Northumbria Research Link

Citation: Victoria, Michele (2017) Developing decision support models for early stage embodied carbon management in buildings. Doctoral thesis, Northumbria University.

This version was downloaded from Northumbria Research Link: http://nrl.northumbria.ac.uk/id/eprint/36182/

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: http://nrl.northumbria.ac.uk/policies.html

Northumbria University NEWCASTLE

UniversityLibrary

DEVELOPING DECISION SUPPORT MODELS FOR EARLY STAGE EMBODIED CARBON MANAGEMENT IN BUILDINGS

MICHELE FLORENCIA VICTORIA

PhD

2017

DEVELOPING DECISION SUPPORT MODELS FOR EARLY STAGE EMBODIED CARBON MANAGEMENT IN BUILDINGS

MICHELE FLORENCIA VICTORIA

A thesis submitted in partial fulfilment of the requirements of the University of Northumbria at Newcastle for the degree of
Doctor of Philosophy

Research undertaken in the Faculty of Engineering and Environment

March 2017

Abstract

The regulatory requirement for improving operational energy efficiency in buildings make the unregulated Embodied Carbon (EC) of buildings relatively significant. The reduction potential of EC is high during the early stages of design while estimating EC during the early stage is challenging due to the unavailability of detailed design information. Similar to building costs, EC is also influenced by morphological and quality parameters of buildings. However, there is little evidence in the literature concerning the relationship between EC and design variables. Further, the increasing significance of the dual currency of construction projects emphasises the need for optimisation of cost and carbon of building designs. However, it is not easy to attain the best combination of cost and carbon without the adequate knowledge and expertise supported by decision support tools. Therefore, the research reported in this thesis addresses this knowledge gap by firstly identifying the relationship of the dual currency (cost and carbon) with building morphological and quality related parameters (referred to as 'design variables'). Later, developing Capital Cost (CC) and EC prediction models to assist in the dual currency estimating during the early stages of designs. The research findings are however, applicable to office buildings of low to medium-rise within a cradle-to-gate system boundary due to data constraints.

The approach involves the development and validation of a heuristic model of cost and carbon, using the statistical simulation of relevant morphological and quality parameters of buildings achieved through regression analysis. Historical project data from primary and secondary sources were collected and processed to develop a complete dataset of 41 buildings. The model variables were identified from a literature review and verified using the hotspot analysis. Finishes and services indices were developed to transform the qualitative variables into quantitative variables for an effective model building. The 'Finishes Quality Index' was developed through a Delphi based expert forum, while the 'Services Quality Index' was developed using price books. The developed EC model had 'Wall to Floor Ratio' and 'Number of Basements' as predictor variables while 'Circulation Space Ratio' and 'Building Height' were the predictor variables of the CC Model. In contrary to the literature findings, finishes and services quality were found to be statistically insignificant in the study, suggesting that finishes and services quality does not hugely influence the prediction of the dual currency of concept designs. However, Services were identified as carbon and cost significant in most of the buildings, while Finishes were identified as carbon and cost significant in some of the buildings of the sample.

The findings of the research have a number of contributions to theory and practice. The contribution to the design economics theory is the addition of the carbon dimension. The relationships analysed between EC and CC at building and element level add new insights to the EC literature. In addition, the methodology adopted for this thesis can form an exemplar for future research in different contexts. The key contribution to practice is the developed dual currency models (EC and CC models) which can be used to predict EC and CC of office buildings during early stages of design. Findings on carbon-critical elements (or carbon hotspots) of office buildings unveil building elements with high EC reduction potential that should be given the most attention during the design.

Acknowledgement

First, I thank the almighty God for his guidance and wisdom throughout this journey, for strengthening me when I am weak and for keeping me at peace during storms.

My heartfelt thanks go to my supervision team, which includes three key people who have been supporting me during this journey at different points of time. My special thanks go to Dr. Seraphim Alvanides who is my current principal supervisor who made the transition to the new supervisory team smooth and pleasant. I am grateful for his immense support and prompt response to all my queries. The key members of the supervisory team include Prof. Srinath Perera and Dr. Alan Davies who guided me from the beginning of the research to date. Prof. Srinath has not only been my former principal supervisor and current external supervisor but also a mentor. I am grateful for his consistent guidance and feedback on my work. I extend my thanks to Dr. Nirodha Fernando who acted as an advisor in this research for a short period. I would also like to thank Dr. Lei Zhou who was a former member of staff at Northumbria University and acted as a co-supervisor during the first year of my research. I am also thankful to the university administration team and support staff that made this journey less tiring.

I am indebted to the industry practitioners and academics who supplied me with the required data to accomplish this research who stay anonymous due to confidentiality reasons. Last but not the least my sincere thanks go to my family and friends back home and my church family in Newcastle who have been my moral support and cheer during my highs and lows.

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved by the School's Research Ethics Committee.

Word count of the main body of the thesis: 77,031

Signed :

Date :

Table of Contents

List of Figures

List of Tables

List of Publications

1. Victoria, M., Perera, S., Davies, A. and Fernando, N. 2017. Carbon and cost critical elements of buildings: an investigation of the UK office buildings. *Built Environment Project and Asset Management* (accepted for publication).

2. Victoria, M., Perera, S. and Davies, A. 2017. Carbon and cost hotspots of office buildings in the UK. In *The 6th World Construction Symposium*, 30 June - 02 July 2017, Colombo, Sri Lanka.

3. Perera, S. and Victoria, M. 2017. Role of carbon in sustainable development. In *Future challenges for sustainable development within the built environment. P. Lombardi*, G. Q. Shen, P. S. Brandon (Eds). Wiley's publication.

4. Victoria, M. and Perera, S. 2017. The application of elemental EC prediction model for buildings. In *WSBE17 Hong Kong*, 5-7 June 2017, Hong Kong.

5. Victoria, M., Perera, S. and Davies, A. 2016. Design economics for dual currency management in construction projects. In *RICS COBRA 2016*, September 2016, Toronto, Canada.

6. Victoria, M., Perera, S., Davies, A. and Fernando, N. 2016. Carbon and cost critical elements of buildings: a case study. In *The 5th World Construction Symposium*, 29 - 31 July 2016, Colombo, Sri Lanka.

7. Victoria, M., Perera, S., and Davies, A. 2016. A pragmatic approach for EC estimating in buildings. In *SBE16 Torino*, 18-19 February 2016, Torino, Italy.

8. Victoria, M. F., Perera, S., Zhou, L., & Davies, A. 2015. Estimating EC: A dual currency approach. In *Sustainable Buildings and Structures: Proceedings of the 1st International Conference on Sustainable Buildings and Structures*, 29 October-1 November 2015, Suzhou, PR China, CRC Press, 223-230.

9. Victoria, M, Perera, S and Davies, A. 2015. Developing an early design stage EC prediction model: A case study In Raidén, A B and Aboagye-Nimo, E (Eds) Proceedings of 31st Annual ARCOM Conference, 7-9 September 2015, Lincoln, UK, Association of Researchers in Construction Management, 267-276.

Papers under review:

10. Victoria, M. and Perera, S. 2017. Carbon hotspots: an EC mitigation approach during early stages of design, In *EC in Buildings: Measurement, Management and Mitigation*. Springer: UK (Submitted the revised paper).

Abbreviations

AI	Artificial Intelligence
BCIS	Building Cost Information Services
BER	Building Emissions Rate
BIM	Building Information Model
BoQ	Bills of Quantities
BRE	Building Research Establishment
CBR	Case-Based Reasoning
CC	Capital Cost
CIBSE	Chartered Institution of Building Services Engineers
CLT	Central Limit Theorem
CV	Coefficient of Variation
DEFRA	Department for Environment Food & Rural Affairs
EC	Embodied Carbon
ECE	EC Efficiency
ECO	Energy Company Obligation
EE	Embodied Energy
EPBD	Energy Performance of Buildings Directives
EU	European Union
EUQ	Element Unit Quantity
EUR	Element Unit Rates
FEES	Fabric Energy Efficiency Standard
GHG	Greenhouse Gas
GIFA	Gross Internal Floor area
HVAC	Heating Ventilating and Air Conditioning
ICE	Inventory of Carbon and Energy
ICT	Information and Communication Technology
LCA	Life Cycle Assessment
LCI	Life Cycle Inventory
LED	Light-Emitting Diode
MACC	Marginal Abatement Cost Curve
NN	Neural Network
NRM	New Rules of Measurement
NZEB	Nearly Zero-Energy Buildings

OC	Operational Carbon
OE	Operational Energy
PV	Photo Voltaic
QS	Quantity Surveyor
RHI	Renewable Heat Incentive
RIBA	Royal Institute of British Architects
RICS	Royal Institution of Chartered Surveyors
RQ	Research Question
SAP	Standard Assessment Procedure
SBEM	Simplified Building Energy Model
SMM	Standard Method of Measurements
TER	Target CO_2 Emission Rate
UK-GBC	UK Green Building Council
UNFCCC	United Nations Framework Convention on Climate Change
VIF	Variance Inflation Factor
WRAP	The Waste and Resources Action Programme

1. Introduction

1.1. Background

The construction industry is one of the largest consumers of both renewable and non-renewable resources (Dixit et al., 2010) and responsible for 30% of global Greenhouse Gas (GHG) emissions which creates a major impact on the environment (UK-GBC, 2014b). The UK Government mentioned in a White Paper on Energy, that GHG emissions challenge the stability of the world's climate, economy and population (RICS, 2008b). A more recent study reported that global food production is likely to decline by 0.5% in 2020 and by 2.3% in 2050 due to climate change (Calzadilla et al., 2013). Therefore, the emissions from the UK construction industry are regulated by stringent statutory requirements to minimise damage to the environment. For instance, the Energy Performance of Buildings Directive (2002) and the UK Building Regulations Part L (2006) use carbon emissions as a metric to measure building performance.

Emissions from buildings are mainly categorised into two types namely Operational Carbon (OC) and Embodied Carbon (EC) also known as capital carbon (HM Treasury, 2013). OC of buildings encapsulates emissions related to the energy consumption during the operation of the building. EC in buildings refers to the emissions involved in the construction of the built asset (including raw material extraction, manufacture, transport, construction, repair, replacement and demolition of materials or products). Of the two, OC has been given more attention as the contribution of OC emissions is generally higher than EC emissions. In fact, OC accounted for approximately 70-80% of total emissions until the introduction of Part L of the Building Regulations (RICS, 2014, Anderson, 2011) which set benchmark values for the acceptable amount of OC of typical building designs. However, the proportion of OC to EC varies depending on the location, building type and the life cycle of the building considered in the analysis (See for example, Ibn-Mohammed et al., 2013).

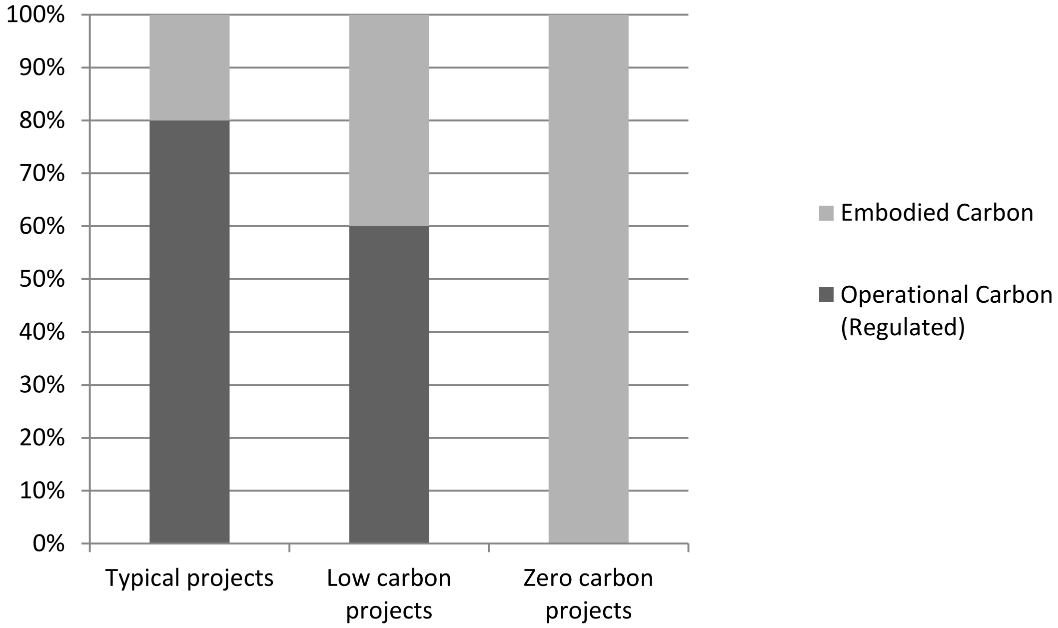


Figure 1.1: Increasing significance of EC in buildings

Modified from: RICS (2014)

The regulatory move of the UK government towards zero carbon buildings by 2019 aims at making OC almost nil which is shifting the focus of the government and regulatory bodies towards EC (Anderson, 2011, Sansom and Pope, 2012, Rawlinson and Weight, 2007) as there is no any legislative control over EC. As per Figure 1.1, if zero OC targets are to be met, 100% of the carbon emissions are projected to be from EC although the current deadlines of the UK are considered to be ambitious (Osmani and O'Reilly, 2009). Despite the debates on the achievement of zero carbon deadlines, there is a need to control EC. In fact, Rawlinson and Weight (2007) suggest that Embodied Energy (EE) might be ten times the annual operational energy in domestic buildings while this ratio could be as high as thirty times in commercial building. This demonstrates the need to manage EC.

Management requires measurement. In order to manage EC, there should be a standard method to quantify it. EC emissions can be calculated from cradle (earth)-to-gate, cradle-to-site, cradle-to-end of construction, cradle-to-grave, or even cradle-to-cradle which is called the system boundary of the calculation (Hammond and Jones, 2011) (Each of the system boundary mentioned here is explained in detail in Section 2.3). In particular, estimating EC is affected by several factors including system boundary, the method of estimating, location, accounted energy (primary or delivered energy), assumptions and data used (Dixit et al., 2010, Clark, 2013, Ekundayo et al., 2012, Hammond and Jones, 2008b). Hence, existing EC datasets and tools have deficiencies and are inconsistent. Further, the lack of agreement in the definition of EC and absence of a uniform method for quantification (Lockie, 2012, Dixit et al., 2012) signify the issues in estimating EC. Further, a recent survey conducted in the UK suggests that EC estimating is likely to be one of the future trends in the construction industry (Perera et al. 2015). Hence, it is important that the issue of standardisation is addressed and a standard measurement protocol is in place for EC estimating.

Carbon emissions reduction measures deliver a range of benefits to various stakeholders in addition to combating global threats. Reducing carbon emissions enriches competitive advantage and export potential of organisations; drives resource efficiency and better business solutions; leads to innovation; and provides health benefits (HM Treasury, 2013, Woodcock et al., 2009). Commercial buildings require more attention (See, Rawlinson and Weight, 2007) to reap the benefits and in particular, office buildings are in the forefront due to changing clients' perspectives to enhance their corporate social responsibility through legislatively complied sustainable offices (Target zero, 2012). This suggests that office buildings developers have started realising the benefits, hence, demanding carbon compliant designs.

Selecting the best design solution involves a systematic process as a design is developed from a conceptual to a more detailed design. In particular, RICS (2014) claims that the carbon reduction potential is high during the early stages of a project. Focusing on intensive emission sources would be one good approach for achieving high carbon reduction or to reap benefits (Carbon Trust, 2010, RICS, 2014, Halcrow Yolles, 2010b). For instance, the Pareto principle proposes that 80% of effects are attributable to 20% of the causes, in most cases. In the context of EC emissions of a building, it can be argued that 80% of the EC emissions are attributable to 20% of the building elements. These elements are referred to as the carbon-intensive elements of a building or 'carbon hotspots'. RICS (2014) proposes two conditions that should be met in order for an element to be classed as a carbon hotspot: (1) Measurement data is more easily available; and (2) Carbon reductions are possible (RICS, 2014).

However, the knowledge on carbon hotspots or the carbon-intensive elements is yet to be developed.

High Capital Cost (CC) of low and zero carbon designs, in comparison to conventional designs, used to be a major concern to clients (Catto, 2008), but it is now accepted that low and zero carbon buildings are attainable at an efficient cost on a par with conventional buildings (Sturgis and Roberts, 2010, Target zero, 2012) or at marginally higher cost (Department of Energy & Climate Change, 2012). In fact, Langston and Langston (2008) found that there is a positive linear relationship between EE and CC of projects. However, the findings of Langston and Langston (2008) cannot be substituted with EC as there are differences between EE and EC due to the process related emissions and sequestrations (Lélé, 1991, Brandt, 2012, Ayaz and Yang, 2009, see, Section 2.3 for more details). Therefore, the knowledge gap concerning the relationship between EC and CC needs to be explored. Especially, with increasing awareness towards the dual currency of construction projects (cost and carbon) the need to estimate, control and manage carbon alongside construction cost becomes fundamental for construction professionals and businesses to be sustainable (Ashworth and Perera, 2015). However, it is not easy to attain the best combination of cost and carbon without adequate knowledge and expertise supported by decision-making tools.

A range of online tools exists to help estimate the carbon accountability of building designs and some tools propose recommendations to reduce emissions. However, most of the estimating tools lack the transparency of the underlying methodology of calculations which is one of the major reasons for the inconsistency in outcomes of different tools (Čuček et al., 2012). In addition, some studies have pointed out that the inconsistency is basically due to the lack of EC measurement protocol and have stressed the immediate need for developing such a protocol (Dixit et al., 2010, Dixit et al., 2012). While there are estimating practices, tools and techniques pertinent to estimating EC, these are still in the early stages of development. Therefore, the need for an early stage EC estimating tool, which also generates CC estimates, was identified.

1.2. The Research Problem

The background of this research suggests that there is an urgent need to reduce not only the operational carbon but also the embodied carbon in buildings, which can be achieved with minimal time and effort by employing carbon management tools during the early stages of design. Further, cost effective low and zero carbon designs are demanded by current construction clients. Designers are therefore under the pressure to produce cost effective yet carbon optimum designs.The existing tools are limited in their capacity to perform both embodied carbon and cost estimating at the same time, however. This knowledge gap for adopting early stage carbon management in buildings will be addressed by the following Research Questions (RQs):

RQ1: How significant are embodied and OC in building projects and how are they regulated?

RQ2: What are the existing EC estimating tools, methods, their functions, outputs and limitations?

RQ3: What are the carbon-intensive elements or carbon hotspots in buildings?

RQ4: Are there statistically significant associations between EC and design variables and CC and design variables of buildings?

RQ5: Is there a statistically significant association between EC and CC of buildings?

RQ6: How can an early design stage EC prediction model be developed using design variables of buildings?

RQ7: How can the developed models of EC and CC be validated?

It has been proven that the cost of buildings is influenced by building morphological and quality parameters (or design variables), and this knowledge (i.e. design economics) yields economic benefits to construction clients (Ashworth and Perera, 2015, Seeley, 1996). Similarly, design economics with respect to the EC in buildings will contribute to knowledge and unlock novel thinking of designers to help design carbon efficient buildings. The knowledge about relationships between EC and building design variables will enable designers to produce cost and carbon optimum designs, even during the early stages of design. Therefore, by adopting the theory underlying parametric capital cost model prototypes, predicting embodied carbon in relation to design variables during the early stages of design, will be considered as an effective method. Accordingly, the research approach can be illustrated as follows:

$$Embodied\ Carbon\ per\ Gross\ Internal\ Floor\ area\ (GIFA) \left\lfloor \frac{\mathrm{EC}}{m^2} \right\rfloor \propto Morphological\ Parameters\ (M_P)$$

$$Embodied\ Carbon\ per\ GIFA \left\lfloor \frac{\mathrm{EC}}{m^2} \right\rfloor = \mathrm{a}\left(\frac{\mathrm{Wall}}{\mathrm{Floor}}\right) + \mathrm{b}(\mathrm{Storey\ Height}) + \mathrm{c}(\mathrm{Building\ Height}) + \cdots + \mathrm{k}$$

(Where, a, b, c…k = model coefficients)

Similarly, the effects of other design parameters such as the quality of services and the quality of finishes upon EC remain unexplored which are deemed to have a significant influence on emissions (Cole and Kernan, 1996). The relationship can be hypothesised as,

$$Embodied\ Carbon\ per\ GIFA \left\lfloor \frac{\mathrm{EC}}{m^2} \right\rfloor \propto Level\ of\ Sevices\ (L_S)$$

$$Embodied\ Carbon\ per\ GIFA \left\lfloor \frac{\mathrm{EC}}{m^2} \right\rfloor \propto Level\ of\ Finishes\ (L_F)$$

Consequently, EC per Gross Internal Floor Area (GIFA) can be expressed as follows:

$$Embodied\ Carbon\ per\ GIFA \left\lfloor \frac{\mathrm{EC}}{m^2} \right\rfloor = f(M_P, L_S, L_F)$$

Subsequently, the relationship between EC and CC can be compared and inferences can be made. This knowledge can support designers in decision-making with minimal time and effort and is likely to yield greater benefits.

1.3. Aim and Objectives

The aim of this research is to develop decision support models (EC and CC models) to aid design decision-making for early stage carbon management of building projects.

The following objectives were formulated in order to achieve the above aim:

- Review the significance of embodied and OC in building construction projects and relevant regulatory requirements and conventions.
- Evaluate the existing EC estimating tools, methods, their functions, outputs and limitations.
- Identify and analyse the carbon-intensive elements (hotspots) in buildings.
- Investigate the relationship between EC and building design variables and CC and building design variables.
- Investigate the relationship between EC and CC of buildings.
- Develop models for predicting EC and CC during early design stages.
- Validate the decision support models with real-time construction projects.

1.4. Scope and Limitations

The term early design stage refers to the first three stages from the Royal Institute of British Architects (RIBA) plan of work 2013, strategic brief, preparation and brief and concept design (RIBA, 2013). Particularly, the developed models best cater to the estimating needs of the 2-Concept Design stage of RIBA plan of work 2013. Further, the system boundary of EC estimating was selected as 'Cradle-to-Gate' (emissions associated with raw material extraction up to the manufacturing factory gate, see Section 2.3 for more details) due to the limitations in the availability of EC data.

Non-domestic buildings are responsible for higher EC emissions than domestic buildings and infrastructure (The Green Construction Board, 2013). It is also predicted that non-domestic floor area is expected to increase by 35% by 2050 (UK-GBC, 2014a). Hence, the focus of the study was on non-domestic buildings. In particular, office buildings are expected to grow at a rate of 2.7%

which is higher compared to other types of non-domestic building (The Green Construction Board, 2013). Further, The Green Construction Board (2013) states that the commercial office buildings are superior to other types of building in terms of the clarity of definition and availability of data which reduces the risk of uncertainty in modelling. In addition to that office buildings are the key focus of many scholars and an extensive work has been undertaken to improve the energy efficiency of office buildings (Halcrow Yolles, 2010a, Yohanis and Norton, 2002, Halcrow Yolles, 2010b, Cole and Kernan, 1996, Wu et al., 2012). For these reasons, office buildings were selected as the scope of the study.

The focus of the study was on low to medium-rise office buildings due to the limited availability of data. A building is classed as a high-rise building if it is not feasible to deploy external firefighting equipment and rescue operations due to its height or position (Department of Communities and Local Government, 2014). Even though there is no set absolute value of height to distinguish high-rise buildings from low to medium rise buildings, generally, buildings that are 30m and above (in other words, 10 storeys and above) are generally classified as high-rise buildings (Khoukhi and Al-Maqbali, 2011, Craighead, 2009, Anderson and Hammarberg, 2015). Accordingly, the boundary of the study was established to consider buildings up to nine storeys and the findings are applicable to the buildings falling within this range.

1.5. Structure of the Thesis

This thesis is presented in nine chapters. This chapter introduces the research background, the aim, the objectives and the scope and limitations of the study. The next two chapters are dedicated to the literature review. The legislations and conventions with regards to the carbon emissions, operational and EC in building and their relationships, low and zero carbon buildings, carbon hotspots and the cost and EC relationships are reviewed in Chapter 2 which sets the platform for EC research. A detailed review of the estimating practices of operational and EC including the methods, tools and databases in use are presented in Chapter 3 highlighting the deficiencies and limitations of the existing EC estimating practices. Chapter 3 concludes by proposing and conceptualising a solution to manage EC during the early stages of design.

The methodology adopted in the study is explained and justified in Chapter 4 by reviewing past research on cost modelling. The research philosophy, approach, and design are presented in detail in this chapter. The process of data collection, the description of different datasets used in the study, the process followed in the study sample development and the data collected through an expert forum are explained in Chapter 5. The carbon and cost hotspots analyses, the relationship between CC and EC and the regression analyses are presented in Chapter 6 followed by the model validations in Chapter 7. Key research findings and implications of these findings are presented in three major headings: the carbon and cost models, the carbon and cost hotspots, EC and CC relationships in Chapter 8. The way the study objectives have been achieved is reviewed and recommendations to the industry are made in Chapter 9. In addition, contributions to knowledge about the theory and to practice are discussed and further research directions are proposed in Chapter 9.

2. Carbon Emissions

2.1. Introduction

This chapter investigates the background of carbon management by reviewing international climate change regimes and the UK specific carbon management policies that are in place to combat climate change by stabilising the global temperature rise. A detailed carbon control trajectory for the UK is presented by capturing the emission reduction targets from international and national climate policies, conveying the seriousness of the climate change problem and the need for rigorous carbon emissions cuts. This chapter also introduces the concepts of energy and carbon in the built environment and reviews the relationship between energy and carbon. In addition, the concepts of operational and EC of buildings are introduced and a few EC case studies are presented to give an idea about the EC values of different types of buildings. In addition, the relationships between EC and OC is explored and the increasing significance of EC in a low and zero carbon built environment is emphasised. Finally, the literature on the relationship between carbon and cost of buildings is reviewed.

2.2. Carbon Management

The industrial revolution between 18th and19th centuries was a major milestone in the human history. Many countries experienced exceptional improvement in the economy due to new inventions and ideas that uplifted the status of countries. The Quality of Life Policy Group (2007) states that the industrial revolution was mainly driven by fossil fuels (coal, oil and gas), while BBC (2013) reported that coal was the major fuel which triggered the industrial revolution and Britain had plenty of coal that could be easily mined. Two centuries of industrial revolution brought material progress into the quality of human life while environmental burden caused by those aggressive developments was often overlooked (The Quality of Life Policy Group, 2007).

Intergovernmental Panel on Climate Change (2013) reported that 40% increase in CO_2 levels and 150% increase in methane levels were noticed as at 2011 since the industrial revolution. The increased amount of GHGs in the atmosphere leads to a steep global temperature rise causing the world climate to change. Impacts of the climate change include but are not limited to: decline in the food production such as agriculture and fisheries; substantially lower economic growth in low-income countries; energy sectors becoming sensitive and suffering due to excess demand; tourism to be affected in some parts of the world; occurrence of extreme weather events; human health problems; increase of civil conflicts; social and economic inequalities in poor populations (Intergovernmental Panel on Climate Change, 2013); and even heat-related deaths (Intergovernmental Panel on Climate Change, 2012). Hence, climate change is a fundamental challenge facing world regions.

Figure 2.1 illustrates the latest prediction of the mean global temperature rise at four different emission scenarios (from low to high – bars on the right-hand side of the figure indicates the range of variability of the predicted levels of the temperature rise) modelled by Intergovernmental Panel on Climate Change (2013). According to Figure 2.1, the mean global temperature at high emission scenario (Representative Concentration Pathway (RCP) 8.5, red line) tends to reach 4°C at a sharp rate by 2100 and there is a possibility of going beyond that level depending on the sensitivity of the assumptions. Further, RICS (2011), predicted that in the UK, summer temperatures are likely to increase by 4°C to 5°C, while rainfall levels may increase by 10 to 30% in the winter and decrease by 20% to 30% in the summer by 2080. Moreover, it was also pointed out that in the UK opening windows during summer will no longer be a workable solution as a cooling mechanism in the near future (RICS, 2011, Intergovernmental Panel on Climate Change, 2012). Therefore, the UK has stronger grounds to act as fast as possible upon global temperature rise.

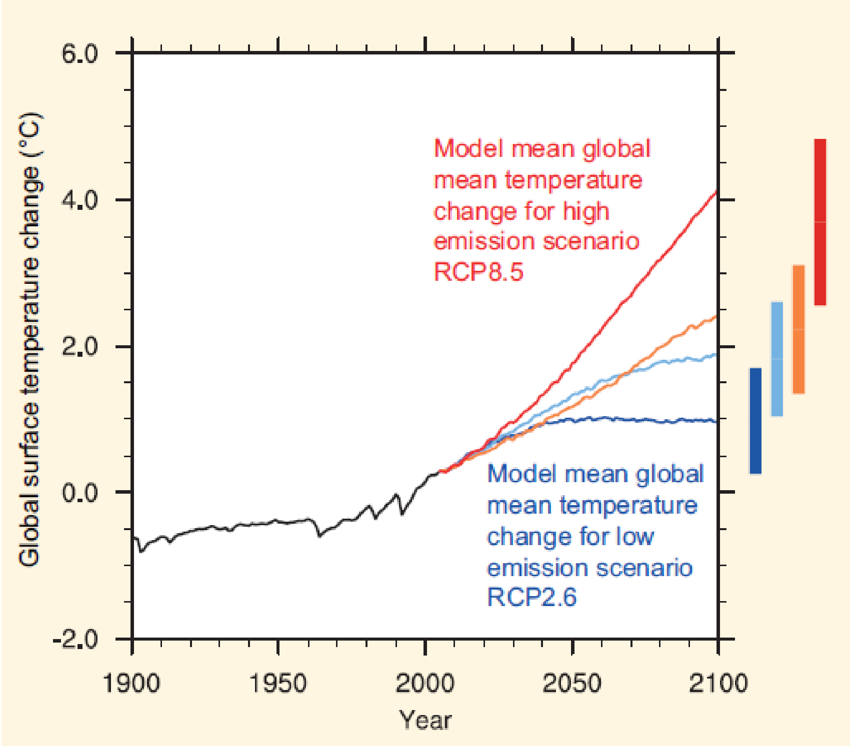


Figure 2.1: Global mean temperature change at four different emission scenarios

Source: Intergovernmental Panel on Climate Change (2013)

The Article 2 of the United Nations Framework Convention on Climate Change (UNFCCC) echoes the need to achieve a stabilisation level for GHGs that would prevent dangerous anthropogenic interference with the climate system (United Nations, 1992). Intergovernmental Panel on Climate Change (2001) projects that an average temperature rise over 2.5°C will have negative impacts on the biodiversity of the earth. Further, the report warned that when the temperature starts to rise over 2°C future warming of the earth would accelerate. According to that, the concentration of GHGs in the atmosphere has to be stabilised to not to cross the threshold of 2°C. Stern (2007) concluded in his review on the economics of climate change that most feasible and effective range of stabilising Carbon dioxide equivalent (CO_2e) levels in the atmosphere would be 450 to 550ppm (particles per million) CO_2e as cutting below 450ppm would be costly and above 550ppm would lead to high climatic risks. Table 2.1 lists the stabilisation levels and their respective temperature thresholds presented. It is almost in line with the prediction of the Stern review, reconfirming that the concentration of GHGs should not exceed 540ppm CO_2e to be within the maximum allowable threshold of 2°C.

Table 2.1: Relationship of atmospheric concentrations of CO_2 to temperature

Source: National Research Council (2011)

Stabilization CO_2-Equivalent Concentration (ppmv): Range and Best Estimate	Equilibrium Global Average Warming (°C)
320 ⟵ 340 ⟶ 380	1
370 ⟵ 430 ⟶ 540	2
440 ⟵ 540 ⟶ 760	3
530 ⟵ 670 ⟶ 1060	4
620 ⟵ 840 ⟶ 1490	5
Note: Green and red numbers represent low and high ends of ranges, respectively; **black bolded** numbers represent best estimates.	

Reduction in the burning of fossil fuels is recommended not only to manage the climate change but also to address scarcity of fossil fuels. Shafiee and Topal (2009) predicted the time depletion for oil, coal and gas to be 35,107 and 37 years, respectively. Findings of Singh and Singh (2012) also coincide with same depletion times predicted by Shafiee and Topal (2009) whilst different predictions are presented in other studies (see for example, Lior, 2008, International Energy Agency, 2006, Eco Info, 2012). However, the literature suggests that only coal will last for the next century and beyond that any activity relying on fossil fuels will be distorted. This gave rise to energy and emission related policies and regulations to meet the future energy demand and combat the climate change, which is also a key conclusion of the 2003 UK White Paper on Energy (DTI, 2003).

2.2.1. Energy and Emission policies

The 2007 UK White Paper on Energy argues that most of the world's carbon dioxide emissions are due to the inefficient production methods and user patterns of the energy. Hence, energy policies having a major role in regulating emission levels (DTI, 2007). Further, Stern (2007) argues that the emissions tend to grow as the global economy grows if stringent policies are not in place. In addition, the reduction of carbon emission levels is the cornerstone of energy policies, without which the policies will fail (The Quality of Life Policy Group, 2007).

Following the first World Climate Conference in 1979, many policies and agreements emerged eventually to manage carbon emissions to combat climate change. The key policies and regulations pertinent to the UK include Kyoto protocol in 1997 (amended in 2015), Part L of Buildings regulations in 2006 (which underwent revisions in 2010, 2013 and expect a further revision in the future), Stern review in 2007 and Climate Change Act in 2008. All of which encompasses high emission reduction targets for the UK. However, a clear vision and roadmap towards the target are important in order to achieve it. The crucial target now for the UK is 80% emissions reduction by 2050 at 1990 levels. Therefore, the UK being a signatory to Kyoto Protocol, it is important to understand the targets of various policies and integrate them into a clear trajectory to achieve them cost effectively.

Figure 2.2 illustrates the key targets of emissions reduction in the UK. Accordingly, all new homes were expected to be zero-carbon from 2016 in the UK until recently, which was abandoned by the UK government and this target is now under review. Further, the inclusion of existing stock in the target remains undecided where the Energy Company Obligation and Green Deals are a few schemes, which were introduced to improve the energy efficiency of existing stock. The next most ambitious target is 2019 zero carbon buildings which is again likely to be scrapped. Then, the 2020 target of the UK Climate Change Act, 30% reductions followed by 50% reduction and 80% reduction in 2025 and 2050 below 1990 levels, respectively (the increasing size of circles in Figure 2.2 represents the increasing emission reduction targets).

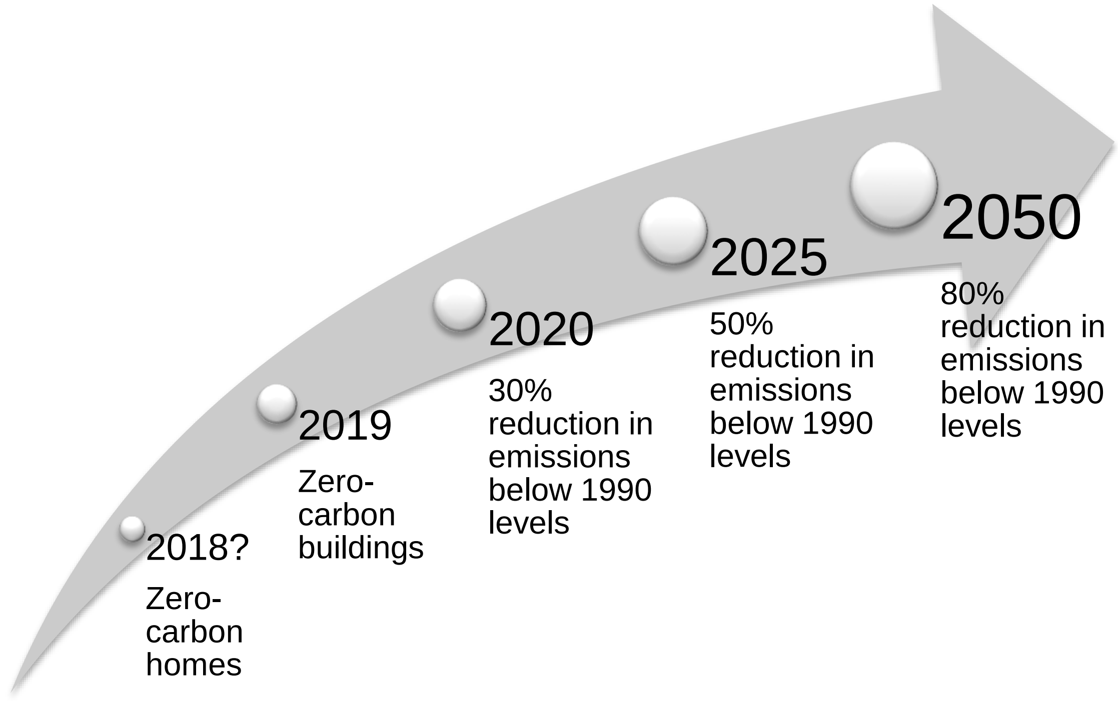


Figure 2.2: Key milestones in the UK carbon roadmap

However, there is a gap between current policy standards and targets (see, Figure 2.3) according to the European Union (EU) Commission's report named 'Communication on the development of a Roadmap for a Low Carbon Economy by 2050' to the Council and the European Parliament. Consequently, the report outlines a strategy to enable and steer the transition and explore the most effective options for "decarbonising" the European economy (European Commission, 2011). Table 2.2 lists the strategies proposed for the building sector to achieve the main and intermediate milestones to achieve the 2050 target as buildings being the focus of the study. The 'Actions' column of the table is separated using broken line to indicate that the actions will be continued in the subsequent years.

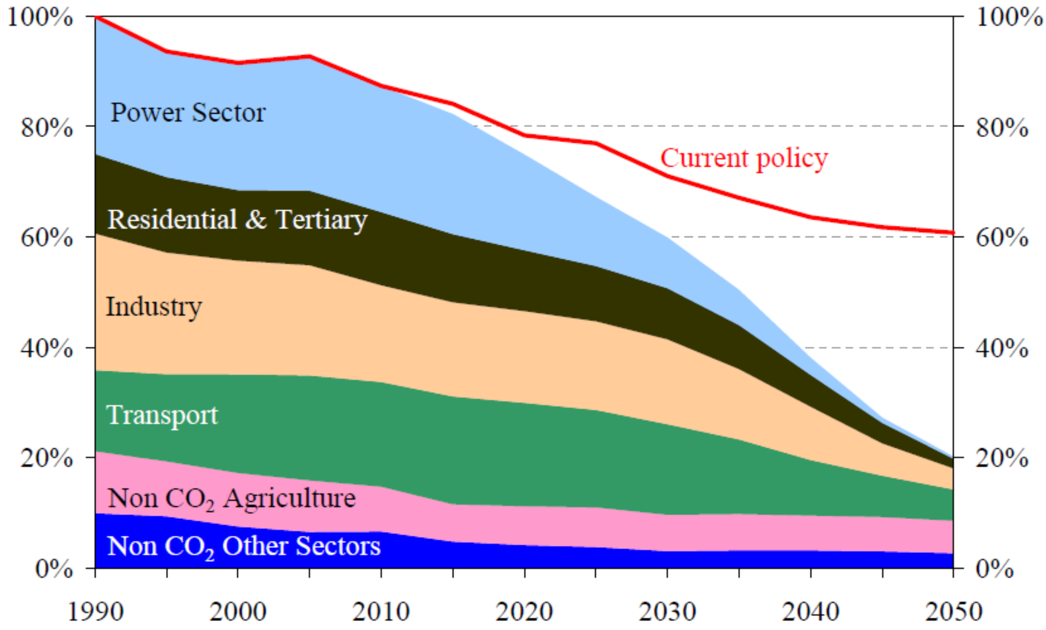


Figure 2.3: EU GHG emissions towards 80% domestic reduction (100%=1990)

Source: European Commission (2011)

All plans and actions mostly focus on energy consumed and carbon emitted during the operational phase of buildings while less attention is given to the emissions associated with the production, maintenance and demolition of buildings. Nevertheless, energy consumption and carbon emissions associated with these stages are receiving significant attention now.

Table 2.2: Detailed trajectory of carbon control in the building sector

Timeline	Target	Actions	Reference
2016	Zero carbon homes	Renewable Heat Incentive, tightening of Part L regulations, green deal, ECO, loft and solid and cavity wall insulation, switching from halogens to Light-Emitting Diodes (LEDs), improvement in efficiency of cold and wet appliances (e.g. fridges and dishwashers)	Gambhir and Vallejo (2011), Committee on Climate Change (2013), Committee on Climate Change (2014), The Green Construction Board (2013)
2017	29% reduction in emissions below 1990 levels – 2nd carbon budget	Solid wall insulation in residential sector (3.5 million by 2030) , investment in a technology portfolio including renewables, nuclear and Carbon Capture and Storage applied to coal and gas, switching from halogens to LEDs, improvement in efficiency of cold and wet appliances (e.g. fridges and dishwashers), 13% of homes to have heat pumps in 2030, changes in ECO and green deal schemes	
2019	Zero carbon buildings	Renewable Heat Incentive (RHI), tightening Part L regulations, Carbon Reduction Commitment, Display Energy Certificates and	

Timeline	Target	Actions	Reference
		Energy Performance Certificates, smart metering, UK Green Investment Bank, Landfill tax escalator, deploy ground source and air source heat, switching from halogens to LEDs	
2020	30% reduction in emissions below 1990 levels	Use of low carbon technologies (e.g. PV) , renewable energy sources, supply chain management	
2022	35% reduction in emissions below 1990 levels– 3rd carbon budget	Install heat networks in dense urban areas, supply chain management	
2027	50% reduction in emissions below 1990 levels – 4th carbon budget	Full ramp up of hard to treat properties	
2050	80% reduction in emissions below 1990 levels	Mandatory measuring and reporting of whole life carbon for all buildings, aligned to carbon budgets, Promoting large demonstration of carbon capture and storage projects	

2.3. Energy and Carbon

Energy is one of the key concerns of the world at the moment. Meeting energy demand is considered a huge challenge, hence, a range of measures exist to regulate the energy demand. It is also predicted that the global primary energy demand will rise by 53%, by 2050 (DTI, 2007). Built environment is one of the main reasons for the rise in the predicted energy demand as the global built environment is responsible for 30-40% of global energy consumption (UK-GBC, 2014b). Hence, Energy Performance of Building Directive (2002) and Part L of the Building Regulations (2013) are in place to control the energy demand in the UK through set benchmarks for building performance.

Energy usage during the operational phase of buildings is significant in the building sector, which is referred to as “Operational Energy” (OE). It is the primary energy (the direct energy used at the source without any transformation or conversion) consumed for space heating and cooling, hot water and fixed lighting. In other words, it is also called the regulated energy as per the UK Building Regulations Part L. Meanwhile, energy consumed in Information and Communication Technology (ICT) equipment, cooking and refrigeration appliances is not regulated or included under the Standard Assessment Procedure (SAP) or Life Cycle Assessment (LCA) calculation which outlines the method of quantifying energy and carbon of domestic and non-domestic buildings (PRé Consultants, 2014, Wan Omar et al., 2014). However, the other part of the energy consumed by the building sector known as “EE” (EE) is not regulated presently.

In response to the requirements of national schemes to assess the environmental impacts of buildings and to prevent any technical barriers to trade within the EU, TC350 Standard was introduced in 2011-12, through a process LCA in which both the OE and EE are included in assessments. In accordance with the TC350 Standards, EE is the total primary energy consumed from direct and indirect processes associated with a building including material extraction, manufacturing, transportation, construction, refurbishment and replacement, and disposal activities at the end of the building’s life. It also includes the impacts from all material that is lost at every stage (Mebratu, 1998).

Table 2.3: A review of EE definitions

	Initial EE	Recurring EE	Demolition Energy	Direct Energy	Indirect Energy	Remarks
Shafiee and Topal (2009)	√	√	√			Demolition energy is included in EE
Brandt (2012)	√	√	√			Demolition energy is excluded from EE
Dixit et al. (2010)				√	√	Direct energy includes construction and assembly on site, prefabrication, transportation and administration Indirect energy includes initial EE, recurring EE and demolition energy

Nevertheless, the definitions on EE are still evolving and it is common to see that EE is further classified as initial EE and recurring EE (and in some instances, demolition energy, see, Table 2.3). Dixit et al. (2010) reviewed various definitions of EE and interpreted EE as the energy consumed during the life cycle stages of buildings such as processes of production, on-site construction, and final demolition and disposal. Dixit et al. (2010) classifies EE as direct and indirect energy where direct energy consists of construction and assembly on site, prefabrication, transportation and administration and indirect energy consists of initial EE, recurring EE and demolition energy. On the other hand, Brandt (2012) interprets EE as the energy consumed during the manufacturing phase of the building where manufacturing phase is perceived as raw material extraction, material production, transportation, construction and renovation. While Brandt (2012) classified EE as the initial and recurring EE, energy sequestered during the demolition and disposal of buildings at the end of the lifespan is excluded from EE and referred to as demolition energy (Brandt, 2012).

Even though there are similarities and contradictions within the above interpretations of definitions, all of the above definitions lack another important phase of the lifecycle, which is gaining importance at present, namely “benefits beyond the life cycle”. This phase includes reuse, recovery and recycle and is considered in TC350 standards (see, Figure 2.4). Recent studies also suggest that this phase should also be taken into account in EE computations, as the end of life benefits might be significant for some projects (Clark, 2013, Wu et al., 2012, Anderson et al., 2002). However, there is a risk of double counting EE during the lifecycle analysis, which should be avoided.

All these studies attempt to define EE with reference to different stages of building lifecycle, which is referred to as the “system boundary” of the analysis. Hence, it is common to see EE studies with different system boundaries. For instance, EE can be quantified in the following ways (Hammond and Jones, 2011, Hammond and Jones, 2008a, RICS, 2014) where “cradle” here is the earth:

- Cradle-to-gate: includes total energy consumed for all the processes from cradle up to the factory gate of the material manufacturing factory.
- Cradle-to-site: includes total energy consumed in Cradle-to-gate plus delivery to the installation site.
- Cradle-to-construction: includes total energy consumed in Cradle-to-site plus the construction.
- Cradle-to-grave: a complete study that includes total energy consumed in Cradle-to-gate, operation and end of life processes.
- Cradle-to-cradle: the process of making a component or a product and converting it into a new component or product of the same or lesser quality at the end of its life

These system boundaries are mapped on to the TC350 (EN 15978:2011) Standard for assessing lifecycle impacts of a product as illustrated in Figure 2.4. This standard is widely accepted for EE calculations and commonly cited in studies related to EE and carbon.

Figure 2.4: Scope of a life cycle assessment based on the impacts agreed by CEN/TC350 and set out in BS EN 15978:2011

Hammond and Jones (2008a) defined EE in the ICE (the very first energy and carbon database of building materials in the UK) as the total primary energy consumed throughout its life cycle including raw material extraction, manufacturing and transport, energy to manufacture capital equipment heating and lighting of factory maintenance, disposal etc. This falls into the cradle-to-grave system boundary. However, Hammond and Jones (2008a) pointed out that it is a common practice to define EE as ‘cradle-gate’. Hence, the definition of EE was revised in the second version of ICE report of Hammond and Jones (2011) as ‘the total primary energy consumed from direct and indirect processes associated with a product or service and within the boundaries of cradle-to-gate’.

Eventually, it can be interpreted from the aforementioned discussion that the OE is the energy consumed during heating, cooling, hot water, lighting and the operations of all other energy appliances. On the other hand, EE of buildings implies the total energy consumed by the building materials which form the building after deducting for any savings (such as sequestration, reuse and recycling), including all the phase from raw material extraction until the demolition and disposal of buildings. This cycle includes material manufacturing (including extraction of raw materials), transportation to site, construction, repair and maintenance, replacement, demolition, disposal, reuse and recycle.

Similarly, the definitions and concepts of carbon are derived from energy. OC is the GHG emissions associated with operational energy consumption while EC is the GHG emissions associated with the EE consumption. EC is also referred to as capital carbon in studies related to infrastructure (HM Treasury, 2013, Ahrens, 2007, Intergovernmental Panel on Climate Change, 2007). In the earlier definitions proposed by Hammond and Jones (2008a), EC was interchanged with EE. However, the definition of EC was modified in the ICE version 2.0 as the sum of fuel related carbon emissions (i.e. EE which is combusted – but not the feedstock energy which is retained within the material) and process related carbon emissions (i.e. non-fuel related emissions which may arise, for example, from chemical reactions) which can be measured from cradle-to-gate, cradle-to-site, or cradle-to-grave. (Hammond and Jones, 2011). Based on this discussion, energy and carbon are compared and contrasted in Table 2.4.

Table 2.4: Energy vs. Carbon

	Energy	**Carbon**
Embodied	Total primary energy consumed from direct and indirect processes throughout the life cycle of a product excluding operational or use phase.	Net carbon emissions resulting from EE consumption and chemical processes (after deducting any emissions sequestrations) during the life cycle of a product excluding operational or use phase.
Operational	Total primary energy consumed during the operational phase of a facility.	Total carbon emissions resulting from operational energy consumption of a facility.

Therefore, there is a close relationship between energy and carbon. Both energy and carbon can be classified as EE/carbon and operational energy/carbon. In fact, both terms can be interchanged and energy can be interpreted as carbon in most occasions. However, Lélé (1991) noted that EE and EC are improperly interchanged. Lélé (1991) argues that operational energy can be interchanged with OC while EE cannot be interchanged. This is because OC is roughly proportional to operational energy and the magnitude depends on the type of energy (or fuel) whereas EC cannot be directly interchanged with EE at all times as material production processes emit or sequester carbon. For example, cement production emits about half of its EC during the chemical process and the timber sequester carbon during its growth (Ayaz and Yang, 2009). Therefore, it is important that a clear distinction be maintained when attempting to interchange carbon in the place of energy. With this understanding, operational and EC literature is reviewed in the context of buildings in the following sections.

2.4. Operational Carbon in Buildings

OC of buildings gained substantial attention from building owners, construction professionals and regulatory bodies as the operational emissions were higher than the embodied emissions and it was identified that operational emissions account for nearly 70-80% of total emissions from buildings (RICS, 2012b, Anderson, 2011). However, the percentage contribution of OC varies for different types of buildings as shown in Figure 2.5. Accordingly, less energy intensive buildings like warehouses need considerable attention during other phases (i.e. EC emission).

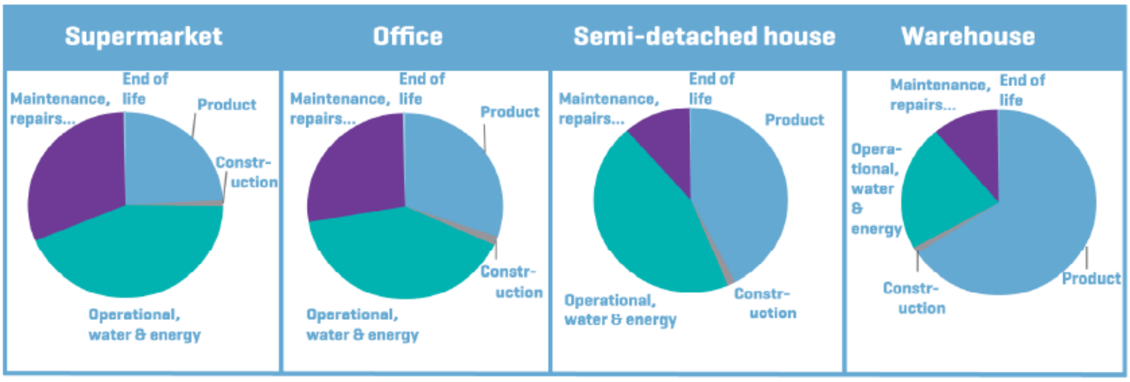


Figure 2.5: Carbon emissions in different phases of buildings' life in different types of buildings

Source: RICS (2014, p. 10)

RICS (2014) defines OC in buildings as emissions related to energy consumption during the operation of buildings. These emissions include both regulated load (e.g. heating, cooling, ventilation and lighting) and unregulated/plug load (e.g. ICT equipment, cooking and refrigeration appliances). The Part L of Building Regulations has provisions of controlling regulated OC in buildings as the unregulated emissions are entirely depended on the occupants' behaviour of the building.

As per the Part L of the Building Regulations, the operational emissions or the Target CO_2 Emission Rate (TER) for a notional building design is calculated using either the Simplified Building Energy Model (SBEM) or other approved software tools where actual the Building CO_2 Emission Rate (BER) should be less than the TER for the building design to be approved. The operational emissions are expressed in mass of CO_2 emitted per year per square meter of the usable floor area of the building (kg/m^2/year) (see, Section 3.2 and 3.5.1 for more details on OC estimating).

Furthermore, zero carbon agenda has increased the concern on EC emissions, because, theoretically total emissions will be equal to the total EC emissions in a zero carbon building (see, Figure 1.1). Therefore, EC emissions require special attention in a low or zero carbon environment and need to be controlled to attain the Kyoto goal of 80% reduction by 2050 and a carbon free economy in long run.

2.5. Embodied Carbon in Buildings

Reviews on definitions and interpretations of EE and carbon highlight the variations in the type of energy considered (primary energy – fossil fuels such as coal, oil and gas; and renewable energy like wind, waves, bio fuels etc.) and the scope or system boundary defined (Ibn-Mohammed et al., 2013; Dixit et al., 2012). Dixit et al. (2012) question the inclusion of only non-renewable energy sources in the EE calculation, which can be answered based on the definition of EC. It is the carbon emitted as a result of the fuel consumption and thus, it is sensible only to include fuel related energy consumption (or emissions) in the definitions. Further, EC can be calculated from cradle (earth)-to-gate (material manufacturing factory gate), cradle-to-site (construction site), cradle-to-end of construction, cradle-to-grave (demolition), or even cradle-to-cradle (includes savings from reuse, recovery and recycle) which is termed the system boundary of the EC calculations as discussed in Section 2.3. System boundary can be selected based on the needs of the beneficiary (or the client); therefore, it is unwise to confine the definition by including system boundary. However, the revised definition proposed by Hammond and Jones (2011) defines EC as "the sum of fuel related carbon emissions and process related carbon emissions" resolving the above-mentioned dilemmas.

The scope of EC is illustrated in Figure 2.6. EC can be categorised into mainly two types: initial EC and recurring EC (Chen et al., 2001a, Ramesh et al., 2010). Initial EC is the emissions associated with the production of the building including raw material extraction, manufacturing, transport and construction; recurring EC includes emissions during use of the building such as repair, maintenance and replacement due to the difference in the life spans of building elements and the overall building. In addition, a third type called demolition EC or end-of-life EC that includes emissions associated with the demolition of the building. Furthermore, EC saved because of reuse, recovery and recycle at the end of buildings' life cycle are referred to as the 'benefits beyond system boundary' (as identified in TC350, BS EN 15978 standard). A cradle-to-cradle system boundary includes the embodied impacts from all of the four types discussed above (please note that the length of the arrows in Figure 2.6 does not represent the magnitude of EC emissions).

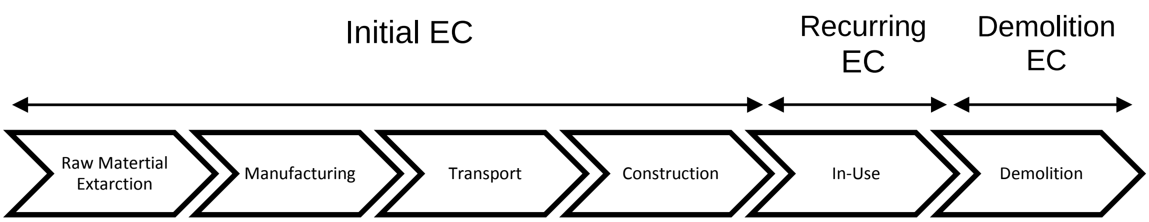


Figure 2.6: Scope of EC in a building life cycle

Hammond and Jones (2011) and Sansom and Pope (2012) noted that many EC datasets available are cradle-to-gate and fail to include emissions from latter stages of life cycle (such as construction, operation and maintenance and demolition and disposal) due to project specific emissions (see, Section 3.6 for the discussion on EC databases). However, transport of materials to the site can be significant for materials with lower EC emissions in other phases (Hammond and Jones, 2008a). Furthermore, lesser transport distance not necessarily means lesser carbon emissions; mode of transport and type of fuel also plays a significant role other than the distance of travel (RICS, 2014, Sundarakani et al., 2010).

Figure 2.7 summarises the EC values of different types of buildings obtained from various studies. It should be noted that the EC values presented in the graph is for the building structure only. The values of semi-detached houses were obtained from Hacker et al. (2008) and Monahan and Powell (2011). A two storeyed semi-detached house was studied in both cases and alternative structural options were simulated to analyse the impact of design decisions on the EC of the building. Both studies concluded that the EC of the residential building increases when moving form a lightweight timber framed building to a heavy weight concrete building and proved that the EC can be reduced by 51% from the structure of the building alone. The EC values of other types of buildings were obtained from a study conducted by Sansom and Pope (2012) which again includes the EC analysis of the structural form of the buildings. Single case studies were employed for each type of the building and the impact of alternative structural forms on the EC of each building was studied. Further, Sansom and Pope (2012) adopted a cradle-to-grave system boundary which includes the emissions associated with the raw material extraction up to the demolition of the building (however, the study excluded recurring EC which covers repair, maintenance and replacement during the use phase of the building). Estimating EC using cradle-to-grave approach provides a more holistic view though cradle-to-grave EC analysis is hugely influenced by project specific assumptions.

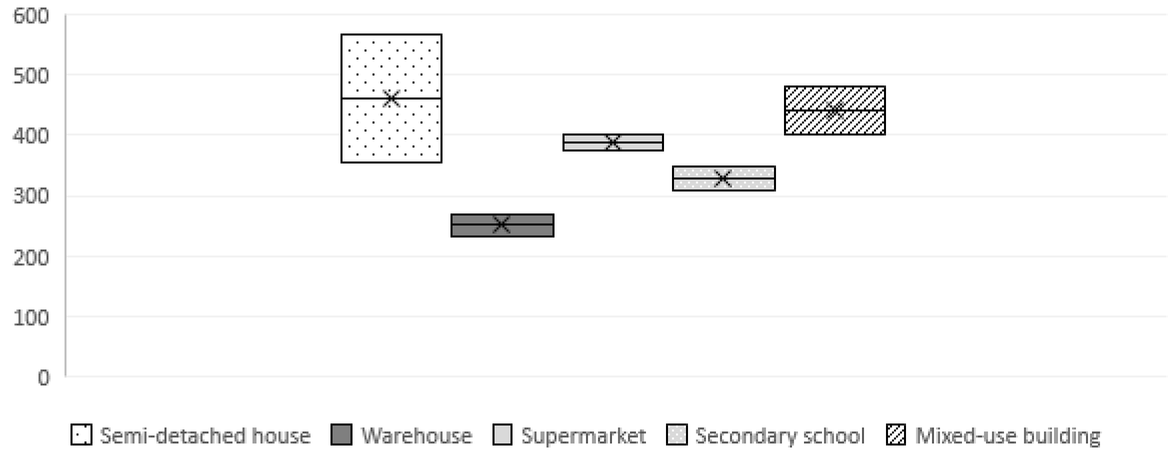


Figure 2.7: EC values of different types of buildings from the literature

On the other hand, EC studies on office buildings are prevalent in the literature. Hence, EC values of office buildings are presented separately in Figure 2.8. Findings of four studies were mapped onto a spider web diagram to demonstrate the variation in the EC values of office buildings. Clark (2013) reported EC analyses of office buildings ranging from low to high-rise buildings, structure only analyses to whole building analyses and cradle-to-gate analyses to cradle-to-grave analyses. Hence, the reported EC values range from 300 $kgCO_2/m^2$ to 1,650 $kgCO_2/m^2$. As explained before, the findings of Sansom and Pope (2012) covers cradle-to-grave EC analysis of the structure of an office building excluding recurring embodied emissions. The change in the EC influenced by the change in the structural form of the building was investigated by Sansom and Pope (2012). Hence, the variation is small and it was shown that 11% reduction in the EC is achievable (structure only) in that particular building. Victoria et al. (2015) reported cradle-to-gate EC analyses of seven office buildings and the EC ranges from 271 $kgCO_2/m^2$ to 706 $kgCO_2/m^2$. However, the EC values reported by Victoria et al. (2015) excludes some of the major building services, hence, not holistic. Halcrow Yolles (2010b) studied three low-rise office buildings within a cradle-to-gate system boundary. The EC of the three office building ranges from 538 $kgCO_2/m^2$ to 924 $kgCO_2/m^2$ (excluding major building services). Further, Halcrow Yolles (2010b) found that improvement to operational energy can escalate the EC of up to 25% (Halcrow Yolles, 2010b).

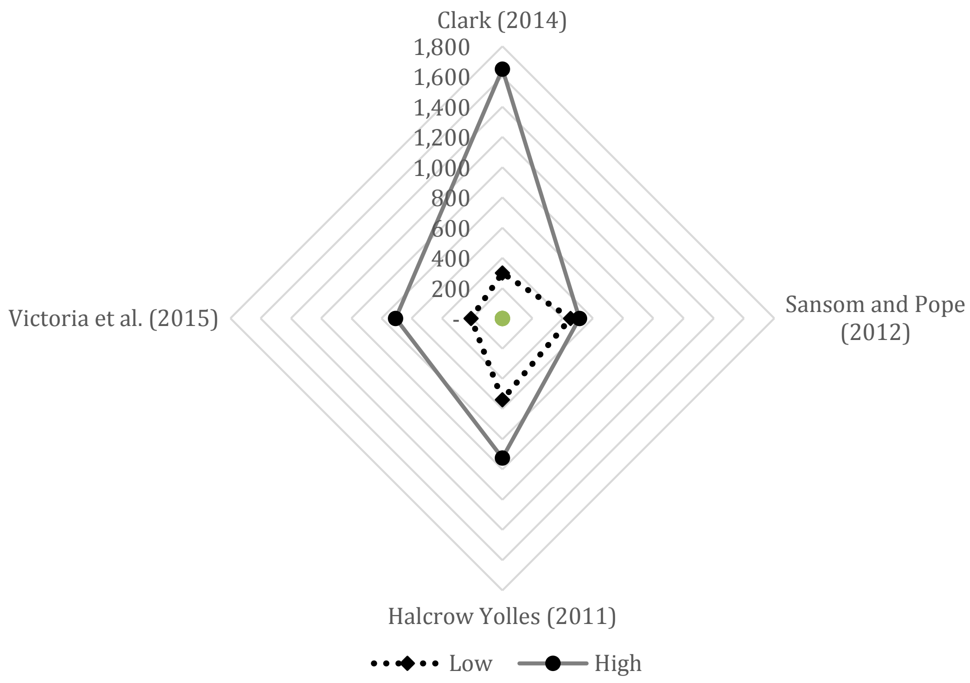


Figure 2.8: EC studies on office buildings

EC management requires a great deal of understanding and attention to detail. Measures to minimise EC of the building has to be taken during the early stages of the design to yield greater savings as the carbon reduction potential is high during the early stages of design (RICS, 2014). Figure 2.9 illustrates the diminishing reduction potential of EC over the project life cycle while approximately 80% of initial EC committed by the end of the design phase (Asiedu and Gu, 1998). As more carbon is committed into the project, the reduction potential decreases increasingly because possible design solutions are constrained by previous design decisions. Then, during the construction phase, the reduction potential can be regarded as nearly zero unless there is a design change. Further, the design becomes static as the project progresses and changing the design at a later stage will result in loss of time and money.

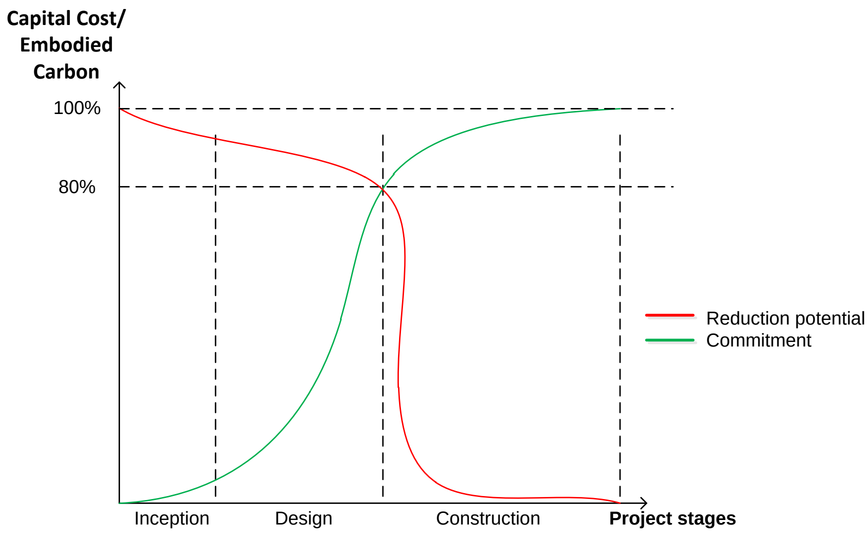


Figure 2.9: Behavioural pattern of EC over project stages

After: RICS (2014) and Halcrow Yolles (2010a)

In fact, RICS (2014) states that investigating EC emissions in different types of buildings is a completely new research avenue, and there are limitations in regulatory standards or academic research to aid decision-making at the initial stages of projects. Nevertheless, carbon hotspots are identified as an ideal way of dealing with this issue.

2.6. Carbon Hotspots

A hotspot may mean different things to people from different disciplines. RICS (2014) defines ‘carbon hotspot’ as the carbon significant aspect of a project, which can be building elements, or other aspects in the supply chain. However, carbon hotspots in this research refer to the carbon critical or significant building elements. RICS (2014) further extends that carbon hotspots are not only carbon-intensive elements but also the elements that are easily measurable and carbon reduction is possible. The Pareto Principle proposes that 80% of the results (or consequences) are attributable to 20% of the causes which implies an unequal relationship between the inputs and the outputs (Koch, 2011, Delers, 2015). According to 80:20 Pareto rule, it can be assumed that 80% of embodied emissions are caused by 20% of building elements (also see, Section 4.9.1 for the application of the Pareto theory in the research). These carbon hotspots may vary from one building to the other depending on the type or function of the building (Ashworth & Perera, 2015).

Monahan and Powell (2011) highlighted the importance of identifying hotspots in buildings by modelling a two-storeyed residential building (in the UK) in three different scenarios – timber frame and larch cladding, timber frame and brick cladding, conventional masonry cavity wall. The substructure (including foundation and ground floor) accounted for 50% of EC in timber frame and larch cladding building and substructure, external walls and roof were identified as the carbon hotspots in the building (elements responsible for 81% of EC, however, not all the building elements were included in the accounting). Further, the same building (timber frame with larch cladding) substituted with timber frame and brick cladding and conventional masonry resulted in additional EC of 32% and 51% respectively. The majority of the difference in EC was found to be attributed to the difference in foundations and external walls. The findings of the study (Monahan & Powell, 2011) reveal substructure and external walls as 'carbon hotspots' in the particular residential building and highlight the potential for EC reduction.

Shafiq et al. (2015) studied a two-storied office building in Malaysia by modelling six different scenarios for structural composition using a Building Information Model (BIM). However, Shafiq et al. (2015) used UK databases to estimate EC due to lack of EC databases in Malaysia. Different grades or classes of concrete and steel were combined to generate different composition, which resulted in different material quantities producing varying EC impacts. Only a few elements were studied including foundation, beams, slabs, columns and staircases, which can be related to the substructure, frame, upper floors and stairs as per the New Rules of Measurement (NRM) element classification. Shafiq et al. (2015) found that it was possible to reduce up to 31% of EC by designing these elements with different classes of concrete and steel to meet the given design criteria. However, it should be noted that only the elements that constitute concrete and steel are considered because concrete and steel are considered as the main structural building materials and emit high EC during production. Particularly, upper floors were identified as the key carbon hotspot followed by substructure, frame and stairs.

It is clear that EC studies in different types of buildings highlighted above (Monahan & Powell, 2011; Shafiq et al, 2015) have different focuses and hence, limit the analysis to few elements. However, an analysis of the whole building will provide a holistic picture on the EC contribution of each element and will highlight the potential areas for carbon reduction. Generally, floors (ground and upper floors), frame, external wall and roof are identified as carbon hotspots in buildings (Clark, 2013; Davies, Emmitt, & Firth, 2014; Halcrow Yolles, 2010a). It was noticed that the element classification differs from one study to the other due to incompatible element classification standards (for example, NRM, Standard Method of Measurements (SMM)/ Building Cost Information Services (BCIS) - older version, British Council of Offices 2011, some studies did not follow any standards). Therefore, literature findings were organised in accordance with the NRM element definition, which are presented in Table 2.5. Most studies lack transparency of the methodology adopted in the study. This questions the validity and applicability of those findings.

Table 2.5: Carbon profile of building elements of office buildings from published studies

<table>
<tr><th rowspan="2"></th><th colspan="4">Published Studies</th></tr>
<tr><th>Halcrow and Yolles (Average of 3 case studies)</th><th>Sturgis Associates</th><th>WRAP</th><th>Davis Langdon (30 case studies) from Clark (2013)</th></tr>
<tr><td>Substructure</td><td rowspan="3">89% (some elements are combined)</td><td>25%</td><td>18.3%</td><td rowspan="2">Structure - 45%-85%, Facade - 5%-25%</td></tr>
<tr><td>Superstructure</td><td>56%</td><td>58.24%</td></tr>
<tr><td>Internal Finishes</td><td rowspan="3">Fit-out (shell & core) - 8%, Fit-out (Cat B) - 8%</td><td>8.619%</td><td rowspan="2">4%-25% (Internal walls included)</td></tr>
<tr><td>Fittings & Furnishings</td><td>Not given</td><td>Not given</td></tr>
<tr><td>Services</td><td>3%</td><td>11.96%</td><td>2-25%</td></tr>
<tr><td>Others</td><td>8% (External works)</td><td>4% (Waste)</td><td>2.9% (External works)</td><td></td></tr>
</table>

Even though services account for 10-25% of total EC emissions, it is not widely considered as a carbon hotspot due to difficulty in the measurement of services during the early stages of design and lower EC reduction potential (Hitchin, 2013; RICS, 2014). However, Cole and Kernan (1996) found that cladding finishes and services are the biggest components of recurring EC emissions of an office building in Canada. Especially, it was highlighted that in a 50-year life cycle, recurring EE is almost the same as the initial EE and for a longer life cycle it would be greater than the initial EE (However, the findings were subject to the assumptions and energy data available at that time). Hence, the quality of services and finishes cannot be disregarded when making initial design decisions, as the contribution is significant. Therefore, it is important that an indication of the likely EC of building services and finishes be given at the early stages of design to understand the total carbon accountability of the building.

Hitchin (2013) investigated services EC for a typical office building in London and found that ‘space heating and air treatment’ and ‘electrical installations’ were the most EC intensive building services. However, this is an incomplete picture painted by most scholars due to limiting the building services EC analysis to only fundamental services such as water, sanitary and drainage installations, electrical and HVAC installations. Sophisticated services such as communication installations and building management system are excluded in most studies due to their complex nature and limited EC data. However, these services constitute cost significant items in office buildings.

Clark (2013) proposed benchmarks for EC values of a typical UK office building based on findings of a range of reported studies, which are listed in Table 2.6. However, Clark (2013) admits that the proposed benchmark values are not subject to a detailed scrutiny and it should only be regarded as a rule of thumb for EC calculations of office buildings in the UK. Further, Clark (2013) insists that further research is needed to develop robust carbon benchmarks. The lack of scientific evidences on EC hotspots of buildings and the non-conformity of existing EC analyses to a standard element classification drive this research and highlight the knowledge gap in the existing body of literature.

Table 2.6: Indicative EC values for the UK office buildings

Source: Clark (2013)

Components	**Indicative EC ($kgCO_2e/m^2$ GIFA)**		
	Typical design	**Low carbon design**	**High carbon design**
New build (shell and core)	600	400	900
Fit-out (Category A)	400	70	150
Fit-out (Category B)	200	100	300
Minor refurbishment (excluding fit-out)	25	15	40
Major refurbishment (excluding fit-out)	100	70	150
Reclad	100	70	150
Demolition and disposal	30	30	30

Given that there is no empirical evidence on the carbon hotspots of buildings, it can be argued that the hierarchy of building elements in terms of carbon intensity will change for different types of buildings/projects due to different element intensities. Dixit et al. (2010) identified a list of factors that affects the EC measurements. However, diversity of assumptions, the source of EC data and the methodology adopted (Clark, 2013) can be regarded as the most significant factors for the reported variations. Furthermore, element classification also highly alters the findings of the studies. Especially, analysis of EC of building services remains a mystery due to lack of comprehensive published dataset and hence, services represent a small percentage in some of the reported studies (see, Table 2.5).

2.7. Operational Vs. Embodied Carbon in Buildings

Figure 2.10 illustrates the contribution of operational and embodied (capital) carbon in different sectors of the built environment (The Green Construction Board, 2013). The Figure has been derived using the operational energy consumption data from the Digest of UK Energy Statistics for each sector (which is produced by the Department for Energy and Climate Change) which has been converted into emissions by applying appropriate GHG emission factors from DEFRA. On the other hand, capital carbon has been estimated using a Multi Regional Input Output model, which is the most comprehensive inventory of historical annual data on capital carbon, though it does not provide sectoral distributions. Hence, data from the Office for National Statistics on construction outputs by sectors has been used to decide on the percentage contribution of each sector owing to the fact that there is a relationship between the value of the construction output and the embodied emissions (The Green Construction Board, 2013). Accordingly, the domestic sector contributing more than 50% of the operational emissions highlights the importance of the 'zero carbon home' target. Similarly, the zero carbon target for non-domestic buildings is also considered equally important, as the non-domestic building stock is accountable for one-fourth of the total emissions during the operational phase according to Figure 2.10. Given that the targets are achieved by 2019 (even though these are ambiguous and questionable now), the remaining component of carbon to be managed will be EC of domestic and non-domestic buildings and infrastructure. Further, it is clear from Figure 2.10 that non-domestic capital carbon (EC) is higher than the domestic and infrastructure EC.

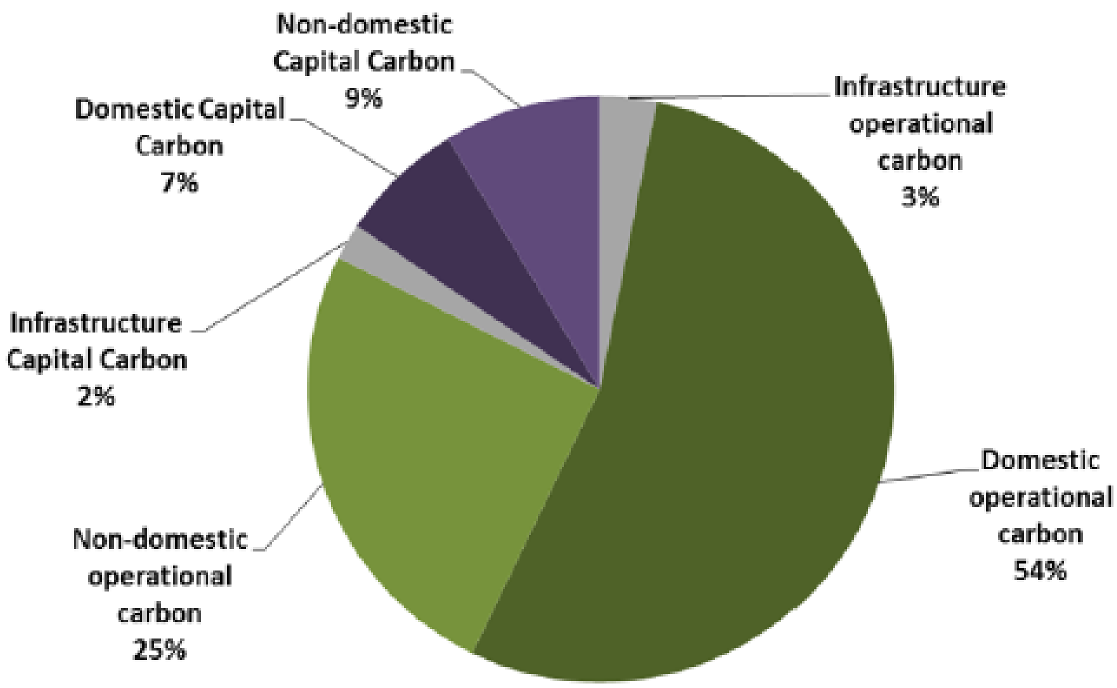


Figure 2.10: Breakdown carbon emissions in built environment 2010

Source: The Green Construction Board (2013)

Figure 2.11 demonstrates the contribution of embodied and operation carbon in buildings and infrastructure in different countries. The reasons for varying proportions may include and not limited to the type of building being assessed, the use of the building, the type of building materials used, construction methods employed, the period of analysis considered and geographical differences. Therefore, it is important to be aware of these variations when developing models and benchmarks so that these variations be normalised. Further, it is important to understand the relationship between operational and EC since both are interdependent. Ramesh et al. (2010) noted that operational energy/carbon has been drastically reduced in low-energy buildings and EE/carbon increases as the operational energy/carbon decreases. The Green Construction Board (2013) further extends that there is a strong linkage between operational and EC as the efforts to reduce OC tends to increase EC and vice versa. This affirms the inverse relationship noted by Ramesh et al. (2010). Moreover, in a zero carbon environment, EE/carbon can be expected to be higher than in low carbon buildings, which necessitates the need for controlling the EC instantly.

Country	Author	Relationship between embodied and operational emissions in different buildings and infrastructure
UK	Lee & White (2008)	Embodied energy is 3-35% of 100 year life-cycle energy demand
	Yohanis & Norton(2002)	Embodied energy is 67%of operational energy over a 25year period
	Eaton &Amaton (2005)	Embodied carbon is 37-43%of 60 year life-cycle carbon
	Smith (2008)	Up to 80% of life-cycle carbon emission is embodied carbon
	CIBSE (2010)	Embodied carbon is 42-68%of 60 year life-cycle carbon
US & Canada	Engin& Francis (2010)	Embodied energy is 11-50% of 60 year life-cycle carbon emissions
	Webster (2004)	Embodied energy is 2-22%of 50 year life-cycle energy demand
	Athena (2007)	Embodied energy is 9-12%of 60 year life-cycle energy demand
	Build Carbon Neutral (2007)	Embodied energy is 13-18%of 66 year life-cycle energy demand
Australia	CSIRO (2006)	Over 10%of 100 year life-cycle energy demand is embodied carbon
Sweden	Thormark (2002)	Embodied emission is 45% of 50 years life-cycle emissions
Israel	Huberman &Pearlmutter (2008)	Embodied emission is 60% of 50 years life span
Key:		Emissions Embodied Operational

Figure 2.11: Operational vs. embodied energy and carbon emissions in different countries

Source: Ibn-Mohammed et al. (2013)

2.8. Zero Carbon Buildings

Zero carbon policy for homes simply means that all emissions arising as a result of the operational energy use of buildings on site should be offset by possible means (Zero Carbon Hub, 2014b). Figure 2.12 illustrates the concept more clearly. According to Figure 2.12, three key criteria should be met for a home to be regarded as a zero carbon home. Firstly, energy efficiency should be achieved at the minimum specified standard (Fabric Energy Efficiency Standard (FEES)) thorough fabric performance; secondly, the remaining carbon emissions should not exceed the carbon compliance level set in Part L1; finally, after meeting the first two requirements (i.e. carbon compliance) the remaining emissions should be offset to reach the zero carbon standard (Zero Carbon Hub, 2014b). Further, the zero carbon target for homes in the UK was set as 2016 and it was expected that all new homes should be zero carbon from 2016 though it went through several twists and turns in the recent past and now under another review to push the deadline forward to 2018.

A survey and semi-structured interviews conducted by Osmani and O'Reilly (2009) with major house builders in the UK identified a number of barriers to achieve the target of zero carbon homes by 2016. These barriers include legislative (lack of clarity in requirement and expected outcomes in the policy), cultural (lack of customer demand), financial (lack of data on cost of achieving zero carbon homes) and technical barriers (moving from conventional design and technology, though it was the considered as the least significant barrier). Further, the study (Osmani and O'Reilly, 2009) pointed out that even though the target of zero carbon homes seems technically feasible, it requires proper strategies and a plan in place and effective implementation. Apparently, the zero carbon target for homes has not been achieved yet and is still under review.

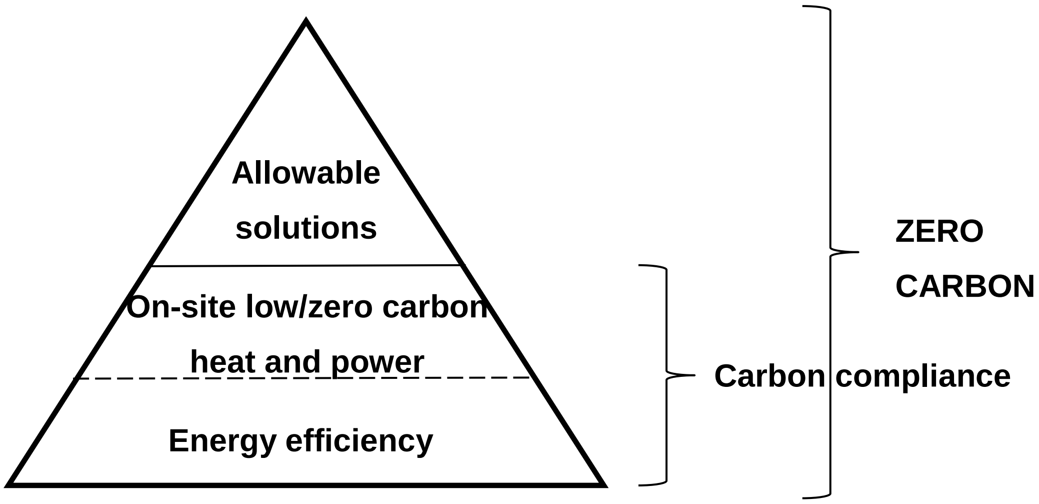


Figure 2.12: Zero carbon policy for homes

Source: Zero Carbon Hub (2014b)

While there is no firm definition or policy for zero carbon non-domestic buildings due to varying energy usage pattern for different types of buildings (especially unregulated), government has stated in a consultation report (Department for Communities and Local Government, 2008) that the minimum compliance for non-domestic buildings will be all regulated emissions should be zero. Even though the target for non-domestic zero carbon buildings was set as 2019 the industry assumes that this target has also been scrapped similar to the target of zero carbon homes imposing a challenge on 2050 emission reduction target.

The Zero Carbon Hub (2013) has proposed three strategies to comply with zero carbon definition for homes, namely, Balanced, Extreme Fabric and Extreme Low Carbon Technologies. These imply meeting the minimum standard for FEES and carbon compliance through a moderate but sensible focus on low-carbon technologies; reaching extremely high FEES and less on-site low carbon technology; fabric performance beyond FEES and maximum use of on-site low/zero carbon technologies to reduce emissions well beyond carbon compliance, respectively. On the other hand, meeting non-domestic zero carbon lacks strong stance due to the absence of a firm definition.

It is also interesting to see another concept of Nearly Zero-Energy Buildings (NZEB) from 2020 in EU in accordance with Energy Performance of Buildings Directives (EPBD) Article 2, alongside with zero carbon (or energy) buildings. EPBD defines NZEB as a building which has very high energy performance that produces the energy it required from renewable sources on-site or nearby. The major difference between NZEB and zero carbon buildings concept is the metric. FEES is measured in kWh/m^2/year energy demand, carbon compliance in kg/m^2/year of CO_2 and allowable solution in £s whereas NZEB is measured in primary energy consumption units, kWh/m^2/year. Furthermore, zero carbon policy aims at delivering zero carbon domestic buildings from 2016 whereas NZEB policy focuses on all types of buildings. The lack of a proper zero carbon definition is the reason for this variance (Zero Carbon Hub, 2014a).

