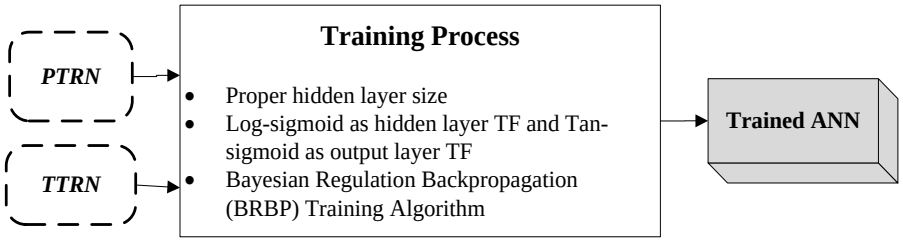


Figure 5.9. Flow chart and ANN settings of the training process

As the training process considers all combinations of voltages and weighting factors, once trained, the ANN can be applied to extract the load compositions at any time of any day in the future, as long as the measurement of bus voltages and the demand are available. Therefore, a robust and accurate training algorithm is required, although it does not have to be very fast. According to the description in Section 4.2.2.4, Chapter 4, Bayesian Regulation backpropagation (BRBP) with high accuracy and high robustness is adopted as the training algorithm in this study. Other parameters are configured as MATLAB default value because their effects are not as significant as main parameters mentioned in Chapter 4. The ANN trained in this way

captures the relationship among the load supply voltage, total demand consumption and the compositions of different load categories. This relationship can be used for validation purpose and for further application.

### 5.2.3 ANN Validation with Generated Data

The developed ANN tool will first be validated with generated data. The validation process using measured steady-state and dynamic demand data is illustrated in Chapter 6.

#### 5.2.3.1 Generation of the Validation Data

A similar process for training data generation as described in Section 5.2.2 is implemented for validation data generation. The input and the target in the validation process can be written as an input matrix ***PTST*** according to (5.7) and a target matrix ***TTST*** according to (5.8). The target for validation process is used as a standard to assess the performance of the trained ANN.

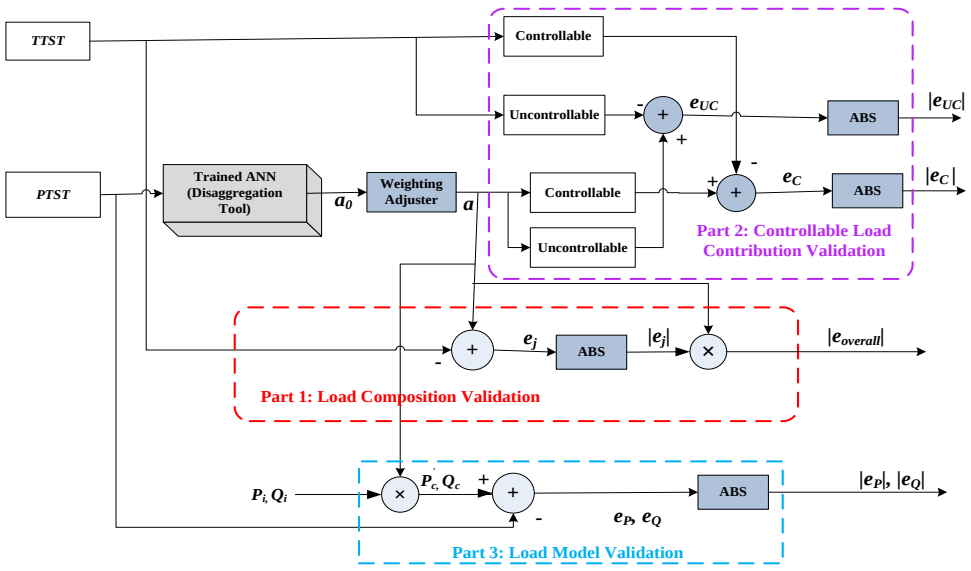


Figure 5.10. Validation process block diagram

#### 5.2.3.2 Validation Process

The block diagram of the validation process is shown in Figure 5.10. It consists

of three parts: i) validation of load composition weighting factor (Part 1), ii) validation of controllable load weighting factors (Part 2), iii) validation of load model (*P* and *Q*) calculated by ANN computed composition weighting factors and the individual *P* and *Q* consumption (Part 3).

At the beginning, the input data ***PTST*** (*P&Q&V* or a subset of them) is processed by the trained ANN to create an output vector $\boldsymbol{a_0}$, which contains weighting factors for individual load categories computed by the trained ANN. Due to the error produced during the computation process, the sum of elements inside $\boldsymbol{a_0}$ deviates slightly from 1. To ensure the sum of the weighting factors is exactly equal to 1, a weighting adjuster is installed after the trained ANN to normalise $\boldsymbol{a_0}$. The normalisation process is represented as

$$a_i = \frac{a_{0i}}{\sum_{i=1}^{N} a_{0i}} \tag{5.9}$$

where $a_i$ is the adjusted weighting factor for load category $i$, $a_{0i}$ is the weighting factor for load category $i$ before being adjusted, $N$ is the number of elements in vector $\boldsymbol{a}$ and $\boldsymbol{a_0}$. Therefore, the weighting adjuster output, vector $\boldsymbol{a}$, is regarded as the final computed weighting factor vector.

### *A. Weighting Factor Validation*

In Part 1, the weighting adjuster output is compared with the target ***TTST***, the initial weighting factors used to generate ***PTST***. The weighting factor error (WFE) for different individual load categories, Vector $\boldsymbol{e_j}$ in Figure 5.10, is obtained by taking the difference between $\boldsymbol{a}$ and ***TTST***. An absolute value operator denoted by **ABS** is installed after $\boldsymbol{e_j}$ in Figure 5.10 to take the absolute value. To obtain the overall error ($|e_{overall}|$) of the disaggregation approach, an aggregation approach is needed to combine the absolute errors of different load categories into an aggregate error. In [135], series of probability combination approaches are proposed, such as average or

weighted average, multiplicative average, and Bayesian approaches. It is concluded that simple average or weighted average can work better on probability combination problem than more complicated rules which always induce abundant computation tasks and over-performance. As a result, considering the effects of the weights of different load categories on the error aggregation process, weighted average approach is adopted to aggregate individual errors into an overall error, shown as

$$|e_{overall}| = \sum_{i=1}^{N} a_i |e_i| \tag{5.10}$$

where $N$ is the total number of load categories, $a_i$ is the computed weighting factor for load category $i$, $e_i$ is the weighting factor estimation error for load category $i$ in each "Vector $\boldsymbol{e_j}$", $|e_{overall}|$ is the aggregate load disaggregation error, also named absolute weighting factor error (AWFE).

### *B. Validation of Controllable Load Contribution*

After weighting factors of individual load categories are calculated, for both $\boldsymbol{a}$ and ***TTST***, the weighting factors that correspond to controllable loads are summed up to create the WFs of the controllable load, shown in Part 2 of Figure 5.10. A similar task applies to the derivation of the weighting factors of the uncontrollable load. Then, the weighting factors of the controllable load calculated from $\boldsymbol{a}$ and ***TTST*** are subtracted, and the difference is defined as the weighting factor error (WFE) of the controllable load ($\boldsymbol{e_C}$ in Figure 5.10). Sometimes, absolute weighting factor error (AWFE, the absolute value of WFE) is also of interest. Therefore, similar to the process of weighting factor validation, an absolute value operator is installed after $\boldsymbol{e_C}$ to obtain AWFE, $|\boldsymbol{e_C}|$. A similar approach is used to obtain WFE and AWFE of the uncontrollable load.

### *C. Load Model Validation*

In Part 3 of Figure 5.10, the computed weighting factors are used to calculate

the total real and reactive power at the bulk supply point ($P_c$, $Q_c$) via component-based load modelling approach using Equation (2.2) and (1.2). The calculated real and reactive power, $P_c$ and $Q_c$, respectively, are compared with the ANN input $P$ and $Q$. The absolute difference between $P_c$ and ANN input $P$, denoted as $e_P$, is given by

$$e_P = \left| \sum_{i=1}^{N} P_0 \left( \frac{V}{V_0} \right)^{\alpha_i} (a_i - w_i) \right| = P_0 \left| \sum_{i=1}^{N} \left( \frac{V}{V_0} \right)^{\alpha_i} \Delta w_i \right| \tag{5.11}$$

and the absolute difference between $Q_c$ and ANN input $Q$, denoted as $e_Q$, is given by

$$e_Q = P_0 \left| \sum_{i=1}^{N} \sqrt{\left( \frac{1}{pf_{0i}} \right)^2 - 1} \cdot \left( \frac{V}{V_0} \right)^{\beta_i} \Delta w_i \right| \tag{5.12}$$

where $N$ is the total number of load categories, $w_i$ is the weighting factor for load category $i$ used to calculate $P$ and $Q$ in ***PTST***, $a_i$ is the computed weighting factor for load category $i$, $pf_{0i}$ is the nominal power factor of load category $i$, $\alpha_i$ and $\beta_i$ are voltage-dependent exponents of load category $i$ for $P$ and $Q$, respectively. Both $e_P$ and $e_Q$ are defined as absolute load model error (ALME), the former for $P$ and the latter for $Q$.

There are two reasons to compare the computed load characteristics with the input $P$ and $Q$ of ANN. One reason is that in reality, the load composition data of different categories of loads at any time of the day are generally not available, but the total demand data can be obtained easily from the bulk supply point. Therefore, comparing the computed load characteristics with the measured total demand is a more straightforward way to validate the load disaggregation than comparing the computed percentages with the "acquired percentage" of individual load categories (if there is any). Another reason is that to some extent, the computed load characteristics are indeed able to provide an assessment of the accuracy of the load disaggregation. As the individual load characteristics (provided in Chapter 2) are accurate, the conditional statement "if the computed load composition percentages are accurate,

then the computed aggregate load characteristic is accurate" is true. Therefore, the contrapositive of the conditional statement "if the computed aggregate load characteristic is not accurate, then the computed load composition percentages is not accurate" is also true. In other words, a large absolute load model error (ALME) indicates a disaggregation with a lower accuracy or a lower confidence level under some given accuracy; a relatively small ALME, on the contrary, can indicate that the disaggregation result is to some extent accurate, or at least reasonable.

#### 5.2.3.3 Probabilistic Distribution, Confidence Interval and Confidence Level of the Errors

As ***PTST*** contains large numbers of randomly generated validation data, the weighting factor error (WFE), the absolute weighting factor error (AWFE) and the absolute load model error (ALME) are probabilistically distributed. Thus, probability distribution function (PDF) is adopted to present the distribution of the errors. The confidence interval and the confidence level of the error indicate the reliability of the load disaggregation approach. A confidence interval (CI) is an observed interval that frequently includes the parameter of interest if the experiment is repeated, and confidence level (CL) is defined as the percentage of all possible samples that can be expected to include the true population parameter [125]. To assess the reliability, cumulative distribution function (CDF) is adopted to obtain the confidence levels in different confidence intervals of the errors.

### 5.2.4 Illustrative Results of Load Disaggregation

In this study, there are 22500 *P&Q&V* data sets for training and 5625 *P&Q&V* sets for validation. Implemented in MATLAB 2013a on an Intel(R) Core (TM) i5-2400 CPU @ 3.10GHz computer installed with 64-bit Windows 7 system, the training process takes up to 30 minutes when all inputs (*P&Q&V*) are available (Case 1), and the validation process takes up to 1 sec. The estimation error of weighting factors (i.e.

both WFE and AWFE), controllable load contributions (i.e. both WFE and AWFE) and load models (i.e. ALME) when all inputs are available are presented and discussed in this section. The cases in which the input is missing are discussed in Section 5.2.5.

### 5.2.4.1 Estimation Error of Weighting Factors

#### *A. WFE of Individual Load Categories*

Figure 5.11(a) shows PDFs of the weighting factor errors (WFEs) of different load categories when all inputs are available. The letter 'C' or 'UC' inside the bracket in the legend stands for 'controllable' or 'uncontrollable' load. It shows that the most probable values of WFEs (i.e. WFE with highest probability density) for all categories range from -1% to 6%. Figure 5.11(b) shows CDFs of WFE for all load categories. The confidence level (CL) of estimation with required WFE for each load category can be read from it, and it shows that for different load categories, there are from 60% (i.e. 0.8-0.2=0.6=60%) to 95% (i.e. 1-0.05=0.95=95%) of WFEs at the interval [-10%, 10%]. After estimating the percentages, those CDFs can be used to generate probabilistic decomposed daily loading curves (mentioned and defined in Chapter 1 and Chapter 3) which indicate probabilistically distributed participation of different load categories. For the same reason, the weighting factor error (WFE) instead of the absolute weighting factor error (AWFE) is sometimes used for error analysis of individual load categories.

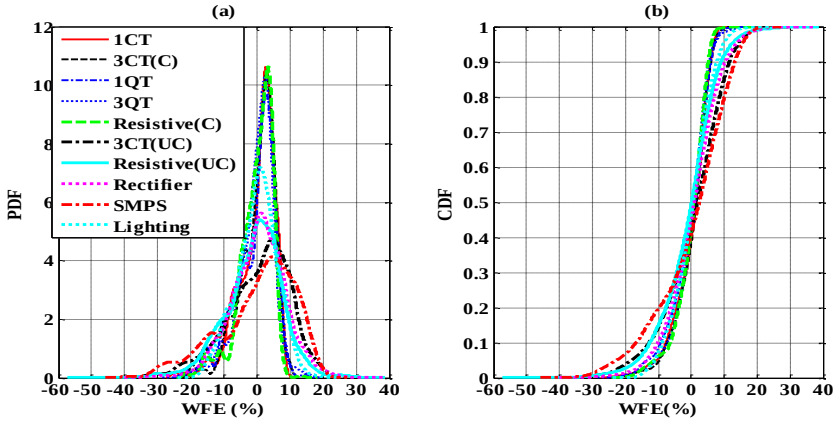


Figure 5.11. (a) PDF, (b) CDF of WFE for individual load categories when all inputs are available

### *B. AWFE of Individual Load Categories*

Figure 5.12 shows PDFs and CDFs of absolute weighting factor errors (AWFEs) of different load categories when all inputs are available. It shows that the most probable values of AWFEs for all categories range from 0 to 5%, mostly around 2%. Figure 5.12(b) shows CDFs of AWFE for all load categories. It shows that for different load categories, there are from 30% to 80% of AWFEs within [0, 5%], from 60% to 95% within [0, 10%], and from 80% to 99% within [0, 15%].

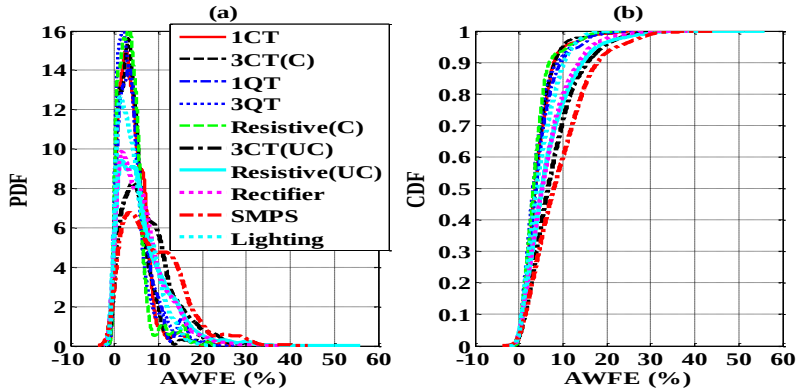


Figure 5.12. (a) PDF, (b) CDF of AWFE for individual load categories when all inputs are available

### *C. Overall AWFE of the Disaggregation Approach*

The probability density functions (PDFs) or the cumulative distribution functions (CDFs) of different load categories need to be aggregated to an overall PDF or CDF to assess the performance of the disaggregation approach via the weighted average process defined by Equation (5.10). However, the most probable absolute weighting factor errors (AWFEs) of all individual load categories are not likely to occur simultaneously. In other words, the AWFEs of different load categories for each *P&Q&V* input set are probably correlated. Considering this possible circumstance, the probability aggregation process is implemented for every *P&Q&V* set immediately after AWFE of each individual load category is obtained ($\boldsymbol{e_i}$ in Part 1, Figure 5.10), rather than after PDFs and CDFs for all individual load categories are produced.

Figure 5.13 shows aggregate PDF and CDF of the absolute weighting factor

error (AWFE) for the load disaggregation approach. From PDF, it is observed that the error is most likely to be about 6%. On the data cursor of CDF, it can be read that the confidence level of the disaggregation approach is about 23% at the confidence interval [0, 5%], 86% at the confidence interval [0, 10%], and 99% at the confidence interval [0, 15%]; the confidence interval with a confidence level of 90% is [0, 10.5%]. In other words, for every *P&Q&V* measurement set, the error of the load disaggregation approach is no more than 15% in 99% of the cases, no more than 10% in 86% of the cases, and no more than 5% in 23% of the cases; 90% of all estimations are with AWFE less than 10.5%.

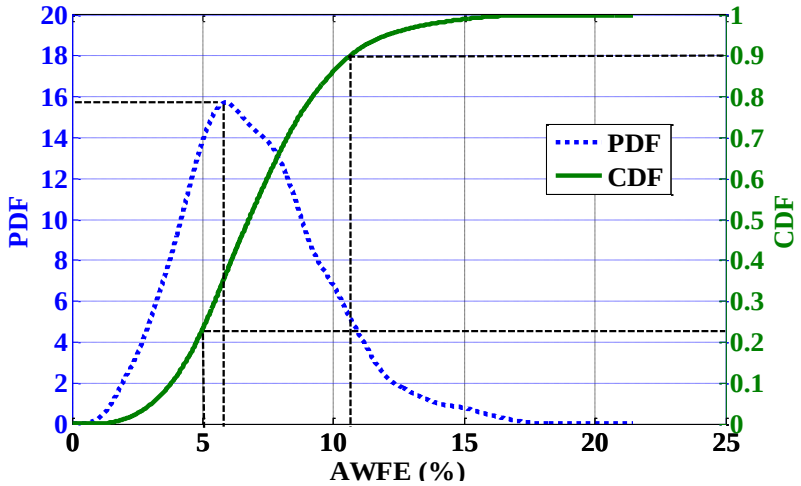


Figure 5.13. PDF and CDF of AWFEs for the disaggregation approach when all inputs are available

By comparing CDFs in Figure 5.12(b), it can be seen that at 5% the cumulative probability of the aggregate absolute weighting factor error (AWFE) is lower than any of the individual AWFE; and at 10%, the cumulative probability of the aggregate AWFE is in the middle among all the individual AWFEs. Simultaneously, from PDFs in Figure 5.12(a), it can be seen that the most probable value of the aggregate AWFE is about 6% and of any of the individual AWFEs is within the interval [0, 5%]. These facts verify the consideration of correlations among different AWFEs, justify the error aggregation for every input set immediately after the individual AWFEs are computed, and repudiate the routine of individual PDF and CDF aggregation, all discussed in Part ***A*** of Section 5.2.3.2.

### 5.2.4.2 The Error of Estimation of the Contribution of the Controllable Load to the Total Load Mix

Figure 5.14(a) shows PDFs and CDFs of the absolute weighting factor errors (AWFEs) of the controllable and uncontrollable load when all inputs are available. They are found to be coincident, which was expected, because for each calculation, the weighting factor error (WFE) of the controllable load is the additive inverse of that of the uncontrollable load, and vice versa. The most probable absolute weighting factor error (AWFE) for the controllable load (as well as the uncontrollable load) is around 4.3%. Figure 5.14(b) presents PDFs and CDFs of WFEs of the controllable and uncontrollable load. It shows that PDFs of WFEs of the controllable and uncontrollable load are symmetrical with respect to the *y* axis. This phenomenon vindicates the coincidence of two PDF curves or two CDF curves of AWFEs shown in Figure 5.14(a). It can be read from Figure 5.14(a) that the confidence level (CL) with an AWFE below 5%, 10%, 15%, 20% is about 32%, 60%, 78%, 90% respectively.

