



City Research Online

City St George's, University of London

Citation: Leung, Ming (2017). A novel approach to determine building occupancy for cooling energy consumption prediction. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/17403/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

**A NOVEL APPROACH TO DETERMINE BUILDING OCCUPANCY FOR
COOLING ENERGY CONSUMPTION PREDICTION**

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

By

LEUNG MING CHIU

School of Mathematical Computer Science and Engineering
Electrical and Electronic Engineering Department
Systems and Control Research Centre

City, University of London
Northampton Square
London EC1V 0HB
United Kingdom

May 2017



Abstract

Building cooling load prediction is one of the key elements in the energy conservation achievements. Most of the mathematical models using in the industry nowadays include forward and inverse modeling approaches. However, these models consume much computer resources and require a longer computational time.

Multi-layer perceptron (MLP) model of artificial neural network (ANN) is adopted in this thesis. The model is widely used in engineering approaches that render good performance in adaptability, nonlinearity and mapping. It also has good ability in predicting the cooling energy consumption of buildings. It is reported that the occupants' activities inside the buildings can have significant impact on the accuracy of the model. The existing input parameters used for the ANN models could not represent the complexity of the activities inside the buildings well. Most of the traditional ANN models adopted fixed profile or historic load data to represent building occupancy in simulating building cooling energy consumption. However, building occupancy is never still. The dynamic changes occurred in the occupancy of the buildings therefore make the forecasting of building cooling load difficult and less accurate. This thesis aims at (i) introducing a novel model to represent occupants' presence and activities; and (ii) investigating the effect of using the novel model on improving the predictive accuracy of building cooling energy consumption.

The simulation results demonstrate that building occupancy data play a significant role in building cooling energy consumption prediction and the use of the novel approach significantly improves the predictive accuracy of building cooling energy consumption model.

Keywords: Artificial Neural Network, Building Cooling Load, Building Occupancy.

Acknowledgments

I would like to thank Professor Lai Loi Lei and local supervisor Dr. Tse Chung Fai, Norman for their excellent guidance and provided me the opportunity to carry out the research study. Without their encouragement and valuable advices, it is not possible for me to complete this research study. I would like to thank Professor Lai and his family for their warm reception in London when I arrived United Kingdom for school registration.

Thanks also give to Dr. Chan Yau Chung for his advice in program writing and my working department for collecting data on university building in Hong Kong.

Last but not the least, I have to express my appreciation to my wife, Hanny and my son, Karsten. Without their understanding, patience and encouragement in past six years, it would have not been possible for me to complete my research study.

Copyright Declaration

I, “Leung Ming Chiu”, of Hong Kong, “the Depositor”, would like to deposit “A Novel Approach to Determine Building Occupancy for Cooling Energy Consumption Prediction”, hereafter referred to as the “Work”, in the City University Institutional Repository and agree to the following:

NON-EXCLUSIVE RIGHTS

Right granted to the City University Institutional Repository through this agreement are entirely non-exclusive and royalty free. I am free to publish the Work in its present version or future versions elsewhere. I agree that the City University Institutional Repository administrators or any third party with whom the City University Institutional Repository has an agreement to do so may, without changing content, translate the Work to any medium or format for the purpose of future preservation and accessibility.

DEPOSIT IN THE CITY UNIVERSITY INSTITUTIONAL REPOSITORY

I understand that the Work deposited in the City University Institutional Repository will be accessible to wide variety of people and institutions – including automated agents – via the World Wide Web. I also agree to an electronic copy of my thesis being included in the British Library Electronic Theses On-line System (EThOS).

I understand that once the Work is deposited, a citation to the Work will always remain visible. Removal of the Work can be made after discussion with the City University

Intuitional Repository, who shall make best efforts to ensure removal of the Work from any third party with whom the City University Institutional Repository has an agreement.

I AGREE AS FOLLOWS:

- That I am the author or co-author of the Work and have the authority on behalf of the author or authors to make this agreement and to hereby give the City University Institutional Repository administrators the Right to make available the Work in the way described above.
- That I have exercised reasonable care to ensure that the Work is original, and does not to the best of my knowledge break any UK Law or infringe any third party's copyright or other Intellectual Property Right. Where I have included third party copyright material, I have fully acknowledged its source.
- The administrators of the City University Institutional Repository do not hold any obligation to take legal action on behalf of the Depositor, or other Rights holders, in the event of breach of Intellectual Property Rights, or any other Right, in the material deposited.

Leung Ming Chiu

May 2017

Table of Contents

Abstract.....	i
Acknowledgements.....	iii
Copyright Declaration.....	iv
Table of Contents.....	vi
Glossary.....	x
Abbreviations.....	x
Mathematical Notations.....	xii
List of Figures and Tables.....	xiii
List of Figures.....	xiii
List of Tables.....	xv
Chapter 1 Introduction.....	1
1.1 Motivation.....	1
1.2 Building cooling energy consumption prediction.....	3
1.3 From state-of-art approach to intelligent approach of building cooling energy consumption prediction modelling	4
1.4 ANN models for building cooling energy consumption prediction.....	6
1.5 Applications of building cooling energy consumption prediction using ANN models.....	6
1.6 A novel approach to determine building occupancy in ANN model.....	7
1.7 Objectives	9
1.8 Organisation of the thesis.....	10
Chapter 2 Review of existing load forecasting technologies	13
2.1 General.....	13
2.2 Types of models	15
2.2.1 Forward approach models.....	15

2.2.2	Inverse approach models.....	16
2.3	Types of load forecasting.....	17
2.3.1	Short-term load forecasting (STLF).....	18
2.3.2	Medium-term load forecasting (MTLF).....	18
2.3.3	Long-term load forecasting (LTLF).....	18
2.4	Intelligent approaches to load forecasting.....	18
2.5	ANN models for cooling load forecasting.....	20
2.6	Conclusion.....	22
Chapter 3	Load forecasting using artificial neural networks.....	23
3.1	General.....	23
3.2	Model architecture.....	23
3.3	Input parameters.....	25
3.3.1	Sources of building cooling load.....	25
3.3.2	Components that influence building cooling energy.....	28
3.4	Hidden and output layers.....	31
3.5	Output parameter.....	32
3.6	Network training.....	32
3.6.1	Backpropagation algorithm.....	32
3.6.2	Early stop validation approach.....	33
3.7	Validation of results	35
3.7.1	Performance measures criteria.....	35
3.7.2	Training schedule.....	36
3.8	Conclusion.....	37
Chapter 4	Current building occupancy modelling.....	39
4.1	General.....	39
4.2	Interactions between building occupancy and energy consumption	40
4.3	Modelling building occupancy.....	41
4.3.1	Standard profile and diversity factors.....	42
4.3.2	Empirical models of occupants' presence and actions in buildings.....	43

4.4	Current ANN load forecasting models for measuring building occupancy.....	47
4.4.1	Factors affecting the internal load.....	47
4.4.2	Limitations of the current ANN approach to determine building occupancy.....	50
4.5	Conclusion.....	50
Chapter 5	A novel approach to determine building occupancy.....	52
5.1	General.....	52
5.2	Current approaches.....	54
5.3	Occupants' presence and activity.....	55
5.3.1	Background.....	55
5.3.2	User interactions with environmental control systems.....	56
5.3.3	Stochastic processes of occupants' presence and activity.....	58
5.4	Measuring building occupancy via stochastic process.....	58
5.4.1	Pollutants emission and dilution.....	59
5.4.2	Metabolic heat gain and removal.....	59
5.4.3	Comfort and demand.....	60
5.4.4	Arrivals, departures and inter-floor movements.....	61
5.5	Novel approach to determine building occupancy data.....	62
5.5.1	Total space electrical consumption of each floor.....	62
5.5.2	Data for the PAU operation schedule in a building.....	63
5.5.3	Day/hour type.....	63
5.6	Stochastic model of occupancy and activity inside an air-conditioned university building.....	63
5.7	Conclusion.....	64
Chapter 6	Application of the novel approach in ANN model.....	66
6.1	General	66
6.2	Characteristics of the sample university building	66
6.3	Development and training of the ANN model.....	68
6.4	Model training and prediction results.....	70

6.4.1	Input parameters and algorithm development.....	70
6.5	Hourly and daily cooling load prediction.....	72
6.6	Discussion.....	74
6.6.1	Overall prediction performance.....	74
6.6.2	Variability in prediction performance.....	80
6.6.3	Daily peak cooling demand prediction performance.....	83
6.7	Conclusion.....	84
Chapter 7	Conclusions and recommendations.....	85
7.1	Conclusions.....	85
7.2	Recommendations for future research.....	86
List of Publications		88
Journal paper		88
Conference papers		88
References		89
Appendix I	–A brief introduction to ANNs	103
Appendix II	–Configuration of the computational machine.....	105
Appendix III	–The use of occupancy space electrical power demand in building cooling load prediction.....	106
Appendix IV	–Application program scripts for the ANN model.....	120
Appendix V	–Building data for simulation model.....	127

Glossary

Abbreviations

AHU	Air handling unit
ASHRAE	American Society of Heating, Refrigerating and Air-conditioning Engineers
ANN	Artificial neural network
BLAST	Building loads analysis and systems thermodynamics
BP	Backpropagation
BPMS	Building power monitoring system
CCMS	Central control and monitoring system
CO ₂	Carbon dioxide
HVAC	Heating, ventilating and air-conditioning
GA	Genetic algorithms
KWH	Kilowatt hour
LTLF	Long-term load forecasting
MAPE	Mean absolute percentage error
MLP	Multi-layer perceptron
MTLF	Medium-term load forecasting
PAU	Primary air handling unit
RMS	Root-mean-square
RMSPE	Root-mean-square-percentage-error
STLF	Short-term load forecasting
SVM	Support vector machines

VAV Variable air volume

Mathematical Notations

E	RMS error
N_h	number of hidden neurons
N_p	sample size
o	desired output
p	pattern
r	coefficient-of-correlation
$\bar{p}$	predicted values
$\bar{t}$	target value
t	network output target
k	subscript indicating the sum of all of the nodes in the downstream layer
j	subscript indicating the weight position of each node
ε	learning rate and size of the weight and adjustments during each training iteration
m	momentum factor that is applied to the weight change used in the previous training iteration, w_{ij} (old)
$^{\circ}\text{C}$	degree Celsius
%	percentage

List of Figures and Tables

List of Figures

Figure 2-1	Typical forward approach model.....	16
Figure 2-2	Typical inverse approach model.....	17
Figure 3-1	Typical artificial neural network computational structure.....	24
Figure 3-2	Sources of building cooling load.....	26
Figure 3-3	Components for estimating building cooling load	29
Figure 3-4	Input parameters for building cooling load prediction.....	30
Figure 3-5	Early-stop validation approach of the backpropagation training process.....	35
Figure 3-6	Training scheme designed to minimise the effect of the random data extraction process.....	37
Figure 4-1	Interactions between occupants' presence and behaviour and building load	41
Figure 5-1	Dynamic interaction of systems in an air-conditioned university building.....	57
Figure 5-2	Occupants' arrival pattern in a multi-tenant office building in Paris	62
Figure 5-3	Stochastic model of occupants' presence and activity inside an air-conditioned university building.....	64
Figure 6-1	Architecture of the prediction model.....	69
Figure 6-2	2010 hourly simulation results.....	73
Figure 6-3	2011 hourly simulation results.....	73

Figure 6-4	2010 daily simulation results.....	74
Figure 6-5	2011 daily simulation results.....	74
Figure 6-6	Comparison of the performance indices for the four simulations...	75
Figure 6-7	Monthly performance indices of the 2011 hourly prediction results (Simulation D) R (upper) and MAPE and CV (lower)	77
Figure 6-8	Monthly performance indices of the 2011 daily prediction results (Simulation D) R (upper) and MAPE and CV (lower)	78
Figure 6-9	2011 prediction results with six estimations (Simulation D) hourly (upper) and daily (lower)	81

List of Tables

Table 2-1	Type of models for load forecasting.....	18
Table 6-1	Design criteria of the cooling system.....	68
Table 6-2	Representation of day-type.....	71
Table 6-3	Four simulations with different input parameters.....	71
Table 6-4	2010 hourly and daily simulation results.....	72
Table 6-5	2011 hourly and daily simulation results.....	72
Table 6-6	Monthly performance indices of 2011 hourly prediction results (Simulation D).....	76
Table 6-7	Monthly performance indices of 2011 daily prediction results (Simulation D).....	76
Table 6-8	Comparison with previous study.....	80
Table 6-9	2011 hourly prediction - mean performance indices of the six estimations (Simulation D).....	82
Table 6-10	2011 daily prediction - mean performance indices of the six estimations (Simulation D).....	82
Table 6-11	Monthly performance indices of 2010 daily peak prediction results (Simulation D).....	84
Table 6-12	Monthly performance indices of 2011 daily peak prediction results (Simulation D).....	84

**THE FOLLOWING PARTS OF THIS THESIS HAVE BEEN REDACTED
FOR COPYRIGHT REASONS:**

p. 107-119, Journal Article: Leung, M., Tse, N., Lai, L. & Chow, T.
2012, "The use of occupancy space electrical power demand in building
cooling load prediction", ENERGY AND BUILDINGS, vol. 55, pp. 151-
163.

Chapter 1

Introduction

1.1 Motivation

Rising global temperatures and sea levels are evidence of climate change. Greenhouse gas emissions are believed to be major contributors to climate change, as greenhouse gases such as water vapour, carbon dioxide, methane and ozone have a warming effect on the atmosphere.

Since the beginning of the industrial revolution in 1750, humans have, through extensive fossil fuel combustion, released an enormous amount of carbon dioxide into the atmosphere. According to Blasing (2015), the concentration of atmospheric carbon dioxide increased from 280 ppm in 1750 to 400 ppm in 2015 [1].

Modernisation brings costs. Electricity is an essential requirement for a high standard of living. The main consumption of electricity comes from buildings in modern cities. However, electrical power plants rely mainly on coal and other fossil fuels; as a result, the process of generating electricity leads to a continuously increasing volume of carbon dioxide emissions.

To improve this situation, carbon emission reduction organisations have been established in many countries. They encourage gas and electricity suppliers to adopt greenhouse gases emission protocols and to reduce greenhouse gases emission by a certain amount. They also promote energy conservation and efficiency.

Due to the increasing demand for energy and rising tariffs on electricity, energy efficiency has become a global concern. People in many fields are examining ways to reduce the energy loss in power transmission to improve energy efficiency and reduce carbon emissions. According to the Census and Statistics Department of the Hong Kong SAR Government, the domestic and commercial sectors in Hong Kong accounted for 90% of the total electricity consumption in 2011 [2]. As one of the most densely populated metropolises in the world, there are thousands of skyscrapers in Hong Kong. Energy efficiency in these buildings has been a concern since the 90s.

In Hong Kong, 60% of carbon emissions are produced by electricity generation, and buildings consume 89% of the electricity produced. Improving the electrical energy efficiency of existing buildings may therefore be the fastest route to carbon emission reduction. Thus, the recently established Smart Grid technology can be used to combat climate change. Specifically, demand side management and dynamic demand response functions enable a more efficient use of electrical energy. Smart Grid technology has important implications for building cooling systems, as building cooling loads account for approximately 45% of buildings' electrical energy consumption. To implement these developments efficiently, it is necessary to accurately predict building cooling load demand. This is particularly important in Hong Kong, where most commercial buildings are fully air-conditioned and mechanically ventilated. As heating, ventilation, air conditioning and lighting account for the majority of a building's energy use, there has been considerable research on how to effectively manage building energy demand, especially the demand of air-conditioning systems, in sub-tropical regions like Hong Kong.

The four main factors contributing to building cooling load demand are as follows:

- i. building characteristics (location, orientation and type of building);
- ii. building service systems installed for sustaining human activities (heating, ventilating and air-conditioning (HVAC), lighting and electricity);
- iii. outdoor environment and meteorological factors (temperature, humidity, solar radiation and etc.); and
- iv. occupants' activities (planned operations and human behaviour).

1.2 Building cooling energy consumption prediction

Building designers have gained access to high-end computer hardware and software programs and computer simulation models that can evaluate building energy consumption levels [3]. These models also provide building managers with energy demand profiles and information on energy consumption levels, which can be used to estimate buildings' energy use conditions and running costs. With this piece of information, building managers can take steps to ensure the energy consumption and operating costs of a building are maintained at an acceptable level [4]. Energy simulation programs are essential tools in building energy audits. Energy auditing experts use building energy simulation programs to precisely predict the energy consumption conditions of actual buildings. In sub-tropical regions, energy audits focus on reducing the space cooling loads and the energy used in air-conditioning systems, as air conditioning systems account for at least 45% of the total building energy consumption in these regions [5]. It has been suggested that the building industry should pay more attention to the energy requirements of air-conditioning systems in existing buildings and in building design. Comparing the energy use and operating costs of alternative systems enables industry experts to choose the optimal design for a building's air-conditioning system. This thesis will focus on the simulations of a

building's cooling load under various internal loads. With the improved building energy simulations made possible with current software, building designers can identify poor building envelopes or oversized HVAC equipment and chiller plants, thus reducing maintenance and running costs. Building cooling load demand prediction has many applications [6], which are listed as follows:

- i) comparing the predicted operational cooling load demand to the actual load and exploring the building operational problems in the early stage of building occupation; and
- ii) comparing the predicted energy use after a building retrofit project to the actual load of the existing building, in order to estimate the energy savings that could be achieved.

1.3 From state-of-art approach to intelligent approach of building cooling energy consumption prediction modelling

Many state-of-art building energy simulation models can accurately simulate building energy performance, as long as the building operation is well planned. Building energy simulation and analysis has been practiced for decades. However, traditional analytical approaches adopt a mathematical model, which requires advanced analytical skills to implement. Building designers and management teams may have difficulty applying such complicated models and achieving accurate simulation results. Furthermore, the computer simulation processes require a large amount of time. The limited use of analytical approaches in building energy studies also discourages the development of this expertise. To address these problems, artificial intelligence (AI) techniques have been widely adopted in recent analytical approaches. A model with AI techniques can be used to manipulate historical, incomplete or noisy data and to perform non-linear

data analysis. It can also be used to make predictions and generalisations at high speed [7]. A model with AI techniques can also be used for pattern recognition, system optimisation, system identification, system mapping or signal processing. Building energy prediction programs nowadays are usually constructed using artificial neural networks (ANN), genetic algorithms (GA), fuzzy logic (FL) or support vector machines (SVM). The input parameters of an ANN model include difference variables, historical data and external variables that are usually non-linear. The output of an ANN model is the best values for building energy consumption, and ANN models have recently been adopted for load prediction [8,9,10,11]. Some of the input data are used to train the model and the rest are used to validate the ANN model by comparing the actual output values with the predicted output values. ANN models also use an approximate function to evaluate the output values [12]. GA is a robust approach to developing heuristics for large-scale combinational optimisation problems [13]. Studies of short-term load forecasting have reported good results from GA models. However, GA models usually require a long period of computational time [13]. The FL simulates human decision-making, which is characterised by uncertainty and imprecision. As FL can only provide simple outputs of either true or false in its final stage [14], the model is seldom to be used in building energy load prediction. The SVM approach is commonly used with a regression, which makes it possible to retain all of the main features that characterise the maximal margin algorithm in the output; a non-linear function is trained by a linear training machine in a kernel-induced feature space, and the capacity of the system is controlled by a parameter that does not depend on the dimensionality of the feature space [15]. In the classification problem approach, there is motivation to seek and optimise the generalisation bounds for a given regression. The approach defines the loss

function that ignores errors, which is situated a certain distance from the true value. This type of function is often called the epsilon-intensive loss function.

1.4 ANN models for building cooling energy consumption prediction

ANN, GA, FL and SVM models with AI techniques have been adopted by a large number of researchers to simulate building energy use [26,28,29,30,32]. ANN models are capable of modelling highly nonlinear systems. They are also black boxes and work well for large datasets. ANN is a powerful data-driven, self-adaptive and flexible computational tool, which is capable of capturing a high degree of accuracy of the nonlinear and complex underlying characteristics of any physical process. ANN models accurately predict building cooling loads and give forecasts for different times of the year. They are universal function estimators that can evaluate a continuous function and achieve an acceptable level of accuracy [16,17,18]. Traditional prediction models assume a certain relationship between input and output parameters and thus are not suitable for determine complex relationships in real situations. ANN models have produced better results than traditional models in the analysis of complex and nonlinear relationships [18]. The data-driven technique of ANN models can perform nonlinear modelling without prior knowledge of the relationships between the input and output parameters. ANN models have been widely used in energy analysis and predictions as they are formulated by dynamic-inverse models, whereas the traditional models are formulated by programmed rules. In other words, ANN models are capable of learning from historical data rather than simply following programmed rules.

1.5 Applications of building cooling energy consumption prediction using ANN models

Building energy simulations can be used to estimate the energy use and operating costs of buildings, especially those related to air-conditioning systems. A simulation compares alternative HVAC system designs, and selects the optimal design for a particular building. ANN models, which are an inverse type of model, are suitable for such building energy simulations for two reasons.

- i. The ANN simulation models provide building management teams with information about the past operating strategies, thus enabling them to rearrange and operate the system in the most conservative way. The models can also diagnose the actual energy use of a building and compare it with a benchmarking system, so as to find out the optimal operating strategy.
- ii. ANN models are trained using the existing energy use data of a building. This can be used to compare the energy use data after the building is retrofitted. The difference between the existing and the retrofitted energy use profile is a measure of the energy conservation achievements of the newly installed equipment. ANN dynamic models also include weather data and occupancy profiles as input parameters. Thus making the models a useful tool for the building designers and the building management teams.

In conclusion, ANN models are capable of handling complex systems and nonlinear parameters. As a result, users do not need to obtain the details of the buildings' physical systems or even a comprehensive understanding of ANN models. The energy performance of buildings can thus be easily obtained by a wide variety of users.

1.6 A novel approach to determine building occupancy in ANN model

Modern buildings have complex building service systems, and control strategies and building energy prediction techniques are becoming an important part of building

systems design [19]. Citizens seeking high quality living standards desire fresh indoor air and a pleasant living environment. As a result, building energy simulation is an important part of building design and operation [20]. Traditional ANN building energy simulations use fixed operational schedules as occupancy input data. However, flexible working hours and interactive control systems nowadays may affect the accuracy of energy use simulations if the models are based on simple assumptions on the behaviour of occupants. The study has reported that as users' behaviour affected building energy load, it needed to be included in simulation models as an input parameter [19].

The accuracy of the simulation results is determined by the model's algorithms and the robustness; and also relevance of the input parameters, such as the external weather and building occupancy conditions. The building occupancy condition reflects the fixed schedule and type of human activities inside the building [21]. In buildings with the same physical characteristics, such as the same building envelope and building service systems, the interactions of the buildings' occupants with the building systems may result in very different energy demand profiles for similar buildings [22]. Traditional ANN simulation models adopt a fixed schedule for occupancy rate, resulting in poor simulation performance. Modern building management teams recognise that due to flexible working hours and interactive control of building service systems, models must include more sophisticated estimates of occupancy rate to achieve higher accuracy [23]. Managers perform energy audits and analyses of buildings to determine the optimal operational strategy and to identify energy conservation measures (Energy Efficiency Advisory Committee, 1995). To achieve this aim, managers require data on the precise energy consumption of each building service system; however, metering devices for monitoring such systems have seldom been installed in Hong Kong in the past decade. Only monthly electricity bills are available [24]. In addition, only air-conditioning plant

equipment log sheets can be obtained in Hong Kong. The low-end metering devices provide insufficiently accurate data for the model.

However, energy audits are becoming mandatory in Hong Kong. Building regulations require alternations to existing buildings and require new buildings to install metering devices. The university building targeted in this thesis needed to be upgraded with an enhanced metering system, to meet the new building regulations. A sophisticated metering system was built in 2010. A datum is collected every 15 minutes and is stored on a high quality server.

Therefore, total space electrical consumption of each floor, data for the primary air handling unit (PAU) operation schedule in a building and day/hour type, which are a novel approach to determine the building occupancy, can be retrieved and becomes one of the input parameters used in building occupancy for ANN model.

1.7 Objectives

Both the indoor microclimate and the energy demand of a large building complex are affected by occupants' activities, which are stochastic in nature [25]. A number of studies have assessed people's interactions with building environmental systems, through user-controlled actions or user-controlled devices that switch on or off. Although there are practical difficulties in predicting individual human behaviours, trends in user-controlled related behaviours and group patterns of building occupants can be obtained from long-term observational data. This thesis uses a novel approach to establish dynamic occupants' presence and behaviour models. An ANN model is used to simulate the total cooling load demand of a university building in Hong Kong based on the building's occupancy space electrical power demand profile. Real-time power data from the building power monitoring system (BPMS), hourly weather data from the

Hong Kong Observatory and time factors are used as input parameters for the ANN model. The total electrical demand of the building cooling system is selected as the model output. The occupancy space electrical power demand from the BPMS is used to determine the human behaviours inside the university building, and is formed as an input parameter of ANN model – this is a novel approach to determine building occupancy. The proposed ANN model has been validated with actual data retrieved from the university building. The simulation results show that the input parameters, including the proposed building occupancy, significantly improve the accuracy of ANN models in predicting building cooling load. The proposed model, which uses an occupancy space electrical power demand profile, is found to be capable of improving the accuracy of the cooling energy simulation.

Therefore, this thesis aims at i) developing a novel model to represent occupancy' presence and behaviours; and ii) investigating the effect of using the developed model on improving the predictive accuracy of building cooling energy consumption.

The novel model includes three areas that are usually overlooked in previous studies in determining occupancy's presence and behaviours. They are i) total space electrical consumption of each floor; ii) data for the primary air handling unit (PAU) operation schedule in a building; and iii) day/hour type.

The proposed model is verified and validated by real building data.

1.8 Organisation of the thesis

This thesis consists of seven chapters. The first chapter provide a general description of building cooling simulation.

Chapter 1 explains the need for building cooling load predictions and reviews the energy prediction and modelling methods currently used by the building industry and researchers. The constraints of the traditional approaches are discussed in Chapter 2 and an intelligent approach based on ANN methodology is proposed and will be discussed in Chapter 3. The need for building occupancy data – internal parameter – is recognised in the new ANN model.

Chapter 2 introduces the theoretical foundations of load forecasting. Different load forecasting techniques was discussed and the applications of short-term load forecasting are presented. Traditional approaches are used to compare with the ANN approach – one of the most popular new applied techniques.

Chapter 3 describes the development of the ANN model for building cooling load prediction. The background ANN models is introduced before the architecture of the specific model is described. The development of the model's three major components – the input layer, the hidden layer and the output layer – is then explained. Each component is introduced; and the critical external and internal load factors are determined. The methods used to determine the numbers of hidden layers are presented. Then, the chapter give an overview of the multi-layer perceptron algorithm used in this thesis and presented the early stop validation approach for network training. The last section of the chapter discusses the details of network validation by means of a performance measure and training schedule.

Chapter 4 shows how building occupancy is modelled to simulate energy use. A discussion of how building occupancy interacts with building energy consumption is presented before the current methodologies for modelling building occupancy are reviewed. The most widely used approaches – the standard profile and diversity

approach – are discussed before the details of more empirically based models of people's presence and actions are given. The final two sections discuss how current ANN models handle building occupancy data and the limitations encountered if the building occupancy data are not correctly incorporated into simulations of building energy use.