It is apparent that definitions and solutions to achieve zero carbon buildings totally focus on OC of the buildings. However, the need to incorporate EC within the zero carbon definition has been manifested in many studies. A recent case study of a residential building in Norway found that even though the building was constructed with energy efficient building envelope and PV system, the net environmental impact after considering EC was inadequate (Lützkendorf et al., 2014). Further, a few studies and consultation reports have reviewed existing definitions (Marszal et al., 2011, McLeod et al., 2012, UK-GBC, 2008) and proposed that embodied impacts should be included in the zero carbon definition (Lützkendorf et al., 2014, Hernandez and Kenny, 2010, McLeod et al., 2012, UK-GBC, 2014a).

Figure 2.13 illustrates the state of different buildings, ranging from conventional, low and zero energy to energy producing buildings, against a life cycle zero energy building in a real sense (this can be applied to carbon as well) (Hernandez and Kenny, 2010). It is clear from Figure 2.13 that buildings referred to as “energy producing buildings” are not even close to a real zero energy/carbon building. This raises a serious concern about the zero carbon definition.

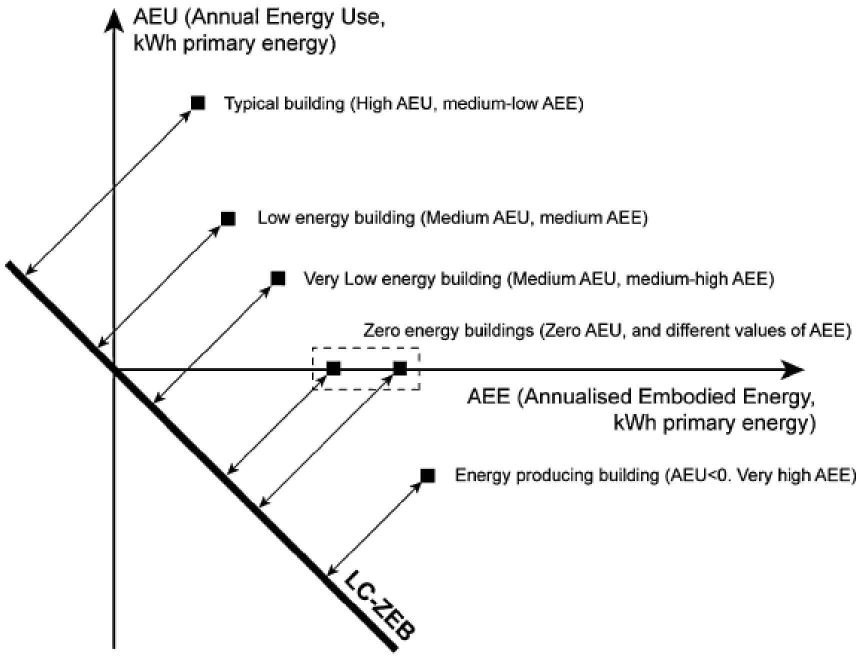


Figure 2.13: Annualised life cycle energy vs. annualised embodied energy of generic buildings against life cycle zero energy buildings

Source: Hernandez and Kenny (2010)

All these emerging studies and recommendations increase the awareness on embodied impacts of buildings and highlight the necessity to incorporate EC within the zero carbon definition. Therefore, it can be expected that embodied impacts could be incorporated into the zero carbon definition in the near future. In that case, designers will be under the pressure to design buildings with minimal or zero embodied impacts. Hence, choosing energy/carbon efficient designs during the early stages of projects will play an important part in reaching the zero carbon targets.

Even though the carbon accountability of projects is considered important due to the climate risk, the cost of construction projects is also of a greater concern for clients who initiate these projects. Therefore, it is important to strike a balance between carbon and cost when selecting a design to satisfy the needs of both the planet and the client.

2.9. Carbon and Cost

An often criticised fact about low and zero carbon buildings is that the cost of achieving it. Developing zero carbon designs used to be a bottleneck to designers due to the requirement of advanced technologies and high CC involvement (Catto, 2008). However, now designers are handling this challenge tactfully and inventing intelligent technologies to design a passive design with active solution to address climate changes and it is believed and proved that low and zero carbon buildings are attainable at an efficient cost similar to conventional buildings (Sturgis and Roberts, 2010, Target zero, 2012, CB Richard Ellis, 2009) or at a little higher cost. A recent study by the Sweett group (Zero Carbon Hub and Sweett, 2014) validates the above claim by modelling different house types and the findings suggest that zero carbon homes can be achieved at an additional cost of between £34/m^2 and £53/m^2 by 2020. Moreover, a recent case study on a commercial building (Torcellini et al., 2014) proved that zero energy building can be attained at no additional cost when best cost controlling practices are implemented. Therefore, the cost can no longer be a barrier to the development of zero carbon buildings.

However, it is not easy to attain a low level of carbon emissions at an efficient cost. It demands expert knowledge and structured decision-making. For instance, selection of a carbon efficient material might increase the cost while not reducing the carbon significantly. Therefore, the decision to choose such a material would not yield the desired value for money. Hence, taking a crucial decision on designs requires expert knowledge and information.

Ibn-Mohammed et al. (2013) explain this issue through a Marginal Abatement Cost Curve (MACC) illustrated in Figure 2.14. Accordingly, the width of the bar indicates the net savings in emissions and the height of the bar denotes the cost per unit of CO_2e saved. Consequently, it can be seen from the graph that most likely the height of the bars increases (cost per net CO_2e emissions savings) as the width (net CO_2e emissions savings) increases, which means both operational emissions savings and embodied emissions incurred increase. EC emissions increases as the operational emissions savings increase. This supports the claims of Ramesh et al. (2010) and the Green Construction Board (2013). On the other hand, the more important relationship between emission savings and cost is established through this graph. As the magnitude of emissions saved increases, the costs also increase (see the bars, E, F, and H). It should be noted that this discussion applies only to the positive cost curves due to the perverse behaviour of MACC (Ibn-Mohammed et al., 2013).

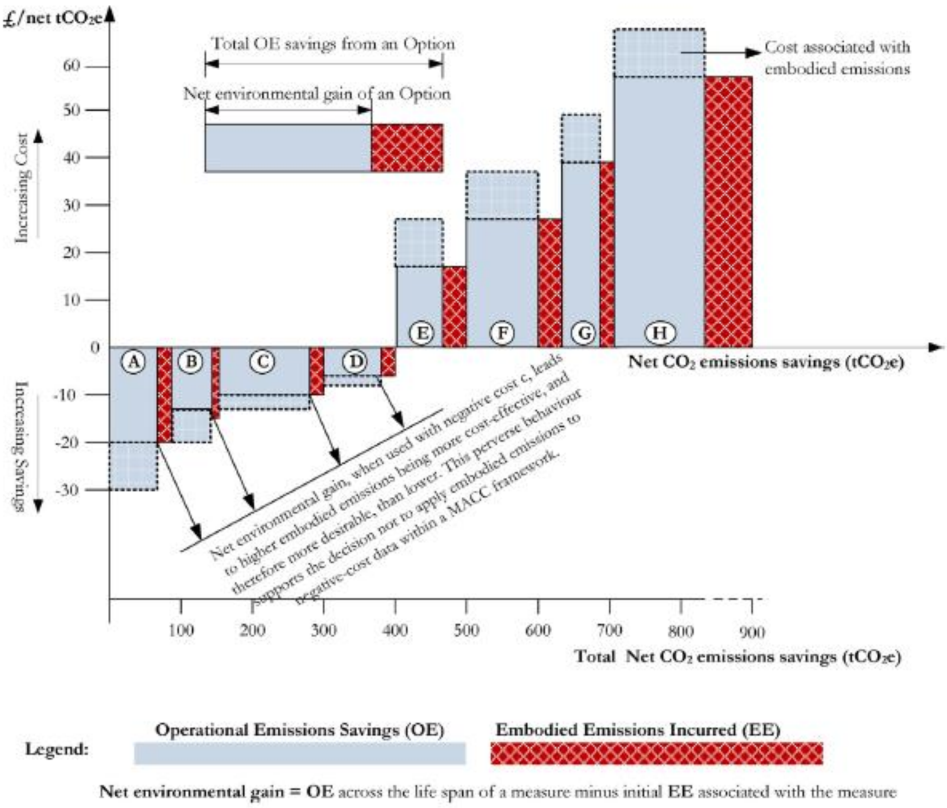


Figure 2.14: MACC integrating economic considerations (positive measures only)

Source: Ibn-Mohammed et al. (2013)

On the other hand, Langston and Langston (2008) studied the relationship between EE and CC at various levels of details (such as projects, elemental groups, elements and selected items of work) with the goal of predicting EE based on CC. Langston and Langston (2008) found a strong positive correlation between EE and CC of the buildings. However, this relationship is most likely to have caused by a third variable (GIFA) which was responsible for the causality between the two variables (EE and CC). However, this third variable was not explored in the study of Langston and Langston (2008). The authors also noted that the correlation between EE and CC drops as the level of detail increases from project level to individual work item level. This means that all work items collectively at the project level demonstrates a correlation (between EC and CC) rather than individually which indirectly conveys that differences in rates (cost and energy) of work items are neutralised when analysed at the project level.

Even though there is a close association between EC and EE, both are distinguishable and cannot be interchangeable (see, Section 2.3). Therefore, EC and CC relationship could be different to EE and CC relationship. Further, the study sample of Langston and Langston (2008) includes buildings with different functions and both new build and redevelopment. This is a major drawback of this study as these factors might possibly alter the findings. For instance, it is reported that generally 20-30% of total emissions from buildings are associated with EC while EC of warehouses can account for up to 80% (See, RICS, 2014). This is also supported by the study findings of Sansom and Pope (2012). Therefore, it is important that the sample is homogeneous in terms of the building function (i.e. houses, apartments, offices, retail, hospitals etc.) and the type of work (i.e. new build or renovation).

As per the discussion above, it can be deduced that building costs tend to increase as emission reduction measures are applied. However, the relationship between cost and carbon is under-explored and only a few studies have focused on both low cost and low carbon buildings. Further, it is also evident that the relationship between carbon and cost might vary due to varying element intensities of different types of buildings. Therefore, it is important that this knowledge is captured and disseminated, so that designers can produce carbon and cost-efficient design solutions. This calls for a tool that supplies the design team with the necessary information on carbon and cost accountability of a given design, as decision-making is made easier with the use of decision support tools.

2.10. Summary

This chapter discussed the background of carbon management beginning with the industrial revolution, which triggered the fossil fuel production. This lead to significant rise in the global temperature causing unanticipated and radical climate change due to the excessive presence of heat-trapping gases like carbon dioxide in the atmosphere. As a result, the economic, environmental and social conditions of the world regions are severely impacted. Further to that, stringent national targets are set by the UK government to meet the 2050 emission reduction goal through a carbon control trajectory. These actions plans are continually reviewed and reported to the government periodically to ensure compatibility with the projected climate change. Carbon control in the building sector is identified as one of the significant action plans to reach the goal, as the building sector is one of the major energy consumers. However, in the action plans more focus is given to reducing carbon emissions during the operation of the building, which contributes a significant proportion of total emissions while emissions associated with the production, maintenance and demolition of buildings given less focus. However, EC gained popularity with the introduction of the concept of zero carbon buildings.

It was also found that due to the development of low and zero carbon buildings and the zero carbon agenda of the UK government, the OC component was significantly reduced to reach the national targets. However, OC reduction measures tend to increase EC. It is envisaged that EC might be regulated in the future. Therefore, the management of EC is becoming significant. On the other hand, low-carbon options affect the cost of buildings. Hence, the need for a tool that predicts both carbon and cost to aid decision-making at the early design stages was identified.

3. Carbon Estimating

3.1. Introduction

A famous quote of Lord Kelvin, a mathematical physicist, is ‘If you cannot measure it, you cannot improve it’. Accordingly, carbon of a design has to be measured in order to reduce it and improve the environmental performance of the design. This chapter reviews a range of carbon estimating tools namely: OC tools, EC tools, life cycle analysis tools and multi-functional tools for early design stage and detail design stage, to establish the case for the development of decision support models for early stage carbon management.

3.2. Operational Carbon Estimating

OC of a proposed building in the UK is estimated using SAP (for domestic buildings), Simplified Building Energy Model (SBEM - for non-domestic buildings) or other approved software tools. The software calculates the monthly building energy consumption and the carbon emissions for given inputs. Inputs include general information about the building, description of the building geometry, construction, use, Heating Ventilating and Air Conditioning (HVAC) and lighting equipment (Building Research Establishment, 2014). Accordingly, Building CO_2 Emission Rate (BER) should be less than the Target CO_2 Emission Rate (TER) to approve the building design (as per the Part L of the Building Regulations compliant). Operational emissions are expressed in mass of CO_2 emitted per year per square meter of the usable floor area of the building (kg/m^2/year). As discussed earlier, OC is proportional to operational energy, thus, OC is based on the energy consumption of the building. According to Part L of the Building Regulations (in the UK), OC estimating is compulsory and is straightforward.

Further, OC estimating forms an important part of cradle-to-grave and cradle-to-cradle LCA of buildings. As presented in Figure 2.4, main energy consumption of in-use stage is the operation of the buildings (HVAC and lighting). Hence, OC is calculated by estimating the annual energy consumption of the building and converting it using carbon conversion factors for fuels, available for the UK reporting at the Department for Environment Food & Rural Affairs (2015).

While OC of a proposed building is calculated using a modelling software, actual OC (from actual operational energy consumption) can be calculated from a meter reading during the use of the building (Ekundayo et al., 2012). Some studies (Pan and Garmston, 2012, UK-GBC, 2008, UK-GBC, 2014a) pointed out that usually there is a gap between the predicted and the actual performance of buildings and stress the importance of attending to this issue. As a result, CIBSE developed a platform named CarbonBuzz to manage the gap between predicted and the actual performances. This emphasises the importance of harmonising predicted and actual emissions in reaching zero carbon target.

3.3. Embodied Carbon Estimating

Estimating EC follows a completely different process to that of OC. Measurement of EC has evolved during the recent past. The Inventory of Carbon and Energy (ICE) became the fundamental source of reference for EC estimating (cradle to gate) (Hammond and Jones, 2008a; Hammond and Jones, 2011). It is a database of construction materials containing energy and carbon data in the form of mass CO_2 emissions per mass of materials. Hence, the mass of materials that constitute a building needs to be quantified to estimate the amount of EC of a building. RICS (2014) guidance notes clearly state the steps in estimating EC based on a bottom-up approach – deconstructing a building element up until the material, labour and plant components and applying ICE EC factors to arrive at the total amount of EC of the building. This is a tedious task, as a building constitutes numerous items, which needed to be decomposed to follow this method.

The process is simplified to a certain extent by Franklin & Andrews (2011) with the introduction of the UK Building Blackbook which consists of itemised EC dataset for standard building items that are in accordance with SMM6/SMM7. The UK Building Blackbook presents data in a similar fashion to building price books used for cost estimating though Blackbook presents data in a dual currency format (EC and cost). Refer Table 3.1 for the basic two methods of EC estimating.

Table 3.1: EC estimating methods

Estimating EC of Concrete in a building							
Weight/ICE Method				**Unit of Measurement/Blackbook method**			
Material	**Quantity (Weight)**	**ICE factor**	**EC**	**Item**	**Quantity (units)**	**Blackbook factor**	**EC**
Cement	kg	$KgCO_2/kg$	$KgCO_2$	Concrete	m^3	$KgCO_2/m^3$	$KgCO_2$
Sand	kg	$KgCO_2/kg$	$KgCO_2$				
Aggregates	kg	$KgCO_2/kg$	$KgCO_2$				
Plant/Fuel	l	$KgCO_2/l$	$KgCO_2$				

3.4. Carbon Estimating Tools for Early Design Stage

The phrase 'early design stage' in this research refers to the first three stages of RIBA plan of work 2013 (strategic brief, preparation and brief and concept design) (RIBA, 2013). These stages merely hold less amount of design information, making the carbon estimating challenging, vague and less accurate. However, the calculations are less time-consuming.

3.4.1. Operational Carbon Tools

a) Carbon Critical Buildings

An early stage carbon prediction tool developed by Atkins (2014) to determine how space plan, primary system selection (heating and cooling), orientation and form assessment and envelope performance will affect the OC of a given design using built-in regional data from different countries. Also, this tool allows sensitivity analysis of different variables on the carbon footprint of buildings enabling better decision-making. In addition, comparison of OC against cost can also be generated. (More details can be found at http://www.atkinsglobal.co.uk/~/media/Files/A/Atkins-Corporate/group/cr/buildings-product-sheet-final-july10-tcm12-8458.pdf).

There seems to be a lack of standalone early design stage OC estimating tools and most tools are life cycle carbon estimating tools, which are discussed in section 3.4.3.

3.4.2. Embodied Carbon Tools

a) Construction Carbon Calculator

The calculator is developed by Build Carbon Neutral organisation and is a simple web-based tool. The basic concept of the calculator is Reduce, Renew and Offset, which means reducing through efficient building design, renew through renewable energy sources, locally sourced and recycled material, and achieve the maximum reduction possible. Then, focus on offsetting remaining project carbon through other available options like carbon trading and investing in low-carbon development projects.

The scope of the calculator is cradle to construction and the inputs required by the calculator include floor area, the number of floors above ground and below ground, primary structural system, ecoregion, vegetation and landscape. The output is measured in tonnes CO_2. The accuracy of the calculator ranges from -25% to +25%. The major limitation of the calculator is that it is applicable to US context and commercial or multi-family projects. The underlying database of the calculator is ICE Version1.6. (Build Carbon Neutral, 2007).

Tool is available at http://buildcarbonneutral.org/

b) Embodied CO_2 Estimator

A simple web based tool developed by Phlorum, in collaboration with the University of Brighton as part of a Knowledge Transfer Partnership. The tool requires inputs of floor areas, a number of floors, building perimeter, glazing ratio and brief elemental specifications to calculate EC from cradle to construction (excluding transport). The output of the calculator is given in tonnes CO_2e as well as the particular outcome is compared with a typical construction outcome and presented in a graphical format. Further, the tool is being developed to include cradle to grave impacts of a project (Phlorum, 2011).

Tool is available at http://eco2.phlorum.com/calculator/index

c) TATA steel - Steel Construction EC Tool

A simple web-based tool developed by TATA Steel (2014) to calculate EC of the superstructure of steel framed buildings. The tool has two modes namely: auto-generate and ‘manual. The auto-generate mode will generate the quantities through built-in algorithms and when data is input to the tool, relevant carbon factors are applied to derive the EC figure. The Same process takes place in manual mode except the quantities are input by the user manually if the quantities are known. The inputs required by the tool includes in auto-generate mode are upper floor areas, number of storeys, upper floor construction type, structural grid size (primary span and secondary span), roof structure, fire protection columns, upper floor concrete type, vertical bracing, voids in upper floors, % of void in upper floors and void walls. The outcome of the tool will be in different forms such as total EC - CO_2e figure; EC per one unit floor area - CO_2e/m^2; EC contribution of each element illustrated by a bar chart. A limitation of this tool is to be that it can be used only for steel buildings.

3.4.3. Life Cycle Assessment Tools

a) Green Footstep

A web-based tool developed by Rocky Mountain Institute. Requires design inputs including location, the size of the site, building type, floor area, expected life, project completion year. The tool gives three different outputs namely, ‘Site carbon storage’ in tonnes CO_2e, ‘Construction emissions’ in tonnes CO_2e and ‘Operational emissions’ in tonnes CO_2e/year.Limitation of the tool is to be the data sources of the tool are US based (coefficients from the U.S. Environmental Protection Agency), hence, applicable to US context only. Further, the output graph seems less user-friendly, leading to difficulties in interpreting the results (Rocky Mountain Institute 2009).

Tool is available at http://www.greenfootstep.org/

b) Building Carbon Calculator

An excel tool developed by the University of Minnesota. Calculations are linked in a separate excel sheet, hence, the user can determine the system boundary for the analyses depending on the availability of data. The tool requiresinputs on Operating energy, Potable water, Wastewater, Solid waste, Materials, Transportation, Soils and Vegetation. However, all the inputs are based on predictions. The outputs of the tool are ‘Immediate construction impact’ given in CO_2e (embodied impact), ‘Recurring annual impact’ given in CO_2e/year and the sum of the above two, ‘Total over building life cycle’ given in CO_2e. Major limitations include the tool was developed for Minnesota context, more focused on OC so requires users to input EC impact and depends on lots of predictions on energy usage and the like.A major source of data of the tool is to be the Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990 –2005 (University of Minnesota, 2014).

Tool is available at http://www.csbr.umn.edu/research/carboncalc.htm

3.5. Carbon Estimating Tools for Detailed Design Stage

The phrase ‘detailed design stage’ in this research refers to the developed design and technical design of RIBA plan of work 2013. The early stage design will be groomed by material selection. During these stages, the crucial decisions on the specification and services arrangements are made. Therefore, the carbon accounting becomes more accurate than the early stage estimating with the clearly defined elements and services components. On the other hand, the calculation becomes more complex than the early stage due to the extensive amount of design information.

3.5.1. Operationa Carbon Tools

Commonly accepted and legislated OC accounting tools include SAP for dwellings (incorporated in the Building Regulations Part L1) and SBEM (incorporated in the Building Regulations Part L2) for non-domestic buildings as discussed in Section 2.4. Given the scope of the study focuses on office buildings, this section will review only SBEM tool.

a) SBEM

SBEM is a computer-based tool, which analyses the energy consumption of a given building and thereby calculates the carbon dioxide emissions for a given period of time (per annum) which is called as the Building Emissions Rate (BER) measured in CO_2e/m^2/year. The tool requires comprehensive list of inputs which includes the shape & orientation of the building, HVAC efficiency and type of fuel used, thermal efficiency of building elements (floors, walls & roofs – U values), control of heating and cooling systems, renewable technologies, ventilation of the building, airtightness of the building, types and control of lighting. BRE (2009) states that SBEM could assist in the design process, though; it is not a design tool.

Tool can be downloaded at http://www.ncm.bre.co.uk/disclaimer.jsp

b) CarbonBuzz

CarbonBuzz is a platform developed by Chartered Institution of Building Services Engineers (CIBSE) for post-occupancy review of the building, to compares the designed energy use with actual energy due to the reported gap between predicted and actual performance of the building designs (UK-GBC, 2008, UK-GBC, 2014a, Pan and Garmston, 2012). The platform works based on the display energy certificate (which is mandatory for public buildings) to capture the actual energy usage rather than the forecasted energy usage. The platform enables to enter inputs for electricity and fuel consumption in kWh/m^2/year and then converts them into emission profile of the building (more information can be found at http://www.carbonbuzz.org/). This way the platform helps the building owners to compare the results with the benchmarks and take necessary actions to close the gap.

3.5.2. Embodied Carbon Tools

a) Carbon calculator for construction projects

The calculator is an excel tool developed by Environment Agency. The scope of calculator includes cradle to construction allowing transportation related EC into calculations and the emissions are calculated in tonnes CO_2e. The inputs required by the calculator include material quantity, waste disposal, plant and equipment, site accommodation, transport distance, mode of transport and personnel travel. The main data sources used by the calculator are Hammond and Jones (2006), ICE version 2.0 and DEFRA (2011) for carbon coefficients.

The carbon calculator enables appraisal of different designs in terms of the material section. The major limitation of the calculator is that the inputs are to be entered manually by the user (Environment Agency, 2012).

Tool is available at https://www.gov.uk/government/publications/carbon-calculator-for-construction-projects

b) The Green Guide Calculator

A web-based tool developed by BRE in compliance with 'The Green Guide to Specification' with the database of extensive building specifications. The tool helps designers to make choices between materials and specifications to achieve better environmental rating (e.g. BREEAM, CSH etc.). The Green Guide covers six common building types (commercial, educational, healthcare, retail, residential and industrial) and eight building components including ground floors, upper floors, roofs, external walls, windows, internal walls and partitions, insulation and landscaping. However, few window types (i.e. domestic windows and commercial windows), insulation, floor finishes and landscaping are not included in the calculator and may be included in the future. Access is available only to registered users and users are able to upload the design information by selecting relevant element and sub-element of a given design so that the tool calculates the environmental rating, ranging from A+ to E (lower environmental impact to higher impact), as well as the embodied impact of the element in $kgCO_2/m^2$. The major drawback of the calculator was that it had limited predefined specifications, though the latest version of the calculator enables users to submit a new specification as a query to BRE Environmental Assessment Method (BREEAM) and the new specification will be added and made available to all users (BREEAM, 2013). Hence, the calculator is becoming more responsive and therefore, is a good tool to select building specification. However, the major limitation of the tool is the exclusion of significant carbon hotspot elements such as services and finishes in the specification database, which is to have a huge impact on design decisions.

c) Interoperable Carbon Information Modelling (iCIM)

The iCIM is a case study by “OpenBIM” (Open Building Information Modelling) to assist designers to make informed design decision in terms of carbon and enable to choose low carbon specification (Moncaster and Symons, 2013). iCIM is a well-advanced tool developed in a BIM platform to allow easy and faster calculation of EC of a design. Further, the tool is integrated with a database such as NRM2 and ICE to ensure consistency between cost and carbon data. The tool is effectively used with detailed designs as the tool indicates the alternative specification available for a given element so that allows running what-if analyses of different specifications for a particular element. This enables the designers to choose the most carbon efficient specification for each element and achieve the desired carbon footprint for the entire design.

However, BIM lacks early stage carbon estimating models so that the outcome of this research (early stage carbon prediction model) can be integrated into a BIM environment will lead to effective decision-making at early stages of design.

3.5.3. Life Cycle Assessment Tools

It is common to see that most of the life cycle assessment tools available are very complex tools developed by software developers and available at a high cost. Further, very little information is available about the tools. Some of the available tools are discussed below:

a) GaBi Software

A software solution developed by PE Internationals for life cycle assessment of the product and building designs and many other applications including certification, EPD generation, design solutions, water footprint, resources and energy efficiency solutions and the like (more details can be found at http://www.gabi-software.com/solutions/). Contains GaBi database developed by PE Internationals, ecoinvent database, U.S Life Cycle Inventory (LCI) database and in addition to those PE International provides data on demand which is not included in the underlying databases (PE International, 2014). Further, a recent review of DEKRA (a leading service provider in auditing and certification) on GaBi software pronounces that the methodology used in the software for modelling is thorough in terms of LCA best practice and continuous improvement and data maintenance seems to be very much coherent and transparent. However, the review does not certify the correctness of the outcome, but the focus was on the development and continuous management process of key technology dataset of the GaBi database.

b) SimaPro LCA Software

A software developed by PRé Consultants to include a cradle to gate as well as a cradle-to-grave system boundary that uses regularly updated databases like ecoinvent v3 LCI database, other EU, US and Swiss databases which increase the accuracy of the predictions. However, it is available at a cost (PRé Consultants, 2014).

c) Carbon Critical Knowledgebase

A web-based tool developed by Atkins (2014) which evaluates alternative options in terms of embodied and OC and indicates the ways of minimising carbon footprint (for more details refer the product sheet at http://www.atkinsglobal.co.uk/~/media/Files/A/Atkins-Corporate/group/cr/knowledgebase-product-sheet-final-july10-tcm12-8459.pdf).

d) Sturgis Carbon Profile Model

While most studies treated EC and OC as per the general rule, this model combines both operational and EC into one unit and proposed a methodology to measure lifecycle carbon of a building in $kgCO_2/m^2/year$. Figure 3.1 illustrates the way that single metric for life cycle carbon of a building is achieved. Accordingly, OC profiling follows the industry accepted model, sBEM while Sturgis and Roberts (2010) define EC prediction model as Sturgis compatible metric and the EC output is called as EC Efficiency (ECE).

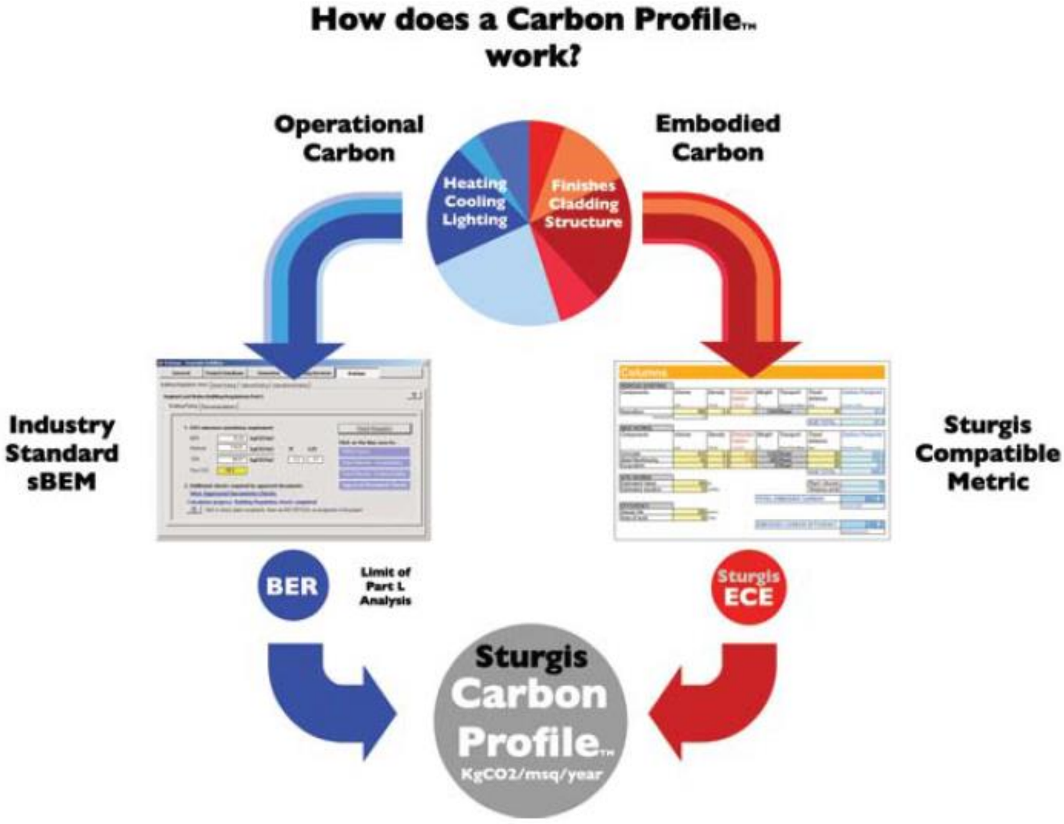


Figure 3.1: Sturgis Carbon Profiling Model

Source: Sturgis and Roberts (2010)

Consequently, Sturgis Carbon Profile of a given building is derived as,

$$\underbrace{A^{-1}\left[\sum_{n=1}^{N} x_n\right]}_{\text{BER}} + \underbrace{A^{-1}\left[\overbrace{\left[\sum_{j=1}^{J}\frac{y_j}{l_j}\right]}^{\text{Independent Components}} + \overbrace{\left[\sum_{b=1}^{B}\left(\frac{\sum_{t=1}^{T} y_t}{\min l_t}\right)_b\right]}^{\text{Linked Component Systems}}\right]}_{\text{ECE}}$$

A - Net Internal Area of building

N - Set of elements giving rise to all operational emissions

X - Element giving rise to operational emissions

J - Set of all independent components

y- Component giving rise to EC emissions

B- Set of all linked component systems

l - Lifespan of component

T - Set of all components comprising an individual linked system

The methodology followed by Sturgis to achieve a compatible unit for ECE is unique and follows five distinct steps in calculations as indicated in Figure 3.7.

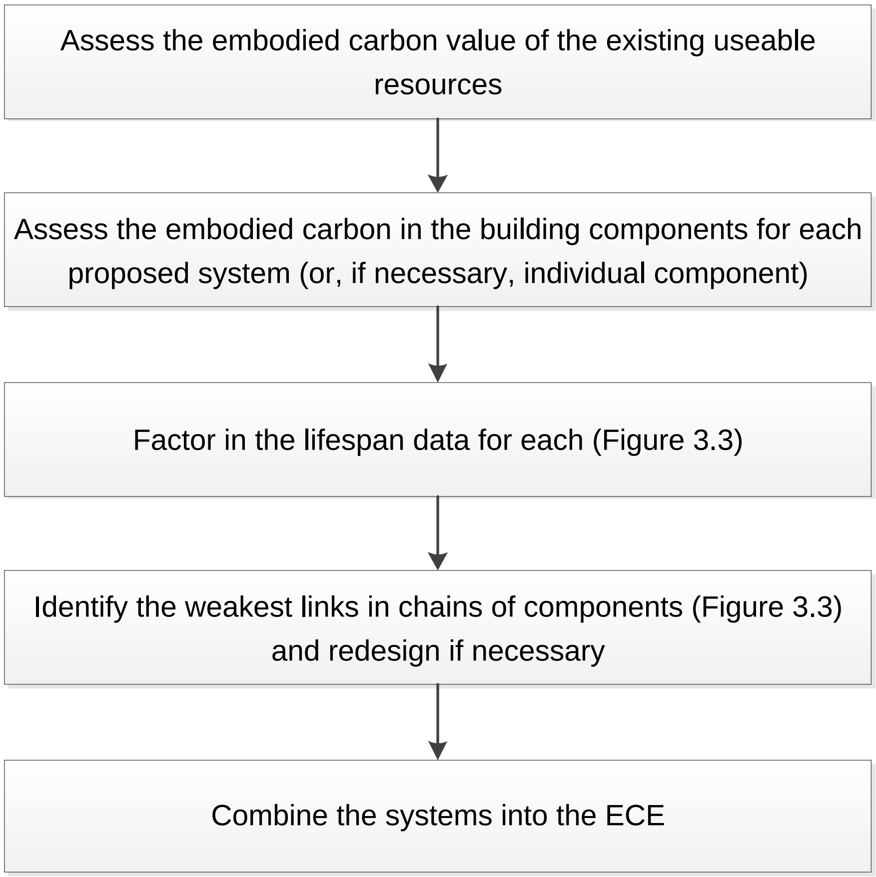


Figure 3.2: Five steps in calculating ECE

The most challenging part of the calculation is the factoring of life span data for each element and identifying the weakest link in the system to arrive at ECE, which is briefly illustrated in Figure 3.3. Sturgis and Roberts (2010) admit that the outcome of the model is subject to the quality of available data on EC, lifespan, building quantities and other subjective interpretations.

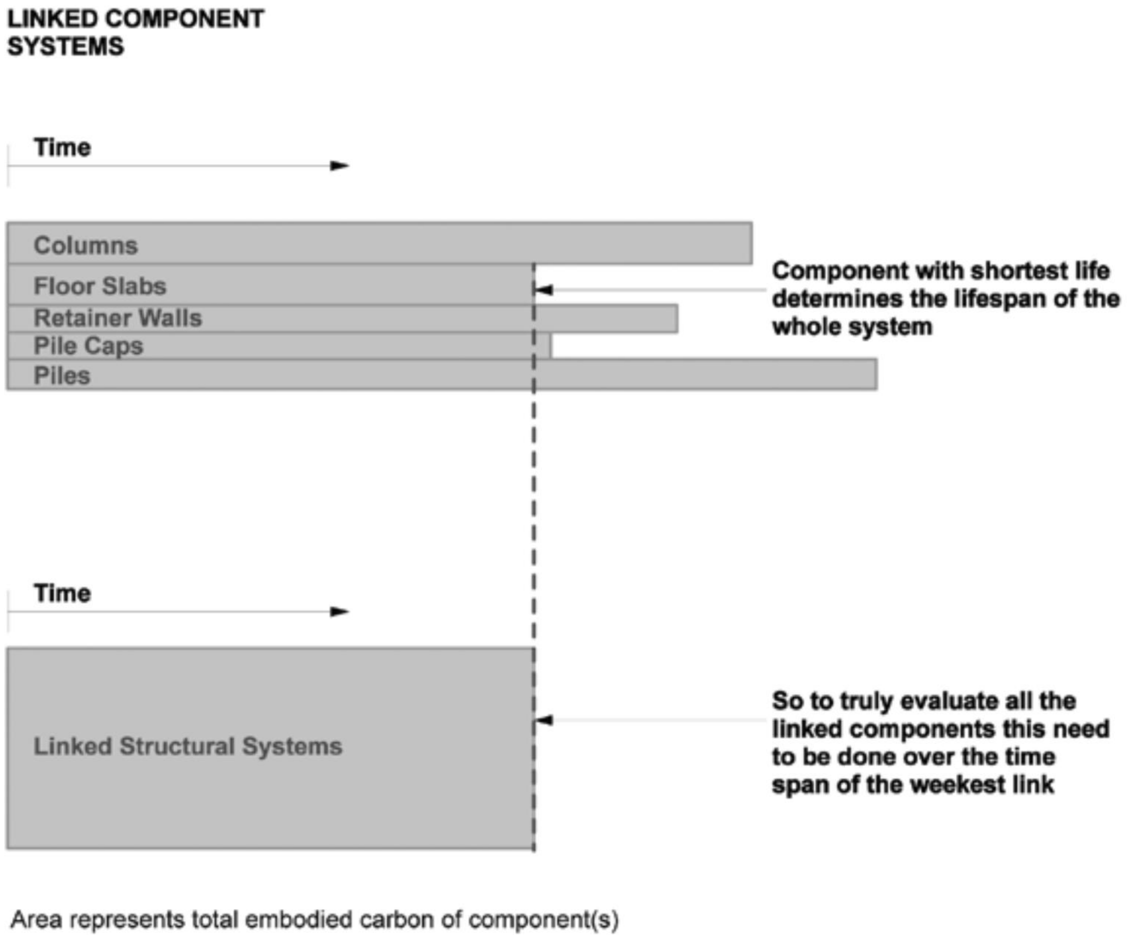


Figure 3.3: Evaluating linked components

Source: Sturgis and Roberts (2010)

Nevertheless, the usefulness of the model in decision-making is not very prominent and being a model for detailed design stage the outcome is not helpful for the selection of materials. Further, the lifespan of linked component relies on the weakest link or system with the shortest life span, which might influence the outcome.

3.6. Carbon Databases

A range of carbon databases is available to aid design decision-making. These databases form the key knowledge base of many tools and platforms discussed above. However, it should be noted that the embodied impact factor of a material should not be compared directly with another material as the material databases give the impact of one unit of mass of each material (kgCO_2/kg). Hence, the total quantity of the material is required to find the total impact of a particular material on the design, which should be taken into consideration before making a decision.

a) ecoinvent Database

The ecoinvent database is developed by the Centre for Life Cycle Inventories. It is an international life cycle inventory database with an updated inventory of data from several disciplines, in addition to carbon inventory. The database form as the underlying source of data in many design tools with LCA calculations. The latest version of ecoinvent database is 3.1 with new updates to the inventory and changes to the underlying methodologies. Access to the database is allowed only for the registered users (ecoinvent Association, 2015).

b) Inventory of Carbon and Energy (ICE)

ICE is an extensive database of carbon and energy of building materials, which was developed by Professor Geoff Hammond (University of Bath) and Dr Craig Jones (Circular Ecology). The first version was made available in 2006 for free download which then underwent several revisions and version 1.6 was published in 2008 then with significant improvement to the previous version the second version was published in 2011 (version 2.0). One of the most important revisions includes, the data had been converted from CO_2 to CO_2e in the latest version allowing accountability of other GHG emissions (Hammond and Jones, 2011).

The system boundary of the database is said to be cradle to gate. Further, the emissions from primary energy had been accounted though feedstock energy is considered in special circumstances. Further, carbon sequestration is not included in the data. EC values were derived from including foreign sources while most data were sourced within the UK and claimed to be fairly recent (Hammond and Jones, 2008a).

Further, Hammond and Jones (2011) highlighting the growing concern on EC, recommends the government and the industry to agree on a standard to measure EC to use as a design tool. Hence, enabling carbon appraisal to be included in the feasibility studies of a project which will help construction industry to meet the low carbon agenda more effectively (Hammond and Jones, 2011). Moreover, the inventory also recognises the uncertainty in the carbon data due to various fuel types. Nevertheless, this is the most widely used energy and carbon database for calculations, especially within the UK context, and most tools have ICE database as underlying data source.

c) Hutchins UK Building Blackbook - Small and Major Works

Franklin+Andrews previously published this book to help industry professionals to get the updated knowledge about the cost of doing business which is now covering two aspects, cost and carbon accountability of doing a business. This book presents the cost and carbon in an itemised pattern in accordance with SMM7 (for major works) and SMM6 (for small works). The following is an extract from Blackbook listing the resource requirement for one unit quantity of the items and the cost and carbon of one unit of the respective items.

Excavation, Earthwork and Concrete Work

Small Works 2011		Unit	Labour Hours	Labour Net £	Plant Net £	Materials Net £	Unit Net £	Unit with 10% £	CO_2 Kg
101	**NEW WORK**								
10101	**SITE PREPARATION**								
1010101	**Form temporary site road; 150 mm hardcore; maintain during period of contract**								
1010101A	3.00 m wide	m	2.70	56.30	-	10.78	67.08	73.79	6.950
1010102	**Break up and remove temporary site road 150 mm hardcore**								
1010102A	3.00 m wide	m	2.25	46.91	-	-	46.91	51.60	-
1010103	**Temporarily enclose site; fencing up to twenty times used**								
1010103A	1.35 m chestnut fencing	m	0.25	5.21	-	-	5.21	5.73	-
1010103B	2.70 m chainlink fencing	m	0.45	9.38	-	-	9.38	10.32	-

Figure 3.4: An extract from the Blackbook

Source: Hutchins (2011)

This type of dataset is very useful in detailed design stage because when the bills of quantities are produced to ascertain the cost aspect of projects the carbon accountability can also be established. Provided that carbon plans are developed at the early stage, carbon checks checks can be performed in parellel to cost checks using this dataset. Eventually, the design can be revised to match the established target what is called 'designing to cost' can be extended to the dual currency – 'designing to cost and carbon'. A problem with the structure of this dataset is that it complies with the SMM. There is a need for a NRM compliant dataset as it is considered to be the latest measurement standard of the construction industry.

d) EC Database - WRAP

The database is developed by WRAP in collaboration with the UK Green Building Council to capture the EC data for the whole building. WRAP and UK-GBC have created a closed database in order to prevent misuse of the project information uploaded into the database. Further, the database requires construction professionals and academics (including students), those who seek carbon information or wish to share carbon information, to register in order to gain access to the database. Then the data can be freely accessible by registered users. There are more than 300 registered users and more than 200 projects are stored at present in the database.

The database allows comparison within the registered projects, in anticipation that the designers will use the data to develop more carbon efficient designs. The database allows the registered users to upload the project-specific carbon data themselves. It follows the definition of life cycle stages stipulated in BS EN 15804 (see, Figure 2.4) and allows filtering data depending on the extent of the analysis (system boundary of the analyses) required by the user namely: product stage (A1-3), construction process stage (A4-5), use stage (B1-7), end of life stage (C1-4), benefits and loads beyond system boundary (D). The database also allows the data to be filtered in terms of CO_2 and CO_2e.

In addition, EC analyses are presented in an elemental fashion in accordance with NRM element definition. Building elements are grouped into six categories including Substructure, Superstructure Structural, Superstructure Non-

Structural, Envelope, Internal Finishes and External Works. Superstructure Structural includes Frame, Upper Floors and Roof; Superstructure non-structural includes Internal Walls and Partitions and Internal Doors; Envelope includes External Walls and Windows and External Doors; Internal Finishes cover Wall Finishes, Floor Finishes and Ceiling Finishes.

Currently, only a few projects are that are available in the database cover Cradle-to-Grave system boundary. Nevertheless, it is a worthwhile attempt in bringing the concern of the construction professionals and academics on EC of projects and provide with more updated information on the real-time projects. Moreover, the success of the database entirely depends on the users because findings can be generalised when the number of projects in the database is large. The database can be accessed through http://ecdb.wrap.org.uk/ (see for more details, WRAP and UK-GBC (2014)).

e) End of Life Dataset of framing materials

PE International (an international market leader in sustainability-related consultancy and software solutions) developed an end of life (during and after demolition and disposal - C and D modules in BS EN 15804) dataset for common framing materials of buildings. However, end of life EC is a less researched area and suffers from limited data.

This dataset is useful in deriving life cycle embodied impact of the given framing materials (brickwork, blockwork, concrete and steel) so that a holistic picture can be seen before taking decisions on the type of framing materials for the proposed building.

f) Department for Environment Food & Rural Affairs (DEFRA) carbon conversion factors

This is an online repository with up to date carbon conversion factors for fuels to the calculated carbon footprint of business operations and products. However, this repository is suitable only for UK businesses, researchers and international organisations reporting on the UK operations. This repository allows three options in downloading the factors as an excel file as follows (Department for Environment Food & Rural Affairs, 2015):

- Specific data demanded by the user: this option allows users to filter data depending on the scope, fuel or activity type and by the data type that needs conversion. DEFRA also recommends this option as it eases the process of locating relevant data. However, this option is only available for the dataset from 2012.

- DEFRA's frequently used data: this allows users to download pre-filtered factors used by DEFRA frequently for estimating purposes. This includes a range of factors, which are adequate for average footprint calculations of businesses.

- All available data: this option allows users to download all the factors for a respective year. This option is not recommended by DEFRA for usual carbon accounting while users may be interested in this option for advance use.

This data becomes useful when estimating EC during construction, use stage and end-of-life stage.

3.7. Carbon Guides

RICS has been a pioneer in carbon profiling research and published many research reports and guidance notes related to the topic. The key guidance note on quantifying EC is discussed below.

a) RICS guideline - Methodology to Calculate EC

The latest guide on EC calculation of construction project during different stages of the project was published in 2014. The initial guide on EC calculations was published in 2012 titled 'Methodology to calculate EC of materials' covering the cradle to gate system boundary. Later, RICS developed the guidance note to cover cradle-to-grave system boundary for EC calculations, which remains as the latest guidance note. RICS (2014) classifies the project into four main stages namely: product, construction process, use and end-of-life stages. The methodology to be followed in EC calculations on each stage as per the guidance note is listed in Table.

Table 3.2: EC counting guide in different stages of project

Stage	Methodology	Data source
Product	$EC_{product}$ = ∑ Quantity of material constituents in each item/element x EC factor of the respective material	ICE (UK) (Hammond and Jones 2011), SimaPro, GaBi
Construction Process	$EC_{construction}$ = ∑ Quantity of energy used for the activity x EC factor for respective energy source	DEFRA Greenhouse Gas Conversion Factor Repository, GHG Protocol calculation tools
Use Stage	EC_{use} = ∑ Quantity of materials to be replaced x No. of replacements x EC factor of the respective material	BCIS Life Expectancy of Building Components (BCIS 2006) + product stage sources
End-of-Life	$EC_{end\text{-}of\ life}$ = ∑ Quantity of energy used for the activity x EC factor for respective energy source	Construction stage sources

This guidance note allows Quantity Surveyors to calculate the EC manually while quantifying building element quantities and pricing the project. Hence, the guidance enables the competencies of QS to be utilised for EC calculations without spending on expensive tools.

3.8. Review of Carbon Estimating Tools

According to Figure 3.5, it is clear that there are plenty of carbon assessment tools for both early design and detail design stages. However, the prediction accuracy of most tools has a wide band, making them less reliable. Moreover, all these early stage tools are perceived as an indicator of the carbon accountability of designs rather than a design decision tool. Sturgis and Roberts (2010) pointed out that though most of these tools are helpful in giving an overall picture of the emissions, they fail to address the issue of mitigation measures of one component of emissions affecting the other component. Furthermore, most tools fail to integrate cost, which is another important aspect of designs. As it can be seen from Figure 3.5, iCIM is intersecting both carbon and cost while iCIM is designed to work on a BIM platform during detailed design stages. Therefore, there is a need for a simple tool to aid design decision-making in terms of carbon and cost at early design stages as discussed previously. Hence, the research outcome intends to fit into the patterned area in Figure 3.5, which is vacant at the moment.

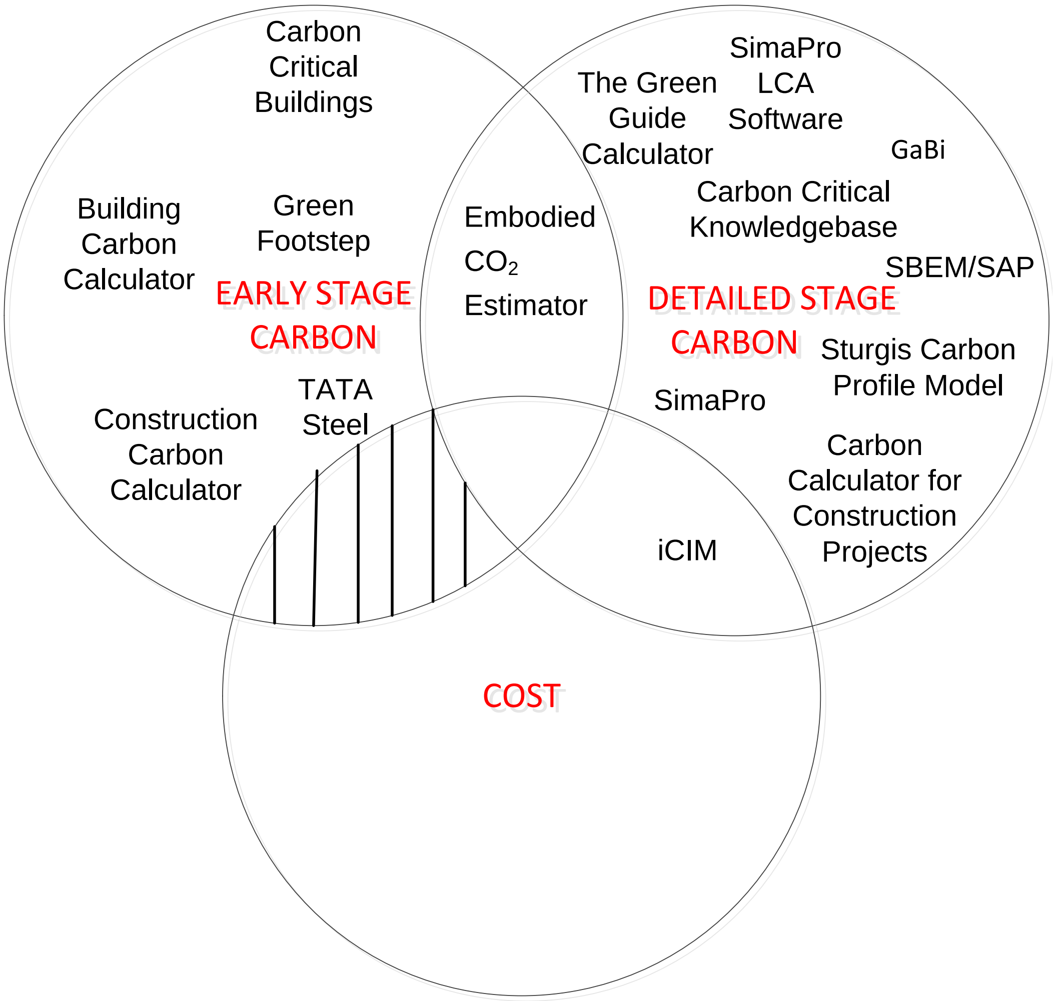


Figure 3.5: Overview of the carbon tools

3.9. Factors Affecting EC Estimating

A major concern in carbon estimating is the accuracy of estimates. It is not surprising to find variations in estimates produced for the same building by different estimators (Clark, 2013). A few scholars (Dixit et al., 2010, Clark, 2013, Ekundayo et al., 2012) identified that variations in EC measurements among which five key factors are system boundary, the method of estimating, assumptions, data sources used and element classification, which are discussed below.

a) System Boundary

The EC estimate can be based upon any one of five system boundaries as discussed in Section 2.5. Therefore, an estimate with a cradle-to-grave boundary will have higher figures than an estimate with a cradle-to-gate boundary. Therefore, the system boundary is one factor to be considered when comparing studies and using data from other studies for analysis purposes.

b) Method of Estimating

There can be two main possible methods in carbon estimating,

- Manual estimating: this can be either a bottom-up approach of estimating using ICE data source and other relevant sources or itemised estimating approach using Blackbook data. Even though the Blackbook is developed using ICE data, new data were also sourced by the Blackbook team to develop the book. Therefore, there are possibilities of variations. Furthermore, missing data in ICE and Blackbook need to be sourced from local manufacturers, suppliers or contractors, which can vary from project to project. Hence, this needs to be taken into consideration.
- Automated: automated systems will have a unique built-in program for extracting quantities and retrieving carbon data. Hence, different standards adopted by the system for the method of measurement will result in variations. In addition, most software use ecoinvent database, which is updated time to time. These can lead to varying result from manual measurements.

c) Assumptions

Assumptions are an important cause of variations. Because EC estimates are mainly produced from detailed cost plans or Bills of Quantities (BoQ), if an item description is imprecise then assumptions have to be made to proceed with the estimate. For instance, a staircase measured in ‘Nr’ has to be broken into concrete, formwork, reinforcement, balustrades and finishes to get the carbon estimate of that element. In this case, assumptions play a major role in the carbon estimate. Further, assumptions vary from a person to person, a project to project and it cannot be standardised. Therefore, this is a major drawback in EC estimating.

d) Data Sources

As explained under the method of measurement, data sources other than ICE and Blackbook might vary from study to study due to the difference in manufactures, suppliers, contractors, the age of data source and the like. This will result in different EC figures.

e) Element Classification

Element classification is a common variation among studies. Different studies (Halcrow Yolles, 2010b, WRAP, Halcrow Yolles, 2010a, Clark, 2013, Sturgis Associates, 2010) adopt different element classifications such as NRM, SMM/BCIS - older version, British Council of Offices 2011 and some studies did not follow any standard. This inconsistency in element classification makes the comparison of findings difficult.

3.10.The Research Context

Given that the quantification of carbon is not that easy at early stages of design, this research adopts a unique concept of predicting carbon at early stages of design. The research idea is explained below:

Firstly, it could be hypothesised following the Pareto Principle (80:20 rule) that 80% of the EC emissions come from 20% of the building components. These can be referred to as the carbon-intensive elements or hotspots of a building as discussed in section 2.6. Hence, EC of a building can be calculated using the hotspots, which is presented in Equation 3.1: Accordingly, it is crucial to identify the carbon-intensive elements (A, B, C…) of a building. The EC of carbon-intensive elements added with the minor EC components (k) will result in the total EC emissions (C_E) of a building. This implies the selection of material determines the EC of buildings. However, selection of building materials and specification is to be carried out at the detail design stage. Hence, calculating EC is challenging during the early design stages.

Equation 3.1: Conceptual model to calculate EC using hotspots

$$\mathrm{C_E} = \mathrm{A_{CE}} + \mathrm{B_{CE}} + \mathrm{C_{CE}} + \cdots + \mathrm{k}$$

Nevertheless, carbon-intensive elements can be captured by obtaining historical project data and the building morphology parameters (plan shape, storey height, total height, and the like) related to the carbon hotspots can be modelled to predict EC at early design stage with low error margin. Further, the influence of services and finishes quality on EC was taken into consideration. Consequently, services and finishes can be identified as quality parameters in the mathematical model, though, assessing the quality of services during the early stages of design is challenging (RICS, 2014).

Finally, the research idea can be presented as a conceptual regression model as follows:

$$Carbon\ Factor \left\lfloor \frac{\mathrm{Carbon}}{\mathrm{m2}} \right\rfloor \propto Morphology\ Parametrs\ (M_P)$$

$$Carbon\ Factor \left\lfloor \frac{\mathrm{Carbon}}{\mathrm{m2}} \right\rfloor \propto Level\ of\ Sevices\ (L_S)$$

$$Carbon\ Factor \left\lfloor \frac{\mathrm{Carbon}}{\mathrm{m2}} \right\rfloor \propto Level\ of\ Finishes\ (L_F)$$

$$Carbon\ Factor \left\lfloor \frac{\mathrm{Carbon}}{\mathrm{m2}} \right\rfloor = f(M_P, L_S, L_F)$$

$$Carbon\ Factor \left\lfloor \frac{\mathrm{Carbon}}{\mathrm{m2}} \right\rfloor = \mathrm{a}\left(\frac{\mathrm{Wall}}{\mathrm{Floor}}\right) + \mathrm{b(Storey\ Height)} + \mathrm{c(Building\ Height)} +$$

$$\ldots\ldots + \mathrm{Service\ Index} + \mathrm{Finshes\ Index} + \mathrm{k}$$

(Where, a, b, c…k = regression coefficients)

Developing the model requires identification of influential design variables that affect EC. Since there are no reported studies on the relationship between such variables and EC relationships, the study capitalises the existing body of literature on cost and design variable relationships to deduce the variables affecting EC, as there is a connection between carbon and cost. The following section explores the design variables affecting cost.

3.10.1. Design Variables Affecting Cost

Figure 3.6 presents the variables identified in the past studies affecting construction cost, both from the theoretical knowledge base and practical application of parametric cost models. However, not all the variables identified can be considered for analysis due to the scope of the study that focuses only on design variables of alternative designs. Variables that are eliminated from further consideration include:

- **Life of the building and end use of the building** are important in the case of life cycle costing calculations. However, these have no benefit if incorporated when only CC is analysed. Further, this variable remains constant for alternative designs.
- **Quality of workmanship and specification related variables** have an impact on CC though during early design stages information is not likely to be available.
- **Region or location of the site and site considerations** are important variables that affect the CC though the aim of the model is to choose an optimum design from alternatives, where the site is not a variable.
- **Number of occupants** is usually reflected from the building size.
- **Contract duration** affects the project overheads, which are covered under preliminaries in most of the cases. Also, preliminaries vary from project to project depending on the client's or the contractor's requirements. Moreover, the duration is not a design variable but a project variable. Hence, the model excludes preliminaries.
- **Amount of liquidated damage** has no implications on the design.

- **Buildability** is another variable that determines the CC. However, the impact of buildability is considered irrelevant for this study. Further, it is also not easily quantifiable.
- **Refurbishment** is an options appraisal and decision to refurbish will eliminate the need for a new build.

After the initial screening, seven design variables were identified as most significant and applicable for the study and listed in Table 3.3 with implications of each variable on cost described.

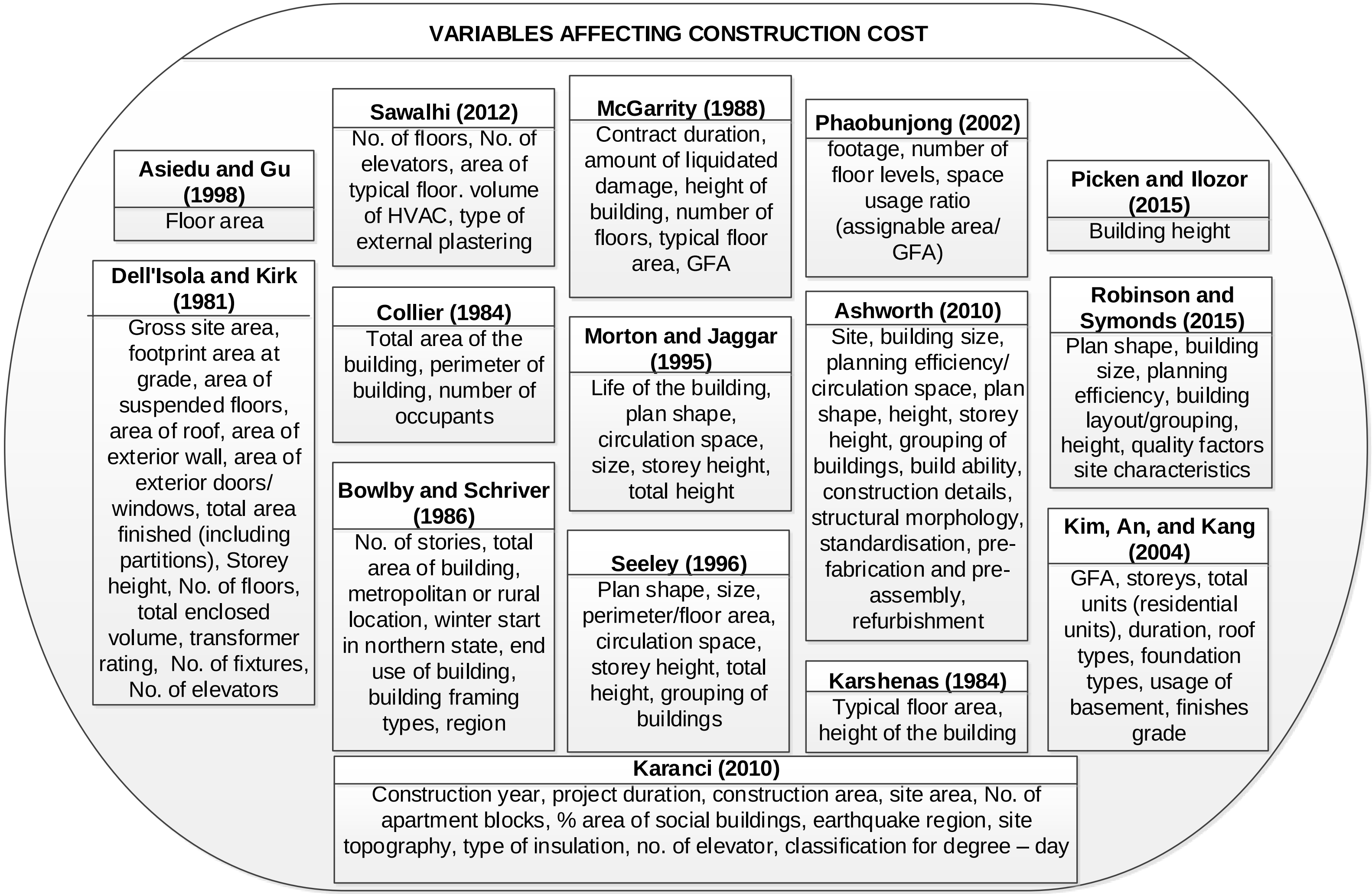


Figure 3.6: Variables influencing construction cost from past studies

Table 3.3: Cost influential design variables and its implications on the CC

Design variables	Sources	Comments
Plan shape or Wall/Floor area	Dell'Isola and Kirk 1981 Collier (1984) Morton and Jaggar (1995) Seeley (1996) Ashworth (2010) Robinson and Symonds (2015)	Plan shape is usually quantified by External wall area/Gross Floor Area. Design with the lowest ratio is economical in terms of plan shape. However, sometimes site layout dictates the plan shape where alternative design solutions will be limited.
Building size	Asiedu and Gu (1998) Dell'Isola and Kirk 1981 Collier (1984) Karshenas 1984 Bowlby and Schriver (1986) Phaobunjong (2002) McGarrity (1988) Morton and Jaggar (1995) Seeley (1996) Ashworth (2010) Robinson and Symonds (2015)	As the project size increases, project overheads tend to decrease due to economies of scale. Also discounts on bulk purchase will result in reduced cost.
Planning efficiency/circulation space	Phaobunjong (2002) Morton and Jaggar (1995) Seeley (1996) Ashworth (2010) Robinson and Symonds (2015)	Lower non-usable space will save energy cost. However, it is subject to planning requirements.
Building layout/grouping of buildings	Seeley (1996) Ashworth (2010) Robinson and Symonds (2015)	The advantage of common elements reduces the cost.
Storey height	Morton and Jaggar (1995) Seeley (1996) Ashworth (2010)	More the storey height more the cost.
Total height/No. of floors	Sawalhi (2012) Karshenas 1984 Bowlby and Schriver (1986) Phaobunjong (2002) McGarrity (1988) Morton and Jaggar (1995) Seeley (1996) Picken and Ilozor (2015) Ashworth (2010) Robinson and Symonds (2015)	Relationship with total height and cost is slightly complex. Different studies at different locations report different results. Generally, cost expected to increase with building height.
Quality factors	Sawalhi (2012) Dell'Isola and Kirk 1981 Robinson and Symonds (2015)	The quality of finishes and services affect the cost.

The identified cost influential variables are an indication of potential EC influential variables. However, the knowledge on carbon influential variables is not readily available but can be captured by identifying ‘carbon hotspots’ in buildings, as each variable defines one or more building elements. Table 3.4 lists the design variables that affect the building element/s. For instance, if the frame was identified as the biggest carbon hotspot then the height of the building would become the most carbon influential design variable of the building.

Table 3.4: Building parameters affecting building elements (After Dell'Isola and Kirk (1981) and Collier (1984))

Building Parameters	Building Elements
Footprint area	Substructure
Area of suspended floors	Floors
Area of roof	Roof
Area of exterior wall	External walls
Area of exterior doors/windows	External doors and windows
Total area finished (including partitions)	Finishes
Total enclosed volume	Services – mechanical
Transformer rating	Services - electrical
Gross site area	External works
Gross floor area	Substructure, upper floors, roof, internal partitions, mechanical, electrical
No. of storeys/total height of the building	Frame, stairs
Storey height	Frame, stairs
Plan shape or Wall/Floor area	External walls, external doors and windows, upper floors
Planning efficiency/circulation space	Internal partitions, finishes, services
Building layout/grouping of buildings	External wall
Quality factors	Finishes, services

In this way, carbon and cost influential variables will be identified by building case studies and models will be developed to predict EC and cost during early stages of design based on design variables of the buildings.

3.11.Summary

This chapter presented a comprehensive review of the available carbon estimating tools during early stages and detailed stages of designs. Among early stage and detailed stages of design, it can be noted that there exist a number of simple web-based early stage carbon estimating tools. However, it is obvious that the predictions are vague and the accuracy band is wide, as these tools require the minimum amount of information. On the other hand, detailed design stage tools are more complex and require more information as inputs resulting in predictions that are more accurate. However, most of the detailed stage tools are in the form of expensive software packages available for purchase. Some advanced tools only work under BIM platforms. Further, none of the tools presented here take account of the most important aspect of construction projects that is cost, especially during the early design stage, which could lead to decisions that are more rational. This gap laid a strong foundation for the current study. Subsequently, the need for developing an early design stage prediction tool that uses design variables of buildings to predict EC and CC of building designs was identified. Use of parametric cost models to estimate cost during the early stages of projects has proven successful application. Therefore, the same approach is attempted in this study in estimating carbon to make it more approachable and concurrent to cost estimates.

4. Research Methodology

4.1. Introduction

The first two objectives of the study were achieved through the discussion of the existing literature on EC management and measurement. Literature findings highlighted the need for an early design stage EC prediction model. Accordingly, this chapter unfolds the methodology followed in investigating the problem established in the literature review (see, Sections 1.2 and 3.10). The research methodology is a systematic procedure followed by the researcher based on logical thinking to achieve the research objectives and ultimately the aim. A sound methodology has the power to uphold the research and find the best possible answer for the identified research problem in the literature review. Hence, different worldviews and methodological choices available to develop an EC prediction model are reviewed in this chapter and the most appropriate methods are selected to answer each of the research questions posed in the introduction chapter (see, Section 1.2) to achieve the remaining five objectives.

4.2. Research Philosophy

A researcher has the key role to establish his or her philosophical stance towards the problem researched based on certain assumptions or belief systems. Two basic questions needed to be brought into the discussion to establish the philosophical stance of the researcher, namely ontological question (what is the nature of the reality?), epistemological question (what is acceptable knowledge?) which typically dictates the methodological question (how the knowledge can be acquired?) and axiological question (what is the role of values of the researcher?) (Saunders et al., 2009, Corbetta, 2003, Guba and Lincoln, 1994). Sutrisna (2009) argues that ‘Objectivism’ and ‘Subjectivism’ are popular examples of ontology while ‘Positivism’ and ‘Interpretivism’ are of epistemology. Objectivism assumes that the reality exists independent of human conscience and experiences while Subjectivism assumes that the existence of reality is conceived through human conscience and experiences (Saunders et al., 2009). On the other hand, ‘Positivism’ suggests that the reality can be observed and measured by the researcher in an objective way while ‘interpretivism’ assumes that the knowable reality is influenced by the individuals (Gray, 2014).

Positivists prefer working in an observable reality where the outcome of a research is generalizable and believe that the researcher is independent of the subjects of the research (Saunders et al., 2009). However, the positivist theory was criticised for its assumption of objectivity in knowing the reality (Robson, 2011, Guba and Lincoln, 1994, Corbetta, 2003). For instance, values of the researcher have an influence on research designs which leads to subjective outcomes. This gave rise to ‘Postpositivism’ theory which is also known as ‘Realism’ that addresses the limitations of the epistemological positivism. Realism claims that the reality is conceived through our senses. Furthermore, Realism also takes two positions including ‘Direct Realism’ and ‘Critical Realism’. Direct Realism claims that our senses show us the true reality while Critical Realism suggests that our sensations are not a true representation of the reality and objects have an existence independent of our human minds (Levers, 2013, Saunders et al., 2009). A clear distinction of the three paradigms (Positivism, Realism and Interpretivism) and the philosophical questions are explained in Table 4.1. However, sometimes it is hard to fit into one of the three positions due to the nature of the research questions. In such cases, ‘Pragmatism’ allows researchers to work with variations in the branches of epistemology, ontology, methodology and axiology for different research questions (Saunders et al., 2009).

Table 4.1: Comparison of research philosophies (After: Corbetta (2003), Saunders et al. (2009), Guba and Lincoln (1994))

		Positivism	Postpositivism/ Realism	Interpretivism
Ontology		Naïve realism: Reality is real and knowable	Critical realism: Reality is knowable only in an imperfect and probabilistic manner	Constructivism: the knowable world is made of meanings attributed to individuals Relativism: constructed realities vary
Epistemology		Dualism/objectivity	Modified dualism/objectivity	Non-dualism/subjectivity
		True findings	Probabilistically true results	Created findings
		Explanation generalisations: natural immutable laws	Explanation generalisations: provisional laws, open to revisions	Comprehension generalisation
Methodology		Experimental manipulative	Modified experimental manipulative	The empathetic interaction between scholar and object studied.
		Mostly induction	Mostly deduction	Induction
		Quantitative techniques	Quantitative techniques with some qualitative	Qualitative techniques
		Analysis by 'variables'	Analysis by 'variables'	Analysis by 'cases'
Axiology		Research is undertaken in a value-free way	Research is value laden	Research is value bound
		Researcher is independent of the data	Researcher is biased by world views, cultural experience and upbringing	Researcher is part of what is being researched and cannot be separated

Table 4.2 presents the ontological and epistemological stands of the research questions. Accordingly, the researcher believes that "the reality" exists independent of human conscience and experiences (carbon-intensive elements and the association between the variables – EC and design variables; EC and CC), and the reality can be modelled (induced by collecting data) probabilistically by employing appropriate research methods. However, the outcome does not represent the perfect reality but the best-conceived reality by human senses with the selected sample and the selected analysis techniques. These assumptions and belief

system can be matched with the ontological objectivism and the epistemological critical realism. Hence, the research fits better within the post-positivist paradigm.

Table 4.2: The ontological and epistemological positions of the research questions

The RQ	Ontology	Epistemology	Axiology	Comment
RQ3	Objectivism	Critical realism	Value laden	Carbon-intensive elements and the association between variables exist independent of human conscience and experiences; this knowledge is knowable only in an imperfect and probabilistic manner, and the research is affected by the values of the researcher.
RQ4	Objectivism	Critical realism	Value laden	
RQ5	Objectivism	Critical realism	Value laden	
RQ6	Objectivism	Critical realism	Value laden	
RQ7	Objectivism	Critical realism	Value laden	

4.3. Research Approach

The research approach explains the logic of the research, the role of the literature, the purpose the data collection and the data analysis (Sutrisna, 2009). There are two types of research approaches including: Deductive and Inductive (see Figure 4.1). Deductive research uses the existing body of knowledge to deduce a hypothesis and test the hypothesis by collecting data to affirm or contradict the existing knowledge. Popper (1975) argues that theories cannot be proved true but can only be falsified. Therefore, with deductive research, hypotheses are tested for falsification and if proved to be false, the hypothesis is rejected (Gray, 2014). Whereas, inductive research involves observing the reality and gathering data to develop a hypothesis or to create a theory (Sutrisna, 2009, Gray, 2014).

Accordingly, the existing body of knowledge around the problem considered plays an important role in determining the research approach. Deductive approach can be adopted if there is a wealth of knowledge concerning the problem investigated where a hypothesis can be formed from the existing knowledge, which will be eventually tested to confirm or to contradict the existing theory. On the other hand, a topic which is under researcher, up for debate and with little existing literature is well coped with an inductive approach (Saunders et al., 2009). In view of that, the literature on relationships between EC and design variables is scarce and the topic is relatively new. The absence of such theory or hypothesis for EC and design variable relationship dictated the research approach to be inductive. Hence, data were collected to understand the relationship between EC and design variables of buildings and to induce a model to predict EC during the early design stage.

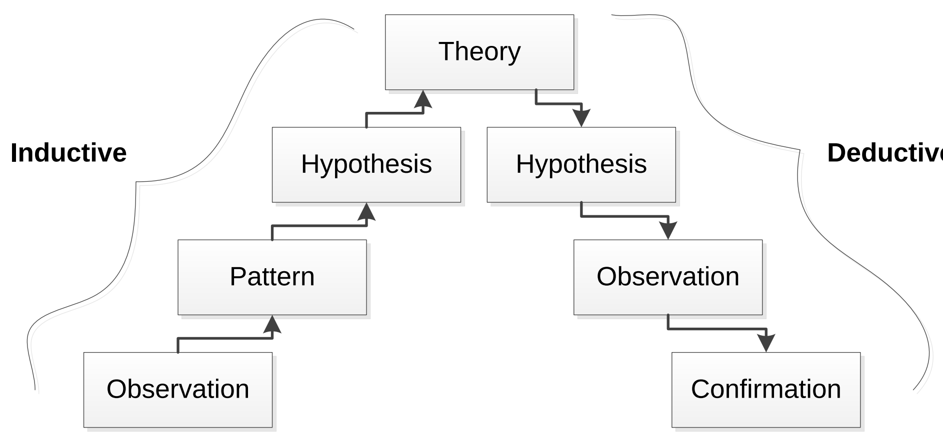


Figure 4.1: Inductive and deductive research

Adapted from: Trochim (2006)

Meanwhile, the relationship between the EC and the CC will also be analysed inductively, which is one of the research questions. The thought process of combining and comparing EC and cost emerged because both EC and cost are determined by the quantity of materials and plant (labour only for cost). However, the relationship between CC and EC is not explored in the literature (while EE and cost relationships are reported by Langston and Langston (2008)). Both CC and EC can be reduced simultaneously if a positive linear relationship is found to exist between CC and EC, which can be an important contribution to knowledge to the construction industry.

4.4. Research Strategy

There are several research strategies that one can employ in a research to answer the research questions such as experiments, survey, case study, action research, grounded theory, ethnography, archival research and history, to name a few. It is possible to combine research strategies to answer the research questions while some research strategies clearly belong to one or the other research approach (e.g., grounded theory and ethnography are strongly rooted in the inductive approach). Yin (2014) proposes three questions that would help to select a research strategy including (1) the type of research question, (2) the extent of control the researcher has over behavioural events, (3) the degree of focus on contemporary events. In line with that, Yin (2014) suggests that experiments, history and case study are appropriate to deal with 'how' and 'why' form of research questions while experiments require control of behavioural events and history does not deal with contemporary events. Hence, a combination of 'how' or 'why' form of research questions which does not require the control of behavioural events and focusing on contemporary events will employ case study research strategy. On the other hand, surveys and archival analysis are good at answering 'who', 'what', 'where', 'how many', 'how much' types of research questions (Yin, 2014).

The form of research questions indicates whether a research is exploratory, explanatory or descriptive. For instance, 'what' questions can be either exploratory or explanatory; 'who' and 'where' questions are descriptive; 'how' and 'why' questions are explanatory (Yin, 2014). The research questions of the study are of exploratory in nature as opposed to explanatory and descriptive. Hence, surveys and archival analysis can be shortlisted based on the form of research questions. Among the two (survey and archival analysis), the archival analysis is the better option as the research deals with tangible subjects which are buildings and data can be obtained from the archives of construction practices.

4.5. Research Choice

A research can be designed with only one method or multiple methods depending on the nature of the research questions. Saunders et al. (2009) named the former as ‘mono method’ and the latter as ‘multiple methods’. Multiple methods research choice is subdivided into two levels as presented in Figure 4.2. Accordingly, multi-method research choice involves considering either quantitative data collection and analysis techniques or qualitative data collection and analysis techniques separately. On the other hand, mixed-method research allows both quantitative and qualitative data collection and analysis techniques to be employed but the approaches are not combined in mixed-method research. However, mixed-model research allows combininig the qualitative and quantitative approaches in answering the research questions (Saunders et al., 2009).

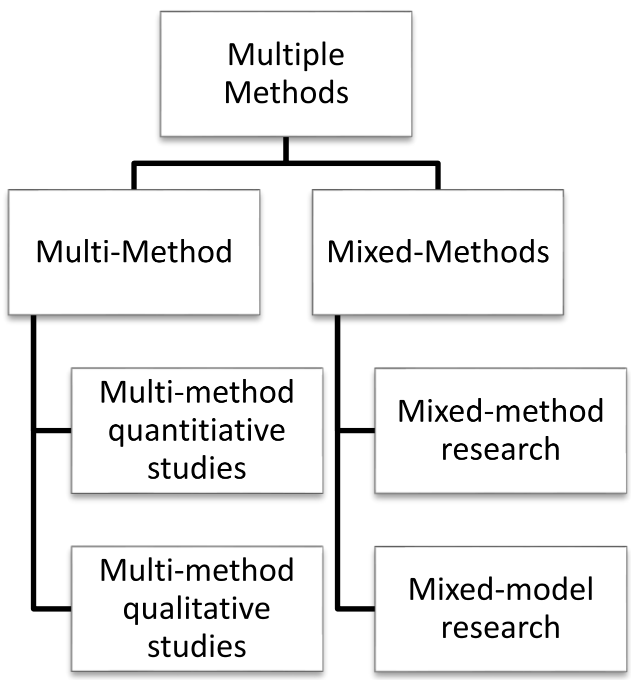


Figure 4.2: Subdivisions of multiple methods research choice

Modified from : Saunders et al. (2009)

As can be seen from the conceptual model presented in Section 3.10, the predictor variables of the model include qualitative design variables such as Finishes Quality Index and Services Quality Index. Hence, both qualitative and quantitative data collection and analysis techniques had to be employed to deal with quantitative and qualitative variables separately. Eventually, the outcomes of the qualitative data analysis will serve as inputs (finishes quality index and services quality index) for the model formulation. Therefore, the research adopts a mixed-model research as per the classification presented by Saunders et al. (2009).

4.6. Time Horizon

Cross-sectional studies provide a snapshot of a problem concerned at a particular time while longitudinal studies investigate an issue over a period of time. Accordingly, the research fits within the cross-sectional time horizon category as it tries to capture the relationships between EC of buildings and design variables at a given time.

4.7. Modelling Techniques

Table 4.3 presents a range of modelling techniques used in the development of parametric cost model in different studies. Among which statistical and Artificial Intelligence (AI) techniques are used repeatedly to predict cost at early stages of designs. In particular, regression and Neural Networks (NNs) can be found as the most preferred techniques for parametric cost model development. While many view regression and NN as competing model building techniques, Paliwal and Kumar (2009) suggest that better performance can be achieved if researchers use the models to complement each other rather than to compete. Nevertheless, it is useful to review the pros and cons of each technique critically to select the appropriate technique for the study.

Table 4.3: Techniques employed in the development of parametric cost models

Source	Technique	Type of Project
Hegazy and Ayed (1998)	Neural network	Highway projects
Cheng et al. (2009)	Combination of Genetic Algorithms, Fuzzy Logic and Neural Networks	Buildings
Kim et al. (2004a)	Regression, Neural network, Case-based reasoning	Residential
Adeli and Wu (1998)	Neural network	Highway construction
Seo et al. (2002)	Neural network	Product
Wilmot and Cheng (2003)	Linear model	Highway construction
Yu (2006)	No- linear mapping technique	Civil structures and buildings
Sonmez (2004)	Regression, Neural network	Residential, healthcare and commons buildings
Sawalhi (2012)	Fuzzy logic	Buildings
Karshenas (1984)	Regression - power exponential function	Multi-storey office buildings (steel framed)
Phaobunjong (2002)	Multiple linear regression analysis	Buildings (various functions)
McGarrity (1988)	Multiple regression analysis - power exponential function	Buildings (steel framed office buildings)
Kouskoulas and Koehn (2005)	Multiple linear regression analysis	Buildings (various functions)
Karanci (2010)	Multiple linear regression analysis, Neural networks, Case-based reasoning	Mass housing projects
Kim et al. (2004b)	Neural network model incorporating genetic algorithm	Residential buildings
Alshamrani (2016)	Multiple linear regression analysis	College buildings

4.7.1.Regression

Regression is the conventional technique of developing mathematical models based on the relationships between variables. Regression models present the relationships between input variables (independent variables) and an output variable (dependent variable) in the form of a mathematical equation. The relationships are captured using the correlation coefficient constant. For instance, cost of the project can be estimated as follows:

$Y = ax_1 + bx_2 + \ldots. + k$

Cost = f(GFA, Height, Wall Area,….)

Cost = a.GFA + b.Height + c.Wall Area + ……… + k

However, the above equation assumes linear regression while there can be non-linear relationships as studied by Yu (2006), McGarrity (1988) and Karshenas (1984). In fact, these authors claim that exponential function produces better results than linear regression.

Key advantages of regression models are that these are transparent, easy to understand, reasoning is possible and development procedure is less tedious compared to all other models. On the other hand, few studies highlighted that NN models outperform regression models (Kim et al., 2004a, Sonmez, 2004, Karanci, 2010). Another key issue with regression models is that the modelling becomes difficult with a large number of variables (Kim et al., 2004a). Therefore, the choice has to be made depending on the study objectives. Nevertheless, regression models have been used abundantly in construction cost estimating since 1970s due to its well-defined mathematical basis (Kim et al., 2004a). Similarly, regression could be a starting point for the relatively blooming carbon estimating field.

4.7.2.Neural Network

NN is an AI technique, which was developed, based on the metaphor of brain. It is actually inspired from brain functions rather than a replication (French et al., 2009). A simple NN architecture is illustrated in Figure 4.3. Accordingly, there can be more than one hidden layers and neurons depending on the complexity of the problem. Neurons are the basic processing unit of the NN with mathematical functions. Further, the outputs in Figure 4.3 could be the final outputs or inputs to another neuron (Turban et al., 2011).