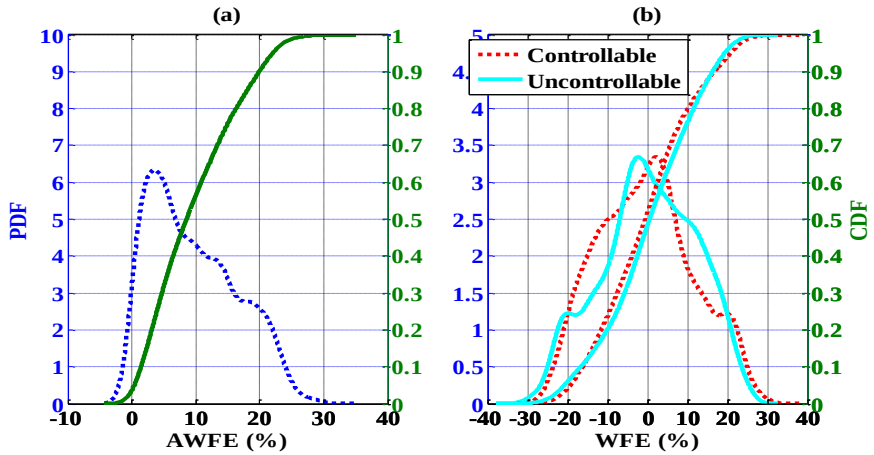


Figure 5.14. (a) PDF and CDF of AWFE of controllable or uncontrollable load, and (b) PDF and CDF of WFE of controllable and uncontrollable load, when all inputs are available

### 5.2.4.3 ALME for both *P* and *Q*

Figure 5.15 shows the PDFs and CDFs of the absolute load model error (ALME) for the real and reactive power (*P* and *Q*).

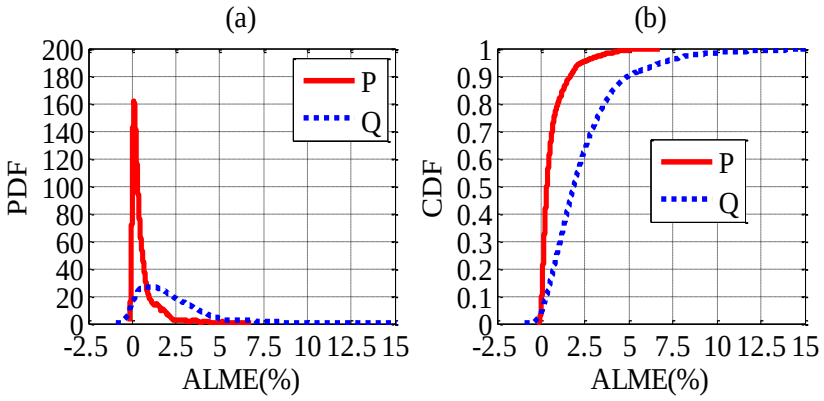


Figure 5.15. (a) PDFs, and (b) CDFs of ALME of *P* and *Q* when all inputs are available

From PDF, ALME with highest probability density occurs at about 0.11% for *P* and 1.2% for *Q*. From CDF, it can be seen that ALME of *P* is less than 2.5% in over 95% of the cases. While ALME of *Q* is less than 7.5% in over 95% of the cases, i.e., ALME of *Q* is more affected by weighting factor deviations than ALME *P* is. This can be explained with the sensitivity of the ALME to the WFE. From Equation (5.11) and (5.12), the *complete differential* of $e_P$ and $e_Q$ at Point ($\Delta w_1$, ...,$\Delta w_i$,... ,$\Delta w_N$ ) can be written as

$$\partial e_P = P_0 \left| \sum_{i=1}^{N} \left( \frac{V}{V_0} \right)^{\alpha_i} \partial \Delta w_i \right| \tag{5.13}$$

and

$$\partial e_Q = P_0 \left| \sum_{i=1}^{N} \sqrt{\left( \frac{1}{pf_{0i}} \right)^2 - 1} \cdot \left( \frac{V}{V_0} \right)^{\beta_i} \partial \Delta w_i \right| \tag{5.14}$$

respectively, where $\Delta w_i$ is the WFE of load category *i* and it can be positive, zero or negative.

On the one hand, for individual load categories, the absolute value of $\alpha$ is smaller than $\beta$ under the same voltage, which contributes to the larger sensitivity of ALME of *Q* than that of *P*. On the other hand, when there is a large number of devices with a nominal power factor lower than 0.707 (i.e., rectifiers with passive power

factor correction circuits, some induction motors, etc.), the sensitivity of ALME of $Q$ to WFE will increase faster, especially when the actual voltage is higher than 1p.u and when the WFE for these individual load categories are positive. However, when the WFE of some individual load categories is negative and their sensitivities are relatively large, the ALME of $Q$ could be smaller than that of $P$. Figure 5.16 shows that most of the points ($e_{Pi}$, $e_{Qi}$) fall above the line $e_Q = e_P$. In other words, ALME of $P$ is smaller than that of $Q$ in most cases.

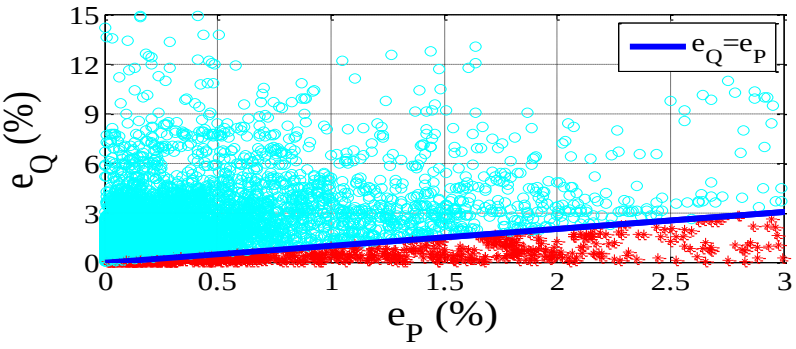


Figure 5.16. $e_P$-$e_Q$ plot against Line $e_Q$=$e_P$

Compared with Figure 5.13, Figure 5.15(b) also shows that within the same confidence interval, ALME has a higher confidence level than AWFE. This means that the moderate variation in load composition estimation does not significantly affect the estimation of aggregate load characteristics.

## 5.2.5 Effect of Missing Input Data on the Accuracy of Load Disaggregation

In this section, the case when $V$ is unavailable (Case 2) and when $Q$ is unavailable (Case 3) are discussed, and they are compared with the case when all inputs are available (Case 1). The rules for ANN parameter configuration is similar as mentioned for Case 1.

### 5.2.5.1 Voltage Input Data are unavailable (Case 2)

#### *A. WFE of Individual Load Categories*

Figure 5.17 shows PDFs and CDFs of weighting factor errors (WFEs) of different load categories when *V* measurement is missing. It can be seen that the most probable values of WFEs now range from -1%-7%, which shows a slight increase in magnitude when compared with -1%-6% in Figure 5.11(a). Figure 5.17(b) indicates that for different load categories, there are from 58% to 93% of WFEs within [-10%, 10%], which shows a slight decrease when compared with Figure 5.11(b).

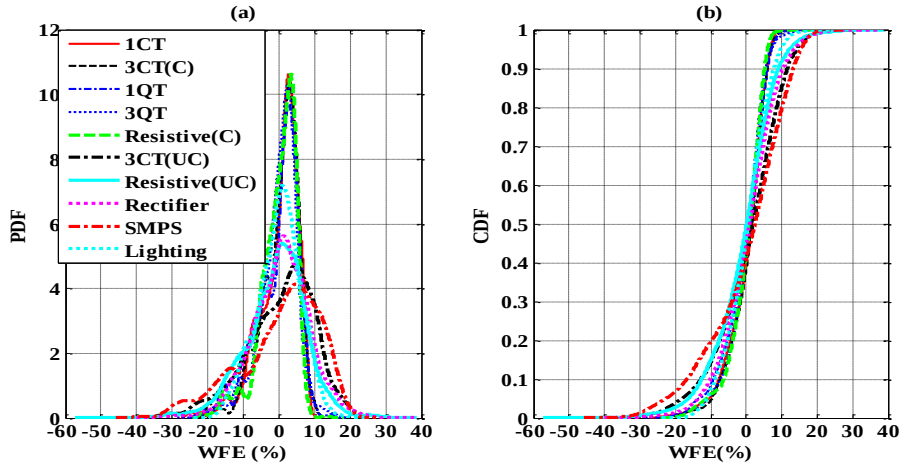


Figure 5.17. (a) PDF, and (b) CDF of WFE of different load categories when *V* measurement is missing

### *B. AWFE of Individual Load Categories*

Figure 5.18 shows PDFs and CDFs of AWFEs of different load categories when *V* measurement is missing.

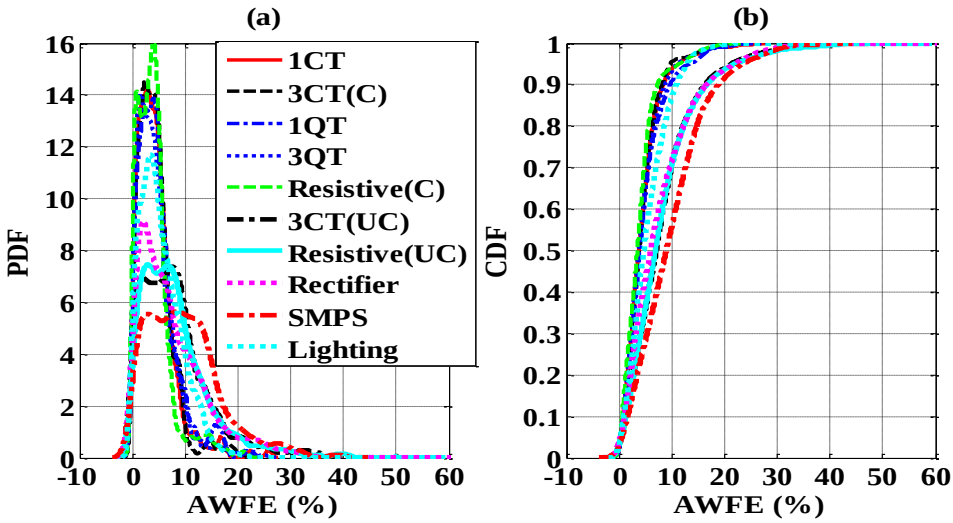


Figure 5.18. (a) PDF, and (b) CDF of AWFEs of different load categories when *V* measurement is missing

Compared with Figure 5.12(a), it can be seen in Figure 5.18(a) that the most probable values of WFEs increase from 0-5% to 2%-7%, mostly around 4% or more. Figure 5.18(b) shows CDFs of AWFE for all load categories. It shows that for different load categories, there are from 30% to 78% of AWFEs within [0, 5%], from 55% to 95% within [0, 10%], from 80% to 97% within [0, 15%], which shows a slight decrease when compared with Figure 5.12.

### *C. Overall AWFE of the Disaggregation Approach*

Similar to Case 1, the aggregate PDF and CDF of the absolute weighting factor error (AWFE) for the load disaggregation approach in Case 2 are shown in Figure 5.19.

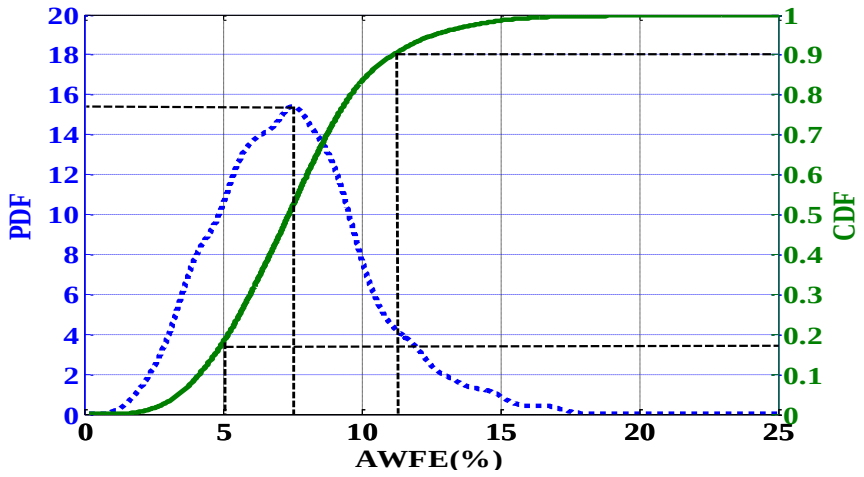


Figure 5.19. PDF and CDF of AWFEs for the disaggregation approach when *V* measurements are unavailable

From the given PDFs, it can be seen that the error is most likely to be about 7.5%. On the data cursor of CDF, it can be read that the confidence level of the disaggregation approach is about 18% at the confidence interval [0, 5%], 83% at the confidence interval [0, 10%], and 99% at the confidence interval [0, 15%]; the confidence interval with a confidence level of 90% is [0, 11.5%]. Therefore, in Case 2, the error of the load disaggregation approach is no more than 15% in 99% of the cases, no more than 10% in 83% of the cases and no more than 5% in 18% of the cases; 90% of all

estimations are with AWFEs lower than 11.5%. Compared with Case 1, it can be seen that at the same given confidence interval, the confidence level in Case 2 is slightly lower than that in Case 1.

### *D. The Error of Estimation of the Contribution of the Controllable Load to the Total Load Mix*

Figure 5.20(a) shows the PDF and CDF of the absolute weighting factor error (AWFE) of the controllable load (as well as the uncontrollable load) in Case 2. The most probable AWFE of the controllable and uncontrollable load contribution is around 5.2%. Figure 5.20(b) presents PDFs and CDFs of WFEs of the controllable and uncontrollable load. Again, similar to Case 1, it shows that PDFs of WFEs of the controllable and uncontrollable load are symmetrical with respect to the *y* axis, and the reasons are given in Section 5.2.4.2. It can be read from Figure 5.20(a) that the confidence level with an AWFE of the controllable and uncontrollable load contribution below 5%, 10%, 15%, 20% is about 29%, 54%, 75%, 89% respectively. When compared with Case 1, at the same given confidence interval, AWFE of the controllable and uncontrollable load also decreases, although the reduction is moderate.

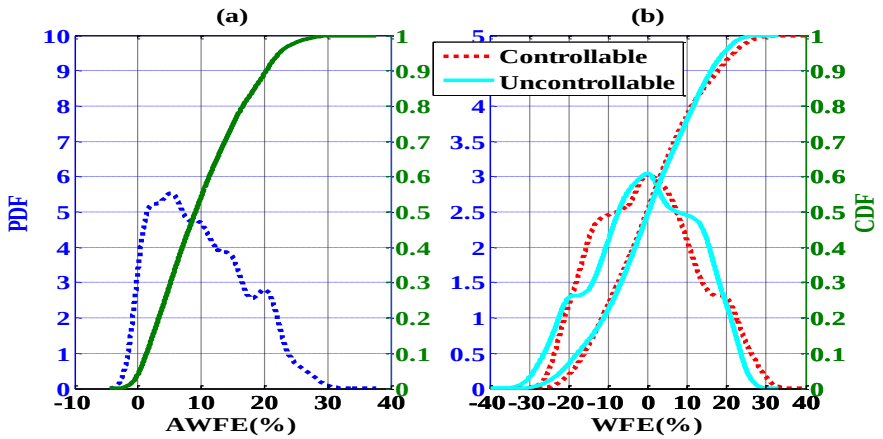


Figure 5.20. (a) PDF and CDF of AWFE of controllable or uncontrollable load, and (b) PDF and CDF of WFE of controllable and uncontrollable load in Case 2

The results of Case 2 show that the missing *V* measurements only slightly

reduce the confidence level of the same estimation errors. The absolute load model error (ALME) is not analysed in Case 2 due to the absence of *V* measurements.

#### 5.2.5.2 Reactive Power Input Data are unavailable (Case 3)

In some cases, RMS measurements of the reactive power (*Q*) are not available (Case 3). Therefore, the results for the weighting factor error (WFE), the absolute weighting factor error (AWFE) and the absolute load model error (ALME) of *P* in this case are obtained as well and compared with the case when the voltage, the real and reactive power are all available (Case 1). The rules of ANN parameter configuration, the ANN training algorithm, the training data and the testing data (apart from *Q*) are exactly the same as those for Case 1 and Case 2.

***A. WFE of Individual Load Categories***

Figure 5.21(a) and Figure 5.21(b) show PDFs and CDFs of the weighting factor errors (WFEs) of different load categories when the measurements of *Q* are missing. Compared with Figure 5.17(a), it can be seen that the most probable values of WFEs increase slightly, but the increase is moderate. Figure 5.17(b) indicates that for different load categories, there are from 57% to 93% of WFEs within [-10%, 10%], which shows a slight decrease of 1% when compared with Figure 5.17(b).

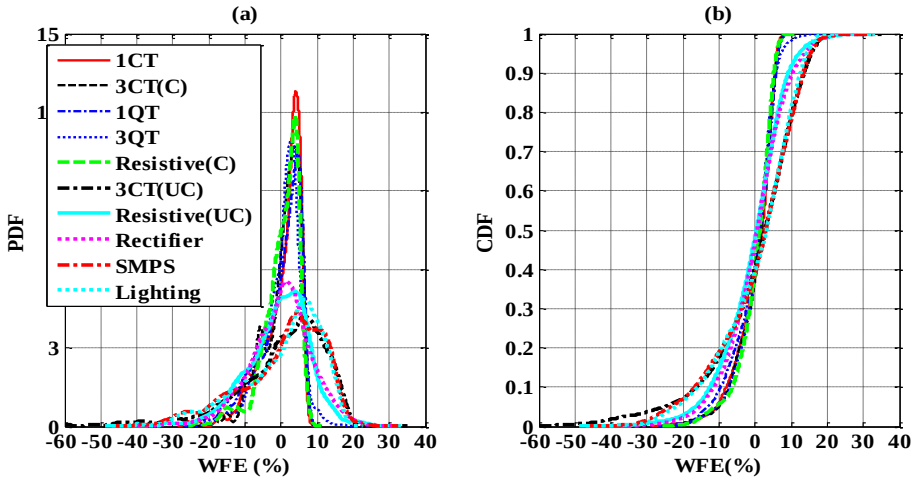


Figure 5.21. (a) PDFs, and (b) CDFs of WFEs of different load categories when the measurements of *Q* are missing.

### *B. AWFE of Individual Load Categories*

Figure 5.22(a) and Figure 5.22(b) show PDFs and CDFs of the absolute weighting factor errors (AWFEs) of different load categories when the measurements of $Q$ are missing. It can be seen from Figure 5.22(a) that the most probable values of WFEs lie between 2%-8%. From Figure 5.22(b), it can be seen that for different load categories, there are from 30% to 76% of AWFEs within [0, 5%], from 59% to 95% within [0, 10%], and from 83% to 97% within [0, 15%]. The results are similar to those obtained in Case 2.

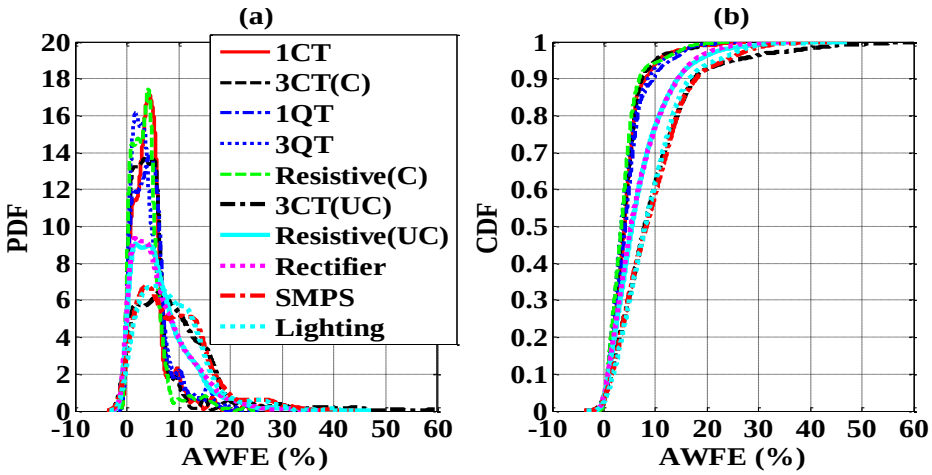


Figure 5.22. (a) PDFs, and (b) CDFs of AWFEs of different load categories when $Q$ measurement is missing.