Chapter 5 explores a novel approach to the acquisition of building occupancy data. The concepts underlying stochastic models of people's presence and activities are elaborated in detail. The details of such stochastic models are then presented, including the modelling of the presence of occupants using the fresh air supply rate; modelling of occupants' presence and behaviour via the conditioned air supply rate and the energy consumption rate due to tenant demand; and modelling of occupants' presence and activity via the lift and escalator traffic rate. The chapter concludes by proposing three approaches to measuring building occupancy: i) the data for PAU operation schedule; ii) day/hour type; and iii) occupancy space power demand.

Chapter 6 verifies the ANN model developed and incorporates building occupancy into the model by using real building data from university buildings in Hong Kong. The application of the crucial factor – space electrical load profile – is introduced and discussed.

Chapter 7 concludes this thesis by summarising the findings and suggesting recommendations for the use of novel occupancy models, in which the building occupancy factors that affect building cooling load are used as inputs for an ANN model.

Chapter 2

Review of existing load forecasting technologies

2.1 General

Facilities managers need to estimate and predict the energy demand according to the use of buildings, especially the energy demand for the air conditioning systems. Both external climatic conditions and occupancy rate vary the heat load inside buildings [26]. Traditional methods of predicting the energy demands of buildings treat buildings characteristics, such as building materials, air-conditioning systems, heating systems and building functions, as more or less ideal physical sub-processes [27]. These methods consider a large number of physical parameters with varying degrees of idealisation and simplification. Time series analytical and statistical methods are adopted in these modelling techniques. These methods ignore the influence of climatic factors such as atmospheric temperature and relative humidity. Moreover, extensive computation time and resources are required for the energy demand prediction, especially when the data are collected on a daily or hourly basis [28].

Existing methods also make a large number of assumptions about occupancy rates. A forward and inverse approach is usually adopted in current energy demand estimations and analyses since then.

Forward approach models are used in building designing, in order to optimise the heating, ventilating and air-conditioning (HVAC) systems by comparing the energy use of different building system designs. The most common software programs for energy

analysis are DOE-2.2 (J. J. Hirsch & Associates, 2005), TRNSYS (Solar Energy Laboratory, 2006), BLAST (BLAST Support Office, 1992) and ENERGYplus (US Department of Energy, 2007). They all adopt the forward dynamic approach, in which energy predictions are based on the physical parameters of building systems, such as their geometry, location, construction details and the HVAC type and operation [28]. The forward dynamic approach uses numerical or analytical processes to determine energy flow between buildings. It requires complicated software programs and lengthy computing times. Forward models could not accurately predict the energy demand of each building due to varying factors, such as external weather, construction materials, internal load of occupants, building control strategies, unconditional responses between systems and various modes of heat transfer (ASHRAE, 2005)[6].

Inverse approach models are used to compare the actual performance data on energy use and the energy demand prediction in buildings (ASHRAE, 2009). The inverse approach derives representative building parameters, such as the building load coefficient, building time constant from data on energy use, weather and relevant performance data [28]. Dynamic inverse models, such as artificial neural networks (ANNs) are used for energy analysis and have the ability to learn from the data. The models provide output based on generalised input data. They do not require knowledge of an explicit relationship between the input and output parameters. The models do not need to be set up or operate by a skilful programmer. ANN models can be trained on a dataset with or without a designated solution for each case. The models are able to identify the output solutions after training and provide solutions for new datasets. The ANN models are suitable for energy analysis because they can model the behaviour of complex building systems with several independent input parameters. Therefore, a building professional can make an energy demand prediction effortlessly and evaluate the performance of the

buildings. Furthermore, ANN models are suitable for load forecasting [26,28,29,30,31,32].

The advantages of the ANN approach have been demonstrated in previous studies [26,28,29,30,31,32]. Current load forecasting types and techniques are summarised in the next section.

2.2 Types of models

Two major types of models, which are named as Forward Approach and Inverse Approach, are mainly used to estimate and analyse building energy. The forward approach is used by building designers to estimate the energy use of buildings and the HVAC systems, in order to optimise energy savings. The inverse approach is used to estimate the actual energy use of buildings and HVAC systems [6].

2.2.1 Forward approach models

Forward approach models use the physical design of building energy systems to estimate the energy demand of end users and to predict the energy saved through conservation. Steady state forward models and dynamic forward models are commonly used in building energy prediction and there are several software programs based on these approaches [28], such as DOE-2, building loads analysis and systems thermodynamics (BLAST) [36], ENERGYplus (US Department, 2007) and the TRNSYS simulation tool [35].

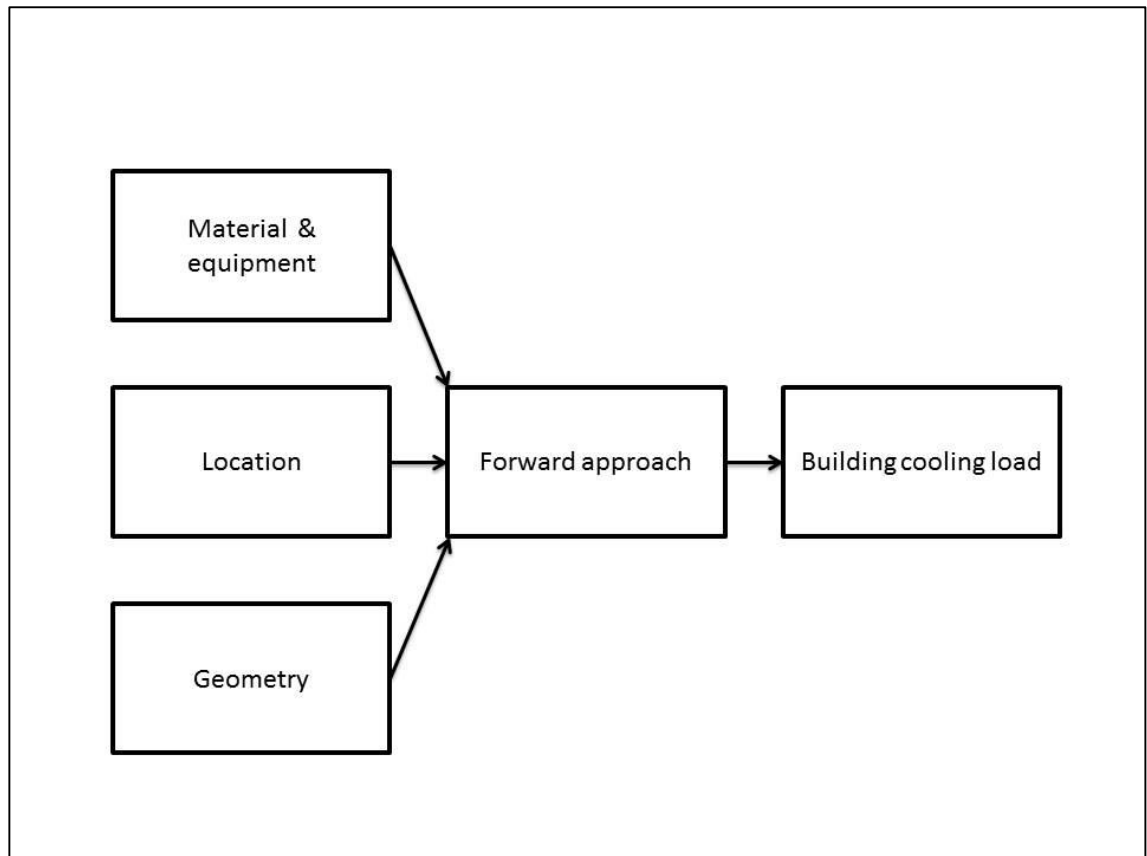


Figure 2-1 Typical forward approach model

2.2.2 Inverse approach models

Inverse approach models are used to estimate building energy efficiency after the adoption of energy conservation measures and energy saving measures, such as energy retrofits and building energy efficiency improvements. Steady state inverse models combine different models such as the similar-day approach, regression and time series, but they have limitations in the analysis of transient effects such as thermal mass effects and seasonal changes in the efficiency of HVAC systems. Dynamic inverse models are capable of simulating the thermal mass dynamic effects of buildings and complex systems with several independent parameters; however, the models require large

amount of resources for user interaction and realising the prediction building model or system [28,33,37].

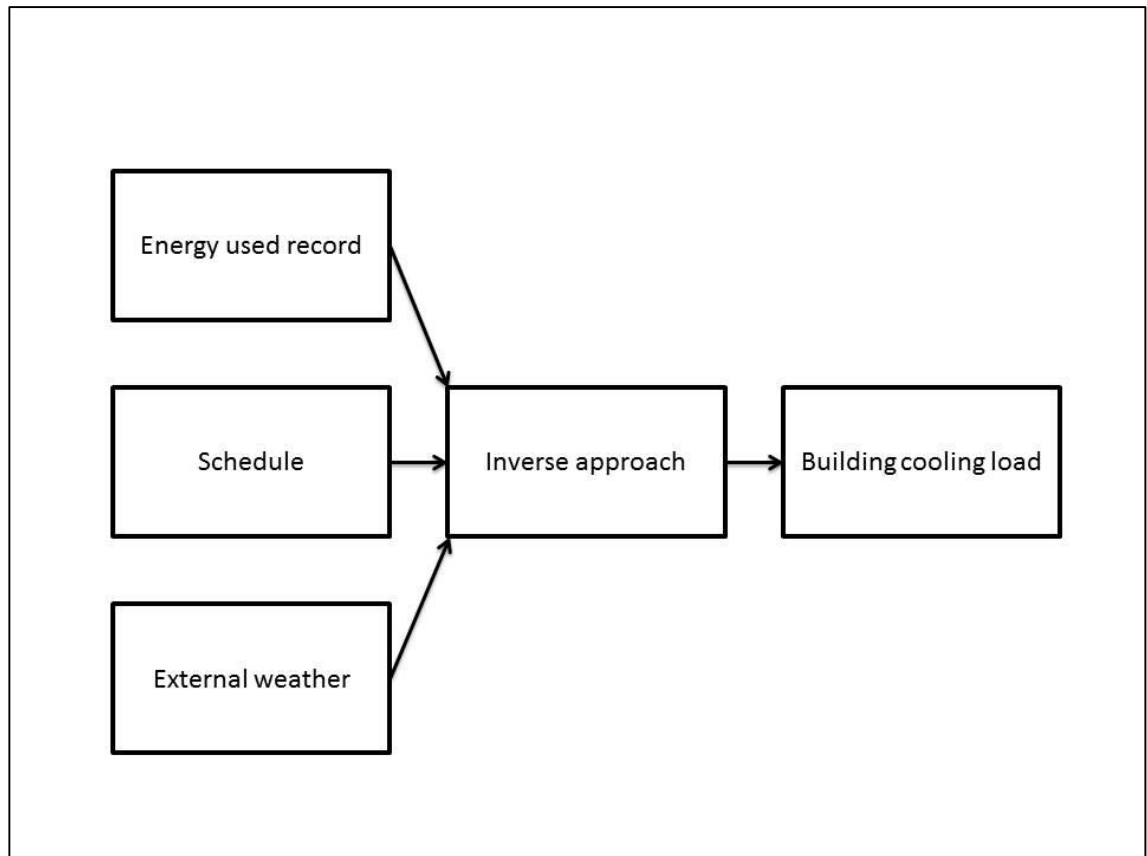


Figure 2-2 Typical inverse approach model

2.3 Types of load forecasting

Load forecasting is very important for the building and utilities industries. In building management, the optimisation of control strategies and the energy conservation of HVAC systems are essential applications. In the utilities industry, load forecasting has important applications in power generation, grid development and load shedding [33]. The different types of load forecasting are shown in table 2-1.

Short-term load Forecasting (STLF)	Medium-term load Forecasting (MTLF)	Long-term load Forecasting (LTLF)
Hourly to weekly	Monthly to yearly	Yearly basis

Table 2-1 Type of models for load forecasting

2.3.1 Short-term load forecasting (STLF)

STLF predicts the future hourly load for periods of up to one week [12]. It is an essential part of the daily operation of utilities systems [34].

2.3.2 Medium-term load forecasting (MTLF)

MTLF predicts future hourly loads for periods ranging from one week to one year; such forecasts allow utilities companies to estimate energy demands over longer periods and assist them in contract negotiations with other companies [34].

2.3.3 Long-term load forecasting (LTLF)

LTLF predicts the load for periods of over a year. The duration can be even more than twenty years [12].

2.4 Intelligent approaches to load forecasting

Genetic algorithm (GA), fuzzy logic, artificial neural network (ANN) and support vector machine (SVM) approaches have been widely adopted in building energy prediction.

The GA approach is a powerful and robust tool for developing heuristics for large-scale combinational optimisation problems [13], [38]. Originally inspired by biological evolution, a genetic algorithm, which is a type of evolutionary computation, is a method for solving optimisation problems. The GA approach encodes a potential solution to a

specific problem to a chromosome-like structure and applies recombination operators to these structures to preserve critical information [14]. This method has been used to forecast power system load demand. The advantages of GA over traditional techniques are as follows [13].

- i) GA only needs simple information about the objective function and does not impose restrictions such as differentiability and convexity on the objective function.
- ii) It is a method with a set of solutions from one generation to the next, rather than a single solution. Therefore, it is less likely to converge into local minima.
- iii) Solutions are developed randomly based on the probability rate of the genetic operators such as mutation and cross-over. Thus, the initial solutions do not dictate the searching direction of GA.

However, the main disadvantage of GA is that it requires a tremendous amount of computation time.

Fuzzy logic makes decisions by simulating human reasoning. It is characterised by uncertainty and imprecision, and is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth – i.e., truth values between ‘absolutely true’ and ‘absolutely false’ [13], [39]. A FL methodology for combining information has been developed for spatial load predictions of the magnitudes and locations of future electric loads. These methods recognise that the load growth in different locations depends on factors such as distance to highway, electric poles and costs [40]. A fuzzy inference model for STLF in power systems, which uses a tabu search with supervised learning to optimise the inference structure (i.e. number and location of fuzzy membership functions) and minimise forecast errors, has been

proposed. Srinivasan & et al (1999) [41] has combined three techniques – FL, neural networks and expert systems – in a highly automated hybrid STLF approach with unsupervised learning. However, as FL can only provide simple outputs of either true or false in its final stage [14], the model is seldom to be used in building energy load prediction.

Support vector machines (SVMs) methods, introduced by Vapnik in the 1960s, are based on statistical learning theory [42]. SVMs, which contain a set of related supervised machine learning methods used for classification, have recently become an active area of intense research; and have been extended to regression and density estimation [43]. SVMs use the structural risk minimisation (SRM) principle to overcome the intrinsic limitations of ANNs. Support vector regression (SVR) in SVMs can be used for time series prediction, which is useful for problems characterised by non-linearity, high dimension and local minima. SVRs have been successfully used to solve regression problems such as time series modelling [44,45], financial forecasting [46], electricity load forecasting [47-54] and non-linear control systems [55]. Nevertheless, SVM models are complicated in operation, making the models less user-friendly [18].

2.5 ANN models for cooling load forecasting

The elements of an ANN model are similar to brain neurons, which can process massive amounts of information in parallel. An input and a weight value are connected to the model and the output of the model is a function of the summed value. The model is trained with respective datasets and learns its patterns as inputs. The model can be used for unknown and non-linear functions [12]. To predict building energy, the ANN models use historical data and external climate data in the training stage and then

deliver an output value that corresponds to the input data. ANN models are fast, robust and appropriate for nonlinear functions with sound learning ability [13].

A number of researchers and engineers have used ANNs for modelling and predicting in the field of building services engineering [56]. Some of the recent studies that have applied ANN models to building cooling load prediction are reviewed [56]. A multi-layer perceptron model (MLP) is a type of ANN model that was used to determine the total chiller plant power of a 42-storey commercial building in downtown Honolulu, Hawaii. The independent input variables consisted of climate data, and the model output was the chiller plant's power consumption. The input parameters used in the MLP model included dry bulb temperature, wet bulb temperature, dew point temperature, relative humidity percentage, wind speed and wind direction. It should be noted that none of the input variables were building-related. The hour of the day, when the data were collected, was also recorded to account for variations in occupancy throughout the day. However, data for the hourly power consumption in the chiller plant were not available. The total number of matching data items was only 121 out of 312 for the 13 days in the study period. A further study used [57] an ANN model to predict the energy savings of building equipment retrofits. The Levenberg-Marquardt back-propagation algorithm was used and the input layer included the weather variables at particular hours of the day. The weather data consisted of dry bulb temperature, dew point temperature, wind speed, wind direction, air pressure and visibility. The output was the hourly electricity measurements from the retrofitted equipment. Again, the input parameters used were not building-related parameters [58]. An ANN model was used to predict the pre-retrofit energy consumption of a building and this was then compared to the measured energy consumption of the retrofitted building. The input layer consisted of eight types of input data including weather factors such as ambient dry-bulb

temperature, humidity ratio, horizontal insulation and wind speed, and occupancy-related factors such as hour (00:00 to 23:00), a weekday/weekend binary flag (i.e. 0 and 1), chilled water consumption in the past hour and chilled water consumption in the previous hour [26]. An ANN model that used the Levenberg-Marquardt training algorithm was introduced to predict the heating/cooling load consumption. The input parameters for the ANN model included hourly weather data such as the outdoor temperature, relative humidity and set-point temperatures. The occupancy schedule, which was the key parameter in energy consumption, was used as one of the model's input parameters.

2.6 Conclusion

All the existing energy demand prediction models adopted in previous studies were either forward or inverse approach models. Load forecasting software programs of forward approach model, such as TRNSYS, BLAST and ENERGYplus, were introduced. They were performed based on the physical arrangement and assumptions in steady-state condition. The disadvantage of the forward models is the low accuracy on the energy demand prediction with the changes of external weather and the internal load of occupants in various situations.

This thesis focuses on the inverse approach models, which is data-driven. ANN models have been adopted in previous studies [28] and are found to be able to handle more complicated situations. ANN models are also more user-friendly comparing with other intelligent approach models like GA, SVM and FL. The application of ANN models on load forecasting will be introduced in the next chapter.

Chapter 3

Load forecasting using artificial neural networks

3.1 General

ANN models are created with data-driven and self-adaptive methods, which can be used to perform non-linear modelling without prior knowledge of the relationship between the input and output parameters. The models had been proven with universal function approximators [59] and could be used to predict nonlinear system behaviour by constructing the behaviour on the basis of historical system data. An MLP model supervises a neural network consisting of a number of neurons arranged in layers. Each neuron is a multi-input-single-output computational unit in the model. Neurons in one layer are interconnected with neurons in adjacent layers. The model is designed to determine the learning processes of the human brain and to simulate the relationships between input and output parameters based on historical system data. The MLP model has been widely adopted for a variety of energy forecasting problems due to its inherent and superior input-output mapping capability.

3.2 Model architecture

Backpropagation is a common learning algorithm in building services engineering. Backpropagation networks have been adopted to solve problems in many areas, such as load forecasting, fault detection and pattern recognition. The model training of backpropagation consists of three stages: the feedforward stage of the input training

pattern; the calculation; and the backpropagation of associated error and weight adjustment. A multilayer can only learn input patterns to an arbitrary accuracy and the weight adjustment is based on the generalised delta rule. Weight in a neural network is a stored segment of the information about the input signal. The neural network used in this thesis consists of three layers of interconnected neurons, as shown in Figure 3-1. Figure 3-1 indicates a typical configuration consisting of three layers and one output parameter.

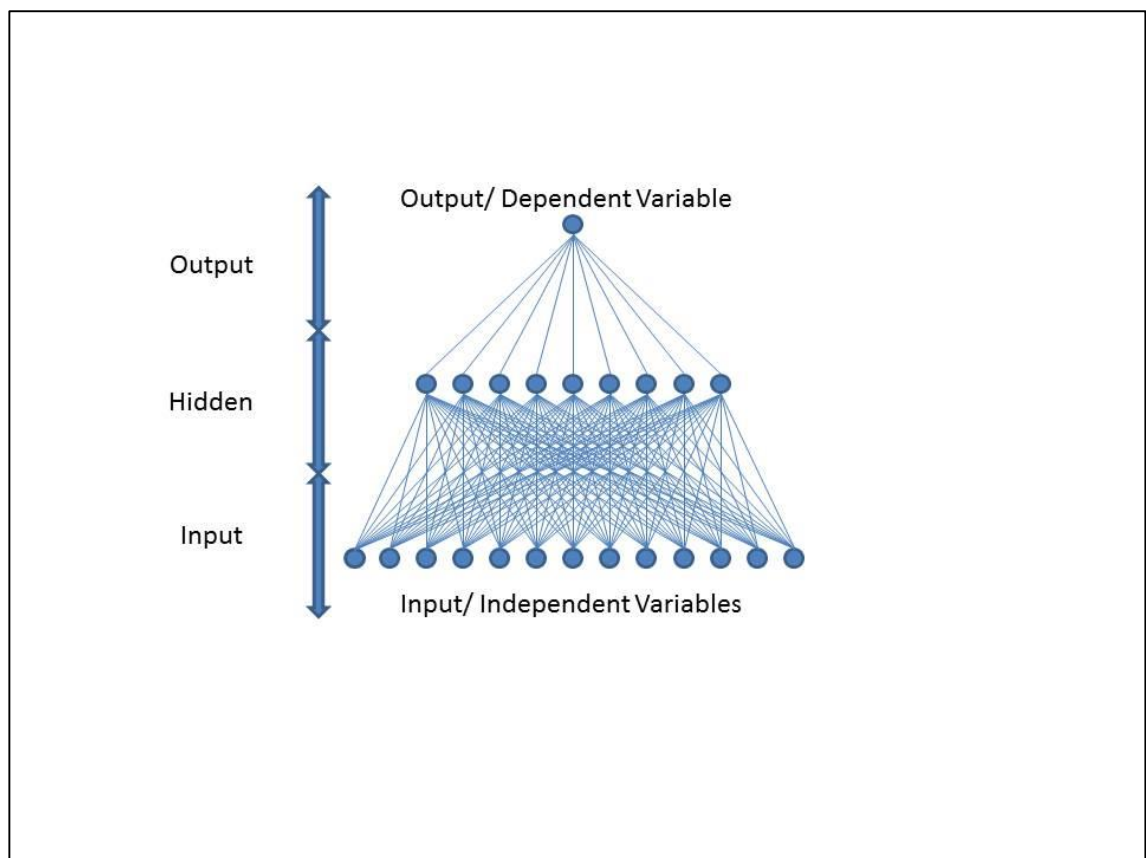


Figure 3-1 Typical artificial neural network computational structure

The backpropagation neural network consists of a multilayer, feedforward neural network with an input layer, an output layer and a hidden layer. The neurons in the hidden and output layers tend to connect with units whose output is always 1. The inputs are fed to the backpropagation net and the output obtained is either binary (0,1)

or bipolar $(-1,+1)$. The activation functions are any function that increases monotonically and is also differentiable. The generalised delta rule is implemented in backpropagation networks. The gradient descent method is adopted to minimise the total squared error of the output of the network. The layers are classified as input, hidden and output layers. The nodes in the input layer collect information from external sources, and the neurons in the hidden layers that act as the computational nodes in the neural network, transmit and transform the information from the input layer to the output layer. The output layer neurons transmit information out of the network. An MLP model with one hidden layer and a sufficient number of hidden neurons has been proven to be a universal function approximator [59]. This type of model is adopted in this thesis.

3.3 Input parameters

A number of input neurons corresponding to the number of parameters in the input vector are adopted to forecast future values. Although there is no systematic way to determine this number, the selection of this parameter should include all of the critical components that can significantly affect the future value.

3.3.1 Sources of building cooling load

Heat is generated from occupants, equipment, artificial lighting, IT devices and other sources inside buildings. Heat is then transferred from elements of the building envelope, such as the walls, windows, roofs and doors. They are the common sources of building heat load, as illustrated in Figure 3-2. Heat load can be classified into two categories: i) external heat load, which is a result of heat transfer from outdoor climatic conditions through the building envelope into the conditioned area; and ii) internal heat load, which is the heat released from the heat sources inside the indoor environment.

The relative proportion of internal and external heat load depends on the types of buildings, climatic conditions, building orientation, building design and so on. The total heat load of a building consists of sensible and latent heat load. The sensible heat load affects dry bulb temperature, whereas the latent heat load affects the moisture content of the air-conditioned area. Internal, latent and solar heat loads are considered in common design practice [60].

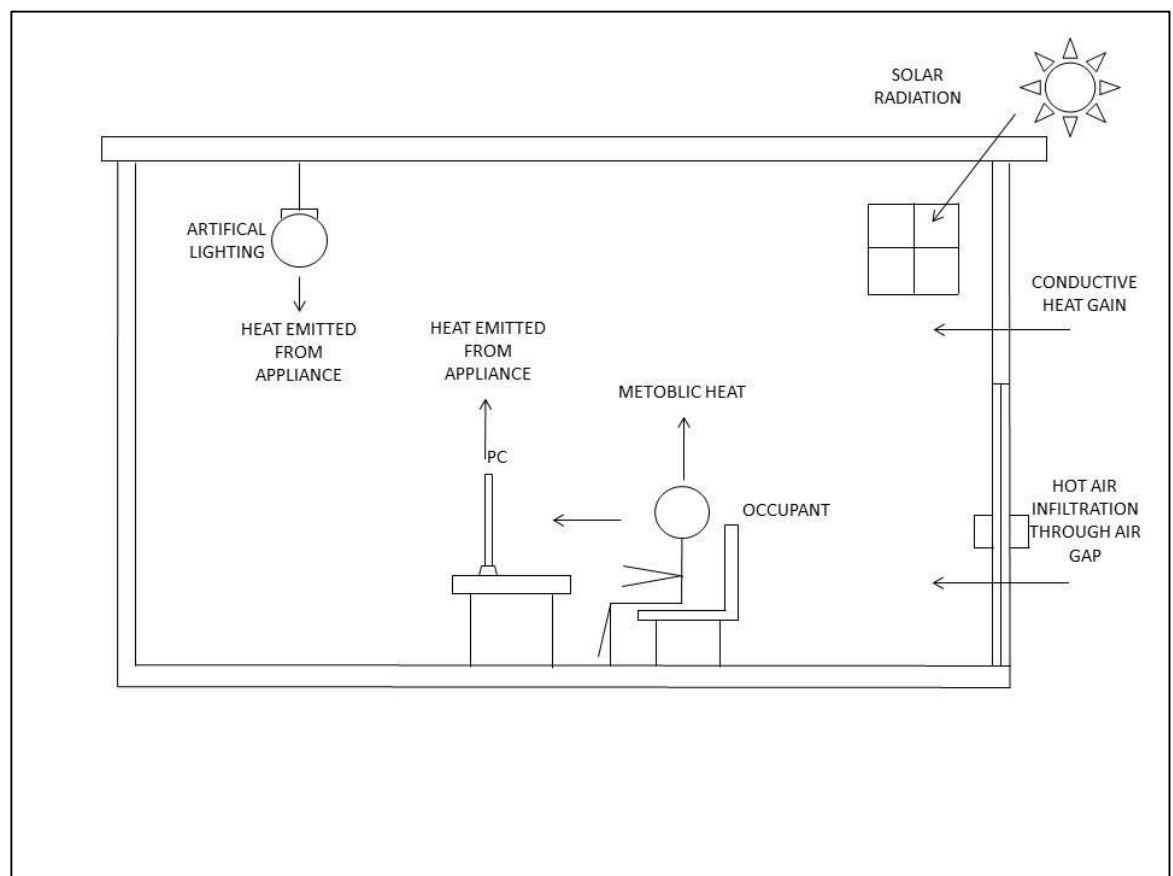


Figure 3-2 Sources of building cooling load

i) External factors

The heat transferred from external sources through the building envelope and partition walls is called the external heat load. The major sources of external heat loads [61] are listed as follows:

- a. sensible heat gain through an exterior wall or roof;
- b. sensible heat gain (solar or conductive heat gain) through fenestrations such as window glass;
- c. heat gain through partitions, ceilings and interior doors; and
- d. infiltration of outdoor air into the conditioned space.