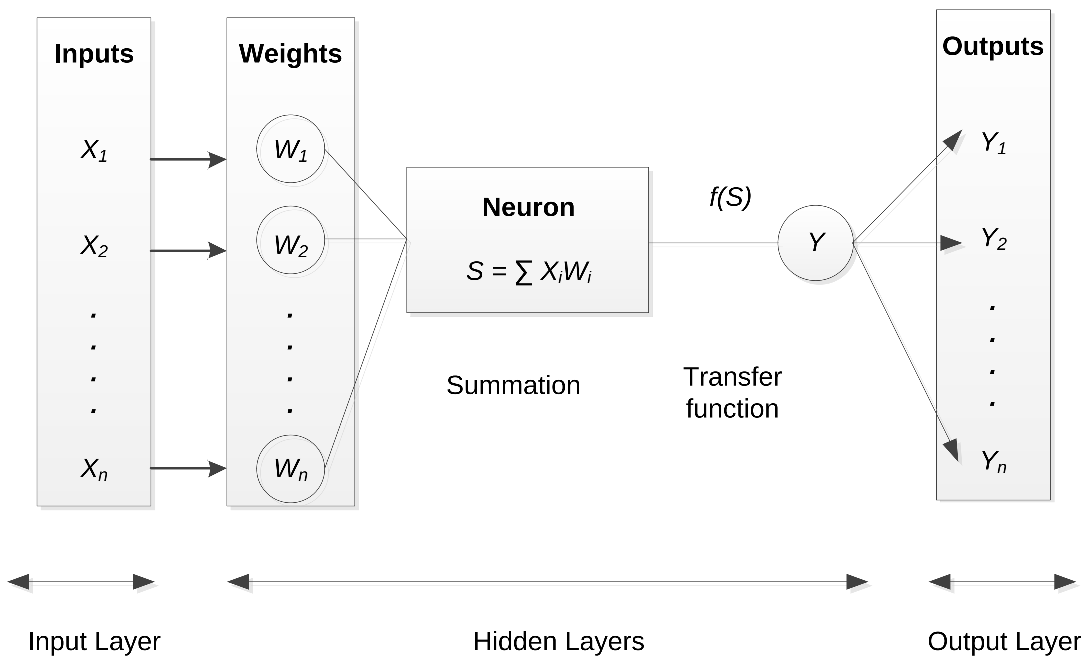


Figure 4.3: Neural network architecture

Source: Turban et al. (2011)

NN has the ability to learn from the dataset and capture the relationship that is either parametric or non-parametric which is considered as the major advantage of the model. Also, many studies witness the promising results of this type of models. Especially NN models mostly outperform other types of models in the test for 'closeness of fit' as NN can capture the best relationship between variables. However, a common criticism faced by this model is that it is a 'black box', lacking explanations on their capabilities. As with regression, the relationships are not transparent in NN and all the learning happens within the model itself. Nevertheless, according to French et al. (2009) sensitivity analysis illuminates the black box and gives the user an idea of the behaviour of each variable within the model.

Another shortfall of the model is that identifying the best NN model is challenging and time-consuming. Normally, the best model is chosen by trial and error method. A great example of this is, the best NN model was chosen among seventy-five NN models in the study of Kim et al. (2004a). Therefore, time is an important deciding factor in experimenting with this type of models. However, later few studies proposed that incorporating GA into NN systems eliminates trial and error method and allow optimisation of the system (Kim et al., 2004b, Cheng et al., 2009). On the other hand, GA said to be more successful with a large range of optimisation problems than a simple one (Michalewicz, 2013).

4.7.3.Fuzzy Logic

Human reasoning is mostly approximate rather than precise. Eventually, fuzzy systems are developed in a similar concept, thus, fuzzy logic provides a model for approximate reasoning rather than precise (Zadeh, 1994, Zadeh, 1975). Zadeh (1975) lists three distinct features of fuzzy logic as follows:

1. Fuzzy truth values are expressed in linguistic terms, i.e. true, very true, more or less true etc.
2. Imprecise truth tables
3. Rules of inference whose validity is approximate

Even though fuzzy systems claimed to mimic human behaviour, the decisions and methods of choosing the decisions are substituted by fuzzy sets and rules in which fuzzy rules operate base on if-then statements (Sawalhi, 2012). Fuzzy control systems are applied in a number of fields including industrial process control, medical diagnosis and securities trading (Lin and Lee, 1991) while very few studies (Cheng et al., 2009, Sawalhi, 2012) are found in the construction context. Also, fuzzy logic is commonly combined with other types of models, especially with NN (Lin and Lee, 1991, Cheng et al., 2009).

4.7.4.Case-Based Reasoning

While most identify Case-Based Reasoning (CBR) as an AI technology some claim that it is only a methodology (Watson, 1999). CBR works based on the lessons learned from the past projects that lie within the database. It matches the features of the input data with that of the historical data in the database and provides a tailor-made solution to the problem with reasoning (Aamodt and Plaza, 1994, Gupta, 1994, Watson and Marir, 1994, Xu, 1994). Xu (1994) defines two types of CBR namely, a) Problem-solving systems: provide new solutions by modifying historic case solutions and b) Interpretive systems: evaluate and justifies new case based on the similarities and differences with the historic case. Xu (1994) also mentions that both the systems are required to solve most real-world problems.

CBR logic involves four main steps as follows (Aamodt and Plaza, 1994, Xu, 1994):

- **Retrieve** the most similar case/s
- **Reuse** the knowledge in the similar case/s to solve the problem
- **Revise** the proposed solution
- **Retain** the proposed solution as a lesson learned for future problem solving

The process is illustrated in Figure 4.4.

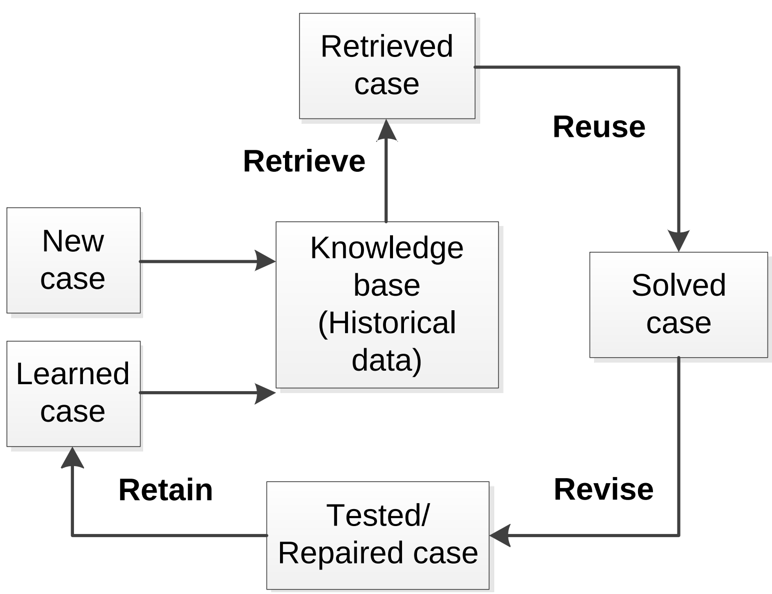


Figure 4.4: Case base reasoning logic

Modified from: Aamodt and Plaza (1994)

A key benefit of CBR is that it operates like a human mind or an estimator who performs the estimating function from his/her previously acquired experience. Karanci (2010) believes that CBR is capable of estimating the cost of conceptual designs in a similar fashion to a human mind by using the data stored in its knowledge base. Hence, CBR tends to gain superiority over all other models and outperforms other models in most of the tests which are supported by Kim et al. (2004a) and Karanci (2010). However, Karanci (2010) also pointed out that CBR demonstrates better results in the closeness of model fit due to the ability to locate the same project from the case library. Another benefit is that the CBR stores the new solved project in its case library as a historic project accumulating the library. However, the predictions could deviate from the actual cost when the project is implemented which is not usually captured. Nevertheless, it is possible to update the library if needed (Aamodt and Plaza, 1994).

Other benefits include (Gupta, 1994, Xu, 1994):

- suitable for the domains which are experience rich but knowledge poor (many past cases)
- efficient reasoning on the outcome
- unique explanation capability
- Faster knowledge acquisition

However, one drawback of CBR is that the solution or the prediction is entirely dependent on the case library. For instance, if the closest similar case does not represent a satisfactory match, it could lead to untrustworthy reuse solution. Further, other expert systems are preferred over CBR if the sample size (past projects in the knowledge base) is small (Gupta, 1994).

4.8. Data Collection Techniques

EC, CC and design variables of building are the required data for the research (detailed discussion on the data requirement is presented in Section 5.2). Variables selected for the study include GIFA, the number of storeys, average storey height (or building height), façade area, wall to floor ratio, circulation ratio, the number of basements, finishes quality and services quality (see, Section 4.9.3 and 6.4.1 for more details on the selection process of variables). Accordingly, historical project data (Bill of Quantities and layout drawings) were collected to obtain quantitative design variables and to estimate EC and CC. On the other hand, data to develop objective finishes and services quality indices were collected using qualitative techniques and these nominal variables were methodically converted into ordinal variables to perform the regression analysis (see, Section 4.9.2 for an explanation on the level of measurement of variables). Hence, data collection techniques used for the historical project data collection and the development of finishes and services quality indices are presented separately.

4.8.1.Data Collection Techniques Used for Historical Project Data Collection

a) Data Sources

The problem entailed by historical project data is that the researcher having less control over the data contents, quality and quantity which is unavoidable in the research context as pointed out by Saunders et al. (2009). However, data are more objective and free from external influences or opinions as data are extracted from documents or repositories (Phaobunjong, 2002). Data in a research can take two forms: primary and secondary. Primary data refers to the data that are specifically obtained for the study and the secondary data refers to the data that are collected for a different purpose and are reused by another study (Hox and Boeije, 2005). In the present context, primary historical project data can be collected from construction consultancy practices and secondary historical project data can be collected from public databases. Both primary and secondary sources of data were surveyed to obtain as much data as possible. The identified key secondary databases included BCIS (RICS, 2016) and WRAP EC Database (WRAP and UK-GBC, 2014). BCIS is an online cost database maintained by RICS and WRAP EC Database is maintained by UK-GBCSL and WRAP. Both databases designed with multi-faceted search facilities to cater varying user requirements and contain advanced search options to filter the most appropriate data to suit different study requirements. In addition to these data sources, another type of data is also required for the research to facilitate CC and EC estimating which are data books.

These are referred to as ‘supporting data’ in this research context and include the ICE (Hammond and Jones, 2011), UK Building Blackbook (Franklin & Andrews, 2011), EC data from manufacturers and Spon’s price book (Davis Langdon Consultancy, 2014).

b) Sample selection

The sample size is a key decision to be made to achieve statistical credibility and generalise findings. Patton (2015) claims that the sample size of a particular study is dependent on the research questions, time and resource availability. It is commonly said that the sampling error will be small if the sample size is large. However, there are also problems with extremely large samples due to diminishing returns where larger sample size results in smaller benefit at a higher cost of time and money (Miles and Shevlin, 2001). Therefore, there are methods to assist in sample size estimation for a given experiment such as central limit theorem and power analysis.

According to the Central Limit Theorem (CLT), the mean of a sufficiently large number (i.e. 30 or more) of independent random variables (each with finite mean and variance) will be approximately normally distributed. However, the words ‘sufficiently large’ are still debatable. Rule of thumb suggests that a sample size of a minimum of 30 and a maximum of 500 is appropriate for most research (Roscoe, 1975). However, Chakrapani (2011) highlights that the minimum sample size of 30 should not be taken for granted as the research context also plays a major role in determining the sample size. Others suggest that a minimum of 50 subjects is appropriate for regression analysis (Kelly et al., 2012, VanVoorhis and Morgan, 2007) while Miles and Shevlin (2001) recommend calculation of sample size using power analysis (which takes into account of the value of alpha or the significance level, the effect size, the power and the number of predictors in the model).

Similar research on cost model development (cost is predicted using design and project variables) have obtained a larger sample of up to 2827 historical projects as a result of publicly available databases (Kim et al., 2004a, Adeli and Wu, 1998, Seo et al., 2002, Wilmot and Cheng, 2003, Sawalhi, 2012). On the other hand, other studies were reported to have smaller sample ranging from 18 to 41 (Karanci, 2010, Kouskoulas and Koehn, 2005, McGarrity, 1988, Karshenas, 1984, Sonmez, 2004, Cheng et al., 2009, Hegazy and Ayed, 1998). VanVoorhis and Morgan (2007) explain that time, access to samples and cost involved are some of the practical limitations in obtaining a larger sample. Therefore, studies with smaller sample size are not anomalous. Consequently, a minimum sample size of the study was set as 30 due to the lack of EC databases and insufficiency of the existing EC databases.

4.8.2. Data Collection Techniques Used for the Development of Design Quality Indices

Finishes and services quality levels vary a lot in commercial buildings compared to domestic buildings. Hence, an objective index for finishes and services quality of office buildings needed to be developed to be integrated into the model. Finishes index covers the quality level of the internal walls, floor and ceiling finishes of office buildings. Subsequently, literature was surveyed to assess the adaptability of the existing finishes and services quality indices for the study.

a) Finishes Index Development

Kouskoulas and Koehn (2005) developed a tailor-made overall quality index of the building to be integrated into their cost model. The quality index identifies eight (8) components of buildings namely use of the building, design load, exterior wall, plumbing, flooring, electrical, HVAC and elevator. Each component is categorised into four quality levels such as Fair, Average, Good and Very Good where each category is defined. The overall quality of the building is derived by calculating the mean quality index from all eight components. Table 4.4 presents the quality index of flooring developed by Kouskoulas and Koehn (2005). Problems with the flooring quality index propsoed by Kouskoulas and Koehn (2005) are that they were confined to only a few types of floor finishes. In addition, ceiling finishes were included with electrical component while wall finishes were not included in the quality index at all. Hence, the quality index developed by Kouskoulas and Koehn (2005) is inadequate.

Table 4.4: Floor finishes index developed by Kouskoulas and Koehn (2005) (Modified from: Kouskoulas and Koehn (2005))

Component	Fair	Average	Good	Very Good
Flooring	Resilient, ceramics	Resilient, ceramics and terrazzo	Vinyl, ceramic, terrazzo	Rug. Terrazzo, marble

Indices developed by AACE (formerly known as the Association for the Advancement of Cost Engineering) International (2015) and Kim et al. (2004a) follows the second type of index. AACE International (2015) proposed an interior finishes quality index ranging from one (1) to ten (10) which is listed in Table 4.5. This index also suffers from non-objective definitions of the quality levels similar to the quality index developed by Kouskoulas and Koehn (2005). This type of quality index increases ambiguity in ascertaining the quality level of finishes. Another drawback of the index is the possibility of having different combinations of traffic in a building was not considered. On the other hand, Kim et al. (2004a) classified finishes into five grades, ranging from Grade I to Grade V. Residential buildings for rental that are owned by the public sector are classed as Grade I (due to the fact that they are to be government-supported non-profit construction). Private owned residence are categorised into poor (grade II), average (grade III), good (grade IV), and luxury (grade V) where luxury grade implies buildings with imported finishes. The index developed by Kim et al. (2004a) also has issues with subjectivity even though, it has more clarity than the index developed by AACE International (2015).

Table 4.5: Interior finishes quality classification of a cost model (AACE International, 2015)

1 – functional, unattractive	6 – moderate duty, attractive
2 – functional, passable	7 – heavy duty, passable
3 – light duty, passable	8 – heavy duty, attractive
4 – light duty, attractive	9 – moderate duty, luxury
5 – moderate duty, passable	10 – heavy duty, luxury

The model developed by Sawalhi (2012) integrated finishes quality by capturing the type of external plastering while disregarding internal finishes. Specifically, the model of Sawalhi (2012) will not be useful in case of curtain walling which is currently the most common type of facade of office buildings in the UK.

Mainly, two types of quality indices were identified in the literature. First type of indices list the common types of finishes and classify the identified finishes into defined quality levels (See, Kouskoulas and Koehn, 2005); second type of indices consist of different quality levels defined as numerical values which are vague and not objectively defined (See, AACE International, 2015, Kim et al., 2004a). Further, finishes indices surveyed from literature lack objectivity and are not comprehensive. Hence, a tailor-made finishes quality index for the study was developed due to the inadequacies found in the existing finishes quality indices. Between the two types of finishes quality indices identified in the literature, the first type was selected to be developed (identifying the common types of finishes and classifying the identified finishes into prescribed quality levels) as it was more objective than the second type. Consequently, three levels of finishes quality were established:

1. Basic finishes (Finishes Index – 1)
2. Moderate finishes (Finishes Index – 2)
3. Luxury finishes (Finishes Index – 3)

The reason for having only three levels of quality is because the model facilitates early design stage estimating and during early design stage, detailed specification is seldom thought about. Therefore, it is easier to choose the finishes quality of the building from three levels rather than more detailed levels.

Further, literature provides no evidence of any verification process employed in the development of these finishes quality indices. However, verification process improves the rigour of the proposed finishes index and eliminates the bias of the researcher as it involves the judgement of more than one individual. Therefore, a two-step process was adopted in developing the finishes quality index for the study where a conceptual finishes index was developed and was verified through experts (more details follow in Section 5.6.3 and 5.7.3).

Clayton (1997) suggests that Delphi technique is appropriate when seeking the consensus of the experts on content validity, which in this research context is the validation, and verification of the proposed finishes quality levels in office buildings. Further, Delphi technique allows rigorous and systematic data collection and dissemination without the need of the experts to travel and meet as a group at a particular time and a place (Clayton, 1997). Further, responses are isolated (independent of the other experts) and anonymised which is an advantage of Delphi-based approach over other group based decision-making such as Nominal Group Technique and Interacting Group Method (Clayton, 1997, Van de Ven and Delbecq, 1974).

An expert is someone who possesses the knowledge and experience in a particular field (Oxford Dictionary, 2010). In this context, construction professionals who are competent in early stage cost advising and having experience in office building projects were considered as experts for the Delphi-based expert forum formed for the study. Further, being a chartered surveyor and having a minimum of ten (10) years of professional experience are two key criteria required by RICS (2009) for a person to qualify to be considered as an expert. Accordingly, construction professionals with a chartered qualification and with more than 10 years of experience in the UK construction industry were selected using purposive or judgmental sampling technique. Purposive sampling enables the researcher to select the cases or respondents based on the research questions and the knowledge about the population (Polit-O'Hara et al., 2001, Saunders et al., 2009). Purposive sampling is usually adopted in studies with very small samples, though Saunders et al. (2009) warn that such samples should not be considered to be statistically representative of the population. The use of a Delphi-based expert forum in the study aims at verifying a conceptual finishes index developed for the study, which is not a study objective or research question, but it contributes towards the model development, which is a key objective of the study. Hence, purposive sampling technique was adopted to select experts for the expert forum to verify the conceptual finishes index due to the time constraint of the study.

The size of the panel depends on the purpose of the study and availability of resources (Patton, 2015). A panel size of 15 to 30 is suggested for a homogeneous population and 5 to 10 is suggested for a heterogeneous population as a rule of thumb (Clayton, 1997). The reason for proposing a larger panel size for homogeneous population compared to heterogeneous population is because it can be expected that experts in a homogeneous population to be like-minded, hence, requiring a larger panel size to ensure an acceptable representation of the population to validate the content. In contrary, a heterogeneous population will consists of people from different disciplines, hence, a smaller panel size will be adequate to scrutinise the content due to the diversity of the experts. Accordingly, a panel size of five (5) to ten (10) was decided to be employed as the construction industry is composed of a heterogeneous population (such as Architect, Engineer, QS and the like).

Further, it is also recommended that at least one opportunity is given to the respondents to re-evaluate their responses based on the examination of the response of the group (Clayton, 1997). Further, Williams and Webb (1994) state that the researcher must be aware of when to stop collecting data which is when the consensus among the experts' judgement is achieved and most studies fall short in defining consensus. The consensus in the research context was considered to be achieved when four (4) of the five (5) respondents agreed on a particular quality level.

Despite the benefits, Clayton (1997) identified following limitations in Delphi technique:

1. The background of the experts might have an influence on the judgement, which is beyond the control of the researcher.
2. Personal and profession obligations might limit the experts to invest more time and effort in arriving at rational judgements.
3. It cannot be measured whether the experts, judge based on their experience and work towards consistency of their previous judgement or if they are pressurised to conform to the group's judgement due to an iterative process.
4. The value of the information presented is subjective to the reader and might be limited due to the constraints in panel selection and the background of the experts.
5. Biases of the researcher in arriving at a final decision.

Accordingly, the developed finishes quality index is custom made for the study and influenced by the above-mentioned factors. Therefore, the readers should be mindful of these limitations when trying to re-use the proposed finishes quality index.

b) Services Index Development

Building services EC can account up to 25% of the total EC emissions (Hitchin, 2013) while the cost of services can contribute up to 40% of total building cost (RICS, 2016). However, building services are paid less attention during early stage cost planning and estimating of projects due to the complex nature of the element (RICS, 2014), hence, little literature evidence was found on services quality indices. Kouskoulas and Koehn (2005) developed quality indices for plumbing, electrical, HVAC and lift installations separately under four quality levels such as Fair, Average, Good and Very Good (see, Table 4.6). Plumbing, electrical and HVAC quality were defined in a way that is subjective to the user of the model where a question arises as to 'what is an average quality?'. Further, both Good and Very Good quality levels of plumbing and HVAC were defined as 'above average quality' causing more confusion in human judgement.

AACE International (2015) has developed a services index (see, Table 4.6) similar to finishes index presented in Table 4.5 for mechanical and electrical services, which have the same issues of subjectivity as discussed in the finishes quality index development. On the other hand, Sawalhi (2012) captured inputs like the number of elevators and volume of HVAC to predict the cost where other types of building services such as electrical, plumbing, protective installations, communication and IT had not been considered.

Table 4.6: Services index developed by Kouskoulas and Koehn (2005) (Modified from: Kouskoulas and Koehn (2005))

Component	Fair	Average	Good	Very Good
Plumbing	Below average quality	Average quality	Above average quality	Above average quality
Electrical	Fluorescent light, poor quality ceiling	Fluorescent light, average quality suspended ceiling	Fluorescent light, above average quality ceiling	Fluorescent light, excellent quality ceiling
HVAC	Below average quality	Average quality	Above average quality	Above average quality
Elevator	Minimum required	Above required minimum	High speed	High speed deluxe

The existing services quality indices show inadequacies and could not be adapted to the study as the historical project data obtained for the study had limited or no detailed specification of Services. Further, the nuances of Service quality were not explored due to the lack of detailed measurements and specification of the sample projects and limited EC data on Services. In addition, the models are intended to assist early design stage estimating. Hence, a service quality index that represents the provision of different services (sub-elements of Services as per the NRM such as Sanitary Installations, Services Equipment, Drainage Installation and the like) is considered appropriate to meet the study requirements and to cater the estimating need of early stages of designs. Consequently, the quality levels of Services presented in price books were surveyed to develop a simple, yet an objective, service quality index for early design stage estimating (see, Section 6.9).

4.9. Data Analysis Techniques

The collected data have to be analysed quantitatively and qualitatively as discussed in Section 4.5 in light of answering the research question presented at the beginning of the chapter. Accordingly, the research questions include identifying the carbon-intensive elements of office buildings; investigating the relationships between the EC and design variables of buildings; investigating the relationship between EC and CC of buildings; and finally, formulating a model to predict EC using design variables of buildings. Hence, the data analysis section is divided into three subsections namely analysis of carbon and cost hotspots, analysis of relationships between variables and the formulation of the EC prediction model.

4.9.1. Analysis of Carbon and Cost Hotspots

Identifying carbon and cost-intensive elements or hotspots of buildings depends on the definition of hotspots. The Pareto Principle was adopted to define the carbon (or cost) hotspots due to its popularity and applicability, especially in, economics, business and management related areas. Vilfredo Pareto (1848 – 1943), an economist, found that the 80% of the wealth of his country was owned by the 20% of the people. Then, Pareto applied the same theory to other states like Russia, France and Switzerland and found the same results. However, it was in 1940s Joseph Juran (1904 - 2008), an American engineer, recognised the 80:20 theory and named it after Vilfredo Pareto. Pareto Principle defines that 80% of the results (or consequences) are attributable to 20% of the causes which implies an unequal relationship between the inputs and the outputs (Koch, 2011, Delers, 2015).

Munns and Al-Haimus (2000) noted that seminal texts in the cost management literature (Ashworth and Perera, 2015, Seeley, 1996, Ashworth and Skitmore, 1983) approving the applicability of Pareto Principle to identify the cost significant items. The works of Munns and Al-Haimus (2000) and Tas and Yaman (2005) are examples of embracing 80:20 Pareto Principle to identify the cost significant items in a BoQ and eventually, developing prediction models using cost significant modelling technique. Hence, it is evident that 80:20 Pareto Principle is widely accepted as the popular method of capturing cost significant items in a BoQ. However, to identify the cost significant items, the BoQ items have to be grouped (to minimise complexity by reducing the number of items) according to the work packages (trades) or functional elements as done in previous studies (See, Munns and Al-Haimus, 2000, Tas and Yaman, 2005). Accordingly, items were grouped as elements as the focus the study was to aid design decision-making during the early stages of design as opposed to detailed stages of design where trade wise analysis would have been appropriate otherwise. Further, the grouping of elements prescribe in NRM standards (RICS, 2012a) was adopted in the study as it is the prevailing standard of measurements in the UK at present.

Consequently, it can be hypothesised that 80% of the EC is emitted by 20% of the building elements. The building elements responsible for 80% of EC emissions are referred to as the carbon hotspots in the context of the research. Even though 80:20 is accepted as the universal ratio, the Pareto Principle neither dictates that the 80:20 ratio is applied to all situations nor should the two figures add up to 100 (say, it could be 90:50 or 80:30). Therefore, this ratio was tested in the case of the relationship between EC (and cost) and building elements. Figure 4.5 illustrates the process followed in identifying the EC and CC hotspots.

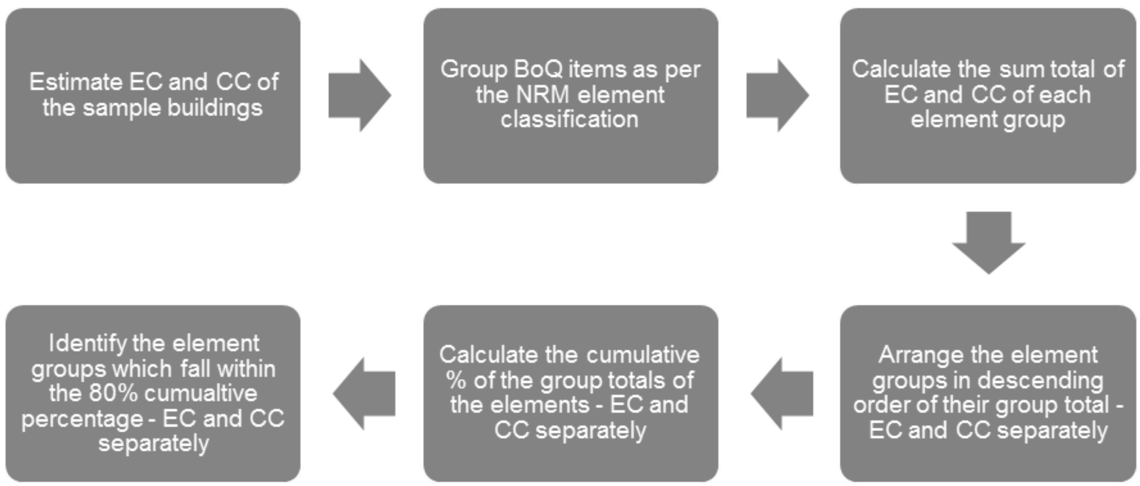


Figure 4.5: The process of identifying EC and CC hotspots of the sample buildings

As per Figure 4.5 EC and CC were estimated using the UK Building Blackbook, ICE, manufacturers' data and price books for the sample buildings and the BoQ items were grouped in accordance with the NRM elements classification. Sum total of EC and CC of Each element group were obtained and the element groups were arranged in a descending order of their group totals. Cumulative percentage of the element group totals were calculated to identify the elements contributing up to 80% of the total EC and total CC separately, which are referred to as the carbon or cost significant elements or the hotspots of office buildings.

4.9.2. Analysis of the Relationship between Variables

The technique used to analyse the relationship between variables is the correlation coefficient followed by the examples set by previous studies including Langston and Langston (2008) and Luo et al. (2015). Langston and Langston (2008) analysed the relationship between EE and CC while Luo et al. (2015) investigated the relationship between the EC and the number of storeys of buildings. The correlation coefficient is denoted by ‘r’ and it measures to what extent two variables are linearly related. Equation 4.1 presents the formula to calculate the correlation coefficient. Miles and Shevlin (2001) advise that it is always useful to check if the correlation is statistically significant. This implies that the correlation between two variables is unlikely to be zero if the probability value associated with the correlation is less than 0.05 (usually referred to as the ‘significance’ value).

Equation 4.1: Formula to calculate correlation coefficient

$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$

Where x and y are the two variables considered, S_x is the standard deviation of variable x, S_y is the standard deviation of variable y and S_{xy} is the covariance of x and y of the sample.

Further, the correlation can take any value ranging from -1.00 to +1.00, -1 represents perfect negative linear correlation while +1 implies perfect positive linear correlation. The sign of the correlation coefficient dictates the direction of the relationship while the magnitude conveys the strength of the relationship between two variables. There are guides available in the literature to interpret the size of the effect of correlation. Cohen (1988) proposed an absolute value of 0.1 represents a small effect, 0.3 represents a medium effect and 0.5 or more represents a large effect. Evans (1996) suggests values between 0 and 0.19 to be “very weak”, between 0.20 and 0.39 to be “weak”, between 0.40 and 0.59 to be “moderate”, between 0.60 and 0.79 to be “strong” and 0.80 and 1.0 to be “very strong”. Even though these benchmarks are useful, Field (2013) suggests that it is important to interpret the correlation in the context of the research.

Furthermore, a conclusion about causality cannot be made from the correlation between two variables. Causality between two variables simply means that the observed change in one variable is caused by the other variable. However, the existence of a correlation between two variables does not prove causality, instead, the correlation between the two variables could have been caused by a third variable, which could be unknown or not studied. In addition, correlation does not convey the direction of causality, that is which variable causes the change in the other (Field, 2013). Hence, it is important to prove causality and the direction of causality between the two variables considered in order for the correlation to be meaningful.

There are four techniques to calculate the correlation coefficients as follows:

1. Pearson's correlation coefficient
2. Spearman's correlation coefficient
3. Kendall's tau
4. Biserial and point-biserial correlation

In order to understand the use of each technique, it is important to understand the different levels of measurement of variables such as:

1. Nominal – a numerical scale used to identify different categories of a variable, but the magnitude of the number does not have any value (i.e. Male – 0, Female – 1)
2. Ordinal – a numerical scale used to represent an order in the categories of a variable unlike nominal measurement, however, does not tell anything about the difference between the two categories (i.e. strongly disagree -0, disagree – 1, agree – 2, strongly agree – 3)
3. Interval - a numerical scale which has the same intervals throughout the scale (i.e. temperature measurements)
4. Ratio – a numerical scale which has a true and meaningful zero point (i.e. scores in a test)

Pearson's correlation is used only if variables are measured at the interval level of measurement. Spearman's correlation and Kendall's tau are known as non-parametric correlation, which can be used with ordinal level variables (ranked data). Kendall's tau is used for a small sample with a large number of tied ranks.

Biserial or point-biserial correlation is used when one of the variables is dichotomous (categorical variables with two categories). Accordingly, appropriate correlation technique was used to analyse the correlation of each pair of variables as presented in Table 4.7.

Table 4.7: Suggested correlation techniques depending on the level of measurements of variables

Variables pairs	Measurement scale	Correlation Technique
GIFA/ EC per GIFA	Ratio/Ratio	Pearson
Building height/ EC per GIFA	Ratio/Ratio	Pearson
Average storey height/ EC per GIFA	Ratio/Ratio	Pearson
Wall to Floor ratio/ EC per GIFA	Ratio/Ratio	Pearson
Circulation space ratio/ EC per GIFA	Ratio/Ratio	Pearson

4.9.3. Formulation of the EC Prediction Model

Formulation of the EC prediction model was preceded by sub-processes including identification of independent variables of the model (or the predictor variables), development of finishes and services indices, development of the dataset for the formulation of the model and the examination of data which are discussed herein.

a) Identification of the variables or model predictors

Design variables or model predictors were identified through an extensive literature search for both CC and EC. Evidences were found in the literature for CC and design variable relationship while literature supporting EC and design variables relationships was scarce (See Section 3.10.1). However, the EC model was conceptualised based on the well-established CC and design variable relationships due to the fact that CC and EC are affected by the same design variables (though the strength and direction of relationship could be different which is yet to be found and this gap became the driver of the research). Even though the conceptual model is presented using the key design variables that are likely to be available during the early stages of design based on the literature, another step of verification is designed to identify the most influential design variables in predicting EC (and CC). This verification was performed using the data collected for the study. The design variables affecting the carbon (and cost) hotspots were identified through the carbon (and cost) hotspot analysis (see, Table 6.5 in Section 6.2.1) though, this type of verification is not found in the literature. However, the selection of the most influential design variables is beneficial in terms of fitting the model with only (statistically) significant variables.

b) Development of the finishes and services quality indices

Two of the cost influential qualitative design variables include finishes quality and services quality of the buildings. As a result, quantitative indices were developed for finishes and services quality of office buildings by collecting data through a Delphi-based expert forum and a document review respectively as explained in Section 4.8.2. Experts verified the conceptual finishes quality index in an iterative process until consensus is reached. The conceptual finishes index was then modified by incorporating the comments of the experts and presenting the modified finishes index in the next round, repeating the process until consensus was reached among the experts. In this way, the finishes quality index for the study was developed. On the other hand, services quality classification adopted in the price books was content analysed and matched with the specification of the study sample and an applicable services quality index was developed by modifying the standard classifications in the price book.

c) Development of the dataset for the formulation of the model

An important step in the process of the model formulation is the development of the dataset. The dataset of the study was developed with the information obtained from historical projects, the EC and cost estimates of the projects and the finishes and services quality indices. The raw data (quantitative design variables) and the processed data (CC, EC, Finishes Index and Services Index) of the sample buildings were entered into a spreadsheet to facilitate analysis. The information that was captured and their unit of measurements are as follows:

1. Project Identification code
2. Frame type – Concrete, Steel, Timber, Masonry or Hybrid
3. GIFA – m^2
4. Number of storeys – Nr
5. Number of basements – Nr
6. Average story height - m
7. Building height – m
8. Façade area – m^2
9. Wall to Floor ratio (Façade area/GIFA) – %
10. Circulation Ratio (Non-usable area/GIFA) - %
11. Finishes Index - no units
12. Services Index - no units
13. CC - £1000s
14. EC - tCO_2
15. CC per GIFA - £/m^2 GIFA
16. EC per GIFA - $kgCO_2/m^2$ GIFA

d) Data examination

Prior to data analysis, a careful examination of data ensures better model development (Hair, 1998). Therefore, data (values of design variables, CC and EC) were examined by producing histograms and box plots. Histograms give a visual indication of the normality of the distribution, though, histograms can be misleading when the sample size is small (Miles and Shevlin, 2001). However, box plots are useful in identifying non-normality even when the sample size is small. In addition to that, outliers and extremes in data can be spotted through box plots. Therefore, histograms and box plots were used to study the distribution, normality and outliers in the data for each variable of the model (dependent and independent).

Further, descriptive statistics were used to describe the variables and to discover problematic data distributions that deviate significantly from a normal distribution, which is a key assumption in regression analysis. Hence, mean, standard deviation, minimum, maximum and skewness of the variables were calculated and interpreted. A Linear relationship between dependent and independent variables is another key assumption in regression. Scatterplots were produced to identify linear relationships between the dependent variable and all the independent variables and correlations were calculated for each pair – CC, EC, CC per GIFA and EC per GIFA separately. Further, log transformations were applied to the values of the variables which deviated from normality and linear assumption to make the variables comply with regression assumptions (Field, 2013).

e) The model development process

The process of formulating the EC prediction model follows the basic structure of cost modelling research which involves three main stages including conceptualization of the model, model formulation (by collecting data) and validation of the model (Ashworth & Perera, 2015). Figure 4.6 illustrates the basic process and sub-process involved in the model development. Accordingly, the basic process involves the conceptualization of the model through an extensive literature review (see Section 2.10); estimating the model parameters by obtaining a study sample of historical office building projects; and validating the refined model for applicability and generalisability of the model. In terms of the sub-processes presented in Figure 4.6, the EC of buildings were estimated using the supporting data as explained in Section 4.8.1 and grouped as NRM compliant element form (refer to Section 5.3.1 for details on different datasets presented in Figure 4.6 – Dataset 1, 2 and 3). Finishes and services quality indices were developed through a Delphi-based expert forum and price books as discussed in Section 4.8.2. The finishes and services quality level of each building in the sample were identified quantitatively using the finishes and service quality indices developed for the study (refer to Section 5.8 and 5.9 for the discussion on the development of finishes and services quality indices). Consequently, the model dataset was developed by collating design data, CC and EC data in a spreadsheet. This was followed by the statistical analysis, which is discussed in the next subsection (f).

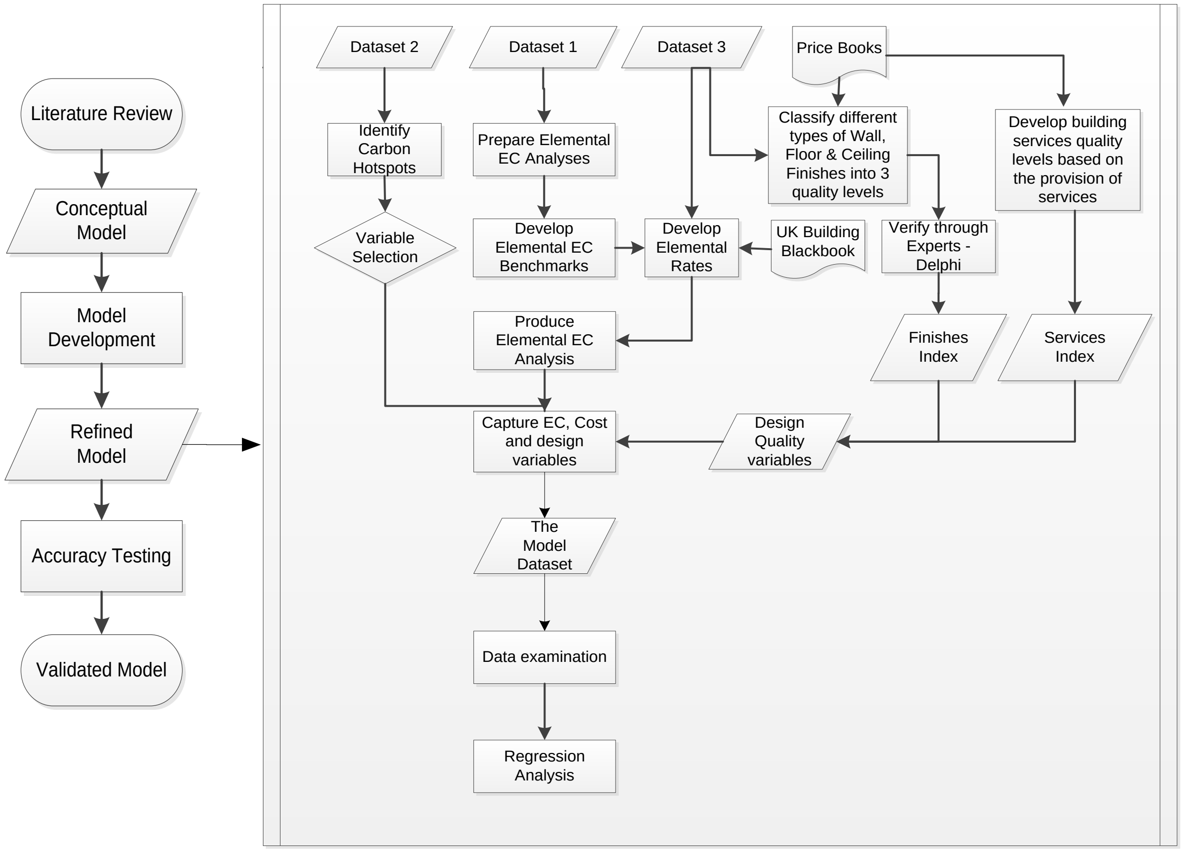


Figure 4.6: The model development process

f) Multiple regression analysis

Multiple regression analysis was selected over the other decision approaches due to its well-defined mathematical basis (Kim et al., 2004a) and transparency. Even though NNs were found to be outperforming regression models by many scholars, NN was not selected because of its 'black box' nature making it difficult to interpret the outcome, update and maintain the database. Further, identifying the optimal network is a time-consuming process compared to other models. CBR was rejected due to small sample size which will affect the trustworthiness of CBR as pointed out by Gupta (1994) (see, section 4.5 for detail review of the modelling approaches).

Consequently, the conceptualised EC regression model can be presented as follows:

Equation 4.2: Multiple regression model of the research

$$y = \beta_0 + \beta_1 . x_1 + \beta_2 . x_2 + \cdots + \beta_n . x_n + \varepsilon$$

Where, $\boldsymbol{y}$ is the EC *per GIFA* of the building considered, $\boldsymbol{\beta_0}$ is the intercept, $\boldsymbol{\beta_{1 \ldots n}}$ represents the coefficient parameters of the predictor variables $\boldsymbol{x_{1 \ldots n}}$ where $\boldsymbol{x_{1 \ldots n}}$ represents the influential design variables and $\boldsymbol{\varepsilon}$ represents the error term.

However, the multiple regression equation is presented as follows without the error term:

Equation 4.3: Multiple regression equation of the research

$$E(y) = \beta_0 + \beta_1 . x_1 + \beta_2 . x_2 + \cdots + \beta_n . x_n$$

Here, the error term is assumed to be zero.

Yet, the model developed in the research can be presented as follows:

Equation 4.4: Estimated multiple regression equation of the research

$$\hat{y} = b_0 + b_1 . x_1 + b_2 . x_2 + \cdots + b_n . x_n$$

Where, $\boldsymbol{\hat{y}}$ is an estimate of $\boldsymbol{y}$, $\boldsymbol{b_0}$ is an estimate of $\boldsymbol{\beta_0}$, $\boldsymbol{b_{1 \ldots n}}$ are estimates of $\boldsymbol{\beta_{1 \ldots n}}$. The coefficients of Equation 4.4 are estimates of the parameters presented in Equation 4.3, which will be calculated from the study sample.

Further, key assumptions underlying regression analysis including normality and linearity assumptions, have to be checked to ensure the validity of the derived regression equation. Key assumptions made in the multiple regression analysis include (Miles and Shevlin, 2001, Field, 2013):

1. The dependent variable is normally distributed - histograms and box plots along with descriptive statistics (skews) will be used to test this assumption (as discussed in subsection (d) of Section 4.9.3).
2. Relationships between dependent and independent variables are linear – this can be checked through scatterplots (as discussed in subsection (d) of Section 4.9.3) before performing the regression analysis and residual plot will also be

used to check this assumption. This assumption is considered to be met if the residuals in the standardised residual plot are randomly distributed.

3. There is little or no multicollinearity between data - correlation matrix was produced for all the independent variables and Pearson's correlation will be calculated between each pair of the independent variables. A correlation coefficient greater than 0.7 (irrespective of the direction + or -) signposts multicollinearity. The pairs of the independent variables having a strong correlation (> 0.7) will be detected and either of the variables in the pair detected with multicollinearity will be used in the model formulation. In addition to that, Variance Inflation Factor (VIF) will also be calculated for the model to detect multicollinearity where VIF between 5 and 10 indicates high correlation and VIF beyond 10 reveals that the regression correlations are poorly estimated.
4. The variance of residuals is equal across all values of independent variables (Homoscedasticity) - scatterplots will be produced between residuals and predicted values to test this assumption where residuals are expected to be randomly distributed and not demonstrate any patterns.
5. There is little or no autocorrelation in the data - this assumption was tested using Durbin-Watson score. The Durbin-Watson test statistics (d) is calculated by the following equation (Montgomery et al., 2012):

Equation 4.5: Durbin-Watson test statistics

$$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}$$

Where, n is the number of observations and e is the residual. d = 2 indicates no autocorrelation as the value of d will always lies between 0 and 4. d is compared to the lower and the upper critical values ($d_{L,\alpha}$ and $d_{U,\alpha}$) at significance α (See Table in Appendix 3 for critical values of the Durbin Watson Score).

To test for positive autocorrelation:

- If d < $d_{L,\alpha,}$ - there is statistical evidence that the error terms are positively autocorrelated.

- If $d > d_{U,\alpha,}$ - there is no statistical evidence that the error terms are positively autocorrelated.
- If $d_{L,\alpha} < d < d_{U,\alpha,}$ - the test is inconclusive.

To test for negative autocorrelation:

- If $(4 - d) < d_{L,\alpha,}$ - there is statistical evidence that the error terms are negatively autocorrelated.
- If $(4 - d) > d_{U,\alpha,}$ - there is no statistical evidence that the error terms are negatively autocorrelated.
- If $d_{L,\alpha} < (4 - d) < d_{U,\alpha,}$ - the test is inconclusive.

The process to be followed in fitting a regression model to the collected data is illustrated in Figure 4.7 which was modified from Field (2013). Accordingly, compliance of the developed model to the regression assumptions should be checked as it is fundamental to consider the model as statistically valid. Different strategies (such as weighted least squares regression, bootstrap confidence intervals and multilevel modelling) can be used and the regression analysis can be rerun when any assumption is not met as indicated in Figure 4.7. Further, a different process is followed in the development of a multilevel model which is not presented in Figure 4.7 (See, Chapter 19 in Field (2013) more information on multilevel models).

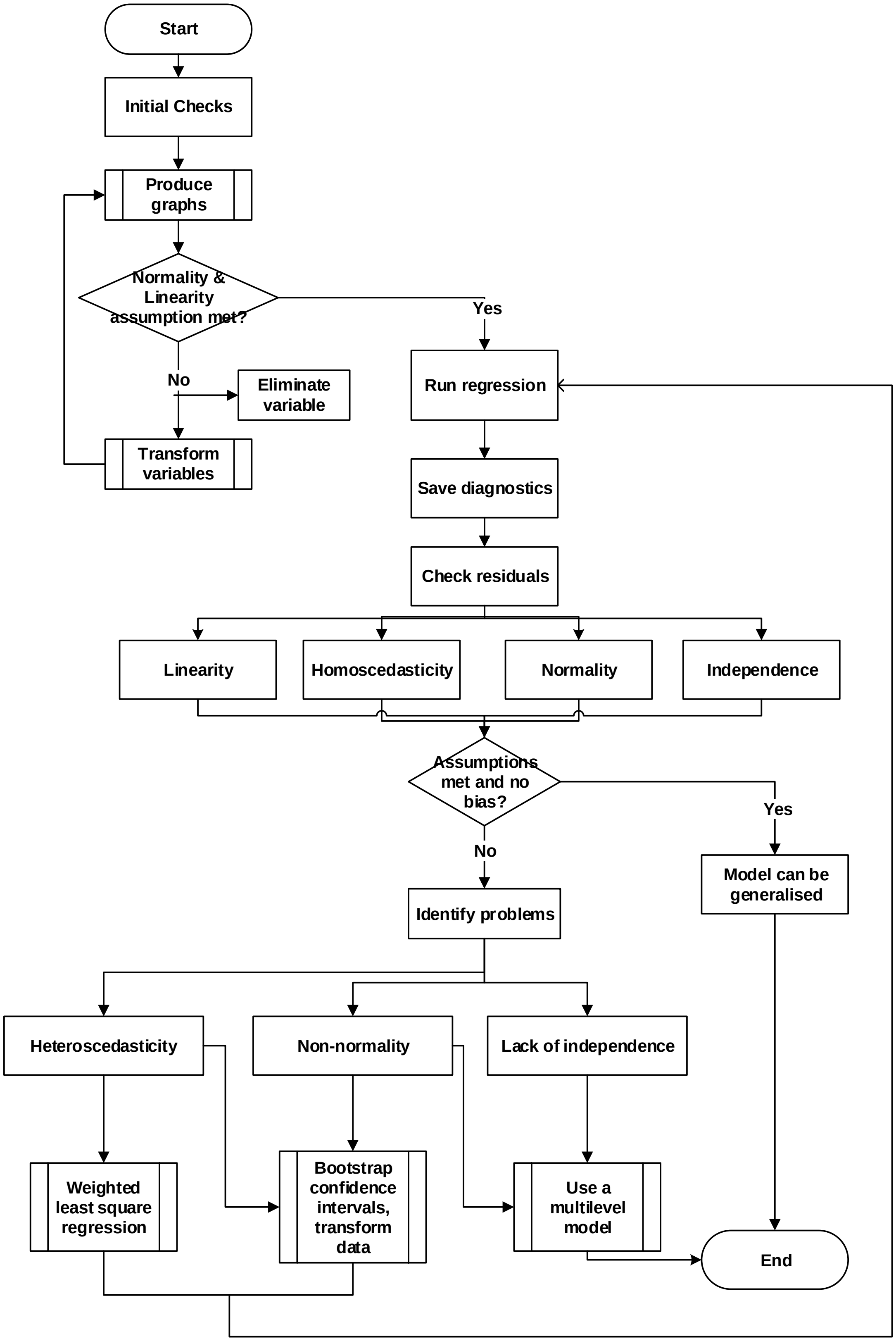


Figure 4.7: The process of fitting a regression model

Modified from: Field (2013)

4.10.Credibility of Research Findings

Reliability and validity are two key aspects of a research design to ensure the credibility of research findings. Saunders et al. (2009) define reliability as the ability of the data collection and analysis techniques to produce consistent results when iterated while validity as being able to produce the result what it intended to produce. The intended key research findings include:

1. Carbon and cost hotspots of office buildings
2. Relationship between EC/CC and design variables
3. Relationship between EC and CC
4. The EC prediction model

All of which are objective for a given sample and will yield the same results when repeated using the same data collection and analysis techniques ensuring the reliability of research findings. However, the validity of the models was confirmed by assessing two metrics including coefficient of determination of the regression model (R^2) and coefficient of variation of the regression model (CV).

The coefficient of determination (R^2) also referred to as the 'model fit' presents the percentage change in the dependent variable explained by the model (independent variables). R^2 is calculated as follows:

Equation 4.6: Formula to calculate coefficient of determination

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

Where, y_i denotes the observed value of the depended variable, $\hat{y}_i$ denotes the predicted value of the depended variable and $\bar{y}_i$ denotes mean of the observed values. There is another estimate of R^2 called adjusted R^2, which attempts to estimate the R^2 of the population instead of the sample itself. It is calculated as follows:

Equation 4.7: Formula to calculate the adjusted coefficient of determination

$$R^2_{adj} = 1 - (1 - R^2)\,\frac{n-1}{n-p-1}$$

Where n denotes the number of observations and p denotes the number of independent variables in the model. The rationale for this adjustment is that it is less likely that the correlation between the dependent variable and a newly added independent variable is zero even if it is zero in the population due to sampling error and R^2 always increases when a new independent variable is added. Hence, R^2 is adjusted to compensate for this possible error (Miles and Shevlin, 2001). Higher R^2 implies a better model fit.

The CV is the metric used to check the accuracy of the models. CV is calculated as a percentage of the standard deviation of the residuals divided by the mean of the observed values of the dependent variable, which is presented in Equation 4.2.

Equation 4.8: Formula to calculate coefficient of variation

$$CV = \frac{\sqrt{\frac{\sum(y_i - \bar{y}_i)^2}{n-1}}}{\bar{y}_i} X\ 100$$

For instance, a CV of 15% implies that the accuracy of the prediction of most of the cases (68%) would fall between ±15%. Hence, a smaller CV is desirable. However, Ashworth and Skitmore (1999) from a thorough analysis of the past studies suggested that CV of 15% to 20% of prediction accuracy is acceptable for early design stage cost estimates while Peurifoy and Oberlender (2002) proposed that an accuracy between +25% to -5% is acceptable for a conceptual estimate.

However, a lower CV implies a better model prediction. Therefore, a prediction accuracy of CV ±20% is considered sufficient to validate the models. Nevertheless, the CV of a model will most likely deteriorate when the model tends to predict cases outside its database (McCaffer, 1999). Hence, the prediction accuracy of the model was assessed using internal (data that will be used to formulate the model) and external data (data that will not be used to formulate the model).

4.11.Research Design

Based on previous discussions, the positioning of the research is presented in the form of the Sauder's research process 'Onion' in Figure 4.8. As discussed in Section 4.2 the research philosophy takes the post-positivist critical realist perspective and an inductive approach (Section 4.3) was used to formulate the EC prediction model by collecting data. Mixed methods of data collection and analysis techniques were employed due to the use of qualitative predictor variables in the EC prediction model. The research falls within the mixed-model research as the findings of the qualitative analysis were used as inputs to the model formulation (Section 4.5). In addition, the research design in terms of research objectives is presented in Table 4.8. Table 4.8 summarises the data collection and analysis techniques employed in order to achieve each objective and provide references to the respective chapters.

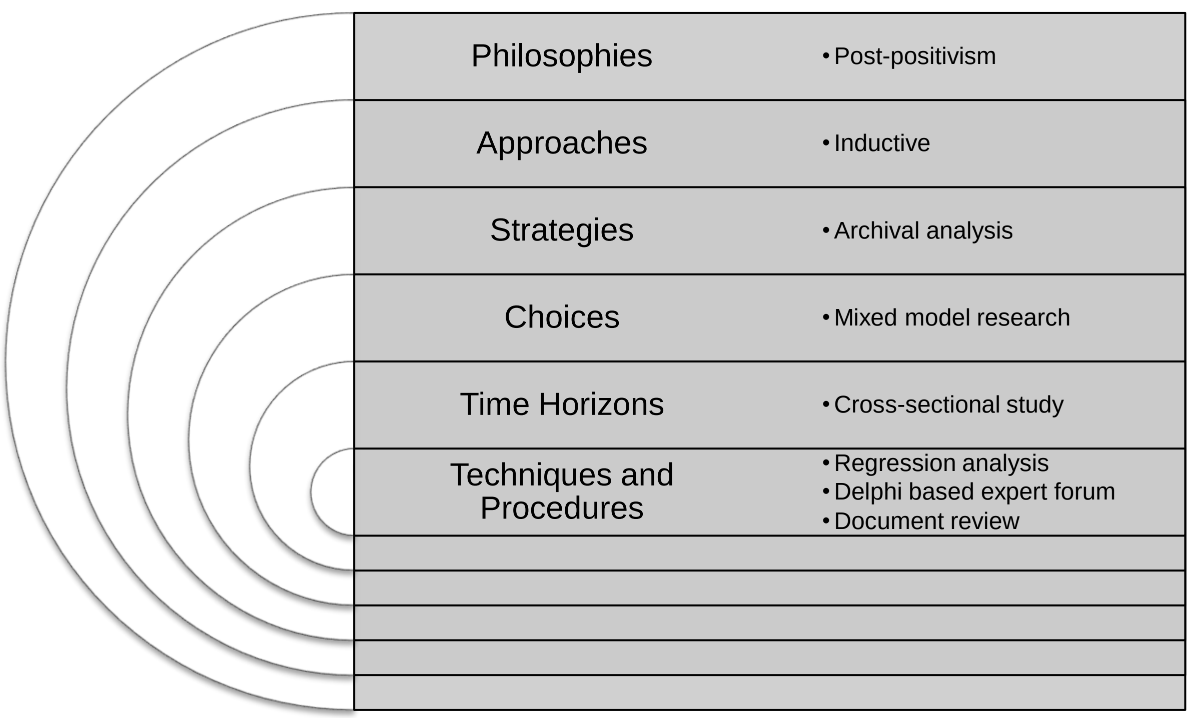


Figure 4.8: The positioning of research based on Sauder's research process 'Onion'

Modified from: Saunders et al. (2009)

Table 4.8: Research design in term of research objective

Objectives	Research Methods		Chapter Reference
	Data Collection	Data Analysis	
1. Review of EC and OC	Literature review	Synthesis and critiquing of the literature	2
2. Review of EC estimating practices	Literature review	Synthesis and critiquing of the literature	3
3. Analysis of carbon hotspots	Literature review, historical project data collection	Pareto Principle	2.6, 4.8.1 and 4.9.1
4. Relationship between EC/CC and design variables	Literature review, historical project data collection	Correlation coefficient	3.10.1, 4.8.1, 4.8.2 and 4.9.2
5. Relationship between EC and CC	Literature review, historical project data collection	Correlation coefficient	2.9, 4.8.1 and 4.9.2
6. Developing EC and CC models	Literature review, historical project data collection, expert forum, document review	Regression analysis, content analysis of the qualitative data	4.7, 4.8.1, 4.8.2 and 4.9.3
7. Validation of EC and CC models	Literature review, historical project data that were not used for the model formulation	Coefficient of Determination, Coefficient of Variation	4.10

4.12. Summary

The research was designed to develop an EC prediction model by capturing relationships between EC and design variables; and EC and CC. The research design begins with the conceptualization of the model from an extensive literature review. The parameters of the conceptual were estimated by collecting primary and secondary data. Both quantitative and qualitative data collection and analysis techniques were employed to develop the model. Historical data were collected from online cost databases and QS organisations and the EC and cost were estimated using published secondary cost and carbon databases. Meanwhile, design data were also captured from historical projects. However, the need to develop numerical indices for the finishes and services quality of buildings was identified to include quality of buildings as predictors in the model. Consequently, a Delphi-based expert forum was selected to develop the finishes quality indicator with two rounds of the verification process in order to develop an objective index. On the other hand, document review was employed for the development of services quality index based on the service provision of buildings due to the lack of detailed measurements and specification of services of the collected data and the lack of EC data of services.

Histograms, boxplots and descriptive statistics were used to examine the data to identify problems in the dataset before performing the regression analysis. Regression analysis was chosen to develop the model due to its strong mathematical basis, transparency and the use of it in similar past research. However, regression outcome will become invalid if the regression assumptions are violated. Hence, various techniques were used to test the regression assumptions. Finally, the validity of the model was measured using the coefficient of determination and coefficient of variation and a CV of ±20% was considered satisfactory as the model is an early design stage prediction model.

5. Data Collection and Processing

5.1. Introduction

This chapter unveils the intricacies involved in the data collection of the study in employing the techniques introduced in the Methodology chapter. Data requirement was defined at the beginning of the chapter and the sources of data were identified. The information available in each of the data source were mapped onto the data requirement of the study to indicate the inadequacies found in each dataset. The study sample was developed by combining the available datasets using different techniques which are presented in the data collection process of this chapter (Section 5.3). Section 5.3 gives a snapshot of the overall data collection and processing and set the scene for a detailed discussion. A pilot study was conducted initially to assess the feasibility of obtaining data from BCIS online cost database, which was identified as the most appropriate data source to obtain a large sample. A detailed description and the results of the pilot study are presented to demonstrate the inadequacy of BCIS online cost database. The findings of the pilot study paved the way to explore additional data sources such as data from QS consultancy practices and other special databases. Data were obtained from QS consultancy practices, which were then used in conjunction with the data obtained from BCIS online cost database to develop the study sample. The process of developing the study data involved some data processing including the calculation of descriptive statistics and the comparison of means of the samples to identify differences between low to medium rise and high-rise buildings. Further, the developed sample for the regression analysis was validated using an independent dataset. In addition, the data collected from the expert forum for the finishes quality index development and the data collected from published documents for services quality index development are also discussed under separate sections.

5.2. Data Requirement

EC, CC and design variables (quantitative like GIFA, building height, wall to floor ratio; and qualitative like finishes quality and services quality indicators) of buildings are the basic building blocks to formulate the models. However, it is unusual to find data sources where all the information is readily available. In most of the cases, at least one of the three can be found to be missing. For instance, historical data available in the online databases contain either cost or EC and design variables (BCIS – holds cost and most of the design variables data; WRAP EC Database contains EC data of most of the building elements and some design variables). Therefore, the best available data were obtained and different techniques were employed to replace the missing data systematically. This is discussed later in the chapter. Figure 5.1 illustrates types and sources of the data obtained for the study. Primary data refers to historical construction project data specifically obtained for the study and the secondary data refers to historical

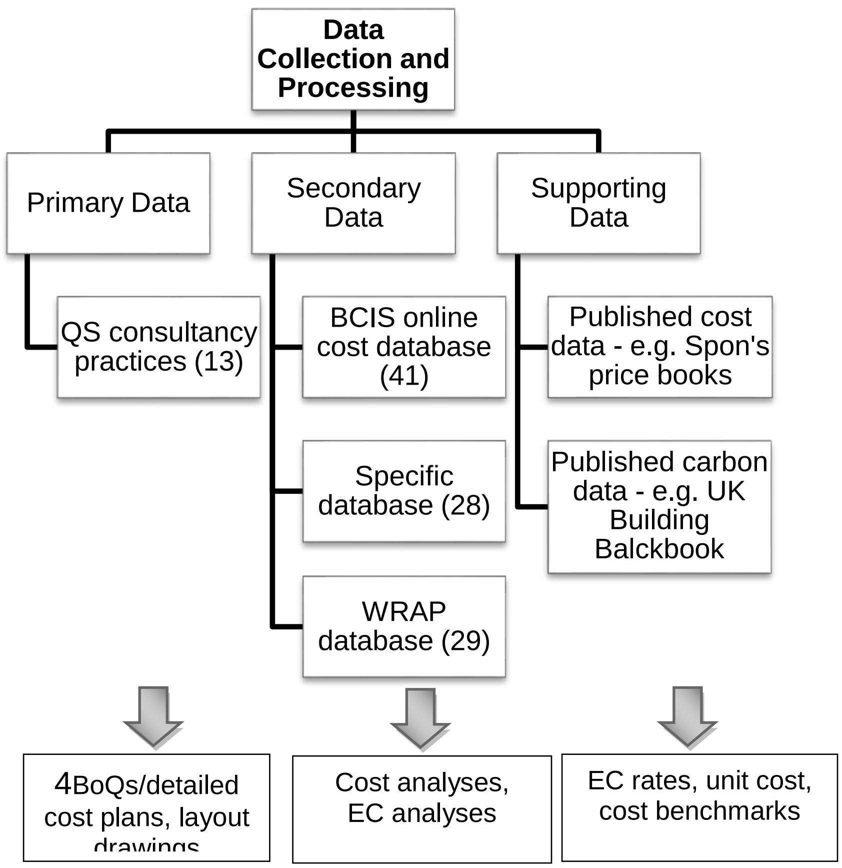


Figure 5.1: Types and sources of data obtained

construction project data that are made available to researchers and community for further investigation of a problem (See, Hox and Boeije, 2005). Hence, primary data include unpriced BoQ or detailed cost plans and layout drawings obtained from QS consultancy practices; secondary data include cost analysis obtained from BCIS online database and EC analysis obtained from a specific database from a QS consultancy practice. In addition, another category of data was identified as 'supporting data' for the study which include published cost and EC databases consist of cost and EC benchmarks. These benchmarks are essentials to estimate the cost and EC of primary and secondary data.

5.3. The Data Collection Process

Data collection of the study involved collection of both quantitative (historical project data) and qualitative data (finishes and services indices development). Hence, an overview of the data collection process is provided in this section before discussing each step in detail.

5.3.1. Historical Project Data Collection

The collection process of historical project data consists of three successive stages (see, Figure 5.2) which are discussed briefly here and a detailed discussion is presented in subsequent sections.

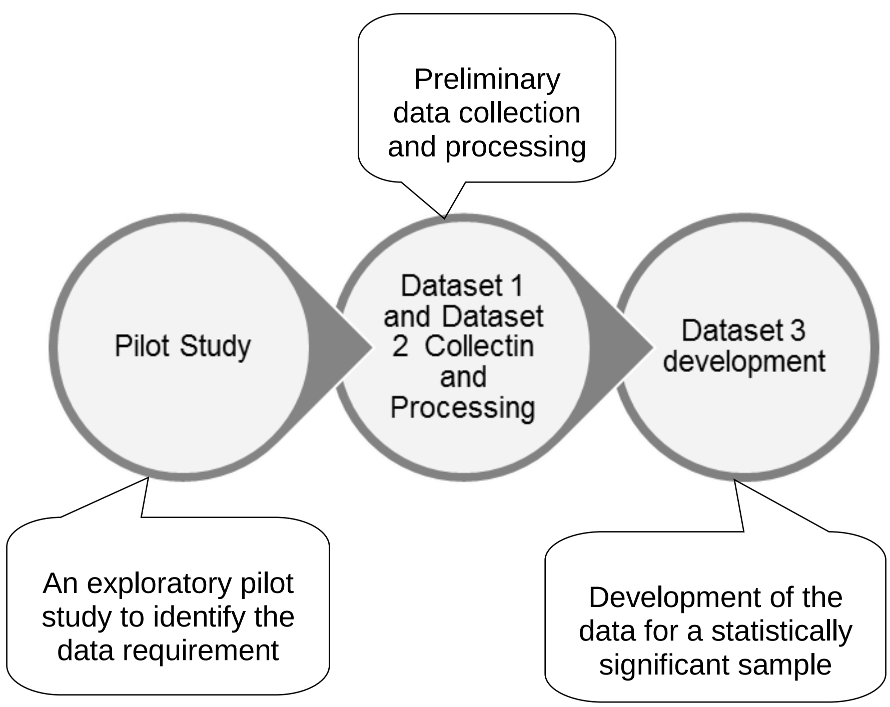


Figure 5.2: Data collection process - overview

Stage 1: Pilot study

An exploratory pilot study was conducted to determine the feasibility of obtaining data from BCIS online cost database to develop EC plans. BCIS was chosen as it is a first-hand data source, which contains historical project data. Subsequently, a pilot case was employed and effort was made to develop EC plans from data that are available in BCIS online database. However, EC of all building elements could not be calculated due to insufficient design data. Hence, it was decided to collected data from alternative data sources, which include data from QS consultancy practices. Detailed discussion of the pilot study is presented in Section 5.4.

Stage 2: Processing of Dataset 1 and Dataset 2 from QS consultancy practices

Two types of data were from QS practices. The first type includes unpriced BoQs/detailed cost plans and drawings (Dataset 1), and the second type includes elemental EC analysis of buildings with limited design data (Dataset 2). Shortcomings identified in each dataset (see Table 5.1) disqualify the datasets to be considered for the statistical analysis. However, these datasets were used to develop the study sample, which is discussed in Stage 3.

Table 5.1: Overview of the data obtained from the QS consultancy practices

DATASET 1		**DATASET 2**	
Source : 7 QS consultancy practices		Source : A QS consultancy practice	
No of projects: 13		No of projects: 28	
Available data	**Unavailable data**	**Available data**	**Unavailable data**
• Measurement of quantities for most of the items • Design variables • Specification	• Measurement of quantities for Fitting, Furnishing & Equipment and Services	• Elemental EC analysis • Some design variables – GIFA and no. of storeys	• Measurement of quantities • Specification • Other design variables

Stage 3: Development of Dataset 3

Dataset 3 consists of 41 historical project data (cost analyses) obtained from BCIS. Only 41 projects were selected for the statistical analysis due to their conformity with the required data. However, data gaps were identified in Dataset 3, which were eventually filled by the information derived from Dataset 1 and Dataset 2. A major shortcoming of Dataset 3 is inadequacies in element specifications. On the other hand, there is no industry developed elemental EC benchmarks to produce EC plans for cost analyses obtained from BCIS. Therefore, EC elemental rates were developed from the datasets obtained in Stage 2 to feed into the data obtained from BCIS whenever detailed specification was lacking. Particularly, elemental EC rates were used for building elements such as Substructure, Frame, Upper Floors (only for the in-situ concrete floor), Roof, Fittings, Furnishings and Equipment and Services. Figure 5.3 illustrates the inputs received from Dataset 1, Dataset 2 and the published sources to develop Dataset 3. Accordingly, the gaps in Dataset 3 were filled by Dataset 1 and Dataset 2 to develop the EC estimates for projects in Dataset 3.

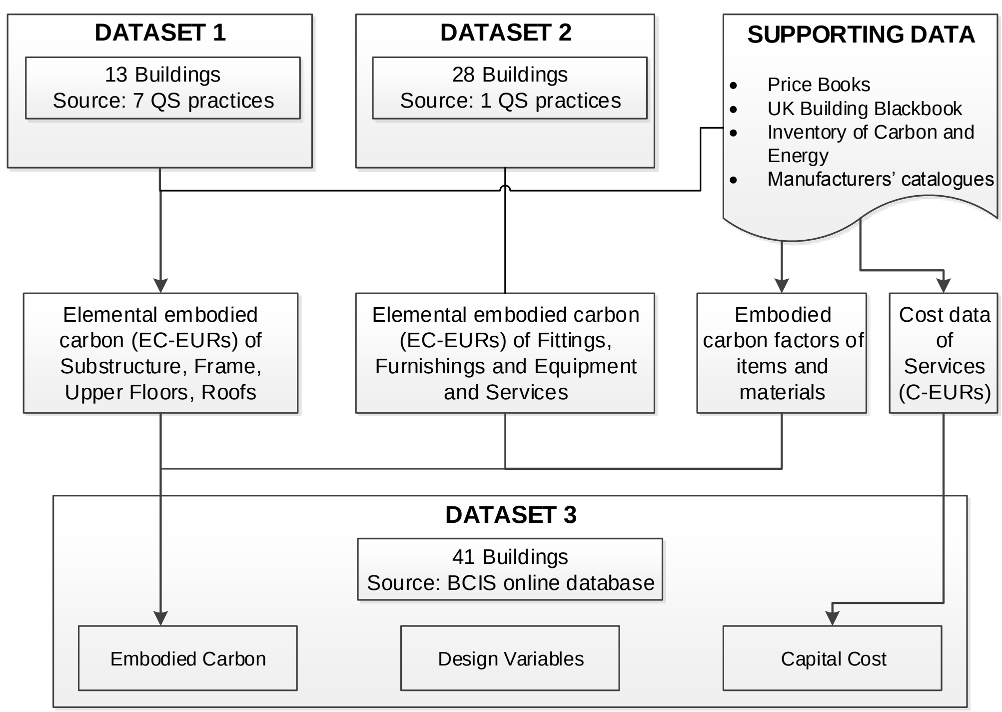


Figure 5.3: Inputs from Dataset 1, Dataset 2 and the published sources to the development of Dataset 3

Further, the need was identified to validate Dataset 3 as it was developed from multiple sources of data. Hence, WRAP EC Database (WRAP and UK-GBC, 2014) was used to check the validity of Dataset 3 (see, Section 3.6 (d) for more details on WRAP database) because the data obtained from WRAP database could not be used for the study due to inadequate design information and the EC data available in the database lacks detailed elemental analysis. The validation of Dataset 3 is presented in sub-section 5.7.3.

Table 5.2: Data available in each of the dataset

Required data	Measurement scale	Dataset 1	Dataset 2	Dataset 3
1. Measurement of quantities/ element unit quantities of the buildings	Ratio	Yes (some elements are not measured)	-	Yes
2. Specification of the buildings	Nominal	Yes	-	Yes
3. Design variables of the buildings – quantitative: GIFA, Building height/No. of storeys, Wall to floor ratio, Circulation space ratio, No. of basements	Ratio	Yes	Yes (Only GIFA & no. of storeys)	Yes
4. Design variables of the buildings – qualitative finishes quality, services quality	Ordinal	Yes	-	Yes
5. EC	Ratio	Yes (Excludes Fittings & Services)	Yes	Yes
6. CC	Ratio	Yes (Excludes Fittings & Services)	-	Yes

In summary, the checklist of the required data and the data available in each of the dataset is presented in Table 5.2. Even though Dataset 1 contains most of the required data except for detailed measurements of Fittings, Furnishings and Equipment and Services, Dataset 1 could not be used for the model formulation, as the sample size is small and not statistically significant. Dataset 2 could not be used because it does not contain most of the required data. Subsequently, Dataset 3 was developed using both Dataset 1 and Dataset 2, eliminating deficiencies in the original data sets to meet the data requirements.

5.3.2. Finishes and Services Index Development

Finishes and services quality of buildings were required to be identified in an objective and a systematic way as finishes and services quality were identified as variables affecting cost and carbon (see, Table 3.3). Hence, finishes and services quality of buildings become predictors in the cost and EC models. Therefore, the finishes quality index was developed from a Delphi-based expert forum while services quality index was developed from a review of published price books, which are discussed in detail in Section 5.8 and 5.9 respectively.

5.4. Pilot Study

A pilot study was conducted to determine the feasibility of developing EC estimates from the data obtained from BCIS online database as briefly mentioned in Section 5.3. BCIS database is designed with multi-faceted search facilities and allows users to refine search criteria to obtain specific data. A pilot case was employed to proceed with the pilot study from a refined sample produced by BCIS, which met the set search criteria, which is discussed as follows.

5.4.1. Selection of the Pilot Case

Selection of the pilot case was a step by step process. Data requirement has to be fed into the system by defining basic parameters of the buildings, building specification and rebase date. Accordingly, 'Building function' was selected as office buildings under the main category '300 Administrative, commercial, protective facilities' as shown in Figure 5.4. 'Age of analyses' was chosen to

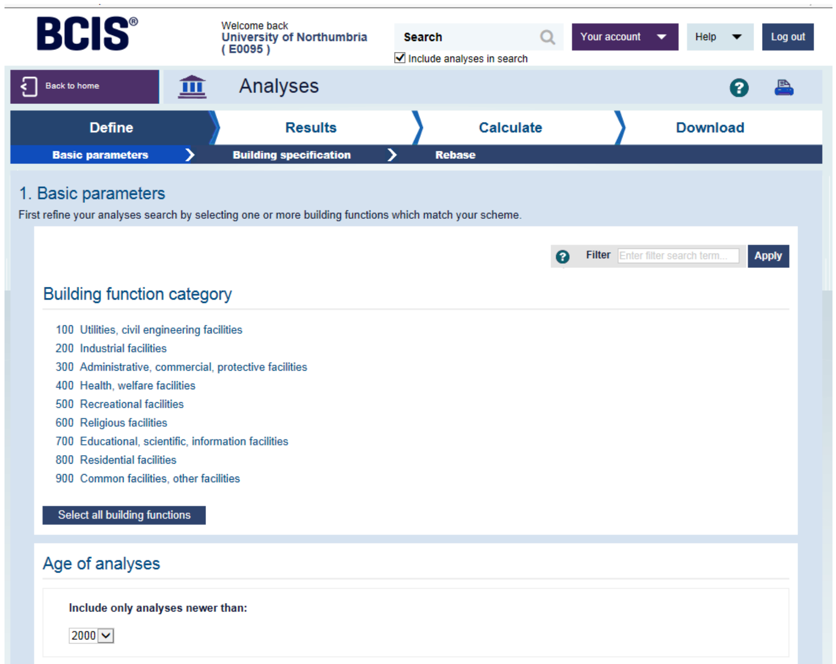


Figure 5.5: Defining building specification

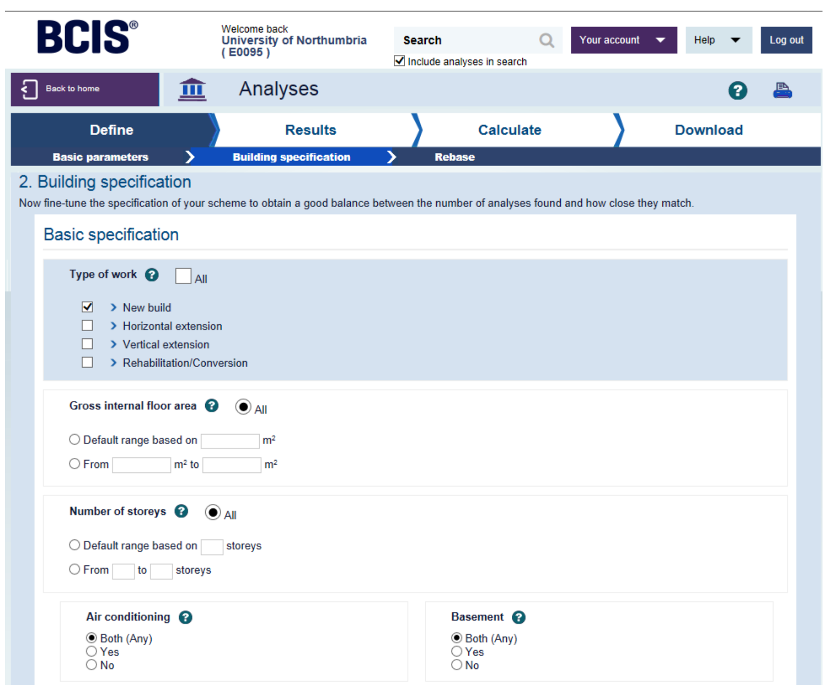


Figure 5.4: Defining basic parameters of the required data

include the analyses from 2006 to obtain most recently completed buildings. ‘Type of Work’ was filtered to include only ‘New build’ (see, Figure 5.5) and no constraints were imposed on GIFA, the number of storeys, air-conditioning and basements. Finally, all the analyses were rebased to 2016 1Q and a location index of 100 (see, Figure 5.6) to be in line with the base of the UK Building Blackbook.

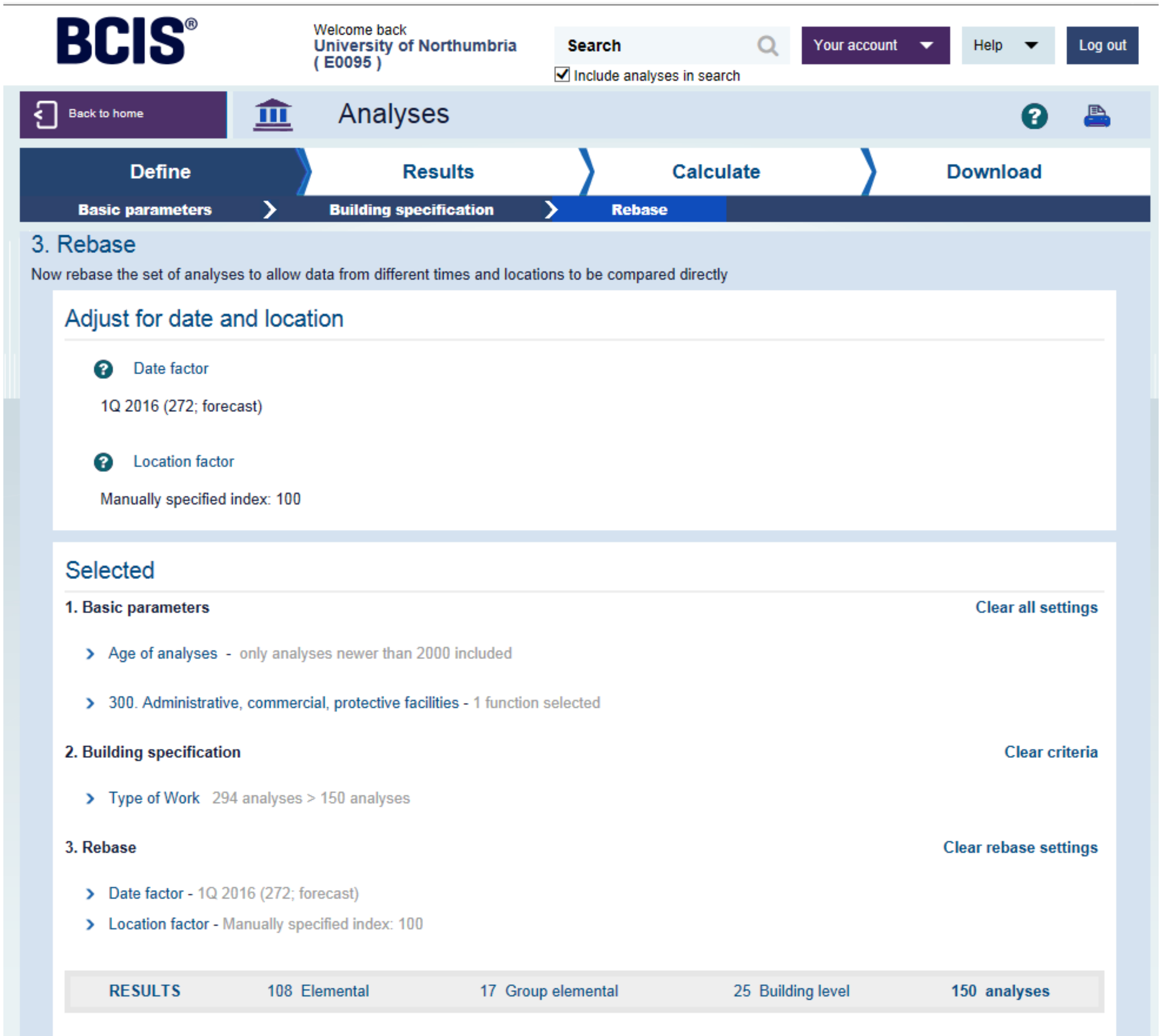


Figure 5.6: Defining base

An additional filter was applied to show only the buildings that contain elemental analysis including Element Unit Quantity (EUQ – is the total quantity of the element expressed in a suitable unit and measurement convention as defined by BCIS) which is a fundamental data to produce EC estimates. However, the first search did not provide any data that meet all the criteria defined and filters applied. Subsequently, 'Age of analyses' was modified to include analyses from the year

2000 where seven projects were found to meet the defined criteria. Consequently, a pilot case building was selected randomly from the seven analyses. The pilot case was a customer service centre located in Bridgend, Mid Glamorgan and constructed in 2002. It was a two-storied steel framed building with a GIFA of 3,080m^2 (See Appendix 1 for more information on the selected pilot case).

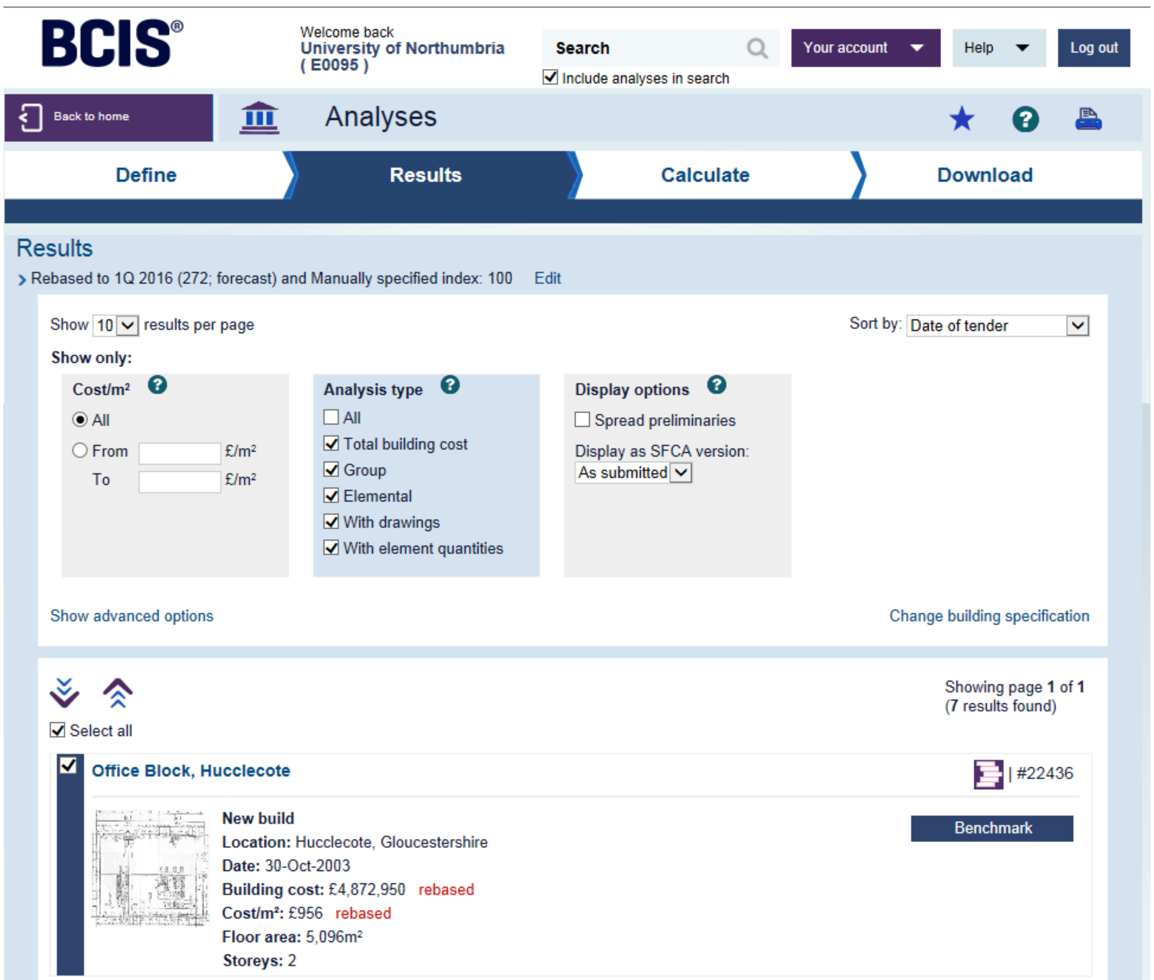


Figure 5.7: Applying additional filters

5.4.2. Estimating Embodied Carbon of the Pilot Case

The steps involved in estimating the EC of a building element for data obtained from BCIS is presented in Figure 5.8. Accordingly, EUQ was extracted from cost analysis of the building, as EUQs were readily available. However, EC-EUR is not readily available. Therefore, the corresponding EC Element Unit Rate (EC-EURs) was developed for the building element under concern using published EC factors in ICE (Hammond and Jones, 2011), Blackbook (Franklin & Andrews, 2011) and manufacturers' data. EC-EUR is the quantity of carbon embodied in one unit of an element. However, the unit of measurement of the elements vary. For instance, Substructure EUQ is the internal area of the lowest floor while External Wall EUQ is the façade area without the area of windows and external doors. Finally, EC of the element was calculated by multiplying EUQ by EC-EUR. In this way EC of all building elements in the building were calculated and the total EC of the building was derived by adding the EC of elements together.