### *C. Overall AWFE of the Disaggregation Approach*

The aggregate PDF and CDF of the absolute weighting factor error (AWFE) of the load disaggregation approach, when measurements of $Q$ are not available, are shown in Figure 5.23. From the PDF, it can be seen that similar to Case 2, the error is most likely to be about 7.5%. On the data cursor of the CDF, it can be read that the confidence level of the disaggregation approach is about 15% at the confidence interval [0, 5%], 82% at the confidence interval [0, 10%], and 98% at the confidence interval [0, 15%]; the confidence interval with a confidence level of 90% is [0, 12%]. Therefore, in Case 3, the error of the load disaggregation approach is up to 15% in

98% of the cases, is up to 10% in 82% of the cases and is up to 5% in 15% of the cases; 90% of all estimations are with AWFEs less than 12%. Compared with Case 2, it could be found that at the same given confidence interval, the confidence level in Case 3 is slightly lower than that in Case 2, although the decrease is moderate.

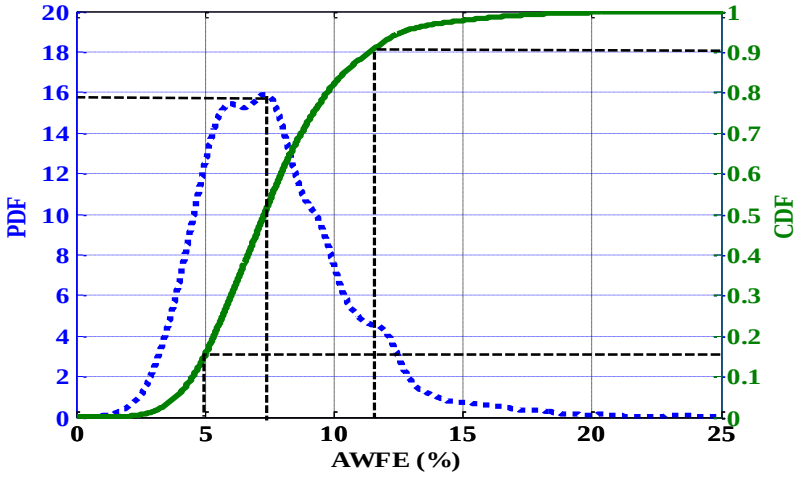


Figure 5.23. PDF and CDF of AWFEs for the disaggregation approach when $Q$ measurements are unavailable

### *D. The Error of Estimation of the Contribution of the Controllable Load to the Total Load Mix*

Although only moderate changes are observed in PDFs and CDFs of WFE and AWFE of individual load categories in Case 3, this does not mean that the WFE or AWFE of the controllable and uncontrollable load will not significantly change. The changes of WFE or AWFE of the controllable and uncontrollable load could significantly affect demand side control action.

Figure 5.24(a) shows the PDF and CDF of AWFE of the controllable and uncontrollable load when measurements of $Q$ are missing. Although the most probable AWFE is about 2.7% (smaller than that in Case 2), the peak of PDF is not as sharp as that in Case 1 or Case 2, and AWFEs lying between 0 and 15% have almost the same probability density. The confidence level of an AWFE below 5%, 10%, 15%, 20% is about 25%, 48%, 70%, 90% respectively. It shows that when $Q$ measurement is

missing, the confidence level of AWFE under the same error interval will considerably decrease.

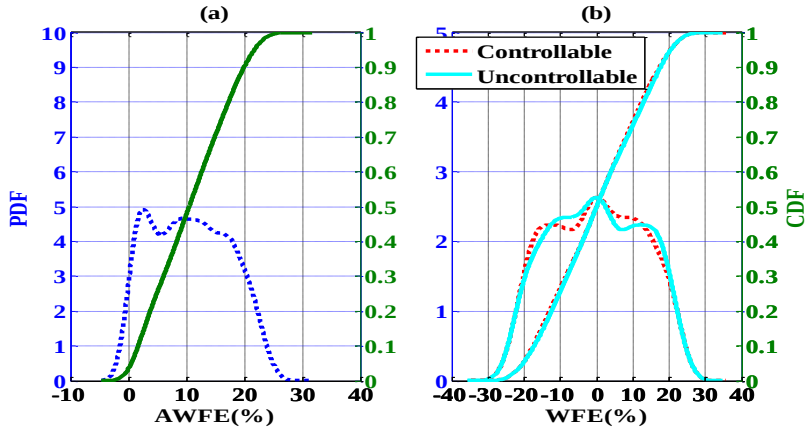


Figure 5.24. (a) PDF and CDF of AWFE of controllable or uncontrollable load, and (b) PDF and CDF of WFE of controllable and uncontrollable load when $Q$ is missing.

Therefore, the availability of $Q$ measurements plays a significant role in high-accuracy and high-reliability estimation of load participation. The absence of the measurements of $Q$ could significantly distort the result. Figure 5.24(b) shows the PDFs and CDFs of WFE of the controllable and uncontrollable load when measurements of $Q$ are missing. Similar as before, PDFs of WFE of the controllable and uncontrollable load are symmetrical with respect to the $y$ axis, and the PDF of AWFE of the controllable and uncontrollable loads are coincident. The reasons for this are explained in Section 5.2.4.2.

### *E. ALME for P*

Figure 5.25(a) shows the PDF of the absolute load model error (ALME) for the aggregate real power ($P$). The absolute load model error (ALME) for the aggregate $Q$ is not assessed due to the absence of the measurements of $Q$. It shows that the most probable ALME for $P$ is about 1.27%. Figure 5.25(b) shows the CDF of ALME for $P$. It indicates that only 31% of all ALMEs for $P$ are within 2.5%, 51% within 5%, 66% within 7.5%, and 81% within 10%, which is a considerable decrease when compared with Case 1. Like the AWFE of the controllable and uncontrollable loads, the results

of ALME also show that the absence of measurements of $Q$ measurements will significantly degrade the performance of the disaggregation approach.

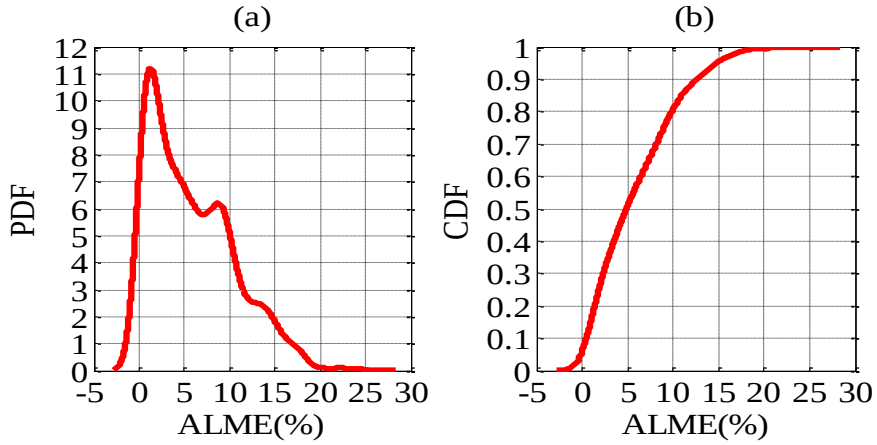


Figure 5.25. (a) PDFs, and (b) CDFs of ALMEs for aggregate $P$ when measurements of $Q$ are missing.

### 5.2.6 Influence of the Prior Knowledge on the Contribution of the Controllable Load

As mentioned in Section 5.1.3.2, the percentage of the controllable load in the UK power systems ranges from 10% to 50% based on the customer surveys. Therefore, in the ANN training process, when the weighting factors of different load categories are being generated, the variation range of the controllable load can be set to [10%, 50%]. However, the information about the share of the controllable load in the total demand is not always known. In such a case, the controllable load should be set to vary within the range [0, 100%], and the number of different combinations of the controllable and uncontrollable loads increases from 9 (i.e. explained in Section 5.1.3.2 and 5.2.3.2) to 21. For ease of comparison, the total number of training sets and validation sets used in this case is similar to the case where the controllable load range is known.

The PDFs and CDFs of the aggregate absolute weighting factor errors (AWFEs) of Case 1, Case 2 and Case 3, when the information of the share of the controllable load in the total demand is unknown, are shown in Figure 5.26.

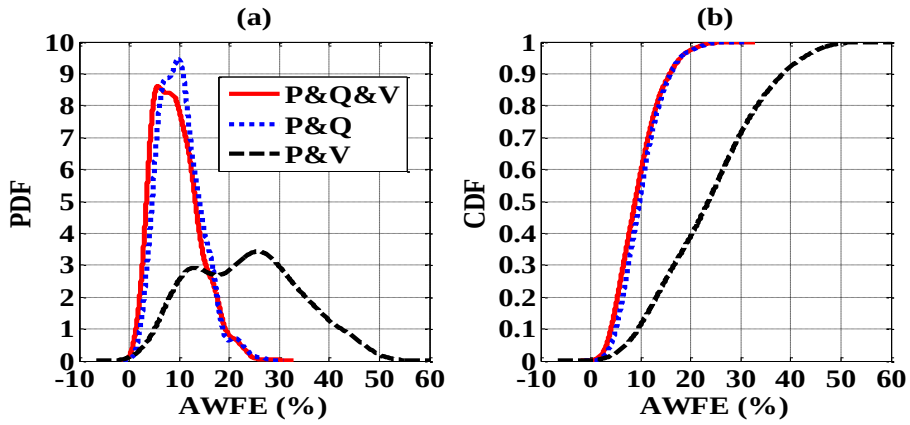


Figure 5.26. (a) PDF, and (b) CDF of AWFE in Case 1-3 when the share of the controllable load is unknown

### 5.2.6.1 Comparison of the Absolute Weighting Factor Error (AWFE) from Cases 1 to 3 when the Share of the Controllable Load is Unknown

#### *A. Absolute Weighting Factor Error of Case 1*

The comparative results of the overall absolute weighting factor error (AWFE) of the disaggregation approach in Case 1 can be read from the corresponding PDFs and CDFs and are shown in Table 5.3.

Table 5.3. Comparison of the Overall AWFE of the Disaggregation Approach in Case 1 when the Share of the Controllable Load is 10%-50% and when It is Unknown (MP: most probable; CL: confidence level)

| Share of the Controllable Load | MP AWFE | CL of AWFE≤5% | CL of AWFE≤10% | CL of AWFE≤15% |
|---|---|---|---|---|
| [10%, 50%] | 6% | 23% | 86% | 99% |
| Unknown | 6% | 17% | 59% | 87.5% |

When all inputs are available, the most overall AWFE for the disaggregation approach is about 6%, which is almost the same as the case when the information of the controllable load variation range is 10%-50%. The confidence level (CL) of the disaggregation approach is about 17% at the confidence interval [0, 5%], 60% at the confidence interval [0, 10%], and 87.5% at the confidence interval [0, 15%].However, when the information of the controllable load participation is known, the confidence levels at the three confidence intervals are 23%, 86% and 99%, respectively. The result indicates that the preceding acquisition of the information on controllable load

participation can increase the confidence level of disaggregation under a specific disaggregation error. It is probably because the number of controllable-uncontrollable load combinations is reduced from 21 to 9, and if a similar number of training sets and validation sets are taken, then the number of voltages and weighting factors generated by MCS in each combination increases. This results in ANN being trained effectively with more detailed information, and the trained ANN thus works better.

### *B. Absolute Weighting Factor Error of Case 2*

The comparative results of the overall absolute weighting factor error (AWFE) of the disaggregation approach in Case 2 are shown in Table 5.4.

Table 5.4. Comparison of the Overall AWFE of the Disaggregation Approach in Case 2 when the Share of the Controllable Load is 10%-50% and when It is Unknown (MP: most probable; CL: confidence level)

| Share of the Controllable Load | MP AWFE | CL of AWFE≤5% | CL of AWFE≤10% | CL of AWFE≤15% |
|---|---|---|---|---|
| [10%, 50%] | 7.5% | 15% | 82% | 98% |
| Unknown | 10% | 10% | 52.5% | 86% |

If the measurements of voltages are unavailable, the most probable overall AWFE for the disaggregation approach is about 7.5%, which represents an increase of 1.5% when compared with the case where the information of the controllable load variation range is known as 10%-50%. The confidence levels of the disaggregation approach at confidence intervals [0, 5%], [0, 10%], and [0, 15%] are 10%, 52.5% and 86%, respectively. However, when the information on the controllable load participation is known, the confidence levels at the three confidence intervals are 15%, 82%, 98% respectively. The result strengthens the conclusion drawn from Part ***A*** of this subsection (i.e. Section 5.2.6.1) that the availability of the information on controllable load participation in the total demand can increase the confidence level of disaggregation under a specific disaggregation error. Moreover, the result also indicates that the availability of the information on controllable load participation can help reduce the most probable overall AWFE for the disaggregation approach.

### *C. Absolute Weighting Factor Error of Case 3*

The comparative results of the overall AWFE of the disaggregation approach in Case 3 are shown in Table 5.5. When $Q$ measurements are unavailable, the most probable overall AWFE for the disaggregation approach is about 25%, which dramatically increases by 17.5% when compared with the case where the information of the controllable load variation range is known as 10%-50%. The confidence levels of the disaggregation approach at confidence intervals [0, 5%], [0, 10%], and [0, 15%] are 2%, 11% and 25.5%, respectively. However, when the information of controllable load participation is known, the confidence levels at the three confidence intervals are 15%, 82%, 98%, respectively. The result strengthens the conclusions drawn from Part ***A*** and Part ***B*** of this subsection (i.e. Section 5.2.6.1).

Table 5.5. Comparison of the Overall AWFE of the Disaggregation Approach in Case 3 when the Share of the Controllable Load is 10%-50% and when It is Unknown (MP: most probable; CL: confidence level)

| Share of the Controllable Load | MP AWFE | CL of AWFE≤5% | CL of AWFE≤10% | CL of AWFE≤15% |
|---|---|---|---|---|
| [10%, 50%] | 7.5% | 15% | 82% | 98% |
| Unknown | 25% | 2% | 11% | 25.5% |

## 5.2.6.2 Comparative Results of Cases 1 to 3 when the share of the controllable load is unknown

### 5.2.6.2.1 Overall Absolute Weighting Factor Error (AWFE)

The comparative results of the overall absolute weighting factor error (AWFE) of the disaggregation approach in Case 1-3 when the share of the controllable load is unknown are shown in Table 5.6.

Table 5.6. Comparison of the Overall AWFE of the Disaggregation Approach in Case 1-3 when the Share of the Controllable Load is Unknown (MP: most probable; CL: confidence level)

| Case No. | MP AWFE | CL of AWFE≤5% | CL of AWFE≤10% | CL of AWFE≤15% |
|---|---|---|---|---|
| 1 | 6% | 17% | 59% | 87.5% |
| 2 | 10% | 10% | 52.5% | 86% |
| 3 | 25% | 2% | 11% | 25.5% |

The most probable overall AWFE for the disaggregation approach in Case 1, 2 and 3

are 6%, 10% and 25%, respectively. The confidence levels in Case 1, 2 and 3 are 17%, 10% and 2%, respectively at confidence interval [0, 5%], 59%, 52.5% and 11% respectively at confidence interval [0, 10%], 87.5%, 86% and 25.5% respectively at confidence interval [0, 15%].

From the comparative result, it can be seen that under the same conditions, the overall AWFE of Case 1 and Case 2 are similar, although the result of Case 1 is slightly better than that of Case 2 probably because in Case 1, more information is available. This can also be observed from Figure 5.26, where the CDF of the overall AWFE in Case 1 and Case 2 are very close to each other. However, the result of Case 3 is much less accurate than the result of Case 1 and Case 2. From Figure 5.26, it can be seen that the CDF of the overall AWFE in Case 3 is far from CDFs for Case 1 and 2, and obviously, it has a much lower confidence level at the same given confidence interval when compared with Case 1 and 2. Additionally, it has a most probable AWFE of 25%. This also indicates that the reactive power measurement plays a significant role in accurate and reliable load disaggregation, especially when the information of the controllable load contribution is unknown.

#### 5.2.6.2.2 Absolute Load Model Error (ALME) for the real power

Figure 5.27 shows PDFs and CDFs of the absolute load model errors (ALMEs) for the real power ($P$) in Case 1 and Case 3. ALME for the reactive power ($Q$) is not discussed because the measurements of it are not available in Case 3. Case 2 is not discussed here due to the absence of measurements of the voltage ($V$). The most probable ALME for $P$ are much smaller in Case 1 than in Case 3. The confidence levels within the same confidence interval are much higher in Case 1 than in Case 3. The visible difference between the two cases again shows that the ANN will perform the disaggregation approach better with all $P$, $Q$, and $V$ RMS measurements available than with only $P$ and $V$ available, especially when the participation information of

controllable loads are unknown. It can be inferred that for *P&V* input only, there is a large number of possible combinations of different load compositions that can describe the *P-V* relationship, while the addition of *Q* measurement narrows the range of the possibilities as it determines the power factor of the total load and the approximate participation of different load categories.

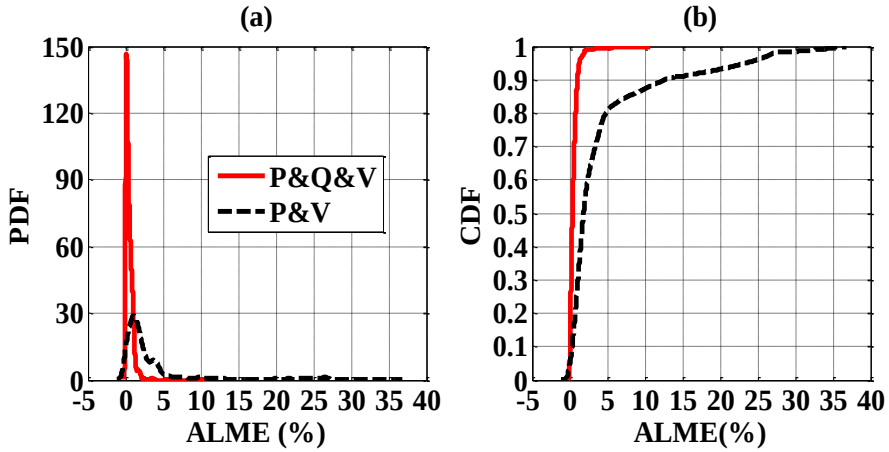


Figure 5.27. (a) PDF, and (b) CDF of ALME for *P* in Case 1 and Case 3 when the share of the controllable load is unknown

## 5.3 Integration with Total Demand Forecasting Tool

It has been proved in Section 5.2.5 and 5.2.6 that no matter whether the contributions of controllable loads are known or not, the absence of voltage measurements will not dramatically reduce the disaggregation accuracy. Therefore, once validated, the load disaggregation tool could be integrated with total demand forecasting tool in order to predict the contribution of different load compositions in the future. The only task needed is to transform the predicted MW and MVAR into per-unit values, and the way of doing it has been discussed in Section 5.1.

### 5.3.1 Joint Probability of Total Demand Forecasting Error and AWFE

In the disaggregation approach, the confidence level of the accuracy of the input (i.e. *P&Q&V* or a subset of them) is 100%. However, when the disaggregation tool is

integrated with the total demand forecasting tool to predict load composition, it is highly unlikely to guarantee that the accuracy of the input (i.e. predicted MW and MVAR) has a confidence level of 100%. As a result, a joint probability of total demand forecasting error and the absolute weighting factor error (AWFE) is introduced in the case where the uncertainty assessment is required.

The probability density function (PDF) and the cumulative distribution function (CDF) of the absolute weighting factor errors (AWFEs) in the disaggregation approach are denoted as $f_D(e_D)$ and $F_D(e_D)$, where $e_D$ is the notation of the overall AWFE for the disaggregation approach. They can also be interpreted as PDF and CDF of load composition forecasting error when CAPE of total demand forecasting, $e_F$, has a confidence level of 100%. The probability density function (PDF) and the cumulative distribution function (CDF) of CAPE of total demand forecasting are denoted as $f_F(e_F)$ and $F_F(e_F)$.

In reality, the confidence level of specific total load forecasting error interval is not likely to reach 100% under most circumstances. Thus, the actual confidence level of load composition forecasting error, which is the joint confidence level of total load forecasting error and AWFE, is lower than the confidence level of AWFE. Because the forecasting process and the disaggregation process are independent, the joint confidence level is the product of the confidence level (i.e. read from CDF) of the total load forecasting error and AWFE, shown as

$$F_{FD}\left(e_F, e_D\right) = F_F\left(e_F\right) F_D\left(e_D\right) \tag{5.15}$$

A similar relationship applies to the joint PDF $f_{FD}(e_F, e_D)$, shown as

$$f_{FD}\left(e_F, e_D\right) = f_F\left(e_F\right) f_D\left(e_D\right) \tag{5.16}$$

### 5.3.2 Illustrative Results of Load Composition Forecasting

The probability density function (PDF) and the cumulative distribution function

(CDF) of the cumulative absolute percentage error (CAPE, defined in Chapter 4) of total demand forecasting shown in Figure 4.10 and those of the absolute weighting factor error (AWFE) when the measurements of the voltage are not available shown in Figure 5.19 are used for illustration purpose, and the simulation result is shown in Figure 5.28.