Infiltration is the uncontrolled inward flow of outdoor air through cracks and openings in the building envelope due to the pressure difference across the envelope [61]. It is very difficult to estimate the exact amount of infiltration as it depends on several factors, such as the type of the building, the age of the building, indoor and outdoor conditions. Indoor and outdoor conditions include wind velocity and direction, outdoor temperature, humidity etc.

ii) Internal factors

Heat released from the heat sources inside the conditioned space consists of sensible and latent heat sources, which collectively make up the internal heat load. Heat is generated by occupants, equipment, lighting, appliances and IT devices. The sensible heat load is only generated by lighting. However, the conversion of sensible heat gain from the three factors discussed below to space cooling load is influenced by the thermal storage characteristics of the space.

a) Electric lights

The heat load generated from lighting includes light-emitting elements and ballasts. It is part of the sensible heat load and is depended on the types of installations inside the conditioned area and the devices used [61].

b) People

Occupants inside a building produce different levels of heat and moisture content when they engaged in different activities. The sensible and latent heat loads from occupants significantly affect the total heat load in a building. The sensible heat rate increases slightly with activities. Moreover, the latent heat rate increases significantly with a greater perspiration rate, which is used to maintain the body temperature. In such cases, additional heat and moisture are generated by occupants and this would significantly affect the short-term predictions.

c) Equipment and appliances

Common office equipment and appliances, such as microcomputers, display panels, printers, copy machines and communication tools convert electrical energy to heat energy. The heat released by these machines has increased remarkably in recent years [61]. The heat generated by the office equipment and other electric appliances contributes to the space surrounding as a by-product of their operation. All of the electric energy used for heat generation in office equipment would become a waste eventually [62]. The energy consumption of information technology (IT) equipment has increased remarkably in recent years. Electrical equipment has not only become one of the most important energy consuming sectors in office buildings (ranging between 20% and 40% of energy use), but also one of the main heat sources loading the cooling system.

3.3.2 Components that influence building cooling energy

There are 4 factors that play vital parts in the cooling load consumption of a building:

- i. the buildings' physical properties;

- ii. the equipment installed to maintain the desired internal environment, such as the ventilation and air conditioning system;
- iii. the building occupants' schedules, activities and behaviour; and
- iv. the outdoor climatic conditions [63].

The cooling load energy requirement of a building is governed by the complex interactions between the space cooling load, air handling system and the cooling plant. These interactions are influenced by time-varying parameters such as internal loads, heat gains through the building envelope, occupancy patterns, operating schedules and external weather conditions. The interactions are illustrated in Figure 3-3.

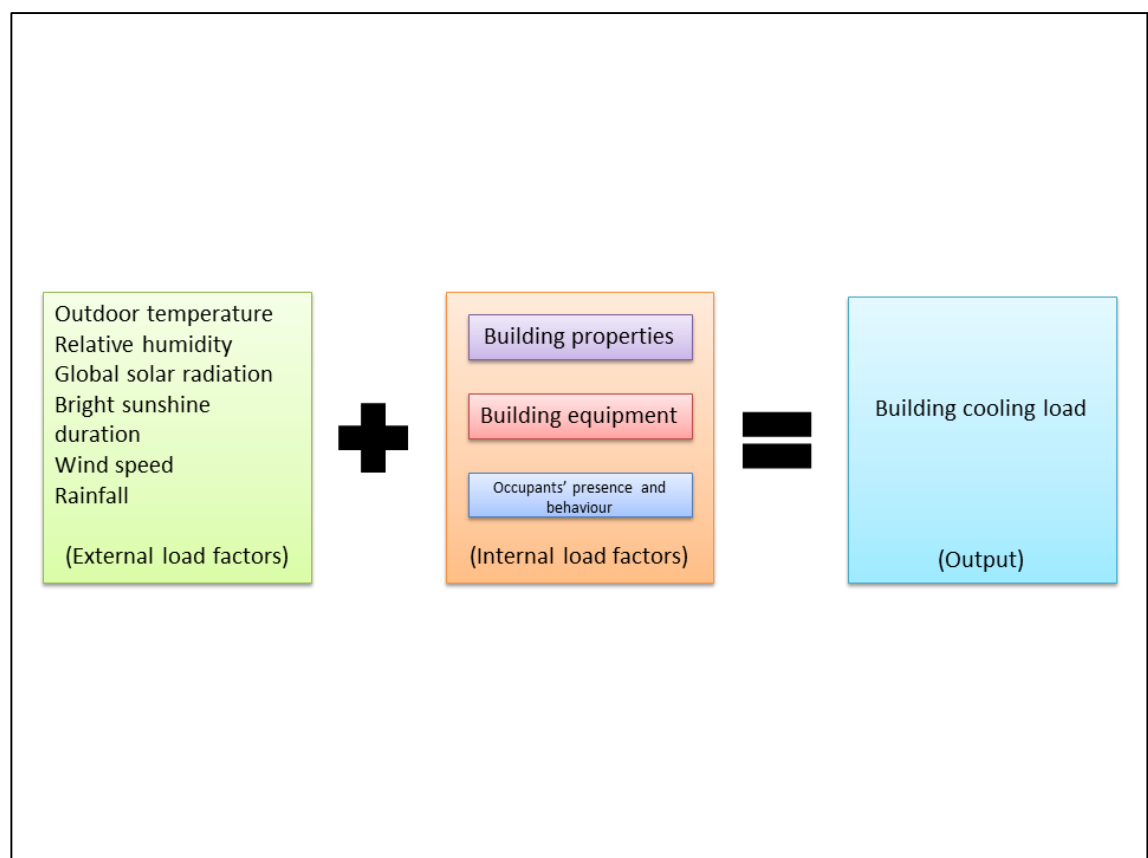


Figure 3-3 Components for estimating building cooling load

The building's physical properties (building envelope) are critical factors in the cooling load. They are assumed to be constant at a particular time and are not considered as

inputs in this simulation. The second set of critical factors (the equipment installed to maintain the desired internal environment) varies according to the changes in the occupant's presence and behaviour. The outdoor climatic conditions are represented by a number of external climate parameters, such as outdoor temperature and relative humidity. Therefore, representing building occupancy precisely is the most critical task in the development of building cooling load forecasting models. The input parameters for ANN models of cooling load prediction are shown in Figure 3-4.

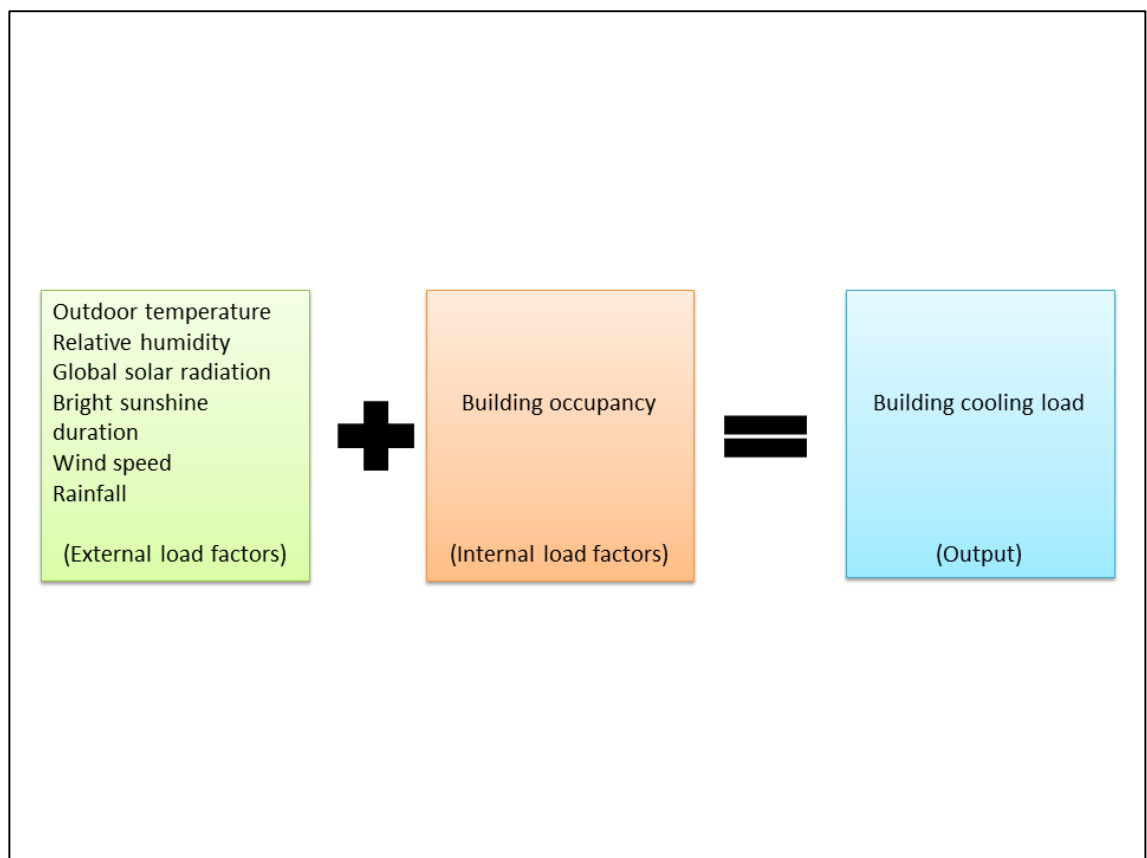


Figure 3-4 Input parameters for building cooling load prediction

Compared with the external climate parameters, it is relatively difficult to obtain data of time-varying parameters, such as internal loading, operating schedule and occupancy behaviour, since it is almost impossible to count the number of occupants and to determine their behaviour inside a building at a given time. As a result, the randomness

linked to occupants, i.e. the differences in behaviour between occupants and the time variation of behaviour, is a major contributor to the discrepancy between the simulated and actual energy demand of buildings [63]. This might be the reason that building occupancy has seldom been adopted as one of the input load parameters in ANN cooling load simulation models [26,35,56,57]. The situation becomes even complicated if the simulated building is a university building with a number of multi-use facilities, which operate 24 hours a day on working days and sometimes on weekends. Laboratories inside the university building may require a 24-hour air-conditioning load supply. The available timed input parameters are far from adequate for evaluating the dynamic changes in cooling loads in such buildings. Chapter 4 will discuss current approaches to determine building occupancy and accessing building occupancy using a stochastic occupancy model.

3.4 Hidden and output layers

i. Rule of thumb

The number of hidden neurons for a three-layer network is estimated by the rule-of-thumb as suggested in [64], which adopted the equation (3-1). N_h and N_p are the number of hidden neurons, and N_i and N_o are the number of input and output parameters, respectively.

$$\text{ii. } Nh = \frac{Ni+No}{2} + \sqrt{Np} \quad (3-1)$$

ii) Sensitivity test

A sensitivity test is also carried out to test the validity of the number of hidden neurons. It is determined by observing the change in the prediction error when the number of

hidden neurons is varied by ± 5 from the number of hidden neurons determined by equation (3-1).

3.5 Output parameter

The number of neurons in the output layer of an MLP model is equal to the number of system outputs. As the aim of this thesis aim is to predict the building cooling load under different internal and external load factors, the output of the MLP model is the building cooling load. Therefore, the output layer of the MLP model has only one neuron, representing the output of the prediction (i.e. building cooling load).

3.6 Network Training

3.6.1 Backpropagation algorithm

The backpropagation algorithm is one of the most powerful learning algorithms in ANN models. The training of all the patterns in a training datum is called the Epoch. The training set has to be a representative collection of input-output examples. Backpropagation training is a gradient descent algorithm. It tries to improve the performance of the neural networks by reducing the total error by changing the weight along its gradient. The error is expressed by the root-mean-square value (RMS), which can be calculated as follows:

$$E = \frac{1}{2} \left[\sum_p \sum_i |t_{ip} - o_{ip}|^2 \right]^{\frac{1}{2}} . \quad (3-2)$$

E is the RMS error, t is the network output (target) and o is the desired output vectors in the entire pattern (p) computed backward through the network. This error term is the product of the error function E , and the derivative of the activation function, and hence is a measure of the change in the network output produced by an incremental change in

the node weight values. Therefore, for the output layer nodes and the case of the logistic-sigmoid activation, the error term is computed as follows:

$$\delta_{pi} = (t_{pi} - \alpha_{pi})\alpha_{pi}(1 - \alpha_{pi}). \quad (3-3)$$

For a node in a hidden layer,

$$\delta_{pi} = \delta_{pi}(1 - \alpha_{pi}) \sum_k \delta_{pk} w_{kj}, \quad (3-4)$$

where the k subscript indicates a summation over all nodes in the downstream layer and the j subscript indicates the weight position in each node.

Finally, the δ and α terms for each node are used to compute an incremental change to each weight term via

$$\Delta w_{ij} = \varepsilon(\alpha_{pi}\delta_{pi}) + mw_{ij} (old), \quad (3-5)$$

where ε refers to the learning rate and determines the size of the weight adjustments during each training iteration. m is the momentum factor, which is applied to the weight change used in the previous training iteration, w_{ij} (old).

Both of these constant terms are specified at the start of the training cycle and determine the speed and stability of the network.

3.6.2 Early stop validation approach

Backpropagation [65] (BP) is a traditional training algorithm used in MLP models. It feeds back the prediction errors from the output layer to the input layer and the weights of the links between the neurons are adjusted according to the BP algorithm. Upon the completion of weight adjustments, a new prediction is carried out to evaluate a new prediction error for the next epoch of weight adjustments. These procedures are

repeated until a satisfactory prediction result is achieved. In this thesis, the early-stop validation approach is adopted to monitor and stop the BP training. Figure 3-6 illustrates the concept of the early-stop training approach. For model training and evaluation, the first 80% of the samples are used for network training, while the remaining 20% are hidden during the network-training phase and kept in the reserve as a testing set to evaluate the performance of the trained network. 80% of the samples used for network training are then further divided randomly into proportions of 75% and 25% as the training samples and validation samples, respectively. The training set is used to train the model with the BP algorithm, whereas the validation set is used to monitor and stop the BP training using the early-stop validation approach. The testing set does not play a role in the training of the MLP model. Upon completion of the model training, the testing set is used to evaluate the performance of the trained model. To prevent an ‘over-fitted training’, the intermediate-state trained model in every training epoch is applied to the validation set to evaluate the prediction error (i.e. the validation error). The network training process stops when the validation error reaches the minimum value. With no prior knowledge of the trend of the validation error, ‘early-stop’ training is adopted. This records the status of the model continuously during the training. When there is no reduction in validation error over a predefined number of epochs (in this thesis, the number of epochs selected is 1000), the model state with the minimum validation error is taken to be the trained model. Figure 3-5 illustrates the process.

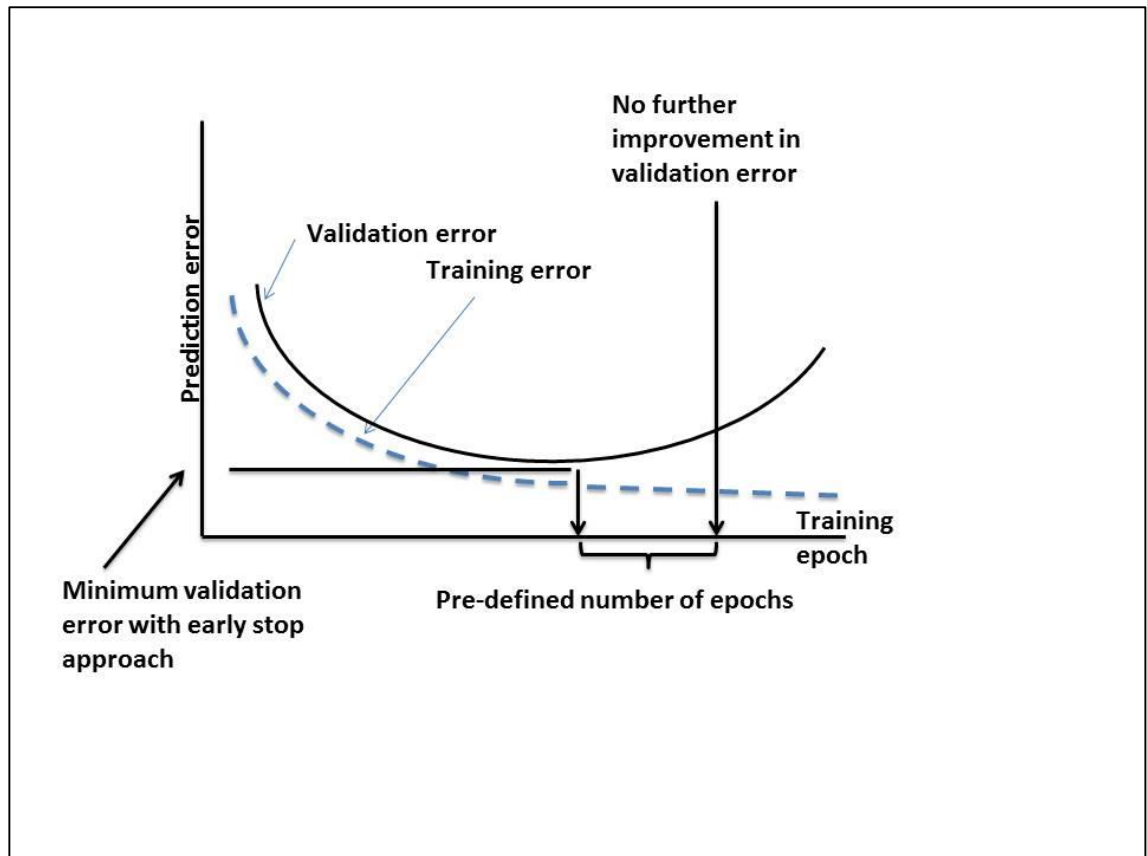


Figure 3-5 Early-stop validation approach of the backpropagation training process

In Figure 3-5, the early-stop validation approach stops the backpropagation training when there is no further improvement in the validation error over a pre-defined number of epochs after it has reached its minimum level. The intermediate state of the model with the minimum validation error is selected as the trained model.

3.7 Validation of results

3.7.1 Performance measures criteria

Upon completion of the network training, the trained MLP is applied to the testing dataset and the performance indices are evaluated by comparing the target values of the testing set and the values predicted by the trained model. The performance indices used in this thesis are the coefficient of variation (CV), mean absolute percentage error

(MAPE) and the coefficient-of-correlation (r), as defined, respectively, in equations (3-6), (3-7) and (3-8), where N is the total number of samples and $\{t_i p_i\}_{i=1}^N$ are the target values and the predicted values, respectively. $\bar{p}$ and $\bar{t}$ are the mean value of the predicted values and the mean value of the target values, respectively.

$$CV = \frac{\frac{1}{N} \sqrt{\sum_{i=1}^N (p_i - t_i)^2}}{\bar{t}} \times 100\% \quad (3-6)$$

$$MAPE = \frac{\frac{1}{N} \sum_{i=1}^N |p_i - t_i|}{\bar{t}} \times 100\% \quad (3-7)$$

$$r = \frac{\sum_{i=1}^N (t_i - \bar{t})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^N (t_i - \bar{t})^2 \sum_{i=1}^N (p_i - \bar{p})^2}} \quad (3-8)$$

3.7.2 Training schedule

It should be noted that a random process is normally involved in the network training of an MLP model, especially when the available samples are divided into training and validation sets. It is thus possible for the random process to result in ‘fortuitous’ samples showing that the evaluated performance indices are good. Instead of reporting only the best simulation result, a less-prejudiced statistical approach is adopted to minimise the effect of randomisation. The network training and performance evaluation process are carried out 100 times. The 100 results are statistically analysed by evaluating the mean of the results and the limits of the 95% confidence intervals. This approach, which differs from those in previous studies [56,57], assesses the performance of the ANN model when this less-prejudiced approach is taken, as illustrated in Figure 3-6.

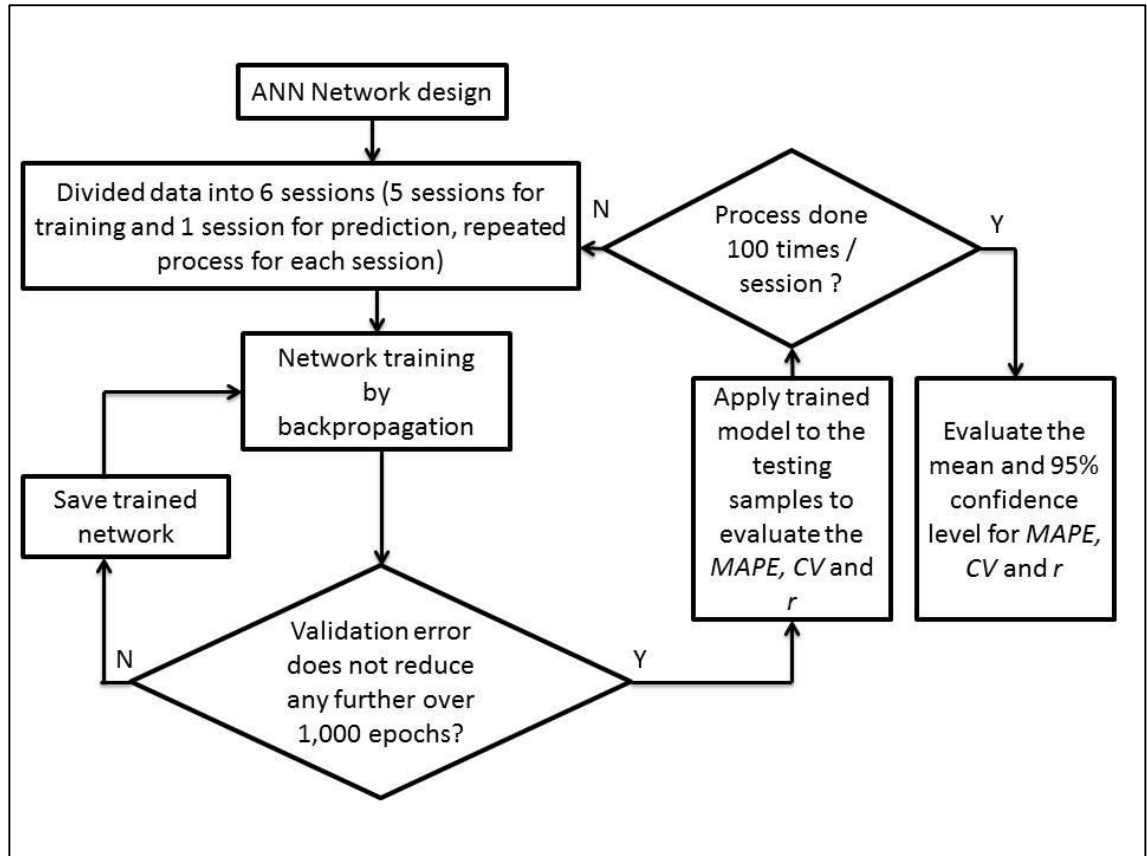


Figure 3-6 Training scheme designed to minimise the effect of the random data extraction process.

3.8 Conclusion

The application of ANN models on load forecasting is introduced in this chapter. The model architecture includes an input layer, a hidden layer and an output layer. The input parameters can be divided into two major categories: internal and external cooling load. External cooling load is affected by outdoor climatic change, whereas internal cooling load is affected by the occupants' presence and behaviour. The hidden layer is determined by the rule of thumb and the sensitivity test suggested by the study [64]. The output layer represents the prediction output, which is the building cooling load. The backpropagation algorithm and early stop validation approach are adopted in this ANN model and also in the study mentioned [65]. The evaluation is based on the performance

indices: the coefficient of variation (CV), mean absolute percentage error (MAPE) and coefficient of variation (r). The training schedule and evaluation algorithm with the limits of 95% confidence intervals are introduced and discussed [56,57]. The importance of internal cooling load and the approach used in this thesis – building occupancy modelling will be described in the next chapter.

Chapter 4

Current building occupancy modelling

4.1 General

The building cooling load is affected by four external load factors and the internal load. The external load factors include solar radiation, temperature, transmission load and ventilation/infiltration load [66]. The reliability of the building cooling load prediction depends on the accuracy of the input data. In fact, significant improvements have been made in techniques for estimating building geometry, building material properties and external weather conditions. However, the improvement in techniques for collecting accurate building occupancy information (i.e. people's presence and behaviour in buildings) is relatively slow [21].

The presence of occupants and their activities have significant effects on the performance of buildings, including its energy efficiency and indoor environment [21]. The indoor environment is affected by the occupants' body temperature and the pollutants they emit, such as water vapour, odours and carbon dioxide. These factors vary according to the activities occupants are engaged in. As a result, the heat load adds to the internal load and this is reported to influence the human comfort level [63]. The accumulated heat and pollutants generated by occupants have adverse effects on the indoor environment. To ensure their personal comfort level, occupants interact with the building systems by adjusting the warming, cooling or ventilation settings and the lighting systems. The occupants in an office building may use diverse electrical

appliances, which tend to increase the internal heat gains and the electricity consumption [68]. Recent studies have been focused on investigating the relationship between occupants and their adjustments to the building systems.

4.2 Interactions between building occupancy and energy consumption

The most complex processes in buildings energy demand prediction are related to human behaviour because people's behaviour are intrinsically unpredictable. As a result, it affects the indoor environment and energy demands in ways that have important implications for building energy consumption [68].

People enter and leave a building in different entrances and at different timing. They can work from 9am to 5pm on weekdays, or they can have flexible working hours or even work on weekends. They can also work overtime, take vacations and even sick leaves in stochastic rhythms [68]. The influence of occupants on buildings energy consumption can be classified into six types of interaction, as illustrated in Figure 4-1 [63].

People produce heat and pollutants such as carbon dioxide, water vapour, odours, etc. that affect the indoor environment. Windows, window blinds, luminaries, radiators, fans and other appliances can be adopted by building occupants, in order to maintain the level of comfort and the quality of indoor conditions [69]. For example, an occupant may switch the lighting systems on or off, or adjust window blinds and the thermostat to create a visual- and thermal- comfortable environment. All of these interactions in turn affect the building's HVAC system and the thus building energy consumption. If such changes do not produce a satisfactory outcome, people will take further actions to restore their expected level of comfort [70].

