Energy Efficient Operation of Data Centres – Technical, Computational and Political Challenges

Morgan Rhys Tatchell-Evans

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Chemical and Process Engineering

School of Mechanical Engineering

August 2017

Work formed from jointly authored publications

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Sections 1.1, 1.2.1, 1.3, and Chapters 2 and 3 are based on work published in:

M. Tatchell-Evans, N. Kapur, J. Summers, H. Thompson, and D. Oldham, "An experimental and theoretical investigation of the extent of bypass air within data centres employing aisle containment, and its impact on power consumption," Applied Energy, vol. 186, no. 3, pp. 457–469, 2017.

All laboratory experiments referred to in this article were carried out by the candidate. The experimental methods and the system model described in the article were developed by the candidate under the guidance of the other named authors. The candidate initially produced a draft copy of the entire article, with the other authors providing guidance and taking part in editing successive drafts.

Acknowledgements

I would like to take this opportunity to thank my supervisors Professor Nik Kapur, Dr Jon Summers, Professor Harvey Thompson and Professor Peter Taylor. It has been a pleasure and a privilege to work with all of you, and your support and guidance has been invaluable. I also thank Dr Gemma Brady, Dr Yaser Al-Anii and Mustafa Khadim for sharing their experience and insights with me.

Besides my colleagues at the University, I have had the good fortune to receive the support of many industrial partners. The guidance of Dan Oldham and Geoff Fox at Digiplex has helped to ensure that the outcomes of this research are of direct value to the data centre sector. Patrick France, Lance Davis and Amy Pollard at Teledata provided me with a fantastic opportunity to immerse myself in a live data centre when this project was in its infancy. Mark Seymour at Future Facilities has been generous in sharing his vast knowledge in the field of CFD modelling of data centres.

I owe a great debt of gratitude to my fellow PhD students on the Low Carbon Technologies DTC, who have enabled me to get through the past few years largely with a smile on my face. They are perhaps too numerous to name in full, but I would like to especially thank my beloved cohort, Dave, Jamie, Stephen, Josh, Lloyd, Ben, Hattie and Eddy. In addition to the DTC crowd, I thank Kat for her friendship and company.

I am also fortunate enough to benefit from the love and support of a wonderful family. In particular I thank my parents, who have given me the strength and stability without which it would be difficult to contemplate taking on a project such as this.

Finally I thank Gina for bringing me great joy and contentment, and for giving me a good reason to come home every evening and leave work behind.

Abstract

The aim of this study was to investigate the technologies and policy instruments available to improve efficiency in data centres. Data centres consume a significant and increasing proportion of the world's electricity, and much of this electricity is consumed in cooling the computing equipment housed in these facilities. Significant potential exists to improve the efficiency of cooling in data centres.

One popular method for improving efficiency in data centre cooling is to physically separate the hot and cold air streams using 'aisle containment' systems. This has been shown to reduce 'bypass' (cold air returning to the air conditioning system without having passed through any computing equipment) and 'recirculation' (rejected hot air returning to computing equipment inlets, leading to over-heating). However, the benefits of aisle containment have not previously been extensively quantified, nor have the optimal operational conditions been investigated.

Experimental investigations were undertaken to determine the extents of bypass and recirculation in data centres employing aisle containment. Effective measures for minimising bypass and recirculation in such data centres were identified. A system model was developed to predict the impacts of this bypass and recirculation on data centre electricity consumption. The system model results showed that taking action to minimise bypass could reduce electricity consumption by up to 36%, whilst minimising the pressurisation of the cold aisles could reduce electricity consumption by up to 58%.

Computational fluid dynamics models were developed to further investigate the implications of aisle containment, for both electricity consumption and cooling efficacy. Significant advancements in the techniques used to model bypass and recirculation within contained systems have been made.

Finally, interviews were undertaken with data centre operators, in order to enable an assessment of the current policy environment pertaining to energy efficiency in UK data centres. Recommendations have been made for potential improvements to this policy environment.

Table of Contents

Work formed from jointly authored publications ii

Acknowledgements iii

Abstract.. iv

Table of Contents... v

List of Tables xi

List of Figures xiii

Nomenclature xxi

1. The Data Centre and Energy Efficiency.. 1

   1.1   Introduction... 1

   1.2   Electricity consumption and energy efficiency in data centres 1

      1.2.1   How much electricity do data centres consume?.. 1

      1.2.2   Attitudes towards energy efficiency in data centres 3

      1.2.3   How is electricity consumed in data centres? 4

   1.3   Technologies and Practices for Efficient Data Centres 5

      1.3.1   Data centre cooling.. 5

      1.3.2   Layout and Management... 6

      1.3.3   Supply air temperature.. 7

      1.3.4   Minimizing IT Power Consumption... 8

   1.4   Data Centre Efficiency Metrics 9

   1.5   Research aims and objectives 15

   1.6   Thesis outline 16

2. Bypass and Recirculation in Data Centres Employing Aisle Containment 18

   2.1   Introduction.. 18

      2.1.1   Air handling in aisle contained data centres.. 18

      2.1.2   Experimental and computational investigations of aisle containment published to date 20

      2.1.3   Server rack design 23

ix

List of Tables

List of Figures

Nomenclature

Abbreviations

ASHRAE	American Society of Heating, Refrigeration and Air-conditioning Engineers
BAU	Business as usual
BREEAM	Building Research Establishment Environmental Assessment Method
CADE	Corporate Average Data Center Efficiency
CAGR	Compound annual growth rate
CCA	Climate change agreement
CCL	Climate change levy
CFD	Computational fluid dynamics
CO_2	Carbon dioxide
CO_2e	Carbon dioxide equivalent
COP	Coefficient of performance
CPU	Central processing unit
CRAC	Computer room air conditioner
CRAH	Computer room air handler
CRC	Climate reduction commitment
CSR	Corporate social responsibility
DC CCA	Umbrella Climate Change Agreement for the standalone Data Centres
DCA	Data Centre Alliance
DCP	Data Centre Productivity
EEM	Energy efficiency measure
EPBD	Energy Performance of Buildings Directive
ESOS	Energy Savings Opportunity Scheme
EU	European Union
EU CoC	EU Code of Conduct on Data Centre Energy Efficiency
EU-ETS	European Union Emissions Trading Scheme
EVM	Eddy viscosity model
GFLOPS/W	Giga-floating point operations per Watt
GHG	Greenhouse gas
GWh	Gigawatt hours
HACA	Hot aisle-cold aisle

HVAC	Heating, ventilation and air conditioning	
IT	Information technology	
itEUE	IT Energy Usage Effectiveness	
M&E	Mechanical and electrical	
NHER	National Home Energy Rating	
NS	Navier-Stokes	
OHAD	Overhead air distribution	
PF	Potential flow	
pPUE	Partial power usage effectiveness	
PUE	Power usage effectiveness	
RANS	Reynolds averaged Navier-Stokes	
Re	Reynolds number	
RHI	Return heat index	
RTI	Return temperature index	
SHI	Supply heat index	
SLA	Service level agreement	
TPFR	Toulouse potential flow replica	
TWh	Terawatt hours	
U	Rack unit, 44.45 mm	
UPS	Uninterruptible power supply	
US EPA	United States Environmental Protection Agency	
VFD	Variable frequency drive	

Symbols

a	Width of channel	m
A	Area	m^2
b	Height of channel	m
c_p	Specific heat capacity	$kJ.kg^{-1}.K^{-1}$
$C_{in/out}$	Product of c_p and mass flow rate into/out of a cell	$kW.K^{-1}$
D	Length of cell side	m
D_h	Hydraulic diameter	m

e	Specific internal energy	$J.kg^{-1}$
E	Electricity consumption	KWh
$\dot{E}$	Rate of electricity consumption	W
f	Friction factor	-
F	Body force	$kg.m^{-2}.s^{-2}$
$F_{\delta m}$	Buoyancy force	N
h	Heat transfer coefficient of heat exchanger	$kW.K^{-1}$
H	Height	m
k	Loss coefficient	-
k_t	Rate of production of turbulent kinetic energy	$kW.kg^{-1}$
K	Thermal conductivity	$kW.m^{-1}.K^{-1}$
L	Length	m
$LMTD$	Logarithmic mean temperature difference	K
$\dot{m}$	Mass flow rate	$kg.s^{1}$
m_R	Roughness	-
n	Index	-
N	Number of nodes	-
$\boldsymbol{N}$	Number of servers	-
p	Pressure	Pa
P_t	Rate of production of turbulent kinetic energy	$kW.kg^{-1}$
q	Scalar velocity	$m.s^{-1}$
Q	Heat dissipation	J
$\dot{Q}$	Rate of heat dissipation/transfer/load	kW
r	Distance from centre of circle	m
R	Radius of circle	m
R_D^2	Coefficient of determination	-
$R_\emptyset$	Ø residual	$m^2.s^{-1}$

Re	Reynolds number	-
S_T	Heat source	$K.s^{-1}$
T	Temperature	K
u	Velocity in the x direction	$m.s^{-1}$
$\boldsymbol{u}$	Velocity vector	$m.s^{-1}$
U	Heat transfer coefficient of Test Data Centre	$kW.m^{-2}.K^{-1}$
v	Velocity in the y direction	$m.s^{-1}$
v_t	Kinematic viscosity	$m^2.s^{-1}$
V	Volume	m^3
$\dot{V}$	Volumetric flow rate	$m^3.s^{-1}$
w	Velocity in the z direction	$m.s^{-1}$
W	Width	m
W_{SOR}	Relaxation factor	-

Greek symbols

β	Volumetric thermal expansion coefficient	K^{-1}
Δ	Difference	-
ε	Rate of turbulent dissipation of kinetic energy	$kW.kg^{-1}$
ε_T	Maximum error on inlet temperature predictions	K
λ_v	Bulk viscosity coefficient	$Pa.s$
η	Efficiency	-
μ	Dynamic viscosity	$Pa.s$
μ_t	Eddy viscosity	$Pa.s$
ρ	Density	$kg.m^{-3}$
τ	Stress	Pa

ϕ	Velocity potential	$m^2.s^{-1}$

Subscripts

amb	Ambient
BP	Bypass
C	Characteristic
con	Contraction
CA	Cold aisle
CH	Between the cold and hot aisles
CW	Chilled water
eco	Economiser
exp	Expansion
HA	Hot aisle
i	Index in the x direction
I	Interface
j	Index in the y direction
k	Index in the z direction
l	Laminar
LB	Load bank
$leak$	Leakage
max	Maximum
o	Interface open
p	Periphery
ref	Reference
req	Required
s	Interface sealed

$spec$ Manufacturer specification

t Turbulent

tdc Test Data Centre

T Total

θ Tangential

1. THE DATA CENTRE AND ENERGY EFFICIENCY

1.1 Introduction

Data centres are buildings which facilitate the operation of large quantities of computing equipment, and form the backbone of today's digital infrastructure [1]. The information technology (IT) equipment housed in these facilities is used to process, store and transmit data on behalf of a broad range of end users, such as internet service providers, banks, corporations, educational institutions, governmental bodies and individuals [1]. They are energy intensive facilities, with typical power densities of 540-2200 $W.m^{-2}$, and extreme cases exceeding 10 $kW.m^{-2}$ [2]. The sector has grown rapidly in recent years [3], and its electricity consumption has attracted the concern and attention of policy makers around the world [4] [5]. The electricity consumption is largely attributable to the IT equipment itself, and to the cooling infrastructure required to maintain an appropriate environment for operation of this equipment [6].

The data centre sector may broadly be divided into two distinct groups: enterprise data centres and co-location providers [7]. Enterprise data centres are owned and operated by the company for which they are providing services. Co-location providers are specialist data centre operators, who lease space within their facilities to customers seeking a suitable environment in which to operate their IT equipment.

1.2 Electricity consumption and energy efficiency in data centres

1.2.1 How much electricity do data centres consume?

Whilst the total energy consumption of the data centre sector is unknown, indications suggest that it is large, and has undergone significant growth in recent years. Estimating the electricity consumption of data centres is inherently difficult, partly because much of the sector is made up of enterprise data centres, in which case their power consumption may be reported collectively with office electricity consumption [8]. However, some attempts have been made to estimate electricity consumption in the sector, with one report from as long ago as 2006 suggesting that data centres in the United States consumed around 42 gigawatt hours (GWh) per year [9].

More recently, research supported by the United States Department of Energy estimated that by 2014 the US data centre sector's annual electricity consumption was as high as 70

terawatt hours (TWh) per year, or 1.8% of total US electricity consumption [10]. The research also found that there had been a 4% increase in the sector's consumption between 2010 and 2014, contrasted with a 24% increase from 2005 to 2010. The slowdown in growth was attributed to a move towards virtualisation (discussed in section 1.3.4), and transfer of services from small, inefficient, enterprise data centres to larger, more efficient, co-location facilities. The report also stresses that demand for data centre services has increased rapidly since 2010, and that this would have led to very large increases in electricity consumption if not for advances in energy efficiency.

Data centre electricity consumption within Europe was estimated at around 56 TWh per year in 2007, with this figure expected to rise to 104 TWh per year by 2020 (Bertoldi 2010, cited in [11]). A more recent report into EU data centre power consumption, carried out by Germany's Borderstep Institute for the country's Federal Ministry of Economic Affairs, estimated that the sector consumed around 65 TWh in 2015 [12]. This is 2.4% of the 2,742 TWh of EU electricity consumption in 2015 [13]. EU data centre electricity consumption has risen more rapidly than that in the US in recent years, with a rise of approximately 20% from 2010 to 2015 [12].

The Borderstep Institute's report also estimated the UK data centre sector's electricity consumption at 11 TWh per year in 2015. TechUK's Emma Fryer [14] in 2016 made a much lower estimate of 6 TWh of electricity per year, or 2.0% of total UK electricity consumption. Fryer noted however that there is great uncertainty around any estimates, due to the limited information available. In the UK, data centres account for a particularly large proportion of total electricity consumption, as shown in Table 1-1. This is thought to result from the UK being an attractive destination for data centre operators and organisations dependent upon information technology services, due to its having a highly skilled work force, good data connections with the rest of the world and a stable political environment (at least prior to the country's EU membership referendum of 2016) [15].

Global data centre electricity consumption for the year 2012 was recently estimated as 268 TWh [16], or 1.4% of total global electricity consumption during that year [17]. The figures for data centre electricity consumption in different regions, as discussed in the preceding paragraphs, are collected in Table 1-1.

The compound annual growth rate (CAGR) of electricity consumption in the world data centre sector from 2007-2012 was recently estimated as 4.4% [16], much higher than the 2.1% projected for total global electricity demand from 2012-2040 [18]. If the 4.4% CAGR for data centre electricity consumption were to persist, this would result in an annual

electricity consumption of 855 TWh per year by 2040. This is 2.5% of the 34,250 TWh of predicted total global electricity consumption in 2040, if current policies, aims and intentions are carried out [18]. Hence, if the growth rates of recent years persist, the importance of data centre electricity consumption will grow, and will have serious implications for efforts to reduce carbon emissions and tackle climate change over the coming decades [19], [20]. Accordingly, governments have begun to take action over recent years to drive efficiency improvements in data centres, with the aim of reducing costs and environmental impact [5], [21], [22]. This action has come in the form of policy instruments such as tax incentives [22], voluntary codes of conduct [4], and training and education programmes [23].

Country/ region	Electricity Consumption (TWh/year)	Proportion of total electricity consumption (%)	Year of data
UK	6-11 [12], [14]	2.0-3.7 [12], [14]	2015-16
EU	59 [12]	1.9 [24]	2014
US	70 [10]	1.8 [10]	2014
World	268 [16]	1.4 [16] [17]	2012

Table 1-1. Estimates of data centre electricity consumption by region.

When discussing the electricity consumption of data centres, it is important to note that the services provided by these facilities can themselves help to reduce energy consumption. Macroeconomic studies have suggested that IT services can help to reduce greenhouse gas emissions by enabling activities such as building management, smart grids, virtual goods and teleconferencing [25]–[28]. However, as has been shown in this section, the current extent of and trends in data centre electricity consumption mean that allowing growth to continue unchecked would have serious consequences for attempts to limit global energy consumption and GHG emissions.

1.2.2 Attitudes towards energy efficiency in data centres

The Uptime Institute's [29] 2013 survey of data centre professionals found that the majority of respondents regard reduction of energy consumption as being "very important", with the percentage varying from 50% in North America to 71% in Asia. However, in the same survey energy efficiency is ranked as only the 4th most important factor in consideration of expansion efforts, behind reliability, up-front cost and long-term cost. The report notes that respondents to their survey are generally from larger, more advanced data centres, which are likely to be more focused on energy consumption than others within the industry.

The importance of reliability relates to the need to avoid 'downtime', or breaks in the service provided by the data centre. Data centres often supply services to organisations which require them 24 hours a day, with any downtime leading to a loss of business for these organisations. For an enterprise data centre, this directly impacts profits. Co-location facilities often have contracts in place with their customers, called Service Level Agreements (SLAs), which detail the level of service to be provided by the data centre [30]. SLAs specify factors such as the time to respond to queries and the amount of downtime allowable, as well as what penalties will be incurred by the data centre if these stipulations are not met. Hence downtime can directly cause a loss of revenue, as well as being damaging to the co-location provider's reputation. Resultantly, the avoidance of downtime is of great importance to all data centre operators, with energy efficiency only recently being seen as of comparable importance [6].

Downtime may be caused by power failures, but also cooling failures, since IT equipment may fail or shut down automatically if overheating occurs [31]. This leads data centre operators to strive for uninterruptible power and cooling infrastructure [32]. Data centres have varying levels of redundancy in place in their power and cooling systems in order to achieve this, the costs of which must be balanced against the potential costs of downtime. Energy Efficiency Measures (EEMs) often have knock-on effects on reliability, which must be taken into account when making decisions regarding the technology to be used in a facility [30], [32].

1.2.3   How is electricity consumed in data centres?

In data centres, electrical power is used for two main purposes. Firstly, it is consumed directly by the computing equipment. Secondly, since this equipment converts the electricity into heat, electricity is consumed in removing the heat from the data centre. Audits suggest that cooling typically accounts for 20-50% of electricity consumption [6]. In addition, significant amounts of electricity may be consumed by the power delivery system, although an efficient, modern system can be up to 99% efficient [33]. Efficiencies of >90% are common, although this can be compromised where a facility is working at well below its power capacity [33].

Electricity consumption in data centre cooling systems will form the focus of this thesis. The efficiency of cooling is a major focus of efforts to reduce data centre electricity consumption, with good practice regarding air management, cooling equipment operating conditions and selection of efficient equipment receiving the attention of academics, industry representatives and best practice guidelines [34], [35].

1.3 Technologies and Practices for Efficient Data Centres

1.3.1 Data centre cooling

The primary sources of heat in data centres are the central processing units (CPUs) and hard drives of the computer servers. Servers usually have internal fans to direct air flow past the hottest components [1]. CPUs are usually the hottest components, and usually reject heat to a heat sink made from extruded metal, which is bonded to the CPU using a thermal interface material. CPUs can also be cooled using a 'heat pipe' – a pipe filled with a liquid, with one end placed near the CPU and another at a remote location within the server casing. The liquid at the CPU end is heated to vaporization and is thus transported to the colder end, where it condenses and is returned by capillary action. Another solution is the use of a 'cold plate', which is a liquid filled chamber bonded to the CPU, through which cold liquid flows [1], [32].

Air enters the servers via their inlets (typically at the front), and exits at the outlets (typically at the back) at an increased temperature. Hot air leaving the servers is typically cooled by computer room air conditioning (CRAC) units using mechanical cooling [36]. A more efficient alternative (or supplement) to mechanical cooling is economizer cooling. This can be in the form of either water-side or air-side economization [32]. Water-side economisation involves the rejection of heat from a chilled water loop to ambient air, either directly or via a mechanical refrigeration loop [32]. The chilled water loop is then used to cool the air supplied to the data centre. Air-side economizers can either bring ambient air directly into the data centre, or pass it through a heat exchanger with air from the data centre [36]. The former solution requires filtration of the ambient air to remove contaminants, and even then can lead to premature failure of IT equipment. Humidity control can also be expensive in this method. Economizers tend to increase capital costs, with implementation being especially expensive in retrofit projects [32], [36]. In climates where the economizer will not meet the cooling demand all year round, controlling the transition from economizer to mechanical cooling can be problematic, with the possibility of sharp fluctuations in supply air conditions [36]. Where a CRAC is not used, air is supplied to the data hall by a computer room air handler (CRAH), which contains a heat exchanger enabling the rejection of heat from the process air to the ambient air or chilled water.

Whether mechanical cooling or economizers are used, cool air is usually supplied to a raised floor plenum and conducted into the server room through perforated floor grilles [32]. Less

commonly, cold air may be supplied via overhead ducting. The disadvantage of the latter approach is inflexibility: it is relatively simple to move or add grilles in a plenum based system in order to react to changes in IT provision or layout, whereas rearranging ducting is time-consuming and expensive [31]. Overhead Air Distribution (OHAD) has the advantage that there is greater control over the pressure and temperature of air supplied to different regions, due to the possibility for branched distribution [36].

The use of CRAC or CRAH units with variable frequency drives (VFDs) can lead to reduced power consumption since they allow the speed of CRAC/CRAH fans to be varied according to the dynamic cooling requirement of the data centre. This advantage must be balanced against the greater capital costs of such units [36].

Having conveyed the cold air to the server room, it must then find its way to the equipment requiring cooling. In order to maximize efficiency, the aim is to avoid "bypass" – cold air being returned to the cooling infrastructure without having passed through IT equipment in need of cooling [36]. In addition, the reliability of the IT equipment requires the avoidance of "recirculation" – whereby hot air from server outlets mixes with supplied cold air and enters a server inlet [36]. Bypass is known to impair the efficiency of a data centre's cooling system [37]–[40], since it increases the rate at which cold air must be supplied in order to ensure that sufficient air is available to cool the servers. Efforts to minimise bypass must be balanced against the need to avoid recirculation, since this can cause servers to fail due to over-heating [36]. Where recirculation is prevalent, data centre managers are likely to compensate by reducing the supply temperature of conditioned air, which reduces the efficiency of operation of the cooling infrastructure [41]. The goal is thus to distribute the cold air in such a way as to minimise the supply flow rate of cold air which is required in order to achieve an acceptably small level of recirculation, keeping server inlet temperatures within acceptable limits. Improvements in air management have been highlighted both as a cause of recent improvements in data centre cooling efficiency and as an area in which further efficiency improvements can be made [29].

1.3.2 Layout and Management

The extents of bypass and recirculation in a data centre are affected by many factors, with the key considerations being discussed in this section. Essentially, the aim is to ensure that sufficient cool air is available where and when it is needed, whilst minimizing the extent of over-provisioning and bypass.

One method for reducing bypass and recirculation is to adopt a hot aisle-cold aisle (HACA) arrangement. This involves aligning adjacent rows of racks in opposite directions, such that

server inlets face each other into aisles in which cold air is supplied (termed cold aisles). The servers contain fans which draw the cold air over the heat generating components. Server outlets then face each other in hot aisles, from which hot air returns to the CRAC/CRAHs [36]. CRAC/CRAHs are typically positioned at the perimeter of the room.

Segregation of hot and cold air streams is increasingly being further enhanced through the introduction of solid barriers separating hot and cold aisles, commonly referred to as aisle containment systems [29], [37], [39]. Aisle containment is discussed further in Chapter 2.

Where cool air is distributed via a raised floor plenum, the air flows in this plenum are critical to the effectiveness and efficiency of the cooling. If air is distributed in such a way as to ensure sufficient airflow is achieved at the foot of each rack, recirculation is unlikely. If the flow rates through the servers in the rack exceed the rate of supply of cold air, recirculation is assured, regardless of what measures are undertaken above the plenum [31].

Even if sufficient cooled air is supplied through the plenum, some recirculation of warm air can occur at the ends of the aisle, and over the tops of the racks [31], [37], [42], [43]. There has been considerable research undertaken into the potential to control the distribution of cold air from an underfloor plenum via measures such as adding carefully positioned obstructions in the plenum to guide flow [31], managing the positioning of pre-existing obstructions in the plenum (such as pipes and cables) [44], varying the plenum height [31], tuning the grille open areas [31], and varying the ceiling height [45]. The relative merits of overhead and underfloor distribution have also received some attention [38], [46].

Ideally, the distribution of IT equipment in the facility should take into account the availability of cool air in different parts of the room. Distributing high power racks amongst those with lower consumption can help to make it easier to supply sufficient cool air across the whole data centre. Additionally, where the capacity exists to supply air at different temperatures in different regions of the data centre, equipment with especially stringent supply air temperature requirements may be segregated to avoid having to meet these requirements across the whole facility [36].

1.3.3 Supply air temperature

The reliability of equipment housed in data centres is affected by the inlet air temperatures (T_{inlet}) and humidities to which they are exposed [47]. These parameters are largely guided by recommendations from the American Society of Heating, Refrigeration and Air-conditioning Engineers (ASHRAE), which are themselves guided by information from IT equipment manufacturers regarding the effect of T_{inlet} on failure rates [48]. In 2011,

ASHRAE's recommendations were relaxed, so that a wider range of inlet temperatures is now deemed safe. This has allowed many facilities to increase the supply air temperature, which can improve the coefficient of performance (COP) of the cooling system [41]. The most recent recommendations are shown in Table 1-2. The 'recommended' temperature ranges refer to temperatures which are permissible as a long term state. The 'allowable' range is deemed to be acceptable for limited periods. The 'Class' column refers to the level of reliability offered by the facility, with A1 offering the greatest reliability.

Class	IT Equipment	Recommended / Allowable Temperature range (°C)	Allowable Relative Humidity (%)
A1	Enterprise servers, storage products	18-27 / 15-32	8-80
A2	Volume servers, storage products, personal computers, workstations	18-27 / 10-35	8-80
A3	Volume servers, storage products, personal computers, workstations	18-27 / 5-40	8-85
A4	Volume servers, storage products, personal computers, workstations	18-27 / 5-40	8-90

Table 1-2. ASHRAE data centre IT inlet temperature and humidity ranges [49].

The reductions in cooling power consumption afforded by increased supply air temperature must be offset against the increase in IT power consumption which results from increasing T_{inlet}. This increase is caused by increased server fan speeds, and increased leakage current [50]–[54]. The overall effect of increasing T_{inlet} depends on parameters such as the overall efficiency of cooling and power infrastructure and the algorithms used to control server fans, but generally increasing supply air temperature towards the higher end of the ASHRAE limits reduces total data centre electricity consumption (E_T) [51], [54].

1.3.4   Minimizing IT Power Consumption

Poor utilisation of IT resources can be a major source of inefficiency [55]. Utilisation here refers to the proportion of time during which servers are performing useful work [56]. Hence, poor utilisation implies the wasteful consumption of energy within idling servers, which also presents a burden to the cooling and power distribution infrastructure [56]. One

study from VMware [57] has suggested that the majority of data centre servers may have processor utilisation rates of less than 5%, whilst the Uptime Institute estimates that around 20% of servers may be obsolete, outdated or unused [29]. Improvements in utilisation rates allow the IT, cooling and power distribution infrastructure to be reduced, thus reducing energy consumption [58]. Capital costs of this infrastructure are also reduced.

One method for improving utilisation is virtualisation. Virtualisation is the creation of virtual versions of servers, storage devices and network resources [30]. For example, a virtualised server can run on an isolated partition within a physical server, and can be moved between physical servers [59]. Without virtualisation, each piece of IT equipment may be assigned to a certain task, which may not need to be completed continuously, which leads to time spent idling.

The European Commission's Code of Conduct on Data Centre Efficiency recommends virtualisation, shutdown of idle IT equipment and minimisation of data storage in order to reduce IT power consumption [21]. It goes on to recommend that efficiency, operating temperature and humidity ranges should be considered during selection of new IT equipment, as well as its air flow direction and how compatible this is with the air flow design of the facility [21].

The computational efficiency of servers also varies widely between models, with more efficient servers tending to incur greater capital cost [60]. Trends in computational efficiency have tended to approximate an exponential improvement over time, analogous to Moore's Law for the density of transistors on integrated circuits, and sometimes referred to as Koomey's law [60], [61].

1.4 Data Centre Efficiency Metrics

Efficiency metrics can be useful for self-improvement, comparison between data centres, or site selection/design [62]. For a metric to be useful, improvement of performance against the metric must result in "measurable gains", and it must be possible to measure or at least accurately estimate all parameters required to evaluate the metric [62]. The measurement of the metric should ideally not increase the power consumption of the equipment being measured [63].

The cost of measurement against a metric is also important; if this cost is greater than the expected saving from efficiency improvements, clearly measurement is unattractive [62]. Another important issue is the potential for manipulation, i.e. the implementation of

measures which improve performance against a metric without giving the desired efficiency improvement [62].

The most widely used metric for energy efficiency in data centres is Power Usage Effectiveness (PUE), which is defined as the total energy consumption of the facility over a period of time divided by the IT energy consumption [64]. The PUE was initially proposed by Belady and Malone in 2006 (cited in [64]), before being further developed by the Green Grid [64]. Its mathematical definition is shown in Eq. 1-1 (note that a PUE of unity indicates perfect performance against this metric).

$$PUE = \frac{Total\ Facility\ Energy}{IT\ Equipment\ Energy} = \frac{E_T}{E_{IT}}$$ Eq. 1-1

The Green Grid provides detailed instructions for how PUE should be calculated [64]. E_T should be measured at the electricity meter of the facility, and hence should include not only energy consumption in IT equipment and cooling infrastructure, but also losses in power delivery systems, energy used in lighting and in offices supporting the data centre etc. Where the data centre exists within a mixed use facility, the electricity should be measured at the meter for the data centre itself. E_{IT} is defined as "the energy consumed by equipment that is used to manage, process, store or route data within the compute space" [64]. It should be measured as the energy actually delivered to the IT equipment, i.e. it should not include energy lost in power switching and conversion. PUE has now also been formalised as an International Organization for Standardization (ISO) standard, which is implemented in the UK by the British Standards Institution as BS ISO/IEC 30134-2:2016 [65].

The PUE is intended to be useful in identifying opportunities to improve efficiency and in comparing between efficiencies of similar data centres [64]. The Green Grid [64] notes that the type of processing carried out by a data centre will affect the PUE, meaning that comparisons between PUE's of data centres carrying out different types of jobs are of limited value.

The Uptime Institute's [29] recent survey found an industry average PUE of 1.65 in 2013, which had fallen from 2.5 in 2007. The Institute stresses that these figures are self-reported and hence not entirely reliable, with 6% of respondents claiming PUEs of less than 1, which is impossible according to the measure's definition [64]. The Institute also stresses the likelihood of greater than average focus on efficiency within their respondents than generally within the data centre sector, which could lead these figures to be lower than the true average for the sector. It should also be stressed that, even with the likely skew of

respondents towards those particularly interested in efficiency, 13% reported PUEs greater than 2.

Despite the limitations in the data collection, the reduction in PUE shown in the Uptime Institute's survey results [29] is thought to represent a genuine improvement, brought about by improvements in air management, as well as improved efficiency of uninterruptible power supplies (UPSs) and power distribution systems [29]. It is suggested that in many cases, the low-hanging fruit have been taken, with the remaining improvements often requiring considerable capital expense and/or in-house expertise. There is said to be a gulf between large companies, or those for whom IT operations make up a large portion of expenditure, who have the capital and expertise required to achieve these more expensive measures, and others who are lagging behind. The key measures required to further reduce PUEs are highlighted as economisation, aisle containment and raising server inlet air temperatures [29].

Salim and Tozer reported in 2010 on data collected from over 40 data centres, finding that PUEs ranged from 1.67 to 3.57, at an average of 2.34 [6]. It may be expected that these figures are more accurate than those collected by the Uptime Institute, having been obtained from independent analysis. In any case, the figures are roughly consistent with The Uptime Institute's reported fall from an average of 2.5 in 2007 to 1.65 in 2013 [29]. Salim and Tozer also noted that PUE correlated strongly with floor space, as shown in Table 1-3.

Floor Space (ft^2)	Mean PUE
<10000	2.8
10000-30000	2.2
>30000	2.1

Table 1-3. Correlation between floor space and PUE [6].

TechUK reported in 2015 that the average PUE for participants of the Climate Change Agreement (CCA) for data centres was 1.93 [28]. The CCA is a scheme by which UK co-location data centres receive tax reductions provided that they achieve specified reductions in PUE, and is discussed further in section 6.1.6.1. Note that participants of the CCA are likely to be more focussed than others on energy efficiency, and specifically on reducing PUE, since they have a direct financial incentive to do so.

A 2013 survey of data centres in the Asia-Pacific region found much higher mean PUEs, ranging from 2.2 to 2.6 for the different countries surveyed (Digital Realty, cited in [66]).

Meanwhile, market leaders such as Facebook and Google claim PUEs of around 1.1 [67], [68].

Whilst there is uncertainty both regarding the accuracy of self-reported data, and around the extent to which surveys are able to obtain data representing the entire sector, the PUE figures quoted suggest that a significant proportion of the data centre sector operates well below the achievable standards of efficiency.

One major flaw of the PUE metric is that a reduction in IT power consumption without a corresponding reduction in cooling power will increase PUE [62], [64]. Hence, the use of PUE could discourage investment in efficient servers, or attempts to improve IT utilisation. Similarly, as discussed in Section 1.3.3, increasing the supply air temperature may reduce cooling power consumption, whilst increasing IT power consumption. This could potentially improve PUE whilst leading to an overall increase in E_T. This is clearly problematic, since this could provide an incentive to increase E_T if a data centre manager is concerned about their PUE. In light of this, PUE is mainly useful in assessing the efficiency of the cooling and power delivery infrastructure, rather than that of the data centre as a whole [62].

Cho & Kim [38] described 2 metrics which quantify the extent of mixing of hot and cold air streams. These are the Supply Heat Index (SHI) and Return Heat Index (RHI), and are given by Eq. 1-2 and Eq. 1-3 [38], respectively, where Q is the total heat dissipation of the IT equipment and δQ is the enthalpy rise of the cold air before entering the racks. SHI and RHI sum to 1. Note that perfect air management, i.e. zero mixing, would be indicated by $SHI = 0$ and $RHI = 1$. Clearly, measuring performance against these metrics requires the measurement of server and CRAC/CRAH inlet and outlet air temperatures and flow rates, as well as IT power consumption [6]. Both are essentially measures of recirculation.

$$SHI = \frac{\delta Q}{Q + \delta Q}$$

Eq. 1-2

$$RHI = \frac{Q}{Q + \delta Q}$$

Eq. 1-3

Herrlin [69] defined a similar metric, the Return Temperature Index (RTI), whose value is affected by both recirculation and bypass. It is given by Eq. 1-4 [69], where T_{Return} is the temperature of the return air to the cooling infrastructure, T_{Supply} is the supply air temperature and ΔT is the temperature rise across the IT equipment. An RTI of greater than 1 implies a prevalence of recirculation, whereas an RTI of less than 1 implies more bypass. Clearly this metric does not individually measure bypass and recirculation, and a value of 1 could be achieved where bypass and recirculation cancel each other out.

$$RTI = \frac{T_{Return} - T_{Supply}}{\Delta T}$$

Eq. 1-4

SHI, RHI and RTI are all more specific in scope than PUE, and provide measures of the effectiveness of air handling in a data centre.

A true measure of efficiency divides the useful work done by a system by the energy input to the system [62]. Since PUE does not include a measure of the useful work done by the data centre, it can only indicate the efficiency with which heat is rejected. Patterson [62] notes the intrinsic difficulty in measuring data centre efficiency, which is the difficulty in defining useful work. The tasks being performed by a data centre may change over time, and vary greatly between facilities. CPU utilization gives some indication of output, but does not consider the efficiency of the software being used. One server's work may be usefully quantified by the number of emails it processes, whereas another's may be best assessed by the number of iterations completed on a computational fluid dynamics (CFD) simulation. Even an individual server may be used for different kinds of work at different times. The Green Grid [63] goes further, suggesting that the value assigned to a piece of useful work completed should also take into account the time taken to complete the work, and the effect that this has on its value to the end user.

The Uptime Institute and McKinsey & Company have described a metric which seeks to include consideration of the computational efficiency [70]. It is called the Corporate Average Data Center Efficiency metric (CADE), and is given by Eq. 1-5 [62].

$$CADE = Infrastructure\ utilization$$
$$\times Infrastructure\ energy\ efficiency \times IT\ utilization$$
$$\times IT\ energy\ efficiency$$

Eq. 1-5

The infrastructure utilization is given by the IT load divided by the total facility capacity, whilst the infrastructure energy efficiency is simply 1/PUE [70]. The IT utilization refers to the average CPU server utilization [70]. IT energy efficiency was not defined by the Uptime Institute and McKinsey & Company, but Patterson [62] suggests using the reciprocal of the IT Energy Usage Effectiveness (itEUE). This is equal to the total IT energy consumption divided by the computational energy consumption. The computational energy consumption is composed of the energy used by the CPU, memory and storage, with the total IT energy including power usage of server fans, power supplies and voltage regulators. itEUE can only realistically be measured in a laboratory setting due to the range of measurements required.

Patterson [62] suggests that the lack of use of the CADE metric results from its creators not having fully defined the terms involved at the time of its announcement. It is also clear from

the above description that it is much more complex than the other metrics mentioned, which is likely to be another factor in its lack of use. The main advantage of CADE is that it combines factors assessing the level of overprovision of infrastructure of the data centre with a measure of how efficiently energy is used within the data centre, both in the cooling and power infrastructure, and in the servers. The CADE value does not give any indication of which aspect of efficiency has the most room for improvement, meaning that its constituent parts are still individually important. The value of the metric will typically be below 10%, even for modern, efficient data centres [62].

The Green Grid [63] proposed another metric incorporating the useful work undertaken by a data centre, called the Data Centre Productivity (DCP) metric. This is defined broadly as the useful work done in the data centre, divided by the quantity of a resource consumed in completing this work (e.g. electricity, water, floor space etc.). The useful work will be a function of the number of tasks completed during the specified time window, with value attributed to different tasks (e.g. loading a new record into a data base, or satisfying a query) as deemed appropriate. The value attributed to each completed task should be related to the stipulations of the SLA, taking into account any penalties for failure to meet specified completion times [63].

In a later document, The Green Grid [71] defines eight proxies for useful work done, to be used in the calculation of DCP. In the most detailed proxy, each individual instance of a piece of software running in the data centre logs and reports the number of units of useful work completed during a defined assessment window. The value attributed to each piece of software's "unit" of useful work must then be defined by the data centre operator. This proxy is flexible, and allows data centre operators the freedom to appraise the value of work completed by different applications. However, the measure would require amendments to software in order to automate the required data logging. In addition, the setting of weighting values for different pieces of software may be time consuming, and may require regular review if services provided by the data centre change over time. Two other of the proposed proxies are similar, but log the useful work of only a subset of the servers. This simplifies the data collection process, whilst requiring assumptions to be made around the applicability of the measured data to the performance of the remainder of the data centre. The remaining proxies involve logging either the number of bits of information leaving the data centre or the CPU utilisation of the data centre's servers during the assessment window, or logging the number of operating system instances supported by the data centre at a particular point in time. Clearly these proxies are likely to give a less reliable assessment

of the useful work output than those for which the units of work are defined individually for each piece of software. However, they are simpler to measure.

In conclusion, it is clear that measuring efficiency in data centres is intrinsically difficult, due primarily to the difficulty in defining useful work. Comparing the efficiencies of different data centres with each other is also difficult, due to the variation in the kinds of work carried out by different data centres. The current most widely used metric, PUE, has limited scope, but is useful in measuring the efficiency with which heat is removed from a data centre. CADE and DCP provide more holistic measures of data centre efficiency, but are more time consuming to calculate.

1.5   Research aims and objectives

The aim of this study is:

To investigate the technologies and policy instruments available to improve efficiency in data centre cooling, with a particular focus on air management in data centres employing aisle containment.

The study uses a combination of experimental, modelling and social research methods. The objectives of the study are:

1) **To investigate the extent of bypass and recirculation in data centres employing aisle containment**

As will be shown in Chapter 2, aisle containment is an increasingly common energy efficiency measure, and has the potential to improve efficiency in data centre cooling. However, there are limitations in the existing literature on the subject, specifically in the absence of robust investigations into the extent of bypass and recirculation in data centres employing aisle containment. This objective will be achieved through experimental methods, which will determine the magnitude of bypass and recirculation in these data centres under a range of conditions.

2) **To investigate the implications of bypass and recirculation for electricity consumption in data centres employing aisle containment**

As will be shown in Chapter 2, the impact on electricity consumption of bypass and recirculation in data centre centres employing aisle containment has received little attention in the research literature. A system model will be developed to utilise the results of the experiments quantifying bypass and recirculation in predicting data centre power consumption under a range of conditions.

3) **To investigate the potential for using CFD models to aid in the efficient design and management of data centres employing aisle containment**

As will be shown in the literature review in Chapter 4, the CFD models presented in the literature to date largely focus on air-cooled data centres which do not employ aisle containment. Authors presenting models which do investigate data centres employing aisle containment have not extensively investigated or criticised the methods used to determine bypass and recirculation under contained conditions.

A commercially available Navier-Stokes CFD software package will be used in addition to a new potential flow CFD model. The models will utilise the findings of the experiments conducted in relation to objective (1). The implications of the CFD models' results in relation to the assumptions made in the system model will be investigated.

4) **To investigate the potential for policy instruments to drive energy efficiency improvements in the data centre sector**

In section 1.4 it has been noted that there is thought to be a gulf between the most and least efficient data centres. The literature review presented in Chapter 6 will provide evidence that there are significant opportunities for improvement in data centre energy efficiency, and that progress is hampered by both technical and political challenges. Chapter 6 will also show that policy instruments can play a key role in encouraging energy efficiency in various industries and sectors, and that various existing policy instruments impact on the data centre sector. No academic study thus far has investigated the role of policy instruments in the data centre sector.

Semi-structured interviews will be undertaken with people working within the data centre sector, in order to investigate opinions within the sector towards existing policy instruments. The findings of these interviews will be analysed in the context of the findings of the literature review and of the experimental and computational work relating to objectives (1) to (3). This will enable recommendations to be made for potential improvements to the current policy environment in the UK.

1.6 Thesis outline

Chapter 2 firstly provides a literature review summarising the research conducted to date regarding the impact of aisle containment on bypass, recirculation and electricity consumption. The remainder of the chapter presents an experimental investigation into the extents of bypass and recirculation in data centres employing aisle containment. Chapter 3 presents a new, physics-based, system model, which utilises the results presented in Chapter

2 to predict the impact of aisle containment on electricity consumption, under a range of conditions. Chapter 4 provides a literature review of CFD methods, with a focus on methods used in modelling of data centre air flows. Chapter 5 presents a new data centre potential flow CFD model, validated against the results presented in Chapter 2. This is compared with the results of a Navier-Stokes CFD model, produced using commercially available software. Chapter 6 presents a literature review of the use of policy instruments to drive energy efficiency improvements in the data centre sector. The remainder of the chapter presents an investigation into attitudes within the data centre sector towards such policy instruments. Chapter 7 summarises the key findings of the work, and makes recommendations for further work.

2. BYPASS AND RECIRCULATION IN DATA CENTRES EMPLOYING AISLE CONTAINMENT

2.1 Introduction

Chapter 1 has highlighted that improvements in air management present great potential for improving both the efficiency of data centre cooling, and the reliability of operation of servers. The present chapter focuses on aisle containment, which is an approach which can be used to improve data centre air management. A literature review is firstly presented, detailing the present knowledge on the impacts of aisle containment. The chapter goes on to detail experiments undertaken by the author to quantify the impacts of aisle containment, and the potential to minimise bypass and recirculation in aisle contained systems through practical measures. These experiments seek to fulfil objective (1), as defined in section 1.5, which is:

'To investigate the extent of bypass and recirculation in data centres employing aisle containment.'

The results of experiments undertaken to investigate air flows and temperature distributions in a Test Data Centre employing aisle containment are also reported, which will be used in Chapter 5 to validate computational fluid dynamics (CFD) models.

2.1.1 Air handling in aisle contained data centres

Bypass and recirculation may be reduced by introducing physical barriers separating hot and cold aisles. This approach is referred to as hot or cold aisle containment [36], and can be an inexpensive way of reducing recirculation and/or bypass [31].

Aisle containment is an increasingly popular design characteristic in data centres [29], since it can improve both cooling efficacy and energy efficiency [37], [39]. Figure 2-1 shows diagrammatical representations of data centres employing hot aisle-cold aisle (HACA) formation and cold aisle containment, with the air supplied from an underfloor plenum. An alternative arrangement is hot aisle containment, in which ducting is introduced to enclose the flow leaving the rack, transporting this air directly to the CRAH inlets, usually via a ceiling plenum. In either case, the aim is to force the supplied cold air to pass through the racks, rather than bypassing them and returning to the CRAC/CRAH without having provided any cooling.
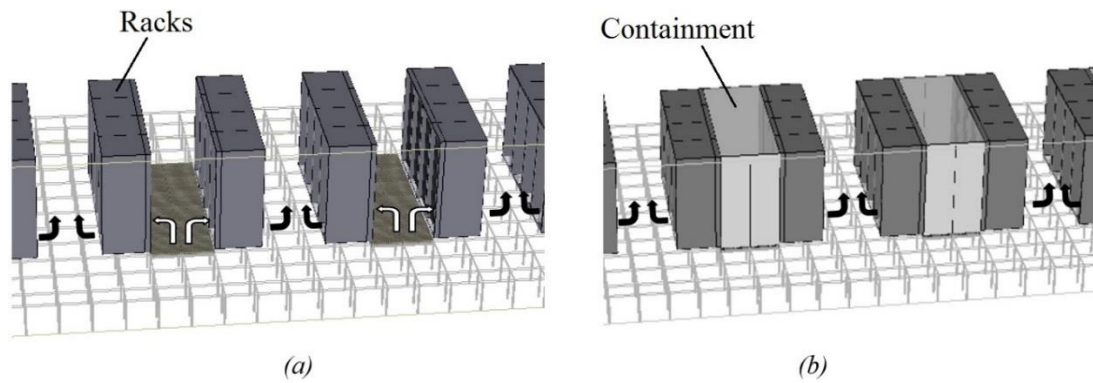
Figure 2-1: (a) Diagram of data centre in HACA formation and (b) diagram of cold aisle containment system. Paths of cold and hot air are shown in white and black arrows, respectively.

Some bypass still occurs in contained systems, since "over supply" of cold air is required to minimize recirculation [37], [72]. Here, "over supply" refers to setting the CRAC/CRAH flow rates in excess of the air flow required through the servers. This produces a pressurisation of the cold aisle, causing conditioned air to escape into the hot aisle without passing through servers. White papers, equipment specifications and trade journals indicate common pressure differentials between the cold and hot aisles, Δp_{CH}, ranging from 2-20 Pa [73]–[77], although this has not been studied extensively from an academic perspective. The latest (2017) Best Practice Guidelines from the EU Code of Conduct on Data Centre Energy Efficiency[1] (EU CoC) recommend that Δp_{CH} not exceed 5 Pa [78], although the 2014 edition of the guidelines included no such recommendation [79]. This recommendation will become a requirement for Participants of the EU CoC in 2018. This indicates the relative infancy of the debate around pressurisation of cold aisles. Note that, over the course of this thesis, positive Δp_{CH} indicates that the static pressure in the cold aisle exceeds that in the hot aisle.

In principle, in a data centre employing containment, bypass and recirculation may result from air leaking through the containment structure, or from air passing through a rack without passing through a server. In this thesis, these two broad leakage routes are identified as **containment leakage** and **rack leakage**, respectively.

---

[1] See Chapter 6 for a detailed description of the EU Code of Conduct on Data Centre Energy Efficiency.

2.1.2 Experimental and computational investigations of aisle containment published to date

The effect of aisle containment on data centre cooling has been investigated experimentally by Arghode et al. [37] and Alkharabsheh et al. [72], and in CFD simulations by numerous authors [5], [72] [12]–[15] [80], [83]–[92].

Arghode et al. [37] measured flow rates through the racks and CRAHs in a data centre with and without aisle containment. Rack leakage was neglected, with all air passing through the racks assumed to pass through the servers. With containment installed, the results show that roughly 10% of the air supplied to the cold aisle bypassed the racks at $\Delta p_{CH} = 6.2\ Pa$, and 11% of air passing through the racks was recirculated with $\Delta p_{CH} = -9.1\ Pa$. No other levels of Δp_{CH} were investigated [37]. In absolute terms, this amounted to 0.81 m^3.s^{-1} bypass and 0.84 m^3.s^{-1} of recirculation for the cases with $\Delta p_{CH} = 6.2\ Pa$ and $\Delta p_{CH} = -9.1\ Pa$, respectively, for a single cold aisle containing 14 racks. Cold aisle containment was shown to dramatically improve uniformity of cold aisle temperatures. With CRAH fan speed held constant, installing containment was found to reduce the cold aisle temperature range (i.e. the difference between the lowest and highest cold aisle temperature) from 7 Kelvin (K) to 2 K. Note that this was at the CRAH fan speed corresponding to the case with $\Delta p_{CH} = 6.2\ Pa$ with containment installed.

Alkharabsheh et al. [72] also took experimental measurements of a data centre with and without aisle containment, however the extent of bypass and recirculation was not discussed. The results showed that, with containment in place, a very small cold aisle temperature range of 1.5 K was achieved. The level of Δp_{CH} was not disclosed, nor was the equivalent cold aisle temperature range in uncontained conditions.

Two authors have presented measured temperatures from data centres employing aisle containment, without investigating the corresponding uncontained case. Firstly, Ham & Jeong [86] found a return temperature index (RTI) of 109% in a data centre employing aisle containment, indicating some recirculation. However, Δp_{CH} was not measured. In addition, Tradat et al. [93] have presented data from a cold aisle contained data centre showing that negative Δp_{CH} can lead to server inlet temperatures significantly exceeding supply temperatures, indicating recirculation. It is not possible to deduce the rate of recirculation from the data presented.

The only computational model to be reported which explicitly considers rack leakage is from Alkharabsheh et al. [72]. Here, rack leakage was allowed through 5 cm wide channels down the sides of the equipment rails in each rack, and containment leakage through channels at

various potential leakage points identified visually in the aisle containment structure of the data centre being modelled (see section 2.1.3 for a description of typical rack designs and leakage paths). All of these channels were assigned the same percentage open area in a CFD model, which was selected via a calibration process in which simulation results were compared with experimental measurements. The effect of bypass on energy consumption was not discussed, with the focus of the analysis being the air temperatures within the cold aisle. The model predicted a cold aisle temperature range of around 7 K in the uncontained condition [72]. A later paper from the same lead author used the same model to predict a cold aisle temperature range of around 5 K in the contained condition, although this greatly exceeded the range recorded in experimental measurement, which was less than 1.5 K [94]. The CFD results presented in the later paper also showed that increasing the percentage open area reduced the flow rate through servers and increased recirculation (at constant CRAH flow rate), which implies that bypass also increased [94]. This paper also reported that 4.6-13.4% of conditioned air supplied to the cold aisle bypassed the servers, depending on the CRAH fan speed. However, there was no discussion of Δp_{CH}, nor any validation of the relative importance attached to bypass through the racks and containment system.

The preceding results may be compared with typical levels of bypass in uncontained data centres. Monitoring of server and CRAC/CRAH inlet and outlet temperatures carried out by Salim & Tozer [6] at 40 data centres not employing aisle containment found that, on average, 50% of cold air supplied by the CRAC/CRAHs bypassed the servers. This is significantly greater than the rates of bypass recorded experimentally by Arghode et al. [37], and predicted in simulations by Alkharabsheh et al. [72].

Other researchers have used CFD simulations to demonstrate the potential reductions in electricity consumption and bypass resulting from the implementation of aisle containment, whilst either not considering rack leakage, or not discussing the methods used to govern it, and without reporting Δp_{CH}. Some have also neglected containment leakage. The findings of these studies will be summarised in the remainder of this section.

Ham & Jeong [86] used a CFD model of a laboratory data centre to predict a 15 to 33% reduction in power consumption resulting from the introduction of containment, resulting from improvements in return heat index (RHI) and supply heat index (SHI), allowing more efficient operation of the cooling infrastructure. Containment leakage was modelled using methods similar to those used by Alkharabsheh et al. [72]. Rack leakage was neglected, with all flow through the racks assumed to pass through the servers [86]. Δp_{CH} was set to 3 Pa. It is not possible to determine the rates of bypass from the data presented.

Schmidt et al. [81] used CFD simulations to predict a reduction in CRAH power consumption of 59% resulting from the installation of aisle containment, while Shrivastava et al. [82] similarly predicted a 33% reduction in the electricity consumption of the cooling infrastructure. Schmidt et al. [81] and Shrivastava et al. [82] report bypass percentages within contained systems of 3.1 and 13% respectively, although neither disclose any information regarding the model detail governing rack or containment leakage, and neither report Δp_{CH} [81], [82].

Choo et al. [39] used a CFD model to predict that the introduction of containment could reduce the cold aisle temperature range (i.e. the difference between the lowest and highest temperature within the cold aisle) from around 16 K to around 4 K in a legacy university data centre. This was due to the reduction in recirculation, which in turn allowed the supply air temperature to be increased. An energy balance was used to show that this would allow a 58% reduction in the flow rate of chilled water (which in this case had a fixed supply temperature) to the CRAHs, leading to significant energy savings and payback within 3.6 months of installation of containment. The methods used to model bypass and recirculation were not discussed, and it was not clear whether rack leakage was included.

Fouladi et al. [92] used a CFD model to show that containment can worsen performance against the PUE metric in some instances, due to increased pressure in the cold aisle, and the impact of this on CRAH fan power consumption. However, under all arrangements, the minimum PUE coincided with the contained arrangement if supply air flow rate were set so as to minimise PUE, with PUE falling by 7.1% in the most extreme case. Containment was shown to dramatically improve performance against the SHI metric. The results show that modulation of CRAH fan speed is essential to ensure that the benefits of aisle containment are fully realised. Again, the methods used to model bypass and recirculation were not discussed, and it was not clear whether rack leakage was included.

Cho & Kim [38] used CFD models to assess the effects of installing a vertical partition between hot and cold aisles, finding that the partition significantly reduced both bypass and recirculation. The ends of the aisles were left open, i.e. this configuration did not represent full containment. Similarly, Nada et al. [84] used a CFD model to show that installing a ceiling in the cold aisles (with the ends of the aisles left open) could significantly improve SHI and RHI, demonstrating reductions in recirculation and bypass, respectively.

The aforementioned investigations have demonstrated the potential for aisle containment to reduce bypass and recirculation. However, there is still a need for further investigation

to better understand the extent to which aisle containment can reduce electricity consumption, and how these systems can be optimised.

2.1.3   Server rack design

Server racks are fitted with equipment rails sized to accommodate IT equipment of standard 483 mm (19 inch) width [95]. The rails contain 'slots' into which IT equipment may be installed. Racks usually contain some empty slots, which should be filled with blanking panels to prevent the slots from providing paths for rack leakage (leakage paths). Figure 2-2 shows the positions of the equipment rails and servers/blanking panels, as well as indicating other components referred to later in this chapter. Each slot has a height of 44.45 mm, commonly referred to as one rack unit, or "1 U", which corresponds to the height of a standard server. Space exists on either side of the equipment rails, as well as above and below them, which may provide undesired, additional leakage paths. There may also be the potential for air to escape through the top, bottom or sides of the rack, after entering the rack front, but prior to passing into the server inlets. These leakage paths are shown in Figure 2-3, illustrated by the arrows.
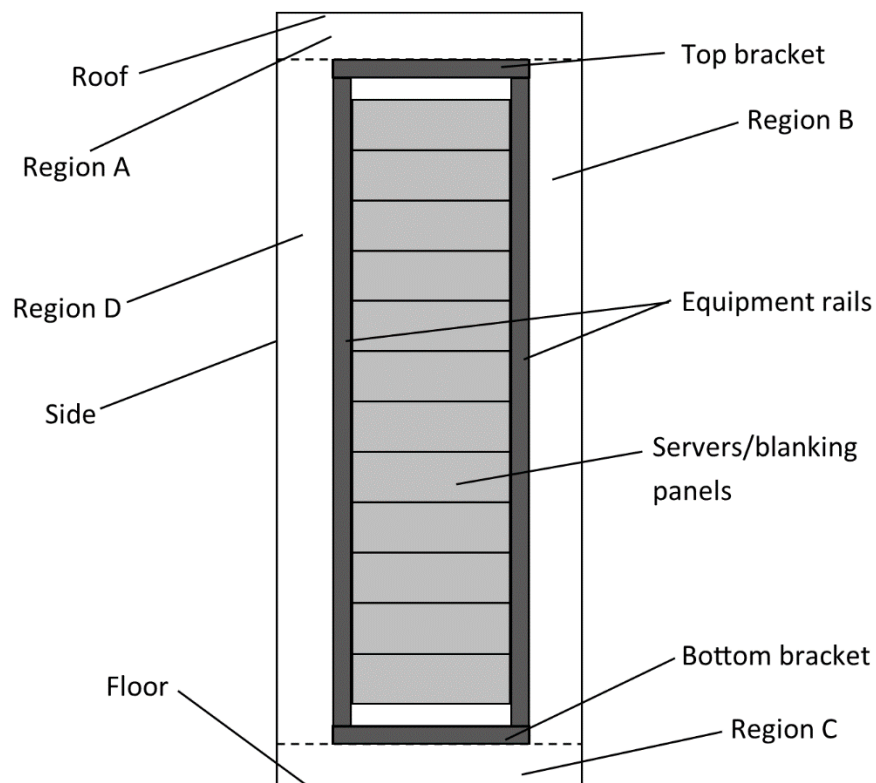
Figure 2-2. Definitions of leakage paths and rack components.

(a) 3D view   *(b) Plan view*   *(c) Side view*
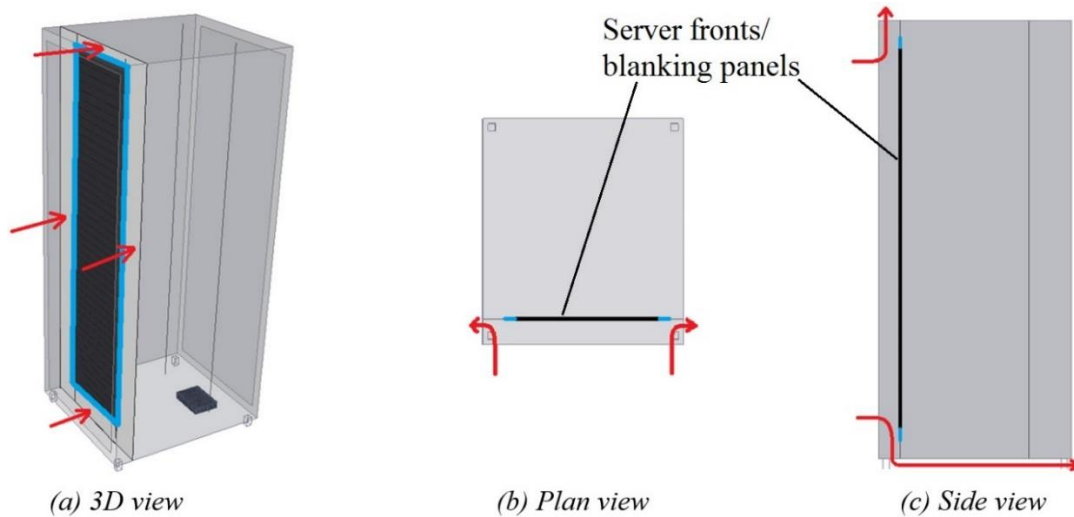
Figure 2-3. Rack leakage paths (a) above, below and at sides of equipment rails, (b) through sides of rack and (c) through top and bottom of rack. The black lines in (b) and (c) represent the plane coinciding with server fronts in occupied slots, and blanking panels in unoccupied slots.

Simple, inexpensive measures which reduce the potential for air transport from hot to cold aisles (and vice versa), such as installing blanking panels in empty rack slots, and covering empty spaces within rows of racks, are still in some cases overlooked [29] despite being recommended in the EU CoC best practice guidelines, for example [78]. It is generally believed that adoption of these kinds of measures is becoming more widespread [29].

2.1.4   Estimation of flow through an empty slot

Makwana et al. [96] carried out experiments in a data centre thermal laboratory employing aisle containment, in which the impact of the introduction of empty slots was investigated. The total required supply airflow was found to be increased by 9% by the introduction of one empty slot in each rack, with each rack operating with a heat load of 14.6 kW. Industry best practices were said to be applied to prevent other leakage through racks and through the containment structure. The level of Δp_{CH} was not discussed, nor was the overall extent of bypass and/or recirculation.

For comparison, the volumetric flow rate, $\dot{V}_{slot}$, through an empty slot may be estimated by modelling the flow as being through a rectangular channel which is 430 mm wide, 1 U (44.45 mm) deep and 724 mm long, which represents the space between 2 typical servers separated by an empty slot [97]. Such a flow is depicted in Figure 2-4. The resistance offered by the entrance to and exit from the slot may be estimated by recognising that these flow features represent flow from a reservoir into a pipe (a contraction) and from a pipe into a

reservoir (an expansion), respectively. The pressure drops across these features can be calculated using Eq. 2-1 and Eq. 2-2, respectively [98]. Here, p_{CA}, p_{con}, p_{exp} and p_{HA} are the gauge pressures in the cold aisle, immediately after the contraction, immediately before the expansion and in the hot aisle, respectively; ρ_{air} is the density of the air; u is the average air velocity within the slot; and k_{con} and k_{exp} are the loss coefficients relating to the contraction and expansion, respectively. k_{con} and k_{exp} may be approximated as 0.5 and 1, respectively, for this kind of contraction and expansion [98].

$$p_{CA} - p_{con} = 0.5\rho_{air}k_{con}u^2 \qquad \text{Eq. 2-1}$$

$$p_{exp} - p_{HA} = 0.5\rho_{air}k_{exp}u^2 \qquad \text{Eq. 2-2}$$
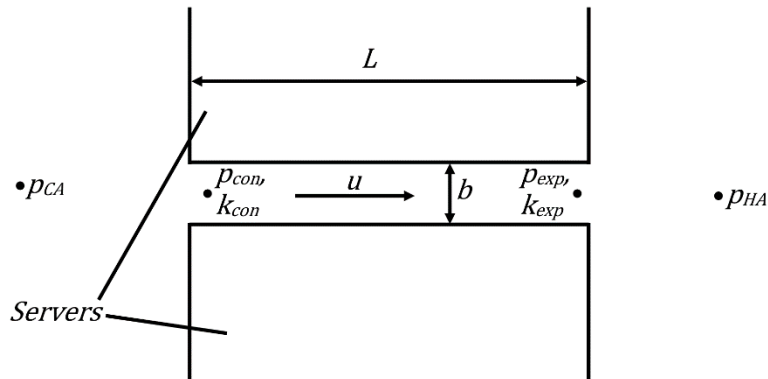
Figure 2-4. Side view of flow through an empty slot.

Losses within a rectangular channel are affected strongly by whether the flow regime is laminar or turbulent [99]. For laminar flow, the pressure drop along the length of the channel may be estimated using Eq. 2-3, where f is the friction factor, L is the length of the channel and D_h is the hydraulic diameter [99]. D_h is given by Eq. 2-4, where a and b are the width and depth of the channel, respectively [99]. f is dependent on the aspect ratio of the channel, a/b, and for a channel with $a = 0.0445\ m$ and $b = 0.437\ m$ is given by Eq. 2-5 [99]. Re, the Reynolds number, is given by Eq. 2-6, where μ is the dynamic viscosity [99]. Combining Eq. 2-3 to Eq. 2-6 gives Eq. 2-7, where u_l is the average velocity within the slot in the laminar case.

$$p_{con} - p_{exp} = 0.5 f \rho_{air} L u^2 / D_h \qquad\qquad \text{Eq. 2-3}$$

$$D_h = 2ab/(a+b) \qquad\qquad \text{Eq. 2-4}$$

$$f = 84.4/Re \qquad\qquad \text{Eq. 2-5}$$

$$Re = \rho u D_h / \mu \qquad\qquad \text{Eq. 2-6}$$

$$p_{con} - p_{exp} = 10.55 \mu L u_l (a+b)^2 / a^2 b^2 \qquad\qquad \text{Eq. 2-7}$$

Laminar flow generally dominates in such channels where $Re < 2000$, with turbulence dominating where $Re > 4000$ [99]. An initial assumption of laminar flow allows us to combine Eq. 2-1, Eq. 2-2 and Eq. 2-7, whilst noting that $\Delta p_{CH} = (p_{CA} - p_{con}) + (p_{con} - p_{exp}) + (p_{exp} - p_{HA})$ and $u = u_l$, giving Eq. 2-8. Thus, u_l may be determined for a given Δp_{CH} using the quadratic formula (assuming that $\mu = 2.2 \times 10^{-5} \, Pa.s$ and $\rho_{air} = 1.2 \, kg.m^{-3}$ [100]).

$$0.5\rho\left(k_{con} + k_{exp}\right)u_l^2 + \frac{10.55\mu L(a+b)^2}{a^2 b^2} u_l - \Delta p_{CH} = 0 \qquad\qquad \text{Eq. 2-8}$$

Having determined u_l, Re can be found using Eq. 2-6. This allows the validity of the assumption of laminar flow to be determined. If $Re > 2000$, the calculation must be repeated using a different method. Jones [101] reviewed the results of tests carried out in rectangular channels of various aspect ratios, finding that Eq. 2-9 fitted the results well for turbulent flow. Here, u_t is the average velocity within the slot for the turbulent case, Re is again given by Eq. 2-6, but D_h is given by Eq. 2-10. f is again related to the pressure drop via Eq. 2-3 [101]. Combining Eq. 2-1 to Eq. 2-3, noting that $u = u_t$ and re-arranging to make u_t the subject, gives Eq. 2-11. Combining Eq. 2-6, Eq. 2-9 and Eq. 2-11 gives Eq. 2-12.

$$f^{-0.5} = 2\log_{10}(f^{0.5}Re) - 0.8 \qquad\qquad\qquad \text{Eq. 2-9}$$

$$D_h = 2\left[\frac{2}{3} + \frac{11}{24}\frac{a}{b}\left(2 - \frac{a}{b}\right)\right]ab/(a+b) \qquad\qquad \text{Eq. 2-10}$$

$$u_t = \left(\frac{2\Delta p_{CH}}{\rho\left(k_{con} + k_{exp} + \frac{fL}{D_h}\right)}\right)^{0.5} \qquad\qquad \text{Eq. 2-11}$$

$$f^{-0.5} - 2\log_{10}\left[\frac{f^{0.5}\rho_{air}D_h}{\mu}\left(\frac{2\Delta p_{CH}}{\rho_{air}\left(k_{con} + k_{exp} + \frac{fL}{D_h}\right)}\right)^{0.5}\right] + 0.8 = 0 \qquad \text{Eq. 2-12}$$

So for a given Δp_{CH}, f may be calculated by applying the bi-section method to Eq. 2-12 (after calculating D_h using Eq. 2-10). u_t may then be calculated using Eq. 2-11. Whether the flow is laminar or turbulent, $\dot{V}_{slot}$ may then be calculated according to $\dot{V}_{slot} = uab$, after setting u equal to u_l or u_t as appropriate. For flow in the transition region, u may be approximated using a linear interpolation between u_l at $Re = 2000$ and u_t at $Re = 4000$, taking Re for the given Δp_{CH} as the average of those predicted using the laminar and turbulent methods. For $2 < \Delta p_{CH} < 20$, it was found that $Re > 4000$, indicating turbulent flow, and that $\dot{V}_{slot}$ ranged from 25 to 83 $l.s^{-1}$. The calculation process is illustrated in Figure 2-5.

28
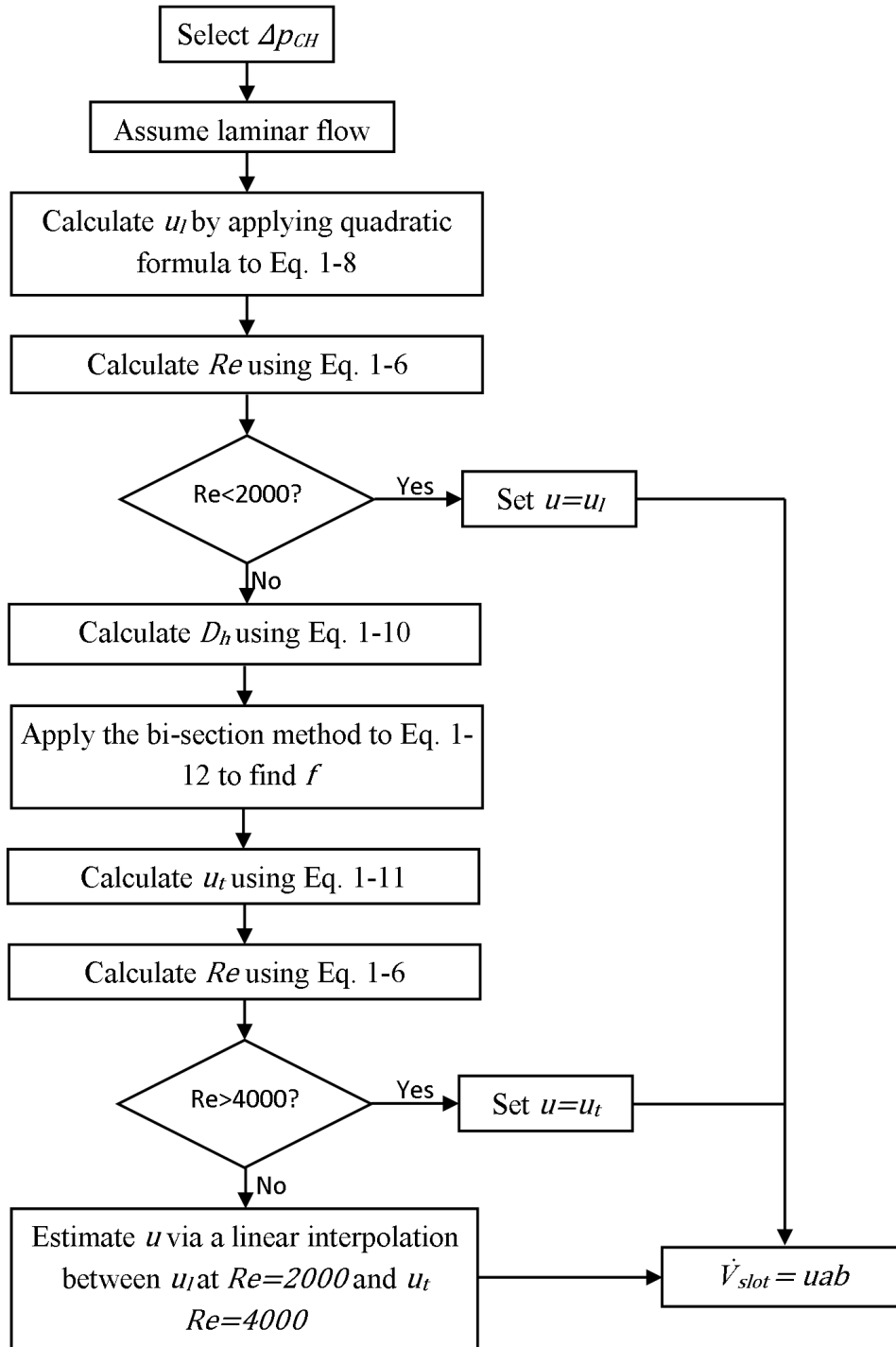
Figure 2-5. Flow chart for $\dot{V}_{slot}$ calculation.

To put these figures in context, at typical rack power densities of 12 kW per rack [102] and temperature rise across the servers, ΔT, of 12.5 K [41], [55], the flow rate of air required to cool the servers may be calculated by applying conservation of energy via Eq. 2-13 (assuming that all of the server power consumption is converted to heat, which is removed via convection). Here, $\dot{E}_{server}$ is the rack power consumption, $\dot{m}_{server}$ is the mass flow rate through the servers and $c_{p,air}$ is the specific heat capacity of air, held constant at

$1.005 \; kJ.kg^{-1}.K^{-1}$[100]. Since $\dot{m}_{server} = \dot{V}_{server}\rho_{air}$, (where $\dot{V}_{server}$ is the volumetric flow rate through the servers), $\dot{V}_{server}$ may be calculated, enabling the flow through a single empty slot within a rack to be quantified as representing 3.1-10.4% of supplied air flow for $2 \; Pa < \Delta p_{CH} < 20 \; Pa$. Note that the increase in required air flow reported by Makwana et al. [96] to result from the introduction of an empty slot into each rack within a laboratory data centre was 9% (as discussed earlier in this section), which falls within this range.

$$\dot{E}_{server} = \dot{m}_{server}c_{p,air}\Delta T$$

<div align="right">Eq. 2-13</div>

2.2 Methods

The literature review presented in section 2.1 has demonstrated that, whilst aisle containment can be effective in reducing bypass and recirculation, the potential extent of such reductions has not been rigorously quantified. The potential for rack leakage in particular has garnered little attention. The experiments detailed in this section were intended to address these omissions by measuring bypass and recirculation in contained systems under a range of conditions. In essence, the experiments involved driving air through racks, and measuring the resulting relationship between pressure and flow rate.

2.2.1 Experimental apparatus

Two fans were used in the experiments, which are referred to as Fan 1 and Fan 2. Fan 1 was a centrifugal fan with a maximum flow rate of 250 l.s^{-1}, manufactured by ebm-papst (EC Centrifugal Blower D3G133-BF03-06, [103]). Fan 2 was an axial fan with a maximum flow rate of 420 l.s^{-1}, manufactured by Vent-Axia (ACP31512HP, [104]).

Two hot wire anemometers were used to measure air speeds, from which volumetric flow rates were calculated. Anemometer 1 was an Omega HHF2005HW [105]. Its resolution was 0.1 m.s^{-1}, and its accuracy was $\pm(0.1 \times reading + 0.1) \; m.s^{-1}$. Anemometer 2 was an Airflow 'TA-2 2' [106], and was used on measurements below 2.0 m.s^{-1} when available. Its resolution was ±0.02 m.s^{-1} on speeds below 0.5 m.s^{-1}. However, readings at this range were observed to fluctuate considerably, such that only a resolution of ±0.05 m.s^{-1} could be given with a reasonable level of confidence. At higher speeds, its resolution was 0.1 m.s^{-1}. The manufacturer specifications for Anemometer 2 do not define the instrument's accuracy [106]. It will be demonstrated in section 2.4.1 that Anemometers 1 and 2 give very similar results, indicating similar accuracies.

In the various experiments, air speed measurements were taken in a range of ducts. The specifications of these ducts are given in Table 2-1.

Duct	Material	Shape	Internal Dimensions
Measurement Duct 1	Corrugated plastic	Rectangular	232 by 97 mm
Measurement Duct 2	Rigid aluminium	Circular	Diameter=315 mm
Measurement Duct 3	Corrugated plastic	Rectangular	117 by 135 mm
Measurement Duct 4	Corrugated plastic	Rectangular	120 by 134 mm
Measurement Duct 5	Corrugated plastic	Rectangular	104 by 106 mm

Table 2-1. Specifications of volumetric flow rate measurement ducts.

For **Measurement Duct 1**, a single velocity measurement was taken at the centre of the duct for each test, which was converted to a volumetric flow rate by multiplying by the cross sectional area of the duct. The air speed was assumed not to vary across the section, with preliminary measurements having shown any variation to be within the range of the resolution of the anemometer plus 10% of the velocity measured at the centre.

For **Measurement Duct 2**, the recommendations of BS ISO 3966:2008 [107] for velocity measurement positions used to determine fluid flow rates in circular ducts were followed. One air velocity measurement was taken at the centre of the duct, with 8 further measurements being taken along each of 2 perpendicular diameters of the section. That is to say that a total of 17 velocity measurements were taken for each calculated flow rate. The measurement points were evenly distributed along each diameter, with the measurements farthest from the centre of the duct being taken 9mm from the wall of the duct. The conversion of velocity measurements into volumetric flow rate also utilised the methods described in BS ISO 3966:2008 [107], and is described in the following. The flow rate within the area bounded by a circumference passing through the measurement points farthest from the centre of the duct was calculated by integrating along each diameter with respect to distance from the centre of the section, assuming velocity varied linearly between measurements. The flow rate, $\dot{V}_{1,2}$, between two radii, r_1 and r_2, must be given by Eq. 2-14.

$$\dot{V}_{1,2} = \int_{r1}^{r2} 2\pi r v\, dr \qquad\qquad \text{Eq. 2-14}$$

Assuming that the mean velocity, v_i, along a given radius, r_i, on which v measurements were taken, is equal to the mean of the measurements taken on that radius, and that this mean varies linearly between r_i and $r_{i\pm1}$, we have $v = v_1 + \frac{(v_2 - v_1)(r - r_1)}{r_2 - r_1}$. We can therefore develop Eq. 2-15, via the following manipulations:

$$\dot{V}_{1,2} = 2\pi \int_{r1}^{r2} r \left[v_1 + \frac{(v_2 - v_1)(r - r_1)}{r_2 - r_1} \right] dr$$

$$\dot{V}_{1,2} = 2\pi \int_{r1}^{r2} r\left[\left(v_1 + \frac{r_1(v_1 - v_2)}{r_2 - r_1}\right) + \frac{r(v_2 - v_1)}{r_2 - r_1}\right] dr$$

$$\dot{V}_{1,2} = 2\pi \left[\frac{r^2}{2}\left(v_1 + \frac{r_1(v_1 - v_2)}{r_2 - r_1}\right) + \frac{r^3}{3}\left(\frac{v_2 - v_1}{r_2 - r_1}\right)\right]_{r=r_1}^{r=r_2}$$

$$\dot{V}_{1,2} = 2\pi \left[\frac{r_2^2 - r_1^2}{2}\left(v_1 + \frac{r_1(v_1 - v_2)}{r_2 - r_1}\right) + \frac{r_2^3 - r_1^3}{3}\left(\frac{v_2 - v_1}{r_2 - r_1}\right)\right] \qquad \text{Eq. 2-15}$$

Hence the flow rate through the area bound by each pair of adjacent radii passing through measurement points may be calculated. The flow rate at the periphery of a circular duct, $\dot{V}_p$, can be calculated using Eq. 2-16, where r is the distance from the centre of the duct, r_0 is the distance from the centre of the duct to the measurement points farthest from the centre, $\dot{V}_p$ is the flow rate in the annulus bounded by the circle at r_0 and the inside wall of the duct, v is the mean velocity at r, and R is the inner radius of the duct.

$$\dot{V}_p = \int_{r_0}^{R} 2\pi r v\, dr \qquad \text{Eq. 2-16}$$

Karman's conventional law for the variation of the fluid velocities in the peripheral zone is given by Eq. 2-17 [107], where v_0 is the mean velocity at r_0, and m_R is a constant describing the roughness of the duct wall. m_R was set equal to 8 for these calculations, which is suitable in most instances [107].

$$v = v_0 \left(\frac{R - r}{R - r_0}\right)^{\frac{1}{m_R}} \qquad \text{Eq. 2-17}$$

Combining Eq. 2-16 and Eq. 2-17 gives Eq. 2-18.

$$\dot{V}_p = 2\pi v_0 (R - r_0)^{-\frac{1}{m_R}} \int_{r_0}^{R} r(R - r)^{\frac{1}{m_R}} dr \qquad \text{Eq. 2-18}$$

Eq. 2-18 ultimately yields Eq. 2-19. The full derivation of Eq. 2-19 is shown in Appendix 1.

$$\dot{V}_p = 2\pi m v_0 (R - r_0)\left[\frac{m_R r_0 + m_R R + r_0}{(m_R + 1)(2m_R + 1)}\right] \qquad \text{Eq. 2-19}$$

Hence, $\dot{V}_p$ can then be added to the flow calculated through the area bound by the radius at r_0, giving the total flow rate through the duct.

For **Measurement Ducts 3, 4 and 5**, nine velocity measurements were taken across a cross section, at the positions depicted in Figure 2-6.
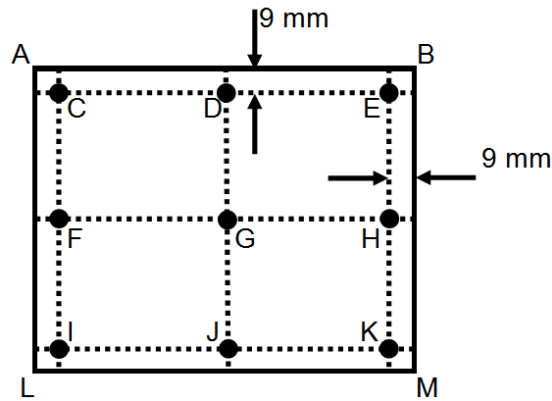
Figure 2-6. Positions of velocity measurement points in Measurement Ducts 3, 4 and 5.

The flow rate through these ducts was calculated in two stages, again following methods described in BS ISO 3966:2008 for conversion of velocities measured in rectangular ducts into volumetric flow rates [107], which are described in the following. The flow rate within the area bound by points C, E, K and I (see Figure 2-6) was calculated by dividing the total area enclosed by the measurement points into four rectangles, enclosed by the lines CDGFC, DEHGD, GHKJG and FGJIF. The flow rate through each rectangle was calculated by multiplying the mean of the 4 velocity measurements at its corners by its area.

The flow at the periphery of the duct is enclosed by 4 trapeziums, enclosed by the lines ABECA, BMKEB, MKILM and LICAL, respectively. The flow through trapezium $ABECA$, must be given by $\dot{V}_{ABECA} = \int_0^a lvdx$, where $l = W - 2x$, W is the internal width of the duct, x is the distance from the measurement nearest to the duct wall, $v = v_0 \left(\frac{a-x}{a}\right)^{\frac{1}{m_R}}$ (from Karman's law of the periphery, [107]), $v_{0,i}$ is the average velocity measured at a distance a from duct wall i, and m_R is a constant describing the roughness of the duct wall (again, set equal to 8 for these calculations). These expressions may be manipulated as follows to give Eq. 2-20:

$$\dot{V}_{ABECA} = v_0 \int_0^a (W - 2x) \left(\frac{a-x}{a}\right)^{\frac{1}{m_R}} dx$$

$$\dot{V}_{ABECA} = v_0 \int_0^a (W - 2x) \left(1 - \frac{1}{a}x\right)^{\frac{1}{m_R}} dx \qquad \text{Eq. 2-20}$$

Manipulating Eq. 2-20 and applying the same approach to the other three walls ultimately yields Eq. 2-21, where H is the height of the duct and $\dot{V}_P$ is the total flow through the four

trapeziums. The full derivation is shown in Appendix 2. The flow rate through the duct may then be calculated as the sum of $\dot{V}_P$ and the flow bounded by points C, E, K and I.

$$\dot{V}_P = \left(v_{0,1} + v_{0,3}\right)\frac{am_R}{m_R + 1}\left(W - \frac{2am_R}{2m_R + 1}\right) + (v_{02} + v_{04})\frac{am_R}{m_R + 1}\left(H - \frac{2am_R}{2m_R + 1}\right)$$

Eq. 2-21

Three different manometers were used to measure pressure, according to availability at the time of the experiments. Manometer 1 (KIMO MG 50 [108]) and Manometer 2 (KIMO HP 15 [109]) were inclined plane manometers with resolutions of 5 Pa and 1 Pa, respectively. However, since for each of these manometers, the distance between the 5 and 1 Pa markers was 1.5 mm, it was possible to record pressures to the nearest 2.5 and 0.5 Pa, respectively. Manometer 3 was a digital manometer (Digitron 2080p, [110]) with a resolution of ±0.1 Pa. The accuracy of Manometer 3 was equal to 0.3% of the reading, plus 0.3 Pa [110]. No accuracy information was available for Manometers 1 and 2 [108], [109].

Some temperature measurements were also taken. These measurements were taken using k-type thermocouples, attached to Anemometer 1 [105], as well as using the thermistor integrated into the anemometer's probe. The thermistor has an accuracy of ±0.8 K, and the thermocouples have an accuracy of ±(0.2%+0.5 K) [105]. In both cases, readings have a resolution of 0.1 K [105].

A calibration of the thermistor and thermocouples used was carried out by placing the sensors close together for a period of 5 minutes, before taking readings from each sensor. This calibration was undertaken at room temperature. The readings all fell within a range of ±0.4 K, which is within the range of accuracies of the sensors [105]. For all measurements, readings from each sensor were adjusted to account for their deviation from the mean measurement in the calibration, i.e. readings from a sensor which gave a reading 0.1 K below the mean temperature during the calibration would be increased by 0.1 K in the final results.

2.2.2 Single rack tests

The single rack tests were designed to determine the relationship between Δp_{CH} and bypass flow rate for a number of different racks, in a controlled setting with no servers installed in the racks. The tests also investigated the potential to reduce rack leakage by taking practical measures to seal leakage paths, allowing recommendations to be made with respect to how to set up racks in data centres employing aisle containment.

Four 800 mm wide server racks were obtained, each from a different manufacturer. All racks were 42 U high, meaning that they could accommodate 42 pieces of IT equipment each with a standard height of 44.45 mm.

A duct was constructed from corrugated plastic, to enclose an air flow driven through the rack. One end of the duct was attached to Fan 1, using adhesive tape. The other end was attached to the rack front, again using the tape. A diagram of the duct is shown in Figure 2-7. The first section is Measurement Duct 1 (as described in section 2.2.1), which was 2 m long and of constant cross-section, allowing a uniform flow to develop. Section A had an expanding width, up to 800mm. Section B had an expanding height, up to 2010mm. Consequently, the end of section B was large enough to accommodate the rack. The individual pieces of plastic used to form the duct were joined using the tape.
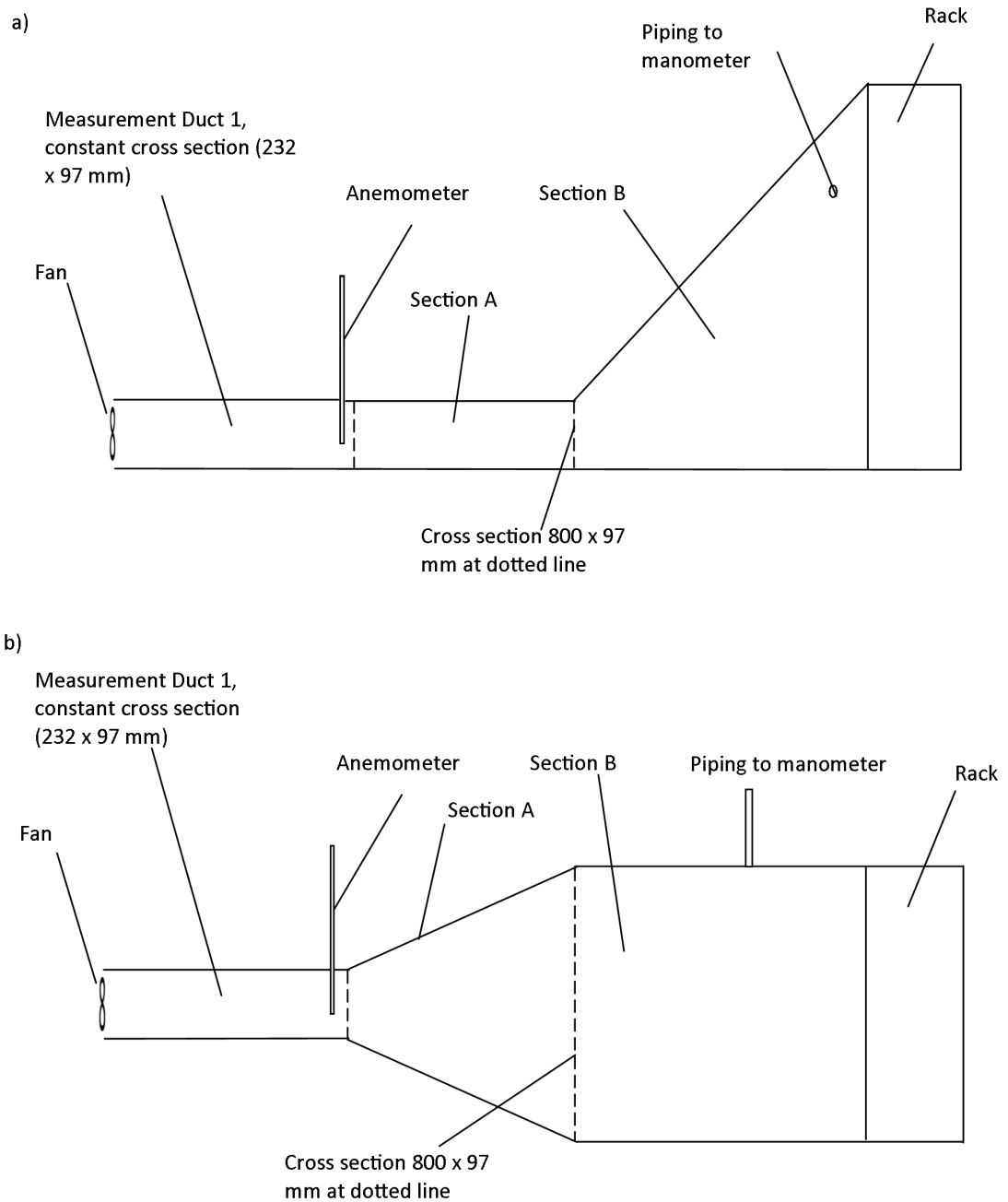
Figure 2-7. Experimental set-up for single rack tests in (a) side and (b) plan view.

A hole was drilled in the duct towards the end of Measurement Duct 1, in order to enable an anemometer probe to be inserted at this point. The space around the probe's point of entry was sealed with adhesive tape.

Another hole was drilled in section B, 200 mm from the rack front, 150 mm from the floor. A length of tubing was inserted into this hole, with its other end attached to a manometer, which allowed the differential pressure between the room and the front of the rack to be measured. This differential pressure is analogous to Δp_{CH}, with the duct supplying air to the

rack front representing the cold aisle and the room to which air is expelled from the rack representing the hot aisle. The tubing within the duct was positioned so as to be perpendicular to the direction of flow, such that the static pressure could be measured.

Potential leakage points in the duct (joins between sheets, entries for probes/tubes, attachment to rack and fan) were observed closely throughout the tests, but no significant leakage was found.

For each rack, an iterative process was undertaken whereby during tests measuring flow rate and Δp_{CH}, the key leakage paths were identified through the use of a smoke pen. Measures were taken to seal the identified leakage paths before repeating the process. Thus, leakage was gradually reduced. No IT equipment was installed in the racks, hence all air flow through the rack represents air which would bypass the IT equipment if any were installed. The gaps between the equipment rails were filled using two 27 U PlenaFill blanking panels [111], cut to size to give a total of 42 U. The meeting point of the two panels was sealed with adhesive tape, as were the points at which the panels met the top bracket and bottom bracket (as defined in Figure 2-2).

For each rack configuration, anemometer and manometer readings were taken over a range of fan speeds. The fan was first set to the desired speed, and the manometer reading recorded after it was judged to have settled. An anemometer reading was then taken.