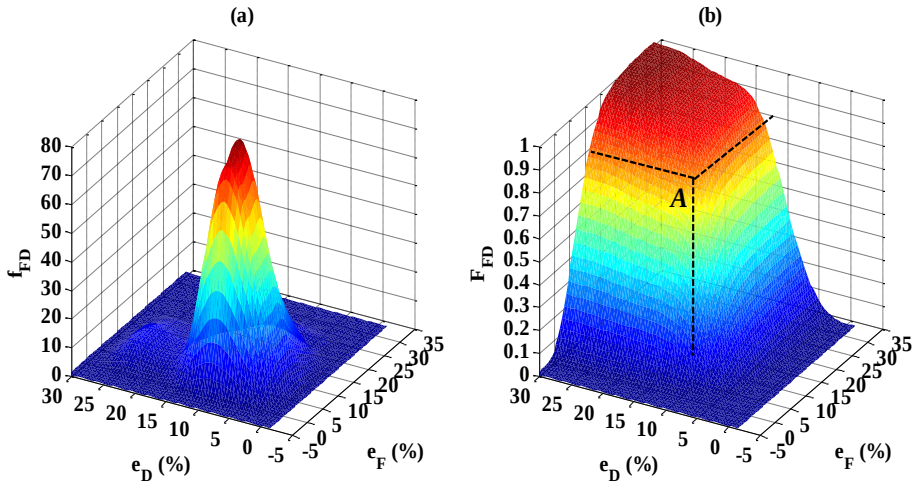


Figure 5.28. (a) PDF, and (b) CDF of load composition forecasting error when total load forecasting error is less than 100%

Figure 5.28(a) shows the joint PDF of load composition forecasting error when the confidence level of total load forecasting error is less than 100%. From Figure 5.28(a), it can be seen that the most likely load composition forecasting error ($e_D$) is about 10% if the total load forecasting error ($e_F$) is about 9.2%, which matches the result shown in Figure 4.10 and Figure 5.19. Figure 5.28(b) provides the confidence level of load composition forecasting error when the confidence level of total load forecasting error is less than 100%. For example, for a load composition forecasting error ($e_D$) lower than 15%, if the total load forecasting error ($e_F$) has a confidence level of 100%, then according to Figure 5.19(b), the confidence level for the load composition forecasting in this case is approximately 90%. While, if the total load forecasting error is also required to be lower than 15%, which has a confidence level of approximately 88% according to Figure 4.10, then the confidence level of the load

composition forecasting becomes 90%×88%=79.2%. The point which represents this example in the order of ($e_D$, $e_F$, $F_{FD}$) is the Point $A$(15, 15, 0.792) clearly marked on the surface plot given as Figure 5.28(b).

## 5.4 Summary

This chapter develops an effective ANN-based methodology for load disaggregation. For convenience, the methodology is developed in per-unit. As a result, another methodology is developed to convert the voltage, the real and reactive power given in kV, MW and MVAR, respectively, to per-unit value before load disaggregation so that they can be applied in the developed disaggregation methodology. The developed disaggregation methodology employs the artificial neural network (ANN) as the tool, and only requires RMS measurements of the voltage, the real and reactive power from substations to identify the contribution of different load categories to the total demand. In most cases, the developed disaggregation methodology can provide an overall absolute weighting factor error (AWFE) less than 10%.

The effect of missing input data on the accuracy of the estimation is also discussed. From the discussion, it can be seen that missing of voltage measurement does not significantly affect the result, while missing of reactive power data can affect the result significantly. As expected, the disaggregation algorithm provides the highest accuracy and confidence level when all inputs are available. Additionally, the effect of prior knowledge on the share of the controllable load to the total demand mix is also investigated. It is found that prior knowledge on the contribution of the controllable load can help select a relatively precise range for the weighting factor used in the ANN training process, and this yields estimation of load compositions with higher accuracy (or higher confidence level for the same accuracy).

Furthermore, the disaggregation approach can be integrated with total demand

forecasting approach to predict the load composition in the future. The developed methodology enables the prediction of the percentage of different load categories (i.e. induction motors, lighting etc.) in the total load mix in advance, with forecasting error typically less than 15%.

# 6 Validation of Developed Methodologies and Illustration of Shaping of Dynamic Responses of Demand

Prediction and shaping of dynamic response of the demand (DRD) will enable advanced control algorithms for active demand management as well as improved stability assessment of the power system. Previous chapters develop methodologies for each stage of the overall flowchart as shown in Figure 1.6, Chapter 1. This chapter integrates the stages and uses measured data from real UK power networks to validate the overall approach. The required input data include standard RMS measurements at bulk supply points and actual and day-ahead forecasted weather data, and the overall approach does not rely on having access to detailed customer surveys or high-resolution load signatures. Measured steady-state and dynamic responses of the demand from the substations of the local utility are used for validation. In addition, based on the prediction results, load shifting and shaping of dynamic responses of the demand are also illustrated.

## 6.1 Validation Results and Discussion

Demand and voltage data used for load forecasting and validation of prediction of dynamic responses of the demand (DRD) are collected from the substation in UK real distribution networks. There are two sampling rates, 1 sample per second and 1 sample per minute. The data collected with the former sampling rate are used for validation of estimation/prediction of load compositions and dynamic responses of the demand, and those with the latter sampling rate can be used for validation of estimation/prediction of load compositions and steady-state load characteristics. Both actual and forecasted weather data, including temperature ($°C$), humidity (p.u) and wind speed ($km/h$), are collected from the official website of Weather Underground [132], with a resolution of 30min. As a result, in total load forecasting, the demand is predicted every 30min.

Data collected at "Bus BRW" with mixed load sectors [1] on 24 June, 2014 are used to predict the demand and DRD at the same bus on 25 June, 2014. Besides, disaggregation based on measurements on 25 June, 2014 is also implemented and validated for comparison purposes.

The validation result of total load forecasting is shown in Figure 4.8, Chapter 4. The most probable values of per-unit real and reactive power derived based on measurements at "Bus BRW" on 24 June, 2014 are given in Figure 5.7(c) and Figure 5.7(d), respectively. Based on these results, validation of estimation/prediction of load compositions and dynamic responses of the demand are given as follows.

### 6.1.1 Load Disaggregation and Validation

In Chapter 3, the framework for prediction/estimation of dynamic responses of the demand is built probabilistically. In the disaggregation approach developed in Chapter 5, the absolute weighting factor errors and the absolute load model errors are

also analysed probabilistically. However, it is worth noting that in the reality, the measured dynamic response of the demand is a curve rather than a range composed of many curves; the curve can be regarded as the most probable response of the aggregate demand. In Chapter 3, it is found that the most probable individual dynamic responses and percentages of different load categories can be used to calculate the most probable dynamic responses of the aggregate load. As a result, instead of directly using the probabilistic approach or framework in Chapter 5 or Chapter 3, the validation process in the reality can be simplified in the following way:

1. Take the most probable values of the load composition percentages (i.e. directly derived using the disaggregation approach in Chapter 5) and the dynamic responses of the demand of individual load categories (i.e. derived in Chapter 3) to calculate the dynamic responses of the demand for validation;
2. If the calculated response matches the measured one reasonably well, the process can be terminated;
3. If the calculated response is far from the measured one, then probabilistic approach will be used to inspect whether the measured response falls within the range of calculated responses.

#### 6.1.1.1 Load Disaggregation on 25 June, 2014

Using the developed disaggregation methodology, with the calculation shown in Figure 5.7(c) and Figure 5.7(d), the decomposed daily loading curve for 25 June, 2014 at “Bus BRW” can be estimated, as shown in Figure 6.1. In the legend of Figure 6.1, 1CT, 3CT, 1QT, 3QT represent different categories of induction motors, R represents resistive load, SMPS represents switch-mode power supply, C and UC represent controllable and uncontrollable load, respectively.

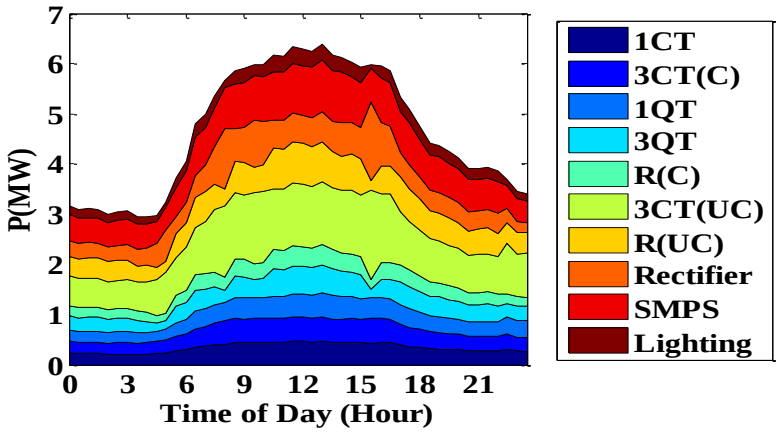


Figure 6.1. Estimated DDLC of 25 June, 2014 based on measurement of the same day

### 6.1.1.2 Validation of Load Disaggregation

#### 6.1.1.2.1 Validation Using Steady-state Load Characteristics

Figure 6.2(a) shows the voltage verses voltage level (i.e. 6.6kV). Figure 6.2(b) and Figure 6.2(c) show the calculated steady-state load characteristics for real and reactive power respectively. It can be seen that the estimated real and reactive power (i.e. *P* and *Q*) characteristics (the red dashed line) match the measured ones (the blue solid line). Therefore, according to the description in Chapter 5, it can be concluded that the disaggregation result shown in Figure 6.1 is reasonable.

#### 6.1.1.2.2 Validation Using Measured DRD

A 0.7% voltage drop (i.e. for UK distribution transformers, each tap causes 0.7% voltage change) occurs at *t=0s* (i.e. at about 3:30, with the lowest demand), shown as Figure 6.2(d). The estimated dynamic responses of the demand (DRD) for real and reactive power are shown as Figure 6.2(e) and Figure 6.2(f), respectively. From the shape of DRD, it can be seen that the estimated dynamic response of the demand matches reasonably well the measured one for both real and reactive power (i.e. the probabilistic framework is therefore not needed in this case). This means that the disaggregation result shown in Figure 6.1 is reasonable. Besides, from Figure 6.2(d)-

(f), when the voltage changes, the real and reactive power remains fairly the same. Therefore, according to load model characteristics provided in [1] and Chapter 3, the overall demand could be modelled by a constant power load model, and the dominant loads could be electronic loads or three-phase induction motors. This also matches derived decomposed daily loading curves shown as Figure 6.1. The mean absolute percentage error (MAPE) of the estimated dynamic response of the demand is 0.30% for real power and 4.74% for reactive power.

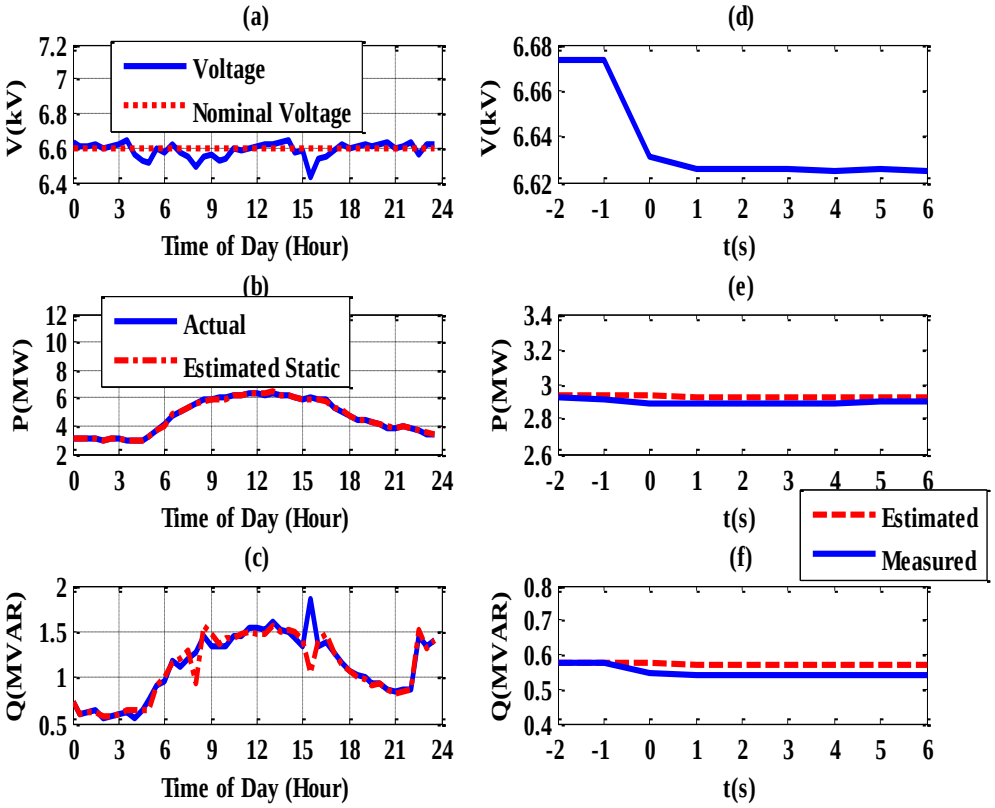


Figure 6.2. (a) Voltage verses voltage level; (b) estimated static *P* verses measured *P*; (c) estimated static *Q* verses measured *Q*; (d) voltage drop at 3:30; (e) estimated dynamic response of the demand of *P* verses measured dynamic response of the demand of *P* at 3:30; (f) estimated dynamic response of the demand of *Q* verses measured DRD of *Q* at 3:30.

### 6.1.2 Prediction of Load Composition on 25 June, 2014

As the disaggregation process has been validated via both the measured steady-state load characteristics and dynamic responses of the demand, it could be connected

with the load forecasting process to predict day-ahead load compositions at different times of the day. The procedure of deriving most probable per-unit values of the predicted real and reactive power and the disaggregation process (i.e. shown as Section 6.1.1.1) are repeated, but this time, the predicted demand for 25 June, 2014 replaces the actual measured demand. With the inputs shown in Figure 1.6, a day-ahead predicted decomposed daily loading curve can be obtained and is shown in Figure 6.3.

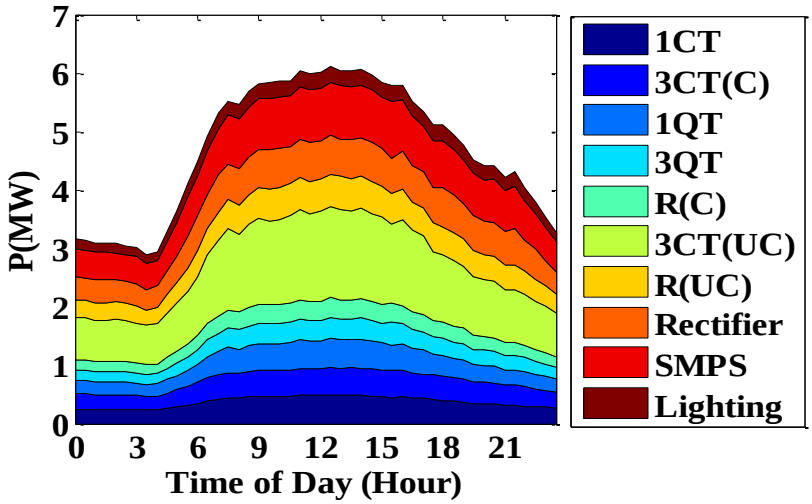


Figure 6.3. Predicted DDLC of 25 June, 2014 based on measured data of 24 June, 2014

From the figure, it can be seen that the lines of the predicted decomposed daily loading curve look smoother than those of the estimated one shown in Figure 6.1. In other words, Figure 6.1, as expected, presents the contribution of different load categories at different times of the day slightly more clearly as it was produced using the measured voltages, real and reactive powers, while Figure 6.3 was produced using the predicted real and reactive power only. Same as Figure 6.1, though, Figure 6.3 indicates that the dominant loads are electronic loads and 3-phase induction motors.

### 6.1.3 Day-ahead Prediction and Validation of DRD

With the predicted decomposed daily loading curve shown in Figure 6.3 and the dynamic response of the demand (DRD) of individual load categories provided in

Chapter 3, predicted DRD of the total demand to the same 0.7% voltage drop is obtained and shown as Figure 6.4(b) and Figure 6.4(c) for *P* and *Q*, respectively.

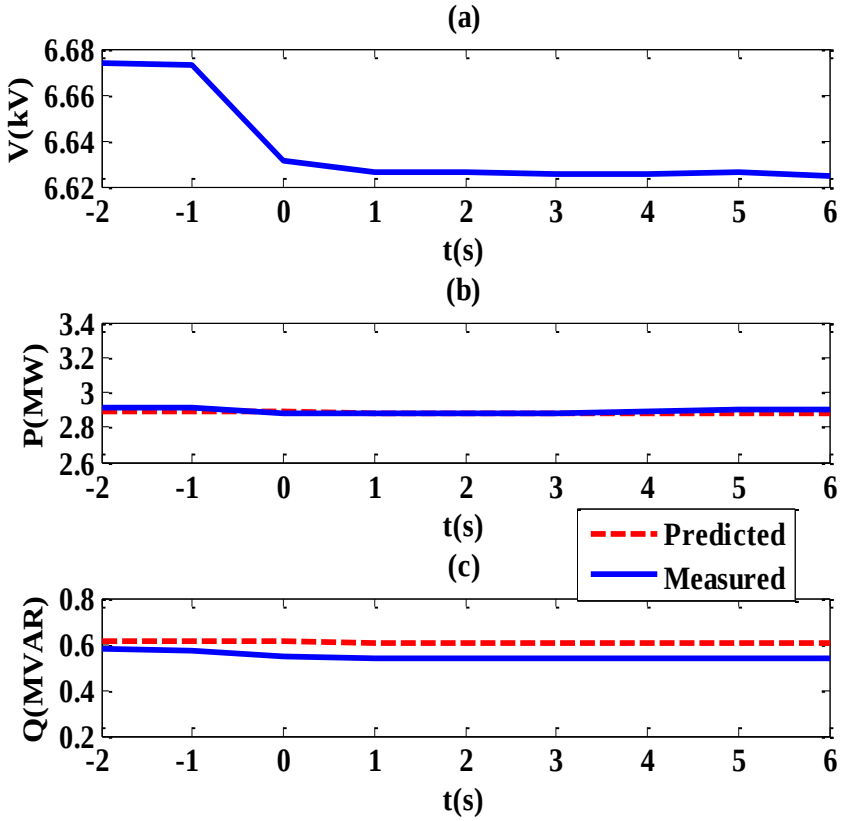


Figure 6.4. (a) A 0.7% voltage drop; (b) predicted DRD of P verses measured DRD of P; (c) predicted DRD of Q verses measured DRD of Q.

Similar to the estimated DRD, the new steady-state value of *P* DRD remains fairly the same after the voltage drop and can be modelled by constant power load model. Therefore, the dominant customer could be 3-phase induction motors or electronic loads, which matches the predicted decomposed daily loading curve in Figure 6.3 and the estimated decomposed daily loading curve in Figure 6.1. The prediction of DRD of *Q* however, is not as good as prediction of *P*, although the difference is moderate. Similar to validation of estimated DRD, probabilistic framework is not needed in this case.

The mean absolute percentage error (MAPE) of DRD prediction is about 0.33%

for *P* and 10.49% for *Q*. MAPE of prediction of DRD of *P* is slightly higher than MAPE of estimated DRD of *P*. For *Q*, MAPE of predicted DRD of *Q* is larger than the estimated one. There are three possible reasons for this "inaccuracy" in prediction of DRD of *Q*:

1. The base value of *Q* is small (in the order of $10^{-1}$);
2. The total demand prediction for *Q* is not as accurate as that for *P*;
3. There are more accurate inputs for the DRD estimation case than for the prediction case.