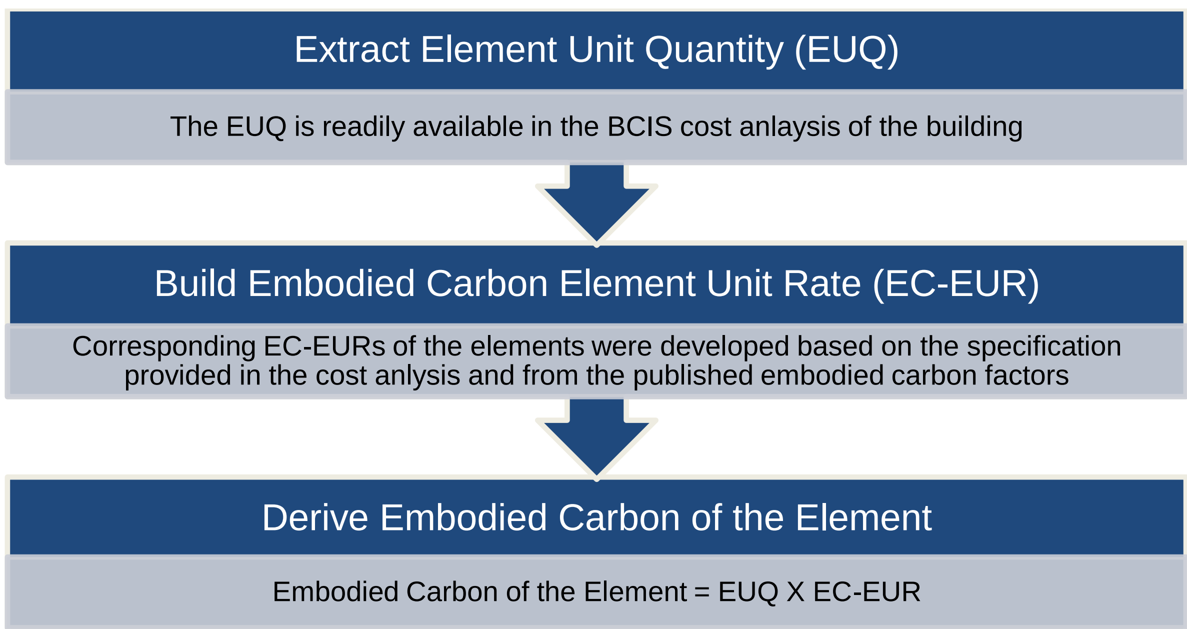


Figure 5.8: Steps in estimating EC of elements for BCIS data

The method used in calculating EC-EUR and EC of an element is presented in Table 5.3 by using Wall Finishes of the selected customer service office building as an example. The element was sub-divided into BoQ items and the quantities of each item were extracted from the cost analysis. The EC factors were obtained from the UK Building Blackbook and item quantities were multiplied by the respective EC factors and the resultants were added together to derive the EC of the element (See Appendix 2 for the EC calculations of the rest of the elements). Even though the EC estimating process appears to be explicit, challenges were faced in extracting quantities of items and EC factors, which, are listed in Table 5.4, and the approaches used to overcome the challenges are presented with examples.

Table 5.3: Calculating EC of wall finishes for the selected building

Blackbook Item Nr.	Wall Finishes	Unit	Qty	EC per Unit	EC
M201202A	Cement, sand (1:3) screed finish, 13mm thick, over 300m wide	m^2	3057	4.950	15,131.386
M601001B	Emulsion paint, one mist coat and two full coats, plastered background, 3.5-5.0 m	m^2	3057	1.068	3,264.876
	Total of Wall Finishes	**m^2**	**3057**	**6.018**	**18,396.262**

Table 5.4: Approaches used to overcome the challenges in EC estimating

Challenges	Approaches used to overcome challenges	Examples
Missing details/specification	Obtain possible details from drawings	Size and the number of Brise Soleil were read from the drawings.
	Making assumptions	Number of coats of painting was assumed.
No matching EC factors from published sources	Pro-rata	EC factor for 15mm render was not present so pro-rata applied to 12mm render rate.
	Use average of the EC factors	Average of the carbon factors for varying sizes of glazed screen was used.
	Find close match	Carbon factor of 3.5 N/mm^2 Block walls was used in the place of for 4 N/mm^2 block walls

EC of the selected building was calculated in this way using the available design information, published EC factors and different approaches as presented in Table 5.4. The same approach was followed in estimating EC of Dataset 3, which is discussed in Section 5.7.2.

5.4.3. Results of the Embodied Carbon Estimating of the Pilot Case

Table 5.5 presents the elemental EC profile of the building. However, the EC of all the elements could not be calculated due to insufficient specification and lack of detailed measurements as mentioned already. EC calculation becomes challenging especially when there is a difference in the unit of measurement between the main element (such as Substructure) and the components within the element (concrete, formwork, reinforcement etc.) where no detailed specification is available. For instance, elements such as Substructure and Frame which are measured in m^2 while the components constituting the elements are measured in m^3 (concrete), m^2 (formwork, surface treatments), t (reinforcement), nr (piles) and the like. Therefore, EC of Substructure, Frame, Roof, Fittings, Furnishing and Equipment and Services could not be calculated. However, similar insufficiencies in Dataset 3 were addressed by developing elemental benchmarks from Dataset 1 and Dataset 2, which is discussed in Section 5.7.2.

Table 5.5: EC of the selected elements of the customer service centre

Building Elements	Unit EC ($kgCO_2$/Unit)	Total EC ($kgCO_2$)	Comments
2B Upper Floors	86.387	140,983.58	Upper Floor rate is an average of two EC factors
2D Stairs	1,680.950	6,723.80	Stairs rate was built based on the assumptions on the sizes and the specification.
2E External Walls	26.011	59,565.19	
2F Windows and External Doors	26.971	22,655.60	
2G Internal Walls & Partitions	16.541	36,853.35	
2H Internal Doors	15.887	1,270.96	The EC rate is an average.
3A Wall Finishes	6.018	18,396.26	The number of coats of paint was assumed.
3B Floor Finishes	60.684	54,750.82	
3C Ceiling Finishes	25.027	67,948.31	

5.4.4. Outcome of the Pilot Study

Details of measurement, specification and EC factors are the key inputs for EC estimating. However, data available in BCIS data suffers due to insufficiently detailed cost analyses. Even though there are EC factors for materials (ICE) and BoQ items (Blackbook), there is no published source of data for EC-EURs (However, industry published EURs for the cost (C-EURs) are available to assist early design stage cost estimating in the form of price books). EC calculation becomes challenging when there is a difference in the unit of measurement between the main element and its components when there is no detailed design data. Therefore, insufficiently detailed cost analyses in BCIS database and lack of industry benchmarks for EC-EURs disqualifies the BCIS online database to be used as a standalone data source to develop EC estimates.

5.5. Dataset 1

Dataset 1 comprises of thirteen (13) historical project data obtained from seven (7) QS consultancy practices. The collected data include blank BoQs or detailed cost plans of office buildings and layout drawings for some of the buildings.

5.5.1. Data Description

The majority of buildings were steel framed; one building was a hybrid; and the rest were concrete framed. GIFA ranges from 2,374 m^2 to 63,246 m^2 and number of storeys ranges from three (3) to eighteen (18). Fittings, Furnishings and Equipment and Services were not measured in detail and identified, as ‘Item’ or ‘Lump Sum’ was the major problem with Dataset 1. Further, measurement of the quantities of Finishes in two of buildings was found to be ambiguous (D1001 & D1005). Data description of the thirteen (13) buildings is presented in Table 5.6.

Table 5.6: Design data of Dataset 1

Building ID	Frame Type	GIFA (m^2)	No. of Storeys
D1001	Steel	33,663	18
D1002	Steel	11,320	8
D1003	concrete	2,859	3
D1004	Steel	15,120	7
D1005	Hybrid	63,246	16
D1006	Steel	1,949	4
D1007	Steel	22,288	10
D1008	Steel	3,289	4
D1009	Concrete	3,262	3
D1010	Concrete	4,959	3
D1011	Steel	21,300	13
D1012	Concrete	21,300	12
D1013	Steel	2,374	4

5.5.2. Estimating Embodied Carbon and Capital Cost

CC and EC estimates were prepared from scratch using unpriced BoQs or detailed cost plans where unit rates and EC factors were obtained from the UK Building BlackBook (Franklin & Andrews, 2011) (see Table 5.7). Unit rates are the cost including mark-up per unit of a BoQ item, given in £ per unit and EC, factors are the EC per unit of a BoQ item, given in $kgCO_2$ per unit. The data presented in the Blackbook have a base date of 2010 2Q (price index - 218) and a location index of 100. Therefore, the cost estimates of Dataset 1 were updated to 2016 1Q (price index – 276 (forecast)) and the location index was maintained as 100 to normalise the base of all the estimates. Further, the unit rates include 10% of mark-up (head office overhead and profit).

On the other hand, EC in buildings are influenced by the manufacturing process of building materials, transport, the method of construction, recycling potential and the like (Chen et al., 2001b, Ramesh et al., 2010). However, the manufacturing process of a material is the key determinant of its EC when cradle-to-gate system boundary (EC emissions associated with the fuel consumption from raw material extraction up to the manufacturing factory gate) is considered (as in the Blackbook data). Therefore, the study assumes the same method of manufacturing for all building materials used in the Blackbook by default.

Table 5.7: Extract of an estimate of the Upper Floors of a building

Items	Qty (a)	Unit	Unit Rate (b)	EC factor (c)	CC (a x b)	EC (a x c)
Upper Floors						
RC concrete in suspended floor slab 300 thick	598	m^3	106.650	336.444	63,776.70	201,193.51
Formwork to soffit of suspended slabs	1,992	m^2	49.630	5.733	98,862.96	11,420.14
RC concrete in upstands to suspended slabs	14	m^3	131.640	353.418	1,842.96	4,947.85
Formwork to edge of suspended slab 250 - 500 high	338	m	18.450	2.783	6,236.10	940.65
Formwork to edge of suspended slab 500 - 750 high	338	m	33.180	4.656	11,214.84	1,573.73
Tamped finish to concrete	1,992	m^2	2.030	0	4,043.76	0
Rebar - bar reinforcement in suspended slabs	73.39	t	1013.940	1722.160	74,413.06	126,389.32
Total of Upper Floors					**260,390.38**	**346,465.20**

From the collected BoQs/cost plans

From the UK Building Blackbook

Based on the above mentioned assumptions, EC and cost estimates were prepared for all the buildings. However, the element classification was inconsistent among the BoQs in Dataset 1. Eventually, each building analysis was arranged in accordance with NRM compliant BCIS cost analysis standard. NRM provides a standard set of measurement rules and essential guidance for the cost management of construction projects and maintenance works. It also superseded the old measurement standard Standard Method of Measurements 7 (SMM7). NRM suite contains three parts – NRM1: Provides guidance on preparing cost estimates and cost plans; NRM 2: Facilitates preparation of BoQs and quantified schedules of works; NRM 3: Provides guidance on the quantification and description of maintenance works to prepare an initial order of cost estimates. Each BoQ item was mapped on to NRM1 element classification system (see, Table 5.8) which is the latest measurement standard prevailing in the UK. Subsequently, NRM compliant elemental cost and carbon plans for each building were produced. Even though the carbon plans are not catered for in NRM, it is anticipated that

NRM will cater for carbon plans in the future due to the increasing emphasis on dual currency appraisal of construction projects (See, Ashworth and Perera, 2015). Therefore, the same practices proposed for cost planning and estimating in NRM were adopted in carbon planning and estimating to allow like-for-like comparisons.

Table 5.8: Mapping BoQ items on to NRM compliant element classification

Item Description	Quantity	Unit	NRM Main Element Group	NRM Sub Element Group
Masonry paint; to blockwork walls	500	m^2	3	3.1
Dry lining and paint	55	m^2	3	3.1
Entrance matting	17	m^2	3	3.2
Carpet tiles	100	m^2	3	3.2
Painted soffit	74	m^2	3	3.3
Suspended plasterboard ceiling 1 x 12.5mm	37	m^2	3	3.3

The estimates exclude Preliminaries (the site overheads and the costs that are not directly associated with the building but to the project) and External Works (works outside the building including site works, road works within the site, landscaping, fencing, external fixtures, drainage and services.). The reason for this is that Preliminaries and External Works do not form part of the main building structure and are influenced by clients and contractors. Further, exclusion of Preliminaries and External Works will allow capturing only the impacts of building design variables. The major problem with Dataset 1 was that Fittings, Fixtures and Equipment and Services were often identified as ‘Item’ or ‘Lump Sum’, hence, not measured. As a result, the produced EC and cost estimates of Dataset 1 exclude Fittings, Fixtures and Equipment and Services. The summary of the EC and CC of the 13 buildings is presented in Table 5.9. The cost and carbon estimates exclude Preliminaries, External Works, Fittings, Fixtures and Equipment and Services. The estimates have a base date of 2016 1Q and a location index of 100.

Among the 13 buildings, EC and cost estimates of one building (D1006) varied significantly from the other buildings. This is highlighted in Table 5.9. The elemental EC values were also identified as anomalies in that particular building.

Therefore, it was concluded that the use of the estimates of this building will affect the credibility of the study findings. Hence, it was decided to eliminate building D1006 from further analysis. Similarly, Frame EC-EUR of D1009 was also identified as an anomaly as the building had similar features to D1010 while has a smaller EC-EUR for Frame. Hence, it was not used to estimate Frame EC in Dataset 3 (see, Table 5.10).

Table 5.9: Summary of the CC and the EC of the buildings in Dataset 1

Building Code	GIFA (m^2)	CC (£)	EC ($kgCO_2$)	CC per GIFA (£/m^2)	EC per GIFA ($kgCO_2/m^2$)
D1001	33,663	38,354,147	27,007,531	1,139	802
D1002	11,320	8,123,613	6,798,939	718	601
D1003	2,859	2,639,247	1,692,852	923	592
D1004	15,120	12,847,259	8,825,578	850	584
D1005	63,246	77,333,327	46,977,344	1,223	743
D1006	1,949	1,104,199	509,197	567	261
D1007	22,288	18,353,008	13,283,008	823	596
D1008	3,289	3,120,328	1,787,778	949	544
D1009	3,262	2,602,018	1,577,015	798	483
D1010	4,959	5,636,879	2,944,681	1,137	594
D1011	21,300	20,156,774	13,251,877	946	622
D1012	21,300	16,570,313	9,944,685	778	467
D1013	2,374	1,659,586	1,101,503	699	464

Table 5.10: EC-EURs of the selected elements

Building Code	Substructure EC-EUR	Frame EC-EUR	Upper Floors EC-EUR	Roof EC-EUR
	$kgCO_2/m^2$ EUQ			
D1001	1027.5	167.0	199.6	368.6
D1002	1439.4	203.9	111.4	148.0
D1003	731.3	126.3	131.9	175.6
D1004	2462.9	143.9	71.1	171.9
D1005	1318.9	345.2	57.6	199.5
D1006	117.9	23.3	24.6	68.6
D1007	1602.7	193.3	116.1	155.9
D1008	259.2	84.1	211.1	212.1
D1009	453.0	19.3	154.4	237.8
D1010	377.2	59.2	191.8	225.8
D1011	1002.5	236.2	166.5	220.0
D1012	1167.5	96.3	164.9	257.1
D1013	366.1	108.8	118.6	81.9

5.5.3. Comparison of the Means of Low to Medium Rise and High Rise Buildings

Elemental EC rates (EC-EURs) of Substructure, Frame, Upper Floors and Roof are used from Dataset 1 to deduce EC of these elements in Dataset 3. However, Dataset 3 consists of low to medium rise buildings while Dataset 1 consists of low, medium and high-rise building. Due to the shortage of data, it was decided to utilise the elemental EC rates from all of the selected data from Dataset 1. However, to ensure homogeneity of the data, mean values of low to medium rise buildings and high-rise buildings were compared to identify any significant difference in the EC values. Dataset 1 was categorised into two groups such as low to medium rise buildings (sample size - 7) and high rise buildings (sample size - 5) and two independent samples were produced. A t-Test for two independent samples was run to compare the elemental EC rates of Substructure, Frame, Upper Floors and Roof of low to medium and high rise buildings. The test hypothesis is as follows:

H_0: $\mu_{EC\text{- low to medium}} = \mu_{EC\text{-high}}$ (Null hypothesis)

H_1: $\mu_{EC\text{- low to medium}} \neq \mu_{EC\text{- high}}$ (Alternative hypothesis)

where, $\mu_{EC\text{ - low to medium}}$ is the mean EC of low to medium rise buildings (of Substructure, Frame, Upper Floors and Roof), $\mu_{EC\text{- high}}$ is the mean EC of high rise buildings (of Substructure, Frame, Upper Floors and Roof).

The null hypothesis is tested for falsification at $\alpha = 0.05$. If the significance value (p-value) is,

1. $\leq \alpha$ then, H_0 will be rejected, which implies that there is sufficient evidence to conclude that the means of the two samples are significantly different OR the difference between the two samples is statistically significant.
2. $> \alpha$ then, H_0 cannot be rejected, which implies that there is no sufficient evidence to conclude that the means of the two samples are significantly different OR the difference between the two samples is not statistically significant.

However, the test statistics of Levene's Test for equality of variances need to be examined to investigate the right t-Test statistics. Accordingly, if the variances of the two independent samples are found to be significantly different (Sig. > α) then, the respective t-Test statistics need to be checked.

The descriptive statistics of the two samples and the t-Test statistics are presented in Table 5.11 and Table 5.12 respectively. According to the descriptive statistics, except for Substructure EC values rest of the elements have mean values that are not significantly different for the two groups. Also, Substructure EC has a high standard deviation in the low to medium rise buildings group compared to the high rise buildings group. This difference in Substructure is assumed to be attributable to not only the building design features but also the ground conditions. Nevertheless, the t-Test provides more certain conclusions about the mean values.

The t-Test results suggest with 95% confidence that the variances of the population of the two groups are equal. However, it could not be concluded whether there is a significant difference in the mean values of EC of Substructure, Frame, Upper Floors and Roof of low to medium rise and high rise buildings due to insufficient evidence (sig. > 0.05, hence, fail to reject H_0). Hence, the EC values of both low to medium rise and high rise buildings were used to estimate the EC of Substructure, Frame, Upper Floors and Roof in Dataset 3. Nevertheless, the test results are considered valid based on the assumption that the data follows a normal distribution.

Table 5.11: Descriptive statistics of the two samples from Dataset 1

	Group	Sample size	Mean	Std. Deviation	Std. Error Mean
Substructure	Low to medium	7	869.87	808.76	305.68
	High	5	1223.82	246.62	110.29
Frame	Low to medium	6	121.03	50.51	20.62
	High	5	207.60	92.18	41.23
Upper Floors	Low to medium	7	141.47	48.34	18.27
	High	5	140.94	55.30	24.73
Roof	Low to medium	7	179.01	53.53	20.23
	High	5	240.22	80.53	36.02

Table 5.12: t-Test statistics of the two samples from Dataset 1

Elements		Levene's Test for Equality of Variances		t-test for Equality of Means					Test Outcome
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
Substructure	Equal variances assumed	4.066	.071	-.936	10	.371	-353.9486	378.0160	Fail to reject H_0
	Equal variances not assumed			-1.089	7.474	.310	-353.9486	324.9695	
Frame	Equal variances assumed	1.298	.284	-1.984	9	.079	-86.5667	43.6406	Fail to reject H_0
	Equal variances not assumed			-1.878	5.953	.110	-86.5667	46.0952	
Upper Floors	Equal variances assumed	.121	.735	.018	10	.986	.5314	30.0002	Fail to reject H_0
	Equal variances not assumed			.017	7.974	.987	.5314	30.7464	
Roof	Equal variances assumed	.653	.438	-1.591	10	.143	-61.2057	38.4579	Fail to reject H_0
	Equal variances not assumed			-1.482	6.492	.185	-61.2057	41.3105	

5.6. Dataset 2

Dataset 2 comprised EC estimates of 28 office buildings obtained from a special database of a QS consultancy practice. The elemental analyses were prepared to an NRM compliant standard.

5.6.1. Data Description

The sample contains three (3) hybrid framed buildings, one (1) concrete framed building and twenty-four (24) steel framed buildings. The GIFA of buildings ranges from 1,788 m^2 to 130,930 m^2. The number of storeys ranges from 1 to 36. The major shortcoming of Dataset 2 is that it lacks design and specification data and cost data of the buildings.

Table 5.13: Design data of Dataset 2

Building ID	Frame Type	GIFA (m^2)	No. of Storeys
D2001	Steel	95,945	36
D2002	Steel	54,101	19
D2003	Steel	65,414	18
D2004	Steel	29,806	18
D2005	Steel	86,211	17
D2006	Steel	126,872	15
D2007	Steel	130,930	14
D2008	Steel	54,550	14
D2009	Steel	48,509	13
D2010	Steel	31,833	12
D2011	Steel	35,760	12
D2012	Steel	13,209	12
D2013	Steel	23,156	12
D2014	Steel	19,764	9
D2015	Steel	19,600	9
D2016	Steel	27,940	8
D2017	Steel	66,093	8
D2018	Steel	9,587	7
D2019	Steel	19,125	7
D2020	Steel	11,170	5
D2021	Steel	7,472	5
D2022	Steel	11,117	5
D2023	Steel	9,645	3
D2024	Steel	1,788	1
D2025	Hybrid	9,372	5
D2026	Hybrid	12,470	4
D2027	Hybrid	2,776	3
D2028	Concrete	15,192	9

5.6.2. Obtaining Embodied Carbon Element Unit Rates (EC-EURs)

Dataset 2 contains EC estimates of 28 office buildings in the form of NRM compliant elemental analysis structure. Hence, EC per GIFA of each element was calculated and the descriptive statistics of the elemental EC per GIFA of the building elements is presented in Table 5.14. However, EC-EURs could not be calculated due to the unavailability of EUQs of the elements. Nevertheless, when EUQ of an element is supposed to be GIFA of a building (in accordance with BCIS definition) then EC-EUR and EC per GIFA rates will be the same. Hence, EC-EURs of Frame, Fittings, Furnishings and Equipment and Services were obtained from Dataset 2, as the EUQs of these elements are equal to GIFA.

Table 5.14: Descriptive statistics of elemental EC (per GIFA) of Dataset 2

Element	Average of the EC per GIFA ($kgCO_2$ per m^2)	Minimum	Maximum	Standard Deviation
1A Substructures	137.20	33.21	320.72	65.31
2A Frame	236.72	98.00	486.41	101.13
2B Upper floors	75.99	1.72	191.08	38.68
2C Roof	25.05	2.88	103.25	19.69
2D Stairs	7.00	2.47	21.46	5.01
2E External walls	111.24	8.37	265.80	63.35
2F Windows and external doors	15.20	0.02	157.64	35.20
2G Internal walls and partitions	20.14	1.19	64.37	15.97
2H Internal doors	1.50	0.12	7.32	1.79
3A Wall finishes	3.65	0.22	18.47	4.23
3B Floor finishes	37.69	0.39	97.77	28.82
3C Ceiling finishes	8.55	0.65	24.62	6.05
4A Fittings and furnishings	0.86	0.02	3.39	1.15
5 Services	106.81	6.63	192.88	50.16

5.6.3. Comparison of the Means of Low to Medium Rise and High Rise Buildings

Similar to Dataset 1, Dataset 2 also contains low, medium and high-rise buildings. Hence, it is required to compare the mean values of EC of Fittings, Furnishings and Equipment and Services of low to medium rise and high rise buildings. Hence, similar hypothesis established in Section 5.5.3 was tested in this case, the only difference is that the test was performed for EC values of Fittings, Furnishings and Equipment and Services, as follows:

H_0: $\mu_{EC\text{- low to medium}} = \mu_{EC\text{-high}}$ (Null hypothesis)

H_1: $\mu_{EC\text{- low to medium}} \neq \mu_{EC\text{- high}}$ (Alternative hypothesis)

where, $\mu_{EC\text{ - low to medium}}$ is the mean EC of low to medium rise buildings (of Fittings, Furnishings and Equipment and Services), $\mu_{EC\text{- high}}$ is the mean EC of high rise buildings (of Fittings, Furnishings and Equipment and Services).

Descriptive statistics and the t-Test statistics are presented in Table 5.15 and Table 5.16 respectively.

Table 5.15: Descriptive statistics of the two samples from Dataset 2

	Group	Sample size	Mean	Std. Deviation	Std. Error Mean
Fittings, Furnishings and Equipment	Low to medium	10	1.15	1.18	.37
	High	9	.52	1.09	.36
Services	Low to medium	13	110.86	50.02	13.87
	High	15	103.31	51.76	13.36

Table 5.16: t-Test statistics of the two samples from Dataset 2

Elements		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Fittings, Furnishings and Equipment	Equal variances assumed	1.502	.237	1.205	17	.245	.62956	.52240
	Equal variances not assumed			1.211	16.984	.243	.62956	.52004
Services	Equal variances assumed	.332	.569	.391	26	.699	7.54938	19.31226
	Equal variances not assumed			.392	25.663	.698	7.54938	19.26326

The test statistics suggests with 95% confidence that the variances of the population of the two groups are equal and there is insufficient evidence to conclude that the mean of the EC of Fittings, Furnishings and Equipment and Services are significantly different for low to medium rise and high rise buildings. Hence, the EC data of both low to medium rise and high rise buildings were used to develop the elemental EC rates of Furnishings and Equipment and Services for Dataset 3. However, it should also be noted that the test outcomes are valid based on the normality assumption.

5.7. Dataset 3

Dataset 3 consists of 41 historical project data obtained from BCIS online database which contains projects since 1987. Adjusting cost analyses to the same base (date and location) is an important step before obtaining data from BCIS online database as the database contains data of buildings located in different locations within the UK and constructed at different times. Therefore, cost analyses of Dataset 3 were adjusted to 2016 1Q and a location index of 100 to be in line with Dataset 1.

5.7.1. Data Description

Dataset 3 comprises of one (1) hybrid framed building, eight (8) concrete framed buildings and the rest were steel framed buildings. GIFA ranges from 212 m^2 to 14,652 m^2 while the sample consists of low to medium rise buildings, which vary from one to six storeys (see, Table 5.17). It should be noted that the cost analyses contained in BCIS include mark-up, which varies from project to project, and the mark-up percentage is not explicit. Similarly, the use of different pricing strategies in the estimates such as front loading and back loading is also not knowable. These differences cannot be adjusted without the information on mark-up percentages and pricing strategies. Hence, this is identified as a limitation of the study. The drawback of Dataset 3 is that it does not contain EC data of the buildings which need to be estimated based on the available design and specification details by making sensible assumptions when necessary information are missing.

Table 5.17: Design data of Dataset 3

Building ID	Frame Type	GIFA (m^2)	No. of Storeys
D3001	Hybrid	3,987	3
D3002	Steel	928	2
D3003	Steel	212	4
D3004	Concrete	2,412	3
D3005	Steel	1,028	2
D3006	Steel	9,007	2
D3007	Concrete	1,930	2
D3008	Concrete	9,653	3
D3009	Concrete	1,136	3
D3010	Steel	1,896	3
D3011	Steel	1,534	2
D3012	Steel	1,756	2
D3013	Steel	2,432	2
D3014	Steel	10,400	3
D3015	Steel	2,926	3
D3016	Steel	3,797	5
D3017	Steel	1,323	2
D3018	Steel	2,325	2
D3019	Steel	8,444	6
D3020	Steel	5,900	3
D3021	Steel	2,510	3
D3022	Steel	692	1
D3023	Concrete	1,026	2
D3024	Steel	9,900	4
D3025	Steel	3,592	2
D3026	Steel	1,753	3
D3027	Steel	1,266	2
D3028	Steel	2,556	3
D3029	Steel	1,835	2
D3030	Steel	1,376	2
D3031	Steel	1,685	2
D3032	Concrete	5,687	3
D3033	Steel	6,885	3
D3034	Steel	473	2
D3035	Steel	6,643	3
D3036	Concrete	4,538	3
D3037	Concrete	14,652	3
D3038	Steel	3,080	2
D3039	Steel	3,887	3
D3040	Steel	1,545	4
D3041	Steel	718	3

5.7.2. Estimating Embodied Carbon and Capital Cost

EC estimates of the buildings in Dataset 3 were prepared using the inputs from the UK Building Blackbook, Dataset 1 and Dataset 2. Figure 5.3 presented in Section 5.3 illustrate the data inflow and outflow to and from Dataset 3 towards the composition of the study sample. Accordingly, EC-EURs from Dataset 1 (for Substructure, Frame, Upper Floors, and Roof) and Dataset 2 (for Fittings, Furnishings and Equipment and Services) were used to estimate the EC of certain elements and EC-EURs were developed from published sources for the rest of the elements. The intricacies involved in the development of EC-EURs of each element are explained as follows:

Substructure of the sample buildings had three design options including, raft, pile and pad and strip foundations. Also, substructure was measured in m^2 making it difficult to develop EC-EURs from the UK Building Blackbook. Therefore, EC-EUR for Substructure was obtained from Dataset 1 and multiplied by EUQ to arrive at the Substructure EC. Generally, past projects with similar specification will be chosen and adjustments will be applied for differences in the quality based on the estimator's experience and the availability of information to arrive at the cost of the proposed building (Ashworth and Perera, 2015). Similarly, buildings with the closest match to the Substructure specification of the building considered (in Dataset 3) were filtered from Dataset 1. However, a different approach was followed afterwards to arrive at the EC-EUR of the building as shown in Figure 5.9. The most appropriate EC-EUR from Dataset 1 was chosen based on the C-EUR of the Substructure when more than one similar match was found in Dataset 1, because a close relationship was observed between C-EURs and EC-EURs of Substructure in Dataset 1 (see, Figure 5.10). In fact, the correlation coefficient between Substructure cost and EC was extremely strong and significant with a coefficient of 0.955, which justifies the reason for selecting the EC-EUR by matching the C-EUR of the Substructures.

Dataset 1			
Building Code	**Foundation Type**	**Substructure C-EUR (£/m²)**	**Substructure EC-EUR ($kgCO_2/m^2$)**
D1001	Raft	1099	1,028
D1002	Raft	863	1,439
D1003	Pile	521	731
D1004	Raft	2532	2,463
D1005	Raft	612	829
D1007	Pile	1797	1,603
D1008	Pile	230	259
D1009	Pile	529	453
D1010	Pile	482	377
D1011	Pile	1066	1,002
D1012	Pile	1210	1,168
D1013	Pad and Strip	355	366

Dataset 3	
D3001	
Foundation	Pile
Substructure C-EUR	£239
Substructure EC-EUR	259 x (239/230) $kgCO_2/m^2$

Figure 5.9: Choosing the appropriate EC-EUR of the Substructure for the buildings in Dataset 3 from Dataset 1

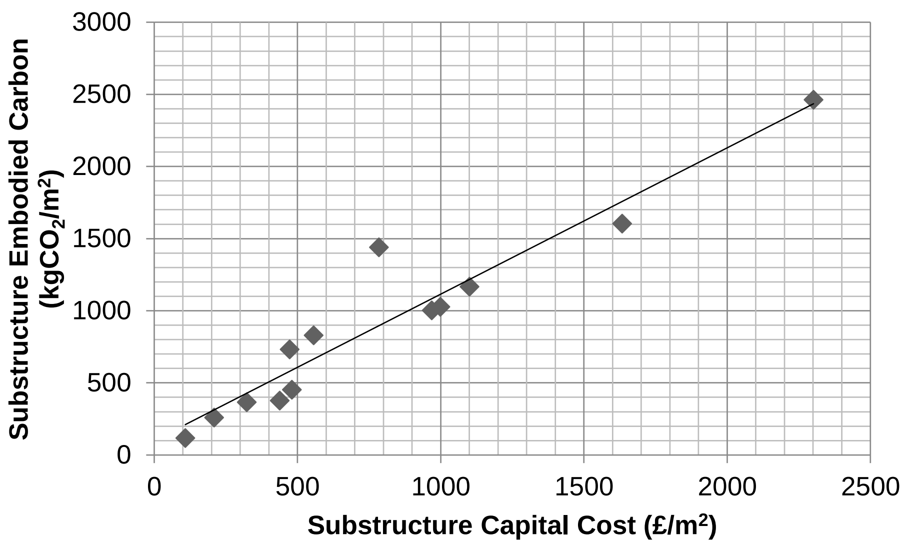


Figure 5.10: Relationship between Substructure capital cost and EC per EUQ – Dataset 1

Frame in the sample had three alternatives. The sample buildings were predominantly steel framed buildings. Rest were concrete frames except for one which was hybrid framed (combination of steel and concrete). Similar to Substructure, Frame is also measured in m^2, which makes it challenging to develop EC-EURs for Frame from Blackbook, as the detailed specification was not available. The averages of the EC-EURs for steel, concrete and hybrid frame was not used to estimate the EC of Frames in Dataset 3 because generally the building cost per GIFA increases as the building height increases (Picken and Ilozor, 2015). Further, BCIS average prices of elements also suggest that the EUR of the Frame increases with the building height (RICS, 2016). However, when the C-EURs of the Frames of the sample buildings in Dataset 3 were plotted in a graph against the building heights (see, Figure 5.11) no significant relationship (p value > 0.05) was found between Frame cost and building height. The building with the highest Frame cost was not the tallest building in the sample. Therefore, it can be deduced from Figure 5.11 that even though it is said that taller buildings have higher Frame cost there is no enough evidence to prove that there is a significant relationship between the Frame C-EURs and the building heights from the given sample.

However, it was anticipated that there could be a relationship between C-EUR and EC-EUR of Frames. Hence, C-EUR and EC-EUR of Frames of Dataset 1 were plotted in a graph as shown Figure 5.12. The correlation coefficient between EUR and EC-EUR of Frames was found to be significant and strong with a correlation coefficient of 0.744. The quantity of steel or concrete primarily determines the cost and EC of frames. Therefore, when quantity of steel or concrete is low, cost tends to be low; thus, EC also tends to be low. Due to the strong positive relationship between C-EUR and EC-EUR, comparable EC-EUR values from Dataset 1 were obtained to estimate EC of Frames in Dataset 3, similar to the EC estimating approach followed for Substructure.

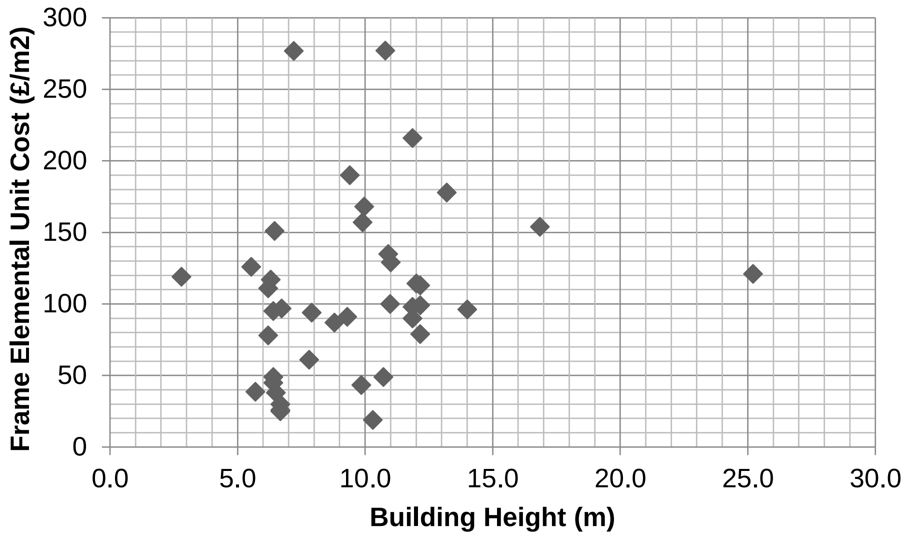


Figure 5.11: Mapping frame element unit cost against building height

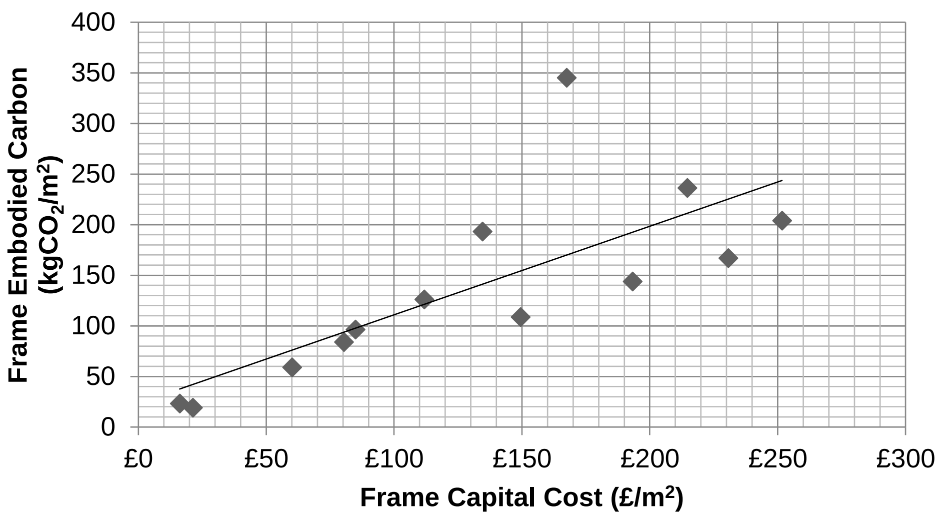


Figure 5.12: Relationship between frame capital cost and EC per element unit quantity – Dataset 1

.

Upper Floors in the sample had following alternatives including in-situ concrete floors, pre-cast concrete floors, metal decking, timber floors and the combination of two or more of the above. Similar EC-EUR values were noticed for in-situ concrete floors of the buildings in Dataset 1; hence, an average of the EC-EURs for in-situ floors was used to estimate the EC of the in-situ concrete floors. The average EC-EUR of in-situ floors was found to be 160.76 $kgCO_2/m^2$ with a standard deviation of 24.84 $kgCO_2/m^2$. Pre-cast floor, metal decking and timber floor EC were calculated using the Blackbook. Further, a composite rate was developed when there were combinations of more than one type of floor as shown in Table 5.18.

Table 5.18: Method of estimating composite rate for upper floors

Building A	**Quantity (m^2)**	**Elemental EC Rate ($kgCO_2/m^2$)**	**Total EC ($kgCO_2$)**	**Source**
In-situ concrete floors	500	160.76	80,380	Dataset 1
Pre-cast floors	300	98.73	29,619	Blackbook
Timber floor	200	12.23	2,446	Blackbook
Total EC of Upper Floors	**1,000**	**112.445**	**112,445**	

Roof of the sample had many alternatives and combinations of different types of roofs. Dataset 1 also had different specifications for roof including concrete flat roof, timber pitched roof, timber flat roof, timber mansard roof, metal decking, green roof, atrium glass roof, steel trussed roof with various roof finishes. Individual rates were developed for all the different types of roof from Dataset 1 and used as the basis to develop elemental EC rates for roof for Dataset 3. A similar approach used in upper floors (see, Table 5.18) to develop element EC rate for a combination of specification was adopted for roof as well.

Stairs had two alternatives – concrete and steel stairs. However, no detailed specification was available. Therefore, EC was estimated based on the assumptions on the dimensions for treads, risers, the width of the stairs, depth of the landing, reinforcement factor and the like and measuring quantities approximately. Finishes to stairs were also included in this element as defined by NRM compliant BCIS element classification.

External Walls were mainly cavity walls and curtain walls. However, other types of cladding were also formed part of the external wall. EC of the external walls was developed using the Blackbook rates.

Windows and External Doors included aluminium, steel, timber and glazed (single and double glazed) windows. EC factors for the windows and doors were obtained from Blackbook. While windows and external doors were measured in m^2 in most of the cases, it was measured in numbers in some buildings. Hence, when it was measured in numbers, standard sizes of windows and doors specified in the Blackbook were assumed for the purpose of estimating.

Internal Walls and Partitions had following alternatives including brick, block, concrete, timber, metal stud, glass, WC cubicles, and a combination of the above. EC of the internal walls and partitions were developed using the Blackbook rates and where there was a combination, the method presented in Table 5.18 was followed.

Internal Doors were mainly timber, steel, aluminium or glazed. The Same strategy followed in windows and external doors was followed in estimating the EC of the internal doors.

Wall Finishes had six alternatives include plastering, painting, tiling, wallpapers, board linings, claddings and combinations of the above. EC of the wall finishes was estimated using the Blackbook data and EC of the different combinations of wall finishes was estimated as explained in Table 5.18.

Floors Finishes had a number of alternatives such as exposed concrete, floor paint, screeds, protective finishes, carpet, flexible thin sheets and tiles, rigid tiles, stone finish, access floors, timber finish, and combinations of the above. Floor finishes were handled in a way, which was similar to wall finishes.

Ceiling Finishes also had six alternatives including plastering, painting, paper finish, board finishes, suspended ceiling systems, and combinations of the above. The same method used to estimate EC of the wall and floor finishes was used to estimate the EC of the ceiling finishes.

Fittings, Furnishings and Equipment EC calculation was challenging due to insufficient details. On the other hand, EC of Fittings, Furnishings and Equipment is insignificant which makes it wasteful to invest more time in it. Consequently, average EC-EUR obtained from Dataset 2 was used as the benchmark value. The EC-EUR of Fittings, Furnishings and Equipment was found to be 0.86kgCO_2/m^2, which has a standard deviation of 1.15.

Building Services EC calculation was also challenging due to insufficiently detailed specification. The Blackbook also contains EC factors for fundamental building services and the EC data is not comprehensive. Especially, the EC factors for services like air conditioning, ventilation systems, communications and security installations and special installations are not available in the Blackbook. In addition to that, data for electrical installation is available only for small scale housing development. All these reasons make it challenging to estimate building services EC. Nevertheless, the EC estimates of the building will be incomplete without the inclusion of building services. Hence, EC-EURs were developed from Dataset 2 for each type of services and used to estimate EC of Services in Dataset 3 based on the service provision in the building considered. Average values of each type of building service were calculated and the descriptive statistics is presented in Table 5.19.

Table 5.19: Descriptive statistics of building services from Dataset 2

Sub Elements of Building Services (NRM)	Mean (kgCO_2/m^2)	Standard Deviation	Sample	Minimum	Maximum
5A Sanitary appliances	0.597	0.871	25	0.002	3.542
5B Services equipment	5.224	-	1	-	-
5C Disposal Installations	6.399	4.846	17	0.026	15.557
5D Water installations	1.854	2.214	17	0.015	9.659
5E Heat source	4.487	2.381	15	1.934	10.597
5F Space heating and air treatment	29.769	25.008	20	0.211	89.007
5G Ventilating system	18.678	12.421	19	1.233	39.553
5H Electrical installation	29.782	14.843	27	6.683	55.933
5I Gas installation	1.185	0.116	2	1.103	1.267
5J Lift and conveyor systems	9.241	6.664	27	1.728	32.495
5K Protective installation	11.900	4.382	20	1.796	26.332
5L Communication installations	16.590	13.497	14	0.573	31.707
5M Special installations	13.808	13.076	9	0.481	38.799

As can be seen from Table 5.19 sample size varies from one type of service to the other, which conveys that not all 28 buildings had all types of services EC estimated.

5.7.3. Validation of Dataset 3

EC analyses of twenty-nine (29) office buildings were obtained from WRAP database, which conforms to cradle-to-gate system boundary. The EC analyses available in WRAP database are presented as six element categories (Substructure, Superstructure Structural, Superstructure Non-Structural, Envelope, Internal Finishes and External Works - see, Section 3.6 (d)) rather than individual elements as defined in NRM. However, only five categories were considered for this validation as External Works were excluded from all analyses. Further, the dataset obtained from WRAP database did not have Envelope EC values included in the analyses. As a result, Envelope EC values could not be validated. Therefore, only four element categories were validated including Substructure, Superstructure Structural, Superstructure Non-Structural and Internal Finishes.

EC analyses of Dataset 3 were altered to suit WRAP analyses to allow comparisons of the EC values. Accordingly, EC of Superstructure Structural of the buildings in Dataset 3 was derived by adding the EC of Frame, Upper Floors and Roof. The sum of the EC of Internal Walls and Partitions and Internal Doors gave the EC of Superstructure Non-Structural. EC of Internal Finishes was calculated by adding the EC of Wall Finishes, Floor Finishes and Ceiling Finishes. In this way, two independent samples of EC of office buildings were produced as presented in Table 5.20. Two sample t-Test allows comparisons between two independent samples and helps to decide if there is a significant difference between the means of the two groups. Hence, two-sample t-Test was conducted to compare the mean EC values of the four element categories of WRAP dataset and Dataset 3.

Table 5.20: Datasets for two-sample t-Test

Data	Sample size
WRAP dataset	29
Dataset 3	41

Test hypothesis has to be established before performing the t-Test. Accordingly,

H_0: $\mu_{EC\text{- WRAP}} = \mu_{EC\text{-Dataset 3}}$ (Null hypothesis)

H_1: $\mu_{EC\text{- WRAP}} \neq \mu_{EC\text{-Dataset 3}}$ (Alternative hypothesis)

where, $\mu_{EC\text{- WRAP}}$ is the mean EC of WRAP dataset (of each element), $\mu_{EC\text{-Dataset 3}}$ is the mean EC of Dataset 3 (of each element).

Descriptive statistics of the element groups are presented in Table 5.21and the test statistics of Levene's Test for equality of variances and t-Test are presented in Table 5.22. The test statistics suggest that the variances of Substructure and Superstructure Non-Structural are not significantly different (sig. > 0.05) and the variances of Superstructure Structural and Internal Finishes are significantly different (sig. < 0.05) of the two samples. Accordingly, the relevant t-test statistics were examined to arrive at a conclusion about the means of the two groups, which are shown in greyscale in Table 5.22. The t-Test statistics display with 95% confidence that there is no sufficient evidence to say that the means of the Substructure, Superstructure Non-Structural and Internal Finishes of the two samples are significantly different. In other words, the difference between the EC values of Substructure, Superstructure Non-Structural and Internal Finishes of the two samples are not statistically significant. On the other hand, there is sufficient evidence (sig. < 0.05) to conclude that there is a significant difference between the mean EC values of Superstructure Structural of the two samples, i.e. the difference between the EC values of Superstructure Structural of the two samples is statistically significant.

Table 5.21: Group statistics for individual element categories

Element Group	Group	Sample size	Mean	Std. Deviation	Std. Error Mean
Substructure	WRAP	29	146.08	74.13	13.77
	Dataset 3	41	161.158	57.53	8.98
Superstructure Structural	WRAP	29	363.84	116.01	21.54
	Dataset 3	41	219.45	63.80	9.96
Superstructure Non-Structural	WRAP	29	34.67	49.77	9.24
	Dataset 3	41	25.40	33.76	5.27
Internal Finishes	WRAP	29	55.68	36.87	6.85
	Dataset 3	41	54.64	16.06	2.51

Table 5.22: t-Test statistics of the two samples from Dataset 3 and WRAP database

Element Category		Levene's Test for Equality of Variances		t-test for Equality of Means					Test Outcome
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
Substructure	Equal variances assumed	1.416	.238	-.957	68	.342	-15.07294	15.74265	Fail to reject H_0
	Equal variances not assumed			-.917	50.522	.364	-15.07294	16.43799	
Superstructure Structural	Equal variances assumed	5.030	.028	6.680	68	.000	144.38844	21.61510	Reject H_0
	Equal variances not assumed			6.083	39.979	.000	144.38844	23.73529	
Superstructure Non-Structural	Equal variances assumed	1.123	.293	.928	68	.357	9.26064	9.97663	Fail to reject H_0
	Equal variances not assumed			.870	45.797	.389	9.26064	10.64089	
Internal Finishes	Equal variances assumed	32.743	.000	.161	68	.873	1.04135	6.47304	Fail to reject H_0
	Equal variances not assumed			.143	35.573	.887	1.04135	7.29264	

However, the WRAP sample also contained high-rise office buildings (10 storeys and above) to obtain a statistically significant sample. Hence, the same tests were conducted after removing the high-rise buildings from WRAP sample to compare only the means of low to medium rise office buildings, which are the scope of the study. The test statistics are presented in Table 5.23 and Table 5.24. Even though a different observation was noticed regarding the variance of Substructure and Superstructure Structural, the outcome was in line with the previous test statistics. Variances of Superstructure Structural and Superstructure Non-Structural are not significantly different (sig. > 0.05) and the variances of Substructure and Internal Finishes are significantly different (sig. < 0.05) of the two samples. Further, there is no sufficient evidence to say that the means of the Substructure, Superstructure Non-Structural and Internal Finishes of the two samples are significantly different. Meanwhile, there is sufficient evidence (sig. < 0.05) to conclude that there is a significant difference between the mean EC values of Superstructure Structural of the two samples.

Table 5.23: Group statistics for individual element categories – reduced sample

	Group	Sample size	Mean	Std. Deviation	Std. Error Mean
Substructure	WRAP	18	146.34	86.68	20.43
	Dataset 3	41	161.15	57.53	8.98
Superstructure Structural	WRAP	18	325.44	88.26	20.80
	Dataset 3	41	219.45	63.80	9.96
Superstructure Non-Structural	WRAP	18	31.29	52.20	12.30
	Dataset 3	41	25.40	33.76	5.27
Internal Finishes	WRAP	18	63.79	33.27	7.84
	Dataset 3	41	54.64	16.06	2.51

Table 5.24: Independent sample t-Test statistics

Element Category		Levene's Test for Equality of Variances		t-test for Equality of Means					Test Outcome
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
Substructure	Equal variances assumed	4.855	.032	-.775	57	.441	-14.81011	19.10033	Fail to reject H_0
	Equal variances not assumed			-.664	23.83	.513	-14.81011	22.31827	
Superstructure Structural	Equal variances assumed	2.046	.158	5.209	57	.000	105.98489	20.34836	Reject H_0
	Equal variances not assumed			4.595	25.13	.000	105.98489	23.06522	
Superstructure Non-Structural	Equal variances assumed	.531	.469	.518	57	.607	5.88012	11.35451	Fail to reject H_0
	Equal variances not assumed			.439	23.48	.664	5.88012	13.38595	
Internal Finishes	Equal variances assumed	16.088	.000	1.432	57	.158	9.15619	6.39276	Fail to reject H_0
	Equal variances not assumed			1.112	20.57	.279	9.15619	8.23295	

Based on the test statistics it can be concluded that there is no significant difference in the means of EC of Substructure, Superstructure Non-Structural and Internal Finishes of the two samples, which validate the reliability of the EC values of these elements in Dataset 3. However, a significant difference in EC values of Superstructure Structural (consists of Frame, Upper Floors and Roof) is revealed. Possible reason for this difference could be attributable to Roof EC as there are several design options available for Roof and the specification is not available for WRAP data to study the differences in the specification of the two samples. Similarly, EC of Upper Floors could also have influenced the identified difference if WRAP data sample predominantly consists of timber floors and pre-cast floors. Hence, it is concluded that most of the estimates of the EC of Dataset 3 are reliable. However, it is acknowledged that there is some ambiguity about the estimate of Superstructure Structural EC and the reliability of the estimate cannot be warranted without any additional information on element specification, which is unknown in this case. This is identified as a limitation of the study.

5.8. Development of the Finishes Quality Index

The need was identified to develop a finishes quality index to capture the finishes quality of the buildings in a uniform and systematic way as explained in the Section 4.8.2. Data was collected from the expert forum in order to develop finishes quality index. The process of data collection is presented in Figure 5.13. Commonly used types of wall, floor and ceiling finishes in office buildings in the UK were surveyed from Dataset 1, Dataset 3 (the sample comprises 13 buildings from Dataset 1 and 41 buildings from Dataset 3) and the UK price books and a list of finishes were prepared. The identified types of finishes were classified into one of the three categories: Basic, Moderate and Luxury, for wall, floor and ceiling separately and a conceptual finishes index was developed as presented in Figure 5.13. The developed conceptual finishes quality index was verified through a Delphi-based expert forum to improve the rigorousness of the proposed finishes quality index of the study.

As discussed in the methodology chapter (see, section 4.8.2) an expert in a Delphi-based expert forum is someone who possesses knowledge and has experience in the particular field of study. Accordingly, construction professionals with more than 10 years of industry experience and with a Royal Institution of Chartered Surveyors (RICS) membership were chosen purposively to be the experts on the panel. RICS is a leading professional body in land, real estate, infrastructure and construction and RICS membership has an international recognition. The profile of the experts is presented in Table 5.26. The expert panel consisted of four QSs and an architect. One of the core duties of QSs is early stage estimating and advising clients on design solutions (Ashworth and Perera, 2015, RICS, 2015). Therefore, QSs are expected to be competent in determining the quality level of finishes and RICS membership ensures that its members are equipped with necessary competencies. For instance, QSs are examined on competencies falling under three categories including mandatory, core and optional competencies where 'Design economics and cost planning' is a core competency expected to be demonstrated by QSs at the Level 2 (knowledge and understanding into practice) or Level 3 (providing professional advice to clients) competency level defined by

RICS. This confirms the suitability of QSs being identified as experts to verify the conceptual finishes quality index. However, it was also decided to include an Architect in the panel to eliminate homogeneity and ensure consistency in cross-disciplinary judgement. Only one Architect was employed mainly due to the time constraint. Accordingly, the selected respondents/experts for the panel demonstrated a strong work profile of handling office projects and delivering cost advice to the client during early stages of design. Hence, the judgement of the experts can be considered contemporaneous and applicable to the present construction industry standards.

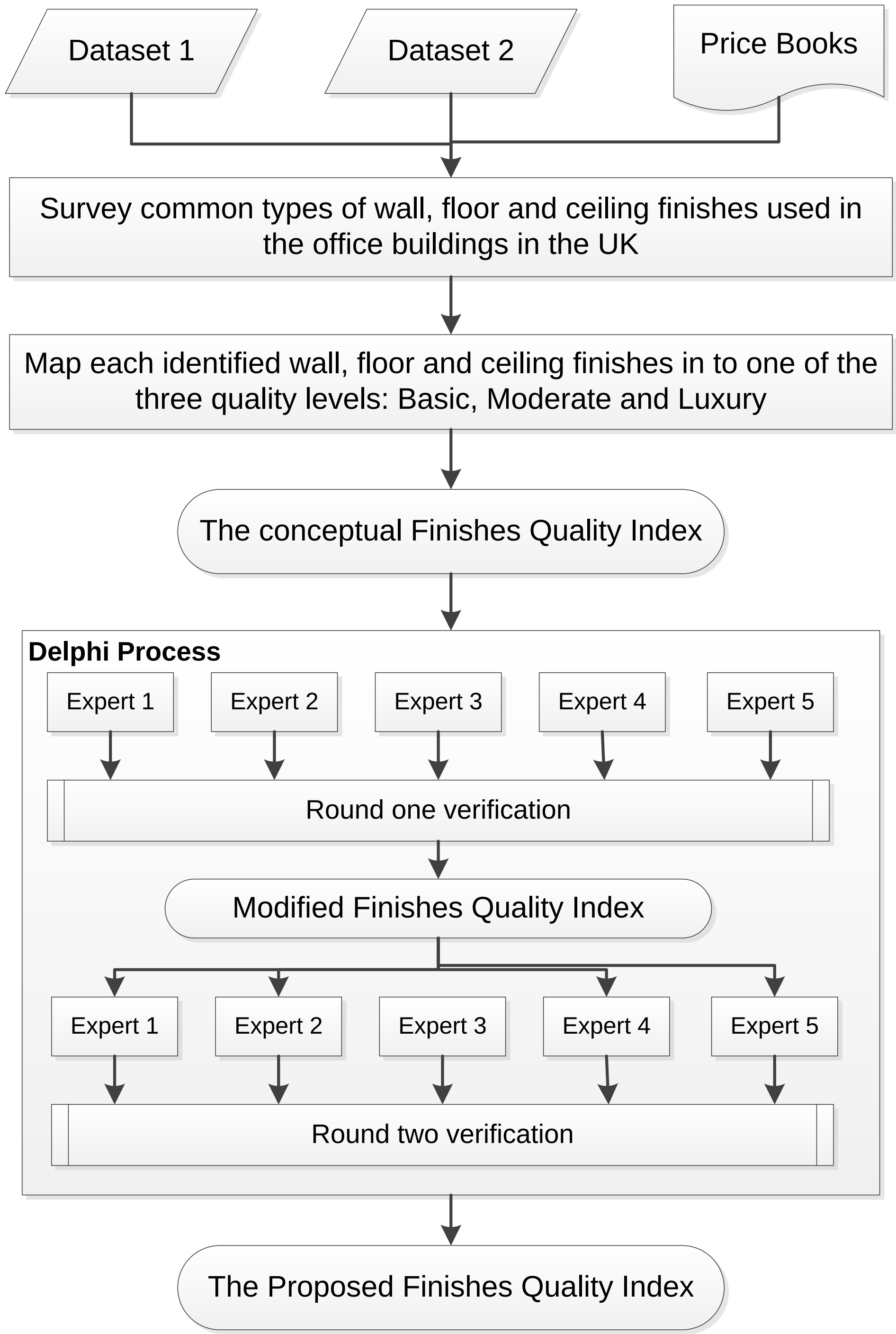


Figure 5.13: The development of the finishes quality index

Table 5.25: The conceptual finishes quality index

Element	Quality Level	Main Categories	Sub categories
Wall Finishes	**Basic**	Plaster and paint	Cement sand plaster and paint
			Lime plaster and paint
		Wallpaper	Lining paper
			Vinyl paper
	Moderate	Plaster and paint	Gypsum plaster and paint
			Carlite lightweight plaster
		Board linings	Plasterboard and paint
			Insulated plasterboard and paint
			Plywood wall panels
			Wallboards
			Hardboard
			Softwood boarding
			Chipboard
			Veneered MD panels
		Wall tiles	Ceramic tiles
	Luxury	Wall tiles	Mosaic tiles
		Stone cladding	Natural granite
			Marble
		Other claddings	Composite aluminium
			Glass
Floor Finishes	**Basic**	Exposed concrete	Concrete hardener
		Screeds	Cement sand
			Latex screed
		Protective finish	Mastic asphalt floor
		Carpet	Medium duty
		Rigid Tiles	Slate tiles
			Cement tiles
	Moderate	Screeds	Granolithic
		Resin based finish	Epoxy floor
		Flexible thin sheets and tiles	Rubber floor tiles
			Linoleum sheet
			Marmoleum
			Linoleum tiles
			Vinyl sheet
			Vinyl tiles
			Cork tiles
			Carpet tiles
		Rigid tiles	Ceramic tiles
			Clay tiles
		Carpet	Heavy duty
		Stone finish	Terrazzo

		Access floor	Metal access floors
	Luxury	Rigid tiles	Quarry tiles
			Porcelain tiles
			Mosaic tiles
		Timber floor	Woodblock floor
			Woodstrip floor
			Veneered laminated floor
			Parquet floor
		Stone finish	Natural granite
			Marble
Ceiling Finishes	**Basic**	In-situ finishes	Sealer
			Skim coat and paint
			Cement sand plaster and paint
			Lime plaster and paint
		Paper finish	Lining paper
			Vinyl paper
	Moderate	In-situ finishes	Gypsum plaster and paint
			Carlite lightweight plaster
		Board finishes	Plasterboard and paint
		Suspended ceiling systems	Metal frame plasterboard ceilings
			Plasterboard acoustic ceilings
			Moisture resistant ceilings
			Metal suspended ceilings
	Luxury	Integrated/composite ceiling	Coffered ceilings

Table 5.26: Profile of the experts

Expert ID	Role	Years of Industry Experience
Expert 1	Senior Quantity Surveyor	35
Expert 2	Senior Quantity Surveyor	30
Expert 3	Executive Quantity Surveyor	14
Expert 4	Project Architect and Construction Project Manager	15
Expert 5	Senior Quantity Surveyor	50

As discussed in the Methodology (see, Section 4.8.2 and 4.9.3(b)), Delphi technique was employed to verify the conceptual finishes quality index by allowing each expert to comment on the identified finishes quality levels. Experts were consulted virtually via email communications individually and their comments were recorded. The comments of the experts were allowed to take three different forms as follows:

1. Addition of new types of finishes
2. Removal of any of the proposed finishes
3. Changes to proposed quality level (Basic, Moderate, Luxury)

The conceptual finishes quality index was modified to accommodate the comments of all the experts from the first round and a summary of the finishes quality levels to which consensus was not reached was sent to the experts via email. Round two was conducted to reconcile contradicting comments and to achieve consensus among the experts. This was done by compiling the comments of experts and producing a summary of comments and sharing with all the experts to allow them to re-evaluate their responses based on the examination of the response of the group. The consensus in the research context was considered to be achieved when 4 of the 5 respondents agreed on a particular quality level (refer Section 4.8.2). Therefore, iteration was stopped at the end of two rounds because the consensus was achieved among the experts in two rounds.

5.9. Development of the Services Quality Index

Price books were reviewed to study different services quality levels proposed in the industry published guides. These guides are intended to assist in the estimating of the cost of construction projects across different stages of the project. The summary of the reviewed sources is presented in Table 5.27. None of the price books except for Spon's Mechanical and Electrical Services Price Book 2014 (Davis Langdon Consultancy, 2014) have provision for identifying the quality level of services installation for early design stages.

Table 5.27: Proposed services quality levels from literature

Price Book	Proposed Quality Levels For Services	Reference
Spon's Mechanical and Electrical Services Price Book 2014	Non air-conditioned buildings; non air-conditioned automated buildings; air-conditioned automated buildings	Davis Langdon Consultancy (2014)
Comprehensive Building Price Book Major Work 2013	No abstract level quality classification is proposed. Detailed level specifications are presented to allow detailed stage estimates.	BCIS (2013)
Griffiths Price Book 2012		Franklin and Andrews (2012)
Spon's First Estimating Handbook		Spain (2010)
Estimating Price Book SMM7 2008		RICS (2008a)
Laxton's Building Price Book Major and Small Works		Johnson (2008)

5.10.Summary

A pilot study was conducted to determine the feasibility of obtaining historical project data from BCIS online cost database to develop EC estimates. However, the pilot study proved the use of BCIS alone to be unsuccessful due to inadequate building data found in BCIS and the lack of industry developed elemental EC benchmarks. Consequently, historical project data were collected from both primary and secondary data sources. Published EC and the cost databases were used as supporting data sources to develop EC and cost estimates. Primary data consists of 13 office buildings (Dataset 1) and secondary data were collected from three different sources: special database from a QS consultancy practice (Dataset 2), BCIS online cost database (Dataset 3), and WRAP EC databases.

Dataset 3 was developed for statistical analysis using the inputs from Dataset 1, Dataset 2 and published data books. EC-EURs of Substructure, Frame, Upper Floor (only in-situ) and Roof were obtained from Dataset 1 while EC-EURs of Fittings, Furnishing and Equipment and Services were obtained from Dataset 2 to develop the EC estimates of the respective elements of Dataset 3. On the other hand, EC of the rest of the building elements was calculated using the UK Building Blackbook, ICE and manufacturers' EC data. Further, two sample t-Tests were conducted for Dataset 1 and Dataset 2 separately to compare the means of low to medium rise and high rise buildings within each dataset to ensure that the mean values do not differ significantly (at $\alpha = 0.05$). The t-Test results suggested with 95% confidence that the variances of the population of the two groups are equal and there is no sufficient evidence to conclude that there is a significant difference in the values of EC of Substructure, Frame, Upper Floors, Roof, Furnishing and Equipment and Services of low to medium rise and high rise buildings.

Similarly, the reliability of Dataset 3 was confirmed by comparing EC values of Dataset 3 with an independent dataset sourced from the WRAP EC database using two sample t-Test. Based on the test statistics it was concluded that there is no significant difference in the means of EC of Substructure, Internal Walls and Partitions, Internal Doors and Internal Finishes of the two samples, which validate the reliability of the EC values of these elements in Dataset 3. However, a significant difference in EC values of Superstructure Structural (consisting of Frame, Upper Floors and Roof) was detected which could be attributable to Roof or Upper Floor EC due to distinctive element specifications. It was concluded that most of the estimates of the EC of Dataset 3 are reliable with the caution that there is an ambiguity about the estimate of Superstructure Structural EC.

Meanwhile, qualitative data were collected to develop finishes and services quality indices in an objective way to incorporate finishes and services quality as predictors in the model as these were identified as cost and carbon influential design variables. Accordingly, data for the development of finishes quality index was collected through a Delphi-based expert forum where five experts including four QSs and an Architect formed the expert. The conceptual finishes index was presented to each expert separately and a summary of comments of experts were presented to all experts and a chance was given to experts to re-evaluate their responses based on the examination of the response of the group. Consensus was reached at the end of the second round, hence, the data collection was stopped at this point. On the other hand, data for services quality index development was collected by reviewing proposed services quality levels in published price books.

6. Data Analysis

6.1. Introduction

The analysis of data is presented in five main themes including the carbon and cost hotspot analysis, the development of services and finishes indices, pre-regression analysis, regression analysis and EC and cost relationships. Carbon hotspots analysis of Dataset 2 and Dataset 3 and cost hotspots analysis of Dataset 3 are presented initially which became the basis for the selection of the most influential design variables of the models. As suggested in the literature, finishes and services quality of buildings were identified as cost and carbon influential design variables from the hotspot analysis. Hence, finishes and services quality indices were decided to be developed. In light of developing finishes and services quality indices, data collected from expert forum for finishes quality index development and data collected from documents for services index development were content analysed and the outcomes are presented in separate sections. Pre-regression analysis was performed before the actual regression analysis to ensure better model building, which includes univariate and bivariate analysis. Univariate analysis is used to describe the distributions of variables and identify outliers and bivariate analysis is used to find correlations and multicollinearity between independent variables. After the diagnostics, regression analysis was performed with the selected variables to formulate EC and CC models which are presented in separate sub-sections. The outcomes of the regression analysis are presented with and without outliers and the better model among the two was selected as the final model. In addition, the relationship between the EC and the CC was explored at building level and individual element level using correlation analysis.

6.2. Analysis of Carbon and Cost Hotspots

Carbon and cost hotspots in the sample office buildings were identified through elemental EC analysis using the Pareto 80:20 rule (as discussed in the methodology chapter section 4.9.1). The elements that are responsible for 80% of EC and CC were identified for each building separately in the manner presented in

Table 6.1. Firstly, EC of individual elements was estimated and the percentage contribution was found. Then, the elements were arranged from the largest to smallest in terms of EC (and CC separately). Then, the cumulative percentage was calculated to draw a cut-off point at 80% as shown in Table 6.1. Accordingly, Frame, External Walls, Services and Substructure are found to be the carbon hotspots of the particular building presented in Table 6.1. In addition to the individual building analysis, hotspots were analysed for the whole sample and the carbon and cost hotspots in office buildings were identified and presented in Section 6.2.1 and Section 6.2.2 respectively.

Table 6.1: Identifying carbon hotspots of a building – an example

Building Elements	**EC % (in descending order)**	**Cumulative EC%**
2A Frame	38.54	38.5
2E External walls	20.30	58.8
5 Services	13.82	72.7
1A Substructures	9.90	82.6
2B Upper floors	6.71	89.3
2C Roof	3.94	93.2
2D Stairs	2.44	95.7
2G Internal walls and partitions	1.66	97.3
3B Floor finishes	1.50	98.8
4A Fittings and furnishings	0.43	99.2
3A Wall finishes	0.34	99.6
2H Internal doors	0.32	99.9
3C Ceiling finishes	0.09	100.0
2F Windows and external doors	0.01	100.0

Later, the building elements were classified into three categories namely: elements that are identified as a carbon (or cost) hotspot in:

1. most of the buildings/Lead positions (Probability of occurrence $\geq$ 0.8)
2. some of the buildings/Special positions (0 < Probability of occurrence < 0.8)
3. none of the buildings/Remainder positions (Probability of occurrence = 0)

Of the three types, first two types needed to be considered in modelling as those elements clearly have an impact on the EC and the CC of the building. Eventually, design variables that influence the hotspot elements were identified which became the predictors of the models.

6.2.1. Carbon Hotspot Analysis

Carbon hotspots were analysed for the buildings in Dataset 2 and Dataset 3 separately as Dataset 3 was inspired partially by Dataset 1 and Dataset 2. As mentioned before, buildings were analysed both individually and as a whole. Table 6.2 presents the carbon hotspots of the 28 buildings. For instance, carbon hotspots of the building #D1001 were Substructure, Frame, External Walls and Services. Afterwards, the probability of each element being found as a hotspot in the given sample was calculated and presented in the bottom of the table. Accordingly, Frame found to be a hotspot in all the buildings; Substructure and Services found to be a hotspot in 90% of the buildings, and External Walls found to be a hotspot in 80% of the buildings in the sample. On the other hand, elements like Stairs, Internal Doors, Wall Finishes, Ceiling Finishes and Fittings and Furnishings were not found as hotspots in any of the sample buildings. Rest of the elements were found to be hotspots in some of the buildings.

Table 6.2: Carbon hotspot analysis of Dataset 2

Building ID	1A Substructures	2A Frame	2B Upper floors	2C Roof	2D Stairs	2E External walls	2F Windows and external doors	2G Internal walls and partitions	2H Internal doors	3A Wall finishes	3B Floor finishes	3C Ceiling finishes	4A Fittings and furnishings	5 Services
#D2001	x	x				x								x
#D2002	x	x				x								x
#D2003	x	x	x			x								
#D2004	x	x				x								x
#D2005	x	x				x								x
#D2006	x	x	x			x								x
#D2007	x	x				x								x
#D2008		x	x			x					x			x
#D2009	x	x				x					x			x
#D2010	x	x	x			x								x
#D2011	x	x	x											
#D2012	x	x	x			x								x
#D2013	x	x	x			x								
#D2014	x	x	x			x								x
#D2015	x	x	x											x
#D2016	x	x	x			x								x
#D2017	x	x				x								x
#D2018	x	x				x								x
#D2019	x	x				x								x
#D2020		x	x	X			x				x			x
#D2021		x	x	X		x					x			x
#D2022	x	x		X			x				x			x
#D2023	x	x	x			x		x						x
#D2024	x	x	x	X										x
#D2025	x	x												x
#D2026	x	x	x			x	x							x
#D2027	x	x	x											
#D2028	x	x	x			x								x
Probability of occurrence	0.9	1	0.6	0.1	0	0.8	0.11	0	0	0	0.2	0	0	0.9

	Lead Positions
	Special Positions
	Remainder Positions

In the same way, Dataset 3 was analysed and the results are presented in Table 6.3. Carbon hotspots of Dataset 3 overlap with the carbon hotspots of Dataset 2. Substructure, Frame, Upper Floors, Services were identified as hotspots in most of the buildings in Dataset 3 while Stairs, Internal Doors, and Fittings and Furnishings were not identified as carbon hotspots in any of the buildings. The analysis of Dataset 3 identifies more elements as carbon hotspots; hence, increasing the uncertainty.

Based on the observation of carbon hotspots in both the samples, the building elements were classified into three categories such as ‘Lead Positions’, ‘Special Positions’ and ‘Remainder Positions’. Lead positions were the elements that always or mostly found as carbon hotspots in the buildings. Special positions were the building elements that were occasionally found to be a carbon hotspot in office buildings. Subsequently, the design variables influencing the identified carbon hotspots are presented in Table 6.5. However, not all the design variables identified in Table 6.5 are likely to be available during the early stages of design. Hence, the prediction models needed to be based on the variables that are most likely to be available during the early stages of design which includes GIFA (≈ footprint area + upper floor area), building height, average height, no. of basements, façade area or Wall to Floor ratio and circulation space ratio.

Table 6.3: Carbon hotspot analysis of Dataset 3

Building ID	1 Substructure	2A Frame	2B Upper Floors	2C Roof	2D Stairs	2E External Walls	2F External Windows and Doors	2G Internal Walls and Partitions	2H Internal Doors	3A Wall Finishes	3B Floor Finishes	3C Ceiling Finishes	4 Fittings and Furnishings	5 Services
#D3001	x		x	x		x		x						x
#D3002	x	x		x		x		x						x
#D3003	x	x		x		x								x
#D3004	x	x	x			x								x
#D3005	x	x	x			x								x
#D3006	x	x	x	x										x
#D3007	x	x	x	x		x								x
#D3008	x	x	x											x
#D3009	x	x	x	x		x								x
#D3010	x	x	x			x								x
#D3011	x		x	x		x					x			x
#D3012	x		x	x		x					x			x
#D3013	x		x	x		x					x			x
#D3014	x	x	x											x
#D3015	x	x	x	x		x								x
#D3016	x	x	x			x								x
#D3017	x		x	x		x						x		x
#D3018	x	x	x	x		x								x
#D3019	x	x	x					x						x
#D3020	x	x	x					x						x
#D3021	x	x	x					x						x
#D3022	x	x		x		x								x
#D3023	x	x	x	x										x
#D3024	x	x	x	x		x								x
#D3025	x	x	x									x		x
#D3026	x	x	x			x								x
#D3027	x	x	x			x								x
#D3028	x	x	x			x								x
#D3029	x	x	x			x								x
#D3030	x	x	x			x								x
#D3031	x	x	x			x								x
#D3032	x	x	x			x								x
#D3033	x	x	x			x								x
#D3034	x	x		x		x	x							x
#D3035	x	x	x	x										x
#D3036	x	x	x	x										x
#D3037	x	x	x	x										x
#D3038	x	x	x	x							x			x
#D3039	x	x	x			x								x
#D3040	x	x	x	x		x								x
#D3041	x	x		x		x				x				x
Probability of occurrence	1.0	0.9	0.9	0.5	0	0.7	0.02	0.1	0	0.02	0.1	0.05	0	1.0

Table 6.4: Classification of carbon hotspots

Carbon Hotspot Category	Building Elements
Lead positions	Substructure, Frame, Upper Floors, External Walls, Building Services
Special positions	Roof, Windows and External Doors, Internal Walls and Partitions, Wall Finishes, Floor Finishes, Ceiling Finishes,
Remainder positions	Stairs, Internal Doors, Fittings, Furnishings and Equipment

Table 6.5: EC influential design variables

Building Elements	Influential Design Variable
Substructures	Footprint area, no. of basements
Frame	GIFA, average height, building height
Upper floors	Upper floor area
Roof	Roof area
External walls (including Windows and external doors)	Façade area
Internal walls and partitions	GIFA, internal wall area, usable floor area
Wall Finishes	Wall finish area
Floor finishes	Floor finish area
Ceiling Finishes	Ceiling finish area
Building Services	GIFA

In addition to the above analysis, carbon hotspots for the whole sample were analysed and presented in Table 6.6 and Figure 6.1. Accordingly, Substructure, Services, Frame, Upper Floors, External Walls and Roof were identified as the most carbon significant building elements in descending order. On the other hand, it was also noticed that the same building elements were accountable for 72% of the CC.

Table 6.6: Carbon hotspots – Dataset 3

Elements (NRM Classification)	Average EC per GIFA ($kgCO_2/m^2$ GIFA)	Average CC per GIFA (£/m^2 GIFA)	Cumulative Carbon %	Cumulative Cost %
1 Substructure	161	89	23.6	7.0
5 Services	145	419	44.9	39.6
2A Frame	100	102	59.6	47.6
2B Upper Floors	69	57	69.7	52.0
2E External Walls	60	159	78.5	64.5
2C Roof	43	91	84.8	71.6
3B Floor Finishes	26	75	88.6	77.4
2G Internal Walls and Partitions	23	39	92.0	80.5
3C Ceiling Finishes	19	36	94.8	83.3
2F External Windows and Doors	16	94	97.2	90.7
3A Wall Finishes	9	34	98.6	93.3
2D Stairs	8	27	99.7	95.4
2H Internal Doors	1	31	99.9	97.8
4 Fittings and Furnishings	1	28	100.0	100.0

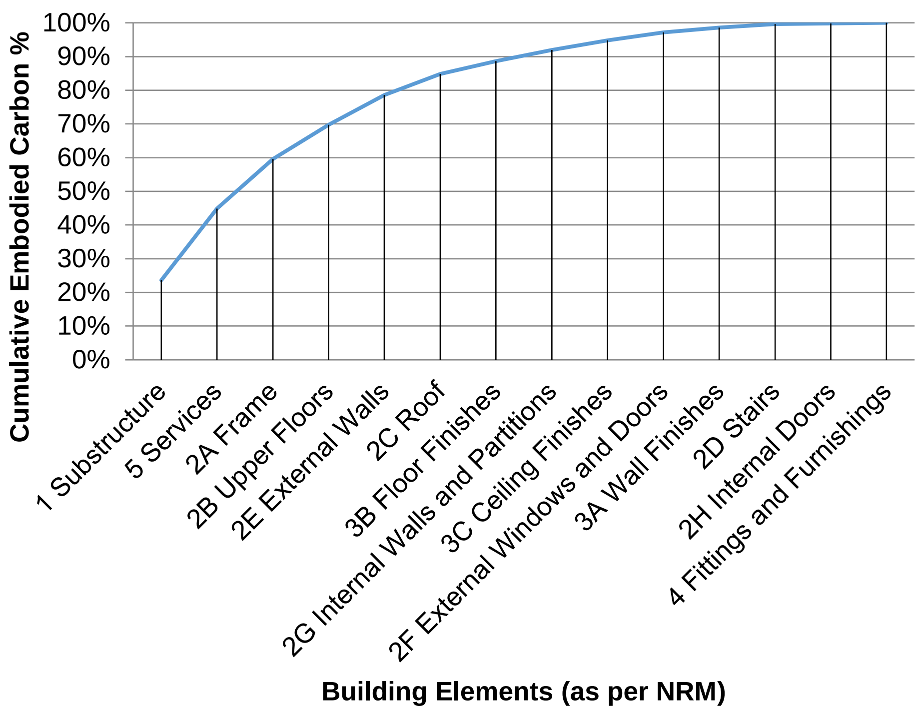


Figure 6.1: Pareto curve for EC in Dataset 3

6.2.2. Cost Hotspot Analysis

Cost hotspots were identified only for Dataset 3 due to unavailability of cost data for Dataset 2. Substructure, Frame, Roof, External Walls, External Windows and Doors and Services were identified as cost hotspots in most of the buildings (Services in all of the buildings). Rest of the elements were found to be cost hotspots in at least one of the sample buildings in Dataset 3.

Cost hotspots for the whole sample are presented in Table 6.8 and Figure 6.2. Accordingly, Services, External Walls, Frame, External Windows and Doors, Roof, Substructure, and Floor Finishes were identified as the most cost significant building elements in descending order. Further, these building elements are also identified to be accountable for 81% of the EC of the buildings on average.

Table 6.7: Cost hotspot analysis of Dataset 3

Project Id	1 Substructure	2A Frame	2B Upper Floors	2C Roof	2D Stairs	2E External Walls	2F External Windows and Doors	2G Internal Walls and Partitions	2H Internal Doors	3A Wall Finishes	3B Floor Finishes	3C Ceiling Finishes	4 Fittings and Furnishings	5 Services
#D3001	x		x	x		x	x				x	x		x
#D3002	x	x	x	x		x	x	x		x		x		x
#D3003	x	x		x	x	x	x		x	x			x	x
#D3004	x		x	x		x	x				x			x
#D3005	x	x	x	x		x	x		x					x
#D3006	x	x		x		x		x						x
#D3007	x		x			x					x			x
#D3008	x	x	x	x		x								x
#D3009	x		x	x		x	x				x			x
#D3010	x	x	x	x		x				x	x			x
#D3011	x			x		x	x				x			x
#D3012	x			x		x	x				x			x
#D3013	x			x	x	x	x				x			x
#D3014	x	x		x			x				x			x
#D3015	x	x		x		x	x				x			x
#D3016	x	x		x		x	x				x			x
#D3017	x	x		x	x	x	x					x		x
#D3018	x	x	x	x	x	x	x					x		x
#D3019		x	x			x	x	x				x		x
#D3020	x	x	x	x		x	x				x			x
#D3021	x	x		x		x	x	x						x
#D3022	x	x		x		x			x					x
#D3023	x	x	x	x			x				x	x		x
#D3024		x	x			x	x				x			x
#D3025	x	x		x		x	x	x						x
#D3026	x	x	x	x			x				x			x
#D3027	x	x		x		x	x				x			x
#D3028	x	x	x			x	x				x			x
#D3029		x		x		x	x				x			x
#D3030	x	x		x		x	x				x			x
#D3031	x	x		x		x	x				x			x
#D3032		x		x		x	x				x			x
#D3033		x	x			x					x			x
#D3034	x	x		x		x	x	x			x		x	x
#D3035		x	x			x					x			x
#D3036	x	x	x			x	x				x			x
#D3037	x	x		x		x	x				x			x
#D3038	x	x		x		x	x				x		x	x
#D3039		x	x			x					x			x
#D3040	x	x	x		x	x	x							x
#D3041	x	x		x		x	x			x				x
Probability of occurrence	0.8	0.8	0.5	0.8	0.1	0.9	0.8	0.1	0.1	0.1	0.7	0.1	0.1	1.0

Table 6.8: Cost hotspots – Dataset 3

Elements (NRM Classification)	Average CC (£/m^2 GIFA)	Average EC ($kgCO_2/m^2$ GIFA)	Cumulative CC %	Cumulative EC %
5 Services	419	145	32.7	21.3
2E External Walls	159	60	45.1	30.1
2A Frame	102	100	53.1	44.8
2F External Windows and Doors	94	16	60.4	47.1
2C Roof	91	43	67.5	53.4
1 Substructure	89	161	74.5	77.0
3B Floor Finishes	75	26	80.3	80.8
2B Upper Floors	57	69	84.7	90.9
2G Internal Walls and Partitions	39	23	87.8	94.3
3C Ceiling Finishes	36	19	90.7	97.2
3A Wall Finishes	34	9	93.3	98.6
2H Internal Doors	31	1	95.7	98.8
4 Fittings and Furnishings	28	1	97.9	98.9
2D Stairs	27	8	100.0	100.0

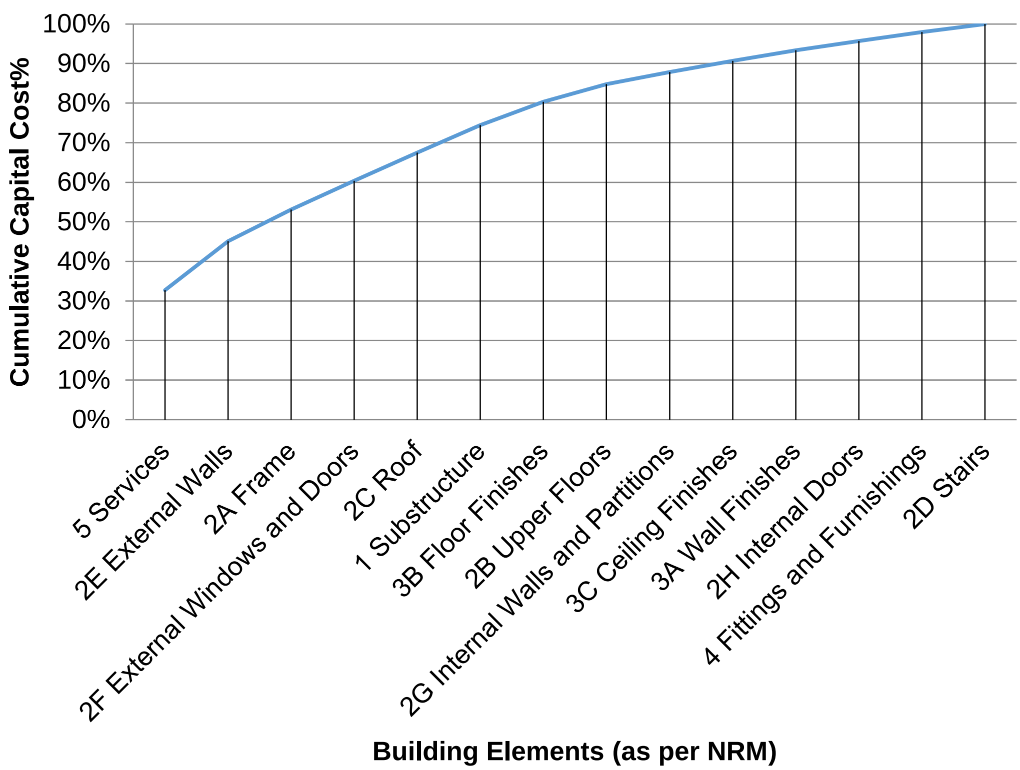


Figure 6.2: Pareto curve for capital cost in Dataset 3

Based on the cost hotspot analysis presented in Table 6.7 building elements were categorised into three types similar to carbon hotspots, which are presented in Table 6.9. Accordingly, Substructure, Frame, Roof, External Walls, Windows and External Doors, and Services were found to be lead positions while all other elements were identified as special positions (which were found to be hotspots in some of the buildings). All elements being identified as cost hotspots in one building or the other increases the uncertainty in decision-making.