2.2.3   The Test Data Centre

2.2.3.1   Setup of racks and enclosure

Having completed the single rack experiments, one rack (identified as Rack C in section 2.3.1) was selected to be used in experiments investigating bypass and recirculation in a contained aisle system. Four of the racks were bayed together within an enclosure with internal dimensions of 3.97 x 3.23 x 3.4 metres, referred to hereafter as the Test Data Centre. The Test Data Centre was composed of polycarbonate sheets connected with aluminium profiles, and incorporated 2 doors, allowing access to both the hot and cold aisles. On the wall opposite the doors were two circular holes enabling air to be supplied to and rejected from the cold and the hot aisle, respectively. Photographs of the Test Data Centre are shown in Figure 2-8. Schematics of the Test Data Centre setup are shown in Figure 2-9 to Figure 2-12.
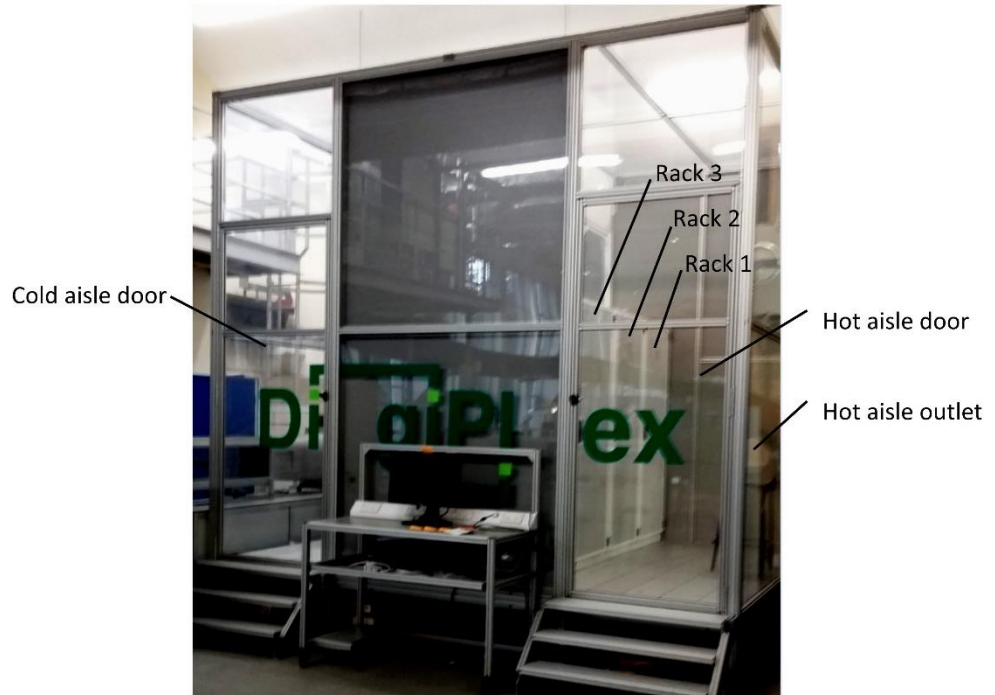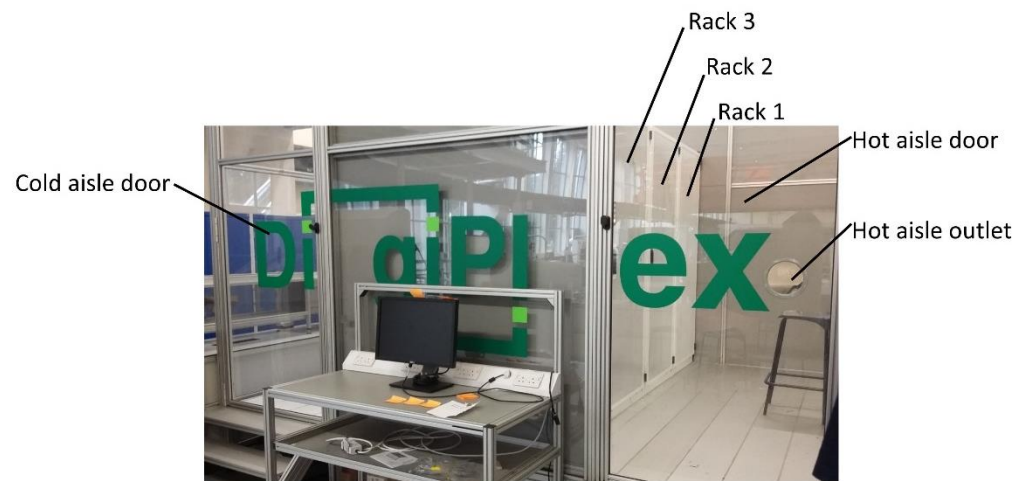
Figure 2-8. Photographs of Test Data Centre, shown a) zoomed out and b) close up for clarity.
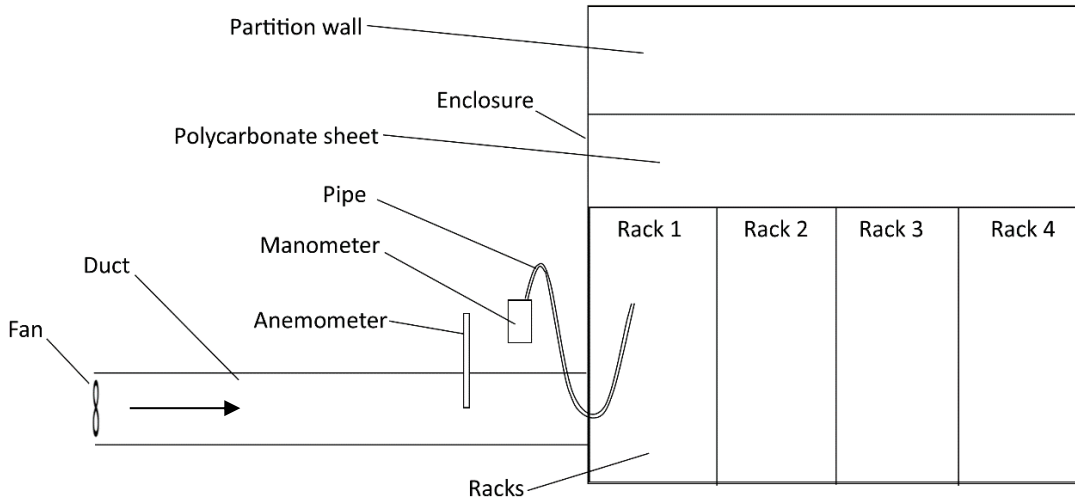
Figure 2-9. Schematic of Test Data Centre – front view.
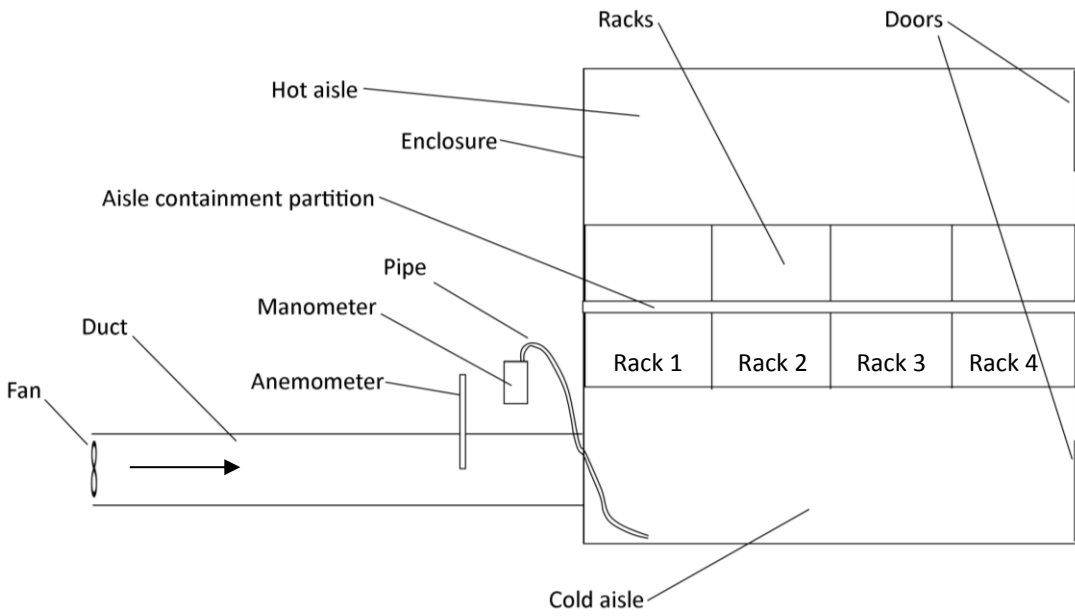
Figure 2-10. Schematic of Test Data Centre – plan view.
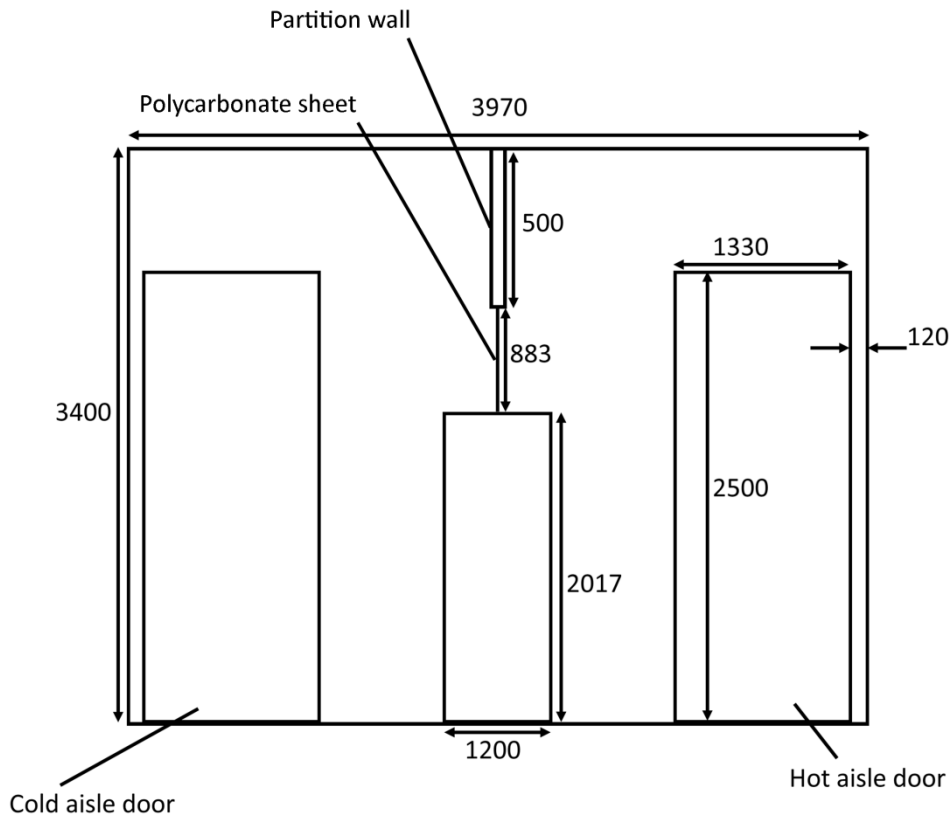
Figure 2-11. Schematic of Test Data Centre showing doors (dimensions in mm).
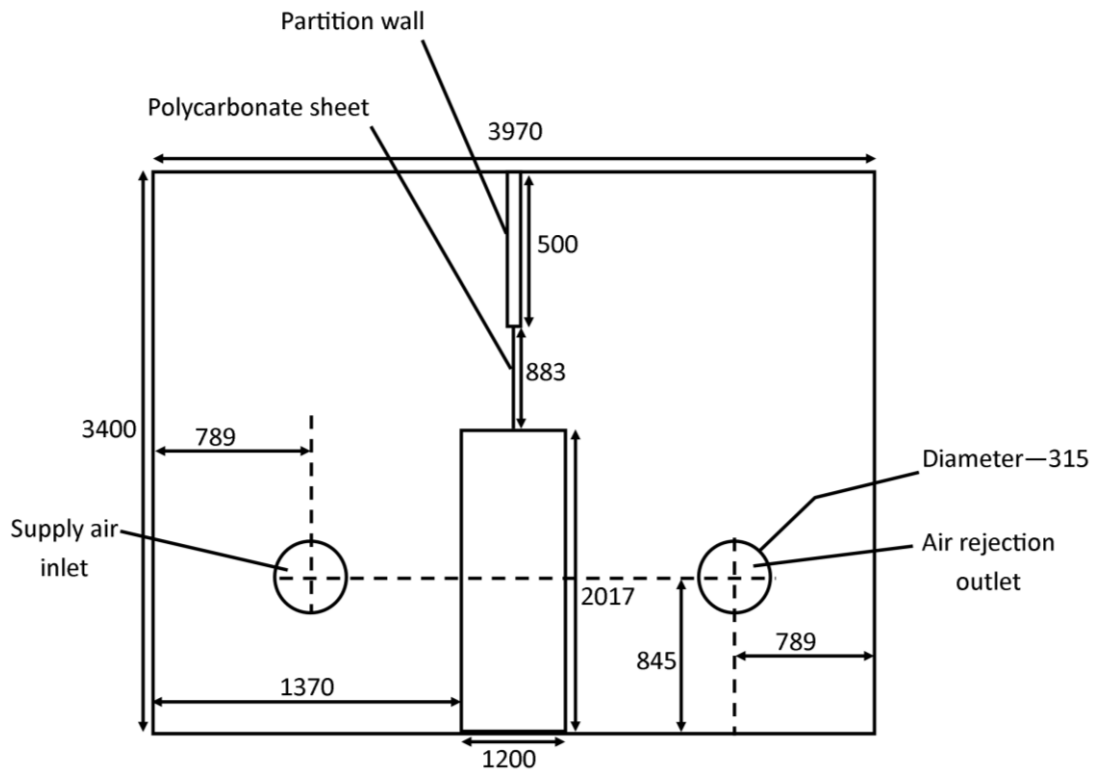
Figure 2-12. Schematic of Test Data Centre showing air inlet and rejection ducts (dimensions in mm).

The space between the bottom of each rack and the floor was bridged with a folded metal section. This was a standard part, supplied by the rack manufacturer for that purpose. The space between the tops of the racks and the ceiling of the enclosed space was bridged in two separate sections. The upper section was a partition wall, constructed of folded steel enclosing a cardboard honeycomb structure. The interfaces between separate sections of this wall, between the wall and the ceiling, and between the wall and the sides of the enclosure, were sealed using silicone sealant.

The space between the bottom of the partition wall and the tops of the racks was bridged with polycarbonate sheets, connected with plastic profiles. Silicone sealant was used to seal the gap between this section and the tops of the racks, and between the sides of this section and the walls of the Test Data Centre. The interface between this panel and the partition wall was also sealed with silicone sealant.

Each pair of adjacent racks were bayed together by bolting a short piece of sheet metal to the top of each rack in the pair. The interfaces between the racks were sealed using silicone sealant. The interface between rack 4 and the wall of the enclosure was sealed using silicone sealant (the rack numbering scheme is shown in Figure 2-9). Between rack 1 and the adjacent wall there was a gap of around 20 mm. This was filled using a sheet of plastic, with silicone sealant applied at its interfaces with the rack and the wall.

The gaps between the equipment rails were filled using 27 U PlenaFill blanking panels [111]. The upper panel was positioned such that 1 U overlapped the section joining the 2 equipment rails at the top of the rack. The lower blanking panel was cut to a length of 18 U and positioned such that 1 U could be folded through 90 degrees to lay flat against the floor of the rack. The interface between the two blanking panels was sealed with adhesive tape, as were the interfaces between the blanking panels and the top and bottom brackets.

Other potential leakage paths were filled variously using rubber strips and pieces of expanding foam, provided by the rack manufacturer. Care was taken to ensure that the methods used to seal the racks would be repeatable in a live data centre, and used only parts available from the rack manufacturer.

2.2.3.2   Determination of bypass and recirculation

These tests were designed to determine the relationship between Δp_{CH} and bypass flow within a data centre employing aisle containment, with best practice undertaken with respect to minimisation of rack and containment leakage. For the first tests conducted in the Test Data Centre, no IT equipment was installed in the racks, hence these tests allowed

bypass to be measured in a controlled setting with no air flow through IT equipment. Further tests were then carried out with equipment installed in the racks, to enable the impacts of the presence of IT equipment on bypass in an operational data centre to be investigated.

Air was initially supplied to the enclosure using Fan 1 [103], which was connected to Measurement Duct 2 using flexible aluminium ducting with a diameter of 315 mm. Measurement Duct 2 was connected to the inlet to the cold aisle using similar flexible ducting. Air speed measurements were taken within Measurement Duct 2, using Anemometer 1, in order to determine the volumetric flow rate supplied to the cold aisle ($\dot{V}_{supply}$). Manometer 3 was used to measure the pressure differential, Δp_{CH}. This was achieved by feeding a pipe attached to the manometer through the flexible ducting and into the cold aisle, with the other side of the manometer open to the air outside the enclosure. The free end of the pipe was fixed to the wall of the cold aisle throughout the tests during which flow rates and temperatures were measured. Preliminary tests showed that opening the hot aisle door allowed the pressure in the hot aisle to fall so as to equal the pressure outside the Test Data Centre (within the resolution of the manometer used). This confirmed that measuring Δp_{CH} as the difference between the cold aisle pressure and that outside the Test Data Centre was appropriate. Further pressure measurements were also undertaken during some tests, with the free end of the pipe fixed in different positions within the cold aisle, in order for the variation of pressure within the cold aisle to be investigated.

Prior to measuring the rack and containment leakage, the potential for leakage from the Test Data Centre to the surrounding environment was assessed. This was achieved by closing both doors, sealing the hot aisle air rejection outlet, removing blanking panels from the racks (to allow free flow between the aisles), and measuring $\dot{V}_{supply}$ and the pressure differential between the interior and exterior of the Test Data Centre, Δp_{tdc}, at a variety of fan speeds. Such tests were carried out both with positive and negative $\dot{V}_{supply}$. Negative $\dot{V}_{supply}$ was achieved by placing Fan 1 [103] inside the cold aisle, and connecting it to the inlet to the cold aisle using flexible ducting. Hence, the relationship between the leakage from the enclosure, $\dot{V}_{leak}$, and Δp_{tdc}, could be determined.

After determining the enclosure's permeability, tests were carried out to assess the leakage through the racks and containment system. This was achieved using the same method, except that blanking panels were in place, and air was rejected from the hot aisle through the door, which was left open, while the air rejection outlet in the hot aisle was sealed. In this way, a pressure differential could be built up between the two aisles.

After measuring the flow rate from the cold to the hot aisle with no equipment present in the racks, tests were undertaken with a 3 U Sun Fire v40z server [112], [113] added to the Test Data Centre, installed in slots 12-14 of rack 4 (where slot 1 is the lowest slot in the rack). These tests were intended to determine the impact of the presence of the server on the magnitude of bypass. Initially, the entire front of the server and the interface between the server and the rack were sealed with adhesive tape, and the server was switched off. This allowed a test to be undertaken which was essentially a repeat of the test prior to the introduction of the server. Following this, the interface between the server and the rack was unsealed, with small gaps present between the top and bottom of the server and the adjacent blanking panels, as well as at the sides of the server where it was attached to the rails. This allowed the previous test to be repeated, with the inclusion of leakage at the interface between the rack and the server. Since these tests were undertaken with no flow allowed through the server, $\dot{V}_{supply} = \dot{V}_{BP}$ in this instance (where $\dot{V}_{BP}$ is the total bypass flow rate, and after accounting for $\dot{V}_{leak}$ as described in section 2.3.2.1). Hence, the relationship between $\dot{V}_{BP}$ and Δp_{CH} was established with no flow through the server itself.

The server front was then unsealed, and a duct constructed from corrugated plastic fixed to the rear of the server using adhesive tape (this arrangement is depicted in Figure 2-13). The server duct channelled the server outlet flow into Measurement Duct 4. This allowed the flow rate through the server, $\dot{V}_{server}$, to be measured, by taking air velocity measurements within Measurement Duct 4 using Anemometer 1 [105]. This in turn allowed $\dot{V}_{BP}$ to be determined as the difference between $\dot{V}_{supply}$ and $\dot{V}_{server}$ (after accounting for $\dot{V}_{leak}$).

Tests were undertaken with the server in 3 conditions: 'switched off', 'idle' and 'stressed'. The 'stressed' condition entailed running a program called stresslinux [114], which is designed to test the operation of a server under a computational load. A 100% load was applied to each of the server's CPUs and its random access memory (RAM), concurrently. The 'idle' condition entailed simply booting up the stresslinux program without running a stress test.

Having installed a heat load in the Test Data Centre, i.e. the server, tests were undertaken to characterise the temperature field. This was undertaken in order to enable validation of CFD models, as will be described in Chapter 5. The methods used to investigate the temperature field are described in section 2.2.3.3.
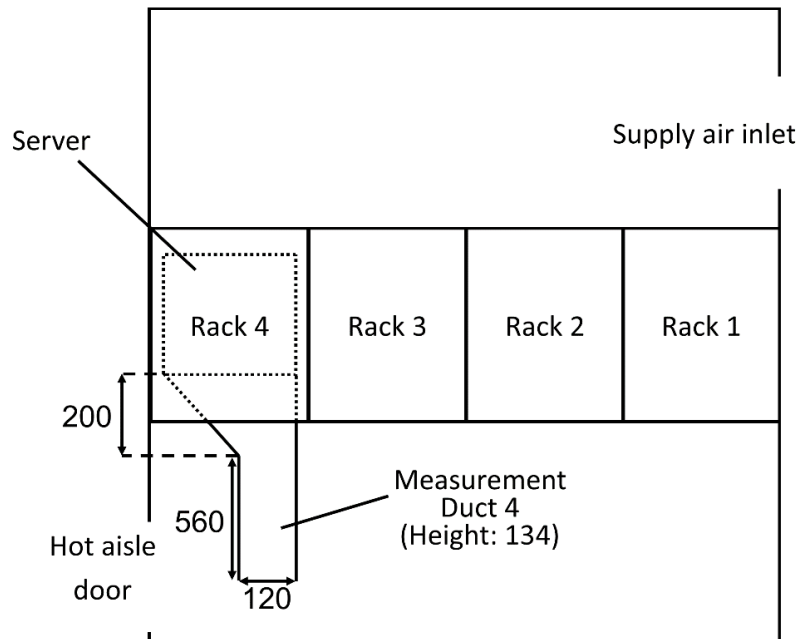
Figure 2-13. Plan view of Test Data Centre with the v40z server in place (dimensions in mm).

Further tests were conducted after removing the server from the rack, and installing eight 4 U load banks, in slots 1-16 of racks 1 and 4, where slot 1 is the lowest of the 42 slots. The load banks used were HAC230-6RM units, manufactured by Hillstone Loadbanks [115]. Such load banks consist simply of heating elements and fans, and are commonly used in commissioning of data centres to test the ability of the cooling and power infrastructure to cope with a facility's design IT power capacity. The load banks each had a variable heat load of 1.1-6.7 kW, and a flow rate of up to 120 $l.s^{-1}$ produced by 3 fans, according to manufacturer specifications.

Initially, the inlets of all load banks were covered over with a piece of corrugated plastic, which was sealed to the fronts of the load banks using adhesive tape. The interfaces between the load banks and the racks were also sealed. A test was then undertaken to establish the level of bypass with the load banks installed, in order to confirm that the introduction of the load banks did not impact bypass.

After undertaking the tests to establish the extent of bypass with no flow through the load banks, the inlets of the load banks in slots 9-16 of rack 4 were opened up, and the test repeated with the following conditions:

(i) the load bank fans switched off

(ii) all 3 fans in the 2 open load banks switched on

Test (ii) was then repeated with the load banks in rack 1, slots 9-16 uncovered, and with all load banks in rack 4 covered.

For the tests with flow allowed through the load banks, Fan 2 [104] was used to supply air into the cold aisle instead of Fan 1, since Fan 1 could not generate the larger flow rates required in these tests. During initial tests using Fan 2 it was noted that the velocity field across Measurement Duct 2 showed very large local variations, leading to poor repeatability in the results. A decision was then taken to include a flow straightener in the supply duct and to introduce a duct of smaller cross section in which to take the air speed measurements (Measurement Duct 3), since such measures have been shown to improve the accuracy of flow rate measurements [107]. The flow straightener consisted of plastic drinking straws of 150 mm in length and 10 mm in diameter, packed into a section of this duct. A schematic of this assembly is shown in Figure 2-14. The flow out of the load banks was channelled into Measurement Duct 5, in the same way as for the v40z server.
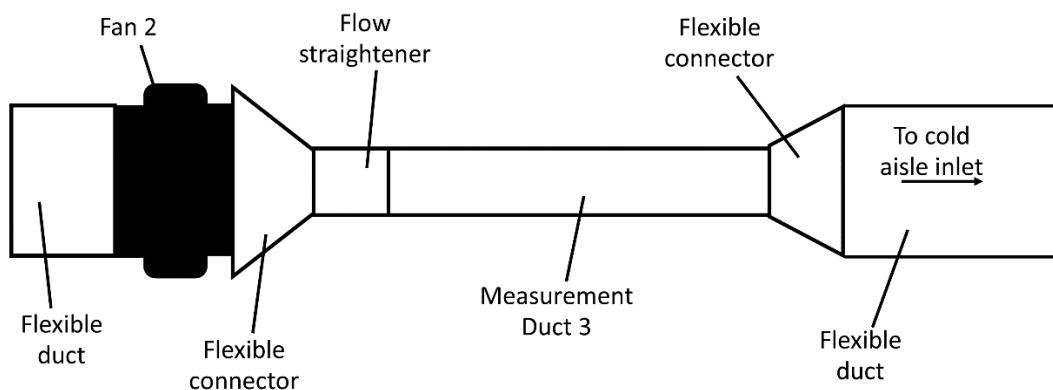
Figure 2-14. Schematic of supply duct after introduction of Measurement Duct 3 and flow straightener.

2.2.3.3   Determination of temperature field

A series of tests were undertaken in the Test Data Centre in which temperature measurements were taken, with electricity-consuming components installed in the racks. These tests were undertaken to provide data for validation of the CFD models described in Chapter 5. Firstly, temperature measurements were taken with the v40z server installed in the Test Data Centre, during the 'idle' and 'stressed' tests. The thermocouples were positioned in the air supply duct, in the server outlet duct, at the inlet to the server (near to the centre of the server's largest inlet), and at the front and back doors of rack 2, (1.8m above the floor and at the horizontal centre in each case).

Further temperature measurements were taken with the server in the 'stressed' condition after removing the server duct, over a range of supply fan speeds. The intention of removing the server duct was to give a velocity profile at the server outlet more akin to normal operating conditions, enabling validation of the CFD models described in Chapter 5. At each supply fan speed, six temperature measurements were taken at the rear door of rack 4, three at the rear door of rack 2, three within rack 4 and one at the server inlet (Figure 2-15 shows the exact positions of these measurements). This arrangement did not enable the flow rate through the server to be calculated, due to the lack of uniformity in the velocity field at the server outlet.
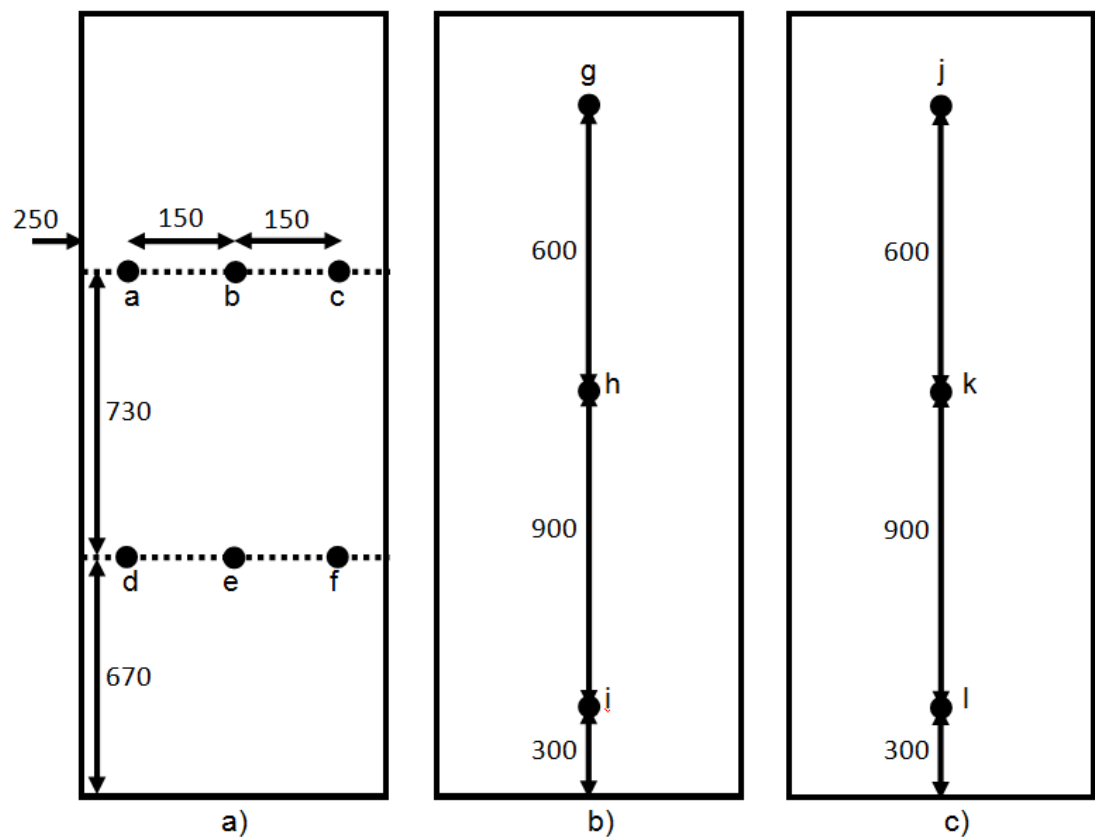
Figure 2-15. Temperature measurement points for tests using v40z server, with the server duct detached, a) at the rear door of rack 4, b) within rack 4, at the rack's horizontal centre and c) at the rear door of rack 2 (dimensions in mm). The measurement points are labelled with letters for ease of reference.

Further temperature measurements were taken in the Test Data Centre with the load banks in place. In order to more closely replicate the flow regime produced by servers, the duct used to measure the flow rate at the outlet of the load banks was removed, and replaced with a duct of the same cross-sectional area as the block of 4 load banks (700 by 408 mm), and with a length of 507 mm, such that the total distance from the inlets of the load banks

to the outlet of the duct was equal to the length of the v40z server (757 mm) [113]. It was not possible to measure the flow rate through the load banks with this setup due to having a short duct with a large cross sectional area, leading to a very non-uniform flow field with a considerable component of swirl. Two tests were undertaken with the load banks set up in this way as described in Table 2-3. The positions of the temperature measurements taken are shown in Figure 2-16 to Figure 2-19.
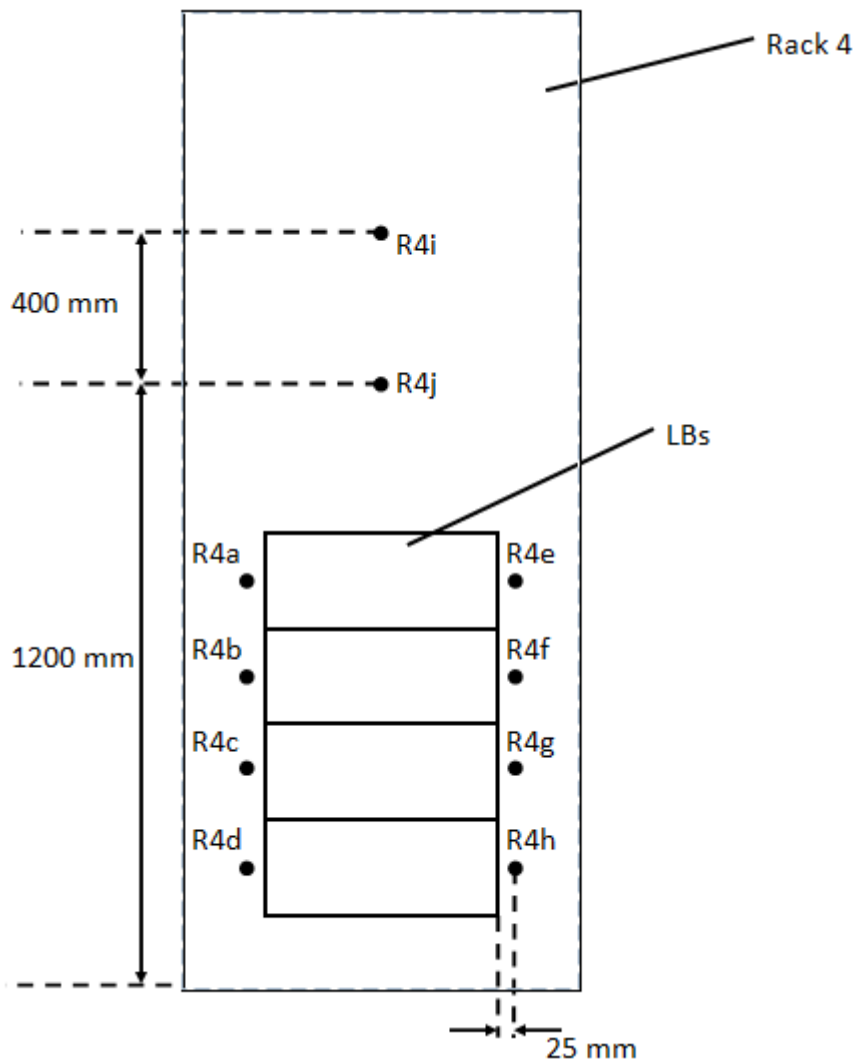
Figure 2-16. **Measurement points with the load banks operating in rack 4.** **Measurement points R4a-R4h were all positioned 40 mm back from the load** **bank inlets, and at the centre of the respective adjacent load bank by height.** **R4i and R4j were positioned 530 mm back from the load bank inlets and at the** **centres of the load banks by width.**
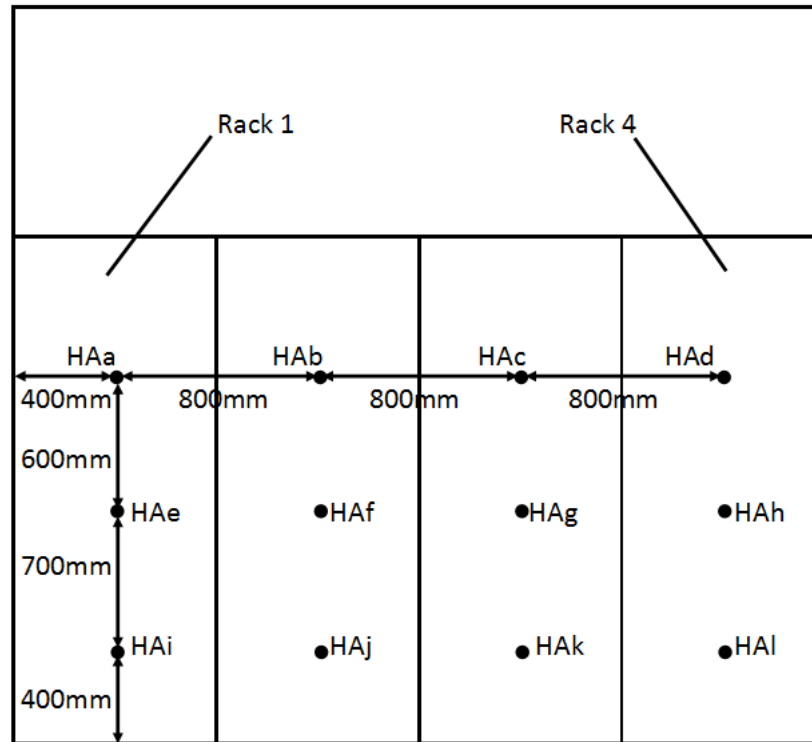
Figure 2-17. Measurement points in the hot aisle during the tests with the load banks installed. Points HAa-HAl are positioned 950 mm from the side wall of the hot aisle. A further set of points, HAm-HAx, are positioned 550 mm from the side wall of the hot aisle, in a similar grid format.
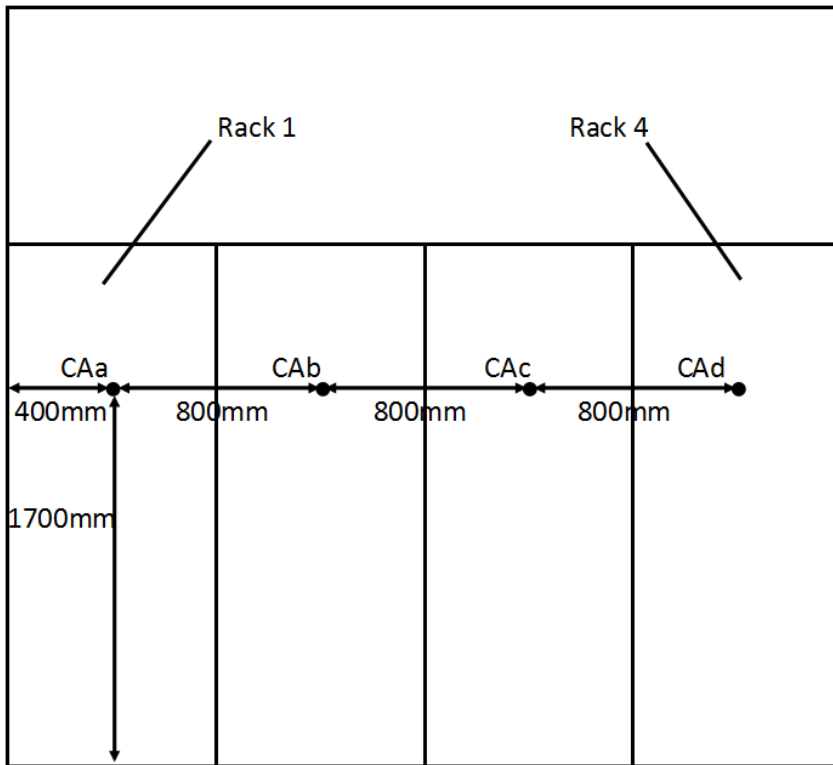
Figure 2-18. Measurement points in the cold aisle for tests with the load banks installed. All points are positioned 666 mm from the side wall of the cold aisle.
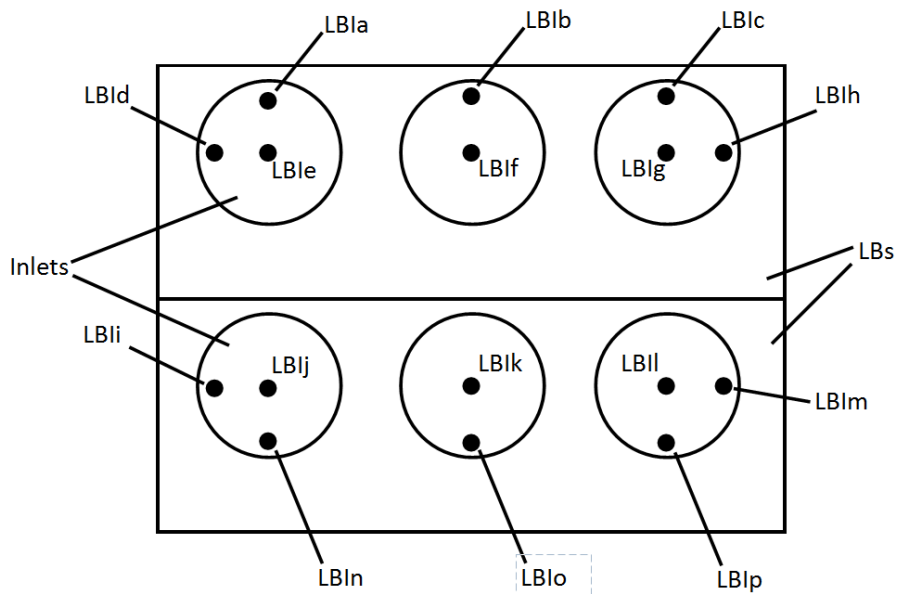
Figure 2-19. Measurement points at the load bank inlets. Measurement points LBIe-LBIg and LBIj-LBII were at the centre of each load bank inlet, with the other measurement points positioned 5 mm from the edge of the respective inlet.

For all temperature measurements, the respective sensors were kept in place for a minimum of 5 minutes before a reading was taken. Preliminary testing showed that this was sufficient time for the reading to have settled.

For the tests with heat loads in operation in the load banks, the hot aisle door was closed, and air allowed to escape from the hot aisle through the air rejection outlet (depicted in Figure 2-12). This change was made due to concerns that air was entering the Test Data Centre through the hot aisle door, which in turn affected the temperature field. The change resulted in pressures within the hot aisle exceeding the pressure outside the Test Data Centre. This necessitated a change to the process of measuring Δp_{CH}. For these tests, Δp_{CH} was measured by feeding a pipe attached to the manometer through the hot aisle outlet, and fixing it to the wall of the hot aisle. As with previous Δp_{CH} measurements, a pipe attached to the other end of the manometer was fed into the cold aisle. Hence, the reading on the manometer showed the pressure differential between the two aisles. Although for periods during which temperature measurements were being taken, the positions of the pipes within each aisle were fixed, further pressure measurements were also undertaken to investigate any variation in pressure in each aisle.

2.3   Results

2.3.1   Individual rack tests

Table 2-2 summarises the methods used to seal the various key leakage paths for each rack, both in the 'as delivered' state (i.e. after setting them up as recommended by the manufacturers, and with blanking panels installed), and after leakage minimisation. See Figure 2-2 for definitions of leakage path regions and key rack components.

As delivered				
Leakage path	**Rack A**	**Rack B**	**Rack C**	**Rack D**
Region A	Folded metal	Folded metal	Plastic strip	Foam
Regions B & D	Foam (poorly fitting)	Foam (poorly fitting)	Rubber	Foam (poorly fitting)
Region C	Foam	Folded metal	Folded metal	Corrugated plastic
Other paths	Considerable leakage out of the sides and bottom of the rack		Some leakage out of the bottom of the rack	Considerable leakage from sides and bottom of rack
Dominant leakage paths	Regions B & D, bottom and sides of rack	Regions B & D	Holes in folded metal sections	Regions B & D, bottom and sides of rack
After leakage minimisation				
Leakage path	**Rack A**	**Rack B**	**Rack C**	**Rack D**
Region A	Holes in, and gaps between, folded metal sections sealed with tape	Holes in folded metal sections sealed with tape	Holes in folded metal sections sealed with tape	Holes in folded metal sections sealed with tape
Regions B & D	Second layer of foam added	Second layer of foam added	A layer of foam added	Fixed foam in place with tape
Region C	Various holes in, and gaps between, folded metal sections sealed with tape	Holes in folded metal sections sealed with tape	Holes in folded metal sections sealed with tape	Layer of foam added to corrugated plastic
Other paths sealed	Rack sides sealed with tape, holes in the bottom plugged with foam		Bottom of rack sealed with tape	Bottom and sides of rack sealed with tape
Dominant leakage paths	Regions B & D	Regions B & D	Regions B & D	Regions B & D

Table 2-2. Methods used to seal leakage paths. See Figure 2-2 for definitions of leakage path Regions A-D.

Figure 2-20 shows the performance of each of the racks in their 'as delivered' states. There was considerable difference between the extents to which rack leakage had been minimized in the 4 racks, which can be seen from the differing performances. The error bars included in Figure 2-20, and in the remainder of section 2.3, represent the accuracies and resolutions of the instruments used, as detailed in section 2.2.1. In the case of flow rates, errors

represent the accuracy of the anemometer multiplied by the cross sectional area of the relevant measurement duct. No information on accuracy was available for Anemometer 2. Its accuracy has therefore been assumed to be the same as Anemometer 1's, since both are thermal anemometers.
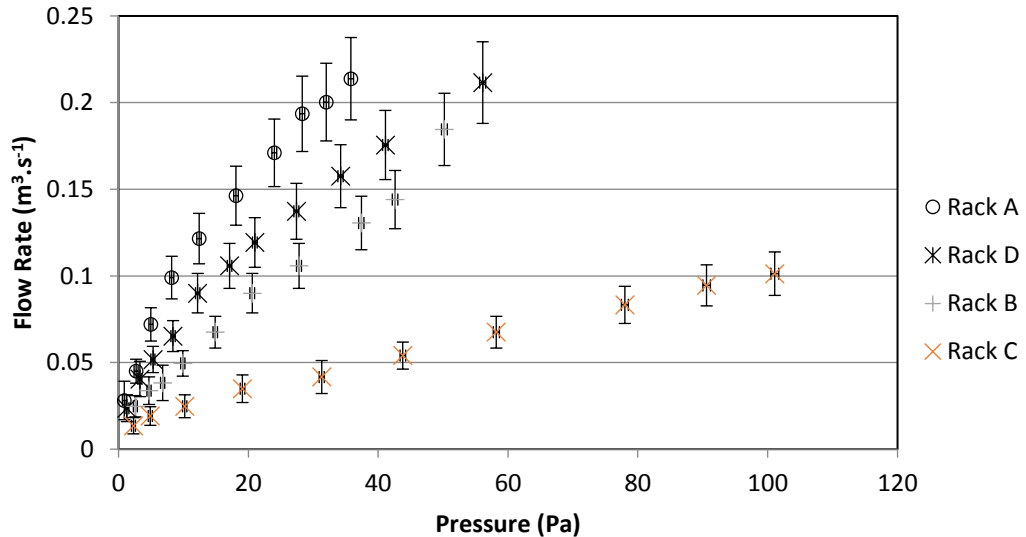
Figure 2-20. Comparison of flow rates through the 4 racks in their 'as delivered' states.

As described in section 2.2.2, a series of tests was undertaken for each rack, with attempts being made to progressively seal the most significant leakage paths. Figure 2-21 shows the results of these tests for Rack B. In the figure, 'Test B' represents the results shown in Figure 2-20 for Rack B. For 'Test A', the rack was set up in the same way as for Test B, except that the space between the uppermost blanking panel and the top bracket was left open, as was the space between the lower blanking panel and the bottom bracket (note that for all of the 'as delivered' tests, these spaces were sealed with tape). For Test C, new foam strips were added down the sides of the equipment rails (since those supplied by the manufacturer were poorly fitted). For Test D, perforations in the top and bottom brackets were sealed with adhesive tape. The results show a clear improvement resulting from taking simple, practical measures to reduce leakage. In particular, comparing the results of Tests A and B shows that leakage is approximately halved simply by sealing the spaces between the blanking panels and the top and bottom brackets. This demonstrates the potential to significantly reduce rack leakage in data centres employing aisle containment by taking practical action to seal the key leakage paths. Recommendations for approaches to this will be made in section 2.4.2.
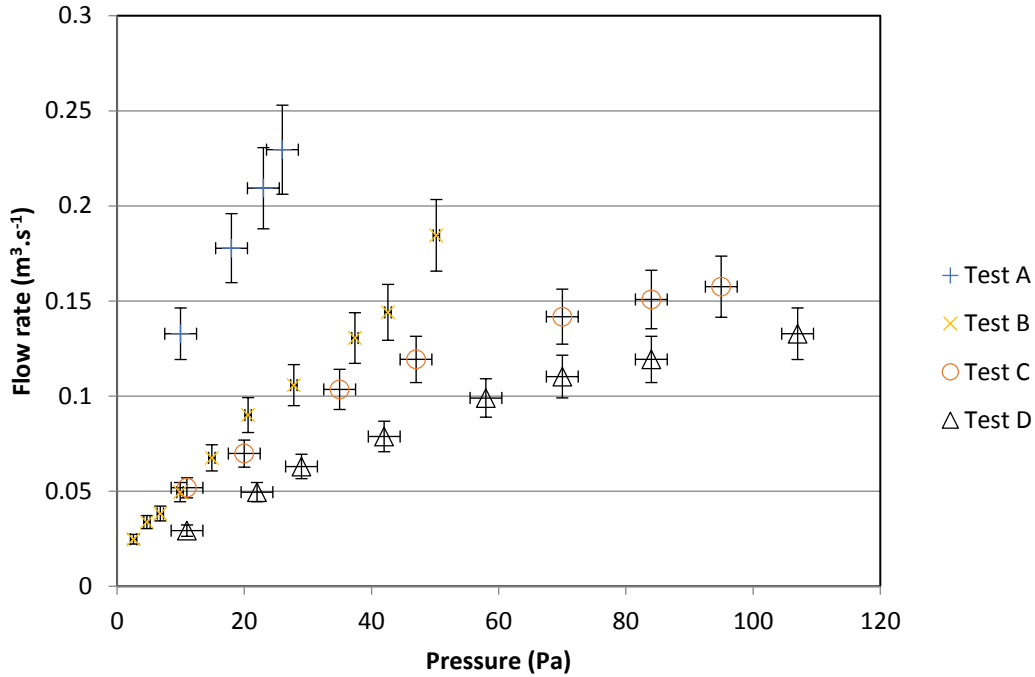
Figure 2-21. Results of iterative testing of Rack B.

Figure 2-22 shows the best performance of each of the 4 racks tested. In each case, numerous measures had been undertaken to minimize the flow at the key leakage paths identified, as described in Table 2-2. The range of performances reflects the variation in rack design, i.e. it was easier with some racks than others to minimise rack leakage.
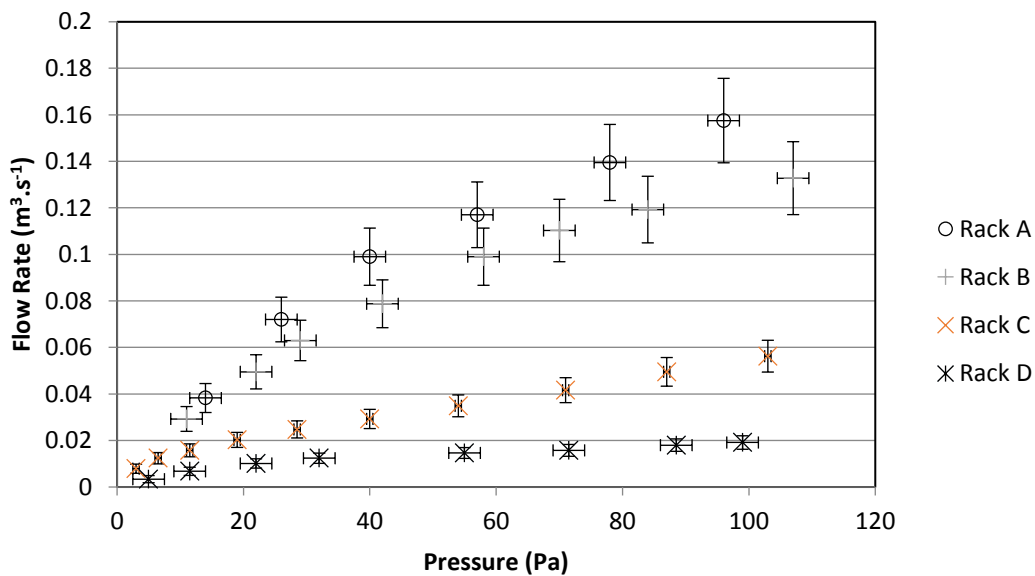
Figure 2-22. Comparison of flow rates through the 4 racks over a range of pressures, after minimizing flow through key leakage paths.

2.3.2 The Test Data Centre

2.3.2.1 Bypass and recirculation

As described in sections 2.2.3.2 and 2.2.3.3, a number of pressure measurements were undertaken to investigate any variation in pressure within each of the hot and cold aisles. For each aisle, variation between individual measurements and the mean pressure measured was within ±0.2 Pa, provided that areas of high velocity (such as the inlets to and outlets from the Test Data Centre, servers and load banks) were avoided. This range was within the reported accuracy of the manometer used [110].

Figure 2-23 shows the results of tests undertaken to determine the permeability of the Test Data Centre. Note that Δp_{tdc} is the pressure differential between the inside of the Test Data Centre and the laboratory.
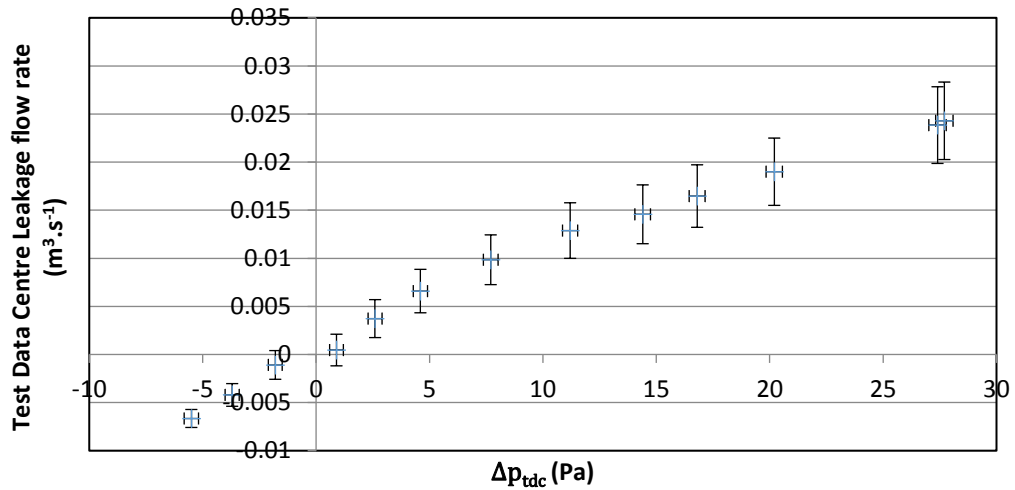
Figure 2-23. Leakage measured from the Test Data Centre, with air velocity measurements taken in the rectangular supply duct.

Figure 2-24 shows the flow rates from the cold to the hot aisle with the racks occupied with different equipment, but with no flow enabled through this equipment. Since no flow is allowed through the equipment, the flow rate from the cold to the hot aisle is equal to the bypass flow rate ($\dot{V}_{BP}$). The flow rates shown in Figure 2-24 represent the flow rate through the supply duct ($\dot{V}_{supply}$), minus the expected flow rate of air leaking from the Test Data Centre, $\dot{V}_{leak}$. $\dot{V}_{leak}$ is calculated by carrying out a linear interpolation on the results presented in Figure 2-23 to predict the expected leakage at the relevant value of Δp_{tdc}. This is then halved, since in the aisle leakage tests only half of the Test Data Centre was pressurised, whereas the entire Test Data Centre was pressurised during the tests whose results are presented in Figure 2-23.

The error bars included on flow rates in Figure 2-24 represent the numeric sum of the errors on the measurements of $\dot{V}_{supply}$ and $\dot{V}_{leak}$ (each calculated in the same way as described for Figure 2-20). The results of each test shown in Figure 2-24 fall within the same range when the error bars are taken into consideration. However, for the test with the server-rack interface left open, $\dot{V}_{BP}$ is greater than in the test with this interface sealed, over the entire pressure range. It should be considered that the error bars represent the reported accuracy of the instruments used, but that they are likely to refer both to systematic and random errors. If systematic errors make up a significant proportion of the reported errors, the results shown may still be taken to demonstrate that a significant contribution to $\dot{V}_{BP}$ is made by the interface between the server and the rack. Repeat measurements reported in section 2.4.1 will indicate the limited extent of random errors in the results. Where the server is introduced, but its inlet and the interface between the server and the rack are both sealed, there is no discernible change in $\dot{V}_{BP}$.
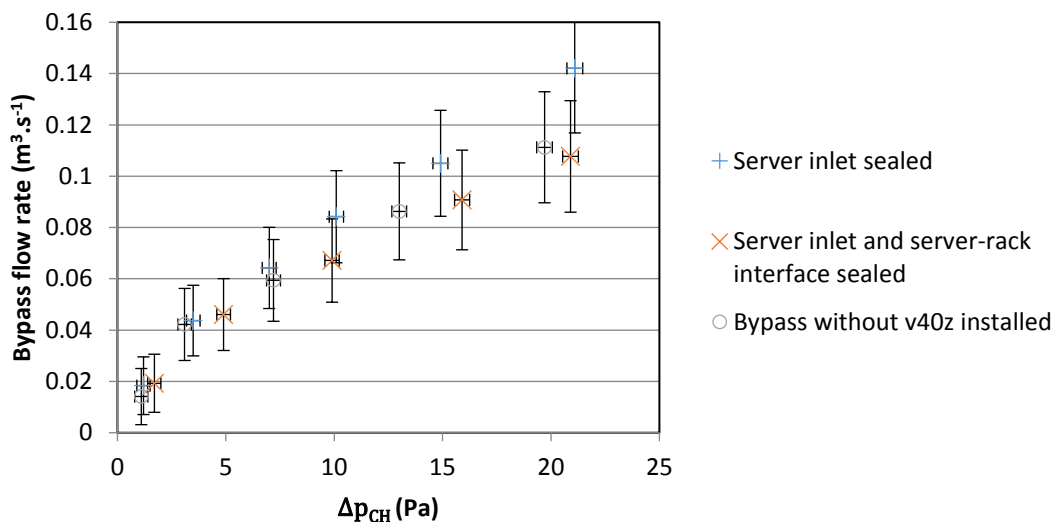
Figure 2-24. Relationship between $\dot{V}_{BP}$ and Δp_{CH} with and without the v40z server in place, with no flow allowed through the server.

Figure 2-25 shows $\dot{V}_{BP}$ for the experiments with flow allowed through the v40z server, compared with $\dot{V}_{BP}$ with the server front sealed, but the interface between the server and the rack and blanking panels left open. $\dot{V}_{BP}$ is calculated by subtracting $\dot{V}_{leak}$ and the flow through the server, $\dot{V}_{server}$, from $\dot{V}_{supply}$. Accordingly, the error bars for flow rates represent the sum of the error in the measurements of $\dot{V}_{supply}$, $\dot{V}_{leak}$ and $\dot{V}_{server}$.

The results show very little change in $\dot{V}_{BP}$ after the introduction of flow through the server. This indicates that bypass is not affected by the presence of the server, or by flow being

driven through it by the server fans. The results shown with negative Δp_{CH} represent tests with the supply air fan switched off, with the negative $\dot{V}_{BP}$ values indicating recirculation.

a)

b)
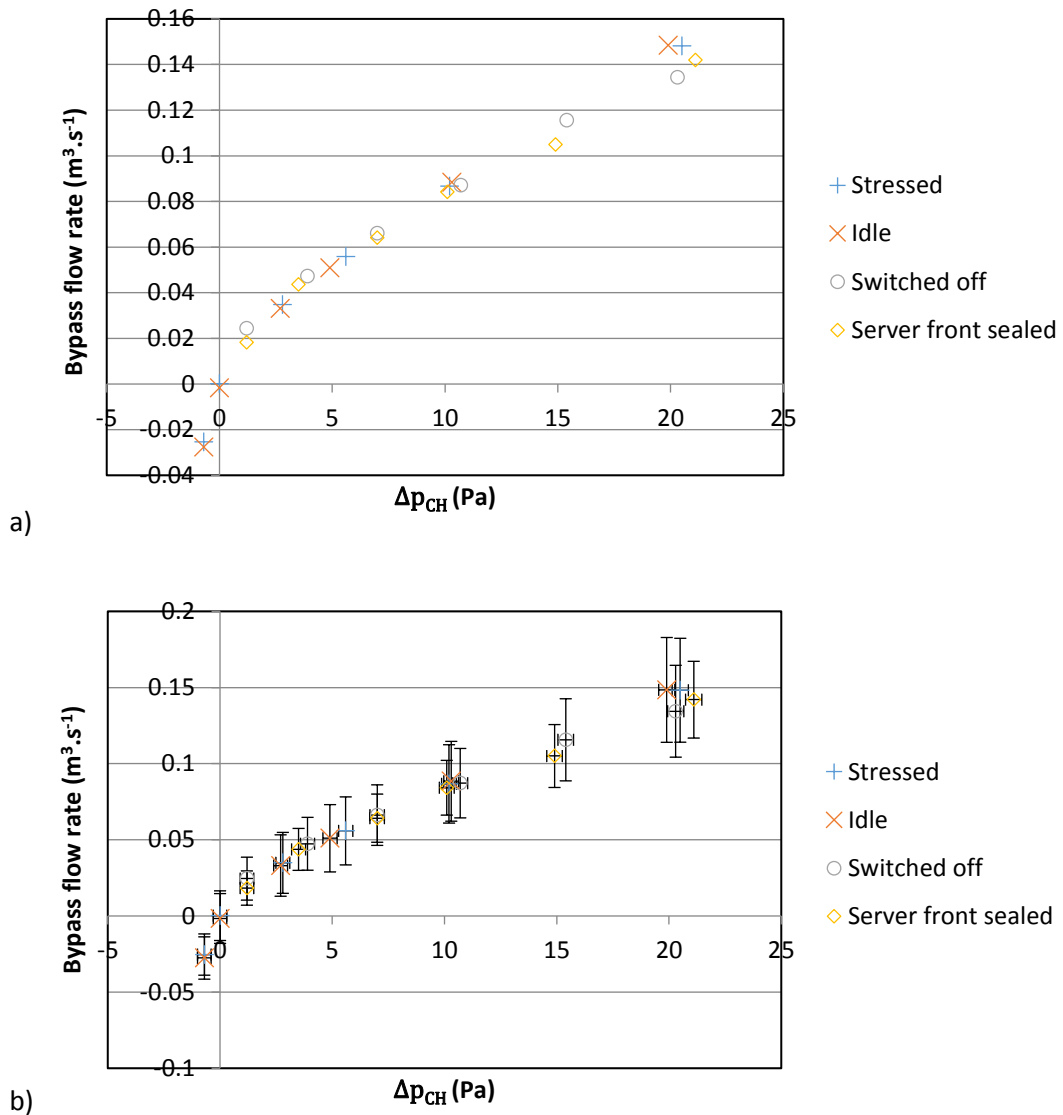
Figure 2-25. Bypass flow rates for tests with flow allowed through v40z server, compared with bypass measured with server inlet sealed. The results are shown (a) without, and (b) with error bars, for ease of interpretation.

Figure 2-26 shows the flow rates through the server under different conditions. As expected, the flow rate through the server is greater where the server is switched on, due to the action of the server fans. Surprisingly however, the flow rates are consistently lower where a computational load is applied to the server, than where the server is idle. This implies a greater server fan speed in the idle condition. However, since the difference is considerably smaller than the error bars, this is not a conclusive result.
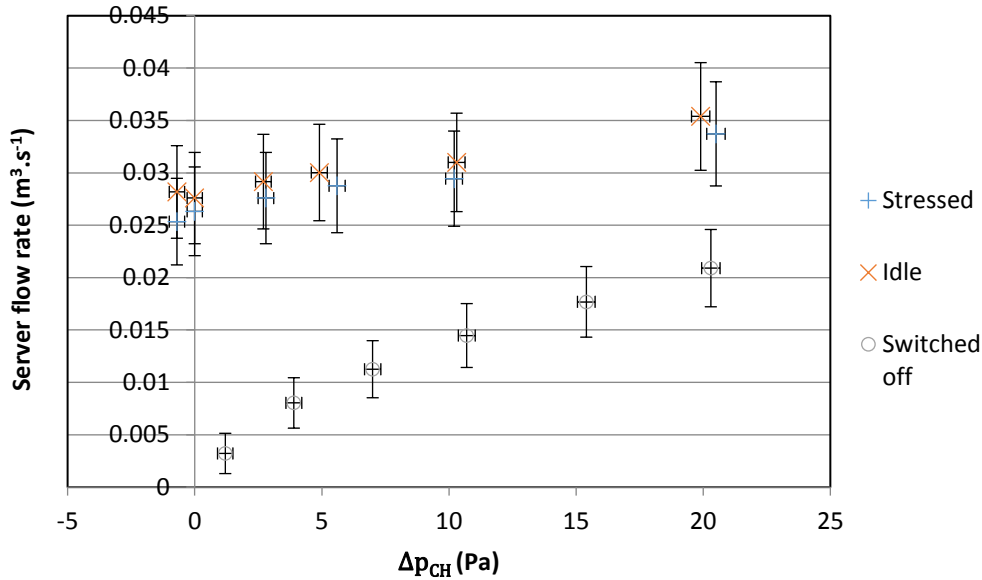
Figure 2-26. Flow rate through the v40z server under various stress conditions.

Figure 2-27 shows the bypass measured with no flow allowed through any equipment, with and without a single equipment slot left open. The error bars for flow rates represent the sum of the error in the measurements of $\dot{V}_{supply}$ and $\dot{V}_{leak}$. The introduction of the empty slot clearly makes a large contribution to the overall level of bypass.
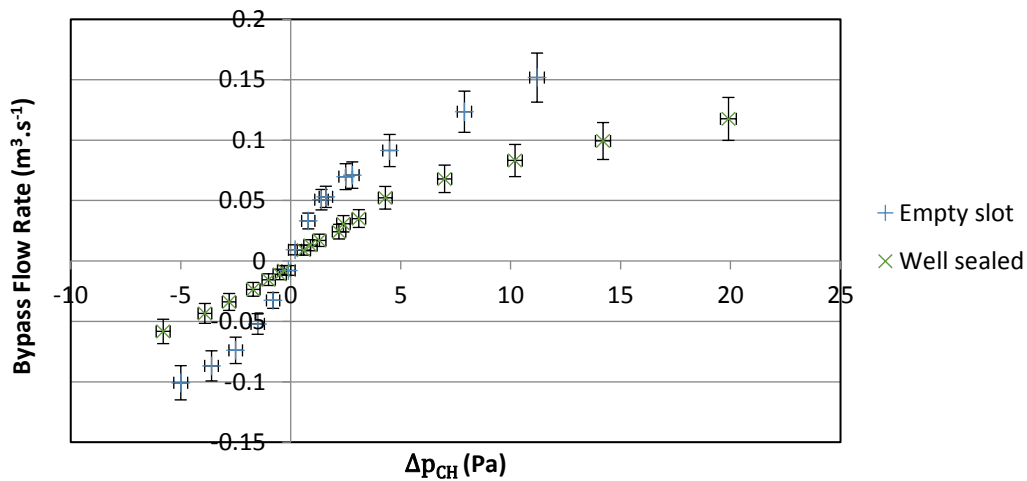
Figure 2-27. Relationship between $\dot{V}_{BP}$ and Δp_{CH} with and without one empty slot.

Figure 2-28 shows the effect of the introduction of load banks on bypass. Note that, in each case, the interface between the load banks and the rack is sealed with adhesive tape. For flow rates, the error bars represent the sum of the errors on the measurements of $\dot{V}_{supply}$, $\dot{V}_{leak}$ and $\dot{V}_{LB}$.

The results suggest that bypass was slightly increased by the introduction of load banks when their fans were switched on, regardless of which rack they were placed in. Note that, for both tests with the load bank fans switched on, a significant bypass flow rate was recorded where $\Delta p_{CH} = 0$. This is discussed further in section 2.4.3.
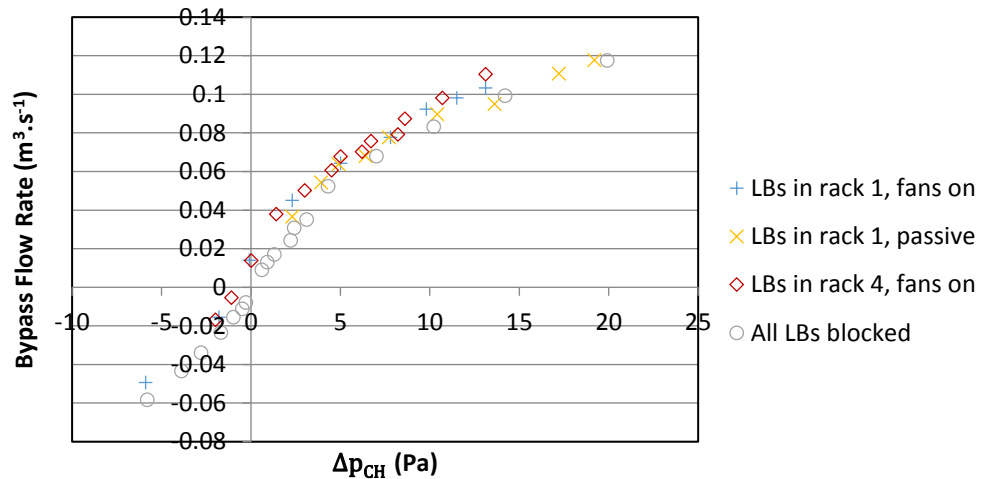
a)
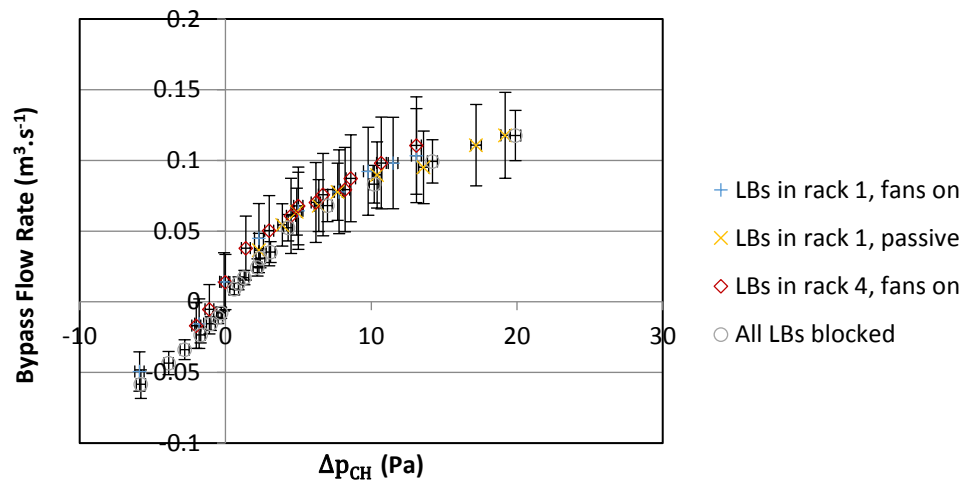
b)

Figure 2-28. The effect on bypass of the introduction of load banks, with and without their fans active. The results are shown (a) without, and (b) with error bars, for ease of interpretation.

In order to assess the relative importance of rack leakage and containment leakage, tests were undertaken using a smoke pen to investigate the dominant leakage paths, with a pressure differential applied across the partition. Whilst smoke could clearly be seen passing through the racks (primarily through Regions B & D), it was not possible to detect smoke

passing through the containment structure. Hence, it was concluded that bypass was dominated by rack leakage, rather than containment leakage.

2.3.2.2 Temperature measurements

Figure 2-29 shows the results of the temperature measurements taken with the v40z server in place, with the server duct removed. The measurements were taken at 2 different levels of Δp_{CH}. The error bars represent the accuracy of the instrument used in each position. The temperature measurements in the supply duct were taken using Anemometer 1's in built thermistor, with the other measurements taken using the thermocouples. There are no clear trends shown in the results, with, for example, a higher server inlet temperature corresponding to the test with highest Δp_{CH}, despite this test also having a lower supply duct temperature. Also, there was very little temperature rise shown in the hot aisle (measurement points a to l) in comparison to the inlet temperature, other than at positions very close to the server outlet (positions d to f). This implies that spatial and temporal variation in temperature within the laboratory in which the Test Data Centre was situated, in addition to errors in the temperature measurements, exceed variation in temperature resulting from the paths taken by hot air exiting the server. The temperature measurements undertaken with the v40z server flow measurement duct attached similarly showed no clear trends between Δp_{CH} and server inlet temperature, and small temperature rises in the hot aisle. Since the main function of the temperature measurements was to validate the CFD models, and since the main function of the CFD models is to predict inlet temperatures, it was concluded that a greater heat load was required in order to produce greater variation in the temperature field, and specifically, higher inlet temperatures. This led to the introduction of load banks, as described in section 2.2.3.2.

Temperature measurements were taken under 2 conditions with the load banks installed. In each case, 2 load banks were under operation, positioned in slots 9-16 of rack 4. In Test 1, all other equipment slots were closed. In Test 2, slot 17 in rack 4 was left open. The total power consumption of the 2 load banks was 6.76 kW. Table 2-3 shows the values of $\dot{V}_{supply}$, $\dot{V}_{LB}$, and Δp_{CH} recorded during these tests.

As shown in Figure 2-16 to Figure 2-19, a large number of temperature measurements were taken. These measurements are not all reproduced here, since their main function is to validate the CFD models described in Chapter 5. The results are shown in full in Appendix 3. Figure 2-30 shows the load bank inlet temperatures recorded during Tests 1 and 2. The figure shows the temperature rise at each point, that is, the difference between the temperatures measured at the load bank inlets and that measured in the supply duct. Both

tests showed raised temperatures at the load bank inlets, demonstrating recirculation (i.e. rack leakage from the hot to the cold aisle allows convective transport of heat from the load bank outlets to their inlets). It is clear that inlet temperatures are significantly higher for Test 2, particularly for the upper measurement positions (LBIe to LBIg). This is due to the open slot positioned immediately above the load banks in Test 2. The error bars represent the accuracy of the instrument used in each position, plus the accuracy of the instrument used to measure supply air temperature.
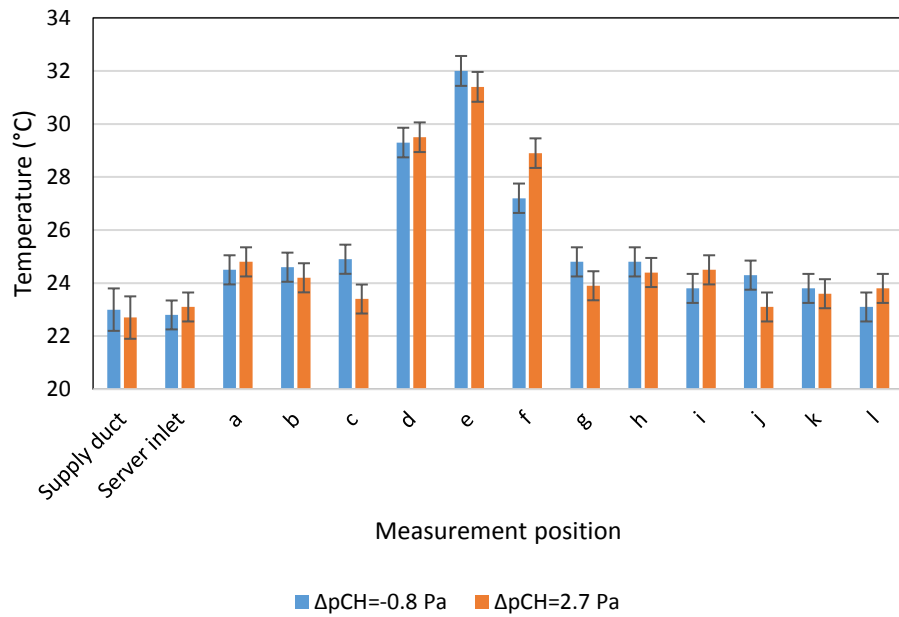
Figure 2-29. Temperature measurements from tests with v40z server installed (letters refer to sensor positions detailed in Figure 2-15).

Test	$\dot{V}_{supply}$ $(l.s^{-1})$	$\dot{V}_{LB}$ $(l.s^{-1})$	Δp_{CH} (Pa)
1	141.9	196.9	-4.49
2	98.3	177.8	-3.50

Table 2-3. Details of temperature measurement tests undertaken in the test data centre.
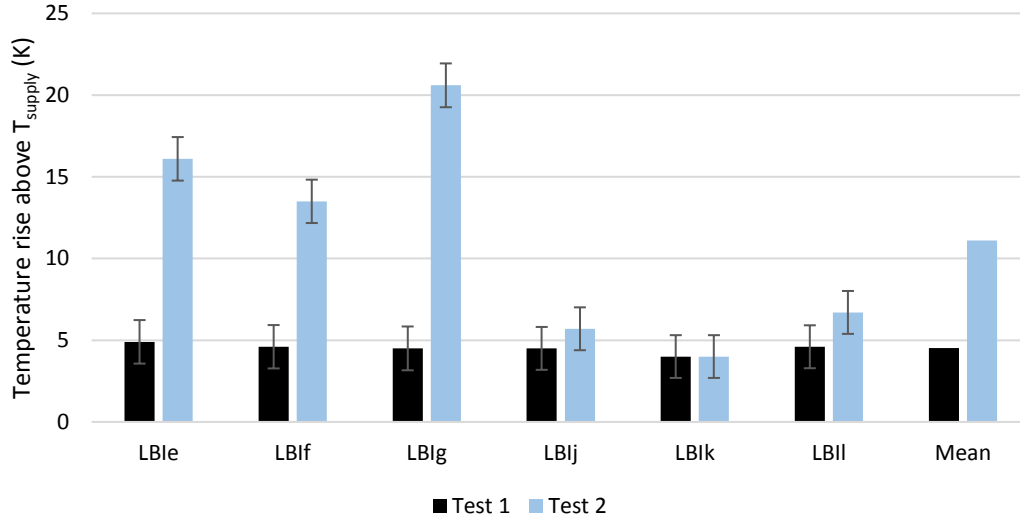
Figure 2-30. Temperatures at load bank inlets in Tests 1 and 2. Note that measurement positions are as depicted in Figure 2-19.

A calculation was also undertaken to determine the extent of heat loss from the hot aisle to the laboratory in which the Test Data Centre was housed. Since the hot aisle outlet temperature was not measured, this was assumed to be equal to the mean of the measurements made at points HAa to HAx, as defined in Figure 2-17. A conservation of heat calculation was undertaken, using Eq. 2-22. Here, U_{HA} is the heat transfer coefficient of the hot aisle walls ($kW.m^{-2}.K^{-1}$), $\dot{E}_{LB}$ is the load bank power consumption, T_{HA} is the mean temperature measured at points HAa to HAx, T_{supply} is the temperature measured in the supply duct, and A_{HA} is the combined surface area of the walls, ceiling and floor of the hot aisle. It is assumed that T_{supply} is equal to the ambient temperature in the laboratory in which the Test Data Centre is situated. Applying this calculation to the results from Test 1 yielded $U_{HA}A_{HA} = 0.192\ kW.K^{-1}$.

$$U_{HA}A_{HA}\left(T_{HA} - T_{supply}\right) = \dot{E}_{LB} - \dot{V}_{supply}\rho_{air}c_p(T_{HA} - T_{supply})$$

Eq. 2-22

2.4 Discussion

2.4.1 Repeatability

As discussed in section 2.3, the potential errors associated with the reported accuracies of the instruments used to measure velocity are fairly large. The significance of this is partially dependent on the form which these errors take – specifically whether they are random or systematic. Systematic errors would cast doubt over the exact magnitude of the flow rates reported, whilst allowing for relatively precise comparison between the results of individual tests, proportionally speaking.

In order to investigate the repeatability of the single rack tests, one of the tests undertaken on Rack D was repeated. Figure 2-31 shows the results of the tests. The differences between the air speeds recorded in the two tests are much smaller than would be expected if random errors of the magnitudes described by the instrument specifications were prevalent. Comparing the results numerically, after carrying out linear interpolations to estimate the velocity which would have been recorded in the original test at the pressures for which velocity measurements were taken in the repeat test, showed errors of up to 0.17 m.s^{-1}, with a mean of 0.08 m.s^{-1}. This corresponds to mean errors in calculated flow rates of 0.0018 m^3.s^{-1}. This is small in relation to the range of flow rates reported for the different racks and over the range of pressures reported in Figure 2-22 and Figure 2-20. Hence, conclusions over the differences in performances of different racks, and the impact of Δp_{CH}, may be drawn with considerable confidence. The exact magnitude of bypass for the single rack tests should be considered to be subject to the error bars displayed in Figure 2-20 and Figure 2-22. Note that the original test used Anemometer 1 for all measurements, whereas the repeat test used Anemometer 2 for velocities below 2 m.s^{-1}. The similarity of the results for the two tests lends credibility to the assumption made previously that both anemometers had similar accuracies (accuracy data was not available in the manufacturer specifications for Anemometer 2, as noted in section 2.2.1).
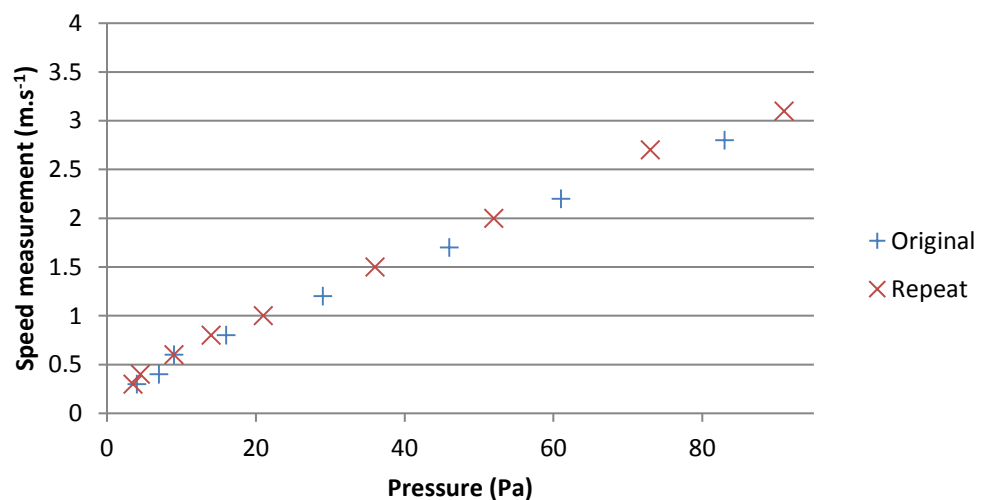
Figure 2-31. Rack D repeated test compared with original.

As with the single rack test results, the potential errors associated with the reported accuracies of the instruments used to measure velocity in the Test Data Centre experiments are fairly large. Again, the significance of these errors is partially dependent on whether these errors are random or systematic.

Figure 2-32 shows $\dot{V}_{BP}$ as calculated from two tests in which no flow was allowed through load banks or servers, both of which used Measurement Duct 2. This comparison acts as an investigation into the repeatability of the bypass measurements with this measurement setup, since although different equipment was installed in the racks during each of these tests, no airflow was allowed through this equipment, meaning that $\dot{V}_{BP}$ would not be expected to change.
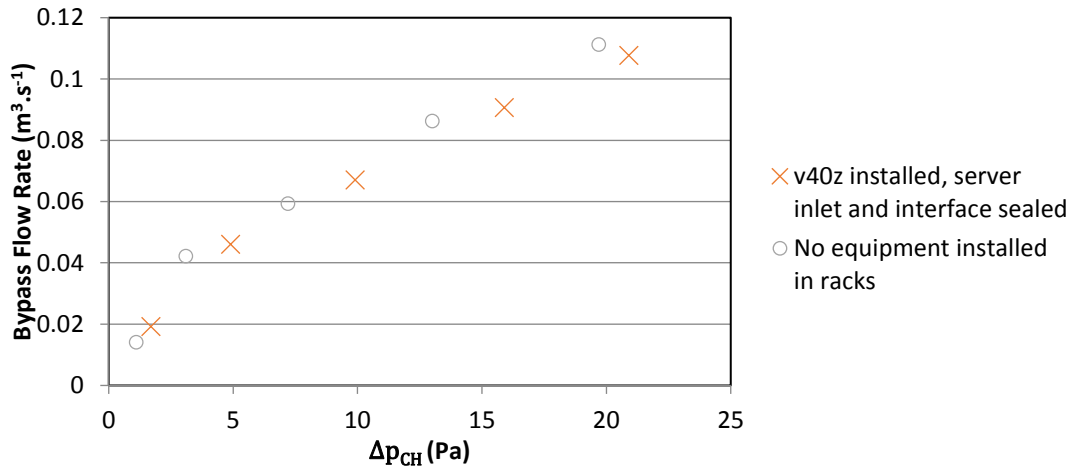
Figure 2-32. Results for bypass from 2 tests using Measurement Duct 2, in which no flow was allowed through any load banks or servers.

The results shown in Figure 2-32 were also compared numerically, as shown in Table 2-4. The results are shown at the levels of Δp_{CH} for which tests were undertaken with no servers/load banks installed. Linear interpolations were undertaken to approximate the likely measurements of $\dot{V}_{BP}$ at these levels of Δp_{CH} in the experiments with the server installed. The average magnitude of percentage difference was 13.7%, with a maximum of 36.2%. This shows that there is a significant level of variation in the results of tests using Measurement Duct 2.

Figure 2-33 shows $\dot{V}_{BP}$ as calculated from two tests in which no flow was allowed through load banks or servers, both of which used Measurement Duct 3. This comparison acts as an investigation into the repeatability of the bypass measurements with this measurement setup.

Δp_{CH} (Pa)	$\dot{V}_{BP}$ (m³.s⁻¹)		% difference	Absolute difference (m³.s⁻¹)
	No servers/load banks installed	v40z server installed		
1.1	0.0141	0.0125	11.4	0.0016
3.1	0.0422	0.0269	36.2	0.0153
7.2	0.0594	0.0510	6.1	0.0036
13.0	0.0863	0.0717	8.1	0.0070
19.7	0.1113	0.0938	6.9	0.0076

Table 2-4. Repeatability of $\dot{V}_{BP}$ measurements with rigid, circular supply duct.
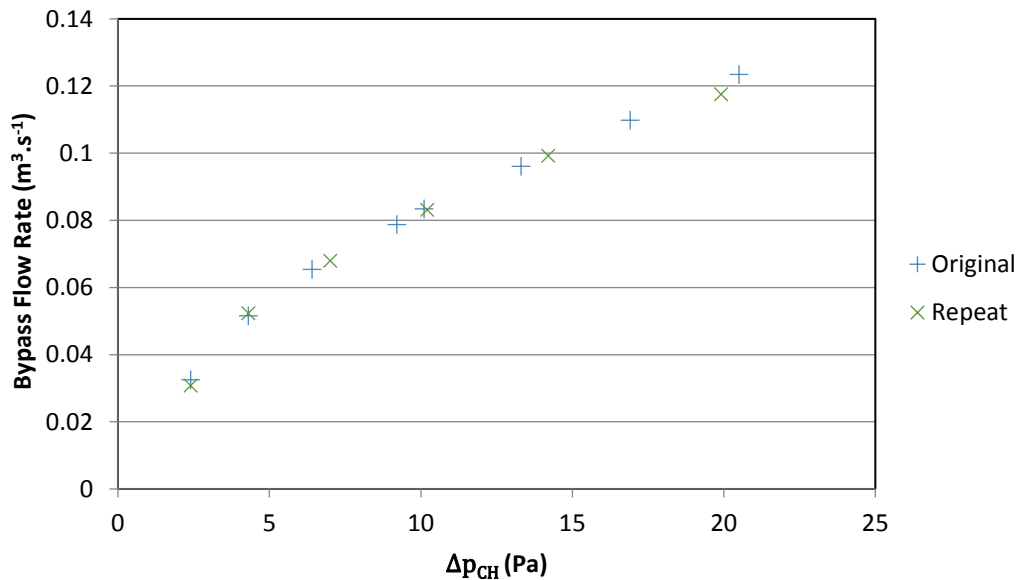
Figure 2-33. Results for bypass from 2 tests using Measurement Duct 3, in which no flow was allowed through load banks or servers.

The results shown in Figure 2-33 were also compared numerically, as shown in Table 2-5. The results are at the levels of Δp_{CH} for which tests were undertaken in the original test. Linear interpolations were undertaken to approximate the likely values of $\dot{V}_{BP}$ at these levels of Δp_{CH} in the repeat experiment. The average absolute percentage difference was 1.9%, with a maximum of 5.4%. This shows good repeatability with this measurement setup, well within the accuracy levels of the instruments used, and shows that comparisons between results of different tests using this technique can be made with confidence. However, with respect to the exact magnitude of flow rates reported, the reported accuracy of the instruments used should be considered, as discussed in relation to Figure 2-28.

Pressure (Pa)	$\dot{V}_{BP}$ ($m^3.s^{-1}$)		% difference	Absolute difference (m³.s⁻¹)
	Original	Repeat		
2.4	0.0326	0.0308	-5.4	-0.0018
4.3	0.0516	0.0523	1.4	0.0007
6.4	0.0654	0.0645	-1.3	-0.0009
9.2	0.0788	0.0784	-0.4	-0.0003
10.1	0.0834	0.0827	-0.9	-0.0007
13.3	0.0961	0.0957	-0.4	-0.0004
16.9	0.1099	0.1080	-1.8	-0.0019
20.5	0.1235	0.1195	-3.2	-0.0040

Table 2-5. Repeatability of $\dot{V}_{BP}$ measurements with rectangular supply duct.

Error bars shown on the temperature measurements in Figure 2-30 represent the reported accuracies of the instruments used to measure temperature. As with air speed measurements, these reported accuracies may represent some combination of systematic and random errors. Repeat temperature measurements were made during Tests 1 and 2 to further investigate the reliability of the results. The variation found during repeat temperature measurements taken in Test 1 is summarised in Table 2-6. Note that the data refers to the temperature rise at each point, i.e. the difference between the temperature measurement at that point and the most recent T_{supply} measurement. In absolute terms, the variation is greater for measurements taken in the hot aisle. This is due to the much greater spatial variation in temperature rises in this region, meaning that any temporal variability in temperature at a particular point due to turbulence will be magnified. The percentage differences in temperature rises are relatively small, and are small enough for their purpose – validation of the CFD models.

	Absolute difference (K)		Percentage difference (%)	
	Mean	Maximum	Mean	Maximum
Hot aisle	0.8	2.3	4.2	11.2
Cold aisle	0.3	0.5	7.2	11.6

Table 2-6. Summary of differences between repeat and original temperature rise, for measurements taken during Test 1.

2.4.2 Single rack tests

Comparing Figure 2-22 with Figure 2-20, it is clear that, for each rack, it was possible to significantly reduce rack leakage through the measures undertaken during testing. Figure 2-34 shows the flow rates for each rack at $\Delta p_{CH} = 20\ Pa$, before and after applying leakage reduction measures. Since measurements were not taken with Δp_{CH} set at exactly 20 Pa for each test, linear interpolations were undertaken to estimate the respective flow rates at this pressure. Whilst the relatively large error bars indicate considerable uncertainty over the exact magnitudes of the flow rates, the accuracy is sufficient to show that there is large variation in the flow rates recorded for the different racks, and that large reductions in flow rates were achieved through the measures undertaken to reduce leakage. In the 'as delivered' state, Rack A displays more than four times the flow rate of Rack C (at $\Delta p_{CH} = 20\ Pa$), and the reductions in leakage achieved for each rack range from 41 to 92%.
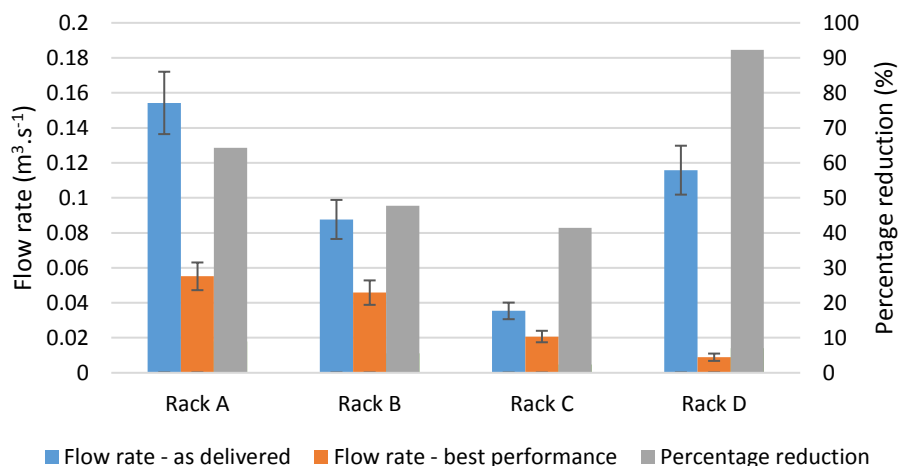
Figure 2-34. Reduction in flow rate achieved for each rack after undertaking measures to reduce leakage (all flow rates shown for $\Delta p_{CH} = 20\ Pa$).

Through the iterative testing of each rack, during which measures were taken to progressively tackle the most important leakage paths, it was possible to draw some general

conclusions regarding common features of rack design which serve to facilitate bypass. With reference to the information in Table 2-2, it is clear that Regions B & D were the most important leakage paths overall, both before and after leakage minimisation (see Figure 2-2 for definitions of leakage path regions). In all cases it was possible to significantly reduce leakage through these regions by using well-fitted foam strips. Leakage down through the bottom or out of the sides of racks (as illustrated in Figure 2-3) was also important in cases where joins between separate folded metal pieces in the rack walls or floor were not sealed. Again, this could be minimised by the application of tape and/or foam, depending on the geometry of the holes. Where these two leakage paths had been effectively minimised, any holes in the various folded metal sections became important, but were easy to seal with tape.