## 6.2 Load Shifting and Shaping of Dynamic Responses of the Demand

The shape of dynamic responses of demand (DRD) depends on the load types participating in the total demand. In demand side management (DSM), when the load curve is shaped via series of DSM techniques [107] such as load shifting (i.e. shifting certain amount of demand from the peak time to the off-peak time to help balancing the generation and the demand in the network), the load compositions constituting the total demand during those periods will change at the same time. This ultimately changes the dynamic load characteristic (i.e. represented as the shape of DRD), the key factor that influences the voltage stability [117]. For transient voltage stability and for the final stage of a slower occurring voltage collapse, the dynamic characteristic of loads such as induction motors is important [4]. Therefore, shaping DRD should be considered as a compulsory task in DSM planning. It is achieved via altering the percentage of different load categories under the total demand at different times during the day so that the shaped DRD looks as close as a defined "favourable" DRD. It can enhance more effective DSM, provide greater savings in power system planning, and protect power system from potential voltage stability issues.

### 6.2.1 Theory of Load Shifting

Based on day-ahead predicted total demand, day-ahead predicted load compositions and dynamic response of the demand (DRD), it is possible to make plans for load shifting in advance. In load shifting, the controllable load at peak times are shifted and properly allocated to off-peak times so that not only predicted load curve is flattened and total cost is reduced, but also DRDs after load shifting at both peak times and off-peak times remain fairly the same as DRDs before load shifting (or look as similar as a given DRD). The load shifting task in this study can be formulated as two optimisation problems as shown in (6.1) and (6.2)

$$Minimising \quad Obj = \sum_{i=1}^{N}\left(P_{2i} - \frac{1}{N}\sum_{i=1}^{N}P_{1i}\right)^2 \tag{6.1}$$

$$Minimising \quad Cost = \sum_{i=1}^{N} c_i P_{2i} \tag{6.2}$$

where $N$ is the total sampling number of the load curve (usually 48 for sampled or predicted load curves with 30min resolution and 24 for those with 60min resolution), $P_{2i}$ is the $i^{th}$ sampling of the load curve after load shifting and $P_{1i}$ is the $i^{th}$ sampling of the load curve before load shifting, *Obj* is the objective function, *Cost* is the total cost after load shifting, $c_i$ is the price of the electricity at the $i^{th}$ measured moment during the day. The three main constraints are:

$$\sum_{i=1}^{N}\left(P_{2i} - P_{1i}\right) = 0 \tag{6.3}$$

i.e., the total consumption before and after load shifting remains unchanged. For any $i$ during peak times before load shifting

$$P_{1i} - P_{2i} \le k_i P_{1i} \tag{6.4}$$

i.e., the demand shifted from the peak time should not exceed the total amount of the controllable load (i.e. $k_i$ is the percentage of controllable load in the total demand), and

for the total price change

$$\sum_{i=1}^{N} c_i \left( P_{2i} - P_{1i} \right) \leq 0 \tag{6.5}$$

i.e., the total cost after load shifting is less than or equal to the total cost before load shifting. Other constraints, not considered here, include the willingness of customers to have their loads shifted, DNO's ability to shift the loads, generation capacity limits, etc.

After the controllable loads at peak times are shifted and allocated to off-peak times, the component-based load modelling approach is used to predict dynamic responses of the demand after load shifting at both peak times and off-peak times. At the beginning, desired predict dynamic responses of the demand during peak times are defined so that the amount and type of controllable load to be shifted can be determined. Then, these shifted loads are allocated to off-peak times to satisfy corresponding defined dynamic responses of the demand at off-peak times. The process is repeated until all defined dynamic responses of the demand during both peak and off-peak times are satisfied.

### 6.2.2 Illustration of Load Shifting and Shaping of Dynamic Responses of the Demand

Considering the objective functions and constraints described in Section 6.2.1, the load curve after load shifting are determined as the red dotted line shown in Figure 6.5(a). The blue solid line is the predicted load curve before load shifting. The predicted peak demand (at 12:30) is reduced by 0.93MW. Figure 6.5(b) shows dynamic response of the demand (DRD) to a 0.7% voltage drop at 3:30 (lower demand) following the redistribution of the same amount (MW) but different types of devices available for shifting, i.e., resistive load only (R), induction motor only (IM),

half IM and half resistive load (R+IM). As the total demand at given hour before and after the shifting changes, for ease of comparison, Figure 6.5(b) uses the power change after the voltage drop to represent DRD. It demonstrates that different DRDs can be obtained based on the availability and redistribution of controllable devices and load shifting decisions should be made based on predicted DRD before and after load shifting as a potentially unfavourable DRD could endanger system stability.

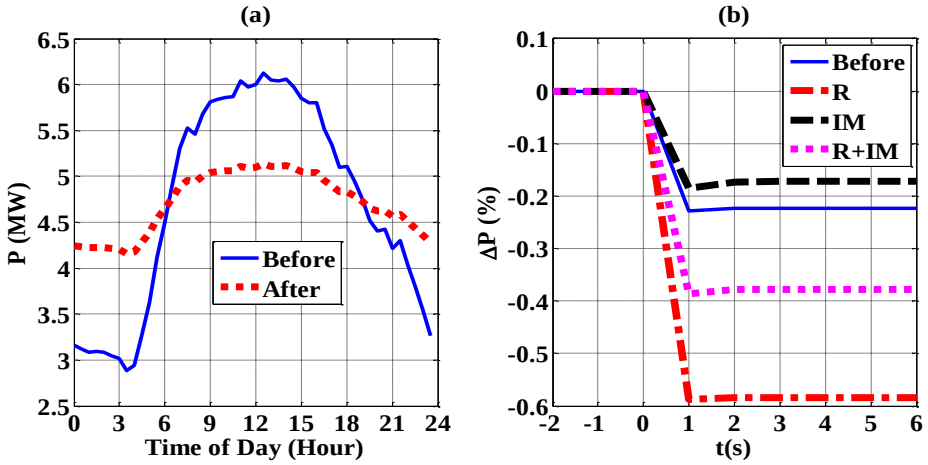


Figure 6.5. (a) Load curves before and after load shifting; (b) predicted DRD of P at 03:30 with different categories allocated but the same total demand shifted

## 6.3 Summary

In this chapter, total load forecasting, load disaggregation and load aggregation (i.e. the component-based load modelling approach) are integrated to form the overall approach to enable prediction as well as shaping of dynamic responses of the demand (DRD) without field measurements, and the measured data from real UK distribution networks are used to validate the overall approach. Artificial neural network, Monte Carlo simulation and optimisation techniques are used to forecast the total demand, convert the demand in MW and MVAR to per-unit value, disaggregate the total demand into different load categories, predict the decomposed daily loading curve for the next day, and finally predict and shape the dynamic response of the demand for the

next day. The overall approach does not require measured load signatures with high sampling rates (except in validation process) or results of customer surveys on consumptions. From the figures plotting the predicted/estimated dynamic response of the demand verses the measured one, it can be seen that the overall approach works very well.

Based on the predicted decomposed daily loading curves and dynamic responses of the demand (DRDs), shaping of DRDs can be achieved by load shifting, and it is found that different DRDs can be obtained if the contributions of different load categories to the total demand are different.

# 7 Conclusions and Future Work

## 7.1 Conclusions

This thesis has developed an approach for the prediction and shaping of dynamic responses of the demand to voltage disturbances at any given time of the day. It consists of the following main stages: total load forecasting, load disaggregation (i.e. including the sub-stage that transforms measurements in real values into per-unit values), estimation and prediction of dynamic response of the demand, and load shifting plans. Artificial neural network together with Monte Carlo Simulation are used to develop the approach. Several original contributions have been made in the thesis.

The thesis first presents a comprehensive overview of some past events and the effect of different dynamic load characteristics (i.e. load models) on power system stability studies. This overview concluded that the selection of load models can affect the result of the power system analysis. Afterwards, the most widely used load models, load modelling approaches (i.e. the measurement-based approach and the component-based approach) and their advantages and disadvantages are reviewed in detail.

In the majority of the short-term dispatch and control activities such as load shifting, the knowledge of the predicted dynamic load characteristic is important as it can help avoiding the potential system stability issues. However, no past research on the prediction of dynamic load characteristics has been found. As a result, a new approach should be developed to achieve this objective. As it is not possible to obtain recorded dynamic responses of the demand in the future from field tests, the dynamic load characteristics should be predicted without measurements and the computer simulation should be used. In order to facilitate the prediction of dynamic load characteristics, the percentage of different load categories (and their controllability if load shifting is required) should be estimated or predicted in advance.

### 7.1.1 Development of the Probabilistic Framework for Estimation/Prediction of Dynamic Responses of Demand

Based on the above considerations and the investigation of available consumption data of different load categories, a framework for estimation or prediction of dynamic responses of the demand without having to perform field tests is developed in this thesis (Chapter 3). Different from past work which usually uses deterministic decomposed daily loading curves and deterministic individual load models, in this thesis, both decomposed daily loading curves and individual load models are presented probabilistically. Therefore, dynamic responses of the demand are also presented probabilistically, with the range and the most probable responses indicated. It was confirmed that the share of different load categories can affect the dynamic load characteristics of the aggregate load. This strengthens the importance of the knowledge of dynamic response of the demand. It was also found that the most probable dynamic response of the aggregate load can be derived using the following two ways: (i) considering all possible responses using probabilistic analysis techniques, and (ii) considering the most probable responses of individual load categories and participation of individual categories. The results derived in the above

two ways are very close. Therefore, in the validation process of the overall approach using data from the real network (Chapter 6), the most probable load compositions and dynamic responses of the demand are predicted first; then the predicted most probable dynamic response of the demand is compared with the measured one. The proposed methodology contributes to the development of advanced control algorithms for active demand management as the decisions about demand control action can be based not only on the requirement to reduce consumption at different times of day (or to "match" demand and renewable generation output), but also on achieving desired dynamic responses of the demand (induced by appropriate price signals, for example) at given hours. This is the *first original contribution* of the thesis.

### 7.1.2 Comparison of Load Forecasting Approaches and Applications to both Real and Reactive Power Forecasting

Typically used approaches for load forecasting are reviewed in Chapter 1. The review concluded that advanced methodologies such as fuzzy logics and the artificial-intelligence (AI) based approach predict the demand more accurately than traditional approaches such as regression and stochastic time series. However, comparisons between different AI-based approaches (such as artificial neural network and adaptive neuro-based fuzzy inference system) used for load forecasting in the same operating environment have rarely been made in the past. As a result, the thesis compared the performance of two most widely used approach, artificial neural network (ANN) and adaptive neuro-based fuzzy inference system (ANFIS), used for load forecasting in the same operating environment (Chapter 4). The results indicate that the two methodologies can have comparable performance if the parameters of ANN and ANFIS are set properly. Therefore, either of them can be used for load forecasting. The artificial neural network (ANN) based approach is selected for total load forecasting in this research due to its stronger adaptation and its applicability in the case with multiple outputs. The ANN-based approach developed by EPRI is upgraded

and applied to both real and reactive power forecasting. The mean absolute percentage error of total demand prediction is about 4.27% for real power and 15.29% for reactive power. In addition, probabilistic characteristics of absolute percentage error for load forecasting are analysed. The accurate prediction of real and reactive power could lead to greater saving in power system planning and enhance the accuracy of load composition prediction. The comprehensive comparison of the two artificial-intelligence-based methodologies and the recommendation of the suitable parameter settings for either of them as well as the application of the modified ANN approach to both real and reactive power forecasting represent the *second original contribution* of the thesis.

### 7.1.3 Development of ANN-based Load Disaggregation Approach

A detailed review of past work on load disaggregation is presented in Chapter 1. The two most widely used methodologies in the past were identified as the intrusive load monitoring and the non-intrusive load monitoring, and they require either direct access to appliances (and large quantities of customer surveys) or load signatures of both aggregate loads and individual appliances with high sampling rates. These, however, are highly unlikely to be available in most cases. Therefore, in Chapter 5, an artificial neural network (ANN) based load disaggregation methodology is developed, where only the substation $P$, $Q$ and $V$ RMS measurements are needed to perform load disaggregation. The methodology identifies the percentages of different load categories in the total demand mix as well as their controllability with certain confidence levels considering different absolute weighting factor errors (AWFEs). The resulting load compositions can be used to estimate the real and reactive power characteristics using the component-based load modelling approach. The load disaggregation can be performed with high confidence levels for different requirements of absolute load model errors (ALMEs). Absolute load model errors (ALMEs) can be used as an index for accuracy (or reasonability) of the disaggregation

approach in reality as the actual data on load compositions are not always available. Generally, when all the inputs are available, the developed disaggregation methodology can provide estimations with an overall absolute weighting factor error (AWFE) less than 10% for the estimation of load compositions and less than 15% for controllable load estimation; and the absolute load model error (ALME) is less than 2.5% for $P$ and less than 5% for $Q$ with a confidence level (CL) over 80%. The developed ANN-based disaggregation approach can additionally provide close to real time estimation of different customer energy consumption at the bulk supply point without direct or indirect contact with the customers and the real time estimation of dynamic load characteristics at the bulk supply point without voltage disturbance experiments. It can also contribute to demand side management, more effective integration of the renewable generation and contribute to close to real time control of the network.

In addition to the development of the new load disaggregation approach, its robustness is also investigated and discussed. It is found that absence of voltage measurements does not significantly affect the result; this means that the disaggregation algorithm can be integrated with the total demand forecasting algorithm to predict the contribution of load compositions in the future. The absence of the reactive power data, however, significantly affects the result; thus, the reactive power measurement plays a significant role in the load disaggregation approach. There is no doubt that the disaggregation algorithm works the best when all inputs are available. The effect of the availability of prior information about controllable load on load disaggregation is also discussed. Knowing controllable load contribution in advance helps determine the range of the weighting factor used in the ANN training process and results in disaggregation performance with a higher level of accuracy or a higher confidence level under the same level of accuracy. The development of this ANN-based load disaggregation approach is the *third original contribution* of the

thesis.

### 7.1.4 Integration of Load Forecasting and Load Disaggregation to Enable Prediction of Load Compositions

The developed disaggregation approach can be integrated with the total demand forecasting approach to predict load composition in future. It is possible to predict the percentage of different load categories (i.e. induction motors, lighting etc.) in advance, with a load composition forecasting error typically less than 15% in the majority of the cases. It is also found that the confidence level of the total load forecasting accuracy can affect the confidence level of the accuracy of load composition forecasting. The integration of the total load forecasting approach and the developed load disaggregation approach to enable prediction of load compositions represents the *fourth original contribution* of the thesis.

### 7.1.5 Calculation of Rated Demand

Load disaggregation is typically developed in per-unit for convenience. If it is to be integrated with the total demand forecasting tool, whose outputs are predicted demand in MW and MVAR, it is necessary to convert the demand into per-unit values. This, however, may not be very straightforward as the demand continuously varies even though the voltage may remain constant (as it is regulated). A Monte-Carlo based approach is therefore developed in this thesis to obtain the rated demand at different times during the day and use the obtained rated demand to convert the demand given in MW and MVAR to per-unit values. The definition of the rated demand is given as follows: if the actual voltage at a bulk supply point is the same as its rated voltage, the demand of the bulk supply point at this moment is referred to as rated demand. The required data are measured voltage, real and reactive power at bulk supply point. When deriving the per-unit value of the one-day ahead predicted demand, the voltages at the corresponding times are assumed to be the same as on the day before. This

conversion provides an appropriate "base value" to transform the demand in MW and MVAR into a per-unit value so that the measured demand could be used for disaggregation. It further enables the validation of load disaggregation and dynamic load response prediction, as the demand collected at the bulk supply point is in MW and MVAR. Additionally, this helps significantly with load modelling in the case of loads with long-time recovery time constants and in the case of natural demand change as it helps distinguish between the load changes caused by voltage change and those induced by variations in customer behaviour. The methodology for calculating the rated demand is the *fifth original contribution* of the thesis.

### 7.1.6 Methodology for Prediction and Shaping of Dynamic Responses of Demand

Using the approaches developed for load disaggregation and the prediction of load compositions, together with the component-based approach, the dynamic response of the demand to a voltage disturbance can be predicted. Chapter 6 focusses on describing and validating the methodology for prediction of dynamic responses of the demand at a given point in time in the future using the data form real UK network. The validation process shows that the mean absolute percentage error (MAPE) of the prediction of dynamic responses of the demand is about 0.33% for real power, and 10.49% for reactive power. The smaller value of reactive power leads to higher MAPE of prediction. Based on the predicted load compositions and the controllable part of the demand, load shifting can be planned in advance. Past research on load shifting focussed almost exclusively on total cost minimisation and peak demand reduction. There was no mention at all that the change of the load composition caused by load shifting can lead to changes in dynamic responses of demand during the corresponding periods, which could eventually influence power system voltage and angular stability. There was no work at all on prediction and shaping of the dynamic response of the demand, either. The methodology developed in this thesis enables not

only estimation of dynamic responses of demand, but also modifying of dynamic load characteristics to assume desired shapes. A load shifting example is given in the thesis considering the prediction of load compositions, load controllability and dynamic responses of the demand. The methodology for prediction and shaping of dynamic responses of the demand is the *sixth original contribution* of the thesis.

### 7.1.7 Benefits of the Research

The prediction and shaping of dynamic responses of the demand without field measurements could enhance the effectiveness of demand side management application and reduce the risk of potential power system stability issues. The proposed approach will contribute to the development of advanced control algorithms for active demand management, effective renewable generation integration and real time electricity price regulation. It should be pointed out that this approach enables the decisions of demand control to be made, not only on requirements to balance demand & generation and reduce the total cost, but also on requirements to ensure the stable and secure operation of the whole power system (or at least minimise the risk of potential unfavourable dynamic interactions).

## 7.2 Future Work

The work presented in the thesis has achieved the aim and all objectives specified in Chapter 1. There are still a few areas, however, where future improvements could be made, and directions that could be followed in future research in the field.

***A. Improvement of the Confidence Level of Load Disaggregation***

The *first area* for future work can be trying to improve the confidence level of load disaggregation and in particular, identification of controllable load. It can be achieved by, exploring the effectiveness of other techniques such as support vector

machine (SVM), or by defining and conducting more detailed but less time-consuming customer surveys. One way of achieving this in the future would probably be by using the data provided by "smart meters", whose application is going to expand.

### *B. Application to Large Voltage Disturbance Cases (ΔV>20%)*

The *second area* for future work can be the application of the methodology to voltage disturbances larger than 20% both at bulk supply points and in different load sectors, which may involve load disconnection. Considering that the methodology in this thesis focuses on load responses at bulk supply points (typically mixed demand), the potential disconnection of individual component(s) following relatively large voltage disturbance may not significantly affect overall load response. The issue of load disconnection and subsequent change in demand composition and its dynamic response is more relevant to single class loads (e.g., industrial or commercial load), where disconnection (and subsequent restart) of variable speed drives or other large power electronics interfaced loads or energy efficient lighting may significantly affect dynamic responses of the demand. Due to its probabilistic nature, the methodology developed in this thesis can be extended to incorporate variation in both, demand composition and dynamic responses of individual load categories, resulting from disconnection of part of demand following large system disturbances. This, however, needs to be properly tested and validated.

### *C. Inclusion of Renewable Energy Technologies in Load Mix*

The *third area* for future work can be the incorporation of renewable devices in the methodology. Due to the requirements for energy saving and emission reduction, renewable devices which produce environmental-friendly energy are gaining more and more popularity and will become an important part in the future power system. Therefore, it will be beneficial to study the steady-state and dynamic load

characteristics of bulk supply points with significant participation of renewable generations. Although the approach used could be the same as that described in this thesis (at least initially), greater variability in demand and particularly faster responses could result in different requirements and challenges in predicting the dynamic response of the demand.

### *D. Development of Commercial Tools for Prediction of Load Compositions and Dynamic Responses of Demand*

*Last but not least*, the overall approach developed in this thesis can be upgraded to a commercial package with suitable user interfaces and additional functionalities to meet the requirements of both utilities and customers.

# References


[1] CIGRE WG C4.605, "Modelling and aggregation of loads in flexible power networks," (566), ISBN: 978-2-85873-261-6, February 2014.

[2] Dmitry N. Kosterev, C. W. Taylor, and W. A. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Transactions on Power Systems,* vol. 14, no. 3, pp. 967-979, 1999.

[3] Kao Wen-Shiow, "The effect of load models on unstable low-frequency oscillation damping in Taipower system experience w/wo power system stabilizers," *IEEE Transactions on Power Systems,* vol. 16, no. 3, pp. 463-472, 2001.