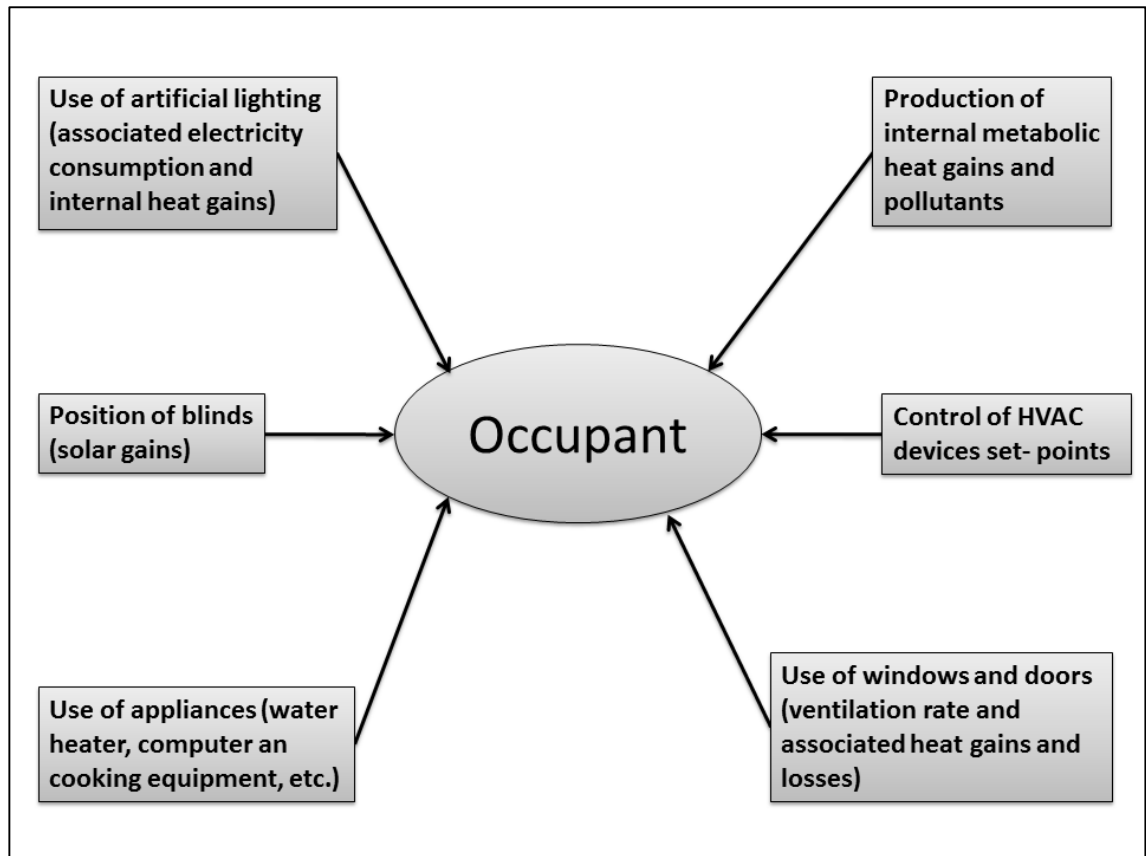


Figure 4-1 Interactions between occupants' presence and behaviour and the internal load

4.3 Modelling building occupancy

Current approaches do not have the level of sophistication required to reflect the complicated relations between people's passive and active actions and building performance. Only rough approximations are available according to the general building type (residential or commercial), the environmental system (open area or air conditioned area) and even the organisational information (working hours) [21]. In fact, previous studies have shown that occupants' behaviour and interactions have a great effect on the energy consumption of lighting, electrical appliances, space heating, cooling, ventilation and building controls [63]. Over the past 40 years, the simulation of thermal processes

in building energy performance has become matured [71]. Simulations have shown that occupants' behaviour is more influential on building energy performance than the thermal processes within the building façade. Therefore, accurate simulation results can only be obtained if occupants' behaviour is predictable and follows a routine pattern. Current simulation approaches adopt 'standard profile and diversity factors' to represent these internal load factors. Alternately, some researchers have developed 'empirically-based models of people's presence and actions in buildings' to simulate building occupancy.

4.3.1 Standard profile and diversity factors

The building simulation models assume that there are a number of fixed metabolic heat generators, which passively experience the indoor environment [72]. Such models ignore the fact that office users are not passive and static. In fact, occupants influence their immediate working environment by operating the artificial lighting systems, the window blinds and glare protection devices and/or the air-conditioning systems.

Diversity factors are adopted in some simulation models to account for the presence and behaviour of occupants [73]. Such models represent the internal heat gains generated by occupants, office appliances, artificial lighting and consider the resulting cooling load in the simulation results. These models may also categorise buildings into residential or commercial types of building. The size and composition of a household can also be represented.

Weekdays and weekends generate two different profiles of energy consumption, due to the variations in metabolic heat gain, receptacle load and light load in different hours. These differences may be identified in historical data collected from monitored buildings or simply by common sense or national guidelines [67].

A standard profile or diversity factors approach can be used to assess the effect of occupants' behaviour on building performance independent of climatic conditions and operational schedule. However, one of the significant disadvantages of using diversity factors or other similar profiles is that they are derived independent of meteorological data [73]. Therefore, energy modelling simulations using standard profiles provide results of varying accuracy. In [74], the researchers found that in various US locations, the use of average profiles overestimated the electrical energy savings, and that demand reduction through occupancy-sensing controls underestimated the heating loads.

4.3.2 Empirical models of occupants' presence and actions in buildings

i) Background

The presence and actions of building occupants affects the energy efficiency and indoor environment. Researchers have investigated how building occupants interact with buildings' environment control systems. They have attempted to determine the relationship between occupants' actions and monitored indoor or outdoor environmental conditions such as dry and wet bulb temperature or solar radiation [21]. The following studies have offered valuable insights into the circumstances and potential triggers of occupancy control actions in buildings.

ii) Lighting

Occupants switch on the lights when they are in the office [76]. Moreover, they rarely switch off them until the last person has left. Generally, the lightings are switched on at the beginning of the workday and off at the end of the workday. In most cases, the lights are either all switched on or all switched off. Similar functions were subsequently suggested by [76] [77].

The first version of the LIGHTSWITCH stochastic occupancy model was developed in [79]; it used field data to predict the arrival, departure and temporary absence probabilities of individual occupants in office environments at five-minute intervals. The relationship between the propensity toward turning off the lights and the length of absence from the room was established [80]; in general, people are more likely to switch off the lights when they leave the room for prolonged periods. As suggested in [76], there are two important factors, which influence this tendency. They are i) occupants switch on the lighting during working hours and seldom switching it off when they leave the room temporarily; and ii) occupants switch on the artificial lighting only when the luminance level in the room is low.

The stochastic model of occupant presence in the LIGHTSWITCH model [77,79] considers the arrival and departure times of occupants and fine-tunes the standard profiles accordingly. Poisson distributions were adopted in [81] to formulate daily occupancy profiles for single-occupied offices and lighting patterns were used to represent the typical periods of presence and absence in a room. In [82], a sub-hourly occupancy model based on Reinhart's algorithm [78] was integrated into ESP-r to investigate the lighting use. Also, Bourgeois developed a method for integrating sub-hourly stochastic processes into Reinhart's building simulation. The technique is quite applicable in estimating routine occupancy patterns, i.e. a single occupancy office or a classroom, but it may not be suitable for the complicated occupancy patterns found in many environments.

iii) Blind operation

Past studies have shown a strong correlation between the operation of venetian blinds and solar radiation intensity [83]. A strong correlation between blind operation rates and

building orientation was found [84,85]. Blinds were operated more frequently on southern facades. In [85], a specific pattern was observed between blind operation and the incident illumination on the façade. The researchers concluded that occupants largely ignore short-term irradiance dynamics. Beyond a certain threshold of vertical solar irradiance on a façade (50 W.m^{-2}), the deployment level of shades was proportional to the depth of solar penetration into a room. This conjecture was corroborated in [77]. Once closed, shades seem to remain deployed until the end of the working day or when visual conditions become intolerable. Other researchers [84] observed that there was a rather low rate of blinds operation throughout the day, implying that occupants' perception of solar irradiance was in long term. Study also suggested that occupants manipulate shades mainly to avoid direct sunlight and overheating [86].

iv) Window opening

A model based on Markov chains has been established to model the random opening of windows by occupants [87]. Records and observations were made of the operation of 21 south-facing single offices in Freiburg, Germany. The data were used to correlate with parameters such as indoor and outdoor temperatures, occupancy rate, solar radiation and window status [88]. The results were as follows.

- i) In summer, 60% to 80% of smaller windows were opened, whereas only 10% were opened in the winter.
- ii) A high window operation rate was found in swing seasons, i.e. spring and autumn.
- iii) A strong correlation between window opening rate and outdoor temperature was found.

- iv) 80% of small windows were opened and 60% of large windows were tilted when the temperature was above 20°C.
 - v) The window operation rate was found to be high in the morning (09:00) and afternoon (15:00).
 - vi) The window operation rate was found to be high at the time of occupants' arrival and departure.
 - vii) Windows were closed at the end of the working day.
 - viii) There was a strong correlation between the percentage of open windows and the time of a year, outdoor temperature and building occupancy patterns.
- v) Lights and blinds

A model for predicting occupants' personal control of electric lighting systems and window blinds was established in [72] and [77]. The prediction model was based on stochastic functionality and dynamic response to short-term changes in luminance levels and occupancy patterns. The presence of occupants was represented by a simple term such as arrival and departure. Different models were introduced to predict whether lights were switched on upon arrival and switched off upon departure. A stochastic simulation model was established in [89] to predict the behaviour of occupants in buildings. It also determined the correlation between outdoor temperature and the windows open/close rate, heating and window blinds. The correlation between solar radiation intensity and window blind operation rate has also been studied. In [78], a stochastic model of arrival and departure was adopted to formulate the interaction of occupants with window blinds and lighting systems. The user control actions pertaining to lighting and shading systems under different indoor and outdoor conditions occurred at the same time [21]. Individuals' interactions are difficult to predict, but control-

related actions, trends and general patterns of building occupants can be determined from long-term observational data [21].

4.4 Current ANN load forecasting model for measuring building occupancy

4.4.1 Factors affecting the internal load

Time and historical parameters are common features of internal loading models. Researchers have tried to identify the best internal loading parameters for modelling short-term load prediction. However, loading data for model load estimation and forecasting cannot be retrieved directly. Rather, load modelling and prediction require basic knowledge of the factors that influence the building load [90]. Both external and internal factors are used in load forecasting. The three main types of parameters, which have been adopted for prediction models [91], are listed as follows:

i) Time factors

Time factors are described in units ranging from time of day, day of week and time of year. A building load varies over time according to occupants' activities. There is usually a high building load during the daytime, and a low building load at night. Building loads on weekends are lower than on weekdays. There are two peaks of building loads during the daytime in an office building: one in the morning and one in the afternoon. This piece of information is important for short-term load forecasting. Therefore, hourly indicators are important in building load prediction. The varying load profiles between weekdays and weekends has also made day indicators important. In [92], an ANN model was used to predict short-term load. The internal factors used were a day-of-week indicator and the current hour. In [93],

an ANN model was applied for sub-station load forecasting. Two different day types, representing weekdays and Sundays, were adopted in the load prediction model, and each day was divided into 24 data files as an hour indicator. In [94], an ANN model was used to model and predict building energy. The internal factors were derived from an office located in Athens, Greece. The hourly energy was predicted by an ANN model with different inputs parameters including hour type and previous six numbers of hourly energy output. In order to investigate the strong effect of building occupancy on the building load, the date data were divided into weekday, weekends and holiday datasets. In [95], the abovementioned factors were used in an ANN model to predict energy consumption. In [96], hour and day indicators and weather-related factors were used to efficiently predict short-term load forecasting of a local load, in Mascouche, Canada. In [97], the month of the year and the day of the week were used as input parameters to develop an ANN model for evaluating the Hellenic daily electricity demand load.

ii) Historical load factors

Historical load factors include previous load consumption and curve patterns. Building loads have periodic patterns that vary from hour to hour and from day to day; these load can be predicted by models with good accuracy. [98] conducted short-term loading predictions for building loads. The temperature and historical temperatures and loadings were used as input parameters in the model used to predict power consumption. The model also considered the load patterns from previous days and weeks. Researchers used an ANN model to predict the power distribution over the short term and the medium term [99]. The input parameters for the ANN model included the daily consumption of the previous five days and

the climatic data. To predict the consumption at a particular hour, the model used the load consumption of the previous 3 hours on the same day [100].

iii) Hybrid factors

An ANN network is a black-box model that does not require the user to know the relationships between the inputs and output. It can also map nonlinear inputs to an output. Time factors and historical load data are commonly used in ANN models to predict short-term loads. An ANN model can predict the next 24-hour load profile by adopting the load profiles of the two previous days and the day of the week as input parameters [101]. To obtain the average hourly load consumption, a model averages the previous three weeks' load consumption for the same day type, and predicts one day ahead hourly load consumption. The model uses the average load consumption for the previous hour obtained by averaging the data for three weeks. [41] predicted one day ahead hourly loads using average hourly load values obtained by averaging the previous three weeks' load data for the same day type and the average load value for the previous hour obtained by averaging the data for three weeks. The multi-layer ANN model developed for power systems load forecasting in [102] adopted a 24-hour load profile of two days and the day prior to the forecast day, day of the week and holiday as input parameters. The load consumption for the previous 24 hours and the day of the week were used as the input parameters of an ANN model for predicting the short-term electricity load in [103]. In [104], an ANN model was developed to predict hourly building energy consumption based on feedback that measured time factors (the hour of the day and the day of the week) and the current load. Date, time, weather data and six indices representing special events, such as examinations, public holidays etc., and historical hourly load consumption data were used as the input parameters, in order to predict the next 24 hours load consumption by the ANN model [105]. Past electric load

consumption and day type were adopted as the input parameters in the ANN model for forecasting electricity load consumption in [90].

4.4.2 Limitations of the current ANN approach to determine building occupancy

In previous studies, the building operation conditions have been treated as fixed conditions. The presence and interactions of occupants are assumed to be deterministic in current simulation programs, which use predefined schedules and day types that assume that occupants' arrive and depart according to a fixed schedule. Great differences are found in actual conditions due to the variety in occupants' behaviour. In [106], the occupants were away from their offices for 25% to 30% of the nominally occupied hours of the day. Adopting time and historical data as internal load factors is sometimes acceptable and may produce reasonable results. However, such an approach is likely to have low accuracy in some conditions, for example, the dynamic short-term load prediction of a university building with different types of use on each floor.

4.5 Conclusion

Current building occupancy modelling is illustrated and discussed. The interactions between building occupancy and the cooling energy consumption in the building have been investigated. Standard profile and diversity factors are considered to exert an effect on occupants' behaviour and building performance, independent of climate conditions and operational schedule. Low accuracy was found when the standard profiles are applied in various locations [74]. Empirical models using lighting, blind operation and window openings are described. Studies have showed that there is a correlation between the ANN models adopted and energy consumption. Current ANN load forecasting model for measuring building occupancy is discussed. The model includes time factors,

historical load factors and hybrid factors. Studies using the standard profile, diversity factors and empirical model are discussed and the limitations of the models are suggested. Therefore, a novel approach to determine building occupancy for cooling energy consumption prediction is developed in order to improve the accuracy of current models. The approach will be discussed in the next chapter.

Chapter 5

A novel approach to determine building occupancy

5.1 General

Building simulation has become an essential part of building design due to today's higher performance requirements and increasing design complexity. Building energy simulation models have been used for over 40 years and there are many well-developed models of thermal processes [71]. However, building operational characteristics, which consist of occupants' presence and behaviour, have a stronger effect on buildings' energy loads than the thermal processes of building envelopes. Thus, building energy simulation models are straightforward if a building's operational characteristics are routine and predictable. Four types of factors, which play a vital part in building energy consumption [21], are listed as follows:

- i) physical properties, including type of building, location, orientation, envelope materials, etc.;
- ii) electrical appliances, and the heating and cooling systems installed and equipped to achieve a desired indoor environment and condition;
- iii) outdoor climatic conditions such as dry and wet bulb temperatures, solar radiation, wind speed, etc.; and
- iv) building operational characteristics, including occupants' presence and interactions with the building systems.

In [67], it was found that the effects of physical properties and systems could be accurately measured and calculated, and where accurate climatic data could be retrieved when simulation results were accurate. However, current approaches used to measure occupancy rate are far from accurate [21]. Operational information (e.g. working hours or holidays), building type (e.g. commercial or university) and HVAC equipment (e.g. air conditioning or heating system) have been used to provide general information about the effects of human presence and interactions in buildings [21]. It is suggested that the interactions between occupants and buildings can exert unpredictable effects on building energy performance and the energy demands of the buildings [68]. However, occupants' arrivals, departures and movements between floors are highly variable. People may work overtime or be absent from work due to illness or vacations. Occupants may not spend all of their working hours in their offices due to lunch and breaks, and they may leave their seats anytime in a day [63]. As a result, a simple model is inadequate to represent occupants' presence and interactions in a building.

Traditional ANN energy simulation models are suitable for predicting performance in buildings with routine operational characteristics [71]. However, in complicated cases, low accuracy might be resulted. Therefore, researchers have recently started to examine occupants' presence and interactions; and investigate their relationship with building energy performance. Studies have established a correlation between control actions, such as switching lights on/off, opening/closing windows and operating window blinds, and measureable indoor or outdoor environmental parameters such as temperature or solar radiation [21]. These studies showed that although the interactions of individual occupants are difficult to predict, trends and patterns in the control-related behaviour of groups could be mined from large datasets [21]. Based on this observation, a novel

approach to determine dynamic occupant's presence and behaviour models was developed in this thesis.

5.2 Current approaches

Current simulations adopt the 'diversity profiles' to represent occupant presence [73]. These are based on assumptions about the building energy and the cooling load of the internal gains generated by people, equipment and lighting in general categories of buildings. Residential, commercial and industrial buildings have their own specific diversity profile. Different profiles are created for weekdays and weekends in commercial building models. A number between zero and one is estimated as a multiplier of some user-defined maximum load for lighting and equipment. Many energy standards and codes provide typical diversity profiles for performance-based compliance demonstrations [113]. An overview of existing methods for deriving diversity profiles was provided in [73]. Current approaches using the diversity profiles were also discussed in the ASHRAE Research Project 1093 [114]. One project compiled a library of schedules and diversity factors based on measured electricity use data that could be used in energy simulations and peak cooling load estimations of office buildings [115]. Another project derived multiple sets of diversity factors from measured lighting and receptacle loads in 32 office buildings [116]. These studies reported that when the use of lighting and office equipment was regular and predictable, the building energy use was regular. These parameters were considered to be independent of external weather conditions, which was typical of diversity profile approaches [115]. They only considered the pattern of occupancy and daylight illumination in the core and perimeter zone of a building; hence leading to less accurate results. Applying this simulation approach to a complex university building was likely

to obtain even more inaccurate results. In a study of identical residential units, researchers found that the energy use by different occupants can vary from 200 to 300% [117]. A study of various US locations revealed that the use of average profiles overestimated energy saving, whereas demand reduction through occupancy sensing controls underestimated heating loads [118]. Researchers found that occupants respond differently to various sudden environmental stimuli, which trigger abrupt manual changes in window-blind settings and artificial light use, and these in turn affect electrical energy use and demand [72]. Similarly, many studies have reported that the use of diversity profiles can lead to significant errors when they are applied to control strategies that are sensitive to short-term variations in occupancy [82].

5.3 Occupants' presence and activity

5.3.1 Background

Buildings are becoming more complex in structure, control and operation; at the same time, users' requirements are increasing. Occupants' presence and interactions are important factors in a building's energy balance and they affect the indoor conditions and energy demand [22]. Building energy use is highly correlated with a building's operational and space utilisation characteristics and the behaviour of its occupants. As occupants always want to improve or maintain the indoor environmental conditions such as thermal comfort level, indoor air quality, luminance level and noise levels, they use control devices frequently [19]. Traditional ANN building simulation models use an average or fixed profile to represent the metabolic heat generated in an indoor environment [72] and do not consider the effects of occupants' changes to artificial lighting systems, window blinds and thermal control systems. The internal heat load of a typical office building is an important and inevitable parameter in building energy

performance simulation, and it is correlated with occupants' presence and interactions. However, this parameter is not fully considered in many approaches, although thermal process estimation for building energy performance has greatly improved in recent years [71]. It is believed that current simulation models can obtain accurate estimations if the use of a building is predictable and routine. In these cases, occupants' presence and interactions can be assumed to be static, as in current building simulation models [19]. Both the occupants' presence and interactions tend to be handled in entirely deterministic ways [22]. The occupants are assumed to arrive and depart at the same time and under regular time schedules or patterns. For example, the use of artificial lighting is usually 09:00 to 18:00 for normal office working hours. However, occupants' arrivals, departures and movements inside the building are stochastic. They may need to work overtime, take sick leaves or vacation leaves [22].

5.3.2 User interactions with environmental control systems

Windows, window blinds, artificial lighting, fans and other similar devices are typical control devices in buildings; they can be operated by building occupants to maintain the desired indoor environment [69]. It has been observed that occupants seldom switch off the lights if they only leave the office for a short time. Moreover, users' changes to the lighting and shading systems reflect the indoor and outdoor environmental conditions under which those actions occurred [21]. Previous studies have shown that the actions of individual occupants are difficult to predict. However, control-related trends and patterns of groups of building occupants can be extracted from long-term historical data [21]. As human well-being and productivity are strongly affected by the as-built environment, providing comfortable room conditions is a vital part of office building design [88]. Thermal comfort can be achieved by adjusting the ventilation or air

conditioning settings in the building. Occupants interact dynamically with the building systems at various times and in various numbers. Occupants produce metabolic heat gains and pollutants, which cause uncomfortable indoor conditions and prompt them to adjust the set points of the HVAC systems. Similarly, occupants take action to remove excess heat loads and pollutants. Altering the window blind position maintains comfort by reducing solar heat gain and this reduces the extra cooling load. Internal heat gain is increased by switching on the artificial lighting to maintain visual comfort. Occupants also switch on electrical appliances such as desktop computers, copying machines, office equipment, etc., which increases the internal heat load in the building. Finally, occupants' arrivals, departures and movements change all of these the energy demand over time. The above dynamic interaction of sub-systems in an air-conditioned university building is illustrated in Figure 5-1.

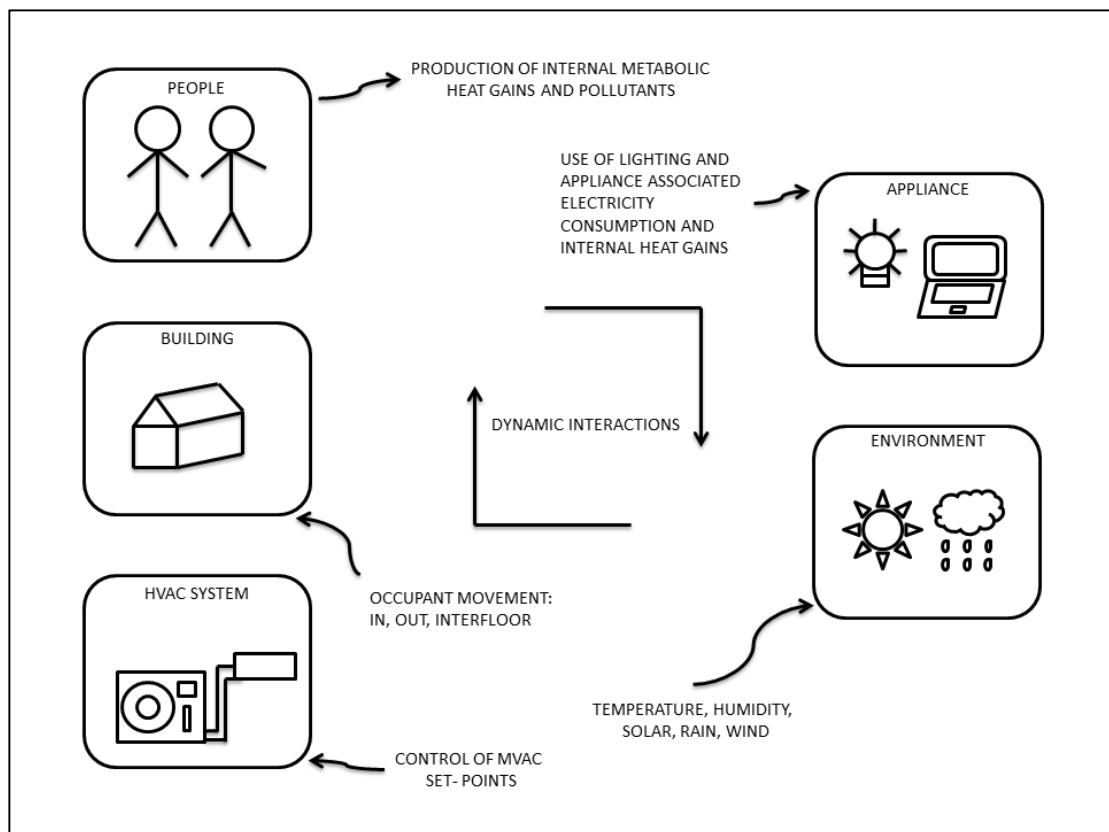


Figure 5-1 Dynamic interaction of systems in an air-conditioned university building

5.3.3 Stochastic processes of occupants' presence and activity

Various types of building systems consume energy in a building, including the equipment that provides building services, such as heating, ventilation, air conditioning and auxiliary production of electricity or hot water. The climatic conditions and occupants' presence and interactions are also major factors in energy consumption [63]. The aggregated presence of occupants is, as discussed above, an essential factor in this process. The presence of occupants is therefore an important input parameter of energy prediction models. The model for occupants' presence is central to the family of stochastic models [67,68]. The presence of occupants and interactions between the occupants and a building are difficult to predict. Improved software tools provide sophisticated simulations of physical properties and equipment installed in the building, yet the estimation of building occupancy relies on simple fixed profiles of typical building occupancy. These have not considered the fact that occupants would emit metabolic heat, water vapour, carbon dioxide and odour, etc., they will adjust the control devices to maintain a comfortable indoor environment.

5.4 Measuring building occupancy via stochastic processes

To maintain the comfort level, occupants operate windows, fans, window blinds, artificial lighting, air conditioning and associated installations. Several studies that investigated building occupants' interactions with building control systems and devices have provided a comprehensive picture on control-oriented user behaviour in buildings [72,75,69,108]. The studies have identified the following features of occupants' presence and interactions in buildings.

5.4.1 Pollutants emission and dilution

Each occupant directly affects the indoor environment by emitting metabolic heat and pollutants such as water vapour, carbon dioxide, odours, etc. Occupants exhale carbon dioxide (CO₂) and a high concentration of carbon dioxide may negatively affect human health. Inside the buildings, ventilation fans are used to intake fresh air into the building to reduce CO₂ concentrations. However, CO₂ can also be diluted by air leakage and windows opening. Traditional HVAC systems provide a fixed amount of fresh air through primary air handling units (PAU). In modern buildings, a variable volume of fresh air can be provided by variable speed primary air handling units, which are controlled by CO₂ sensors. This saves energy. High occupancy rates can lead to a high level of CO₂ and in these conditions, the PAU adjusts the fan speed using a variable speed drive and supplies the required amount of fresh air into the building. Such systems vary the volume of fresh air supplied to the air-conditioning system according to building occupancy. Thus, the fresh air supply rate can be used to represent the building occupancy rate of the building.

5.4.2 Metabolic heat gain and removal

Occupants influence and modify indoor environments. They open windows, switch on cooling, heating and ventilation equipment, and adjust window blinds and artificial lighting. They can also improve their thermal comfort level by adjusting the variable air volume (VAV) in an air conditioning system. VAV systems are becoming more popular as they can save energy better than the traditional constant volume systems. Low noise levels and no condensation are also the advantages of VAV systems. The basic principle of a VAV system is that it dynamically supplies adequate air volume to each room under different load conditions, and controls the total air volume supplied to each VAV

box through two main components: air handling units (AHU) and VAV boxes. It is difficult to measure the power of each VAV box in a building, as the number of VAV box is enormous in a single building. It is not practical to install so many electricity meters. However, the total power consumption of AHUs can be recorded and used to illustrate the load profile of the air supply required by different cooling load conditions. The presence and departure of occupants influences the supply air volume and this in turn affects the total supply air volume of the AHUs. Hence, the total power consumption of the AHUs varies with different air supply volume conditions and can reflect the occupants' presence, departures and activity intensity in the building.