2.4.3   Test Data Centre

The results shown in Figure 2-23 represent a very low level of bypass, achievable under best practice conditions, i.e. with well-designed racks, and with all empty equipment slots sealed with blanking panels. A least squares regression analysis was applied to this data using Microsoft Excel's built in fitting function. The data was found to fit well to a 2nd order polynomial, with a condition applied to set $\Delta p_{CH} = 0$ where $\dot{V}_{BP} = 0$. Separate equations were produced for $\dot{V}_{BP} \geq 0$ and $\dot{V}_{BP} < 0$, as displayed in Eq. 2-23. The coefficients of determination, R_D^2, for the 2 equations, were 0.999 and 0.996, respectively. Here, R_D^2 is given by $1 - \frac{SS_{res}}{SS_{tot}}$, where SS_{res} is the sum of the differences between each $\dot{V}_{BP}$ data point and the respective $\dot{V}_{BP}$ predicted by the equation, and SS_{tot} is the sum of the differences between each $\dot{V}_{BP}$ data point and the mean value of $\dot{V}_{BP}$ across the data set [116]. R_D^2 is a measure of agreement between the equation and the experimental data, with $R_D^2 = 1$ indicating perfect agreement.

The 2nd order equations may be taken to demonstrate a combination of laminar and turbulent behaviour for leakage flows, since, analogously, the pressure drop in a pipe is directly proportional to flow rate through it in the laminar case, and is proportional to the square of the flow rate in the turbulent case [98], [117].

$$\Delta p_{CH} = 1121\dot{V}_{BP}^2 + 28.74\dot{V}_{BP} \qquad - \text{Where } \dot{V}_{BP} \geq 0$$

$$\Delta p_{CH} = -887.7\dot{V}_{BP}^2 + 35.21\dot{V}_{BP} \qquad - \text{Where } \dot{V}_{BP} < 0$$

Eq. 2-23

It should be acknowledged that this equation does not include leakage at the interfaces between servers and racks. The results displayed in Figure 2-24 suggested that this could contribute significantly to bypass. Whilst the error bars shown in Figure 2-24 suggest that this bypass was too small to be measured within the accuracies of the instruments used, the repeatability tests reported in section 2.4.1 suggested that the difference between $\dot{V}_{BP}$ with and without the interface sealed exceeded the random variation in the repeat measurements, at least at higher pressures.

After applying the same regression analysis method to the data shown in Figure 2-24 for the cases with and without the interface between the v40z server and the rack sealed, Eq. 2-24 and Eq. 2-25 were found. These equations had R_D^2 values of 0.999 and 0.994, respectively. Eq. 2-24 and Eq. 2-25 may be rearranged to make $\dot{V}_{BP}$ the subject (by applying the quadratic formula), giving Eq. 2-26 and Eq. 2-27, respectively. Here, $\dot{V}_{BP,s}$ is the bypass with the interface sealed, and $\dot{V}_{BP,o}$ is the bypass with the interface open. Subtracting $\dot{V}_{BP,s}$ from $\dot{V}_{BP,o}$ should then give an estimate of the bypass at the interface itself, $\dot{V}_{BP,I}$. We may then imagine an operational data centre, in which best practice is applied with regards containment and leakage minimisation, with each rack occupied by a block of servers, at whose perimeter there is bypass as described by $\dot{V}_{BP,I}$. The total bypass per rack in such a data centre would be given by $\dot{V}_{BP,I} + 0.25\dot{V}_{BP,s}$ (since $\dot{V}_{BP,s}$ represents bypass in a 4-rack aisle, and assuming that one block of servers introduces the same potential for leakage as a single server in isolation). This leads to equation Eq. 2-28, which is used in the remainder of this thesis to represent a **low bypass scenario**.

$$\Delta p_{CH} = 1289\dot{V}_{BP}^2 + 56.59\dot{V}_{BP} \qquad \text{- Server-rack interface sealed} \qquad \text{Eq. 2-24}$$

$$\Delta p_{CH} = 453\dot{V}_{BP}^2 + 80.04\dot{V}_{BP} \qquad \text{- Server-rack interface open} \qquad \text{Eq. 2-25}$$

$$\dot{V}_{BP,s} = \frac{-56.59 + \sqrt{56.59^2 + 4 \times 1289 \Delta p_{CH}}}{2 \times 1289}$$

Eq. 2-26

$$\dot{V}_{BP,o} = \frac{-80.04 + \sqrt{80.04^2 + 4 \times 453 \Delta p_{CH}}}{2 \times 453}$$

Eq. 2-27

$$\dot{V}_{BP} = \frac{-80.04 + \sqrt{80.04^2 + 4 \times 453 \Delta p_{CH}}}{2 \times 453}$$

Eq. 2-28

$$+ 0.75 \times \frac{-56.59 + \sqrt{56.59^2 + 4 \times 1289 \Delta p_{CH}}}{2 \times 1289}$$

A similar analysis can be applied to the data shown in Figure 2-27 to produce an equation for bypass through an empty slot, $\dot{V}_{slot}$. Again, a regression analysis was applied to the data in Figure 2-27 for the total bypass with one empty slot, $\dot{V}_{BP+slot}$. The resulting equations, for the negative and positive case, are shown as Eq. 2-29, and have $R_D^2 = 0.995$ and 0.999, respectively. Rearranging each of the components of Eq. 2-29 to make $\dot{V}_{BP+slot}$ the subject gives Eq. 2-30. Subtracting $\dot{V}_{BP}$ as described in Eq. 2-23 from $\dot{V}_{BP+slot}$ as described in Eq. 2-30 then gives Eq. 2-31, an estimation of $\dot{V}_{slot}$.

$$\Delta p_{CH} = 442 \dot{V}_{BP+slot}^2 + 7.6 \dot{V}_{BP+slot} \qquad - \text{Where } \dot{V}_{BP} \geq 0$$

$$\Delta p_{CH} = -422 \dot{V}_{BP+slot}^2 + 5.7 \dot{V}_{BP+slot} \qquad - \text{Where } \dot{V}_{BP} < 0$$

Eq. 2-29

$$\dot{V}_{BP+slot} = \frac{-7.6 + \sqrt{7.6^2 + 4 \times 442 \Delta p_{CH}}}{2 \times 442} \qquad - \text{Where } \dot{V}_{BP} \geq 0$$

$$\dot{V}_{BP+slot} = \frac{-5.7 + \sqrt{5.7^2 - 4 \times 422 \Delta p_{CH}}}{2 \times 422} \qquad - \text{Where } \dot{V}_{BP} < 0$$

Eq. 2-30

$$\dot{V}_{slot} = \frac{-7.6 + \sqrt{7.6^2 + 4 \times 442 \Delta p_{CH}}}{2 \times (-\Delta p_{CH})} - \frac{-28.7 + \sqrt{28.7^2 + 4 \times 1121 \Delta p_{CH}}}{2 \times 1121} \qquad \dot{V}_{BP} \geq 0$$

$$\dot{V}_{slot} = \frac{-5.7 + \sqrt{5.7^2 - 4 \times 422 \Delta p_{CH}}}{2 \times 422} - \frac{-35.2 + \sqrt{35.2^2 - 4 \times 888 \Delta p_{CH}}}{2 \times 888} \qquad \dot{V}_{BP} < 0$$

Eq. 2-31

A worst case scenario for bypass within a contained environment may be established using the results for Rack A from Figure 2-20. Applying the same regression analysis to this data to determine a 2nd order relationship gives Eq. 2-32, with $R_D^2 = 0.999$. Rearranging this equation to make $\dot{V}_{BP}$ the subject, using the quadratic formula, gives Eq. 2-33. This is used in the remainder of this thesis to represent a **high bypass scenario**, since Rack A was the worst-performing rack.

$$\Delta p_{CH} = 737\dot{V}_{BP}^2 + 14.76\dot{V}_{BP}$$

Eq. 2-32

$$\dot{V}_{BP} = \frac{-14.76 + \sqrt{14.76^2 + 4 \times 737\Delta p_{CH}}}{2 \times 737}$$

Eq. 2-33

Figure 2-35 shows the bypass for the low bypass scenario (Eq. 2-28), high bypass scenario (Eq. 2-33), and bypass through an empty slot (Eq. 2-31), in each case as a percentage of the total required supply flow rate, $\dot{V}_{supply}$, in an operational data centre. Here, $\dot{V}_{supply}$ is calculated as the sum of $\dot{V}_{BP}$ and $\dot{V}_{IT}$. $\dot{V}_{IT}$ is the total flow rate through the servers, and is estimated according to conservation of energy, using Eq. 2-13, where $\dot{E}_{server}$ and $\dot{V}_{server}$ are replaced with $\dot{E}_{IT}$ (the total IT power consumption) and $\dot{V}_{IT}$, respectively. This approach assumes that 100% of server power consumption is converted to heat, and that 100% of the heat is rejected from the server via convection. Further assumptions are that $\dot{m} = \rho_{air}\dot{V}_{IT}$, $\rho_{air} = 1.2\ kg.m^{-3}$, $c_{p,air} = 1.005\ kJ.kg^{-1}.K^{-1}$, $\Delta T = 12.5\ K$ [41], [55]. $\dot{E}_{IT}$ is set to either 6 or 12 kW per rack, as indicated, which represent typical industry rack power densities [118].
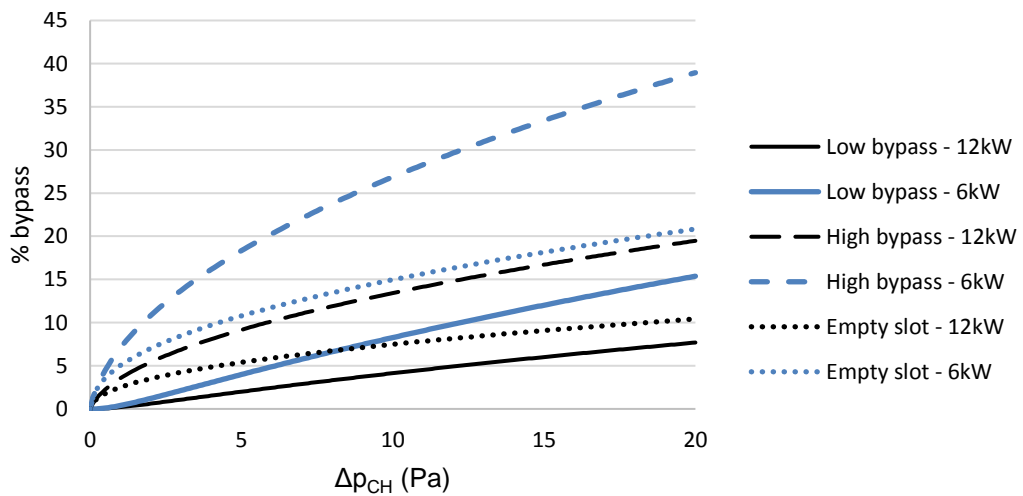
Figure 2-35. Bypass associated with 2 different levels of rack sealing, and with empty slots, as a percentage of total required supply air flow.

The results displayed in Figure 2-35 show that, even with modest levels of Δp_{CH} and high IT power densities, bypass within aisle contained data centres can be significant. For example, with the high bypass condition, with 12 kW of IT per rack, and with $\Delta p_{CH} = 5\ Pa$, around 9.2% of the supplied air bypasses the servers. With $\Delta p_{CH} = 20\ Pa$, this increases to 19.5%.

Note that $\Delta p_{CH} = 5\ Pa$ is the upper limit recommended in the EU CoC best practice guidelines [78].

However, the results also show that levels of bypass can be greatly reduced by taking practical measures to minimise the permeability of racks, by minimising Δp_{CH}, and by maximising IT power density. With low bypass, $\Delta p_{CH} = 2\ Pa$, and an IT power density of 12 kW per rack, bypass is estimated to be 0.6%. With high bypass, $\Delta p_{CH} = 20\ Pa$, and an IT power density of 6 kW per rack, this increases to 38.9%.

It is interesting to compare the results shown in Figure 2-35 with other, related results, reported in section 2.1. For example, Arghode et al. [37] reported that, with $\Delta p_{CH} = 6.2\ Pa$, 0.81 m³.s⁻¹ of air bypassed the racks. Since this was measured across a cold aisle containing 14 racks, this indicates bypass of 0.058 m³.s⁻¹ per rack. Eq. 2-23, which shows the relationship between Δp_{CH} in the Test Data Centre when well sealed, predicts $\dot{V}_{BP} = 0.0626\ m^3.s^{-1}$ at $\Delta p_{CH} = 6.2\ Pa$, or $0.0157\ m^3.s^{-1}$ per rack. Smoke pen tests suggested that the majority of this bypass was attributable to rack leakage, rather than containment leakage. Since Arghode et al.'s [37] results do not include rack leakage, this comparison shows that containment leakage may vary greatly between different configurations, and specifically that containment leakage may in some cases greatly exceed that observed in the Test Data Centre.

Comparing the results in Figure 2-35 with Salim & Tozer's [6] finding that bypass in data centres not employing aisle containment averages 50% tells us that containment is likely to reduce bypass, although effective rack sealing, low Δp_{CH}, and high IT power densities are needed to maximise the benefit.

The theoretical calculation of flow through an empty slot presented in section 2.1.4 compares remarkably well with the equivalent experimental data shown in Figure 2-35, with the former predicting 3.1 to 10.4 % bypass corresponding to $2 < \Delta p_{CH} < 20$, 12 kW of IT per rack. The data also broadly corroborates Makwana et al.'s [96] finding that introducing an empty slot to each rack in a data centre with a heat load of 14.6 kW per rack increased the required supply airflow by 9%, with Δp_{CH} not disclosed.

The results presented in Figure 2-28 show that bypass was slightly increased by the introduction of load banks to the Test Data Centre, if their fans were switched on. Importantly, a significant level of $\dot{V}_{BP}$ was detected at $\Delta p_{CH} = 0$, i.e. with no pressure gradient across the partition. Since this was unexpected, an investigation was undertaken to determine the cause of this increased bypass. During tests with $\Delta p_{CH} = 0$, with load bank

fans switched on, and with the duct used to measure the flow through the load banks in place, significant bypass was observed through the equipment rails in the rack containing the load banks (the air movement through this region could be felt by hand). Repeating this test with the load bank duct removed showed no significant levels of bypass in this region. It was concluded that the high air velocities prevalent at the outlet of the load bank duct resulted in low pressures, causing air to be entrained from the cold to the hot aisle even in the absence of a prevailing pressure differential between the hot and cold aisles. Since velocities at server outlets are likely to be significantly less than velocities at the outlet of the load bank duct, it may be assumed that bypass is unlikely to be elevated by the presence of servers in a real data centre. With lower equipment flow rates (i.e. with the v40z server in use rather than the load banks), there was no measurable effect on bypass (see Figure 2-25). The results displayed in Figure 2-28 also show that the position of the load banks within the aisle made no measurable difference to the extent of bypass. Hence, it may be concluded that the magnitudes of bypass and recirculation are governed largely by Δp_{CH} and the extent to which leakage paths are sealed, with the density of and computational loads assigned to servers having little impact.

2.4.4   Temperature measurements

The results presented in Figure 2-30 demonstrate that, where $\Delta p_{CH} < 0$ (i.e. where hot aisle pressure exceeds cold aisle pressure), significant levels of recirculation can occur within data centres employing aisle containment, even where best practice is applied to sealing leakage paths. This is consistent with the findings of Tradat et al. [93], as discussed in section 2.1.2, and demonstrates the need to maintain suitable pressure conditions. It also demonstrates the need to incorporate rack and containment leakage into CFD models of data centres employing aisle containment, particularly where there is an interest in predicting inlet temperatures with $\Delta p_{CH} < 0$. This would include modelling of conditions with CRAC/CRAH failures, during which supply flow rates could fall sufficiently to cause negative pressures in the cold aisle.

2.5   Summary

The literature review presented in section 2.1 has shown that there is evidence within the academic literature that aisle containment can improve thermal conditions, reduce bypass and recirculation, and reduce electricity consumption. However, section 2.1 also highlighted limitations in previous work undertaken to quantify the impact of aisle containment on data centre air management and energy consumption. Specifically, there has previously been

very limited discussion of the influence of the pressure differential between cold and hot aisles (Δp_{CH}) on the benefits of aisle containment, and of the potential for rack leakage.

The experimental results and analysis presented in sections 2.3 and 2.4 have shown that significant levels of bypass can occur through server racks in data centres employing aisle containment. This is the first time that such bypass has been measured in the research literature as a function of Δp_{CH}. The results show that the minimisation of Δp_{CH} can dramatically reduce rack leakage. The results also show that practical measures undertaken to seal leakage paths within racks can dramatically reduce rack leakage, and approaches to achieving this have been identified in section 2.4.2. Bypass flow rate as a percentage of total air flow supplied to the data centre was predicted to range from 0.6 to 38.9%, depending on Δp_{CH}, IT power density, and the permeability of the racks and containment structure. The potential for bypass through empty slots has also been quantified for the first time as a function of Δp_{CH}. The work presented in this chapter represents the most complete analysis of bypass and recirculation in data centres employing aisle containment published to date. In obtaining these experimental results, objective (1) has been fulfilled (as defined in section 1.5).

3. A SYSTEM MODEL INVESTIGATING POWER CONSUMPTION IN DATA CENTRES EMPLOYING AISLE CONTAINMENT

Chapter 2 has investigated the extent of bypass and recirculation in data centres employing aisle containment. The impact of aisle containment on total data centre electricity consumption, E_T, depends on the responses of the various electricity consuming components of the data centre to the changing conditioned air flow rates and temperatures caused by bypass and recirculation. The work presented in this chapter seeks to address this issue, thus fulfilling objective (2), as defined in section 1.5, which is as follows:

'To investigate the implications of bypass and recirculation for electricity consumption in data centres employing aisle containment'.

3.1 System models in data centres and other heating, ventilation and air conditioning applications

Numerous models of heating, ventilation and air conditioning (HVAC) systems in thermal comfort applications are available which predict the energy consumption of such systems under different conditions or arrangements. Nguyen et al. [119] produced an extensive review of these models, which are typically grouped into data-driven, physics-based and grey box models. Data-driven models require the availability of substantial quantities of data regarding the performance of the system under different conditions, from which mathematical descriptions of system response to the key variables can be deduced empirically. Physics-based models describe a system using scientific principles, and typically include modelling of conductive heat transfers through elements such as walls and windows, heat gains from the sun and electrical equipment, convective heat transfers, heat transfers at heating and cooling coils, control systems for coolant and chilled water flow rates, fans and pumps, boilers, chillers, heat pumps and cooling towers. Grey box models use measured data pertaining to the system's true performance to inform the parameters used in the equations underpinning a physics-based model.

A number of system models have been developed for data centres, typically with the aim of enabling the facility power consumption associated with various configurations and conditions to be predicted [41], [120]–[125]. In all cases, they fall into the category of physics-based models. The components of the models typically include computer servers, heat exchangers, computer room air handlers (CRAHs) or computer room air conditioners (CRACs), chillers, cooling towers and chilled water pumps. The various models use a

combination of manufacturer specifications, thermodynamic analysis and empirical models to define the performance of the components.

Various approaches have been used to model the heat transfers in heat exchangers in data centre system models, including fundamental analysis to determine the thermal resistances at the heat exchangers [122], heat transfer effectiveness-NTU methods [120] and empirical relationships [41]. Most of the models assume no bypass occurs. Of the 2 models that do include bypass, one uses CFD simulations to predict levels of bypass and recirculation within a data centre not employing aisle containment [124], whilst the other defines bypass as a model parameter without justification of the values used [122]. The discussions around these models fail to examine the effect of bypass on energy consumption. Since the extent of bypass varies significantly between data centres, and has a large impact on energy consumption, bypass should be considered an important parameter in any data centre model. Efforts within the industry in recent years to reduce bypass through measures such as the introduction of hot aisle-cold aisle (HACA) formation and aisle containment provide further incentive to quantify the likely benefits of bypass reduction.

No data centre system models previously presented in the research literature have considered the influence of the external static pressure drop across the server on the server's power consumption, or on flow rate through the server [41], [120]–[125].

3.2   Factors affecting server and bypass flow rates, and server power consumption

As shown by the results presented in Chapter 2, the magnitude of bypass flow, $\dot{V}_{bypass}$, is determined largely by Δp_{CH} and the extent to which leakage paths have been sealed. Determination of the proportion of the air supplied to the data centre which bypasses servers requires an estimation to be made of the flow rate through servers.

The relationships between server flow rates, server power consumptions, pressure differential between cold and hot aisles (Δp_{CH}) and server inlet temperatures (T_{inlet}) are complex. For example, increasing Δp_{CH} would tend to increase flow through the server ($\dot{V}_{server}$), reducing its core temperature. Since server fan speeds are typically controlled in response to server core temperatures, this could lead to a reduction in server fan speeds in response to lower cooling demand [95], which in turn could reduce $\dot{V}_{server}$.

Experiments reported by Brady [126] and in an industry white paper [127] have given an indication of the flow rates through idle and switched off servers subjected to external static

pressures (analogous to Δp_{CH}). The results for switched off servers are collected in Figure 3-1, with the measured flow rates having been divided by the respective servers' nameplate power consumptions. Assuming $\rho_{air} = 1.2\ kgm^{-3}$, $c_{p,air} = 1.005\ kJkg^{-1}K^{-1}$ and $\Delta T = 12.5\ K$ [41] [55], the flow rate required to cool a server per kW of server power consumption can be calculated via conservation of heat as $\frac{\dot{V}_{server}}{\dot{E}_{server}} = \frac{1}{c_{p,air}\rho_{air}\Delta T} = 0.0663\ m^3.s^{-1}.kW^{-1}$. This is plotted as a horizontal line in Figure 3-1, and shows that the majority of servers would require Δp_{CH} to exceed 30 Pa in order for sufficient air flow to occur without the action of server fans. This implies that typical values of Δp_{CH} will not drive sufficient airflow through servers, and that server fans are required to ensure adequate cooling. However, this analysis used the servers' nameplate power consumptions, which are usually higher than actual power consumption [57]. It is therefore likely that air flow rates per kW of actual power consumption would be greater than the displayed air flow rates per kW of nameplate power consumption.
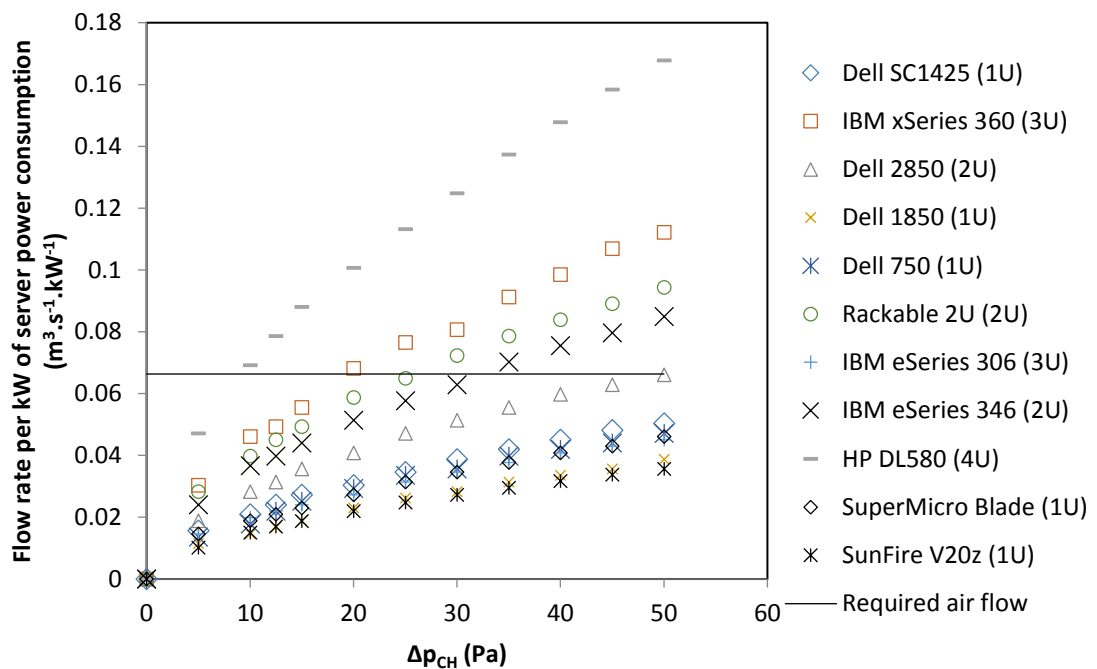
Figure 3-1: Measured flow rates per kW of nameplate power consumption through switched off servers, compared with air flow required for adequate cooling. Data from [127] except Sun Fire V20z from [126].

Brady [126] also investigated the potential to cool servers using only an external static pressure, with server fans removed. The tests showed that under some conditions this resulted in overheating of servers. Server fan speeds are generally controlled to maintain appropriate central processing unit (CPU) temperatures [50], [95]. Brady's test results demonstrate that while cold aisle pressurisation may be effective in cooling the server to

some extent, server fans may still be needed to direct air to the hottest components if failures are to be avoided.

Alissa et al. used both experimental methods and CFD simulations to investigate the flow rates through operational servers in a data centre employing aisle containment, with different levels of Δp_{CH} [83]. The results showed that flow rates through servers were strongly affected by Δp_{CH}, and that IT equipment such as switches which do not contain internal fans can act as routes for recirculation during CRAH failure events (i.e. where Δp_{CH} is negative). The curves derived experimentally by Alissa et al. [83] for flow through a number of different pieces of IT equipment are shown in Figure 3-2. Note that, in Figure 3-2, positive pressure denotes that hot aisle pressures exceed cold aisle pressures, which is the reverse of the definition of Δp_{CH} in this thesis. Note also that conditions in which hot aisle pressures exceed cold aisle pressures are generally avoided in order to minimise recirculation, but were investigated by Alissa et al. [83] in relation to instances of CRAC/CRAH fan failure, which can lead to these pressure conditions occurring. The power consumptions and identities of the servers used were not reported. The tests were undertaken with the server fans fixed at their maximum speeds, i.e. the servers' fan speed algorithms were over-ridden.

Tradat et al. [93] conducted an experimental investigation of the effect of supply air temperature on the behaviour of servers in a data centre employing aisle containment. Their results (shown in Figure 3-3) showed that increasing the supply air temperature tended to reduce the pressure in the cold aisle, indicating increasing server fan speeds, presumably in response to rising server core temperatures. The CRAH's were consistently operated at maximum blower speed, which, at $T_{supply} = 15°C$, resulted in $\Delta p_{CH} = 0\ Pa$. Δp_{CH} fell dramatically as T_{supply} exceeded 24°C, reaching $-12\ Pa$ at $T_{supply} = 29°C$. This also led to increasing server inlet temperatures, indicating recirculation. Unfortunately the methodology does not allow the individual impacts of falling Δp_{CH} and rising server inlet temperatures on server fan speeds, power consumptions and flow rates to be deduced. There was no discussion of the relative importance of rack and containment leakage in contributing to the recirculation.
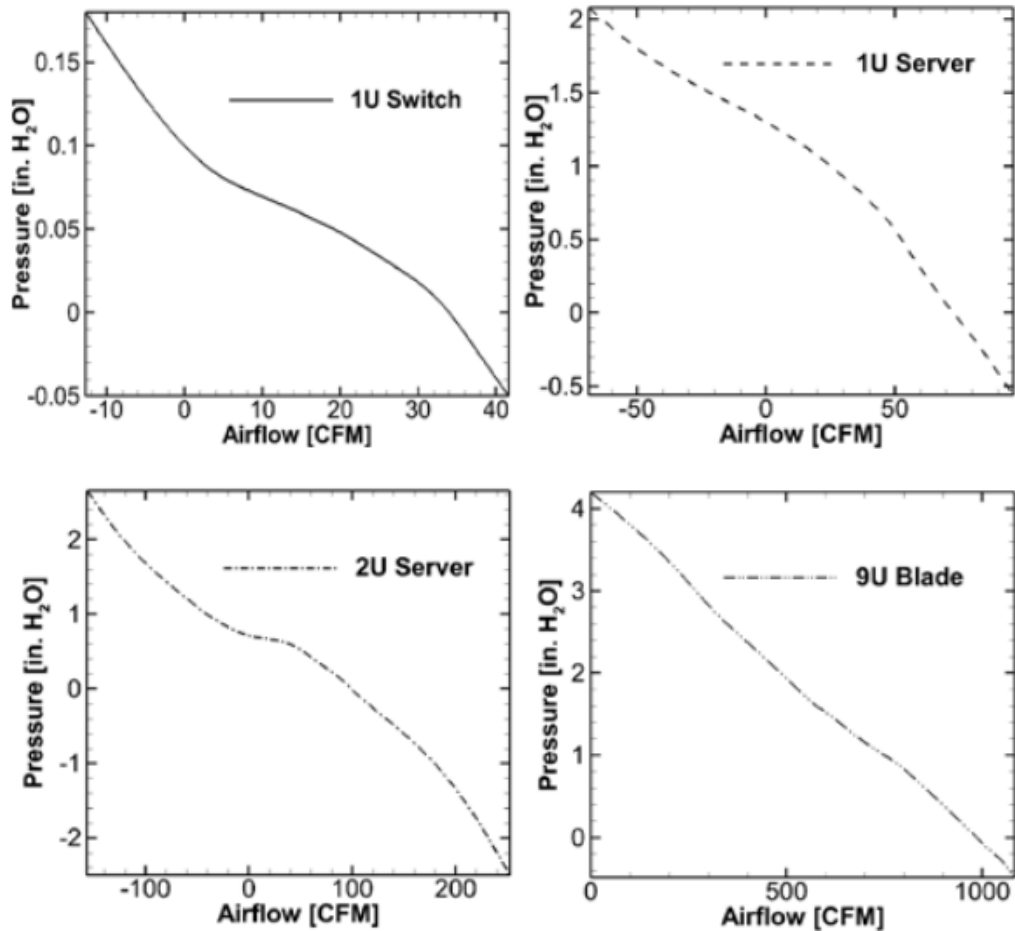
Figure 3-2. Relationship between flow rate and external static pressure for different types of IT equipment under operation, from Alissa et al. [83]. Note that positive pressure here indicates that hot aisle pressure exceeds cold aisle pressure, which is the reverse of the definition of Δp_{CH} in this thesis.

Only Tatchell-Evans et al. [128] have presented results of experiments investigating the relationship between Δp_{CH} and $\dot{V}_{server}$ for a server operating under a computational load, with the server's own fan control algorithm in operation. These results showed a slight reduction in $\dot{V}_{server}$ as Δp_{CH} increased from 5 to 10 Pa, suggesting a reduction in server fan speeds. However, there was no clear evidence of a corresponding reduction in $\dot{E}_{server}$, which would be expected for a reduction in server fan speeds. Tatchell-Evans et al. [128] noted that the apparent reduction in $\dot{V}_{server}$ could have been due to errors in measurement of $\dot{V}_{server}$, since the reduction was within the range of reported experimental errors. Additionally, only one server was studied, making it difficult to draw conclusions about server behaviour more generally.
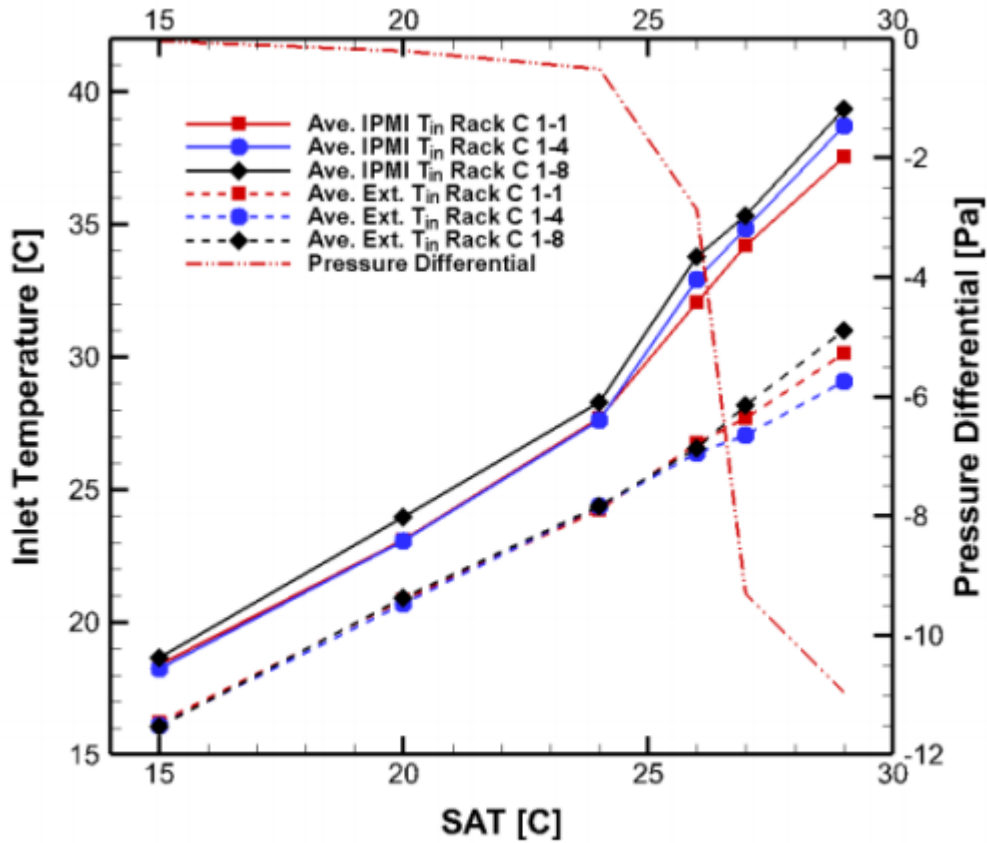
Figure 3-3. Results from Tradat et al.'s study [93]. Note here that 'Inlet Temperature' indicates server inlet temperature, 'SAT' indicates T_{supply} and 'Pressure Differential' is equal to Δp_{CH}.

$\dot{V}_{server}$ must also be affected by the rate at which heat is generated by the server ($\dot{Q}_{server}$), since this affects the air flow rate required to maintain the required internal temperatures. Server power consumptions are usually lower than nameplate values due to low CPU utilisation, with many servers housed in data centres being obsolete, outdated or unused [57]. Utilisation rates vary between data centres, with cloud computing and high performance computing usually enabling higher utilisation [58], [129], [130]. Barroso et al. [56] reported data showing the CPU utilisation distribution against time from their measurements of more than 5000 servers in use in DCs, finding that they operated at a utilisation of between 10 and 50% for more than 70% of the sampling time. The same study stated that a typical energy efficient server used around 50% of its peak power consumption at 0% CPU utilisation, with this increasing linearly to 100% at full CPU utilisation.

In addition to the utilisation rates of IT equipment, the density of IT within the facility must be a determining factor in the total IT flow rate, and must subsequently affect the relative

significance of bypass flow rates. Low IT densities correspond to low flow rates of air required for server cooling, increasing the significance of bypass air flow rates. Data centres often operate well below their design IT capacity since they are designed to accommodate future expansion [131] and to cope with peak workloads which may occur infrequently [132], [133].

Server power consumption is also affected by T_{inlet}. Increasing the temperature of air supplied to the data centre from the CRAC/CRAH, T_{supply}, can reduce electricity consumption in the cooling infrastructure, as discussed in section 1.3.3. However, these savings must be offset against the increase in IT power consumption which results from increasing T_{inlet}. This increase is caused by increased server fan speeds, and increased leakage current [50]–[53]. The overall effect of increasing T_{inlet} depends on parameters such as the overall efficiency of cooling and power infrastructure and the algorithms used to control server fans, but generally increasing supply air temperature to the higher end of the ASHRAE limits is thought to reduce E_T [51].

Within the peer-reviewed research literature, only Tatchell-Evans et al. [128] have experimentally investigated the impact of inlet temperature on server power consumption, finding no significant change in $\dot{E}_{server}$ for a Sun Fire v20z server with T_{inlet} ranging from 23 to 28°C. Conversely, a white paper from Schneider Electric [54] found that the power consumptions of 2 unidentified servers rose by 2 and 5%, respectively, with this being due to increases in fan power consumption. Considering the vast number and variety of different servers in use in data centres, this represents a very limited resource of experimental data from which to draw conclusions regarding server response to T_{inlet}.

3.3   Methods

The literature review presented in section 3.1 has demonstrated that, whilst a number of data centre system models have been presented in the research literature, none have accounted for the impact of Δp_{CH} on bypass and IT flow rates. Δp_{CH} should be seen as a key parameter in such models due to the increasing prevalence of aisle containment, and because of the strong impact of Δp_{CH} on bypass flow rates (as shown in Chapter 2) and IT flow rates (as discussed in section 3.2). This section describes a new system model, intended to address these issues. Mathworks' Matlab software was used to compute the system model [134], and the Matlab script is included in the supplementary material provided with this thesis.

Figure 3-4 shows a schematic of a physics-based system model constructed to predict the impact of the pressure differential between the cold and hot aisle (Δp_{CH}) on total data centre electricity consumption (E_T). Note that the symbols used in the figure are defined in the remainder of section 3.3. The solution process of the system model is illustrated in Figure 3-7 and Figure 3-8, as well as being fully described in the remainder of section 3.3. The model assumes a data centre employing aisle containment, with servers cooled using a closed loop of process air, which rejects heat to a chilled water loop in a CRAH. The pressure field within each aisle is assumed to be uniform, hence, for positive Δp_{CH}, there is assumed to be no recirculation. The total process air flow rate is the sum of the bypass ($\dot{V}_{BP}$) and server flow rates, whose calculation as functions of Δp_{CH} and server power consumption are described later in this section ($\dot{V}_{BP}$) and in section 3.3.1, respectively. The model was run with an ambient temperature (T_{amb}) of either 11°C (the average T_{amb} for London [135]) or 30°C. In the former case the chilled water loop used to cool the process air rejects heat to the ambient air, in an economiser consisting of a heat exchanger and fan. At $T_{amb} = 30°C$ the chilled water is cooled mechanically in a chiller whose working fluid is cooled using ambient air. Process air is supplied to the cold aisle at 25°C, which is at the higher end of the range recommended by ASHRAE (18-27°C) [49].

a)

b)
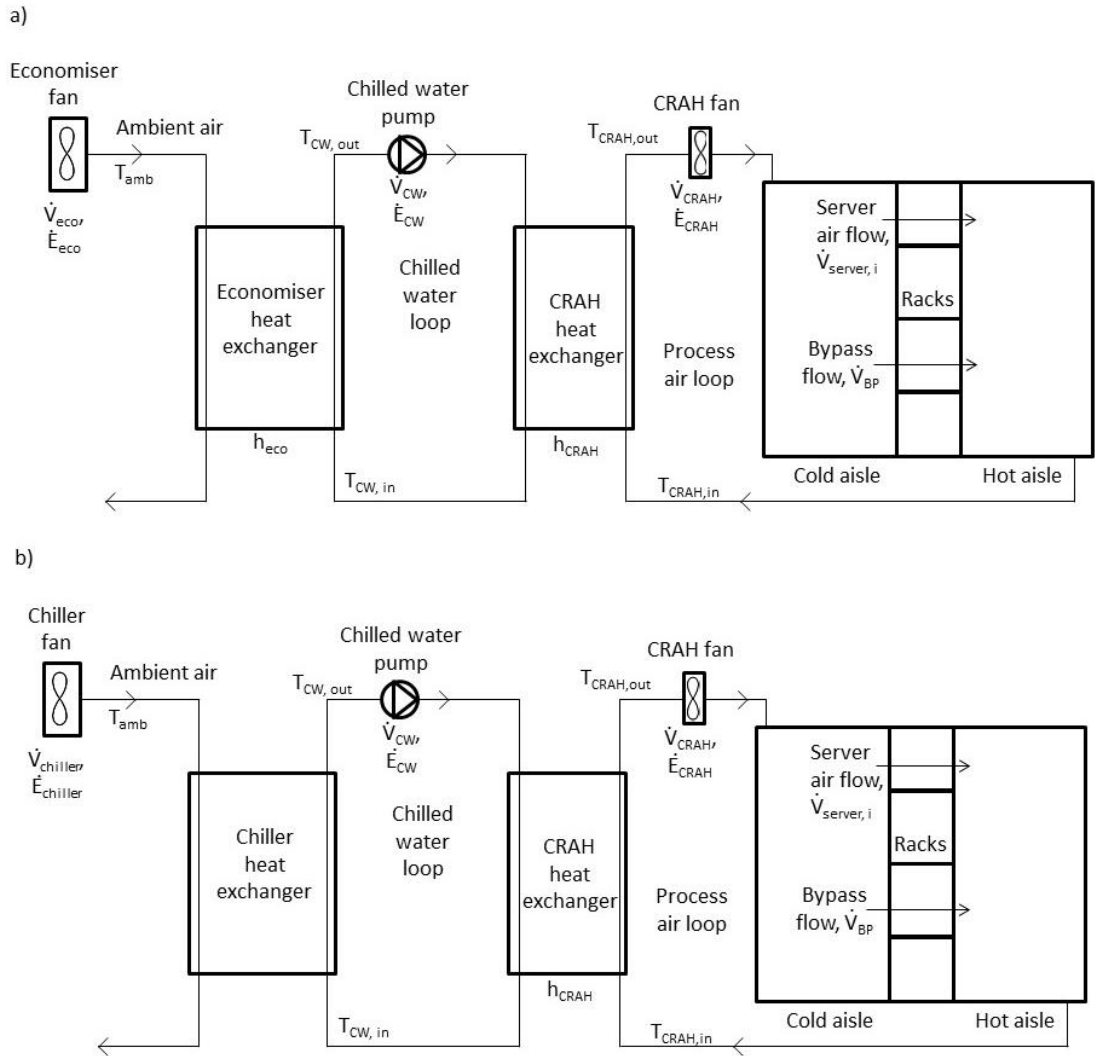
Figure 3-4: Schematic of the new system model with a) $T_{amb} = 11°C$ and b) $T_{amb} = 30°C$.

Δp_{CH}, T_{amb} and the number of servers housed are set as inputs to the model. The bypass flow rate ($\dot{V}_{BP}$) is determined using Eq. 2-28 for low bypass conditions and Eq. 2-33 for high bypass conditions.

3.3.1 Calculation of server power consumption and flow rate

The power consumption attributed to servers within the model is based on the data reported by Barroso et al. [56], discussed in Section 3.2. The servers are divided into 10 blocks, each comprising 10% of the total number of servers. The power consumption of servers within each block as a percentage of maximum power consumption is shown in Figure 3-5, with the distribution being set to match the findings of Barroso et al. Each rack is populated with sufficient servers that the total power consumption within the rack would be 12 kW, were each server to operate at its full capacity (representing a typical design rack power density [102]). The simulations were also repeated with 6 kW of IT capacity per rack.

Maximum power consumption is taken as 236 W per server, as measured by Brady [136] for a Sun Fire V20z server operating under a high computational load. This allows the baseline server power consumptions, $\dot{E}_{server,0,i}$, to be calculated for each block of servers, with each block indicated by the index i, which takes a value of 1 to 10. Hence the total baseline IT power consumption is $\sum_{i=1}^{i=10}\dot{E}_{server,0,i}$.
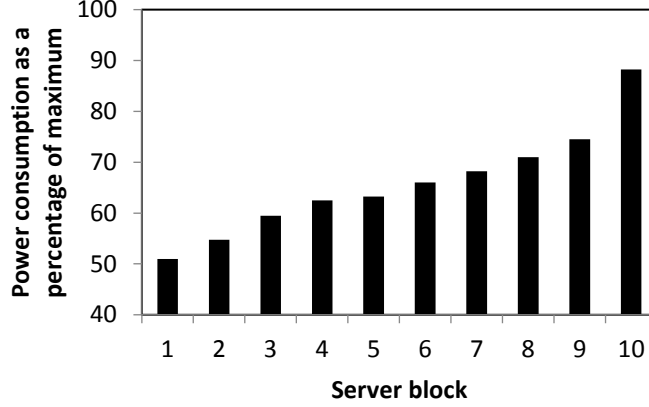
Figure 3-5: Distribution of server power consumptions, with each block accounting for 10% of the total number of servers. After Barroso et al. [56].

The volumetric flow rate of air required through each server block, $\dot{V}_{server,0,i}$, in order to absorb the baseline heat generated by the server ($\dot{E}_{server,0,i}$), is calculated using Eq. 3-1, assuming a temperature rise across the server, ΔT, of 12.5 K [41], [55]. The specific heat capacity of air, $c_{p,air}$, and air density, ρ_{air}, were held constant at $1.005\ kJ.kg^{-1}.K^{-1}$ and $1.2\ kg.m^{-3}$, respectively. N is the number of servers in the block. The static pressure required to act across the server block, $\Delta p_{req,i}$, in order to achieve a flow rate of $\dot{V}_{server,0,i}$, and the flow rate forced through the server block, $\dot{V}_{server,forced,i}$, as a result of the cold aisle pressure (Δp_{CH}), are then calculated using a system curve determined experimentally for the Sun Fire V20z server by Brady [137]. The curve is given by Eq. 3-2, where $\Delta p = \Delta p_{req,i}$ or $\Delta p = \Delta p_{CH}$, and $\dot{V}_{server,i} = \dot{V}_{server,0,i}$ or $\dot{V}_{server,i} = \dot{V}_{server,forced,i}$ as appropriate.

$$\dot{V}_{server,0,i} = \frac{\dot{E}_{server,0,i}}{\rho_{air}c_{p,air}\Delta T} \qquad \text{Eq. 3-1}$$

$$\Delta p = 38289\left(\frac{\dot{V}_{server,i}}{N}\right)^2 \qquad \text{Eq. 3-2}$$

Two options are considered for server fan speed control. With Server Fan Speed Option 1, the fan speed for each block of servers is controlled so as to produce a pressure drop, Δp_{excess}, which, when added to Δp_{CH}, will produce $\Delta p_{req,i}$. With Option 2, the server fan

speed is affected only by the server heat load, effectively assuming that the forced air flow does not take the path required for it to cool the CPU. The server fan speed in this case is controlled so as to produce $\Delta p_{req,i}$ in conditions with $\Delta p_{CH} = 0$. The two Server Fan Options are fully described in sections 3.3.1.1 and 3.3.1.2.

3.3.1.1   Server Fan Option 1

Where Option 1 is selected for server fan speed control, the actual flow rate through the server block, $\dot{V}_{server,i}$, is set equal to $\dot{V}_{server,forced,i}$. The server power consumptions assigned under the Barroso distribution must be reduced to account for the reduction in server fan power from that which would be required with $\Delta p_{CH} = 0$. This reduction in server fan power is determined using data from Brady [138], who undertook measurements to determine the relationship between flow rate and pressure developed by fans used in the Sun Fire v20z server, at a range of power consumptions. Brady's data is reproduced in Figure 3-6. The fan power consumptions required to produce $\dot{V}_{server,0,i}$ must be determined in the cases with (i) $\Delta p_{CH} = 0$, and with (ii) Δp_{CH} as specified in the model inputs. In order to achieve this, the data series showing the relationships between developed pressure and flow rate at four different fan power consumptions, shown in Figure 3-6, were imported into Matlab after visually extracting each data point from the figure. The flow rates shown in the figure were increased by a factor of 6, since this is the number of fans in each V20z server. 100% fan power consumption corresponds to 36 W per server, since each fan consumes 6 W at maximum power.

To find the power consumptions required with $\Delta p_{CH} = 0$ and with Δp_{CH} at the set level, the flow rates generated by the server fans operating at 25, 50, 75 and 100% power consumption were calculated using Eq. 3-3, for each level of Δp_{CH}. Note that Eq. 3-3 shows the equation as implemented for the 25% power condition, and that $\dot{V}_{25,2}$ is the flow rate from the data set at 25% power consumption measured at the lowest pressure at which a measurement was taken for which the pressure exceeded $\Delta p_{req,i} - \Delta p_{CH}$, and $\dot{V}_{25,1}$ is the flow rate for the measurement at the next lowest pressure from the same data set.

After carrying out similar interpolations for each data set, the percentage power consumption required to give $\dot{V}_{server,0,i}/6N$ for each level of Δp_{CH} is calculated using another interpolation. Eq. 3-4 shows an example of this interpolation, in the case where $\dot{V}_{25,0} \leq \dot{V}_{server,0,i}/6N < \dot{V}_{50,0}$ (note that $\dot{V}_{server,0,i}$ is divided by $6N$ since there are N servers per block, and 6 fans per server). The server fan power consumption required at $\Delta p_{CH} = 0$ can thus be calculated by multiplying the resulting percentage by 36 W. The difference between calculated server fan power consumptions for each level of Δp_{CH} was

multiplied by N, then subtracted from $\dot{E}_{server,0,i}$ to give $\dot{E}_{server,i}$ (i.e. the actual power consumption of this server block). $\dot{V}_{server,i}$ was then set equal to $\dot{V}_{server,0,i}$. Where $\Delta p_{CH} > \Delta p_{req,i}$, the server fan power consumption was assumed to be zero, and $\dot{V}_{server,i}$ was set using Eq. 3-2, after replacing Δp with Δp_{CH}.
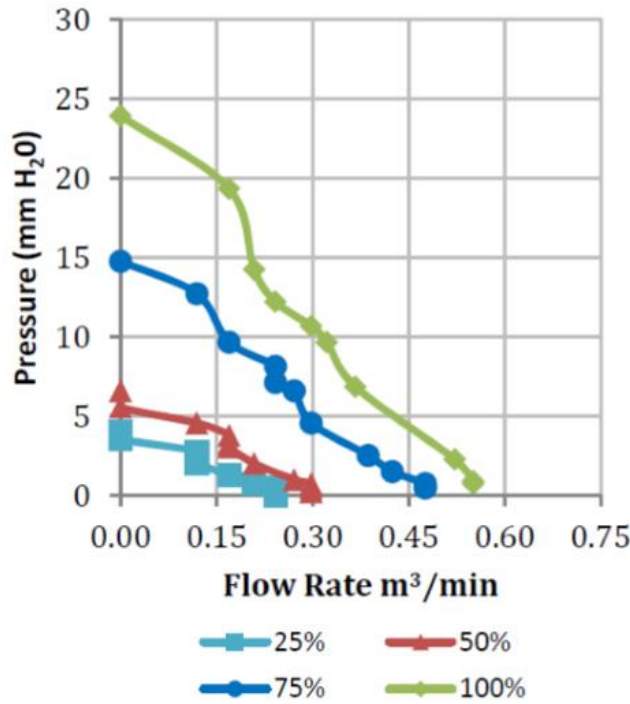
Figure 3-6: Server fan curves for fans housed in the Sun Fire v20z at 25, 50, 75 and 100% of total fan power consumption, from Brady [138].

$$\dot{V}_{25,0} = \dot{V}_{25,1} + \frac{(\Delta p_{req,i} - \Delta p_{CH}) - p_{25,1}}{p_{25,2} - p_{25,1}}(\dot{V}_{25,2} - \dot{V}_{25,1}) \qquad \text{Eq. 3-3}$$

$$Percentage = 25 + \frac{\dfrac{\dot{V}_{server,0,i}}{6N} - \dot{V}_{25,0}}{\dot{V}_{50,0} - \dot{V}_{25,0}}(50 - 25) \qquad \text{Eq. 3-4}$$

3.3.1.2   Server Fan Option 2

For Server Fan Option 2, $\dot{E}_{server,i} = \dot{E}_{server,0,i}$, since server fan speed is assumed to not be affected by Δp_{CH}. $\dot{V}_{server,i}$ is then set using Eq. 3-2, with Δp set equal to $\Delta p_{req} + \Delta p_{CH}$.

3.3.2   Modelling the cooling infrastructure

In estimating the power consumptions of the chilled water pumps and CRAH and economiser fans ($\dot{E}_{CW}$, $\dot{E}_{CRAH}$ and $\dot{E}_{eco}$, respectively), reference power consumptions ($\dot{E}_{CW,ref}$,

$\dot{E}_{CRAH,ref}$ and $\dot{E}_{eco,ref}$, respectively) are first calculated for reference chilled water, process air and economiser air flow rates ($\dot{V}_{CW,ref}$, $\dot{V}_{CRAH,ref}$ and $\dot{V}_{eco,ref}$, respectively).

$\dot{V}_{CRAH,ref}$ is set according to an assumption that the CRAH is sized to provide the flow rate required to cool all of the servers with a total power consumption of $\sum_{i=1}^{i=10} \dot{E}_{server,0,i}$, with $\Delta T = 12.5\ K$, increased by a factor of 1.2 in order to account for an assumed 20% bypass. Hence, $\dot{V}_{CRAH,ref}$ may be calculated using Eq. 3-5. $\dot{V}_{eco,ref}$ is set to double $\dot{V}_{CRAH,ref}$.

$\dot{E}_{CRAH,ref}$ is then calculated with reference to the manufacturer specification data for Schneider Electric's Uniflair TD/UCV 0700 ACU [139]. This is a CRAH designed for use in data centres, and consumes 2.2 kW with a flow rate of 1.66 m³.s⁻¹. It is assumed that the CRAH used in the data centre being modelled has the same efficiency as this unit at the reference flow rate. Hence, $\dot{E}_{CRAH,ref}$ is given by Eq. 3-6. The economiser is assumed to have the same efficiency, hence $\dot{E}_{eco,ref}$ is also given by Eq. 3-6.

The actual flow rate required of the CRAH is given by Eq. 3-7, where $\sum_{i=1}^{i=10} \dot{V}_{server,,i}$ is the sum of the flow rates through the server blocks. The power consumptions of the CRAH and economiser fans, $\dot{E}_{CRAH/eco}$, at different flow rates, are assumed to scale with the cube of their respective flow rates [41], such that they are given by Eq. 3-8. The heat load on the CRAH ($\dot{Q}_{CRAH}$) is calculated using Eq. 3-9. Here, $\dot{E}_{server\ fans,i}$ and $\dot{Q}_{server\ fans,i}$ are the power consumptions and heat dissipation rates of the fans in each server block, respectively. $\dot{Q}_{server\ fans,i}$ is calculated by subtracting the rate of air flow kinetic energy generated by the server fans from $\dot{E}_{server\ fans,i}$. The rate of air flow kinetic energy is given by the product of the pressure and flow rate developed by the fans at $\dot{E}_{server\ fans,i}$. The pressure in this calculation is given by $\Delta p_{req,i} - \Delta p_{CH}$ for Server Fan Option 1, and by $\Delta p_{req,i}$ for Server Fan Option 2.

The CRAH return temperature, $T_{CRAH,in}$, may then be calculated via conservation of energy, as shown in Eq. 3-10.

$$\dot{V}_{CRAH,ref} = \frac{1.2 \sum_{i=1}^{i=10} \dot{E}_{server,0,i}}{\rho_{air} c_{p,air} \Delta T} \qquad \text{Eq. 3-5}$$

$$\dot{E}_{CRAH/eco,ref} = 2.2 \times \frac{\dot{V}_{CRAH/eco,ref}}{1.66} \qquad \text{Eq. 3-6}$$

$$\dot{V}_{CRAH} = \sum_{i=1}^{i=10} \dot{V}_{server,i} + \dot{V}_{BP} \qquad \text{Eq. 3-7}$$

$$\dot{E}_{CRAH/eco} = \dot{E}_{CRAH/eco,ref} \left(\frac{\dot{V}_{CRAH/eco}}{\dot{V}_{CRAH/eco,ref}} \right)^3 \qquad \text{Eq. 3-8}$$

$$\dot{Q}_{CRAH} = \sum_{i=1}^{i=10} (\dot{E}_{server,i} - \dot{E}_{server\ fans,i} + \dot{Q}_{server\ fans,i}) \qquad \text{Eq. 3-9}$$

$$T_{CRAH,in} = T_{CRAH,out} + \frac{\dot{Q}_{CRAH}}{\dot{V}_{CRAH}\rho_{air}c_{p,air}} \qquad \text{Eq. 3-10}$$

For the chilled water pump, a reference pressure drop, $\Delta p_{CW} = 60910\ Pa$, is again taken from the Uniflair TD/UCV 0700 specifications [139]. This pressure drop is assumed to act in the heat exchangers at the CRAH and at the economiser/chiller, with the two being summed to give the total pressure drop for the chilled water loop, $121,820\ Pa$. The manufacturer specifications cite that $\Delta p_{CW} = 60910\ Pa$ at $\dot{V}_{CW} = 0.0016\ m^3.s^{-1}$, and that the pump provides 31.2 kW of cooling at this flow rate. It is assumed that the pump used in the data centre being modelled is sized to meet an anticipated heat load of $1.1 \sum_{i=1}^{i=10} \dot{E}_{server,0,i}$, and that the reference flow rate scales linearly with heat load. $\dot{V}_{CW,ref}$ is therefore given by Eq. 3-11. The reference pump power consumption is given by Eq. 3-12. The efficiency of the chilled water pump, η_{CW}, is taken to be 0.9, as is reported to be typical for such pumps by Salim & Tozer [6]. The power consumption of the chilled water pump at different flow rates is assumed to scale with the cube of the flow rate [41], and is thus given by Eq. 3-13.

$$\dot{V}_{CW,ref} = 0.0016 \times \frac{1.1 \sum_{i=1}^{i=10} \dot{E}_{server,0,i}}{31.2} \qquad \text{Eq. 3-11}$$

$$\dot{E}_{CW,ref} = 2\eta_{CW}\Delta p_{CW}\dot{V}_{CW,ref} \qquad \text{Eq. 3-12}$$

$$\dot{E}_{CW} = \dot{E}_{CW,ref} \left(\frac{\dot{V}_{CW}}{\dot{V}_{CW,ref}} \right)^3 \qquad \text{Eq. 3-13}$$

The heat dissipation rates of the CRAH fans and the chilled water pumps, $\dot{Q}_{fan/pump}$, are calculated using Eq. 3-14. The efficiency, η, is in both cases taken to be 0.9, which is again consistent with efficiencies reported by Salim & Tozer [6].

$$\dot{Q}_{fan/pump} = E_{CW/CRAH}(1 - \eta) \qquad \text{Eq. 3-14}$$

A heat transfer rate at the CRAH heat exchanger, $\dot{Q}_{spec}$, and associated air and water flow rates, $\dot{V}_{spec,air}$, and $\dot{V}_{spec,CW}$, and temperatures, are obtained from manufacturer data relating to the Uniflair TD/UCV 0700 ACU [139]. This allows a reference heat transfer coefficient, $h_{CRAH,ref}$, to be calculated, using Eq. 3-15 and Eq. 3-16 [140]. Here, $LMTD$ refers to the logarithmic mean temperature difference across the heat exchanger. As with

$\dot{V}_{CRAH,ref}$, $h_{CRAH,ref}$ is assumed to scale linearly with power consumption, and by extension with $\dot{V}_{CRAH,ref}$ itself. In Eq. 3-16, the temperatures, T, are identified by the subscripts air and CW for air and chilled water, respectively, and in and out, for the temperature at entrance to and exit from the CRAH heat exchanger, respectively. The reference heat transfer rate of the economiser is calculated in the same way.

$$h_{CRAH/eco,ref} = \frac{\dot{V}_{CRAH/eco,ref}}{1.66} \times \dot{Q}_{spec}/(LMTD_{spec}) \qquad \text{Eq. 3-15}$$

$$LMTD = \frac{\left(T_{air,in} - T_{CW,in}\right) - \left(T_{air,out} - T_{CW,out}\right)}{\ln\left[\dfrac{T_{air,in} - T_{CW,in}}{T_{air,out} - T_{CW,out}}\right]} \qquad \text{Eq. 3-16}$$

Heat transfer coefficients in the CRAH and economiser are modelled using Eq. 3-17. This is derived from experimentally verified relationships whose accuracy has been confirmed by numerous authors for gas to liquid fin and tube heat exchangers (commonly used in CRAHs). For the gas side flow rate, n varies from 0.581-0.681, depending on the heat exchanger design [141], [142] (0.631 was used as the default value). For heat transfers from liquids flowing in a pipe, h has been shown to be proportional to the liquid flow rate raised to the power 0.8 [143]. This can be applied to the liquid side of gas to liquid fin and tube heat exchangers.

$$h_{CRAH/eco} = h_{CRAH/eco,ref}\left(\dot{V}_{CW}/\dot{V}_{CW,ref}\right)^{0.8}\left(\dot{V}_{CRAH/eco}/\dot{V}_{CRAH/eco,ref}\right)^{n} \qquad \text{Eq. 3-17}$$

The chilled water flow rate is controlled so as to minimise E_T, which is calculated according to $E_T = E_{servers} + E_{CRAH} + E_{CW} + E_{eco/chiller}$ (where $E_{eco/chiller}$ is the power consumption of the economiser or chiller). This is achieved by calculating power consumptions of each component at a range of temperature rises across the chilled water loop, ΔT_{CW}, from 1 to 10 K, at 1 K intervals. The temperature rise and corresponding flow rate which gives the lowest E_T is then selected. Having selected ΔT_{CW} for a given iteration, $\dot{V}_{CW}$ may be determined via conservation of energy, as in Eq. 3-18. Here, ρ_{water} and $c_{p,water}$ are the density and specific heat capacity of water (taken as 1,000 $kg.m^{-3}$ and 4.18 $kJ.kg^{-1}.K^{-1}$, respectively). $\dot{E}_{CW}$ may then be calculated using Eq. 3-13. The heat load on the economiser or chiller, $\dot{Q}_{eco/chiller}$, is the sum of $\dot{Q}_{CRAH}$ and the heat generated rate of the CRAH fans ($\dot{Q}_{CRAH\ fans}$) and chilled water pumps ($\dot{Q}_{CW\ pumps}$), which are calculated using Eq. 3-14. Hence, $\dot{Q}_{eco/chiller}$ is given by Eq. 3-19. The required LMTD at the CRAH, $LMTD_{CRAH,req}$, must then be calculated using Eq. 3-20. Numerical methods can then be used to find the value of $T_{CW,out}$ which satisfies Eq. 3-16. In this case, Matlab's 'vpasolve' function is used to solve the equation.

$$\dot{V}_{CW} = \frac{\dot{Q}_{CRAH}}{\rho_{water} c_{p,water} \Delta T_{CW}} \qquad \text{Eq. 3-18}$$

$$\dot{Q}_{eco/chiller} = \dot{Q}_{CRAH} + \dot{Q}_{CRAH\,fans} + \dot{Q}_{CW\,pumps} \qquad \text{Eq. 3-19}$$

$$LMTD_{CRAH,req} = \dot{Q}_{CRAH}/h_{CRAH} \qquad \text{Eq. 3-20}$$

For $T_{amb} = 11°C$ (i.e. where the economiser is in use), numerical methods are then used again to find a value of $\dot{V}_{eco}$ which satisfies both Eq. 3-16 and Eq. 3-17. $\dot{E}_{eco}$ may then be calculated using Eq. 3-8, allowing E_T to be calculated.

For $T_{amb} = 30°C$ (i.e. where the chiller is in use), manufacturer data for Airedale's DeltaChill DCF046DR-07DXY0 unit is used to determine the power input (for the compressor and fans) required to achieve the required cooling, which varies with T_{amb} and with the difference between the chilled water supply temperature and T_{amb} [144]. This chiller is designed for use in data centres, and has a refrigerant loop designed to be cooled using ambient air. The manufacturer data shows how the input power ($\dot{E}_{chiller}$) and heat transfer rate ($\dot{Q}_{chiller}$) vary with chilled water supply temperature ($T_{CW,out}$), at $T_{amb} = 30°C$. The relationship between $\frac{\dot{Q}_{chiller}}{\dot{E}_{chiller}}$ and $(T_{amb} - T_{CW,out})$ was found to approximate a straight line, which was identified using a least squares regression analysis. The equation defining the straight line is given by Eq. 3-21, and can be used to find $\dot{E}_{chiller}$.

$$\frac{\dot{Q}_{chiller}}{\dot{E}_{chiller}} = 4.40 - 0.0587(T_{amb} - T_{CW,out}) \qquad \text{Eq. 3-21}$$

Having determined $\dot{E}_{chiller}$ or $\dot{E}_{eco}$ as appropriate, for each value of ΔT_{CW}, the value of ΔT_{CW} for which E_T is minimised may be selected.

It should be noted that the total power consumption calculation neglects power consumption in the power distribution systems, which, as discussed in section 1.2.3, can account for up to 10% of a data centre's total power consumption, although it is usually less with modern systems.

The solution process is summarised in Figure 3-7 and Figure 3-8.
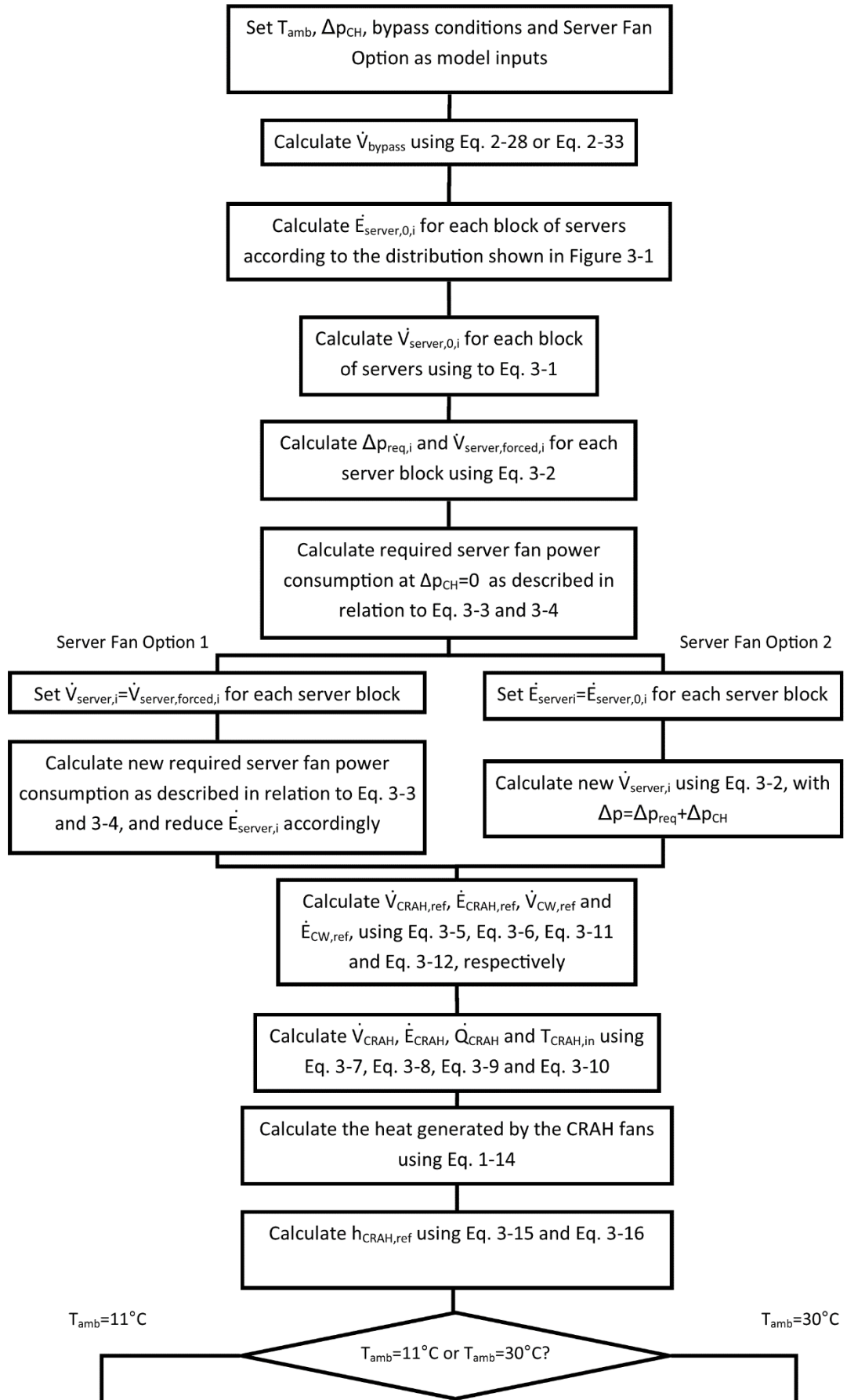
Figure 3-7. System model solution procedure, part 1.
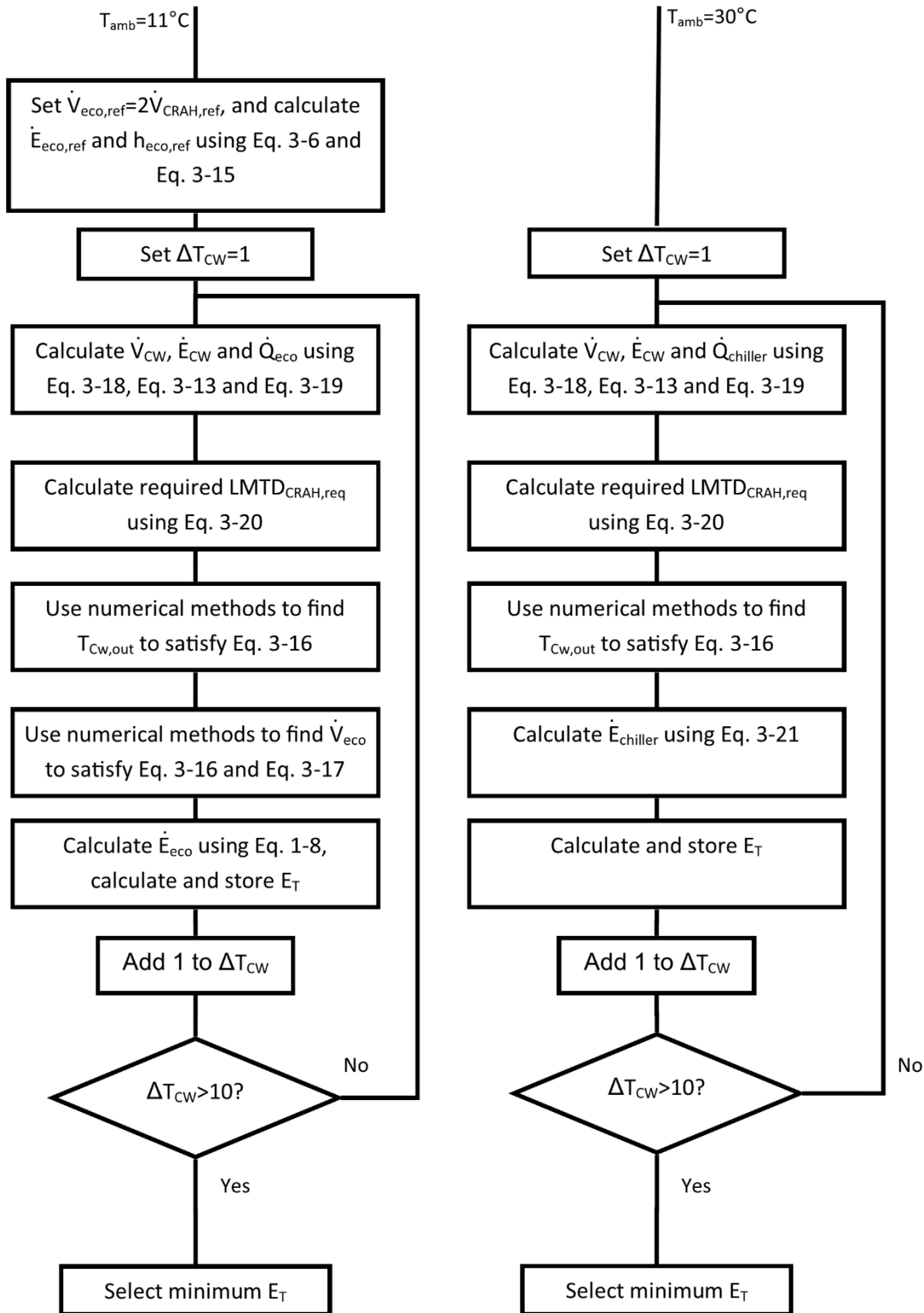
Figure 3-8. System model solution procedure, part 2.

3.4 Results

Figure 3-9 shows the results from the system model, with the effect of the pressure differential between the cold and hot aisles (Δp_{CH}) on total data centre electricity consumption (E_T) being investigated with different combinations of Server Fan Options, IT

power densities, bypass conditions and T_{amb}. For each pair of simulations with the same IT power density, T_{amb} and Server Fan Option, the power consumption at $\Delta p_{CH} = 2$ Pa with low bypass is set at unity, with the remaining results being shown proportional to this figure. Each series broadly shows E_T rising with Δp_{CH}, although the extent of the increase is strongly affected by the other variables. There are two instances in which the opposite trend is observed. With Server Fan Option 1, 12 kW per rack, low bypass, $T_{amb} = 11°C$, E_T falls by 0.1% as Δp_{CH} increases from 2 to 5 Pa, and by 0.3% over the same range with Server Fan Option 1, 12 kW per rack, high bypass and $T_{amb} = 30°C$. This small reduction is due to the reduction in server fan power consumption outweighing the corresponding increase in $\dot{E}_{CRAH}$ over this range. Note that the two server fan options are described in sections 3.3.1.1 and 3.3.1.2.

The increases in E_T shown in Figure 3-9 are dominated by increasing CRAH electricity consumption ($\dot{E}_{CRAH}$) in all cases. Since $\dot{E}_{CRAH}$ is sensitive to the value of the reference CRAH flow rate ($\dot{V}_{CRAH,ref}$), a further set of simulations were run with $\dot{V}_{CRAH,ref}$ increased by a factor of 2. This change simulates an increase in the size of the CRAH. The results of these simulations are shown in Figure 3-10. Whilst the trends are broadly similar, with increasing Δp_{CH} tending to lead to increased E_T, the extents of the increases in E_T are much smaller. Also, an increased number of series show slight reductions in E_T over the range of $2 \leq \Delta p_{CH} \leq 5\ Pa$, where Server Fan Option 1 is used. The largest reduction is in the case of the series with Server Fan Option 1, 12 kW of IT capacity per rack, low bypass and $T_{amb} = 11°C$, where E_T falls by 1.5% over this pressure range.

Figure 3-9. Variation of total power consumption, E_T, with a) Server Fan Option 1, T_{amb}=11°C, b) Server Fan Option 2, T_{amb}=11°C, c) Server Fan Option 1, T_{amb}=30°C, and d) Server Fan Option 2, T_{amb}=30°C, for simulations with $\dot{V}_{CRAH,ref}$ set at its original level.

93

○ High bypass, 12 kW/rack; ◇ Low bypass, 12 kW/rack;

✕ High bypass, 6 kW/rack; ✛ Low bypass, 6 kW/rack.
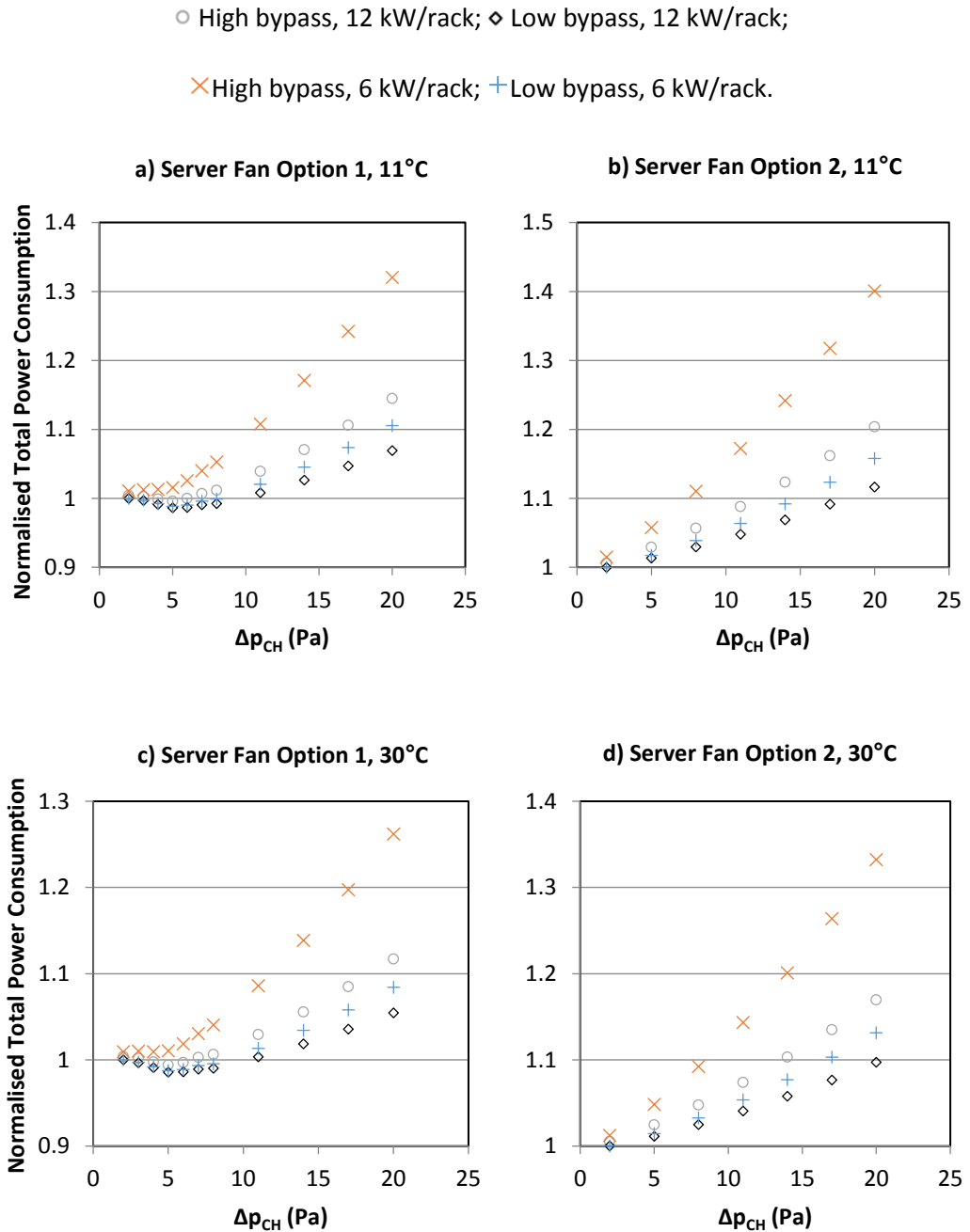
Figure 3-10. Variation of total power consumption, E_T, with a) Server Fan Option 1, T_{amb}=11°C, b) Server Fan Option 2, T_{amb}=11°C, c) Server Fan Option 1, T_{amb}=30°C, and d) Server Fan Option 2, T_{amb}=30°C, for simulations with $\dot{V}_{CRAH,ref}$ doubled from its original level.

3.5   Discussion

Figure 3-9 and Figure 3-10 show that bypass through racks in contained aisles can significantly increase E_T. This is primarily due to an increase in $\dot{E}_{CRAH}$. The reduction in E_T

achievable by minimising rack leakage depends strongly on Δp_{CH}, the response of server fans to Δp_{CH}, IT power density and CRAH sizing, ranging from 0.4% (with Server Fan Option 1, 12 kW of IT per rack, $T_{amb} = 30°C$, $\Delta p_{CH} = 2\ Pa$, and after doubling $\dot{V}_{CRAH,ref}$) to 36.5% (with Server Fan Option 2, 6 kW of IT per rack, $T_{amb} = 11°C$, $\Delta p_{CH} = 20\ Pa$, and with the original $\dot{V}_{CRAH,ref}$).

The results also indicate the potential for control of Δp_{CH} to minimise E_T. The increase in E_T which occurs when increasing Δp_{CH} from 2 to 20 Pa ranges from 5.4% (with Server Fan Option 1, 12 kW of IT per rack, $T_{amb} = 30°C$, low bypass conditions and after doubling $\dot{V}_{CRAH,ref}$) to 138% (with Server Fan Option 2, 6 kW of IT per rack, $T_{amb} = 11°C$, high bypass conditions and with the original $\dot{V}_{CRAH,ref}$). The results displayed in Figure 3-9 show that the optimum Δp_{CH} depends strongly on how server fan speeds respond to Δp_{CH}. This is because the increase in $\dot{E}_{CRAH}$ resulting from increasing Δp_{CH} can be offset by the potential reduction in $\dot{E}_{server\ fans}$. The results of Brady's experiments [126], discussed in Section 3.2, suggested that air driven through a server by aisle pressure may not cool the server effectively, hence server fans may still be required to maintain the desired CPU temperature. More experimental work is required to determine the exact effect of Δp_{CH} on server fan speeds, but at present it may be assumed that the behaviour falls somewhere between Options 1 and 2 as implemented in the described model. The results displayed in Figure 3-9 show that bypass and Δp_{CH} have a stronger impact on E_T when the data centre is operating at a lower IT load. This is because the bypass percentage increases under these conditions due to falling server air flow rates, increasing the importance of $\dot{E}_{CRAH}$.

Comparing the results shown in Figure 3-9 with those shown in Figure 3-10, it is clear that the size of the cooling infrastructure strongly affects the response of E_T to Δp_{CH}. Specifically, increases in E_T with Δp_{CH} are reduced when the size of the CRAH is increased. With Server Fan Option 2 selected, it is still the case that E_T consistently increases with Δp_{CH}, albeit at a lesser rate than was found with the smaller CRAH (as shown in Figure 3-9). However, with Server Fan Option 1 selected, E_T falls as Δp_{CH} increases from 2 to 5 Pa (by up to 1.5%), before beginning to rise.

It is useful at this point to reflect on the likely levels of Δp_{CH} prevalent in operational data centres. As noted in section 2.1.1, there are indications that Δp_{CH} may range from 2 to 20 Pa [73]–[77]. The results clearly show that maintaining Δp_{CH} at the higher end of this range could significantly increase E_T. The EU Code of Conduct on Data Centre Energy Efficiency's best practice recommendation to maintain $\Delta p_{CH} < 5\ Pa$ should keep E_T fairly close to its

minimum, although lower pressures would be beneficial if server fan behaviour replicated Server Fan Option 2.

Since the results are dominated by CRAH and server power consumptions, the impact of T_{amb} may be seen to be very small. Tests were also undertaken using the model to investigate the effect of the value of n used in the heat exchanger modelling (see Eq. 3-17), with the finding that this had no significant impact on the results. Similarly, the model was modified so that the chilled water flow rate or temperature drop could be held constant, rather than being controlled to minimise E_T. Again, this was found not to have a significant impact on the results.

3.6 Summary

The literature review presented in section 3.1 has shown that the data centre system models presented previously in the research literature have either not included bypass, or have not accounted for the impacts of factors such as the pressure differential between cold and hot aisles (Δp_{CH}) and IT power density on bypass percentages. The system model described in section 3.3 represents the most complete such model described to date, and is the first to investigate the effect of bypass and Δp_{CH} on the total data centre power consumption, E_T. Thus, objective (2) has been fulfilled (as defined in section 1.5).

The simulation results have shown that E_T can be strongly affected by Δp_{CH} and by the efforts made to minimise rack leakage. The exact trends seen are dependent upon the assumed response of server fans to changes in Δp_{CH}, and on the sizing of the cooling infrastructure and IT power density. The former issue highlights the need for greater understanding of server behaviour under pressurised conditions, particularly as aisle containment becomes more prevalent. The impact of IT power density and sizing of cooling infrastructure highlights that the impacts of Δp_{CH} on E_T will vary from facility to facility. However, the results clearly show that Δp_{CH} should generally be minimised to reduce E_T, and that IT power density should be maximised. In light of these findings, the following recommendations can be made to data centre operators seeking to minimise electricity consumption in data centres employing aisle containment:

- Leakage paths within server racks should be sealed as far as possible, using the guidance presented in section 2.4.2.
- The pressure differential between cold and hot aisles, Δp_{CH}, should not exceed 5 Pa.

4. DATA CENTRE AIR FLOW MODELLING

Chapter 3 has investigated the impact of bypass on electricity consumption in data centres employing aisle containment, using a system model. No air flow modelling was undertaken, hence it was necessary to assume that each aisle had uniform temperature and pressure conditions. In reality, pressure will vary within the aisles due to variation in air speeds, raising the possibility of variation in leakage and server flow rates across the length of an aisle, and of recirculation rather than bypass where local pressure conditions permit. In addition, CRAC/CRAH failures can lead to negative cold aisle pressures, causing widespread recirculation. Investigating these issues requires consideration of the fluid dynamics of the air flows within the data centre. This chapter presents a literature review of air flow modelling in data centres. This forms part of the process of fulfilling objective (3), which was identified in section 1.5 as follows:

'To investigate the potential for using CFD models to aid in the efficient design and management of data centres employing aisle containment.'

4.1 Background

Data centres are complicated thermodynamic environments, often housing hundreds of pieces of equipment containing heat sources and fans [145]. This equipment operates across several length scales, ranging from ~mms at the chip level to ~10s of metres at the room level [1]. In addition, the spatial distribution of the equipment contained in a data centre is constantly changing, with IT equipment being upgraded as often as every 2 years [36], and typically around 10% of IT equipment being replaced each month [31]. Hence, it is difficult for data centre managers to predict the required capacity of cooling and power distribution systems, and the optimum operating conditions of CRACs/CRAHs, both in the initial design and in response to ongoing changes. The response is often to err on the side of caution via overprovisioning, at the expense of energy efficiency [31], as discussed in section 1.2.2.

The development of computational models of proposed data centres during the design stage (or in the planning of retrofits) can be useful in improving the accuracy of estimates of the required cooling capacity and power supply [36]. Computational models can be used to predict temperature, pressure and velocity distributions in data centres whilst avoiding, or at least minimizing, intrusive measurement activities. This can enable cooling efficiency and the presence of hot spots to be predicted for both existing and proposed data centre projects, and for changes to existing data centres [36]. Thus, such models may help with

design optimization. Hamann and Lopez [146] have summarised the three main types of models used in relation to data centres as Navier-Stokes computational fluid dynamics (NS-CFD) models, simplified physics models and statistical models.

Statistical models do not use any explicit fluid modelling, rather relying on measurement of parameters such as temperature, pressure and velocity in an existing data centre under a range of conditions [146]. The data collected is then analysed and mathematical models developed empirically to describe how those parameters vary with changes in the conditions investigated [147]. Their accuracy can vary greatly [146]. Typically they contain tens or hundreds of degrees of freedom, as opposed to tens or hundreds of thousands for NS-CFD models [146]. Statistical methods can give very accurate results, very quickly, if sufficient detail is used with regards input data. However, such models are very sensitive to any changes in the data centre [146]. Hence they are primarily useful for monitoring data centre operation, i.e. to identify from real-time sensor data whether any equipment is in danger of failure due to over-heating. It should also be noted that taking sufficient temperature measurements to generate the large data sets needed to represent the cooling regime in a data centre is made challenging by the facility's need to maintain constant operation [45].

NS-CFD models determine the distribution of pressures, temperatures and air velocities by numerically solving the Navier-Stokes equations for fluid flow, alongside heat transfer equations [146] (described in section 4.3). Airflows within data centres tend to be complex, with numerous inlets and outlets, and some turbulence and buoyancy [145]. Hence, the resulting models are very computationally expensive, and time consuming to set up [145].

Simplified physics models (also referred to as "compact" or "reduced order" models) can sometimes give results of acceptable accuracy with less computation, allowing the response of the system to variation in design parameters of interest to be investigated rapidly [36] [145]. Such models operate on a similar basis to NS-CFD models, except that some simplifications are made in the physics of the fluid interactions, for example by neglecting viscosity or compressibility.

The generic approach to physics-based CFD models of fluid flows is to divide the domain of interest into discrete elements, and to apply mechanical principles to each element in turn in order to find a solution for the whole domain [148]. The dimensions of the elements must be small enough in relation to the length scales relevant to the variables of interest that variation across an element is approximately linear. This enables the discretisation of the differential equations which define the flow. A given CFD modelling project generally consists of three distinct stages [149]:

- **Pre-processing**

  This involves defining the geometry of the problem, dividing that geometry into discrete elements to create a 'mesh', determining the relevant physical properties, and setting appropriate boundary conditions representing solid walls, sources/sinks of heat/matter etc. [149].

- **Numerical solution**

  Here numerical calculations are carried out to seek a solution which satisfies the governing equations and boundary conditions. Prior to calculation, an initial guess must be made at the values of the variables of interest across the domain, and decisions must be made about the degrees of convergence and stability which are required in the final solution [149].

- **Post-processing**

  This stage describes the process of converting the numerical results into a format which is easy to understand and interpret. The most common methods are contour plots displaying a temperature or pressure field across a plane using a colour scale, and vector plots showing velocities across a plane using arrows to denote the magnitude and direction [149].

4.2   The Potential Flow Model

One example of a simplified physics model which has been used in data centre modelling is the potential flow CFD (PF-CFD) model [145], [146], [150]–[152]. This model is based on potential flow theory, which assumes that flow is incompressible, irrotational, inviscid and laminar [153]. The assumption of incompressibility, coupled with conservation of mass, enables the use of the Laplace equation (Eq. 4-1). Note that '$\emptyset$' is the velocity potential, defined by Eq. 4-2, ∇ is the vector gradient operator, $\boldsymbol{u}$ is the velocity vector and x, y and z represent the Cartesian coordinates.

$$\frac{\partial^2 \emptyset}{\partial x^2} + \frac{\partial^2 \emptyset}{\partial y^2} + \frac{\partial^2 \emptyset}{\partial z^2} = 0 \qquad\qquad \text{Eq. 4-1}$$

$$\boldsymbol{u} = -\nabla \emptyset \qquad\qquad \text{Eq. 4-2}$$

Eq. 4-1 must be disctretised in order to be used in a CFD model. This involves approximating the derivatives. One popular method of discretisation is the finite-difference method, which assumes linear variation of $\emptyset$ between nodes, and uses the Taylor series expansions [154]. Firstly, the second order derivatives must be approximated, for example $\frac{\partial^2 \phi}{\partial x^2} \sim \left(\frac{\partial \phi_{i+0.5,j,k}}{\Delta x} - \right.$ $\left. \frac{\partial \phi_{i-0.5,j,k}}{\Delta x} \right) / \Delta x$. Here, i, j and k identify the position of the ϕ value in the x, y and z

dimensions, respectively. This is demonstrated for a 2D domain in Figure 4-1, with the distance between e.g. nodes $\phi_{i,j}$ and $\phi_{i+1,j}$ being Δx. The first order derivatives can then be approximated using a similar method, giving $\frac{\partial^2 \phi}{\partial x^2} \sim \frac{\left(\frac{\phi_{i+1,j,k}-\phi_{i,j,k}}{\Delta x}-\frac{\phi_{i,j,k}-\phi_{i-1,j,k}}{\Delta x}\right)}{\Delta x} = \frac{\phi_{i+1,j,k}+\phi_{i-1,j,k}-2\phi_{i,j,k}}{(\Delta x)^2}$. Applying the same approach to the other second order derivatives in Eq. 4-1 ultimately yields Eq. 4-3 [154].

$$\phi_{i+1,j,k} + \phi_{i-1,j,k} + \phi_{i,j+1,k} + \phi_{i,j-1,k} + \phi_{i,j,k+1} + \phi_{i,j,k-1} = 6\phi_{i,j,k} \qquad \text{Eq. 4-3}$$
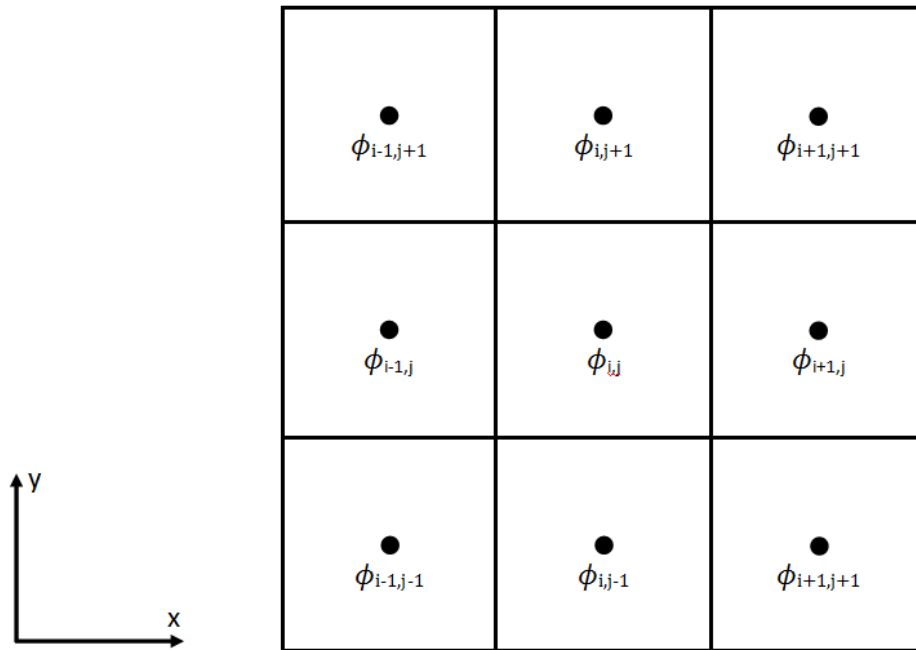
Figure 4-1. Indexing system for ϕ nodes.