[4] Carson W. Taylor, Neal J. Balu, and Dominic Maratukulam, *Power system voltage stability*. New York: McGraw-Hill, 1994

[5] EPRI, "Advanced load modeling-entergy pilot study," 1011391, 2004.

[6] K.N.Hasan and J.V.Milanovic. Interim Profile Modelling Study [Online]. Available: http://www.enwl.co.uk/docs/default-source/class-documents/university-of-manchester-interim-report-wp1.pdf?sfvrsn=4

[7] William Koenigsberg, *Recursive least squares parameter estimation and associated computer programs*. Cambridge: Charles Stark Draper Laboratory, 1975

[8] Jin Ma, Zhao-yang Dong, Ren-mu He, and D. J. Hill, "Measurement-based Load Modeling using Genetic Algorithms," in *Proc*. *IEEE Congress on Evolutionary Computation* 2007, pp. 2909-2916.

[9] S. Z. Zhu, Z. Y. Dong, K. P. Wong, and Z. H. Wang, "Power system dynamic load identification and stability," in *Proc*. *International Conference on Power System Technology (PowerCon)*, Dec. 2000, vol. 1, pp. 13-18.

[10] H. Niska, "Extracting controllable heating loads from aggregated smart meter data using clustering and predictive modelling," in *Proc*. *IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, 2013, pp. 368-373.

[11] V. Knyazkin, C. A. Canizares, and L. H. Soder, "On the parameter estimation and modeling of aggregate power system loads," *IEEE Transactions on Power*

*Systems,* vol. 19, no. 2, pp. 1023-1031, 2004.

[12] EPRI, "Measurement-based load modeling," 2006.

[13] L. M. Korunovic, D. P. Stojanovic, and J. V. Milanovic, "Identification of static load characteristics based on measurements in medium-voltage distribution network," *IET Generation, Transmission & Distribution,* vol. 2, no. 2, pp. 227-234, 2008.

[14] Y. Li, H. D. Chiang, B. K. Choi, Y. T. Chen, D. H. Huang, and M. G. Lauby, "Representative static load models for transient stability analysis: development and examination," *IET Generation, Transmission & Distribution,* vol. 1, no. 3, pp. 422-431, 2007.

[15] C. J. Lin, A. Y. T. Chen, C. Y. Chiou, C. H. Huang, H. D. Chiang, J. C. Wang, and L. Fekih-Ahmed, "Dynamic load models in power systems using the measurement approach," *IEEE Transactions on Power Systems,* vol. 8, no. 1, pp. 309-315, 1993.

[16] H. Renmu, Jin Ma, and D. J. Hill, "Composite load modeling via measurement approach," *IEEE Transactions on Power Systems,* vol. 21, no. 2, pp. 663-672, 2006.

[17] Han Dong, Ma Jin, He Ren-mu, and Dong Zhao-yang, "A real application of measurement-based load modeling in large-scale power grids and its validation," *IEEE Transactions on Power Systems,* vol. 24, no. 4, pp. 1756-1764, 2009.

[18] Dobrivoje P. Stojanović, Lidija M. Korunović, and J. V. Milanović, "Dynamic load modelling based on measurements in medium voltage distribution network," *Electric Power Systems Research,* vol. 78, no. 2, pp. 228-238, 2008.

[19] L.M.Korunovic, S.Sterpu, S.Djokic, K.Yamashita, S.M.Villanueva, and J.V.Milanovic, "Processing of load parameters based on existing load models," in *Proc. 3rd IEEE International Conference On Innovative Smart Grid Technologies (ISGT Europe)*, 2012.

[20] Lidija M.Korunovic and Dobrivoje P.Stojanovic, "Dynamic load modelling of some low voltage devices," *Facta Universitatis (NIS), SER.: Elec. Energ.,* vol. 22, no. 1, pp. 61-70, 2009.

[21] P. Ju, F. Wu, Z. Y. Shao, X. P. Zhang, H. J. Fu, P. F. Zhang, N. Q. He, and J. D. Han, "Composite load models based on field measurements and their applications in dynamic analysis," *IET Generation, Transmission & Distribution,* vol. 1, no. 5, pp. 724-730, 2007.

[22] Fangtao Dai, "Load modeling for power system stability studies," Doctor of Philosophy, Electrical Engineering and Electronics, UMIST, 2002

[23] C. Concordia and S. Ihara, "Load Representation in Power System Stability Studies," *IEEE Transactions on Power Apparatus and Systems,* vol. PAS-101, no. 4, pp. 969-977, 1982.

[24] A. J. Collin, I. Hernando-Gil, J. L. Acosta, I. S. Ilie, and S. Z. Djokic, "Realising the potential of smart grids in LV networks. Part 1: demand-side management," in *Proc*. *Innovative Smart Grid Technologies (ISGT Europe)*, 2011, pp. 1-7.

[25] A. J. Collin, J. L. Acosta, I. Hernando-Gil, and S. Z. Djokic, "An 11 kV steady state residential aggregate load model. Part 2: Microgeneration and demand-side management," in *Proc*. *IEEE PowerTech*, Trondheim, 2011, pp. 1-8.

[26] C. Cresswell and S. Djokic, "Representation of directly connected and drive-controlled induction motors. Part 1: Single-phase load models," in *Proc*. *18th International Conference on Electrical Machines*, 2008, pp. 1-6.

[27] C. Cresswell and S. Djokic, "Representation of directly connected and drive-controlled induction motors. Part 2: Three-phase load models," in *Proc*. *18th International Conference on Electrical Machines*, 2008, pp. 1-6.

[28] C Cresswell and Sasa Djokic, "Steady-state models of low energy consumption light sources," in *Proc*. *16th PSCC*, 2008, pp. 1-7.

[29] C. Cresswell, S. Djokic, and S. Munshi, "Analytical modeling of adjustable speed drive load for power system studies," in *Proc*. *IEEE Power Tech*, Lausanne, 2007, pp. 1899-1904.

[30] Charles Cresswell, Sasa Djokic, K. Ochije, and Ewen Macpherson, "Modelling of Non-Linear Electronic Loads for Power System Studies: A Qualitative Approach," in *Proc*. *19th Int. Conf. on Electricity Distribution CIRED*, 2007, pp. 1-4.

[31] D. Kosterev, A. Meklin, J. Undrill, B. Lesieutre, W. Price, D. Chassin, R. Bravo, and S. Yang, "Load modeling in power system studies: WECC progress update," in *Proc*. *IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, 2008, pp. 1-8.

[32] A. J. Collin, I. Hernando-Gil, J. L. Acosta, and S. Z. Djokic, "An 11 kV steady state residential aggregate load model. Part 1: aggregation methodology," in *Proc*. *IEEE PowerTech*, Trondheim, 2011, pp. 1-8.

[33] Xiaodong Liang, Wilsun Xu, C. Y. Chung, W. Freitas, and Kun Xiong, "Dynamic Load Models for Industrial Facilities," *IEEE Transactions on Power Systems,* vol. 27, no. 1, pp. 69-80, 2012.

[34] B. Khodabakhchian and G. T. Vuong, "Modeling a mixed residential-

commercial load for simulations involving large disturbances," *IEEE Transactions on Power Systems,* vol. 12, no. 2, pp. 791-796, 1997.

[35] WECC. (14 March, 2012). Load Composition Data [Online]. Available: http://en.sourceforge.jp/projects/sfnet_lcm-data/downloads/loadcomposition-excel-1_7_8.zip/

[36] Intertek, "R66141 Household Electricity Survey_A study of domestic electrical product usage," 2012.

[37] Serena Hesmondhalgh, "GB energy demand-2010 and 2025. Initial brattle electricity demand-side model-scope for demand reduction and flexible response," 2012.

[38] Rainer Stamminger and Rheinische Friedrich-Wilhelms, "Synergy potential of smart appliances," University of Bonn, Bonn, Germany, Nov. 2008.

[39] Lars Dittmar, "Intergration of energy saving technologies and learning curves into the basic model of regional electricity liberalisation," Diplom-Wirtschaftsingenieur, Dept. Elect. Eng., University of Flensburg, Germany, 2006

[40] Richard E. Brown and Jonathan G. Koomey, "Electricity use in California: past trends and present usage patterns," *Energy Policy,* vol. 31, no. 9, pp. 849-864, 2003.

[41] Jean Paul ZIMMERMANN, "End-use metering campaign in 400 households in Sweden assessment of the potential electricity savings," Ingenierie Energetique et Fluides, Sep. 2009.

[42] ADEME, CCE, and CRES, "End-use metering campaign in 400 households of the European Community-assessment of the potential electricity savings," Commisson of the European Communities, Jan. 2002.

[43] Liang Jian, S. Ng, G. Kendall, and J. Cheng, "Load Signature Study-Part I: Basic Concept, Structure, and Methodology," *IEEE Transactions on Power Delivery,* vol. 25, no. 2, pp. 551-560, 2010.

[44] Liang Jian, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, "Load Signature Study-Part II: Disaggregation Framework, Simulation, and Applications," *IEEE Transactions on Power Delivery,* vol. 25, no. 2, pp. 561-569, 2010.

[45] K. Suzuki, S. Inagaki, T. Suzuki, H. Nakamura, and K. Ito, "Nonintrusive appliance load monitoring based on integer programming," in *Proc*. *SICE Annual Conference*, 2008, pp. 2742-2747.

[46] Soon Lee and Jung-Wook Park, "Estimation of Electric Load Composition on a Utility Side," in *Proc*. *IEEE Industry Applications Conference*, 2006, vol. 3,

pp. 1425-1431.

[47] H. Yan, Z. Zabar, D. Czarkowski, L. Birenbaum, E. Levi, and J. Hajagos, "Experimental test of a load model in the presence of harmonics," *IEE Proceedings-Generation, Transmission and Distribution,* vol. 146, no. 2, pp. 186-192, 1999.

[48] J. Duan, D. Czarkowski, and Z. Zabar, "Neural network approach for estimation of load composition," in *Proc*. *International Symposium on Circuits and Systems*, 2004, vol. 5, pp. V-988-V-991.

[49] J.A.Fuentes, A.Gabaldon, E.Gomez, A.Molina, and F.Ruz, "Object oriented architecture of a load composition identification system at distribution level," in *Proc*. *14th PSCC*, Sevilla, 2002, pp. 1-7.

[50] Fan Li and Leon Walker, "A quadratic programme for MW demand composition analysis at national grid supply points," in *Proc*. *14th PSCC*, Sevilla, 2002, pp. 1-6.

[51] C. Reese and L. Hofmann, "Determination of load and generation composition in distribution grids," in *Proc*. *IEEE PES Power Systems Conference and Exposition (PSCE)*, 2009, pp. 1-7.

[52] D. R. Sagi, S. J. Ranade, and A. Ellis, "Evaluation of a load composition estimation method using synthetic data," in *Proc*. *37th Annual North American Power Symposium*, 2005, pp. 582-588.

[53] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE,* vol. 80, no. 12, pp. 1870-1891, 1992.

[54] Wilsun Xu and Ming Dong, "Tracking energy consumptions of home appliances using electrical signature data," in *Proc*. *Power and Energy Society General Meeting, 2012 IEEE*, 2012, pp. 1-5.

[55] Marisa Figueiredo, Ana de Almeida, and Bernardete Ribeiro, "Home electrical signal disaggregation for non-intrusive load monitoring (NILM) systems," *Neurocomputing,* vol. 96, no. 0, pp. 66-73, 2012.

[56] Leslie K. Norford and Steven B. Leeb, "Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms," *Energy and Buildings,* vol. 24, no. 1, pp. 51-64, 1996.

[57] Yi Du, Liang Du, Bin Lu, R. Harley, and T. Habetler, "A review of identification and monitoring methods for electric loads in commercial and residential buildings," in *Proc*. *Energy Conversion Congress and Exposition (ECCE), 2010 IEEE*, 2010, pp. 4527-4533.

[58] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review

and outlook," *IEEE Transactions on Consumer Electronics,* vol. 57, no. 1, pp. 76-84, 2011.

[59] Hesham K. Alfares and Mohammad Nazeeruddin, "Electric load forecasting: Literature survey and classification of methods," *International Journal of Systems Science,* vol. 33, no. 1, pp. 23-34, 2002.

[60] W. Charytoniuk, M. S. Chen, and P. Van Olinda, "Nonparametric regression based short-term load forecasting," *IEEE Transactions on Power Systems,* vol. 13, no. 3, pp. 725-730, 1998.

[61] Ahmed Z. Al-Garni, Yaagoub N. Al-Nassar, Syed M. Zubair, and Abdallah Al-Shehri, "Model for Electric Energy Consumption in Eastern Saudi Arabia," *Energy Sources,* vol. 19, no. 4, pp. 325-334, 1997.

[62] I. S. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *IEEE Transactions on Power Systems,* vol. 4, no. 4, pp. 1484-1491, 1989.

[63] A. A. El-Keib, X. Ma, and H. Ma, "Advancement of statistical based modeling techniques for short-term load forecasting," *Electric Power Systems Research,* vol. 35, no. 1, pp. 51-58, 1995.

[64] G. A. N. Mbamalu and M. E. El-Hawary, "Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation," *IEEE Transactions on Power Systems,* vol. 8, no. 1, pp. 343-348, 1993.

[65] Q. C. Lu, W. M. Grady, M. M. Crawford, and G. M. Anderson, "An adaptive nonlinear predictor with orthogonal escalator structure for short-term load forecasting," *IEEE Transactions on Power Systems,* vol. 4, no. 1, pp. 158-164, 1989.

[66] J. H. Park, Y. M. Park, and K. Y. Lee, "Composite modeling for adaptive short-term load forecasting," *IEEE Transactions on Power Systems,* vol. 6, no. 2, pp. 450-457, 1991.

[67] E. H. Barakat, J. M. Al-Qassim, and S. A. Al-Rashed, "New model for peak demand forecasting applied to highly complex load characteristics of a fast developing area," *IEE Proceedings C: Generation, Transmission and Distribution,* vol. 139, no. 2, pp. 136-140, 1992.

[68] E. H. Barakat, M. A. Qayyum, M. N. Hamed, and S. A. Al-Rashed, "Short-term peak demand forecasting in fast developing utility with inherit dynamic load characteristics," *IEEE Transactions on Power Systems,* vol. 5, no. 3, pp. 813-824, 1990.

[69] S. R. Huang, "Short-term load forecasting using threshold autoregressive

models," *IEE Proceedings-Generation, Transmission and Distribution,* vol. 144, no. 5, pp. 477-481, 1997.

[70] Jiann-Fuh Chen, Wei-Ming Wang, and Chao-Ming Huang, "Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting," *Electric Power Systems Research,* vol. 34, no. 3, pp. 187-196, 1995.

[71] X. Ma, A. A. El-Keib, R. E. Smith, and H. Ma, "A genetic algorithm based approach to thermal unit commitment of electric power systems," *Electric Power Systems Research,* vol. 34, no. 1, pp. 29-36, 1995.

[72] Mo-Yuen Chow and H. Tram, "Application of fuzzy logic technology for spatial load forecasting," in *Proc. IEEE Transmission and Distribution Conference Proceedings*, 1996, pp. 608-614.

[73] T. Senjyu, S. Higa, and K. Uezato, "Future load curve shaping based on similarity using fuzzy logic approach," *IEE Proceedings Generation, Transmission and Distribution,* vol. 145, no. 4, pp. 375-380, 1998.

[74] H. Mori, Y. Sone, D. Moridera, and T. Kondo, "Fuzzy inference models for short-term load forecasting with tabu search," in *Proc. IEEE International Conference on Systems, Man, and Cybernetics*, 1999, vol. 6, pp. 551-556.

[75] A. Al-Anbuky, S. Bataineh, and S. Al-Aqtash, "Power demand prediction using fuzzy logic," *Control Engineering Practice,* vol. 3, no. 9, pp. 1291-1298, 1995.

[76] D. Srinivasan, A. C. Liew, and J. S. P. Chen, "A novel approach to electrical load forecasting based on a neural network," in *Proc. IEEE International Joint Conference on Neural Networks*, 1991, vol. 2, pp. 1172-1177.

[77] T. M. Peng, N. F. Hubele, and G. G. Karady, "An adaptive neural network approach to one-week ahead load forecasting," *IEEE Transactions on Power Systems,* vol. 8, no. 3, pp. 1195-1203, 1993.

[78] Ku-Long Ho, Yuan-Yih Hsu, and Chien-Chuen Yang, "Short term load forecasting using a multilayer neural network with an adaptive learning algorithm," *IEEE Transactions on Power Systems,* vol. 7, no. 1, pp. 141-149, 1992.

[79] D. C. Park, M. A. El-Sharkawi, R. J. Marks, II, L. E. Atlas, and M. J. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Transactions on Power Systems,* vol. 6, no. 2, pp. 442-449, 1991.

[80] A. Oonsivilai and M. E. El-Hawary, "Wavelet neural network based short term load forecasting of electric power system commercial load," in *Proc. IEEE Canadian Conference on Electrical and Computer Engineering*, 1999, vol. 3,

pp. 1223-1228.

[81] S. Rahman and O. Hazim, "A generalized knowledge-based short-term load-forecasting technique," *IEEE Transactions on Power Systems,* vol. 8, no. 2, pp. 508-514, 1993.

[82] S. Rahman and O. Hazim, "Load forecasting for multiple sites development of an expert system-based technique," *Electric Power Systems Research,* vol. 39, pp. 161-169, 1996.

[83] R. E. Brown, A. P. Hanson, and D. L. Hagan, "Long range spatial load forecasting using non-uniform areas," in *Proc. IEEE Transmission and Distribution Conference* 1999, vol. 1, pp. 369-373.

[84] Robert L. Sullivan, *Power system planning*. United State: McGraw-Hill, Inc., 1997

[85] K. Kouzelis, B. Bak-Jensen, P. Mahat, and J. R. Pillai, "A simplified short term load forecasting method based on sequential patterns," in *Proc. IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2014, pp. 1-5.

[86] Mathworks, "MATLAB Fuzzy Logic Toolbox User's Guide," 2013.

[87] M. Mourad, B. Bouzid, and B. Mohamed, "A hybrid wavelet transform and ANFIS model for short term electric load prediction," in *Proc. 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)*, 2012, pp. 292-295.

[88] A. Khotanzad, R. Afkhami-Rohani, T. L. Lu, A. Abaye, M. Davis, and D. J. Maratukulam, "ANNSTLF-a neural-network-based electric load forecasting system," *IEEE Transactions on Neural Networks,* vol. 8, no. 4, pp. 835-846, 1997.

[89] A. Khotanzad, R. Afkhami-Rohani, and D. Maratukulam, "ANNSTLF-Artificial Neural Network Short-Term Load Forecaster generation three," *IEEE Transactions on Power Systems,* vol. 13, no. 4, pp. 1413-1422, 1998.

[90] EPRI, "Enhancements to ANNSTLF, EPRI's short term load forecaster," TR-109482 1997.

[91] J. W. Taylor and R. Buizza, "Neural network load forecasting with weather ensemble predictions," *IEEE Transactions on Power Systems,* vol. 17, no. 3, pp. 626-632, 2002.

[92] J. N. Fidalgo and J. A. Pecas Lopes, "Forecasting active and reactive power at substations' transformers," in *Proc. IEEE PES PowerTech*, Bologna, 2003, vol. 1, pp. 1-6.

[93] Aditya Kumar Bhatt, Priyanka Solanki, Aditi Bhatt, and Ravindranath Cherukuri, "A fast and efficient back propagation algorithm to forecast active and reactive power drawn by various capacity Induction Motors," in *Proc. International Conference on Circuits, Power and Computing Technologies (ICCPCT)*, 2013, pp. 553-557.

[94] J. S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man and Cybernetics,* vol. 23, no. 3, pp. 665-685, 1993.

[95] A. Ghanbari, S. F. Ghaderi, and M. Ali Azadeh, "Adaptive Neuro-Fuzzy Inference System vs. Regression based approaches for annual electricity load forecasting," in *Proc. 2nd International Conference on Computer and Automation Engineering (ICCAE)*, 2010, vol. 5, pp. 26-30.

[96] Zohreh Souzanchi-K, Hadi Fanaee-T, Mahdi Yaghoubi, and Mohammad-R Akbarzadeh-T, "A multi adaptive neuro fuzzy inference system for short term load forecasting by using previous day features," in *Proc. International Conference on Electronics and Information Engineering*, 2010.