5.4.3 Comfort and demand

Occupants switch lights on or off to maintain their visual comfort level. The power consumption of lights and appliances reflects their activities [63]. When occupants arrive in the office, they switch on the lights and associated equipment such as computers, copying machines, printers, etc. Thus power consumption correlates with occupants' presence and departure. Artificial lighting provides adequate lux levels and visual comfort to the occupants and this affects a building's energy requirement. For instance, it was observed that people were more likely to switch off the lights when they left the office for longer periods of time [80]. In [76], it was concluded that people switched on the lights during the working hours and kept them on during temporary absences. The frequency with which lights are switched on and off depends on each occupant's behaviour, habits and attitudes towards energy conservation. These factors play an essential role in energy consumption. The electrical energy is converted into light and heat energy at the same time. The heat energy is evacuated by the air conditioning, affecting the indoor environment and building energy consumption. The

energy consumed inside buildings is related to human activities. Energy resources can be split into those used by the HVAC and lighting systems to ensure a comfortable indoor environment [63]. As a result, the occupants' presence determines the power consumption of lighting and appliances.

5.4.4 Arrivals, departures and inter-floor movements

Studies have found that commercial buildings are not fully occupied for a large percentage of working hours [81]. It has been reported that workers spend 25% to 30% of business hours away from the office [106]. Three common events were identified in studies of building use: arrivals, departures and inter-floor movements in the building. During arrivals, occupants arrive at the main entrance floor or lobby, and then travel to the desired floors. During departures, occupants leave the office and then travel to the lobby or main entrance and exit the building. Occupants may also travel from floor to floor within the building. Figure 5-2 demonstrated the characteristics of these three types of movement. The three temporal patterns of travel correlate with the occupants' presence and activities in the building. Occupants travel within a building by a vertical transportation system – lifts and escalators. Occupants' traffic reflects the stochastic movement of building occupancy. Each lift shows the traffic pattern and the unique characteristics of building occupancy, which varies even between buildings with identical specifications such as load capacity. As the power consumption of lifts and escalators reflects the frequency of occupants' traffic, a higher power consumption of lifts and escalators represents more occupants inside a building.

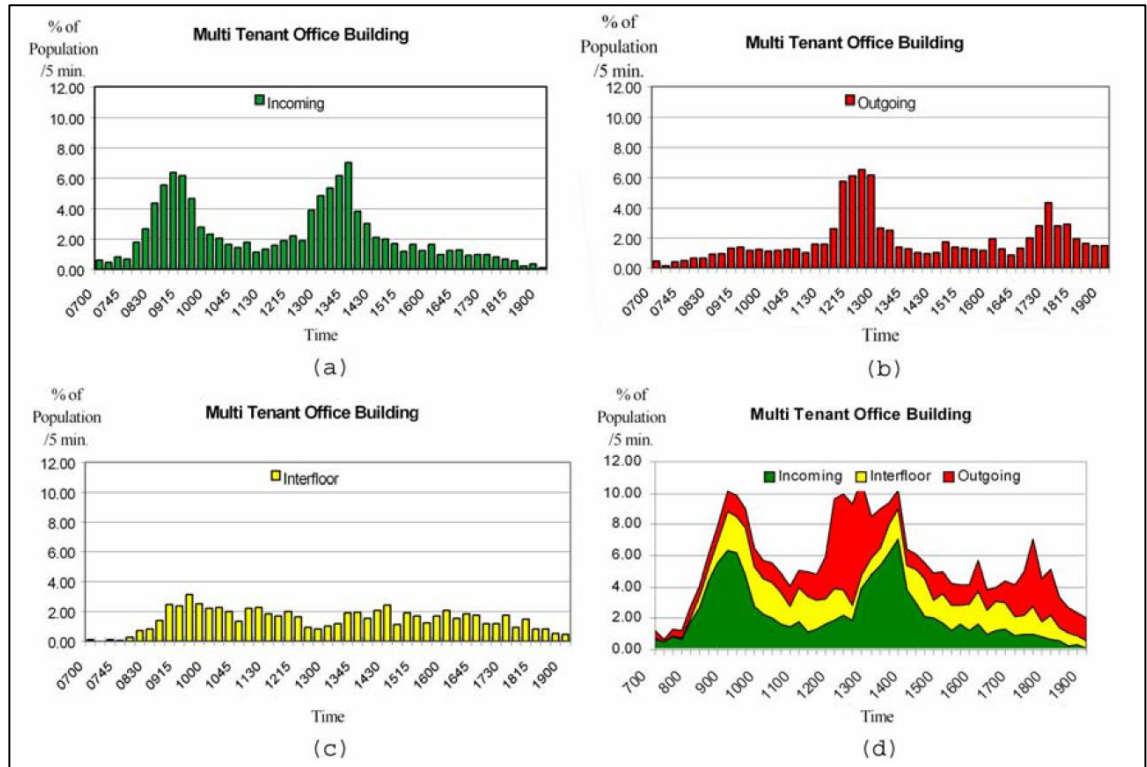


Figure 5-2 Occupants' arrival pattern in a multi-tenant office building in Paris: (a) up-peak, (b) down-peak, (c) inter-floor traffic, and (d) stacked total traffic [107].

5.5 Novel approach to determine building occupancy data

Although the above four approaches would give approximate building occupancy rates, it is still difficult to collect quantitative data to represent building occupancy. The following sub-sections propose a novel approach to collecting data on building occupancy that can be used as input parameters for building energy predictions.

5.5.1 Total space electrical consumption of each floor

Lighting and appliances are switched on and off by each occupant, indicating their arrival and departure. The power consumption of each interaction can be measured by the electrical energy meter of each floor, which can be installed at the cut-out switch of each floor. Electrical energy consumption is measured by the building power monitoring system through the energy meters installed for each zone of floors. The

system is able to record energy data in 15 minutes intervals. The electrical consumption of a floor could be used to represent occupancy.

5.5.2 Data for the PAU operation schedule in a building

The PAU operation is an indicator of the fresh air supply rate and CO₂ concentration inside the building. In this thesis, an alternative approach to this measure is to count the central control and monitoring (CCMS) schedule for the PAUs inside a building.

5.5.3 Day/hour data

Day/hour data can be used to improve the model prediction accuracy. These parameters have been reported to make significant improvements to prediction accuracy. The details are discussed in the next chapter.

5.6 Stochastic model of occupancy and activity inside an air-conditioned university building

Two important parameters, which are the outdoor climatic condition and the indoor occupancy rate, are significantly influential on cooling energy consumption. Stochastic processes are introduced in the above sessions. Stochastic model to determine the building occupants' presence and activities is established as Figure 5-3. This thesis makes use of the parameters stated to perform energy consumption prediction. Figure 5-3 shows the relationship between occupants' presence and activity and the building cooling load.

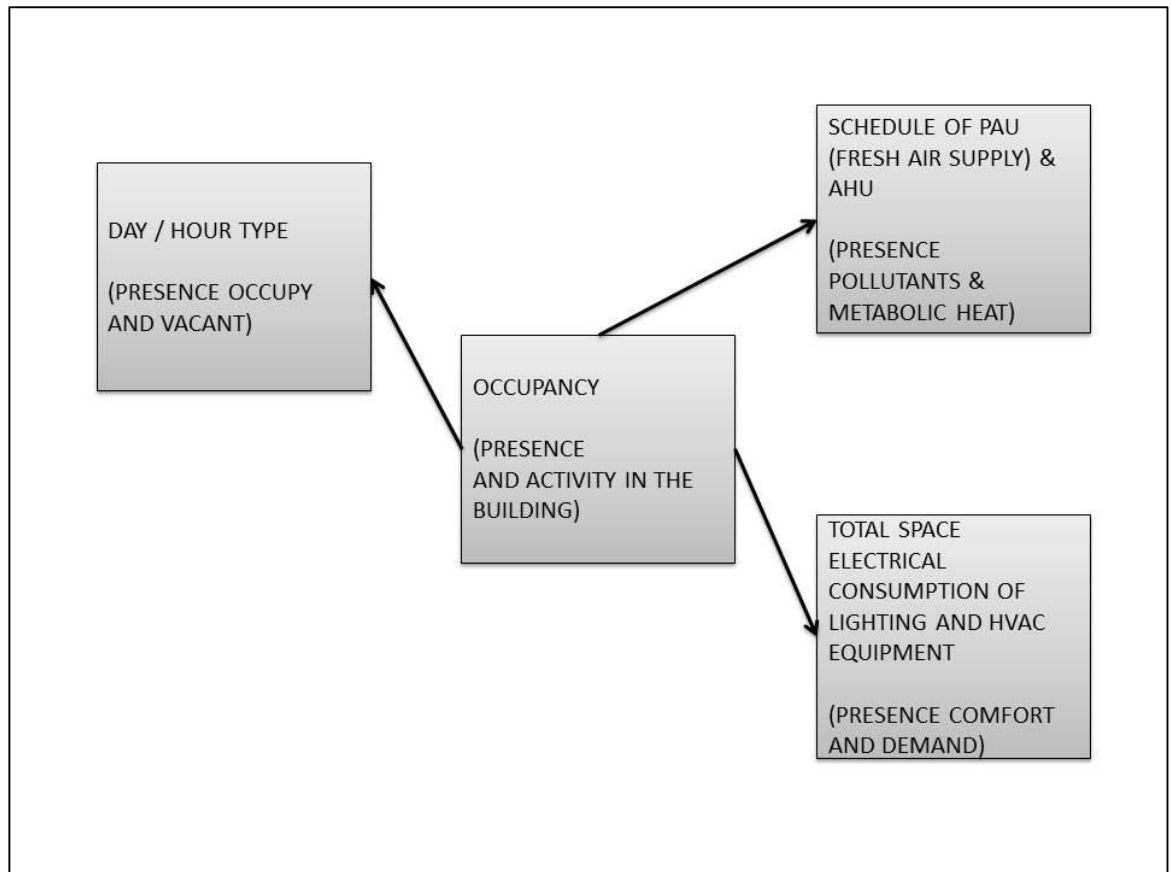


Figure 5-3 Stochastic model of occupants' presence and activity inside an air-conditioned university building

5.7 Conclusion

Two major input parameters for the building energy simulation are introduced: the external and the internal cooling load. External load can be represented by climatic weather data, whereas internal cooling load can be indicated by the building occupancy data. Current approaches to determine building occupancy adopted diversity profiles based on the assumptions of performance-based compliance demonstrations, including types of buildings, lighting operation and window blind operation. The results showed that the standard profile could not accurately predict the actual energy consumption. The dynamic changes of occupants' activities and the environmental condition were not

comprehensively considered. As a result, improvement on building occupancy model needs to be achieved. Researchers have adopted various stochastic processes of occupants' presence and activity, including measuring pollutants emission and dilution, heat gain and loss, demand for comfort level, occupants' arrivals and departures, and also their inter-floor movements. However, only four approaches reported approximation of building occupancy rates and it was also challenging to collect quantitative data in actual situation. A novel approach to collect building occupancy data for stochastic model is developed. Three major parameters of the model are listed as follows:

- i) total space electrical consumption of each floor;
- ii) data for the PAU operation schedule in a building; and
- iii) day/hour data.

The parameters are able to represent the internal building occupancy in the simulation model. The simulations with different input parameters based on this novel approach will be described in the next chapter.

Chapter 6

Application of the novel approach in ANN model

6.1 General

Currently, the ANN approach is the most widely used approach for simulating the energy consumption of buildings' electrical and HVAC systems. It is a sophisticated and useful tool for evaluating the performances of different systems. In this thesis, the ANN model discussed in previous chapters would be applied to a university building. This chapter introduces a novel model for cooling energy prediction that uses a space electrical power consumption profile, operation schedule of fresh air units and the day or hour type.

6.2 Characteristics of the sample university building

The cooling load in a building is affected by several parameters. The two main components identified in the previous chapters are outdoor climatic condition and the presence and activity of building occupants. However, the degree of influence of these factors varies with forecasting type (e.g. short term/medium term), climatic zone, building orientation, building type (e.g. commercial/residential building), building occupancy (e.g. presence and behaviour), operational schedule (e.g. fixed / dynamic schedule) and etc.

In this thesis, a university building in Hong Kong is studied and analysed. The thesis introduces a neural networks technique that uses dynamic operation area and occupancy

rate as internal load parameters. The approach adopts the multi-layer perceptron (MLP) model, which is one of the ANN models that widely adopted in engineering approaches for cooling energy consumption estimation of buildings. The training samples includes climatic data obtained from the Hong Kong Observatory and building operational data retrieved from the existing university building.

The university building had a gross floor area of 70000 m². There are two blocks, each with seven floors. Both blocks are northwest oriented. The building accommodates 570 employees and 10,800 students. The building envelope is mainly reinforced concrete and it has fixed windows with absorptive glazing. The window area occupies approximately 50% of the building facade. Manually controlled blinds are provided for all of the windows. The university building operates from 07:00 to 23:00 on weekdays, and 07:00 to 18:00 on Saturdays. The building cooling system needs to operate in winter to cater to the cooling requirements of critical laboratory areas and computer centres, and to maintain occupants' expected comfort level. The building cooling system is a central air-cooled chiller system. Air-cooled screw-type chillers are adopted with a total installed plant capacity of 1800 tonnes. Each chiller consists of two refrigeration circuits with four screw compressors. The refrigerant used is R134a. The shell-and-tube flooded type evaporator is designed to produce chilled water at a constant flow rate, with supply and return temperatures of 7°C and 12.5°C, respectively. The cooling output is by stepwise control from 25% to 100%; the system adjusts the number of compressors while controlling the chilled water temperature supplies at 7°C. There is a two-level pumping design, defined as primary and secondary pumps, to distribute chilled water to various areas. Fan coil units are individually controlled by the occupants. The design criteria are summarised in Table 6-1. The central chiller plant is controlled by a CCMS equipped with an auto-sequencing program. As fume extraction

systems are installed in some laboratory areas, fresh air is supplied by PAUs to compensate the large amount of indoor air discharged by the exhaust fans. The lighting systems, fume extraction systems and the HVAC system are all controlled by the CCMS. Their operation schedules can be readily obtained.

Total floor area : 70000 m ²			No. of occupants : 11370		
Ambient condition :	Summer	33°C dry bulb 66% saturation	Indoor design condition :	Summer	25.5°C dry bulb 54% saturation
	Winter	10°dry bulb 40% saturation		Winter	20°C dry bulb 50% saturation
Chilled water temperature :	Supply temp.	7°C	Normal operation schedule :	07:00 to 23:00 (Weekdays)	
	Return temp.	12.5°C		07:00 to 18:00 (Saturdays)	
			Total cooling capacity : 1800 tonnes		

Table 6-1 Design criteria of the cooling system

6.3 Development and training of the ANN model

The well-known feedforward backpropagation (BP) function is used as the primary test. The function can be used to perform nonlinear modelling without prior knowledge of the relationship between the input and output parameters. The Levenberg-Marquardt algorithm significantly outperforms the basic backpropagation and BP's variations with variable learning rate in terms of training accuracy, convergence properties and overall training time. The Levenberg-Marquardt algorithm can be considered a trust-region modification to the Gauss-Newton algorithm [109]. The convergence rate of the Levenberg-Marquardt algorithm is known to be super-linear. The only disadvantage of the algorithm is that the computation solution and memory requirement are comparatively higher per iteration.

The training process uses the gradient descent method. This is a network training function that updates the weight and bias values according to Levenberg-Marquardt

optimisation. It minimises a given combination of squared errors and weights, and then determines the correct combination to produce a network that can be generalised.

A three-layer feedforward model, which consists of an input layer, a hidden layer and an output layer, is adopted. As discussed in Chapter 4, the external and internal load factors are the model inputs, and the output is the electrical power consumption of the chiller plant, as depicted in Figure 6-1. MATLAB's neural network function model is used in the simulation. The activation function adopted in the ANN model is a sigmoid function. To avoid over-fitting and to provide the best generalisation, an automated regularisation feature is built into the MATLAB Levenberg-Marquardt algorithm.

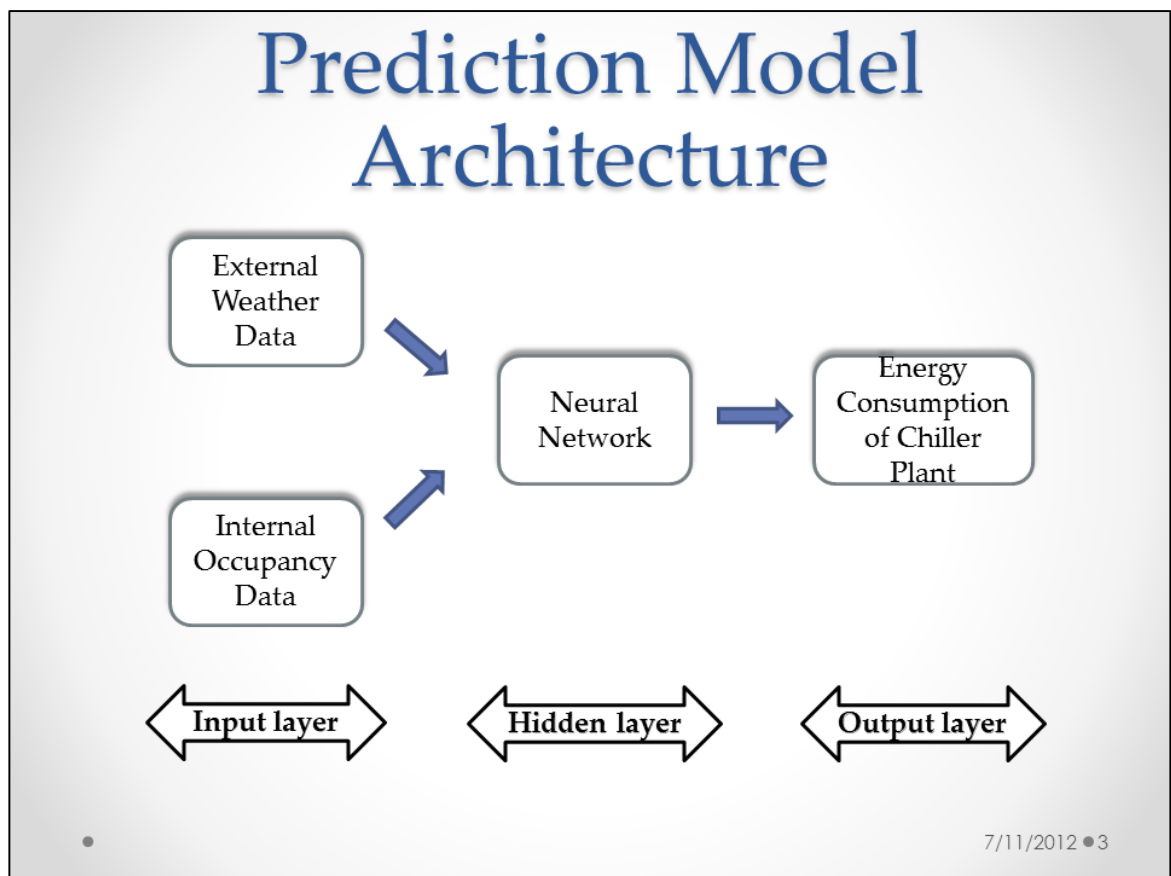


Figure 6-1 Architecture of the prediction model

6.4 Model training and prediction results

6.4.1 Input parameters and algorithm development

The ANN model adopted in this thesis is based on the Levenberg-Marquardt back-propagation algorithm. A three-layer feedforward configuration is adopted that includes an input layer, a hidden layer and an output layer. The external and internal load factors are the model inputs. The output of the model is the electricity energy consumption of the central chiller plant. Seven climatic variables, dry bulb temperature, wet bulb temperature, rainfall, mean wind speed, cloud condition, solar radiation and visibility, are adopted as the external load factors. Hourly weather data are obtained from the Hong Kong Observatory for the 1 January 2010 to 31 December 2011 period. The PAU operation schedule, hour-type/day-type and the occupancy space electrical power demand are adopted as the internal load factors. The actual building cooling load demand (kWh) is obtained from the BPMS for ANN model training and validation. The electrical power consumption data are retrieved from the BPMS at 15-minute intervals and are aggregated into hourly data and daily data for the hourly and daily estimations, respectively. For the 1 January 2010 to 31 December 2011 period, 17520 samples of hourly data are collected from the BPMS.

The data for the PAU operation schedule is estimated as

$$\text{hourly value} = \sum_{n=1}^N \left(\frac{\text{flow rate of PAU}_n \times \text{operating hours of PAU}_n}{\text{total flow rate} \times 24 \text{ hours}} \right); \quad (4)$$

$$\text{daily value} = \text{sum of hourly value}, \quad (5)$$

where N is the total number of PAUs and n represents the nth PAU.

The daily value is equal to 1 when all of the PAUs is switched on for 24 hours, and is equal to 0 when all of the PAUs are switched off.

Hour-type is indicated as a value between 1 and 24 to represent the hours in a day. The representation of day-type is as shown in Table 6-2.

Day-type	Day of the week
0	Public holiday
1	Monday
2	Tuesday
3	Wednesday
4	Thursday
5	Friday
6	Saturday
7	Sunday

Table 6-2 Representation of day-type

This thesis focuses on hourly and daily cooling load predictions for a whole year. Therefore, the 2010 and 2011 yearly data are adopted. In each study, simulations with different input parameters are conducted for comparison purposes, as described in Table 6-3. Moreover, a simulation study for evaluating variability is also performed.

Load factor	Input parameter	Simulation			
		A	B	C	D
External	Weather data ¹	✓	✓	✓	✓
Internal	Operation schedule of PAU		✓	✓	✓
	Hour-type/day-type ²			✓	✓
	Occupancy space power demand				✓

Note:

(1) Weather data includes dry bulb temperature, wet bulb temperature, global radiation, wind speed, rainfall, visibility and cloud condition

(2) Hour-type is used for hourly predictions and day-type is used for daily predictions

Table 6-3 Four simulations with different input parameters

The total computation number conducted is $2 \times 2 \times 4 \times 6 \times 100 = 9600$, as estimated as follows:

- hourly and daily predictions are conducted for 2010 and 2011;
- four simulations with different input parameters are performed;
- there are six estimations in each simulation; and
- 100 trials are performed for each estimation.

6.5 Hourly and daily cooling load prediction

The simulations are performed separately on the 2010 and 2011 datasets. For each year, the yearly data are divided into 12 months. The data from each month are further divided into six sessions. For each estimation, five sessions in each month are adopted for training the ANN model and the remaining session is reserved for verifying the model output. Six estimations are conducted in each simulation, which is useful for assessing the variability in the predictions. The overall accuracy of the overall cooling load prediction is assessed by averaging the six estimations. Tables 6-4 and 6-5 summarises the performance indices of the simulation results for 2010 and 2011, respectively. The variability reflected in the six estimations is discussed in Section 6.6.2. The performance index is statistically analysed by evaluating the upper and lower limits of 95% confidence intervals.

Performance index	Hourly simulations				Daily simulations			
	A	B	C	D	A	B	C	D
R (95% upper limit)	0.769	0.861	0.890	0.927	0.819	0.910	0.919	0.935
R (95% lower limit)	0.615	0.744	0.790	0.841	0.540	0.687	0.684	0.726
Mean R	0.692	0.803	0.840	0.884	0.680	0.799	0.802	0.831
MAPE (95% upper limit)	18.128	14.005	12.662	10.431	8.886	7.307	7.269	6.208
MAPE (95% lower limit)	13.691	10.286	9.296	7.435	5.293	3.764	3.883	2.965
Mean MAPE	15.910	12.146	10.979	8.933	7.090	5.536	5.576	4.587
CV (95% upper limit)	21.438	17.579	16.191	12.961	9.875	8.088	8.200	6.919
CV (95% lower limit)	16.407	12.951	11.727	9.298	6.201	4.655	4.587	3.617
Mean CV	18.923	15.265	13.959	11.13	8.038	6.372	6.394	5.268

Table 6-4 2010 hourly and daily simulation results

Performance index	Hourly simulations				Daily simulations			
	A	B	C	D	A	B	C	D
R (95% upper limit)	0.746	0.820	0.860	0.911	0.810	0.925	0.924	0.978
R (95% lower limit)	0.594	0.696	0.750	0.821	0.462	0.736	0.717	0.776
Mean R	0.67	0.758	0.805	0.866	0.636	0.831	0.821	0.877
MAPE (95% upper limit)	16.856	12.547	11.293	9.874	7.741	6.167	6.289	6.819
MAPE (95% lower limit)	12.915	9.767	8.418	6.916	4.355	3.007	3.167	2.782
Mean MAPE	14.886	11.157	9.856	8.395	6.048	4.587	4.728	4.801
CV (95% upper limit)	20.195	16.278	14.872	13.080	8.823	6.986	7.140	7.701
CV (95% lower limit)	15.751	12.709	11.098	8.957	5.106	3.705	3.721	3.326
Mean CV	17.973	14.494	12.985	11.019	6.965	5.346	5.431	5.514

Table 6-5 2011 hourly and daily simulation results

Figures 6-2 and 6-3 shows graphically the actual and predicted load profiles for the hourly simulations in a typical week from 2010 and 2011, respectively. Figures 6-4 and 6-5 shows graphically the actual and predicted load profiles for daily simulations in a typical month from 2010 and 2011, respectively.