Other discretisation methods used in CFD include finite-volume and finite-element methods, both of which are described by Tu [154]. The finite-volume method can be applied to a wide variety of element shapes more easily than finite-element and finite-difference methods. The finite-volume method divides the domain into discrete control volumes, with the conservation equations discretised so as to preserve continuity across the control volume boundaries. The finite-element method is more flexible than both the finite-difference and finite-volume methods since it allows 2nd or higher order discretisation. It can therefore be more readily applied to complex geometries. However, its computational intensity is relatively high. It is more commonly used in structural engineering problems than in CFD [154].

In order to solve the Laplace equation across the domain, boundary conditions representative of the data centre environment must be determined. Neumann boundary conditions, as in Eq. 4-4, can be used to represent sources and sinks of air within the data centre, for example CRAC/CRAH and server inlets and outlets [146]. Note that these may be considered sources and sinks since the insides of the CRACs/CRAHs and servers are typically not included in the domain. In Eq. 4-4, n is the unit outward normal vector at a point on the boundary of the domain, and $\boldsymbol{u}$ is the known velocity vector at this same point. Note that the velocity may not be normal to the boundary, and the direction must be specified. Note also that, in order for conservation of mass to be achieved across the domain, the sum of sources of flow must equal the sum of sinks for a steady state problem.

$$n.\nabla\phi = \boldsymbol{u}$$

<div align="right">Eq. 4-4</div>

Similarly, zero flow boundary conditions are applied at solid surfaces such as walls [146]. Such boundary conditions simply represent a special case of Eq. 4-4, where the velocity normal to the boundary is equal to zero. The values specified at the boundary conditions may be obtained from experimental measurement or from manufacturer specifications e.g. for CRAC/CRAH units or servers [146]. Algorithms representing the control systems of CRAC/CRAH units may also be included [146].

In order to solve for the temperatures across the domain, boundary conditions such as measured temperatures, controlled CRAC/CRAH supply or return temperatures, and heat sources and sinks (e.g. IT equipment and CRACs/CRAHs) are implemented, along with the consideration of advection and diffusion of heat and conservation of energy [155], which gives Eq. 4-5 (where t is time, K, ρ and c_p are the thermal conductivity, density and specific heat capacity of the fluid, T is the temperature and S_T represents any heat source or sink within the cell). This equation can be discretised in the same way as Eq. 4-1. Note that, for steady state conditions, the left hand side of Eq. 4-5 is equal to zero.

$$\frac{\partial T}{\partial t} = \frac{K}{\rho c_p}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) - \boldsymbol{u}.\nabla T + S_T$$

<div align="right">Eq. 4-5</div>

Having developed a complete set of equations for the domain, they are solved numerically to give the velocity and temperature fields. The discretised equations are linear, and may be solved via iterative methods [154]. This involves selecting the initial values of the variable at each node (known as the 'initial conditions'), before applying the equations to each node in succession. This process is repeated over a number of iterations, until convergence is achieved, as discussed in section 4.4. In the commonly-used Gauss-Siedel method, the

values of variables at nodes adjacent to that for which calculations are currently being undertaken, are taken from calculations undertaken during the present iteration where available, and are otherwise taken from the previous iteration. This tends to speed up the solution process in comparison with the Jacobi method, which exclusively uses values of variables from the previous iteration [154]. Direct, analytical methods such as Gaussian elimination may also be used as an alternative to iterative methods, and may reach a solution more quickly in circumstances where large numbers of iterations are required [154]. However, they are only applicable to problems governed by linear equations, and may not readily be applied to domains with complex geometries [156].

The assumption within the potential flow model of irrotationality is key to the validity of a simulation, and relies upon there being no tendency for 'separation' to occur within the domain [157]. Here, 'separation' refers to the bulk fluid flow becoming separated from the flow near to a solid surface due to the presence of an adverse pressure gradient at that surface [158]. Such adverse pressure gradients are caused by viscosity, as the fluid close to the surface is slowed by friction, and thus this behaviour is not captured by the potential flow model. Separation introduces both vorticity and turbulence at the intersection between the stagnating region and the bulk of the flow [159]. Since viscosity, vorticity and turbulence are neglected in potential flow theory, the propensity for separation to occur in a flow is a key factor in determining whether or not the flow can be replicated in a potential flow model.

The solution of the pressure field in potential flow may be determined by applying the Bernoulli equation to each node [157]. For steady, irrotational, inviscid flow, the Bernoulli equation is given by Eq. 4-6, where u, v and w are the velocities in the x, y and z directions, respectively, at the point of interest, and u_0, v_0 and w_0 are the velocities at a reference point at which $p = 0$.

$$p = \frac{\rho \left((u_0^2 + v_0^2 + w_0^2) - (u^2 + v^2 + w^2) \right)}{2}$$
<div align="right">Eq. 4-6</div>

4.3   NS-CFD Models

NS-CFD models are commonly used to simulate air flows in data centres [160]. Eq. 4-7 [155] shows the equation used to enforce conservation of mass in the NS equations. Note that D represents the total derivative, such that $\frac{D\rho}{Dt} = \frac{\partial \rho}{\partial t} + u \frac{\partial \rho}{\partial x} + v \frac{\partial \rho}{\partial y} + w \frac{\partial \rho}{\partial z}$. Note also that Eq. 4-7 differs from Eq. 4-1 only in that it allows for changes in density [155].

$$\frac{D\rho}{Dt} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0 \qquad \text{Eq. 4-7}$$

The NS equations also include the enforcement of conservation of momentum, which is achieved using Eq. 4-8, Eq. 4-9 and Eq. 4-10 for the x-, y- and z-directions, respectively [153], [160]–[162]. Here, μ is the dynamic viscosity, λ_v is the bulk viscosity coefficient (usually assumed to equal $-\frac{2}{3}\mu$), and F_x, F_y and F_z are any body forces acting on the element in the x-, y- and z-directions, respectively [162]. These equations represent Newton's 2nd law, equating the resultant force on each element with the rate of change of momentum. The resultant force includes viscous forces and forces due to the pressures applied to a fluid element, as well as body forces such as gravity [148]. Note that these equations (Eq. 4-8 to Eq. 4-10) are non-linear, hence the much greater computational expense of NS-CFD models in comparison with those based on the potential flow model's central linear equation (Eq. 4-1) [157].

$$\frac{\partial(\rho u)}{\partial t} + \frac{\partial(\rho u^2)}{\partial x} + \frac{\partial(\rho uv)}{\partial y} + \frac{\partial(\rho uw)}{\partial z} =$$

$$-\frac{\partial p}{\partial x} + \frac{\partial}{\partial x}\left(\lambda_v \nabla \cdot \boldsymbol{u} + 2\mu\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left[\mu\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)\right] + \frac{\partial}{\partial z}\left[\mu\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)\right] + \rho F_x$$

Eq. 4-8

$$\frac{\partial(\rho v)}{\partial t} + \frac{\partial(\rho uv)}{\partial x} + \frac{\partial(\rho v^2)}{\partial y} + \frac{\partial(\rho vw)}{\partial z} =$$

$$-\frac{\partial p}{\partial y} + \frac{\partial}{\partial x}\left[\mu\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)\right] + \frac{\partial}{\partial y}\left(\lambda_v \nabla \cdot \boldsymbol{u} + 2\mu\frac{\partial v}{\partial y}\right) + \frac{\partial}{\partial z}\left[\mu\left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right)\right] + \rho F_y$$

Eq. 4-9

$$\frac{\partial(\rho w)}{\partial t} + \frac{\partial(\rho uw)}{\partial x} + \frac{\partial(\rho vw)}{\partial y} + \frac{\partial(\rho w^2)}{\partial z} =$$

$$-\frac{\partial p}{\partial z} + \frac{\partial}{\partial x}\left[\mu\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)\right] + \frac{\partial}{\partial y}\left[\mu\left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right)\right] + \frac{\partial}{\partial z}\left(\lambda_v \nabla \cdot \boldsymbol{u} + 2\mu\frac{\partial w}{\partial z}\right)$$

$$+ \rho F_z$$

Eq. 4-10

The last of the NS equations ensures compliance with the first law of thermodynamics - that energy must be conserved. The application of this law to an element within the domain is described in words by Anderson, Jr. as follows [162]:

$$
\begin{array}{ccc}
\text{Rate of change of energy} & = & \text{Net flux of heat} & + & \text{Rate of work done on the} \\
\text{inside the fluid element} & & \text{into the} & & \text{element due to body and surface} \\
& & \text{element} & & \text{forces}
\end{array}
$$

Here, "surface forces" refers to pressure, shear stresses and normal stresses. The heat flux into the element comprises radiative, conductive, and convective heat transfer. The rate of change of energy in the fluid element includes internal energy and kinetic energy. Eq. 4-11 shows the conservation of energy equation [162]. Here, e is the specific internal energy, $\dot{Q}$ is the sum of the rates of radiative and convective heat transfer into the cell and any sources/sinks of energy within the cell, τ_{xy} is the stress acting in the y-direction on the element faces in the x-plane, and F is the body force vector. Note that D again denotes the total derivative, and that, for steady state conditions, the left hand side of the equation is equal to zero.

$$
\rho \frac{\mathrm{D}}{\mathrm{Dt}}\left(e + \frac{u^2 + v^2 + w^2}{2}\right)
$$

$$
= \rho\dot{Q} + \frac{\partial}{\partial x}\left(K\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(K\frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(K\frac{\partial T}{\partial z}\right) - \frac{\partial(up)}{\partial x}
$$

$$
- \frac{\partial(vp)}{\partial y} - \frac{\partial(wp)}{\partial z} + \frac{\partial(u\tau_{xx})}{\partial x} + \frac{\partial(u\tau_{yx})}{\partial y} + \frac{\partial(u\tau_{zx})}{\partial z} + \frac{\partial(v\tau_{xy})}{\partial x} \quad \text{Eq. 4-11}
$$

$$
+ \frac{\partial(v\tau_{yy})}{\partial y} + \frac{\partial(v\tau_{zy})}{\partial z} + \frac{\partial(w\tau_{xz})}{\partial x} + \frac{\partial(w\tau_{yz})}{\partial y} + \frac{\partial(w\tau_{zz})}{\partial z} + \rho F \cdot \boldsymbol{u}
$$

Full derivations of the conservation of mass, momentum and energy equations may be found in Anderson, Jr. [162].

Unlike the PF-CFD model, NS-CFD models can include turbulence. Turbulence occurs in fluid flows with high Reynolds numbers (Re), and is characterised by disordered and irregular motion which causes exchange of momentum from one portion of a flow to another [163]. Re indicates the relative importance of inertial and viscous forces, and is given by Eq. 4-12 [163]. Here, q_C and L_C are a scalar velocity and length, respectively, which characterise the flow (for example, in a pipe flow, q_C is the mean velocity through the pipe, and L_c is its diameter) [158]. With small Re, viscous forces dominate and any disturbances in the flow are dissipated away. With large Re, inertial forces dominate, allowing disturbances to propagate. This causes chaotic, turbulent behaviour, which needs to be incorporated into numerical models. In data centres, Re is likely to be sufficiently small in the bulk of the

domain that flow remains laminar [145]. However, turbulent behaviour is likely to occur close to surfaces [145].

$$Re = \rho q_C L_C / \mu$$
<div align="right">Eq. 4-12</div>

Turbulence is usually modelled in NS-CFD using the Reynolds Averaged Navier-Stokes (RANS) equations [160]. This involves decomposing the instantaneous magnitudes of the flow variables into a mean and a fluctuating component, as shown in Eq. 4-13 for the velocity field. Here, $\bar{u}$ is the long term average velocity vector and u' is the velocity fluctuation vector. The Navier-Stokes equations must be modified to incorporate this decomposition, creating the RANS equations.

$$\boldsymbol{u} = \boldsymbol{\bar{u}} + \boldsymbol{u'}$$
<div align="right">Eq. 4-13</div>

A model must then be used to determine the fluctuations due to turbulence. In data centre NS-CFD models, the $k - \varepsilon$ version of the eddy-viscosity model (EVM) is the most commonly used due to its simplicity, stability, robustness and low computational expense [160] [164]. EVMs are based on the assumption that the eddy viscosity, μ_t, is proportional to the velocity gradient, as shown in Eq. 4-14 for the 2-dimensional case [160]. Here, u' and v' are the x and y components of the velocity fluctuation, respectively. μ_t is defined in terms of the kinematic eddy viscosity, v_t, by Eq. 4-15 [160].

$$-\rho \overline{u'v'} = \mu_t \frac{\partial u}{\partial y}$$
<div align="right">Eq. 4-14</div>

$$v_t = \mu_t / \rho$$
<div align="right">Eq. 4-15</div>

Solution of the RANS equations then requires the specification of μ_t and v_t [160]. The equations governing the $k - \varepsilon$ model account for the dissipation of energy via turbulence, and must be incorporated into the equations for conservation of mass, momentum and energy. In the $k - \varepsilon$ model, v_t is defined by Eq. 4-16, where C_μ is a constant, usually set at 0.09, k_t is the rate of production of turbulent kinetic energy and ε is the rate of turbulent dissipation of kinetic energy [160]. In order to determine k_t and ε, Eq. 4-17 and Eq. 4-18 must be solved [155], [160]. The constants σ_k, σ_ε, $C_{\varepsilon1}$ and $C_{\varepsilon2}$ are typically set at 1.0, 1.3, 1.44 and 1.92, respectively, whilst P_t represents the rate of production of turbulent kinetic energy, and is given by Eq. 4-19 [155]. Note that Eq. 4-17 and Eq. 4-18 are shown for the 2-dimensional case.

$$v_t = C_\mu \frac{k_t^2}{\varepsilon}$$

<div align="right">Eq. 4-16</div>

$$\frac{\partial k_t}{\partial t} + u\frac{\partial k_t}{\partial x} + \frac{v \partial k_t}{\partial y} = \frac{\partial}{\partial x}\left(\frac{\mu_t}{\sigma_k}\frac{\partial k_t}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\mu_t}{\sigma_k}\frac{\partial k_t}{\partial y}\right) + P_t - \varepsilon$$

<div align="right">Eq. 4-17</div>

$$\frac{\partial \varepsilon}{\partial t} + u\frac{\partial \varepsilon}{\partial x} + v\frac{\partial \varepsilon}{\partial y} = \frac{\partial}{\partial x}\left(\frac{\mu_t}{\sigma_\varepsilon}\frac{\partial \varepsilon}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\mu_t}{\sigma_\varepsilon}\frac{\partial \varepsilon}{\partial y}\right) + \frac{\varepsilon}{k_t}(C_{\varepsilon 1}P_t - C_{\varepsilon 2}\varepsilon)$$

<div align="right">Eq. 4-18</div>

$$P_t = 2\mu_t \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2\right] + \mu_t\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)^2$$

<div align="right">Eq. 4-19</div>

It should be noted that the prevalence of the $k - \varepsilon$ model in data centre NS-CFD models has received little critical attention. One study undertaken by Cruz et al. [165] found that the inclusion of the $k - \varepsilon$ model provided little improvement in accuracy when compared with simulations which assumed laminar flow. The zero equation [166] and Spalart-Allmaras [167] turbulence models were found to slightly improve accuracy, whilst increasing solution time [165]. This study is discussed in more detail in section 4.6.2.

Having established the governing equations, as with the PF-CFD model, appropriate boundary conditions must be identified. A solution may then be achieved by applying the governing equations to each cell in turn, using iterative methods.

4.4   Stability and convergence

The number of iterations required to achieve convergence in CFD models may be reduced using a technique known as successive over-relaxation. Here, convergence refers to the tendency for changes in the model results in successive iterations to diminish as more iterations are completed. In successive over-relaxation, the value of a variable at the current node is influenced not only by the governing equations, but also by its value in the previous iteration. For example, ϕ as calculated for the current iteration using the governing equations, ϕ^n, may be updated at the end of each iteration using Eq. 4-20, where W_{SOR} is the relaxation factor, which is set between 1 and 2 for the Gauss-Siedel method [154]. Here, ϕ^{n-1} refers to the value of ϕ at the previous iteration. Raising W_{SOR} tends to increase the solution speed, but can induce instability. One method for minimising the number of iterations required to achieve convergence, whilst avoiding instability, is to initially set W_{SOR} at a high level, before progressively reducing it to ensure stability as a solution is neared [168].

$$\phi^n = (1 - W_{SOR})\phi^{n-1} + W_{SOR}\phi^n \qquad\qquad \text{Eq. 4-20}$$

Another key factor determining the number of iterations required to achieve convergence is the accuracy of the initial conditions. Using the results of previous simulations or experimental measurements as initial conditions can improve the suitability, and hence reduce the number of iterations required [168].

Another method for reducing the number of required iterations is "upwinding". This method may be used to increase the convergence speed of variables such as temperature, by increasing the influence of temperatures upstream of the current node on that node's heat transfer calculations. For example, when applying Eq. 4-5 to node (i, j, k), the term $\frac{\partial^2 T}{\partial x^2}$ would normally be approximated as $\frac{T_{i+1,j,k} + T_{i-1,j,k} - 2T_{i,j,k}}{(\Delta x)^2}$ (in the same way as described in relation to Eq. 4-1). With upwinding applied, where $u(i, j, k) > 0$, $T_{i+1,j,k}$ may be replaced with $T_{i,j,k}$, and $T_{i,j,k}$ with $T_{i-1,j,k}$.

4.5   Verification and validation of CFD models

In order to assess the value of a given model, steps must be taken to investigate the extent to which the model's results are consistent with reality. This generally involves examining the behaviour of the model through two distinct processes: verification and validation.

Verification is concerned with assessing the extent to which the simulation results represent an exact solution of the governing equations [168]. Hence, verification involves testing the sensitivity of the model to element size, as well as to other input parameters such as geometry, initial conditions and boundary conditions. Verification also involves monitoring of the "residuals", which describe the extent to which the solution of the discretised equations is consistent with the governing equations at the element level. In the 3D application of the potential flow model, the ϕ residual, R_ϕ, is given by Eq. 4-21 [168]. Convergence is deemed to have been achieved when R_ϕ falls below some pre-defined level.

$$R_\phi = \phi_{i+1,j,k} + \phi_{i-1,j,k} + \phi_{i,j+1,k} + \phi_{i,j-1,k} + \phi_{i,j,k+1} + \phi_{i,j,k-1} - 6\phi_{i,j,k} \qquad \text{Eq. 4-21}$$

Validation is concerned with comparing the results of the model with results of experimental investigations, or previous numerical simulations where experimental data is unavailable [168]. A quantitative comparison is usually made between the two sets of results. The extent and types of data used to validate data centre CFD models varies, but can include:

- grille flow measurement [31] [37],

- air temperatures and velocities at IT and CRAC/CRAH inlets and outlets [31], [37], [39], [45],

- power consumption of CRAC/CRAHs, fans, pumps, power distribution systems and IT equipment [39].

4.6 Summary of data centre CFD models demonstrated to date

Numerous CFD models of data centres have been described in the research literature, using both PF and NS methods.

4.6.1 Potential Flow CFD models

Whilst the majority of data centre CFD models reported in the research literature use the full NS equations, PF-CFD models have also received some attention. Seymour [169] notes that the exclusion of viscosity and conservation of momentum from data centre PF-CFD models allows unrealistically abrupt changes in direction. Toulouse et al. [145] argue that whilst the exclusion of viscosity from these models limits the accuracy of predictions close to solid surfaces, the predictions of large scale flow features may still be sufficiently accurate to achieve the purposes of the model. Some investigations have found that acceptable levels of accuracy can only be achieved using PF-CFD models if data from either experimental measurements or NS-CFD simulations are used as inputs to the models [150]. The speed of solution in comparison with NS-CFD can make these models attractive for some purposes. For example, Toulouse et al. [145] and Healey et al. [152] suggest using PF-CFD models during the early stages of data centre design, as a precursor to NS-CFD modelling. None of the PF-CFD models presented to date have represented data centres employing aisle containment.

Lopez & Hamann [150] developed a PF-CFD model of a real data centre for which detailed air temperature and velocity measurements had been taken. The model was designed to use real time measured data to enable a quick, accurate solution. The accuracy of predicted rack inlet temperatures was found to vary from ±2 K to ±4.5 K depending on the number of measurements used to inform the model. This level of accuracy compares favourably with results of NS-CFD simulations, discussed in Section 4.6.2. However, achieving the most accurate solution required several temperature measurements from each server inlet to be inputted to the model. The least accurate solution required one temperature measurement for every two servers to be supplied to the model. Clearly the reliance on measurements

limits the usefulness of the model, particularly when predicting the impact of changes made to the data centre, due to the effect of changes on the measured data used as inputs. The model is most obviously useful in predicting the positions of hot spots in a current data centre configuration, where a substantial amount of input data is available.

Toulouse et al. [145] developed a PF-CFD model for a simple data centre test case with a single cold air supply, hot air return and server rack. The speed of solution for this model was increased by determining the convective heat transfers via the tracing of flows from heat sources along streamlines, avoiding the need to calculate heat transfers over successive iterations. Conductive heat transfers were ignored.

A later paper by Lettieri et al. [170] modified Toulouse et al.'s [145] model to predict flows within a real data centre, with the results differing significantly from experimental results. The mean deviation between predicted and measured temperatures at server inlets was around 5 K, with a maximum deviation of 25 K [170]. However, amending the model to include buoyancy using a "Vortex Superposition" method led to mean and maximum deviations of 2 and 4.5 K, respectively, which was similar to deviation corresponding to the results of a NS-CFD model with the same geometry, produced using Ansys [170].

Healey et al. [152] produced PF- and NS-CFD models of 8 different example data centre layouts. They found good agreement between the two model types for mean server inlet temperatures, with most being within 2 K.

4.6.2 Navier-Stokes CFD models

A variety of software packages have been used to complete NS-CFD simulations of data centres. This includes software designed specifically for data centre applications, such as Future Facilities' 6Sigma package [39], [87], [90], [92], [171] and Innovative Research's 'Tileflow' [42]. More general NS-CFD packages, such as Star-CD [46], [91], FLOVENT [45], ANSYS CFX [172] and Fluent [40], [37], [84], [173] have also been used.

Researchers have used NS-CFD models to investigate the impact on energy consumption and/or thermal environment of measures such as:

- Installation of full aisle containment [46], [171];
- Adding a ceiling to the cold aisles, whilst leaving the aisle ends open [84];
- Changing the height of the data centre's ceiling [45];
- Supplying air from above the cold aisle, rather than from an underfloor plenum [46];
- Allowing some air leaving the servers to bypass the CRAHs, before being mixed with conditioned air [40];

- Providing cooling via air to liquid heat exchangers in each individual rack, rather than at the room level [87];

- Installing fan-assisted floor grilles to better direct supply air [89];

- Reducing redundancy in the cold air supply, increasing the set point return air temperature, introducing aisle containment, and implementing air-side economisation [39].

Other researchers have studied the impacts of variations in computational methods on the model predictions. Arghode et al. [37] demonstrated that, for a data centre with cold air supplied via an underfloor plenum, consideration of the increased momentum induced in air passing through floor grilles helped to improve the accuracy of temperature predictions made by a NS-CFD model. Samadiani et al. [173] investigated the importance of including obstructions within the underfloor plenum in the model geometry, and of including pressure losses incurred across floor grilles, finding that both measures significantly improved the accuracy of predicted tile flow rates. Alissa et al. [171] demonstrated the accuracy and low computational expense of using experimentally derived flow curves for servers and CRAHs. Cruz et al. [165] compared the accuracy of NS-CFD models using a variety of turbulence models, when compared with experimental results. The $k - \varepsilon$ model was found to produce results with accuracy similar to the laminar model (with mean errors on temperature measurements ranging from 4 to 6 K), whilst the Spalart-Allmaras turbulence model performed slightly better (with mean errors ranging from 3 to 4.5 K). De Boer et al. [174] (in press) compared the performance of finite-element and finite-volume based NS-CFD models of an example data centre layout, finding that both methods predicted similar temperature distributions. Siriwardana et al. [172] used results of a NS-CFD model to train a statistics-based model, which was then used to optimise the placements of servers in order to minimise cooling power consumption.

Cruz & Joshi [175] developed a model which combined regions of the domain in which the full RANS equations were used, with other regions in which viscosity was neglected. An extensive design of experiments was required in order to determine appropriate parameters for the partitioning of the 2 regions, as well as for the convergence criteria. Overall, the results of the combined RANS-inviscid model were very similar to those of a model using the full RANS equations for the entire domain. The solve time was significantly reduced, although it must be assumed that the time required to develop the model was larger for the combined RANS-inviscid model than for the standard NS-CFD model.

As discussed in Chapter 2, numerous authors have used NS-CFD simulations to demonstrate the benefits of aisle containment for improving uniformity of cold aisle temperatures, reducing bypass and/or recirculation, and reducing electricity consumption [46], [84], [86], [90], [92]. The majority of authors describing models utilising containment have neglected bypass and recirculation under contained conditions [88]–[91], or have not disclosed the methods used to model bypass and recirculation [46], [87], [92]. Where bypass and recirculation are considered, this is often restricted to containment leakage, with rack leakage (air passing through the racks whilst avoiding the servers) being neglected.

Ham & Jeong [86] have presented one of the few models of a data centre employing aisle containment for which the methodology governing bypass and recirculation has been disclosed. Rack leakage was neglected, with all flow through the racks assumed to pass through the servers. Containment leakage was modelled by estimating the likely effective leakage area, based on reference leakage test data obtained from the indoor air quality and ventilation software, CONTAM (NIST, cited in [86]). Leakage paths with areas summing to this effective leakage area were positioned along the perimeter of the containment system, which was judged to be the most likely leakage path. Gondipalli et al. [176] have also used a similar approach, although with a much larger area left open to allow containment leakage. No justification was provided for the sizing or positioning of this area.

As discussed in section 2.1.2, only Alkharabsheh et al. [72] have provided any detail of methods used to model rack leakage in data centres employing containment. Their model applied percentage open areas to various potential leakage paths, identified via visual inspection of a data centre employing containment. Specifically, containment leakage was allowed around the perimeter of the containment system, and rack leakage along the lengths of the equipment rails. No direct validation of the percentage open areas was undertaken, although other aspects of the model, specifically CRAH and load bank flow rates, were validated successfully. The model's predictions of temperatures at the rack inlets were presented, and differed considerably from the respective experimental measurements. The authors highlight an average error of 0.99 K on rack inlet temperature predictions. However, this is large in comparison with the range of measured rack inlet temperatures (i.e. the difference between the lowest and highest measured rack inlet temperature), which was less than 1.5 K. By contrast, the cold aisle temperatures predicted by the model covered a range of around 5 K, with no obvious correlation between the predictions and measurements. Indeed, by this measure, similar accuracy would have been

achieved by fixing all cold aisle temperatures at the same temperature as the air exiting the CRAH.

This highlights a common issue with assessments of accuracy of data centre CFD models. Where experimental validation is undertaken, the most common data used to make the validation are the rack or server inlet temperatures. The most accurate NS-CFD models generally report errors on predicted inlet temperatures of up to 5 K [31], [37], [165], with others reporting errors of up to 10 K, or even higher [42], [45], [172]. Generally, errors are higher for larger data centres with greater numbers of servers, with high rack power density, and also where recirculation is higher [37], [45]. It is therefore perhaps misleading to compare accuracies of different models simply in terms of an absolute error on inlet temperature predictions. For example, a model of a small data centre with low rack power density and little recirculation may achieve errors much smaller in absolute magnitude than those prevalent when applying the same methods to a large, densely populated data centre with extensive recirculation. It may be more productive to develop a non-dimensional measure of accuracy, for example by showing the maximum error in server inlet temperature predictions (ε_T) divided by the total range of measured server inlet temperatures (i.e. the difference between the maximum and minimum measured inlet temperature, T_{range}). Table 4-1 shows a comparison on this basis between the most accurate models reported in the literature. The best result comes from Lopez & Hamann [150], a PF-CFD model. However, as noted in section 4.6.1, this model required large numbers of temperatures within the domain to be specified in accordance with measured data in order to achieve this level of accuracy.

It should also be noted that errors in temperature predictions of 5 K may be considered large in the context of ASHRAE's recommended server inlet temperatures, which cover a range of 9 K [177]. However, the data presented in Table 4-1 clearly shows that the most accurate CFD models can have considerable success in predicting the patterns of variation in temperatures at server inlets, even if the errors in isolated positions may be significant. Hence, whilst their precise predictions must be considered in the context of the likely errors, the most accurate CFD simulations can be broadly effective in predicting the positions of hotter temperatures, and the impact of changes in geometry and heat load. This makes CFD a useful tool in data centre design and management, whilst real-time monitoring may be required to confirm the precise positions and temperatures of hot spots during operation.

Author	Max inlet temperature error, K (ε_T)	Measured inlet temperature range, K (T_{range})	ε_T/T_{range}
Patankar [31]	5	15	0.33
Cruz et al. [165]	3.5	3	1.17
Arghode et al. [37]	4	7	0.57
Alkharabsheh et al. [72]	2	1.5	1.33
Lopez & Hamann [150]*	2-4.5	12.8	0.16-0.35
Lettieri et al. [170]*	4.5	Not reported	-

Table 4-1. Comparison of errors in the most accurate, validated data centre CFD models reported in the research literature. Note that '*' denotes PF-CFD model, with the remainder being NS-CFD.

Two authors have reported maximum errors on inlet temperature predictions of less than 1 K. However these errors corresponded to validations against a single inlet temperature in a data centre containing 10 racks [89], and two inlet temperatures in a data centre containing 20 racks [86], respectively. In the latter case [86], the difference between the two measured inlet temperatures and the supply temperature was more than double the difference between the predicted inlet temperatures and the supply temperature. The supply air temperature is not disclosed in the former study [89]. These studies have not been included in Table 4-1 due to the limitations in the validation processes.

Some authors have also used grille flow rates and CRAH thermal loads for validation of NS-CFD models. Patankar [31] reported predicted grille flow rates as being within 10% of measured values, whilst Alissa et al. [171] achieved average errors on predicted grille flow rates of less than 5%. Choo, Galante & Ohadi [39] reported predicted CRAH thermal loads as being within 10% of measured values.

A variety of reasons are cited for errors in simulation results, such as:

- too coarse grid spacing [42], [45],
- poor modelling of turbulence, particularly in relation to hot air at server outlets [42],
- fluctuating temperatures in high power density regions, causing inaccuracy in the measured temperatures [45],

- inaccurate representation of obstructions within the plenum [31], [44], [173],

- failure to consider the effect on CRAC/CRAH flow rates of the flow resistance offered by floor grilles [173], and

- failure to account for the impact of floor grille open areas on the momentum of air passing through them [37].

4.7 Summary

Sections 4.1 to 4.5 have described the mathematical basis for CFD modelling techniques commonly used in relation to data centres. A review of data centre CFD models described in the research literature to date has then been presented in Section 4.6. The various models discussed have been used to investigate the impacts of measures such as aisle containment, reduction of supply air flow rate and other changes in geometry on temperature profiles and electricity consumption. The review has highlighted the importance of factors such as accounting for momentum increases and pressure losses in flow through floor grilles, and accurate representation of obstructions in underfloor plenums. Some of the models presented in the research literature have achieved sufficient accuracy in server inlet temperature predictions to demonstrate the value of CFD modelling in aiding data centre design and management. However, others have struggled to achieve accurate results, demonstrating the importance of selecting appropriate boundary conditions and modelling techniques. In section 4.6.2, some issues regarding common approaches to assessing the accuracy of data centre CFD models have also been highlighted, and a new approach proposed.

5. COMPUTATIONAL FLUID DYNAMICS SIMULATIONS OF DATA CENTRES EMPLOYING AISLE CONTAINMENT

The literature review provided in Chapter 4 has shown how computational fluid dynamics (CFD) models can be used to aid in data centre design and management. Some limitations in the existing research literature were highlighted, particularly in the modelling of data centres employing aisle containment. Specifically, potential flow CFD (PF-CFD) methods have not previously been used to model data centres employing aisle containment. In addition, researchers using Navier-Stokes CFD (NS-CFD) methods to model these data centres have not fully demonstrated the validity of approaches to incorporating bypass and recirculation, nor has the influence of Δp_{CH} been thoroughly investigated. It follows that there is value in further developing techniques used in CFD models of data centres employing aisle containment, since improvement of these techniques could increase the potential for using these models to assist in data design and management.

In this chapter, PF- and NS-CFD models are produced, validated, and used to investigate air flows in data centres employing aisle containment. The models utilise the experimental results presented in Chapter 2 to govern bypass and recirculation. This chapter describes the modelling methods used and presents validations of the models against temperature measurements taken in the Test Data Centre, as described in section 2.2.3.3. Finally, the models are expanded to investigate the impacts of Δp_{CH} and efforts to minimise bypass and recirculation on electricity consumption and cooling effectiveness in an example data centre geometry. This work seeks to fulfil objective (3), as defined in section 1.5, which is:

'To investigate the potential for using CFD models to aid in the efficient design and management of data centres employing aisle containment'

5.1 Potential Flow computational fluid dynamics modelling

5.1.1 A Potential Flow model of the Test Data Centre

A PF-CFD model of the Test Data Centre was developed, largely following the methods of Toulouse et al. [145] and Lettieri et al. [170]. This represents the first demonstration of a PF-CFD model of a data centre employing aisle containment. This section describes the methods used. The MATLAB software platform was used to compute the model [134]. The full script is provided in the supplementary material included with this thesis.

5.1.1.1 Geometry

Some minor differences between the geometry of the Test Data Centre (as detailed in section 2.2.3.1) and its representation in the model were allowed, in order to simplify the discretisation of the domain. This allows the positions of features such as inlets, outlets and solid walls to coincide with nodes and cell boundaries. The differences are described in the forthcoming text.

The domain was divided into hexahedral cells of equal size, with each side of each cell also having the same length, D. D was set such that the width of a rack, $0.8\ m$, when divided by D, gave an even integer, allowing the load banks used in the experiments to be positioned centrally within the racks. A grid independence study was completed to determine the appropriate value of D, as described in section 5.5.1.

The convention used in this chapter to define the spatial dimensions is demonstrated in Figure 5-1. The 'x' dimension is designated as representing the direction along the length of the aisle, the 'y' dimension as the vertical direction, and the 'z' dimension as the direction along the length of the racks. The length of the Test Data Centre in the x dimension was set to $3.2\ m$. Note that in the real case, the length was $3.23\ m$ – this was simplified to represent the length as corresponding to an exact number of racks, which are $0.8\ m$ wide. The dimensions of the Test Data Centre in the y and z dimensions were set as close as possible to the true lengths of the Test Data Centre in these dimensions (3.4 and $3.97\ m$, respectively), whilst maintaining a length which could be divided into an exact number of cells. The inside surfaces of the walls, floor and ceiling of the Test Data Centre correspond to cell boundaries. Features protruding from the walls, such as the aluminium sections of the walls, were not included in the model, i.e. the walls were modelled as flat surfaces.

The velocity potential (ϕ) field consists of nodes at the centre of each cell within the domain, with an extra layer of 'virtual nodes' being included just outside of the boundaries of the domain. This allows the boundary conditions at the domain boundaries to be set, as will be described in section 5.1.1.2.
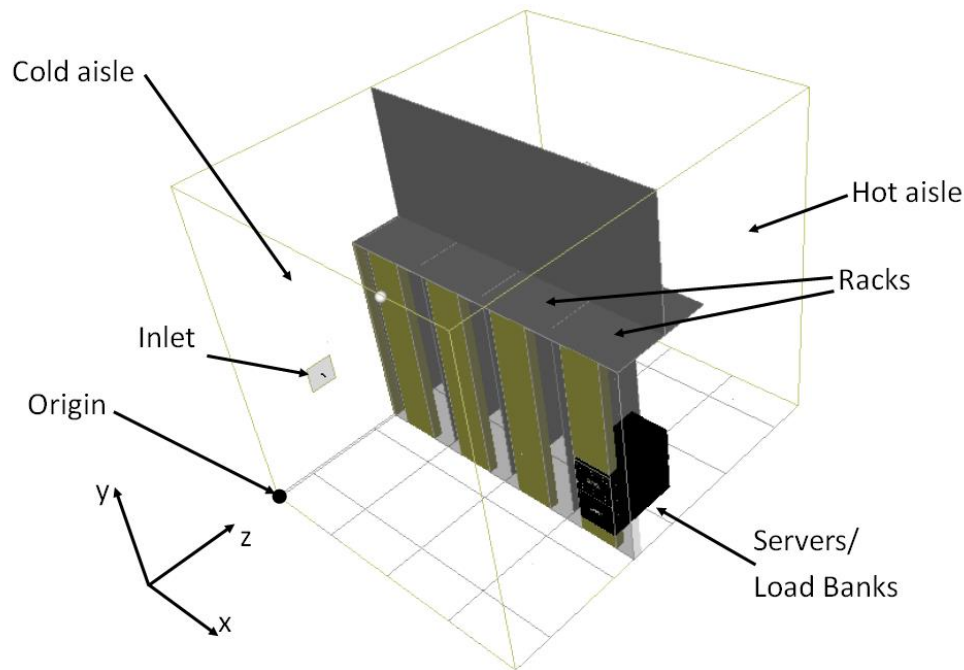
Figure 5-1. Definition of x, y and z dimensions in PF-CFD model (note that the outlet from the hot aisle is on the same wall as the inlet, and is obscured by the racks).

5.1.1.2   Positions of applied boundary conditions

An array was created to store references denoting the boundary condition relating to each ϕ node. This array is subsequently referred to as the BCs array.

The walls enclosing the domain were identified within the BCs array by labelling all nodes coinciding with solid walls, and falling outside of the Test Data Centre, with the number 1. Nodes corresponding to the inlet to the Test Data Centre were identified in the BCs array by the number 2. The geometry of the inlet was simplified to be square in shape, as opposed to circular in the real case. The distance of the centre of the inlet from the origin (shown in Figure 5-1) in both the y and z dimensions was stored in the model, as was the cross sectional area of the true Test Data Centre inlet ($7.79 \times 10^{-2}\ m^2$). Any nodes in the BCs array which fell within the area of a square of side $\sqrt{7.79 \times 10^{-2}}\ m$ centred at the appropriate point were then set to 2 to denote 'inlet'. The outlet was represented in the same way, with the corresponding nodes identified with the number 3.

Nodes at a distance of $0.5\ D$ from the plane separating the hot and cold aisles, on the hot aisle side of the partition, were initially identified with the number 4 in the BCs array. The geometry of the Test Data Centre is here simplified such that these nodes form a single vertical plane coinciding with the server or load bank (subsequently referred to as the heat

load) inlets (as shown in Figure 5-1, the partition above the server racks is in fact set back from the rack fronts).

Some of the nodes falling within this plane were then amended in the BCs array to identify the positions of the heat load inlets, leakage paths, and empty slots where applicable. The heat load inlets, leakage paths and empty slots were identified by the numbers 5, 8 and 11, respectively. The geometry of the heat load inlet was simplified to cover the entire front face of each heat load left open in the conditions being modelled. The heat load width was set to $0.4 \, m$, allowing it to be modelled as occupying, for example, 4 nodes in width where $D = 0.1$. The position in the y dimension of the bottom of the heat load inlet was stored in the model, as was the height of the server. Any nodes falling within vertical positions corresponding to the heat load, and within the appropriate nodes in the x and z dimensions, were then set to 5. The leakage paths were set to cover the areas at either side of the equipment rails, running the entire height of the server racks. The rack heights were set as close to 2.1m as possible, whilst equalling an exact number of cell lengths. Areas which were intended to be represented as impermeable in the model, specifically the aisle partition and blanking panels, retain the identification number, 4. The positions of BCs 4, 5, 8 and 11 are shown in Figure 5-2.
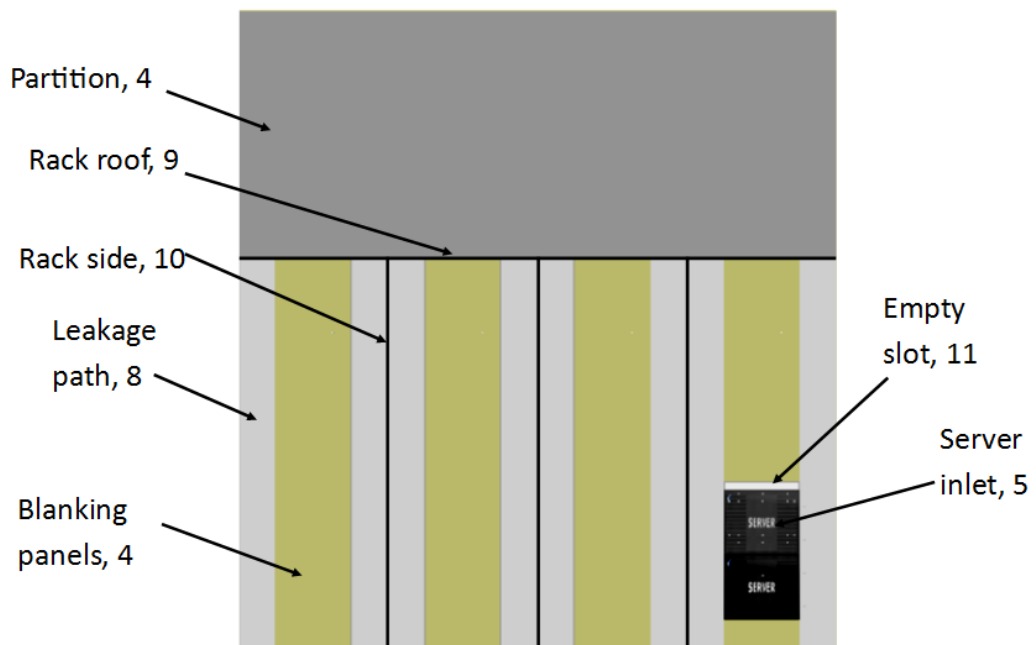
Figure 5-2. Positions and identifications of boundary conditions, shown in a z-plane at the heat load inlets.

Nodes falling within the heat loads are identified within the BCs array with the number 6. Note that the length of the heat load was set to include the length of the duct attached to the back of the load banks in the experiments described in section 2.2.3.3. Nodes corresponding to the heat load outlets are identified with the number 7. Again, the outlet is assumed to cover the entire back face of any heat loads through which flow is allowed.

The rack roofs are identified in the BCs array by the number 9. The roofs are represented as a continuous layer, which is one node thick, and which stretches across the entire length of the aisle, with the depth equal to the true depth of the racks (to the nearest cell length). The rack sides are identified by the number 10, with the sides of 2 adjacent racks represented by a plane which is one node thick. The side of racks 1 and 4 which are adjacent to the Test Data Centre walls are not included in the model. Figure 5-3 shows the boundary conditions stored in an example BCs array for a horizontal slice passing through a heat load.

In some simulations, an empty slot was included in the model (i.e. an equipment slot containing no heat load and no blanking panel). Where this is the case, the slot is represented by the number 11 in the BCs array, occupying a row of nodes one node thick immediately above the server inlet. Note that this gives an empty slot with a height of $D\ m$, which in all simulations was larger than the true height of the empty slot ($0.044\ m$).

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	10	10	10	10	10	10	10	10	10	10	10	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	10	10	10	10	10	10	10	10	10	10	10	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	10	10	10	10	10	10	10	10	10	10	10	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	5	6	6	6	6	6	6	6	7	0	0	0	0	0
0	0	5	6	6	6	6	6	6	6	7	0	0	0	0	0
0	0	5	6	6	6	6	6	6	6	7	0	0	0	0	0
0	0	5	6	6	6	6	6	6	6	7	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 5-3. Slice of BCs array, passing through the heat load and racks. 0=no boundary condition, 1=solid wall, 4=impermeable partition, 5=server inlet, 6=inside server, 7=server outlet, 8=leakage path, 10=rack side.

5.1.1.3   Initialisation

The flow rate through the leakage paths is initialised using data from the experiments detailed in Chapter 2. An initial value of Δp_{CH} (the pressure difference between the cold and hot aisles) is selected, in accordance with the experimental results for the conditions to be modelled, and the expected bypass flow rate calculated using either Eq. 2-28 or Eq. 2-33, depending on the bypass conditions being modelled. Note here that, as in previous chapters, 'pressure' refers to static pressure. The supply flow rate ($\dot{V}_{CRAH}$) is set as an input to the model, and the flow rate through the heat loads ($\dot{V}_{server/LB}$) is initialised in accordance with the experimental results for the conditions to be modelled.

Arrays are created for the Phi (ϕ), pressure (p) and temperature (T) fields, with nodes coinciding with those in the BCs array. The values assigned to the nodes in the ϕ and p fields are 0 and 1.01×10^5 (atmospheric pressure in Pa), respectively.

The values assigned to nodes in the T field were selected to represent the likely average temperatures in each aisle. Nodes within the cold aisle were simply set equal to T_{supply}. As discussed in section 2.3.2.2, conservation of heat calculations suggested that significant heat transfer occurred through the external walls, ceiling and floor of the hot aisle, amounting to a heat transfer rate, $U_{HA}A_{HA}$, of 0.192 $kW.K^{-1}$. Note here that A_{HA} (m^2) is the combined surface area of the external walls, ceiling and floor of the hot aisle. Eq. 2-22, re-printed here as Eq. 5-1 for ease of reference, therefore enables T_{HA} to be determined since all other quantities are known. Hence, T in the nodes in the hot aisle was initialised at T_{HA}.

$$U_{HA}A_{HA}(T_{HA} - T_{supply}) = \dot{E}_{LB} - \dot{V}_{supply}\rho_{air}c_p(T_{HA} - T_{supply}) \qquad \text{Eq. 5-1}$$

5.1.1.4   Satisfaction of the Laplace equation

An iterative process (specifically the Gauss-Siedel method [154]) is used to satisfy the Laplace equation (Eq. 4-1) across the domain, which is discretised via the Finite-Difference method.

Δp_{CH} is first calculated, by summing and averaging the pressure at nodes in the cold aisle and then in the hot aisle, and taking the difference between the two averages. This is used to calculate the flow rates through the leakage paths, as described later in this section.

The values of ϕ at the nodes with $x = -\frac{D}{2}$, $y = -\frac{D}{2}$ or $z = -\frac{D}{2}$ are then determined. For each of these nodes, the BCs array is consulted to determine whether the node corresponds to a solid wall, an inlet or an outlet. For a solid wall, ϕ is set equal to ϕ at the neighbouring node corresponding to open space within the domain. I.e. If node $(1, j, k)$ corresponds to solid wall, $\phi_{1,j,k} = \phi_{2,j,k}$. Here, i, j and k are used as in Figure 5-4, and node (1,1,1) corresponds to $x = y = z = -0.5D$.

If node $(1, j, k)$ is identified as corresponding to the inlet, $\phi_{1,j,k}$ must be set so as to ensure flow into the domain. Accordingly, since the velocity between nodes $(1, j, k)$ and $(2, j, k)$ is given by $\frac{\phi_{1,j,k} - \phi_{2,j,k}}{D}$, in this case, $\phi_{1,j,k} = \phi_{2,j,k} + \frac{\dot{V}_{CRAH}}{N_{inlet}D^2}D$. Here, N_{inlet} is the number of nodes comprising the inlet. $\dot{V}_{CRAH}$ is the flow rate into the Test Data Centre, which is set as an input to the model. Since the area of the inlet in the model may not correspond exactly to the area of the real inlet, this equation ensures that the flow rate across the inlet is equal to $\dot{V}_{CRAH}$, whilst the velocity may deviate from that occurring in the Test Data Centre. It should also be noted that this equation enforces a uniform velocity across the inlet. ϕ at the outlet is calculated in the same way.

ϕ is then calculated at the nodes in the cold aisle. No BCs apply to these nodes, hence ϕ is evaluated using the Laplace equation, discretised via the finite-difference method, as shown in Eq. 5-2. Note that, for each node on the right hand side of the equation, the most recently calculated value of ϕ is used, as in the Gauss-Siedel method [154]. ϕ is then recalculated using the successive over-relaxation method (as described in section 4.4), which helps to speed up convergence of the solution [145], [154]. This process uses Eq. 5-3, where the superscript n refers to the current iteration, $n-1$ refers to the previous iteration, and W_{SOR} is the relaxation factor. Each node is taken in turn, travelling in the positive x and y directions, and the negative z direction.

$$\phi_{i,j,k} = \frac{\phi_{i-1,j,k} + \phi_{i+1,j,k} + \phi_{i,j-1,k} + \phi_{i,j+1,k} + \phi_{i,j,k-1} + \phi_{i,j,k+1}}{6} \qquad \text{Eq. 5-2}$$

$$\phi^n = W_{SOR}\phi^n + (1 - W_{SOR})\phi^{n-1} \qquad \text{Eq. 5-3}$$

Setting w_{SOR} to 1.5 was found to give the fastest solution for most simulations. However in some instances the ϕ field became unstable with $W_{SOR} = 1.5$ (i.e. R_ϕ increased in successive iterations). This was addressed by progressively reducing w_{SOR} where instability was detected in the ϕ field. The method for the detection of instability is discussed at the end of this section (5.1.1.4).

Next, the nodes identified in the BCs array as corresponding to the aisle partition are considered. For a node, (i,j,k), corresponding to the server inlet, the approach is the same as for the Test Data Centre inlet and outlet, and Eq. 5-4 is used. Here, $\dot{V}_{server/LB}$ is the flow rate through the heat loads and N_{server} is the number of nodes comprising the server inlet. $\dot{V}_{server/LB}$ is set as an input to the model.

$$\phi_{i,j,k} = \phi_{i,j,k-1} - \left(\frac{\dot{V}_{server/LB}}{N_{server}D^2}\right)D \qquad \text{Eq. 5-4}$$

For a node, (i,j,k), corresponding to a leakage path, the flow rate is determined according to the local pressure difference across the partition, $\Delta p_{local} = p_{i,j,k-1} - p_{i,j,k}$. The experimental results for bypass described in Chapter 2 are consulted, and the total bypass flow rate, $\dot{V}_{BP}$, which would occur with $\Delta p_{CH} = \Delta p_{local}$ is calculated using Eq. 2-28 for low bypass, and Eq. 2-33 for high bypass conditions. ϕ may then be calculated for the current node, (i,j,k), using Eq. 5-5. Here, $N_{leakage}$ is the number of nodes corresponding to the leakage paths.

$$\phi_{i,j,k} = \phi_{i,j,k-1} - \left(\frac{\dot{V}_{BP}}{N_{leakage}D^2}\right)D \qquad\qquad \text{Eq. 5-5}$$

Where an empty slot is included in the model, the calculation is undertaken in the same way as for the leakage paths, with respect to the relevant experimental results reported in Chapter 2, using Eq. 2-31 instead of Eq. 2-28 or Eq. 2-33.

The remaining nodes in the plane containing the heat load inlets correspond to the partition, which is modelled as a solid wall. Hence, for such a node, (i, j, k), $\phi_{i,j,k} = \phi_{i,j,k-1}$.

ϕ is then calculated for the nodes in the hot aisle. Where no boundary condition applies, this is carried out in the same way as in the cold aisle, using Eq. 5-2. Where the nodes correspond to rack sides or rack roofs, ϕ is set to be equal on each side of the relevant surface, such that there is no flow through it. Similarly, no flow is allowed through the surfaces of the heat loads (other than the inlets and outlets). ϕ at the heat load outlet is calculated in the same way as at the heat load inlet. ϕ at nodes within the heat load is set to 'not a number', such that the heat load inlet and outlet are treated as a sink and a source of ϕ, respectively, with the region within the heat load effectively not forming part of the domain.

The calculations on the ϕ field are completed by calculating ϕ at the remaining nodes falling outside of the Test Data Centre, at the extremities in the x, y and z directions. This is carried out in the same ways as for the nodes with $i = 1$, $j = 1$ and $k = 1$.

The next step is to compute the u, v and w fields, which are the velocities in the x, y and z dimensions, respectively. The nodes in the u, v and w fields coincide with the centres of the $\emptyset$ cell faces in the x, y and z planes, respectively. The equivalent scheme in a 2D field, along with the appropriate indexing scheme, is illustrated in Figure 5-4. Since the velocity fields are given by the rate of change of ϕ with respect to distance, the velocity fields may be approximated as the difference between ϕ at the 2 adjacent nodes divided by the distance between those nodes. For example, with reference to Figure 5-4, $u_{i,j} = (\phi_{i-1,j} - \phi_{i,j})/D$.

The pressure field is then computed using the Bernoulli equation, which can be applied anywhere in the domain in potential flow since losses due to viscosity and vorticity are not considered [157]. For potential flow (i.e. steady, irrotational and inviscid), the Bernoulli equation is given by Eq. 5-6 [157]. Here, p and q are the pressure and scalar velocity, respectively, at the point of interest, and q_0 is the scalar velocity at some reference point with $p = 0$.

$$p = \frac{\rho(q_0^2 - q^2)}{2}$$

Eq. 5-6
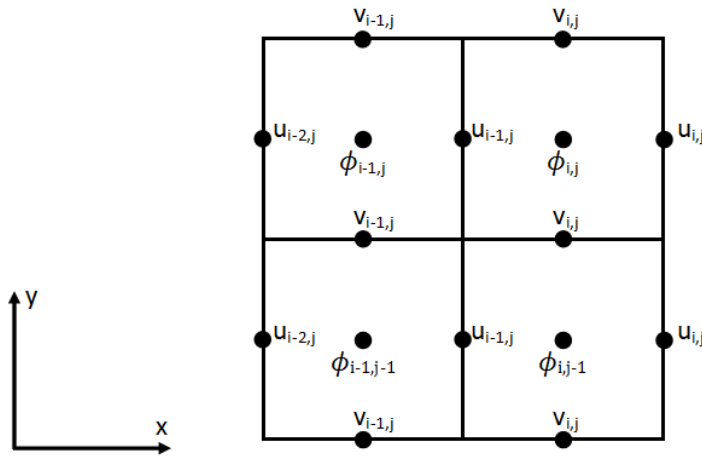
Figure 5-4. Positions of u and v nodes, and the relevant indexing scheme.

The p nodes coincide with the ϕ nodes, so that the magnitudes of u, v and w at each p node must be approximated as the average of the u, v and w values, respectively, at the opposite faces of the cell in question. Since the magnitude of the velocity at a given point is equal to the square root of the sum of the squares of the velocities in 3 perpendicular directions, $p_{i,j,k}$ can be calculated for each node using Eq. 5-7. For nodes in the cold aisle, p_0 is set to p_{CA}, whose calculation is detailed later in this section. For nodes in the hot aisle, $p_0 = 0$. Note that the Bernoulli equation does not apply across the boundary between the cold and hot aisles, since the resistance to flow offered by the leakage paths and sources of momentum offered by the heat loads remove/add energy to the flow, respectively.

$$p_{i,j,k} = p_0 - 0.5\left(\left(\frac{u_{i,j,k} + u_{i-1,j,k}}{2}\right)^2 + \left(\frac{v_{i,j,k} + v_{i,j-1,k}}{2}\right)^2 \right.$$
$$\left. + \left(\frac{w_{i,j,k} + w_{i,j,k-1}}{2}\right)^2\right)$$

Eq. 5-7

The change in pressure in the cold aisle is calculated at the end of each iteration, by noting that, assuming ideal gas behaviour, i.e. $\frac{p}{\rho_{air}} = R_g T$ where R_g is the gas constant, the pressure in the cold aisle must be proportional to the density of air in the cold aisle, which must be proportional to the mass of air in the cold aisle. For the first iteration, the reference cold aisle pressure, p_{CA}, is taken as zero, and the new p_{CA} is calculated using Eq. 5-8, where V_{CA} is the volume of the cold aisle, t_{step} is the time period over which the flows into and out of the cold aisle are calculated. The superscript n refers to the present iteration, and $n - 1$ to the previous iteration. t_{step} is set at the maximum value for which no particle within the

domain could travel a distance greater than the length of the side of a cell. This involves finding the highest absolute value of u, v or w, and setting t_{step} equal to the length of the side of a cell divided by this value.

$$p_{CA}^n = p_{CA}^{n-1}\left(V_{CA} + \left(\dot{V}_{CRAH} - \dot{V}_{BP} - \dot{V}_{servers/LBs}\right)t_{step}\right)/V_{CA} \qquad \text{Eq. 5-8}$$

The final step of the ϕ iteration is to determine whether or not the ϕ field has converged. This is determined by calculation of the ϕ residual, R_ϕ, which is given by Eq. 5-9 (as discussed in section 4.5) [175]. Here, ni, nj and nk are the lengths of the $\emptyset$ array in the x, y and z, respectively, expressed as a number of nodes. Again, the superscript n refers to the present iteration, and $n-1$ to the previous iteration.

$$R_\phi = \frac{\sum_{i=1,j=1,k=1}^{i=ni,j=nj,k=nk}\left(\phi_{i,j,k}^n - \phi_{i,j,k}^{n-1}\right)}{\sum_{i=1,j=1,k=1}^{i=ni,j=nj,k=nk}\left(\phi_{i,j,k}^n\right)} \qquad \text{Eq. 5-9}$$

The ϕ field was judged to have converged when R_ϕ fell below $R_{\phi,target}$, which was set as an input to the model. In every 100th iteration, a test was undertaken to determine whether or not the ϕ field had become unstable. Instability was judged to be present if the sum of R_ϕ for the 50 most recent iterations exceeded the sum of R_ϕ for the previous 50 iterations. Where this was found to be the case, W_{SOR} was reduced by 0.1. However, W_{SOR} was not allowed to fall below unity.

5.1.1.5 Convection of heat

Once the ϕ field has converged, another iterative process begins, which determines the T field.

At the inlet to the Test Data Centre, T is set equal to T_{supply}, which is defined as an input to the model. At the heat load outlet, T is set so as to ensure the conservation of heat. The heat load power consumption, $\dot{E}_{server/LB}$, is assumed to be entirely converted to heat which is convected through the heat load. The mass flow rate of air through the heat load, $\dot{m}_{server/LB}$, is given by the product of the volumetric flow rate, $\dot{V}_{server/LB}$, and the density, ρ_{air}, which is assumed constant at $1.2\ kg.m^{-3}$. The specific heat capacity, $c_{p,air}$, is assumed constant at $1.005\ kJ.m^{-3}.K^{-1}$. The temperature rise across the heat load, ΔT, may then be calculated using Eq. 5-10. Hence for a node (i,j,k) at the server outlet, $T_{i,j,k} = T_{i,j,k_{HL}} + \Delta T$, where node (i,j,k_{HL}) is the corresponding node at the heat load inlet.

$$\dot{E}_{server/LB} = \dot{m}_{server/LB} c_{p,air} \Delta T \qquad \text{Eq. 5-10}$$

The remainder of the T field must be computed by considering the heat transfers impacting on the cell. In addition to convective heat transfers, heat loss through the walls of the hot aisle, as described in section 5.1.1.3, is accounted for. This firstly requires calculating the mean temperature within the hot aisle, T_{HA}, for the current iteration. The rate of heat loss is then calculated for each node as $U_{HA} A_{HA} (T_{HA} - T_{supply})/N_{HA}$, where N_{HA} is the number of nodes in the hot aisle and A_{HA} is the total surface area of the floor, external walls and ceiling of the hot aisle. For the heat transfer calculation, T is assumed not to vary spatially within each cell.

For a steady state solution, the heat transfers into the cell must be equal to the heat transfers out of the cell. This leads to Eq. 5-11 [145]. Here $C_{in/out}$ is the product of $c_{p,air}$ and the mass flow rate into/out of the cell across each face. Note that T_{in} refers to the temperature within the relevant adjacent cell. If T_{out} is taken as the new temperature to be calculated for the current cell, Eq. 5-11 may be re-arranged to solve for this temperature, giving Eq. 5-12. Note that u, v and w are assumed to be constant across the faces of the cells in the x, y and z planes, respectively. Note also that diffusive heat transfers are neglected, as is consistent with the work of Toulouse et al. [145] and Lettieri et al. [170].

$$\sum C_{out} T_{out} - \sum C_{in} T_{in} + \frac{U_{HA} A_{HA} (T_{HA} - T_{supply})}{N_{HA}} = 0 \qquad \text{Eq. 5-11}$$

$$T_{out} = T_{i,j,k} = \frac{\sum C_{in} T_{in} - \dfrac{U_{HA} A_{HA} (T_{HA} - T_{supply})}{N_{HA}}}{\sum C_{out}} \qquad \text{Eq. 5-12}$$

Where available, temperatures from the current iteration are used in this calculation, as is consistent with the Gauss-Siedel method [154]. Otherwise, temperatures from the previous iteration are used.

At this point the T field for the current iteration is compared with that for the previous iteration. If the maximum change in the temperature at a given cell from iteration to iteration, ΔT_{max}, exceeds a specified figure, $\Delta T_{max,target}$, another iteration is undertaken. Otherwise, a solution is judged to have been reached. $\Delta T_{max,target}$ is set as an input to the model, and its specification is discussed in section 5.5.1.

Figure 5-5 summarises the solution process of the PF-CFD model.
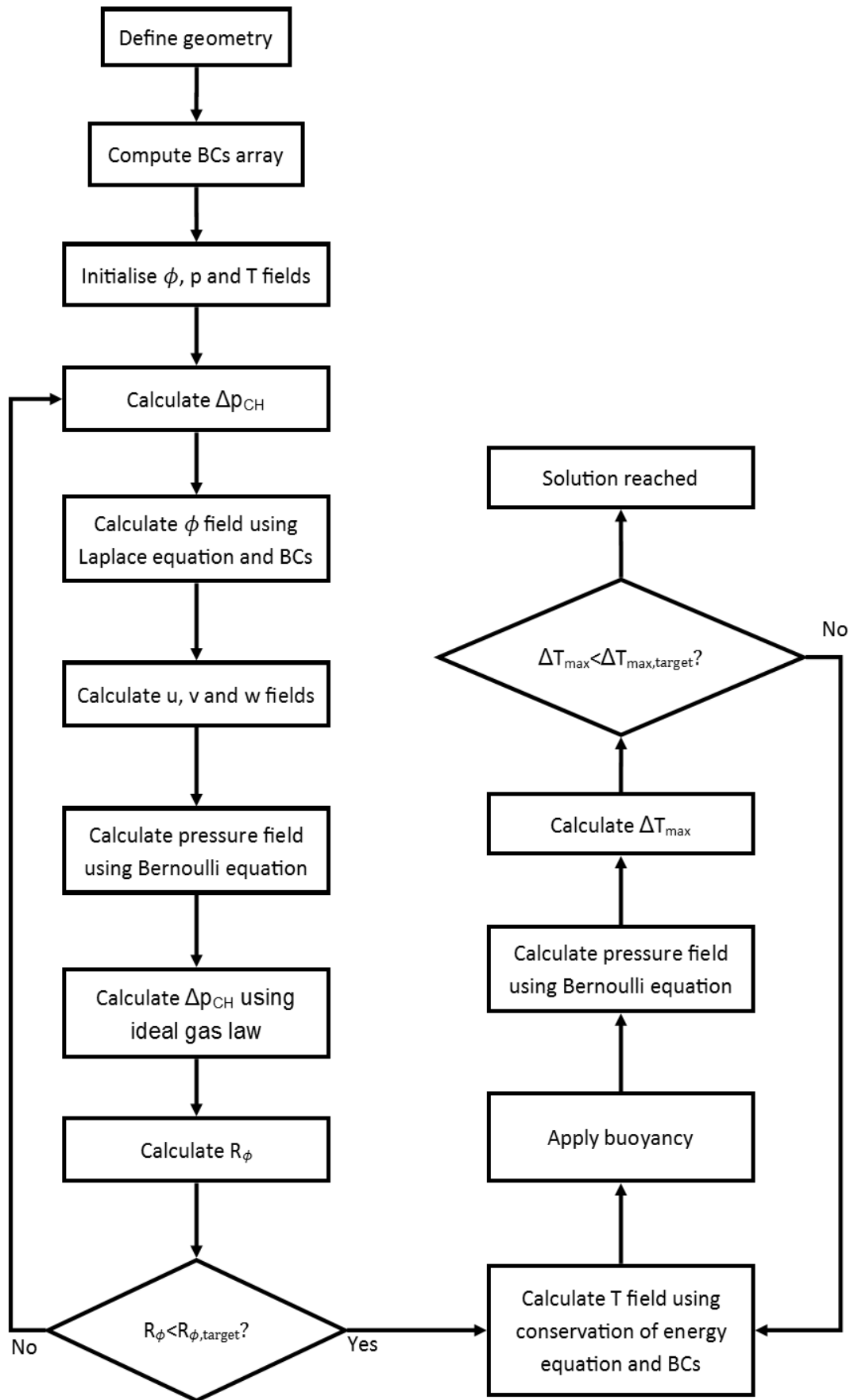
Figure 5-5. Flow chart for PF-CFD model solution process.

5.1.2   Application of the PF-CFD model to an example data centre geometry

The PF-CFD model of the Test Data Centre as described in section 5.1.1 was suitable for validation against the results of experiments described in section 2.2.3.3. This validation is shown in section 5.5.2. In order to enable the model to make predictions of thermal environments in a more typical data centre geometry, some modifications were made.

The geometry of the PF-CFD model was amended to represent a portion of a data centre consisting of aisles each containing 16 racks, each 0.8 m wide, with cold air supplied to and returned from opposite ends of the aisles. Such a data centre is depicted in Figure 5-6, and is based on the design of a commercial data centre which can't be identified due to commercial sensitivity. The domain of the model includes a single row of 16 racks, and half of the depth of each of the adjacent hot and cold aisle. The interior dimensions of the domain are 12.8 x 3.0 x 2.4 m. Figure 5-7 to Figure 5-9 show diagrams of the domain of the expanded PF-CFD model, including the dimensions of the aisles, racks and servers. Each rack was assumed to contain servers occupying a region of height 21 U (0.93345 m), to the nearest cell length, with the bottom of the lowest server being at a height equal to one cell length, D.
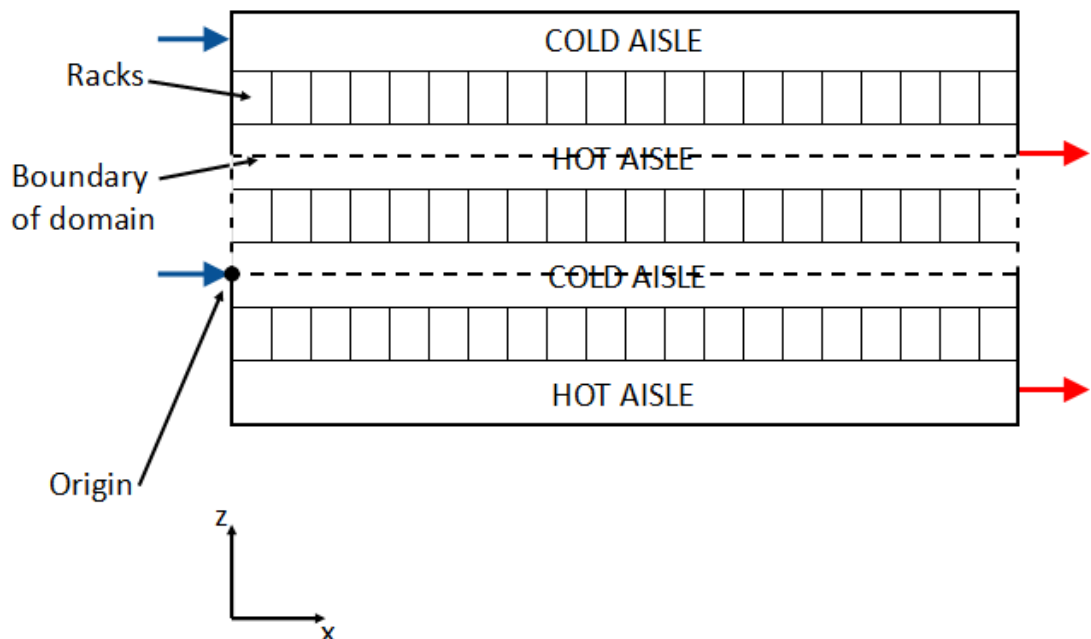
Figure 5-6. Example data centre layout with rows of 16 racks each, with cold air (blue arrows) supplied at the ends of the cold aisles, and hot air (red arrows) returned from the opposite ends of the hot aisles. Note that the vertical dimension is denoted by y, with $y = 0$ coinciding with the floor.
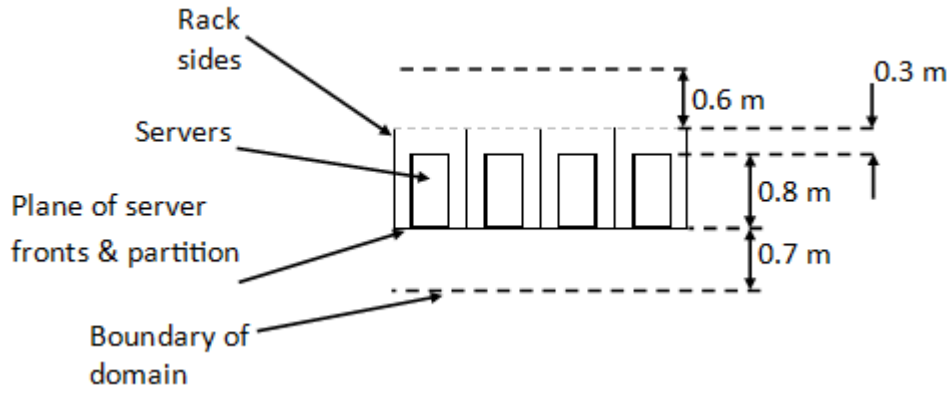
Figure 5-7. Plan view of portion of example data centre.
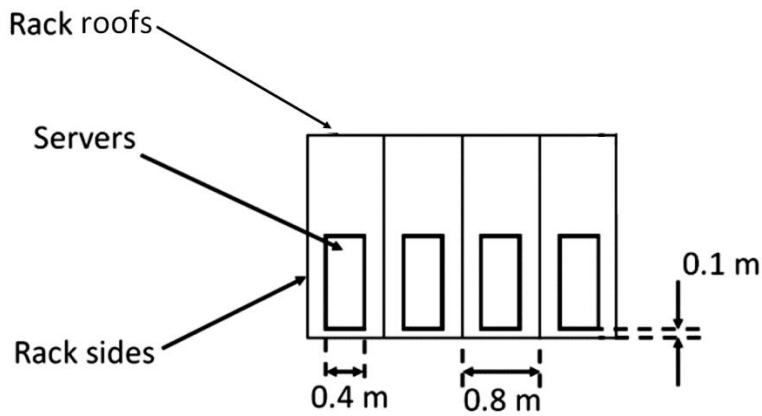
Figure 5-8. Front view of portion of example data centre.
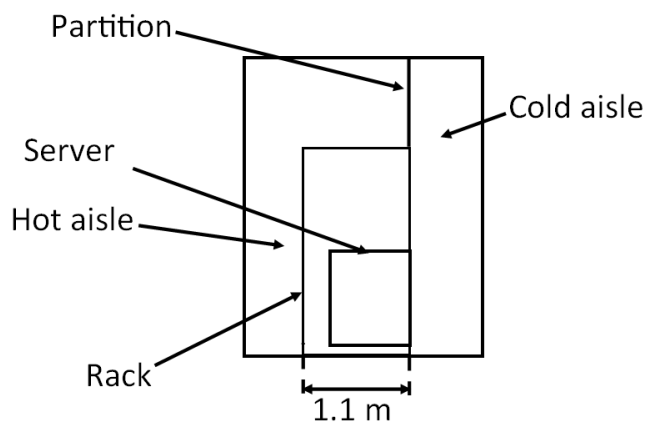
Figure 5-9. Side view of domain of expanded PF-CFD model.

Simulations were run under two bypass scenarios: high bypass and low bypass. As with the PF-CFD model of the Test Data Centre, these two scenarios were governed by Eq. 2-33 and Eq. 2-28, respectively.

The supply air inlet and return air outlet were set to be 0.6 m wide, and to stretch the full height of the domain. The outlet is at the opposite end of the aisle from the inlet, as shown in Figure 5-6.

Since the domain did not incorporate the full widths of the hot and cold aisles, boundary conditions implemented at the boundaries of the domain in the z-dimension were set to give no flow across these boundaries, effectively assuming symmetry across these planes in the full data centre geometry. Hence, in the case of the nodes with $k = 1$, $\phi_{i,j,1} = \phi_{i,j,2}$. Similarly, for the T field, a zero temperature gradient boundary condition was applied at the boundaries of the domain in the z-dimension. This simply required setting $T_{i,j,1} = T_{i,j,2}$.

Δp_{CH} was set as an input to the model, with $\dot{V}_{CRAH}$ set equal to the resultant flow out of the cold aisle at the end of each iteration, i.e. $\dot{V}_{CRAH} = \sum \dot{V}_{server} + \dot{V}_{BP}$.

U_{HA} was set to zero, i.e. the data centre being simulated was assumed to be well insulated.

Each server was assumed to have a height equal to the cell width, D. The power consumption of each server, $\dot{E}_{server}$, was set in accordance with Brady's data for a Sun Fire v20z server [136]. Brady measured power consumption of 0.236 kW with a computational stress applied to the server, and with no pressure gradient applied across the server. Since this describes a 1U server, i.e. 0.04445 m high, the power consumption for each server was scaled according to the height, i.e. $\dot{E}_{server} = \frac{0.236D}{0.04445}\ kW$.

As with the system model, it was assumed that, for a server under a pressure differential $\Delta p_{CH,local} = 0$, the flow rate through the server, $\dot{V}_{server}$, would be sufficient to achieve $\Delta T = 12.5\ K$ (a typical temperature rise according to [41], [55]). Hence, this default flow rate could be calculated using Eq. 5-10, assuming all heat is convected.

The total pressure drop required across the server (Δp_{req}) to achieve $\dot{V}_{server}$ was then calculated using Brady's experimentally derived equation for the system curve of the v20z server, shown in Eq. 5-13 [137], after again scaling this according to the height of the server in the model, and rearranging to give Eq. 5-14.

$$\Delta p_{req} = 38289 \dot{V}_{server}^2 \qquad\qquad\qquad \text{Eq. 5-13}$$

$$\dot{V}_{server} = \frac{236D}{0.04445} \sqrt{\Delta p_{req}/38289} \qquad\qquad\qquad \text{Eq. 5-14}$$

An option was included in the model to have the behaviour of the servers consistent with either Server Fan Option 1 or 2, as described section 3.3. For Server Fan Option 1, the server fan speeds were assumed to modulate in response to the pressure gradient across the server, to ensure $\Delta T = 12.5\,K$. The impact of this modulation on server fan power consumptions was accounted for in the same way as described section 3.3. However, in instances where the maximum server fan power was insufficient to achieve this condition, server power consumption was amended to account for the fan power consumption being at its maximum, and the server flow rate achieved under this condition calculated, and set in the model accordingly. Where the pressure gradient across the server was sufficient to achieve $\Delta T = 12.5\,K$ without the operation of the server fans, the server power consumption was amended to account for the server fans drawing no power, and the server flow rate was set according to Eq. 5-14, after substituting $\Delta p_{CH,local}$ for Δp_{req}.

For Server Fan Option 2, the server fans are assumed to operate at the same speed regardless of $\Delta p_{CH,local}$. Hence, $\dot{E}_{server}$ is unaffected by Δp_{CH}. However, $\dot{V}_{server}$ must be recalculated, which is done using Eq. 5-15, wherein the pressure drop generated by the server fans is assumed to sum with the external pressure. Again this equation is first scaled according to the height of the server in the model. $\Delta p_{CH,local}$ is calculated as the mean pressure drop across the server.

$$\Delta p_{req} + \Delta p_{CH,local} = 38289 \dot{V}_{server}^2 \qquad\qquad\qquad \text{Eq. 5-15}$$

5.2   Navier-Stokes computational fluid dynamics model

5.2.1   Navier-Stokes computational fluid dynamics model of Test Data Centre

An NS-CFD model of the Test Data Centre was developed using Future Facilities' 6SigmaRoom software [178]. This software uses a Reynolds-averaged Navier-Stokes (RANS) solver, with the standard $k - \varepsilon$ model used for turbulence (as described in section 4.3). It operates partially as a 'black box' package, i.e. with the user having limited control over certain aspects of the solution procedure. For example, the user cannot control elements such as the use of upwinding or over-relaxation, and has limited control over the methods

used to discretise the domain. In addition, some details regarding the boundary conditions applied are not made known to the user due to commercial sensitivity.

In the simulations conducted, ρ_{air} was held constant at $1.2 \, kg.m^{-3}$, $c_{p,air}$ was held constant at $1.005 \, kJ.kg^{-1}.K^{-1}$, and the laminar viscosity and thermal conductivity of air were held constant at $1.8 \times 10^{-5} kg.(ms)^{-1}$ and $0.0261 \, W.(mK)^{-1}$, respectively.

The geometry of the model was designed to replicate the Test Data Centre with load banks installed. The internal dimensions of the enclosure were as described in section 5.1.1.1. Since circular inlets/outlets cannot be accommodated in the software, these were replaced with square inlets/outlets having the same area as the respective circular inlets/outlets in the Test Data Centre. The inlet flow was set to have uniform velocity, in a direction normal to the wall. The boundary conditions set at the outlet were not disclosed in the software.

The rack sides were represented as impermeable, two dimensional sheets, measuring 3.2 m in height and 1.192 m in depth, and the rack tops were represented with a single impermeable sheet stretching across the full length of the aisle, with a depth of 1.192 m. In order to limit flow between the cold and hot aisles, an impermeable partition was placed in line with the load bank fronts, with a depth of 0.1 m, occupying the full cross section of the Test Data Centre.

Holes were placed in the impermeable partition, and filled with porous obstructions. These holes and porous obstructions were sized and positioned to represent the regions found to be the most significant leakage paths during the tests described in Chapter 2, i.e. plates at either side of each rack, occupying the full height of the rack, and occupying the full width between the sides of the heat loads/blanking panels and the rack sides. Flow through the porous obstructions was governed by Eq. 2-28 and Eq. 2-33 for the low and high bypass conditions, respectively.

The load banks were represented using the software's generic server module, with the geometry modified to represent the load banks used in the experiments, and with the length increased to include the duct attached to the back of the blocks of load banks. No flow was allowed through the walls of the load banks, with flow allowed at the inlet and outlet through the entire front and back faces of the load banks, respectively. The flow rate and power consumption of the load banks was fixed in the same way as in the PF-CFD model. All heat generated by the load banks was assumed to be convected.

Heat loss was allowed from the walls, floor and ceiling of the Test Data Centre to the environment. The walls, floor and ceiling were set to be maintained at a constant

temperature of $20°C$ (i.e. equal to T_{supply}), and to have very high conductivity, effectively presenting no resistance to heat transfer. The heat transfer coefficient between the walls and the domain was then set to $U_{HA}/A_{NS\ walls}$, where $A_{NS\ walls}$ is the total area of the walls, ceiling and floor in the hot aisle of the NS-CFD domain. Thus, the heat loss found to occur from the Test Data Centre in the experiments detailed in section 2.2.3.3 was incorporated.

5.2.2   Application of the NS-CFD model to an example data centre geometry

As with the PF-CFD model, modifications were made to the NS-CFD model to enable it to make predictions of thermal environments in a more typical, example data centre geometry. The geometry of the room, the inlets and outlets, and the positions and dimensions of the racks were as described for the PF-CFD model in sections 5.1.1.1 and 5.1.1.2. The flow rates through the servers and leakage paths were controlled by the same equations as used in the PF-CFD model, and the server power consumptions were the same as in the version of the PF-CFD model with Server Fan Option 2 implemented. It was not possible to implement Server Fan Option 1 in the NS-CFD model. As described in section 3.3.1.1, in Server Fan Option 1, server power consumption varies in response to changes in Δp_{CH}. Server power consumption could not be made to vary with respect to pressure drop across the server within the software used. Accordingly, Server Fan Option 2 was used for all NS-CFD simulations of the example data centre geometry.

5.3   A modified system model

A modified version of the system model described in Chapter 3 was created to enable the impact of Δp_{CH} and bypass conditions on total electricity consumption, E_T, to be predicted, with consideration of the results of the CFD models of the example data centre geometry.

The server flow rates and power consumptions, and $\dot{V}_{CRAH}$, were set as inputs to the model, using the results of the CFD models. T_{supply} was set in the model such as to ensure that the mean inlet temperature of any given server did not exceed 27°C. Hence T_{supply} was set equal to $27 - T_{server,max}$, where $T_{server,max}$ is the highest mean server inlet temperature recorded in the respective CFD simulation. The ambient temperature, T_{amb}, was set to $11°C$. The power consumptions of the CRAH, chilled water pumps and economiser were calculated in the same way as described in section 3.3.

5.4   Initial validation of potential flow methodology

Prior to developing a potential flow model representing the geometry of the Test Data Centre, a potential flow model replicating the geometry of Toulouse et al.'s model [145], discussed in Chapter 4, was developed with the aim of validating the methods used in the Test Data Centre model. The model created is referred to as TPFR (Toulouse Potential Flow Replica) for ease of reference. The only difference between Toulouse et al.'s model and the TPFR model was that in the former, convection of heat was computed only for nodes which were found to be downstream of heat-generating nodes. This served to reduce the number of computations required to reach a solution. By comparison, in the TPFR model, convection of heat was computed for every node within the domain. The methods used in the TPFR model are as described in section 5.1.1, with the geometry and other input parameters set so as to be consistent with Toulouse et al.

Figure 5-10 and Figure 5-11 show comparisons of Toulouse et al.'s [145] results with the results of the TPFR model. Note that Toulouse et al. did not publish numerical results. However, a visual analysis of Figure 5-10 and Figure 5-11 reveals very good agreement between the two models. This demonstrates that the implementation of the physics in the PF-CFD model of the Test Data Centre described in section 5.1.1 is consistent with the work of Toulouse et al.

(a)

(b)

(c)
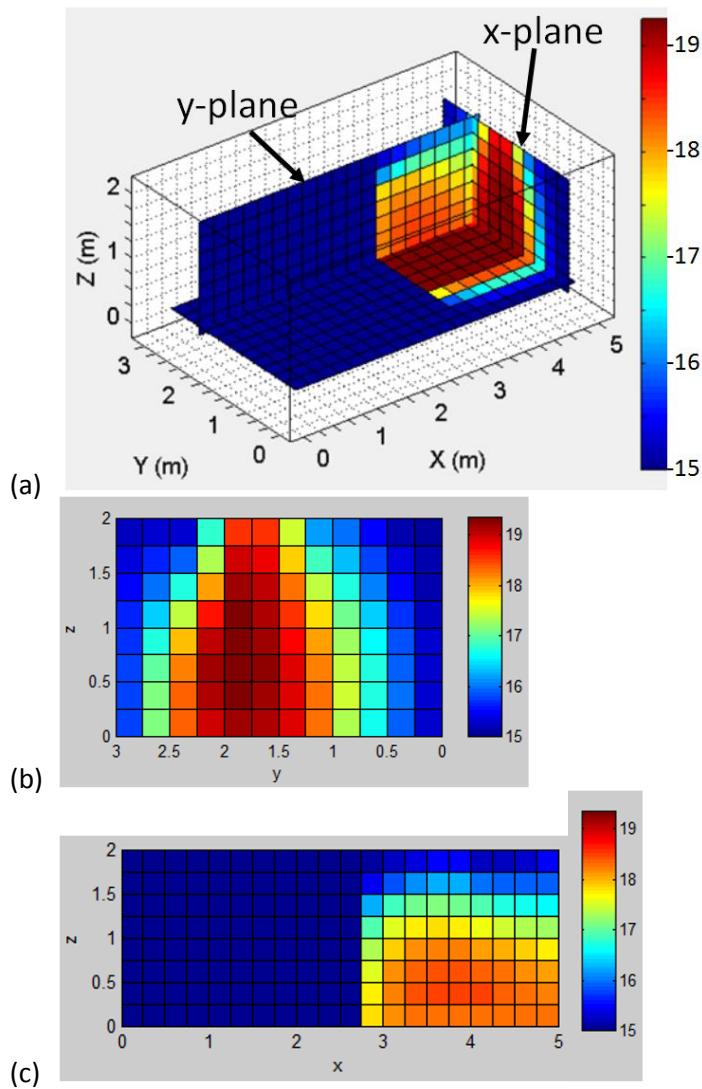
Figure 5-10. Temperature planes in $°C$ from (a) Toulouse et al. [145] and (b, c) TPFR. The results from the TPFR model show planes at (b) constant x and (c) constant y, at the same positions as the constant x and y planes shown in the image from Toulouse et al.

(a)

(b)
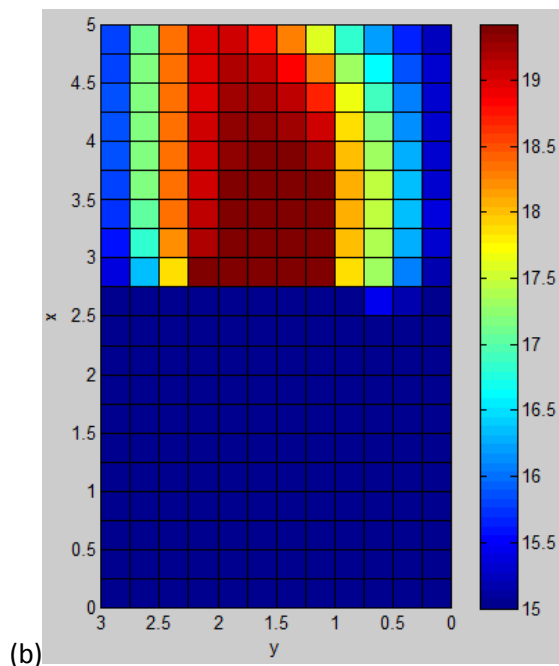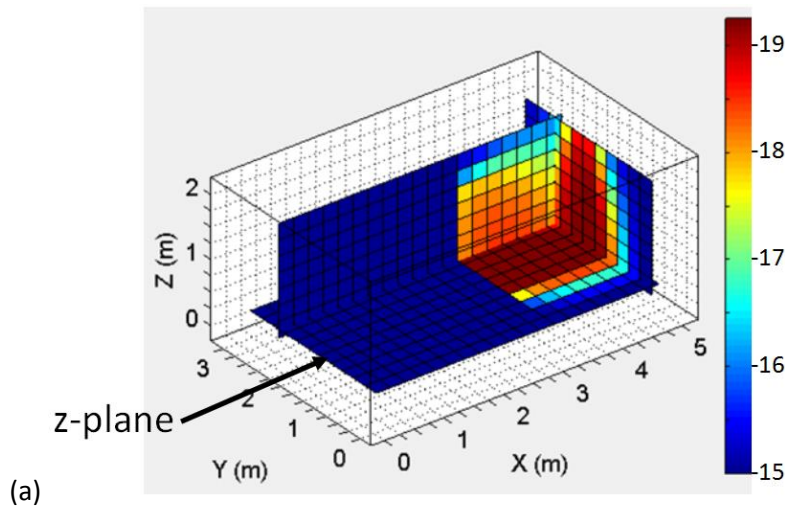
Figure 5-11. Temperature planes from (a) Toulouse et al. [145] and (b) TPFR. The results from the TPFR model show a plane at constant z, at the same position as the constant z plane shown in the image from Toulouse et al.

5.5 Results – Test Data Centre

5.5.1 Verification of PF-CFD model

A series of simulations were run using the PF-CFD model of the Test Data Centre, whilst varying $R_{\phi,target}$, $\Delta T_{max,target}$ and D. Note here that $R_{\phi,target}$ is the input parameter used to define the magnitude of residuals in the velocity potential field at which convergence was assumed to have been achieved (see section 5.1.1.4), $\Delta T_{max,target}$ is the maximum change in T at any node from iteration to iteration at which convergence was judged to have been achieved (see section 5.1.1.5), and D is the cell side length. These simulations enabled

appropriate values of $R_{\phi,target}$, $\Delta T_{max,target}$ and D to be selected, ensuring that the results were independent of these parameters, whilst minimising time to solution. The geometry and BCs were set to be consistent with Test 1 (as described in section 2.3.2.2). Except where otherwise indicated, $D = 0.1\ m$, $R_{\phi,target} = 10^{-6}$ and $\Delta T_{max,target} = 0.005\ K$.

Figure 5-12 to Figure 5-14 show the impact of D, $R_{\phi,target}$, and $\Delta T_{max,target}$, respectively, on temperature predictions and time to solution. Temperatures are shown in the form of 'temperature rise', i.e. the difference between the measured temperature and T_{supply}. Temperatures are shown at a selection of positions at the load back inlets (positions LBIe and LBIl) and in the hot aisle (positions HAi and HAp), with positions as defined in Figure 2-19 and Figure 2-17. The full analysis incorporated not only the measurement points shown in the figures, but all 24 measurement points in the hot aisle.

The simulations with varying D represent a grid independence study for the PF-CFD model. A full analysis of the predicted temperatures at each of the load bank inlets in each simulation found that reducing D from 0.133 to 0.1m reduced the mean inlet temperature by 1.4 K, whereas reducing D from 0.1 to 0.057m incurred changes in the mean load bank inlet temperature of less than 1 K. The maximum variation for a single load bank inlet was 3.1 K with D reducing from 0.133 to 0.1 m, and 1.9 K when reducing D from 0.1 to 0.057 m. Hot aisle temperature predictions were more variable in absolute terms, although all varied by less than 15% of the minimum temperature rise recorded in the hot aisle with D ranging from 0.067 to 0.133m. Figure 5-12 shows that the solution time increased by 2 orders of magnitude over the range of D considered. Preliminary simulations using the PF-CFD model of the example data centre geometry found solution times of around 16 hours and around 46 hours for cell side lengths of 0.1m and 0.08m, respectively. Hence, it was decided that to give a reasonable degree of grid independence with an acceptable solution time, D would be set to 0.1 m for the remainder of the simulations. Note that all simulations were run on a desktop computer with a processor speed of 3 GHz, and with 8 GB of RAM.

The maximum absolute change in load bank inlet temperature incurred by reducing $R_{\phi,target}$ from 10^{-5} to 10^{-6} was 2.6 K, with the maximum change as a percentage of minimum load bank inlet temperature rise being 29%. When reducing $R_{\phi,target}$ from 10^{-6} to 10^{-7}, the maximum changes were 0.06 K and 0.7%, respectively. For the hot aisle temperatures, there were significant changes in temperature predictions when reducing $R_{\phi,target}$ from 10^{-5} to 10^{-6} (up to 67% change in temperature rise). When reducing $R_{\phi,target}$ from 10^{-6} to 10^{-7}, the maximum variation was just 0.2%. Figure 5-13 shows that solution time is less strongly affected by reducing $R_{\phi,target}$ than by reducing D, as it

increases by less than a factor of 2 whilst $R_{\phi,target}$ fell by 2 orders of magnitude. Hence, it was decided that $R_{\phi,target}$ should be set to 10^{-6} in order to give a good degree of independence from this factor, with minimal penalty in terms of solution time.