[97] O. E. Dragomir, F. Dragomir, R. Gouriveau, and E. Minca, "Medium term load forecasting using ANFIS predictor," in *Proc. 18th Mediterranean Conference on Control & Automation (MED)*, 2010, pp. 551-556.

[98] H. Lee Willis and J. E. D. Northcote-Green, "Comparison Tests of Fourteen Distribution load Forecasting Methods," *IEEE Transactions on Power Apparatus and Systems,* vol. PAS-103, no. 6, pp. 1190-1197, 1984.

[99] H. Lee Willis, *Spatial electric load forecasting*, 2nd ed. New York: Marcel Dekker, 2002

[100] K. Liu, S. Subbarayan, R. R. Shoults, M. T. Manry, C. Kwan, F. L. Lewis, and J. Naccarino, "Comparison of very short-term load forecasting techniques," *IEEE Transactions on Power Systems,* vol. 11, no. 2, pp. 877-882, 1996.

[101] P. K. Dash, A. C. Liew, S. Rahman, and S. Dash, "Fuzzy and neuro-fuzzy computing models for electric load forecasting," *Engineering Applications of Artificial Intelligence,* vol. 8, no. 4, pp. 423-433, 1995.

[102] P. K. Dash, A. C. Liew, and S. Rahman, "Fuzzy neural network and fuzzy expert system for load forecasting," *IEE Proceedings-Generation, Transmission and Distribution,* vol. 143, no. 1, pp. 106-114, 1996.

[103] S. Kuusisto, M. Lehtokangas, J. Saarinen, and K. Kaski, "Short term electric load forecasting using a neural network with fuzzy hidden neurons," *Neural Computing & Applications,* vol. 6, no. 1, pp. 42-56, 1997.

[104] Ruey-Hsun Liang and Ching-Chi Cheng, "Short-term load forecasting by a

neuro-fuzzy based approach," *International Journal of Electrical Power & Energy Systems,* vol. 24, no. 2, pp. 103-111, 2002.

[105] D. Srinivasan and M. A. Lee, "Survey of hybrid fuzzy neural approaches to electric load forecasting," in *Proc*. *IEEE International Conference on Systems, Man and Cybernetics*, 1995, vol. 5, pp. 4004-4008.

[106] Suhartono, I. Puspitasari, M. S. Akbar, and M. H. Lee, "Two-level seasonal model based on hybrid ARIMA-ANFIS for forecasting short-term electricity load in Indonesia," in *Proc*. *International Conference on Statistics in Science, Business, and Engineering (ICSSBE)*, 2012, pp. 1-5.

[107] T. Logenthiran, D. Srinivasan, and Shun Tan Zong, "Demand Side Management in Smart Grid Using Heuristic Optimization," *IEEE Transactions on Smart Grid,* vol. 3, no. 3, pp. 1244-1252, 2012.

[108] A. J. Van Staden, J. Zhang, and X. Xia, "A model predictive control strategy for load shifting in a water pumping scheme with maximum demand charges," in *Proc*. *IEEE PowerTech*, Bucharest, 2009, pp. 1-7.

[109] M. C. Vlot, J. D. Knigge, and J. G. Slootweg, "Economical Regulation Power Through Load Shifting With Smart Energy Appliances," *IEEE Transactions on Smart Grid,* vol. 4, no. 3, pp. 1705-1712, 2013.

[110] I. L. Sarioglu, O. P. Klein, H. Schroder, and F. Kucukay, "Energy Management for Fuel-Cell Hybrid Vehicles Based on Specific Fuel Consumption Due to Load Shifting," *IEEE Transactions on Intelligent Transportation Systems,* vol. 13, no. 4, pp. 1772-1781, 2012.

[111] H. J. Belitz, S. Winter, and C. Rehtanz, "Load shifting of the households in the E-Energy project E-DeMa," in *Proc*. *IEEE PES PowerTech*, Grenoble, 2013, pp. 1-6.

[112] A. G. Paetz, T. Kaschub, P. Jochem, and W. Fichtner, "Load-shifting potentials in households including electric mobility - A comparison of user behaviour with modelling results," in *Proc*. *10th International Conference on the European Energy Market (EEM)*, 2013, pp. 1-7.

[113] Bingying Wu, Buhan Zhang, Biao Mao, and Jiajun Zhai, "Optimal capacity of flow battery and economic dispatch used in peak load shifting," in *Proc*. *4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, 2011, pp. 1395-1400.

[114] National Grid, "Frequency response obligations-statutory, code and operational standards," 2002.

[115] J.V.Milanović, Koji Yamashita, S. Martinez Villanueva, S.Ž.Djokić, and L.M.Korunović, "International industry practice on power system load

modelling," *IEEE Transactions on Power Systems,* vol. 28, no. 3, pp. 3038-3046, 2013.

[116] Jan Machowski, Janusz W. Bialek, and J. R. Bumby, *Power system dynamics and stability*. Chichester ; New York: John Wiley, 1997

[117] P. Kundur, Neal J. Balu, and Mark G. Lauby, *Power system stability and control*. New York: McGraw-Hill, 1994

[118] Chongru Liu, "Power system stability," Beijing: North China Electric Power University, 2010.

[119] Paul C. Krause, Oleg Wasynczuk, and Scott D. Sudhoff, *Analysis of electric machinery and drive systems*: Wiley-IEEE Press, 2002

[120] T. Y. J. Lem and R. T. H. Alden, "Comparison of experimental and aggregate induction motor responses," *IEEE Transactions on Power Systems,* vol. 9, no. 4, pp. 1895-1900, 1994.

[121] K.Yamashita, S.M.Villanuva, and J.V.Milanovic, "Initial Results of International Survey on Industrial Practice on Power System Load Modelling Conducted by CIGRE WG C4.605," in *Proc*. *CIGRE Symp.*, Bologna, Italy, 2011, vol. C4-333.

[122] F. S. Chassin, E. T. Mayhorn, M. A. Elizondo, and Lu Shuai, "Load modeling and calibration techniques for power system studies," in *Proc*. *North American Power Symposium (NAPS)*, 2011, pp. 1-7.

[123] L. Du, D. HE, Y. Yang, J. A. Restrepo, B. Lu, R. G. Harley, and T. G. Habetler, "Self-organizing classification and identification of miscellaneous electric loads," in *Proc*. *IEEE PES General Meeting*, San Diego, CA USA, 2012, pp. 1-6.

[124] Pouyan Pourbeik and Anish Gaikwad, "A summary of EPRI load modeling efforts," Electric Power Research Institute, Knoxcille, USA 2012.

[125] Ju Sheng, Shiqian Xie, and Chengyi Pan, *Probability theory and mathematical statistics*, 4th ed. Beijing: Higher Education Press, 2008

[126] Mustafa Tahsin Yilmaz, "Comparison of effectiveness of adaptive neuro-fuzzy inference system and artificial neural networks for estimation of linear creep and recovery properties of model meat emulsions," *Journal of Texture Studies,* vol. 43, no. 5, pp. 384-399, 2012.

[127] Howard Demuth, Mark Beale, and Martin Hagan, "MATLAB Neural network tookbox user's guide," 2009.

[128] Kevin Gurney, *An introduction to neural network*: University College London Press, 1997

[129] Shuxiang Xu and Ling Chen, "A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining," in *Proc*. *5th ICITA*, 2008, pp. 683-686.

[130] Yinyin Liu, Janusz A. Starzyk, and Zhen Zhu, "Optimizing number of hidden neurons in neural network," in *Proc*. *25th IASTED International Multi-Conference: Artificial Intelligence and Applications*, 2007, pp. 121-126.

[131] I Ileană, C Rotar, and A Incze, "The optimization of feed forward neural netowrks structure using genetic algorithms," in *Proc*. *ICTAMI*, Thessalonili, Greece, 2004, pp. 223-230.

[132] Weather History @ Weather.org [Online]. Available: http://weather.org/weatherorg_records_and_averages.htm

[133] D. Karlsson and D. J. Hill, "Modelling and identification of nonlinear dynamic loads in power systems," *IEEE Transactions on Power Systems,* vol. 9, no. 1, pp. 157-166, 1994.

[134] System Dynamic Performance Subcommittee and Power System Engineering Committee, "Standard load models for power flow and dynamic performance simulation," *IEEE Transactions on Power Systems,* vol. 10, no. 3, pp. 1302-1313, 1995.

[135] Rober T. Clemen and Robert L. Winkler, "Combining probability distributions from experts in risk analysis," *Risk Analysis,* vol. 19, no. 2, pp. 187-203, 1999.

# Appendix A Moving Average Filtering

Moving average is a widely used technique for filtering as the noise in the measured data is generally the Gaussian noise. In the moving average (MA) filtering technique, an individual sample is replaced with the average value of the neighbouring data points over a specific data span. Data span is the design parameter of the moving average filter. If the data span is small (i.e. only a few samples are considered), the filtered signal looks similar to the original measured data; if it is large and contains adequate samples, the filtered signal looks smoother than the original one.

The moving average filtering technique can be represented mathematically as

$$y_s(i) = \frac{1}{2N+1} \sum_{k=-N}^{N} y(i+k) \tag{A.1}$$

where $y_s(i)$ is the value averaged for the $i^{th}$ sample, $y(i+k)$ is the $k^{th}$ sample on either side of the $i^{th}$ sample, $N$ is the total number of the samples on either side of the $i^{th}$ sample, and $2N+1$ is the data span over which the values are averaged.

# Appendix B Decomposed Daily Loading Curves for Different Load Sectors in Different Regions

In Chapter 1, a decomposed daily loading curve (DDLC) in the UK residential load sector in winter obtained from the customer survey and one in the commercial load sector of California in summer obtained from measured demand data of different end-users are given as Figure 1.4 and Figure 1.5 for illustration. This appendix provides DDLCs for different load sectors at different locations in different seasons.

## B.1 Residential Load Sector

In Section B.1, the decomposed daily loading curves (DDLC) of the residential load sector in different areas and different seasons are shown.

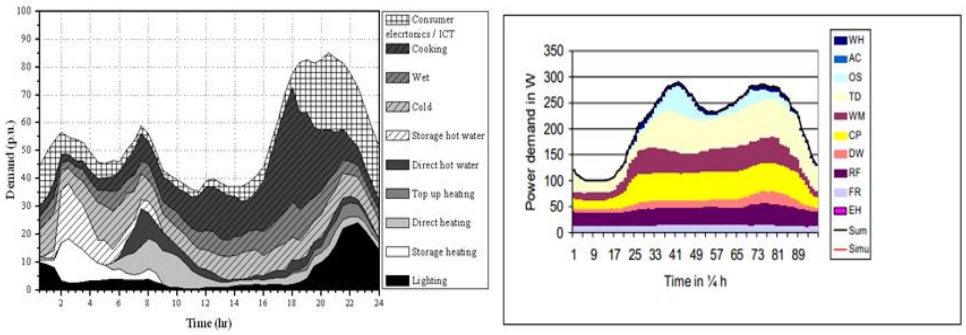


Figure B.1. DDLC of the residential load sector during the winter (left, adopted from [1]) and the summer (right, adopted from [38]) in the UK

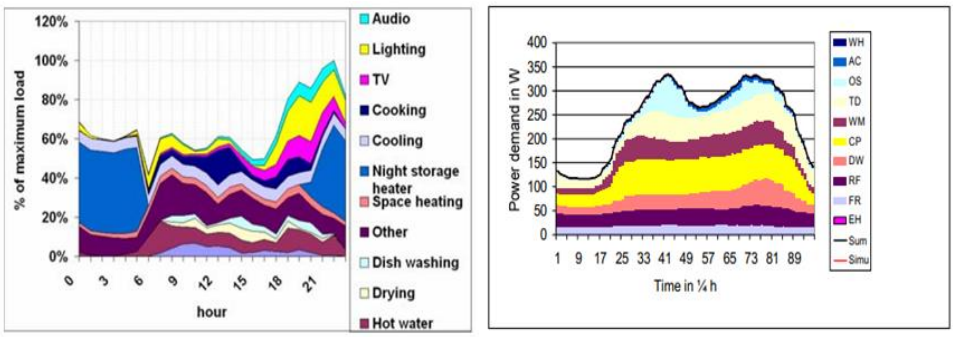


Figure B.2. DDLC of the residential load sector during the winter (left, adopted from [39]) and the summer (right, adopted from [38]) in Germany

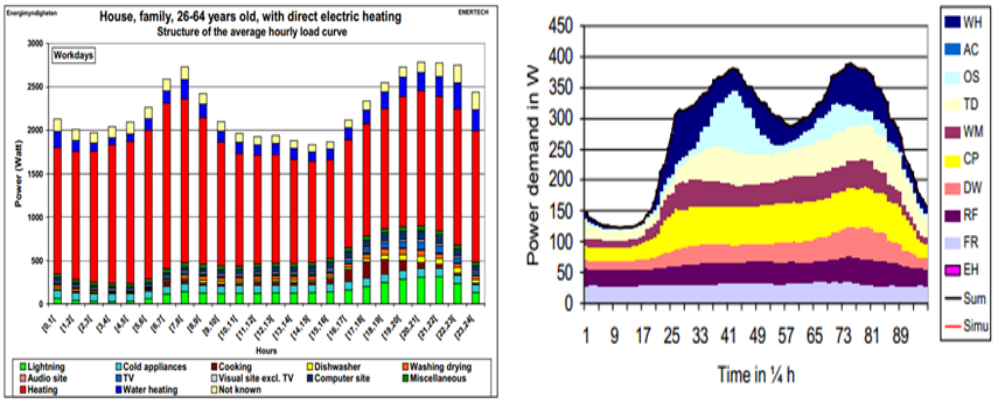


Figure B.3. DDLC of the residential load sector during the winter (left, adopted from [41]) and the summer (right, adopted from [38]) in Sweden

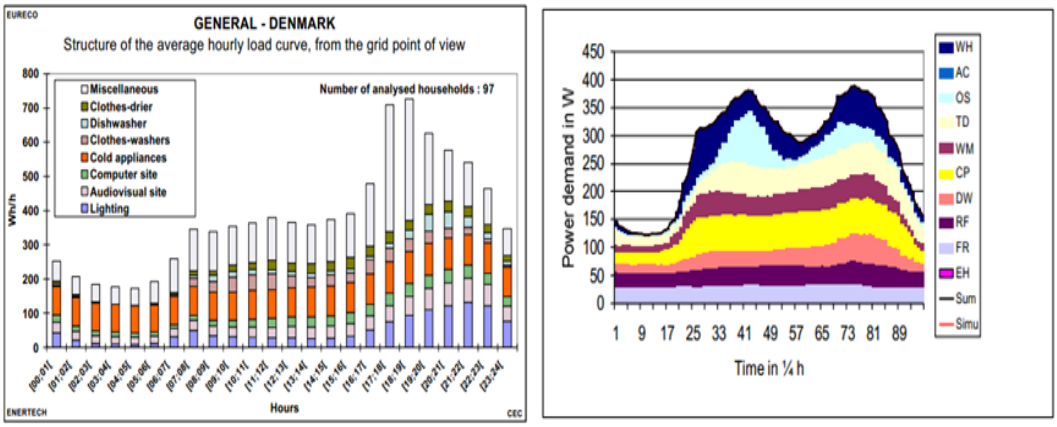


Figure B.4. DDLC of the residential load sector during the winter (left, adopted from [42]) and summer (right, adopted from [38]) in Denmark

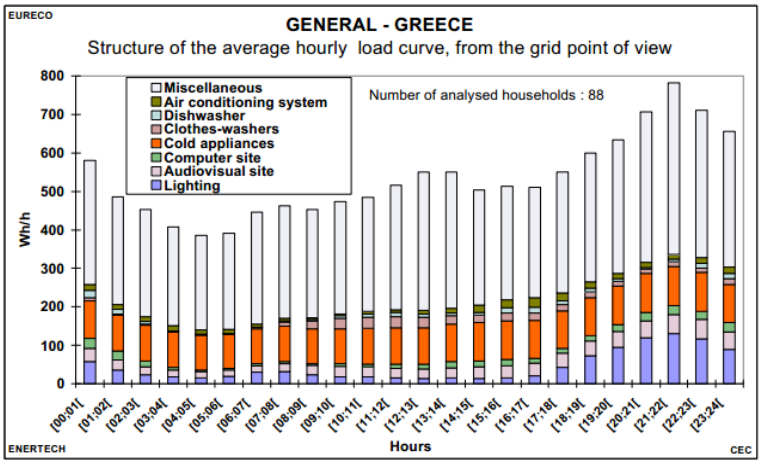


Figure B.5. DDLC of the residential load sector during the winter in Greece (adopted from [42])

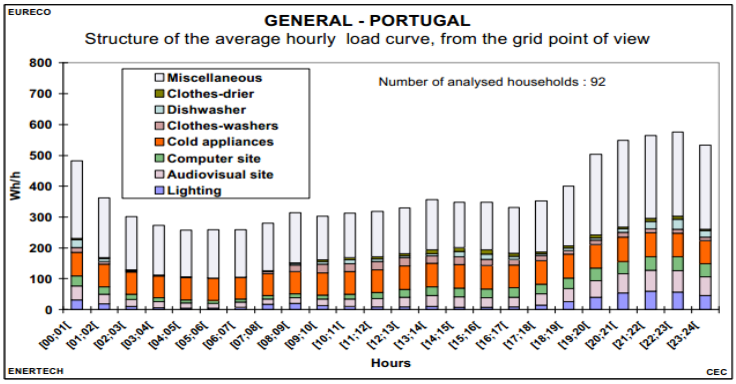


Figure B.6. DDLC of the residential load sector during the winter in Portugal (adopted from [42])

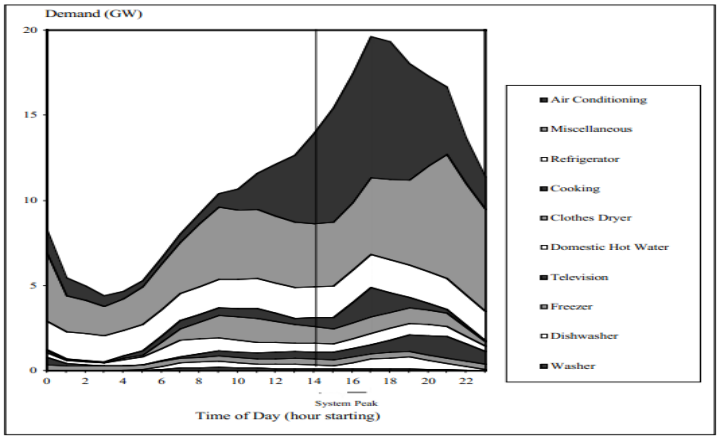


Figure B.7. DDLC of the residential load sector during the summer in California (adopted from [40])

## B.2 Commercial Load Sector

In Section B.2, the decomposed daily loading curves (DDLC) of the commercial load sector in different areas and different seasons are shown.

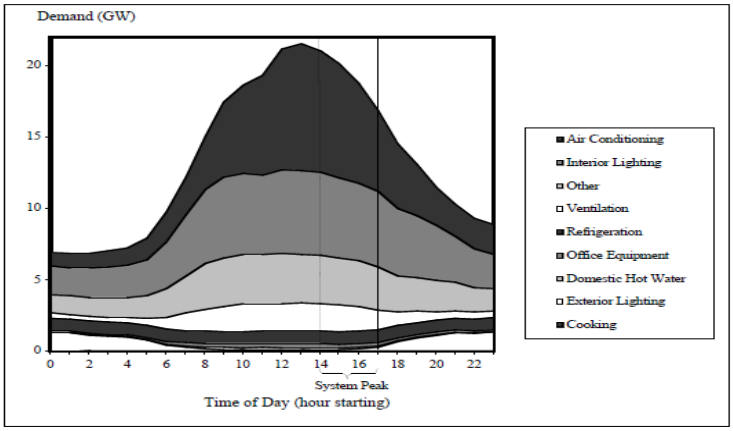


Figure B.8. DDLC of commercial load sector during the summer in California (adopted from [40])

## B.3 Industrial Load Sector

In Section B.3, the decomposed daily loading curves (DDLC) of the industrial load sector in different areas and different seasons are shown.