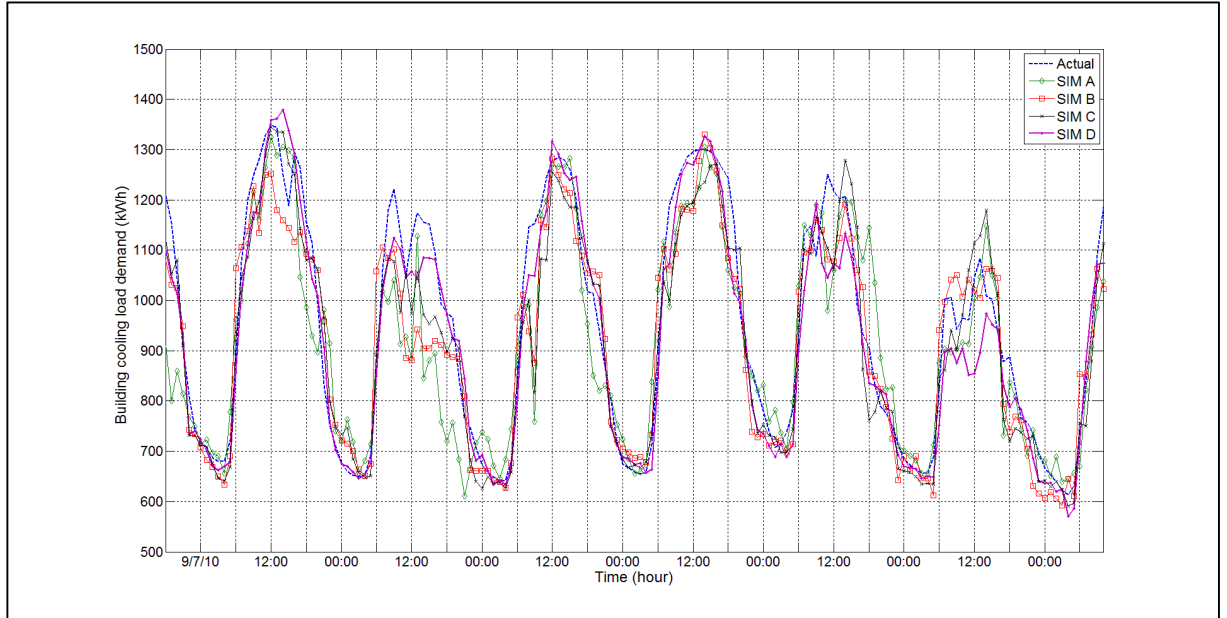


Figure 6-2 2010 hourly simulation results

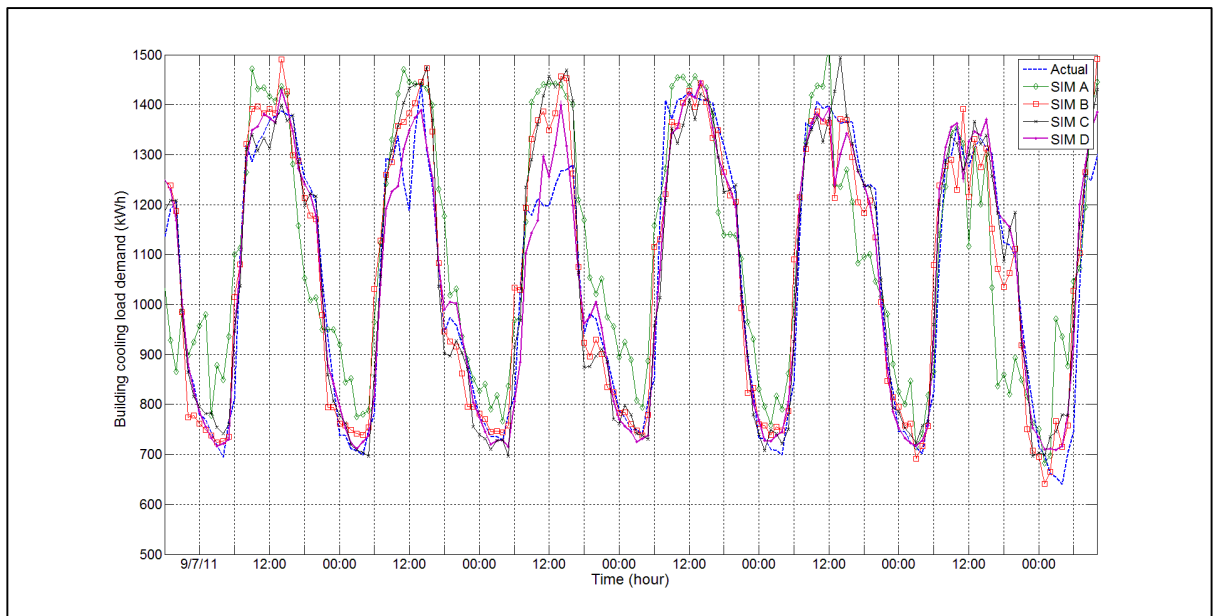


Figure 6-3 2011 hourly simulation results

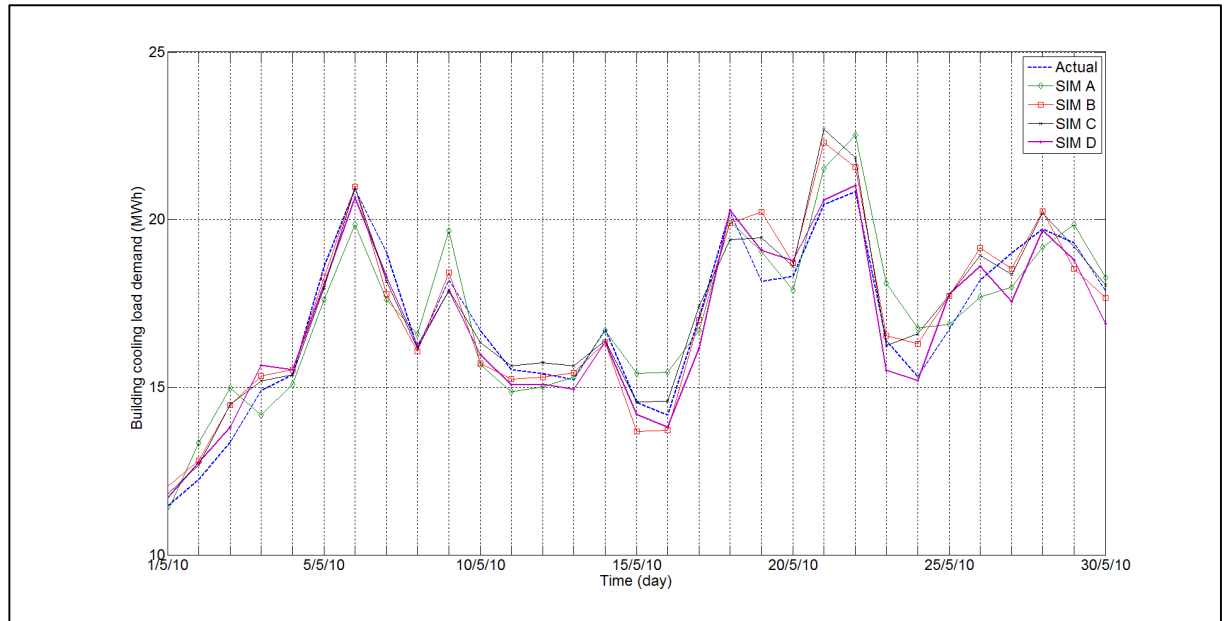


Figure 6-4 2010 daily simulation results

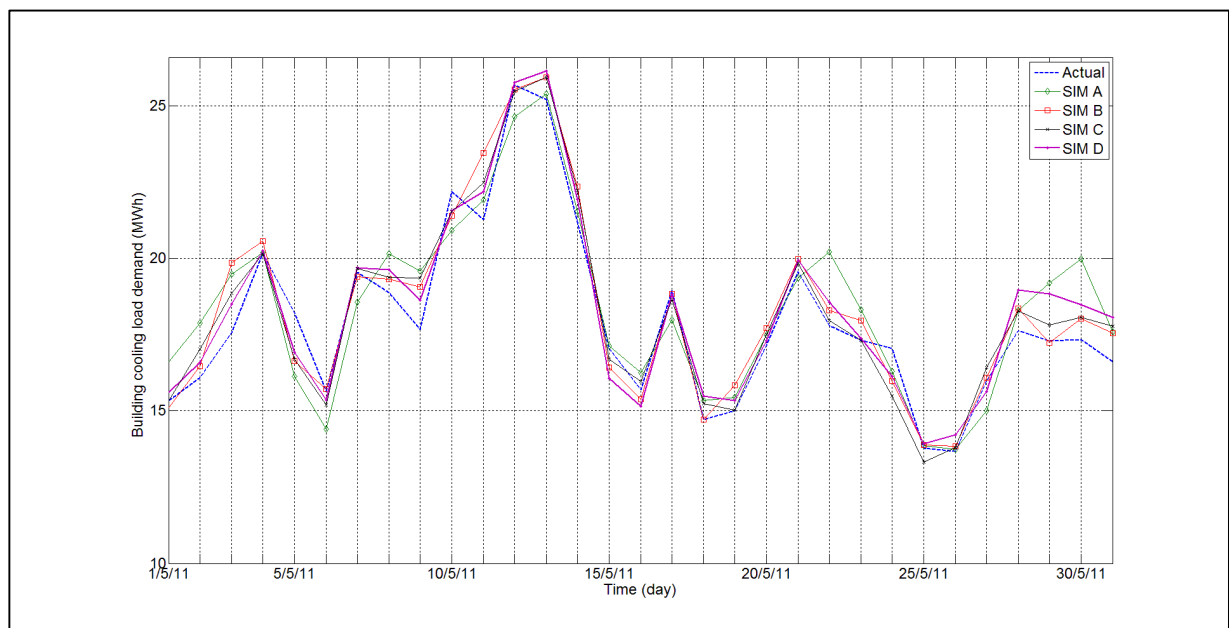


Figure 6-5 2011 daily simulation results

6.6 Discussion

6.6.1 Overall prediction performance

Figure 6-6 presents a comparison of the performance indices of the four simulations for hourly and daily cooling load predictions in 2010 and 2011.

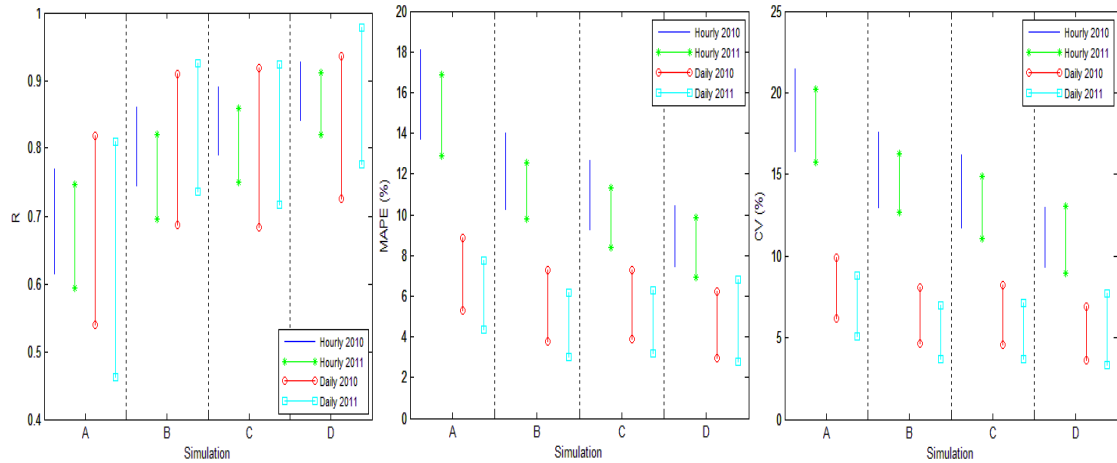


Figure 6-6 Comparison of the performance indices for the four simulations

The performance of Simulation A is far from satisfactory. This is as expected since external weather variables are the only input parameter. The prediction performance improves when the PAU operation schedule (Simulation B) and hour-type/day-type (Simulation C) are introduced. For the hourly prediction, Simulation C has a better R value than Simulation B. For the daily prediction, the R values for Simulations B and C are similar. The same pattern is observed in the MAPE and CV values for Simulations B and C. As a result, it can be concluded that the use of hour-type as an input parameter improves the hourly prediction performance, but the use of day-type as an input parameter do not improve the daily prediction performance, as the variation in cooling loads between weekdays is small. It was evident that Simulation D has best performance in both hourly and daily predictions. The improvement is significantly obvious for hourly predictions. As the daily variation of energy consumption is less rapid than the hourly variation, the daily prediction is expected to be more accurate. The simulation results, which are presented in Tables 6-4 and 6-5, and Figure 6-6 confirms this conclusion.

The prediction results for the 2011 data are selected for detailed study. The prediction results for 2010 show similar patterns and therefore are not discussed further. Tables 6-6

and 6-7 show the performance indices for the hourly and daily simulation results for Simulation D in each month of 2011. The graphical representations are shown in Figures 6-7 and 6-8, respectively. The performance index is statistically analysed by evaluating the upper and lower limits of 95% confidence intervals.

	Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Mean
R	upper limit	0.754	0.670	0.884	0.936	0.965	0.972	0.970	0.977	0.974	0.971	0.955	0.908	0.911
	lower limit	0.538	0.313	0.764	0.892	0.904	0.950	0.942	0.956	0.945	0.927	0.909	0.816	0.821
MAPE	upper limit	10.244	16.021	13.378	9.393	8.030	8.333	7.525	6.207	8.489	7.990	10.054	12.822	9.874
	lower limit	7.105	10.020	10.046	7.165	5.303	5.847	5.634	4.494	5.871	5.338	6.892	9.276	6.916
CV	upper limit	12.628	24.222	17.498	12.515	12.511	10.356	10.162	7.902	10.563	10.082	12.201	16.324	13.080
	lower limit	9.034	13.391	12.697	9.328	7.758	7.576	7.385	5.808	7.536	6.674	8.495	11.797	8.957

Table 6-6 Monthly performance indices of the 2011 hourly prediction results (Simulation D)

	Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Mean
R	upper limit	0.956	0.967	0.984	0.988	0.966	0.985	0.979	0.986	0.967	0.984	0.995	0.980	0.978
	lower limit	0.419	0.302	0.854	0.885	0.882	0.812	0.862	0.845	0.817	0.837	0.946	0.858	0.776
MAPE	upper limit	9.656	12.666	9.283	5.586	5.122	4.548	4.418	4.727	4.107	5.002	5.942	10.776	6.819
	lower limit	3.211	5.601	4.126	1.965	2.389	1.875	1.935	2.174	1.977	2.137	2.130	3.860	2.782
CV	upper limit	11.417	14.452	10.258	6.423	5.828	5.275	4.842	5.217	4.949	5.787	6.546	11.414	7.701
	lower limit	3.938	6.748	4.822	2.449	2.880	2.283	2.345	2.454	2.394	2.571	2.480	4.543	3.326

Table 6-7 Monthly performance indices of the 2011 daily prediction results (Simulation D)

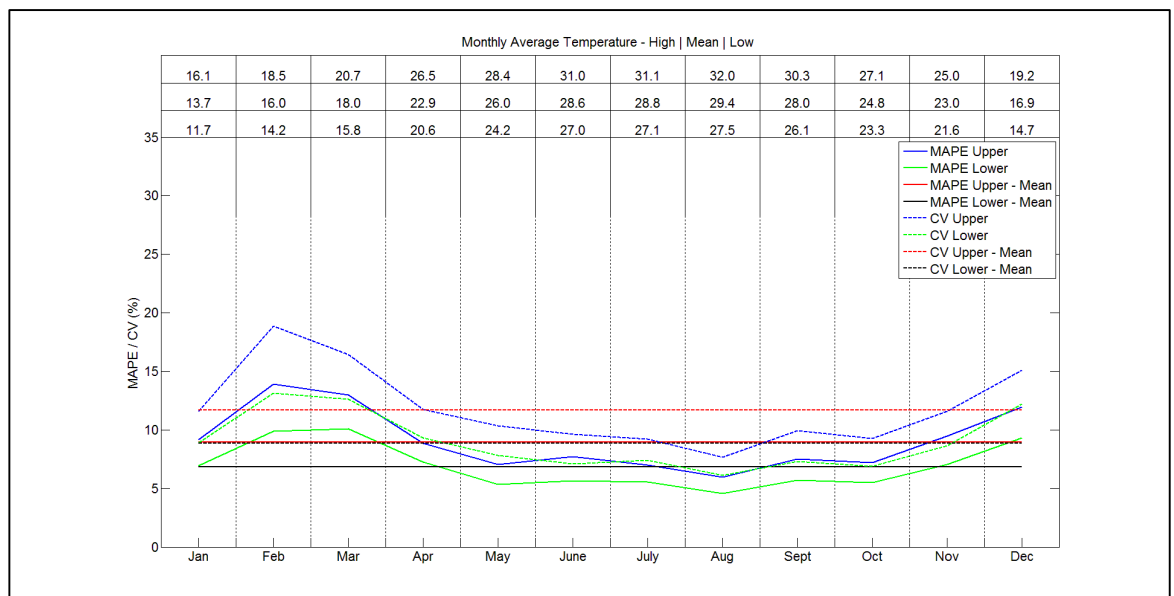
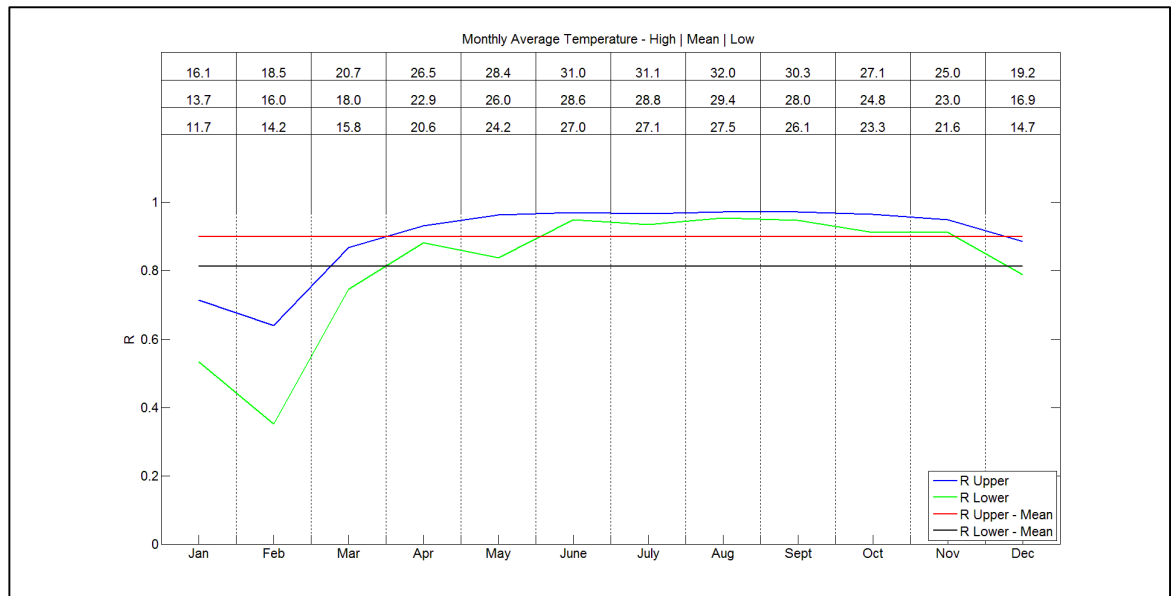


Figure 6-7 Monthly performance indices of the 2011 hourly prediction results (Simulation D) R (upper) and MAPE and CV (lower)

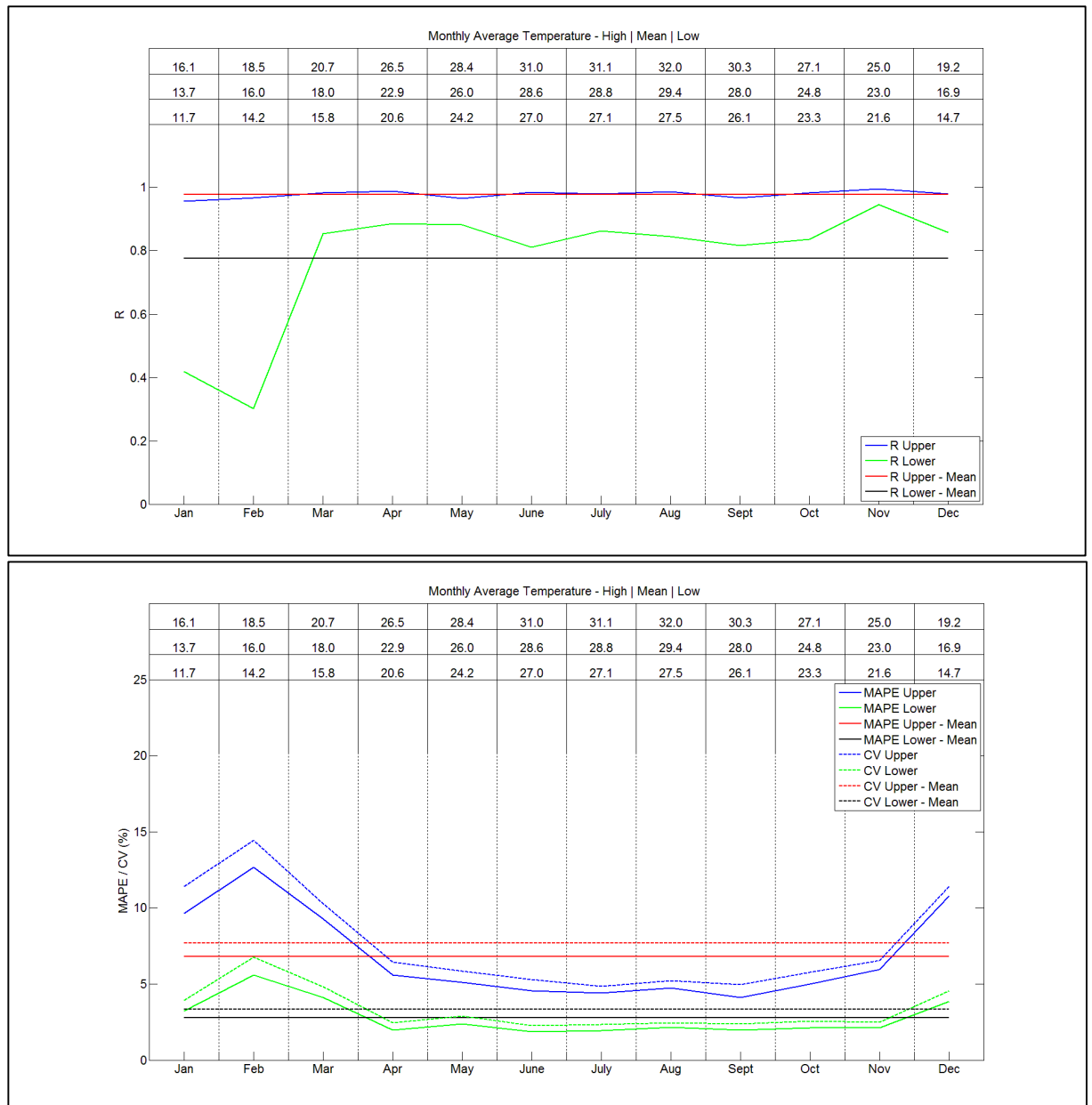


Figure 6-8 Monthly performance indices of 2011 daily prediction results (Simulation D) R (upper) and MAPE and CV (lower)

As observed, both the hourly and daily prediction performances are less accurate in December, January, February and March. The spread of performance indices in these months is also large. These months are in winter season in sub-tropical Hong Kong, when the monthly average outdoor temperature is below 20°C. The winter indoor design condition of the building cooling system is set at 20°C (Table 6-1). However, the building cooling system still needs to operate in winter to cater the cooling requirements

of critical laboratory areas and computer centres, and to maintain the occupants' comfort level. As in winter, heat generated by human activities is removed by the cooling system and absorbed by the building, the correlation between the building cooling load and occupancy space power demand is not strong in these months.

In contrast, both the hourly and daily prediction performances from April to November are better than the yearly mean. The spread of the performance indices in these months is also small. The summer in sub-tropical Hong Kong includes June, July, August and September. The monthly average outdoor temperature is around 29°C. The summer indoor design condition of the building cooling system is set at 25.5°C (Table 6-1). The prediction performance in the summer is very satisfactory. The best 95% lower limits of MAPE and CV in the hourly predictions are 4.494% and 5.808%, respectively, and in the daily predictions they are 1.935% and 2.345%, respectively. An inspection of the actual electricity consumption reveals that the prediction performance is satisfactory when the building cooling load demand is higher than the occupancy space power demand.

A similar study conducted for an office building in Hong Kong is selected for comparison [110]. That study used CO₂ concentration as a proxy for human behaviour. The study adopted an ANN approach to perform short-term (one week) hourly building cooling load predictions. A direct comparison may not be fully appropriate as the building used and the prediction periods are very different in the two studies. Nevertheless, a comparison of the prediction accuracy is shown in Table 6-8. It shows that the ANN model proposed in this study performs better in terms of the best result and variations in various trials.

	ANN model in [26]	Proposed ANN model	
	One-week hourly prediction (RMSPE)	Hourly prediction in 2011 (CV)	
		Yearly average	Best month
95% lower limit	12.12	8.957	5.808
95% upper limit	16.36	13.080	7.902

Table 6-8

Comparison with previous study

6.6.2 Variability in prediction performance

As mentioned in Section 6.4, six estimations are conducted in each simulation. Figure 6-9 shows the prediction results of the six estimations in Simulation D for hourly and daily predictions in 2011. Only the R and CV values are shown. Tables 6-9 and 6-10 present the mean performance indices of the six estimations for hourly predictions and daily predictions, respectively.