Table 6.9: Cost hotspots of office buildings

Cost Hotspot Category	Building Elements
Lead positions	Substructure, Frame, Roof, External Walls, Windows and External Doors, Services
Special positions	Upper Floors, Stairs, Internal Walls and Partitions, Internal Doors, Wall Finishes, Floor Finishes, Ceiling Finishes, Fittings, Furnishings and Equipment
Remainder positions	Nil

6.2.3. Cost and Carbon Hotspots Relationships

Based on the above analysis cost and carbon hotspots were mapped onto the hotspot category as shown in Table 6.10 to gain a better understanding and infer relationships. As discussed in the beginning of the chapter, lead positions are the building elements that are identified as hotspots in most of the buildings; special positions are the building elements that are identified as hotspots in some of the buildings, and remainder position refers to the elements that never found to be hotspots. Accordingly, Substructure, Frame, External Walls and Building Services were found to be lead carbon and cost hotspot in office buildings. Roof and Windows and External Doors were identified as lead cost hotspots and special carbon hotspots. Similarly, Upper Floors identified as lead carbon hotspot and special cost hotspots. Internal Walls and Partitions, Wall Finishes, Floor Finishes and Ceiling Finishes were found to be special carbon and cost hotspots. Further, Stairs, Internal Doors and Fittings, Furnishing and Equipment were identified as the remainder carbon hotspots and special cost hotspots.

Table 6.10: Mapping carbon hotspots and cost hotspots onto the hotspot category

Lead Positions	Special Positions	Remainder Positions
Upper Floors	Roof	Stairs
Substructure	Windows and External Doors	Internal Doors
Frame	Internal Walls and Partitions	Fittings, Furnishing and Equipment
External Walls	Wall Finishes	
Services	Floor Finishes	
	Ceiling Finishes	
Roof	Upper Floors	
Windows and External Doors	Stairs	
	Internal Doors	
	Fittings, Furnishing and Equipment	

CARBON HOTSPOTS

COST HOTSPOTS

6.3. Development of the Design Quality Indices

6.3.1. Finishes Quality Index

Finishes quality index was developed from a Delphi-based expert forum as discussed in Section 5.8 in the Data Collection chapter. A conceptual finishes quality index was developed by surveying the types of internal finishes applied in office buildings and classifying the commonly used finishes type under three quality levels (Basic, Moderate and Luxury). The developed conceptual finishes quality index was then verified by receiving the inputs from the experts.

Most of the comments were to add or remove a certain type of finishes. Under additions, floor painting was suggested to be considered as a Basic type of floor finish; fair face masonry as Basic wall finish; moisture resistant painting as a Moderate type of wall finish; heavily embossed wallpapers as Luxury wall finish; and moisture resistant ceilings with high sound proofing and timber boarded ceilings as Luxury ceiling finish. Under deletions, lime plaster was suggested to be removed from the list as lime plaster is generally used in historic properties while not very common in modern offices. Hardboard and chipboards were not considered suitable as wall finishes and hence, suggested to be removed. However, they were not removed from the list as some of the projects use hardboard and chipboards as finishes. Similarly, screeds were also suggested to be removed from the list, as it was not considered suitable as a finish but a build up for an applied finish. Further, it was suggested that screeds are applied to structural concrete to make a surface for tiling, carpeting, sheeting, etc. However, NRM classifies screeds as finishes and screeds form part of the finish when used as a build up for an applied finish on top. For these reasons, screeds were not removed from the list; vinyl paper was considered not suitable for ceiling, so it was removed from the list of ceiling finishes.

On the other hand, there were some comments on the quality levels of the proposed finishes. The most controversial quality level was of ceramic tiles as the experts indicated that ceramic tiles could be found in all categories ranging from Basic to Luxury at various prices, as it is more dependent on the manufacturer, size, grout used and the like. Subsequently, ceramic tiles were classed under each quality level as Basic, Moderate and Luxury ceramic tiles. Similarly, porcelain tiles were classed under Moderate and Luxury and vinyl sheet, vinyl tiles, carpet tiles were classed under Basic and Moderate followed by the comments of the experts. Marble was identified as another problematic finish as there are Chinese marble and European marble and some projects opt for Chinese marble due to lower cost. Therefore, it is clarified in the finishes quality index by classifying Chinese marble under Moderate floor finish and European Marble as a Luxury floor finish.

Furthermore, there were contradicting comments from one expert on the quality level of finishes including lightweight plaster, plasterboard, mastic asphalt floor, slate tiles, terrazzo floor and veneer laminated floor. This was then highlighted in the second round of verification to other experts and opportunity was given to vary their judgment on the quality levels of the above-mentioned types of finishes. Three of the experts confirmed that slate could also be considered as Luxury finish and one suggested terrazzo can be considered as Luxury in its high-value ranges while the judgement of the experts remained unchanged for the rest of the controversial finishes. Subsequently, slate was moved from Moderate to Luxury and terrazzo was left under Moderate as it was stated that terrazzo could not be classified under Luxury when marble, parquet floor and the like are identified as Luxury. Finally, the verified finishes quality index is presented in Table 6.11.

Table 6.11: Finishes quality index

Element	Quality Index	Main Category	Sub Category
Wall Finishes	**Basic**	Fair face finish	Paint to fair face
		Plaster/render	Cement sand plaster
		Paint	Emulsion/eggshell
		Wallpaper	Lining paper
			Vinyl paper
		Wall tiles	Basic ceramic tiles
	Moderate	Plaster	Thistle plaster
			Carlite plaster
		Paint	Moisture resistant paint
		Board linings	Plasterboard
			Plywood wall panels and treatment
			Wallboards
			Softwood boarding and treatment
			Hardboard
			Chipboard
			Veneered MD panels
		Wall tiles	Moderate ceramic tiles
			Moderate porcelain tiles
		Stone cladding	Chinese marble
	Luxury	Wall tiles	Mosaic tiles
			Luxury ceramic tiles
			Luxury porcelain tiles

Element	Quality Index	Main Category	Sub Category
		Wallpaper	Heavily embossed wallpapers
		Stone cladding	Natural granite
			European marble
		Other claddings	Composite aluminium
			Glass
Floor Finishes	**Basic**	Exposed concrete	Concrete hardener
		Floor paint	Regular floor paint
		Screeds	Cement sand
			Latex screed
		Protective finish	Mastic asphalt floor
		Flexible thin sheets and tiles	Linoleum sheet
			Linoleum tiles
			Basic vinyl sheet
			Basic vinyl tiles
			Basic carpet tiles
		Rigid Tiles	Cement tiles
			Basic ceramic tiles
		Carpet	Medium duty carpet
	Moderate	Screeds	Granolithic
		Resin based finish	Epoxy floor
		Flexible thin sheets and tiles	Rubber floor tiles
			Marmoleum
			Moderate vinyl sheet
			Moderate vinyl tiles
			Cork tiles
			Moderate carpet tiles
		Rigid tiles	Moderate ceramic tiles
			Moderate porcelain tiles
			Clay tiles

Element	Quality Index	Main Category	Sub Category
			Quarry tiles
		Carpet	Heavy duty carpet
		Stone finish	Terrazzo
			Chinese marble
		Access floor	Metal access floors
		Timber floor	Veneered laminated floor
			Redwood floor
	Luxury	Rigid tiles	Mosaic tiles
			Slate tiles
			Luxury ceramic tiles
			Luxury porcelain tiles
		Timber floor	Woodblock floor (Oak etc.)
			Woodstrip floor (Oak etc.)
			Parquet floor
		Stone finish	Natural granite
			European marble
Ceiling Finishes	**Basic**	In-situ finish	Sealer
			Skim coat
			Cement sand plaster
		Paint finish	Emulsion/eggshell
		Paper finish	Lining Paper
	Moderate	In-situ finish	Thistle plaster
			Carlite plaster
		Paint finish	Moisture resistant paint
		Board finish	Plasterboard
		Suspended ceiling systems	Metal frame plasterboard ceilings
			Plasterboard acoustic ceilings

Element	Quality Index	Main Category	Sub Category
			Moisture resistant ceilings
			Metal suspended ceilings
	Luxury	Timber ceiling	Timber boarded ceilings
		Suspended ceiling systems	Moisture resistant ceilings with high sound proofing
		Integrated/composite ceiling	Coffered ceilings

It should also be noted that sometimes materials might be imported from other countries due to lower cost while EC in this case, will be higher than the locally sourced materials. However, if the system boundary is cradle-to-gate then it will not be highlighted in EC values. In conclusion, it was identified that it is difficult to categorise many finishes under a particular quality level as the choices spans from Basic to Luxury. Nevertheless, an objective finishes quality index had to be adopted in the study to assess the finishes quality of the sample building in a consistent way, which is satisfied by the developed finishes quality index. The verified finishes quality classification system was used as a benchmark to assess the finishes quality of the buildings in Dataset 1 and Dataset 3.

Table 6.12 illustrates the method followed in determining the finishes quality of the building considered. Firstly, the percentage of each type of floor finish used in the building was calculated. Then, each type of floor finish was classified based on the finishes quality index developed for the study. Thirdly, the cumulative percentage of Basic, Moderate and Luxury finishes were derived. Finally, the weighted quality index was calculated for the building and the final value was rounded off to arrive at the finishes quality index of the building. In this way, wall, floor, ceiling finish indices were calculated for each of the sample buildings.

Table 6.12: Method of determining finishes index of a building

Building A

Floor Finish	Quantity	Finishes Index
Porcelain tiles	20%	Luxury
Carpet – heavy duty	30%	Moderate
Cement sand screed	10%	Basic
Access floor	30%	Moderate
Clay tiles	5%	Moderate
Oak timber floor	5%	Luxury

Floor Finish Category	Quantity	Floor Finishes Index	Weighted Index
Basic	10%	1	0.10
Moderate	65%	2	1.30
Luxury	25%	3	0.75
Sum			2.15
Floor Finishes Index of the building	**2 (Moderate)**		

Then, wall, floor and ceiling finishes quality indices were combined into an overall finishes quality index of the building to avoid several predictor variables in the model. Consequently, a weighted average approach was used to calculate the overall finishes quality index of the building as shown in Table 6.13. Firstly, the weighted quantity of the wall, floor and ceiling finishes of the building were calculated as a percentage of the total area finished. Then, the derived wall, floor and ceiling indices were multiplied by the respective weighted quantity and summed up to find the overall finishes index of the building. In this way finishes index of the sample buildings were calculated.

Table 6.13: Calculating overall finishes index of the building

Internal Finish Category	% of the area finished	Finishes Index	Weighted Index
Wall	40%	1	0.40
Floor	30%	3	0.90
Ceiling	30%	2	0.60
Sum			1.90
Overall Finishes Index of the building			**2 (Moderate)**

6.3.2. Services Quality Index

Services specifications are less likely available during early design stages. Hence, UK price books were surveyed for a more practical way of classifying quality for approximate estimates. Accordingly, a three-tiered classification system was proposed in Spon's Mechanical and Electrical Services Price Book 2014 (Davis Langdon Consultancy, 2014) for owner occupied office buildings namely: non air-conditioned buildings, non-air-conditioned automated buildings and air-conditioned automated buildings, which was identified as the most appropriate classification of all (see, Section 5.9). however, one type was found to be missing in the above classification – air-conditioned non automated buildings. Consequently, based on the provision of Services (sub-elements of Services installed in the building) in buildings, a four-tiered quality classification system was proposed for the study as follows:

- Level 1 - Non air-conditioned buildings (Essential building services)
- Level 2 - Air-conditioned buildings (Level 1 + A/C)
- Level 3 - Non air-conditioned automated buildings (Level 1 + BMS)
- Level 4 - Air-conditioned automated buildings (Level 2 + BMS)

(Note: Essential building services in office buildings include sanitary appliance, water installations, disposal installations, space heating systems, ventilation

systems, electrical installations, fire and lighting protection, communication and security installations)

Furthermore, under each category buildings with and without lift were also present in the sample as the sample buildings were primarily low to medium rise. Hence, the proposed services index had to accommodate the difference of having lift in the building. Considering all of the above-mentioned points the services quality index proposed for the study is presented in Table 6.14.

Table 6.14: Services quality index

Services Quality Index	
Level 1 - Non air-conditioned buildings (Essential building services)	
1.1	Without lift
1.2	With lift
Level 2 - Air-conditioned buildings (Level 1 + A/C)	
2.1	Without lift
2.2	With lift
Level 3 - Non air-conditioned automated buildings (Level 1 + BMS)	
3.1	Without lift
3.2	With lift
Level 4 - Air-conditioned automated buildings (Level 2 + BMS)	
4.1	Without lift
4.2	With lift

6.4. Pre-Regression Analysis

A detailed analysis of the variables forming the model is a pre-requisite of a regression analysis. Therefore, this section outlines the variable selection, the description of each variable and the relationship between the dependent and the independent variables. The analysis of the individual variables is presented under the univariate analysis and the paired analysis is presented under the bivariate analysis.

6.4.1. Variable Selection

Regression analysis works based on the relationship between the dependent variable and the independent variables. Two regression models were developed to predict EC and CC for early stages of design. Hence, there were two dependent variables (because of the two models) in the study namely: EC and CC. However, each model was presented in two ways such as not normalising EC and CC for GIFA (EC and CC) and normalising EC and CC for GIFA (EC per GIFA and CC per GIFA). The reason for performing regression analysis with each dependent variable by normalising and not normalising for GIFA is that the best performing model can be selected from the two, as there was a significant difference in prediction performance was noted. In addition, some of the previous studies use CC as the dependent variable while others use CC per GIFA. It is more convincing to compare the performance of the two models and choose the better model rather than providing justification for the selection of one of the two forms of the dependent variable (EC or EC per GIFA) and not exploring the other. Therefore, both versions of the models were analysed herein.

Design variables affecting cost was studies and presented in the literature review (Chapter 3). Accordingly, previous studies regressed GIFA, building height, the number of storeys, circulation space, building quality, technology and other non-design related variables like project duration, liquidated damage, location and the like with construction cost (McGarrity, 1988, Kouskoulas and Koehn, 2005, Karanci, 2010, Phaobunjong, 2002). In addition to that, literature also suggests that plan shape, grouping of buildings and average storey height can have an influence on construction cost.

However, carbon and cost hotspot analyses were used as the basis for variable selection in the study. As explained in Section 6.2, the most influential design variables that need to be modelled as presented in Table 6.5 includes footprint area, GIFA, average height, total height, upper floor area, roof area, façade area, internal wall area, useable floor area/circulation space and internal finish area. Of which some of the variables are also represented by others, for instance, the summation of footprint and upper floor area gives GIFA; floor and ceiling finish areas are approximately equal to GIFA. Therefore, variables were shortlisted to a minimum number of variables because fitting more variables in the regression model might not be effective (Kim et al., 2004a) and the change in the dependent variable will less likely be explained by all the independent variables. Further, more variables also cause the problem of multicollinearity (the relationship between independent variables which should be eliminated in a regression model). In addition to that, the selected design variables should also meet the requirement of the availability during the early stages of design.

Finally, the variables selected for the study includes GIFA, the number of storeys, average storey height (or building height), façade area, wall to floor ratio, circulation ratio, the number of basements, finishes quality and services quality. Internal wall area was not selected as an independent variable as it is less likely to be available during the early stages of design. Further, an objective and consistent way of incorporating finishes and services quality of the building was enabled by the use of finishes and services quality, which is specific to the study.

6.4.2. Univariate Analysis of Variables

It is important that data be examined before the analysis to identify any outliers or extremes in the dataset and check for normality of the dependent variables. Histograms and boxplots together with descriptive statistics such as minimum, maximum, mean and skewness are used to examine each variable separately. Histogram presents how the data are distributed and gives a visual indication of the distribution of the data, however, when the sample size is small, histogram may not represent the normality clearly, hence, boxplots were used to complement histograms to understand the distribution of each variable and to identify outliers and extremes in the dataset. The summary of the descriptive statistics of the variables is presented in Table 6.15.

Table 6.15: Descriptive statistics of the variables

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
						Statistic	Std. Error
GIFA (m^2)	41	212	14652	3642.07	3329.495	1.535	.369
Storeys (No)	41	1	6	2.73	.923	1.379	.369
Average Height (m)	41	2.5	4.4	3.45	.461	.162	.369
Building Height (m)	41	2.8	25.2	9.50	3.828	1.756	.369
Façade (m^2)	41	148	6682	2261.32	1747.189	1.101	.369
Wall to Floor Ratio	41	.24	1.50	.71	.243	.926	.369
Circulation Ratio	33	.09	.46	.24	.092	.477	.409
CC (£1000s)	41	392	17928	4915.15	4870.790	1.439	.369
EC (tCO_2)	41	177	9383	2469.75	2311.514	1.512	.369
CC per GIFA (£/m^2)	41	698	2285	1301.13	324.321	1.301	.369
EC per GIFA ($kgCO_2/m^2$)	41	551	916	680.44	95.581	.696	.369

As can be seen from Table 6.15, data values for the selected variables are present for all 41 buildings except for circulation space. Circulation space had 8 missing data points. Further, one of the basic assumptions in performing regression analysis is that the variables are normally distributed. The measurements for skewness and kurtosis give an indication of the normality of the data distribution of the variables. Skewness of a normally distributed variable will have a value of 0. Miles and Shevlin (2001) suggest that there is little problem if the skewness statistics is less than 1.0 and skewness statistics between 1.0 and 2.0 is also cautiously acceptable attributing to the fact that it might have an impact on the estimates. However, skewness statistics above 2.0 indicates a serious problem with normality. According to the skewness statistics presented in Table 6.15, average height, wall to floor ratio, circulation ratio and EC per GIFA are less than 1.0 and the rest lies between 1.0 and 2.0, which ensures no major violation of the assumption of normality occurs.

a) Gross Internal Floor Area

Figure 6.3 presents the distribution of the sample data points of the variable GIFA, which ranges from 212m^2 to 14,652m^2. Mean of GIFA of the sample is 3,642m^2 and skewness is 1.535 indicating a positive skew where more than half of the buildings have a GIFA between 0 and 3,000m^2. This is because the scope of the study covers only low to medium-rise office buildings. Even though the skewness statistics is less than 2.0, the skewness of the distribution looks prominent. Boxplot indicates that 4 data points fall out of the normal curve, of which 3 data points were identified as outliers and 1 as extreme.

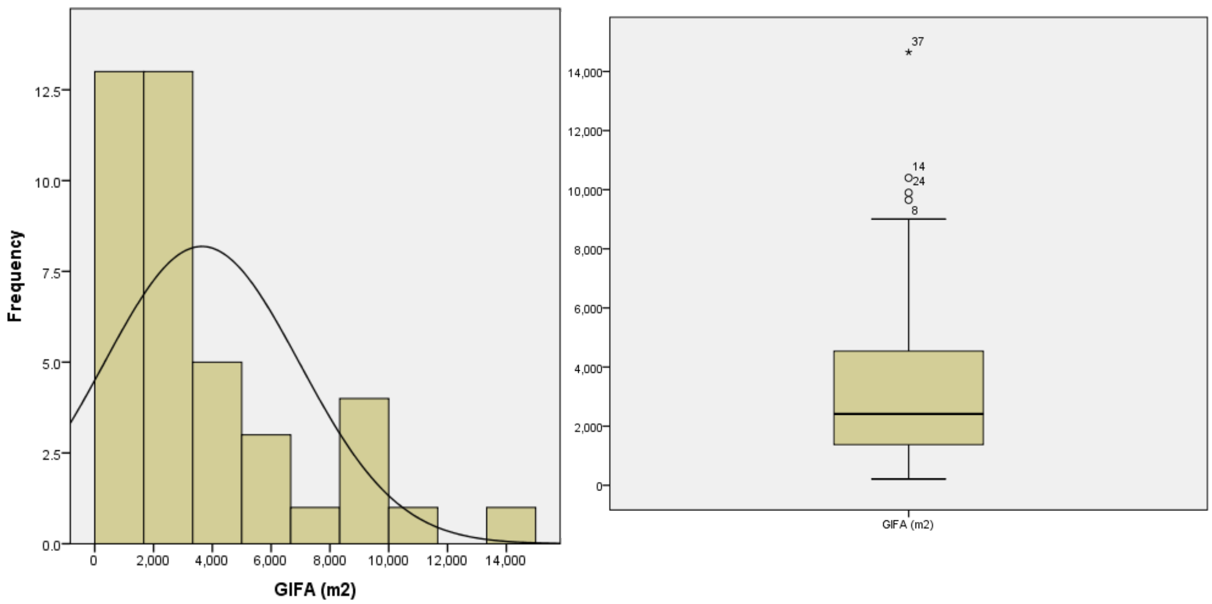


Figure 6.3: Histogram and boxplot for GIFA of the sample buildings

b) Number of Storeys/Building Height

Distribution of data points of the number of storeys is presented in Figure 6.4. The number of storeys ranges from one (1) to six (6) with a mean of three (3) and a positive skew of 1.379 was found. The majority of the buildings in the sample were 2 to 3 storied buildings. Boxplot identified two data points as outliers, the two buildings with five (5) and six (6) storeys while these buildings were not identified as outliers or extremes with regards to GIFA. However, they are true data points.

Figure 6.5 presents data distribution of building height of the sample ranges from 2.8m to 25.2m, with a mean of 9.50 and a positive skew of 1.756. Analysis of building height suggests that the building with six (6) storeys is an outlier while five (5) storeyed building was also found as an outlier in the analysis of number of storeys.

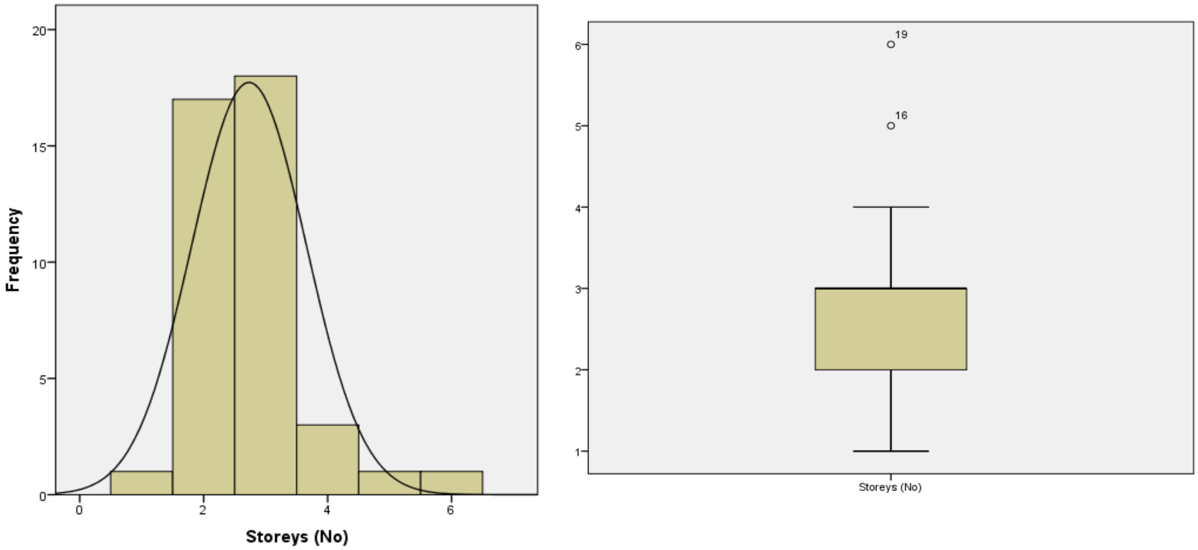


Figure 6.4: Histogram and boxplot for number of storeys in the sample buildings

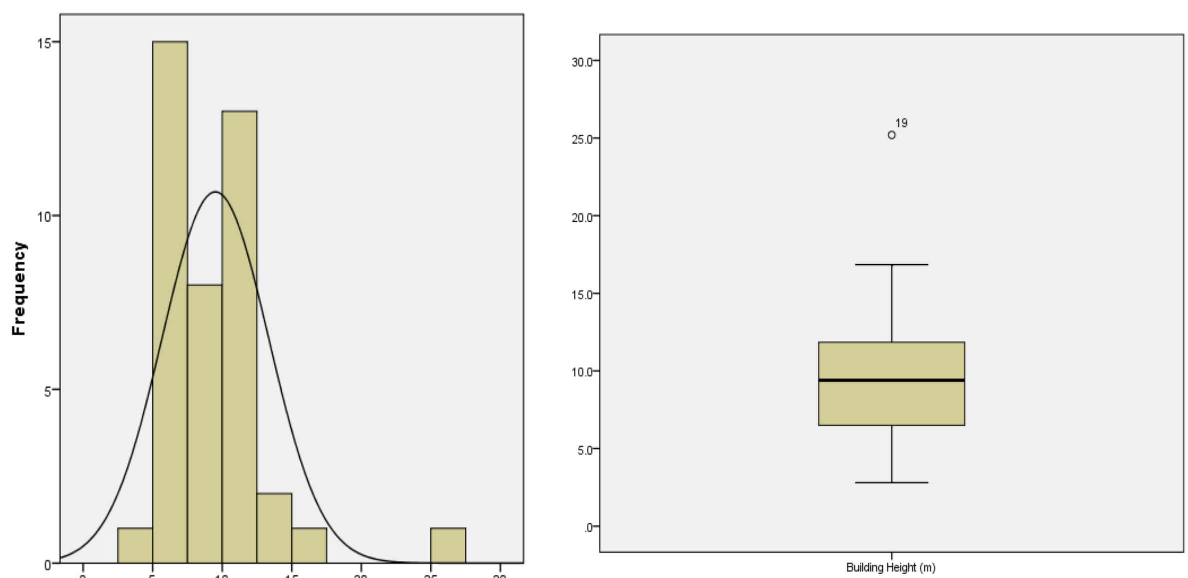


Figure 6.5: Histogram and boxplot for building height of the sample buildings

c) Average Storey Height

Figure 6.6 illustrates the data distribution of the average storey height of the sample buildings, which ranges from 2.5m to 4.4m, with a mean storey height of 3.451m displaying almost perfect normality. Further, boxplot also shows that there are no outliers in the dataset for average storey height.

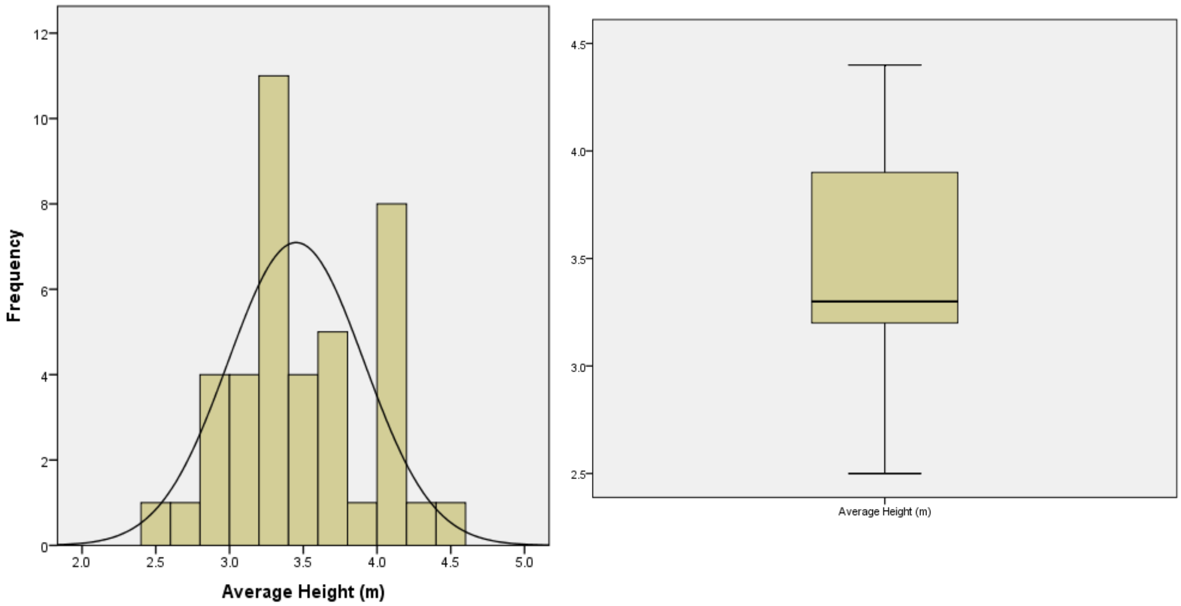


Figure 6.6: Histogram and boxplot for average storey height of the sample buildings

d) Wall to floor ratio

Distribution of wall to floor ratio is presented in Figure 6.7. The Wall to Floor ratio of the sample ranges from 0.24 to 1.50, with a mean value of 0.71 and the skewness is 0.926. Half of the sample buildings have a wall to floor ratio between 0.35 and 0.75. One data point was identified as an outlier, which was identified as an outlier in terms of the number of storeys (5 storeys) too.

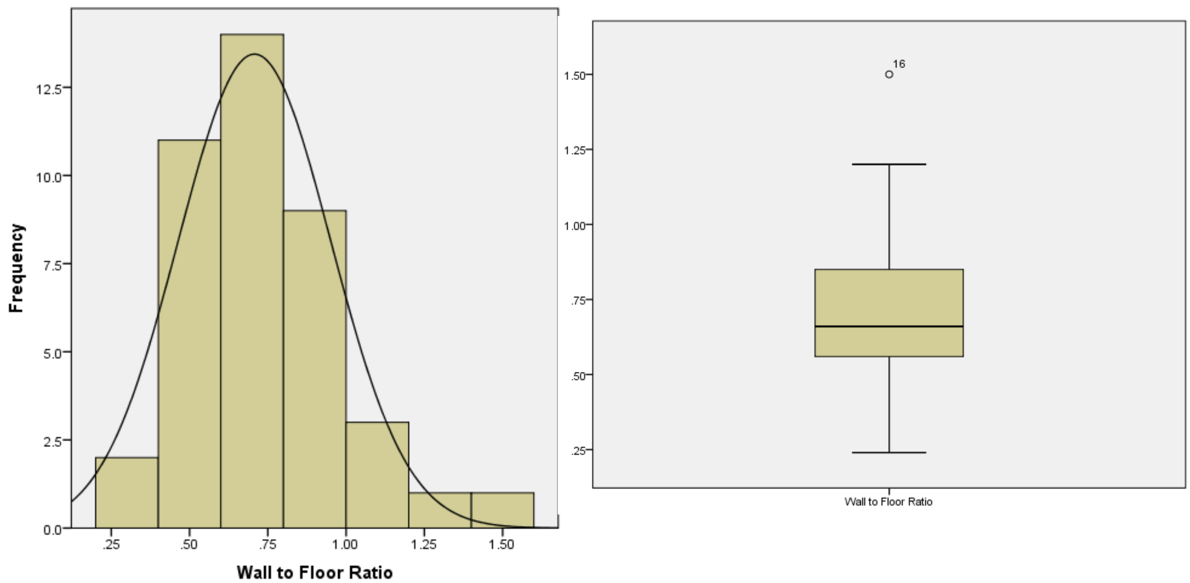


Figure 6.7: Histogram and boxplot for wall to floor ratio of the sample buildings

e) Façade Area

Figure 6.8 illustrates the data distribution of façade area. Façade area ranges from 148m^2 to 6,682m^2 with a mean of 2,261m^2. However, façade area has a positive skew of 1.101 and displays a very similar distribution to GIFA as the façade area is affected by GIFA, plan shape and average storey height. Therefore, two of the outliers identified here are also identified as outliers in GIFA analysis.

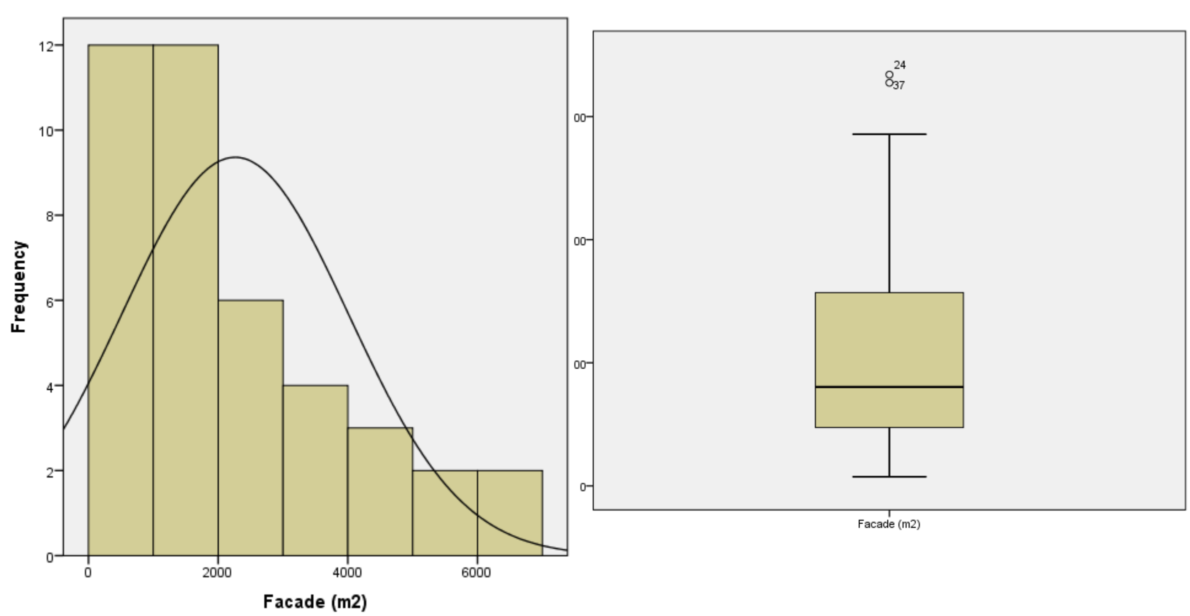


Figure 6.8: Histogram and boxplot for façade area of the sample buildings

f) Circulation Space

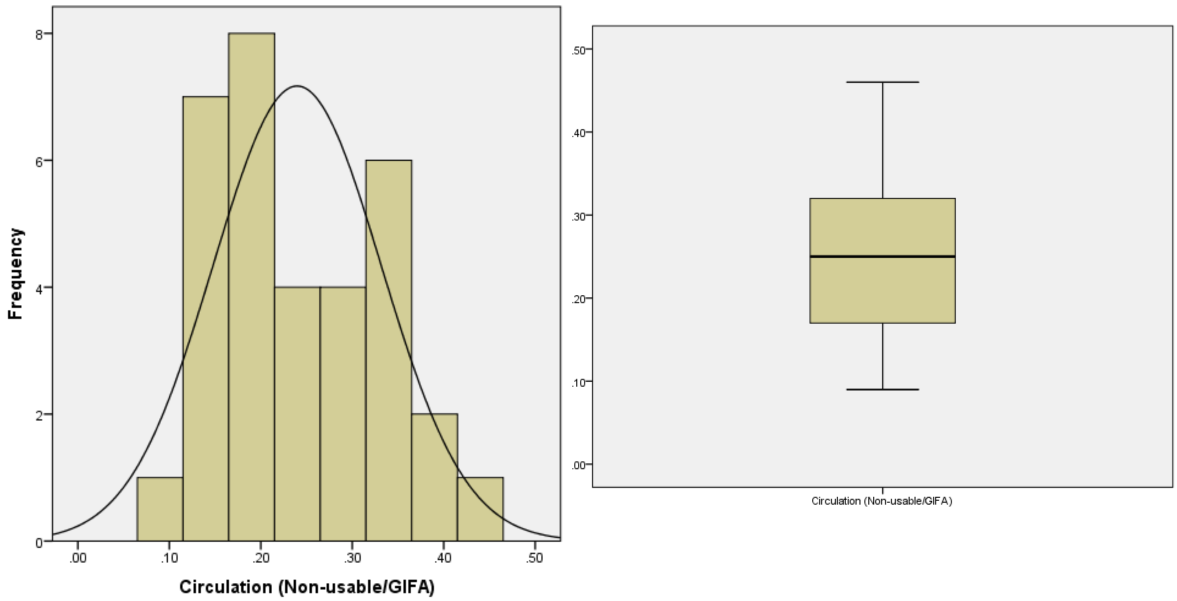


Figure 6.9: Histogram and boxplot for circulation ratio of the sample buildings

Circulation space in the sample ranges from 0.09 to 0.46 with a mean of 0.24. No outliers were identified in the sample and the variable demonstrated perfect normality (See, Figure 6.9).

g) Number of Basements

The number of basements is a discrete variable that contains only whole numbers. Hence, a bar chart is used to illustrate the number of basements in the sample buildings, which is depicted in Figure 6.10. Accordingly, the number of basements ranges from 0 to 2, where most of the buildings do not have basements.

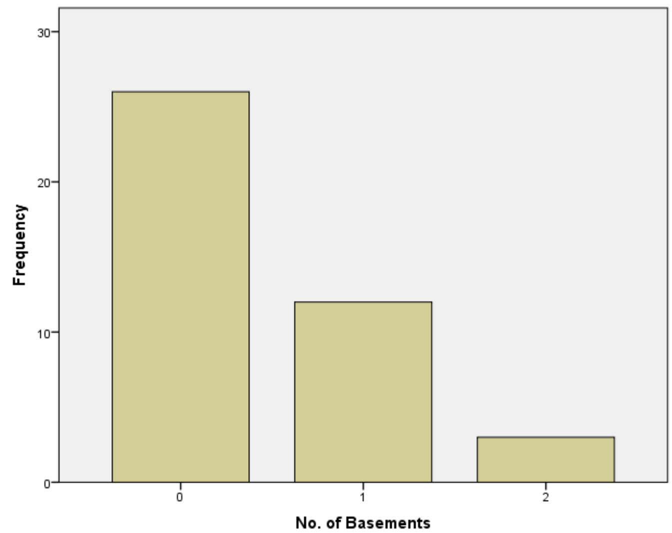


Figure 6.10: Bar chart for no. of basements in the sample buildings

h) Finishes index

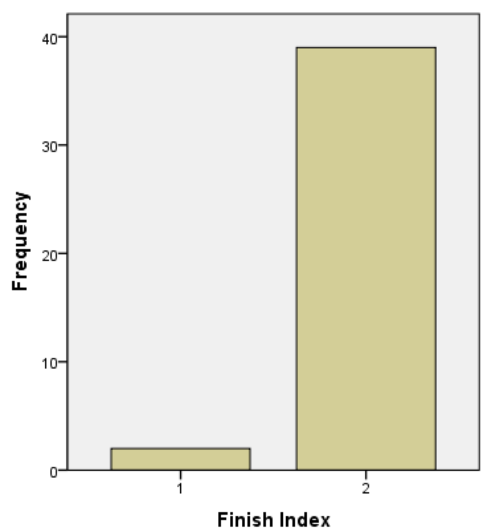


Figure 6.11: Bar chart for finishes index of the buildings in the sample

Figure 6.11 presents finishes index of the buildings in the sample. Finishes index of the sample buildings ranges from 1 (Basic) to 2 (Moderate). Predominantly, the finishes quality of the building in the sample was to be 2 (Moderate) while only two buildings had Basic level of finishes quality. No building had a finishes quality index of 3 (Luxury).

i) Services index

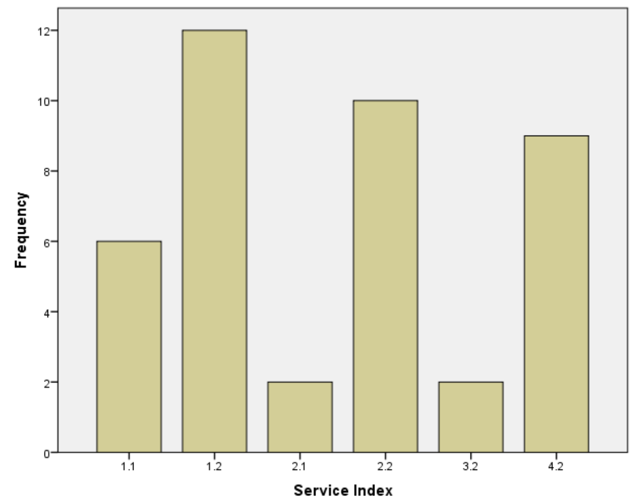


Figure 6.12: Bar chart for services index of the buildings in the sample

Services index of the sample buildings is presented in Figure 6.12. Services index has eight categories including: non A/C (without lift & with lift), A/C (without lift & with lift), non A/C automated (without lift & with lift) and A/C automated (without lift & with lift). Most buildings are non A/C with lift (services index of 1.2) and no buildings in the sample has a services index of 4.1 (A/C automated without lift).

j) EC

EC in the sample ranges from 177 tCO_2 to 9,383 tCO_2, with a mean of 2,470 tCO_2. EC of the sample demonstrates a similar distribution like GIFA with a positive skew of 1.5112 and with the same three outliers identified in GIFA (See, Figure 6.13). This indicates a close relationship between GIFA and total carbon.

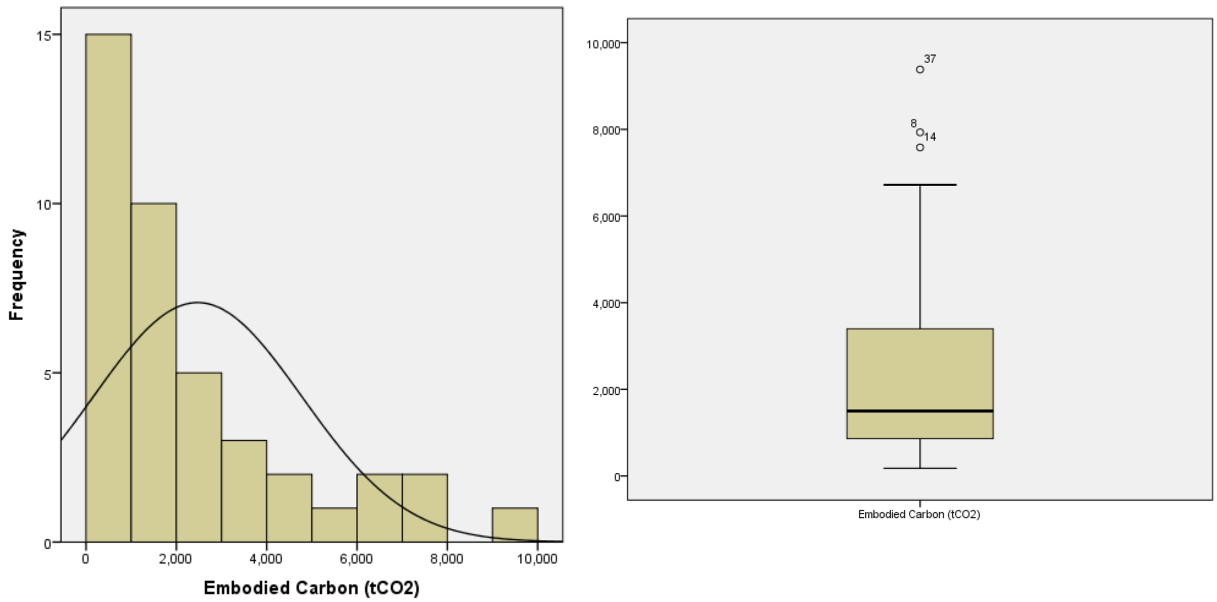


Figure 6.13: Histogram and boxplot for EC of the sample buildings

k) EC per GIFA

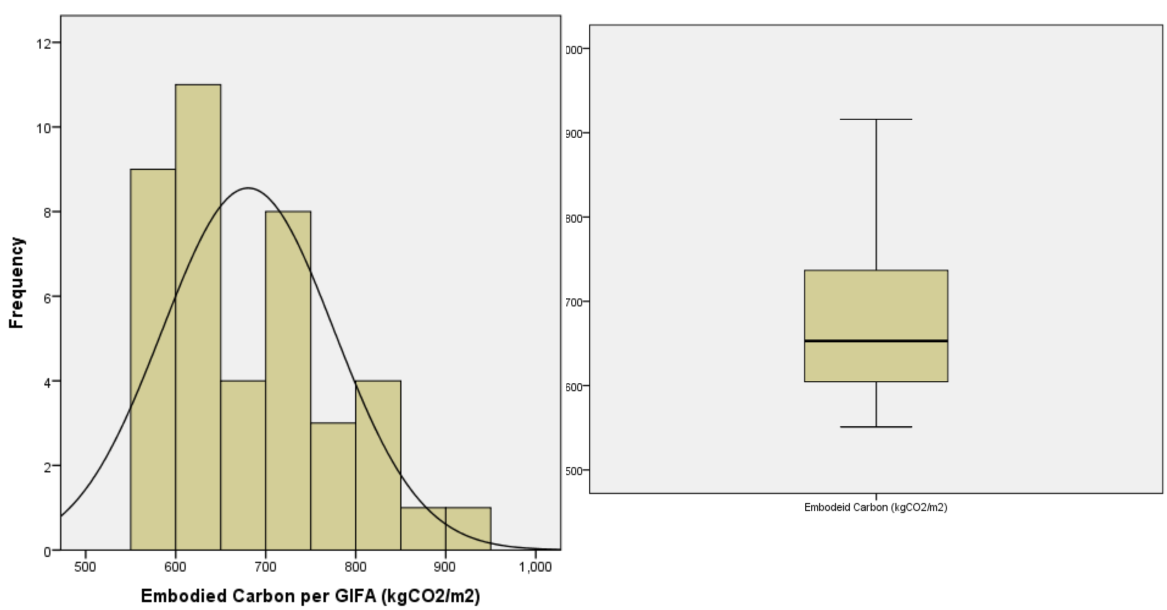


Figure 6.14: Histogram and boxplot for EC per GIFA of the sample buildings

Figure 6.14 represents the data distribution of EC per GIFA of the sample, ranges from 551kgCO_2/m^2 to 916kgCO_2/m^2 with a mean value of 680kgCO_2/m^2. No outliers were identified in the boxplots, which showcase normality.

l) CC

CC in the sample range from £392,000 to £17,918,000 with a mean of £4,915,150. CC of the sample also demonstrates a similar distribution like GIFA with a positive skew of 1.439. Three data points were identified as outliers (see, Figure 6.15) and all these outliers are outliers in GIFA. Of which, two outliers are also identified as outliers in EC. This indicates a close relationship between GIFA, EC and CC.

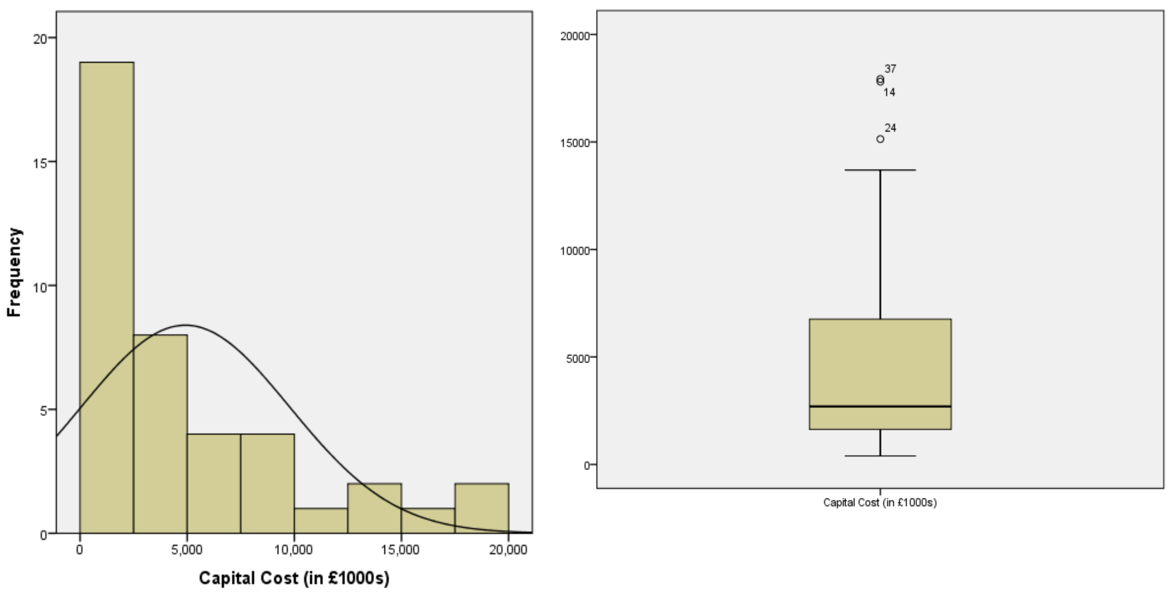


Figure 6.15: Histogram and boxplot for CC of the sample buildings

m) CC per GIFA

Figure 6.16 presents the data distribution of CC per GIFA of the sample. CC per GIFA in the sample ranges from £698 to £2,285 with a mean of £1,301 and a positive skew of 1.301. Three data points were identified as outliers in the data sample.

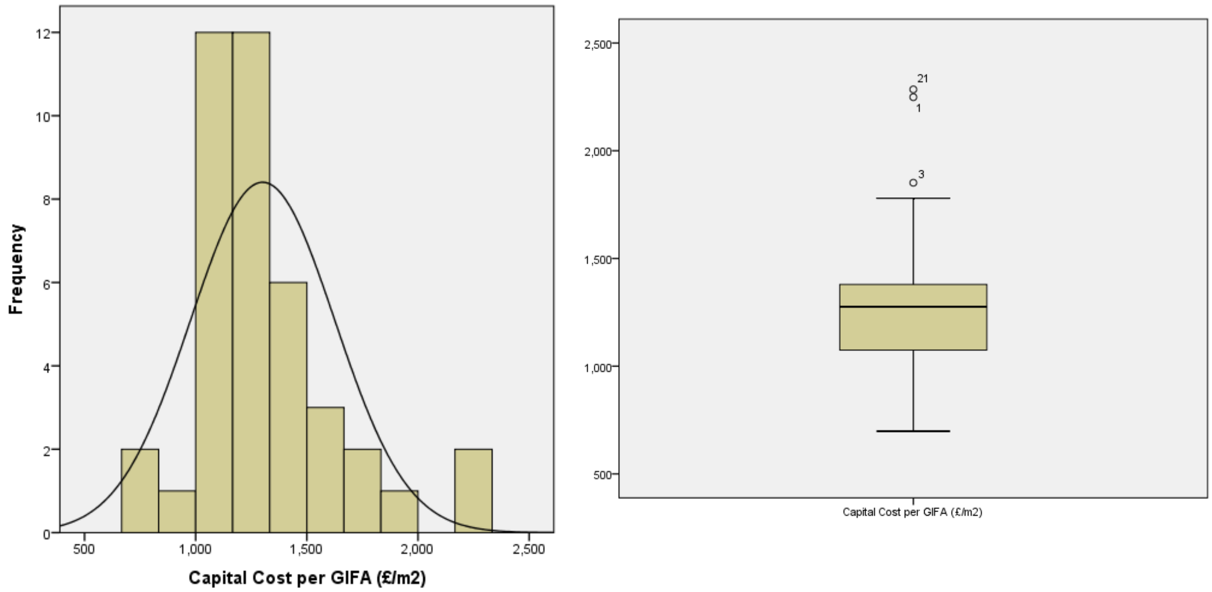


Figure 6.16: Histogram and boxplot for CC per GIFA of the sample buildings

6.4.3. Bivariate Analysis of Variables

Scatterplot matrix was produced to understand the relationships between variables dependent and independent variables before performing regression analysis. Two scatterplot matrices were produced for EC and CC separately.

1. First scatterplot was intended to depict relationship between EC (and CC) and design variables such as GIFA, building height (or number of storeys, average storey height), façade area and circulation ratio – reason for choosing façade area for this model rather than wall to floor ratio is because no clear relationship can be captured between wall to floor ratio and EC or CC as wall to floor ratio simply means how much façade area is required to cover 1m^2 of GIFA and façade area is determined by GIFA. Further, buildings with higher wall to floor ratio might be less expensive and less carbon embodied in it because of lower GIFA compared to building with larger GIFA and lower wall to floor ratio.
2. The second scatterplot was intended to present the relationship between EC per GIFA (and CC per GIFA) and design variables such as building height (or the number of storeys, average storey height), wall to floor ratio and circulation ratio. Figure 6.17 and Table 6.16 presents scatterplot matrix and correlation statistics between selected design variables, EC and CC. Accordingly, GIFA

and façade area demonstrates a strong positive correlation to EC with few data points scattered with a correlation coefficient of 0.985 and 0.862 respectively (correlation is statistically significant at the 0.01 level). However, the relationship between building height and EC is not strong like GIFA and façade area, which has a correlation coefficient of 0.513 (significant at the 0.01). On the other hand, the correlation between circulation space ratio and EC did not yield a statistically significant result. Interestingly, similar behaviour was demonstrated by CC. CC also showed a strong positive correlation with GIFA and façade area with a correlation coefficient of 0.969 and 0.868 (significant at the 0.01 level) while building height is moderately correlated with EC (0.535 - significant at the 0.01 level). Further, circulation ratio did not show a statistically significant relationship with CC same as EC. The relationship between CC and EC is very strong with a positive correlation coefficient of 0.977(significant at the 0.01) explains the similar behaviour of EC and CC. Further, matrices indicate that GIFA and faced are the most influential design variables of EC and CC. However, this conclusion is obvious as bigger buildings cost more and have more carbon embodied in them due to more material, labour and plant inputs. Therefore, it was decided to normalise GIFA and repeat the bivariate analysis.

Further, scatterplot also assists in discovering any collinearity between independent variables. Accordingly, the pair of GIFA and façade area was detected with collinearity (> 0.7) with a correlation coefficient of 0.861 (significant at the 0.01). This is not surprising and is logical as larger building implies higher façade area. In addition to that, the pair of façade area and building height was also demonstrated a positive correlation of 0.670 (significant at the 0.01). This is because façade area is calculated by multiplying building height by the girth of the building; hence, a positive relationship can be expected. However, the correlation between façade area and building height was not of much concern since the correlation was within the collinearity threshold. Collinearity can cause problems in the model building process if both variables are modelled together. Hence, it was decided to exclude façade area in the models.

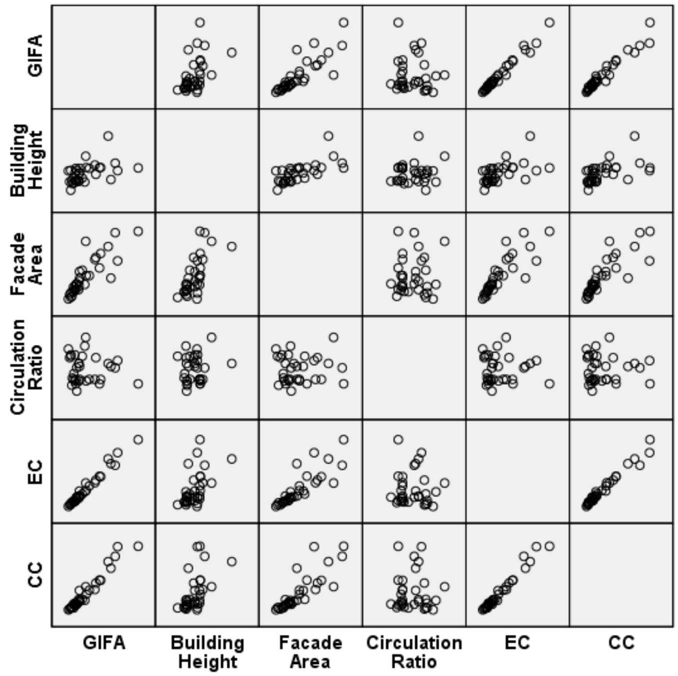


Figure 6.17: Scatterplot matrix of design variables, EC and CC

Table 6.16: Correlations matrix of design variables, EC and CC

		GIFA	Building Height	Façade Area	Circulation Ratio	EC	CC
GIFA	Pearson Correlation	1	.476**	.861**	-.095	.985**	.969**
	Sig. (2-tailed)		.002	.000	.599	.000	.000
	N	41	41	41	33	41	41
Building Height	Pearson Correlation	.476**	1	.670**	.113	.513**	.535**
	Sig. (2-tailed)	.002		.000	.531	.001	.000
	N	41	41	41	33	41	41
Façade Area	Pearson Correlation	.861**	.670**	1	.039	.862**	.868**
	Sig. (2-tailed)	.000	.000		.829	.000	.000
	N	41	41	41	33	41	41
Circulation Ratio	Pearson Correlation	-.095	.113	.039	1	-.041	-.010
	Sig. (2-tailed)	.599	.531	.829		.821	.955
	N	33	33	33	33	33	33
EC	Pearson Correlation	.985**	.513**	.862**	-.041	1	.977**
	Sig. (2-tailed)	.000	.001	.000	.821		.000
	N	41	41	41	33	41	41
CC	Pearson Correlation	.969**	.535**	.868**	-.010	.977**	1
	Sig. (2-tailed)	.000	.000	.000	.955	.000	
	N	41	41	41	33	41	41

**. Correlation is significant at the 0.01 level (2-tailed).

Same scatterplot and correlation matrices were produced for EC per GIFA, CC per GIFA, building height, wall to floor ratio and circulation ratio presented in Figure 6.18 and. Table 6.17. A moderate positive linear relationship was evident between EC per GIFA and wall to floor ratio with a correlation coefficient of 0.523 (significance at the 0.01 level) and a weak positive linear relationship was found between EC per GIFA and circulation ratio with a correlation coefficient of 0.360 (significance at the 0.05 level). Building height did not indicate a statistically significant correlation with EC per GIFA and the relationship was almost neutral. Further, EC per GIFA at a particular building height vary a lot as shown in scatterplot which was surprising as it was expected that EC per GIFA would increase with building height as generally cost per GIFA is expected to increase with building height. On the other hand, CC per GIFA demonstrated weak positive correlations with all the variables - building height, wall to floor ratio and circulation ratio with a correlation coefficient of 0.389, 0.322, 0.391 (significant at the 0.05 level) respectively. However, a moderate linear positive relationship was found between EC per GIFA and CC per GIFA (correlation coefficient of 0.645 significant at the 0.01 level). Furthermore, no significant collinearity between variables was detected.

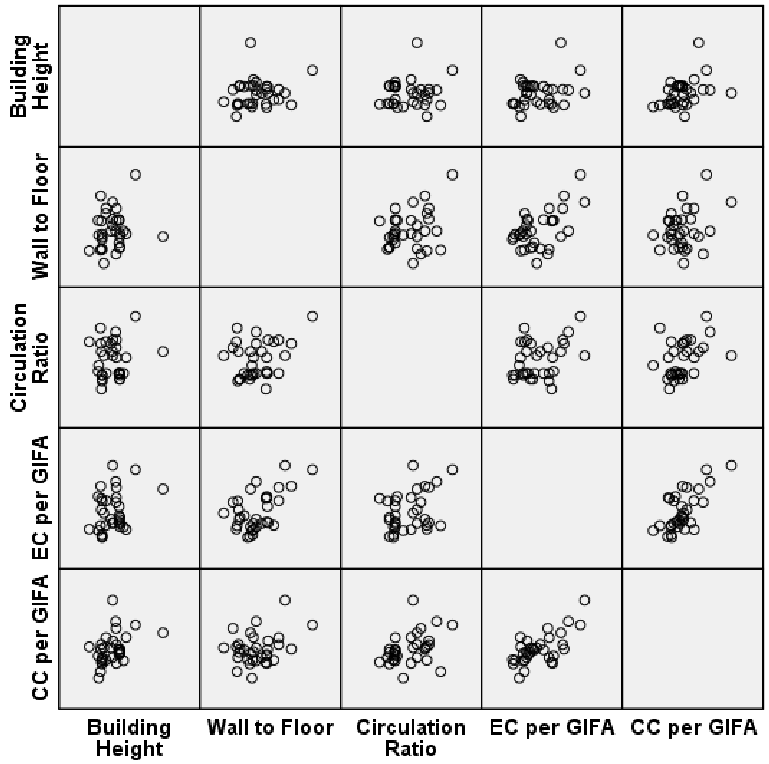


Figure 6.18: Scatterplot matrix of design variables, EC per GIFA and CC per GIFA

Table 6.17: Correlation matrix of design variables, EC per GIFA and CC per GIFA

		Building Height	Wall to Floor Ratio	Circulation Ratio	EC per GIFA	CC per GIFA
Building Height	Pearson Correlation	1	.206	.113	.306	.389*
	Sig. (2-tailed)		.195	.531	.052	.012
	N	41	41	33	41	41
Wall to Floor Ratio	Pearson Correlation	.206	1	.304	.523**	.322*
	Sig. (2-tailed)	.195		.086	.000	.040
	N	41	41	33	41	41
Circulation Ratio	Pearson Correlation	.113	.304	1	.360*	.391*
	Sig. (2-tailed)	.531	.086		.039	.024
	N	33	33	33	33	33
EC per GIFA	Pearson Correlation	.306	.523**	.360*	1	.645**
	Sig. (2-tailed)	.052	.000	.039		.000
	N	41	41	33	41	41
CC per GIFA	Pearson Correlation	.389*	.322*	.391*	.645**	1
	Sig. (2-tailed)	.012	.040	.024	.000	
	N	41	41	33	41	41

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

6.4.4. Outcome of the Pre-Regression Analysis

Univariate analysis helped to discover non-normality in data distribution of the variables where GIFA, façade area, EC and CC are found to be significantly skewed. Therefore, log transformation was applied to each variable to improve the normality of the data distribution through data transformation. Apparently, data transformation reduced skewness and helped to achieve normality in all four variables. The new statistics is presented in Table 6.18 and the boxplots of the variables after log transformation is illustrated in Figure 6.19. Consequently, only one data point was identified as an outlier after log transformations.

Table 6.18: Descriptive statistics of transformed variables

	N	Skewness	
	Statistic	Statistic	Std. Error
GIFA log	41	-.188	.369
Façade log	41	-.485	.369
EC log	41	.010	.369
CC log	41	.140	.369

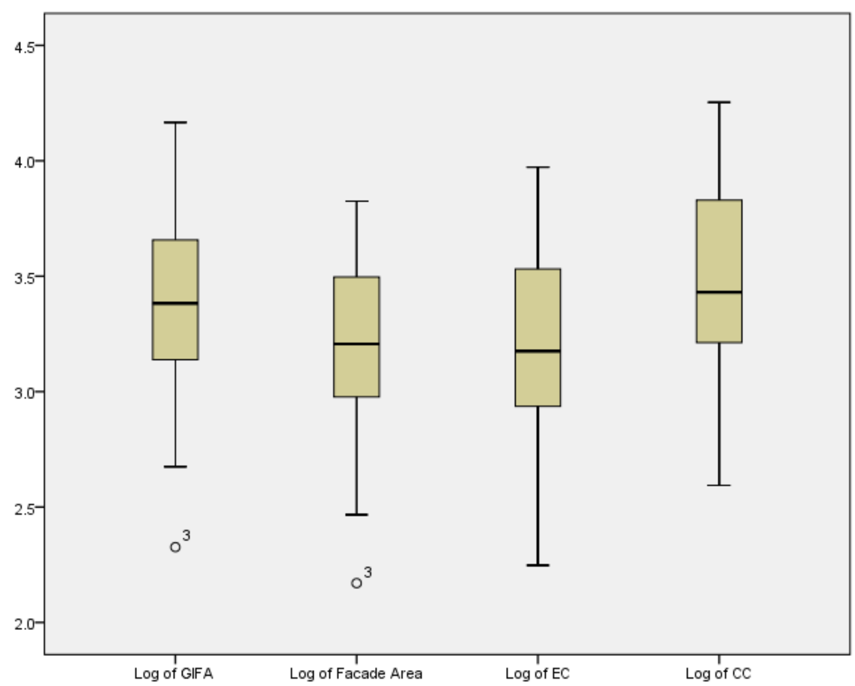


Figure 6.19: Box plots after the log transformation of the selected variables to achieve normality

Scatterplot and correlation matrices enabled to identify the relationships between variables and hence non-linear relationships were detected. As discussed in Section 7.4.3, collinearity exists between GIFA and façade area while no significant relationship was found between circulation space ratio and EC or CC. However, now that the EC and CC were transformed, correlation matrix was produced between circulation space and EC and CC, which is presented in Figure 6.20. Yet, no significant linear relationship was noticed. Then, log transformation was applied to the entire datum in the variable circulation space ratio and the correlation matrix was produced again which is presented in Figure 6.21. Log transformation of the variable (circulation space ratio) did not help to achieve linearity (the data points in Figure 6.21 are randomly scattered). Therefore, the inverse transformation was applied to the data in circulation space ratio to see whether linearity assumption could be met. Inverse transformation also did not yield expected results (see, Figure 6.22). Hence, it was decided to eliminate circulation space ratio as a predictor in EC and CC models.

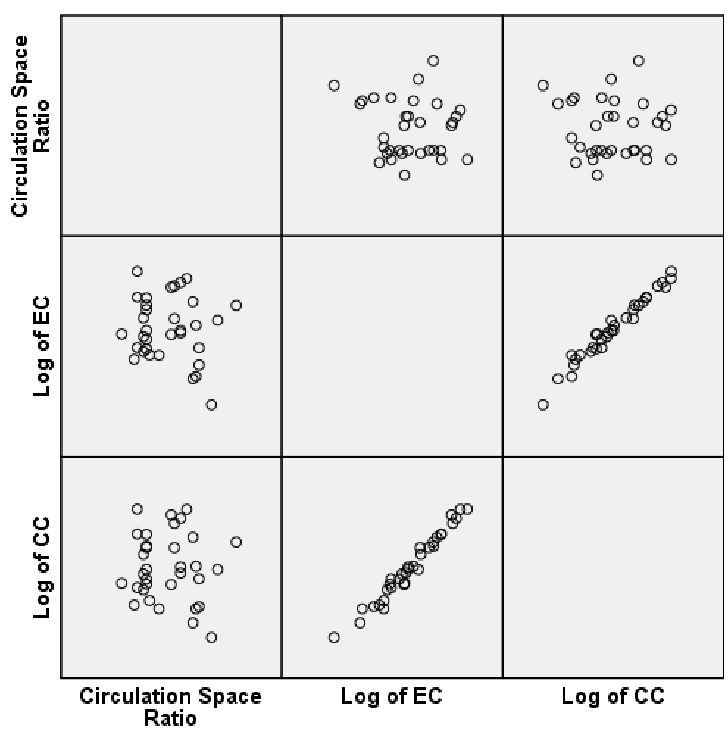


Figure 6.20: Correlation matrix between circulation space ratio, Log of EC and Log of CC

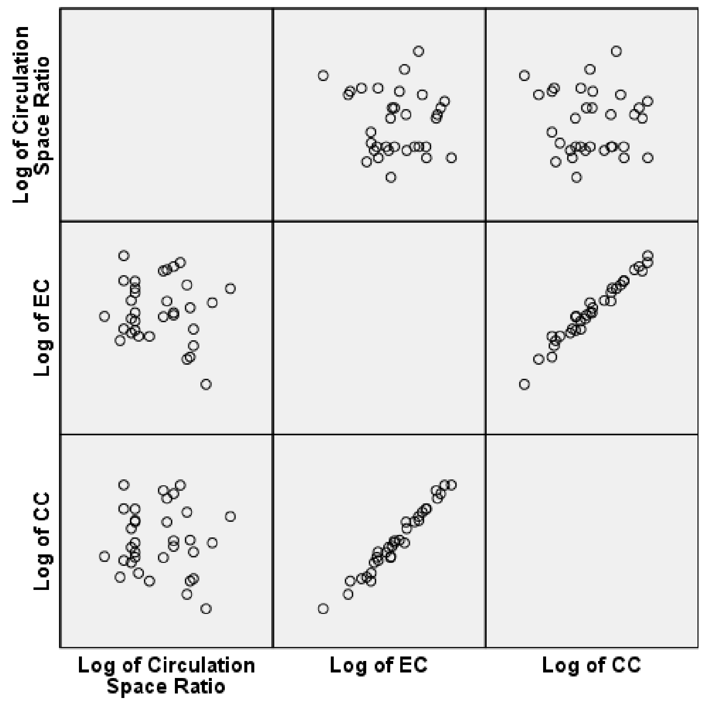


Figure 6.21: Correlation matrix between Log of circulation space ratio, Log of EC and Log of CC

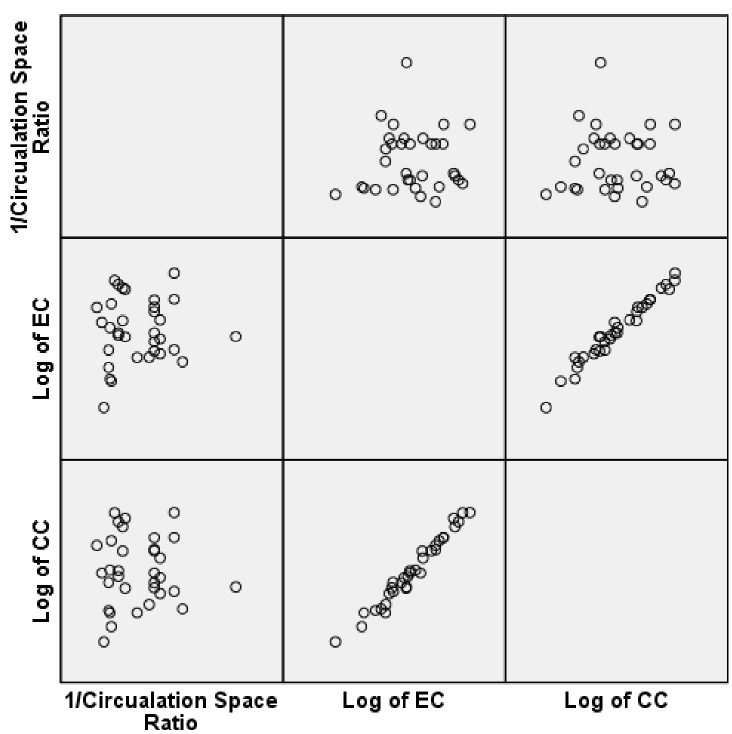


Figure 6.22: Correlation matrix between inverse of circulation space ratio, Log of EC and Log of CC

Further, regression analysis works on the basis that there is relationship between dependent variable and independent variables only and no significant relationship between independent variables should exist in order for the model to perform effectively in real time scenarios. Therefore, one of the variables should be retained and the other should be eliminated to avoid collinearity between independent variables. Collinearity was found only between GIFA and façade area. Therefore, the decision was made to use GIFA only in the EC and CC models as GIFA is more influential and had higher correlation coefficient than façade area.

6.5. Regression Analysis

Regression outputs can be sensitive to outliers. Altogether, 5 data points were found as outliers after log transformation. Subsequently, the effort was made to identify whether the outliers represent true data points or otherwise caused by measurement error, error in transferring data or error in the theory. Raw data were examined to ensure no errors were made during transferring and estimating. In terms of design economics theory, there is less concern as each building is unique and therefore, there are no standard values for a particular variable (though general practices and design norms are adopted by most of the designers – for instance, circulation space might range from 15-25%). Finally, it was concluded that all the outliers represent true data points. Outliers can be dealt in two ways as follows:

1. Modelling with outliers
2. Modelling without outliers

Modelling with outliers might have an influence on the estimate of correlations of the model while modelling without outliers will be unbiased though elimination of the outliers might give rise to other outliers requiring further elimination reducing the sample size. Therefore, the decision has to be made whether to remove outliers and formulate model with less number of data or to include the outliers and modelling data poorly. Pedhazur and Schmelkin (1991) suggested that results with and without outliers shall be reported to give the readers better understanding about the influence of the outliers in modelling. Therefore, the models with outliers and without the outliers were presented and the selection of the model was made based on the R^2, F statistics, standard error and significance statistics.

6.5.1. Regression Models for Embodied Carbon Prediction

Two models were considered to predict EC including EC per GIFA model and EC model. Conceptual models are presented below for EC per GIFA (see, Equation 6.1) and EC (see, Equation 6.2):

Equation 6.1: EC per GIFA conceptual model

$$\widehat{y_1} = a_0 + a_1 x_{BH} + a_2 x_{W:F} + a_3 x_{CR} + a_4 x_B + a_5 x_{FI} + a_6 x_{SI}$$

Where,

$\widehat{y_1}$ $- Estimated\ EC\ per\ GIFA\ of\ the\ building$
a_0 $- Regression\ constant$
a_1 $- Regression\ coefficient\ of\ x_{BH}$
x_{BH} $- Building\ Height$
a_2 $- Regression\ coefficient\ of\ x_{W:F}$
$x_{W:F}$ $- Wall\ to\ Floor\ ratio\ of\ the\ building$
a_3 $- Regression\ coefficient\ of\ x_{CA}$
x_{CR} $- Circulation\ space\ Ratio\ of\ the\ building$
a_4 $- Regression\ coefficient\ of\ x_B$
x_B $- Number\ of\ basements\ in\ the\ building$
a_5 $- Regression\ coefficient\ of\ x_{FI}$
x_{FI} $- Finishes\ Index\ of\ the\ building$
a_6 $- Regression\ coefficient\ of\ x_{SI}$
x_{SI} $- Services\ Index\ of\ the\ building$

Equation 6.2: EC conceptual model

$$\widehat{y_2} = b_0 + b_1 x_{GIFA} + b_2 x_{BH} + b_3 x_{FA} + b_4 x_{CR} + b_5 x_B + b_6 x_{FI} + b_7 x_{SI}$$

Where,

$\widehat{y_2}$ $-$ *Estimated EC of the building*
b_0 $-$ *Regression constant*
b_1 $-$ *Regression coefficient of* x_{GIFA}
x_{GIFA} $-$ *GIFA of the building*
b_2 $-$ *Regression coefficient of* x_{BH}
x_{BH} $-$ *Building Height*
b_3 $-$ *Regression coefficient of* x_{FA}
x_{FA} $-$ *Facade area of the building*
b_4 $-$ *Regression coefficient of* x_{CA}
x_{CR} $-$ *Circulation space Ratio of the building*
b_5 $-$ *Regression coefficient of* x_B
x_B $-$ *Number of basements in the building*
b_6 $-$ *Regression coefficient of* x_{FI}
x_{FI} $-$ *Finishes Index of the building*
b_7 $-$ *Regression coefficient of* x_{SI}
x_{SI} $-$ *Services Index of the building*

However, after the univariate and bivariate analysis, the EC model was modified to address non-normality and collinearity issues, using log values for EC and GIFA to conform to normality and eliminating the independent variable façade area to eliminate collinearity with GIFA. The modified equation is presented as follows:

Equation 6.3: Modified EC conceptual model

$$\hat{y}_2' = b_0 + b_1 x_{GIFA}' + b_2 x_{BH} + b_5 x_B + b_6 x_{FI} + b_7 x_{SI}$$

Where,

$\hat{y}_2'$ $-$ $\log \widehat{y_2}$
x_{GIFA}' $-$ $\log x_{GIFA}$

Consequently, regression analysis was run to identify best predictive EC per GIFA model using the backward method. This method accommodates all input variables in the first run and eventually removes one variable at a time – the variable that is found to be the least significant in the model. The least significant variable in the model is rejected where more than one variable found to meet the elimination criteria. In this way, variables are eliminated one by one until the best model is derived.

a) Regression models with outliers

The backward method produced the best predictive EC per GIFA model in the fifth step with two variables – wall to floor ratio and the number of basements. Model summary, analysis of variance and model coefficients resulting from each step are presented in Table 6.19, Table 6.20 and Table 6.21 respectively.

Table 6.19: Model summary – EC per GIFA Run 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	.747	.559	.457	72.011	Building height, wall to floor ratio, circulation ratio, no. of basements, finishes index, services index
2	.746	.557	.475	70.781	Wall to floor ratio, circulation ratio, no. of basements, finishes index, services index
3	.743	.552	.488	69.922	Wall to floor ratio, no. of basements, finishes index, services index
4	.736	.542	.495	69.456	Wall to floor ratio, no. of basements, finishes index
5	.717	.513	.481	70.386	Wall to floor ratio, no. of basements

Table 6.20: ANOVA table – EC per GIFA Run 1

	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	170638.783	6	28439.797	5.484	.001
	Residual	134824.924	26	5185.574		
	Total	305463.707	32			
2	Regression	170194.758	5	34038.952	6.794	.000
	Residual	135268.949	27	5009.961		
	Total	305463.707	32			
3	Regression	168567.844	4	42141.961	8.620	.000
	Residual	136895.863	28	4889.138		
	Total	305463.707	32			
4	Regression	165565.268	3	55188.423	11.440	.000
	Residual	139898.439	29	4824.084		
	Total	305463.707	32			
5	Regression	156836.038	2	78418.019	15.828	.000
	Residual	148627.669	30	4954.256		
	Total	305463.707	32			

R^2 indicates the percentage change in the dependent variable explained by the independent variables in the model. Model summary displays that no much improvement is achieved in adjusted R^2 when progressing from one step to the other and the standard error of estimate also shows little improvement. However, a drastic drop from R^2 to adjusted R^2 is clearly notable in the first four steps while the drop in less in the fifth model. 48.1% of the change in the dependent variable is explained by wall to floor ratio and number of basements in Model 5 while 48.8% and 49.5% of change is explained by services index and finishes index in Model 3 and Model 4, which is better than Model 5. However, finishes and services indices are found to be insignificant in the models (Sig. < 0.05). Therefore, Model 5 is considered the best predictive EC per GIFA model for the given sample. VIF of the variables in Model 5 is close to 1, which confirms no multicollinearity in the model.

Table 6.21: Coefficient of the variables – EC per GIFA Run 1

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	630.353	119.238		5.286	.000		
	Building Height	1.066	3.641	.045	.293	.772	.725	1.380
	Wall to Floor Ratio	144.233	52.570	.391	2.744	.011	.834	1.199
	Circulation Ratio	85.992	154.656	.081	.556	.583	.804	1.244
	Basements	66.591	23.180	.455	2.873	.008	.678	1.476
	Finish Index	-69.851	54.920	-.173	-1.272	.215	.915	1.093
	Service Index	8.323	10.930	.107	.761	.453	.864	1.157
2	(Constant)	629.572	117.173		5.373	.000		
	Wall to Floor Ratio	145.089	51.592	.394	2.812	.009	.837	1.195
	Circulation Ratio	86.618	152.000	.081	.570	.573	.804	1.244
	Basements	69.569	20.472	.475	3.398	.002	.839	1.191
	Finish Index	-65.985	52.396	-.164	-1.259	.219	.971	1.030
	Service Index	8.952	10.533	.115	.850	.403	.899	1.112
3	(Constant)	654.169	107.611		6.079	.000		
	Wall to Floor Ratio	152.976	49.098	.415	3.116	.004	.901	1.109
	Basements	72.309	19.658	.494	3.678	.001	.888	1.126
	Finish Index	-70.524	51.159	-.175	-1.379	.179	.994	1.006
	Service Index	8.065	10.291	.103	.784	.440	.919	1.088
4	(Constant)	665.044	106.000		6.274	.000		
	Wall to Floor Ratio	160.879	47.730	.437	3.371	.002	.941	1.063
	Basements	68.595	18.951	.468	3.620	.001	.943	1.060
	Finish Index	-68.249	50.736	-.169	-1.345	.189	.997	1.003
5	(Constant)	530.620	35.829		14.810	.000		
	Wall to Floor Ratio	164.079	48.310	.445	3.396	.002	.943	1.060
	Basements	68.147	19.202	.465	3.549	.001	.943	1.060

One of the assumptions in regression is that the residuals should be homoscedastic and not auto correlate. Figure 6.24 depicts the standardised residuals of the regression. Accordingly, histogram displays normality of residuals in of the regression, which is satisfactory. Scatterplot of standardised predicted values against residuals suggests that there is no significant pattern is noticeable and the residuals are randomly distributed. These diagrams approve the assumption of homoscedastic of residuals. The Durbin-Watson test statistics of the model was 1.879 which is greater than $d_{U,\alpha}$ ($d_{U,\alpha,}$ =1.60) indicating no positive autocorrelation among the residuals. Similarly, 4-d (4 – 1.879 = 2.121) is also greater than $d_{U,\alpha}$ confirms no negative autocorrelation.