The maximum absolute change in load bank inlet temperatures incurred by reducing $\Delta T_{max,target}$ from 0.5 to 0.05 K was 1.1 K, with the maximum change as a percentage of minimum load bank inlet temperature rise being 12.7%. When reducing $\Delta T_{max,target}$ from 0.05 to 0.005 K, the maximum changes were 0.12 K and 1.4%, respectively. For the hot aisle temperatures, the maximum absolute and percentage changes when reducing $\Delta T_{max,target}$ from 0.5 to 0.05 K were 0.8 K and 2.4%. When reducing $\Delta T_{max,target}$ from 0.05 to 0.005 K, the maximum changes were 0.003 K and 0.008%. Figure 5-14 shows that solution time does not have a clear negative correlation with $\Delta T_{max,target}$, as would be expected. It must be concluded that natural variation in solution time due to other processes being carried out by the computer performing the simulations over-ride the impact of changes in this variable. Hence, it was decided that $\Delta T_{max,target}$ should be set to $0.05\ K$ in order to give a good degree of independence from this factor, with minimal penalty in terms of solution time.
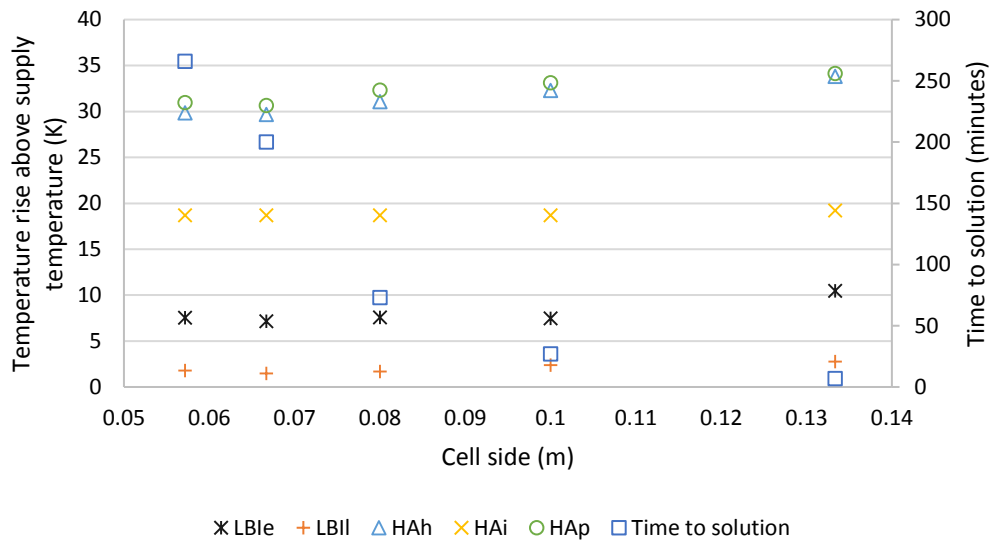
Figure 5-12. Grid independence study for PF-CFD model of Test Data Centre. Results show temperature rise at a selection of points at the load bank inlets (LBIe and LBII) and in the hot aisle (HAh, HAi and HAp) (with positions as defined in Figure 2-19 and Figure 2-17).
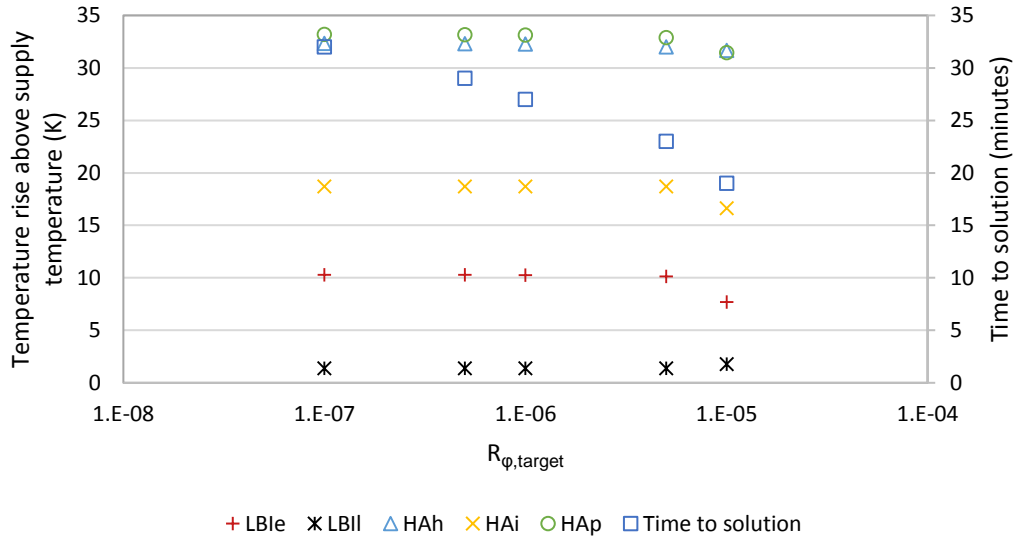
Figure 5-13. Impact of $R_{\phi,target}$ on temperature field predicted by PF-CFD model of Test Data Centre. Temperature rise shown at a selection of points at the load bank inlets (LBIe and LBIl) and in the hot aisle (HAh, HAi and HAp) (with positions as defined in Figure 2-19 and Figure 2-17).
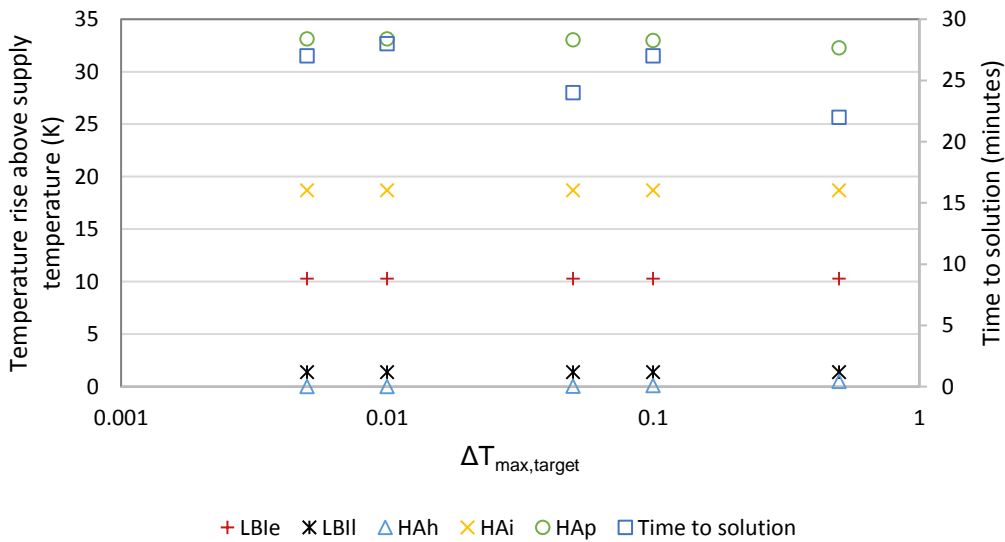
Figure 5-14. Impact of $\Delta T_{max,target}$ on temperature field predicted by PF-CFD model of Test Data Centre. Temperature rise shown at a selection of points at the load bank inlets (LBIe and LBIl) and in the hot aisle (HAh, HAi and HAp) (with positions as defined in Figure 2-19 and Figure 2-17).

5.5.2   Verification of NS-CFD model

A series of simulations were run using the NS-CFD model, with conditions replicating Test 1 as described in section 1. while varying the number of cells and the 'Termination Factor'. This Termination Factor is described in the software's help documentation as determining "the degree of acceptable numerical error", with a Termination Factor of 1 being the default, and 0.5 halving the acceptable level of numerical error. It may be assumed that the Termination Factor refers to the values of the residuals, as described in section 4.5. For these simulations, the temperatures predicted at the 24 measurement points in the hot aisle (as depicted in Figure 2-17) and at the 6 load bank inlets (positions LBIe, LBIf, LBIg, LBIj, LBIk and LBIl, as defined in Figure 2-19) were compared.

The simulations undertaken with varying numbers of cells represent a grid independence study for the NS-CFD model. For these simulations, the Termination Factor was set equal to 1. A selection of the temperature predictions are shown in Figure 5-15. The number of cells in the grid may be controlled by setting a minimum cell side length, with the software then producing a grid automatically. Increasing the number of cells from 396,890 to 577,885 was found to change hot aisle temperatures by up to 1.6 K, and inlet temperatures by up to 0.3 K. Increasing the number of cells from 577,887 to 746,982 changed inlet temperatures by a maximum of 0.2 K, and hot aisle temperatures by a maximum of 0.4K. Figure 5-15 also shows that the time to solution rose by a factor of 1.5 with the increase in cells from 577,887 to 746,982. Hence it was concluded that 577,887 cells was sufficient to provide a reasonable degree of grid independence, whilst minimising solution time. This grid corresponded to the imposition of maximum cell side lengths of 0.045 m, and this maximum cell side length was used for all future simulations. Note that the simulations were carried out on the same computer as the PF-CFD simulations, having a processor speed of 3 GHz, and with 8 GB of RAM.

A selection of the results of the simulations with varying Termination Factor are shown in Figure 5-16. With Termination Factor ranging from 4 to 2, load bank inlet temperature predictions changed by a maximum of 0.1 K, and hot aisle temperature predictions by a maximum of 2.2 K. These maximum changes fell to 0.09K and 1.7K when reducing Termination Factor from 2 to 1. Reducing the Termination Factor from 2 to 1 increased the solution time by a factor of 1.4. Hence it was concluded that a Termination Factor of 2 was sufficient to produce accurate results whilst minimising time to solution. Hence, the Termination Factor was set to 2 for future simulations.
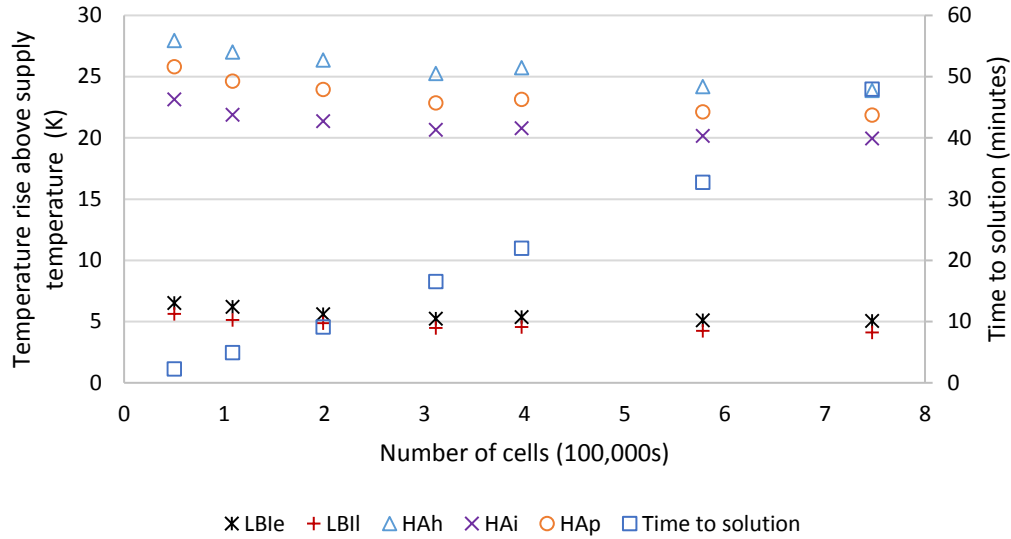
Figure 5-15. Effect of number of cells on temperatures predicted by NS-CFD model of Test Data Centre at a selection of load bank inlet and hot aisle measurement points, as defined in Chapter 2.
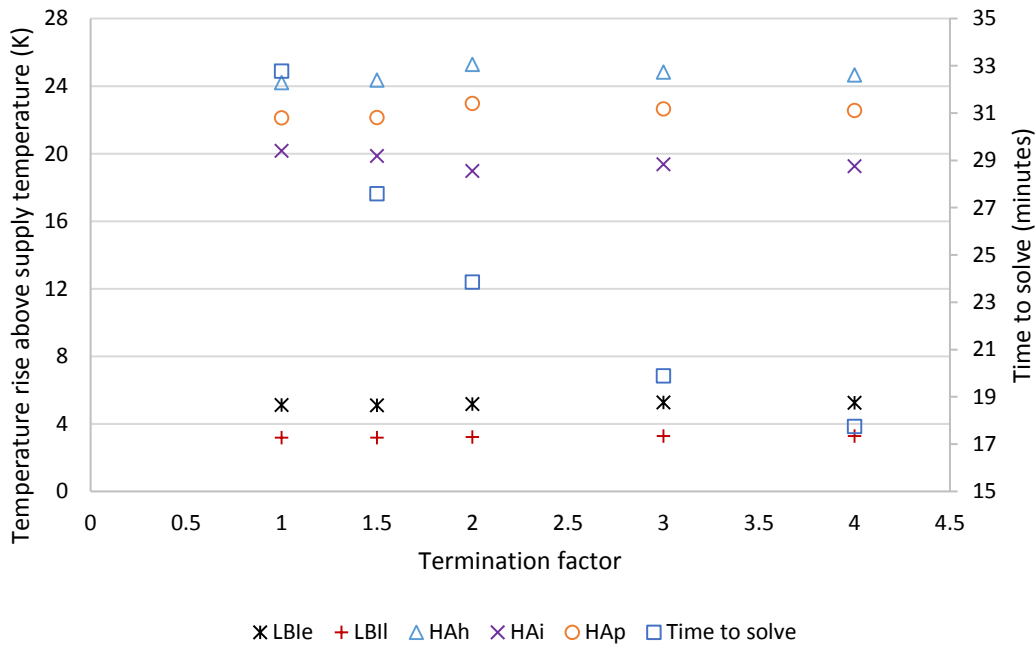
Figure 5-16. Effect of Termination Factor on temperature rise at load bank inlet predicted by NS-CFD model of Test Data Centre.

5.5.3 Validation of the CFD models against results from the Test Data Centre

Having determined appropriate parameters to ensure an acceptable level of independence from grid density and convergence parameters, the predictions of the two CFD models of

the Test Data Centre could be compared with the results of Tests 1 and 2 (as described in section 2.2.3.3). In this section, the extent to which the predictions of the two models correlate with the experimental results is discussed. In section 5.5.4, the discrepancies between the predictions of the two models will be discussed.

The PF-CFD model of the Test Data Centre was firstly run with the input conditions set to be consistent with Test 1. $R_{\phi,target}$ was set to 10^{-6} and $\Delta T_{max,target}$ was set to 0.05 K. Similarly, the NS-CFD model was run with the input conditions set to be consistent with Test 1, with the Termination Factor set to 2, and with cell side lengths limited to 0.045 m.

Figure 5-17 shows the predicted and measured temperature rises for Test 1 at a selection of load bank inlet and hot aisle measurement positions. The error bars on the experimental measurements represent 11.6% of the temperature rise for each measurement, which was found to be the maximum error in repeat measurements, as discussed in section 2.4.1. An analysis of the differences between the predicted and measured temperatures at all load bank inlet and hot aisle measurement positions was also undertaken. The results for all measurement positions are shown in Appendix 3.

Both simulations predicted the mean load bank inlet temperatures to within 0.5 K of the mean temperatures in the experimental results. The simulations showed greater variation in temperatures at the load bank inlets than was recorded in the experiment. Both models show higher temperatures at the top and left and lower temperatures and the bottom and right hand side. However, these trends were not recorded in the experiments. This suggests that both models underestimate the extent of mixing. The NS-CFD model predicts all inlet temperatures to within 1.5 K of the measured value, whereas the PF-CFD model predicts all inlet temperatures to within 2.5 K of the measured value.

Table 5-1 summarises the data for the hot aisle measurement points. Both simulations significantly overestimate the variation in hot aisle temperatures. The ranges of hot aisle temperature rises predicted by the PF- and NS-CFD models are 18.7 to 33.7 K, and 17.4 to 29.0 K, respectively. This compares with a range of 15.5 to 22.6 K in the experimental results. This again shows that both simulations underestimate the extent of mixing in the hot aisle. This could result from the assumption of uniform velocity profile at the load bank outlet. In reality, the flow leaving the load bank is likely to include a considerable swirl component, which would increase mixing. The lack of turbulence in the PF-CFD model could also contribute to this error.

Both simulations predict higher mean hot aisle temperatures than were recorded in the experiment, although the NS-CFD simulation result is much closer to the experimental data than the PF-CFD simulation, as shown in Table 5-1. The maximum errors on hot aisle temperature rise predictions were 15.5 and 6.4 K for the PF- and NS-CFD simulations, respectively. The errors of these predictions show that the models have limited value in predicting temperatures in this part of the domain.
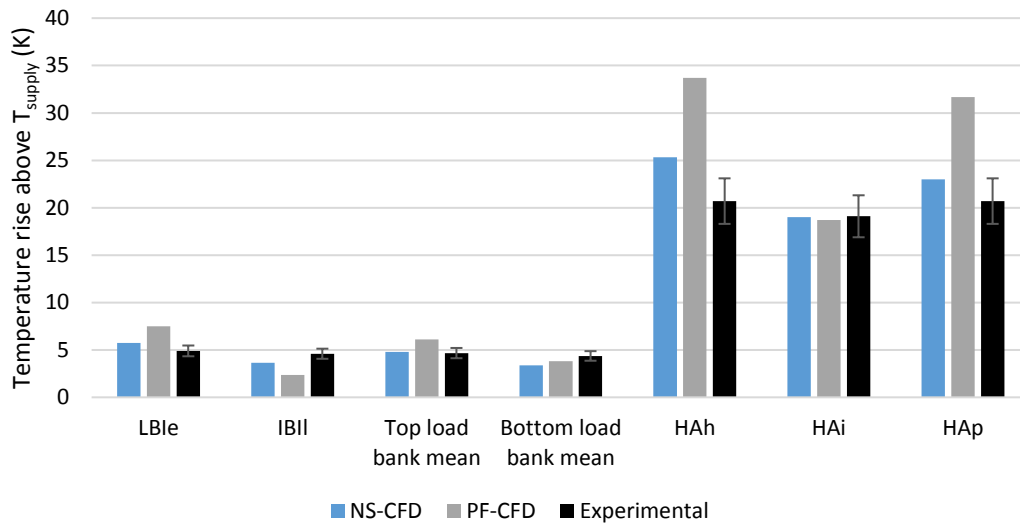
Figure 5-17. Comparison of predicted and measured temperature rise at a selection of measurement points at the load bank inlets (LBle and LBll) and within the hot aisle (HAh, HAi and HAp), as defined in Figure 2-17 and Figure 2-19, for Test 1.

Data set	Temperature rise from T_{Supply} (K)		
	Minimum	**Maximum**	**Mean**
PF-CFD	18.7	33.7	25.9
NS-CFD	17.4	29.0	21.0
Experimental	15.5	22.6	19.6

Table 5-1. Summary of predicted and measured temperature rise data at hot aisle measurement points for Test 1.

Figure 5-18 shows the predicted and measured temperature rises for Test 1 at a selection of measurement positions in the cold aisle and within rack 4 (the rack containing the load banks). The error bars on the experimental measurements again represent 11.6% of the

temperature rise for each measurement. An analysis of the differences between the predicted and measured temperatures at all measurement points in the cold aisle and in rack 4 was also undertaken.

The cold aisle temperature rises predicted by the PF-CFD model were all below 1.5 K, which is well below the range of 2.8 to 5.0 K recorded during the experiments. This is due to the concentration of hot air in the upper part of the cold aisle in the simulations (discussed later in this section in relation to Figure 5-22). The results of the NS-CFD simulation show temperature rises between 7 and 8 K, much higher than those recorded in the experiments. Again, this is due to the concentration of higher temperatures in the upper part of the cold aisle, with the high temperature region coinciding with the measurement points in the NS-CFD simulation, but not in the PF-CFD simulation. Since, as shown in Figure 5-17, all simulations predict load bank inlet temperatures fairly accurately, it may be concluded that whilst the simulations predict recirculation accurately, the mixing of air in the cold aisle which happens in reality (as demonstrated by the relatively consistent measured cold aisle temperatures) is underestimated, particularly in the PF-CFD model.

The NS-CFD model's predictions of temperature rises within rack 4 correlate well with the experimental results, with the maximum error being 3.2 K. Temperatures at the rack sides are accurately predicted to be fairly uniform, with temperatures in the upper part of the rack being higher. The PF-CFD simulation significantly overestimates all temperatures within the rack, other than the highest measurement points at the sides of the racks (R4a and R4e), which are predicted to within 2 K of the measured value. The remaining measurements are overestimated by 3 to 17 K in the PF-CFD model.

The results again show that the PF-CFD model underestimates the extent of mixing. This causes the hot air exiting the load banks to flow back down the sides of the load banks in much more concentrated form than was displayed in the experiments, leading to the discrepancy between measured and predicted temperatures at measurement points within the upper part of the rack (R4h and R4i). The tendency in the NS-CFD model for hot air exiting the load bank to travel in a jet in the z direction (as discussed later in this section in relation to Figure 5-23) helps to prevent this from happening in this model.
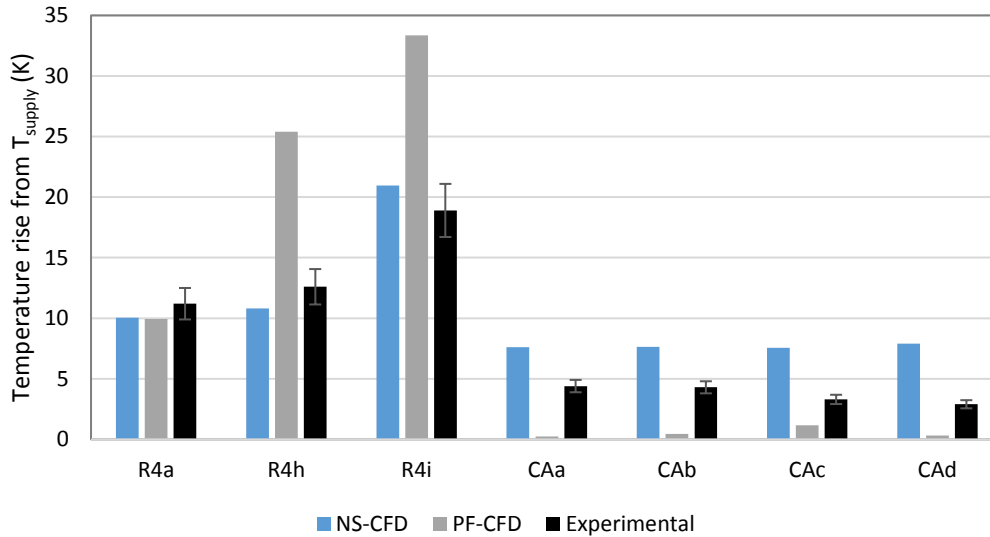
Figure 5-18. Comparison of predicted and measured temperature rise at a selection of measurement points within rack 4 (R4a, R4h and R4i) and in the cold aisle (CAa-CAd), as defined in Figure 2-16 and Figure 2-18, for Test 1.

The PF- and NS-CFD models were then run with the input conditions set to be consistent with Test 2, as described in 1. Note that for this test, the slot immediately above the load banks was left open. For the PF-CFD model, $R_{\phi,target}$ was again set to 10^{-6} and $\Delta T_{max,target}$ was set to 0.05. For the NS-CFD model, the Termination Factor was set to 2, and the maximum cell side length was set to 0.045 m. The empty slot was modelled in the NS-CFD model simply by leaving a hole in the impermeable partition separating the hot and cold aisles, with a height of 0.04445m (1U), positioned immediately above the load bank.

Figure 5-19 compares predicted and measured temperature rises at a selection of load bank inlet and hot aisle measurement points for Test 2. Figure 5-20 compares predicted and measured temperature rises at a selection of measurement points within the cold aisle and within rack 4. The temperature rises at the full set of measurement positions are shown in Appendix 3.

Both models underestimated the mean temperature rise at the load bank inlets - the PF-CFD model by 1.5 and NS-CFD by 2.1 K. However, both models successfully predict higher inlet temperatures for Test 2 than for Test 1, and predict higher temperatures at the upper inlets than the lower inlets, as was recorded in the experiments. This shows that the models are broadly successful in predicting the impact of the introduction of the empty slot on load bank inlet temperatures. This is important since prediction of inlet temperatures is the most important function of a data centre air flow model. The PF- and NS-CFD models predict mean temperatures for the 3 upper inlets 3.8 and 4.4 K less than the measured mean,

respectively. This suggests that the models both underestimate the local recirculation caused by the empty slot.

For the hot aisle, cold aisle, and rack 4 temperatures, the discrepancies between the predictions and the experimental results are very similar to those reported for hot aisle temperatures for Test 1. This is likely to be due to the same issues causing these discrepancies with Test 1. The maximum errors on hot aisle temperature rise predictions were 23.4 and 17.5 K for the PF- and NS-CFD simulations, respectively.
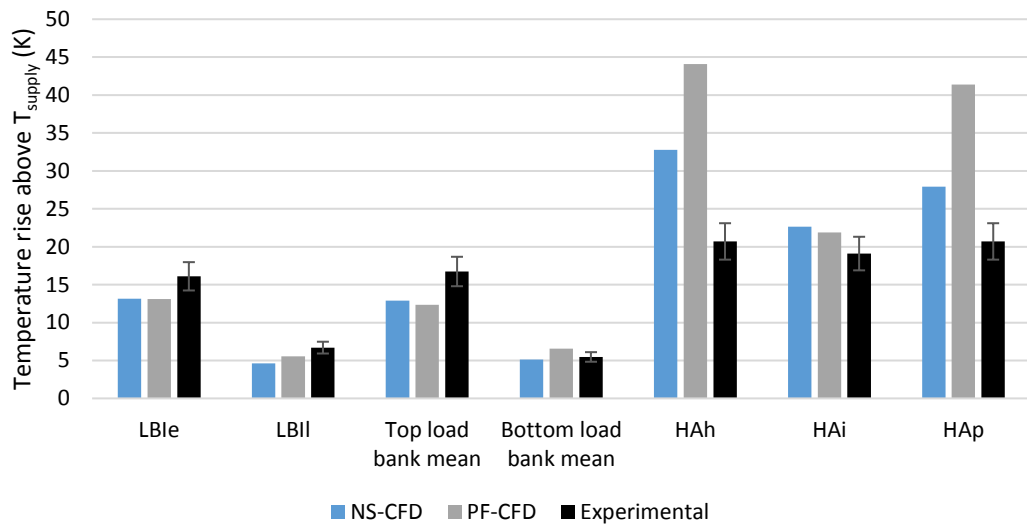
Figure 5-19. Comparison of predicted and measured temperature rise at a selection of measurement points at the load bank inlets (LBIe and LBII) and within the hot aisle (HAh, HAi and HAp), as defined in Figure 2-19 and Figure 2-17, for Test 2.
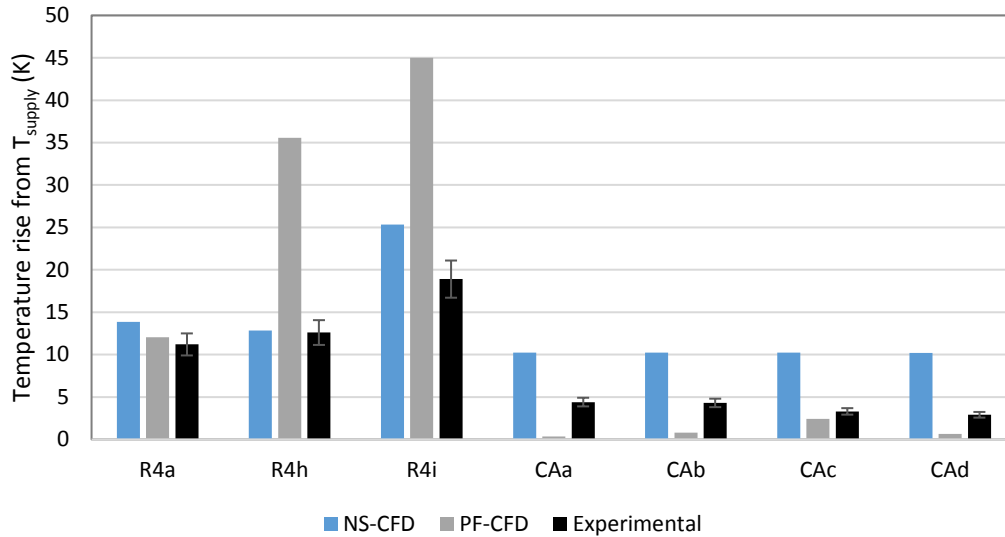
Figure 5-20. Comparison of predicted and measured temperature rise at a selection of measurement points within rack 4 (R4a, R4h and R4i) and within the cold aisle (CAa to CAd), as defined in Figure 2-16 and Figure 2-18, for Test 2.

To summarise, the results presented in Figure 5-17 to Figure 5-20 show that the PF- and NS-CFD models are both quite successful at predicting the temperature rise at the load bank inlets, with all predictions for mean inlet temperatures for individual load banks being within 3.5 K of measured values. Both models successfully predict a significant increase in inlet temperatures resulting from the introduction of an empty slot.

In Chapter 4, the method of assessing a model's accuracy purely by reporting the magnitude of errors in inlet temperature predictions was criticised for its failure to consider the relative importance of errors in comparison with the range of measured temperatures at server or rack inlets. An alternative approach was devised, namely dividing the maximum error in temperature predictions (ε_T, K) by the total range of temperatures measured at the inlets (T_{range}, K). Applying the same approach to this validation gives $\frac{\varepsilon_T}{T_{range}} = 0.39$ for the PF-CFD model, and $\frac{\varepsilon_T}{T_{range}} = 0.34$ for the NS-CFD model. Here, ε_T is the maximum error on the predictions of mean inlet temperature for each load bank across Tests 1 and 2, and T_{range} is given by difference between the highest and lowest measured mean load bank inlet temperature across Tests 1 and 2. These figures compare favourably with the most accurate models reported in the literature, as summarised in Table 4-1. Only Lopez & Hamann [150] have achieved significantly greater accuracy, and could only do so if significant quantities of measured temperature data were used to guide the model.

Even so, it should be noted that the magnitudes of the errors in inlet temperature prediction are significant in comparison with the ASHRAE recommended server inlet temperatures, which cover a range of 18 to 27 K [179]. For example, a predicted mean inlet temperature of 26 K, i.e. within the recommended range, could correspond to an actual temperature of 29.5 K, i.e. outside of the recommended range, for an error of 3.5 K (the maximum recorded error). These errors are significant, although it should be noted that both models succeeded in predicting the trend of increased temperatures for Test 2 in comparison with Test 1 (see Figure 5-17 and Figure 5-19), and correctly showed higher inlet temperatures for the upper than for the lower load bank in Test 2 (see Figure 5-19). Since predictions of inlet temperatures showed errors with higher magnitudes for Test 2 than for Test 1, it may be said that the predictions worsened where the integrity of the containment system was breached (since, for Test 2, an empty slot was introduced above the load banks). This is consistent with the findings of the literature review presented in section 4.6.2.

The models are much less successful at predicting temperatures elsewhere in the domain. In the hot aisle, the maximum discrepancies between simulation and experimental results were 17 K for the NS- and 23.4 K for PF-CFD model, with far greater variation in hot aisle temperatures being predicted by the simulations than was recorded during the experiments. The assumption of uniform velocity profile at the load bank outlets has been cited as one potential reason for this. It may also be that the modelling of turbulence fails to replicate the true extent of mixing in the hot aisle.

5.5.4   Comparing the predictions of the two CFD simulations of the Test Data Centre

Figure 5-21 to Figure 5-23 show the temperature rise (i.e. the increase in comparison with T_{supply}) for various planes through the Test Data Centre, for the two CFD models, in the form of colour plots. The experimental results are also displayed on the plots in numerical form – note that a full comparison of experimental data with simulation results is shown in column graph form in Appendix 3.

The temperature fields within the hot aisle differ considerably between the NS- and PF-CFD models. The NS-CFD model shows a distinct jet of hot air exiting the load bank (see Figure 5-23). Outside of this region, and above the height of the bottom of the load bank, the temperature rise is fairly uniform, ranging from 21 to 28 K outside of the racks. By contrast, the PF-CFD model shows the hottest temperatures concentrated at the end of the aisle opposite racks 3 and 4, and below the level of the rack roofs. Within these areas, the PF-CFD model predicts temperature rises of 25 to 42 K, whereas within the remainder of the

hot aisle, temperature rises are around 18 to 20 K. When referring to the experimental results in each of these figures, it is clear that the more uniform hot aisle temperatures predicted by the NS-CFD model are closer to the experimental results, which show hot aisle temperature rises ranging from 15.5 to 22.6 K.

The differences in the temperature field in the hot aisle likely relate to the different implementations of buoyancy and heat loss to the surrounding environment in the two models. The lack of buoyancy in the PF-CFD model allows the hot air to collect in the lower part of the aisle, whereas the inclusion of buoyancy in the NS-CFD model causes the hot air leaving the load bank to rise, inducing buoyant mixing. As described in section 5.1.1.5, heat loss from the hot aisle to the surrounding environment is incorporated into the PF-CFD model in the form of a uniform heat removal from each cell, as a function of the difference between the supply temperature and the mean hot aisle temperature. In reality, regions closer to the walls would be more heavily impacted by this heat loss. This causes the lower temperatures in the unoccupied racks.

Another key factor could be the behaviour of air leaving the load banks in the two models. Whilst the NS-CFD model shows a distinct jet of hot air leaving the load bank, the PF-CFD model shows this heat being dissipated in all directions (see Figure 5-23). This results from the lack of viscosity and conservation of momentum in the PF-CFD model, which allows unrealistically abrupt changes of direction in airflows [169]. The lack of turbulence in the PF-CFD model is also likely to limit the extent of mixing, leading to less uniform temperatures.

Calculation of the Reynolds number, Re, is useful in estimating the importance of turbulence, as discussed in section 4.3 [158]. Re is given by $Re = \rho_{air} q_C L_C / \mu$ [163], where q_C is the characteristic velocity, L_C is the characteristic length, and μ is the dynamic viscosity. Assuming $\mu = 1.8 \times 10^{-5} \, Pa.s$ and $\rho_{air} = 1.2 \, kg.m^{-3}$ [100], and taking q_C and L_C as the inlet velocity and inlet diameter, respectively, gives Eq. 5-16, where $\dot{V}_{supply}$ is the supply flow rate, and A_{inlet} and D_{inlet} are the area and diameter of the supply air flow inlet, respectively. Re is therefore 3.8×10^4 as shown in Eq. 5-17 (where $\dot{V}_{supply}$ is set so as to be consistent with Test 1). The Reynolds number at which turbulence becomes important varies depending on the geometry of the domain, but generally ranges from 10^4 to 10^6, suggesting that turbulence may play some role in this application. However, since lower velocities and larger length scales dominate for the majority of the domain, it is likely that turbulence only becomes important in regions close to inlets and outlets (of the Test Data Centre and of the heat loads).

$$Re = 1.2 \times \frac{\dot{V}_{supply}}{A_{inlet}} \times D_{inlet}/\mu \qquad \text{Eq. 5-16}$$

$$Re = \frac{1.2 \times 0.1419 \times 0.315}{0.0779 \times 1.8 \times 10^{-5}} \qquad \text{Eq. 5-17}$$

The NS-CFD model shows generally much higher temperatures in the cold aisle, with the PF-CFD model showing warm, recirculated air occupying only the region close to the partition and rack fronts (see Figure 5-21 and Figure 5-22), and particularly above the height of the rack roofs. This is likely to result from the lack of turbulence in the PF-CFD model, which reduces the extent of mixing. The NS-CFD model also predicts much higher temperatures in the upper part of the cold aisle than in the lower part, but the hot air is still distributed more widely in the z direction than is the case in the PF-CFD model.
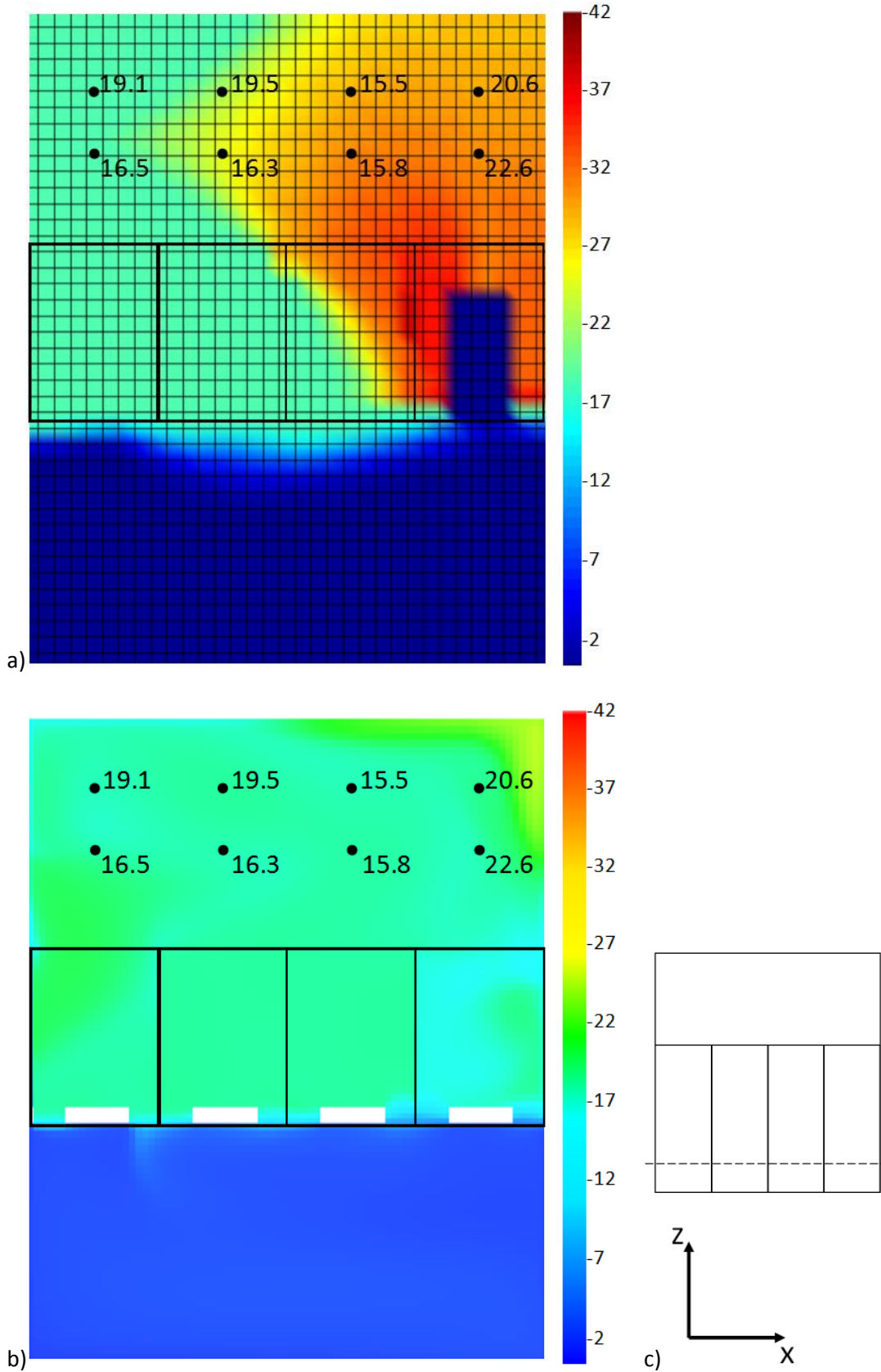
Figure 5-21. Temperature rise (K) from T_{supply} in (a) PF-CFD and (b) NS-CFD at y=0.4 m, for the simulations replicating Test 1. Superimposed numerical data shows the measured temperature rises. Diagram (c) shows the position of the plane within the Test Data Centre.
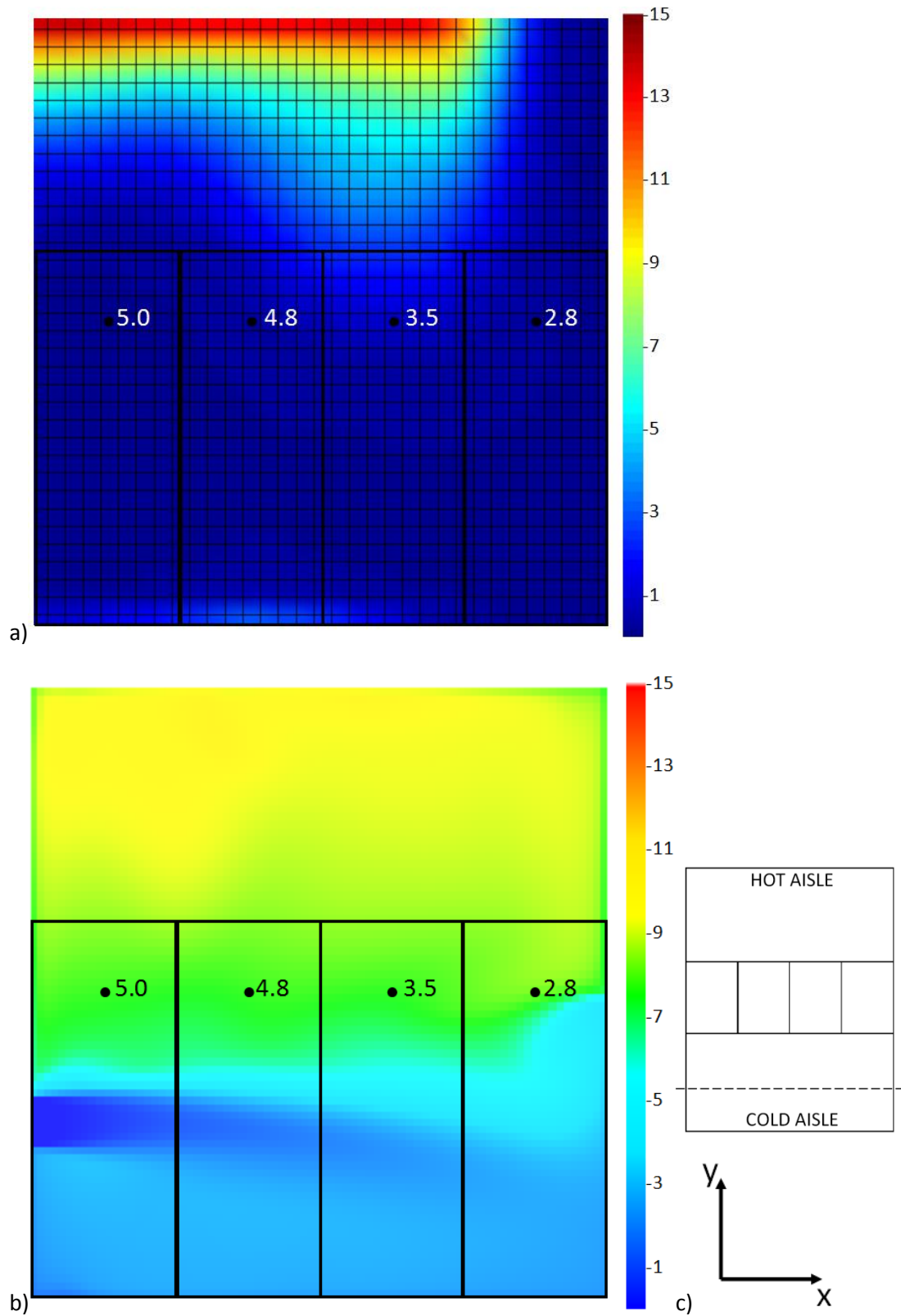
Figure 5-22. Temperature rise (K) from T_{supply} in (a) PF-CFD and (b) NS-CFD at z=0.7 m, for the simulations replicating Test 1. Superimposed numerical data shows the measured temperature rises. Diagram (c) shows the position of the plane within the Test Data Centre.
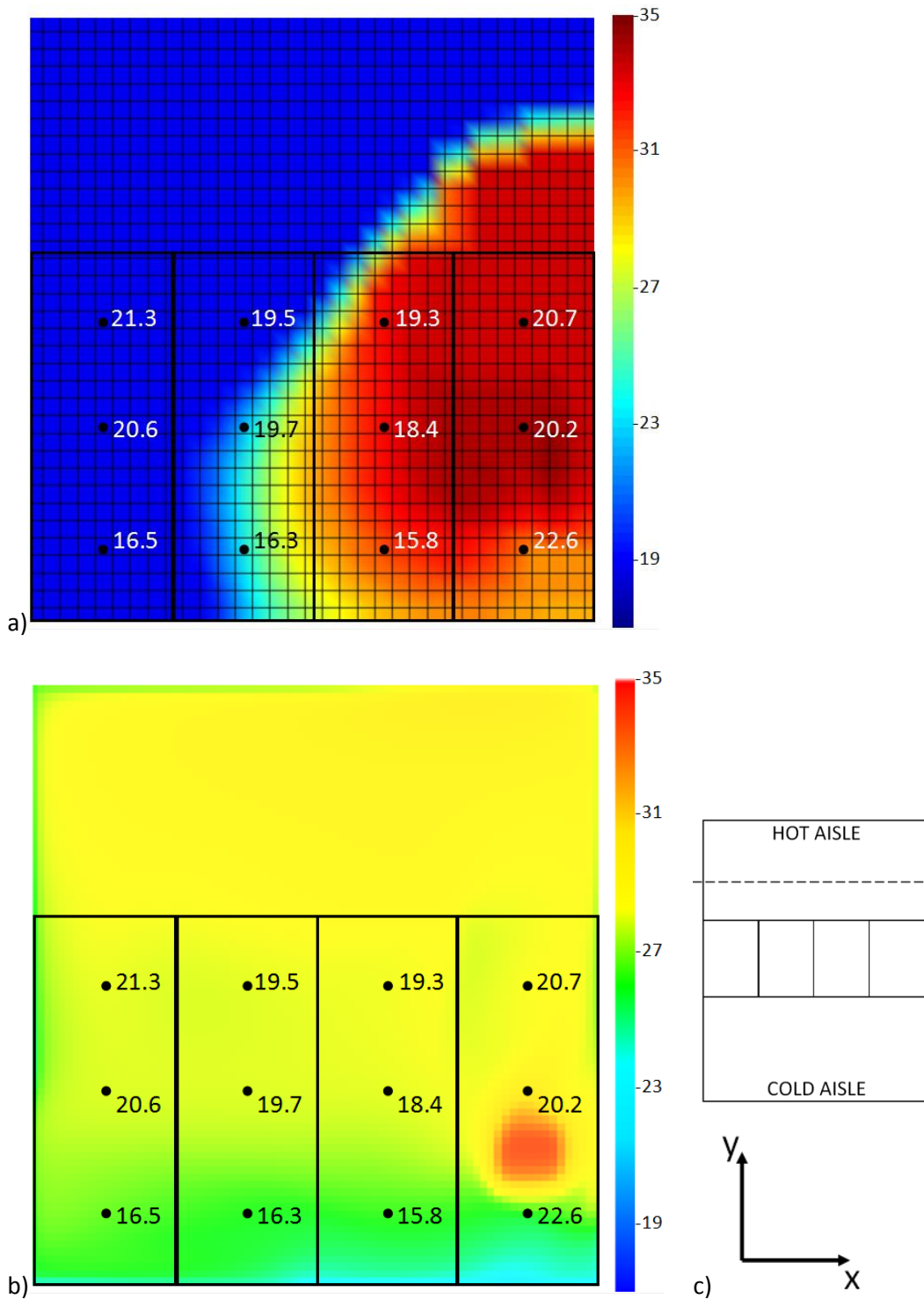
Figure 5-23. Temperature rise (K) from T_{supply} in (a) PF-CFD and (b) NS-CFD at z=3.02 m, for the simulations replicating Test 1. Superimposed numerical data shows the measured temperature rises. Diagram (c) shows the position of the plane within the Test Data Centre.

Figure 5-24 shows the pressure field in the cold aisle for each model for Test 1. Note that, as discussed in section 2.3.2.1, measurements taken during Tests 1 and 2 found variation of pressure in the cold aisle never exceeded ±0.2 Pa, provided that measurements were not taken close to the air supply inlet or load bank inlets. This was within the range of accuracy of the manometer used. Both models broadly replicated this finding, with a range of 0.2 Pa covering almost the entire plane in the NS-CFD model. In the PF-CFD, the same plane is covered almost entirely by a range of 0.05 Pa.

The PF-CFD model shows a clear region of low pressure on the left hand side of Figure 5-24 (a) close to the supply air inlet, which is not replicated in the NS-CFD model. The reasons for this are unclear. Both models showed high velocities in this region in comparison with those in the remainder of this plane, which should lead to low pressures, according to Eq. 5-6 for the PF-CFD model, and via the conservation of energy equation for the NS-CFD model (Eq. 4-11).

The NS-CFD model shows a clear region of high pressure where the jet from the supply air inlet meets the opposite wall (on the right hand side of Figure 5-24 (b)). This is the result of velocity being converted to pressure via the conservation of energy equation (Eq. 4-11) because of the impingement of the jet on the wall. This is not replicated in the PF-CFD model, since the supply air does not flow into the room in such a distinct jet in this model. This difference between the velocity fields in the two simulations is demonstrated in Figure 5-25.
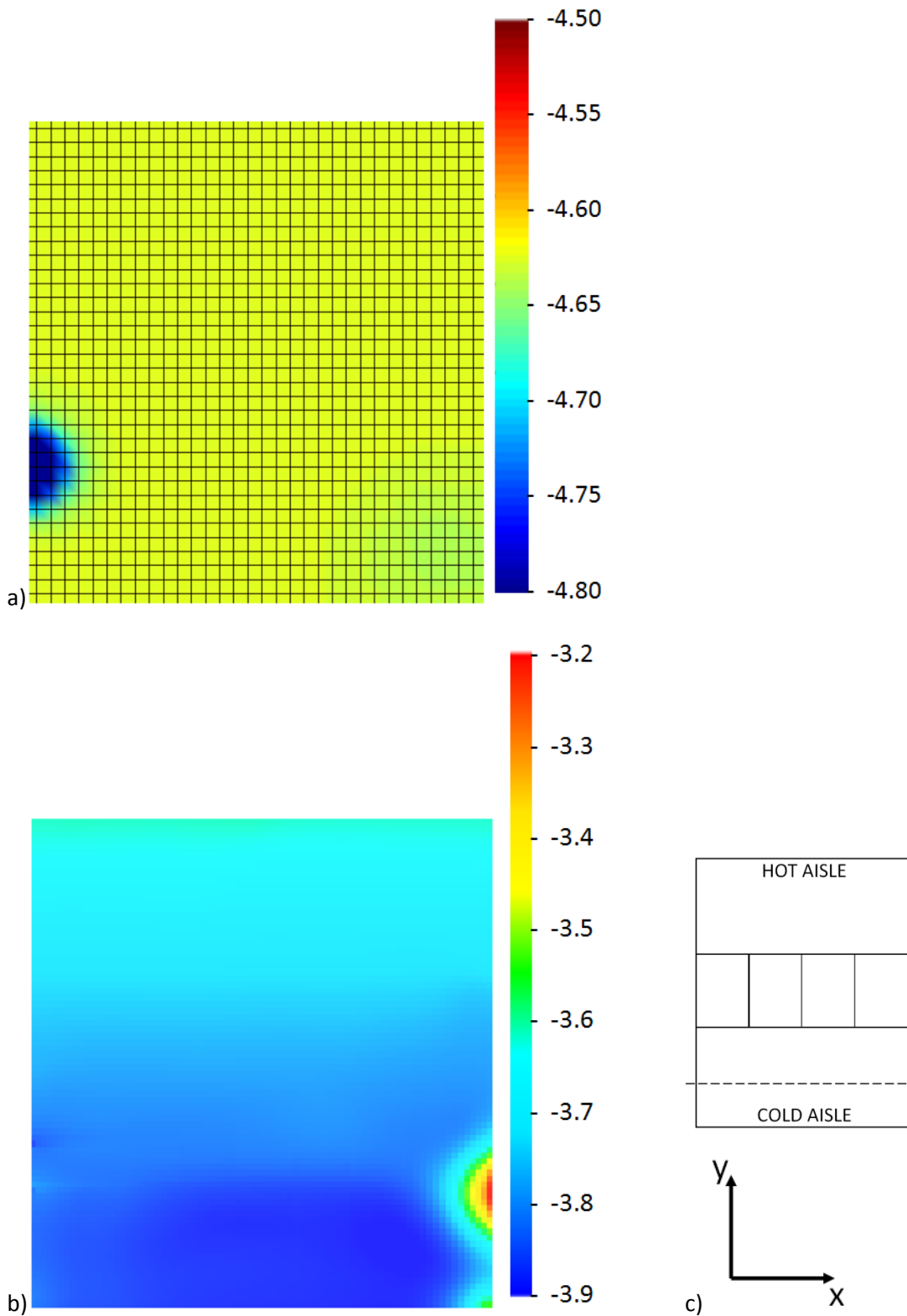
Figure 5-24. Static pressure (Pa) in planes with z=0.8 m, passing through the centre of the supply air inlet, for Test 1, as predicted by the (a) PF- and (b) NS-CFD models. Diagram (c) shows the position of the plane within the Test Data Centre.
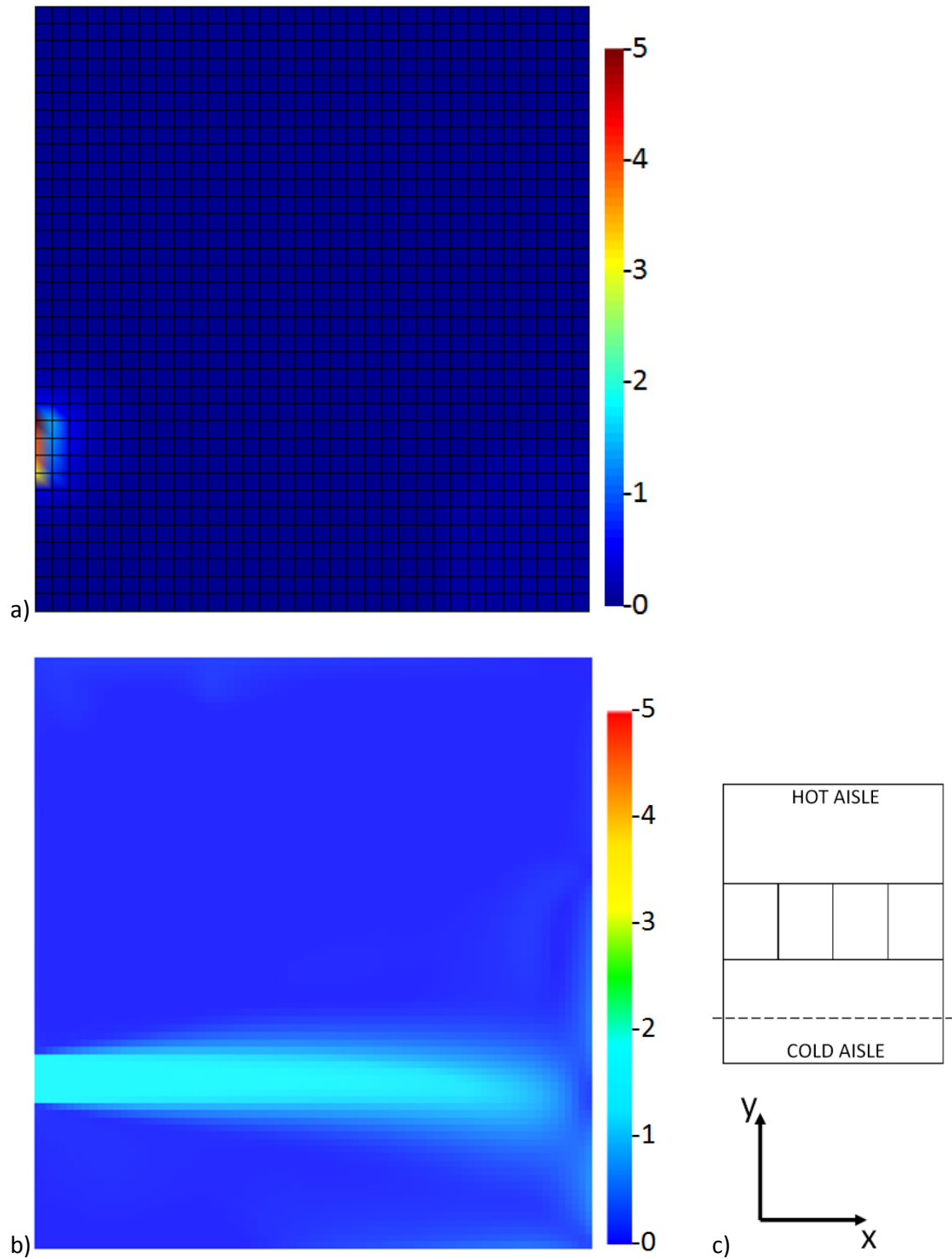
Figure 5-25. Total velocity ($m.s^{-1}$) in planes with z=0.8 m, passing through the centre of the supply duct, for Test 1, as predicted by the (a) PF- and (b) NS-CFD models. Diagram (c) shows the position of the plane within the Test Data Centre.

5.6   Results - example data centre geometry

Section 5.5.3 has shown that the CFD models of the Test Data Centre can predict server inlet temperatures with a reasonable degree of accuracy, comparable with other models presented in the literature. Both models have accurately identified the positions of the hottest inlet temperatures, and broadly predicted the impact of the introduction of an empty slot.

Having validated the two models of the Test Data Centre against experimental data, the predictions of the modified models representing an example data centre geometry, as described in sections 5.1.2 and 5.2.2, were analysed.

5.6.1   Results with $\Delta p_{CH} = 1\ Pa$, Server Fan Option 2

The PF-CFD model was firstly run with the pressure differential between the cold and hot aisles (Δp_{CH}) set at 1 Pa. The supply flow rate required to maintain this pressure differential, $\dot{V}_{supply}$, was $6.413\ m^3.s^{-1}$. The NS-CFD model was then run with $\dot{V}_{supply} = 6.413\ m^3.s^{-1}$. Both of these simulations were run with low bypass conditions and with Server Fan Option 2 (as defined in sections 3.3.1.1 and 3.3.1.2).

Figure 5-26 shows the predicted mean temperature rise at each server inlet for each of the models. The highest mean server inlet temperature rise in the PF-CFD model was 0.06 K, indicating very low levels of recirculation, having no significant impact on inlet temperatures. In the NS-CFD model the highest mean server inlet temperature rise was 0.6 K, again indicating very low levels of recirculation, although higher than in the PF-CFD model.

Figure 5-27 and Figure 5-28 show the temperature rise and pressure fields, respectively, for these two simulations, in horizontal planes passing through the centres of the uppermost server in each rack. Note that for the PF-CFD simulation, Figure 5-28 does not show the full range of pressures. There were some regions of very low pressure in the PF-CFD results, which was not the case in the NS-CFD results, hence a comparison of the two equivalent planes with the full ranges displayed would be unclear. The full range is shown in Figure 5-30, and the reasons for the very low pressures in the PF-CFD model are discussed later in this section.

Both models show server outlet temperatures tending to fall along the length of the aisle with increasing x, although this effect is more pronounced in the NS- than in the PF-CFD model (see Figure 5-27). This effect is caused by the rising pressure drop across the servers with increasing x, shown in Figure 5-28, which causes rising server flow rates and falling server temperature rises. This effect is more pronounced in the NS- than in the PF-CFD

model. Note that many servers experience a negative pressure drop in both models. Negative pressure drops could in practice either cause inadequate cooling of the affected servers, or result in increased server fan power consumption. Note that these results used Server Fan Option 2, i.e. fan speeds were assumed to be unaffected by the pressure drop across the server. As noted in section 3.2, the evidence regarding server fan responses to changing pressure conditions is limited.

Figure 5-28 (b) clearly shows a large negative pressure differential across the partition close to the supply air inlet in the NS-CFD simulation (where a positive pressure differential denotes a higher pressure on the cold aisle side of the partition). This is not replicated in the PF-CFD simulation, and neither model shows sufficiently high velocities in this region to produce such a pressure gradient. Since the local variations in pressure in the PF-CFD model are only affected by velocity, this suggests that the lack of consideration of the impact of viscosity and conservation of momentum on the pressure field here has a considerable impact on the results. These trends are further highlighted in Figure 5-29, which shows the w velocity in the plane of the partition for each simulation. Both simulations show positive w (i.e. flow from the cold to the hot aisle) dominating either side of racks 10 to 16, with both positive and negative w occurring closer to the supply inlet.

When discussing the pressure fields, it is important to note that Δp_{CH} as calculated in the PF-CFD model, and as discussed here, refers to the difference between the static pressures at points in the hot and cold aisle with zero velocity. The results shown in Figure 5-28 indicate that, for an operational data centre, a measurement of pressure differential between the two aisles may vary considerably depending on the position of measurement. Clearly, the variation in pressure within each aisle would depend on the specification of the data centre, with factors such as aisle length, air supply configuration (e.g. side fed as in this example, underfloor or overhead supply) and required supply flow rate having an impact.
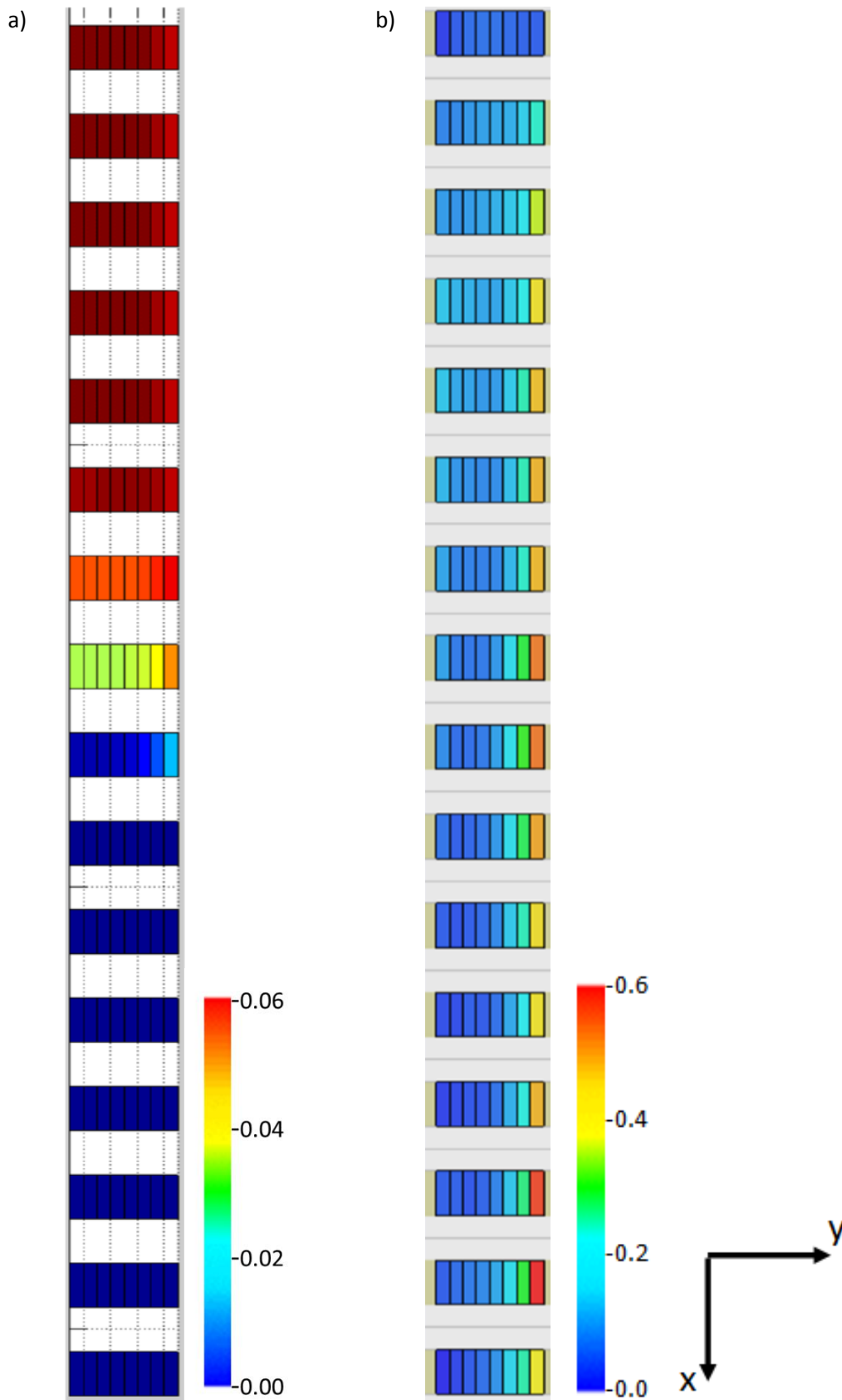
Figure 5-26. Mean temperature rise (K) at server inlets for (a) PF- and (b) NS-CFD simulations with $\Delta p_{CH} = 1$, with low bypass conditions and Server Fan Option 2.
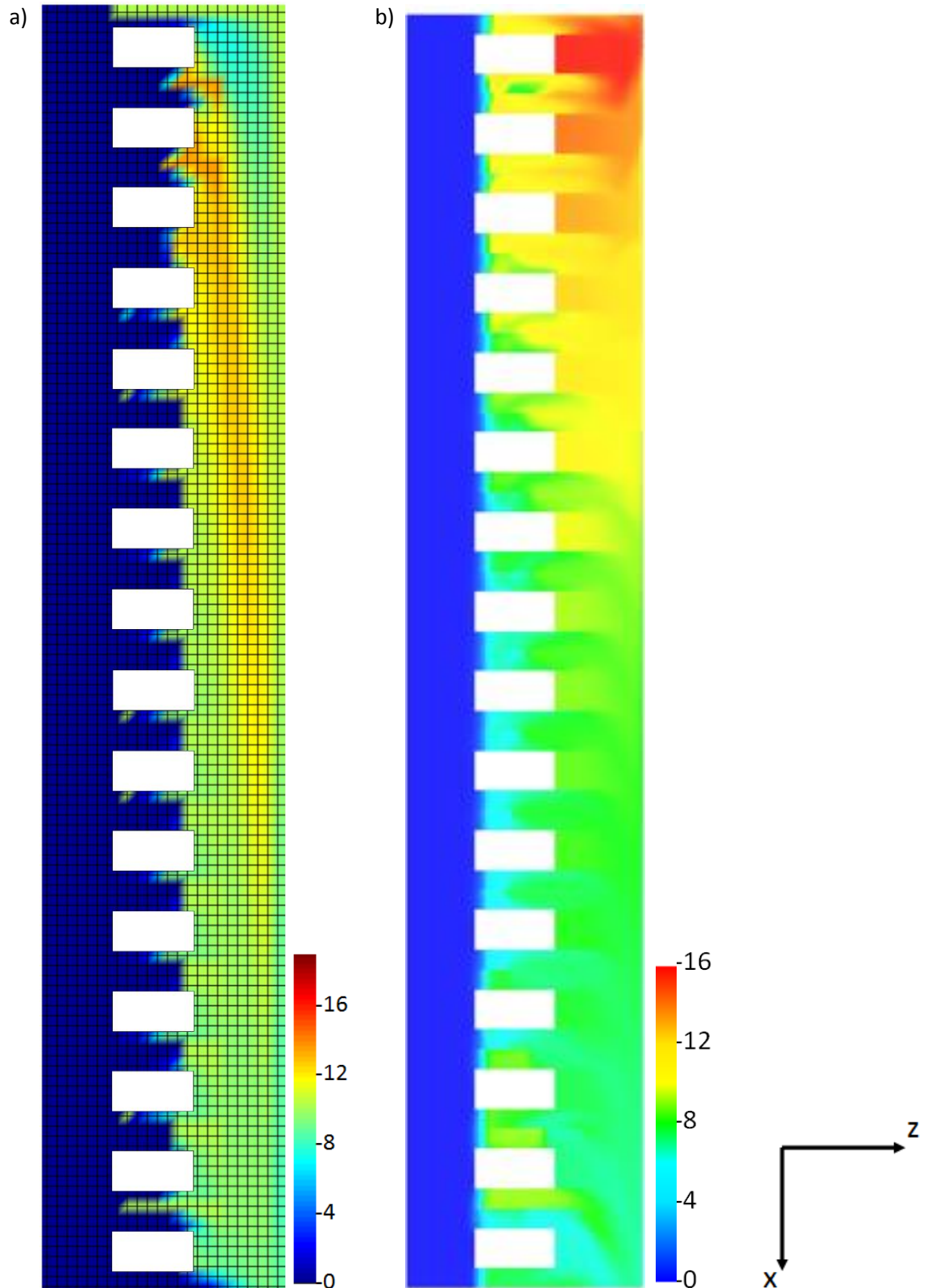
Figure 5-27. Plan view of temperature rise (K) in a horizontal plane passing through the 8th server in each rack, in (a) PF- and (b) NS-CFD, with $\Delta p_{CH} = 1\ Pa$, with low bypass conditions and Server Fan Option 2.
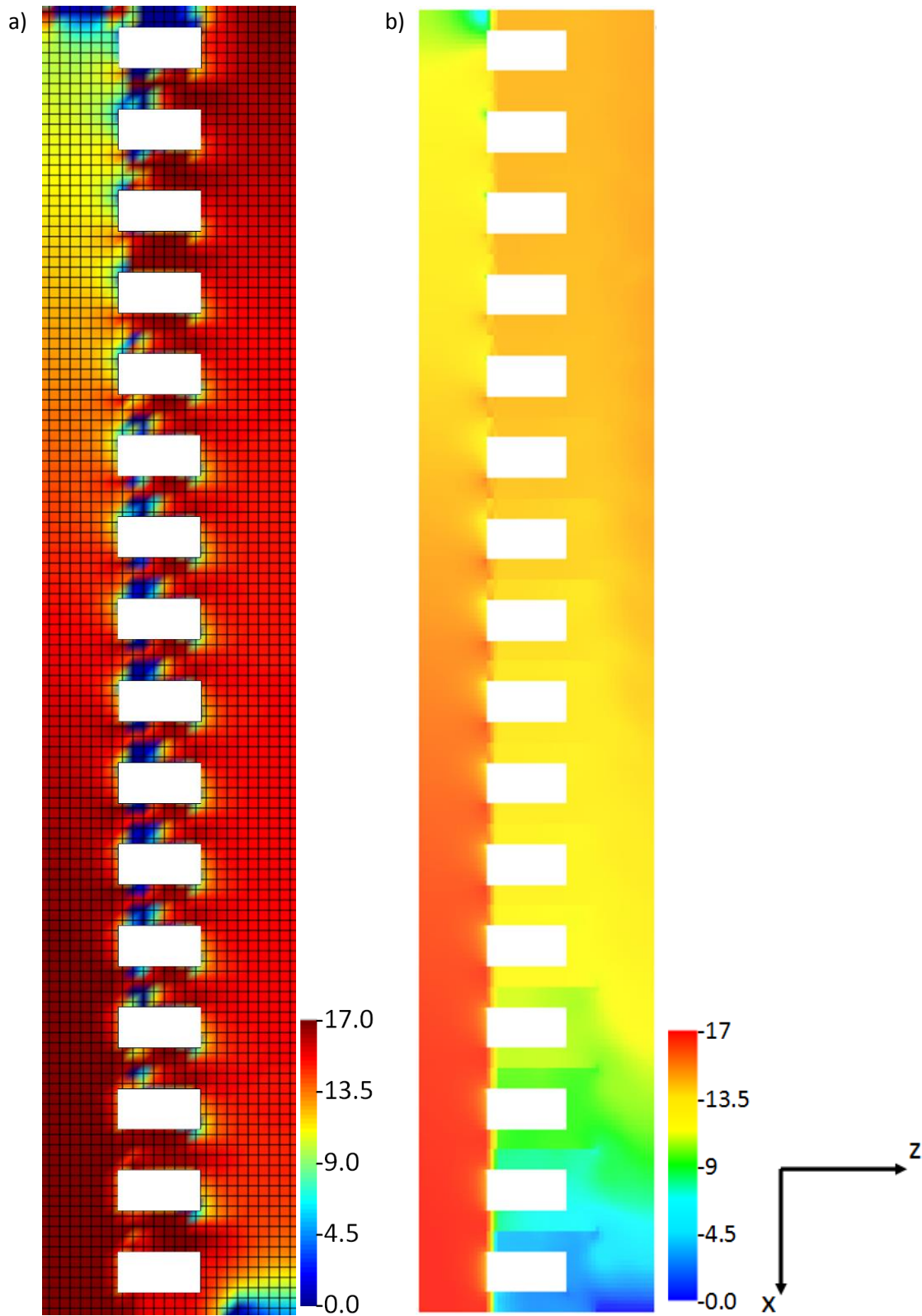
Figure 5-28. Plan view of pressure (Pa) in a horizontal plane passing through the 8ᵗʰ server in each rack, in (a) PF- and (b) NS-CFD, with $\Delta p_{CH} = 1\ Pa$, with low bypass conditions and Server Fan Option 2.

161
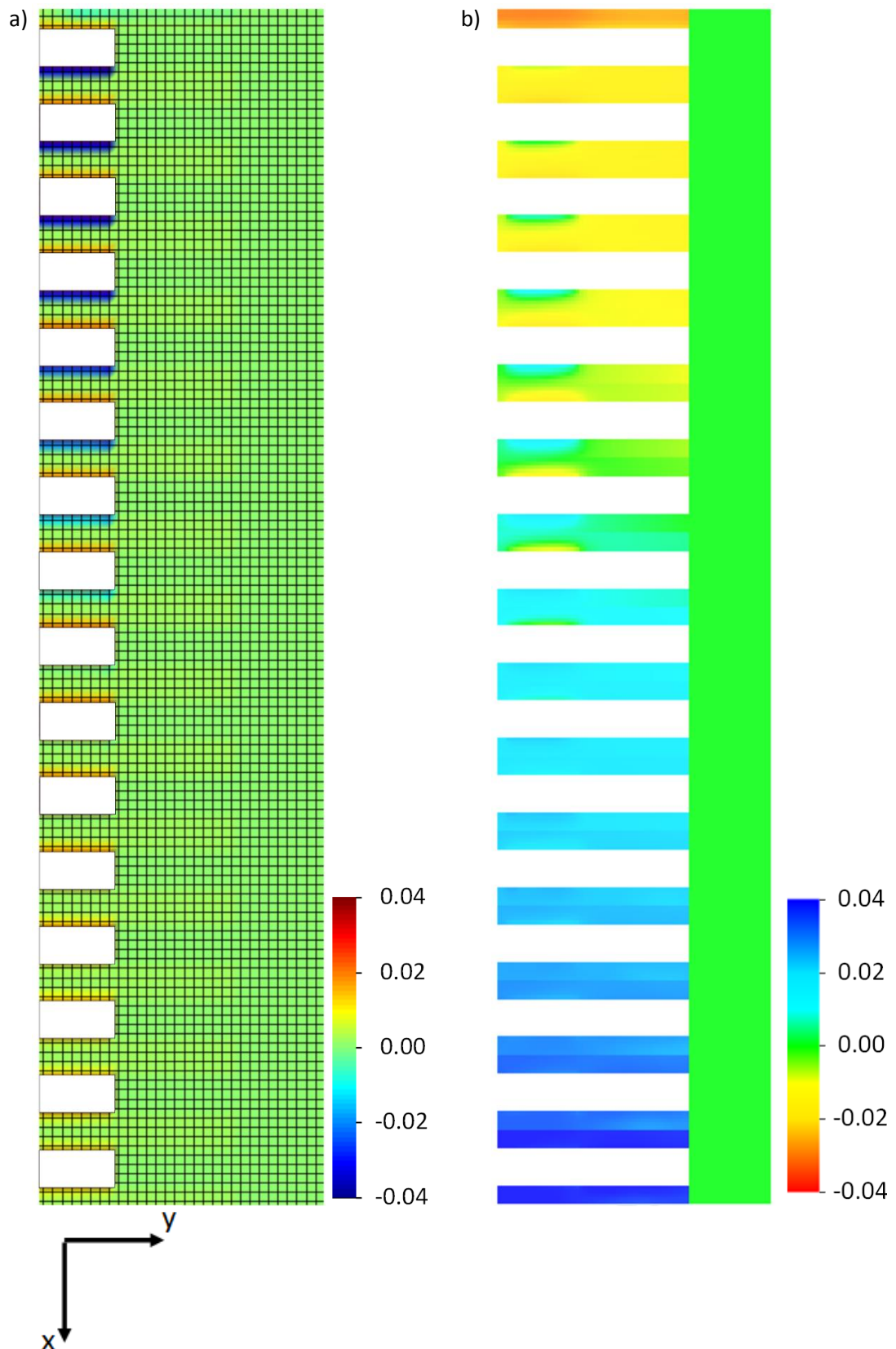
Figure 5-29. w velocity $(m.s^{-1})$ in a vertical plane coincident with the partition separating the hot and cold aisles, in (a) PF- and (b) NS-CFD, with $\Delta p_{CH} = 1\ Pa$, with low bypass conditions and Server Fan Option 2. Note that the colour red depicts positive w in (a) and negative w in (b).

As noted previously, the pressure field shown for the PF-CFD model in Figure 5-28 (a) does not show the full range of pressures predicted. The pressure planes shown in Figure 5-28 are repeated in Figure 5-30, with the full range of pressures shown. This figure shows very low pressures in the PF-CFD simulation in the regions close to the rack sides. This indicates very high velocities in these regions. The high velocities result from the boundary conditions imposed for the $\emptyset$ calculations at the rack sides (described in section 5.1.1.2), which mean that the Laplace equation is not applied for these cells. This leads to discontinuities, and unrealistically high velocities. This trend is not shown in the equivalent NS-CFD simulation. Note that these very low pressures do not occur at the cells adjacent to the partition, hence they do not directly impact bypass and recirculation. Indeed, the only calculations using data from the pressure field relate to the flows through the servers and through the leakage paths, hence the only pressures having an impact on the temperature and velocity fields are those at the server inlets and outlets, and at either side of the partition.

Figure 5-31 shows the pressure drop across each server for each model, with the low bypass condition, Server Fan Option 2, and with $\Delta p_{CH} = 1\ Pa$. Both simulations show increasing pressure drops along the length of the aisle, with negative pressure drops across servers nearer to the conditioned air supply. The NS-CFD simulation shows a greater variation in pressure drops than the PF-CFD simulation, with the total ranges being -7.8 to 28.4 Pa, and -2.5 to 11.8 Pa, respectively. The NS-CFD model predicts 38 servers experiencing negative pressure drops, compared with 22 in the PF-CFD model. By reducing flow rates through the servers, negative pressure drops could result in inadequate cooling being provided to the server, and/or increased server fan power consumption to compensate.
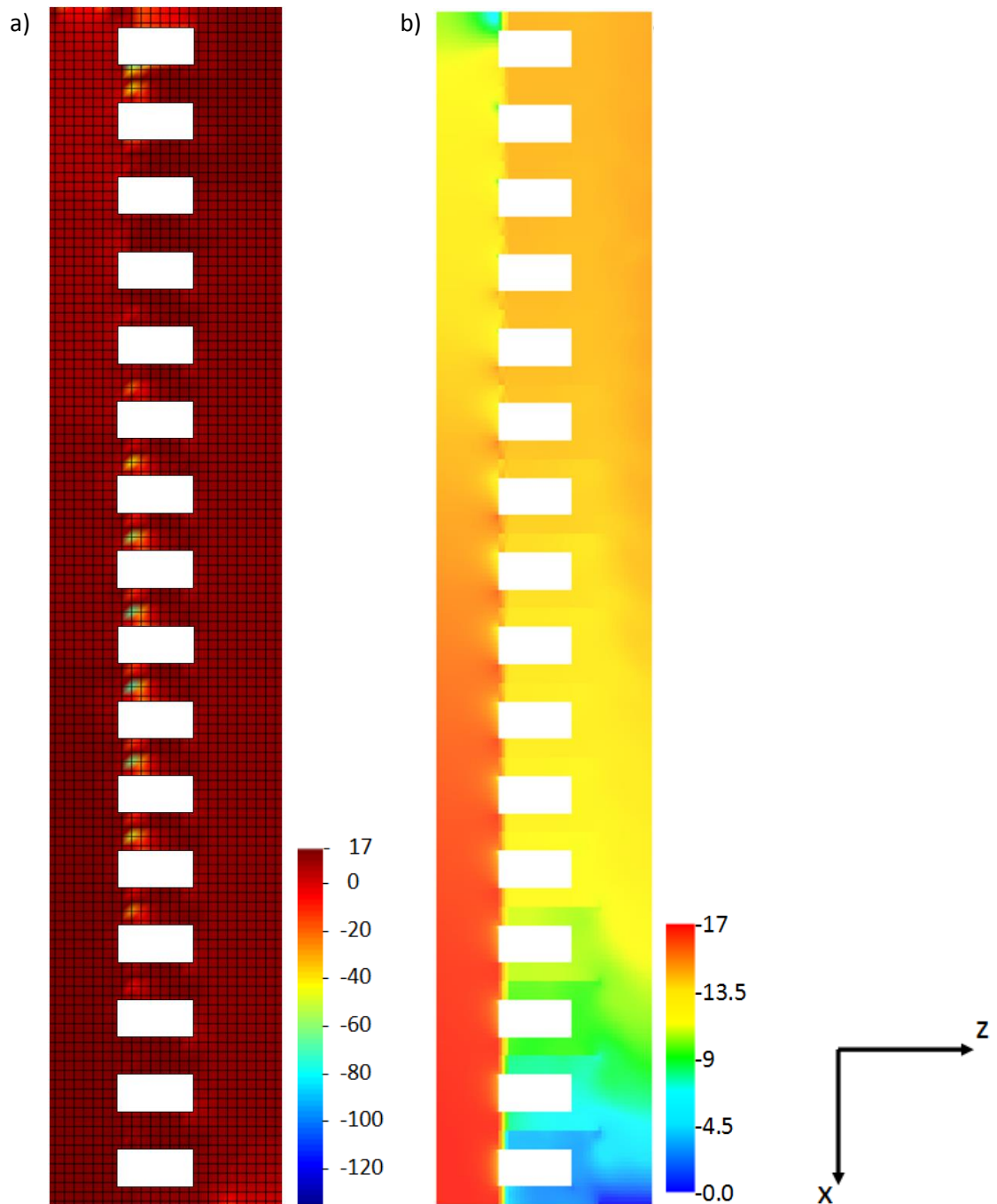
Figure 5-30. Plan view of pressure (Pa) in a horizontal plane passing through the 8th server in each rack, in (a) PF- and (b) NS-CFD, with $\Delta p_{CH} = 1\ Pa$, with low bypass conditions and Server Fan Option 2.
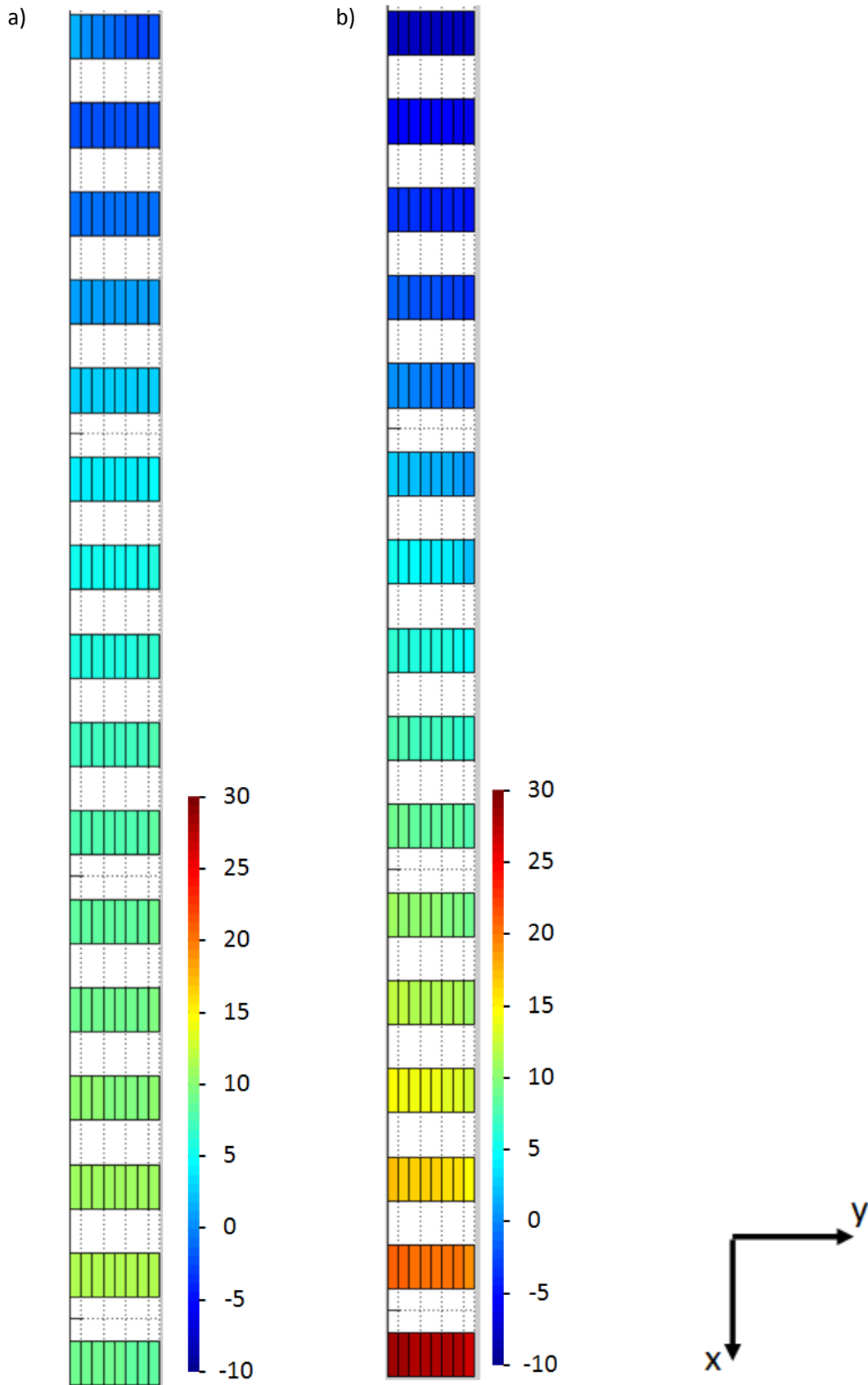
Figure 5-31. Pressure drops (Pa) across servers in (a) PF- and (b) NS-CFD, with $\Delta p_{CH} = 1\ Pa$, with low bypass conditions and Server Fan Option 2.

5.6.2 Trends with Δp_{CH}, with Server Fan Option 2

Further simulations were run with Δp_{CH} ranging from 0 to 20 Pa in the PF-CFD model, and in the NS-CFD model with $\dot{V}_{supply}$ set in accordance with that recorded in each respective PF-CFD simulation. Again, these simulations were run with the low bypass condition and with Server Fan Option 2. Figure 5-32 shows the highest mean server inlet temperature rise for each simulation.

The PF-CFD model failed to converge the $\emptyset$ field to the desired level with $\Delta p_{CH} < 1$. With $\Delta p_{CH} = 0\ Pa$, $R_{\emptyset}$ fell to 4.5×10^{-6} (note that $R_{\emptyset,target} = 10^{-6}$). In order to assess the impact of the failure of convergence of the $\emptyset$ field in the PF-CFD model, the simulation with $\Delta p_{CH} = 0$ was repeated with $R_{\emptyset,target} = 10^{-5}$, i.e. around double the level of $R_{\emptyset}$ originally achieved. This change increased the predicted inlet temperatures, with a maximum mean server inlet temperature rise of 5.6 K, as opposed to 3.5 K in the original simulation. Since with $\Delta p_{CH} = 0\ Pa$, local pressure drops across the leakage paths should all be fairly close to zero, any minor failure to satisfy the Laplace equation for a given cell could have a large impact on recirculation. The results suggest that these errors tend to increase recirculation, which is likely to account for the discrepancies between the results from the NS- and PF-CFD models in this case.

With $\Delta p_{CH} = 0.5\ Pa$, $R_{\emptyset}$ in the PF-CFD simulation fell to a minimum of 2.7×10^{-6}. This PF-CFD simulation showed greatly reduced mean server inlet temperatures in comparison with the simulation with $\Delta p_{CH} = 0\ Pa$. The highest mean inlet temperature in the PF-CFD model was 0.22 K above T_{supply}. The highest mean inlet temperature in the respective NS-CFD model was 0.63 K above T_{supply}. The PF-CFD simulation with $\Delta p_{CH} = 0.5\ Pa$ was repeated with $R_{\emptyset,target} = 5.4 \times 10^{-6}$, i.e. around double the level of $R_{\emptyset}$ achieved in the original simulation. This simulation predicted a higher maximum mean server inlet temperature, 0.39 K above T_{supply}. This adds weight to the hypothesis that limitations of convergence in the $\emptyset$ field tend to increase recirculation and inlet temperatures.

Both models show the highest mean server inlet temperature rise falling as Δp_{CH} rises, due to the reduction in recirculation (see Figure 5-32). The NS-CFD model predicts higher server inlet temperatures. Where convergence was achieved, no simulation predicted inlet temperature rises exceeding $0.65\ K$. Such temperature rises may be considered to be small in comparison with the ASHRAE recommended inlet temperatures, which cover a range of 18-27 °C [177]. Hence, the models may be seen to predict little improvement in thermal conditions with increasing Δp_{CH}. The temperature rises are also small in comparison with the errors in temperature predictions reported with respect to the validation process in

section 5.5.3. It follows that the models cannot be relied upon to accurately predict temperature rises within this range. However, the broader prediction that temperature rises are small in comparison with the allowable range of server inlet temperatures may be considered to be reliable, since the validation process showed that both models were successful in predicting trends in recirculation with changes in pressure conditions, which is key to predicting server inlet temperatures.