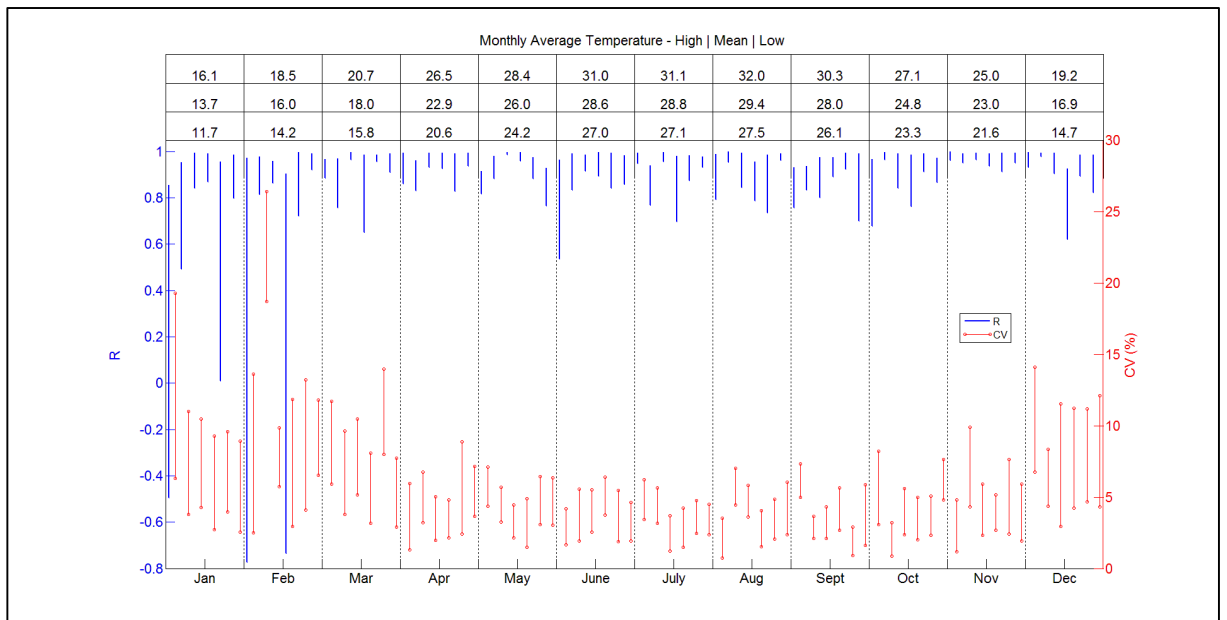
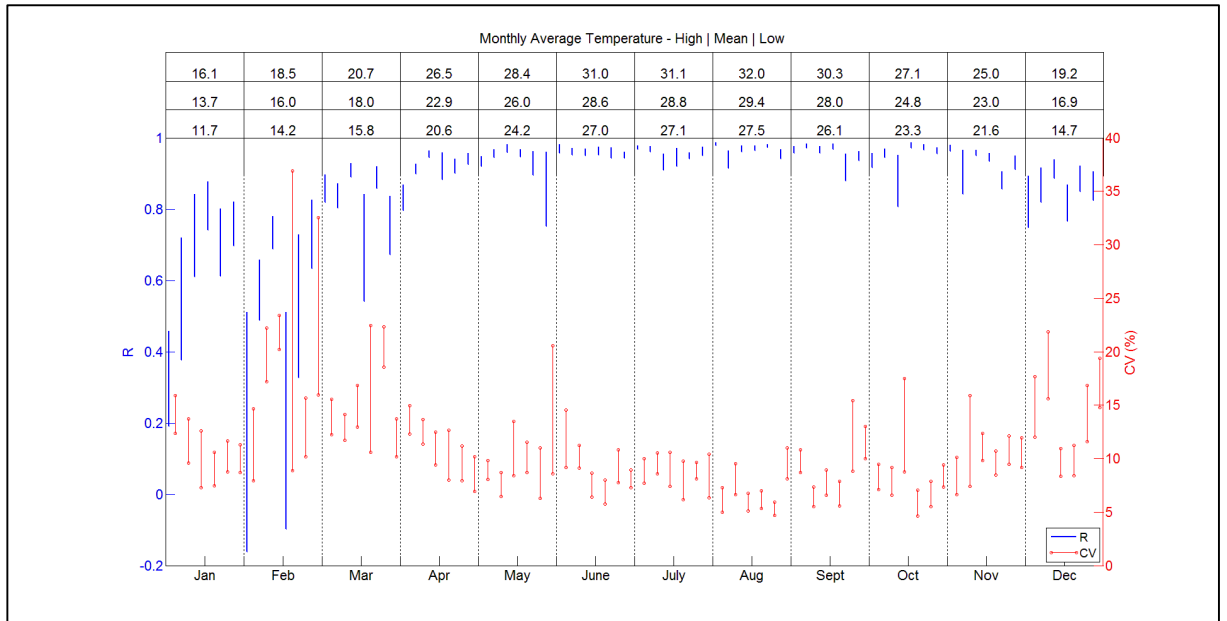


Figure 6-9 2011 prediction results with six estimations (Simulation D) hourly (upper) and daily (lower)

Month Estimation		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
R	1	0.305	0.244	0.827	0.810	0.935	0.967	0.975	0.977	0.962	0.928	0.969	0.751
	2	0.483	0.563	0.802	0.920	0.959	0.961	0.970	0.953	0.981	0.962	0.917	0.856
	3	0.738	0.726	0.904	0.944	0.974	0.961	0.931	0.969	0.965	0.822	0.949	0.907
	4	0.818	0.101	0.678	0.920	0.961	0.959	0.920	0.971	0.961	0.982	0.939	0.772
	5	0.655	0.617	0.883	0.922	0.667	0.961	0.948	0.979	0.933	0.976	0.888	0.889
	6	0.748	0.723	0.751	0.927	0.909	0.950	0.965	0.933	0.951	0.962	0.921	0.848
	mean	0.625	0.496	0.808	0.907	0.901	0.960	0.951	0.963	0.959	0.939	0.931	0.837
	S.D.	0.177	0.239	0.077	0.044	0.107	0.005	0.020	0.016	0.015	0.055	0.026	0.057
CV	1	13.265	9.995	14.076	13.775	8.900	9.887	8.559	6.452	9.450	8.395	7.703	13.764
	2	10.835	19.148	13.343	11.780	7.265	9.889	9.398	7.289	5.790	7.483	11.526	18.674
	3	8.878	21.730	14.376	10.331	10.406	6.988	8.101	5.753	7.842	12.658	10.856	9.689
	4	7.915	16.507	13.645	10.015	9.655	6.604	7.079	6.289	6.873	5.431	9.167	10.212
	5	10.494	11.373	19.731	8.950	6.838	8.739	8.693	5.136	10.579	6.229	10.370	13.411
	6	10.198	17.346	11.971	8.408	11.393	8.048	7.953	10.404	11.229	8.161	11.011	16.244
	mean	10.264	16.017	14.524	10.543	9.076	8.359	8.297	6.887	8.627	8.060	10.105	13.666
	S.D.	1.675	4.129	2.450	1.798	1.622	1.283	0.716	1.704	1.958	2.304	1.299	3.150

Table 6-9 2011 hourly prediction - mean performance indices of the six estimations (Simulation D)

Month Estimation		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
R	1	0.045	0.255	0.926	0.871	0.873	0.653	0.955	0.939	0.856	0.863	0.962	0.972
	2	0.541	0.936	0.851	0.907	0.869	0.913	0.846	0.988	0.898	0.975	0.981	0.987
	3	0.782	0.928	0.756	0.974	0.988	0.952	0.965	0.888	0.888	0.955	0.976	0.241
	4	0.859	0.312	0.905	0.967	0.980	0.936	0.630	0.868	0.930	0.826	0.971	0.765
	5	0.328	0.932	0.970	0.880	0.113	0.899	0.738	0.907	0.959	0.940	0.974	0.934
	6	0.869	0.952	0.964	0.961	0.827	0.939	0.944	0.981	0.925	0.951	0.982	0.623
	mean	0.571	0.719	0.895	0.927	0.775	0.882	0.846	0.928	0.909	0.918	0.974	0.754
	S.D.	0.304	0.309	0.074	0.042	0.302	0.104	0.125	0.045	0.033	0.054	0.007	0.263
CV	1	10.290	5.421	7.692	2.664	5.345	2.849	4.994	1.649	5.797	4.963	2.702	10.748
	2	7.567	21.911	6.923	4.836	4.547	3.846	4.162	5.967	2.485	2.239	5.826	6.123
	3	7.152	7.503	7.621	2.551	2.969	3.698	2.395	4.188	3.139	3.029	3.218	7.655
	4	4.594	4.681	3.720	2.739	2.639	4.752	2.936	2.179	3.231	3.900	3.352	7.433
	5	5.257	5.870	12.036	4.224	4.479	2.956	3.441	3.248	2.036	3.825	5.107	7.375
	6	5.707	8.147	5.061	4.953	5.011	3.284	3.319	4.146	2.653	6.208	3.370	7.304
	mean	6.761	8.922	7.176	3.661	4.165	3.564	3.541	3.563	3.224	4.027	3.929	7.773
	S.D.	1.885	5.929	2.600	1.036	1.009	0.641	0.840	1.425	1.219	1.283	1.129	1.419

Table 6-10 2011 daily prediction - mean performance indices of the six estimations (Simulation D)

In winter (December, January, February and March), there are large variations in the performance indices between the six estimations. Extreme results are noted in the daily predictions. The prediction performance in this season is less accurate for the reasons discussed in Section 6.6.1. For other months, the variation in the prediction performance is modest. In particular, the prediction variation in summer (June, July, August and September) is very modest.

The observed prediction variation in this study is different from that in the simulation results reported in [111], where the cooling demand varied by a factor of two between extreme cases. That large variation was mainly due to variations in individual behaviour, which was reflected in the individual variation in the use of shading devices and the window opening ratio of a small artificial office unit. This thesis is conducted at the building level rather than the office level. The occupants do not individually control window opening and the window blind positions are rarely altered. However, the prediction variation in summer in this thesis matches the simulation results for cooling demand presented in [112].

6.6.3 Daily peak cooling demand prediction performance

The prediction performance of the daily peak cooling load demand during the hot season is of particular interest, as this is a piece of useful information for system dimensioning. Tables 6-11 and 6-12 summarises the prediction performance indices for daily peak load demand in 2010 and 2011, respectively.

The results reveal that the prediction performance in summer (June, July, August and September) is comparable to the overall prediction performance presented in Tables 6-4 and 6-5. Although the variation in prediction performance in other months is moderate, the prediction performance in winter season shows a large variation. The performance index is statistically analysed by evaluating the upper and lower limits of 95% confidence intervals.

	Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
R	upper limit	0.881	0.965	0.971	0.983	0.962	0.944	0.813	0.931	0.990	0.958	0.934	0.963
	lower limit	0.453	0.484	0.894	0.871	0.864	0.747	0.527	0.596	0.052	0.832	0.825	0.782
MAPE	upper limit	13.921	15.359	9.657	9.796	6.287	6.248	4.639	5.612	8.430	7.980	7.634	13.808
	lower limit	6.278	5.750	3.949	4.122	3.156	2.997	2.313	2.786	3.312	3.649	3.193	5.799
CV	upper limit	15.463	16.574	10.583	10.426	7.402	7.018	5.660	6.575	9.366	9.064	8.619	15.577
	lower limit	7.564	6.736	4.787	5.146	3.961	3.618	2.862	3.423	4.064	4.506	4.096	6.961

Table 6-11 Monthly performance indices of 2010 daily peak prediction results (Simulation D)

	Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
R	upper limit	0.881	0.757	0.952	0.979	0.964	0.906	0.938	0.766	0.912	0.937	0.929	0.929
	lower limit	0.378	0.201	0.786	0.841	0.838	0.598	0.790	0.402	0.677	0.663	0.757	0.663
MAPE	upper limit	15.381	19.680	11.657	8.303	7.764	6.237	5.533	5.101	5.958	6.364	7.802	13.622
	lower limit	6.062	9.459	5.712	3.592	4.100	3.320	2.867	2.737	2.958	3.136	3.778	5.930
CV	upper limit	17.795	23.095	13.851	9.232	8.848	7.257	6.539	6.433	6.914	8.092	9.368	15.962
	lower limit	7.718	12.410	7.503	4.254	4.753	3.854	3.538	3.595	3.654	4.144	4.699	7.548

Table 6-12 Monthly performance indices of 2011 daily peak prediction results (Simulation D)

6.7 Conclusion

This chapter describes the application of the novel approach, which includes the building occupants' presence and activities as an input parameter, on cooling load prediction. Occupants' activities are determined by i) the occupancy space power demand, ii) operation schedule of fresh air units and iii) the day/hour type. The simulation results reveal that the novel approach can significantly improve the accuracy of the cooling load prediction. The prediction performance is found to be enhanced compared with the previous similar studies [26].

Chapter 7

Conclusions and Recommendations

7.1 Conclusions

The presence and activities of occupants in the building can exert a significant impact on the cooling energy consumption depending on their energy efficiency measures taken and their expected indoor comfort level. Also, current building control systems have been adopted the principles of engineering and assumed preference averages. As a result, the uncertainties of the presence and interactions of building occupants may lead to questionable performance outcome accuracy in energy prediction models. However, the uncertainties related to the presence and interactions of building occupants have not been investigated intensively.

Many researchers employed the ANN models for energy prediction in the past decade and revealed a high accuracy. As a result, ANN prediction models, which use historic data to determine the energy demand prediction in dynamic short-term loads, are adopted in this thesis.

This thesis reviews different approaches or models for building cooling load prediction, from the state-of-art approaches to the recently intelligent approaches. It also introduces a novel approach that takes the behaviour of building occupants into account. Occupancy space power demand is used as a proxy for occupants' activities. Power data recorded by an intelligent networked building power monitoring system (BPMS) is used to test the new approach. The electrical power consumption of a fully air-conditioned university building in Hong Kong in 2010 and 2011 is used to conduct the simulation study.

The simulation results reveal that the use of occupancy space power demand significantly improves the accuracy of the cooling load prediction. The prediction performance is found to be greater than that in the similar study [26]. Table 6-8 demonstrates the prediction performance improvement as below:

- i) Yearly average prediction (CV) is 8.957% at 95% lower limit of confidence intervals and 13.080% at 95% upper limit of confidence intervals;
- ii) Best month prediction (CV) is 5.808% at 95% lower limit of confidence intervals and 7.902% at 95% upper limit of confidence intervals.

In conclusion, the results of cooling load prediction simulations can be significantly improved by adopting the novel approach suggested in this thesis, which determine the building occupancy by the following areas:

- i) total space electrical consumption of each floor;
- ii) data for the PAU operation schedule in a building; and
- iii) day/hour type.

7.2 Recommendations for future research

A novel approach to determine occupancy activity is developed and validated. The simulation results reveal improved accuracy in cooling load prediction in a university building. However, further validation and verification are required. The proposed ANN model can be applied to other types of buildings that have similar building systems configurations, such as office buildings; commercial complexes and shopping arcades, to further investigate the effect of the approach on improving the cooling load prediction. In the future, ANN models can be implemented in researches that aim at the following goals:

- i) to optimise the chiller operational sequence and mode;

- ii) to improve the model's accuracy in winter;
- iii) to use the model in a smart-grid demand response study;
- iv) to apply the novel model on heating plant of the buildings; and
- v) to verify the novel model by apply in different types of buildings.

List of Publications

Journal Papers

1. Leung, M.C., Tse, Norman C.F., Lai, L.L. and Chow T.T. (2012) The use of occupancy space electrical power demand in building cooling load prediction, *Energy and Buildings* 55 151-163.

Conference Papers

1. Xu, Fang Yuan, Leung, M.C. and Zhou Long (2010) A RBF network for short-term load forecast on microgrid, in: *Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao* 6 3195-3199.
2. Mok, Y.K., Tse, Norman C.F., Lau, W.H. and Leung, M.C. (2011) Experiment-based simulation for distortion behavior in LV networks for microgrid modeling, *Power and Energy Society General Meeting, 2011 IEEE, San Diego, CA* 1-6.
3. Leung, M.C., Mok, Y.K., Tse, Norman C.F., Fong, Alan M.L. and Lai, L.L. (2012) New MVAC control by making use of human behavioural based technique to achieve energy efficiency, *Advance in Power System Control, Operation and Management (APSCOM 2012), 9th IET International Conference on, Hong Kong* 1-5.
4. Chan, John Y.C., So, Jacky M.K., Leung, M.C. Tse, Norman C.F. and Kong, Percy (2012) Power quality and energy usage of a multi-purpose building in Hong Kong, *Advance in Power System Control, Operation and Management (APSCOM 2012), 9th IET International Conference on, Hong Kong* 1-6.

References

- [1] Blasing, T.J. (February 2013) Current greenhouse gas concentrations, doi:10.3334/CDIAC/atg.032, on CDIAC 2013.
- [2] Hong Kong energy statistics. Census and Statistics Department, Hong Kong Government, Hong Kong. Available at <http://www.censtatd.gov.hk/>
- [3] Zhu, Y. (2006) Applying computer-based simulation to energy auditing: A case study, *Energy and Buildings* 38(5) 421-428.
- [4] Neto, A.H. and Fiorelli, F.A.S. (2007) Use of simulation tools for managing buildings energy demand, *Proceedings: Building Simulation*, Beijing, China.
- [5] Yu, P.C.H. and Chow, W.K. (2001) Energy use in commercial buildings in Hong Kong, *Applied Energy* 69(4) 243-255.
- [6] ASHRAE Handbook (2005), Chapter 31: Energy Estimating and Modeling Methods.
- [7] Kalogirou, S.A. (2010) Chapter 1: Artificial Intelligence Techniques for Modern Energy Applications, IG Global.
- [8] Weigend, A.S. and Gershenfeld, N.A. (1994) *Time Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, Reading.
- [9] Master T. (1995) *Neural, Novel & Hybrid Algorithms for Time Series Prediction*, John Wiley & Sons Inc., ISBN 0471130419.
- [10] Petridis, V. and Kehagias, A. (1998) *Predictive Modular Neural Networks*, Springer-Verlag, New York.

- [11] Farrell, J.A. and Polycarpou, M.M. (2006) Adaptive Approximation Based Control: Unifying Neural, Fuzzy and Traditional Approximation Approaches, Wiley, New York.
- [12] Kyriakidesand, E., and Polycarpou, M. (2007) Short term electric load forecasting: A tutorial, Studies in Computational Intelligence (SCI) 35.
- [13] Metaxiotis, K., and Kagiannas, A. (2005) Intelligent computer applications in the energy sector: A literature review from 1990 to 2003, Int. J. Computer Applications in Technology, Vol. 22, Nos. 2/3.
- [14] Lu, X., Clements-Croome, and D., Viljanen, M. (2009) Past, present and future mathematical models for buildings – Focus on intelligent buildings (Part 1), Intelligent Buildings International, 1(1): 23–38.
- [15] Cristianini, N. and Shawe-Taylor, J. (2000) An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, ISBN 0-521-78019-5.
- [16] Irie, B., and Miyake, S. (1988) Capabilities of three-layered perceptrons, Proceedings of IEEE International Conference on Neural Networks, I, 641-648.
- [17] Hornik, K. (1991) Approximation capabilities of multilayer feed-forward networks, Neural Networks 4, 252-257.
- [18] Zhang, G., Eddy, B., Patuwo, and Hu M.Y. (1998) Forecasting with artificial neural networks: The state of the art, International Journal of Forecasting 14 35-62.
- [19] Hoes, P., Hensen, J.L.M., Loomans, M.G.L.C., Vries, de B., and Bourgeois, D. (2008) User behaviour in whole building simulation, Energy and Buildings 41, 2009, 295-302.
- [20] Zimmermann, G. (2007) Modeling and simulation of individual user for building performance predictions, SCSC, 913-920.

- [21] Mahdavi, A. and Pröglhöf, C. (2009) User Behaviour and Energy Performance in Buildings, IEWT, Int. Energy Economics Workshop TUV, Vienna, 2009 1-13.
- [22] Robinson, D., Campbell, N., Gaiser, W., Kabel, K., Le-Mouel, A., Morel, N., Page, J., Stankovic, S., and Stone, A. (2007) SUNtool – A new modeling paradigm for simulating and optimizing urban sustainability, *Solar Energy* 81(9) 1196-1211.
- [23] Song, J., and Tsianaka, E. (2008) Building energy performance considering post occupancy: A case study of an innovative academic research building, 25th Conference on Passive and Low Energy Architecture, Dublin, 22nd to 24th October 2008.
- [24] Yik, F.W.H., Yee, K.F., Sat, S.K., and Chan, W.H. (1998) A detailed energy audit for a commercial office building in Hong Kong, *HKIE Transaction* Vol.5(3) 84-88.
- [25] S. Avdakovic, S., Nuhanovic, A., Kusljugic, M., and Music, M. (2012) Wavelet transform applications in power system dynamics, *Electric Power Systems Research* 83 237-245.
- [26] Yezioro, A., Dong, B., and Leite, F. (2008) An applied artificial intelligence approach towards assessing building performance simulation tools, *Energy and Buildings* 40(4) 612-620.
- [27] Olofsson, T., Andersson, S., and Ostin, R. (1998) A method for predicting the annual building heating demand based on limited performance data, *Energy and Buildings* 28(1) 101-108.
- [28] Krarti, M. (2000) Chapter 16: Energy Audit of Building Systems – An Engineering Approach, CRC Press, ISBN 0-8493-9587-9.

- [29] Neto, A.H., Augusto, F., and Fiorelli, F.A.S. (2008) Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption, *Energy and Buildings* 40(12) 2169-2176.
- [30] Yang, K.T. (2008) Artificial neural networks (ANNs): A new paradigm for thermal science and engineering, *Journal of Heat Transfer-transactions of the ASME* 130(9) Article No. 093001.
- [31] Hou, Z.J., Lian, Z.W., Yao, Y., and Yuan X.J. (2006) Data mining based sensor fault diagnosis and validation for building air conditioning system, *Energy Conversion and Management* 47(15-16) 2479-2490.
- [32] Thomas, B. and Soleimani-Mohseni, M. (2007) Artificial neural network models for indoor temperature prediction: investigations in two buildings, *Neural Computing & Applications* 16(1) 81-89.
- [33] Feinberg, E.A., and Genethliou, D. (2005) Chapter 12: Load forecasting, *Applied Mathematics for Restructured Elections Power Systems*, Springer, New York, U.S., 269-285.
- [34] Gonzalez, P.A., and Zamarreno, J.M. (2005) Prediction of hourly energy consumption in buildings based on a feedback artificial neural network, *Energy and Building* 37(6) 595-601.
- [35] TRNSYS, Transient System Simulation Tool, University of Wisconsin; ENERGYplus, building energy simulation program, U.S. Department of Energy's (DOE) Building Technologies Office (BTO).
- [36] BLAST Support Office (1992) BLAST 3.0 Users' Manual, Department of Mechanical and Industrial Engineering, University of Illinois.
- [37] Engle, R.F., Mustafa, C., and Rice, J. (2003) Modeling peak electricity demand, *Journal of Forecasting* 11 241-251.

- [38] Goldberg, D. (1989) Genetic Algorithms in Search Optimization and Machine Learning, Reading, MA, Addison-Wesley.
- [39] Chow, M., and Tram, H. (1997) Application of fuzzy logic technology for spatial load forecasting, IEEE Transactions on Power Systems 12(3), 1360-1366.
- [40] Mori, H., Sone, Y., Moridera, D., and Kondo, T. (1999) Fuzzy inference models for short-term load forecasting with tabu search, IEEE Systems, Man and Cybernetics Conference Proceedings 6 551-556
- [41] Srinivasan, D., Tan, S.S., Chang, C.S., and Chan, E.K. (1999) Parallel neural network-fuzzy expert system for short-term load forecasting system implementation and performance evaluation, IEEE Transactions on Power Systems, 14(3), 1100-1106.
- [42] Vapnik, V. (1995) The Nature of Statistic Learning Theory, Springer-Verlag, New York.
- [43] Vapnik, V., (1998) Statistical Learning Theory, Wiley, New York.
- [44] Cao, L. (2003) Support vector machines experts for time series forecasting, Neurcomputing 51 321–339.
- [45] L. Cao and Gu, Q., (2002) Dynamic support vector machines for non-stationary time series forecasting, Intelligence Data Analysis 6 67–83.
- [46] Tay, F.E.H. and Cao, L. (2001) Application of support vector machines in financial time series forecasting, Omega 29(4) 309–317.
- [47] Pai, P.-F. and Hong, W.-C. (2005) Forecasting regional electric load based on recurrent support vector machines with genetic algorithms, Electric Power Systems Research 74(3) 417–425.

- [48] Pai, P.-F. and Hong, W.-C. (2005) Support vector machines with simulated annealing algorithms in electricity load forecasting, *Energy Conversion Management* 46(17) 2669–2688.
- [49] Yuancheng, L. Tingjian, F. Erkeng, Y. (2002) Short-term electrical load forecasting using least squares support vector machines, in: *Proceedings of the IEEE International Conference on Power System Technology* 1 230–233.
- [50] Wu, H. and Chang, X. (2006) Power load forecasting with least squares support vector machines and chaos theory, in: *Proceedings of the 6th World Congress on Intelligent Control and Automation*, June 21–23, Dalian, China 4369-4373.
- [51] Niu, D.-X., Wang, Q. and Li, J.-C. (2005) Research on support vector machine model based on grey relation analysis for daily load forecasting, in: *Proceedings of the International Conference on Machine Learning and Cybernetic*, 7 of 9, 4254-4259.
- [52] Niu, D.-X., Wanq, Q. and Li, J.-C. (2005) Short term load forecasting model using support vector machine based on artificial neural network, in: *Proceedings of the International Conference on Machine Learning and Cybernetics*, Guangzhou, 7 of 9 4260-4265.
- [53] Tao, X., Renmu, H., Peng, W., and Dongjie, X. (2004) Input dimension reduction for load forecasting based on support vector machine, in: *Proceedings of the 2004 IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies (DRPT)*, Hong Kong, China 510–513.
- [54] Zhang, M.-G. (2005) Short-term load forecasting based on support vector machines regression, in: *Proceedings of the International Conference on Machine Learning and Cybernetics*, 2005 7 of 9 4310-4314.

- [55] Kulkarni, A., Jayaraman, V.K., and Kulkarni, B.D. (2003) Control of chaotic dynamical systems using support vector machines, *Physics Letters A* 317 429–435.
- [56] Yalcintas, M., and Akkurt, S. (2005) Artificial neural networks applications in building energy predictions and a case study for tropical climates, *International Journal of Energy Research* 29(10) 891-901.
- [57] Yalcintas, M. (2008) Energy-savings predictions for building-equipment retrofits, *Energy and Buildings* 40(12) 2111-2120.
- [58] Kreider, J.E., and Rabl, A. (1994) *Heating and Cooling of Buildings*, Chapter 8 McGraw–Hill, New York, USA.
- [59] Hornik, K., Stinchcombe, M., and White, H. (1989) Multilayer feed forward networks are universal approximators, *Neural Networks* 2, 359-366.
- [60] Wang, S.K., Lavan, Z., and Norton, P. (2000) *Air Conditioning and Refrigeration Engineering*, Chapter 7, CRC Press, Boca Raton.
- [61] Wang, S.K., (2000) *Handbook of Air Conditioning and Refrigeration*, 2nd edition, Chapter 6 - Load Calculation, McGraw-Hill, 2000.
- [62] Brown, G.Z. and DeKay, M. (2000) *Sun, Wind & Light: Architectural Design Strategies*, 2nd Edition, John Wiley & Sons, 2000.
- [63] Page, J. (2007) *Simulating Occupant Presence and Behaviour in Buildings*, Thesis No 3900, École Polytechnique Fédérale de Lausanne, Switzerland.
- [64] Ward Systems Group Inc (1996) *Neuroshell 2 Manual*, Frederick, MA.
- [65] Werbos, P.J. (1974) *Beyond Regression: New Tools for Prediction and Analysis in the Behavioural Sciences*, PhD Dissertation, Applied Mathematics, Harvard University, Boston, MA.

- [66] Ansari, F.A., Mokhtar, A.S., Abbas, K.A., and Adam, N.M. (2005) A simple approach for building cooling load estimation, *American Journal of Environmental Sciences* 1(3) 209-212.
- [67] Page, J., Robinson, D., Morel, N., and Scartezzini, J.-L. (2008) A generalized stochastic model for the simulation of occupant presence, *Energy and Building* 40 83-98.
- [68] Robinson, D. (2006) Some trends and research needs in energy and comfort prediction, *Comfort and Energy Use in Buildings*, Windsor, United Kingdom.
- [69] Mahdavi, A., Kabir, E., Mohammadi, A., Lambeva, L., and Progihof, C. (2006) User interactions with environmental control systems in buildings, *PLEA 2006 – The 23rd Conference on Passive and Low Energy Architecture*, Geneva, Switzerland, 6-8 September I-399.
- [70] Humphreys, M.A., and Nicol, J.F. (1998) Understanding the adaptive approach to thermal comfort, *ASHRAE Transactions* 104(1) 991-1004.
- [71] Degelman, L.O. (1999) A model for simulation of daylighting and occupancy sensors as an energy control strategy for office buildings, *Proceedings of Building Simulation 99(2)*, International Building Performance Simulation Association (IBPSA) Conference, Kyoto, Japan 571-578.
- [72] Newsham, G.R. (1994) Manual control of window blinds and electric lighting: Implications for control and energy consumption, *Indoor Environment* 3 135-144.
- [73] Abushakra, B., Sreshthaputra, A., and Haberl, D. (2001) Compilation of diversity factors and schedules for energy and cooling load calculations, Tech. rep., Energy Systems Laboratory, Texas A and M University.
- [74] Roulet, C-A. (1999) Lesson building – report on monitoring, Tech. rep., EPFL/LESO.

- [75] Hunt, D.R.G. (1980) Predicting artificial lighting use – a method based upon observed patterns of behaviour, *Lighting Research and Technology* 12 (1) 7-14.
- [76] Love, J.A. (1998) Manual switching patterns observed in private offices, *Lighting Research & Technology* 30(1) 45-50.
- [77] Reinhart, C.F. (2001) Daylight Availability and Manual Lighting Control in Office Buildings Simulation Studies and Analysis of measurements, Ph.D. thesis, Technical University of Karlsruhe, Germany.
- [78] Reinhart, C.F. (2004) LIGHTSWITCH-2002: A Model for Manual Control of Electric Lighting and Blinds, *Solar Energy* 77 15-28.
- [79] Newsham, G.R. (1995) Lightswitch: A stochastic model for predicting office energy consumption, *Proceedings of Right Light Three, 3rd European Conference on Energy Efficient Lighting*, Newcastle, UK 59–66.
- [80] Pigg, S., Eilers, M., and Reed, J. (1996) Behaviour aspects of lighting and occupancy sensors in private offices: A case study of a university office building, *Proceedings of the 1996 ACEEE Summer Study on energy Efficiency in Buildings* 8 161-170.
- [81] Wang, D., Federspiel, C.C., and Rubinstein, F. (2005) Modeling occupancy in single person offices, *Energy and Buildings* 37 121-126.
- [82] Bourgeois, D., Reinhart C., and Macdonald I. (2006) Adding advanced behaviour models in whole building energy simulation: A study on the total energy impact of manual and automated lighting control, *Energy and Buildings* 38 814-823.
- [83] Lindsay, C.T.R. and Littlefair, P.J. (1992) Occupant use of Venetian blinds in offices, *Building Research Establishment (BRE), Contract PD233/92*, Garston Library, Watford, UK.