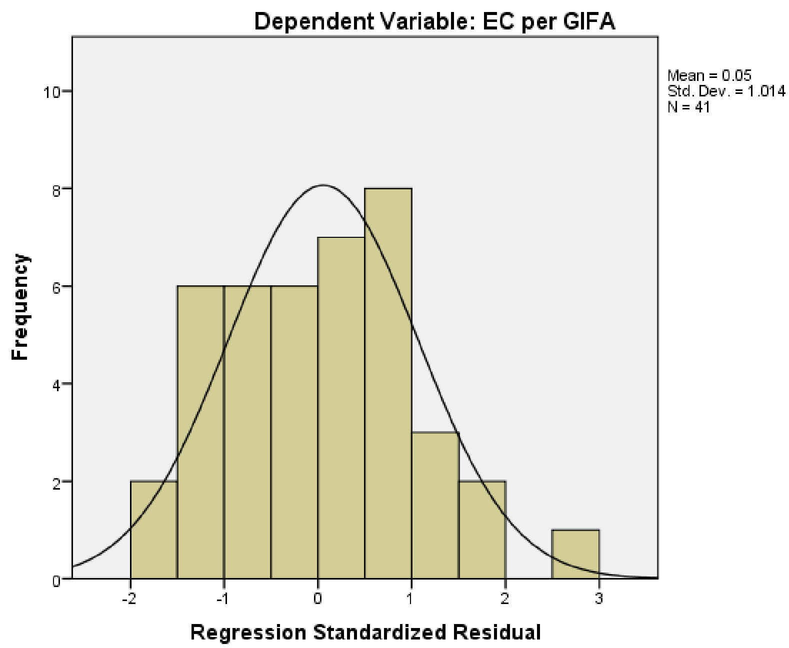


Figure 6.23: Histogram of standardised residual of the regression – EC per GIFA Run 1

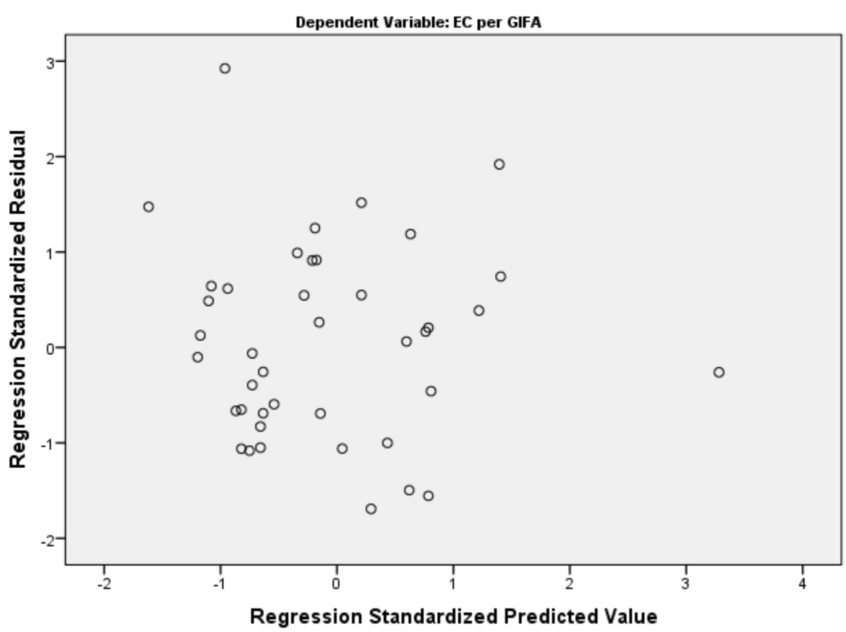


Figure 6.24: Scatterplot of standardised predicted value vs. standardised residuals of regression – EC per GIFA Run 1

Next, the EC model was run to see whether it performs better than the EC per GIFA model. The Model summary, analysis of variance and model coefficients resulting from each step are presented in Table 6.22, Table 6.23 and Table 6.24 respectively.

Table 6.22: Model Summary – EC Model Run 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	.993	.986	.984	.05023	Log of GIFA, building height, no. of basements, finishes index, services index
2	.993	.985	.984	.05039	Log of GIFA, building height, no. of basements, services index
3	.992	.985	.984	.05042	Log of GIFA, no. of basements, services index
4	.992	.984	.983	.05131	Log of GIFA, no. of basements

Table 6.23: ANOVA table – EC Model Run 1

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.199	5	1.240	491.413	.000
	Residual	.088	35	.003		
	Total	6.287	40			
2	Regression	6.196	4	1.549	610.113	.000
	Residual	.091	36	.003		
	Total	6.287	40			
3	Regression	6.193	3	2.064	812.063	.000
	Residual	.094	37	.003		
	Total	6.287	40			
4	Regression	6.187	2	3.093	1175.084	.000
	Residual	.100	38	.003		
	Total	6.287	40			

Table 6.24: Coefficient of the variables – EC Model Run 1

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-.056	.101		-.555	.582		
	Log of GIFA	.969	.025	.981	39.369	.000	.646	1.547
	Building Height	.004	.003	.034	1.196	.240	.499	2.006
	Basements	.040	.016	.064	2.563	.015	.650	1.538
	Finish Index	-.041	.037	-.023	-1.108	.276	.949	1.054
	Service Index	.008	.007	.025	1.215	.232	.920	1.087
2	(Constant)	-.130	.075		-1.726	.093		
	Log of GIFA	.968	.025	.980	39.230	.000	.647	1.546
	Building Height	.003	.003	.029	1.024	.313	.512	1.951
	Basements	.041	.015	.066	2.676	.011	.657	1.523
	Service Index	.009	.007	.026	1.257	.217	.922	1.085
3	(Constant)	-.159	.070		-2.283	.028		
	Log of GIFA	.983	.020	.995	48.935	.000	.978	1.022
	Basements	.051	.013	.081	3.964	.000	.972	1.029
	Service Index	.010	.007	.031	1.534	.134	.974	1.027
4	(Constant)	-.145	.070		-2.060	.046		
	Log of GIFA	.986	.020	.998	48.444	.000	.987	1.013
	Basements	.048	.013	.077	3.739	.001	.987	1.013

The backward method suggests the best predictive EC regression model in four steps. Adjusted R^2 of the four models are almost the same and the standard error is very similar. However, Mode 1 with all the independent variables has the lowest standard error among the four. On the other hand, F statistics is highest in Model 4. Further, the correlation coefficient of building height finishes and services indices are insignificant in the first three models. Therefore, Model 4 was considered as the best of all. No multicollinearity was found between independent variables as VIF of the independent variables are close to 1. Histogram and scatterplot of the standardised residuals of the regression are presented in Figure 6.26 and Figure 6.25 respectively where residuals follow a normal distribution and do not follow any prominent pattern when mapped against standardised predicted value, which meets the assumption of homoscedasticity. The Durbin-Watson test statistics of the model was 1.93 which is greater than $d_{U,\alpha}$ ($d_{U,\alpha,}$ =1.60) indicating no positive autocorrelation among the residuals. Similarly, 4-d (4 – 1.93 = 2.07) is also greater than $d_{U,\alpha}$ confirms no negative autocorrelation.

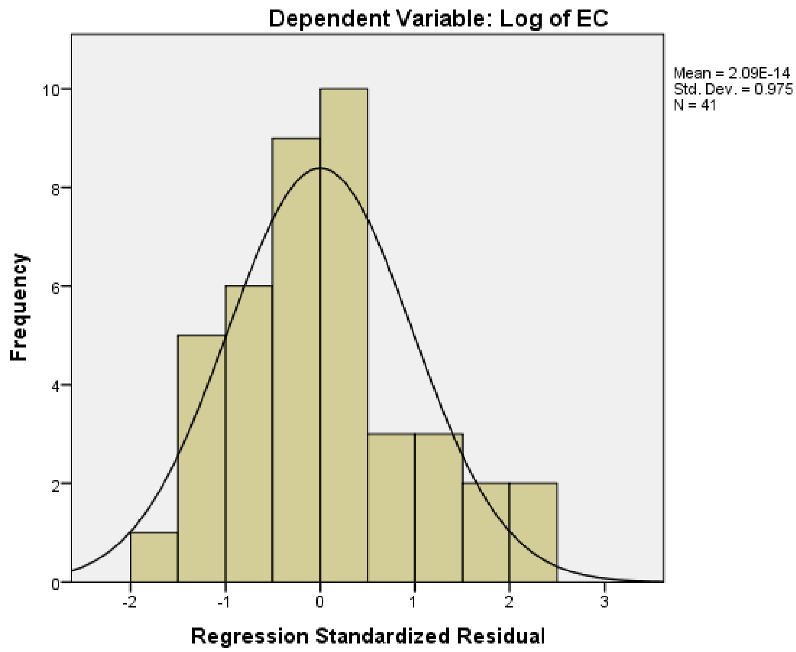


Figure 6.26: Histogram of standardised residual of the regression – EC Run 1

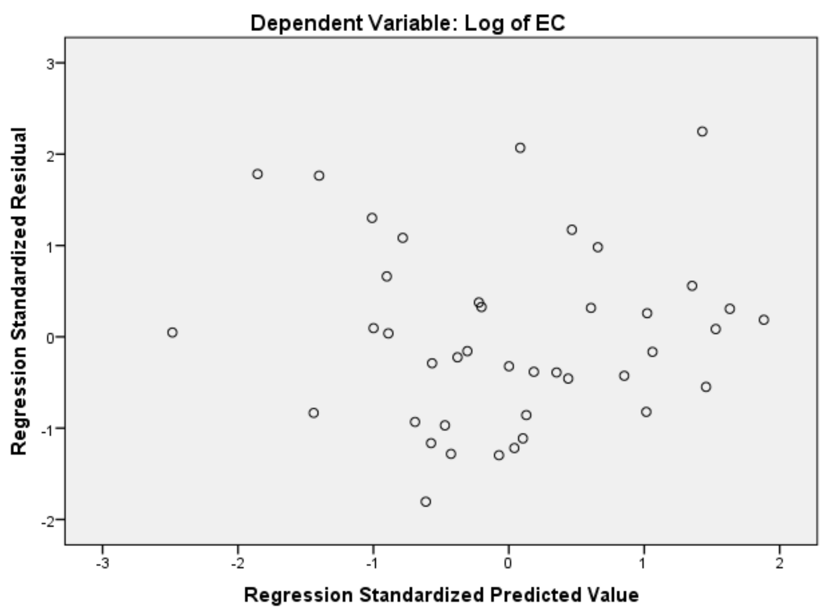


Figure 6.25: Scatterplot of standardised predicted value vs. standardised residuals of regression – EC Run 1

b) Regression models without outliers

Outliers before and after log transformation were identified as a result of verifying the assumptions (See Table 6.25 – numbers indicate the building codes of Dataset 3, for instance, 8 represents the building code D3008). Same regressions were run and the results without outliers are reported here to give the readers better understanding about the influence of the outliers in modelling. Since façade area shows a strong correlation with GIFA as mentioned before façade area is not used as a predictor. Only two data points (D3016 and D3019) were identified as outliers in the sample when formulating EC per GIFA model as the predictor variables include: building height, wall to floor ratio, circulation ratio, no. of basements, finishes index and services index. Similarly, two data points were identified as outliers in EC model (D3003 and D3019). Subsequently, EC per GIFA model and EC model were run again after eliminating the identified outliers from the sample respectively.

Table 6.25: Outliers in the data sample before and after log transformation

Variables	Outliers before log transformations (Building code D30**)	Outliers after log transformations
GIFA	8, 14, 24, 37	3
Building height	19	
Wall to floor ratio	16	
Façade area	24, 37	3
Circulation ratio	None	
basements	None	
EC	8, 14, 37	None
CC	14, 24, 37	None
EC per GIFA	None	
CC per GIFA	1, 3, 21	

The backward method produced EC per GIFA model in the fifth step identifying wall to floor ratio and no. of basements as the predictors (similar to the EC per GIFA model with outliers). Regression summary without the two identified outliers is compared against the output with outliers and presented in Table 6.26. It is clear from all aspects that the model with outlier outperforms the model without outliers – model with outliers has better R^2, lower standard error and higher F statistics. Hence, the model with outliers is identified as the best predictive EC per GIFA model.

Table 6.26: Comparing regression outputs with and without outliers – EC per GIFA models

Summary Statistics	Without Outliers	With Outliers
Predictor Variables	Wall to floor ratio, no. of basements	Wall to floor ratio, no. of basements
R^2	39.8%	51.3%
Adjusted R^2	35.5%	48.1%
Standard error	72.542	70.386
Significance	0.001	0.000
F statistics	9.253	15.828
Durbin-Watson	1.848	1.879
VIF	1.004	1.060

Then, EC model without outliers is compared against EC model with outlier and presented in Table 6.27. EC model without outliers predicts EC using three variables including the log of GIFA, building height and no. of basements. The outputs do not show a drastic difference in the performance of the models. EC model without outliers suggests that 98.2% of the change in the dependent variable (i.e. EC) is explained by GIFA, building height and no. of basements while EC model with outliers suggest that 98.3% of the change in EC is explained by only GIFA and no. of basements. Hence, it is helpful to see the detailed statistics and correlations of the variables, which are presented in Table 6.28 where all three predictor variables have a positive correlation, which is sensible. F statistics suggest that model with outlier outperform the model without outlier while standard error of the estimate is lower in the model without outliers. Analysis of residuals of the model without outliers (Figure 6.27 and Figure 6.28) also conforms to homoscedasticity. Hence, both models have their own merits. However, the p-value of building height in the model without outliers is greater than the 0.05 significance level, which flags a problem in the model. Hence, the model with outliers was selected over the other.

Table 6.27: Comparing regression outputs with and without outliers – EC models

Summary Statistics	Without Outliers	With Outliers
Predictor Variables	Log of GIFA, building, height, no. of basements	Log of GIFA, no. of basements
R^2	98.4%	98.4%
Adjusted R^2	98.2%	98.3%
Standard error	0.04938	0.05131
Significance	0.000	0.000
F statistics	540.996	1175.084
Durbin-Watson	1.917	1.930
VIF	1.098, 1.533, 1.623	1.013

Table 6.28: Coefficient of the variables – EC Model Run 2 – without outliers

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	-.021	.120		-.172	.865		
	Log of GIFA	.945	.031	.968	30.497	.000	.611	1.637
	Building Height	.006	.004	.051	1.490	.149	.533	1.877
	Circulation	.089	.105	.022	.843	.408	.902	1.109
	Basements	.039	.018	.060	2.227	.036	.835	1.198
	Finish Index	-.038	.039	-.026	-.997	.329	.891	1.123
	Service Index	.006	.008	.021	.801	.431	.876	1.142
2	(Constant)	.009	.114		.075	.941		
	Log of GIFA	.940	.030	.962	31.370	.000	.645	1.551
	Building Height	.007	.004	.059	1.845	.077	.590	1.694
	Circulation	.080	.104	.020	.767	.450	.913	1.096
	Basements	.036	.017	.056	2.121	.044	.874	1.145
	Finish Index	-.039	.038	-.027	-1.017	.319	.891	1.122
3	(Constant)	.040	.105		.382	.706		
	Log of GIFA	.938	.030	.961	31.662	.000	.649	1.541
	Building Height	.007	.004	.061	1.942	.063	.596	1.679
	Basements	.039	.017	.060	2.316	.029	.903	1.107
	Finish Index	-.044	.037	-.030	-1.183	.247	.921	1.086
4	(Constant)	-.030	.088		-.341	.736		
	Log of GIFA	.935	.030	.958	31.427	.000	.652	1.533
	Building Height	.006	.004	.055	1.743	.093	.616	1.623
	Basements	.040	.017	.062	2.413	.023	.910	1.098

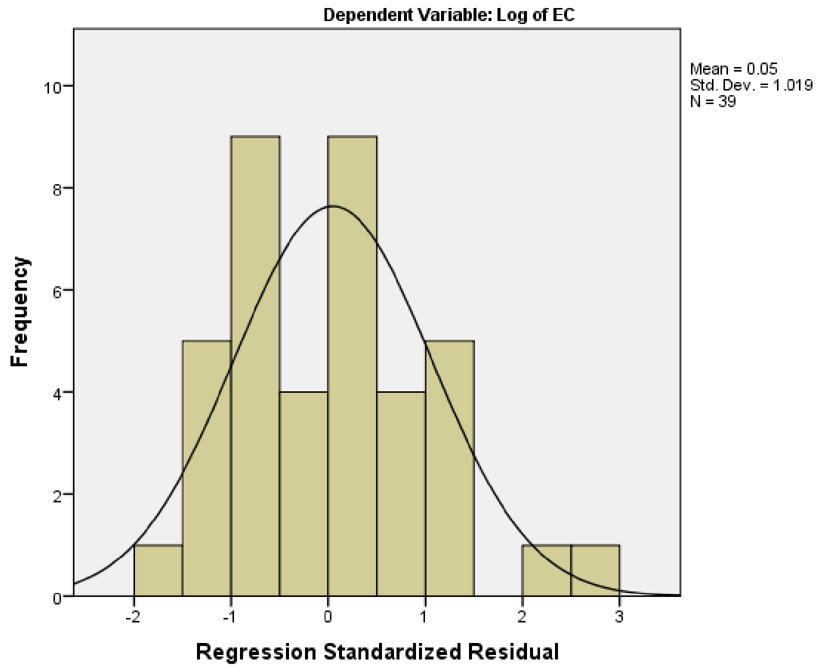


Figure 6.27: Histogram of standardised residual of the regression – EC Run 2

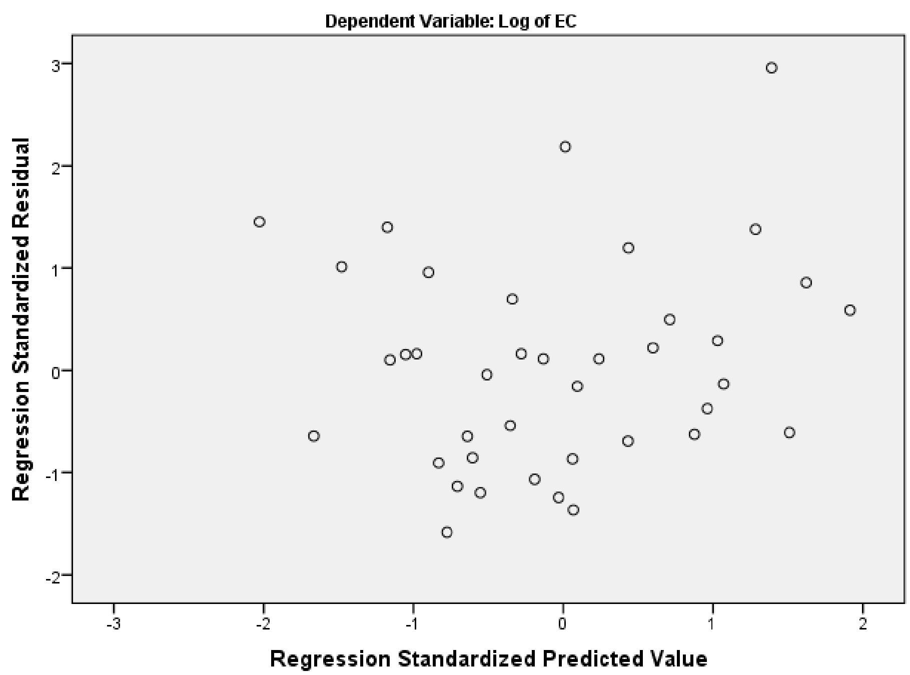


Figure 6.28: Scatterplot of standardised predicted value vs. standardised residuals of regression – EC Run 2

6.5.2. Regression Models for Capital Cost Prediction

Similar to EC models two models were considered to predict CC including CC per GIFA model and CC model. Conceptual models are presented below for CC per GIFA (see, Equation 6.4) and CC (see, Equation 6.5):

Equation 6.4: CC per GIFA conceptual model

$$\widehat{y_3} = c_0 + c_1 x_{BH} + c_2 x_{W:F} + c_3 x_{CR} + c_4 x_B + c_5 x_{FI} + c_6 x_{SI}$$

Where,

$\widehat{y_3}$ $- Estimated\ CC\ per\ GIFA\ of\ the\ building$
c_0 $- Regression\ constant$
c_1 $- Regression\ coefficient\ of\ x_{BH}$
x_{BH} $- Building\ Height$
c_2 $- Regression\ coefficient\ of\ x_{W:F}$
$x_{W:F}$ $- Wall\ to\ Floor\ ratio\ of\ the\ building$
c_3 $- Regression\ coefficient\ of\ x_{CA}$
x_{CR} $- Circulation\ space\ Ratio\ of\ the\ building$
c_4 $- Regression\ coefficient\ of\ x_{CA}$
x_B $- Number\ of\ basements\ in\ the\ building$
c_5 $- Regression\ coefficient\ of\ x_{FI}$
x_{FI} $- Finishes\ Index\ of\ the\ building$
c_6 $- Regression\ coefficient\ of\ x_{SI}$
x_{SI} $- Services\ Index\ of\ the\ building$

Equation 6.5: CC conceptual model

$$\widehat{y_4} = d_0 + d_1 x_{GIFA} + d_2 x_{BH} + d_3 x_{FA} + d_4 x_{CR} + d_5 x_B + d_6 x_{FI} + d_7 x_{SI}$$

Where,

$\widehat{y_4}$ $- Estimated\ CC\ of\ the\ building$
d_0 $- Regression\ constant$
d_1 $- Regression\ coefficient\ of\ x_{GIFA}$
x_{GIFA} $- GIFA\ of\ the\ building$

$d_2 \quad - Regression\ coefficient\ of\ x_{BH}$
$x_{BH} \quad - Building\ Height$
$d_3 \quad - Regression\ coefficient\ of\ x_{FA}$
$x_{FA} \quad - Facade\ area\ of\ the\ building$
$d_4 \quad - Regression\ coefficient\ of\ x_{CA}$
$x_{CR} \quad - Circulation\ space\ Ratio\ of\ the\ building$
$d_5 \quad - Regression\ coefficient\ of\ x_B$
$x_B \quad - Number\ of\ basements\ in\ the\ building$
$d_6 \quad - Regression\ coefficient\ of\ x_{FI}$
$x_{FI} \quad - Finishes\ Index\ of\ the\ building$
$d_7 \quad - Regression\ coefficient\ of\ x_{SI}$
$x_{SI} \quad - Services\ Index\ of\ the\ building$

In the same way, which EC model was modified to address non-normality and collinearity issues after the univariate and bivariate analysis, the CC model was also modified by log transformations applied to the values of CC and GIFA to conform to normality and eliminating the independent variable faced area to eliminate collinearity with GIFA. The modified equation is presented as follows:

Equation 6.6: Modified CC conceptual model

$$\hat{y}_4' = d_0 + d_1 x_{GIFA}' + d_2 x_{BH} + d_5 x_B + d_6 x_{FI} + d_7 x_{SI}$$

Where,

$\hat{y}_4' \quad - \log \widehat{y_4}$
$x_{GIFA}' - \log x_{GIFA}$

a) Regression models with outliers

The backward method produced the best predictive CC per GIFA model in the fourth step with three variables – building height, circulation and finishes index. Model summary, analysis of variance and model coefficients resulting from each step are presented in Table 6.29, Table 6.30 and Table 6.31 respectively.

Table 6.29: Model summary – CC per GIFA Run 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Dependent Variables
1	.734	.538	.432	234.856	Building height, wall to floor ratio, circulation ratio, no. of basements, finishes index, services index
2	.734	.538	.453	230.466	Building height, wall to floor ratio, circulation ratio, no. of basements, finishes index
3	.725	.526	.458	229.339	Building height, wall to floor ratio, circulation ratio, finishes index
4	.714	.510	.459	229.052	Building height, circulation ratio, finishes index

Table 6.30: ANOVA table – CC per GIFA Run 1

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1670984.291	6	278497.382	5.049	.001[b]
	Residual	1434095.635	26	55157.524		
	Total	3105079.926	32			
2	Regression	1670983.637	5	334196.727	6.292	.001[c]
	Residual	1434096.289	27	53114.677		
	Total	3105079.926	32			
3	Regression	1632385.446	4	408096.361	7.759	.000[d]
	Residual	1472694.481	28	52596.231		
	Total	3105079.926	32			
4	Regression	1583605.703	3	527868.568	10.061	.000[e]
	Residual	1521474.223	29	52464.628		
	Total	3105079.926	32			

Table 6.31: Coefficient of the variables – CC per GIFA Run 1

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1904.373	388.884		4.897	.000		
	Building Height	31.155	11.876	.411	2.623	.014	.725	1.380
	Wall to Floor Ratio	142.454	171.450	.121	.831	.414	.834	1.199
	Circulation Ratio	649.002	504.394	.191	1.287	.210	.804	1.244
	Basements	61.136	75.601	.131	.809	.426	.678	1.476
	Finish Index	-622.904	179.115	-.485	-3.478	.002	.915	1.093
	Service Index	.123	35.647	.000	.003	.997	.864	1.157
2	(Constant)	1904.606	375.791		5.068	.000		
	Building Height	31.163	11.426	.411	2.727	.011	.754	1.326
	Wall to Floor Ratio	142.584	164.144	.121	.869	.393	.876	1.141
	Circulation Ratio	648.746	489.536	.191	1.325	.196	.822	1.217
	Basements	61.068	71.637	.131	.852	.401	.727	1.376
	Finish Index	-622.914	175.744	-.485	-3.544	.001	.915	1.093
3	(Constant)	1874.696	372.319		5.035	.000		
	Building Height	35.153	10.372	.463	3.389	.002	.906	1.104
	Wall to Floor Ratio	156.521	162.529	.133	.963	.344	.885	1.130
	Circulation Ratio	755.073	471.064	.223	1.603	.120	.879	1.138
	Finish Index	-629.833	174.698	-.490	-3.605	.001	.917	1.090
4	(Constant)	1953.529	362.754		5.385	.000		
	Building Height	36.739	10.228	.484	3.592	.001	.929	1.076
	Circulation Ratio	881.675	451.782	.260	1.952	.061	.953	1.049
	Finish Index	-636.349	174.348	-.495	-3.650	.001	.919	1.089

The model summary suggests that Model 2, Model 3 and Model 4 have almost similar and better R^2 than Model 1. However, the drop from R^2 to adjusted R^2 is less in Model 4 compared to other models. Further, the standard error of the estimate is also the lowest in Model 4 while not a big difference in standard error is noticeable among models. 45.1% of the change in CC per GIFA is explained by building height, circulation ratio and finishes index while 48.1% of the change in EC per GIFA was explained by wall to floor ratio and number of basements by the best predictive models. However, the correlation coefficient of finishes index was negative in Model 4, which is unusual. It can be expected that higher quality of finishes will increase CC per GIFA rather than decrease it, hence, a positive correlation is anticipated. Therefore, Model 4 seems to have a practical problem though it is statistically significant. Subsequently, regression analysis was performed again after eliminating finishes index as a predictor and the summary statistics are presented in Table 6.32, Table 6.33 and Table 6.34. The new model without finishes index has a lower R^2 and F statistics compared to the previous model with finishes index, which is not impressive.

Table 6.32: Model Summary - CC per GIFA Run 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	.569	.323	.198	278.965	Building height, wall to floor ratio, circulation ratio, no. of basements, services index
2	.569	.323	.227	273.951	Building height, wall to floor ratio, circulation ratio, no. of basements
3	.554	.307	.235	272.496	Building height, circulation ratio, no. of basements
4	.534	.285	.237	272.053	Building height, circulation ratio

Table 6.33: ANOVA table – CC per GIFA Run 2

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1003894.786	5	200778.957	2.580	.049
	Residual	2101185.141	27	77821.672		
	Total	3105079.926	32			
2	Regression	1003703.206	4	250925.802	3.343	.023
	Residual	2101376.720	28	75049.169		
	Total	3105079.926	32			
3	Regression	951710.335	3	317236.778	4.272	.013
	Residual	2153369.591	29	74254.124		
	Total	3105079.926	32			
4	Regression	884692.304	2	442346.152	5.977	.007
	Residual	2220387.622	30	74012.921		
	Total	3105079.926	32			

Table 6.34: Coefficient of the variables – CC per GIFA Run 2

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	706.880	214.673		3.293	.003		
	Building Height	21.220	13.692	.280	1.550	.133	.769	1.300
	Wall to Floor Ratio	160.085	203.562	.136	.786	.438	.835	1.198
	Circulation Ratio	913.621	592.269	.269	1.543	.135	.823	1.216
	Basements	73.950	89.693	.158	.824	.417	.679	1.472
	Service Index	2.101	42.337	.008	.050	.961	.865	1.157
2	(Constant)	710.545	197.938		3.590	.001		
	Building Height	21.355	13.178	.282	1.621	.116	.801	1.249
	Wall to Floor Ratio	162.308	195.003	.138	.832	.412	.877	1.140
	Circulation Ratio	909.301	575.304	.268	1.581	.125	.841	1.189
	Basements	72.795	85.063	.156	.856	.399	.728	1.373
3	(Constant)	783.729	176.395		4.443	.000		
	Building Height	22.432	13.044	.296	1.720	.096	.809	1.237
	Circulation Ratio	1029.600	553.896	.303	1.859	.073	.897	1.114
	Basements	79.969	84.176	.171	.950	.350	.736	1.359
4	(Constant)	734.949	168.482		4.362	.000		
	Building Height	27.702	11.787	.365	2.350	.026	.987	1.013
	Circulation Ratio	1188.330	527.237	.350	2.254	.032	.987	1.013

VIF of the variables in Model 4 is close to one (1), which confirms no multicollinearity in the model. Histogram and scatterplot of standardised residuals of regression are presented in Figure 6.29 and Figure 6.30. Histogram of standardised residuals affirms the normality of the residuals while scatterplot shows that the residuals do not follow any particular pattern. Hence, the assumption of homoscedasticity is met. The Durbin-Watson test statistics of the model was 2.091 which is greater than $d_{U,\alpha}$ ($d_{U,\alpha,}$ =1.60) indicating no positive autocorrelation among the residuals. Similarly, 4-d (4 – 2.091 = 1.909) is also greater than $d_{U,\alpha}$ confirms no negative autocorrelation.

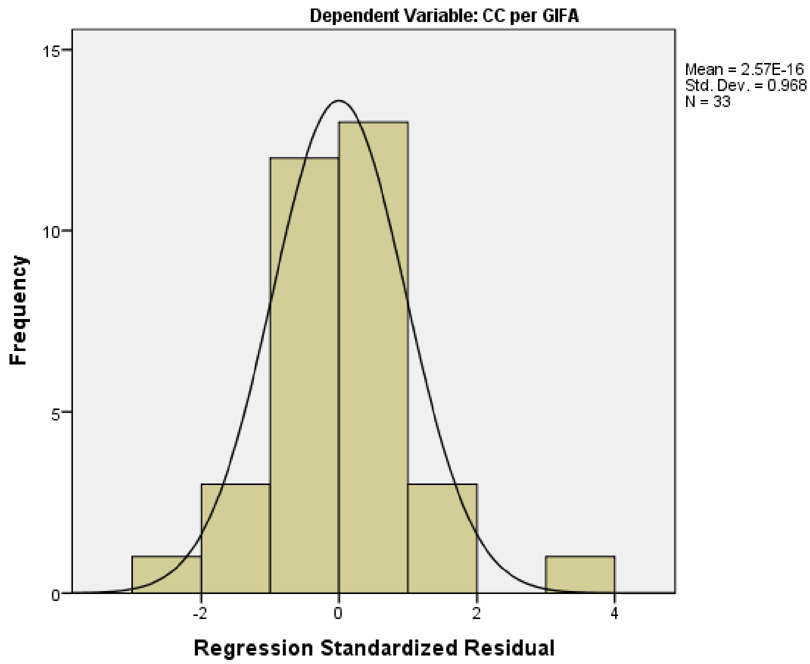


Figure 6.29: Histogram of standardised residual of the regression – CC per GIFA Run 1

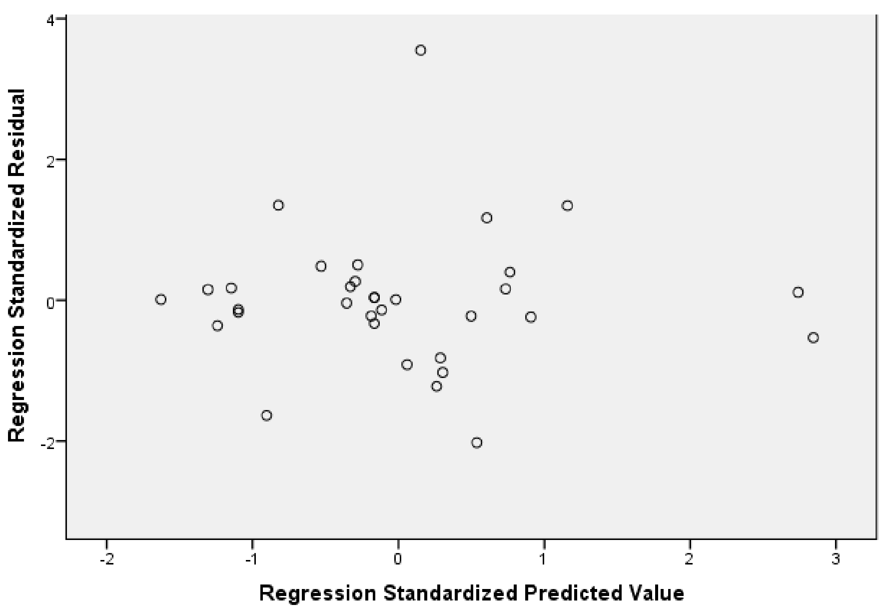


Figure 6.30: Scatterplot of standardised predicted value vs. standardised residuals of regression – CC per GIFA Run 1

Then, regression was run for CC model. The model summary, analysis of variance and model coefficients resulting from each step are presented in Table 6.35, Table 6.36 and Table 6.37.

.

Table 6.35: Model Summary – CC Model Run 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	.981	.962	.957	.08782	Log of GIFA, building Height, basements, finish Index, service Index
2	.981	.962	.958	.08664	Log of GIFA, building Height, basements, finish Index
3	.981	.962	.959	.08556	Log of GIFA, building Height, finish Index

Table 6.36: ANOVA table – CC Model Run 1

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.884	5	1.377	178.502	.000
	Residual	.270	35	.008		
	Total	7.154	40			
2	Regression	6.884	4	1.721	229.250	.000
	Residual	.270	36	.008		
	Total	7.154	40			
3	Regression	6.883	3	2.294	313.377	.000
	Residual	.271	37	.007		
	Total	7.154	40			

Table 6.37: Coefficient of the variables – CC Model Run 1

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.398	.176		2.259	.030		
	Log of GIFA	.980	.043	.929	22.752	.000	.646	1.547
	Building Height	.013	.005	.120	2.585	.014	.499	2.006
	Basements	.009	.027	.013	.326	.746	.650	1.538
	Finish Index	-.185	.065	-.095	-2.830	.008	.949	1.054
	Service Index	.002	.012	.007	.192	.848	.920	1.087
2	(Constant)	.404	.171		2.364	.024		
	Log of GIFA	.979	.042	.929	23.091	.000	.649	1.542
	Building Height	.014	.005	.122	2.742	.009	.527	1.896
	Basements	.008	.026	.011	.294	.771	.687	1.455
	Finish Index	-.185	.064	-.096	-2.879	.007	.951	1.052
3	(Constant)	.419	.160		2.619	.013		
	Log of GIFA	.974	.039	.924	25.289	.000	.766	1.305
	Building Height	.014	.004	.130	3.523	.001	.757	1.320
	Finish Index	-.187	.063	-.097	-2.954	.005	.959	1.043

The best predictive CC model is produced in the third step by the backward method. The adjusted R^2 and the standard error of all three models display no big difference though the third model has the highest adjusted R^2 and the lowest standard error. Further, F statistic is highest in the third model. Even though, the correlation coefficients of the independent variables are found to be statistically significant in the third model finishes index negatively correlated to the log of CC is abnormal, similar to the case explained in the CC per GIFA model. Therefore, regression was run again without the finishes index and the results are presented in Table 6.38, Table 6.39 and Table 6.40.

.

Table 6.38: Model Summary – CC Model Run 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	.977	.954	.948	.09599	Log of GIFA, building Height, basements, service Index
2	.976	.954	.950	.09479	Log of GIFA, building Height, basements
3	.976	.953	.951	.09386	Log of GIFA, building Height

Table 6.39: ANOVA table – CC Model Run 2

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.822	4	1.706	185.099	.000
	Residual	.332	36	.009		
	Total	7.154	40			
2	Regression	6.821	3	2.274	253.069	.000
	Residual	.332	37	.009		
	Total	7.154	40			
3	Regression	6.819	2	3.410	386.994	.000
	Residual	.335	38	.009		
	Total	7.154	40			

Table 6.40: Coefficient of the variables – CC Model Run 2

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.066	.144		.460	.648		
	Log of GIFA	.976	.047	.926	20.745	.000	.647	1.546
	Building Height	.011	.006	.099	1.965	.057	.512	1.951
	Basements	.016	.030	.025	.556	.582	.657	1.523
	Service Index	.004	.013	.011	.282	.780	.922	1.085
2	(Constant)	.074	.139		.537	.594		
	Log of GIFA	.975	.046	.925	21.028	.000	.649	1.540
	Building Height	.011	.005	.102	2.113	.041	.541	1.848
	Basements	.015	.028	.022	.511	.612	.693	1.443
3	(Constant)	.099	.129		.764	.449		
	Log of GIFA	.965	.042	.916	22.914	.000	.771	1.297
	Building Height	.013	.004	.115	2.883	.006	.771	1.297

Run 2 returned the best predictive CC model with two independent variables – log of GIFA and building height. The standardised residuals depict a normal distribution (see, Figure 6.31) and were randomly distributed (see, Figure 6.32), fulfilling the assumption of homoscedasticity. The model also has a VIF statistics of 1.297, which confirms no multicollinearity between independent variables. The Durbin-Watson test statistics of the model was 2.005 which is greater than $d_{U,\alpha}$ ($d_{U,\alpha,}$ =1.60) indicating no positive autocorrelation among the residuals. Similarly, 4-d (4 – 2.005 = 1.995) is also greater than $d_{U,\alpha}$ confirms no negative autocorrelation.

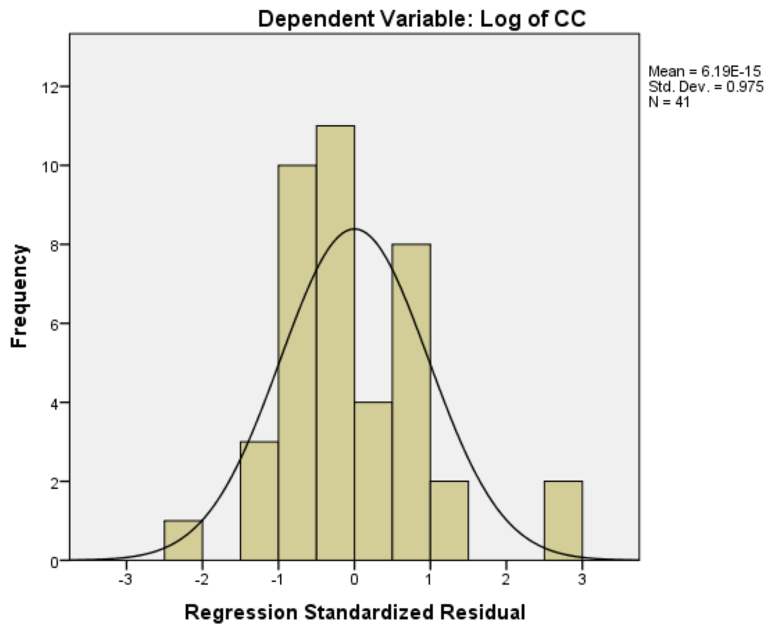


Figure 6.31: Histogram of standardised residual of the regression – CC Run 2

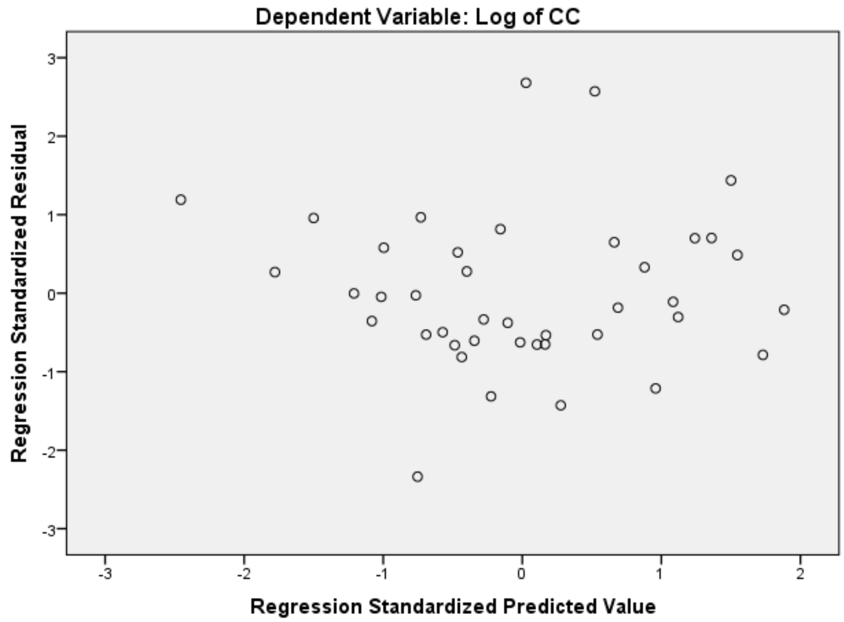


Figure 6.32: Scatterplot of standardised predicted value vs. standardised residuals of regression – CC Run 1

b) Regression models without outliers

Five data points (D3001, D3003, D3016, D3019 and D3021) were identified as outliers in the sample when formulating CC per GIFA model and two data points were identified as outliers in CC model (D3003 and D3019) similar to EC model. Subsequently, CC per GIFA and CC models were run again after eliminating the identified outliers from the sample respectively.

The backward method produced CC per GIFA model in the sixth step identifying only building height as the predictor while CC per GIFA model with outlier identified circulation ratio as another predictor of CC per GIFA. Regression summary with and without the outliers are compared and presented in Table 6.41. Model without outliers is found to be statistically insignificant (sig > 0.05) which disqualifies the model without outliers before comparing with other statistics. Hence, the model with outliers was identified as the best predictive CC per GIFA model.

Table 6.41: Comparing regression outputs with and without outliers – CC per GIFA models

Summary Statistics	Without Outliers	With Outliers
Predictor Variables	Building height	Building height, circulation ratio
R^2	10.6%	28.5%
Adjusted R^2	7.3%	23.7%
Standard error	205.099	272.053
Significance	0.084	0.007
F statistics	3.209	5.977
Durbin-Watson	2.027	2.091
VIF	1.000	1.013

Then, CC model without outliers was compared against CC model with outlier and presented in Table 6.42. CC model without outliers also predicts CC using the same two variables log of GIFA and building height. CC model with outliers outperforms CC model without outliers in terms of adjusted R^2 and F statistics. Hence, CC model with outliers was selected as the best predicting CC model.

Table 6.42: Comparing regression outputs with and without outliers – CC models

Summary Statistics	Without Outliers	With Outliers
Predictor Variables	Log of GIFA, building, height	Log of GIFA, building, height
R^2	95.2%	95.3%
Adjusted R^2	94.8%	95.1%
Standard error	0.09192	0.09386
Significance	0.000	0.000
F statistics	275.133	386.994
Durbin-Watson	1.786	2.005
VIF	1.478	1.297

6.5.3. Final Regression Models

Final EC and CC models were derived from the detailed analysis of models and their constructs based on the best available data sample at the time of the research. EC and EC per GIFA; CC and CC per GIFA models were formulated and the accuracy is tested for all four models to identify the best predictive model in each pair.

a) EC per GIFA Model

Equation 6.7 presents the best predictive EC per GIFA model derived from Table 6.21. The model with outliers was selected as it outperforms the model without outliers in all aspects including adjusted R^2, F statistics and standard error.

Equation 6.7: EC per GIFA model

$$\widehat{y_1} = 530.62 + 164.08x_{W:F} + 68.15x_B$$

Where,

$\widehat{y_1}$ $- Estimated\ EC\ per\ GIFA\ of\ the\ building$
$x_{W:F}$ $- Wall\ to\ Floor\ ratio\ of\ the\ building$
x_B $- Number\ of\ basements\ in\ the\ building$

The regression analysis suggests that wall to floor ratio and the number of basements in the building are the most statistically significant design variables in predicting the EC of the building during early stages of design over the other design variables. The model explains 48.1% of the variation in EC per GIFA accounted by Wall to Floor ratio and the number of basements. The model indicates that increase in one unit of wall to floor ratio (say, 0.3 to 1.3) while maintaining the number of basements will increase EC per GIFA by 164.08 $kgCO_2/m^2$ and adding a basement will increase EC per GIFA by 68.15 $kgCO_2/m^2$ for a given wall to floor ratio. Both of the coefficients are reasonable as a higher wall to floor ratio implies higher façade area and more basements implies more material and plant inputs for a given GIFA. In addition, the addition of basements in building design influence the substructure of the building and substructure is identified as a predominant carbon hotspot in office buildings. Therefore, the significance of basement as a variable in the model can be explained by the fact that Substructure is identified as a predominant carbon hotspot. Further, it can be noticed that the constant is high compared to other coefficients. This can be explained by the descriptive statistics of the sample data as the EC per GIFA ranges from 551 $kgCO_2/m^2$ to 916 $kgCO_2/m^2$. Even the smallest building has an EC per GIFA value of 834 $kgCO_2/m^2$ GIFA. Therefore, it is clear from the coefficient that the minimum EC per GIFA of a building will be more than 530.62 kgCO2/m^2 as per the results.

However, it was surprising to find that building height has not been identified as a significant predictor as Frame is identified as a predominant building element and literature (Luo et al., 2015) also suggest that building height (no. of storeys) and EC per GIFA has a strong positive correlation while the relationship found in the study was not stronger (0.392 at the 0.05 level) compared to the findings of Luo et al. (2015). It can be articulated that when fitting into the regression model other variables (wall to floor ratio and basements) override building height. This may be due to the selected sample and with a larger sample different result can be expected.

b) EC Model

The best predictive EC model derived from the sample data is presented in Equation 6.8 derived from Table 6.28. The model with outliers was selected because both of the independent variables in the models were statistically significant. Further, the adjusted R^2 and F statistics were also higher in this model than the model without outliers.

Equation 6.8: EC model

$$\hat{y}_2' = -0.145 + 0.986x_{GIFA}' + 0.048x_B$$

Where,

$\hat{y}_2' \quad - \log \hat{y}_2$

$\hat{y}_2 \quad - Estimated\ EC\ of\ the\ building$

$x_{GIFA}' - \log x_{GIFA}$

$x_{GIFA} - GIFA\ of\ the\ building$

$x_B \quad - Number\ of\ basements\ in\ the\ building$

EC model identifies GIFA and number of basements as the most statistically significant design variables to predict EC for early stages building designs. The model explains 98.3% of the variation in EC accounted by GIFA and number of basements. EC is highly influenced by GIFA than any other variables because as the building becomes bigger the EC content will also become higher due to more material and plant inputs. However, the model, does not predict the estimated EC of buildings but the log of estimated EC. Therefore, the model prediction has to be converted to get the estimated EC of the building (See, Table 6.43). Accordingly, one unit increase in Log of GIFA (given the no. of basements) will increase EC by 9.683 tCO_2. Similarly, adding another basement (given the GIFA) will increase EC by 1.10 tCO_2.

Table 6.43: Transforming Log of EC to EC

Predictor Variables	Increase in Log EC	Increase in EC (tCO_2)
Log GIFA	0.986	9.68278
Basements (No)	0.048	1.11686

c) CC per GIFA Model

Equation 6.9 presents the best predictive CC per GIFA model derived from Table 6.34. The model with outliers was selected as it outperforms the model without outliers in terms of R^2 and F statistics even though the standard error was higher.

Equation 6.9: CC per GIFA model

$$\widehat{y_3} = 734.95 + 27.7x_{BH} + 1188.33x_{CR}$$

Where,

$\widehat{y_3} \quad - Estimated\ CC\ per\ GIFA\ of\ the\ building$

Proposed CC per GIFA model identifies building height and circulation ratio as the statistically significant predictors over the others. The model explains 23.7% of the variation in CC per GIFA accounted by building height and circulation ratio, which indicates a poor fit of the model. The model suggests that increase in one unit of building height (say, 10m to 11m) will increase CC per GIFA by £27.2/m^2 for a given circulation ratio and increasing circulation area by 1% for a given building height will increase CC per GIFA by £11.88/m^2. Both coefficients are sensible as taller buildings generally have higher CC per GIFA due to plants involved in hoisting materials and operations and higher circulation ratio increases services cost resulting in higher CC per GIFA. Services are the most cost significant building element in office buildings contributing up to 40% of the CC. Hence, it makes sense when circulation ratio is identified as a significant predictor as more circulation space increases Services cost. Phaobunjong (2002) also found a negative correlation coefficient for usable space ratio (usable space/GIFA) in his parametric cost model, which implies lower usable space (higher circulation space) increase Services cost which supports the study findings. Further, the constant is 734.95, which implies the minimum CC per GIFA of a building will be more than £734.95 /m^2. This can be explained by the descriptive statistics of the sample data as the EC per GIFA ranges from £698/m^2 to £2,285/m^2.

d) CC Model

The best predictive CC model derived from the sample data is presented in Equation 6.10 derived from Table 6.40. The model with outliers was selected as it outperforms the model without outliers similar to CC per GIFA model in terms of R^2 and F statistics even though the standard error was higher.

Equation 6.10: CC model

$$\hat{y}_4' = -0.099 + 0.965x_{GIFA}' + 0.013x_{BH}$$

Where,

$\hat{y}_4'$ $- \log \hat{y}_4$
$\hat{y}_4$ $- Estimated\ CC\ of\ the\ building$

The proposed CC model suggests that GIFA and building height are the statistically significant design variables that predict CC during early stages of design while EC model identified the number of basements as a significant design variable in addition to GIFA and building height in predicting EC. The model explains 95.1% of the variation in CC accounted by GIFA and building height. Similar to EC model, CC model also predicts log of estimated CC. Therefore, the model prediction has to be converted to get the estimated CC of the building (see, Table 6.44). As per the model, one unit increase in Log of GIFA (given the building height) will increase CC by £9226. Similarly, increase in one unit of building height (given the GIFA) will increase CC by £1030.

Table 6.44: Transforming Log of CC to CC

Predictor Variables	Increase in Log CC	Increase in CC (£1000s)
Log GIFA	0.965	9.22571
Building Height (m)	0.013	1.03039

6.5.4. Assumptions in the multiple regression analysis

The outcomes of the multiple regression analysis rely on five assumptions (Statistics Solutions, 2016, Miles and Shevlin, 2001) as discussed in Subsection 4.9.3 (f) and how the models comply with these assumptions are discussed herein.

a) Assumption 1 - Normality of data

The dependent variables of the four models include EC per GIFA, Log of EC, CC per GIFA and Log of CC. The histograms, boxplots and descriptive statistics confirm that the variables are normally distributed as discussed in Section 6.4.2 and 6.4.4.

b) Assumption 2 - Linear relationship between dependent and independent variables

This assumption is met if the residuals in the standardised residual plots are randomly distributed which is satisfied by all the four models.

c) Assumption 3 - No multicollinearity between independent variables

VIF values of the models were less than 5 in all the cases, which assure no multicollinearity between the independent variables in the model.

d) Assumption 4 - Residuals are homoscedastic

Histograms and scatterplots for standardised residuals of the regressions for the models confirm that the residuals are homoscedastic (randomly distributed) and do not demonstrate any significant patterns.

e) Assumption 5 - Residuals are not autocorrelated

Durbin-Watson score was used to test this assumption and all four models satisfied this assumption. The summary of Durbin-Watson score is presented in Table 6.45.

Table 6.45: Summary of Durbin-Watson statistics of the models

Model	Durbin-Watson Score (d)	$d_{U,\alpha,}$	$d > d_{U,\alpha,}$ $(4-d > d_{U,\alpha,})$
EC per GIFA	1.879	1.60	Yes
EC	1.930	1.60	Yes
CC per GIFA	2.091	1.60	Yes
CC	2.005	1.60	Yes

6.6. Embodied Carbon and Capital Cost Relationships

Summary of the predictor variables of all four models is presented in Table 6.46. Accordingly, EC and CC models have the same set of the predictor variables (GIFA and building height) except for the number of basements in EC model. The number of basements has been identified as a significant design variable in EC model because Substructure was found to be a predominant carbon significant building element. On the other hand, EC per GIFA and CC per GIFA models have two distinct predictor variables. Even though the variables are distinctive some of the variables are interrelated. For instance, Wall to Floor ratio implicitly captures building height. Subsequently, it was decided to study the relationship between EC and CC due to the similarities found in the model predictors and the carbon and cost hotspots.

Pearson's correlation was calculated to identify the relationship between EC and CC (see, Table 6.47), and EC per GIFA and CC per GIFA (see, Table 6.48) as it can be expected that correlation between EC and CC is caused by a common third factor which is GIFA in this case. A very strong positive correlation was found between CC and EC as expected (0.977). On the other hand, EC per GIFA and CC per GIFA was also found to be strongly correlated with a correlation coefficient of 0.645.

Table 6.46: Summary of dependent and independent variables of the models

Dependent variable	Independent variables
EC per GIFA	Wall to floor ratio, no. of basements
CC per GIFA	Building height, circulation ratio
EC	GIFA, building height, no. of basements
CC	GIFA, building height

Table 6.47: Correlation between CC and EC

		CC
EC	Pearson Correlation	.977
	Sig. (2-tailed)	.000
	N	41

Table 6.48: Correlation between CC per GIFA and EC per GIFA

		CC per GIFA
EC per GIFA	Pearson Correlation	.645
	Sig. (2-tailed)	.000
	N	41

Further, the data points were also mapped in a scatterplot to study the relationship in detail. CC and EC relationship and CC per GIFA and EC per GIFA relationship are presented in Figure 6.33 and Figure 6.34 respectively. Accordingly, CC and EC showcase a perfect linear correlation and there are only a few data points scattered from a straight line while the graph of CC per GIFA and EC per GIFA is not perfectly linear and many data points are scattered. However, a strong positive linear relationship is noticeable. EC and CC relationships suggest that both EC and CC tend to move in the same direction.

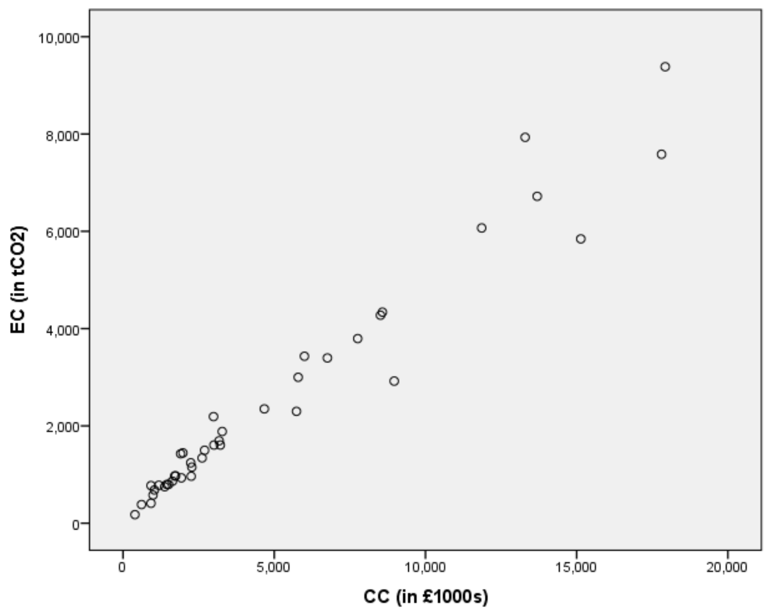


Figure 6.33: Scatterplot of CC and EC

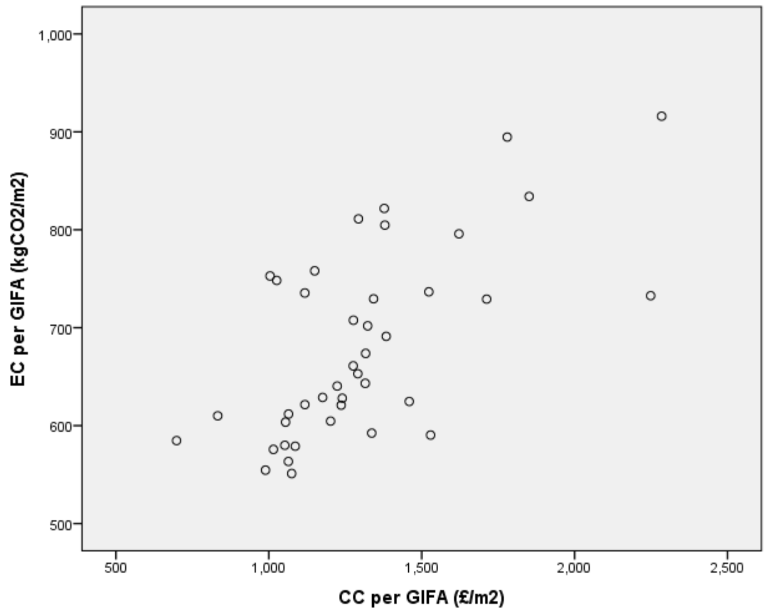


Figure 6.34: Scatterplot of CC per GIFA and EC per GIFA

Nevertheless, it is important to analyse at an elemental level to understand the intricacies of the relationships. There is a lack of reported study especially on the relationship between EC per GIFA and CC per GIFA. Consequently, EC per GIFA and CC per GIFA of the building elements were analysed and presented in Figure 6.35. Accordingly, both EC per GIFA and CC per GIFA follows a similar pattern across the elements expect for Substructure, Frame and Upper Floors where the EC per GIFA exceed the CC per GIFA values. This explains the fact that Substructure, Frame and Upper Floors being identified as lead carbon hotspots in the sample. Even though Services was identified as the second most carbon significant element, CC per GIFA of Services was extremely higher than the EC per GIFA of Services, which is apparent in Figure 6.35.

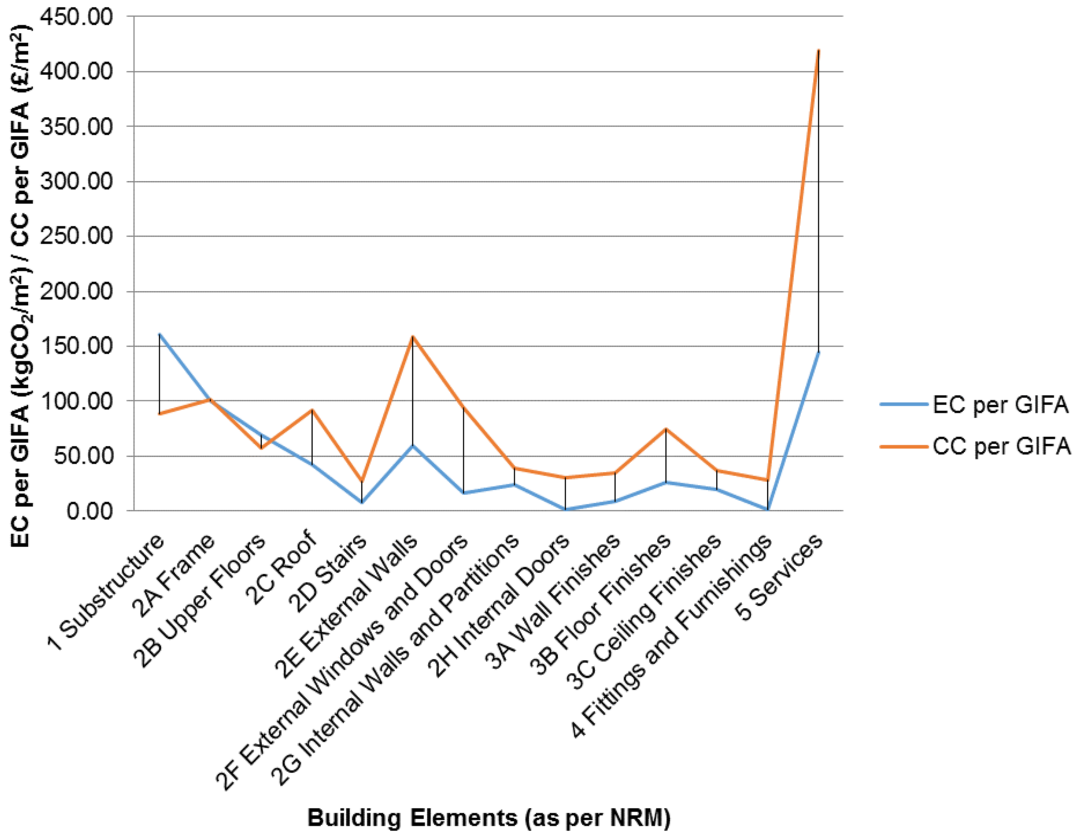


Figure 6.35: Comparing the EC per GIFA and the CC per GIFA of the building elements of the sample

Table 6.49 presents the descriptive statistics of EC per GIFA and CC per GIFA. In terms of EC per GIFA, Substructure, Upper Floors, Floor Finishes, Ceiling Finishes and Services have lower standard deviation, which implies that 68% of the data are closely clustered around the meanwhile EC per GIFA of Stairs, Windows and External Doors, Internal Walls and Partitions and Wall Finishes demonstrate high dispersion of data. On the other hand, CC per GIFA values of Roof and Services have a lower standard deviation (close to the means) while Stairs, External Windows and Doors, Internal Walls and Partitions, Wall Finishes and Fittings, Furnishings and Equipment demonstrate high standard deviation (high dispersion from the means).

Table 6.49: Descriptive statistics of EC per GIFA and CC per GIFA of the sample

Building Element	EC per GIFA ($kgCO_2/m^2$)				CC per GIFA (£/m^2)			
	Average	Standard Deviation	Minimum	Maximum	Average	Standard Deviation	Minimum	Maximum
1 Substructure	161.15	57.54	54.04	329.00	89.10	43.00	38.00	222.00
2A Frame	100.21	50.99	21.29	230.36	102.02	53.21	18.00	277.00
2B Upper Floors	68.84	25.16	8.14	131.10	56.90	29.86	2.00	125.00
2C Roof	42.81	21.57	14.66	113.76	91.41	36.26	24.00	163.00
2D Stairs	7.59	7.44	0.00	45.99	26.73	18.06	0.00	77.00
2E External Walls	59.80	25.76	13.62	120.64	159.12	104.73	46.00	506.00
2F Windows and External Doors	15.97	11.44	0.00	53.92	93.63	61.74	0.00	281.00
2G Internal Walls and Partitions	24.04	33.58	2.31	176.40	39.32	31.52	2.00	149.00
2H Internal Doors	1.37	0.89	0.23	4.00	30.51	22.44	6.00	118.00
3A Wall Finishes	9.36	15.50	2.21	103.71	34.24	21.42	11.00	111.00
3B Floor Finishes	25.89	9.13	7.18	38.92	74.90	28.92	15.00	148.00
3C Ceiling Finishes	19.40	5.29	3.98	26.03	36.39	14.77	9.00	75.00
4 Fittings and Furnishings	0.86	0.00	0.86	0.86	28.31	37.22	0.87	218.00
5 Services	145.09	19.38	92.67	177.69	418.93	144.77	164.00	864.00

Table 6.50: Element level analysis of EC and CC relationships

Element	Pearson's Correlation between EC per GIFA and CC per GIFA	P - value	No. of observations
Substructure	0.639	0.000	41
Frame	0.707	0.000	41
Upper Floors	0.816	0.000	41
Roof	-0.068	0.672	41
External Walls	0.741	0.000	41
Stairs	0.086	0.592	41
External Windows and Doors	0.442	0.004	41
Internal Walls and Partitions	0.872	0.000	41
Internal Doors	0.769	0.000	41
Wall Finishes	0.288	0.067	41
Floor Finishes	0.457	0.003	41
Ceiling Finishes	0.015	0.927	41
Fittings, Furnishings and Equipment	-	-	-
Services	0.277	0.080	41

The relationships between the EC per GIFA and CC per GIFA were also analysed to get insights into elemental relationships. Pearson's correlation coefficients between EC per GIFA and CC per GIFA were calculated and presented in Table 6.50. Accordingly, most of the elements demonstrated a statistically significant (at 99% confidence) positive correlation between EC per GIFA and CC per GIFA except for Roof, Wall Finishes, Ceiling Finishes and Services. Especially, EC per GIFA and CC per GIFA were very strongly correlated (> 0.80) in Upper Floors and Internal Walls and Partitions and strongly correlated (between 0.60 and 0.79) in Substructure, Frame, External Walls, and Internal Doors. The correlation was moderate (between 0.40 and 0.59) in External Windows and Doors and Floor Finishes. Further, no correlation was found between EC per GIFA and CC per GIFA of Fittings, Furnishings and Equipment due to the use of an average EC per GIFA value. The elemental analysis suggests that not only there is a relationship between EC and CC at building level but also in elemental level.

6.7. Summary

Carbon and cost analysis was undertaken to identify the most carbon and cost significant building elements in office buildings and to select the most cost and carbon influential design variables in light of achieving the third objective. Substructure, Services, Frame, Upper Floors, External Walls and Roof were identified as the most carbon significant building elements. On the other hand, it was noticed that the same building elements are accountable for 72% of the CC. Alongside, 80:20 Pareto ratio was also verified which was not supported in the case of EC but the findings propose a ratio of 80:43, which implies that 43% of building elements are responsible for 80% of EC emissions. Services, External Walls, Frame, External Windows and Doors, Roof, Substructure, and Floor Finishes were identified as the most cost significant building elements and these elements are identified to be accountable for 80% of the CC of the buildings on average. Further, the cost hotspot analysis proposed a ratio of 80:50.

In addition, building elements were categorised into three types namely: 'Lead Positions', 'Special Positions' and 'Remainder Positions'. Lead positions are the building elements which were identified as carbon/cost hotspot in most of the buildings (≥80%) in the samples and Frame, External Walls and Services were identified as lead carbon and cost hotspots. Remainder positions are the building elements that were seldom identified as hotspots and Stairs, Internal Doors and Fittings, Furnishing and Equipment were identified as remainder carbon hotspots. Wall, Floor and Ceiling Finishes were identified as special carbon and cost hotspot, which are building elements that are identified as hotspots in some of the buildings in the sample (0-80%, both numbers exclusive). Based on the hotspot analysis, cost and carbon significant elements were captured which are ‘Lead’ and ‘Special’ positions, and the design variables affecting those elements were identified.

Finishes and services quality indices were decided to be developed to represent the quality level of the building in the model after finishes and services being identified as cost and carbon significant building elements. Finishes index was developed from a conceptual finishes index developed for the study which had three levels including basic, Moderate and Luxury for wall, floor and ceiling finishes. The conceptual finishes index was content verified through a Delphi-based expert forum consisting of five experts in two rounds. The conceptual finishes index was improved by the experts’ inputs and the final finishes index for the study was derived. On the other hand, services index was developed from Spon’s Mechanical and Electrical Services price book to suit the study data. Consequently, finishes quality and services quality of each building in the sample was denoted in accordance with the developed finishes and services indices. However, for the finishes quality of the building, the overall finishes quality index of each building was calculated from individual wall, floor and ceiling finishes indices using weighted average method based on the area finished.

Bivariate analysis was performed to find correlations between EC and design variables and CC and design variables to achieve the fourth objective. Statistically significant (α=0.05) relationships were found between EC and certain design variables including GIFA, Building Height and Faced Area. These correlation coefficients also remain significant at 0.01 significance level. Very similar results were obtained for CC, which makes it comparable. On the other hand, EC per GIFA correlate with Wall to Floor ratio and Circulation Ratio. The correlations between EC per GIFA and Wall to Floor Ratio was also significant at 0.01 significance level. CC per GIFA correlate with Building Height, Wall to Floor ratio and Circulation Ratio at 0.05 significance level. However, the correlations were not significant at 0.01 significance level.

Further, non-normal distributions, outliers and non-linear relationships in the data were identified from the univariate and the bivariate analyses of the variables. GIFA, EC and CC were found to be positively skewed and the need for data transformation was realised. Hence, log transformation was applied to GIFA, EC and CC. Further, collinearity between GIFA and façade area caused the elimination of façade area as a predictor variable in the models. Also, statistically significant linear relationship was not found between circulation space and EC and CC, hence, circulation space was eliminated from being a predictor in the models. Consequently, regressions were run for EC, EC per GIFA, CC and CC per GIFA, which were the dependent variables of the four different models. Regressions were performed in two rounds for each type with and without outliers to compare the results. Except for the EC model, all models with outliers outperformed the models without outliers. EC and CC models demonstrated much better coefficients of determination (R^2) or better model fit compared to EC per GIFA and CC per GIFA models. In addition, the models were tested against all the regression assumptions to ensure that no major violation of assumptions had occurred, as the regression outcomes are sensitive to the assumptions. Accordingly, all models satisfied the regression assumptions. In this way, the sixth objective of formulating models was achieved.

Pearson's correlation was calculated and analysed between EC and CC at building level and elemental level in order to achieve the fifth objective. Pearson's correlation between EC and CC found to be 0.977 (p-value = 0.000) which was expected to have caused by a third variable GIFA, hence, EC and CC were normalised for GIFA. EC per GIFA and CC per GIFA also demonstrated a strong positive correlation of 0.645 (p-value = 0.000) which implies that both EC and CC can be reduced at the same time. However, to investigate at a deeper level, EC per GIFA and CC per GIFA of individual elements were also analysed. Results suggested that most of the elements display a strong positive correlation between EC per GIFA and CC per GIFA, which implies that there is a strong relationship between EC and CC at both building and element levels.

7. Model Validation

7.1. Introduction

The validation of models is an important step in any model development. The model validation ensures that the model is a reasonable representation of the actual system with sufficient fidelity. In the research context, model validation implies the examination of the model fitness and the prediction performance of the models. The model validation process is illustrated in Figure 7.1. As introduced in the methodology chapter, R^2 and CV were used to assess models' fit and prediction accuracy respectively. Accordingly, all four selected models (EC per GIFA, EC, CC per GIFA, CC) listed below, were assessed based on the two metrics mentioned above.

- EC per GIFA Model: $\widehat{y_1} = 530.62 + 164.08x_{W:F} + 68.15x_B$
- EC Model: $\widehat{y_2'} = 530.62 + 164.08x_{W:F} + 68.15x_B$
- CC per GIFA Model: $\widehat{y_3} = 734.95 + 27.7x_{BH} + 1188.33x_{CR}$
- CC Model: $\hat{y}_4' = -0.099 + 0.965x_{GIFA}' + 0.013x_{BH}$

The R^2 reveals the percentage change in the dependent variable explained by the model (independent variables). Hence, higher R^2 implies lower uncertainty in the prediction. On the other hand, CV gives an estimation of the average deviation in the model predictions from observed values. Hence, higher CV implies higher uncertainty in the prediction. Therefore, a model with high R^2 and low CV is desirable. In this context, there is no established cut-off point for R^2 value, however, past research report R^2 ranging from 26.1% (Phaobunjong, 2002) to 99.8% (Kouskoulas and Koehn, 2005). Conversely, CV is expected to be less than 20% for early design stage estimating models as discussed in the methodology chapter (see, Section 5.8.4).

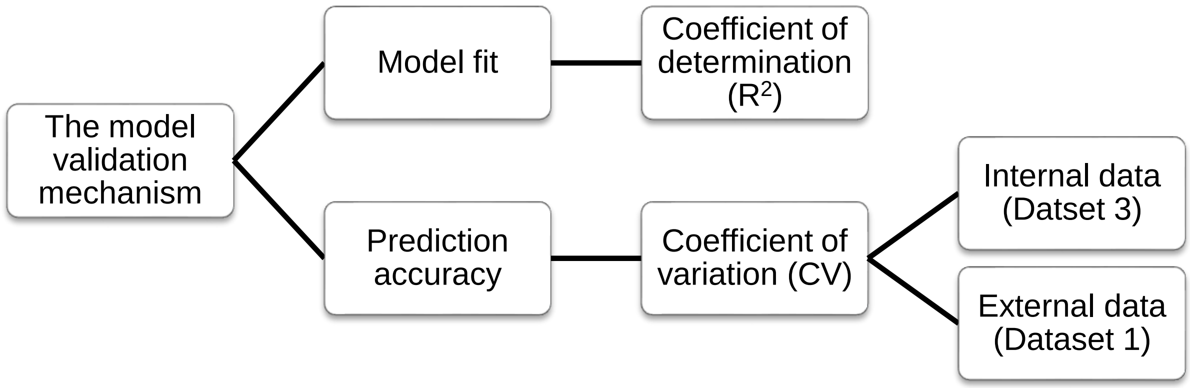


Figure 7.1: Model validation process

Internal validity is achieved when the model accurately captures causal relationship between variables while external validity is achieved when the study findings can be generalised to other similar settings (Saunders et al., 2009). Accordingly, internal validity was ensured by testing the prediction accuracy of the models for Dataset 3 which was used to develop the model. A good model prediction for internal data (data used to develop the model) suggests that the causal relationship between variables have been captured accurately. Similarly, external validity was ensured by testing the prediction accuracy of the models for Dataset 1, where an acceptable prediction accuracy affirms the generalisability of the models for data outside the model.

Further, the models were compared in pairs (EC per GIFA against EC Model and CC per GIFA against CC Model) and the better model from each pair was identified based on R^2 and CV for internal and external data. Additionally, the model performance in different storey clusters were analysed to identify the best predictive zone for the models.

7.2. Model Validation Dataset

The model validation data includes two datasets: internal and external data to validate the models internally and externally. Internal validation ensures effective model building while external validation displays the applicability of the developed models to real world problems.

7.2.1. Dataset for Internal Validation

Dataset 3, which was used to develop the models, was used for internal validation. Data comprising Dataset 3 are presented in Table 7.1. Further, the data were clustered into three groups such as 1-2 storeys, 3-5 storeys and more than 6 storeys to study the model performance in different input ranges. The adopted clustering of building data was inspired by the classification used in BCIS (see, BCIS average price analysis available at RICS (2016)).

Table 7.1: Data for internal validation (Dataset 3)

Building Code	GIFA (m^2)	No. of Storeys	Building Height	Wall to Floor ratio	Circulation Space Ratio	No. of Basements	CC (in £1000s)	EC (in tCO_2)	CC per GIFA (£/m^2)	EC per GIFA ($kgCO_2/m^2$)
1-2 Stories										
D3022	692	1	2.8	0.42	0.33	0	925	410	1,337	592
D3002	928	2	6.7	0.44	0.13	1	1,037	683	1,118	735
D3005	1,028	2	5.5	0.85	0.18	0	1,182	779	1,150	758
D3006	9,007	2	7.2	0.24	0.26	0	11,859	6,309	1,317	700
D3007	1,930	2	6.5	0.84	0.09	1	1,980	1,444	1,026	748
D3011	1,534	2	6.7	0.65	0.16	0	1,632	864	1,064	563
D3012	1,756	2	6.7	0.61	0.14	0	1,737	974	989	555
D3013	2,432	2	6.7	0.58	0.16	0	2,614	1,340	1,075	551
D3017	1,323	2	5.7	0.64	0.21	1	924	774	698	585
D3018	2,325	2	6.5	0.45	0.28	1	3,217	1,607	1,384	691
D3023	1,026	2	6.2	0.77	-	1	1,377	748	1,343	730
D3025	3,592	2	6.2	0.43	0.40	0	2,992	2,191	833	610
D3027	1,266	2	7.9	1.02	0.17	0	1,929	933	1,523	737
D3029	1,835	2	6.4	0.62	-	0	2,275	1,152	1,240	628
D3030	1,376	2	6.4	0.58	-	0	1,448	798	1,052	580
D3031	1,685	2	6.4	0.56	-	0	1,710	970	1,015	576
D3034	473	2	6.3	1.20	0.32	0	612	384	1,293	811
D3038	3,080	2	7.8	0.95	0.33	0	3,279	1,885	1,065	612
3-5 Stories										
D3004	2,412	3	9.9	0.78	0.28	1	2,697	1,499	1,118	621
D3001	3,987	3	10.7	0.84	-	0	8,966	2,921	2,249	733

Building Code	GIFA (m^2)	No. of Storeys	Building Height	Wall to Floor ratio	Circulation Space Ratio	No. of Basements	CC (in £1000s)	EC (in tCO_2)	CC per GIFA (£/m^2)	EC per GIFA ($kgCO_2/m^2$)
D3008	9,653	3	8.8	0.52	-	0	13,296	7,933	1,377	822
D3009	1,136	3	10.3	0.81	-	0	1,503	797	1,323	702
D3010	1,896	3	11.0	0.85	0.25	1	1,904	1,427	1,004	753
D3014	10,400	3	10.9	0.37	0.30	1	17,807	7,583	1,712	729
D3015	2,926	3	9.3	0.65	-	0	3,179	1,694	1,086	579
D3020	5,900	3	9.4	0.70	0.32	1	7,760	3,795	1,315	643
D3021	2,510	3	9.9	1.11	0.26	1	5,734	2,299	2,285	916
D3026	1,753	3	11.9	0.86	0.17	1	2,238	1,241	1,277	708
D3028	2,556	3	11.9	0.85	0.17	1	3,006	1,607	1,176	629
D3032	5,687	3	13.2	0.70	0.17	0	5,998	3,433	1,055	604
D3033	6,885	3	12.2	0.66	0.17	0	8,513	4,275	1,236	621
D3035	6,643	3	12.2	0.47	0.14	0	8,577	4,338	1,291	653
D3036	4,538	3	11.9	0.53	0.17	0	5,790	3,000	1,276	661
D3037	14,652	3	12.0	0.46	0.14	0	17,928	9,383	1,224	640
D3039	3,887	3	12.2	0.62	0.16	0	4,671	2,350	1,202	605
D3041	718	3	11.0	1.02	0.34	0	990	578	1,379	805
D3003	212	4	10.8	0.70	0.38	2	392	177	1,851	834
D3024	9,900	4	14.0	0.66	0.25	0	15,138	5,845	1,529	590
D3040	1,545	4	10.0	0.87	0.34	0	2,255	965	1,459	625
D3016	3,797	5	16.9	1.50	0.46	2	6,758	3,397	1,780	895
6+ Stories										
D3019	8,444	6	25.2	0.62	0.28	2	13,695	6,720	1,622	796

7.2.2. Dataset for External Validation

External validation is crucial to ensure the model can predict the CC and EC of new building designs at an acceptable accuracy range. Building data from Dataset 1 were used to validate the model. However, Dataset 1 underwent an initial screening process to identify inadequacies in the data, which is listed in Table 7.2. The dataset for external validation after eliminating inadequate data is presented in Table 7.3. The cost of the buildings has a base price index of the second quarter of 2010 and a location index of 1.0.

Table 7.2: Screening of Dataset 1

Building Code	No. of Storeys	Adequate/ Inadequate	Comments
D1001	18	Inadequate	Errors in measurements were detected.
D1002	8	Adequate	Out of the scope of the model prediction (>6 storeys).
D1003	3	Adequate	Out of the scope of the model prediction (>6 storeys).
D1004	7	Adequate	Out of the scope of the model prediction (>6 storeys).
D1005	16	Inadequate	Out of the scope of the model prediction (>6 storeys).
D1006	4	Adequate	Errors in measurements were detected and hence, the cost and carbon values were identified as anomalies.
D1007	10	Adequate	Out of the scope of the model prediction (>6 storeys).
D1008	4	Adequate	Within the scope of the model prediction.
D1009	3	Adequate	Within the scope of the model prediction.
D1010	3	Adequate	Within the scope of the model prediction.
D1011	13	Adequate	Out of the scope of the model prediction (>6 storeys).
D1012	12	Adequate	Out of the scope of the model prediction. (>6 storeys).
D1013	4	Adequate	Within the scope of the model prediction.

Table 7.3: External validation dataset from Dataset 1

Building Code	GIFA (m^2)	No. of Storeys	Building Height	Wall to Floor ratio	Circulation Space Ratio	No. of Basements	CC (in £1000s)	EC (in tCO_2)	CC per GIFA (£/m^2)	EC per GIFA ($kgCO_2/m^2$)
1-2 Stories										
None										
3-5 Stories										
D1003	2,859	3	6.5	0.64	0.27	1	2,085	1,693	729	592
D1009	3,262	3	10.7	0.49	0.20	0	2,055	1,577	630	483
D1010	4,959	3	11.5	0.48	0.26	0	4,452	2,944	898	594
D1008	3,289	4	14.1	0.66	0.35	0	2,465	1,788	716	519
D1013	2,374	4	10.8	0.61	0.42	1	1,311	1,102	552	464
6+ Stories										
D1004	15,120	7	27.5	0.31	0.35	2	10,147	8,826	671	584
D1002	11,320	8	29.7	0.30	0.34	2	6,416	6,799	567	601
D1007	22,288	10	40.0	0.43	0.18	0	14,455	13,256	649	595
D1012	21,300	12	48.0	0.37	0.25	1	13,088	9,945	614	467
D1011	21,300	13	56.0	0.40	0.28	1	15,921	13,252	747	622
D1005	63,246	16	63.4	0.31	0.39	2	61,082	46,977	966	743

Later, the remaining data were clustered into three groups as discussed before. The entire screening and clustering process is presented in Figure 7.2.

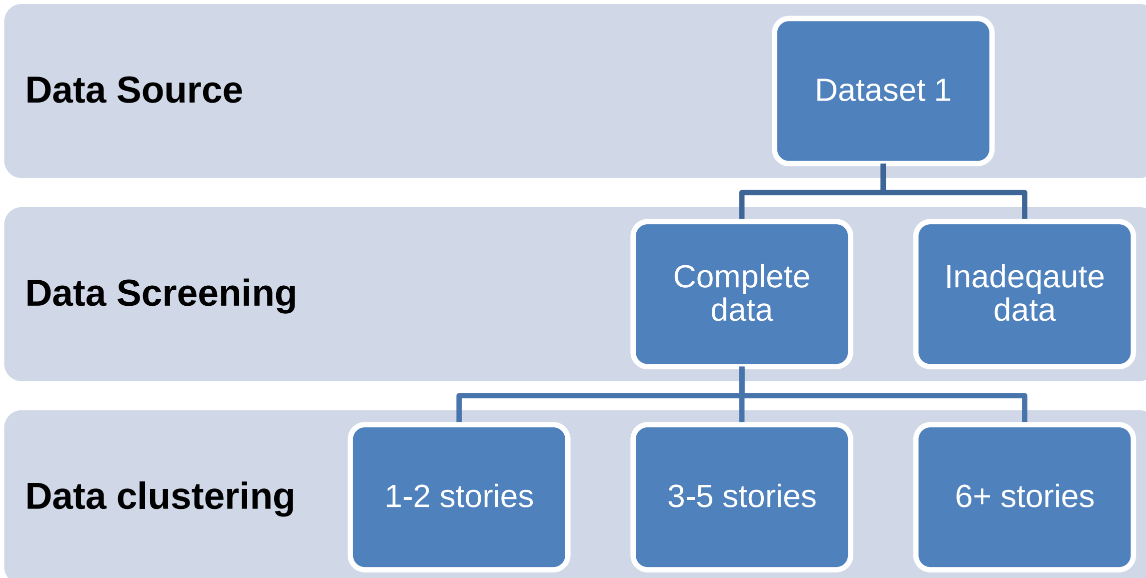


Figure 7.2: Screening and clustering of Dataset 1 for external validation

In addition to the above preparation, another issue in Dataset 1 needed to be addressed. As discussed in the data collection chapter (see, Section 6.5.1), estimates of Dataset 1 excludes Fittings, Furnishing and Equipment and Services due to lack of detailed measurements. However, the model will predict embodied carbon and capital cost including Fittings, Furnishing and Equipment and Services. Hence, one of the two adjustment options mentioned below needed to be applied:

1. Adjust the data to include Fittings, Furnishing and Equipment and Services embodied carbon (OR capital cost) in their estimates and compare with model predictions.

OR

2. Adjust the model prediction by subtracting Fittings, Furnishing and Equipment and Services embodied carbon (OR capital cost) from the prediction and compare with the estimates.