One major difference between the server inlet temperature predictions of the two models is that the PF-CFD model predicts no server inlet temperature rise where $\Delta p_{CH} > 2$, whereas the NS-CFD model predicts a highest mean server inlet temperature rise of $0.37\ K$ even with $\Delta p_{CH} = 20\ Pa$. This is due to differences in the pressure fields of the two simulations, as demonstrated in Figure 5-34. As with the pressure fields with $\Delta p_{CH} = 1\ Pa$ (as shown in Figure 5-28), the NS-CFD model predicts a significant pressure gradient from the hot to the cold aisle, in the region adjacent to rack 1 (i.e. close to the conditioned air supply), even at $\Delta p_{CH} = 20\ Pa$. This leads to recirculation in this region, even though the pressure gradient acts from the cold to the hot aisle over the rest of the partition. Conversely, in the PF-CFD model, a positive pressure gradient from the cold to the hot aisle acts over the entire partition with $\Delta p_{CH} = 20\ Pa$. As identified with regards to Figure 5-28, this difference is likely to be attributable to the NS-CFD model's accounting for viscosity and momentum, which are not included in the PF-CFD model.One other possible explanation for the higher server inlet temperatures in the NS-CFD model than in the PF-CFD model is the inclusion of diffusion of heat in the former model. However, a repeat of the NS-CFD model with $\Delta p_{CH} = 20\ Pa$, with low bypass and with Server Fan Option 2, but with diffusion not included, found that this made no difference to server inlet temperatures. Similarly, when repeating the PF-CFD simulation with $\Delta p_{CH} = 0\ Pa$, with low bypass and with Server Fan Option 2, with diffusion included, server inlet temperatures changed by no more than $0.02\ K$.
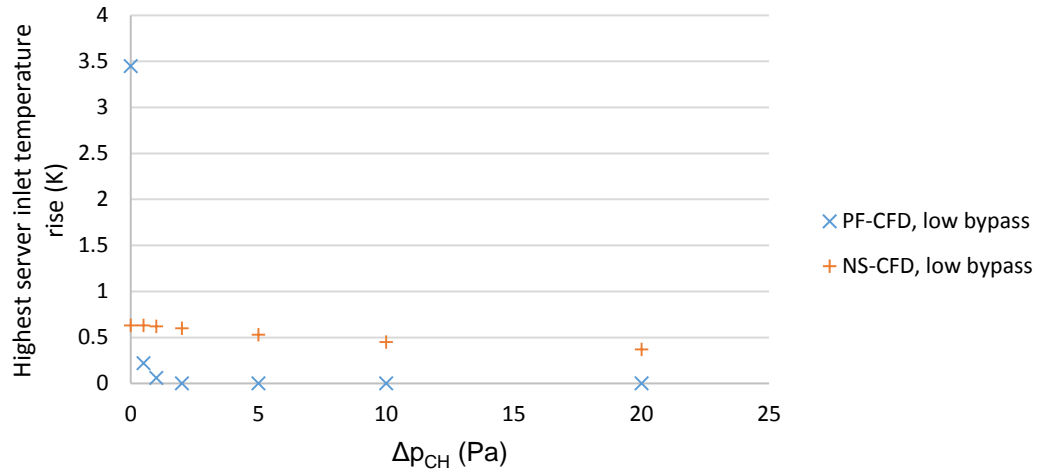
Figure 5-32. Variation of highest mean server inlet temperature rise with Δp_{CH} in the PF- and NS-CFD simulation results, with low bypass and Server Fan Option 2.
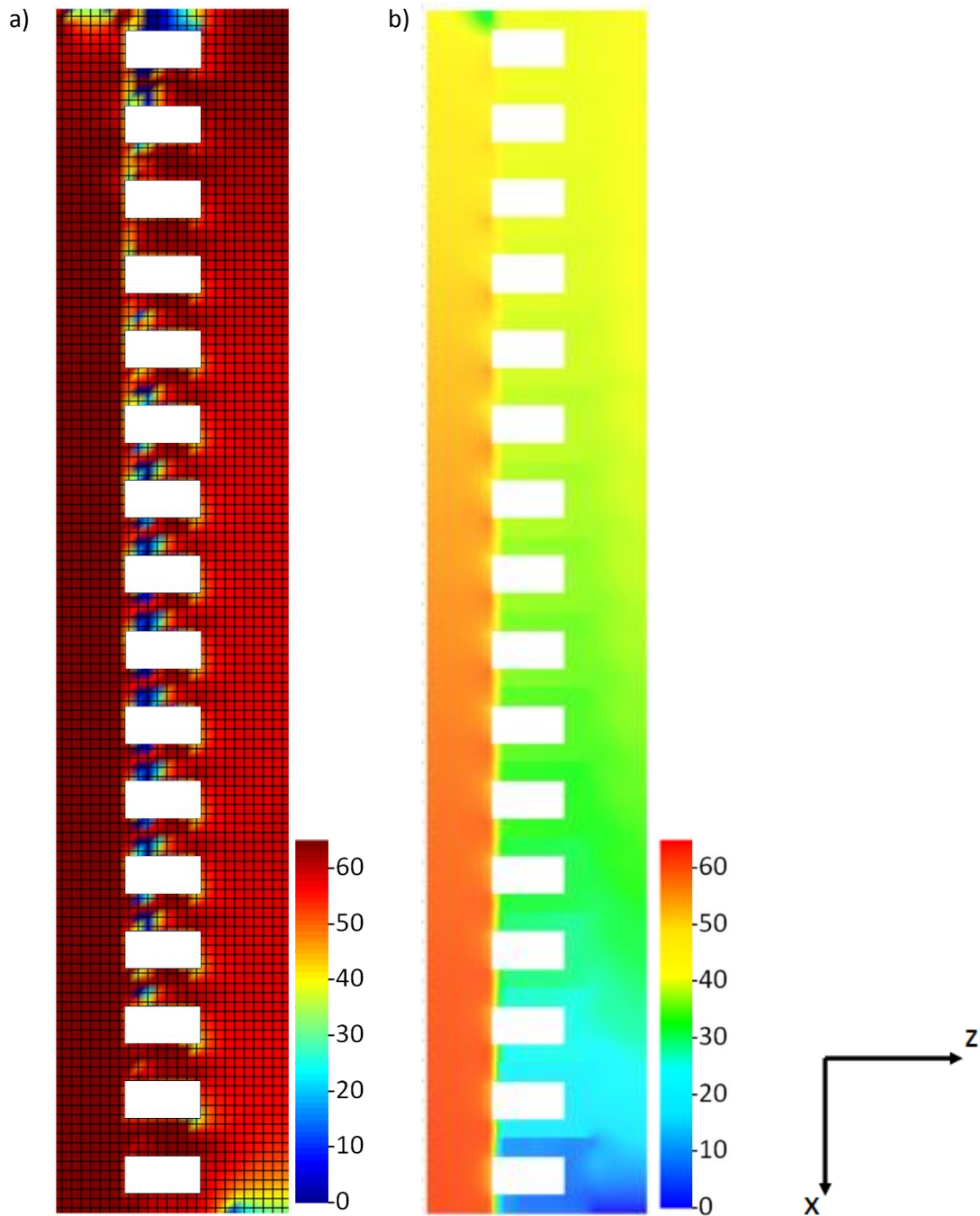
Figure 5-33. Plan view of pressure (Pa) in a horizontal plane passing through the 8th server in each rack, in (a) PF- and (b) NS-CFD, with $\Delta p_{CH} = 20\ Pa$, with low bypass conditions and Server Fan Option 2.

The modified system model, as described in section 5.3, was then used to predict E_T for each of these simulations. The results are shown in Figure 5-34. The results generally show E_T increasing with Δp_{CH}. The only exception is with the PF-CFD model at $\Delta p_{CH} < 1\ Pa$. This is due to the impact of high server inlet temperatures predicted in these simulations, which, as discussed earlier in this section, are likely to be erroneous. The remaining results show similar trends to those shown in Figure 3-9 and Figure 3-10, with Server Fan Option 2. This

is expected, since the only difference between the system model as used here and as used in Chapter 3 is the consideration of hot air recirculation, and since this recirculation has had little impact on inlet temperatures. Hence, the changes in E_T are dominated by the impact of Δp_{CH} on the supply air flow rate required, and by extension on CRAH fan power consumption. Since bypass increases with Δp_{CH}, so does the supply flow rate, and therefore E_T. Also note that, since the supply flow rate in the NS-CFD model is fixed at the same rate as that predicted by the PF-CFD model for each value of Δp_{CH}, and since variation in the CRAH power consumption dominates variation in E_T (as discussed in section 3.4), the two models predict very similar results.
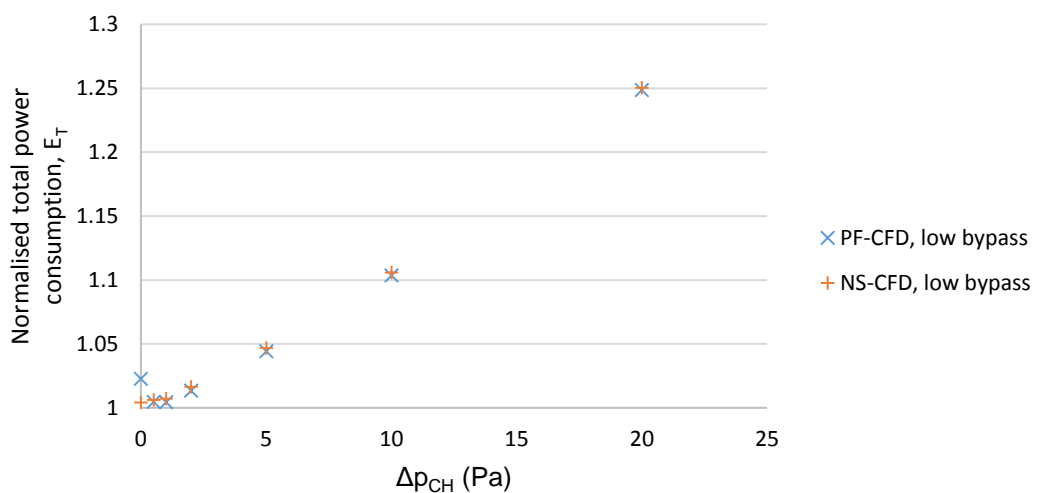
Figure 5-34. Variation of E_T with Δp_{CH} in the PF- and NS-CFD simulation results, with low bypass conditions and Server Fan Option 2.

Figure 5-35 shows the percentage of servers across which there is a negative pressure differential (i.e. a pressure gradient from the hot to the cold aisle), for the simulations undertaken with low bypass conditions and Server Fan Option 2. Both models show the percentage of affected servers falling as Δp_{CH} increases, with the PF-CFD model predicting less servers being affected. The results suggest that cooling effectiveness may be impaired at low Δp_{CH}, such that there will be a trade-off between reducing Δp_{CH} in order to reduce E_T, and increasing Δp_{CH} to ensure adequate cooling. This is in spite of the very minimal levels of recirculation, even at very low levels of Δp_{CH}, as discussed in relation to Figure 5-32. This suggests that it may in practice be necessary to monitor pressure at numerous positions within each aisle in order to ensure that $\dot{V}_{CRAH}$ is set sufficiently high to ensure that $\Delta p_{CH} > 0$ for all or most of the racks. Of course, such negative pressure gradients may in reality also increase server fan power consumption, which was not considered in the simulations using Server Fan Option 2. It is also important to note that, as mentioned in section 5.6.1, the

extent to which pressure varies within each aisle in a given data centre will be dependent upon the specifications of the data centre.
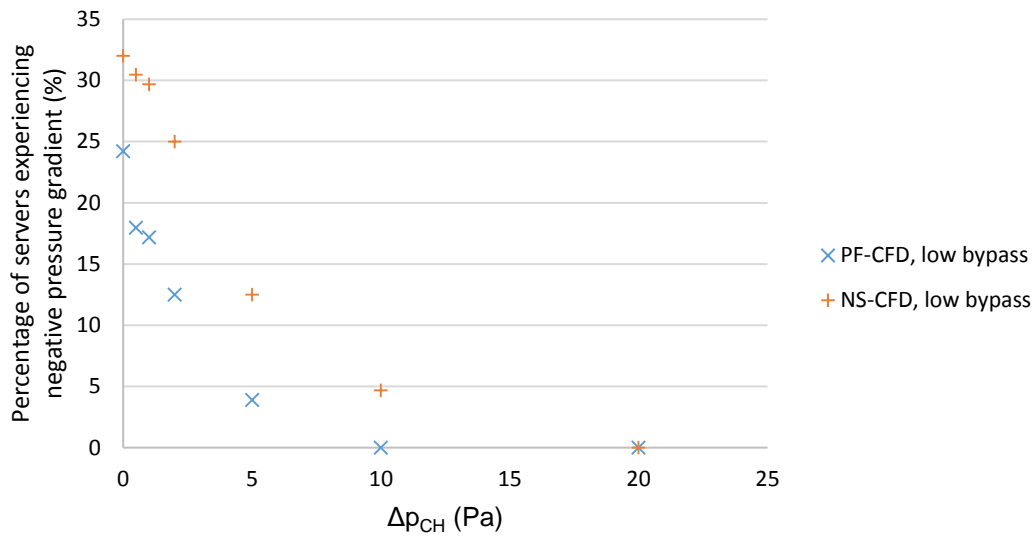
Figure 5-35. Variation of percentage of servers experiencing negative pressure gradient with Δp_{CH} in the PF- and NS-CFD simulation results, with low bypass and Server Fan Option 2.

5.6.3   Effect of Server Fan Model

Further simulations were then undertaken using the PF-CFD model, to investigate behaviour with high bypass, and with Server Fan Option 1. Figure 5-36 collects the highest server inlet temperature rise for these simulations, with varying Δp_{CH}. Note that the figure only shows the results with $\Delta p_{CH} \leq 5\ Pa$, since there were no server inlet temperature rises predicted with $\Delta p_{CH} > 5\ Pa$. Note also that, in the figure's legend, "SF" denotes Server Fan. Other than the simulations where the $\emptyset$ field failed to converge, as discussed earlier in this section, no temperature rise exceeding $0.5\ K$ was predicted, even under high bypass conditions. Hence, the model predicts that very uniform temperature profiles can be achieved so long as $\Delta p_{CH} \geq 0\ Pa$, for both low and high bypass conditions. However, with different data centre geometries this may not always be the case, particularly if empty slots or other imperfections in the containment system are present, or during failures in the cooling infrastructure which can cause Δp_{CH} to fall below zero. In such cases, the CFD models described, in conjunction with the modified system model, could be useful in predicting the impact of recirculation on both server inlet temperatures and E_T.
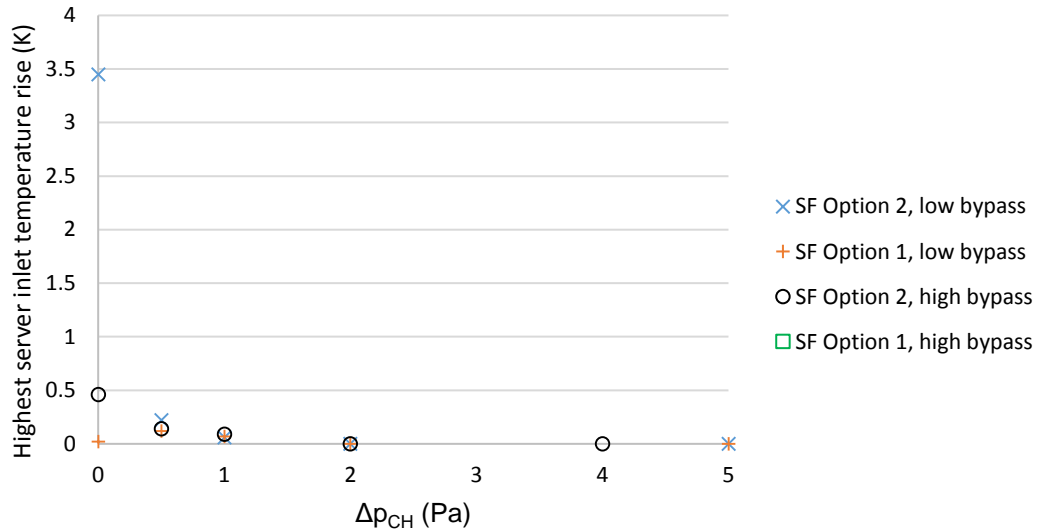
Figure 5-36. Variation of highest server inlet temperature rise with Δp_{CH} in the PF-CFD model.

Figure 5-37 shows the percentage of servers experiencing negative pressure gradients for the PF-CFD simulations with Server Fan Option 2. Note that, with Server Fan Option 1, the allowance of variation in server fan speeds enabled all servers to achieve an adequate flow rate, regardless of external pressure differential. By contrast, the results presented in Figure 5-37 show that, with Server Fan Option 2 applied, in order for all servers to achieve adequate cooling (assuming this requires $\Delta T_{server} \leq 12.5\ K$), Δp_{CH} needed to be greater than or equal to $10\ Pa$ for low bypass conditions, and greater than or equal to $20\ Pa$ for high bypass conditions. Generally, more servers are affected by this issue under the high bypass condition. This is due to the increased levels of bypass, which lead to greater supply flow rates, and higher velocities near to the air supply inlet. These higher velocities cause low pressures close to the server inlets in the affected region.
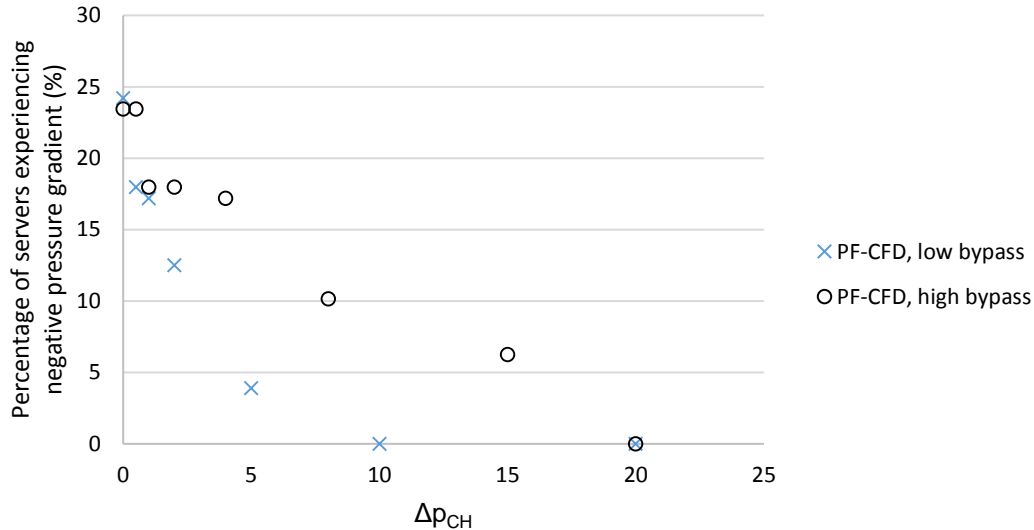
Figure 5-37. Variation of percentage of servers experiencing negative pressure gradients with Δp_{CH} in the PF-CFD model, with Server Fan Option 2.

Figure 5-38 shows the variation in E_T with Δp_{CH} under different conditions. Note that, for simulations with Server Fan Option 1, the value of E_T at a particular Δp_{CH} is shown relative to E_T at $\Delta p_{CH} = 0$ and with low bypass. For simulations with Server Fan Option 2, the value of E_T at a particular Δp_{CH} is shown relative to E_T at $\Delta p_{CH} = 0$ and with low bypass. Absolute values of E_T are not shown. Note also that in this figure's legend, "SF" denotes Server Fan.

With Server Fan Option 2, increasing Δp_{CH} can be seen to increase E_T. The high bypass condition exaggerates this trend. This is due to the amplified impact of Δp_{CH} on bypass, and therefore CRAH fan power consumption, under these conditions.

For Server Fan Option 1, with low bypass, E_T falls by 3.6% when increasing Δp_{CH} from 0 to 5 Pa. This reflects the fall in server fan power consumption which is enabled by increasing Δp_{CH}. Effectively, the more efficient CRAH fans provide an increasing proportion of the required flow energy, as the less efficient server fans provide less. Note that, even at $\Delta p_{CH} = 0$, all servers were able to achieve adequate air flow to achieve a ΔT of $12.5\ K$, without exceeding the maximum power consumption of the server fans. Increasing Δp_{CH} beyond $5\ Pa$ increases E_T. Specifically, E_T rises by 16.2% with Δp_{CH} increasing from 5 to $20\ Pa$.

With Server Fan Option 1, with high bypass, increasing Δp_{CH} from 0 to $8\ Pa$ has very little impact on E_T, as the increasing CRAH fan power consumption roughly cancels out the falling server fan power consumption. Increasing Δp_{CH} from 8 to $20\ Pa$ increases E_T by 23.9%.

Comparing these results with those presented in Figure 3-9 and Figure 3-10, it can be seen that the trends are very similar to the original system model predictions with Server Fan Option 1.

The results presented in Figure 5-35 suggested that a Δp_{CH} of $10\ Pa$ was required to ensure that all servers received adequate cooling under Server Fan Option 2, with low bypass. With high bypass, this figure was $20\ Pa$. It may therefore be concluded that, were Server Fan Option 2 an accurate model of server fan behaviour, Δp_{CH} should be maintained at 10 to $20\ Pa$ for the minimum power consumption at which adequate cooling is achieved, depending on the extent to which bypass has been minimised. If Server Fan Option 1 is taken as the accurate model, the minimum E_T occurs at pressures of around $5\ Pa$ for the low bypass condition, and 0 to $8\ Pa$ for high bypass, with adequate cooling achieved at all pressures.

However, as discussed in section 5.6.1, the definition of Δp_{CH} here masks the variation of pressure along the aisle length. In practice, variation in pressure will depend on the geometry of the data centre. The system model results show that, where a more consistent pressure field is present (i.e. in the PF-CFD model), adequate cooling can be achieved at lower E_T. There could therefore be benefits in introducing measures to encourage a more consistent pressure drop along the length of the aisle. For example, shorter aisles could be advantageous, as could under floor or overhead air supply. Reducing the rack power density would also tend to reduce variation in the pressure field, by reducing the required supply flow rate. However, the other impacts of these measures, such as impacts on capital costs and capacity per square metre of floor space, would have to be considered. The CFD and system models described in this chapter could be used to estimate the impact of such changes in geometry and operation. The server fan behaviour is also key here, particularly the extent to which adequate cooling can be achieved at negative external pressure drop, and the increased power consumption associated with this.

For comparison, it is interesting to note that the EU Code of Conduct on Data Centre Energy Efficiency recommends that Δp_{CH} should not exceed 5 Pa. The recommendations include no discussion of the possibility of variation of pressure within the containment system. In reality, it may be necessary to monitor pressure in numerous positions within each aisle in order to ensure that Δp_{CH} is always within the desired range.

Regardless of the Server Fan Option and pressure distribution, the results clearly show that minimising bypass reduces E_T. The only exception to this is for the simulations in which $\emptyset$ failed to fully converge (those with Server Fan Option 2, low bypass, and $\Delta p_{CH} < 1$). The

reduction in E_T resulting from minimising bypass ranges from 2.0 to 12.5% depending on the Server Fan Option and Δp_{CH}.
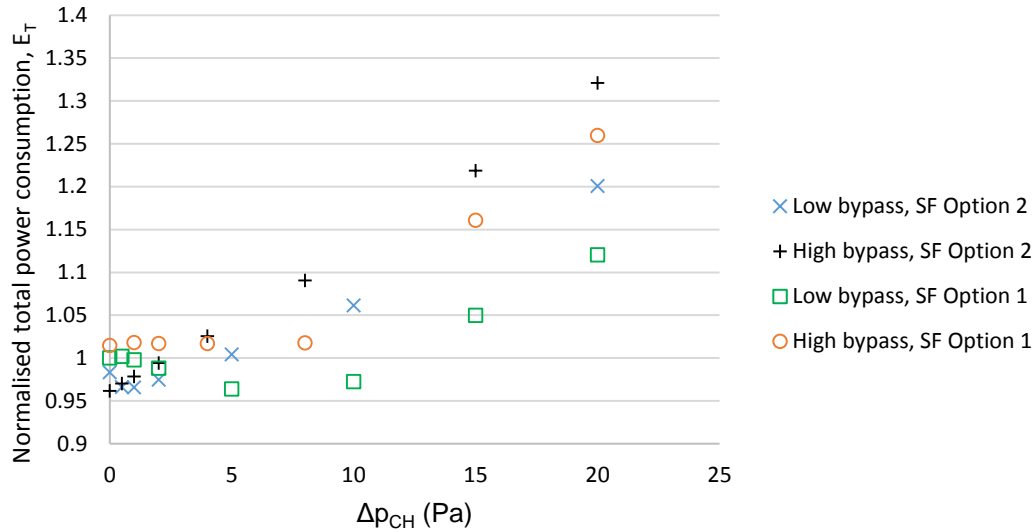
Figure 5-38. Variation of E_T with Δp_{CH} in the PF-CFD simulation results, with different bypass conditions and Server Fan Options.

5.6.4 Computational expense

The solution time for the PF-CFD simulations averaged around 16 hours, and did not appear to be affected by the Server Fan Option, bypass conditions or Δp_{CH}. Each simulation required around 31,000 iterations to solve. By comparison, Lettieri et al.'s [170] potential flow model took 7 minutes 12 seconds to solve a simulation of a domain composed of 6.6 million elements. The PF-CFD model contained only 1.3 million elements. The number of iterations required to achieve convergence in Lettieri et al.'s simulation was only 250, 2 orders of magnitude less than in the PF-CFD model, at which point the maximum residual in the $\emptyset$ field was 2.5×10^{-3}, far greater than the target for the $\emptyset$ residual in the PF-CFD model.

The solution time for the NS-CFD simulations averaged around 22 minutes, and did not appear to be affected by Δp_{CH}. Note that this solution time is around 2% of that required by the PF-CFD model. Each simulation required only around 180 iterations to solve, which is around 0.6% of the number of iterations required in the PF-CFD simulations.

The results indicate that the time taken to perform each iteration was greater in the NS- than the PF-CFD model. This is unsurprising as the equations being solved by the NS-CFD model are non-linear, and also because the NS-CFD domain has a greater number of elements than the PF-CFD domain (1.3×10^6 rather than 1.1×10^5). However, the NS-CFD

model requires far fewer iterations to achieve convergence. Speculation on the reasons for this is to some extent limited by the limited information available regarding the methods used by the 6Sigma software package. However, it is likely that the use of more sophisticated grid design could be utilised to speed up the solution in the PF-CFD model. This could enable a coarser grid across the bulk of the domain, and a finer grid in regions such as the server inlets and outlets where smaller scale flow features become more important.

The PF-CFD model performs poorly with regards the number of iterations required to achieve convergence, in comparison with Lettieri et al.'s model [170]. Like the PF-CFD model, Lettieri et al.'s model utilises a simple tetrahedral grid, hence the discretisation procedure does not account for the discrepancy. It may be assumed that the geometry of the problem modelled by Lettieri et al., coupled with the more relaxed convergence criteria, allows convergence to be achieved with fewer iterations. In addition, Lettieri et al.'s model applies fixed server flow rates and power consumptions, and does not include containment. The calculations associated with leakage flow rates and server flow rates and power consumptions in the PF-CFD model may therefore also introduce challenges to achieving convergence and stability. Since a large number of iterations was required in the iterative procedure used to satisfy the governing equations of the PF-CFD model, it is likely that the use of a direct solution method such as Gaussian elimination would reduce the solution time [154], as discussed in section 4.2.

5.7   Summary

This chapter has described the development of potential flow and Navier-Stokes CFD models of data centres employing aisle containment. A simple potential flow model was successfully validated against results of a similar model described in the academic literature (Toulouse et al. [145]). The geometry of this model was then modified to represent the Test Data Centre, and a NS-CFD model of the Test Data Centre was also developed. These models have been validated using the experimental results from the Test Data Centre, as presented in Chapter 2. The validation has shown that the models are fairly successful at predicting inlet temperatures, with accuracy being comparable with that of the most accurate models presented previously within the academic literature. Temperatures elsewhere in the domain are not predicted particularly accurately. The NS-CFD model was generally more accurate than the PF-CFD model in predicting temperatures elsewhere in the domain than at the load bank inlets. However, since prediction of server inlet temperatures is the most important function of a CFD model of a data centre, failure to accurately predict

temperatures elsewhere in the domain does not fundamentally compromise the value of the models.

As discussed in Chapter 4, there has been only one CFD model presented in the research literature previously which represents a data centre utilising aisle containment, whilst incorporating bypass/recirculation within racks (rack leakage) [72]. The factors governing the extent of rack bypass were not experimentally verified, and there was no discussion of ΔP_{CH} or of the relative importance of rack and containment bypass. The work presented here furthers the existing knowledge on data centre CFD modelling by demonstrating an appropriate method for modelling rack bypass. However, noting the likely variation of the success with which rack sealing is achieved in different facilities, it is likely that some calibration of the equations governing rack bypass (for example by adjusting the constants in the equations governing bypass, Eq. 2-28 and Eq. 2-33) would be necessary in individual applications.

The PF-CFD model described here is the first to represent a data centre employing aisle containment. As discussed in section 4.2, PF-CFD models typically take less time to solve than NS-CFD models. This has not been the case here (the PF-CFD simulations of the example data centre geometry took around 16 hours, as opposed to 22 minutes for the NS-CFD simulations). However, it is likely that the solution time of the PF-CFD model could be reduced, for example by employing a direct solution method to satisfy the governing equations, such as Gaussian elimination [154] (as discussed in section 4.2). Hence, demonstrating that PF-CFD models can accurately predict inlet temperatures in data centres employing aisle containment is significant, as it implies a potential to produce less computationally expensive CFD models of these data centres than have been demonstrated previously.

The modification of the models to represent an example data centre geometry has enabled a more detailed discussion around the appropriate level of Δp_{CH} in data centres employing aisle containment. The system model presented in Chapter 3, when used in isolation, is unable to account for the effect of recirculation on server inlet temperatures, and of variations in the pressure field on flow rates through and power consumptions of servers. These factors affect the supply air flow rate and temperature required for adequate cooling, as well as server power consumptions. This chapter represents the most detailed investigation of the appropriate level of Δp_{CH} to date, and the methods used in the models presented could be used to guide the appropriate setting of this parameter in an operational data centre. The appropriate level of Δp_{CH} was shown to depend strongly on the pressure

field and on the model adopted for server behaviour. The pressure fields predicted by the two models differed significantly from one another, with the NS-CFD model showing much greater variation in pressure differential across the partition along the length of the aisle. Experimental validation would be required to determine with certainty which model predicted the pressure field more accurately. However, it is likely that the lack of viscosity and conservation of momentum within the PF-CFD model are the main causes of this discrepancy. Hence, since the NS-CFD model in these respects gives a more accurate representation of the behaviour of the fluid, the author suggests that the NS-CFD model's predictions of the pressure field are likely to be the most accurate. Should this be the case, it suggests that there may be limitations to the value of PF-CFD simulations of data centres employing aisle containment, at least for data centres with geometries similar to that investigated. However, it is worth noting that the PF-CFD model achieved similar accuracy to the NS-CFD model when used to predict load bank inlet temperatures in the Test Data Centre (as demonstrated in section 5.5.3).

The strong impact of the choice of Server Fan Option on the relationship between Δp_{CH} and E_T (see Figure 5-38) demonstrates that further work to determine the behaviour of servers in response to changing pressure and temperature conditions is required to further the discussion around identifying the optimum level of Δp_{CH}.

The large variation in pressure within each aisle predicted by the CFD models of the example data centre geometry indicates that, in practice, it may be necessary to measure pressures at numerous positions within each aisle in order to ensure appropriate pressure conditions are achieved. This result also raises the potential importance of consideration of the impact of data centre design on the variation of pressure within each aisle. Designs encouraging minimal variation in pressure would be beneficial since they would enable minimisation of bypass and excess air flow through servers (caused by excessive positive pressures), whilst also ensuring that adequate cooling is achieved.

Regardless of the details of the pressure field and server behaviour, the results from the modified system model show that minimising leakage reduces the level of Δp_{CH} required to achieve adequate cooling for all servers, and therefore reduces CRAH fan power consumption and E_T. This is consistent with the findings of the system model as presented in Chapter 3, and shows that these findings hold where air flows are modelled.

The NS-CFD model of the example data centre geometry achieved convergence within a much shorter time than the equivalent PF-CFD model. This was due to the NS-CFD model requiring far less iterations to achieve convergence. The time required for solution of the

PF-CFD model is relatively large, and this would be increased for the larger domains prevalent in real data centres. Hence, improving the speed of solution would be useful in expanding the range of uses for the PF-CFD model.

By contributing to the ongoing development of methods in CFD modelling of data centres employing aisle containment, the work presented in this chapter has fulfilled objective (3), as identified in section 1.5. The analysis of the results of the CFD models of the example data centre geometry has also contributed to the fulfilment of objective (2), that is:

'To investigate the implications of bypass and recirculation for electricity consumption in data centres employing aisle containment.'

6. POLICY INSTRUMENTS FOR DATA CENTRE ENERGY EFFICIENCY

This chapter investigates the use of policy instruments to drive progress on energy efficiency in data centres in the UK. Section 6.1 provides a review of the literature regarding barriers to, and drivers for, energy efficiency, in the data centre sector and in other industries. Section 6.2 provides a literature review of social research methods which are relevant to the investigation of barriers and drivers. Section 6.3 describes the methods used in a series of semi-structured interviews conducted to investigate attitudes towards policy instruments impacting the UK data centre sector. The findings of these interviews are presented in section 6.4, and in section 6.5 the findings are analysed with respect to the information presented in section 6.1. In section 6.6, some recommendations are made for potential changes to the current policy environment relating to energy efficiency in the UK data centre sector. This work seeks to fulfil objective (4), as defined in section 1.5, which is:

'To investigate the potential for policy instruments to drive energy efficiency improvements in the data centre sector.'

6.1 Literature review – barriers and drivers

6.1.1 Defining barriers and drivers

The preceding chapters have investigated some of the technical challenges related to energy efficient management of data centres, particularly in relation to air management and aisle containment. As discussed in section 2.1.1, aisle containment has long been understood within the industry to be a cost effective method for improving the efficiency of cooling. Despite this, uptake of this technology and other energy efficiency measures (EEMs) has been slow, as will be discussed in section 6.1.6.2. The concept of 'barriers' is useful in understanding the reasons for this.

Sorrell et al. [180] define a barrier as "a postulated mechanism whose outcome is an organisation's neglect of (apparently) cost effective energy efficiency opportunities". Addressing these barriers is useful both to an industry and its customers, and to society as a whole due to the potential to reduce energy consumption and greenhouse gas (GHG) emissions.

The word "driver" is typically used simply to denote the opposite of a barrier, i.e. any factor which increases the likelihood of adoption of energy efficient behaviour [181]. Thollander and Ottosson [181] further specify that a factor should be seen as a driver if it encourages

investment in technologies which are both energy efficient and cost-effective. Cagno & Trianni [182] stress that a driver may encourage the adoption of either energy efficient technologies or of energy efficient practices. Sorrell et al. [180] have noted the importance of considering the nature of the key barriers in a particular instance when designing public policy drivers, to ensure the success of the latter.

Traditionally, barriers to energy efficiency are seen as only demanding policy intervention if they can be classed as market failures, i.e. if they prevent the implementation of genuinely cost-effective EEMs [183]. Note that an EEM may be cost-effective from the perspective of the economy as a whole, despite not being so from the perspective of an individual organisation. This is an important consideration with regards to the environmental impacts of energy generation, which may incur costs borne by society, rather than directly by the energy consumer. Hence, an organisation may not be responsible for paying the full cost of energy consumption, making societally beneficial EEMs unattractive to the organisation [183]. Environmental costs in this context are referred to as externalised costs.

It should be noted that there are instances when the 'market failure' model may be inadequate for assessing the need for policy intervention. This is particularly important where there is a need to foster innovation and create new markets, as opposed to creating policy to encourage particular behaviours within an existing paradigm [184].

6.1.2   The energy efficiency gap

There is widespread agreement that barriers preventing cost-effective EEMs do exist, as evidenced by legislation and strategies implemented by governments around the globe to overcome such barriers [185]–[187]. The problem is also widely acknowledged in the academic literature [180], [183], although some comment that the potential for truly cost-effective EEMs may in some cases be exaggerated due to underestimation of the full costs of an EEM, such as opportunity costs, high discount rates resulting from uncertainty, and person-hours associated with implementation [183], [188]–[191]. Another area of debate and uncertainty is how to identify the true cost of energy, so as to account for externalised costs such as climate change [183].

As a result, there is no agreement on the true size of the "energy efficiency gap" – the improvements in energy efficiency which are achievable through cost-effective means [188]. However, some estimates have been made. Granade et al. [192] have estimated that the total demand for energy in the United States in 2020 could be reduced by 23% from a business as usual (BAU) scenario if EEMs with a net positive economic value were fully exploited across all sectors. Note here that externalised costs such as environmental

impacts were not included in the calculations, which would likely have increased the extent of the identified opportunities for EEMs. Similarly in the EU, a target has been set to achieve a 20% reduction in energy consumption by 2020 through EEMs which are expected to be economically beneficial (here, externalities such as environmental impacts are to be accounted for as far as possible) [193], [194].

No specific estimates have been made of the size of the energy efficiency gap in the data centre sector. However, there are indications that a gap exists. For instance, surveys have suggested a great gulf between the more efficient and less efficient sites, as discussed in section 1.4 [6], [29]. However, it is theoretically feasible that this variation represents the different business models involved, with some facilities finding that less efficient technology and practices, which may be more reliable, are financially preferable.

One indication of the perceived potential for efficiency improvements in the data centre sector is given by the UK's Climate Change Agreement covering the sector (DC CCA), which came into force in 2014 [22]. The programme is available only to co-location facilities, and offers participating sites a 90% rebate on the Climate Change Levy (CCL) [22]. The first phase of the CCA (prior to the introduction of the DC CCA) also granted participants exemption from the Climate Reduction Commitment (CRC), although this exemption has since been revoked. Both the CCL and CRC are essentially taxes on carbon emissions, and are discussed in more detail in section 6.1.4. The rebate is conditional upon a 15% reduction in PUE across the sector by 2020, from a 2011 baseline.

The aim of the broader CCA programme is to bring each participating sector's energy efficiency 60% of the way from a projected BAU scenario to one representing all cost-effective measures having been implemented [195]. However the scheme has been criticised for underestimating BAU energy efficiency improvements, and for being vulnerable to industry representatives playing down the potential for improvements, leading to unambitious targets [195]. These criticisms have been backed up by the prevalence of sectors exceeding the targets given during the first stage of the program, which ended in 2010 [195]. Hence, the possibility exists that the energy efficiency gap for the data centre sector exceeds that represented by a 15% reduction in PUE. It should also be noted that, as discussed in section 1.4, the PUE metric only represents the efficiency with which heat is removed, and that there are also likely to be opportunities for cost-effective EEMs related to IT efficiency.

6.1.3   The rebound effect

The rebound effect refers to the tendency for the rate of consumption (and therefore production) of a product or service to change in response to changes in the efficiency of the provision of that product or service [196]. Most commonly, it refers to the tendency for increases in efficiency to reduce the cost of a product or service, therefore increasing demand for it. Greening et al. [197] provide a review of evidence of the size of the rebound effect, concluding that it varies widely, from 0-100% depending on the product or service of interest, but rarely exceeds 50%. Greening et al. also refer to the potential for efficiency improvements relating to one product or service affecting demand for other related products or services (secondary and economy-wide effects). The authors concede that measurement of the effects of energy efficiency programs on energy consumption beyond the field in which the efficiency measures were implemented has so far yielded no reliable results.

Related to the rebound effect is Jevons's paradox (or the Khazzoom-Brookes postulate). Jevons's paradox suggests that efficiency improvements tend to lead to price reductions for associated goods and services, increasing demand for those goods and services as well as related goods and services in the wider economy, to the extent that the ultimate effect is to increase total energy consumption, rather than reduce it [198]. The theory is supported by neoclassical economic theory, and has been famously observed in the use of steam turbines during the industrial revolution [198]. Efficiency improvements in steam turbines led to advances in coal mining and reduced cost of rail transport, which lowered coal prices and transport costs, leading to rapidly increasing coal consumption [199].

Sorrell [199] notes that commonly cited examples of Jevons's paradox relate to energy-intensive technologies in an early stage of development, and which have a wide range of applications. He suggests that the same results may not be seen in other instances of energy efficiency improvements, for example where energy is not a major cost, or where there is little potential for expansion of use of the good or service into other sectors. The proportion of data centre expenditure accounted for by energy consumption is similar to that for traditional energy intensive industries, as will be shown in section 6.1.5. In addition, the data centre sector provides an increasingly broad range of services to a broad range of customers. Hence, it seems that Jevons's paradox is relevant to this sector.

Nguyen et al. [200] and Figuerola et al. [201] have both cited Jevons's paradox as a reason to focus on greening the generation of electricity used by data centres, rather than improving the efficiency of facilities. However, if recent growth trends in the data centre

sector continue unchecked, the sector could consume 855 TWh of electricity per year by 2040, which is likely to amount to around 2.5% of total global electricity consumption, as shown in section 1.2.1. Whilst the proportion of electricity generated from renewables is likely to increase, there are limits to the likely growth, with the IEA predicting that 37% of electricity demand could be accounted for by renewables by 2040 [18]. Hence it is clear that, were this approach pursued, data centres could become a significant consumer of the available renewable electricity resource. The limits to the likely capacity for renewable energy generation in the coming decades mean that unchecked growth in data centre electricity consumption, renewable or otherwise, is not compatible with reducing emissions sufficiently to avoid dangerous climate change [202], [203].

However, it should also be noted here that, as discussed in section 1.2.1, data centres provide services which enable reductions in the environmental impact of other sectors [25]–[28], [204]. It has been suggested that it would make economic sense to subsidise companies operating these kinds of services, since this may represent an efficient path to meeting energy conservation targets [27]. Although the extent of the relevant supported energy savings has not been quantified, it is clear that this issue further complicates efforts to understand the full impact of energy efficiency improvements in the sector.

One potential approach to preserving the energy savings produced directly by industrial energy efficiency programs is to complement them with rises in energy prices, whether through taxation or trading schemes [199], [205].

6.1.4 Barriers and drivers in industry

Extensive research exists into barriers to the introduction of EEMs in the energy intensive industries. The identified barriers vary somewhat between industries and in different studies, but often include:

- The cost of process disruption during introduction of EEMs [206]–[208]
- Prioritisation of operational control over energy efficiency [206], [208], [209]
- Uncertainty surrounding future energy prices [206]
- The need for front line staff wanting to implement changes to gain approval from management [206]
- Slow return on investment/high capital costs [206]–[210]
- Lack of knowledge of new technologies [206], [207], [210]

Some authors have noted that the importance of different barriers varies between organisations with different characteristics. For example, larger firms have been shown to

be less affected by barriers relating to time, accessing information, staff/skills and financial resources [207], [210], [211]. This is likely due to larger companies having more specialised staff with skills relevant to the technologies in question and/or the time available to devote to researching them. On the other hand, smaller companies have an advantage in being better able to keep personnel aware of energy efficiency issues, with larger companies perceiving a lack of awareness among personnel to be a more significant barrier [210].

Generally, the key drivers for energy efficiency identified within the literature include:

- Provision of finance [182], [212]
- Access to efficiency experts [182], [212]
- Information sharing [182], [212]
- Energy performance contracts [182]
- Staff with ambition [182]
- Access to low cost consultancy and auditing [182], [212]

It is clear that these drivers relate directly to the key barriers identified earlier in this section. For example, the 'provision of finance' relates to the 'slow return on investment/high capital costs' barrier, and 'information sharing' relates to the 'lack of knowledge of new technologies' barrier.

Cagno & Trianni [182] note that these drivers are usually identified via questionnaires or interviews with stakeholders, meaning that they show only the opinions of these stakeholders. The true importance of each driver is hard to discern.

Tanaka [213] notes that policy interventions aimed at driving adoption of EEMs can come in three forms:

- **Prescriptive policies**, such as regulations and voluntary agreements,
- **Economic policies**, such as taxes, tax exemptions, cap and trade schemes, subsidies and loans, and
- **Supportive policies**, such as provision of audits and distribution of technical information.

Note that a voluntary agreement may be seen as a prescriptive policy in the sense that it typically prescribes a specific set of standards to be met, whilst participation in the scheme is voluntary.

One example of **prescriptive policies** is legislation to set minimum efficiency standards for energy intensive equipment. Such legislation is only suitable for situations where differences between individual models of a piece of equipment are small, such that the

efficiencies of different models can be easily compared. The potential of such legislation to reduce energy use can be high for suitable products. The implementation and maintenance of these standards can be costly due to the level of technical knowledge required and the need for revision as technology changes [213].

Prescriptive policies can also come in the form of negotiated agreements, whereby targets for energy use intensity are set after consultation between government and industry. Compliance is usually rewarded with tax breaks and/or failure to comply is punished by public identification. The impacts of these measures vary greatly, depending on the level of ambition shown in targets, and in the strength of rewards/punishments.

One example of a negotiated agreement used in the UK is CCAs. These are agreements which grant participating companies a reduction in payments under the CCL if they are able to meet an agreed target related to energy consumption [214]. The CCL is a tax payable on energy consumption in commerce, agriculture and the public and service sectors, and is an example of an economic policy [215]. It was set at 0.559 pence per kWh for electricity from April 2016 [216]. The first phase of the CCAs also granted participants exemption from the CRC, which is a policy requiring participants to buy allowances for every tonne of carbon they emit, and targets emissions not covered by the EU Emissions Trading Scheme (EU ETS, discussed later in this section) or CCA [214].

The aim of the CCAs is to protect energy-intensive industries from becoming unable to compete with international competitors due to increased energy costs associated with the CCL and CRC [217]. The energy use targets defined by the CCA are agreed in negotiations between representatives of the relevant sector and the Environment Agency (originally the Department for the Environment and Climate Change) [217], [218]. The targets may relate either to absolute energy consumption or to efficiency metrics, but are ultimately designed to foster energy efficiency improvements [219], [220]. The CCA scheme is set to run until 2023 [220]. Since July 2014 there has been a CCA in place for the UK's co-location data centres, which will be discussed specifically in section 6.1.6. TechUK (a body representing the interests of the information and communications technology industry) has commented that the data centre sector is an appropriate sector for such an agreement, due to its energy-intensive nature and vulnerability to international competition [28].

Under the first cycle of CCAs, the vast majority of participating companies met their targets, and in most sectors targets were comfortably exceeded [219]. This suggests that the targets were not ambitious, and may not have resulted in the uptake of EEMs beyond that which would have occurred under BAU. The goal of the scheme was to set targets so as to

stimulate energy efficiency improvements which lay 60% of the way from BAU to all cost effective measures being implemented (after ignoring issues of capital availability and management time investment) [221]. An investigation into the performance of organisations within and without the CCA scheme has suggested that the energy efficiency improvements observed under the CCA did not exceed BAU, and that the CCA may have acted to increase energy consumption, whilst the CCL drove improvements where CCAs were not in place [222], [223]. This was in contradiction to earlier research, which compared the actual performance of companies participating in the CCA with the projected performance of those companies had the CCA not been in place. This research concluded that the CCA resulted in a greater improvement in energy efficiency than would have been achieved by the CCL alone, despite unambitious targets being in place [219]. According to this analysis, over-achievement of targets was achieved as a result of the scheme's stimulation of greater sharing of information relating to energy efficiency.

Research has suggested that the CCL, both with and without the CCA, has had either no effect on employment levels and revenue [223] or has increased them [219]. This calls into question one of the stated aims of the CCA, which is to protect the relevant industry sectors from being unable to compete with their international competitors as a result of the impact of energy taxes [218].

The Environment Agency's most recent report on the CCA scheme found that only around half of participants met their targets during the 2013-2014 period [218], which would seem to suggest that targets have become more ambitious since the first cycle of the scheme. The majority of those failing to meet the targets chose to compensate by paying a buyout fee of £12 per tonne of carbon dioxide equivalent (CO_2e), which allowed them to remain in the CCA scheme and avoid paying the full CCL.

Prescriptive policies have also been used effectively in China to drive energy efficiency [224]. Examples include the Energy Efficiency Obligation (in which companies are set targets for energy conservation) [224], and the application of electricity surcharges to companies using inefficient technologies [225]. The latter policy, similar to the combination of CCA and CCL, was estimated to have reduced electricity consumption in the participating companies by almost 20% [225].

Another form of prescriptive policy is the introduction of minimum performance standards for equipment and appliances [226]. Such policies have been shown to be able to achieve a strongly positive environmental impact, sometimes at net-negative cost, particularly when used in conjunction with product labels used to inform customers about the energy

consumption of the product [226]. The EU's Ecodesign includes stipulations relating to server energy consumption which fall into this category, as will be discussed further in section 6.1.6 [227]. Building codes (regulations setting requirements on energy efficiency for buildings) also often have a net-negative cost [226].

One perceived advantage of prescriptive policies is that there is a greater degree of certainty over their impact than with other policies, since the effect of targets being met is relatively easy to quantify compared with the effect of economic and supportive policies [28]. However, in the case of voluntary agreements, the impact depends on take up. In addition, the extent of the rebound effect and impacts on the wider economy may be difficult to predict.

Voluntary agreements have been used extensively since the mid-1990s within the EU to encourage energy efficiency, with their use having been reviewed by Rezessy & Bertoldi [228]. One such agreement is the EU's Code of Conduct On Data Centre Energy Efficiency [4], which is discussed in section 6.1.6. Most of these agreements have focused on industrial processes, although more recently voluntary agreements covering the service and transport sectors have been introduced. Some schemes have included incentives to join such as deferral of regulation or taxation, assistance with funding for energy audits and implementation of EEMs, exemptions or reductions from energy or carbon taxes, provision of information on energy efficiency and training in energy management. Some agreements have been criticised for having unambitious targets [228]. Rezessy & Bertoldi stress that effective systems for monitoring compliance (including either spot checks or third party involvement in monitoring) and for evaluating the results are important to ensure that a voluntary agreement delivers real results [228]. Attempts to assess the success of voluntary agreements in terms of driving efficiency improvements which would not otherwise have occurred have been "virtually non-existent", although reductions in energy consumption have been achieved widely [228]. It has been suggested that voluntary agreements are popular legislative tools due to them receiving far less resistance from industry than mandatory policy instruments, enabling tougher targets to be imposed [228]. The benefits of voluntary agreements include dissemination of information and the reduction in information asymmetry between industry and political organisations [228].

Economic policies have the potential to significantly reduce energy consumption, and allow flexibility in how individual companies choose to achieve compliance [213]. These policies can also be less costly to implement than prescriptive and supportive policies, since they

require little understanding of the routes to improved energy efficiency on the part of policy makers [213].

One example of an economic policy approach is the cap and trade scheme, such as the EU ETS in which a limit was placed on GHG emissions from large emitters, with emissions permits distributed accordingly and traded between emitters [229]. Like other cap and trade schemes, the EU ETS suffered from an initial over-provision of permits, limiting its impact [230], [231]. This can cause uncertainty over the value of EEMs, leading to companies investing in EEMs immediately prior to the introduction of a trading scheme, only to withdraw investments after it becomes clear that permit prices will be lower than expected [230]. In the case of the EU ETS, over-provisioning of permits was possibly caused by concerns within each of the individual EU member states over the potential for under-provisioning to lead to a competitive disadvantage for industries within the country in question, which resulted in each country circulating an abundance of permits. This caused the price of permits to collapse, such that there was little incentive to reduce emissions [231].

Estimates of the effectiveness of the EU-ETS vary. Early studies in which a counter-factual emissions baseline was calculated to estimate the likely emissions without the presence of the EU ETS, and compared with the actual emissions, suggested that significant abatement was achieved within the first stage of the scheme (2005-2007) [232], [233]. A similar analysis of the period from 2008-2009 found that very little abatement was caused by the scheme [234]. A more recent study sought to quantify the importance of various economic and policy factors on GHG emissions under the EU ETS by analysing changes in emissions in response to these factors, from 2005-2012. The EU ETS was found to account for some abatement of emissions (a reduction of around 2%), although it was far less significant than the impact of the economic crisis during this period [235].

In the UK, the EU ETS is complemented by the CRC Energy Efficiency Scheme, which was introduced in 2010 [236]. It applies to organisations with electricity consumption exceeding 6000 MWh per annum, and whose electricity consumption is not covered by the EU ETS [237], and as such applies to any data centres operated by such organisations. Participants must purchase allowances to emit the CO_2 associated with their electricity and gas consumption, with the allowances being available for purchase at a fixed price during 2 annual sales periods.

Supportive policies are low cost, and whilst their effectiveness is generally considered to be limited when working in isolation, they may be useful in maximising the cost-effectiveness

of prescriptive and economic policies used in parallel [213]. Such policies have received little attention in the peer-reviewed literature, although Ürge-Vorsatz et al. [238] found evidence that schemes to label and certify products according to their energy efficiency and to make energy audits mandatory, could achieve significant energy savings in a cost-effective manner. Information campaigns, awareness raising and public leadership programs (whereby organisations in the public sector are developed as examples of best practice) were found to be less effective, whilst still achieving significant savings at acceptable cost [238]. Ürge-Vorsatz et al. [238] note that methods for estimating the impacts of these kinds of programs are not well developed. Others have shown that the introduction of subsidies for energy audits can be successful in increasing their frequency [212].

6.1.5   Barriers to energy efficiency in the data centre sector

Chai & Yeo [207] note that, within industry in general, the motivation to engage with energy efficiency is dependent upon the proportion of expenditure on energy costs. Historically, energy bills for data centres have been small in comparison with IT budgets, although they are becoming increasingly significant [239]. Analysis by APC [240] suggested that around 20% of the total cost of operation may be accounted for by electricity costs, and Salim & Tozer [6] reported anecdotal estimates that 50% of a data centre IT department's budget may be spent on energy. Hence it is likely that electricity costs are now an important consideration for data centres. Indeed these electricity costs are comparable with energy expenditure in traditional energy-intensive industries such as cement, steel and aluminium production, in which spending on energy ranges from 15-40% of total expenditure ([241], [242], IHS Global Insight Inc., cited in [241]).

Recent sources suggest that the key barriers to energy efficiency in data centres are:

- **changing power demands** ([239], Energy Market Innovations, 2009, cited in [243]),
- **prioritisation of reliability over efficiency** (Energy Market Innovations, 2009, cited in [243], Cullen et al. 2009, cited in [243]) [28], [239], [244], [245],
- **split incentives** (Cullen et al. 2009, cited in [243]) [28], [244]–[246],
- **high capital costs** (Cullen et al. 2009, cited in [243]) [244]–[246],
- **poor communication between facilities and IT staff** (Energy Market Innovations, 2009, cited in [243]) [245],
- **requirement for central management approval** (Cullen et al. 2009, cited in [243]),
- **lack of knowledge and understanding** [239] [244],
- **prevalence of inefficient, legacy data centres** [246],

- **short leases on data centre sites**, which reduce the payback period over which EEMs become attractive [28], and

- **disincentives for trade allies** [244].

Note that these are largely similar to the most common barriers cited in relation to other industries (see section 6.1.4). However, issues such as changing power demands, split incentives, poor communication, prevalence of inefficient legacy facilities, short leases on sites and disincentives for trade allies are not commonly cited as barriers in other industries. Each of the identified barriers are discussed in the following text.

The rapid rate of replacement and increasing power density of IT equipment produces rapidly **changing demand for power** distribution and cooling infrastructure, making the specification of energy efficiency challenging. This is exacerbated by long lead times for delivery and installation of new equipment (Energy Market Innovations, 2009, cited in [243]). In addition, data centres tend to deal with their rapidly increasing infrastructural demands, and the attendant uncertainty of future demands, with large tolerances for future expansion. This leads to data centres often running well below capacity, which tends to lead to inefficiencies [239].

The sector's **prioritisation of reliability** is in opposition to attempts to improve efficiency. Daim et al. [8] found that energy efficiency was considered less important than availability and security of data centre services by a majority of managers within various industries. Efficiency improvements can often present a risk of downtime, or even entail planned downtime [247]. Also, measures such as the introduction of redundancy and over-provision of cold air are needed to maximise reliability, both of which impair efficiency (Energy Market Innovations, 2009, cited in [243], Cullen et al. 2009, cited in [243]) [30], [32], [239], [244]. Security and reliability concerns also often lead to restrictions in the potential for surveying data centres in order to determine thermal characteristics, which hinders attempts to monitor and improve efficiency [36].

This focus on reliability also has an impact on the equipment supplied by trade allies. For example, power ratings of IT equipment are often higher than the true peak power consumption, since the manufacturers wish to avoid culpability for equipment failure in case of insufficient power supply. True server power consumption typically averages around 50-60% of the nameplate power [146]. This leads to over-estimation of IT power consumption, and the over-provisioning of power and cooling infrastructures [33].

Split incentives occur where parties responsible for behaviour or decisions which impact on energy consumption are insulated from the effect of this behaviour on energy bills [248],

[249]. Legislation to improve energy efficiency by increasing energy prices may be impaired as a result of this insulation. Split incentives are present in many fields, with researchers having investigated their effects on company car usage [249], utility company relationships with energy users [248], energy use in rented housing [250]–[252], and housing construction [183].

Split incentives exist in data centres because of their organisational structures. For example, the IT department may be responsible for purchasing equipment for a data centre, but not for electricity bills which will be affected by the efficiency of this equipment. Facilities departments are usually responsible for electricity bills. The Uptime Institute [29] reports that IT departments are responsible for data centre electricity bills in only around 16% of cases. The Uptime Institute recommends that this should be addressed, such that those with the responsibility for the running of data centres are more directly accountable for the electricity usage.

In co-location data centres this situation may be exacerbated since the relevant IT department is part of the customer company, and the facilities department part of the co-location provider. However, some co-location providers charge customers for the electricity they use. In this case, the party responsible for maintaining the cooling and power infrastructure is insulated from the impacts of their decisions on energy efficiency; hence, a split incentive still exists (Cullen et al. 2009, cited in [243]) [244]. Cullen et al. (2009, cited in [243]) report that there are 3 main pricing structures used by co-location providers, as follows:

- "Space-based pricing" – tenants pay on a rack or square footage basis, with charges for power only if they exceed an agreed limit.
- "Space- and power-based pricing" – the payments for space and power are separated, with the power cost either based on an agreed limit, or following actual consumption. Clearly where the energy bill relates to actual consumption, there is a greater incentive for the customer to improve efficiency.
- "Cost-plus pricing" – all facility costs are documented, with tenant charges being calculated accordingly.

There is a trend towards incorporating energy bills into pricing schemes (Cullen et al. 2009, cited in [243]). Howard & Holmes [243] suggest that the ideal solution is to charge tenants a fixed rate for space, as well as charging directly for power consumption. Thus, the tenant is directly incentivised to improve efficiency, and the provider is incentivised indirectly, since improving efficiency would enable them to minimize cooling/power infrastructure, or to

maximize IT capacity. Unfortunately most co-location providers charge tenants according to a maximum allowable power consumption, meaning that there is no incentive for the customer to make incremental improvements [243].

Vernon & Meier [248] note that policy responses to split-incentives generally come in three forms:

- Specialized contracts which realign the incentives
- Regulation to enforce efficiency improvements
- "All-in" services, whereby a program is put in place to assist energy consumers with the knowledge and expertise required to reduce energy consumption, as well as providing loans and grants

Blumstein [253] notes that the split-incentive inherent in entrusting utilities to deliver energy efficiency programs is addressed in some US states by adjusting rates such that if sales fall below expected levels, utility earnings remain at the same level. Rewarding utilities according to the results of efficiency programs has also been suggested, although Blumstein [253] notes that the exact effect of a program is impossible to measure, since the outcome of inaction cannot be determined.

Capital costs are high for many EEMs (Cullen et al. 2009, cited in [243]) [244]. This is prohibitive, since data centres can require payback periods as short as 12 months [244], although 2-3 years may be acceptable in some cases [8].

Improving a data centre's efficiency requires co-ordination between facilities and IT departments, since the decisions of one department have an impact on the other [239]. **Poor communication** between the 2 departments has been identified as a barrier to efficiency improvements (Energy Market Innovations, 2009, cited in [243]) [146]. This may be particularly problematic for smaller data centres [6].

Data centre managers are generally familiar with, and understand the importance of, IT monitoring. This is less true of monitoring of the facilities of the data centre. Improved facilities monitoring, particularly when undertaken holistically with IT monitoring, can lead to efficiency gains. This is because it allows the cooling to be optimized based on IT activity, such that total energy usage can be minimized [254].

Chains of command present in data centres often mean that **central management approval is required** before EEMs can be implemented (Cullen et al. 2009, cited in [243]). This can slow down the implementation of EEMs, or even prevent implementation altogether.

Data centre professionals often have a **lack of knowledge and understanding** of the financial and environmental benefits which can result from improvements in energy efficiency [239]. Understanding of the most modern, efficient technology may also be limited, with managers being wary of using unfamiliar technology due to concerns over reliability. Data centre design and management practices are often subsequently based on intuition and/or experience, with a tendency to lean towards excessive conservatism and consequent inefficiency where there is doubt [36].

This barrier is exacerbated by the fact that individual efficiency measures may not necessarily have an additive impact on reduction of electricity usage; indeed, they may cancel each other out. For example, as discussed in section 1.3.3, increasing supply air temperatures may reduce electricity consumption in the cooling infrastructure whilst increasing server power consumption. It is therefore essential for a proposed EEM to be evaluated holistically before implementation [33].

This barrier may be strongest with smaller data centres, which are less likely to participate in seminars at which best practice is shared [6].

Much of the existing stock of data centres consists of **inefficient, legacy** facilities. Such sites often present logistical challenges which can make retrofit measures to improve efficiency expensive or impractical [246].

Disincentives for trade allies, such as companies supplying equipment to data centres, may act to prevent the provision of efficient equipment. For example, server manufacturers may wish to discourage virtualization and consolidation since this could reduce server sales [244].

As identified in section 6.1.4, barriers and drivers may vary between organisations of different sizes, in different countries, as well as on other factors. This is likely to be true in the data centre sector. Specifically, it is likely that co-location facilities will experience different barriers from enterprise data centres, due to their different business models and levels of skills and experience. Hence, it is useful to comment on the relative importance of enterprise and co-location data centres. The United States Environmental Protection Agency (US EPA) (2012) report that, according to industry analysts, utility scale and enterprise facilities (typically those exceeding 5000 square feet) account for approximately 40% of data center electricity consumption. Data Center Dynamics' [255] recent census estimated that the co-location sector is responsible for 23% of data centre electricity consumption in the UK. The Uptime Institute (2013) reports that organizations are increasingly buying in data centre services from co-location providers. 63% of respondents

to their 2013 industry survey reported greater than 10% increases in co-location data centre expenditure, with only 25% reporting such an increase in expenditure on in house facilities. Their report goes on to state that co-location facilities tend to be more focused on efficiency, partially because of more frequent reporting of data centre performance to high level executives. Whilst the foregoing does not provide a definitive identification of the proportions of the data centre sector accounted for by co-location and enterprise facilities, it is clear that both make up a significant proportion of the sector.

6.1.6 Policy instruments for driving efficiency improvements in the data centre sector

A number of policy instruments affect data centres in the UK. The most important are the UK Government's climate change umbrella agreement for the data centre sector (DC CCA), and the EU's Code of Conduct on Data Centre Energy Efficiency (EU CoC). These are examples of economic and prescriptive policies, respectively.

6.1.6.1 The Climate Change Agreement for Data Centres

As described in section 6.1.4, the CCA scheme aims to drive energy efficiency improvements in targeted industries by setting targets either for energy efficiency improvements or reductions in energy consumption. Organisations meeting these targets are granted a 90% rebate on the CCL energy tax. In the case of the DC CCA, the ultimate goal is to achieve a 15% reduction in PUE across the sector over the lifetime of the scheme. Individual facilities have different targets, with data centres which already have a low PUE being granted more modest targets in order to avoid penalisation of early adopters [28]. The scheme applies only to co-location providers.

The primary function of the scheme is to increase the financial incentive for EEMs. Hence, the scheme may be considered to target two of the barriers to EEMs in the data centre sector: **prioritisation of reliability over efficiency**, and **high capital costs**. The reliance of the scheme on PUE as a metric for efficiency means that it only incentivises improvements in the efficiency of heat removal, with IT efficiency being unaffected. Additionally, the restriction of the scheme to co-location providers limits its coverage.

A report by TechUK into the results for the first target period of the DC CCA (01/07/14-31/12/14) found that 86 of the 98 sites participating in the scheme achieved the milestone of a 1% reduction in PUE during this period, whilst accepting that this was necessarily a modest target, due to the short time frame [28]. Those failing to meet the target all chose to purchase carbon credits to make up the shortfall, enabling them to remain in the scheme

[28]. Those facilities which missed the target did so because of either a loss of customers or an expansion of the facility which was not met with increased IT power consumption – TechUK comments that the missed targets therefore do not represent a "lack of good energy management" [28]. Many facilities over-achieved on the target, with the average reduction in PUE across the participants being 4.17%.

TechUK predicts that the CCA will tend to drive data centre services into the co-location market due to the improved efficiency and associated cost reduction which make outsourcing more attractive [28]. This is in contrast with approaches which simply seek to increase energy costs, which may hamper growth in the co-location market, making potential customers of co-location providers more likely to keep data centre services in-house.

TechUK also reports that, according to data they have collected from the participants of the DC CCA, the introduction of the CCA has had a number of benefits:

- an improvement in energy monitoring [28],
- an increase in prominence of energy efficiency in discussions between co-location providers and their customers [28],
- an improved case for investment in energy efficiency [28],
- better sharing of best practice [28], and
- an increased availability of data regarding energy consumption [256].

This suggests that the scheme has had some success in addressing the following barriers: **split incentives**, and **lack of knowledge and understanding**.

Another finding of the TechUK report was that the financial incentives produced by the DC CCA are not strong enough to drive investment in "big ticket" EEMs with payback periods exceeding 3 years [28]. They propose that financial support in the form of Enhanced Capital Allowances or zero interest loans such as are provided under the Energy Technology List would help in this regard [28]. The Energy Technology List and Enhanced Capital Allowance scheme enable businesses to offset the costs of products meeting certain energy saving or efficiency criteria against the business's taxable profits [257]. However, TechUK has also criticised the Energy Technology List scheme, saying that, although most of the equipment used in data centres qualifies for Enhanced Capital Allowances, there has been very little uptake, particularly amongst smaller companies [258]. This is said to be due to the high level of bureaucracy associated with the scheme, which leads to the process being outsourced to specialist consultancies in most cases.

In a response to a call for industry opinions on the UK's energy efficiency policy landscape, TechUK produced another report, providing a summary of opinions within the data centre sector regarding the CCA, in addition to discussing the energy tax landscape more broadly [256]. The body accepts that there is a need to legislate to incentivise energy efficient behaviour, although this is partly contradicted by a statement that the price of energy is sufficiently high to drive efficient behaviour without the need for additional charges. It is also accepted that reporting of energy consumption is an important feature of legislation. The primary issue raised by the report concerns the complexity of the energy tax landscape, with companies finding that energy consumption and GHG emissions had to be reported to multiple bodies under various schemes, with processes, fees and time scales, leading to a large administrative burden. It was proposed that this led to a competitive disadvantage compared with overseas competitors operating under less burdensome legislation, as well as diverting financial and human resources away from efforts to improve energy efficiency. Lack of stability in the tax landscape was also highlighted as an issue, leading to a reluctance to invest in projects requiring large capital investment, since the returns were uncertain.

The overall attitude towards the DC CCA was reported to be very positive, stating that it has been successful in driving energy efficiency improvements without disadvantaging UK data centres in comparison with their overseas competitors [256]. This is due to the tax rebate offered, which effectively makes energy efficiency measures more economically attractive, and to the reputational benefits of achieving a target for energy efficiency improvements.

Some criticisms are made of the scheme [256]. One issue highlighted is that operators bringing new whitespace online but failing to occupy it with sufficient levels of IT will not be able to achieve their targets, despite implementing good "energy stewardship". However, it could be argued that in this way the policy discourages operation at low IT load, which could in itself be described as poor energy stewardship. It is also worth considering that the need to plan for efficient operation at low IT load has been highlighted by the EU CoC, for example [34]. Other suggestions for improvement include requirements for engagement from senior management, longer term targets to enable longer term energy management strategies, and the introduction of complementary measures to assist with the provision of capital for large scale EEMs with long term payback periods. Such changes would perhaps improve the scheme's capacity to address the following barriers: **requirement for central management approval**, and **high capital costs**.

The report proposes that the CRC should be abolished, due to its high administrative burden, its tendency to encourage offshoring of energy intensive processes rather than true

efficiency improvements, and its demands on finances which could be used to invest in efficiency [256].

In this report, TechUK [256] also expressed concerns over the findings of Martin et al.'s paper [222] on the CCL (discussed in section 6.1.4), specifically the finding that the introduction of the CCA had led to smaller reductions in energy consumption than were seen in companies exposed to the full electricity price increase of the CCL. TechUK notes that the analysis used data from 2001-2004, with the first reporting milestone for the CCA's being in early 2003, meaning that the full impacts were yet to become apparent. However, even if it were the case that over time, companies signed up to the CCA would eventually achieve the same energy efficiency as those not signed up, the finding remains that these efficiency gains would be slower to materialise. It is further noted by TechUK that the CCA now represents a stronger driver due to the increased financial penalties associated with failing to meet targets, as well as the fact that results are now published, adding a reputational factor to the policy instrument which was not present at the time of Martin et al.'s analysis [256]. It is also important to note that Martin et al.'s paper does not investigate the DC CCA, which was not in place at the time of the study [222].

6.1.6.2   The EU Code of Conduct on Data Centre Energy Efficiency

In 2008 the European Commission's Joint Research Council launched a Code of Conduct on Data Centre Energy Efficiency (the EU CoC), with the intention of driving efficiency improvements by raising awareness of best practice and improving understanding amongst data centre operators and owners [239]. The intention is to encourage practices which both improve energy efficiency and reduce total cost of ownership. The document is designed to be useful not only for facilities which achieve "Participant" status, but also in providing aspirational targets for facilities which may not wish to become Participants. As of 2014 221 data centres had achieved Participant status, representing 3.2 TWh of total power consumption per year [247]. This compares with a total EU power consumption in 2014 of 2,707 TWh [13], and EU data centre power consumption of around 59 TWh in 2014, according to a study by Germany's Borderstep Institute, carried out for the country's Federal Ministry of Economic Affairs [12]. This implies that, as of 2014, the EU CoC accounted for around 5.4% of EU data centre power consumption, and 0.1% of total EU power consumption. By February 2017, the number of data centres participating in the scheme had risen to 293 [259], with no more recent estimation of power consumption being available. TechUK has reported that over half of those sites included in the DC CCA are also participants of the EU Code of Conduct for Data Centres (as of 2015) [28].

In order to achieve Participant status in the Code of Conduct, data centres must undergo an energy assessment, where major energy saving opportunities are identified. An action plan is then completed, detailing what measures will be taken to improve efficiency [239]. Participants are required to report monthly IT and total facility energy consumption on an annual basis [247]. The scheme also encourages data centre trade allies such as equipment manufacturers, consultancies and customers to seek Endorser status, which requires them to act to support the efficient operation of data centres with whom they interact [34].

A companion document published by the Joint Research Council details the best practices against which prospective Participants must be assessed [78]. This document was contributed to and reviewed by a "broad group of expert reviewers from operators, vendors, consultants, academics, professional and national bodies" [78].

The best practices include some recommendations which relate directly to the work on aisle containment presented in Chapters 2 to 5. For example, the guidelines require aisle containment to be fitted in new data centres and during major retrofits [78]. They also recommend that supply air flow rates be controlled to maintain a pressure differential between the cold and hot aisles, Δp_{CH}, of less than 5 Pa, although this will not become a requirement until 2018.

In addition to technical guidelines relating to air management and installation of efficient infrastructure, the best practices also include guidance for the management of the facility. For example, Participants are expected to establish a system of decision making which ensures that a board representing software, IT and Mechanical & Electrical (M&E) disciplines must approve all significant decisions, thus ensuring that the effects of a decision on each aspect of the facility are considered. It is also recommended that embedded energy be considered, for example by ensuring the use of existing equipment is optimised prior to new material investment. Further, only the resilience level required for the business should be built to. If different functions of the facility have different required resilience levels, dividing the facility into areas with differing resilience levels should be considered. Introducing scalability is also encouraged in order to avoid excessive provision.

The best practices also recommend that IT departments require vendors to supply details of the power consumption of equipment over a range of inlet temperatures, and to use equipment with power and inlet temperature reporting capability where possible [78]. This kind of approach may help in driving manufacturers to produce equipment which is more conducive to efficient operation.

The best practices offer some flexibility in their stipulations, recognising that some measures may be difficult to implement in certain circumstances. Specifically, the best practices document notes that "not all operators will be able to implement all of the expected Practices in their facilities due to physical, logistical, planning or other constraints" [78]. The document also recognises that co-location providers have limited influence over decisions affecting IT efficiency. However, where this is the case, Participant status requires that the operator "act as an Endorser for those Practices outside of their control" [78].

The broad range of measures recommended in the CoC's best practices guidelines enables the scheme to target the following barriers: **poor communication between facilities and IT staff** (and by extension, split incentives between IT and facilities departments), **requirement for central management approval**, **lack of knowledge and understanding**, and **disincentives for trade allies** (through its requirement on operators to request information from suppliers regarding the efficiency of equipment). The major limitation on the impact of the scheme is its voluntary nature, which limits its reach. In addition, the lack of any direct financial incentive to follow its best practice guidelines makes it unsuitable for addressing barriers related to capital costs of EEMs.

The decision to implement a voluntary code was taken due to the perceived success of similar agreements introduced in other sectors [247]. The focus on monitoring and reporting of energy consumption, alongside the adoption of best practices, was favoured over the setting of efficiency targets, due to (i) the difficulty of defining data centre efficiency, (ii) the different levels of responsibilities of data centres operating under different business models (e.g. co-location, enterprise) and (iii) the various operational requirements of data centres [247]. It has been suggested that the code has a positive impact on the data centre industry beyond those organisations which achieve Participant status, with its associated best practices being adopted by organisations which choose not to seek Participant status [247].

As of 2014, the 221 data centres to have submitted data as part of the Code of Conduct were operating with an average PUE of 1.77 [247]. An analysis of which of the best practices were being implemented showed a clear preference for those requiring no capital expenditure, and those which avoided any major changes to business practices.

The UK Council of Data Centre Operators in 2015 published a report making a series of recommendations for better implementation of the EU CoC [260]. The report stated that the technical content of the EU CoC was well suited to improving energy efficiency, if recommendations were followed. However, the report criticised the administration of the

CoC, stating that communication and processing times were very slow, and that there was "a lack of policing regarding data submission". These factors are said to present a barrier to participation. The report supported the (then pending) incorporation of the technical elements of the CoC into the CENELEC EN 50600 series of standards (discussed further in section 6.1.6.3), and suggested that data centre operators be encouraged to undertake audits against standards such as this, or ISO 50001.

The Uptime Institute's 2013 industry survey [29] provides a useful resource for assessing the extent to which the Joint Research Council's recommendations have been implemented, accepting that some progress is likely to have been made since the study's publication. Table 6-1 shows the adoption rates of various EEMs, all of which are recommended in the EU CoC's best practice guidelines, amongst data centres surveyed by the Uptime Institute [29]. As stated in 1.2.2, the respondents of this survey may be in general more energy conscious than others within the industry. The results show that adoption rates vary for the different EEMs, but that none are ubiquitous, and that adoption rates are particularly low amongst smaller data centres. It is interesting to compare the responses for adoption of aisle containment with those for adoption of raised inlet temperature and variable frequency fan drives. Particularly for smaller data centres, the results imply that a large proportion of data centres employing aisle containment have not raised their supply air temperatures or reduced their supply air flow rates. As shown in Chapters 2 to 5, such measures are essential if the full benefits of containment are to be realised.

EEM	Data Centre Size	
	>5000 servers	<1000 servers
Aisle Containment	72%	53%
Raised inlet temperature	64%	39%
Variable frequency fan drives	63%	31%
Air-side economisation	40%	18%
Water-side economisation	36%	12%
Direct Current Power Distribution	14%	7%
Liquid Cooling	12%	6%

Table 6-1. EEM adoption rates [29].

6.1.6.3 Other policy instruments impacting the data centre sector

In the UK, a range of policy instruments impact on the data centre sector, both directly and indirectly.

CENELEC, a designated standards organisation of the European Commission, maintains a standard for data centre management under its EN 50600 series of standards [261], which is transposed into the British Standard Institution's BS EN 50600 series [262]. This includes stipulations on availability, security and energy efficiency. As of 2016, the EN 50600 series of standards has incorporated the best practice guidelines of the EU CoC [263]. Note that the UK's proposed exit from the European Union is unlikely to affect the country's implementation of the CENELEC scheme, assuming the British Standards Institution remains a member of the European Committee for Standardisation, which includes a number of non-EU members [264]. This would provide a route for continued use of the EU CoC's best practice guidelines following the UK's proposed exit from the EU.

Other, more generic, voluntary standards are also used in some cases in the data centre sector, such as BREEAM (Building Research Establishment Environmental Assessment Method, a standard for sustainable building design) [265] and ISO 50001 (a standard for energy management) [266], [267].

The UK government also requires organisations meeting certain criteria to undertake an energy assessment every four years, via the Energy Savings Opportunity Scheme (ESOS) [268]. The scheme applies to companies with over 250 employees, or with an annual turnover exceeding 50 million Euros. Companies achieving the ISO 50001 standard are exempt from the scheme.

The Green Grid provides a service of verifying PUE claims [64]. This requires companies to provide information on the measurement methodology and, for full certification, third party verification.

Bertoldi highlights that the European Energy Performance of Buildings Directive (EPBD) [269], Eco-design directive [227], and the US's Energy Star scheme [270] (which the EU has shared use of), all apply to the data centre sector [247]. However, Bertoldi also notes that measures encouraging or enforcing the selection of efficient equipment (which applies to these schemes) do not guarantee the energy efficient operation of a facility, due to the poor performance of power and cooling infrastructure at low utilisation, and because data centres may run with excessive levels of redundancy [247].

In the USA, a workshop of 150 data centre industry and government representatives held by the US Department of Energy and US EPA [5] led to recommendations being made for drivers for energy efficiency within the US data centre sector. These were detailed in a document produced by the US Department of Energy [23]. The key recommendation was that utilities

companies should develop programmes to facilitate efficiency improvements in data centres. A core element of such programmes would be education, for example to educate co-location data centre account representatives to "sell" efficiency measures to customers. Technical training was also recommended, to improve staff knowledge of efficiency measures. Besides dealing with the data centres directly, it is also recommended that utilities engage with trade allies such as equipment manufacturers and design and consultancy contractors. This could be achieved by structuring incentives to reward the participation of these actors, for example by subsidising energy assessments provided by these actors, or by incentivising the removal of inefficient equipment. These measures primarily relate to the following barriers: **split incentives**, **lack of knowledge and understanding**, and **disincentives for trade allies**.

The recommendations of the US EPA that electrical utilities offer incentives for efficiency measures in data centres resulted in a significant number of programs being undertaken [243]. Some programs were administered by utilities directly, although others were run by regional authorities. The programs have taken various forms, including subsidized engineering and design services, energy audits, incentives for the installation of specific technologies with an assumed per-unit energy saving, and incentives on a more customized basis, with energy savings estimated by engineers. Howard & Holmes [243] suggest that the extent of these programs has been inhibited by the lack of knowledge within utilities companies regarding the technology relevant to the data centre industry, partly due to its relative youth. Another factor is that utilities are usually only able to offer incentives to the bill-paying customer, such that customers of co-location facilities are not eligible for incentives which could encourage the introduction of efficient servers, for example. This is because utilities need to be able to recover incentives paid out if the measures are removed before the allotted time.

One program was carried out by Pacific Gas and Electric from 2006-2008, and was evaluated by Energy Market Innovations (2009, cited in [243]). Interviews with customers of the program revealed concerns that utilities staff had insufficient specialist knowledge of data centres, and that data centre operators subsequently had limited trust in their recommendations. Another difficulty was in the targeting of the program, which covered data centres falling under the "high-tech industries" bracket. This included co-location data centres, but any data centre operators involved in industries such as financial services were not eligible. Howard & Holmes [243] suggest that the specific barriers to energy efficiency in data centres need to be challenged with programs aimed specifically at data centres. This

approach would allow staff involved in the programs to develop specialist knowledge which would improve the chances of success. It would also enable marketing to be focussed on relevant journals and conferences which are seen as trusted sources of information.

Cullen et al. (2009, cited in [243]), reported on a similar program, run by Silicon Valley Power from 2007-2008. This program found it difficult to attract involvement from co-location facilities. This was attributed to such facilities often having no staff with an engineering background, making them incapable of completing application forms for the program.

In addition to these recommendations, the U.S. Department of Energy [23] developed a website which collects information on training opportunities, best practices and case studies, and developed some free software, "DC Pro", designed to help data centre operators to identify opportunities for efficiency improvement. The Energy Star program for computer servers was also introduced, making it easier for operators to select efficient IT equipment [270]. A training programme was also developed, called the "Data Center Certified Energy Practitioner" program [243].

Other actions recommended at the initial workshop remain unimplemented, such as [5]:

- the standardization of metrics for IT equipment, infrastructure and facilities,
- the development of a definition for useful work output,
- the promotion of both internal and external reward systems for employees and organisations implementing efficient practices, and
- the conduction of further market research to improve understanding of the barriers to and drivers for energy efficiency in the industry.

6.2 Social Research Design

6.2.1 Introduction

Surveys and interviews provide a vital tool for harvesting information which is useful both to industry and to policy makers [271]. As will be detailed in Section 6.2.8, these methods have been used extensively in research into barriers and drivers. The collection of useful and accurate information requires that great care is taken in the design of such research [272]–[275].

6.2.2 Grounded theory

The traditional view of the qualitative research process is that of the researcher first constructing a hypothesis or hypotheses derived from the literature or from empirical findings [272]. The hypothesis/hypotheses are then tested via some form of data collection

and analysis. Alternatively, the 'grounded theory' approach advises against the application of a pre-defined theory to the subject of interest. Rather, theories should be formulated through the work undertaken, based on the empirical data found in the course of the work [272]. This approach also allows the methodology for the data collection process to be developed iteratively, with analysis of the findings of one round of data collection being used to inform the approach to data collection in the next round.

6.2.3   Sample selection

Decisions about sampling are made when selecting parties to contact requesting participation in the research, when selecting which interested parties will be used (if there is a surfeit), when results are selected for analysis and when presenting findings [272]. Where accurate information about the population to be investigated is available, sample selection may be tailored to ensure the sample is representative of the population [272]. This approach is termed 'A Priori Determination'. Where this is not the case, random sampling may be favourable, and the research may serve to improve knowledge of the distribution of the population.

'A Priori Determination' requires that the dimensions likely to affect the results of the study be identified. For practical reasons, these dimensions should be kept to a minimum, since the more dimensions considered, the greater the sample size required to accurately represent the population [272]. Having identified these factors, a sample with a distribution of these factors consistent with that of the general population may be selected. This is termed 'stratified sampling', and allows a representative sample to be achieved with fewer participants than can be achieved with simple random sampling [276]. If the sample obtained is not representative of the population, weighting of the results may be used to make the results more representative [276]. Another approach is to seek out deviant or extreme cases to ensure that the total range of possibilities within the study group is investigated. Flick [272] proposes that 'A Priori Determination' is the most appropriate in quantitative research, with the other approaches being useful primarily in qualitative research. Similarly, Arksey & Knight [277] stress the need in qualitative research to select a sample which covers the full range of the population of interest, such that all points of view are heard.

Mason [278] and Marshall [279] note that, when applying grounded theory, there should be no pre-determined sample size, and that gathering of data should cease when further data collection ceases to produce new information. This approach is referred to as theoretical sampling.

Having determined which parties to invite to participate in the research, the next task is to obtain a satisfactory number of responses. Multiple attempts at contact are often required to achieve satisfactory response rates [274]. Non-response can lead to errors if non-respondents have differing characteristics from the rest of the population [274], [276].

Maximising response rates requires that the perceived rewards of participation be maximised, and the perceived costs minimised. Participation is more likely if the invitation is perceived to come from, or have the backing of, a legitimate source [274]. Similarly, invited parties are more likely to respond if they consider the survey to be important or worthwhile. Hence efforts should be made to ensure that the survey and any requests for response are professional in appearance, and do not make technical or typographical errors. The recipient's confidence in the value of the survey may also be improved by the use of technical vocabulary specific to the population being studied. However, a balance must be struck in order to avoid the use of vocabulary unfamiliar to the recipient [274].

Rubin & Rubin [280] note that invitations to participate in research are more likely to be successful if the invitation stresses how and why the prospective interviewee's opinions are useful, and what problem the researcher wishes to solve. Similarly, showing respect and gratitude for the survey recipient could help to increase response rates [274]. For example, ensuring that correspondence is personally addressed, adding phrases such as "we appreciate very much your help" or "many thanks in advance" and offering support with completing the questionnaire are ways to slightly increase the "reward" for completion [274].

Conversely, response rates are negatively impacted by increases in the expected time required of participants. Galesic & Bosnjak [281] found response rates dropped from 24.5 to 17.1% when the predicted length of a web-based questionnaire was increased from 15-30 minutes to 30-45 minutes.

6.2.4   Question 'openness'

One important aspect of a survey or interview question is its 'openness'. Flick [272] stresses that the research design should allow participants to report information which is unexpected but useful. Saris [273] provides useful examples exploring this issue. He notes that broad questions such as "To what extent are you interested in non-governmental politics?" are problematic in that respondents may not consider all of the forms of non-governmental politics that the researcher has in mind. He proposes the more detailed formulation:

"There are many organizations that try to influence political decisions in your country and the world, for example, the trade unions, employers' organizations, environmental organizations. How interested would you say you are in the activities of such organizations?" [273]

This might help to alleviate the problem, but might also lead the respondent to think only of the types of organisations mentioned [273].

Saris [273] goes on to assert that leading questions should be avoided. For example, the formulation "What party are you going to vote for?" is preferable to "Are you going to vote for the Republicans?".

'Balanced' requests for answers (e.g. 'To which extent do you favour or oppose euthanasia?') are generally seen as preferable to 'unbalanced' formulations on the assumption that the latter (e.g. 'To which extent do you favour euthanasia?') will bias respondents towards the answer direction indicated [273]. However, there is no definitive research supporting this assumption [273], [282]. In any case, unbalanced formulations are generally avoided through the removal of indications of direction from the request for an answer (e.g. 'What do you think about euthanasia?') [273].

6.2.5 Interviews, questionnaires and mode effects

Gillham [283] states that the aim of a research interview is "to obtain information and understanding of issues relevant to the general aims and specific questions of a research project". He goes on to state that a questionnaire does not provide the opportunity to understand the answers given, or to explore them. Interviews are required if the "research aims mainly require insight and understanding" whereas questionnaires would suffice where the "research aims are factual and summary in character". Gillham also states that respondents are much more inclined to give detailed answers verbally than in writing [284]. Berg furthers this point, noting that interviews are useful where the researcher is "interested in understanding the perceptions of participants" [285].

Saris [273] asserts that the mode of data collection (e.g. face to face interview, telephone, or postal) affects the responses received, and that the formulation of the questions used should take into account the mode. Self-administered questionnaires are prone to item non-response as well as questions being answered in the 'wrong' order (these occurrences may be prevented in computer based surveys) [273]. These considerations must be weighed against the increased cost and time requirement for face to face interviews. Interviewer administered questionnaires allow for the possibility of further explanation being given to

the respondent if required, although this can lead to responses to questions not being directly comparable [273].

Saris [273] briefly mentions that respondents may be inclined to skew their responses in the direction which they deem to be socially desirable. For example, someone who is in favour of the death penalty may play down their preference if they deem it to be unpopular. Schuman & Presser [286] discuss this in more detail, and report on experiments which found no evidence of a significant impact being caused by this issue. Others *have* noted social desirability effects, particularly in face-to-face and telephone interviews, rather than self-administered surveys [287]. Dillman [288] also notes that more positive responses tend to be obtained from personal interviews than from telephone or mail surveys.

6.2.6 Interview design and technique

Berg [285] proposes that an interview may be defined as a conversation in which the purpose is to gather information. In a 'semi-structured interview', each interviewee is asked a list of predetermined questions in the same order, but the interviewer may digress and probe beyond these questions as deemed appropriate.

Berg [285] proposes that obtaining comprehensive information in an interview requires the use of four different types of questions:

- Essential questions – these are "geared towards eliciting specific desired information" [285].

- Extra questions – these also address the "specific desired information", but are worded differently in order to give some indication of the reliability of the initial response [285].

- Throw-away questions – these are usually placed near the beginning of the interview, and may be used to determine simple, factual information to categorise this particular interviewee within the population of interest, or to "develop a rapport between interviewers and subjects" [285].

- Probing questions – e.g. "could you tell me more about that?", "how come?" etc. [285]. These questions encourage the interviewee to give more detail, following an initial answer giving insufficient information.

Gilham [289] advises that questions used to acquire more detail on a specific point should be "simple, clear, direct and potent". This could be achieved by repeating or rephrasing something of interest which the interviewer has said, or by asking for clarification on a point.

A number of authors have noted the importance of keeping questions as simple as possible [273], [274], [285], [288], [290]. More complex questions tend to result in participants taking less care over their answers [274], misunderstanding questions [273], [285], or giving inaccurate responses [273].

Gillham [291] stresses that interviewers should seek to talk very little, allowing the interviewee to divulge the information which is being sought. The interviewer's function is to encourage response from the interviewee.

Rubin & Rubin note the importance of maintaining a logical structure, with questions on a similar theme being grouped together [292]. They further recommend that an explicit transition from one theme to another should be highlighted to the interviewee, and explained.

The interview design may be improved by undertaking pilot interviews with a sub-section of the population, and/or with people familiar with the subject matter, with interviewers reporting any problems identified [274], [285]. The interviewee should be asked for comments and feedback [293]. This helps to identify poorly worded questions, questions with bias, and any failures to ascertain the desired information [285].

Rubin & Rubin talk about the importance of evaluating each interview afterwards, particularly for a novice interviewer [292]. The findings of the interview should be analysed to ascertain whether the data collected will contribute to answering the research question(s) [294]. The interviewer should try to identify times when a satisfactory answer was not obtained, and to determine what the interviewer could do differently in subsequent interviews to improve the response.

6.2.7 Analysing findings from research interviews

The first stage in analysing the findings of research interviews is transcription, i.e. making a written copy of what was said during the interview. A transcription is usually made from an audio recording of the interview. According to Gillham [293], transcription is easier to complete whilst the interview is still fresh in the mind. This makes it easier to understand words or sentences which are slightly unclear, and to remember gestures etc. which may convey important supplementary information [294].

Analysis of interview transcripts typically involves "coding" the transcripts such that the themes raised by the interviewees can be systematically analysed and contextualised, allowing the findings to be accurately and comprehensively reported [294], [295]. This

structured approach helps to avoid the tendency to give excessive weight to responses which conform to expectations.

The coding process involves studying the transcripts to identify the ideas and opinions raised which are relevant to the research questions [293], [295]. A list of these ideas and opinions is compiled and refined in light of the detail of the answers from the various interviews. The occurrences of the ideas and opinions are then marked in the transcripts and categorised under different headings, such that sections of transcripts relating to specific ideas and opinions can easily be collected and quantified for detailed study [294], [296].

A literature review may help to inform identification of relevant themes and ideas. However, it is important to be open to the possibility of new information being found in the interview process which expands upon or even contradicts information reported by others in the literature [294]. This requires a detailed analysis of the transcripts, and consideration of the subtle differences between the ways in which the different ideas and opinions are expressed by different interviewees.

The researcher should endeavour to ensure that the selected codes cover all important ideas raised by the interviewees, so that no "substantive statements" are left uncoded [296]. Of course, repetitions and irrelevant material may be left uncoded. Gillham also suggests that the codes should be "exclusive", so that no statement could belong to more than one code [296].

Rubin & Rubin [294] stress that coding the scripts requires a good deal of concentration, and recommend that only one or two ideas are coded for in each reading, since the researcher needs to be able to identify where ideas are expressed using unexpected language. If the grounded theory approach is being applied, the transcripts of each interview should ideally be analysed prior to the next interview, such that the findings can be used to inform the questions being asked [294].

Once all of the interviews have taken place and transcripts have been coded, analysis must be undertaken. This includes systematically examining the codes in order to group and compare them, and to "look for patterns and connections" [297]. The individual instances of each code should be analysed as a body, considering the differences and similarities between the thoughts of the various interviewees on the subject [295], [297]. The frequency of occurrence of a particular code may give some indication of the importance of this theme [295]. The coded data may also be grouped into different categories of interviewees, and the differences between the respective findings studied [295], [297].

The final stage is to explain the findings of the analysis [295]. Here the researcher should draw on the findings of related research, and where contradictions occur should propose explanations and/or suggest further research required to provide these explanations [295].

6.2.8 Methods used in research into industrial energy efficiency

Methods commonly used in the identification of barriers to and drivers for energy efficiency in industry include workshops, semi-structured interviews and questionnaires. Participants are usually those under the direct employment of the relevant industry, including energy managers, executives, financial managers and maintenance managers [180], [181], [208], [298], [299]. Less commonly, participation is widened to include other stakeholders, such as academics working within the relevant field, and representatives of trade organisations, local/national government and energy companies [180], [208], [300].

Participants are usually asked questions about what they perceive to be the most important barriers and drivers, and sometimes to numerically quantify the perceived importance of a list of barriers/drivers. The perceived strengths of barriers can be quantified by asking respondents if a barrier is 'very strong', 'quite strong' etc., or by asking respondents to rank barriers [210], [211]. Less commonly, attempts are made to quantify the real strength of a barrier. For example, Trianni, Cagno, Worrell, & Pugliese [301] quantified the barrier of limited availability of capital by dividing the capital required for energy efficiency investments by the total capital available.