- [84] Rea, M.S. (1984) Window blind occlusion: A pilot study, *Building and Environment*, 19(2) 133–137.
- [85] Inoue, T., Kawase, T., Ibamoto, T., Takakusa, S., and Matsuo, Y. (1988) The development of an optimal control system for window shading devices based on investigations in office buildings, *ASHRAE Transaction* 94 1034-1049.
- [86] Rubin, A.I., Collins, B.L., and Tibbott, R.L. (1978) Window blinds as potential energy saver – A case study, *NBS Building Science Series*, 112, National Institute for Standards and Technology, Gaithersburg, MD, USA.
- [87] Fritsch, R., Kohler A., Nygard-Ferguson, M., and Scartezzini, J.-L. (1990) A stochastic model of user behaviour regarding ventilation, *Building and Environment*, 25(2) 173–181.
- [88] Herkel, S., Knapp, U., and Pfafferott, J. (2005) A preliminary model of user behaviour regarding the manual control of windows in office buildings, *Proceedings of The Ninth International IBPSA Conference, Building Simulation 2005*. Montréal, Canada. 403-410.
- [89] Nicol, J.F. (2001) Characterizing occupant behaviour in buildings: Towards a stochastic model of occupant use of windows, lights, blinds, heaters, and fans, *Proceedings of The Seventh International IBPSA Conference, Building Simulation 2001*, Rio de Janeiro, Brazil 1073-1078.
- [90] Afolabi, A.O., Olatunji, B.O., and Ajayi, A.O. (2008) Electricity load forecasting using artificial neural networks, *Journal of Engineering and Applied Sciences* 3(2) 149-154.
- [91] Bhurtun, C., Jahmeerbacus, I, and Jeewooth, C. (2011) Short term load forecasting in Mauritius using Neural Network, , 2011 *Proceedings of the 8th conference on the Industrial and Commercial Use of Energy (ICUE)* 184-191.

- [92] Dillon, T.S., Sestito, S., and Leung, S. (1991) Short-term load forecasting using an adaptive neural network, *Electrical Power & Energy Systems*, 13(4) 186-192.
- [93] Chen, C.S., Tseng, Y.M., and Hwang, J.C. (1996) The application of artificial neural networks to substation load forecasting, *Electric Power Systems Research* 38 153-160.
- [94] Karatasou, S., Santamouris, M., and Geros, V. (2006) Prediction of energy consumption in buildings with artificial intelligent techniques and Chaos time series analysis, *International Workshop on Energy performance and Environmental Quality of Buildings*, Milos island, Greece.
- [95] Karatasou, S., Santamouris, M., and Geros, V. (2006) Modeling and prediction building's energy use with artificial neural networks: Methods and results, *Energy and Buildings* 38(8) 949-958.
- [96] Kandil, N., Wamkeue, R., Saad, M., and Georges, S. (2006) An efficient approach for short-term load forecasting using artificial neural networks, *Electrical Power and Energy Systems* 28 525-530.
- [97] Ekonomou, L., and Oikonomou, D.S. (2008) Application and comparison of several artificial neural networks for forecasting the Hellenic daily electricity demand load, *7th WSEAS Int. Conf. on Artificial Intelligence, knowledge Engineering and Data Bases* ISBN: 978-960-6766-41-1.
- [98] Papalexopoulos, A.D., Hao, S., and Peng, T.M. (1993) Application on neural network technology to short-term system load forecasting, *IEEE/NTUA Athens Power Tech. Conference: Planning, Operation and Control of Today's Electric Power Systems*, Athens, Greece 796-800.

- [99] Yalcinoz, T. and Eminoglu, U. (2005) Short-term and medium-term power distribution load forecasting by neural networks, *Energy Conversion and Management* 46 1393-1405.
- [100] Osman, Z.H., Awad, M.L., and Mahmoud, T.K. (2009) Neural network based approach for short-term load forecasting, *IEEE Xplore*, December 1-8.
- [101] Kiartzis, S.J., Bakirtzis, A.G., and Petridis, V. (1995) Short-term load forecasting using neural networks, *Electric Power Systems Research* 33(1) 1-6.
- [102] Mehta, A.J., Mehta, H.A., Manjunath, T.C., and Ardil, C. (2004) A multi-layer artificial neural network architecture design for load forecasting in power systems, *International Journal of Applied Mathematics and Computer Sciences* 4(4) 227-240.
- [103] Hayati, M. and Karami B. (2005) Application of neural networks in short-term load forecasting, 7th WSEAS Int. Conf. on Mathematical Methods and Computational Techniques in Electrical Engineering, Sofia 37-41.
- [104] Gonzalez, P.A. and Zamarreno, J.M. (2005) Prediction of hourly energy consumption in buildings based on a feedback artificial neural network, *Energy and Building* 37 595-601.
- [105] Al-Shareef, A.S.J., Mohamed, E.A., and Al-Judaibi, E. (2008) Next 24-hours load forecasting using artificial neural network (ANN) for the West Area of Saudi Arabia, *JKAU: Eng. Sci.* 19(2) 25-40.
- [106] Bauman, M. and McClintock, M. (1993) A study of occupant comfort and workstation performance in PG&E's advanced office systems tested final report to PG&E research and development, Center of Environmental Design Research, University of California, Berkeley.

- [107] Siikonen, M.-L. (2000) On traffic planning methodology, International Congress on Vertical Transportation Technologies, ELEVCON Berlin 2000.
- [108] Bourgeois, D., Reinhart C., and Macdonald I. (2006) Adding advanced behaviour models in whole building energy simulation: A study on the total energy impact of manual and automated lighting control, *Energy and Buildings* 38, 814-823.
- [109] Hagan, M.T., and Menhaj, M.B. (1994) Training feedforward networks with the Marquardt algorithm, *IEEE Transactions on Neural Networks* 5(6) 989-993.
- [110] Kwok, S.S.K., and Lee, E.W.M. (2011) A study of the importance of occupancy to building cooling load in prediction by intelligent approach, *Energy Conversion and Management* 52 2555-2564.
- [111] Haldi, F., and Robinson, D. (2011) The impact of occupants' behaviour on building energy demand, *Journal of Building Performance Simulation*, 4(4) 323-338.
- [112] Parys, W. Saelens, D., and Hens, H. (2011) Coupling of dynamic building simulation with stochastic modelling of occupant behaviour in offices – a review-based integrated methodology, *Journal of Building Performance Simulation* 4(4) 339-358.
- [113] Groot, E., Spiekman, M., Opstelten, I. (2008) Dutch research into user behaviour in relation to energy use of residences, *Proceedings PLEA conference – 25th Conference on Passive and Low Energy Architecture*, Dublin, 361.
- [114] Jeeninga, H., Uytendil, M. and Uitzinger, J. (2001) *Energy Use of Energy Efficient Residences*, Report ECN-C-01-072, Energy Research Centre of The Netherlands (ECN), Petten, The Netherlands, 2001.

- [115] Santin, O.G., Itard, L. and Visscher, H. (2009) The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock, *Energy and Buildings* 41 1223-1232.
- [116] Clarke, J.A, Cockroft, J., Conner, S., Hand J.W., Kelly, N.J., Moore, R., O'Brien, T. and Strachan, P. (2001) Control in Building Energy Management Systems: The Role of Simulation, *Proc. Building Simulation 2001*, Rio de Janeiro, 99-106.
- [117] Lutzenhiser, L. (1993) Social and behaviour aspects of energy use, *Annual Review of Energy and the Environment* 18 247-289.
- [118] Linden, A.L., Carlsson-Kanyama, A. and Eriksson, B. (2006) Efficient and inefficient aspects of residential energy behaviour: What are the policy instruments for change?, *Energy Policy* 34 (14) 1918-1927.

Appendix I

A brief introduction to ANNs

Introduction

An ANN model consists of a number of neurons that are interconnected within a neural network. The arrangement of an ANN is as follows:

$$Y = F \sum (p_1w_1 + p_2w_2 + p_3w_3 + p_4w_4),$$

where Y is the sum of p_1w_1 , p_2w_2 , p_3w_3 , and p_4w_4 . F is a transfer function and the sigmoid function is determined by the following equation:

$$F(x) = \frac{1}{1 + e^{-ax}}.$$

Most of the neurons within the network are simple processing units that each take one or more inputs and produce an output. At each neuron, every input has an associated weight that modifies the strength of the input. The neuron simply adds together all of the inputs and calculates an output that it then passes on.

Backpropagation model

The typical neural network shown in Figure 3-1 is a feedforward type of network, in which computations proceed in the forward direction only. There are three layers of neurons: input, hidden, and output. The number of input neurons corresponds to the number of variables in the input vector used to forecast future values. The hidden layer

and neurons play very important roles in the network. The hidden neurons in the hidden layer allow neural networks to perform complicated nonlinear mapping between input and output variables. The output obtained from the output neurons constitutes the network output.

Backpropagation is the most popular and powerful learning algorithm for neural networks. Each step of training on the dataset is called an epoch. The training set has to be a representative of the actual dataset. Backpropagation training is an error-driven algorithm that tries to improve the performance of the neural network by using the weights along its gradient to reduce the total error.

The input data enter into the network via the input layer. Each neuron in the network processes the input data and the resultant values steadily “percolate” through the network, layer by layer, until a result is generated by the output layer. The actual output of the network is then compared to expected output for the given input. This results in an error value. The connection weights in the network are gradually adjusted, working backwards from the output layer, through the hidden layer, to the input layer. This iterative process continues until a reasonable accuracy is achieved. Fine tuning the weights in this way teaches the network how to produce a reasonable prediction for a particular input through network learning.

Appendix II

Configuration of the computational machine

Hardware

Chipset	Intel 965 chipset
CPU	Pentium Core 2, 1.86 GHz, 533 MHz FSB
RAM	2 x 1 GB, DDR 333 MHz
Disk	250 GB, SATA-100, 7200 rpm
DVRW	Read 16x, Write 12x, Rewrite 6x

Software

Operating system	Microsoft Windows 7 (English) Professional
Compiler	MATLAB version R2001a
Documentation	Microsoft Office 2010 Professional
	Adobe Acrobat Reader DC

Appendix III

The use of occupancy space electrical power demand in building cooling load prediction

(Please read following pages)

Appendix IV

Application program scripts for the ANN model

This appendix consists of the program scripts for the ANN model adopted in the study. The listed program scripts are in MATLAB syntax and are executed in the MATLAB version R2011a environment using its internal libraries and ANN toolbox. The program scripts are illustrated in two MATLAB script files.

The script to simulate the hourly cooling energy consumption is as follows:

```
%% Hourly Simulation 2010
% Using feedforward backpropagation function - trainlm
% Script generated by MCL
%
% Few weather data + hr + duct only
% -----
% configurable parameter
%     data_year = Year of the collected data
%     max_partition = Number of partition in each month
%     max_test = Number of test in each partition
% -----
% Output results are stored in:
%     MR = Mean of R (array)
%     MMAPE = Mean of MAPE (array)
%     MRMSPE = Mean of RMSPE (array)
%     sim_min = Min of RMSPE, MAPE, R (array)
%     sim_max = Max of RMSPE, MAPE, R (array)
% -----
clear all

% year of the data (configurable, default = 2010)
data_year = 2010;
% number of partition in each month (configurable, default = 6)
max_partition = 6;
% number of trial for each set of data (configurable, default = 100)
max_test = 100;

% display to screen
fprintf('\n BACKPRO Neural Network - Cooling Load Prediction \n');
fprintf('\n   - Input Parameters: Few weather + hr + duct only \n');
```

```

fprintf('\n - Output Parameters: kwh \n');
fprintf('\n Please wait for a while... creating, training and
simulating Neural Network \n\n');
% display to screen

% load data from file
load wpd2010_2.mat

% input data and target - 365 days x 24 hours = 8760 hours
idata =
[temp;wet_bulb;rad_g;vis;cld;mean_wind;rf;day;occ_duct;occ_pau];
itarget = kwh_lowl ;

% Day Missing in 2010: Day 242 (5785) to 251 (6024)
idata = [idata(:,1:5784), idata(:,6025:end)];
itarget = itarget([1:5784,6025:end]) ;
% Day Missing in 2011: Day 160 (3817) to 161 (3864)
%idata = [idata(:,1:5784), idata(:,6025:end)];
%itarget = itarget([1:5784,6025:end]) ;

% pre-allocation
MR = zeros(max_partition * 12, 1);
MMAPE = zeros(max_partition * 12, 1);
MCV = zeros(max_partition * 12, 1);
sim_min = zeros(max_partition * 12, 3);
sim_max = zeros(max_partition * 12, 3);

% pre-allocation
DayInMonth = zeros(1,12);
% find the number of day in each months (Jan to Nov)
for i = [1 : 11] ;
    DayInMonth(i) = daysact(sprintf('%d/1/%d', i, data_year),
sprintf('%d/1/%d', i + 1, data_year));
end
% find the number of day in December
DayInMonth(12) = daysact(sprintf('12/1/%d', data_year),
sprintf('1/1/%d', data_year + 1));

% two days of Aug and eight days of Sept are missing in 2010
DayInMonth(8) = DayInMonth(8) - 2;
DayInMonth(9) = DayInMonth(9) - 8;
% two days of June are missing in 2011
%DayInMonth(6) = DayInMonth(6) - 2;

%initial for result index and simulation result
rt_index = 1;
sim_Y1 = [];

% initial for loop - 1
m_start_index = 1;
% Loop - 1 - Month
for i = [1 : 12] ;
    % calculate end index
    m_end_index = m_start_index + (DayInMonth(i) * 24) - 1;
    % retrieve monthly data
    mdata = idata(:, m_start_index : m_end_index);
    mtarget = itarget(m_start_index : m_end_index);
    % update month index
    m_start_index = m_end_index + 1;

```

```

% calculate number of day in each partition
max_pday = floor(DayInMonth(i) / max_partition);
% initial number of day in partition (e.g. if number of day is 31
and partition is 6, the result = [5, 5, 5, 5, 5, 6])
DayInPartition(1:max_partition) = max_pday;
DayInPartition(max_partition) = DayInMonth(i) -
sum(DayInPartition(1:(max_partition - 1)));
% initial for loop - 2
p_start_index = 1;
mdata_index = 1 : length(mtarget);
% Loop - 2 - Partition Required
for j = 1 : max_partition
    % calculate end index
    p_end_index = (p_start_index + (DayInPartition(j) * 24) - 1);
    % retrieve simulation data
    P1 = mdata(:, p_start_index : p_end_index);
    T1 = mtarget(p_start_index : p_end_index);
    % retrieve training data
    P = mdata(:, (mdata_index < p_start_index) | (mdata_index >
p_end_index));
    T = mtarget((mdata_index < p_start_index) | (mdata_index >
p_end_index));
    % update Partition index
    p_start_index = p_end_index + 1;
    % loop - 3 - Test Needed
    for k = 1 : max_test
        % create & configure neural network
        %newff(P,T,[S1 S2...S(N-1)],{TF1 TF2...TFN1},
BTF,BLF,PF,IPF,OPF,DDF) takes several arguments
        net = newff(P,T,4,{'tansig','purelin'});
        net.inputs{1}.processFcns =
{'mapminmax','mapstd','processpca'} ;
        net.outputs{2}.processFcns = {'mapminmax','mapstd'} ;
        net.divideFcn = 'dividerand';
        net.divideParam.trainRatio = 60/100;
        net.divideParam.valRatio = 20/100;
        net.divideParam.testRatio = 20/100;
        net.trainFcn = 'trainlm';
        net.performFcn = 'mse';
        %net.trainParam.lr = 0.02 ;
        %net.trainParam.mc = 0.95 ;
        net.trainParam.mu = 0.001;
        net.trainParam.mu_dec = 0.1;
        net.trainParam.mu_inc = 10;
        net.trainParam.mu_max = 1e10;
        net.trainParam.max_fail = 5;
        net.trainParam.min_grad = 1e-10;
        net.trainParam.showWindow = 1;
        net.trainParam.time = inf;
        net.trainParam.goal = 0;
        net.trainParam.epochs = 100;
        net = init(net);
        % training network
        [net,tr] = train(net,P,T);
        % simulate network
        Y = sim(net,P1);

        % display result
        SSE = sse(T1 -Y);
        %RMSE = sqrt(sum((Y-T1).^2)/length(T1));
        CV = (sqrt(sum((Y-T1).^2)/length(T1)))/mean(T1) * 100;

```

```

        %R = 1 - (sum((Y-T1).^2) / sum(T1.^2)) ;
        MAPE = sum(abs((T1-Y)./T1))/length(T1) * 100;
        %RMSPE = ((sqrt(sum((Y-T1).^2)/length(T1)))/mean(Y)) * 100;
        R = sum((-mean(T1) + T1).*(-mean(Y) + Y )) / sqrt( sum((-
mean(T1) + T1).^2) * sum((-mean(Y) + Y ).^2) ) ;

        % display to screen
        fprintf(' Month: %d-%d, SSE : %0.2f%%, CV : %0.2f%%, MAPE :
%0.2f%%, R : %0.2f, Nos. of Sim : %d \n', i, j, SSE,CV,MAPE ,R, k);

        % save result
        if k == 1
            AR = R ;
            AMAPE = MAPE ;
            ACV = CV ;
            sim_Y1 = [sim_Y1, Y];
        else
            AR = cat(2,R,AR) ;
            AMAPE = cat(2,MAPE,AMAPE) ;
            ACV = cat(2,CV,ACV) ;
        end
    end
    % calculate and store average result
    MR(rt_index) = mean(AR) ;
    MMAPE(rt_index) = mean(AMAPE) ;
    MCV(rt_index) = mean (ACV) ;
    % sort simulation result to find max. and min.
    total = [ACV; AMAPE; AR] ;
    total = total(:,1:max_test) ;
    total = total' ;
    totala = sort(total,1) ;
    % store min and max result
    sim_min(rt_index,:) = totala(ceil(0.06 * max_test),:) ;
    sim_max(rt_index,:) = totala(floor(0.95 * max_test),:) ;
    % display result
    fprintf(' Month: %d-%d, mean CV : %0.2f%%, mean MAPE :
%0.2f%% , mean R : %0.2f , nos. of simulation : %d \n', i, j,
MCV(rt_index),MMAPE(rt_index), MR(rt_index), k) ;
    % increment result index by 1
    rt_index = rt_index + 1;
end
end

% plot yearly result using the first simulation of each partition
figure;
plot(itarget,'DisplayName','Actual kwh - 2011','YDataSource','T1');
hold all;
plot(sim_Y1,'DisplayName','Predict kwh - 2011','YDataSource','Y1');
hold all; figure(gcf);
title(' Simulation Result for 2011 ');
xlabel('Day');
ylabel('kWh');

% display to screen
fprintf('\n ~~ Completed ~~ \n\n');

```

The script to simulate the daily cooling energy consumption is as follows:

```
%% Daily Simualtion 2010
% Using feedforward backpropogation function - trainlm
% Script generated by MCL
%
% Few weather data + hr + duct only
% -----
% configurable parameter
%     data_year = Year of the collected data
%     max_partition = Number of partition in each month
%     max_test = Number of test in each partition
% -----
% Output results are stored in:
%     MR = Mean of R (array)
%     MMAPE = Mean of MAPE (array)
%     MRMSPE = Mean of RMSPE (array)
%     sim_min = Min of RMSPE, MAPE, R (array)
%     sim_max = Max of RMSPE, MAPE, R (array)
% -----
clear all

% year of the data (configurable, default = 2010)
data_year = 2010;
% number of partition in each month (configurable, default = 6)
max_partition = 6;
% number of trial for each set of data (configurable, default = 100)
max_test = 100;

% display to screen
fprintf('\n BACKPRO Neural Network - Cooling Load Prediction \n');
fprintf('\n   - Input Parameters: Few weather + hr + duct only \n');
fprintf('\n   - Output Parameters: kwh \n');
fprintf('\n Please wait for a while... creating, training and
simulating Neural Network \n\n');
% display to screen

% load data from file
load wpd2010_2.mat
% input data and target - 365 days x 24 hours = 8760 hours
temp = mean(reshape(temp,24,[]));
wet_bulb = mean(reshape(wet_bulb,24,[]));
rad_g = mean(reshape(rad_g,24,[]));
vis = mean(reshape(vis,24,[]));
cld = mean(reshape(cld,24,[]));
mean_wind = mean(reshape(mean_wind,24,[]));
rf = mean(reshape(rf,24,[]));
day = mean(reshape(day,24,[]));
occ_duct = mean(reshape(occ_duct,24,[]));
occ_pau = mean(reshape(occ_pau,24,[]));
kwh_low1 = sum(reshape(kwh_low1,24,[]));

% input data and target - 365 days x 24 hours = 8760 hours
idata = [temp;wet_bulb;rad_g;vis;cld;mean_wind;rf;day;occ_duct;occ_pau];
itarget = kwh_low1 ;

% Day Missing in 2010: Day 242 (5785) to 251 (6024)
idata = [idata(:,1:241), idata(:,252:end)];
```

```

itarget = itarget([1:241,252:end]) ;

% pre-allocation
MR = zeros(max_partition * 12, 1);
MMAPE = zeros(max_partition * 12, 1);
MCV = zeros(max_partition * 12, 1);
sim_min = zeros(max_partition * 12, 3);
sim_max = zeros(max_partition * 12, 3);

% pre-allocation
DayInMonth = zeros(1,12);
% find the number of day in each months (Jan to Nov)
for i = [1 : 11] ;
    DayInMonth(i) = daysact(sprintf('%d/1/%d', i, data_year),
    sprintf('%d/1/%d', i + 1, data_year));
end
% find the number of day in December
DayInMonth(12) = daysact(sprintf('12/1/%d', data_year),
    sprintf('1/1/%d', data_year + 1));

% two days of Aug and eight days of Sept are missing in 2010
DayInMonth(8) = DayInMonth(8) - 2;
DayInMonth(9) = DayInMonth(9) - 8;

%initial for result index and simulation result
rt_index = 1;
sim_Y1 = [];

% initial for loop - 1
m_start_index = 1;
% Loop - 1 - Month
for i = [1 : 12] ;
    % calculate end index
    m_end_index = m_start_index + (DayInMonth(i) ) - 1;
    % retrieve monthly data
    mdata = idata(:, m_start_index : m_end_index);
    mtarget = itarget(m_start_index : m_end_index);
    % update month index
    m_start_index = m_end_index + 1;
    % calculate number of day in each partition
    max_pday = floor(DayInMonth(i) / max_partition);
    % initial number of day in partition (e.g. if number of day is 31
    and partition is 6, the result = [5, 5, 5, 5, 5, 6])
    DayInPartition(1:max_partition) = max_pday;
    DayInPartition(max_partition) = DayInMonth(i) -
    sum(DayInPartition(1:(max_partition - 1)));
    % initial for loop - 2
    p_start_index = 1;
    idata_index = 1 : length(itarget);
    % Loop - 2 - Partition Required
    for j = 1 : max_partition
        % calculate end index
        p_end_index = (p_start_index + (DayInPartition(j)) - 1);
        % retrieve simulation data
        P1 = mdata(:, p_start_index : p_end_index);
        T1 = mtarget(p_start_index : p_end_index);
        % retrieve training data
        P = idata(:, (idata_index < p_start_index) | (idata_index >
        p_end_index));
        T = itarget((idata_index < p_start_index) | (idata_index >

```

```

p_end_index ));
    % update Partition index
    p_start_index = p_end_index + 1;
    % loop - 3 - Test Needed
    for k = 1 : max_test
        % create & configure neural network
        %newff(P,T,[S1          S2...S(N-1)],{TF1          TF2...TFN1},
BTF,BLF,PF,IPF,OPF,DDF) takes several arguments
        net = newff(P,T,4,{'tansig','purelin'});
        net.inputs{1}.processFcns =
{'mapminmax','mapstd','processpca'} ;
        net.outputs{2}.processFcns = {'mapminmax','mapstd'} ;
        net.divideFcn = 'dividerand';
        net.divideParam.trainRatio = 60/100;
        net.divideParam.valRatio = 20/100;
        net.divideParam.testRatio = 20/100;
        net.trainFcn = 'trainlm';
        net.performFcn = 'mse';
        %net.trainParam.lr = 0.02 ;
        %net.trainParam.mc = 0.95 ;
        net.trainParam.mu = 0.001;
        net.trainParam.mu_dec = 0.1;
        net.trainParam.mu_inc = 10;
        net.trainParam.mu_max = 1e10;
        net.trainParam.max_fail = 5;
        net.trainParam.min_grad = 1e-10;
        net.trainParam.showWindow = 1;
        net.trainParam.time = inf;
        net.trainParam.goal = 0;
        net.trainParam.epochs = 100;
        net = init(net);
        % training network
        [net,tr] = train(net,P,T);
        % simulate network
        Y = sim(net,P1);

        % display result
        SSE = sse(T1 -Y);
        %RMSE = sqrt(sum((Y-T1).^2)/length(T1));
        CV = (sqrt(sum((Y-T1).^2)/length(T1)))/mean(T1) * 100;
        %R = 1 - (sum((Y-T1).^2) / sum(T1.^2)) ;
        MAPE = sum(abs((T1-Y)./T1))/length(T1) * 100;
        %RMSPE = ((sqrt(sum((Y-T1).^2)/length(T1)))/mean(Y)) * 100;
        R = sum((-mean(T1) + T1).*(-mean(Y) + Y )) / sqrt( sum((-
mean(T1) + T1).^2) * sum((-mean(Y) + Y ).^2) ) ;

        % display to screen
        fprintf(' Month: %d-%d, SSE : %0.2f%%, CV : %0.2f%%, MAPE :
%0.2f%%, R : %0.2f, Nos. of Sim : %d \n', i, j, SSE,CV,MAPE ,R, k);

        % save result
        if k == 1
            AR = R ;
            AMAPE = MAPE ;
            ACV = CV ;
            sim_Y1 = [sim_Y1, Y];
        else
            AR = cat(2,R,AR) ;
            AMAPE = cat(2,MAPE,AMAPE) ;

```

```

        ACV = cat(2,CV,ACV) ;
    end
end
% calculate and store average result
MR(rt_index) = mean(AR) ;
MMAPE(rt_index) = mean(AMAPE) ;
MCV(rt_index) = mean (ACV) ;
% sort simulation result to find max. and min.
total = [ACV; AMAPE; AR] ;
total = total(:,1:max_test) ;
total = total' ;
totala = sort(total,1) ;
% store min and max result
sim_min(rt_index,:) = totala(ceil(0.06 * max_test),:) ;
sim_max(rt_index,:) = totala(floor(0.95 * max_test),:) ;
% display result
fprintf(' Month: %d-%d, mean CV : %0.2f%%, mean MAPE :
%0.2f%% , mean R : %0.2f , nos. of simulation : %d \n', i, j,
MCV(rt_index),MMAPE(rt_index), MR(rt_index), k) ;
% increment result index by 1
rt_index = rt_index + 1;
end
end

% plot yearly result using the first simulation of each partition
figure;
plot(itarget,'DisplayName','Actual kwh - 2010','YDataSource','T1');
hold all;
plot(sim_Y1,'DisplayName','Predict kwh - 2010','YDataSource','Y1');
hold all; figure(gcf);
title(' Simulation Result for 2010 ');
xlabel('Day');
ylabel('kWh');

% display to screen
fprintf('\n          ~~          Completed          ~~          \n\n');

```

Appendix V

Building data for simulation model

Hyperlink of the building data 2010 & 2011 is as below:

https://drive.google.com/open?id=0B_4yDKtEJPzJRGVjZkZvUDRSVIU