Option 1 was selected over Option 2 because deducting embodied carbon (OR capital cost) of Fittings, Furnishing and Equipment and Services from the estimates will leave out the error term of embodied carbon (OR capital cost) of Fittings, Furnishing and Equipment and Services in the prediction, leading to higher deviation or lower accuracy in prediction (see, Figure 7.3). Further, in both the cases the estimated residual will remain the same and the observed values (the estimates of Dataset 1) will be higher in option 1 than in the option 2. Hence, the first adjustment (adjustment to Dataset 1) will lead to lower deviation as the deviation is calculated as a percentage of the observed value. Therefore, the Option 1 of the adjustment was selected to maintain the deviation at a lower percentage. Consequently, benchmark values for capital cost and embodied carbon of Fittings, Furnishing and Equipment and Services were surveyed and both cost and EC were treated individually. The steps followed in arriving at cost benchmarks is discussed first, followed by the development of embodied carbon benchmarks.

$$\widehat{EC} = EC_{Fit\&Ser} + EC_{Others} + e_{Fit\&Ser} + e_{Others}$$

$\widehat{EC}$ - Predicted EC

$EC_{Fit\&Ser}$ - Actual EC of Fittings, Furnishing and Equipment and Services

EC_{Others} - Actual EC of the remaining building elements

$e_{Fit\&Ser}$ - Error term of the estimate of EC of Fittings, Furnishing and Equipment and Services

e_{Others} - Error term of the estimate of EC of the remaining building elements

When deducting Fittings, Furnishing and Equipment and Services EC from the prediction the predicted EC will be,

$$\widehat{EC} = EC_{Others} + e_{Fit\&Ser} + e_{Others}$$

Which implies, the error term of the prediction of EC of Fittings, Furnishing and Equipment and Services remains in the prediction.

Figure 7.3: The effect of applying Option 2 of the adjustment

Accordingly, benchmarks for Fittings, Furnishing and Equipment and Services CC were obtained from two data sources including Spon’s Mechanical and Electrical Services Price book 2014 (Davis Langdon Consultancy, 2014) and BCIS average prices (RICS, 2016). Cost benchmarks obtained from BCIS is presented in Table 7.4 and the cost benchmarks for services obtained from Spon’s price book are presented in Table 7.5. As can be seen from Table 7.4, cost benchmarks for services are given based on the number of storeys while the difference in quality levels of services is not reflected in the benchmarks. In addition, the sample used in arriving at the benchmarks for buildings over six (6) storeys is considerably small. Therefore, cost benchmarks for services obtained from BCIS was disregarded and the cost benchmarks available in Spon’s mechanical and electrical services price book 2014 (Davis Langdon Consultancy, 2014) were used to adjust Dataset 1 due to the fact that the quality index adopted in the study was developed from the classification followed in Spon’s price book (see, Section 6.3.2).

Table 7.4: Cost benchmarks obtained from BCIS

Type of building	Element	Category	Mean (£/m^2 GIFA)	Standard deviation	Sample size
Air-conditioned offices	Fittings and Furnishings	Generally	26	53	45
		1-2 storey	19	14	14
		3-5 storey	18	14	23
		6+ storey	64	131	7
	Services	Generally	398	143	50
		1-2 storey	331	157	17
		3-5 storey	425	133	25
		6+ storey	450	94	7
Non air-conditioned offices	Fittings and Furnishings	Generally	22	28	68
		1-2 storey	27	36	36
		3-5 storey	17	15	29
		6+ storey	17	13	4
	Services	Generally	344	168	70
		1-2 storey	315	147	37
		3-5 storey	360	186	30
		6+ storey	573	76	4

Table 7.5: Cost benchmarks obtained from Spon's price book

Quality level of services	Minimum (£/m^2 GIFA)	Maximum (£/m^2 GIFA)	Average (£/m^2 GIFA)
Services without BMS - Non A/C	255	310	283
Services without BMS - A/C	425	515	470
Services with BMS - Non A/C	275	335	305
Services with BMS - A/C	445	540	493

In case of Fittings, Furnishing and Equipment, benchmarks from BCIS were used. General average prices for air-conditioned and non-air-conditioned buildings were used instead of the average prices for different storey clusters due to the lower sample size and higher standard deviation, especially in the values of buildings over six (6) storeys.

Average prices obtained from BCIS has a base price index of the second quarter of 2010 (price index - 218) and a location index of 1.0 which is similar to the base of Dataset 1. Average prices given in Spon's price book have a base price index of the first quarter of 2013 and a location index of 1.03. Hence, all average prices were adjusted to the first quarter of 2016 (price index - 276) and a location index of 1.0 which is the base of the data used to develop the models. Accordingly, the updated cost benchmarks of Fittings, Furnishing and Equipment and Services to the respective base are presented in Table 7.6.

Table 7.6: Average prices updated to the model base (Price index of 1Q 2016 and a location index of 1.0)

Items	**Average (£/m^2 GIFA)**	**Updated Price (£/m^2 GIFA)**
Services without BMS - Non A/C	283	324
Services without BMS - A/C	470	538
Services with BMS - Non A/C	305	349
Services with BMS - A/C	493	564
Fittings and Furnishings (air-conditioned)	26	33
Fittings and Furnishings (Non air-conditioned)	22	28

On the other hand, lack of EC benchmarks lead to the use of average EC values developed from Dataset 2 (see, Table 5.19 in Section 5.6.2) presented in Table 7.7. No adjustments were required for EC values because, the initial EC (cradle-to-gate) depends on the process of manufacturing and it was assumed that no major change in the manufacturing process of materials has occurred as discussed in the data collection chapter (see, Section 6.5.2).

Table 7.7: Average EC values

Items	Average ($kgCO_2/m^2$ GIFA)
Services without BMS - Non A/C	134
Services without BMS - A/C	164
Services with BMS - Non A/C	148
Services with BMS - A/C	178
Fittings and Furnishings	1

Subsequently, Dataset 1 was modified by updating to the model base (price index of the first quarter of 2016 and a location index of 1.0) and by adding Fittings and Services EC and CC to the initial estimates. The modified dataset for external validation is presented in Table 7.8.

Table 7.8: External validation dataset from Dataset 1 adjustment for Fittings and Services (Price index of 1Q 2016 and a location index of 1.0)

Building Code	Quality level	CC per GIFA (£/m^2)	EC per GIFA ($kgCO_2/m^2$)
D1002	Air-conditioned with BMS	597	179
D1003	Air-conditioned with BMS	597	179
D1004	Air-conditioned with BMS	597	179
D1005	Air-conditioned with BMS	597	179
D1007	Air-conditioned with BMS	597	179
D1008	Air-conditioned without BMS	571	165
D1009	Air-conditioned without BMS	571	165
D1010	Air-conditioned without BMS	571	165
D1011	Air-conditioned with BMS	597	179
D1012	Air-conditioned with BMS	597	179
D1013	Air-conditioned without BMS	571	165

7.3. Model Validation Outcome

As discussed in the model validation process, CV is used to assess the accuracy of prediction of the models and R^2 is used to assess the closeness of fit of the models. Hence, R^2, CV of internal and external data of each model have been presented herein along with the graphical representation of residuals for each model.

7.3.1. EC per GIFA Model

a) Closeness of fit

EC per GIFA model has an R^2 value of 48.1%, which is satisfactory. 48.1% of the change in the EC per GIFA is explained by Wall to Floor ratio and no. of basements. The remaining change in the dependent variable can be expected to be explained by other design variables, which were not considered in the study.

b) Prediction performance with internal data

The CV of the model was found to be 10.65%, which is within the desired CV range for early stage estimating. The difference in the estimates to that with the actual EC per m^2 GIFA ranges from -25% to 20% for the overall sample. Except for one building, predictions of all buildings lie within the acceptable ±20% range. Since, the model explains the change in the dependent variable attributable to only Wall to Floor ratio and no. of basements, the observed variations in the estimates can be expected to be attributable to the other design variables, which were not regressed in the model due to insufficient statistical evidence. Figure 7.4 presents the scatterplot of predicted and observed EC per GIFA values, which follow a vague linear relationship. Further, deviations in the predictions were plotted against Wall to Floor ratio and no. of basements as shown in Figure 7.5 and Figure 7.6 (the acceptable deviation region is marked with broken lines in the graphs). Residuals lie between -206 and 119 $kgCO_2/m^2$ GIFA. Residuals of the majority of the buildings lie between -120 and 120 $kgCO_2/m^2$ GIFA. The building with the highest residual has a lower Wall to Floor ratio and has no basements. Further, it can also be noticed that the residuals become smaller as the values for the Wall to Floor ratio and the number of basements increase.

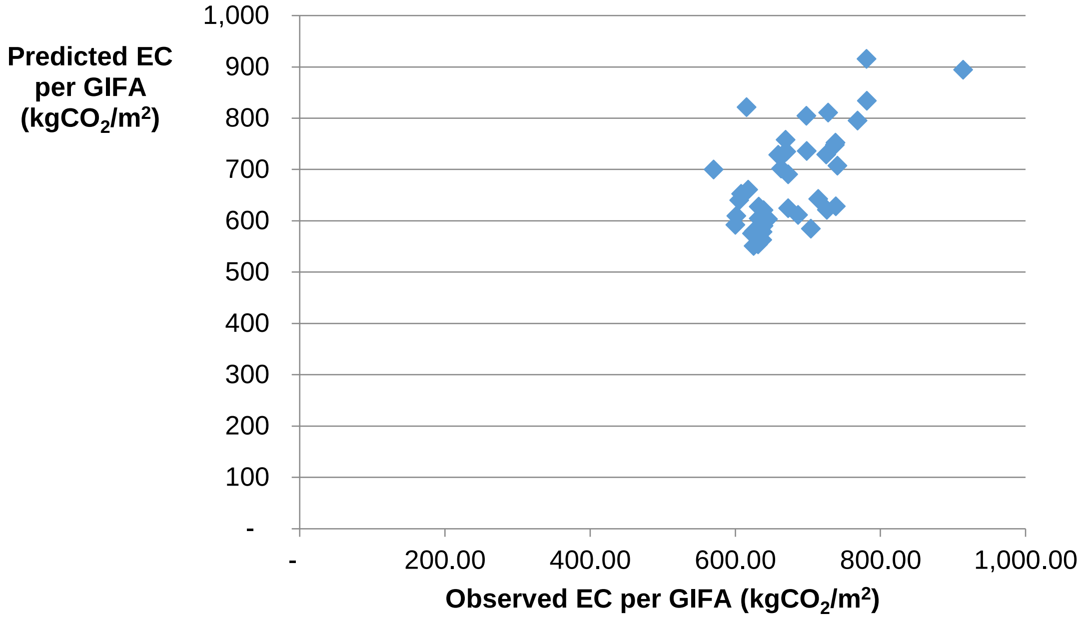


Figure 7.4: Scatterplot of predicted vs. observed EC per GIFA values – internal data

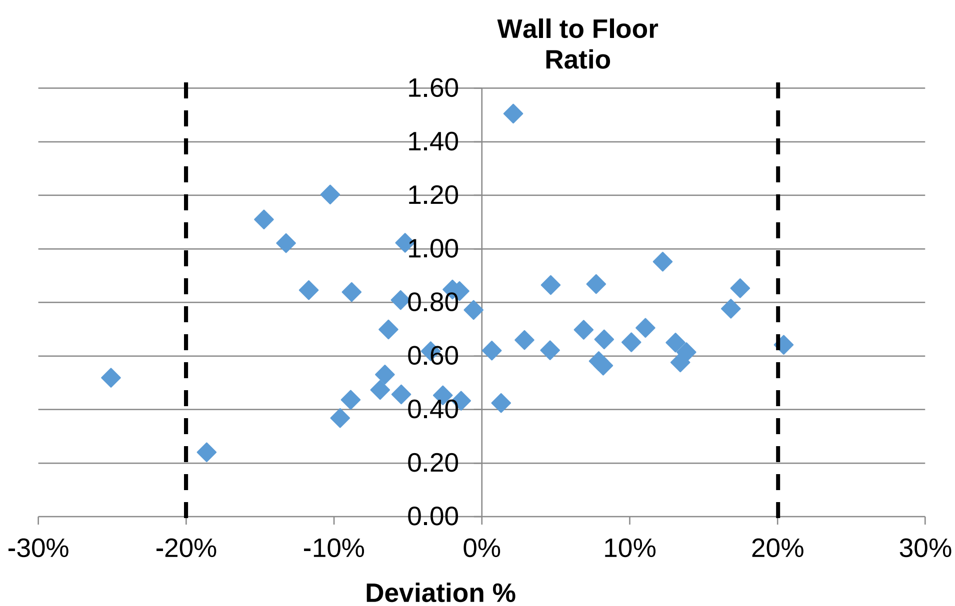


Figure 7.5: Mapping the model prediction deviation against the Wall to Floor ratio – Internal data

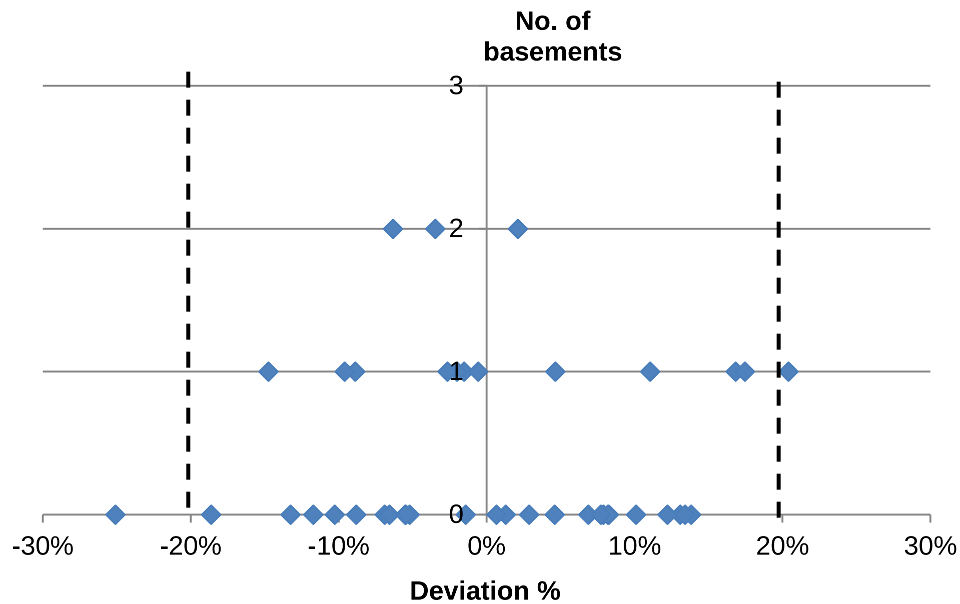


Figure 7.6: Mapping the model prediction deviations against the no. of basements – Internal data

In addition, the performance of the model for different building height clusters was also examined as explained in Section 8.3.1. Figure 7.7 illustrates the model performance at different clusters. Accordingly, the model prediction lies within the 20% margin for all buildings except for one building in the 3-5 storey cluster. However, the prediction performance for 6+ storeys clustered cannot be certainly ascertained as there is only one building in the sample with 6 storeys. Further, the model seems to predict closer to the observed values in 1-2 storey cluster in comparison to 3-5 storey clusters. In addition, the accuracy ranges from -19% to 20% with a CV of 10.4% in 1-2 storey cluster and -25% to 17% with a CV of 11.2% in 3-5 storey cluster. This implies that the model performs at its best in 1-2 storey cluster.

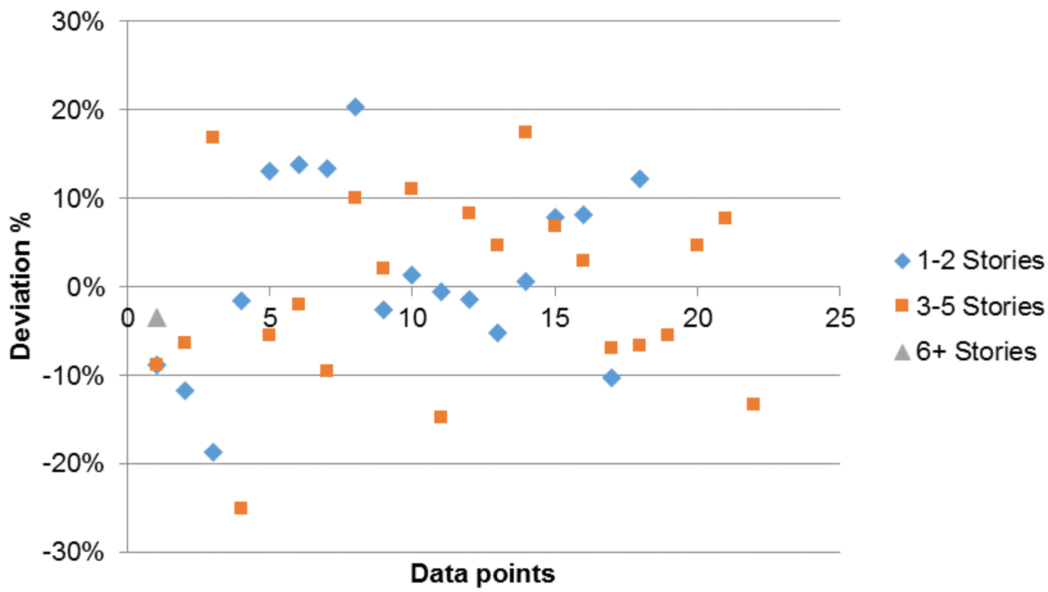


Figure 7.7: The EC per GIFA model prediction at different clusters – Internal data

c) Prediction performance with external data

The model predictions were plotted against the observed EC per GIFA (adjusted Dataset 1) in a graph as presented in Figure 7.8. The spread of deviation in the predictions over the wall to floor ratio and the number of basements are presented in Figure 7.9 and Figure 7.11 respectively. According to the graphs, it is evident that most of the predictions are within the acceptable range, with residuals ranging from -205 $kgCO_2/m^2$ to 70 $kgCO_2/m^2$ (-22% to 11% deviation from the observed values). The overall CV for external data was found to be 11%, which is satisfactory. Table 7.9 presents the model predictions, observed values and the residuals for the external data. Further, the analysis of different storey clusters illustrated in Figure 7.12 reveals that the identified highest deviation is due to the buildings with more than six (6) storeys. The model performs well with 3-5 storey cluster compared to the 6+ storey cluster. This highlights the scope of the developed model. The lack of data points in 1-2 storey cluster in the external data makes it impossible to comment on the model performance within this cluster for external data.

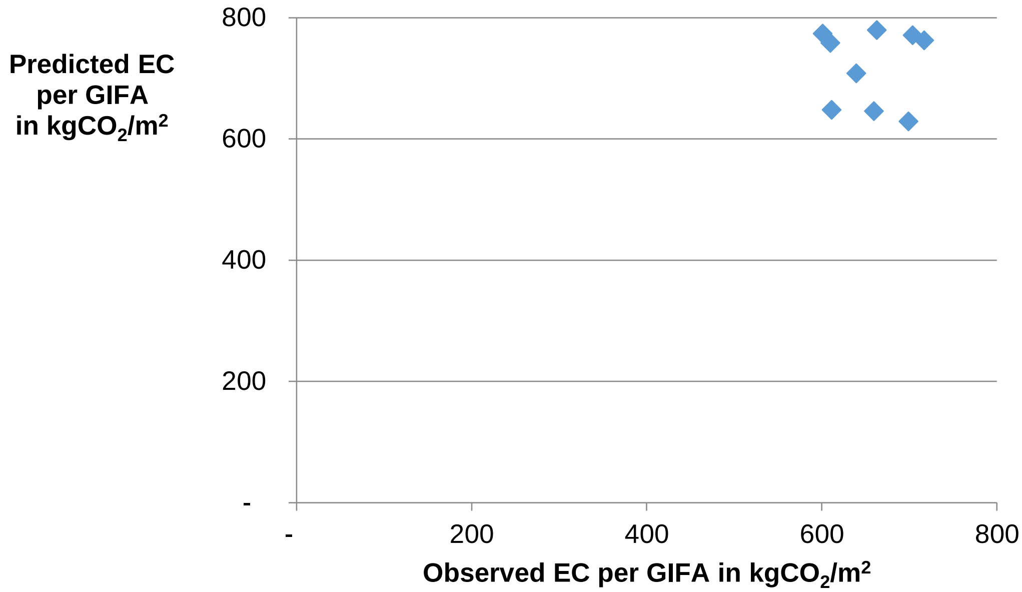


Figure 7.8: Scatterplot of predicted vs. observed EC per GIFA values – external data

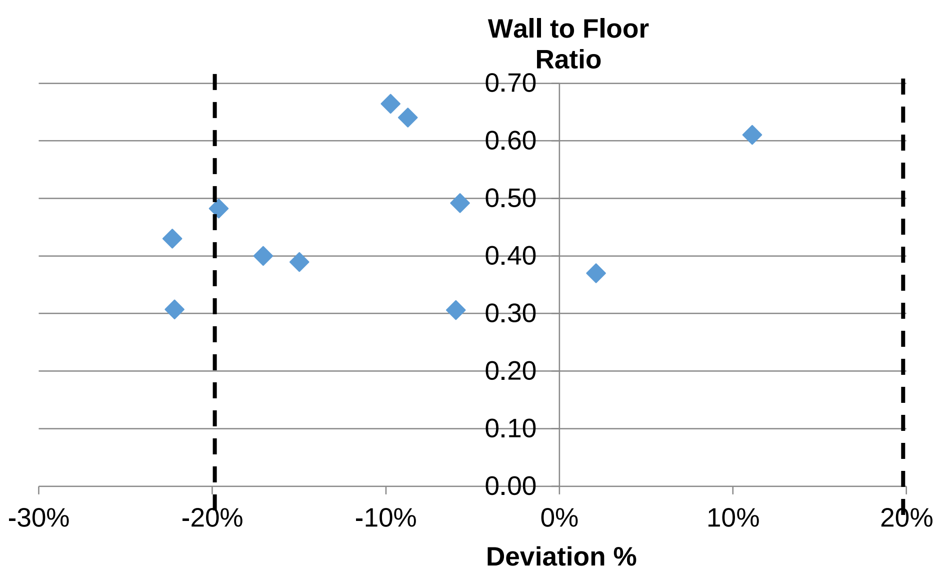


Figure 7.9: Mapping the model prediction deviations against the Wall to Floor ratio – External data

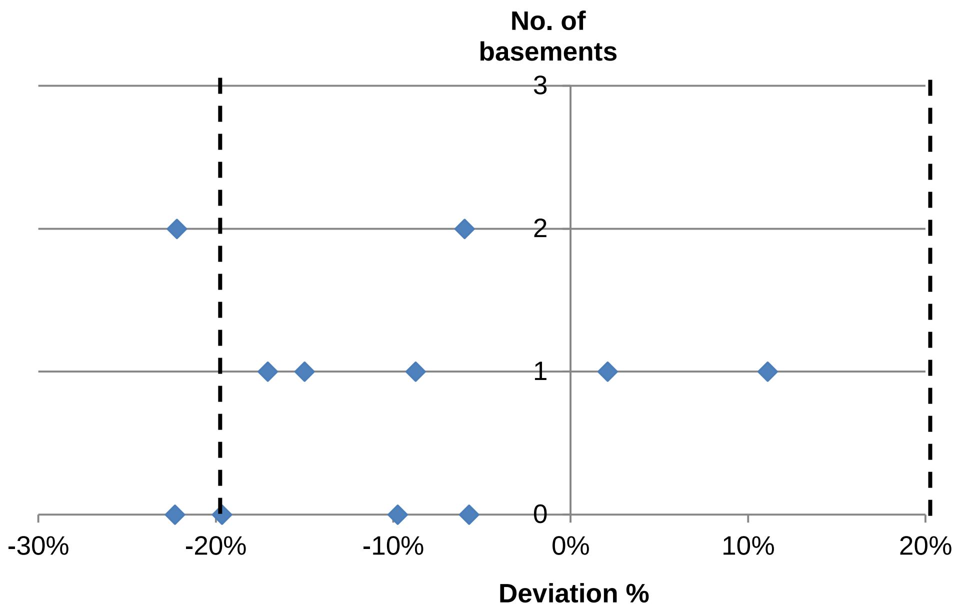


Figure 7.11: Mapping the model prediction deviations against the no. of basements – External data

Table 7.9: Calculation of the EC per GIFA model deviation for the external data

Building ID	Predicted ($kgCO_2/m^2$)	Observed ($kgCO_2/m^2$)	Residual ($kgCO_2/m^2$)	Deviation [(Predicted-Observed)/ Observed]
D1002	663	780	-117	-15%
D1003	704	771	- 67	-9%
D1004	717	763	- 46	-6%
D1005	717	922	-205	-22%
D1007	601	774	-173	-22%
D1008	640	709	- 69	-10%
D1009	611	648	- 37	-6%
D1010	610	759	-149	-20%
D1011	664	801	-137	-17%
D1012	659	646	14	2%
D1013	699	629	70	11%

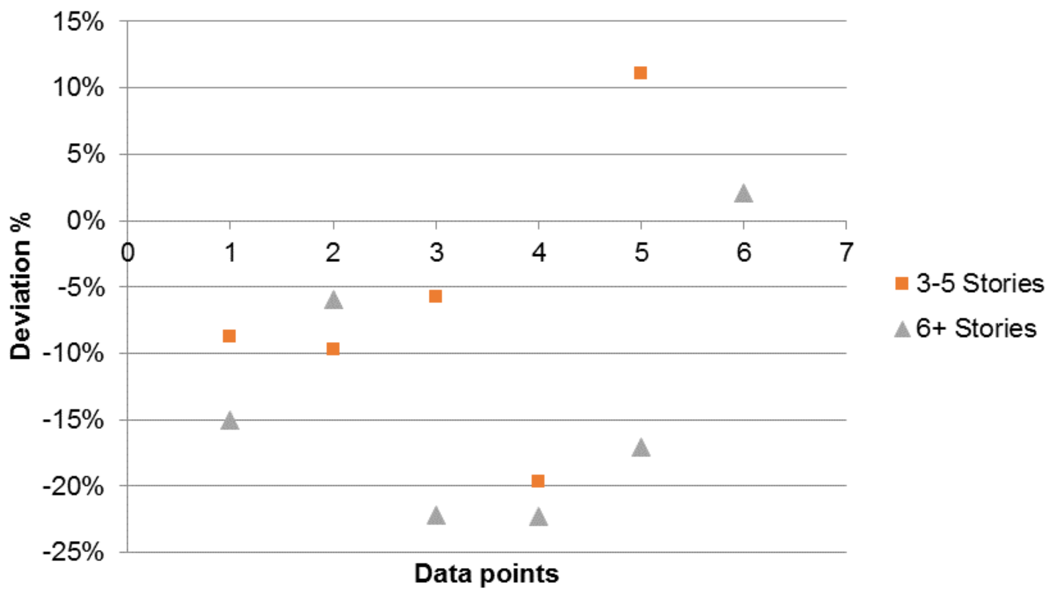


Figure 7.12: The EC per GIFA model prediction at different clusters – External data

7.3.2. EC Model

a) Closeness of fit

EC model had an R^2 of 98.3%, which is extremely high and suggests that change in the EC is explained by GIFA and no. of basements. In addition, the R^2 of the model suggests that the influence of other design variables is almost negligible.

b) Prediction performance with internal data

CV of the model was found to be 25.7%, which is slightly above the maximum threshold (±20%) set for an early stage estimate. The model predictions deviate from the observed values ranging from -0.1% to 61%. Even though the model fit was impressive, the prediction accuracy of the model demonstrates a problem. This could have been caused due to the log transformation. The predicted values were plotted against the observed values, which are illustrated in Figure 7.14, which displays a perfect linear relationship. Further, the deviation of the model predictions against GIFA and the number of basements is presented in Figure 7.13

and Figure 7.15. Only a few predictions fall within the acceptable deviation region, which is ±20%. Further, the deviation ranges from a lower value to the highest between 0 and 4,000m^2 of GIFA (see, Figure 7.13) and the deviation was higher for the buildings with no basements and with one basement compared to the building with two basements (see, Figure 7.15). This is mainly because the prediction was based only on the GIFA when there are no basements. Nevertheless, even with an R^2 value of 98.3% this deviation is unacceptable.

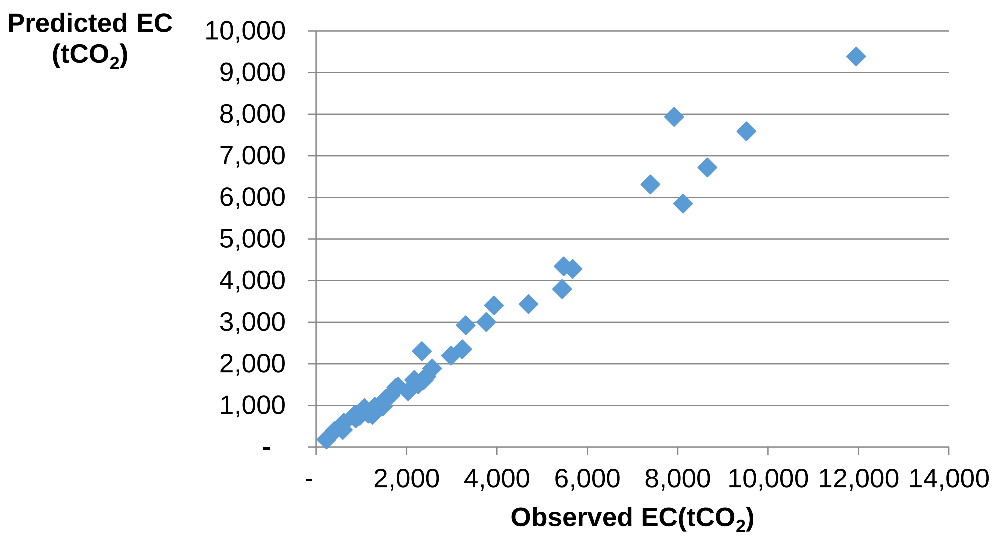


Figure 7.14: Scatterplot of predicted vs. observed EC values – internal data

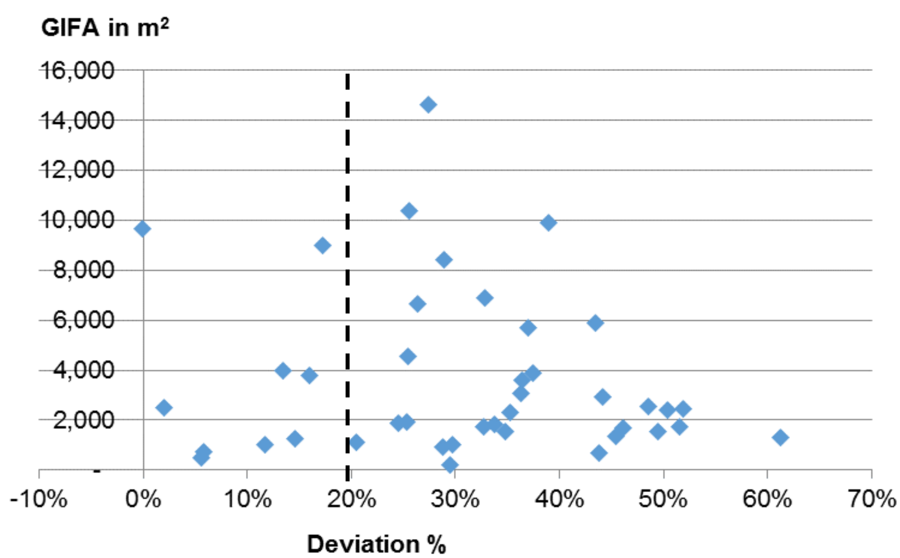


Figure 7.13: Mapping the model prediction deviation against GIFA – Internal data

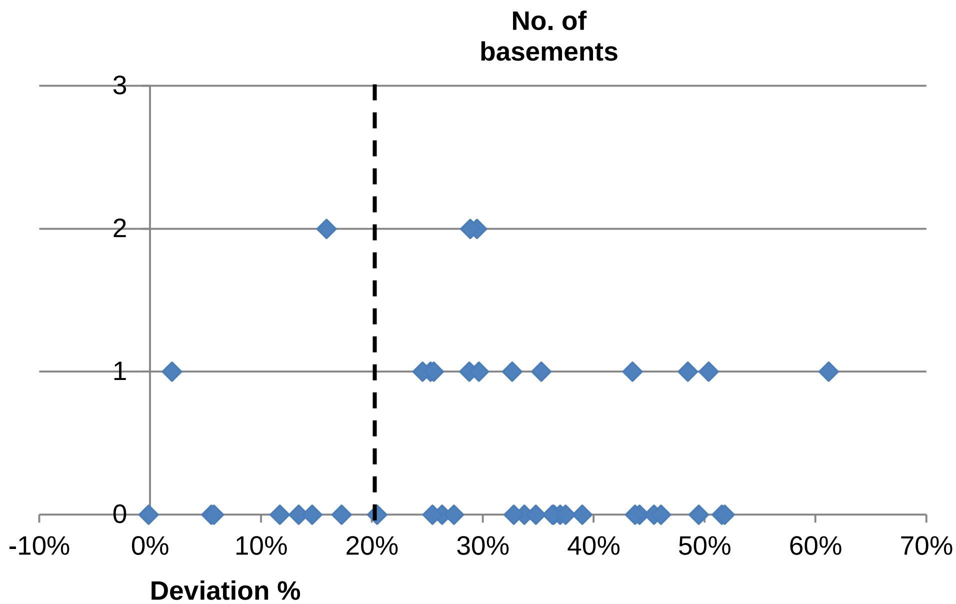


Figure 7.15: Mapping the model prediction deviation against the no. of basements – Internal data

Figure 7.16 illustrates the model performance at different clusters with internal data. Both 1-2 and 3-5 storey clusters show similar deviations while the highest deviation was found in the 1-2 storey cluster.

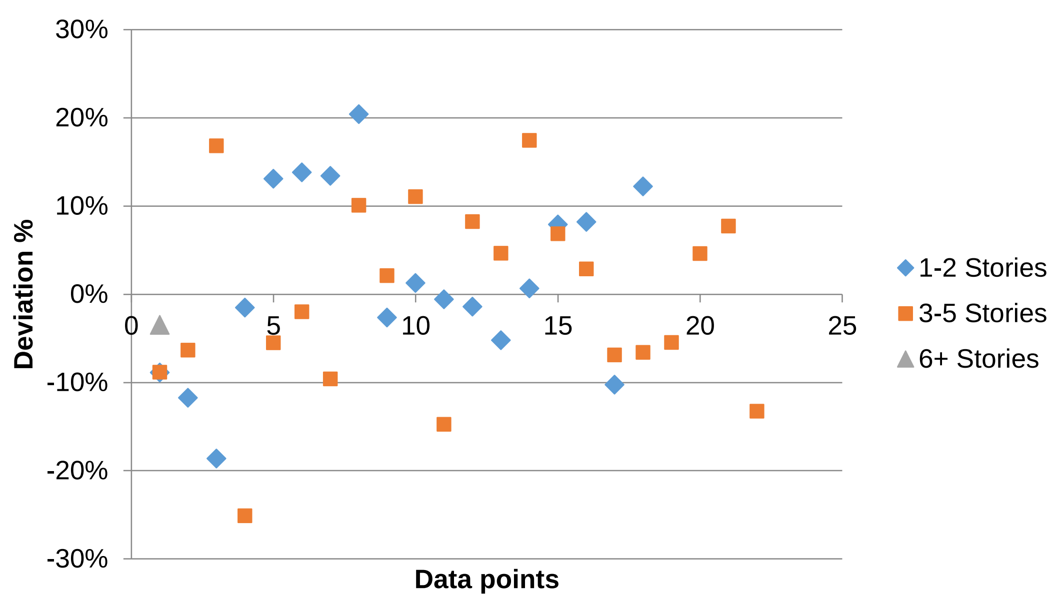


Figure 7.16: The EC model prediction at different clusters – Internal data

c) Prediction performance with external data

Predicted EC and observed EC of external data were plotted in a scatterplot and presented in Figure 7.17. The deviation in predictions against GIFA and the number of basements are presented in Figure 7.18 and Figure 7.19 respectively. Approximately half of the predictions fell within the acceptable deviation region. Further, a deviation within the 20% range and above 20% was noticed for similar input values of GIFA and the number of basements. Table 7.10 lists the deviation calculation of the external data, which ranges from 8% to 49%. A deviation beyond 20% is considered unacceptable even for an early stage prediction model. In addition, predictions were examined based on the storey clusters, which is presented in Figure 7.20. In contrast to the internal data, the model predicts better in the 6+ storey cluster than the 3-5 storey cluster with external data.

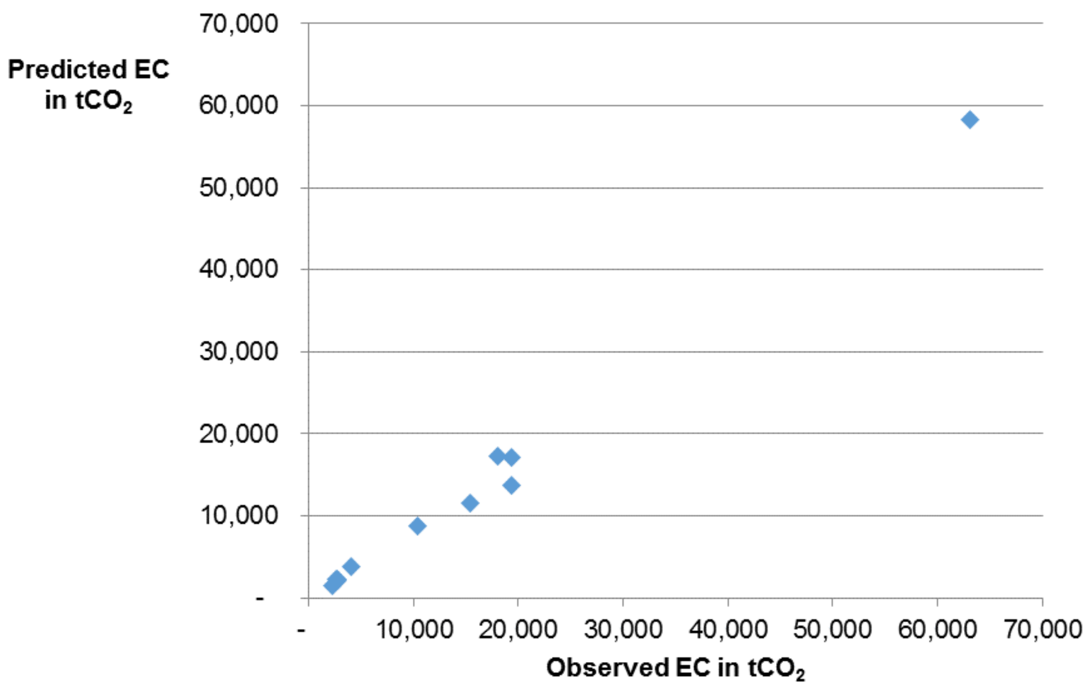


Figure 7.17: Scatterplot of predicted Vs. observed EC values – External data

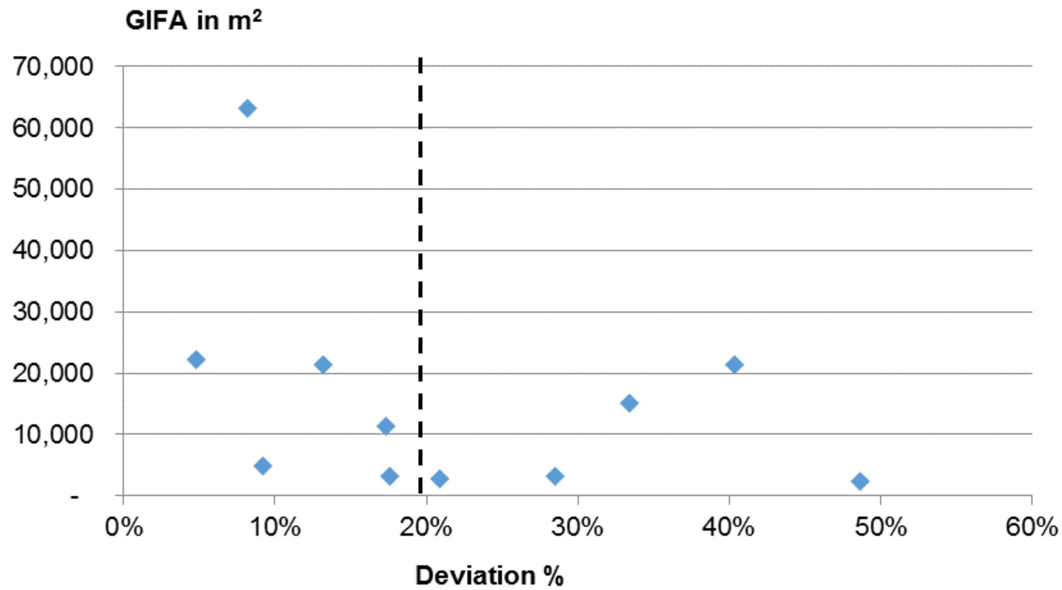


Figure 7.18: Mapping the model prediction deviation against GIFA – External data

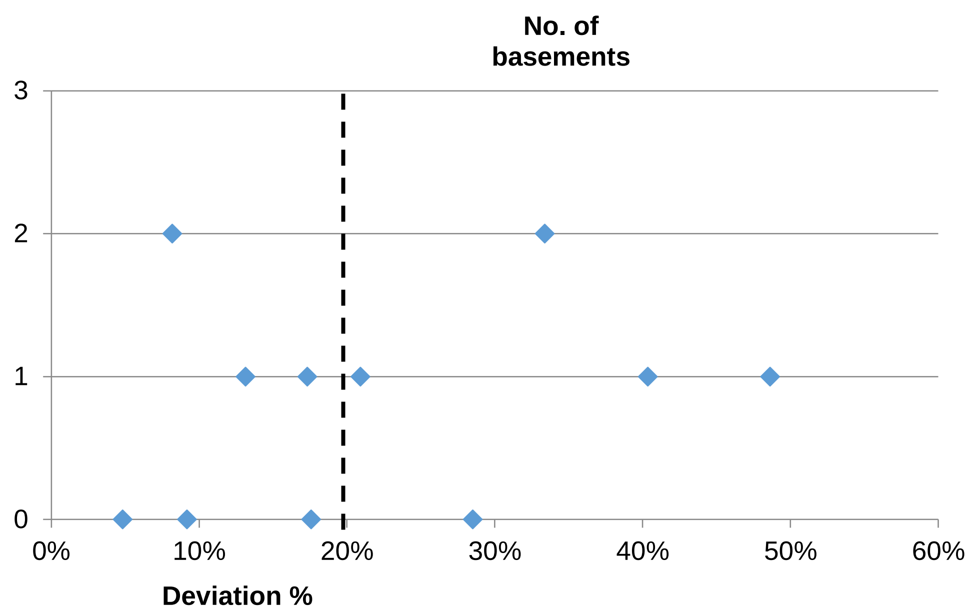


Figure 7.19: Mapping the model prediction deviation against the no. of basements – Internal data

Table 7.10: Calculation of the EC model deviation for the external data

Building ID	Predicted (tCO_2)	Observed (tCO_2)	Residual (tCO_2)	Deviation [(Predicted-Observed)/Observed]
D1002	10,354	8,825	1,528	17%
D1003	2,666	2,205	461	21%
D1004	15,383	11,532	3,851	33%
D1005	63,069	58,298	4,771	8%
D1007	18,080	17,246	834	5%
D1008	2,740	2,330	410	18%
D1009	2,718	2,115	603	29%
D1010	4,108	3,763	345	9%
D1011	19,310	17,065	2,245	13%
D1012	19,310	13,757	5,553	40%
D1013	2,219	1,493	726	49%

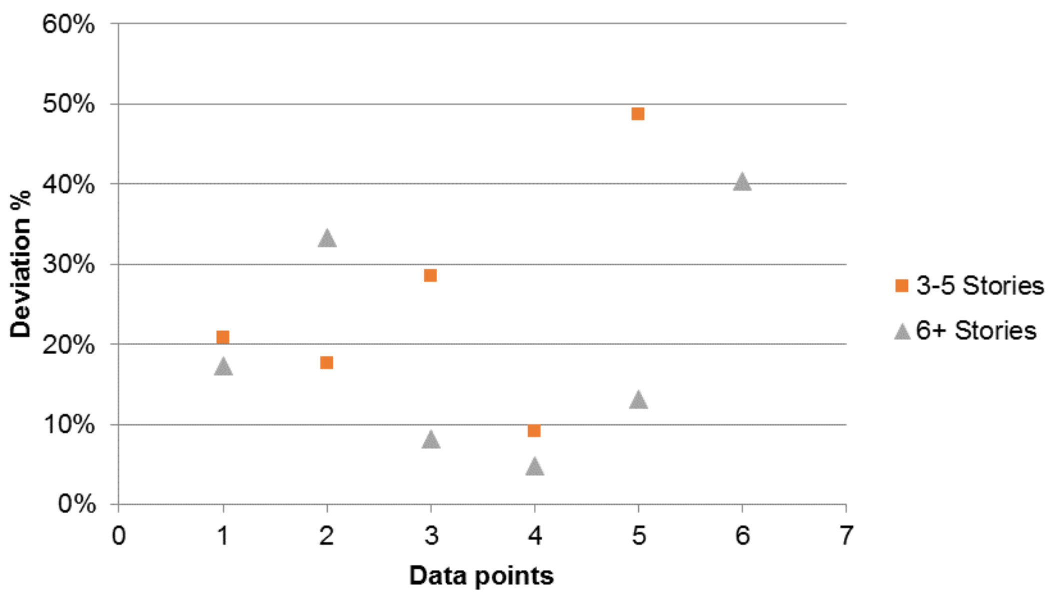


Figure 7.20: The model prediction at different clusters – External data

7.3.3. CC per GIFA Model

a) Closeness of fit

The model fit was found to be 23.7%, which displays a poorly fitted model. Only 23.7% of the variation in CC per GIFA is explained by building height and circulation space ratio. However, other influential design variables like building height, circulation space, finishes and services qualities (according to the literature) were found to be statistically insignificant in the study.

b) Prediction performance with internal data

The CV of the model was 20.3%, which is within the desired CV range for early stage estimating, while, the CV of the CC per GIFA model was lower than the CV of the EC per GIFA model. The scatterplot of predicted and observed CC per GIFA values is presented in Figure 7.21, which follows a vague linear relationship similar to EC per GIFA plot (Figure 7.4). Further, deviations in the model predictions were plotted against building height and circulation space ratio as shown in Figure 7.22 and Figure 7.24. Accordingly, most of the predictions lie within the acceptable accuracy range while three predictions showed high deviations, which are circled in the diagrams (see, Figure 7.22 and Figure 7.24). The storey cluster analysis illustrated in Figure 7.23 highlights that the exceptional three predictions fall within both 1-2 and 3-5 storey clusters. Hence, regression analysis was executed again after eliminating the identified three extreme data points.

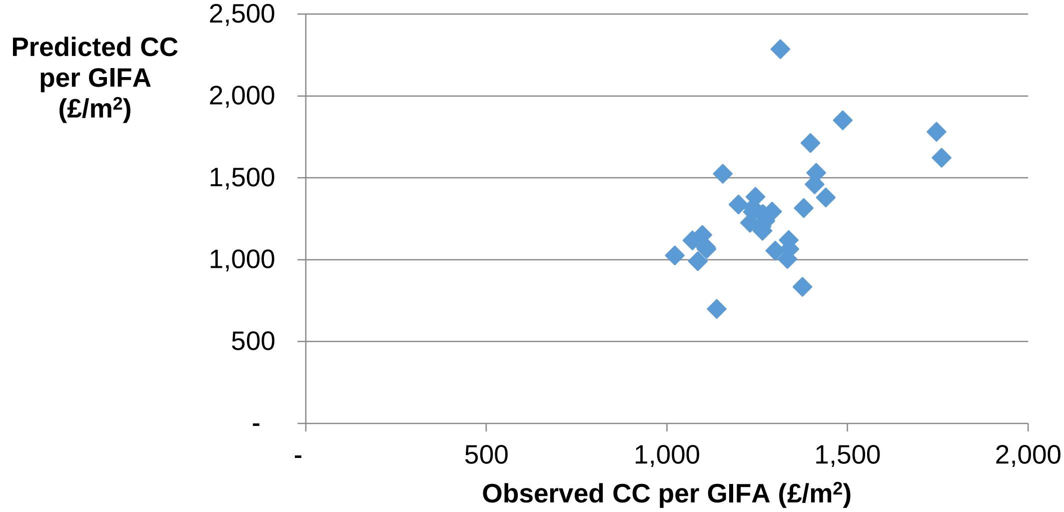


Figure 7.21: Scatterplot of predicted vs. observed CC per GIFA values – internal data

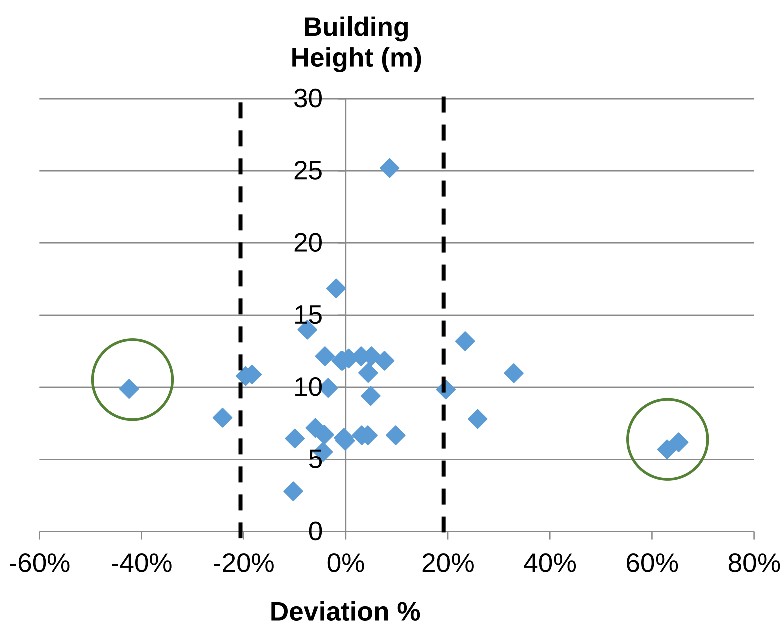


Figure 7.22: Mapping the model prediction deviation against building height – Internal data

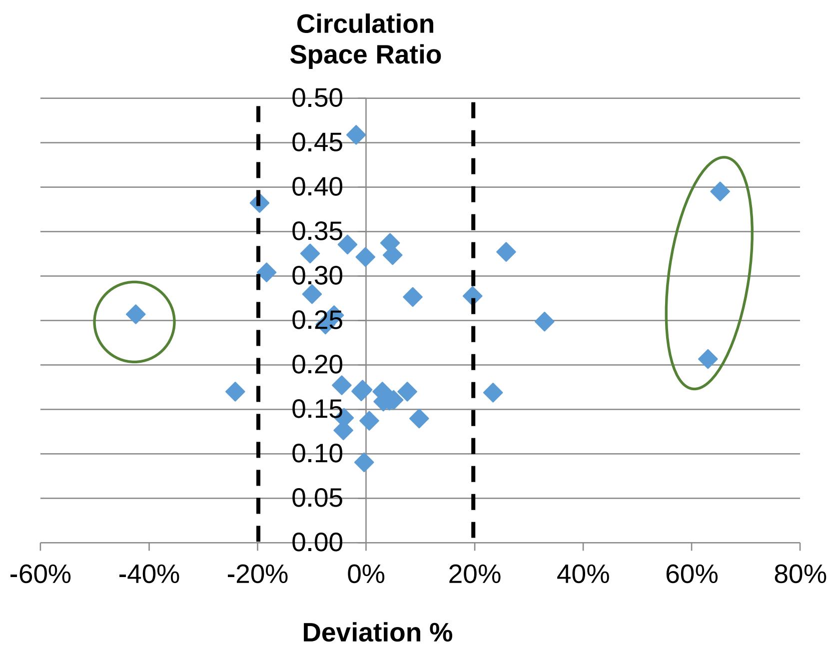


Figure 7.24: Spread of residuals over circulation space ratio

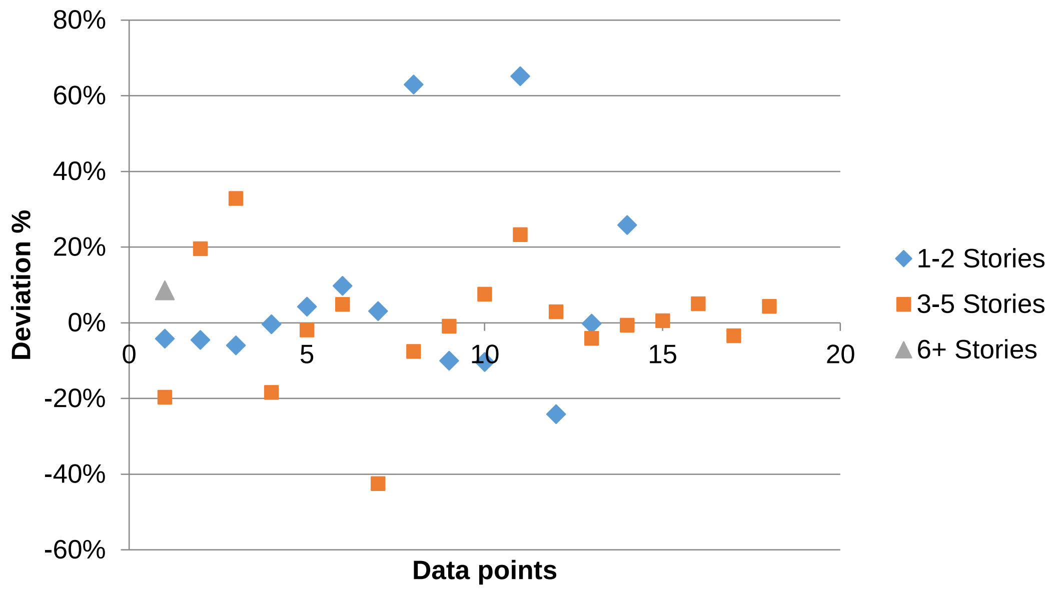


Figure 7.23: The CC per GIFA model prediction at different clusters – Internal data

7.3.4. CC per GIFA Model Recalibrated

Model summary, ANOVA table and the summary of the coefficient of variables are presented in Table 7.11, Table 7.12 and Table 7.13 respectively. The new model also identifies building height and circulation space ratio as significant independent variables. Even though the Model 4 presented in Table 7.11 has better adjusted R^2 value and lower standard error than all other models, the wall to floor ratio in Model 4 is not statistically significant. Further, Model 5 has the highest F statistics and both the variables are statistically significant in the model. In addition, VIF is within the acceptable limit for Model 5 (see, Table 7.13) assuring no multicollinearity between independent variables. Residuals of Model 5 follow a standard normal distribution (see, Figure 7.25) and are randomly distributed (see, Figure 7.26), conforming to the assumption of homoscedasticity. The Durbin-Watson score was 2.118 which is greater than $d_{U,\alpha}$ ($d_{U,\alpha,}$ =1.57) indicating no positive autocorrelation among residuals. Similarly, 4-d (4 – 2.118 = 1.882) is also greater than $d_{U,\alpha}$ confirms no negative autocorrelation, which meets the regression assumptions discussed in the methodology chapter (see, Section 5.8.3). Hence, Model 5 is selected for the validation.

Table 7.11: Model summary – CC per GIFA Run 3

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Independent Variables
1	.771	.595	.490	165.522	Building height, wall to floor ratio, circulation ratio, no. of basements, finishes index, services index
2	.771	.595	.511	162.073	Building height, wall to floor ratio, circulation ratio, no. of basements, services index
3	.767	.588	.522	160.238	Building height, wall to floor ratio, circulation ratio, services index
4	.763	.582	.533	158.234	Building height, wall to floor ratio, circulation ratio
5	.735	.541	.506	162.746	Building height, circulation ratio

Table 7.12: ANOVA table – CC per GIFA Run 3

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	926289.172	6	154381.529	5.635	.001
	Residual	630140.728	23	27397.423		
	Total	1556429.900	29			
2	Regression	926004.381	5	185200.876	7.051	.000
	Residual	630425.519	24	26267.730		
	Total	1556429.900	29			
3	Regression	914524.216	4	228631.054	8.904	.000
	Residual	641905.685	25	25676.227		
	Total	1556429.900	29			
4	Regression	905438.233	3	301812.744	12.054	.000
	Residual	650991.668	26	25038.141		
	Total	1556429.900	29			
5	Regression	841301.117	2	420650.558	15.882	.000[f]
	Residual	715128.784	27	26486.251		
	Total	1556429.900	29			

Table 7.13: Coefficient of the variables – CC per GIFA Run 3

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	895.113	372.463		2.403	.025		
	Building Height	15.799	8.929	.285	1.769	.090	.678	1.476
	Wall to Floor	-233.376	146.222	-.265	-1.596	.124	.638	1.567
	Circulation Ratio	947.062	246.354	.708	3.844	.001	.518	1.930
	Basements	35.947	56.856	.106	.632	.533	.628	1.593
	Finishes Index	19.948	195.654	.016	.102	.920	.740	1.351
	Service Index	18.829	28.067	.101	.671	.509	.782	1.279
2	(Constant)	930.996	119.356		7.800	.000		
	Building Height	16.061	8.372	.290	1.919	.067	.739	1.353
	Wall to Floor	-228.737	136.067	-.260	-1.681	.106	.707	1.415
	Circulation Ratio	938.045	225.144	.702	4.166	.000	.595	1.681
	Basements	36.582	55.336	.108	.661	.515	.635	1.574
	Service Index	18.815	27.482	.101	.685	.500	.782	1.279
3	(Constant)	914.300	115.332		7.928	.000		
	Building Height	18.587	7.365	.335	2.524	.018	.934	1.071
	Wall to Floor	-227.077	134.503	-.258	-1.688	.104	.707	1.414
	Circulation Ratio	983.996	211.722	.736	4.648	.000	.658	1.521
	Service Index	15.962	26.834	.085	.595	.557	.802	1.247
4	(Constant)	941.364	104.654		8.995	.000		
	Building Height	19.321	7.170	.349	2.695	.012	.961	1.041
	Wall to Floor	-200.651	125.368	-.228	-1.600	.122	.794	1.260
	Circulation Ratio	933.046	191.203	.698	4.880	.000	.786	1.272
5	(Constant)	850.167	90.285		9.417	.000		
	Building Height	18.373	7.349	.332	2.500	.019	.967	1.034
	Circulation Ratio	800.646	177.296	.599	4.516	.000	.967	1.034

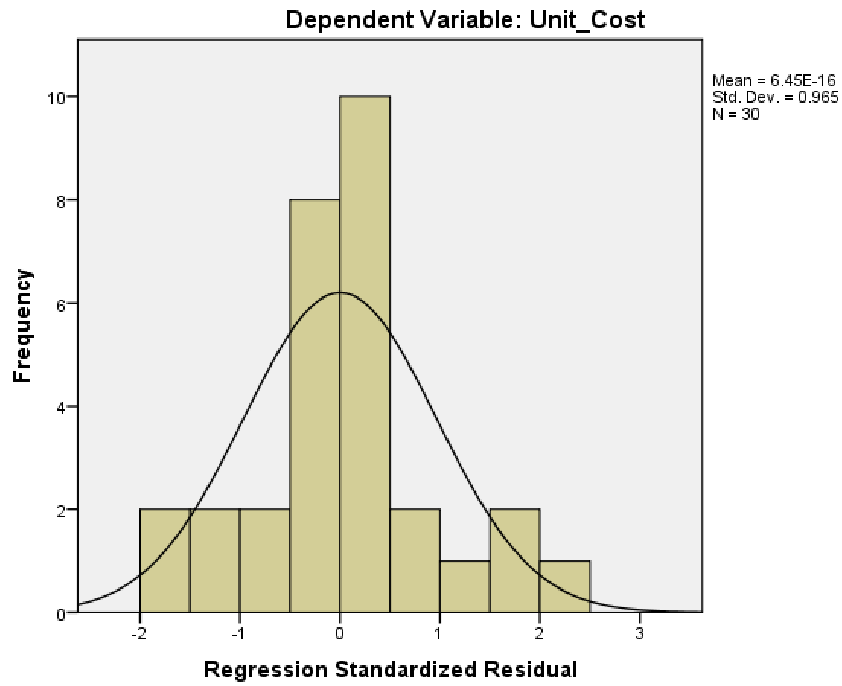


Figure 7.25: Histogram of standardised residual of the regression – CC per GIFA Run 3

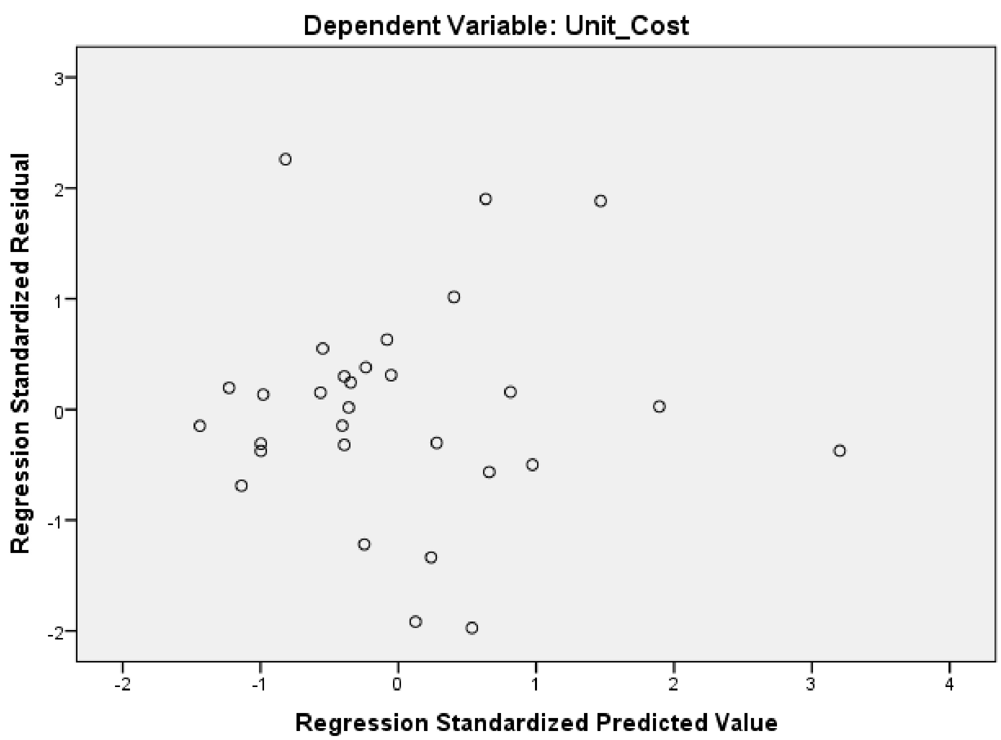


Figure 7.26: Scatterplot of standardised predicted value vs. standardised residuals of regression – CC per GIFA Run 3

The new CC per GIFA model can be presented as follows:

$$\hat{y} = 850.17 + 18.37x_{BH} + 800.65x_{CR}$$

a) Closeness of fit

The model fit was found to be 50.6%, which has improved immensely from the previous model after eliminating the identified three data points. 50.6% of the variation in CC per GIFA is explained by building height and circulation space ratio in the model, which is an acceptable model fit. Similar to the previous model, other design variables were not identified as statistically significant in predicting CC per GIFA.

b) Prediction performance with internal data

The CV of the new model was found to be 13.2%, which has improved, compared to the previous model and within the desired CV range. Yet, the CV of the new CC per GIFA model was lower than the CV of the EC per GIFA model. The scatterplot of predicted and observed CC per GIFA values is presented in Figure 7.27, which follows a vague linear relationship. Further, deviations in the model predictions were plotted against building height and circulation space ratio and presented in Figure 7.28 and Figure 7.29. Most of the predictions lie within the acceptable accuracy range while four predictions were outside the acceptable accuracy range even though those predictions were close to ±25%, which demonstrates a better prediction performance than the previous model. Further, the storey cluster analysis is illustrated in Figure 7.30 point out that the three out of four deviations outside the acceptable region is attributable to 3-5 storey cluster. This implies that the model performs better within the 1-2 storey cluster than the 3-5 storey cluster.

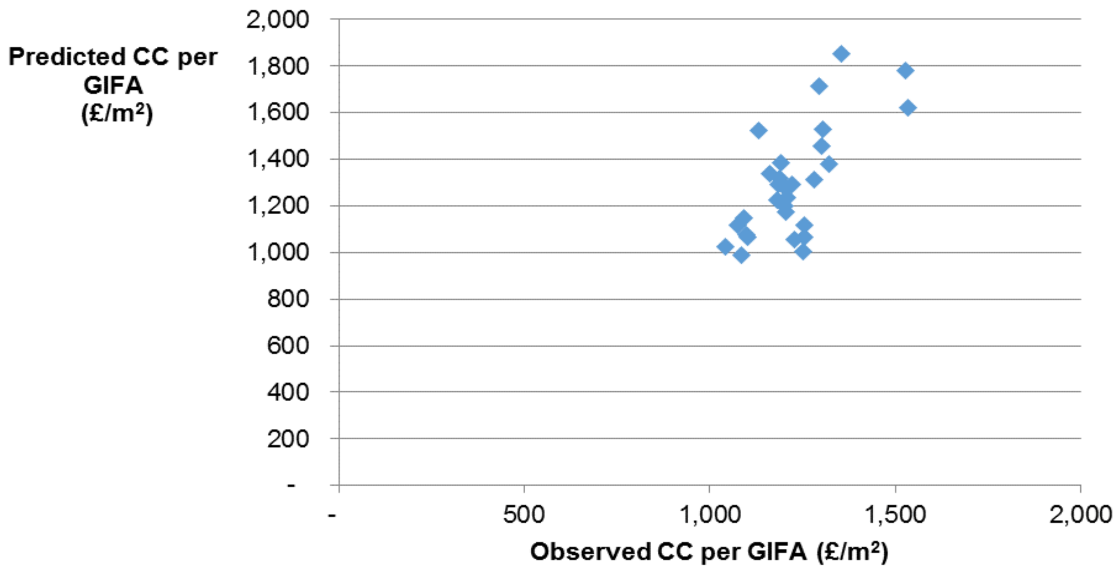


Figure 7.27: Scatterplot of predicted Vs. observed CC per GIFA values – internal data (new CC per GIFA model)

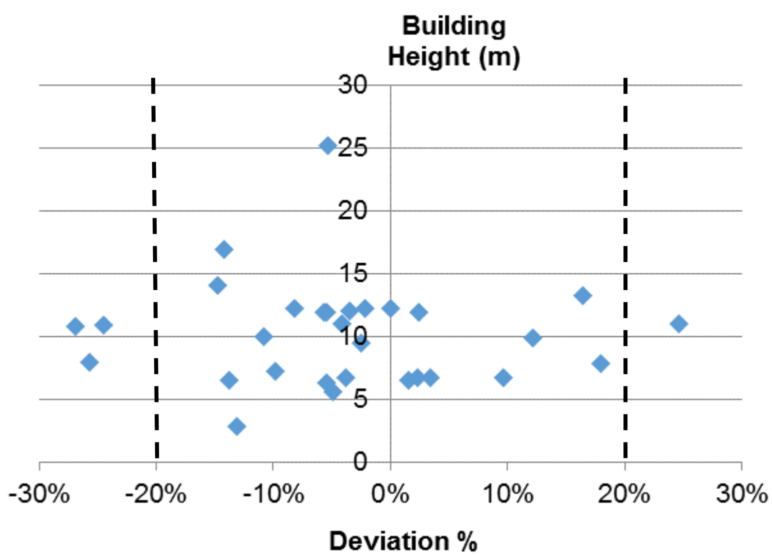


Figure 7.28: Mapping the model prediction deviation against building height – Internal data (new CC per GIFA model)

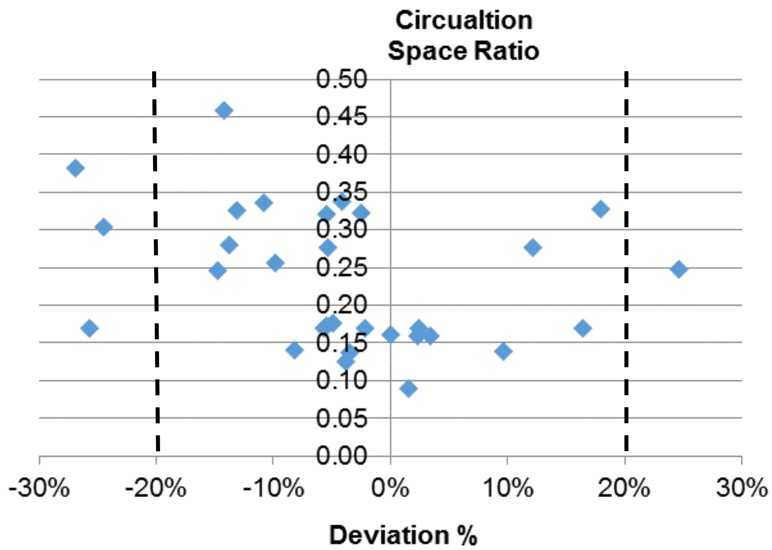


Figure 7.29: Mapping the model prediction deviation against circulation space ratio – Internal data (new CC per GIFA model)

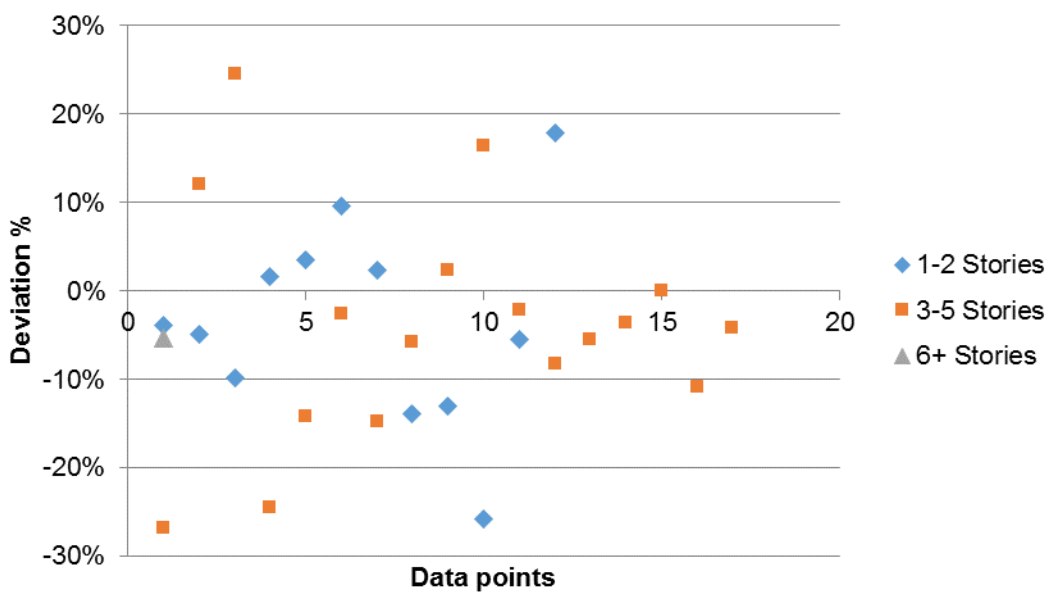


Figure 7.30: The CC per GIFA model prediction at different clusters – Internal data (new CC per GIFA model)

c) Prediction performance with external data

The CV of the model for external data was found to be 24.5% which is outside the acceptable margin established in the study though Peurifoy and Oberlender (2002) suggest 25% accuracy is acceptable for early stage prediction models. The model predictions against the observed values with external data are presented in Figure 7.31. Deviations in predictions are mapped against building height and circulation space ratio is presented in Figure 7.32 and Figure 7.33 where more than half of the predictions fall outside the acceptable region. The summary of the predictions of external data is presented in Table 7.14. Further, the deviation was analysed based on the storey cluster, which is presented in Figure 7.34. Accordingly, the majority of the predictions outside the acceptable accuracy region belong to the 6+ storey cluster. This implies that the model does not work well with buildings more than 6 storeys.

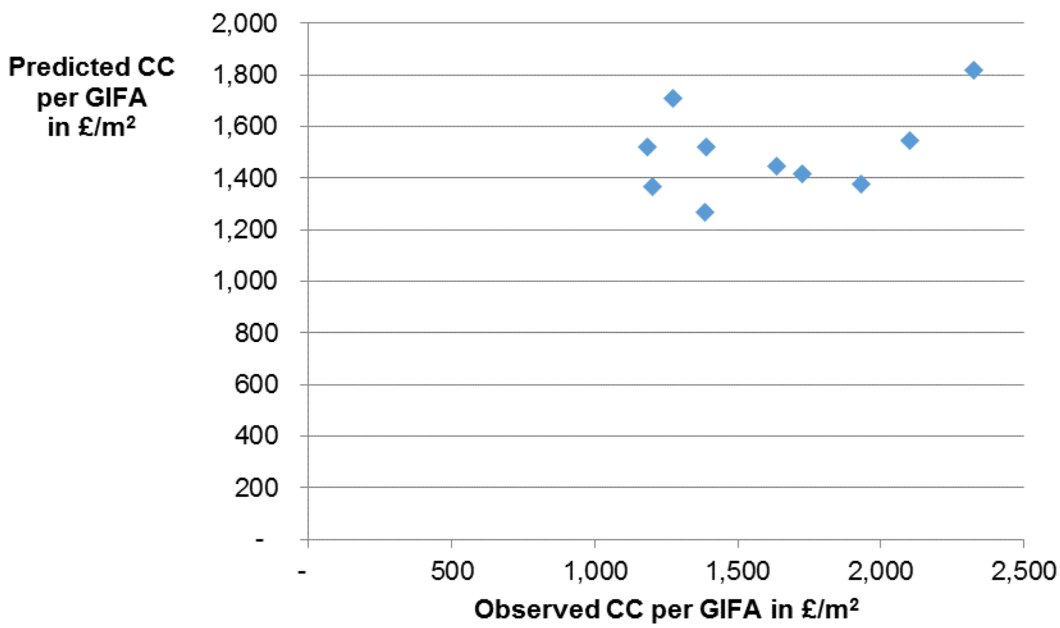


Figure 7.31: Scatterplot of predicted Vs. observed CC per GIFA values – External data (new CC per GIFA model)

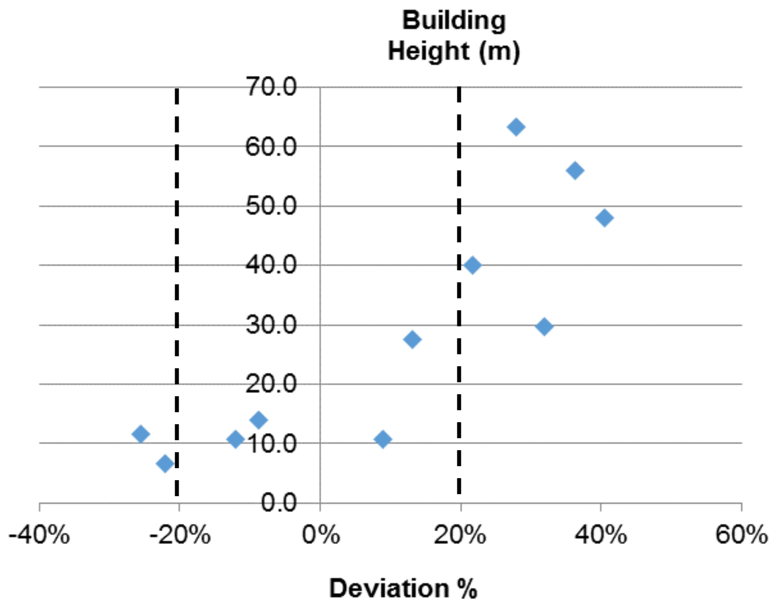


Figure 7.32: Mapping the model prediction deviation against building height – Internal data (new CC per GIFA model)

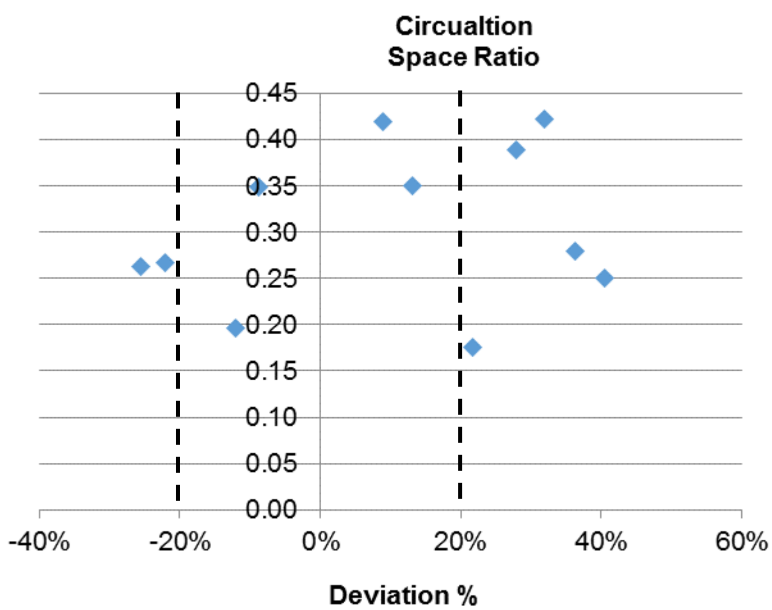


Figure 7.33: Mapping the model prediction deviation against circulation space ratio – Internal data (new CC per GIFA model)

Table 7.14: Calculations of the new CC per GIFA model deviation for the external data

Building ID	Predicted (£/m^2)	Observed (£/m^2)	Residual (£/m^2)	Deviation [(Predicted-Observed)/Observed]
D1002	2,061	1,315	420.04	32%
D1003	1,234	1,520	-335.69	-22%
D1004	1,913	1,447	189.42	13%
D1005	2,953	1,820	506.39	28%
D1006	2,051	1,418	307.18	22%
D1007	1,539	1,520	-132.19	-9%
D1008	1,264	1,369	-165.26	-12%
D1009	1,366	1,708	-435.40	-25%
D1010	2,619	1,543	559.76	36%
D1011	2,362	1,375	557.28	41%
D1013	1,533	1,270	114	9%

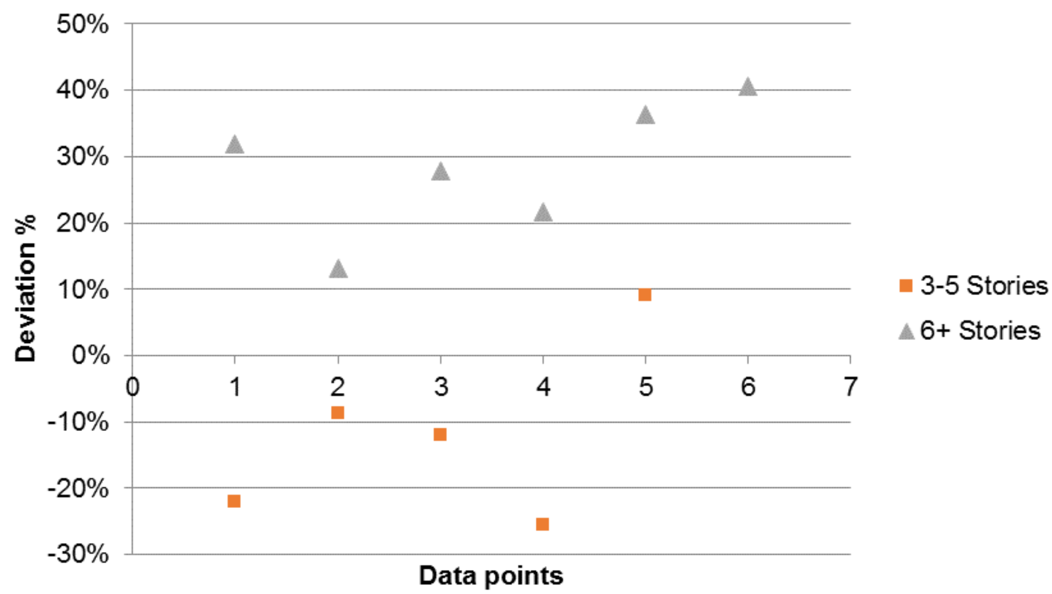


Figure 7.34: The CC per GIFA model prediction at different clusters – External data (new CC per GIFA model)

7.3.5. CC Model

a) Closeness of fit

The model fit was found to be 95.1%, which is high, similar to EC model. The model explains 95.1% of the change in the dependent variable (CC) by GIFA and building height. Remaining change in the dependent variable deemed to be explained by the other design variables, which were not found to be statistically significant in the study.

b) Prediction performance with internal data

The model had a CV of 45.2%, which is very high and unacceptable for an early stage estimate. The model predictions deviate highly from the observed values (from -64% to 5%). Similar to EC model, the prediction accuracy of the model demonstrates a problem. The predicted values were plotted against the observed values, which are illustrated in Figure 7.36, which displays a perfect linear relationship. Further, deviations in predictions were mapped against GIFA and the building height, which are presented in Figure 7.35 and Figure 7.38. There are only a few predictions fall within the acceptable accuracy region showing unsatisfactory performance of the model. The storey cluster analysis presented in Figure 7.37 also suggests that the model prediction is poor in both 1-2 and 3-5 storey clusters while the highest deviation is found within the 3-5 storey range.