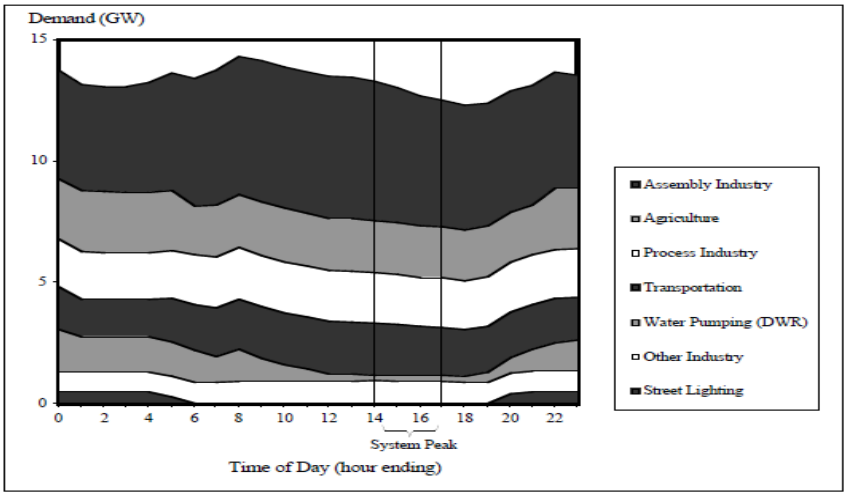


Figure B.9. DDLC of commercial load sector during the summer in California (adopted from [40])

## B.4 Aggregate Load

In Section B.4, the decomposed daily loading curves (DDLC) of the aggregate load in different areas and different seasons are shown.

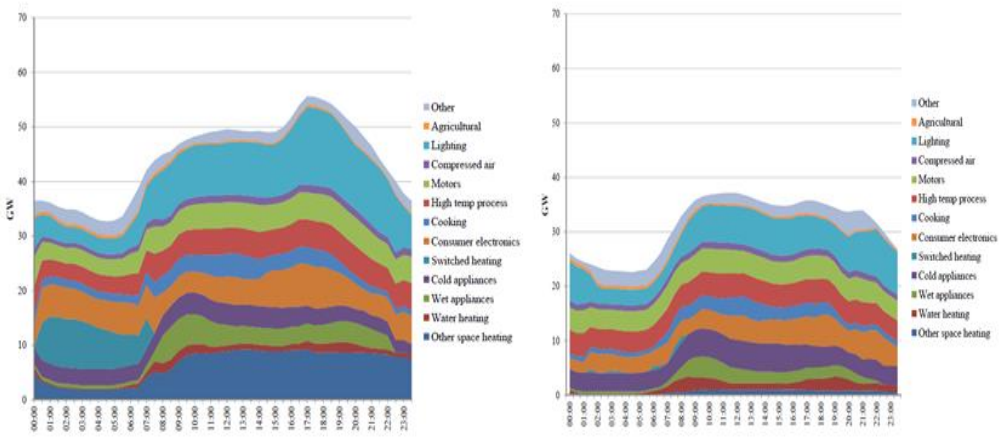


Figure B.10. DDLC of the aggregate load during the winter (left) and the summer (right) in the UK (adopted from [37])

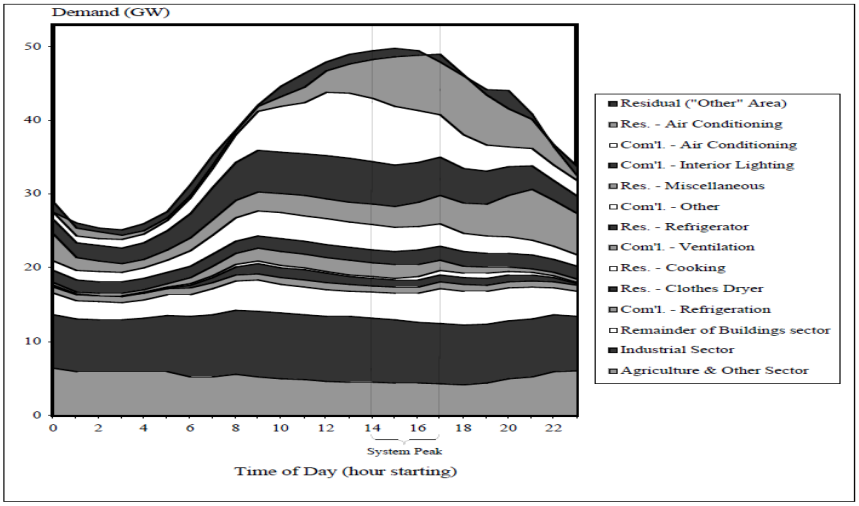


Figure B.11. DDLC of the aggregate load during the summer in California (adopted from [40])

# Appendix C Steady-state Load Characteristics of Induction Motors

## C.1 Swing Equation of the Rotor of Induction Motor

According to [116, 117], the swing equation of the rotor of induction motor is given by

$$J\frac{d\omega}{dt} = M_e - M \tag{C.1}$$

where $J$ is the inertia, $\omega$ is the angular speed of the rotor, $M_e$ and $M$ are the electromagnetic torque and the mechanical load torque, respectively. All the quantities are in per-unit except $J$ and $t$. Substituting (2.5) into (C.1), the swing equation in terms of slip can be written as

$$J\frac{ds}{dt} = J\frac{d(1-\omega)}{dt} = M - M_e \tag{C.2}$$

where all the quantities are in per-unit except $J$ and $t$.

## C.2 Electromagnetic Torque and Mechanical Load Torque

The electromagnetic torque is given as (C.3) [118, 119]

$$M_e = \frac{1}{\omega_s}\frac{V^2}{\left(R_s + \frac{R_r}{s}\right)^2 + \left(X_{\gamma s} + X_{\gamma r}\right)^2}\left(R_s + \frac{R_r}{s}\right) \tag{C.3}$$

and if written in p.u. with $\omega_s$=1, (C.3) becomes identical to (2.6). Since $s$ is generally small, $R_r/s >> R_s$ and the effect of $R_s$ could be neglected. For simplicity, the electromagnetic torque can be therefore approximated as

$$M_e = \frac{1}{\omega_s}\frac{V^2}{\left(R_s + \frac{R_r}{s}\right)^2 + \left(X_{\gamma s} + X_{\gamma r}\right)^2}\left(R_s + \frac{R_r}{s}\right) \approx \frac{sV^2R_r}{R_r^2 + \left(X_{\gamma s} + X_{\gamma r}\right)^2 s^2} \tag{C.4}$$

The mechanical load torque can be represented by

$$M = k\left[\tau_\alpha + (1-\tau_\alpha)(1-s)^{e_\beta}\right] \tag{C.5}$$

where $k$ is the loading rate (defined as the ratio of the actual load and the rated load), $\tau_\alpha$ is the static resisting torque, $e_\beta$ is an exponents relevant to $\omega$. Generally, $e_\beta$=2 [1, 116-118]. Similar to $e_\beta$, other quantities also have typical values for general calculation; they are summarised from [116-118] and shown in Table C.1, where all values are in p.u.

Table C.1. Typical Values of Variables Given in Figure 2.2

| Quantities | $R_s$ | $X_{\gamma s}$ | $R_r$ | $X_{\gamma r}$ | $X_m$ | $k$ | $\tau_\alpha$ | $e_\beta$ |
|---|---|---|---|---|---|---|---|---|
| Value (p.u.) | 0.0465 | 0.295 | 0.02 | 0.12 | 3.5 | 0.56 | 0.15 | 2 |

## C.3 Steady-state Operation

Generally, the induction motor operates under steady state condition. When a voltage drop occurs, the slip will change following the disturbance and reach a new steady state after a transient process. As $s$ is dependent on $V$, in order to derive the $P$-$V$ relationship, $s$-$V$ relationship should be studied first.

If the transient process of the slip is neglected, in the steady state both before and after the voltage disturbance, the slip $s$ should be constants, satisfying $\frac{ds}{dt} = 0$. Therefore, according to (C.2), the electromagnetic torque $M_e$ is equal to the mechanical load torque $M$ in the steady states. Substituting $M_e$ and $M$ by (C.4) and (C.5), the relationship between the slip and the voltage is derived as

$$\frac{sV^2R}{\left(R_r\right)^2 + \left(X_{\gamma s} + X_{\gamma r}\right)^2 s^2} = k\left[\tau_\alpha + (1-\tau_\alpha)(1-s)^{e_\beta}\right] \tag{C.6}$$

As it is complicated to represent $s$ in terms of $V$ and other parameters symbolically in this case, different values are assigned to $V$ to analyse the relationship between $s$ and $V$. If the typical values shown in Table C.1 are used, for any given

voltage $V$, the slip $s$ could be easily derived.

Figure C.1 shows $s$-$V$ relationship derived by MATLAB when voltage varies within the interval [0.8 p.u., 1.2 p.u.], which covers the majority of the operation cases in the reality. It can be observed that the slip increases when the voltage drops, and vice versa. The slip approaches 0 when the voltage approaches infinite, but it never reaches 0.

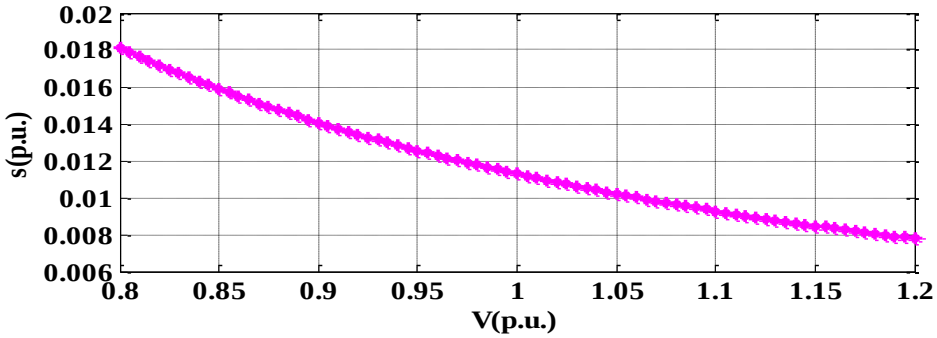


Figure C.1. $s$-$V$ relationship of induction motors with typical parameters

With the simulation result shown in Figure C.1, using Equation (2.6) and (2.7), $P$-$V$ and $Q$-$V$ relationship for typical induction motors are shown in Figure C.2. It can be observed that neither $P$ nor $Q$ change significantly despite the considerable change in voltages.

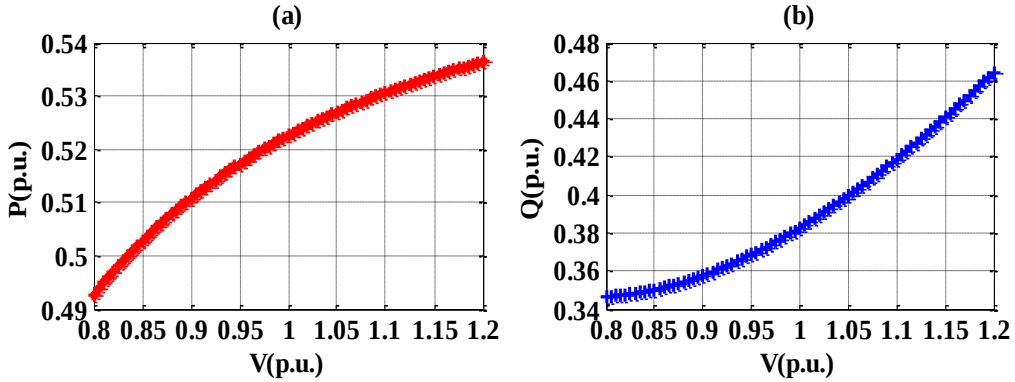


Figure C.2. (a) $P$-$V$, and (b) $Q$-$V$ relationship of induction motors with typical parameters

With the simulation result shown in Figure C.2, and using exponential load model given by (2.2), voltage exponents for both the real and reactive power as a function of the voltage can be calculated and plotted, shown as Figure C.3.

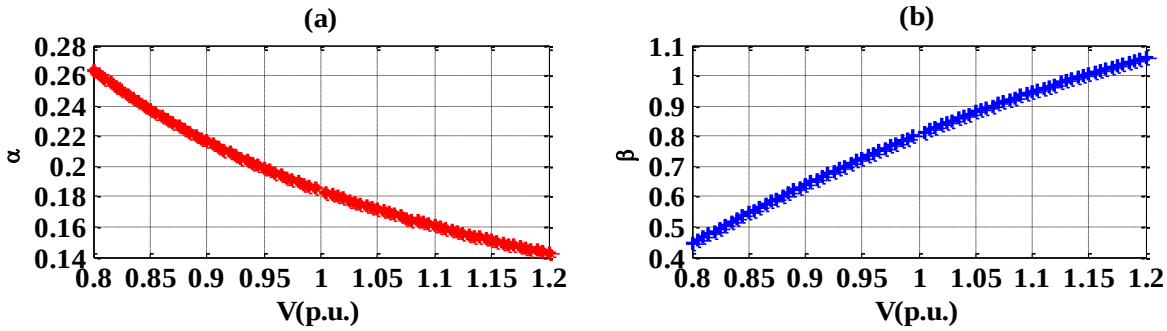


Figure C.3. (a) Exponent $\alpha$, and (b) Exponent $\beta$ of induction motors with typical parameters

It can be seen that the real power exponent $\alpha$ varies between 0.14-0.26, i.e., the real power of the induction motor is approximately proportional to $V^0$ rather than $V^2$. Therefore, the real power characteristics of an induction motor with typical parameters could be approximated by a constant power load model. The reactive power exponent $\beta$ varies from 0.45 to 1.05, therefore the reactive power characteristics could be approximated by either a constant current load model or a mix of constant power and constant current load model.

# Appendix D Typical Values of the Parameters for Load Models

## D.1 Exponential Load Model

Table D.1. Typical Values of Voltage and Frequency Exponents of the Exponential Load Model (adopted from [1])

| Load device/load class | | $k_{pu}$ | $k_{qu}$ | $k_{pf}$ | $k_{qf}$ |
|---|---|---|---|---|---|
| Air conditioner | 3-phase central | 0.088 | 2.5 | 0.98 | -1.3 |
| | 1-phase central | 0.202 | 2.3 | 0.9 | -2.7 |
| | window type | 0.468 | 2.5 | 0.56 | -2.8 |
| Water heaters, oven, deep fryer | | 2 | 0 | 0 | 0 |
| Dish washer | | 1.8 | 3.6 | 0 | -1.4 |
| Clothes washer | | 0.08 | 1.6 | 3.0 | 1.8 |
| Clothes dryer | | 2.0 | 3.2 | 0 | -2.5 |
| Refrigerator | | 0.77 | 2.5 | 0.53 | -1.5 |
| Television | | 2 | 5.1 | 0 | -4.5 |
| Incandescent lights | | 1.55 | 0 | 0 | 0 |
| Fluorescent lights | | 0.96 | 7.4 | 1 | -2.8 |
| Industrial motors | | 0.07 | 0.5 | 2.5 | 1.2 |
| Fan motors | | 0.08 | 1.6 | 2.9 | 1.7 |
| Agricultural pumps | | 1.4 | 1.4 | 5 | 4 |
| Arc furnace | | 2.3 | 1.6 | -1 | -1 |
| Residential load | summer | 1.2 | 2.9 | 0.8 | -2.2 |
| | winter | 1.5 | 3.2 | 1 | -1.5 |
| Commercial | summer | 0.99 | 3.5 | 1.2 | -1.6 |
| | winter | 1.3 | 3.1 | 1.5 | -1.1 |
| Industrial | | 0.18 | 6 | 2.6 | 1.6 |
| Power plant auxiliaries | | 0.1 | 1.6 | 2.9 | 1.8 |

## D.2 Induction Motor Model

Table D.2. Parameters for Models of Single-phase and Three-phase Induction Motors (adopted from [1])

| Index | End-users | Phase number | $R_s$ | $X_{ys}$ | $X_m$ | $R_r$ | $X_{yr}$ |
|---|---|---|---|---|---|---|---|
| 1 | Residential and Commercial Room air conditioner | Single-phase | 0.10 | 0.10 | 1.8 | 0.09 | 0.06 |
| 2 | Refrigerator and freezer | Single-phase | 0.056 | 0.087 | 2.4 | 0.053 | 0.082 |
| 3 | Dish washer | Single-phase | 0.11 | 0.14 | 2.8 | 0.11 | 0.056 |
| 4 | Clothes washer | Single-phase | 0.11 | 0.12 | 2.0 | 0.11 | 0.13 |
| 5 | Clothes dryer | Single-phase | 0.12 | 0.15 | 1.9 | 0.13 | 0.14 |
| 6 | Heat pump residential space heating/central air conditioner | Three-phase | 0.033 | 0.067 | 2.4 | 0.048 | 0.062 |
| 7 | Heat pump commercial central air conditioner | Three-phase | 0.53 | 0.83 | 1.9 | 0.036 | 0.68 |
| 8 | Pumps, fans, other motors | Three-phase | 0.079 | 0.12 | 3.2 | 0.052 | 0.12 |
| 9 | Small industrial motors | Three-phase | 0.031 | 0.1 | 3.2 | 0.018 | 0.18 |
| 10 | Large industrial motors | Three-phase | 0.013 | 0.067 | 3.8 | 0.009 | 0.17 |
| 11 | Agricultural water pumps | Three-phase | 0.025 | 0.088 | 3.2 | 0.016 | 0.17 |
| 12 | Power plant auxiliaries | Three-phase | 0.013 | 0.14 | 2.4 | 0.009 | 0.12 |

# Appendix E Publications from the Thesis

## E.1 International Journal Papers

[E1] J. V. Milanovic and Y. Xu, "Methodology for Estimation of Dynamic Response of Demand Using Limited Data," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1288-1297, 2015.

[E2] Y. Xu and J. V. Milanovic, "Artificial-Intelligence-Based Methodology for Load Disaggregation at Bulk Supply Point," *IEEE Transactions on Power Systems,* vol. 30, pp. 795-803, 2015.

## E.2 International Conference Papers

[E3] Y. Xu and J. V. Milanovic, "Estimation of percentage of controllable load in total demand at bulk supply point," in *Proc. 9th Mediterranean Conference on Power Generation, Transmission Distribution and Energy Conversion (MEDPOWER 2014)*, Athens, Greece, 2014, pp.1-6.

[E4] Y. Xu, J. Cai, and J. V. Milanovic, "On accuracy of demand forecasting and its extension to demand composition forecasting using artificial intelligence based methods," in *Proc. 5th IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE 2014)*, Istanbul, Turkey, 2014, pp. 1-6.

[E5] Y. Xu and J. V. Milanovic, "Accuracy of ANN based methodology for load composition forecasting at bulk supply buses," in *Proc. 13th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS 2014)*, Durham, UK, 2014, pp. 1-6.

[E6] Y. Xu and J. V. Milanovic, "Framework for estimation of daily variation of dynamic response of aggregate load," in *Proc. 4th IEEE PES Innovative Smart*

*Grid Technologies Europe (ISGT EUROPE 2013)*, Copenhagen, Denmark, 2013, pp. 1-5.

[E7] Y. Xu and J. V. Milanovic, "Developmnet of probabilistic daily demand curves for different categories of customers," in *Proc. 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013)*, Stockholm, Sweden, 2013, pp. 1-4.

## E.3 Technical Reports

[E8] CIGRE WG C4.605 : "Modelling and aggregation of loads in flexible power networks", Jovica V. Milanović, (Convenor), Julija Matevosiyan, Anish Gaikwad, Alberto Borghetti, Saša Ž. Djokić, Zhao Yang Dong, Andrew Halley, Lidija M. Korunović, Sergio Martinez Villanueva, Jin Ma, Pouyan Pourbeik, Fernanda Resende, Stefan Sterpu, Fortunato Villella, Koji Yamashita, Odin Auer, Karim Karoui, Dimitry Kosterev, Shu Kwan Leung, Dumisani Mtolo, Samila Mat Zali, Adam Collin, Yizheng Xu, (566), ISBN: 978-2-85873-261-6, February 2014

## E.4 Submitted International Journal Papers

[E9] Y. Xu and J. V. Milanovic, "Day-ahead Prediction and Shaping of Dynamic Responses of the Demand at Bulk Supply Point," *submitted to IEEE Transactions on Power Systems*, 2015 (TPWRS-00371-2015).

## E.5 Accepted International Conference Papers

[E10] Y. Xu and J. V. Milanovic, "Probabilistic Estimation of Rated Demand at Bulk Supply Point," accepted for presentation at *IEEE PES PowerTech Eindhoven 2015*, Eindhoven, Netherland, 2015.