Besides questions directly enquiring about barriers/drivers, participants are sometimes asked questions investigating attitudes towards energy within their organisation [210], their organisation's energy performance [180], and what EEMs they have taken to date [29].

Research into barriers and drivers in the data centre sector specifically has thus far been limited. In 2008 The US Department of Energy and US EPA [5] held a workshop attended by more than 150 industry and government representatives, with the aim of identifying potential routes to improving data centre energy efficiency. The morning session consisted of presentations from government and industry. Attendees were then split into 3 groups, which focused on "defining energy efficient data centers", "advancing energy efficient data centers" and "rewarding energy efficient data centers". The findings of and recommendations resulting from this workshop have been detailed in Section 6.1.6.3.

Besides this work, a survey undertaken by The Uptime Institute [29] canvassed the opinions of data centre professionals on barriers and drivers. Further, Howard & Holmes [243] have carried out an evaluation of utility energy efficiency programs aimed at data centres.

TechUK and the UK Council of Data Centre Operators have canvassed opinions within the UK data centre sector regarding the EU CoC and DC CCA [28], [256], [258], [260]. The findings of these studies have been detailed in Section 6.1.6.

In addition, Brady [246] has conducted semi-structured interviews investigating barriers to implementation of EEMs in the UK data centre sector. Interviews were identified as the appropriate data collection method in Brady's study. This was because the research questions, which focussed on ascertaining stakeholders' views on the key barriers to the introduction of EEMs, were qualitative, rather than quantitative. The semi-structured approach was used in order to allow participants the freedom to bring up themes which had not been anticipated by the researcher.

6.3 A semi-structured interview investigating policy instruments targeting data centre energy efficiency

6.3.1 The research question

Section 6.1.2 has demonstrated that there is likely to be a considerable energy efficiency gap affecting the data centre sector, and has identified some barriers generally agreed to impact the sector. The information regarding the barriers important in the sector comes largely from surveys conducted by trade bodies. The only relevant, recent work within the academic literature is from Brady [246]. Brady's work investigates barriers, as well as the use of energy efficiency metrics, but does not discuss policy instruments more broadly.

Section 6.1 has demonstrated (i) that policy instruments introduced by national governments and international political unions (such as the EU) can help to overcome barriers to energy efficiency, (ii) that these policy instruments come in a variety of forms, and (iii) that the success of policy instruments is dependent on the appropriateness of their design.

The information relating to industry opinions on the appropriate policy responses to the energy efficiency gap in the data centre sector is limited to a small number of reports produced by trade bodies, in addition to the findings of the workshops conducted by the US EPA some time ago, in 2008 [5]. Material produced by trade bodies is certainly of value, but cannot be relied upon to be produced using a rigorous, academic approach, and its findings may be subject to the agenda of the body in question. There is clearly a need to determine the effectiveness of existing policy instruments, and to gain information which may help to improve these policy instruments and to design future instruments effectively. While there is a substantial quantity of academic research relating to the effectiveness of various policy

instruments in driving energy efficiency improvements in industry in general, there is very little relating specifically to data centres.

The following central research question was therefore identified:

'What policy instruments are appropriate for driving energy efficiency improvements in the UK data centre sector?'

The opinions of people working within this sector are an important source of information relating to this research question. Social research methods offer a route to obtaining qualitative information which may be used to inform the design of these policy instruments. Semi-structured interviews undertaken with people with responsibility for energy management have been used extensively in investigating barriers and drivers in industry, as discussed in section 6.2.8. Face to face interviews allow a more detailed investigation of a participant's opinions than other forms of social research, as discussed in section 6.2.5. This is important due to the qualitative nature of the research question, and the multitude of factors which may affect participants' opinions. It is also easier to strike a balance between initial, open questions and guided follow up questions in a face to face interview. Therefore, a decision was taken to undertake a series of semi-structured interviews with people working within the data centre sector, in order to gather their opinions on existing policy instruments impacting on energy efficiency in the data centre sector, and their opinions on other approaches which could be used to drive energy efficiency.

The information gathered during the interviews may then be analysed, critically, with reference to the literature on the effectiveness of different policy instruments relating to energy efficiency in other sectors.

6.3.2   Interview design

Having identified the research question, a script was developed for the semi-structured interview, as shown in Appendix 4. The script begins by informing the interviewee of the structure of the interview. This is followed by some questions about the interviewee and their employer. This information is important since, as discussed in section 6.1.5, the barriers and drivers relevant to a particular data centre are likely to depend on certain characteristics of the data centre.

The interview then moves on to questions specifically regarding the EU CoC, with alternative questions included depending on whether or not the interviewee was a Participant/Endorser of the EU CoC. This section of the interview seeks to assess the suitability of this policy instrument for driving adoption of EEMs in the sector.

The interviewee was then asked for their opinions on the potential for other policies to stimulate improvements in energy efficiency. This was developed into a discussion encompassing the existing CRC, CCL and CCA, and other potential policies. Hence, the interviewee's opinions on the most effective ways to incentivise the adoption of EEMs could be collected, with reference to both existing and other potential policies.

Following the discussion of policy instruments, the script moved on to a question about whether the interviewee's data centre(s) employed aisle containment. This was developed through follow-up questions and probes into a discussion about the drivers for installing aisle containment, and the barriers preventing wider adoption of this approach. This section of the interview provides the opportunity for a discussion around one specific EEM, whose findings may be analysed in light of the findings relating specifically to policy instruments, enabling an assessment of whether the identified policy instruments would be likely to address the barriers identified in relation to this EEM.

Finally, the interviewee was given the opportunity to add anything else which they felt had been left out and to ask any questions they may have.

With all questions, care was taken to ensure brevity and simplicity, and to avoid leading questions, the importance of which is discussed in section 6.2.4. In general, initial questions were kept as open as possible. However, where these questions failed to lead to the required information being obtained, more specific probing questions were used.

Each of the four main sections of the interview (introductory questions about the interviewee's company and job role, questions about the EU CoC, questions about other policy instruments, and questions about aisle containment) were preceded with a short statement describing the focus of the forthcoming section. This was in line with the recommendations of Rubin & Rubin [292], as discussed in section 6.2.6.

6.3.3   Pilot interview

Having produced an initial interview script, the script was tested in a pilot interview, undertaken with a data centre professional. Note that, at this stage, the 'Probes' listed in the interview script included in Appendix 4 were not included. Rather, probes were used on an ad hoc basis, as and when the interviewer felt that they were needed.

Following the pilot interview, the transcript was analysed to determine whether or not the desired information had been obtained. The analysis showed that the script was largely effective in obtaining the desired information.

The pilot interviewee was asked for feedback on the effectiveness of the interview, after being informed of the associated research question. The interviewee largely felt that the interview script was well suited to the research question, with their only criticism being that more detail could have been drawn out on some of the points raised in their answers.

In light of this criticism, the Probes included in the script shown in the Appendix 4 were added. A formal approach of using probes was adopted in subsequent interviews, whereby the interviewer ticked off items which had been fully covered by the interviewee during the interview. This allowed the interviewer to keep track of what information still needed to be elicited, so that the appropriate probes could be used.

Besides the probes included in the Appendix 4, in line with the suggestions of Berg [285], the interviewer also used probing questions where a subject had been touched upon but not fully explored, by asking for more information from the interviewee or asking the interviewee to clarify a particular statement.

6.3.4 Participant recruitment

Since the focus of the interview was to be the use of policy instruments as drivers for energy efficiency in data centres, a key priority in recruitment was to ensure that the sample included both those who could be said to be engaged with such policy instruments, and those who were less so. With this in mind, participation in the EU CoC was used as a proxy representing a data centre's level of engagement with such policy instruments. Hence, it was ensured that the sample contained a significant number of both Participants and non-Participants of the EU CoC.

Interviewees were recruited through contacts of the PhD candidate and his supervisors. This approach was necessary due to difficulties in recruiting interviewees through 'cold calling'. The possibility of these contacts having greater interest in energy efficiency than is typical in the sector must be considered. However, efforts were made to ensure a wide variety of participants were recruited, in terms of the business models and sizes of the data centres represented, since, as discussed in section 6.1.5, these factors are likely to impact what kinds of barriers are experienced. This approach is in line with the recommendations of Arksey & Knight [277] for qualitative research, as discussed in section 6.2.3. Additionally, some of the interviewees were recruited from contacts whose relationships with the researchers were not based on previous research collaborations. It should also be noted that the use of pre-existing contacts helps to avoid the increased level of non-response experienced with 'cold calling', which leads to sampling errors, since the characteristics of those accepting

invitations to participate are likely to differ from the characteristics of non-respondents [274], [276].

The interviews were undertaken with 10 individuals who were involved either directly in data centre management or in data centre consultancy services, and had some responsibility for energy consumption, whether through facilities or IT, or both. Clearly a sample size of 10 is small in comparison with the total number of people working in the sector. Whilst there are no official figures regarding the number of people working in the sector, some indication can be deduced from participation in the UK's DC CCA. As of 2015 there were over 100 data centres signed up to the DC CCA, which of course only represents a small proportion of the UK data centre sector since it is only open to co-location providers. However, the time-consuming nature of carrying out research interviews necessitates small sample sizes. The advantages of this approach with respect to the richness of qualitative data obtainable, as discussed in section 6.2.5, make semi-structured interviews the appropriate choice for this study.

The total sample size was not determined prior to the completion of the research. Instead, data collection was discontinued when it was judged that no significant new information was being discovered (the process undertaken to make this judgement is described in section 6.4.6). This is in line with the recommendations of Mason [278] and Marshall [279] in relation to the grounded theory approach, as discussed in section 6.2.2.

Potential interviewees were contacted initially via email, with an information letter and consent form attached (see Appendices 5 and 6). The email included a brief description of the research being undertaken, with more detail given in the attached information letter. The information letter and consent form included details of how anonymity would be ensured, and how transcripts and other data related to the interviews would be handled to minimise the risk of sensitive data falling into the wrong hands.

In light of the findings of the literature review, specifically section 6.2.3, the information letter was designed to (i) assure recipients of the value of their prospective contribution to the research, (ii) state that the interviews are part of a broader research project, (iii) express gratitude and respect, and (iv) be professional in appearance, since these factors are considered to increase the likelihood of recipients agreeing to participate [274], [280].

6.3.5 Interview process

Interviews were carried out face to face, either at the interviewee's or the interviewer's place of work. Audio recordings were made on the interviewer's mobile telephone, to

enable an accurate transcription to be made following the interview. Following the interview, the recordings were transferred to a computer at the earliest opportunity, where they could be password protected. The audio file was then removed from the mobile telephone, to minimise the risk of data protection being breached. Audio files were not marked with company or interviewee names, with code names being used instead. These code names can be seen on the transcripts included in the supplementary material provided with this thesis, and their meanings are explained in section 6.3.7. Interviewees were informed prior to the interview that it would be recorded, and of the measures in place for data protection.

Transcriptions were made from the audio recordings, and can be seen in the supplementary material provided with this thesis. Again, company and interviewee names were not included in transcripts, with codes being used instead. All company names, names of people, place names, and other sensitive information were redacted from the transcripts to protect the identities of the interviewees. Transcription was carried out within 24 hours of the interview taking place, in line with the recommendations of Gillham [293] and Rubin & Rubin [294].

6.3.6   Refinement of interview script

The interview script included in the Appendix 4 represents the final form of the script. After each interview, the script was reviewed, in line with the grounded theory approach, as discussed in section 6.2.2. Probes were added to the original interview script reflecting issues raised by each interviewee. For example, if one interviewee made reference to a particular barrier not previously mentioned in the script, a probe was added to ensure that future interviewees were asked whether they felt this barrier was important. One Main Question was added over the course of the interviews being carried out. This was Question 13 (see script in Appendix 4), which relates to the CENELEC EN 50600 standards series, as discussed in section 6.1.6. This issue was raised by a number of interviewees, after the EU CoC best practice guidelines were incorporated into this standards series. This occurred after the first interviews had already taken place. The ongoing analysis of interview transcripts also enabled an assessment to be made at regular intervals as to whether further interviews were likely to lead to significant new information being discovered, in line with the grounded theory method.

It became apparent quite quickly that time constraints would have an impact on the interviews. In order to encourage potential interviewees to take part, the interview was described in the Invitation Letter as taking not more than 45 minutes. Achieving this

required that not every probe be used in every interview. Over the course of the undertaking of the interviews, the probes were ordered in terms of their seeming importance to previous interviewees, for example, prioritising issues raised 'unprompted' in previous interviews, and issues which previous interviewees seemed particularly passionate about. In addition, efforts were made to keep interviewees 'on topic', so that minimal time was wasted in discussing matters not relevant to the research question.

6.3.7 Analysis of interview scripts

Coding of transcripts was undertaken using the software package QSR NVivo 10 [302]. This software aids with the organisation of the material from the interview transcripts into codes. Coding was undertaken at 3 distinct levels. The first level divided the transcript into portions relating to the four distinct sections of the questionnaire (introductory questions, the EU CoC, other policy instruments, and aisle containment). The codes at the second and third levels were unique to each of the first level codes. A screenshot from the software is shown in Figure 6-1, and shows some of the second and third level codes. A code was assigned to each substantive piece of text contained in each transcript. The software enables all pieces of text relating to a particular code to be viewed together.

Name		Sources	Referen	Created O	Created	Modified	N
Aisle containment		7	7	27/02/201	MTE	28/02/20	
Barriers to aisle containme		7	11	27/02/201	MTE	28/02/20	N
Awkward legacy faciliti		3	3	01/03/201	MTE	01/03/20	N
Capital cost		5	7	01/03/201	MTE	01/03/20	N
Lack of knowledge and		4	5	01/03/201	MTE	01/03/20	N
Low load so unnecess		3	3	01/03/201	MTE	01/03/20	N
Operational complexity		1	1	01/03/201	MTE	01/03/20	N
Reduced flexibility		1	1	01/03/201	MTE	01/03/20	N
Reluctance from IT ma		2	2	01/03/201	MTE	01/03/20	N
Split incentives		1	2	01/03/201	MTE	01/03/20	N
Time constraints		1	1	01/03/201	MTE	01/03/20	N
Extent of containment		6	9	28/02/201	MTE	28/02/20	N
Reasons for installation		7	9	27/02/201	MTE	28/02/20	N
Code of Conduct		9	13	27/02/201	MTE	28/02/20	N
Introductory Qs		9	14	27/02/201	MTE	28/02/20	N
Other policy instruments		9	16	27/02/201	MTE	28/02/20	N

Figure 6-1. Screenshot from QSR NVivo 10, showing an example of the codes used in the analysis of the interview transcripts.

Individual interviewees are referred to by code names, which indicate characteristics identified during the introductory section of the interview. The codes have 4 characters.

The first character identifies the interviewee's relationship to the EU CoC, the second identifies their business model, the third identifies the combined size of the data centres for which they are responsible and the fourth is an individual identification number. Table 6-2 details the meanings of the interviewee codes. By way of an example, with reference to the information in Table 6-2, the interviewee identified as ABB1 worked for a co-location provider holding Participant status within the EU CoC, with a total power consumption exceeding 1 MW and/or greater than 1,000 m^2 of floor space.

1st character (EU CoC)	A	Participant/Endorser
	B	Non-Participant/Endorser
	C	Not based in the EU
2nd character (business model)	A	Enterprise
	B	Co-location
	C	Consultancy
3rd character (size)	A	Less than 1 MW and less than 1,000 m^2
	B	More than 1 MW and/or more than 1,000 m^2
	C	Not applicable
4th character	-	Individual identification number

Table 6-2. Description of interviewee codes.

6.4   Findings of semi-structured interviews

6.4.1   The interviewees

Figure 6-2 shows the proportions of interviewees having each of the characteristics listed in Table 6-2.

Five of the interviewees either managed data centres with Participant status or (in one case) worked in data centre consultancy and held Endorser status within the EU CoC. One of the interviewees was also actively planning to apply for Participant status. This left four who were not signed up to the CoC, and were not actively seeking Participant status. One of the interviewees not seeking Participant or Endorser status was based in the US, with all other interviewees based in the UK.

Three of the interviewees represented co-location providers, five represented enterprise data centres, and two were involved in data centre consultancy services (of which one was the US-based interviewee).

Three of the interviewees were classed as representing small data centres, and five as representing large data centres. For the two interviewees involved in consultancy, such a designation was deemed not applicable.

The full list of codes for the interviewees is as follows: ABB1, ABB2, AAB1, AAB2, BBB1, BAA1, BAA2, BAA3, CCC1. The interviewee who was in the process of seeking EU CoC Participant status was BAA1.

a) b) c)

- Participant/Endorser
- Non-Participant/Endorser
- Non-EU

- Enterprise
- Co-location
- Consultancy

- Less than 1 MW and less than 1000 m2
- Greater than 1 MW and/or greater than 1000 m2
- Not applicable
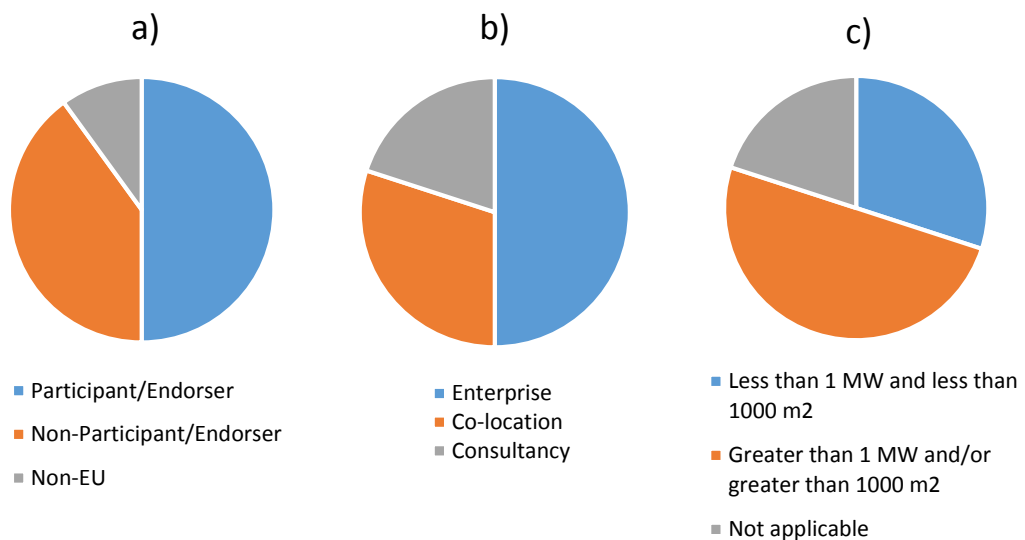
Figure 6-2. Characteristics of interviewees with respect to a) relationship with EU CoC, b) business model and c) size of data centre(s).

6.4.2   The EU Code of Conduct on Data Centre Energy Efficiency

Table 6-3 summarises the responses to the questions regarding the EU CoC. The key themes raised during the interviews are listed in the left hand column in the form of statements, with the numbers of interviewees who tended to agree with or disagree with each statement recorded, as well as the number who seemed unsure, and with whom the issue was not raised. Whilst this table provides an accessible overview of the key issues raised during the interviews, it should be noted that this format does not enable appreciation of the complexity and richness of data which may be collected during research interviews. Further, due to the small sample size, a quantitave analysis of the findings is of limited value. Hence, the text following Table 6-3 investigates the interviewees' views on the EU CoC in finer detail.

Statement	Agree	Disagree	Unsure	Issue not raised or not applicable
CoC best practices are useful	6	0	0	4
Poor administration is a barrier to participation	5	0	1	4
Better promotion of the CoC is needed to increase participation	4	0	0	6
Participation status helps to win customers or improves company image	5	2	0	3
Lack of enforced periodic reporting limits the CoC's impact	2	0	1	7
The CoC has a positive impact in spreading best practice beyond Participants and Endorsers	3	1	0	6
The lack of translation of the document into other languages has hindered its adoption	2	0	0	8
Data centre operators are concerned that CoC assessment may shed light on their own failings	1	0	0	9
The best practices need to be updated more regularly	1	0	0	9

Table 6-3. Summary of responses to questions regarding the EU Code of Conduct.

All of those who were familiar with the EU CoC's set of best practices felt that these did indeed represent an expression of effective management of energy efficient data centres. Interviewee ABB1 commented that the best practices have "a very high reputation", and that "they're all common sense tools that you would use to improve energy efficiency within a data centre". Six interviewees saw the adoption of the CoC's best practices as a route to improving energy efficiency. Interviewee AAB2 specifically felt that the process of seeking

Participant status had forced some staff to "start considering" energy efficiency where previously they hadn't done, particularly staff responsible for IT.

Whilst some of the interviewees not holding Participant/Endorser status seemed to have limited knowledge of the EU CoC's recommended best practices, none made negative comments about them. Interviewee BBB1 felt that, whilst they didn't see adoption of the EU CoC as necessary in their case, in data centres with less experienced staff it could be helpful in providing guidance on energy efficiency. They commented, "I suppose if you went to a data centre that didn't have M&E engineering, I can see that it has a lot of guidance on shall we say non-M&E people on how to approach things".

Five interviewees (ABB1, ABB2, ACC1, AAB2 and BAA1) felt that Participant status either helps co-location data centres to win customers, or is beneficial to a company's image in the case of enterprise data centres. Note that these were all either Participants/Endorsers of the CoC, or were seeking Participant status. Interviewee ABB1 stated that some customers asked about the EU CoC during procurement. Interviewee BAA1, who was seeking Participant status, felt that such status would help them to demonstrate that they were operating their data centre in a "controlled and managed way". Interviewee ACC1 (an Endorser of the EU CoC) said that they recommended customers seek Participant status if the customer was interested in demonstrating to "the industry, and the wider market" that they take energy efficiency seriously, and if they felt it would be beneficial for their image in terms of Corporate Social Responsibility (CSR). Where these issues were not a concern to the customer, this interviewee recommended that the best practices be followed, but Participant status not be sought. On this subject, they commented:

"So if someone was coming to me and saying 'do we want to be a Participant?', well it depends what you want from the code. Do you want to just do the best you can in the data centre in terms of keeping it as efficient as possible? Because you don't need to be a participant to do that, you can apply the code."

CSR was also mentioned as a consideration by one Participant of the CoC, interviewee AAB2, who operated an enterprise data centre for a large company with a significant public profile. They commented that "there is a return in terms of image, you know, it's good from a corporate responsibility point of view to show that you're really serious about your impact on the environment". However, interviewee ABB2 felt that there was a limit to the "perceived value" of the CoC in this respect, stating that "if we can deliver to the industry a perceived value, in other words their clients value it, then I think you'd see uptake rise, and until then some people will wonder why on earth they would bother". The US-based

interviewee (CCC1) commented that co-location providers were generally more focused on ensuring that they operate within their Service Level Agreement (SLA) than they were on accreditation against standards for energy efficiency.

Three interviewees (ABB1, ABB2 and AAB2) felt that the introduction of the EU CoC had had some impact in spreading information about best practices beyond those who sought Participant/Endorser status. Specifically, interviewee AAB2 commented that "certainly there are a lot less ignorant people in the industry than before the code. I think the code sort of made it really really clear and understandable how small changes and…lack of process and lack of attention on some of the things could essentially cost money". The US-based interviewee (CCC1) expressed doubts about the extent to which the EU CoC best practice guidelines were used as a reference in the US.

Five interviewees (ABB1, ABB2, ACC1, AAB2 and BAA1) highlighted issues with administration of the scheme as being a barrier to participation. Of these, four were Participants or Endorsers of the CoC, and the other was in the process of applying for Participant status. Interviewee ABB1 commented that:

"The down side is the way that the scheme is administered…it's totally under resourced at the JRC within the EU that run it…it takes months and months and months from a submission to get feedback on the submission and then get certified…there's been a lot of bad press about the administration…so a lot of people have said, you know, what's the point."

However, one Participant of the CoC, interviewee ABB2, commented that the administrative burden was mostly in the form of slow response times, and did not require a large time investment from the applicant, and therefore "shouldn't" put people off from applying. Interviewee ACC1 felt that the adoption of the CoC best practices into EN 50600 (as discussed in section 6.1.6.3) provides a route for accreditation against these standards with a more efficient system of administration.

Four interviewees (ABB2, AAB2, BBB1 and BAA1) identified a lack of promotion of the EU CoC within the data centre sector and within customers of co-location providers as having inhibited uptake of the scheme. Interviewee AAB2 saw this as the most important barrier to greater uptake of the CoC, commenting that "the promotion and marketing of it is almost non-existent, so people just don't know about it".

Two interviewees (ABB1 and ABB2) noted that, whilst the CoC technically had an element of periodic reporting which was designed to lead to continual improvement, in reality this was

not enforced, meaning that the CoC doesn't necessarily encourage regular auditing. Interviewee ABB1 commented as follows:

"Originally when you applied and got approved, that was valid for 3 years, and at the end of the 3 years they would assess you against…have you implemented what you said you were going to implement 3 years ago. But because of the lack of manpower that's never happened."

Interviewee ABB1 went on to stress the importance of this issue, noting that "what any management system should be all about is continual improvement so you take a set of tools and best practice, and you work with those within a management framework to continually review where you are and set plans and objectives". However, interviewee ABB1 also reported that they used the CoC best practices within their other auditing regimes to provide this element. Interviewee ACC1 highlighted that the adoption of the EU CoC best practices into EN 50600 provides an opportunity to use these best practices within a more strictly enforced environment of continual improvement. Conversely, interviewee ABB1 commented that stricter enforcement would put people off the scheme if they felt Participant status offered insufficient value to them. This interviewee also felt that the adoption of the EU CoC best practices into a compulsory scheme such as the UK government's ESOS mandatory energy assessment scheme could be an effective route to driving improvements.

Interviewees ABB1 and ACC1 felt that the lack of translation of the EU CoC into languages other than English was a barrier to uptake outside of the UK, whilst highlighting that the adoption of its best practices into EN 50600 would lead to translations being made available.

Interviewee BAA2 felt that the scheme did not apply to enterprise data centres, commenting that it would simply be a case of "ticking the box, but we [as an enterprise data centre] don't have anyone who we need to show that tick to". Conversely, interviewee BBB1, based at a co-location facility, felt that the CoC was more applicable to enterprise data centres, since co-location data centres are covered by the DC CCA.

Interviewee BAA1 commented that some data centre managers may be reluctant to go through any process which may expose weaknesses in their management practices. This interviewee suggested that "they're aware that their facilities are probably somewhat lacking…and they don't really wanna expose that…I think some data centre operators are a little bit scared that if they expose their weaknesses, they could be subject to some severe scrutiny from certain areas".

Interviewee AAB2 felt that the best practice recommendations of the CoC should be updated "a bit more frequently". However, this seemed to be a minor point, as they concluded that "in general, I don't have any major suggestions" for improvements to the scheme.

6.4.3   Other policy instruments

6.4.3.1   Energy taxation

Table 6-4 summarises the responses to the questions regarding the CCL, CRC and CCA schemes. As with section 6.4.2, the text following Table 6-4 investigates the interviewees' views in finer detail.

Statement	Agreed	Disagreed	Uncertain	Issue not raised or not applicable
CRC has been effective in driving EEMs	4	3	1	2
CCA has been effective in driving EEMs	5	0	1	4
Increasing energy bills for data centres constrains innovation or results in carbon leakage	1	0	0	9
Reporting related to the current UK energy policy environment is unnecessarily burdensome	1	0	0	9
The use of PUE as a measure of efficiency limits the value of the CCA	1	0	0	9

Table 6-4. Summary of responses to questions regarding energy tax approaches.

There was some disagreement over whether the CRC and CCL had driven adoption of EEMs in data centres. Some felt that any tax on energy tended to focus attention on efficiency, with interviewee BBB1 commenting that "anything where there's a penalty can get people focused on things…it's often difficult to get money to do projects unless there's a [financial] driver behind it". However, a number of interviewees felt that the original conception of the CRC scheme as being revenue neutral would have enabled it to better facilitate genuine improvements. Interviewee ACC1 encapsulated this opinion, commenting that it "started out as well-meaning I think, where the savings would be ploughed back into the best

performers. However it was written badly, because it kind of penalised early movers. And then they scrapped the rewards anyway and it just became a tax". Interviewee ACC1 also felt that increasing energy bills for co-location providers could constrain innovation amongst users of data centre services.

The DC CCA was generally more highly thought of amongst the interviewees than the CRC alone, with most feeling that the scheme allowed growth in the sector whilst incentivising improvements in efficiency, with the threat of financial penalties for failure to improve efficiency being a strong driver. The US-based interviewee, CCC1, also felt that the CCA was "a great structure", but suggested that accredited assessors would be needed to ensure genuine compliance with targets.

Interviewee ABB2 noted the potential for levies on energy bills for data centres to drive potential co-location customers to keep IT services in-house, in relatively inefficient facilities. They commented that:

"Unfortunately, the way in which the taxation regime went, that meant that as your electricity usage increased, although you were doing it efficiently, and therefore as a societal benefit, you suddenly found yourself the recipient of a tax, which you then needed to pass on to your clients, who actually didn't need to pay it because they could in many cases just have a data cupboard. And we remain in the UK a country where 80% of the data centre, of what you would call a data centre, are racks in highly inefficient cupboards."

Interviewee ABB2 felt that this would ultimately lead to greater energy consumption than if no levy was applied to data centre energy consumption. They also noted that the CRC, prior to the introduction of the CCA, penalised growth in data centres, even if efficiency was improved, with higher taxes. They felt that the CCA had gone some way to addressing this, but proposed that a better approach would be to only apply taxes to IT services provided outside of data centres, therefore incentivising the transfer of IT services into more efficient co-location facilities.

Interviewee BBB1 felt that their inclusion in the CCA was "more of a driver" for their efforts in improving efficiency than the existence of standards such as the EU CoC, due to the monetary savings associated with it. They therefore felt that the CoC would not add any value for them.

Interviewee ACC1 commented that for some data centre operators, such as banks, resilience was far more financially valuable than energy efficiency, such that increases in energy prices were unlikely to be an important factor in decision making.

Interviewee ACC1 also felt that the CCA's reliance on PUE was a limitation, stating that PUE "can't be trusted" as a measure of energy efficiency. Note that the limitations of the PUE metric were discussed in section 1.4.

Interviewee ABB1 hesitantly stated that they think the targets of the CCA are ambitious enough. However, no interviewees seemed sure on this point, despite many believing that the scheme was effective in driving improvements.

Interviewee ABB1 also complained that, at present, there were various pieces of policy in the UK which required reporting of similar information multiple times, which resulted in a "massive" administrative burden. This issue has also been raised by TechUK, as noted in section 6.1.6.1 [256].

6.4.3.2   Subsidised or free audits, design and engineering services

Table 6-5 summarises the responses to the questions regarding free audits, design and engineering services. The text following Table 6-5 investigates the interviewees' views in finer detail.

Statement	Agreed	Disagreed	Uncertain	Issue not raised or not applicable
Subsidised/free audits/design services could help to improve energy efficiency in the sector	4	2	3	1
There is a need for the provision of these services from a reliable, unbiased source	1	0	0	9
It is important that those conducting audits etc. are data centre specialists	1	0	0	9
Such schemes are vulnerable to freeridership	1	0	0	9

Table 6-5. Summary of responses to questions regarding subsidised or free audits, design and engineering services.

Interviewees AAB2, BBB1, BAA1 and BAA2 all felt to a greater or lesser extent that the provision of free or subsidised audits, design and engineering services could be useful.

Interviewee AAB2 commented that "anything that improves your knowledge in any of those steps cannot do anything but helping us as an industry get better at what we do". Interviewee BAA2 noted that making consultancy services available through an official program would increase the perceived reliability of the source, making the services provided more attractive. Similarly, interviewee BBB1 noted that there is often some scepticism towards those offering consultancy services, and that "anything that makes it commercially easier for companies to do, they will probably take it up more". However, this interviewee also noted that the costs of auditing, design and engineering services tended to be small in relation to the costs of undertaking EEMs, so that "subsidising the actual construction would probably be more interesting to people".

Interviewee ABB1 mentioned the ESOS when asked about these approaches. They mentioned that this scheme at present lacks impact because "there's no formal requirement to do anything" as a result of the recommendations of the auditor's report. They also noted that the auditors are not data centre specialists, which limits the value of the outcome. They felt that incorporating the CoC best practices into ESOS would be an improvement, stating "that's really the way it needs to go to make good sense".

Interviewee ABB1 accepted that there was a lack of understanding of energy efficiency within the sector, but felt that "a short term consultant's engagement doesn't necessarily solve that problem". They suggested that there is a need to "improve levels of education generally" in order to improve the level of knowledge of those involved in data centre management.

Interviewee ABB2, whilst also accepting that some data centre operators lack the knowledge and expertise required to improve their energy efficiency, felt that this kind of approach would not drive energy efficiency, since data centre managers are already effective at introducing approaches which are in their best interests. Therefore, the best approach would be to "put drivers in place" to encourage operators to prioritise energy efficiency, adding that this "would be far more effective than putting subsidies in place for energy audits".

Interviewee CCC1 had an important perspective on these approaches, having worked in utility programs offering such services in the US. This interviewee felt that these schemes rarely succeeded in encouraging people to introduce EEMs who were not already motivated to do so, and that free-ridership was prevalent within these programs. They felt that free audits were effective in ensuring take up into energy efficiency programs, and that "if you

calculate the cost effectiveness of your services from that point of view, then yes, it's effective, extremely effective. But if you look at it in terms of free ridership it's usually not".

6.4.3.3   Standards schemes

Table 6-6 summarises the responses to the questions regarding free audits, design and engineering services. The text following Table 6-6 investigates the interviewees' views in finer detail.

Statement	Agreed	Disagreed	Uncertain	Issue not raised or not applicable
Adoption of EU CoC best practices into a recognised standards scheme could lead to increased adoption and/or status	3	0	0	7
Obligatory rating of data centres for energy efficiency could drive improvements through the threat of embarrassment	1	0	0	9

Table 6-6. Summary of responses to questions regarding standards schemes.

Three interviewees (ACC1, ABB1 and AAB1) raised the issue of standards schemes in relation to the incorporation of the EU CoC best practices into a recognised standards document. Interviewee ACC1 felt that incorporating the best practices into EN 50600 "gives it formality, also gives it credibility and weight". Interviewee ABB1 commented that the move would result in a respected standard which was "usable as an audit regime", allowing improvement opportunities to be identified and implemented.

Interviewee BAA2 suggested that negative press can be a driver for improvement in energy efficiency, and that something akin to the National Home Energy Rating (NHER) Scheme could be used in data centres to achieve this. The NHER is the system used to rate a house's energy efficiency, in order to produce an Energy Performance Certificate, which is a compulsory requirement for any house prior to its being sold [303]. Interviewee BAA2 illustrated their point, saying "It would have to be something like that where you shamed us, or there was financial penalties if we weren't efficient".

Interviewee ABB2 praised the existing Data Centre Alliance (DCA) Certification Scheme, stating that it "strikes a really good balance" between efficiency and other operational considerations. The DCA is an international industry association, and its Certification Scheme is intended to "provide an industry led, widely adopted recognition of a data centre's designed purpose, its operational integrity, energy efficiency practices and site access security" [304]. Interviewee ABB2 felt that wider acceptance of this certification, such as by making such a certification scheme mandatory, would benefit the sector.

Interviewee CCC1 reported that with standards schemes in the USA, which set minimum efficiencies for infrastructure such as chillers, or in some instances require aisle containment or economisation, there are often loop holes. They stated that "there's a lot of ways to get around implementing it". This interviewee also reported issues with carbon leakage, with companies opting to build new data centres in states with less strict standards schemes.

6.4.4 Aisle containment

6.4.4.1 Reasons for installing aisle containment

Table 6-7 summarises the responses to the questions regarding reasons for installation of aisle containment. The text following Table 6-7 investigates the interviewees' views in finer detail.

Reasons for installing aisle containment	Agreed	Disagreed	Uncertain	Issue not raised or not applicable
Improving energy efficiency	7	0	0	3
Maximising IT capacity	2	0	0	8
Improving the thermal environment	3	0	0	7

Table 6-7. Summary of responses to questions regarding reasons for installing aisle containment.

All interviewees had aisle containment installed in at least some sections of the data centres which they operated in or had operated in. All interviewees who were asked about aisle containment referred to energy efficiency as a key reason for adoption of this strategy. Interviewees ABB1 and ACC1 also mentioned the potential for containment to increase the available capacity for IT load. Interviewees ACC1, AAB2 and BBB1 mentioned the improved thermal environment, and associated reduction in server failures. However, as interviewee

BBB1 commented, the impact of cold aisle temperatures on frequency of server failures is hard to quantify, so that the promise of improved energy efficiency is the key driver.

6.4.4.2 Barriers to aisle containment

Table 6-8 summarises the responses to the questions regarding barriers to installation of aisle containment. The text following Table 6-8 investigates the interviewees' views in finer detail.

Barriers to installation of aisle containment	Agreed	Disagreed	Uncertain	Issue not raised or not applicable
Cost	3	2	0	5
Practical issues in legacy data centres	4	0	0	6
Lack of knowledge or understanding	3	0	0	7
Low IT loads making containment financially unviable	3	0	0	7
Reduced flexibility	1	0	0	9
Split incentives	1	0	0	9
Time available to invest in strategic improvements	1	0	0	9

Table 6-8. Summary of responses to questions regarding barriers to installation of aisle containment.

Interviewees AAB2, BBB1 and BAA3 identified the costs associated with containment as a barrier to installation. Interviewee AAB2 noted that this was particularly an issue for aisle containment projects in small data centres, stating that "the return on investment for that would be many years". Conversely, interviewees ACC1 and BAA1 felt that basic aisle containment could be installed very cheaply in the form of butchers' curtains, with much of the benefit of more robust methods being achievable in this way. However, as shown in Chapters 2, 3 and 5 of this thesis, such an approach would be unlikely to realise the same benefits as more rigorous aisle containment approaches, since greater recirculation and

bypass would be present, leading to higher required air flow rates, higher temperature rises at server inlets, and a subsequent need for lower supply air temperatures.

Interviewees ABB1, ABB2, BAA2 and BAA3 mentioned issues with legacy data centres presenting various practical challenges which put operators off installing containment. Interviewee ABB1 specifically listed issues such as "rooms not originally set up for it, low floor to ceiling heights, congested floor voids". Interviewee ABB2 linked this issue to capital cost, suggesting that issues with retrofitting aisle containment into legacy facilities can make such measures prohibitively expensive.

A lack of knowledge or understanding was cited by interviewees ACC1, AAB2 and BAA1 as slowing uptake of containment, with operators not understanding the benefits. Interviewee ACC1 linked this issue to the tendency for staff responsible for IT to disregard the importance of segregation of hot and cold air streams through the inappropriate installation of equipment, using the following example to illustrate this point:

"I've seen for instance, an attempt at hot and cold aisle arrangement, and then there's the network cabinet that's round the wrong way. And the reason being is they wanted the connections on one particular side. So rather than turn it round, and take the connections round to the back, or through the cabinet, they just turn it round the wrong way."

Interviewee BBB1 stated that some operators do not understand that other changes to the cooling regime, such as reducing air supply flow rates and increasing supply temperatures, are necessary to realise the benefits of aisle containment. This interviewee recalled a particular conversation with a data centre operator planning to install cold aisle containment:

"'We're gonna buy some cold corridors, we're gonna put them on, and it's gonna save us, and it's gonna pay back in 2 years'. And I'm sort of saying, well what other things are you doing? 'Well we're not doing anything else, we're just gonna put the corridors on'. You're not gonna raise the temperatures? 'Oh no we can't do that because of our SLAs', you're not gonna modulate [i.e. reduce supply air flow rate]? 'No because it's fixed'. I'm saying, well I can't be held responsible for if you're really gonna save that energy just because you're holding the air in there. "

These issues are consistent with the findings of the literature review presented in section 2.1, which noted the lack of information available demonstrating how to control the operating conditions of a data centre employing aisle containment so as to maximise the energy saving.

Interviewees ABB1, BBB1 and BAA2 mentioned that at their own sites, there were areas which remained uncontained due to low IT loads, which they felt made containment unnecessary or not financially worthwhile. It may be argued that this does not represent a barrier from the perspective of the data centre, since containment may not result in a financial saving in these instances. However, there may still be a 'market failure' if the total societal cost saved by installing aisle containment exceeds the cost of installation (see discussion in section 6.1.1).

Interviewee ACC1 highlighted the lack of flexibility following the installation of containment as a barrier, noting that "If you want to stick an extra cabinet in, if you've got fixed containment it's quite a lot of work to change that".

Interviewee ABB2 highlighted the issue of split incentives whether in the form of co-location providers passing on energy bills to customers, or in enterprise data centres in which the team responsible for managing the data centre does not pay the energy bill. This reduces the incentive to improve energy efficiency. They illustrated this issue in the context of an enterprise data centre as follows:

"The IT department might own for example the data centre, but the electricity bill's paid by estates. And so some of the inefficiencies and the cost savings have been on different departments, sort of different people's balance sheets. And therefore it's been quite easy to mask those inefficiencies."

Finally, interviewee BBB1 noted that in some smaller data centres the staff tasked with managing the data centre do not have enough time available to invest in strategic improvements regarding energy efficiency, with priority given to deployment of new servers and dealing with customer issues. This was said to be less of an issue where there are dedicated facilities management staff. They commented that:

"You know, they're so busy, and yes, that is something they'd love to do, but to be honest, they've already spent 12 hours dealing with everything else and they've run out of steam."

6.4.5 Other issues raised during interviews

Interviewee ACC1 cited Jevons' paradox, noting that data centres drive "new and innovative new businesses", and that making data centres more efficient would help these businesses to grow, driving expansion in the data centre sector, ultimately increasing energy consumption. They illustrated the issue as follows:

"So basically, the more you digitise, the more you make data centres efficient, effective...unfortunately it's been shown that the Jevons principle kicks in. People say 'Wow! We can do this now! Couldn't do that before because it was inneffective, innefficient. Now we can.', and so they use more and more services, take up more and more power."

This interviewee went on to state that charging consumers for storage of video and images is the only way to constrain growth in data centre energy consumption, whilst conceding that it would be difficult to force companies to do this.

The citation of Jevons' paradox was echoed by another interviewee, who noted that improvements in energy efficiency were unlikely to result in the reduction of energy consumption.

6.4.6   Saturation

In assessing whether saturation had been reached in terms of the information collected during interviews, findings of interviews with EU CoC-Participants/Endorsers and with non-CoC-Participants/Endorsers were considered separately.

The only new issue raised during interview AAB2 (the last interview conducted with a CoC-Participant/Endorser) was the view that the CRC rewards lower energy consumption, rather than improved energy efficiency. However, two other interviewees had previously noted that the CRC and CCL penalised early adopters of energy efficient technology, which is a related point. In addition, there was a general consensus that the CCA was more effective than the CRC and CCL at driving energy efficiency improvements. There were no new ideas raised during interview AAB1, which was the next to last interview undertaken with a Participant/Endorser of the EU CoC. There were also no new ideas raised during interview BAA3, which was the last interview undertaken with a non-Participant/Endorser of the EU CoC, and also the last interview undertaken overall.

It must be conceded that there remains a possibility that further interviews would have yielded additional findings. However, interview-based research is inherently time consuming, and some trade-off must be made between achieving a reasonable degree of confidence that the findings cover the major issues relating to the research question whilst keeping the required time investment to a manageable level. It seems reasonable to conclude from the foregoing that the number of interviews conducted was sufficient to elicit the most pervasive views within the sector relating to the matters discussed.

6.5   Analysis of findings of semi-structured interviews

As discussed in section 6.1.2, there is evidence that an energy efficiency gap exists in the data centre sector, which implies a need for policy intervention to drive adoption of EEMs within the sector. The key policies impacting on the UK data centre sector at present are the Climate Change Agreement for data centres and the EU Code of Conduct on Data Centre Energy Efficiency.

The responses to questions regarding aisle containment largely served to confirm the prevalence of the barriers to EEMs reported in section 6.1.5, with 6 of the 10 barriers reported in section 6.1.5 being raised by interviewees (as summarised in Table 6-8). Those barriers which were not raised during the interviews are either not related directly to aisle containment, or are related to other barriers which were mentioned. For example, **prioritisation of reliability over efficiency** is arguably not relevant, since aisle containment is likely to improve reliability. **Requirement for central management approval** was not mentioned, although this may be related to the idea that staff do not have time to look into EEMs, which was raised by one interviewee. **Disincentives for trade allies** was not mentioned, but the main trade ally relevant to this measure is the installer of the containment system, who is clearly incentivised to recommend installation. **Changing power demands** were not mentioned. However, since such changes are generally in the direction of increasing power consumption and density, they would only tend to strengthen the case for aisle containment.

Interviewees seemed to recognise that different companies faced different issues, which perhaps led to limited differentiation between the responses of interviewees from different kinds of companies. For example, interviewee AAB2 (representing a relatively large enterprise data centre) noted that, for smaller companies, capital cost may present a barrier to installation of aisle containment. However, interviewees representing smaller data centres were no more likely to cite capital cost as a barrier.

All 3 interviewees citing practical issues as a barrier identified this as a problem in enterprise data centres, although only one of these interviewees actually represented an enterprise data centre. Interviewees representing small and large, co-location and enterprise data centres all cited lack of knowledge and understanding as barriers to introduction or effective management of aisle containment, although none felt that these issues applied to their data centres. It should be noted that, due to the limited number of interviews undertaken, the number of interviewees representing each sub-division of data centres (e.g. small or large,

co-location or enterprise) is very small. Hence, the extent to which conclusions can be drawn on the importance of different barriers to different groups is limited.

The findings of the semi-structured interviews show that the best practice guidelines of the EU CoC are highly regarded by data centre professionals. However, criticisms were made regarding the administration and promotion of the CoC, and the lack of enforcement around periodic reporting, with these issues being purported to limit its adoption and impact. These findings are consistent with the key issues raised in the UK Council of Data Centre Operators' report [260] on the EU CoC, as discussed in section 6.1.6.2. Some interviewees felt that the adoption of the CoC's best practice guidelines into CENELEC's EN 50600 standards series would improve the credibility of the scheme by providing a route for stricter enforcement. Again, this is consistent with the views of the UK Council of Data Centre Operators' report [260]. The voluntary nature of the EU CoC was also highlighted as limiting its impact. As noted in section 6.1.4, there have been no major attempts to quantify the impacts of voluntary agreements such as the EU CoC, although their impacts are likely to be limited by take up [28], and may also be impaired by any lack of rigour in monitoring of compliance [228]. Thus, the key findings of the semi-structured interviews and literature review with respect to the EU CoC may be summarised as follows:

- The EU CoC's best practice guidelines are broadly thought to represent an effective approach to achieving a high level of efficiency in data centres.
- The impact of the EU CoC is limited by its inefficient administration, lack of promotion, its voluntary nature, and lack of enforcement of reporting amongst Participants.

The consensus amongst interviewees that the DC CCA has been effective in driving efficiency improvements in the sector is consistent with opinions reported in TechUK's report on the policy, as discussed in section 6.1.6.1 [256]. As reported in section 6.1.4, research has suggested that both the CRC and CCA schemes have been effective at improving efficiency in the industries to which they apply, although there was no clear evidence as to which scheme was the most effective. TechUK's complaints regarding the burdensome nature of reporting under the CCA, CRC, and other policies [256] were only repeated by one interviewee. As discussed in section 6.1.4, prescriptive policies such as the CCA can be highly effective, provided that the targets are sufficiently ambitious [213]. One drawback can be the expense of implementation and maintenance, due to the time investment required to identify suitable targets, to update these targets in response to changing technology, and to monitor performance against these targets [213]. In the case of the DC CCA, the costs have

been minimised by selecting a simple and pre-existing metric, namely the PUE. However, this limits the policy's scope to efficiency of cooling and power provision, as highlighted by interviewees ACC1 and AAB1. There was also no clear indication from the interviewees as to whether or not the targets of the CCA were deemed to be ambitious enough.

One interviewee felt that the application of energy taxes to the data centre sector in any form would tend to encourage companies to keep IT services in house, which could be provided with greater efficiency by co-location facilities. However, this comment should be considered in the context of the research referenced in section 6.1.4 which suggested that the CCL, both with and without the CCA, has had either no effect on employment levels and revenue [223] or has increased them [219]. One interviewee also felt that some data centre operators would not be deterred from inefficient behaviour by an energy tax approach, since for some data centres, resilience is of far greater importance than energy expenditure. Thus, the key findings of the research interviews and literature review relating to the DC CCA, the CRC and the CCL may be summarised as follows:

- The DC CCA is generally more popular within the data centre sector than the CRC and CCL schemes.
- The financial incentive provided by the CCA is likely to have some positive impact in driving improvements in energy efficiency.
- There is no clear indication as to whether or not the targets for reductions in PUE represent 'ambitious' targets.
- The use of the PUE as the measure of improvement in energy efficiency limits the policy's scope to efficiency of cooling and power distribution.
- The impact of the DC CCA is limited by its exclusion of enterprise data centres.

Most interviewees with whom the issue was discussed broadly agreed that subsidised energy audits and engineering design services could help to drive adoption of EEMs. However, no interviewees seemed to be particularly enthusiastic about this approach and some expressed doubts about the extent of its potential impact. These doubts were consistent with failures noted in similar schemes in the US (as discussed in section 6.1.5), and included (i) the relatively low costs of these expenses in comparison with costs of implementation of EEMs, (ii) the importance of knowledgeable personnel being responsible for provision of these services, and (iii) the potential for freeridership. A number of interviewees did agree that there was an issue with a lack of understanding of energy efficiency amongst those working in the sector, although not all felt that free or subsidised audits and design services would help to address this. As noted in section 6.1.4, these kinds

of policies have received little attention in the peer-reviewed literature, although there is evidence that provision of free auditing services can increase the frequency of audits being undertaken [238]. In summary, whilst the evidence regarding the likely impact of these approaches is limited, there is some evidence to suggest that such measures could help to maximise the impacts of other, complementary prescriptive policies and economic policies (such as the EU CoC and DC CCA, respectively).

Three interviewees (ABB1, BAA2 and ABB2) felt that a compulsory energy efficiency standard within the sector, possibly with financial penalties for non-compliance, could be an effective driver for improved efficiency. The existing schemes referred to in this context were the National Home Energy Rating Scheme, the DCA Certification Scheme and the ESOS.

6.6   Proposal for changes to the current policy environment

The literature review and semi-structured interview results presented in this chapter have highlighted a number of benefits of the current policy environment surrounding data centres in the UK.

There is broad agreement that the best practice guidelines associated with the EU CoC are fit for purpose, and result in a high level of efficiency when implemented. Whilst formal Participation in the EU CoC has been limited, there is some indication that the best practices document has an influence on behaviour in the sector more broadly. However, the impact of the scheme is limited by its voluntary nature, as well as issues with its purported poor administration and lack of promotion.

There is also broad agreement that the financial incentive represented by the DC CCA has focussed attention on improving efficiency within the sector. However, there is no clear indication of whether or not the targets of the CCA are ambitious enough to drive an improvement in energy efficiency which goes beyond BAU.

The DC CCA is also limited by its reliance on the PUE metric, which measures only the efficiency of cooling and power infrastructure, ignoring the efficiency of IT. This is a major limitation, due to the potential for energy savings via increased utilisation of servers, as discussed in section 1.3.4. However, it could be argued that since the DC CCA currently only applies to co-location providers, who typically have limited influence over IT management, IT efficiency is inherently beyond the scope of the current scheme. Even so, broadening the scope of the scheme to encompass IT efficiency would be beneficial if possible. In addition,

a worsening performance against PUE may be seen where IT power consumption reduces unexpectedly, or where a facility is expanded in anticipation of increased IT power consumption, despite such changes arguably not representing poor energy management. This could lead to the unfair penalisation of participating data centres.

Another limitation of the DC CCA is its lack of reach, since it only applies to co-location providers. Broadening the scheme to encompass enterprise data centres would make the limitations associated with the PUE metric more pressing, since enterprise data centres have direct control of IT management.

Table 6-9 lists the barriers to energy efficiency in the data centre sector, as identified in section 6.1.5, and identifies which are addressed by the EU CoC and the DC CCA. The table also identifies which barriers could be addressed by a modified version of the DC CCA (referred to in the table as '**DC CCA – proposed**'), which will be described in the remainder of this section.

The key benefits of the DC CCA stem from the financial incentive it provides, through the threat of an increased energy tax for those failing to meet the target. The EU CoC targets a broader range of barriers, through its sophisticated best practice guidelines, which recommend not only practical measures, but also management and decision-making procedures which would address some of the barriers specific to this sector. Amending the current DC CCA scheme to require data centres to work towards, and ultimately achieve, Participant status under the EU CoC, would allow the majority of barriers to be addressed. Such a change would also enable the scheme to target IT efficiency, which is not possible whilst using a PUE-based target. Note that the EU CoC best practice guidelines incorporate flexibility around requirements for specific IT management practices, recognising that, for example, co-location data centres may not be able to force customers to undertake certain measures to improve IT efficiency, or that some legacy data centres may find it impractical to install aisle containment (this was discussed in section 6.1.6). This limits the impact of the standard, but makes it easier to achieve. The level of flexibility with which the best practice guidelines should be enforced would depend on the goal of the policy instrument. For example, the intention may be to drive only EEMs which are cost-effective for the individual company, or it may wish to also account for the externalised costs of environmental impacts. Achieving the latter may require additional financial support in order to avoid making the UK data centre sector unable to compete with its international competitors.

It must be acknowledged that such a change would be likely to increase the cost of the scheme, since monitoring compliance with the EU CoC would be more labour intensive than the simple electricity consumption monitoring required to calculate PUE. Estimating the relative costs and benefits of such a scheme is without the scope of this thesis. However, the clear consensus over the presence of an efficiency gap within the sector and the high regard in which the EU CoC best practice guidelines are held could be taken to indicate that such a requirement could have a net benefit. In addition, the pre-existence of the best practice guidelines, and the structure for assessing compliance against them via the BS EN 50600 standards series, would minimise the associated costs.

Barrier	Policy instrument		
	EU CoC	DC CCA - current	DC CCA – proposed
Changing power demands	✓		✓
Prioritisation of reliability over efficiency		✓	✓
Split incentives	✓	✓	✓
High capital costs		✓	✓
Poor communication between facilities and IT staff	✓		✓
Requirement for central management approval	✓		✓
Lack of knowledge and understanding	✓	✓	✓
Prevalence of inefficient, legacy data centres			
Short leases on data centre sites			
Disincentives for trade allies	✓		✓

Table 6-9. Which policy instruments address which barriers to energy efficiency in the data centre sector?

Another potential route to increasing the scope and impact of the CCA would be to incorporate enterprise data centres into the scheme. This would be challenging, since data centres are inherently difficult to define and identify, as discussed in section 1.2.1. However, the benefits could be considerable. As noted in section 6.1.5, a survey conducted in 2013 estimated that around 77% of UK data centre electricity consumption occurred in enterprise data centres, although this was thought to be falling [255]. Hence, even an only partially

successful integration of enterprise data centres into the DC CCA, for example if only the largest enterprise data centres were targeted, could significantly expand its reach.

As shown in Table 6-9, the suggested change to the DC CCA would not directly address the **prevalence of inefficient, legacy data centres**, or **short leases on data centre sites** barriers. However, the increased financial incentive and information sharing driven by such a scheme could help to lessen the relative strength of these barriers. With regards legacy data centres, the scheme would be likely to speed up the process of moving IT services to newer, more efficient data centres, since legacy data centres would be more likely to incur extra costs.

6.7 Summary

The literature review provided in section 6.1 has demonstrated the need for policy intervention in the data centre sector to drive energy efficiency improvements. Section 6.1 has also shown that there is a lack of academic research into the efficacy of the policy regime in this sector. Section 6.2 has then presented a literature review relating to social research methods relevant to the study of energy efficiency-related policy instruments. The findings of section 6.2 have informed the design of semi-structured interviews used to investigate the attitudes of people working within the data centre sector towards energy efficiency-related policy instruments. The methodology and findings of these semi-structured interviews are shown in sections 6.3 and 6.4, respectively. Taken together with the findings of section 6.1, the results of the semi-structured interviews have demonstrated that there are many positive aspects of the current policy environment. However, there are some major deficiencies, notably (i) the cumbersome administration and voluntary nature of the EU CoC, (ii) the limited application of the DC CCA to co-location providers, and (iii) the DC CCA's reliance upon the PUE metric, which only targets the efficiency of cooling and power distribution. An amendment to the current policy regime has therefore been recommended, whereby participants of the DC CCA must adhere to the well-respected best practice guidelines of the EU CoC in order to avoid the full costs of the CCL and CRC. It is clear then, that the work presented in this chapter represents the fulfilment of objective (4) (as defined in section 1.5).

The findings of the semi-structured interviews included observations from interviewees that aisle containment was not always well understood by data centre operators. Specifically, it was reported that some data centre operators do not understand the need to reduce supply air flow rates and increase supply temperatures in order to reap the benefits of aisle containment. Also, it was reported that poor IT management can compromise the efficacy of containment. The best practice guidelines of the EU CoC go some way to addressing these

issues, by recommending the reduction of supply air flow rates and increase of supply air temperatures where aisle containment is installed. Hence, any policy instruments which increase rates of adherence to these best practice guidelines would help to address these issues. These findings also underline the importance of the work presented in Chapters 2 to 5, which investigated the impact of cold aisle pressurisation on power consumption and recirculation, and which specifically demonstrated the value of controlling the supply flow rate to data centres employing aisle containment.

7. CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

This thesis presents research which seeks to facilitate improvements in energy efficiency in data centres. The importance of this issue has been demonstrated with respect to the data centre sector's large, and growing, electricity consumption. The reviews of the literature presented in Chapters 1 and 2 have demonstrated that well-established energy efficiency measures (EEMs) such as aisle containment are not employed ubiquitously within the sector, and that quantification of the benefits of such measures has been limited. Accordingly, the aim of the study was identified as follows:

'To investigate the technologies and policy instruments available to improve efficiency in data centre cooling, with a particular focus on air management in data centres employing aisle containment.'

To this end, experimental and numerical investigations have been undertaken to further efforts to quantify the impact of aisle containment on electricity consumption and cooling efficacy, and to identify methods for maximising the benefits of this measure. In addition, the suitability of the current policy environment seeking to drive improvements in UK data centre energy efficiency has been assessed, through a literature review in addition to research interviews. Further, some potential improvements to the current policy environment have been proposed.

This chapter seeks to summarise the key findings of the work, make proposals for further work to build on these findings, and offer a broader perspective on the implications of the research for the data centre sector than has been offered elsewhere in this thesis. Specifically, section 7.1 summarises the key novel contributions of this work to the pre-existing research literature. Section 7.2 refers back to the research objectives identified in section 1.5, and outlines how each has been achieved. Section 7.3 details the limitations of the research presented, and outlines the impacts of these limitations on the strength of the conclusions drawn. Section 7.4 makes suggestions for further research to expand on the findings of this thesis. Section 7.5 comments on the broader implications of the work and provides a summary of key recommendations to stakeholders in the sector.

7.1   Novel contributions

This thesis has made a number of novel contributions to the literature regarding energy efficiency in data centres. Firstly, the experimental results presented in Chapter 2 represent the most complete investigation into the extent of bypass and recirculation in aisle contained data centres undertaken to date. No previous researcher has presented data

showing variation of bypass and recirculation with the extent of cold aisle pressurisation (Δp_{CH}), nor have the benefits of improving rack design to minimise leakage been quantified. Chapter 2 represents a valuable resource for data centre operators and rack designers seeking to minimise rack leakage in data centres employing aisle containment.

Secondly, the results from the system model presented in both Chapters 3 and 5 represent the first quantification of the impacts of rack design and Δp_{CH} on electricity consumption in data centres employing aisle containment. This is an important contribution, since it empowers data centre operators to easily make a business case for sealing rack leakage paths and minimising Δp_{CH}.

Thirdly, the CFD investigations presented in Chapter 5 represent a significant contribution to the development of methods used to model air flows in data centres employing aisle containment. Only one CFD model has previously been presented which includes rack leakage for a data centre employing aisle containment [72]. The models described in Chapter 5 are the first to model rack leakage using the results of experimental measurements, and the chapter represents the first CFD investigation of the effect of Δp_{CH} on cooling efficacy and electricity consumption. The validation of the models has shown for the first time that both potential flow and Navier Stokes methods can be used to accurately predict the impact of a change in rack sealing efficacy (in this case, the introduction of an empty slot) on inlet temperatures, using the new method for modelling rack leakage.

Fourthly, the literature review and analysis of research interview findings presented in Chapter 6 represent the first academic study of the policy environment within which UK data centres operate. As demonstrated in the chapter, the success of policy instruments relating to energy efficiency can vary greatly. Hence, a critical, academic analysis of the policy environment is an important contribution to efforts to improve energy efficiency in the sector, and provides a useful resource to those working to refine existing policies, or to develop new ones.

7.2 Achievement of the research objectives

7.2.1 Objective (1)

Objective (1) was identified in section 1.5 as follows:

'To investigate the extent of bypass and recirculation in data centres employing aisle containment'

Previous experimental and numerical investigations have demonstrated the potential to reduce bypass and recirculation in data centres by physically segregating hot and cold air streams via aisle containment. However, efforts to precisely quantify the extent of bypass and recirculation under contained conditions have been limited, with key factors such as rack design and the extent of cold aisle pressurisation having received scant attention. This gap in the literature has been addressed with the experiments described in Chapter 2.

Firstly, an extensive study has been undertaken to measure the relationship between the bypass flow rate ($\dot{V}_{bypass}$) through a rack and the pressure drop across it (Δp_{CH}). This was undertaken for a number of racks from different manufacturers. For each rack, a series of tests was undertaken to determine the key leakage paths and the potential to minimise bypass through simple, practical means. $\dot{V}_{bypass}$ was shown to be strongly dependent on rack design and Δp_{CH}, ranging from 0.6 to 38.9% of supplied air flow depending on rack design, Δp_{CH} and rack power density. The results demonstrate the importance of minimising Δp_{CH}. The key leakage paths within the racks were identified as the spaces at the sides of the equipment rails, any gaps between folded metal sections of the rack walls and floor, and any holes in the folded metal sections of the rack.

Further tests were conducted in a Test Data Centre employing aisle containment, using one of the racks from the single rack tests, with a modified design. The modified rack used standard parts available from the rack manufacturer to minimise leakage, particularly through the key leakage paths identified in the single rack tests. The results demonstrated that this rack design, coupled with a well-designed containment system, could be used to ensure relatively low levels of bypass and recirculation. Tests conducted using a smoke pen showed that leakage through the containment system was small in comparison with leakage through the racks.

It is clear that the results presented in Chapter 2 demonstrate the achievement of objective (1), since bypass and recirculation have been quantified under a wide range of conditions representing air flows in aisle contained data centres.

7.2.2   Objective (2)

Objective (2) was identified in section 1.5 as follows:

'To investigate the implications of bypass and recirculation for electricity consumption in data centres employing aisle containment'

In Chapter 3, a system model was described which used the experimental results presented in Chapter 2, alongside manufacturer data and empirical relationships, to predict the impact

of rack design, Δp_{CH} and IT power density on data centre electricity consumption. The results demonstrated that bypass has a significant impact on electricity consumption in data centres employing aisle containment, and that action taken to minimise Δp_{CH} and to seal leakage paths within the racks is necessary to fully realise the benefits of aisle containment. Specifically, taking action to seal leakage paths could reduce electricity consumption by up to 36%, whilst reducing Δp_{CH} from 20 to 2 Pa could reduce electricity consumption by up to 58%. The results broadly confirm the appropriateness of the EU Code of Conduct on Data Centre Energy Efficiency's (EU CoC) recommendations to maintain $\Delta p_{CH} < 5\ Pa$, and to seal leakage paths in racks as far as is practical.

The predictions of the model were strongly affected by the assumptions made regarding the response of server fans to variation in Δp_{CH}. This demonstrates the importance of improving understanding of server behaviour under pressurised conditions.

The new system model is the first to incorporate Δp_{CH} as an input variable. This is crucial due to the increasing prevalence of aisle containment, and the impact of Δp_{CH} on the required air supply flow rate, as demonstrated in Chapter 2. The results presented in Chapter 3 further demonstrate the importance of consideration of this variable in predicting electricity consumption in data centres employing aisle containment. The results presented in Figure 3-9 and Figure 3-10 make it simple for data centre operators to estimate the benefits of effective management of aisle containment systems in terms of reductions in electricity consumption, enabling a business case to be made for improving rack sealing and minimising Δp_{CH}.

The analysis of the results of the CFD models presented in Chapter 5 has also contributed to objective (2). Since the system model described in Chapter 3 included no air flow modelling, it was necessary to assume uniform pressure and temperature conditions within each aisle. This is a simplification since, as predicted by the CFD models in Chapter 5, pressure variations are caused by variations in air velocity, and by interactions resulting from viscous stresses. These pressure variations can lead to a mixture of bypass and recirculation, even where the pressure in the cold aisle generally exceeds that in the hot aisle. Hence, air flow modelling was necessary in order to fully investigate the implications of the experimental results presented in Chapter 2.

It is clear that the results of the system model presented in Chapters 3 and 5 demonstrate the achievement of objective (2). The impacts of Δp_{CH} and rack design on data centre electricity consumption have been thoroughly investigated, with these two factors having been shown in Chapter 2 to govern bypass and recirculation.

7.2.3 Objective (3)

Objective (3) was identified in section 1.5 as follows:

'To investigate the potential for using CFD models to aid in the efficient design and management of data centres employing aisle containment'

The literature review presented in Chapter 4 demonstrated that most of the CFD models of data centres employing aisle containment previously presented in the research literature have either disregarded bypass and recirculation, or have not described the methods used to model these air flows. Bypass and recirculation through racks in particular has received very little attention, and the importance of Δp_{CH} as an input parameter has not previously been considered in a systematic way. Since the results of Chapters 2 and 3 have demonstrated the importance of bypass and recirculation within data centres employing containment, and of the Δp_{CH} parameter in particular, this was identified as a deficiency in CFD models presented to date.

In Chapter 5, two CFD models of data centres were presented, one based on the potential flow equations, and the other on the Navier-Stokes equations. The models utilised the experimental results presented in Chapter 2 to govern bypass and recirculation, and were validated against experimental measurements undertaken in the Test Data Centre. Both models were able to predict load bank inlet temperatures with accuracies comparable with those of the best performing models described previously in the research literature. The validation demonstrates the accuracy of the new methods used to model bypass and recirculation, which demonstrates an important contribution to CFD modelling of data centres employing aisle containment.

After validation of the two models against results from the Test Data Centre, both models were modified to represent an example data centre geometry. The modified models, in conjunction with a modified version of the system model described in Chapter 3, were used to investigate the impact of Δp_{CH} on electricity consumption and cooling efficacy. The minimum level of Δp_{CH} which was required in order to achieve the necessary flow rate through the servers was shown to be strongly dependent on the assumptions made regarding the response of server fans to variation in Δp_{CH}. This further stresses the importance of research into the behaviour of servers under pressurised conditions.

Both models predicted considerable variation in pressure along the length of each aisle, with the NS-CFD model showing much greater variation than the PF-CFD model. This difference strongly affected the level of Δp_{CH} required to maintain the necessary flow rates through

the servers. Validation of the two models against pressure measurements in data centres larger than the Test Data Centre would be required to determine the accuracy of the predicted pressure fields, and by extension the accuracy of the predictions of the optimum level of Δp_{CH}.

The potential for the pressure to vary within a contained aisle has not been discussed previously within the research literature. The simulation results, if accurate, show that control systems intended to maintain a certain minimum Δp_{CH} may need to take numerous pressure measurements along the length of an aisle in order to ensure that the desired conditions are achieved. This finding also has implications for data centre design. Specifically, a geometry which minimises variation in pressure would be beneficial since it would enable Δp_{CH} to be maintained close to the desired level in all areas of the data centre. Conversely, large variation in Δp_{CH} could require high pressures to be maintained in some regions in order to ensure that adequate server air flows are achieved in low pressure regions. The high pressure regions would correspond to regions of high bypass (as demonstrated by the results presented in Figure 2-20, Figure 2-22 and Figure 2-24), and excessive server flow rates, and would thus increase electricity consumption.

Both models showed that taking steps to seal leakage paths within racks helps to minimise electricity consumption, by reducing the CRAH fan power consumption required to achieve adequate cooling. This is consistent with the findings of the system model as presented in Chapter 3, and shows that these findings hold where air flows are modelled. This underpins the validity of demonstrating a business case for improving rack sealing based on the results of the system model.

Both models predicted very little difference between supply temperature and server inlet temperatures, provided that $\Delta p_{CH} \geq 0$. This demonstrates that very low levels of recirculation can be achieved where aisle containment is installed and sufficient supply air flow rates are provided.

The CFD models presented in Chapter 5 have shown that variations in pressure can occur within individual data centre aisles, casting doubt on the assumption of constant pressure within each aisle which was used in the system model as presented in Chapter 3. It follows that CFD modelling is an important tool in the design and management of data centres employing aisle containment. The validation presented in Chapter 5 has demonstrated that CFD models can make accurate predictions of server inlet temperatures in these data centres. Through the important contributions to the ongoing development of CFD methods

relating to these data centres, and through the analysis of the models' predictions, objective (3) has been achieved.

7.2.4   Objective (4)

Objective (4) was identified in section 1.5 as follows:

'To investigate the potential for policy instruments to drive energy efficiency improvements in the data centre sector'

The literature reviews presented in Chapters 2 to 5 demonstrated that understanding of the impacts of aisle containment on electricity consumption and thermal environment is limited, and particularly that the measures needed to maximise the benefits of aisle containment are not well understood. This is symptomatic of a broader issue in the data centre sector and other sectors, by which beneficial EEMs remain unimplemented due to issues such as lack of information and understanding, and various structural disincentives. It follows that research addressing the technical challenges in data centre energy efficiency must be complemented by research addressing the political challenges.

The literature review presented in Chapter 6 demonstrated that there is likely to be a considerable energy efficiency gap in the data centre sector. The key barriers to progress in energy efficiency were identified, and the policy instruments currently acting to address these barriers were examined. A series of research interviews were carried out with people working in the data centre sector. This enabled a deeper analysis of the impacts of the key policy instruments, and the collation of expert opinions on potential improvements to the current policy environment. The key policy instruments affecting the UK data centre sector were identified as the EU CoC and the UK's Umbrella Climate Change Agreement for the standalone data centres (DC CCA).

The criticisms of the EU CoC and DC CCA obtained in the research interviews were considered in the context of the findings of the literature review. The EU CoC best practice guidelines were generally considered to represent a high standard of efficient data centre management. However, the scheme's impact was found to be limited by its voluntary nature, as well as a lack of promotion and issues with its administration. The financial driver represented by the DC CCA was generally agreed to be effective in incentivising action on energy efficiency. However, its impact was limited by its reliance on the PUE metric (which only targets efficiency of cooling and power delivery), and by its restriction to co-location providers. There was broad agreement on these issues between the findings of the interviews and of the literature review. In light of these findings, an amendment to the DC

CCA was proposed, whereby a financial saving was granted subject to adherence to the EU CoC best practice guidelines, rather than a reduction in PUE. It was also proposed that the scheme be expanded to include enterprise data centres, where possible.

Taking the specific example of aisle containment, the findings of the interviews confirmed that there are issues with a lack of understanding of the benefits of aisle containment, and of the measures required to maximise the benefits. The recommendations within the EU CoC best practice guidelines that data centres install aisle containment, maintain low, positive Δp_{CH}, and seal prominent leakage paths within racks, would help to address these issues. Hence, the proposed amendment to the DC CCA could help to drive the adoption and effective management of aisle containment.

Chapter 6 represents the most thorough analysis of the policy environment relating to data centres since the introduction of key policy instruments such as the EU CoC and the DC CCA. As such, objective (4) has been achieved. The work represents an important contribution to efforts to create effective policy instruments in the data centre sector, such that a balance can be struck between nurturing this strategically important sector and minimising its electricity consumption.

7.3   Limitations of the research

In Chapter 2, it was highlighted that there were significant errors in the measurement of flow rates. However, these errors were small in comparison with (i) the variation in rack leakage flow rates between different racks, (ii) the impact of sealing rack leakage paths and (iii) the impact of Δp_{CH}. Hence, the errors were sufficiently small to ensure that the results provide a useful guide for the impacts of Δp_{CH} and rack design on leakage flow rates.

In Chapters 3 and 5, it was highlighted that the results of the system model and potential flow CFD model were strongly dependent upon the algorithms used to model the response of server fans to the level of Δp_{CH}. The two algorithms used to predict server fan behaviour within the system and CFD models represent two extremes of possible behaviours. Real server fan behaviour is likely to vary from server to server, but to exist somewhere between the two extremes. Regardless of the impact of the methods used to model server fan behaviour, the work presented in Chapters 3 and 5 represents the most detailed investigation of the impacts of Δp_{CH} and rack design on cooling efficacy and electricity consumption in data centres employing aisle containment presented to date.

In the validations of the CFD models presented in Chapter 5, it was highlighted that both models had limited success in predicting temperatures away from the load bank inlets.

However, inlet temperatures were predicted accurately. Since prediction of inlet temperatures is the most important function of a data centre CFD model, the errors in prediction of temperatures elsewhere in the domain do not fundamentally compromise the value of the models.

The recommendation for an amendment to the DC CCA described in Chapter 6 was based on the findings of the ten research interviews conducted, in addition to the literature review presented in the chapter. The number of interviews conducted was small, and interviewees were not asked to give their opinions on the proposed amendment, since it was designed after the completion of the interviews. In addition, there was no attempt to predict the costs of the change to the policy, although it was acknowledged that the change would be likely to increase costs. In any case, the work presented in Chapter 6 represents a significant contribution to research into policy instruments used to drive energy efficiency in the sector, which to date has been very limited. The information collected and analysed in Chapter 6 was sufficient for an initial recommendation for a change in policy to be made, whilst accepting that further work would be required to ascertain the likely costs and impacts of the change.

7.4 Future work

The results and analyses presented in this thesis have raised a number of issues which warrant further research.

The results of the system model and CFD models were strongly dependent upon the assumptions made regarding the behaviour of servers under pressurised conditions. Since it has been demonstrated elsewhere that changing Δp_{CH} changes the flow rate through servers, the rate of heat transfer from servers must be affected by Δp_{CH}. This is likely to affect server fan speeds, as server fans are typically controlled in response to temperatures within the CPU. In addition, the power consumption associated with carrying out computations is affected by the temperatures of components within the server. This leads to a complicated relationship between Δp_{CH}, server power consumption and server flow rate, which is specific to each individual server. Investigations of this behaviour to date have been very limited. Further studies would help to inform system modelling and CFD modelling of data centres employing aisle containment. Specifically, experiments on individual servers within wind tunnels would be useful, as these would enable flow rate through and power consumption of the servers to be measured as functions of Δp_{CH}, in a controlled environment. Since different servers have different geometries and server fan

algorithms, numerous servers would need to be tested in order to assess the variation in behaviour of different models.

Further work is also required to validate the CFD models. Whilst both models were effectively validated against experimental data from the Test Data Centre, applying the potential flow and Navier-Stokes equations to an example data centre geometry with longer aisles resulted in two quite different pressure fields. Since pressure variation within the Test Data Centre was very limited, the models' abilities to predict pressure variation have not been properly tested. This is important since local variations in pressure affect the rates of bypass and recirculation, server flow rates and server power consumptions, and may also determine whether or not each server receives adequate cooling. Hence, a validation against pressure measurements from a data centre with longer aisles is necessary to demonstrate the reliability of the predictions of the models of the example data centre geometry.

Provided that a successful validation of the CFD models can be achieved, the models could be used to investigate the impact of data centre geometry on electricity consumption. As highlighted in Chapter 6, data centre geometries which result in little variation in pressure within each aisle could help to minimise bypass and recirculation, whilst ensuring all servers receive adequate cooling. The CFD models could be used in a parametric study to define the impacts of factors such as aisle length and IT power density on pressure variation, and by extension on electricity consumption and cooling efficacy.

Whilst the proposed change to the DC CCA has been analysed with reference to the existing literature and interview findings, a more detailed investigation into the viability of the proposals would be useful. Specifically, further interviews with stakeholders in the sector could be undertaken to garner opinions on the proposals. Here, stakeholders could include data centre managers, as well as people involved in the administration and design of the EU CoC, DC CCA and the CENELEC EN 50600 standards series. This would enable an informed discussion of the detail of how such a proposal could work in practice to balance the aims of driving efficiency improvements and nurturing this important sector. A full cost benefit analysis of the proposal should also be undertaken.

7.5 Final summary and reflections

Data centres provide digital services without which many aspects of modern life would be unrecognisable. They consume an increasingly significant proportion of the world's electricity supply, and consequently contribute significantly to the world's greenhouse gas

emissions. However, they also provide services which enable greater efficiency and reduced greenhouse gas emissions in other sectors. Hence, there is a need to drive efficiency in the sector, enabling these crucial services to be provided whilst minimising their environmental impact.

During the time in which this thesis was produced, there has been an increase in attention on energy efficiency in data centres, both from academic researchers and policy makers. A large proportion of the experimental and numerical research cited in Chapters 2, 3 and 5 has been published during this period [73], [74], [77], [80], [83]–[94], [96], [128], [137], [171], [175]. In addition, key UK policy instruments such as the DC CCA and BS ISO/IEC 30134-2:2016 PUE standard have been introduced, and the EU CoC best practice guidelines have been updated to include more detailed stipulations regarding management of aisle containment systems, amongst numerous other measures.

These developments have led to improvements in efficiency within the sector. However, there is great potential to further improve efficiency, although there are major technical and political challenges inherent in doing so. This thesis has made significant contributions to addressing both kinds of challenges, and provides a useful resource to data centre operators and policy makers. The thesis has also demonstrated that the technical and political challenges are interconnected, inasmuch as improved understanding of the appropriate implementation of EEMs must be met with greater dissemination of knowledge within the sector, and provision of drivers for implementation of best practice.

The experimental and numerical investigations presented in Chapters 2 to 5 focussed on addressing the technical challenges related to one particular EEM – aisle containment. This measure has been recognised for a number of years as being beneficial to both energy efficiency and reliability. However, its use is not ubiquitous, and there are also issues with inappropriate management of data centres employing aisle containment leading to its benefits going unrealised. The EU CoC's best practice guidelines describe numerous other beneficial measures which similarly are not ubiquitous. Whilst there are numerous people working within the sector who are passionate about improving energy efficiency, support is necessary to enable the sector as a whole to overcome the various barriers to progress. The continued provision of the essential services for which data centres are responsible requires concerted action to continue to (i) improve the state of the art in data centre energy efficiency through experimental and numerical research and (ii) bring an increasing proportion of the data centre sector in line with this state of the art through the

development of new policy measures and the refinement of existing ones. Failure to do so will inevitably result in unsustainable growth in energy consumption within the sector.

To conclude this thesis, and in light of its findings as summarised previously in this chapter, it is appropriate to make the following recommendations to stakeholders in the data centre sector.

i)   Server rack manufacturers and data centre managers utilising aisle containment should endeavour to ensure that leakage paths within server racks are sealed.

ii)   Data centre managers utilising aisle containment should ensure that the pressure differential between cold and hot aisles does not exceed 5 Pa.

iii)   Data centre designers should avoid configurations which produce large pressure variation within individual aisles.

iv)   UK legislators should consider modifying the Climate Change Umbrella Agreement for the Data Centre Sector to require participants to meet the best practice guidelines of the EU Code of Conduct on Data Centre Energy Efficiency.

Bibliography

[1]   Y. Joshi and P. Kumar, "Introduction to Data Center Energy Flow and Thermal Management," in Energy efficient thermal management of data centers, Y. Joshi and P. Kumar, Eds. Springer Verlag, 2012.

[2]   D. L. Beaty, "Internal IT Load Profile Variability," ASHRAE J., vol. 55, no. 2, pp. 72–74, 2013.

[3]   J. Koomey, "Growth in data center electricity use 2005 to 2010," 2011. [Online]. Available: http://analyticspress.com/datacenters.html. [Accessed: 04-Nov-2013].

[4]   European Commission, "Code of Conduct for Energy Efficiency in Data Centres," 2017. [Online]. Available: https://ec.europa.eu/jrc/en/energy-efficiency/code-conduct/datacentres. [Accessed: 14-Jun-2017].

[5]   U.S. Department of Energy and U.S. Environmental Protection Agency, "Energy Efficiency in Data Centers: Recommendations for Government-Industry Coordination," 2008. [Online]. Available: http://www1.eere.energy.gov/manufacturing/tech_assistance/pdfs/data_center_wrkshp_report.pdf. [Accessed: 13-Jan-2014].

[6]   M. Salim and R. Tozer, "Data Centers' Energy Auditing and Benchmarking-Progress Update," in ASHRAE Winter Conference, 2010, pp. 109–117.

[7]   Intellect, "Er, what IS a data centre?," 2013. [Online]. Available: https://www.techuk.org/images/documents/Data_Centres_-_CCA/Note_03_Er_what_is_a_data_centre.pdf. [Accessed: 13-Nov-2015].

[8]   T. Daim, J. Justice, M. Krampits, M. Letts, G. Subramanian, and M. Thirumalai, "Data center metrics: An energy efficiency model for information technology managers," Manag. Environ. Qual., vol. 20, no. 6, pp. 712–731, 2009.

[9]   M. Bailey, M. Eastwood, T. Grieser, L. Borovick, V. Turner, and R. C. Gray, "Data Center of the Future," 2006. [Online]. Available: http://datacenters.lbl.gov/sites/all/files/IDC_Bailey_2006.pdf. [Accessed: 04-Apr-2014].

[10]   A. Shehabi et al., "United States Data Center Energy Usage Report," 2016. [Online]. Available: https://eta.lbl.gov/sites/all/files/publications/lbnl-1005775_v2.pdf. [Accessed: 15-Jun-2017].

[11]   P. Bertoldi, B. Hirl, and N. Labanca, "Energy Efficiency Status Report 2012," 2012. [Online]. Available: http://iet.jrc.ec.europa.eu/energyefficiency/sites/energyefficiency/files/energy-efficiency-status-report-2012.pdf. [Accessed: 05-Aug-2014].

[12]   R. Hintemann, "Energy consumption of data centers continues to increase – 2015 update," 2015. [Online]. Available: https://www.borderstep.de/wp-content/uploads/2015/01/Borderstep_Energy_Consumption_2015_Data_Centers_16_12_2015.pdf. [Accessed: 23-Feb-2017].

[13]   Eurostat, "Supply, transformation and consumption of electricity - annual data," 2017. [Online]. Available: http://ec.europa.eu/eurostat/web/energy/data/database. [Accessed: 23-Nov-2017].

[14]   E. Fryer, "The Donald Rumsfeld Guide to Data Centre Energy Consumption," 2016. [Online]. Available: https://www.techuk.org/insights/news/item/8553-the-donald-rumsfeld-guide-to-data-centre-energy-consumption. [Accessed: 15-Jun-2017].

[15]   techUK, "Silver Linings - the implications of Brexit for the UK data centre sector," 2016. [Online]. Available: http://www.techuk.org/images/programmes/DataCentres/Silver_Linings_The_implications_of_BREXITv2.pdf. [Accessed: 11-Dec-2017].

[16]   W. Van Heddeghem, S. Lambert, B. Lannoo, D. Colle, M. Pickavet, and P. Demeester, "Trends in worldwide {ICT} electricity consumption from 2007 to 2012," Comput. Commun., vol. 50, pp. 64–76, 2014.

[17]   International Energy Agency, "Key world energy statistics," 2016. [Online]. Available: https://www.iea.org/publications/freepublications/publication/KeyWorld2016.pdf. [Accessed: 16-Jun-2017].

[18]   International Energy Agency, "World Energy Outlook 2014," 2014.

[19]   Council of the European Union, "Presidency Conclusions - Brussels, 29/30 October 2009," 2009. [Online]. Available: http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/110889.pdf. [Accessed: 16-Nov-2015].

[20]   The White House, "U.S.-China Joint Announcement on Climate Change," 2014. [Online]. Available: https://www.whitehouse.gov/the-press-office/2014/11/11/us-china-joint-announcement-climate-change. [Accessed: 16-Nov-2015].

[21]   L. Newcombe et al., "2012 Best Practices for the EU Code of Conduct on Data Centres," 2012. [Online]. Available: http://re.jrc.ec.europa.eu/energyefficiency/pdf/CoC/Best Practices v3 0 8 _2_ final release Dec 2011.pdf. [Accessed: 12-Sep-2013].

[22]   Environment Agency, "UMBRELLA CLIMATE CHANGE AGREEMENT FOR THE STANDALONE DATA CENTRES," 2014. [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/336160/LIT_9990.pdf. [Accessed: 15-Sep-2014].

[23]   U.S. Department of Energy, "Saving Energy in Data Centers," 2012. [Online]. Available: http://www1.eere.energy.gov/manufacturing/datacenters/. [Accessed: 13-Jan-2014].

[24]   European Commission, "Electricity production, consumption and market overview," 2017. [Online]. Available: http://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_production,_consumption_and_market_overview. [Accessed: 23-Feb-2017].

[25]   L. Erdmann and L. M. Hilty, "Scenario Analysis: Exploring the Macroeconomic Impacts of Information and Communication Technologies on Greenhouse Gas Emissions," J. Ind. Ecol., vol. 14, no. 5, pp. 826–843, 2010.

[26]   The Climate Group, "SMART 2020: Enabling the low carbon economy in the information age," 2008. [Online]. Available: http://www.smart2020.org/_assets/files/02_Smart2020Report.pdf. [Accessed: 04-Nov-2013].

[27]   M. Catulli and E. Fryer, "Information and Communication Technology-Enabled Low Carbon Technologies," J. Ind. Ecol., vol. 16, no. 3, pp. 296–301, 2012.

[28]   techUK, "Climate Change Agreement (CCA) for Data Centres - Target Period One: Report on findings," 2015. [Online]. Available: https://www.techuk.org/images/CCA_First_Target_Report_final.pdf. [Accessed: 04-Mar-2016].

[29]   Uptime Institute, "2013 Data Center Industry Survey," 2013. [Online]. Available: http://uptimeinstitute.com/2013-survey-results. [Accessed: 28-May-2015].

[30]   S. Klingert, T. Schulze, and C. Bunse, "GreenSLAs for the Energy-efficient Management of Data Centres," in Proceedings of the 2Nd International Conference on Energy-Efficient Computing and Networking, 2011, pp. 21–30.

[31]   S. V Patankar, "Airflow and Cooling in a Data Center," J. Heat Transfer, vol. 132, no. 7, p. 73001, Apr. 2010.

[32]   P. M. Curtis, Maintaining Mission Critical Systems in a 24/7 Environment. Hoboken: Wiley, 2007.

[33]   S. Yeo and H.-H. S. Lee, "Peeling the Power Onion of Data Centers," in Energy Efficient Thermal Management of Data Centers, Y. Joshi and P. Kumar, Eds. New York, 2012.

[34]   European Commission, "Code of Conduct on Data Centre Energy Efficiency: Endorser Guidelines and Registration Form Version 3.0.0," 2015. [Online]. Available: http://iet.jrc.ec.europa.eu/energyefficiency/sites/energyefficiency/files/files/documents/ICT_CoC/endorser_guidelines_v3.0.0.pdf. [Accessed: 18-May-2015].

[35]   E. Oró, V. Depoorter, A. Garcia, and J. Salom, "Energy efficiency and renewable energy integration in data centres. Strategies and modelling review," Renew. Sustain. Energy Rev., vol. 42, pp. 429–445, 2015.

[36]   Y. Joshi and P. Kumar, "Fundamentals of Data Center Airflow Management," in Energy efficient thermal management of data centers, Y. Joshi and P. Kumar, Eds. New York: SPRINGER, 2012.

[37]   V. K. Arghode, V. Sundaralingam, Y. Joshi, and W. Phelps, "Thermal Characteristics of Open and Contained Data Center Cold Aisle," J. Heat Transf. - Trans. ASME, vol. 135, no. 6 Special Issue, 2013.

[38]   J. Cho and B. S. Kim, "Evaluation of air management system's thermal performance for superior cooling efficiency in high-density data centers," Energy Build., vol. 43, no. 9, pp. 2145–2155, Sep. 2011.

[39]   K. Choo, R. M. Galante, and M. M. Ohadi, "Energy consumption analysis of a medium-size primary data center in an academic campus," Energy Build., vol. 76, pp. 414–421, 2014.

[40]   D. W. Demetriou and H. E. Khalifa, "Optimization of Enclosed Aisle Data Centers Using Bypass Recirculation," J. Electron. Packag., vol. 134, no. 2, 2012.

[41]   T. J. Breen, E. J. Walsh, J. Punch, A. J. Shah, and C. E. Bash, "From Chip to Cooling Tower Data Center Modeling: Influence of Server Inlet Temperature and Temperature Rise Across Cabinet," J. Electron. Packag., vol. 133, no. 1, Mar. 2012.

[42]   R. Schmidt, M. Iyengar, and J. Caricari, "Data Center Housing High Performance Supercomputer Cluster: Above Floor Thermal Measurements Compared To CFD Analysis," J. Electron. Packag., vol. 132, no. 2, p. 21009, 2010.

[43]   S. Alkharabsheh, B. Sammakia, S. Shrivastava, and R. Schmidt, "Utilizing Practical Fan Curves in CFD Modeling of a Data Center," in 29th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), 2013, pp. 211–215.

[44]   S. Bhopte, M. Iyengar, R. Schmidt, and B. Sammakia, "Numerical and Experimental Study of the Effect of Underfloor Blockages on Data Center Performance," J. Electron. Packag., vol. 133, no. 1, p. 11007, Mar. 2011.

[45]   S. K. Shrivastava, M. Iyengar, B. G. Sammakia, R. Schmidt, and J. W. Vangilder, "Experimental-Numerical Comparison for a High-Density Data Center : Hot Spot Heat Fluxes in Excess of 500 W / ft 2," IEEE Trans. COMPONENTS Packag. Technol., vol. 32, no. 1, pp. 166–172, 2009.

[46]   J. Cho, T. Lim, and B. S. Kim, "Measurements and predictions of the air distribution systems in high compute density (Internet) data centers," Energy Build., vol. 41, no. 10, pp. 1107–1115, Oct. 2009.

[47]   The Green Grid, "Data Center Efficiency And IT Equipment Reliability at Wider Operating Temperature and Humidity Ranges," 2012. [Online]. Available: https://www.thegreengrid.org/~/media/WhitePapers/WP50-Data Center Efficiency and IT Equipment Reliability at Wider Operating Temperature and Humidity Ranges.pdf?lang=en.

[48]   ASHRAE T.C. 9.9, "2011 Thermal Guidelines for Data Processing Environments – Expanded Data Center Classes and Usage Guidance," Atlanta, 2011.

[49]   The Green Grid, "Updated Air-Side Free Cooling Maps: The Impact of ASHRAE 2011 Allowable Ranges," 2012. [Online]. Available: http://www.thegreengrid.org/~/media/WhitePapers/WP46UpdatedAirsideFreeCoolingMa psTheImpactofASHRAE2011AllowableRanges.pdf?lang=en. [Accessed: 27-Mar-2014].

[50]   M. K. Patterson, "The Effect of Data Center Temperature on Energy Efficiency," 2008. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04544393. [Accessed: 11-Mar-2014].

[51]   M. Zapater et al., "Leakage-Aware Cooling Management for Improving Server Energy Efficiency," IEEE Trans. Parallel Distrib. Syst., vol. 26, no. 10, pp. 2764–2777, 2015.

[52]   M. Iyengar, R. Schmidt, V. Kamath, and P. Singh, "Energy efficient economizer based data centers with air cooled servers," in 13th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2012.

[53]   S. Strutt, "DATA CENTER EFFICIENCY AND IT EQUIPMENT RELIABILITY AT WIDER OPERATING TEMPERATURE AND HUMIDITY RANGES," 2012. [Online]. Available: https://www.thegreengrid.org/resources/library-and-tools/383-Data-Center-Efficiency-and-IT-Equipment-Reliability-at-Wider-Operating-Temperature-and-Humidity-Ranges. [Accessed: 03-Jan-2017].

[54]   D. Moss and J. H. Bean, "Energy Impact of Increased Server Inlet Temperature," 2011. [Online]. Available: http://www.apc.com/salestools/JBEN-7KTR88/JBEN-7KTR88_R1_EN.pdf. [Accessed: 05-Jun-2017].

[55]   D. W. Demetriou and H. E. Khalifa, "Energy Modeling of Air-Cooled Data Centers: Part II—The Effect of Recirculation on the Energy Optimization of Open-Aisle, Air-Cooled Data Centers," in ASME 2011 Pacific Rim Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Systems, 2011.

[56]   L. A. Barroso and U. Holzle, "The Case for Energy-Proportional Computing," Computer (Long. Beach. Calif)., vol. 40, no. 12, p. 33, 2007.

[57]   R. Talaber, T. Brey, and L. Lamers, "USING VIRTUALIZATION TO IMPROVE DATA CENTER EFFICIENCY," 2009. [Online]. Available: http://www.thegreengrid.org/~/media/WhitePapers/White Paper 19 - Using Virtualization to Improve Data Center Efficiency.pdf?lang=en. [Accessed: 05-Aug-2014].

[58]   M. Tighe, G. Keller, M. Bauer, and H. Lutfiyya, "DCSim: A Data Centre Simulation Tool for Evaluating Dynamic Virtualized Resource Management," in 8th International Conference on Network and service management, 2012.

[59]   G. Schulz, The Green and Virtual Data Center. New York: Auerbach, 2009.

[60]   J. G. Koomey, C. Belady, M. Patterson, A. Santos, and K.-D. Lange, "ASSESSING TRENDS OVER TIME IN PERFORMANCE, COSTS, AND ENERGY USE FOR SERVERS," 2009.

[61]   B. Degnan, B. Marr, and J. Hasler, "Assessing Trends in Performance per Watt for Signal Processing Applications," IEEE Trans. Very Large Scale Integr. Syst., vol. 24, no. 1, pp. 58–66, 2016.

[62]   M. K. Patterson, "Energy Efficiency Metrics," in Energy efficient thermal management of data centers, Y. Joshi and P. Kumar, Eds. New York: Springer, 2012, p. 241.

[63]   The Green Grid, "A Framework For Data Center Energy Productivity," 2008. [Online]. Available: http://www.greenbiz.com/sites/default/files/document/GreenGrid-Framework-Data-Center-Energy-Productivity.pdf. [Accessed: 31-Jul-2013].

[64]   The Green Grid, "PUE: A Comprehensive Examination of the Metric," 2012. [Online]. Available: http://www.thegreengrid.org/~/media/WhitePapers/WP49-PUE   A Comprehensive Examination of the Metric_v6.pdf?lang=en. [Accessed: 01-Aug-2013].

[65]   British Standards Institution, BS ISO/IEC 30134-2:2016. Information technology — Data centres — Key performance indicators. Part 2: Power usage effectiveness (PUE). London: BSI, 2016.

[66]   A. McNevin, "APAC data center survey reveals high PUE figures across the region," 2013. [Online]. Available: http://www.datacenterdynamics.com/news/apac-data-center-survey-reveals-high-pue-figures-across-the-region/75116.fullarticle.   [Accessed: 16-Jun-2017].

[67]   Facebook, "Prineville, OR Data Center," 2017. [Online]. Available: https://www.fbpuewue.com/prineville. [Accessed: 16-Jun-2017].

[68]   Google, "Efficiency: How we do it," 2017. [Online]. Available: https://www.google.com/about/datacenters/efficiency/internal/. [Accessed: 16-Jun-2017].

[69]   M. K. Herrlin, "Airflow and Cooling Performance of Data Centers: Two Performance Metrics," ASHRAE Trans., vol. 114, no. 2, pp. 182–187, 2008.

[70]   McKinsey & Company, "Revolutionizing Data Center Energy Efficiency," 2008. [Online].   Available: http://www.ecobaun.com/images/Revolutionizing_Data_Center_Efficiency.pdf. [Accessed: 31-Jul-2013].

[71]   The Green Grid, "Proxy Proposals For Measuring Data Center Productivity," 2009. [Online]. Available: http://www.thegreengrid.org/~/media/WhitePapers/White Paper 17 -

Proxies Proposals for Measuring Data Center Efficiencyv2.pdf?lang=en. [Accessed: 31-Jul-2013].

[72]   S. A. Alkharabsheh, B. Muralidharan, M. Ibrahim, S. K. Shrivastava, and B. G. Sammakia, "Open and Contained Cold Aisle Experimentally Validated CFD Model Implementing CRAC and Server Fan Curves for a Data Center Test Laboratory," in ASME 2013 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, 2013.

[73]   Emerson Network Power, "Optimized energy efficiency with controlled cold aisle containments," 2014. [Online]. Available: http://www.emersonnetworkpower.com/en-EMEA/Latest-Thinking/white-papers/Documents/Optimized-energy-efficiency-with-controlled-cold-aisle-containments.pdf. [Accessed: 14-Aug-2014].

[74]   A. Honeybill, "Air Segregation: How to Measure Effectiveness," 2015. [Online]. Available: http://www.upsite.com/blog/air-segregation-how-to-measure-effectiveness/. [Accessed: 04-Sep-2015].

[75]   D. Kennedy, "Server leakage and cooling," Datacenter Dyn. Focus, vol. 3, no. 29, pp. 50–52, 2013.

[76]   Airedale,   "LogiCool   InRak,"   2012.   [Online].   Available: http://airedalecoolingsolutions.co.uk/updates/downloads/TM_INRAK_UK.pdf.   [Accessed: 04-Sep-2015].

[77]   V. Sundaralingam, V. K. Arghode, Y. Joshi, and W. Phelps, "Experimental Characterization of Various Cold Aisle Containment Configurations for Data Centers," J. Electron. Packag., vol. 137, no. 1, Mar. 2015.

[78]   M. Acton, P. Bertoldi, J. Booth, L. Newcombe, and A. Rouyer, "2017 Best Practice Guidelines for the EU Code of Conduct on Data Centre Energy Efficiency," 2017. [Online]. Available: https://e3p.jrc.ec.europa.eu/publications/2017-best-practice-guidelines-eu-code-conduct-data-centre-energy-efficiency. [Accessed: 25-Apr-2017].

[79]   L. Newcombe et al., "2014 Best Practices for the EU Code of Conduct on Data Centres,"   2014.   [Online].   Available: http://www.bcs.org/upload/pdf/2014_best_practice_guidelines_v5.1.0.pdf. [Accessed: 19-Jun-2017].

[80]   V. K. Arghode and Y. Joshi, "Room Level Modeling of Air Flow in a Contained Data Center Aisle," J. Electron. Packag., vol. 136, no. 1, p. 11011, Feb. 2014.

[81]   R. Schmidt, A. Vallury, and M. Iyengar, "Energy Savings Through Hot and Cold Aisle Containment Configurations for Air Cooled Servers in Data Centers," in ASME 2011 Pacific Rim Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Systems, 2011, pp. 611–616.

[82]   S. K. Shrivastava, A. R. Calder, and M. Ibrahim, "Quantitative Comparison of Air Containment Systems," in Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2012 13th IEEE Intersociety Conference on, 2012, pp. 68–77.

[83]   H. A. Alissa, K. Nemati, B. G. Sammakia, K. Schneebeli, R. R. Schmidt, and M. J. Seymour, "Chip to Facility Ramifications of Containment Solution on IT Airflow and Uptime," IEEE Trans. COMPONENTS Packag. Manuf. Technol., vol. PP, no. 99, pp. 1–12, 2016.

[84]   S. A. Nada, M. A. Rady, M. Elsharnoby, and M. A. Said, "Numerical Investigation of Cooling of Electronic Servers Racks at Different Locations and Spacing from the Data Center Cooling Unit," Int. J. Curr. Eng. Technol., vol. 5, no. 5, pp. 3448–3456, 2015.

[85]   S. A. Nada, M. A. Said, and M. A. Rady, "CFD investigations of data centers' thermal performance for different configurations of CRACs units and aisles separation," Alexandria Eng. J., vol. 55, no. 2, pp. 959–971, 2016.

[86]   S.-W. Ham and J.-W. Jeong, "Impact of aisle containment on energy performance of a data center when using an integrated water-side economizer," Appl. Therm. Eng., vol. 105, pp. 372–384, 2016.

[87]   M. Sahini, E. Kumar, T. Gao, C. Ingalz, A. Heydari, and S. Xiaogang, "Study of Air Flow Energy within Data Center room and sizing of hot aisle Containment for an Active vs Passive cooling design," in 15th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2016, pp. 1453–1457.

[88]   L. Silva-Llanca, M. del Valle, and A. Ortega, "THE EFFECTIVENESS OF DATA CENTER OVERHEAD COOLING IN STEADY AND TRANSIENT SCENARIOS: COMPARISON OF DOWNWARD FLOW TO A COLD AISLE VERSUS UPWARD FLOW FROM A HOT AISLE," in Proceedings of the ASME 2015 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems (InterPACK2015), 2015, pp. 1–10.

[89]   Z. Song, "Numerical investigation for performance indices and categorical designs of a fan-assisted data center cooling system," Appl. Therm. Eng., vol. 118, pp. 714–723, 2017.

[90]   J. R. H. Schaadt, K. Fouladi, P. Aaron Wemhoff, and J. G. Pigeon, "LOAD CAPACITY AND THERMAL EFFICIENCY OPTIMIZATION OF A RESEARCH DATA CENTER USING COMPUTATIONAL MODELING," in ASME 2015 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems (InterPACK2015), 2015, pp. 1–6.

[91]   Y. Fulpagare, Y. Joshi, and A. Bhargav, "TRANSIENT CHARACTERIZATION OF DATA CENTER RACKS," in ASME 2016 International Mechanical Engineering Congress and Exposition (IMECE2016), 2016, pp. 1–5.

[92]   K. Fouladi, J. Schaadt, and A. P. Wemhoff, "A novel approach to the data center hybrid cooling design with containment," Numer. Heat Transf. Part A Appl., vol. 71, no. 5, pp. 477–487, 2017.

[93]   M. I. Tradat et al., "Impact of elevated temperature on data center operation based on internal and external IT instrumentation," in 33rd Thermal Measurement, Modeling Management Symposium (SEMI-THERM), 2017, pp. 108–114.

[94]   S. A. Alkharabsheh, B. G. Sammakia, and S. K. Shrivastava, "Experimentally Validated Computational Fluid Dynamics Model for a Data Center With Cold Aisle Containment," J. Electron. Packag., vol. 137, no. 2, Jun. 2015.

[95]   G. AlLee, "Green Microprocessor and Server Design," in Data Center Handbook, H. Geng, Ed. Hoboken, New Jersey: John Wiley & Sons, Ltd., 2015.

[96]   Y. U. Makwana, A. R. Calder, and S. K. Shrivastava, "Benefits of Properly Sealing a Cold Aisle Containment System," in 14th IEEE ITHERM Conference, 2014, pp. 793–798.

[97]   S. Microsystems, "SunFire V20z Server," 2003. [Online]. Available: http://www.grycap.upv.es/img/repos/Sun Fire V20z.pdf. [Accessed: 07-Aug-2015].

[98]   R. A. Granger, "Laminar Pipe Flow," in Fluid Mechanics, New York, NY, USA: Dover Publications, 1995.

[99]   E. J. Shaughnessy, "Flow In Pipes And Ducts," in Introduction to Fluid Mechanics, Oxford, UK: Oxford University Press, 2005.

[100]   B. E. Poling, G. H. Thomson, D. G. Friend, R. L. Rowley, and W. V. Wilding, "Physical and Chemical Data," in Perry's Chemical Engineers' Handbook2, 8th ed., D. W. Green and R. H. Perry, Eds. London: McGraw-Hill, 2008.

[101]   O. C. Jones Jr, "An Improvement in the Calculation of Turbulent Friction in Rectangular Ducts," J. Fluids Eng., vol. 98, no. 2, pp. 173–180, Jun. 1976.

[102]   R. Miller, "Emerson: 10 Common Data Center Surprises," 2012. [Online]. Available: http://www.datacenterknowledge.com/archives/2012/03/21/emerson-10-common-data-center-surprises/. [Accessed: 13-Aug-2014].

[103]   ebm-papst, "EC Centrifugal Blower." [Online]. Available: http://docs-europe.electrocomponents.com/webdocs/0e1c/0900766b80e1c609.pdf. [Accessed: 23-Aug-2014].

[104]   Vent-Axia, "ACP315HP," 2017. [Online]. Available: http://www.vent-axia.com/product/acp-line-duct-fans-acp315hp.html. [Accessed: 07-Jan-2017].

[105]   Omega Engineering Limited, "HHF2005HW," 2014. [Online]. Available: http://www.omega.co.uk/pptst/HHF2005HW.html. [Accessed: 23-Aug-2014].

[106]   TSI, "Anemometer TA2 Range," 2014. [Online]. Available: http://www.tsi.com/airflow-instruments-thermal-anemometer-ta2-range/. [Accessed: 12-Aug-2014].

[107]   British Standard, "BS ISO 3966:2008 Measurement of fluid flow in closed conduits - Velocity area method using Pitot static tubes." 2008.

[108]   KIMO Instruments, "MG series Inclined Liquid Column Manometers," 2010. [Online]. Available: http://kimouk.com/datasheets-manuals/manometers/Manometers/FT-MG.pdf. [Accessed: 11-Aug-2014].

[109]   KIMO Instruments, "HP Series Inclined Liquid Column Manometers," 2010. [Online]. Available: http://kimouk.com/datasheets-manuals/manometers/Manometers/FT-HP.pdf. [Accessed: 11-Aug-2014].

[110]   Digitron, "2000 Series Technical Data Sheet." [Online]. Available: http://docs-europe.electrocomponents.com/webdocs/0785/0900766b8078507e.pdf. [Accessed: 23-Aug-2014].

[111]   PlenaForm Systems, "PlenaFill 27 U Blanking Panels," 2017. [Online]. Available: http://www.plenaform.com/plenafill-103. [Accessed: 22-May-2017].

[112]   Oracle, "Sun Fire V20z and Sun Fire V40z Servers--User Guide," 2008. [Online]. Available: https://docs.oracle.com/cd/E19121-01/sf.v40z/817-5248-21/chapter1.html. [Accessed: 02-May-2016].

[113]   CNet, "Sun Fire V40z - Dual-Core Opteron 875 2.2 GHz." [Online]. Available: http://www.cnet.com/products/sun-fire-v40z-dual-core-opteron-875-2-2-ghz-monitor-none-series/specs/. [Accessed: 02-May-2016].

[114]   StressLinux, "Welcome to stresslinux," 2015. [Online]. Available: http://www.stresslinux.org/sl/. [Accessed: 02-May-2016].

[115]   Hillstone Loadbanks, "Using the 6RM Server Simulator in IT Cabinets," 2015. [Online]. Available: http://hillstone.co.uk/Download/HAC230-6RM.pdf. [Accessed: 06-Jul-2016].

[116]   S. J. Morrison, "Production," in Statistics for Engineers - an Introduction, Chichester, United Kingdom: John Wiley & Sons, Ltd., 2009.

[117]   R. A. Granger, "Turbulent Pipe Flow," in Fluid Mechanics, New York, NY, USA: Dover Publications, 1995.

[118]   K. Dunlap and N. Rasmussen, "Choosing Between Room, Row, and Rack-based Cooling for Data Centers," 2012. [Online]. Available: http://iisgroupllc.com/wp-content/uploads/2013/02/APC-White-Paper-130-Choosing-Between-Room-Row-and-Rack-based-Cooling-for-Data-Centers.pdf. [Accessed: 04-Jul-2017].

[119]   A.-T. Nguyen, S. Reiter, and P. Rigo, "A review on simulation-based optimization methods applied to building performance analysis," Appl. Energy, vol. 113, pp. 1043–1058, 2014.

[120]   S. Pelley, D. Meisner, T. F. Wenisch, and J. W. VanGilder, "Understanding and abstracting total data center power," in Workshop on Energy-Efficient Design, 2009.

[121]   T. J. Breen et al., "From Chip to Cooling Tower Data Center Modeling: Influence of Air-Stream Containment on Operating Efficiency," J. Electron. Packag., vol. 134, no. 4, 2012.

[122]   A. P. Wemhoff, M. del Valle, K. Abbasi, and A. Ortega, "THERMODYNAMIC MODELING OF DATA CENTER COOLING SYSTEMS," in Proceedings of the ASME 2013 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems (InterPACK2013), 2013.

[123]   C. D. Patel, R. K. Sharma, C. E. Bash, and M. Beitelmal, "Energy Flow in the Information Technology Stack: Coefficient of Performance of the Ensemble and its Impact on the Total Cost of Ownership," 2006. [Online]. Available: http://www.hpl.hp.com/techreports/2006/HPL-2006-55.pdf. [Accessed: 11-Aug-2015].

[124] A. Heydari, "Thermodynamics Energy Efficiency Analysis and Thermal Modeling of Data Center Cooling Using Open and Closed-Loop Cooling Systems," 2007.

[125] M. Iyengar and R. Schmidt, "Analytical Modeling for Thermodynamic Characterization of Data Center Cooling Systems," J. Electron. Packag., vol. 131, no. 2, p. 21009, Apr. 2009.

[126] G. A. Brady, "Energy Efficiency in Data Centres and the Barriers to Further Improvements: An Interdisciplinary Investigation,", p200, University of Leeds, 2016.

[127] D. Kennedy, "Ramification of Server Airflow Leakage in Data Centers with Aisle Containment," 2012. [Online]. Available: https://www.tateinc.com/sites/default/files/support-docs/tate-ramificationleakageaislecontainment.pdf. [Accessed: 14-Aug-2014].

[128] M. Tatchell-Evans, D. Burdett, J. Summers, A. Beaumont, and G. Fox, "An Experimental and Theoretical Investigation of the Effects of Supply Air Conditions on Computational Efficiency in Data Centers Employing Aisle Containment," in SEMI-THERM 32, 2017.

[129] T. Wilkie, "High-Performance Computing - Cool ways to save energy," 2015. [Online]. Available: http://www.scientific-computing.com/features/feature.php?feature_id=460. [Accessed: 17-Feb-2016].

[130] A. Venkatraman, "Enterprises, scientists and academics turn to public cloud for HPC," 2014. [Online]. Available: http://www.computerweekly.com/feature/Enterprises-scientists-and-academics-turn-to-public-cloud-for-HPC. [Accessed: 17-Feb-2016].

[131] L. Newcombe, "Data centre energy efficiency metrics," 2010. [Online]. Available: http://bcs.org/upload/pdf/data-centre-energy.pdf. [Accessed: 04-Sep-2015].

[132] B. Heller et al., "ElasticTree: saving energy in data center networks," in NSDI'10 Proceedings of the 7th USENIX conference on Networked systems design and implementation, 2010, p. 17.

[133] H. Shirayanagi, H. Yamada, and K. Kono, "Honeyguide: A VM Migration-Aware Network Topology for Saving Energy Consumption in Data Center Networks," in 2012 IEEE SYMPOSIUM ON COMPUTERS AND COMMUNICATIONS (ISCC), 2012, pp. 460–467.

[134] MathWorks, "MATLAB," 2017. [Online]. Available: https://www.mathworks.com/products/matlab.html. [Accessed: 15-May-2017].

[135]   The Met Office, "Mean temperature - Annual average: 1971-2000." [Online]. Available: http://www.metoffice.gov.uk/public/weather/climate/gcpuckhb6#?region=southernengland. [Accessed: 24-Aug-2015].

[136]   G. A. Brady, "Energy Efficiency in Data Centres and the Barriers to Further Improvements: An Interdisciplinary Investigation,", p217, The University of Leeds, 2016.

[137]   G. A. Brady, "Energy Efficiency in Data Centres and the Barriers to Further Improvements: An Interdisciplinary Investigation,", p200, University of Leeds, 2016.

[138]   G. A. Brady, "Energy Efficiency in Data Centres and the Barriers to Further Improvements: An Interdisciplinary Investigation,", p206, University of Leeds, 2016.

[139]   Schneider Electric, "Uniflair LE Chilled Water Air Conditioners." [Online]. Available: http://www.schneider-electric.com/ww/en/download/document/APC_KKRZ-8LAPNS_R6_EN_SRC. [Accessed: 07-Sep-2015].

[140]   F. P. Incropera and D. P. Dewitt, "Heat Exchangers," in Fundamentals of Heat and Mass Transfer2, 5th ed., New York: John Wiley & Sons, Ltd., 2002.

[141]   C.-C. Wang and Y.-J. Chang, "Sensible heat and friction characteristics of plate fin-and-tube heat exchangers having plane fins," Int. J. Refrig., vol. 19, no. 4, pp. 223–230, 1996.

[142]   T. RABAS, P. ECKELS, and R. SABATINO, "THE EFFECT OF FIN DENSITY ON THE HEAT-TRANSFER AND PRESSURE-DROP PERFORMANCE OF LOW-FINNED TUBE BANKS," Chem. Eng. Commun., vol. 10, no. 1–3, pp. 127–147, 1981.

[143]   R. H. S. Winterton, "Where did the Dittus and Boelter equation come from?," Int. J. Heat Mass Transf., vol. 41, no. 4, pp. 809–810, 1998.

[144]   Airedale, "DeltaChill Air Cooled," 2014. [Online]. Available: http://www.airedale.com/web/Products/Chillers/DeltaChill-DeltaChill-FreeCool-110kW-1080kW.htm. [Accessed: 08-Sep-2015].

[145]   M. M. Toulouse, G. Doljac, V. P. Carey, and C. Bash, "Exploration of a Potential-Flow-Based Compact Model of Air-Flow Transport in Data Centers," in ASME 2009 International Mechanical Engineering Congress and Exposition, 2009.

[146]   H. F. Hamann and V. Lopez, "Data Center Metrology and Measurement-Based Modeling Methods," in Energy efficient thermal management of data centers, Y. Joshi and P. Kumar, Eds. New York: Springer, 2012, p. 273.

[147]   Y. Hung, P. Z. G. Qian, and C. F. J. Wu, "Statistical Methods for Data Center Thermal Management," in Energy efficient thermal management of data centers, London: SPRINGER, 2012.

[148]   J. Lighthill, "Velocity Fields and Pressure Fields," in An Informal Introduction to Theoretical Fluid Mechanics, New York, NY, USA: Oxford University Press, 1986.

[149]   J. Tu, "CFD Solution Procedure - A Beginning," in Computational Fluid Dynamics - A Practical Approach, 2nd ed., Oxford, UK: Butterworth-Heinemann, 2013.

[150]   V. Lopez and H. F. Hamann, "Measurement-based modeling for data centers," in 12th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2010.

[151]   H. F. Hamann, V. Lopez, and A. Stepanchuk, "Thermal zones for more efficient data center energy management," in Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2010 12th IEEE Intersociety Conference on, 2010, pp. 1–6.

[152]   C. M. Healey, J. W. VanGilder, Z. R. Sheffer, and X. S. Zhang, "Potential-Flow Modeling for Data Center Applications," in ASME 2011 Pacific Rim Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Systems2, 2011.

[153]   G. Falkovich, "Basic equations and steady flows," in Fluid Mechanics, Cambridge University Press, 2011.

[154]   J. Tu, "CFD Techniques - The Basics," in Computational Fluid Dynamics - A Practical Approach, 2nd ed., Oxford, UK: Butterworth-Heinemann, 2013.

[155]   J. Tu, "Governing Equations for CFD - Fundamentals," in Computational Fluid Dynamics - A Practical Approach, Oxford, UK: Butterworth-Heinemann, 2013.

[156]   P. S. Ghoshdastidar, "Introduction," in Computer Simulation of Flow and Heat Transfer, London: Tat McGraw-Hill, 1998.

[157]   J. Lighthill, "Three-dimensional examples of irrotational flows," in An Informal Introduction to Theoretical Fluid Mechanics, New York, NY, USA: Oxford University Press, 1986.

[158]   J. Lighthill, "Vortex dynamics," in An Informal Introduction to Theoretical Fluid Mechanics, New York, NY, USA: Oxford University Press, 1986.

[159]   J. Lighthill, "Principles of mechanics applied to lumps of fluid," in An Informal Introduction to Theoretical Fluid Mechanics, New York, NY, USA: Oxford University Press, 1986.

[160]   B. Sammakia, S. Bhopte, and M. Ibrahim, "Numerical Modeling of Data Center Clusters," in Energy efficient thermal management of data centers, Y. Joshi and P. Kumar, Eds. Springer Verlag, 2012.

[161]   E. Divo and A. J. Kassab, "Localized Meshless Modeling of Natural-Convective Viscous Flows," Numer. Heat Transf. Part B Fundam., vol. 53, no. 6, pp. 487–509, 2008.

[162]   J. D. Anderson Jr., "Governing Equations of Fluid Dynamics," in Computational Fluid Dynamics: An Introduction, 3rd ed., J. Wendt, Ed. Berlin: Springer Verlag, 2009, pp. 15–52.

[163]   J. Law and R. Rennie, Eds., "'turbulence,'" in A Dictionary of Physics, Oxford University Press, 2015.

[164]   J. Tu, "Practical Guidelines for CFD Simulations and Analysis," in Computational Fluid Dynamics - A Practical Approach, 2nd ed., Amsterdam: Elsevier/Butterworth-Heinemann, 2013.

[165]   E. Cruz, Y. Joshi, M. Iyengar, and R. Schmidt, "COMPARISON OF NUMERICAL MODELING TO EXPERIMENTAL DATA IN A SMALL DATA CENTER TEST CELL," in InterPACK2009-89306, 2009, pp. 1–9.

[166]   N. C. Markatos, "The mathematical modelling of turbulent flows," Appl. Math. Model., vol. 10, no. 3, pp. 190–220, 1986.

[167]   P. R. Spalart and S. R. Allmaras, "A one equation turbulence model for aerodynamic flows," Rech. Aerosp. Ed., vol. 5, no. 5, 1994.

[168]   J. Tu, "CFD Solution Analysis - Essentials," in Computational Fluid Dynamics - A Practical Approach, 2nd ed., Oxford, UK: Butterworth-Heinemann, 2013.

[169]   M. Seymour, "Computational Fluid Dynamics Applications in Data Centers," in Data Center Handbook, H. Geng, Ed. Hoboken, New Jersey: John Wiley & Sons, Ltd., 2015, p. 316.

[170]   D. J. Lettieri, M. M. Toulouse, C. E. Bash, A. J. Shah, and V. P. Carey, "Computational and Experimental Validation of a Vortex-Superposition-Based Buoyancy Approximation for the COMPACT Code in Data Centers," J. Electron. Packag., vol. 135, no. 3, pp. 30903-1-30903–8, 2013.

[171]   H. A. Alissa, K. Nemati, B. Sammakia, M. Seymour, K. Schneebli, and R. Schmidt, "EXPERIMENTAL AND NUMERICAL CHARACTERIZATION OF A RAISED FLOOR DATA CENTER USING RAPID OPERATIONAL FLOW CURVES MODEL," in Proceedings of the ASME 2015 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems (InterPACK2015), 2015.

[172]   J. Siriwardana, S. K. Halgamuge, T. Scherer, and W. Schott, "Minimizing the thermal impact of computing equipment upgrades in data centers," Energy Build., vol. 50, pp. 81–92, Jul. 2012.

[173]   E. Samadiani, J. Rambo, and Y. Joshi, "Numerical Modeling of Perforated Tile Flow Distribution in a Raised-Floor Data Center," J. Electron. Packag., vol. 132, no. 2, 2010.

[174]   G. N. de Boer, A. Johns, N. Delbosc, D. Burdett, M. Tatchell-Evans, J. L. Summers, and R. Baudot, "Three computational methods for analysing thermal airflow distributions in the cooling of data centers," Int. J. Numer. Methods Heat Fluid Flow [in press].

[175]   E. Cruz and Y. Joshi, "Coupled inviscid-viscous solution method for bounded domains: Application to data-center thermal management," Int. J. Heat Mass Transf., vol. 85, pp. 181–194, 2015.

[176]   S. Gondipalli, S. Bhopte, B. Sammakia, M. K. Iyengar, and R. Schmidt, "Effect of isolating cold aisles on rack inlet temperature," in 11th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, 2008, pp. 1247–1254.

[177]   ASHRAE T.C. 9.9, "IT Equipment Thermal Management and Controls," 2012. [Online]. Available: http://tc99.ashraetcs.org/documents/ASHRAE 2012 IT Equipment Thermal Management and Controls_V1.0.pdf. [Accessed: 05-Feb-2015].

[178]   Future Facilities, "6SigmaRoom R8," 2013. [Online]. Available: http://www.air-think.com.tw/uploads/6/1/7/6/6176073/6sigmaroom-whatsnew-r8.pdf. [Accessed: 14-Feb-2017].

[179]   ASHRAE, Thermal Guidelines for Data Processing Environments. Atlanta: ASHRAE, 2012.

[180]   S. Sorrell et al., "Barriers to Energy Efficiency in Public and Private Organisations," 2000. [Online]. Available: http://www.sussex.ac.uk/Units/spru/publications/reports/barriers/final.html. [Accessed: 14-Oct-2013].

[181]   P. Thollander and M. Ottosson, "An energy efficient Swedish pulp and paper industry - exploring barriers to and driving forces for cost-effective energy efficiency investments," ENERGY Effic., vol. 1, no. 1, pp. 21–34, Feb. 2008.

[182]   E. Cagno and A. Trianni, "Exploring drivers for energy efficiency within small- and medium-sized enterprises: First evidences from Italian manufacturing enterprises," Appl. Energy, vol. 104, no. 0, pp. 276–285, 2013.

[183]   A. B. Jaffe and R. N. Stavins, "The energy paradox and the diffusion of conservation technology," Resour. Energy Econ., vol. 16, no. 2, pp. 91–122, May 1994.

[184]   M. Mazzucato, "Innovation systems: from fixing market failures to creating markets," Intereconomics, vol. 50, no. 3, pp. 120–155, 2015.

[185]   Department of Energy & Climate Change, "Energy Efficiency Strategy 2013 Update," 2013.   [Online].   Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/266187/ 2901415_EnergyEfficiencyStrategy_acc.pdf. [Accessed: 12-Sep-2014].

[186] Europa, "Energy Efficiency for the 2020 goal," 2009. [Online]. Available: http://europa.eu/legislation_summaries/energy/energy_efficiency/en0002_en.htm. [Accessed: 12-Sep-2014].

[187]   U.S. Environmental Protection Agency, "About Energy Star," 2014. [Online]. Available: http://www.energystar.gov/about/. [Accessed: 12-Sep-2014].

[188]   K. Gillingham and K. Palmer, "Bridging the Energy Efficiency Gap," 2013. [Online]. Available: http://www.rff.org/RFF/documents/RFF-DP-13-02.pdf. [Accessed: 15-Sep-2014].

[189]   D. L. Greene, "Uncertainty, loss aversion, and markets for energy efficiency," Energy Econ., vol. 33, no. 4, pp. 608–616, Jul. 2011.

[190]   R. J. Sutherland, "'No cost' efforts to reduce carbon emissions in the U.S.: An economic perspective," vol. 21, no. 3, pp. 89–112, 2000.

[191]   J. Harris, J. Anderson, and W. Shafron, "Investment in energy efficiency: a survey of Australian firms," Energy Policy, vol. 28, no. 12, pp. 867–876, Oct. 2000.

[192]   H. C. Granade, J. Creyts, P. Farese, and K. Ostrowski, "Energy Efficiency: Unlocking the   US   opportunity,"   2010.   [Online].   Available: https://www.mckinsey.com/~/media/McKinsey/dotcom/client_service/Sustainability/PDFs /A_Compelling_Global_Resource.ashx. [Accessed: 14-Sep-2014].

[193] European Commission, "Energy Efficiency Directive," 2012. [Online]. Available: http://ec.europa.eu/energy/efficiency/eed/eed_en.htm. [Accessed: 15-Sep-2014].

[194] European Commission, "Energy Efficiency Action Plan 2011," 2011. [Online]. Available: http://ec.europa.eu/energy/efficiency/action_plan/action_plan_en.htm. [Accessed: 14-Sep-2014].

[195] R. Martin, L. B. de Preux, and U. J. Wagner, "THE IMPACTS OF THE CLIMATE CHANGE LEVY ON MANUFACTURING: EVIDENCE FROM MICRODATA," 2011. [Online]. Available: http://www.nber.org/papers/w17446.pdf. [Accessed: 15-Sep-2014].

[196] D. F. Vivanco and E. van der Voet, "The rebound effect through industrial ecology's eyes: a review of LCA-based studies," Int. J. LIFE CYCLE Assess., vol. 19, pp. 1933–1947, 2014.

[197] L. A. Greening, D. L. Greene, and C. Difiglio, "Energy efficiency and consumption — the rebound effect — a survey," Energy Policy, vol. 28, no. 6–7, pp. 389–401, 2000.

[198] H. D. Saunders, "The Khazzoom-Brookes Postulate and Neoclassical Growth," Energy J., vol. 13, no. 4, 1992.

[199] S. Sorrell, "Energy, Economic Growth and Environmental Sustainability: Five Propositions," Sustainability, vol. 2, no. 6, p. 1784, 2010.

[200] K. K. Nguyen, M. Cheriet, M. Lemay, V. Reijs, A. Mackarel, and A. Pastrama, "Environmental-aware virtual data center network," Comput. Networks, vol. 56, no. 10, pp. 2538–2550, 2012.

[201] S. Figuerola, M. Lemay, V. Reijs, M. Savoie, and B. St.Arnaud, "Converged Optical Network Infrastructures in Support of Future Internet and Grid Services Using IaaS to Reduce GHG Emissions," J. Light. Technol., vol. 27, no. 12, pp. 1941–1946, 2009.

[202] F. R. Pazheri, M. F. Othman, and N. H. Malik, "A review on global renewable electricity scenario," Renew. Sustain. Energy Rev., vol. 31, pp. 835–845, 2014.

[203] P. Frankl, "World Renewable Energy Outlook 2030-2050," 2013. [Online]. Available: https://www.celluleenergie.cnrs.fr/IMG/pdf/intro_i3_paolo_frankl.pdf. [Accessed: 23-Feb-2017].

[204] Vlad Coroama and L. M. Hilty, "Energy Consumed vs. Energy Saved by ICT – A Closer Look," in Environmental Informatics and Industrial Environmental Protection: Concepts, Methods and Tools, 2009, pp. 347–355.

[205]   J. Freire-González and I. Puig-Ventosa, "Energy Efficiency Policies and the Jevons Paradox," Int. J. Energy Econ. Policy, vol. 5, no. 1, pp. 69–79, 2015.

[206]   C. Walsh and P. Thornley, "Barriers to improving energy efficiency within the process industries with a focus on low grade heat utilisation," J. Clean. Prod., vol. 23, no. 1, pp. 138–146, 2012.

[207]   K.-H. Chai and C. Yeo, "Overcoming energy efficiency barriers through systems approach—A conceptual framework," Energy Policy, vol. 46, no. 0, pp. 460–472, 2012.

[208]   A. Hasanbeigi, C. Menke, and P. du Pont, "Barriers to energy efficiency improvement and decision-making behavior in Thai industry," Energy Effic., vol. 3, no. 1, pp. 33–52, 2009.

[209]   P. Rohdin, P. Thollander, and P. Solding, "Barriers to and drivers for energy efficiency in the Swedish foundry industry," Energy Policy, vol. 35, no. 1, pp. 672–677, 2007.

[210]   A. Trianni and E. Cagno, "Dealing with barriers to energy efficiency and SMEs: Some empirical evidences," Energy, vol. 37, no. 1, pp. 494–504, 2012.

[211]   K. Kounetas, D. Skuras, and K. Tsekouras, "Promoting energy efficiency policies over the information barrier," Inf. Econ. Policy, vol. 23, no. 1, pp. 72–84, 2011.

[212]   P. Thollander, S. Backlund, A. Trianni, and E. Cagno, "Beyond barriers – A case study on driving forces for improved energy efficiency in the foundry industries in Finland, France, Germany, Italy, Poland, Spain, and Sweden," Appl. Energy, vol. 111, no. 0, pp. 636–643, 2013.

[213]   K. Tanaka, "Review of policies and measures for energy efficiency in industry sector," Energy Policy, vol. 39, no. 10, pp. 6532–6550, Oct. 2011.

[214]   DECC, "2010 to 2015 government policy: energy demand reduction in industry, business and the public sector," 2013. [Online]. Available: https://www.gov.uk/government/policies/reducing-demand-for-energy-from-industry-businesses-and-the-public-sector--2/suppohttps://www.gov.uk/government/publications/2010-to-2015-government-policy-energy-demand-reduction-in-industry-business-and-the-public-. [Accessed: 05-Nov-2013].

[215]   DEFRA, "Climate change agreements: the Climate Change Levy," 2008. [Online]. Available: http://webarchive.nationalarchives.gov.uk/20130123162956/http:/www.defra.gov.uk/environment/climatechange/uk/business/cca/levy.htm. [Accessed: 05-Nov-2013].

[216]   HM Revenue & Customs, "Excise Notice CCL1: a general guide to Climate Change Levy," 2015. [Online]. Available: https://www.gov.uk/government/publications/excise-notice-ccl1-a-general-guide-to-climate-change-levy/excise-notice-ccl1-a-general-guide-to-climate-change-levy. [Accessed: 12-Jan-2016].

[217]   DEFRA, "What are climate change agreements?," 2008. [Online]. Available: http://webarchive.nationalarchives.gov.uk/20130123162956/http://www.defra.gov.uk/environment/climatechange/uk/business/cca/agreements.htm. [Accessed: 05-Nov-2013].

[218]   Environment Agency, "Climate Change Agreements: biennial progress report 2013 and 2014," 2015. [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/479679/Biennial_progress_report_2013_and_2014.pdf. [Accessed: 12-Jan-2016].

[219]   P. Ekins and B. Etheridge, "The environmental and economic impacts of the {UK} climate change agreements," Energy Policy, vol. 34, no. 15, pp. 2071–2086, 2006.

[220]   DECC, "Consultation on simplifying the Climate Change Agreements Scheme," 2011. [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/42823/2636-cca-simplification-condoc.pdf. [Accessed: 05-Nov-2013].

[221]   DEFRA, "Memorandum submitted by the Department for Environment, Food and Rural Affairs (Defra)," 2007. [Online]. Available: http://www.publications.parliament.uk/pa/cm200708/cmselect/cmenvaud/354/354we22.htm. [Accessed: 05-Nov-2013].

[222]   R. Martin, L. B. de Preux, and U. J. Wagner, "The Impacts of the Climate Change Levy on Business: evidence from microdata," Cent. Clim. Chang. Econ. Policy Grantham Res. Inst. Clim. Chang. Environ., vol. 6, 2009.

[223]   R. Martin, L. B. de Preux, and U. J. Wagner, "The impact of a carbon tax on manufacturing: Evidence from microdata," J. Public Econ., vol. 117, pp. 1–14, 2014.

[224]   K. Lo, "A critical review of China's rapidly developing renewable energy and energy efficiency policies," Renew. Sustain. Energy Rev., vol. 29, pp. 508–516, 2014.

[225]   J. Hu, F. Kahrl, Q. Yan, and X. Wang, "The impact of China's differential electricity pricing policy on power sector {CO2} emissions," Energy Policy, vol. 45, pp. 412–419, 2012.

[226]   B. Boza-Kiss, S. Moles-Grueso, and D. Urge-Vorsatz, "Evaluating policy instruments to foster energy efficiency for the sustainable transformation of buildings," Curr. Opin. Environ. Sustain., vol. 5, no. 2, pp. 163–176, 2013.

[227]   European Commission, "COMMISSION REGULATION (EU) No 617/2013 of 26 June 2013 - implementing Directive 2009/125/EC of the European Parliament and of the Council with regard to ecodesign requirements for computers and computer servers." 2013.

[228]   S. Rezessy and P. Bertoldi, "Voluntary agreements in the field of energy efficiency and emission reduction: Review and analysis of experiences in the European Union," Energy Policy, vol. 39, no. 11, pp. 7121–7129, 2011.

[229]   European Commission, Directive 2003/87/EC of the European Parliament and of the Council of 13 October 2003, establishing a scheme for greenhouse gas emission allowance trading within the Community and amending Council Directive 96/61/EC. Brussels, 2003.

[230]   M. R. Taylor, "Innovation under cap-and-trade programs," Potsdam Inst. Clim. Impact Res., vol. 109, no. 13, pp. 4804–4809, 2012.

[231]   C. Egenhofer, M. Alessi, A. Georgiev, and N. Fujiwara, "The EU Emissions Trading System and Climate Policy towards 2050: Real incentives to reduce emissions and drive innovation?," 2011.

[232]   A. D. Ellerman and B. K. Buchner, "Over-Allocation or Abatement? A Preliminary Analysis of the EU ETS Based on the 2005--06 Emissions Data," Environ. Resour. Econ., vol. 41, no. 2, pp. 267–287, 2008.

[233]   B. Anderson and C. Di Maria, "Abatement and Allocation in the Pilot Phase of the EU ETS," Environ. Resour. Econ., vol. 48, no. 1, pp. 83–103, 2010.

[234]   B. Declercq, E. Delarue, and W. D'haeseleer, "Impact of the economic recession on the European power sector's {CO2} emissions," Energy Policy, vol. 39, no. 3, pp. 1677–1686, 2011.

[235]   G. Bel and S. Joseph, "Emission abatement: Untangling the impacts of the {EU} {ETS} and the economic crisis," Energy Econ., vol. 49, pp. 531–539, 2015.

[236]   DECC, "Consultation on simplifying the CRC Energy Efficiency Scheme: Government Response,"   2012.   [Online].   Available: http://www.official-documents.gov.uk/document/cm84/8486/8486.pdf. [Accessed: 06-Nov-2013].

[237]   CRC Energy Efficiency Scheme Order 2010. 2010.

[238] D. Ürge-Vorsatz, A. Novikova, S. Köppel, and B. Boza-Kiss, "Bottom–up assessment of potentials and costs of CO2 emission mitigation in the buildings sector: insights into the missing elements," Energy Effic., vol. 2, pp. 293–316, 2009.

[239] European Commission, "Code of Conduct on Data Centres Energy Efficiency Version 2.0 Participant Guidelines and Registration Form," 2009. [Online]. Available: http://iet.jrc.ec.europa.eu/energyefficiency/sites/energyefficiency/files/files/documents/ICT_CoC/participant_guidelines_v2_0-final.pdf. [Accessed: 18-Jan-2016].

[240] APC, "Determining Total Cost of Ownership for Data Center and Network Room Infrastructure," 2005. [Online]. Available: http://www.linuxlabs.com/PDF/Data Center Cost of Ownership.pdf. [Accessed: 06-Aug-2014].

[241] U.S. Energy Information Administration, "International Energy Outlook 2010," 2010.

[242] International Energy Agency, "Energy Technology Transitions for Industry," 2009. [Online]. Available: https://www.iea.org/publications/freepublications/publication/industry2009.pdf. [Accessed: 18-Jan-2016].

[243] A. J. Howard and J. Holmes, "Addressing data center efficiency: lessons learned from process evaluations of utility energy efficiency programs," ENERGY Effic., vol. 5, no. 1, SI, pp. 137–148, Jan. 2012.

[244] United States Environmental Protection Agency, "Understanding and Designing Energy-Efficiency Programs for Data Centers," 2012. [Online]. Available: http://www.energystar.gov/ia/products/power_mgt/ES_Data_Center_Utility_Guide.pdf. [Accessed: 14-Oct-2013].

[245] NRDC, "Data Center Efficiency Assessment - Scaling Up Energy Efficiency Across the Data Center Industry: Evaluating Key Drivers and Barriers," 2014. [Online]. Available: http://www.nrdc.org/energy/files/data-center-efficiency-assessment-IP.pdf. [Accessed: 15-Jan-2016].

[246] G. A. Brady, "Energy Efficiency in Data Centres and the Barriers to Further Improvements: An Interdisciplinary Investigation,", p105-165, University of Leeds, 2016.

[247] P. Bertoldi, "A Market Transformation Programme for Improving Energy Efficiency in Data Centres," in ACEEE Summer Study on Energy Efficiency in Buildings:, 2014, pp. 9-14-26.

[248]   D. Vernon and A. Meier, "Identification and quantification of principal–agent problems affecting energy efficiency investments and use decisions in the trucking industry," Energy Policy, vol. 49, no. 0, pp. 266–273, Oct. 2012.

[249]   W. Graus and E. Worrell, "The principal–agent problem and transport energy use: Case study of company lease cars in the Netherlands," Energy Policy, vol. 36, no. 10, pp. 3745–3753, Oct. 2008.

[250]   L. Maruejols and D. Young, "Split incentives and energy efficiency in Canadian multi-family dwellings," Energy Policy, vol. 39, no. 6, pp. 3655–3668, Jun. 2011.

[251]   A. Levinson and S. Niemann, "Energy use by apartment tenants when landlords pay for utilities," Resour. ENERGY Econ., vol. 26, no. 1, pp. 51–75, Mar. 2004.

[252]   L. Davis, "Evaluating the Slow Adoption of Energy Efficient Investments Are Renters Less Likely to Have Energy Efficient Appliances?," in NBER Conference on Design and Implementation of U.S. Climate Policy Location, 2010.

[253]   C. Blumstein, "Program evaluation and incentives for administrators of energy-efficiency programs: Can evaluation solve the principal/agent problem?," Energy Policy, vol. 38, no. 10, pp. 6232–6239, Oct. 2010.

[254]   P. Bhattacharya, "Data Center Monitoring," in Energy efficient thermal management of data centers, Y. Joshi and P. Kumar, Eds. New York: Springer, 2012.

[255]   C. Drake, "The DCDi 2013 Census - UK Figures," 2013. [Online]. Available: http://www.datacenterdynamics.com/focus/archive/2013/11/dcdi-2013-census-–-uk-figures. [Accessed: 04-Apr-2014].

[256]   techUK, "HMT Consultation 2015: Reforming the business energy efficiency tax landscape: techUK response," 2015.

[257]   E. & I. S. Department for Business, "Energy Technology List (ETL)," 2016. [Online]. Available: https://www.gov.uk/guidance/energy-technology-list#enhanced-capital-allowance-eca-scheme. [Accessed: 23-Feb-2017].

[258]   techUK, "DECC Call for Evidence: the Energy Technology List - Feedback from the UK technology   sector,"   2016.   [Online].   Available: https://www.techuk.org/images/techUK_ETL_response.pdf. [Accessed: 23-Feb-2017].

[259] European Commission, "Joint Research Centre Energy Efficiency - Organisation list," 2017. [Online]. Available: http://iet.jrc.ec.europa.eu/energyefficiency/organisation-list-short/ict_coc_dc_partner. [Accessed: 23-Feb-2017].

[260] UK Council of Data Centre Operators, "Communication: European Code of Conduct for Data Centres," 2015. [Online]. Available: https://www.techuk.org/images/techUK_DCC_Com_EUCOC.pdf. [Accessed: 22-Feb-2017].

[261] CENELEC, "EN 50600-2-5:2016," 2016. [Online]. Available: https://www.cenelec.eu/dyn/www/f?p=104:110:1734924540251401::::FSP_ORG_ID,FSP_PROJECT,FSP_LANG_ID:1258297,56390,25. [Accessed: 23-Feb-2017].

[262] British Standards Institution, "BS EN 50600 - Information technology," 2017. [Online]. Available: http://landingpage.bsigroup.com/LandingPage/Series?UPI=BS EN 50600. [Accessed: 26-Apr-2017].

[263] M. Acton, "The European Code of Conduct, EN50600 and the Data Centre Standards Landscape," 2016. [Online]. Available: https://www.dceureca.eu/workshop/the-4th-eureca-workshop/EURECA_Mark_Acton.pdf. [Accessed: 28-Feb-2017].

[264] British Standards Institution, "Standards policy on the UK leaving the EU," 2017. [Online]. Available: https://www.bsigroup.com/en-GB/about-bsi/uk-national-standards-body/EUReferendum/. [Accessed: 26-Apr-2017].

[265] BREEAM, "Case Studies - Datacentres," 2017. [Online]. Available: http://www.breeam.com/case-studies-datacentres. [Accessed: 23-Feb-2017].

[266] ISO, "ISO 50001 - Energy management." [Online]. Available: http://www.iso.org/iso/home/standards/management-standards/iso50001.htm. [Accessed: 23-Feb-2017].

[267] Blackmores, "Data Centre Energy Management ISO 50001," 2017. [Online]. Available: http://blackmoresuk.com/data-centre-energy-management-iso-50001/. [Accessed: 23-Feb-2017].

[268] Environment Agency, "About ESOS," 2017. [Online]. Available: https://www.gov.uk/guidance/energy-savings-opportunity-scheme-esos. [Accessed: 21-Jun-2017].

[269] European Commission, Energy Performance of Buildings Directive. European Parliament, 2010.

[270]   U.S. Environmental Protection Agency, "Enterprise Servers Specification Version 2.0," 2013. [Online]. Available: http://www.energystar.gov/products/specs/enterprise_servers_specification_version_2_0 _pd. [Accessed: 13-Jan-2014].

[271]   S. G. Heeringa, B. T. West, and P. A. Berglund, Applied Survey Data Analysis. Chapman and Hall/CRC, 2010.

[272]   U. Flick, An Introduction to Qualitative Research, 4th ed. London: SAGE Publications Ltd, 2009.

[273]   W. E. Saris, Design, Evaluation and Analysis of Questionnaires for Survey Research. Hoboken: Wiley-Interscience, 2007.

[274]   D. A. Dillman, Mail and Internet Surveys: The Tailored Design Method, 2nd, 2007 ed. Hoboken: John Wiley & Sons, Ltd., 2007.

[275]   H. Schuman and S. Presser, Questions and answers in attitude surveys: experiments on question form, wording and context. London: SAGE Publications Ltd, 1981.

[276]   S. G. Heeringa, B. T. West, and P. A. Berglund, "Getting to Know the Complex Sample Design," in Applied Survey Data Analysis, Chapman and Hall/CRC, 2010, pp. 13–51.

[277]   H. Arksey and P. Knight, "Designing an Interview-Based Study," in Interviewing for Social Students, London: SAGE Publications Ltd, 1999.

[278]   M. Mason, "Sample Size and Saturation in PhD Studies Using Qualitative Interviews," Forum Qual. Soc. Res., vol. 11, no. 3, p. Article 8, 2010.

[279]   M. N. Marshall, "Sampling for qualitative research," Fam. Pract., vol. 13, no. 6, pp. 522–525, 1996.

[280]   H. J. Rubin and I. S. Rubin, "Conversational Partnerships," in Qualitative Interviewing: The Art of Hearing Data, 2nd ed., London: SAGE Publications Ltd, 2005.

[281]   M. Galesic and M. Bosnjak, "Effects of questionnaire length on participation and indicators of response quality in a web survey," Public Opin. Q., vol. 73, no. 2, pp. 349–360, 2009.

[282]   H. Schuman and S. Presser, "Balance and Imbalance in Questions," in Questions and Answers in Attitude Surveys, London: SAGE Publications Ltd, 1981.

[283]   B. Gillham, "The Nature of The Interview," in The Research Interview, London: Bloomsbury, 2001.

[284]   B. Gillham, "Interviewing: For and Against," in The Research Interview, London: Bloomsbury, 2001.

[285]   B. L. Berg, "A Dramaturgical Look at Interviewing," in Qualitative Research Methods for the Social Sciences, 6th ed., London: Pearson, 2007.

[286]   H. Schuman and S. Presser, "Open Versus Closed Questions," in Questions and Answers in Attitude Surveys, London: SAGE Publications Ltd, 1981, p. 111.

[287]   F. J. Fowler, A. M. Roman, and Z. X. Di, "Mode effects in a survey of Medicare prostate surgery patients," PUBLIC Opin. Q., vol. 62, no. 1, pp. 29–46, 1998.

[288]   D. A. Dillman, "Mixed-Mode Surveys," in Mail and internet surveys: The tailored design method, 2nd ed., Hoboken: John Wiley & Sons, Ltd., 2007.

[289]   B. Gillham, "The Use of Prompts and Probes," in The Research Interview, London: Bloomsbury, 2001.

[290]   J. A. Krosnick, "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys," Appl. Cogn. Psychol., vol. 5, pp. 213–236, 1991.

[291]   B. Gillham, "The Interviewer as the Research Instrument," in The Research Interview, London: Bloomsbury, 2001.

[292]   H. J. Rubin and I. S. Rubin, "The Responsive Interview as an Extended Conversation," in Qualitative Interviewing: The Art of Hearing Data, 2nd ed., London: SAGE Publications Ltd, 2005.

[293]   B. Gillham, "Piloting and Running the Interview," in The Research Interview, London: Bloomsbury, 2001.

[294]   H. J. Rubin and I. S. Rubin, "The First Phase of Analysis: Preparing Transcripts and Coding Data," in Qualitative Interviewing: The Art of Hearing Data, 2nd ed., London: SAGE Publications Ltd, 2005.

[295]   B. L. Berg, "An Introduction to Content Analysis," in Qualitative Research Methods for the Social Sciences, 6th ed., London: Pearson, 2007.

[296]   B. Gillham, "Carrying out a Content Analysis," in The Research Interview, London: Bloomsbury, 2001.

[297]   H. J. Rubin and I. S. Rubin, "Analyzing Coded Data," in Qualitative Interviewing: The Art of Hearing Data2, 2nd ed., London: SAGE Publications Ltd, 2005.

[298]   A. Trianni, E. Cagno, P. Thollander, and S. Backlund, "Barriers to industrial energy efficiency in foundries: a European comparison," J. Clean. Prod., vol. 40, no. 0, pp. 161–176, 2013.

[299]   E. Sardianou, "Barriers to industrial energy efficiency investments in Greece," J. Clean. Prod., vol. 16, no. 13, pp. 1416–1423, 2008.

[300]   P. G. Taylor, R. Bolton, D. Stone, and P. Upham, "Developing pathways for energy storage in the {UK} using a coevolutionary framework," Energy Policy, no. 0, p. , 2013.

[301]   A. Trianni, E. Cagno, E. Worrell, and G. Pugliese, "Empirical investigation of energy efficiency barriers in Italian manufacturing {SMEs}," Energy, vol. 49, no. 0, pp. 444–458, 2013.

[302]   QSR International Pty Ltd., "NVIVO: THE #1 SOFTWARE FOR QUALITATIVE DATA ANALYSIS," 2017. [Online]. Available: http://www.qsrinternational.com/product. [Accessed: 21-Jun-2017].

[303]   HM Government, "Buying or selling your home," 2017. [Online]. Available: https://www.gov.uk/buy-sell-your-home/energy-performance-certificates. [Accessed: 21-Jun-2017].

[304]   Data Center Alliance, "About DCA Certifications," 2017. [Online]. Available: http://www.data-central.org/?page=Certifications. [Accessed: 21-Jun-2017].

Appendix 1 - Calculation of flow rate at the periphery of a circular duct

As established in section 2.2.1, the flow rate at the periphery of a circular duct, $\dot{V}_p$, can be calculated using Eq. A 1, where r is the distance from the centre of the duct, r_0 is the distance from the centre of the duct to the measurement points farthest from the centre, $\dot{V}_p$ is the flow rate in the annulus bounded by the circle at r_0 and the inside wall of the duct, v_0 is the mean velocity measured at $r = r_0$, R is the inner radius of the duct and m_R is a constant describing the roughness of the duct wall.

$$\dot{V}_p = 2\pi v_0 (R - r_0)^{-\frac{1}{m_R}} \int_{r_0}^{R} r(R - r)^{\frac{1}{m_R}} dr \qquad \text{Eq. A 1}$$

If $I_1 = \int_{r_0}^{R} r(R - r)^{\frac{1}{m_R}} dr$, we may integrate by parts, saying that $P_1 = r$, $dP_2 = (R - r)^{1/2}$ and $I_1 = P_1 P_2 - \int_{r_0}^{R} P_2 dP_1$. Differentiating P_1 gives $dP_1 = dr$. To find P_2 requires an integration by substitution, saying that $C = R - r$, and $\frac{dC}{dr} = -1$. P_2 may then be determined by Eq. A 2, after the following manipulations:

$$P_2 = \int_{r_0}^{R} (R - r)^{\frac{1}{m_R}} dr = -\int_{R-r_0}^{0} C^{\frac{1}{m_R}} dC$$

$$P_2 = \left[-\frac{C^{1+\frac{1}{m_R}}}{1+\frac{1}{m_R}} \right]_{C=R-r_0}^{C=0}$$

$$P_2 = \left[-\frac{(R-r)^{1+\frac{1}{m_R}}}{1+\frac{1}{m_R}} \right]_{r=r_0}^{r=R} \qquad \text{Eq. A 2}$$

Hence we may use the following manipulation to show that I_1 is given by Eq. A 3:

$$I_1 = \left[-\frac{r(R-r)^{1+\frac{1}{m_R}}}{1+\frac{1}{m_R}} + \left(1 + \frac{1}{m_R}\right) \int_{r_0}^{R} (R - r)^{1+\frac{1}{m_R}} dr \right]_{r=r_0}^{r=R}$$

$$I_1 = \left[-\frac{r(R-r)^{1+\frac{1}{m_R}}}{1+\frac{1}{m_R}} + \frac{m_R}{m_R + 1} \int_{r_0}^{R} (R - r)^{1+\frac{1}{m_R}} dr \right]_{r=r_0}^{r=R} \qquad \text{Eq. A 3}$$

It is now useful to define $I_2 = \int_{r_0}^{R}(R-r)^{1+\frac{1}{m_R}}dr$. I_2 may be integrated by substitution, now setting $C = R - r$, which yields $\frac{dC}{dr} = -1$ after differentiation. Thus, the following manipulations give I_2 via Eq. A 4:

$$I_2 = -\int_{R-r_0}^{0} C^{1+\frac{1}{m_R}}\, dC$$

$$I_2 = \left[-\frac{C^{2+\frac{1}{m_R}}}{2+\frac{1}{m_R}}\right]_{C=R-r_0}^{C=0}$$

$$I_2 = \left[-\frac{m_R(R-r)^{2+\frac{1}{m_R}}}{2m_R+1}\right]_{r=r_0}^{r=R} \qquad \text{Eq. A 4}$$

Eq. A 3 may therefore be manipulated as follows, giving Eq. A 5:

$$I_1 = \left[-\frac{r(R-r)^{1+\frac{1}{m_R}}}{1+\frac{1}{m_R}} - \frac{m_R^2(R-r)^{2+\frac{1}{m_R}}}{(m_R+1)(2m_R+1)}\right]_{r=r_0}^{r=R}$$

$$I_1 = \left[-\frac{m_R r(R-r)^{1+\frac{1}{m_R}}}{m_R+1} - \frac{m_R^2(R-r)^{1+\frac{1}{m_R}}(R-r)}{(m_R+1)(2m_R+1)}\right]_{r=r_0}^{r=R}$$

$$I_1 = \left[-\frac{m_R r(R-r)^{1+\frac{1}{m_R}}}{m_R+1} - \frac{m_R^2(R-r)^{1+\frac{1}{m_R}}(R-r)}{(m_R+1)(2m_R+1)}\right]_{r=r_0}^{r=R}$$

$$I_1 = \left[-\frac{m_R r(2m_R+1)(R-r)^{1+\frac{1}{m_R}}+m_R^2(R-r)^{1+\frac{1}{m_R}}(R-r)}{(m_R+1)(2m_R+1)}\right]_{r=r_0}^{r=R}$$

$$I_1 = \left[-\frac{m_R(R-r)^{1+\frac{1}{m_R}}(m_R r + r + m_R R)}{(m_R+1)(2m_R+1)}\right]_{r=r_0}^{r=R} \qquad \text{Eq. A 5}$$

Substituting Eq. A 5 back into Eq. A 1, whilst adding the limits, gives Eq. A 6.

$$\dot{V}_p = 2\pi m_R v_0(R-r_0)\left[\frac{m_R r_0 + m_R R + r_0}{(m_R+1)(2m_R+1)}\right] \qquad \text{Eq. A 6}$$

Appendix 2 - Calculation of flow rate at the periphery of a rectangular duct

As shown in section 2.2.3, the flow at the periphery of a rectangular duct is enclosed by 4 trapeziums, with the flow rate through trapezium ABECA (as defined in Figure 2-6) given by Eq. A 7. Here, W is the internal width of the duct, x is the distance from the measurement nearest to the duct wall, $v = v_0 \left(\frac{a-x}{a}\right)^{\frac{1}{m_R}}$ (from Karman's law of the periphery [107]), $v_{0,i}$ is the average velocity measured at a distance, a from duct wall i, and m_R is a constant describing the roughness of the duct wall.

$$\dot{V}_{ABECA} = v_0 \int_0^a (W - 2x)\left(1 - \frac{1}{a}x\right)^{\frac{1}{m_R}} dx \qquad \text{Eq. A 7}$$

Defining $I_1 = \int_0^a (W - 2x)\left(1 - \frac{1}{a}x\right)^{\frac{1}{m_R}} dx$, we must integrate I_1 by parts. Firstly, integrate $\left(1 - \frac{1}{a}x\right)^{\frac{1}{m_R}}$ by substitution, setting $P_1 = W - 2x$, $\frac{dP_1}{dx} = 2$, $dP_2 = \left(1 - \frac{1}{a}x\right)^{\frac{1}{m_R}}$. But in order to integrate dP_2, we must substitute in $C = 1 - \frac{1}{a}x$. Now $dP_2 = C^{\frac{1}{m_R}}$, which gives Eq. A 8 after the following manipulations:

$$P_2 = \int \left(1 - \frac{1}{a}x\right)^{\frac{1}{m_R}} dx = -\int a c^{1/m_R} dC$$

$$P_2 = -\frac{ac^{1+\frac{1}{m_R}}}{1+\frac{1}{m_R}} = -\frac{a}{1+\frac{1}{m_R}}\left(1 - \frac{x}{a}\right)^{1+\frac{1}{m_R}}$$

$$P_2 = -\frac{m_R a}{m_R+1}\left(1 - \frac{x}{a}\right)\left(1 - \frac{x}{a}\right)^{1/m_R}$$

$$P_2 = \frac{m_R(x - a)}{m_R + 1}\left(1 - \frac{x}{a}\right)^{1/m_R} \qquad \text{Eq. A 8}$$

Hence, we may then derive an equation for I_1, Eq. A 9, via the following manipulations:

$$I_1 = \int_0^a (W - 2x)\left(1 - \frac{x}{a}\right)^{\frac{1}{m_R}} dx = P_1 P_2 - \int P_2 dP_1$$

$$I_1 = \frac{(W-2x)a\left(1 - \frac{x}{a}\right)^{1+\frac{1}{m_R}}}{1+\frac{1}{m_R}} - 2\int \frac{a}{1+\frac{1}{m_R}}\left(1 - \frac{x}{a}\right)^{1+1/m_R} dx$$

$$I_1 = \frac{am_R(W-2x)\left(1-\frac{x}{a}\right)^{1+\frac{1}{m_R}}}{m_R+1} - \frac{2am_R}{m_R+1}\int\left(1-\frac{x}{a}\right)^{1+1/m_R}dx \qquad \text{Eq. A 9}$$

Now if we deal with the integral in this term, stating that $I_2 = \int\left(1-\frac{x}{a}\right)^{1+1/m_R}dx$, we can

integrate by substitution, with $C = 1 - \frac{1}{a}x$, hence $dx = -adC$. This enables us to define I_2

by Eq. A 10, via the following manipulations:

$$I_2 = \int C^{1+1/m_R}dx = -a\int C^{1+\frac{1}{m_R}}dC = -\frac{aC^{2+\frac{1}{m_R}}}{2+\frac{1}{m_R}}$$

$$I_2 = -\frac{aC^{2+\frac{1}{m_R}}}{2+\frac{1}{m_R}} = -\frac{a\left(1-\frac{x}{a}\right)^{2+\frac{1}{m_R}}}{2+\frac{1}{m_R}}$$

$$I_2 = -\frac{a^2m_R\left(1-\frac{x}{a}\right)^2\left(1-\frac{x}{a}\right)^{\frac{1}{m_R}}}{a(2m_R+1)}$$

$$I_2 = -\frac{m_R(a^2-ax)\left(1-\frac{x}{a}\right)\left(1-\frac{x}{a}\right)^{\frac{1}{m_R}}}{a(2m_R+1)}$$

$$I_2 = -\frac{m_R(a^2-2ax+x^2)\left(1-\frac{x}{a}\right)^{\frac{1}{m_R}}}{a(2m_R+1)}$$

$$I_2 = -\frac{m_R(a-x)^2\left(1-\frac{x}{a}\right)^{\frac{1}{m_R}}}{a(2m_R+1)} \qquad \text{Eq. A 10}$$

Substituting Eq. A 10 back into Eq. A 9 gives us Eq. A 11.

$$I = \frac{am_R(W-2x)\left(1-\frac{x}{a}\right)^{1+\frac{1}{m_R}}}{m_R+1} + \frac{2am_R}{m_R+1}\frac{m_R(a-x)^2\left(1-\frac{x}{a}\right)^{\frac{1}{m_R}}}{a(2m_R+1)} \qquad \text{Eq. A 11}$$

Applying limits to Eq. A 11 gives Eq. A 12, via the following manipulation:

$$\dot{V}_{P1} = v_{0,1}\left[\frac{am_R}{m_R+1}\left(\frac{2am_R\left(1-\frac{x}{a}\right)^{2+\frac{1}{m_R}}}{2m_R+1} - (W-2x)\left(1-\frac{x}{a}\right)^{1+\frac{1}{m_R}}\right)\right]_{x=0}^{x=a}$$

$$\dot{V}_{P1} = \frac{v_{0,1}am_R}{m_R+1}\left(W - \frac{2am_R}{2m_R+1}\right) \qquad \text{Eq. A 12}$$

Applying the same approach to the other walls gives Eq. A 13:

$$\dot{V}_P = (v_{0,1} + v_{0,3})\frac{am_R}{m_R + 1}\left(W - \frac{2am_R}{2m_R + 1}\right) + (v_{02} + v_{04})\frac{am_R}{m_R + 1}\left(H - \frac{2am_R}{2m_R + 1}\right)$$

Eq. A 13

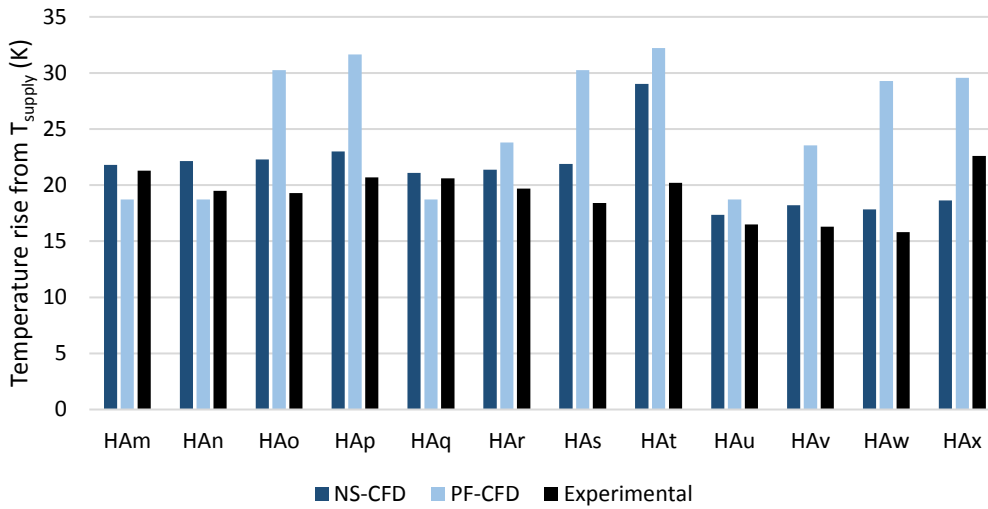# Appendix 3 – Full predicted and measured temperature results for Tests 1 and 2

Figure A 1 to Figure A 10 show the full set of computational fluid dynamics simulation results for temperature, for the simulations replicating Tests 1 and 2, as described in section 5.5.3. These results are compared with the corresponding experimental temperature measurements referred to in section 2.2.3.3.
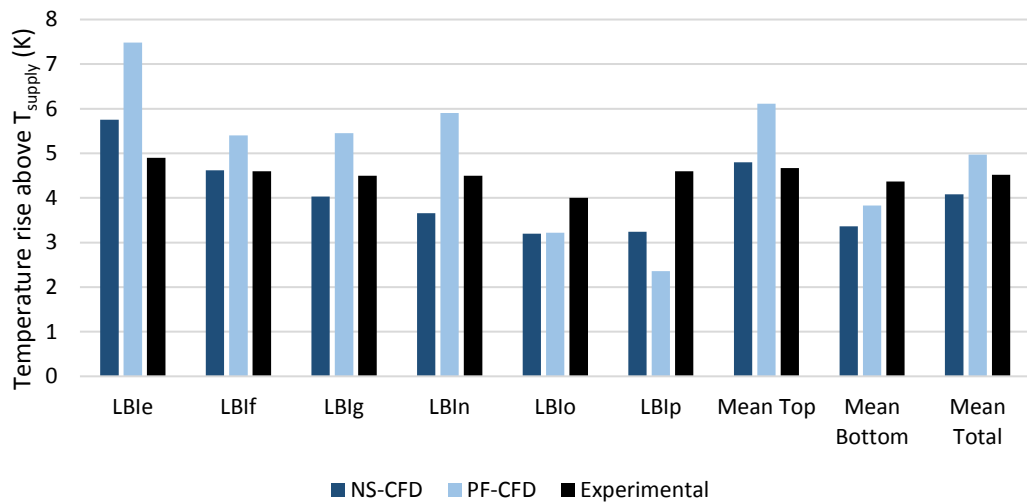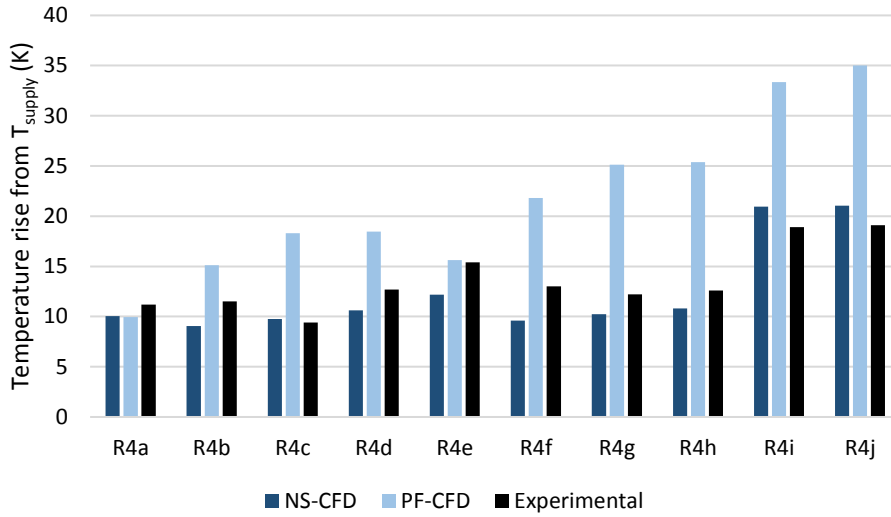

Figure A 1. Comparison of predicted and measured temperature rise at load bank inlets for Test 1 (measurement positions as defined in Figure 2-19).
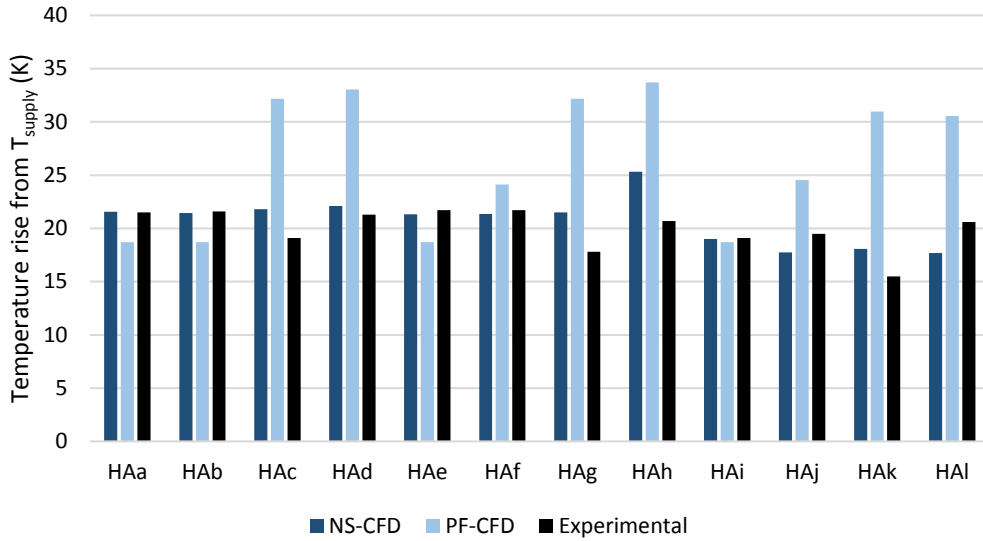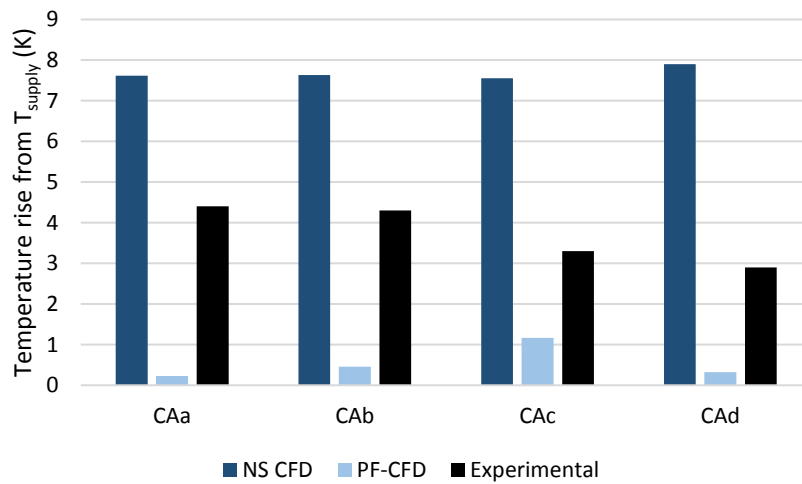
Figure A 2. Comparison of predicted and measured temperature rise at hot aisle measurement positions HAa to HAl (as defined in Figure 2-17) for Test 1.
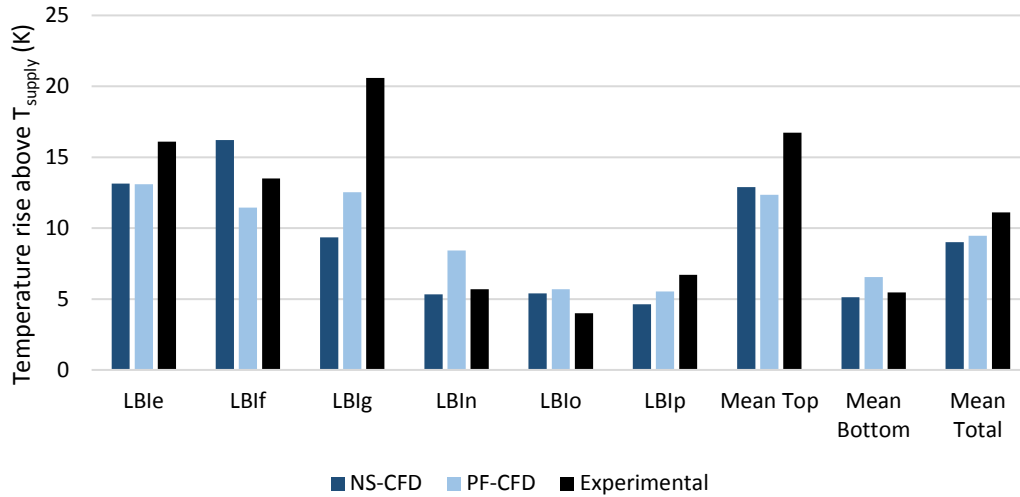

Figure A 3. Comparison of predicted and measured temperature rise at hot aisle measurement positions HAm to HAx (as defined in Figure 2-17) for Test 1.
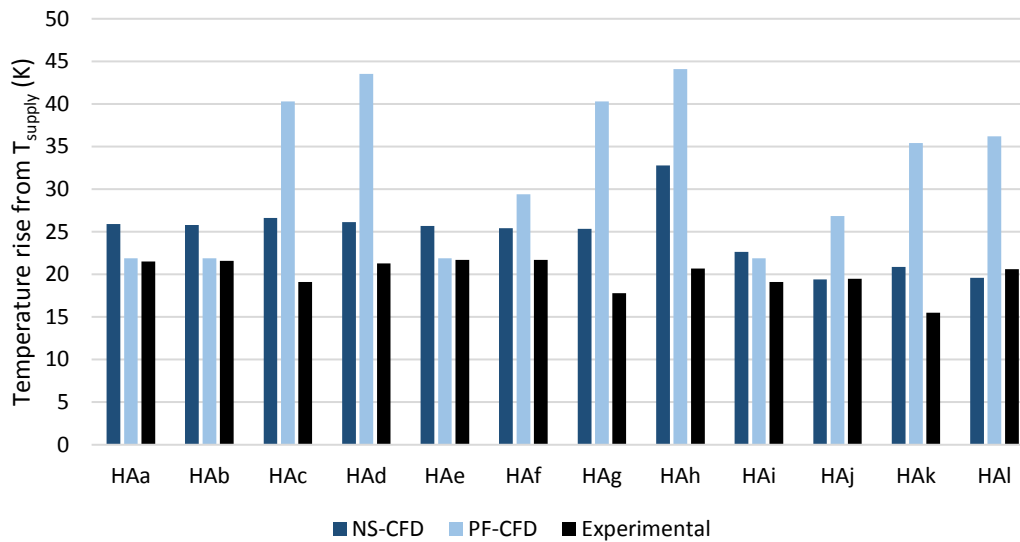
Figure A 4. Comparison of predicted and measured temperature rise at measurement positions within rack 4 (as defined in Figure 2-16) for Test 1.
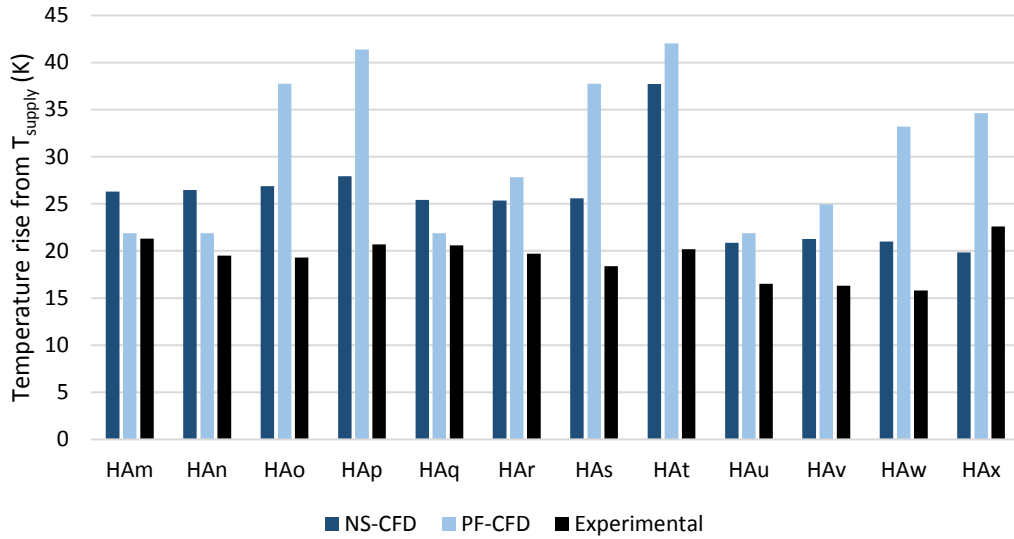

Figure A 5. Comparison of predicted and measured temperature rise at cold aisle measurement positions (as defined in Figure 2-18) for Test 1.
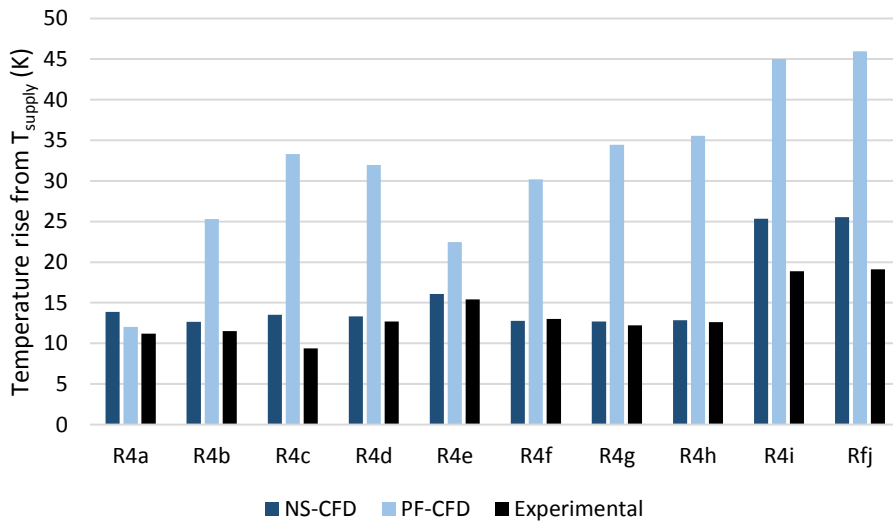
Figure A 6. Comparison of predicted and measured temperature rise at load bank inlets for Test 2 (measurement positions as defined in Figure 2-19).
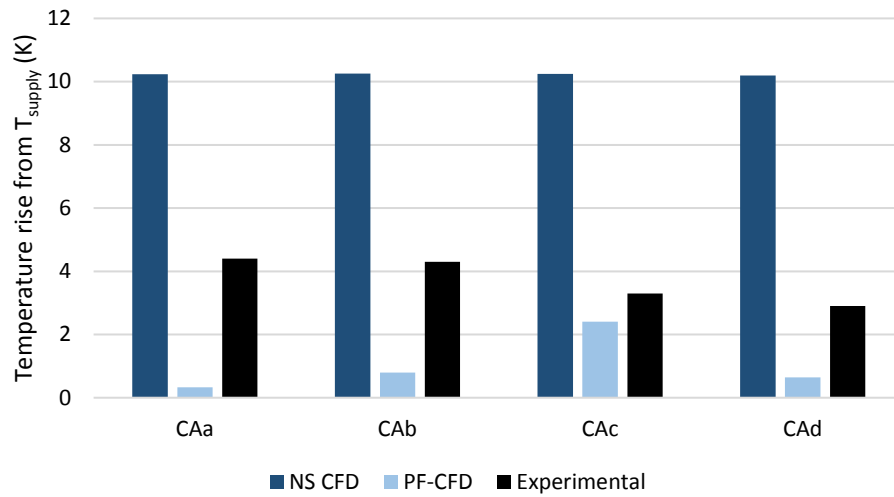

Figure A 7. Comparison of predicted and measured temperature rise at hot aisle measurement positions HAa to HAl (as defined in Figure 2-17) for Test 2.

Figure A 8. Comparison of predicted and measured temperature rise at hot aisle measurement positions HAm to HAx (as defined in Figure 2-17) for Test 2.

Figure A 9. Comparison of predicted and measured temperature rise at measurement positions within rack 4 (as defined in Figure 2-16) for Test 2.

Figure A 10. Comparison of predicted and measured temperature rise at cold aisle measurement positions (as defined in Figure 2-18) for Test 2.

Appendix 4 – Interview Script

<u>Prior to Recording</u>

Prior to the start of the interview the interviewee will be asked to read and sign the Consent Form, and will be invited to ask any questions they may have. The interviewer will reiterate that the interview will be recorded, and that some notes may be taken to help to keep track of what material has been covered. Recording will then commence.

<u>Introduction</u>

The interviewer will state the date and interviewee code, and will thank the interviewee for agreeing to participate in the research. The interviewee will then be informed that the interview will start will a few shorter questions intended to ascertain some information about their job role and the company they work for. The subject will then switch to policy instruments, which is the main focus of the interview.

	Main Questions	Probes
1	Could you describe your job role?	Are you responsible for: - energy consumption - cooling/power provision - IT - winning customers
2	Could you describe the business model of the data centre you represent?	- Is it an enterprise/co-location data centre? - How many individual customers?
3	Which team is responsible for paying the facility's electricity bill?	- Is this passed on to the customer?
4	Could you give some indication of the size of the data centre(s)?	- Approximately what is the power consumption? - What is the total floor space? - Is it currently at full capacity?
5	How long has the data centre been in operation?	

6	What is the data centre's tier classification?	
7	Do you measure PUE?	- If so, what PUE do you achieve? - Do you measure energy efficiency using any other methods?

Code of Conduct

The interviewee will now be informed that the following questions concern the EU's Code of Conduct on Data Centre Energy Efficiency.

	Question	Probes
8	Is your data centre a participant of the EU's Code of Conduct on Data Centre Energy Efficiency?	If yes: How long have you held participant status? Could you explain what your reasons were for applying for participant status? Did you see it as a chance to reduce your energy expenditure? Are the best practices contained in the code useful? Did you expect to see any other benefits, besides saving energy? Are there any ways in which you think the scheme could be improved? Anything that should/shouldn't be included? Does a lack of promotion impair uptake of the CoC? Does a lack of enforced periodic reporting limit the CoC's impact? If no: What would you say prevents you from seeking participation status? Are the best practices contained in the code useful? If the Code of Conduct was managed differently could you imagine seeking participant status? Does a lack of promotion impair uptake of the CoC? Does a lack of enforced periodic reporting limit the CoC's impact?

9	Do you think the Code of Conduct has had much of an effect on the industry as a whole?	Thinking about both participants in the Code and others.
10	What do you think about the rate of uptake of the Code of Conduct?	- Why do you think it hasn't gained wider participation? - Has poor administration inhibited further uptake? - Is better promotion needed to increase uptake? - Would translation into other languages increase uptake? - Do you think the CoC is more important to co-location providers or enterprise data centres?

Other Policy Instruments

The interviewee will now be informed that the next section of the interview will be concerned with the effects of other policy approaches to addressing energy efficiency in the data centre industry.

	Question	**Probes**
11	Are there any other policies which you think would help to stimulate efficiency improvements in the industry?	
12	Do you think the Climate Reduction Commitment and Climate Change Agreement are effective at driving energy efficiency in the sector?	- Do you think these approaches could result in data centre services being kept in house in inefficient server rooms? - Do you think that these policies drive data centre services abroad? - How do you feel about the reporting regime around these policies?

13	Do you think that the integration of the CoC best practices into the CENELEC EN 50600 standards series will help to drive efficiency improvements?	Would enforced auditing against such a standard be effective?
14	In the US the Department of Energy set up a workshop a few years ago to identify potential ways to stimulate efficiency improvements. The recommendations included things like subsidized energy audits and engineering and design services, incentives for efficient technologies, training programs…do you think these kinds of measures could be useful?	- Do you think there's an issue with a lack of understanding of energy efficiency amongst data centre operators? - Do you think it can be difficult to get unpartisan advice from consultants? Could a government-led program help to address this? - Do you think people carrying out these programs would need to be data centre specialists?

Aisle Containment

The interviewee will now be informed that the final section of the interview will focus on issues related to aisle containment.

	Question	**Probes**
15	Do you have aisle containment installed at your data centre(s)?	If yes: - What were your reasons for installing aisle containment? - Was it an expensive or difficult process to install aisle containment? If no: - What are your reasons for not installing aisle containment?
16	What factors do you think have slowed the adoption of aisle containment?	- Is cost a barrier? - Do practical issues reduce adoption? - Is there a lack of understanding of the benefits?

Closing

The interviewee will now be informed that all of the formal questions have been completed, and they will be asked if there is anything else they feel they have not had a chance to say, or if they have any questions.

Finally, the interviewee is thanked for their participation, and the recording is stopped.

Appendix 5 – Information Letter

UNIVERSITY OF LEEDS

Energy Research Institute
Faculty of Engineering
University of Leeds
LS2 9JT

[DATE]

Dear [NAME OF RECIPIENT],

We are writing to ask if a representative of your company would be interested in participating in a research project being undertaken at the University of Leeds, entitled **"Investigating the Role of Policy Instruments in Stimulating Energy Efficiency Improvements in the Data Centre Industry"**. The project is part of a wider research theme within the School of Mechanical Engineering which investigates data centre energy efficiency via laboratory-based experiments, computational modelling and qualitative methods.

Please take the time to read the information contained in this letter as it explains the format of the research and details what would be required of you should you decide to take part. We will be in touch shortly to enquire whether you would like to participate in the study. If you wish to contact us in the meantime please feel free to do so via the details provided at the foot of this letter. **Should you decide to take part, your participation will entail a short interview which may be undertaken at your place of work, at a time which is convenient to you.**

Why have I been invited to take part?

Your company has been identified as operating data centres within the UK, and your views regarding the study's subject are likely to be very useful in informing the conclusions of the research.

What would participation in the project involve?

If you agree to take part, your contribution to the project would extend to a semi-structured interview, which would take approximately 45 minutes to complete. This would be carried out face to face at your place of work, at a time convenient to you. The interview will take the form of a series of pre-written questions, with opportunity for wider discussion. The interview will be recorded to avoid the need for time-consuming note taking.

Will my participation and interview responses be kept confidential?

Great care will be taken to ensure confidentiality and anonymity of participants. All data pertaining to the participants and their interview responses will be anonymised and stored securely on the University's network, with a password known only by the interviewer being required for access.

How will the results of the research be published?

The results of the project will be published as part of a PhD thesis. Publication through an academic journal will also be pursued. Published reports will not name any of the participants.

What are the benefits of taking part?

This research seeks to provide an academically rigorous and impartial investigation into the impacts of policy instruments currently being used to stimulate energy efficiency improvements in the data centre industry. Your participation would be very valuable in helping to achieve this goal, and would give you an opportunity to share your insights relating to this important issue.

Who is the research funded by?

The research is funded by the Engineering and Physical Sciences Research Council, and forms part of a PhD project being undertaken within the Doctoral Training Centre for Low Carbon Technologies at the University of Leeds.

We very much appreciate your taking the time to read this letter/email. We will contact you shortly to ask if you would be willing to participate, but please do not hesitate to contact ourselves via the email address or telephone number below should you have any questions.

Kind regards

Mr Morgan Tatchell-Evans   Prof Nik Kapur
PhD Researcher   Professor of Applied Fluid Mechanics
1st Floor, Energy Building   2.41 School of Mechanical Engineering
Faculty of Engineering   Faculty of Engineering
University of Leeds   University of Leeds
LS2 9JT   LS2 9JT

Email - pm08mrte@leeds.ac.uk

Appendix 6 – Consent Form

UNIVERSITY OF LEEDS

Consent to take part in Research Project: "Investigating the Role of Policy Instruments in Stimulating Energy Efficiency Improvements in the Data Centre Industry"

	Add your initials next to the statements you agree with
I confirm that I have read and understand the letter dated [DATE] explaining the above research project and I have had the opportunity to ask questions about the project.	
I agree for the data collected from me to be stored and used in relevant future research in an anonymised form.	
I agree to take part in the above research project and will inform the lead researcher should my contact details change.	
I understand that it will be possible to withdraw from the research project up until the point at which analysis of information gathered during interviews is undertaken.	

Name of participant	
Participant's signature	
Date	
Name of lead researcher	
Signature	
Date*	

*To be signed and dated in the presence of the participant.