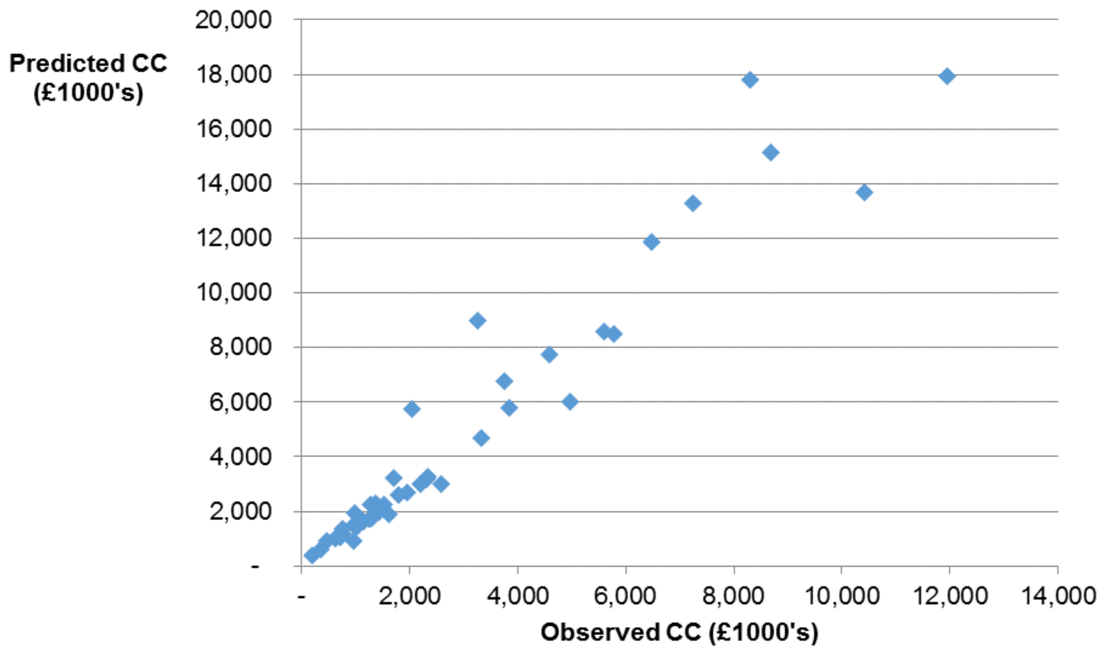


Figure 7.36: Scatterplot of predicted Vs. observed CC values – internal data

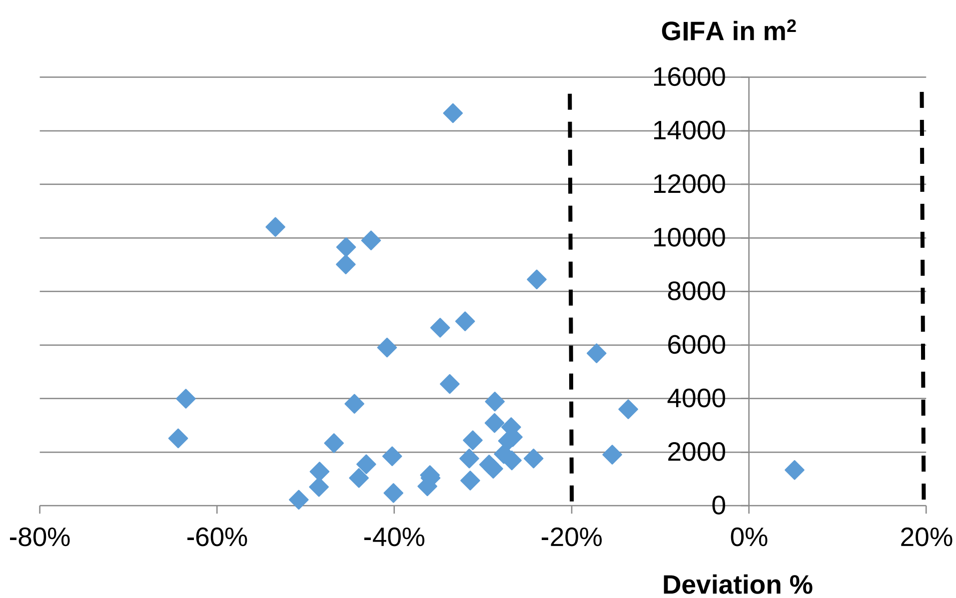


Figure 7.35: Mapping the CC model prediction deviation against GIFA – Internal data

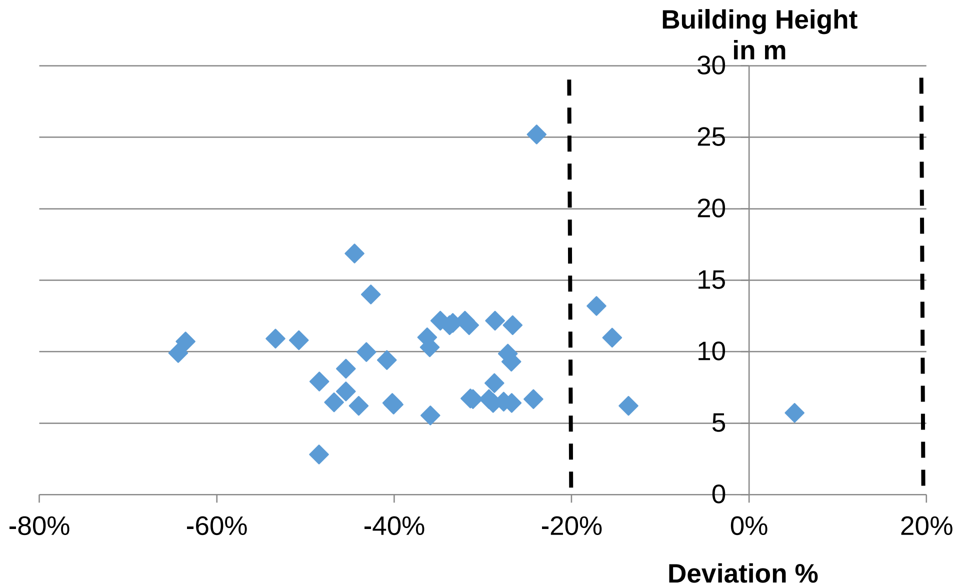


Figure 7.38: Mapping the CC model prediction deviation against building height – Internal data

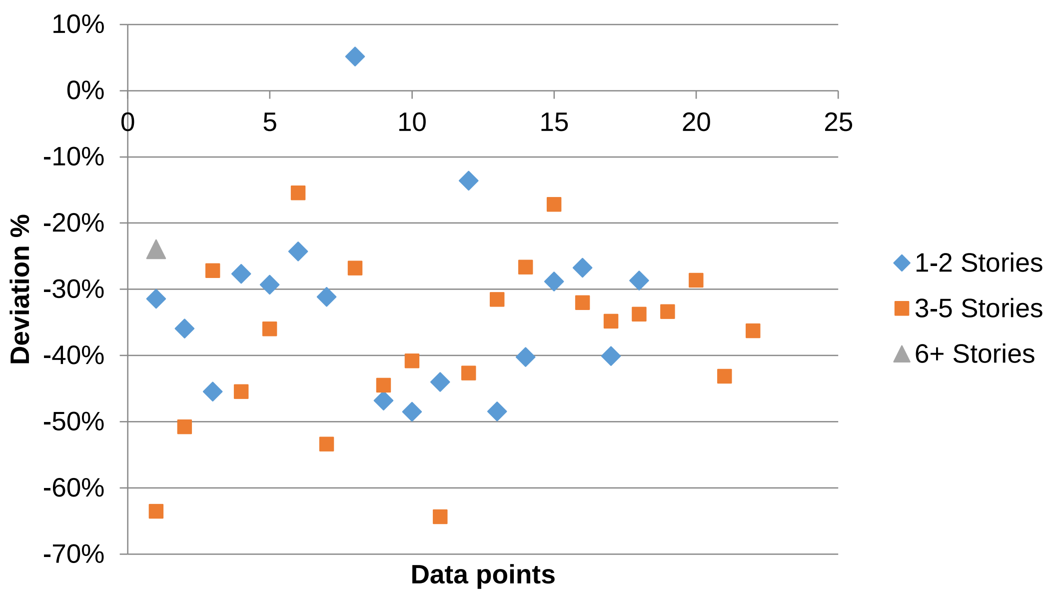


Figure 7.37: The CC model prediction at different clusters – Internal data

c) Prediction performance with internal data

Predicted CC and observed CC for the external data were plotted in a scatterplot and presented in Figure 7.39. The deviation in the model predictions mapped against GIFA and building height are presented in Figure 7.41 and Figure 7.40 where only two predictions fall within the acceptable accuracy region and the deviation tends to increase with the building height.

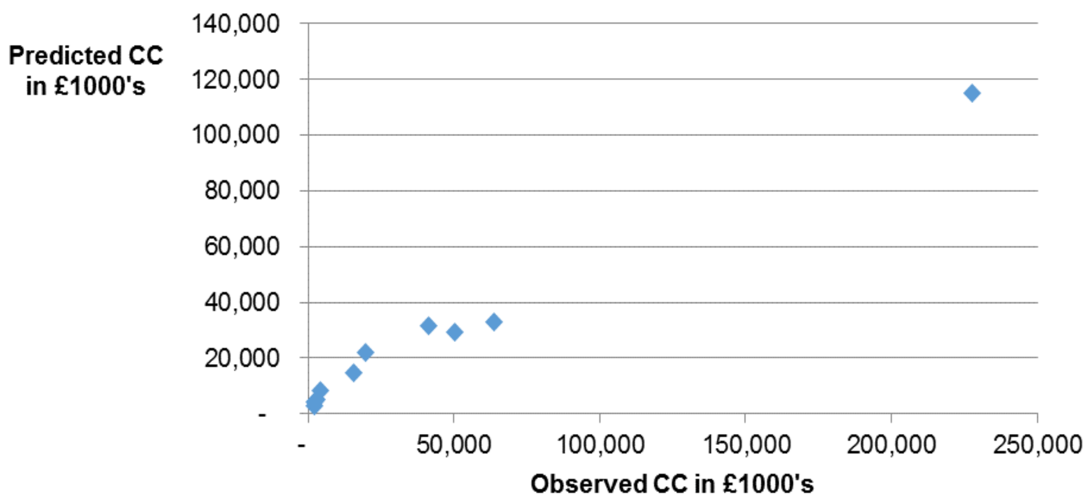


Figure 7.39: Scatterplot of predicted Vs. observed CC values – External data

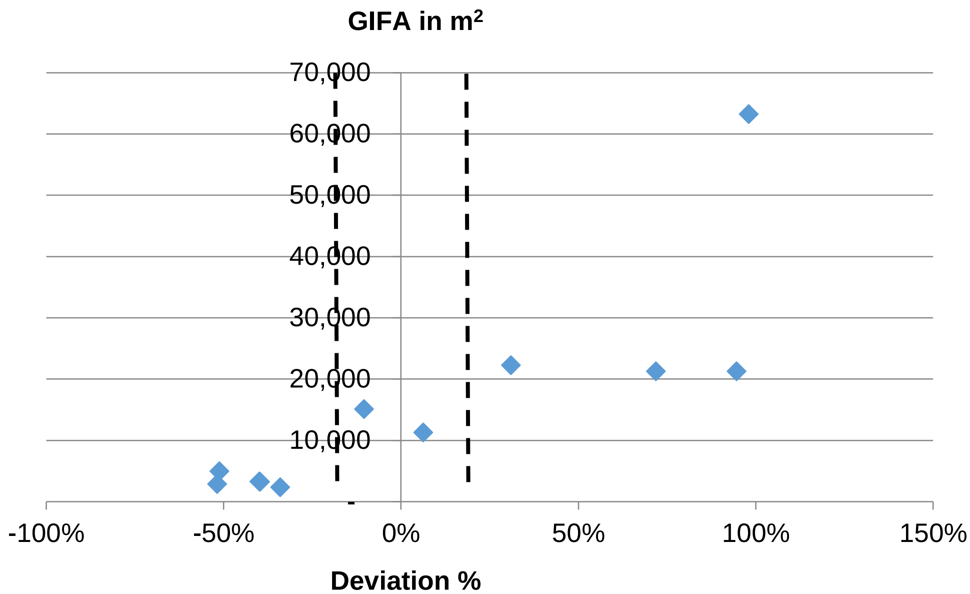


Figure 7.41: Mapping the CC model prediction deviation against GIFA – External data

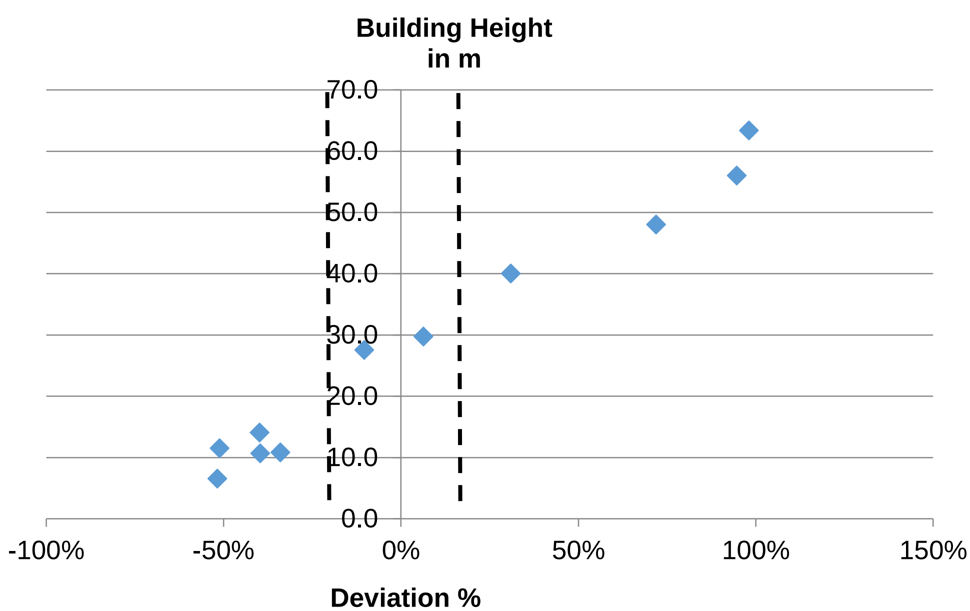


Figure 7.40: Mapping the CC model prediction deviation against building height – External data

The residual analysis of external data is presented Table 7.15. The accuracy ranges from -52% to 98% with a CV of 140% which indicates an extremely poor prediction performance of the model. However, the storey cluster analysis (see, Figure 7.42) suggests that the observed highest deviation falls within the 6+ storey cluster. This reaffirms the model specification that the developed model caters only the estimating need of up to 6 storeys.

Table 7.15: Calculation of the CC model deviation for the external data

Building ID	Predicted (£1000's)	Observed (£1000's)	Residual (£1000's)	Deviation [(Predicted-Observed)/Observed]
D1002	15,819	14,882	938	6%
D1003	2,095	4,346	-2,251	-52%
D1004	19,593	21,874	-2,281	-10%
D1005	227,865	115,091	112,773	98%
D1006	41,390	31,607	9,783	31%
D1007	3,003	4,998	-1,995	-40%
D1008	2,691	4,465	-1,773	-40%
D1009	4,135	8,468	-4,334	-51%
D1010	63,957	32,873	31,084	95%
D1011	50,337	29,286	21,050	72%

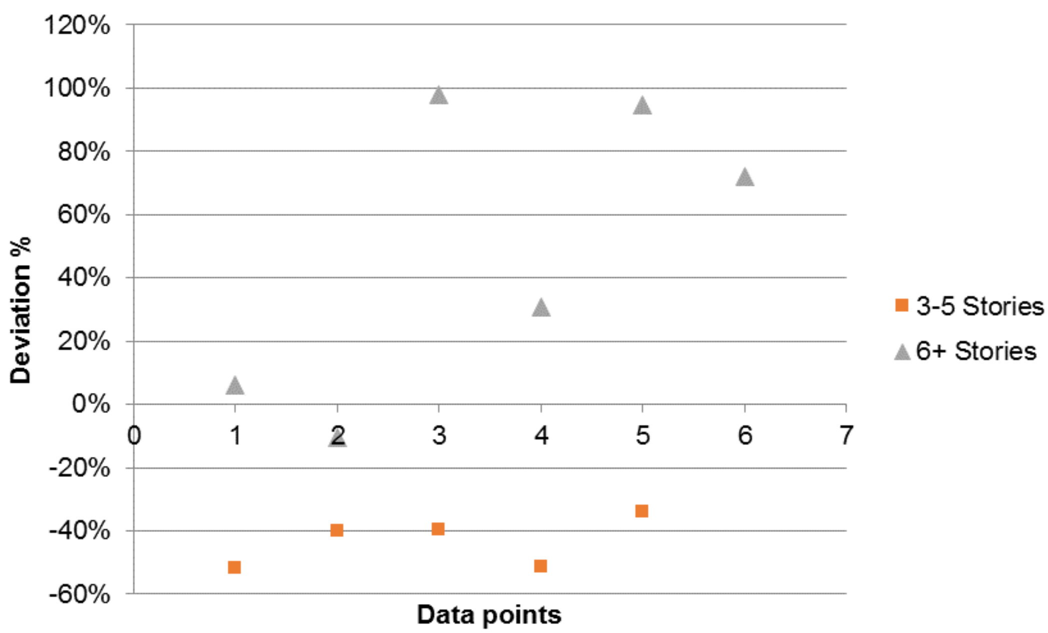


Figure 7.42: The CC model prediction at different clusters – External data

7.4. Discussion of the Model Validation Outcome

7.4.1. Embodied Carbon Models

The EC per GIFA model and EC model are compared and presented in Table 7.16. Accordingly, the EC model outperformed the EC per GIFA model in model fit (R^2) criteria though it did not produce desired outcomes in case of CV and predicting for external data. Therefore, based on the overall performance the EC per GIFA model is considered to be better performing model than the EC model.

Table 7.16: Comparison of EC models

Model Features	EC per GIFA model	EC model
R^2	48.1%	98.3%
CV – internal	10.65%	25.7%
CV – external	11.00%	15.20%
Deviation in prediction for external data	-22% to 11%	9% to 49%

7.4.2. Capital Cost Models

Table 7.17 compares the performance of the CC per GIFA model and the CC model. Even though CC model has a good model fit, the new CC per GIFA model outperformed CC Model in prediction performance of both internal and external data. Even though high deviations in predictions of the CC per GIFA model were found when predicting for external data, the model performs fairly well within the 3-5 storey cluster. Hence, the CC per GIFA model was considered as the better model than the CC model.

Table 7.17: Comparison of CC models

Model Features	CC per GIFA model	CC model
R^2	50.6%	95.1%
CV - internal	13.2%	45.2%
CV - external	24.5%	140.0%
Deviation in prediction for external data	-25% to 41%	-52% to 98%

7.5. Validation of Models with all Variables

As it is evident from the discussion above that none of the models proved to be exemplary, it was decided to validate the models with all the input variables considered in the study to find if the models outperform the previous models. Therefore, this subsection covers the validation of the full models regardless of the statistical significance of the eliminated variables during the model building process.

7.5.1. EC per GIFA Full Model

The derived EC per GIFA model with all the selected design variables is as follows (see, Table 7.21 in Section 7.5.1):

Equation 7.1: EC per GIFA model with all design variables

$$\hat{y} = 630.353 + 1.066x_{BH} + 144.233x_{W:F} + 85.992x_{CR} + 66.591x_{B} - 69.851x_{FI} + 8.323x_{SI}$$

Where,

$\hat{y}$ $- Estimated\ EC\ per\ GIFA\ of\ the\ building$
x_{BH} $- Building\ Height$
$x_{W:F}$ $- Wall\ to\ Floor\ ratio\ of\ the\ building$
x_{CR} $- \%\ of\ Circulation\ Area\ of\ the\ building$
x_{B} $- Number\ of\ basements\ in\ the\ building$
x_{FI} $- Finishes\ Index\ of\ the\ building$
x_{SI} $- Services\ Index\ of\ the\ building$

a) Closeness of fit

The model has an adjusted R^2 value of 45.7%, which implies that 45.7% change in EC per GIFA, is explained by all the independent variables in the model (building height, wall to floor ratio, circulation ratio, no. of basements, finishes index and services index). This is lower than the model considered in Section 7.3.1 where 48.1% variation in EC per GIFA is explained by wall to floor ratio and no. of basements.

b) Prediction performance with internal data

The CV of the model was found to be 9.93%, which is within the desired CV range for early stage estimating and better than the previous model. The overall deviation in the prediction of the internal data ranges from -16% to 20%. The deviation range of this model is smaller compared to the previous model. Figure 7.43 presents the scatterplot of predicted and observed EC per GIFA values, which demonstrate a weak correlation. The model prediction for different numbers of storeys was also examined as illustrate in Figure 7.44. Most of the predictions in 1-2 and 3-5 storey clusters lie within -15% and 15% and the model performs well within both the clusters for internal data.

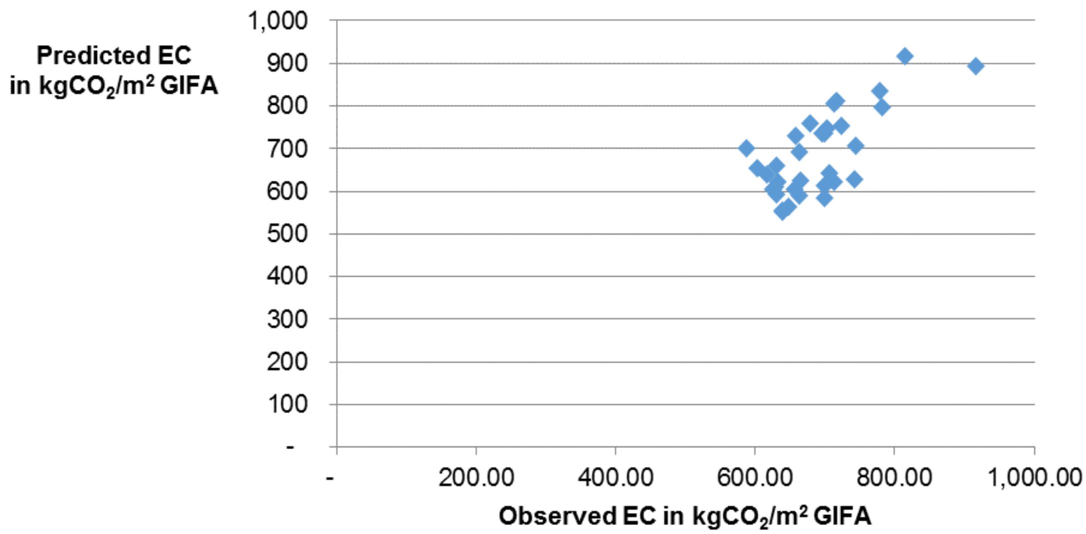


Figure 7.43: Scatterplot of predicted Vs. observed EC per GIFA values for the– internal data (complete EC per GIFA model)

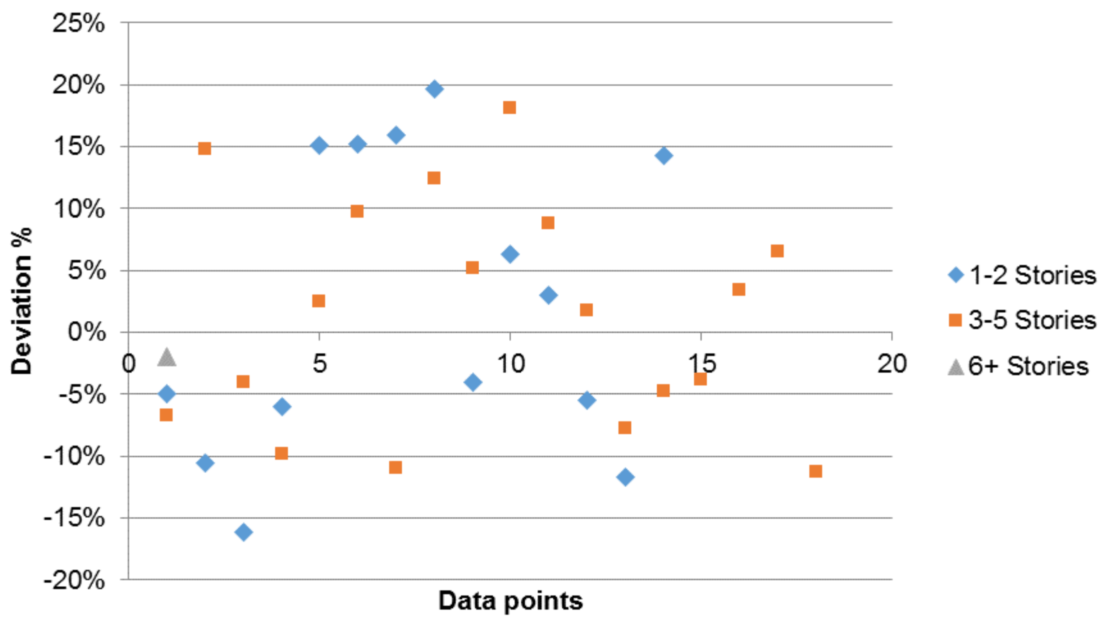


Figure 7.44: The EC per GIFA model prediction at different clusters – Internal data (complete EC per GIFA model)

c) Prediction performance with external data

Predictions were mapped against the observed values for external data, which is presented in Figure 7.45. The accuracy of predictions ranges from -19% to 17% with a CV of 11.4% which is within the desired accuracy range and better than the previous model. Further, the analysis of different storey clusters illustrated in Figure 7.46 reveals that the deviation is smaller (less than ±5%) for most of the predictions within the 3-5 storey cluster compared to 6+ storey cluster. Hence, it can be said that the model performs well within 3-5 storey cluster compared to the 6+ storey cluster similar to the previous model. No conclusions can be drawn about 1-2 storey cluster due to lack of external data within this cluster.

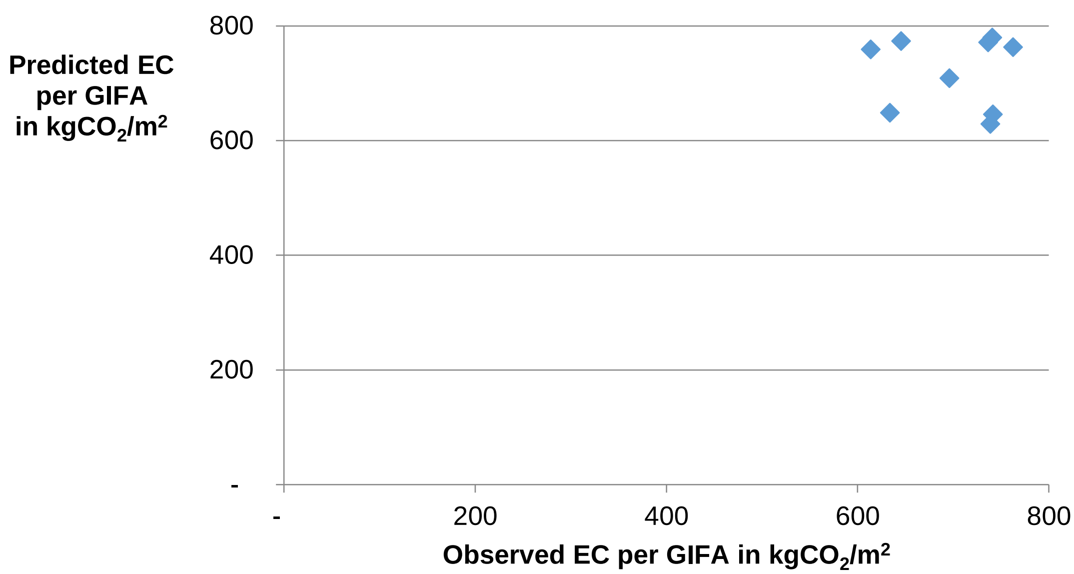


Figure 7.45: Scatterplot of predicted vs. observed EC per GIFA values for the– external data (complete EC per GIFA model)

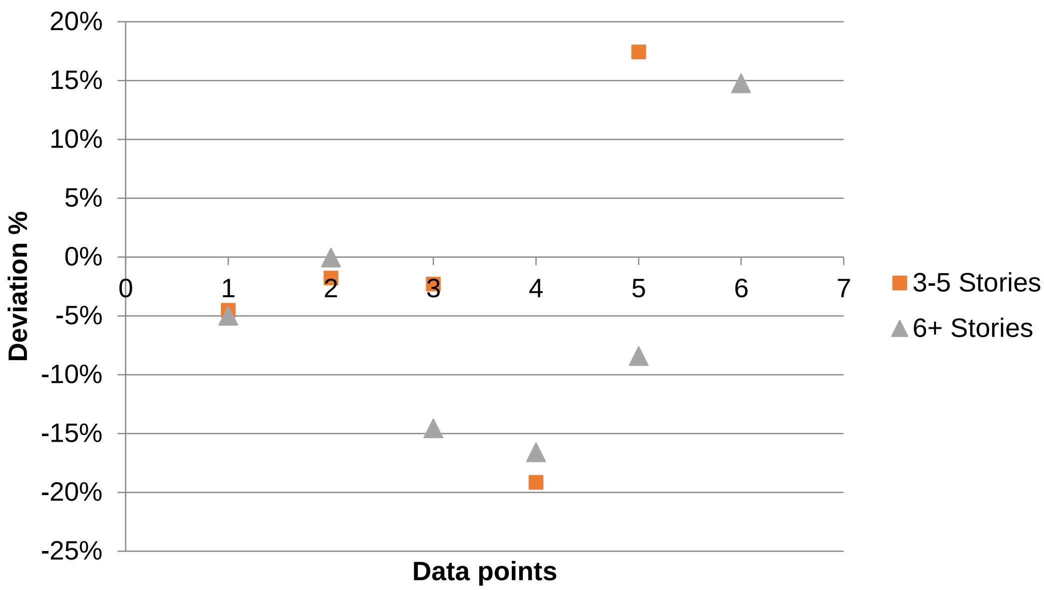


Figure 7.46: The EC per GIFA model prediction at different clusters – External data (complete EC per GIFA model)

7.5.2. CC per GIFA Full Model

The derived EC per GIFA model with all the selected design variables is as follows (see, Table 7.21 in Section 7.5.1):

Equation 7.2: CC per GIFA model with all design variables

$$\hat{y} = 895.113 + 15.799x_{BH} - 233.376x_{W:F} + 947.062x_{CR} + 35.947x_{B} + 19.948x_{FI} + 18.829x_{SI}$$

$\hat{y}$ $\quad - Estimated\ CC\ per\ GIFA\ of\ the\ building$

a) Closeness of fit

The model fit was found to be 49%, which implies that 49% of the change in CC per GIFA is explained by all the design variables considered while 50.6% of the change in CC per GIFA is explained by building height and circulation space ratio, by the previous model.

b) Prediction performance with internal data

In terms of prediction performance, the complete model has a CV of 12.97% where the accuracy of the predictions ranges from -30% to 20% for the whole sample. The model predictions against the observed values presented in Figure 7.47, which shows some degree of correlation. The deviations in predictions are presented for different storey clusters in Figure 7.48 where most predictions fall within the accepted accuracy region except for two predictions which belong to both 1-2 and 3-5 storey clusters (circled in Figure 7.48).

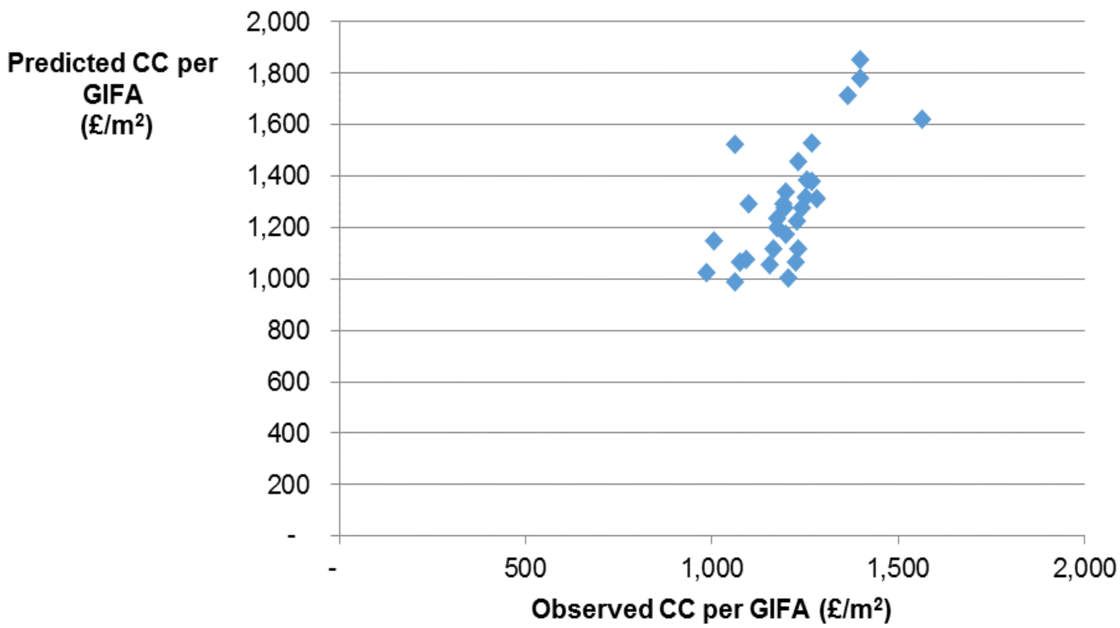


Figure 7.47: Scatterplot of predicted Vs. observed EC per GIFA values for the– Internal data (complete CC per GIFA model)

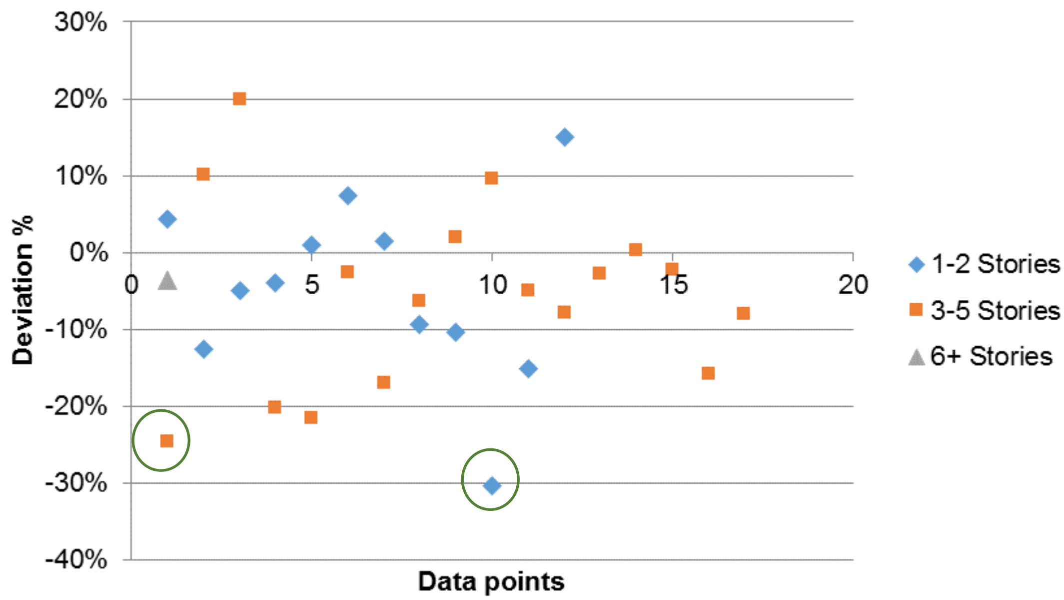


Figure 7.48: The CC per GIFA model prediction at different clusters – Internal data (complete CC per GIFA model)

c) Prediction performance with external data

The model demonstrates a CV of 24.81%, accuracy ranging from -24% to 42% when predicting for the external data. The predictions against the observed values of the external data are presented in Figure 7.49 and the model deviations analysed based on storey cluster is presented in Figure 7.50. Accordingly, all of the predictions, which fall outside the accepted accuracy range, belong to the 6+ storey cluster except for one prediction. Hence, it is clear that the model is not suitable to predict CC per GIFA for the buildings with more than 6 storeys.

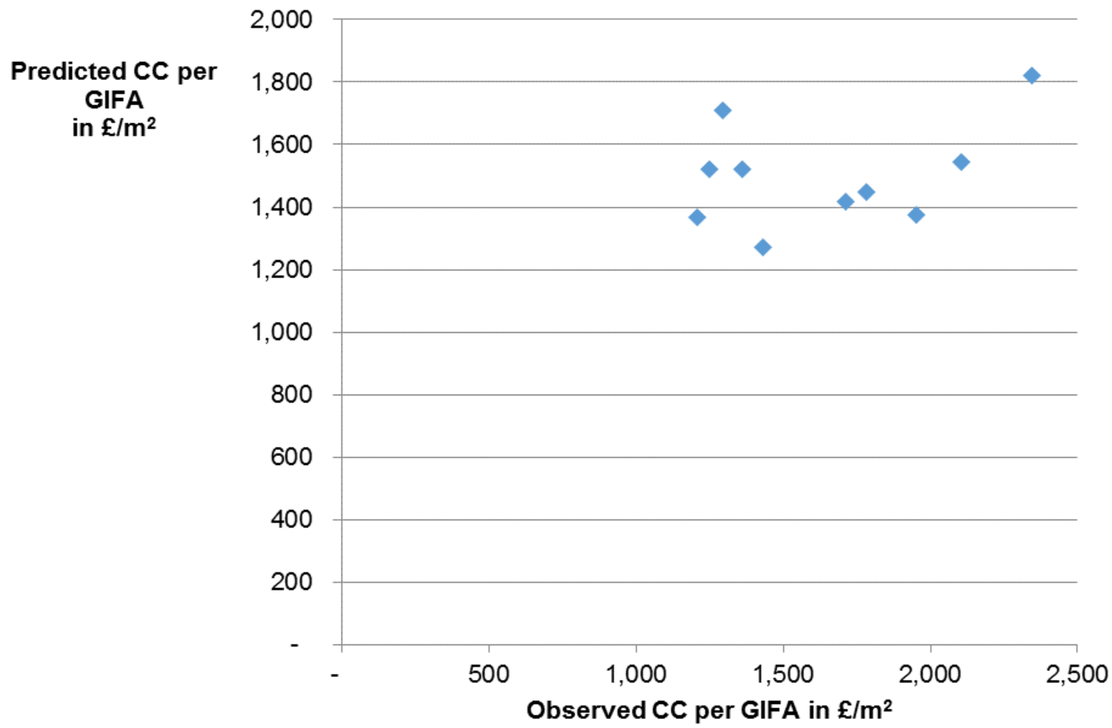


Figure 7.49: Scatterplot of predicted Vs. observed EC per GIFA values for the– External data (complete CC per GIFA model)

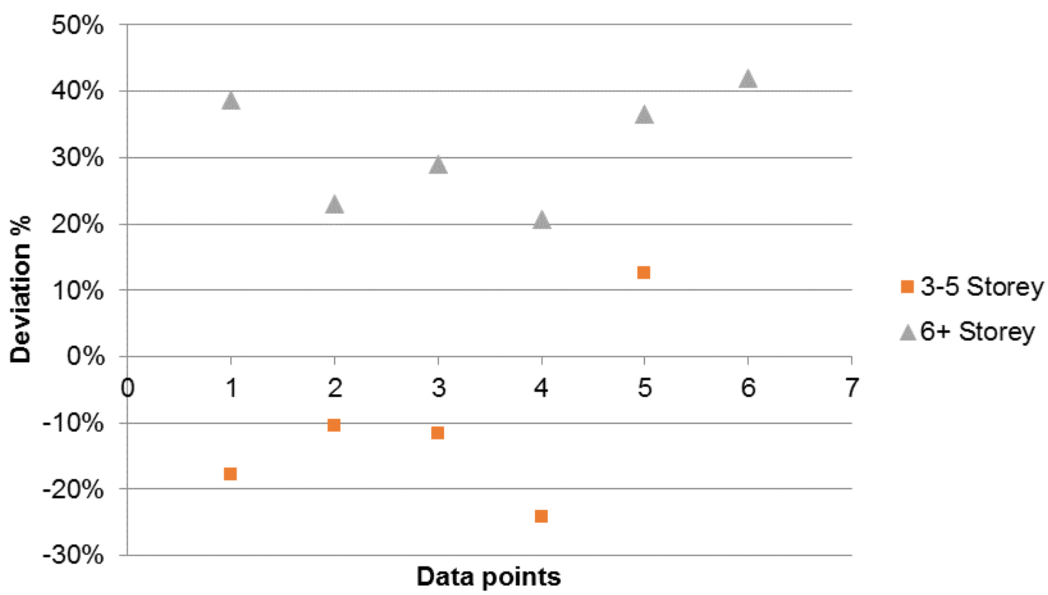


Figure 7.50: The CC per GIFA model prediction at different clusters – External data (complete CC per GIFA model)

7.6. Comparison of Models with Statistically Significant Variables and Models with all Variables

7.6.1. EC per GIFA Models

The model performance measures were compared for the model with statistically significant variables and the model with all variables and presented in Table 7.18. Accordingly, model fit is better in the model with only the statistically significant variables. CV for internal data is better in the model with all the variables while CV for external data is better in the model with only significant variables. Hence, no significant improvement in the predictions of the model with all the variables was found.

Table 7.18: Comparison of EC per GIFA model with statistically significant variables and model with all variables

Performance measures	Model with statistically significant variables	Model with all variables
R^2	48.1%	45.7%
CV – internal data	10.65%	9.93%
CV – external data	11.00%	11.40%

7.6.2. CC per GIFA Models

Similar to the comparison of EC per GIFA model, model fit is better in the model with only the statistically significant variables (see, Table 7.19). CV is better in the model with all the variables for internal data while the model with statistically significant variables outperforms the model with all the variables when performing with external data. Nevertheless, there is no significant difference in the prediction performances of the models.

Table 7.19: Comparison of CC per GIFA model with statistically significant variables and model with all variables

Performance Measures	Model with statistically significant variables	Model with all variables
R^2	50.6%	49%
CV – internal data	13.2%	12.97%
CV – external data	24.5%	24.81%

7.7. Summary

Validating the models is an important step in model development to ensure the applicability of the model. Accordingly, the developed models were validated by assessing their closeness to fit and prediction accuracy with internal and external data. The dataset used to develop the models was used to check the internal validity while eleven buildings out of thirteen from Dataset 1 were used to check the external validity of the models. However, Dataset 1 was adjusted to accommodate Fittings and Services cost and EC in their estimates with the use of benchmarks. The EC model outperformed the EC per GIFA model in model fit (R^2) criterion though it did not produce the desired outcome in the case of CV for both internal and external data. Therefore, based on the overall performance the EC per GIFA model is a better performing model than the EC model. Similarly, the CC model had a good model fit though the CV for both internal and external data were poor. Further, it was also found that the models performed well within the 1-2 storey cluster and poorly for the 6+ storey cluster. The performance within the 3-5 storey cluster is generally within the acceptable accuracy range. Finally, the models with all the variables did not demonstrate any significant improvements in the model predictions. From this analysis and validation, it was concluded that the models with the statistically significant variables are the most satisfactory.

8. Key Findings and Implications

8.1. Introduction

This chapter summarises the key findings of the research presented in the data analysis (Chapter 6) and the model validation (Chapter 7) chapters and the implications of the study findings are discussed by providing examples where necessary. Key findings are presented in three sections including the EC and CC models, the carbon and cost hotspots and the EC and CC relationships. The model descriptions, the applicability of the models and the limitations of the models are discussed. In addition, the developed CC model was compared with other CC models found in the literature while the EC model could not be compared due to the absence of literature on similar models. Further, the application of the knowledge of carbon and cost hotspots is illustrated with examples. EC and cost relationships were explored at the building level and elemental level, which display a close association between the two not only at the building level but also in most of the element levels.

8.2. The Embodied Carbon and Capital Cost Models

As discussed in the model validation chapter, two CC models (CC Model and CC per GIFA model) and two EC models (EC Model and EC per GIFA model) were compared in terms of model fit (coefficient of determination – R^2) and prediction performance (coefficient of variation - CV). Then, the better performing CC and EC model was selected from each pair (see, Section 7.6). The selected EC and CC models, their applicability, usage guidelines and limitations are discussed in detail here. The prediction models were designed specifically for office buildings of up to 6 storeys. The lowest and the highest values of the predictor variables (of the model) that are used to develop the EC and CC models are presented in Table 8.1.

Table 8.1: Ranges of the predictor variables used to develop the model

Design Variable	Lowest Value	Highest Value	The Model
GIFA (m^2)	212	14,652	EC/CC
Building height (m)	2.8	25.2	CC
Wall to floor ratio	0.24	1.50	EC
Circulation space ratio	0.09	0.46	CC
No. of basements (Nr)	0	2	EC

8.2.1. The EC Model

The selected EC model is presented in Equation 8.1 and the model description is presented in Table 8.2. The model has a coefficient of determination (R^2) of 0.48 (higher R^2 implies better model fit) which means that the model explains 48.1% of the variation in EC per GIFA attributable to Wall to Floor ratio and the number of basements. In other words, EC per GIFA is increased by 164.08 $kgCO_2/m^2$ when Wall to Floor ratio is increased by one unit for a given number of basements and adding a basement will increase EC per GIFA by 68.15 $kgCO_2/m^2$ for a given Wall to Floor ratio. The model was statistically significant and all the variables were significant at an α value of 0.05. Further, the model has a CV of 10.65%, which is within the acceptable accuracy range for early stage estimation (±20% see, Section 4.10). However, when predicting outside the database the predictions deviate from -22% to 11% from the observed values, producing a CV of 11%. In addition, CV was improved to 5.94% when predicting for buildings of up to 6 storeys for external data.

Equation 8.1: EC per GIFA model

$$\widehat{y_1} = 530.62 + 164.08x_{W:F} + 68.15x_B$$

Where,

$\hat{y}$ $- Estimated\ EC\ per\ GIFA\ of\ the\ building$

$x_{W:F}$ $- Wall\ to\ Floor\ ratio\ of\ the\ building$

x_B $- Number\ of\ basements\ in\ the\ building$

Table 8.2: EC per GIFA model descriptors

Model Variables	**Description**
EC per GIFA	The EC in $1m^2$ of GIFA of the building
GIFA	The floor area measured from the internal face of the external walls including the areas occupied by the internal elements like walls, partitions, columns and the like
Wall to Floor ratio	The Façade (including area of windows and doors) area divided by GIFA of the building OR the area of the façade covering $1m^2$ of the GIFA
Basements	The number of basements in the building

Accordingly, the estimated model parameters are as follows (see, Equation 6.1 in Section 6.5.1):

$a_0 = 530.62$

$a_2 = 164.08$

$a_4 = 68.15$

Even though the other identified design variables (building height, circulation space ratio, finishes index and services index) were not found to be significant in predicting EC in the model, bivariate analysis (see, Table 6.17 in Section 6.4.3) suggests that building height and circulation space ratio are correlated with EC per GIFA. Building height has a correlation coefficient of 0.306 (p-value - 0.052) and circulation space ratio has a correlation coefficient of 0.360 (p-value - 0.039) which indicate a moderate correlation. Hence, users of the model should be aware that building height and circulation space ratio also have an association with EC and can have an impact on the total EC. Therefore, further investigation of the variables affecting the remaining change in the EC per GIFA is recommended.

8.2.2. The CC Model

The selected CC per GIFA model is presented in Equation 8.2 and the model description is presented in Table 8.3. The model (building height and circulation space ratio) explains 50.6% of the variation in CC per GIFA, which is a better fit than the EC per GIFA model. The model suggests that the CC per GIFA increases by £18.37/m^2 for every meter increase in the building height for a given circulation space ratio. Similarly, the CC per GIFA increases by £800.65/m^2 for every unit increase in circulation space ratio (OR £8.01/m^2 for every percentage increase in circulation space ratio for a given building height). However, the remaining 49.4% of the variation is attributable to other design variables, which were not modelled in the study. The CV of the model was found to be 13.2% when predicting for internal data and it deteriorates to 24.5% when predicting for external data. This suggests that the model does not perform very well with the data outside the model database. However, the data used for external validation consisted of building with more than 6 storeys. Hence, the CV was calculated only for the buildings up to 6 storeys which was 9.85% (deviation ranging from -22% to 9%) and within the acceptable accuracy range.

Equation 8.2: CC per GIFA model

$$\widehat{y_3} = 850.17 + 18.37x_{BH} + 800.65x_{CR}$$

Where,

$\widehat{y_3}$ $- Estimated\ CC\ per\ GIFA\ of\ the\ building$

x_{BH} $- Building\ Height$

x_{CR} $- Circulation\ space\ ratio\ of\ the\ building$

The estimated model parameters are as follows (see, Equation 6.4 in Section 6.5.2):

$c_0 = 850.17$

$c_1 = 18.37$

$c_3 = 800.65$

Table 8.3: CC per GIFA model descriptors

Model Variables	Description
CC per GIFA	CC incurred per m^2 of GIFA of the building
GIFA	Area of the building measured to the internal face of the perimeter walls at each floor level including the areas occupied by the internal elements like walls, partitions, columns and the like
Building Height	Storey height (measured from floor finish to floor finish OR to underside of ceiling finish) multiplied by the number of storeys
Circulation Space Ratio	Non-usable area of the building (total area of all enclosed spaces forming entrance halls, corridors, staircases, lift wells, connecting links and the like) divided by the GIFA of the building

The study findings validate the established theory of cost and design variable relationships, which suggest that CC per GIFA increases with building height and circulation space ratio. However, other design variables were not found to be significant in the model while bivariate analysis (see Table 6.17 in Section 6.4.3) suggests that Wall to Floor ratio is also correlated with CC per GIFA with a correlation coefficient of 0.322 (p-value - 0.040) indicating a moderate correlation.

8.2.3. Applicability of the Models

These models are proposed for the early stages of design where only limited information is likely to be available. This is the conceptual stage according to the RIBA plan of work 2013. The models enable an easy and quick way of estimating EC and CC during the early stages of design and help to optimise the conceptual design. The users can calculate the EC per GIFA and CC per GIFA for a given building by entering the values for the predictive design variables in the models as shown in Table 8.4. The calculations have to be performed manually by the users. Further, the models are only applicable for office buildings of up to 6 storeys. Models for different types of buildings with different height categories can be formulated by collecting data and analysing the data using the same methods proposed in the research.

It should be noted that the developed models are based on a number of assumptions. The prediction of EC model needs not to be adjusted for time and location unless the method of manufacturing of materials is changed. For instance, EC of materials are deemed lower if fossil fuels are substituted by renewable energy sources during the manufacturing process of materials, hence, such global variables need to be considered in the EC estimate. In addition, the predicted EC covers a Cradle-to-Gate boundary, which implies transport is excluded (other than raw material transport to factory gate). However, transport EC could be significant in some projects, which use more of imported materials. Therefore, users should be mindful of such exceptional circumstance and make necessary allowances in the estimate. On the other hand, the prediction of the cost model needs to be adjusted for time and location. The cost model has a base date of 2016 1Q and a location index of 100. Hence, time and location need to be adjusted accordingly when forecasting for a future project. The estimates of EC per GIFA and CC per GIFA include only building work (cost includes mark-up which is unknown) and exclude preliminaries and external works. Estimators should be aware of this and make necessary adjustments to the rates to obtain a holistic estimate. In addition, the models cover only the limited set of specifications listed in Table 8.8 later in this chapter. Therefore, adjustments have to be made to the rates if the specification of a given building differs significantly from the modelled specification.

Table 8.4: Using the model to forecast the EC and CC of a building during the early stages of design

Building design data:

GIFA (m^2)	5000
No. of storeys	4
Building Height (m)	12.40
Wall to Floor ratio	0.62
Circulation space ratio	0.22
No. of basements	1

<u>Calculation of the EC per GIFA of the building</u>

Model components	Correlation coefficient	Value of the design variable	Resultants
Constant	530.62		530.62
Wall to Floor ratio	164.08	0.62	101.73
No. of basements	68.15	1	68.15
EC per GIFA of the building			**700.50 $kgCO_2/m^2$**

EC of the building = 700.50 $kgCO_2/m^2$ x 5000 m^2

= 3,502.5 tCO_2 (accuracy ±11%)

Range of total EC = 3,117.2 tCO_2 to 3,887.8 tCO_2

= 3,100 tCO_2 to 3,900 tCO_2

<u>Calculation of the CC per GIFA of the building</u>

Model components	Correlation coefficient	Value of the design variable	Resultants
Constant	800.17		800.17
Building height	18.37	12.4	227.79
Circulation space ratio	800.65	0.22	176.14
CC per GIFA of the building			**£1204.10**

CC of the building = £1204.10 x 5000 m^2

= £6,020,505 (accuracy ±13%)

Range of total CC = £5,237,839 to £6,803,171

= £5.2 to 6.8 million

8.2.4. Limitations of the Models

The study sample contains only low to medium rise buildings and predominantly 3 to 4 storied buildings. Hence, the developed models work best at predicting EC and CC of up to 6 storeys as the sample comprises buildings from 1 to 6 storeys. Further, the EC data were derived from three different sources including Dataset 1, Dataset 2 and published EC databases. Hence, the EC estimates are influenced by the used data sources. Further, the manufacturing method of the materials was assumed the same as in the published EC databases. Hence, the models should be adjusted to accommodate the changes in the industry, especially for the manufacturing methods of construction materials.

All cost data were rebased to 2016 1Q and a location index of 100. Hence, the cost model predictions need to be adjusted for time and location when predicting the CC of a future project. On the other hand, absence of such indices for EC made it impossible to adjust EC for time and location. However, the difference in time shall be accounted only if there is a difference in the manufacturing method and the adjustment for the location will be crucial for cradle-to-grave (or cradle) system boundary. Accordingly, it was assumed that there is no difference in the manufacturing process of construction materials, hence, there is no need of time adjustment; and only cradle-to-gate boundary is covered by the study which implies that the transport other than from raw material extraction to the manufacturing factory is not included in the estimates, hence, location adjustment is negligible. Nevertheless, cradle-to-gate system boundary is a limitation of the models. Even though, it is desirable for the models to cover cradle-to-grave boundary to provide a holistic perspective on designs, it is challenging due to limited EC data.

Another limitation of regression models is that they are static models. Regression analysis has to be performed again to derive a new model when new data become available. In addition, similar to any other regression model, these models are also dependent on the data used to formulate the models. Hence, with a different set of data different parameter estimates might be obtained. Furthermore, non-linear relationships were not configured in the modelling technique used.

8.2.5. Comparison of the Study Cost Model with Cost Models in the Literature

Comparison between the cost model of the study and other cost models found in the literature is presented in Table 8.5. Accordingly, all models except for one (Phaobunjong, 2002) indicate a better model fit than the study model while the prediction performance of the study model is better than most of the identified models. The study findings closely align with the findings of Phaobunjong (2002) in terms of the predictor variables (independent variables). In comparison, the model developed by Kouskoulas and Koehn (2005) performs better in all aspects although the model encounters some shortfalls. The issue with the model of Kouskoulas and Koehn (2005) was capturing the building height by means of the number of storeys which fail to account for buildings with unusual storey heights as criticised in McGarrity (1988). On the other hand, the model of McGarrity (1988) also suffers from lower sample size, unrealistic correlation coefficients for the predictive variables (negative correlation for number of storeys and GIFA), extreme deviation when predicting outside the database and not considering the building type as a predictive variable as the data sample includes more than one type of buildings. The model developed by Alshamrani (2016) looks almost perfect, but a problem lies in the method used to develop the model. The sample of buildings was formed by considering alternative design scenarios and estimating the cost of each scenario by using national average prices of the construction cost of elements. Two hundred and fifty (250) scenarios were used for the model development and seventy (70) scenarios were used for the model validation of the three hundred and twenty (320) scenarios developed by Alshamrani (2016). Hence, a good prediction performance is obvious due to the use of benchmark rates for all buildings in the sample.

Table 8.5: The cost model of the study compared with the other models in the literature

	The study	**McGarrity (1988)**	**Kouskoulas and Koehn (2005)**	**Phaobunjong (2002)**	**Alshamrani (2016)**
Regression functional form	Linear	Power	Linear	Linear	Linear
Type of building	Offices	Not specific to one type	Not specific to one type	Not specific to one type	College buildings
Dependent variable	Cost per m^2 GIFA	Cost	Cost per ft^2	Cost per GIFA	Cost per ft^2
Independent variables (coefficient)	Building height (18.73), circulation ratio (800.65)	Height (positive), storeys (negative), duration (positive), liquidated damage (positive), GIFA (negative)	Location (23.93), time of realization (10.97), function or type (6.23), height (0.167), quality (5.26), and technology (30.9)	Number of floors (15.74), usage ratio (126.196),	Height (0.666), number of floors (4.498), area of the building (0.000129), sustainability index (6.292), structure type (5.003)
R^2	0.506	0.907	0.998	0.261	0.873
CV - Internal	13.2%	24.27% (ranges from 1.05% to 62.43%)	Ranges from -0.05% to 6.5%	Ranges from <10% to >50%	Not provided
CV - External	9.85% (-22% to 9%)	5.15% to 236.98%	Not specified	-0.8% to 13.5%	5.6%

8.3. Carbon and Cost Hotspots

The carbon and cost hotspots of the sample office buildings are presented in Table 8.6 based on the analysis presented in Section 6.2. Accordingly, Substructure was identified as the most significant carbon hotspot while less significant cost hotspot. Services were identified as the most cost significant hotspot and the second most carbon significant hotspot. The level of significance of Frame as a carbon and cost hotspot was found to be the same (third most cost and carbon significant hotspot). Interestingly, Upper Floors was not identified as cost significant while it was identified as the fourth carbon significant hotspot. External Walls was identified as the second most cost significant building element while carbon significance of External Walls was found to be low. Roof was identified as the least carbon significant hotspot while it was found to be more cost significant than the Substructure. Floor Finishes was identified as the least cost significant hotspot. Further, the concept of cost and carbon hotspot emerged from the Pareto Principle, which suggests that 80% of the EC (or cost) is attributable to 20% of the elements. However, the 80:20 ratio is not supported in this case. The findings suggest that 80% of the EC emissions are caused by 43% of the elements (6 of the 14 elements) and 80% of the building cost is spent on 50% of the elements (7 of the 14 elements) on average.

Table 8.6: Carbon and cost hotspots of the sample office building

Level of Significance	Carbon hotspots	Cost hotspots
1	1 Substructure	5 Services
2	5 Services	2E External Walls
3	2A Frame	2A Frame
4	2B Upper Floors	2F External Windows and Doors
5	2E External Walls	2C Roof
6	2C Roof	1 Substructure
7		3B Floor Finishes

It can also be noticed that the level of significance of each element in terms of carbon and cost vary even though most of the elements were found to be both carbon and cost hotspots. Therefore, achieving optimisation between carbon and cost is not as simple as it first appeared. For instance, an effort to minimise EC in the Substructure might lead to increase in the Frame cost because of wind bracings, which might offset the cost savings achieved in the Substructure. Hence, the findings highlight the need for in-depth case studies of buildings exploring different design options while studying the change in EC and CC for alternative design options. This will inform designers of the impact of different specifications on the EC and CC of building designs and when an optimum point can be achieved. For instance, Figure 8.1 presents average EC and CC values for three different types of foundation in the sample building. Accordingly, both EC and CC values (mean values) are found to be the lowest in raft foundation and the highest in pile foundation. This conveys that the choice of raft foundation could reduce both EC and CC. Similar graphs can be produced for other building elements and an informed decision can be made by the designers. Especially, with this kind of knowledge, cost and carbon reconciliation can be exercised by compromising the design of the elements which do not produce significant savings in cost or carbon and focusing on the design of the most cost and carbon significant elements which are referred to as the 'hotspots'.

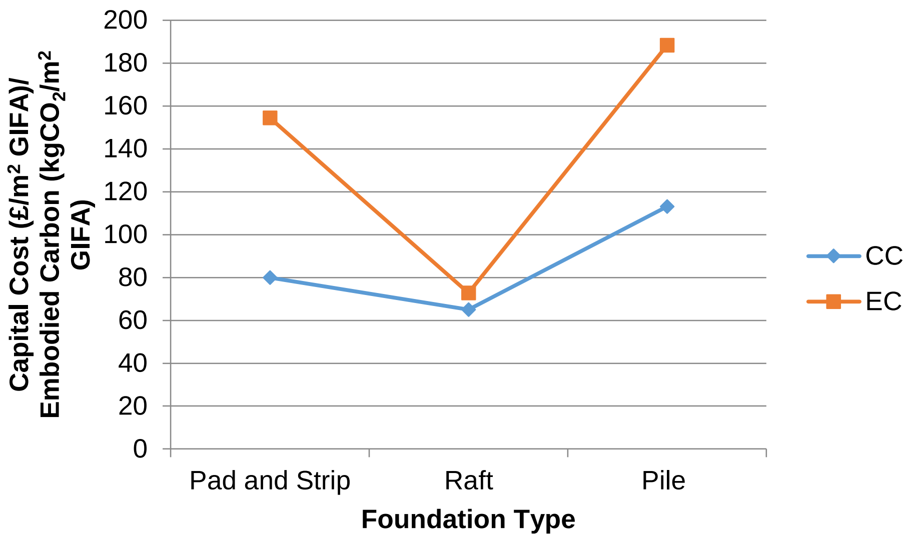


Figure 8.1: EC and CC for different foundation types

Furthermore, some building elements were found to be carbon or cost hotspots in most or all of the sample buildings which were named as 'lead positions'; building elements that were found to be hotspots in some of the buildings were named as 'special positions'; and the building elements that were never identified as hotspots were called as 'remainder positions'. Table 6.4 presents and compares different categories of carbon and cost hotspots of the sample buildings. Substructure, Frame, External Walls and Services were found to be lead carbon and cost hotspots in the office buildings. In addition to the above-mentioned elements Roof and Windows and External Doors were identified as lead cost hotspot while Upper Floors were identified as lead carbon hotspot. On the other hand, Internal Walls and Partitions, Wall Finishes, Floor Finishes and Ceiling Finishes were identified as special carbon and cost hotspots whose identity as a hotspot is ambiguous as these elements were found to be hotspots in some of the buildings. Further, Roof and Windows and External Doors were also identified as special carbon hotspots while both were identified as lead cost hotspots. In addition, Upper Floors, Internal Doors, Fittings, Furnishings and Equipment were identified as special cost hotspots. Interestingly, all of the building elements were found to be a cost hotspot in one or more of the buildings while Stairs, Internal Doors, Fittings, Furnishings and Equipment were never identified as carbon hotspots.

Table 8.7: Classification of carbon and cost hotspots

Hotspot Category	**Carbon Hotspots**	**Cost Hotspots**
Lead positions	Substructure, Frame, Upper Floors, External Walls, Services	Substructure, Frame, Roof, External Walls, Windows and External Doors, Services
Special positions	Roof, Windows and External Doors, Internal Walls and Partitions, Wall Finishes, Floor Finishes, Ceiling Finishes	Upper Floors, Stairs, Internal Walls and Partitions, Internal Doors, Wall Finishes, Floor Finishes, Ceiling Finishes, Fittings, Furnishings and Equipment
Remainder positions	Stairs, Internal Doors, Fittings, Furnishings and Equipment	Nil

It is clear from the findings above that some building elements were identified as hotspots in some buildings, which imply that the building design determines the chances of an element being a hotspot in a particular building. Therefore, the design of 'special positions' can play an important role in influencing carbon and cost accountability of the building. Table 8.8 presents the range of specification for each building element in the sample.

Table 8.8: Alternative design options of the building elements in the sample

Element	Specifications
1 Substructure	Pad and strip, raft, pile
2A Frame	Concrete, steel, hybrid
2B Upper Floors	In-situ concrete, pre-cast concrete, metal decking, timber decking
2C Roof	Concrete flat roof, steel truss, steel mansard, timber truss, timber pitched, aluminium sheet roof, metal decking, glazed atrium roof, Durox roofing units
2D Stairs	Concrete, steel, timber
2E External Walls	Cavity wall, curtain wall, block wall, aluminium cladding, stone ashlar wall, terracotta cladding, pre-cast concrete cladding
2F External Windows and Doors	Double glazed aluminium windows and doors, metal windows and doors, sun screens, shop fronts, softwood doors, curtain wall
2G Internal Walls and Partitions	Brick walls, block walls, metal stud partitions, timber stud partitions, glazed screens
2H Internal Doors	Oak veneered flush doors, Oak veneered solid core doors, ash panelled doors, hardwood doors, softwood panelled doors, aluminium doors, borrowed lights
3A Wall Finishes	Wallboard, MDF panels, gypsum plaster, ceramic tiles, cement plaster, emulsion paint, wallpaper, lightweight plaster, ceramic tiles, eggshell paint, fair face paint, spray paint, stone cladding, laminated chip board, laminated panels, plywood panels, acoustic panels
3B Floor Finishes	Vinyl sheet, ceramic tiles, carpet tiles, raised access floors, asphalt and cement screed, clay tiles, carpet, granolithic paving, slate tiles, chip board, timber floor, quarry tiles, floor paint
3C Ceiling Finishes	Mineral fibre metal suspended ceiling, aluminium PVC composite panels and plasterboard ceiling, Armstrong suspended ceiling, sprayed rendered screed, plaster and paint
4 Fittings and Furnishings	Sanitary fittings, vanity, furniture, kitchen fittings and appliances
5 Services	A/C, non-A/C, A/C automated, non-A/C automated

It can be noticed from the table that some building elements such as Substructure and Frame had minimal design options while elements such as Roof and Internal Finishes were found with many choices. In fact, the special positions have many design options compared to lead positions, highlighting the significance of design decision of special positions. Therefore, further studies and detailed analysis of the impacts of different choices of design in each element will open new avenues for achieving cost and carbon reduction through building designs. For instance, Table 8.9 and Table 8.10 present two design options for Floor Finishes in a particular building. Design option A proposes a combination of vinyl sheet, ceramic tiles (to toilet area) and raised access floor with carpet tiles on top; design option B replaces the area covered by vinyl sheet with raised access floor with carpet tiles. Replacing vinyl sheet with access floor finished with carpet tiles has increased the rates of CC and EC by approximately 100% and 400% respectively, which increased CC per EUQ by 10% and EC per EUQ by 20% (which implies 10% and 20% increase in total CC and EC of the building). Therefore, what-if analysis can be run during detailed design stages and the most efficient design option can be chosen by the designers if this type of analysis is entertained by construction professionals and practices.

Table 8.9: Floor finishes – design option A

Floor Finishes	Qty	Unit	CC	EC	Total cost	Total carbon
Vinyl sheet	797	m^2	28.71	7.69	22,896.65	6,130.41
Ceramic tiles	399	m^2	84.14	15.37	33,547.42	6,127.22
Carpet tiles	2,791	m^2	26.69	10.45	74,484.77	29,159.32
Raised access floor	2,791	m^2	28.30	25.03	78,982.47	69,847.85
					209,911.31	111,264.81
	3,987	m^2	Per m^2		52.65	27.91

Table 8.10: Floor finishes – design option B

Floor Finishes	Qty	Unit	CC	EC	Total cost	Total carbon
Ceramic tiles	399	m^2	84.14	15.37	33,547.42	6,127.22
Carpet tiles	3,588	m^2	26.69	10.45	95,766.13	37,490.56
Raised access floor	3,588	m^2	28.30	25.03	101,862.44	89,804.38
					230,862.44	133,422.16
	3,987	m^2	Per m^2		57.90	33.46

In addition to that, analysis of the whole sample gives a different insight into the problem investigated. It was found that Substructure, Services, Frame, Upper Floors, External Walls and Roof were the most carbon significant building elements (in descending order) which also contribute up to 72% of the CC of the buildings. On the other hand, Services, External Walls, Frame, External Windows and Doors, Roof, Substructure, Floor Finishes (in descending order) were found to be the most cost significant elements which contribute up to 81% of the EC of the building. This finding implies that tackling carbon hotspots also means tackling the building elements that are responsible for 72% of the cost in general. Similarly, tackling the identified cost hotspots implies tackling the elements accountable for 80% of the EC of the building. In comparison, treating cost hotspots seems to be a better option than treating carbon hotspots as it includes the elements responsible for 80% of the CC and EC. Given that, the list of cost hotspots includes all of the identified carbon hotspots except Upper Floors.

8.4. Embodied Carbon and Cost Relationships

EC and CC can be analysed at five different levels including building, elements, components, items and basic inputs (material, labour and plant) as shown in Figure 8.2. However, only the first two levels were analysed in the study due to the limitations in the data obtained. Analysis at Level 1 or the building level suggested that EC and CC are positively correlated, with a strong correlation coefficient of 0.977 at 0.05 significance level. However, it is understood that this correlation was caused by a third variable, which was GIFA. Subsequently, EC and CC were standardised (by dividing the values by the respective GIFA) and correlation coefficient was recalculated. EC per GIFA and CC per GIFA demonstrated a moderately strong correlation with a Pearson's correlation coefficient of 0.645 (p value≤0.05). The finding suggests that it is possible to achieve lower cost and lower EC simultaneously due to the positive association between EC per GIFA and CC per GIFA. However, when investigating EC and CC relationships at different levels different insights were drawn. Level 2 elemental analyses involved the analysis of EC per GIFA and CC per GIFA as presented in Figure 6.35, Table 6.49 and Table 6.50 in Chapter 0. Most of the elements showcase a positive correlation (at 0.05 significance level) between EC per GIFA and CC per GIFA except for Roof, Wall Finishes, Ceiling Finishes and Services which implies that EC and CC can be reduced simultaneously in most of the elements by concentrating on the design.

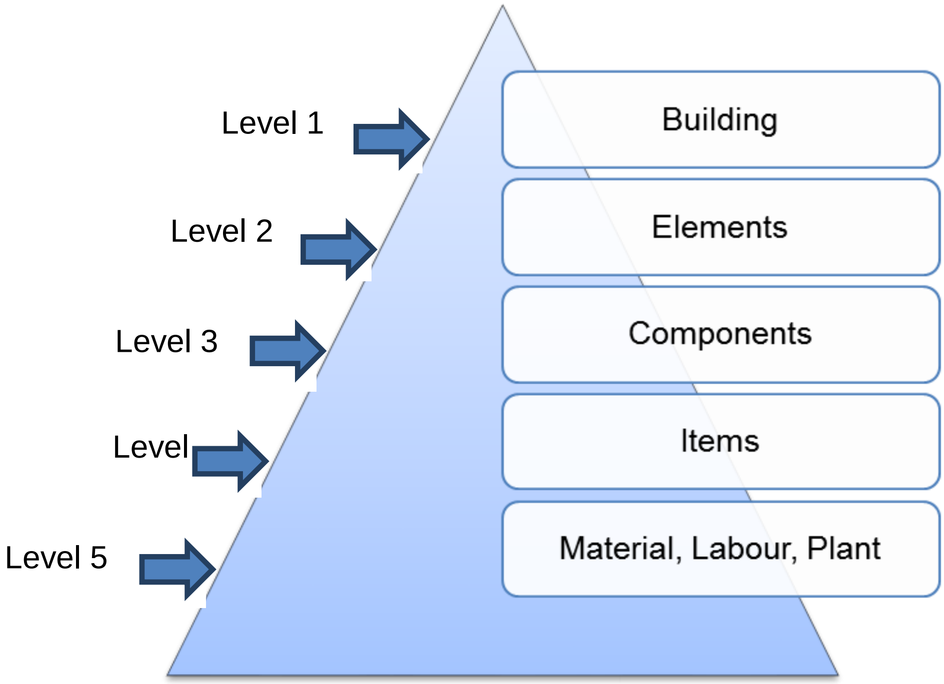


Figure 8.2: Levels of analysis

Further, elements' EC per GIFA values with a lower standard deviation including Substructure, Upper Floors, Floor Finishes, Ceiling Finishes and Services indicates that the EC per GIFA values of the sample buildings hovers closely around the mean, which implies less uncertainty in the prediction of the EC of these elements. Particularly, Substructure and Upper Floors had only up to three design alternatives, which could be the reason for the lower standard deviation. However, elements such as Floor Finishes, Ceiling Finishes and

Services had more than three design alternatives, yet the standard deviation was low. The reason for lower standard deviation in Services EC per GIFA was due to the development of Services EC per GIFA from Dataset 2 (see, Table 5.19 Section 5.7.2). However, lower standard deviation in the EC per GIFA values of Floor finishes and Ceiling Finishes was a true representation of the data. On the other hand, the higher standard deviation was found in the EC per GIFA values of Internal Walls and Partitions and Wall Finishes. Similarly, CC per GIFA of Roof and Services had lower standard deviations even though both of the elements had several design alternatives. However, higher dispersion of data was found in the CC per GIFA values of Stairs, Windows and External Doors, Internal Walls and Partitions, Wall Finishes and Fittings, Furnishings and Equipment.

Table 8.11: Risk or uncertainty matrix of using EC elemental benchmarks for early stage EC estimates

	High Standard Deviation		**Low Standard Deviation**	
	Level of uncertainty/ risk in the estimate	**Elements**	**Level of uncertainty/ risk in the estimate**	**Elements**
Lead Position	High	None	Low	Substructure, Upper Floors, and Services
Special Position	Moderate	Internal Walls and Partitions and Wall Finishes	Very low	Floor Finishes, Ceiling Finishes
Remainder Position	Low	Stairs, Windows and External Doors, and Fittings, Furnishings and Equipment	Negligible	None

The spread of data and the hotspot category together have an influence on the accuracy of the estimate. The risk or uncertainty matrix of using the developed EC elemental benchmarks for early stage EC estimates is presented in Table 8.11 for different combinations of dispersion of data and hotspot category. Accordingly, Substructure, Upper Floors and Services have lower standard deviation for EC per GIFA and are identified as lead positions in the carbon hotspot category. Hence, there is a lower risk or less uncertainty in the EC estimates of these elements. Floor Finishes and Ceiling Finishes have a lower standard deviation and identified as special positions in the carbon hotspot category. This implies very low risk and uncertainty involved in the EC estimates of Floor and Ceiling Finishes. On the other hand, Internal Walls and Partitions and Wall Finishes were identified as 'Special positions' in the carbon hotspot category and have a higher standard deviation which implies that there is moderate risk involved in the EC estimating of these elements.

EC per GIFA and CC per GIFA are the key data to estimate EC and CC using approximate estimating techniques when there is no detailed design of the elements is present. Even though the benchmarks are available for CC per GIFA in the form of published cost data books developed by construction practices and professional bodies, there is no comparable industry developed EC benchmarks to assist early stage EC estimating. The need for developing comparable EC benchmarks (e.g. EC per GIFA and EC per EUQ) is identified and highlighted in the study to facilitate dual currency appraisals (cost and carbon). Further, carbon planning process can be entertained and performed simultaneously by a Quantity Surveyor similar to cost planning process as pointed out by Ashworth and Perera (2015) if such EC benchmarks are available (see, Table 8.12). In this way, cost and carbon management can be achieved simultaneously in a more efficient way.

Table 8.12: EC planning process in parallel to cost planning as per NRM1

Modified from: Ashworth and Perera (2015)

RIBA Plan of Work 2013	Cost Plans (as per NRM1)	Comparable EC plans
1 Preparation	Order of Cost Estimate	EC Estimate for the Building
2 Concept Design	Formal Cost Plan 1	EC Plan 1
3 Developed Design	Formal Cost Plan 2	EC Plan 2
4 Technical Design	Formal Cost Plan 3 Bill of Quantities Post Tender Estimate	Full Pre-tender EC Plan for the Building
5 Specialist Design		Refine EC Plan for specialist design
6 Construction		EC management by the builder
7 Use and Aftercare		EC management by the Facilities Manager

8.5. Summary

The EC model refuted the hypotheses that suggest there is no relationship between EC and wall to floor ratio and EC and number of basements. Similarly, the cost model refuted the hypotheses, which suggest that there is no relationship between the building cost and the building height and cost and the circulation space ratio. The findings suggest that 48.1% of the variation in EC per GIFA is attributable to Wall to Floor ratio and a number of basements and 50.6% of the variation in CC per GIFA is attributable to building height and circulation space ratio. The remaining variation in both models is attributable to other variables that are not modelled here. These models aim at assisting designers during the early design stages of construction projects to select an optimum design solution.

The knowledge of carbon and cost hotspots informs designers about the building elements that need more attention during the design stages that have high reduction potential. The findings suggest that 80% of the EC emissions are caused by 43% of the elements and 80% of the cost is incurred by 50% of the elements on average, which does not comply with the 80:20 Pareto rule. However, it was also found that the cost hotspots are responsible for 80% of the EC emissions while carbon hotspots are responsible for 72% of the CC of the construction on average. Even though the all the carbon hotspots except for Upper Floors were identified as cost hotspots the level of carbon and cost significance of each element is different which makes the cost and carbon optimisation complex. More case studies on alternative design options will provide insights to this issue.

The intensity of the risk or uncertainty in the EC and CC estimates was ascertained based on the hotspot category and the standard deviation of the element rates. Accordingly, it was also found that there is low risk or less uncertainty when estimating EC of Substructure, Upper Floors and Services; very low risk or uncertainty for Floor Finishes and Ceiling Finishes; moderate risk for Internal Walls and Partitions and Wall Finishes. On the other hand, there is low risk or uncertainty in the CC estimates of Roof and Services; moderate risk on Stairs, Internal Walls and Partitions, Wall Finishes and Fittings, Furnishings and Equipment; high risk on Windows and External Doors. These findings are important in presenting the early stage estimates to the client so that necessary allowance for uncertainty is accounted in the estimates to cover insufficient design data.

9. Conclusions and Recommendations

9.1. Introduction

The study aimed at developing cost and carbon models for early design stage decision-making by collecting historical project data. The study objectives are reviewed in this chapter by discussing the method used to achieve each objective and summarising the outcome of each objective. Different data collection and analysis techniques were employed including archival analysis, Delphi-based expert forum, document review and statistical analysis (correlations and linear regressions), to achieve the objectives as discussed in Chapter 5 and Chapter 6. These are discussed briefly in the review of the study objectives. The key findings presented in Chapter 8 are also summarised here leading to the key conclusions of the research. Further, the contribution to knowledge in terms of theory, practice and application is discussed here. This is followed by key limitations of the research and recommendations to the industry and professional bodies to improve research in this area. This chapter and the overall thesis conclude by identifying three key future research directions.

9.2. Review of Objectives of the Study

The aim of the study was achieved through seven objectives, which were presented in Chapter 1.

9.2.1. Review the Significance of Embodied and Operational Carbon in Building Construction Projects and Relevant Regulatory Requirements

This objective was achieved through an extensive literature review (Chapter 2) and answers the RQ1 (How significant are embodied and OC in building projects and how are they regulated?) Literature suggests that generally, OC contributes a significant proportion of the total emissions (70%-80%) from buildings, hence, is regulated (for instance, Part L of Building Regulations of the UK). Further, zero carbon agenda of the UK government aspires to achieve zero OC in all the new building from 2019. On the other hand, case studies by Ramesh et al. (2010) suggest that EC increases when moving from a conventional building to a low or zero carbon building. However, EC is not regulated by any means at present. Therefore, there is a need to manage EC to control the rise in the emission levels to attain the emission reduction targets prescribed in the Kyoto Protocol and the UK Climate Change Act (80% reduction in emission levels by 2050). These targets became more serious with the latest Climate Conference COP21 with 195 countries committing to reduce emission levels. EC management requires EC estimating throughout the project and it is argued that the reduction potential is high during the early stages of design (RICS, 2014) (see, Figure 2.9 in Section 2.5). However, estimating EC during early design stages is challenging and there is no industry developed standards or benchmarks to assist EC estimating during early stages of design. Even though there are estimating practices, tools and techniques pertinent to estimating EC these are still in the early stages of development. In fact, robust early design stage EC estimating tools are scarce.

9.2.2. Evaluate the Existing Carbon Estimating Practices, Tools and Techniques, their Functions, Outputs and Limitations

This objective answers the RQ2 (What are the existing EC estimating tools, methods, their functions, outputs and limitations?) and is achieved through the literature review and the evaluation of the existing EC tools and techniques (Chapter 3). Carbon emissions can be estimated from Cradle to Grave (from the raw material extraction up to the end of life of the building) which is called the system boundary of the estimate (see, Figure 2.4 in Section 2.3). RICS (2014) guidance note assists in estimating EC during different stages of a project by obtaining data from the project and EC and other design specific data from databases such as ICE, DEFRA and BCIS (see, Section 2.6 and 2.7). However, EC estimating is affected by five key factors including system boundary, the method of estimating, assumptions, data sources used and element classification adopted in the analysis (Dixit et al., 2010, Clark, 2013, Ekundayo et al., 2012) (see, Section 0). Hence, findings of the past research are not always consistent and directly comparable. Therefore, the need to define and explicitly state all of the identified five key factors affecting carbon estimating was highlighted for the knowledge to be transferable.

In addition, operational and EC estimating tools were reviewed under two key themes namely early stage (up to the Conceptual Stage of RIBA Plan of Work 2013) and detailed stage (see, Sections 3.4 and 3.5). The review revealed that carbon estimating tools ranges from simple easy to use tools to complex and comprehensive tools when moving from early stages to detailed stages of design. The predictions of the early design stage tools reported to higher CV (lower prediction accuracy) due to high uncertainty of designs during early stages while detailed design stage tools require project-specific inputs and specification information for more accurate predictions. However, most of the detailed stage tools are in the form of software packages and are available for purchase. Yet, tools that integrate both cost estimating and EC estimating rarely exist (unless the tools can operate in a BIM platform), especially during the early design stage, which could lead to more rational decisions.

Consequently, the need to develop early stage estimating tool to predict EC and CC was identified. The use of parametric cost models to estimate cost during early stages of projects has been proven successful. Therefore, the same approach was adopted in EC estimating to make it more approachable and parallel to cost estimates (see, Section 3.10). Integration of theories of design economics with EC estimating was conceptualised into developing a model for EC estimating. Consequently, design parameters of buildings such as morphological parameters (plan shape, storey height, total height and the like) and quality parameters (quality of services and quality of finishes) were used to formulate a linear model for predicting EC.

9.2.3. Identify and Analyse the Carbon-Intensive Elements in Buildings

This objective answers the research question RQ3 (What are the carbon-intensive elements or carbon hotspots in buildings?). In order to rationalise the number of predictor variables to be used in the model, the most influential design variables were identified by analysing the carbon hotspots or the carbon critical elements of the buildings. This was done by collecting data from historical projects (office buildings only) from four different sources (Dataset 1, Dataset 2, Dataset 3 and WRAP dataset – see Section 6.3) and estimating EC of the Dataset 3 using EC data from the UK Building Blackbook, Dataset 1, and Dataset 2 (see, Section 5.7). EUR for Substructure, Frame and Upper Floors were obtained from Dataset 1; EURs for Fitting, Furnishings and Equipment and Services were obtained from Dataset 2; EURs of the rest of the elements were developed from the UK Building Blackbook. Dataset 3 was validated using the WRAP dataset to ensure the consistency of the developed data (see, Section 5.7.3).

Table 9.1: Carbon and cost hotspot categories

Hotspot Category	Lead position	Special position	Remainder position
Both Carbon and Cost	Substructure, Frame, External Walls, Services	Internal Walls and Partitions, Wall Finishes, Floor Finishes, Ceiling Finishes	-
Carbon	Upper Floors	Roof, Windows and External Doors	Stairs, Internal Doors, Fittings, Furnishings and Equipment
Cost	Roof, Windows and External Doors	Upper Floors, Stairs, Internal Doors, Fittings, Furnishings and Equipment	-

The Pareto Principle (80:20 rule) was used to identify the carbon and cost hotspots of the developed sample (Dataset 3). The building elements contributing up to the 80% of EC and CC of the buildings in descending order of intensity were identified and marked as ‘hotspots’ (see, Section 6.2). Accordingly, Substructure, Frame, External Walls, Roof and Services, were identified as both carbon and cost hotspots in the whole sample. Further, Upper Floors was also identified as a carbon hotspot and External Windows and Doors and Floor Finishes were identified as cost hotspots. In addition, elements were classified into three types according to their position, namely ‘Lead Position’ (elements that were identified as hotspots in more than (or equal to) 80% of the buildings in the sample), ‘Special Position’ (elements that were found as hotspots in less than 80% of the buildings in the sample) and ‘Remainder Position’ (elements that were not identified as hotspots in any of the buildings in the sample) which are presented in Table 9.1 (modified from Table 6.10). The findings alert designers of the key building elements, which require focus during the design development to achieve a high reduction in either carbon or cost or to achieve an optimum balance between the both.

9.2.4. Investigate the Relationship between Embodied Carbon and Building Design Variables and Capital Cost and Building Design Variables

The research question RQ4 (Are there statistically significant associations between EC and design variables of buildings?) is addressed by this objective. Correlation between EC and building design variables (quantitative) was analysed using Pearson's correlation at 0.05 significance level (95% confidence). Table 9.2 and Table 9.3 summarises the correlation coefficients obtained from the analysis of Pearson's correlation. Statistically significant relationships were found between EC and certain design variables including GIFA, Building Height and Faced Area. These correlation coefficients remain significant at 0.01 significance level, which is impressive. Very similar results were obtained for CC, which makes it comparable. On the other hand, EC per GIFA correlate with Wall to Floor ratio and Circulation Ratio (the correlation between EC per GIFA and Wall to Floor Ratio was also significant at 0.01 significance level). CC per GIFA correlate with Building Height, Wall to Floor ratio, and Circulation Ratio at 0.05 significance level (however, the correlations were not significant at 0.01 significance level).

Table 9.2: Correlations between design variables and EC and CC

		GIFA	Building Height	Façade Area	Circulation Ratio
EC	Pearson Correlation	.985**	.513**	.862**	-.041
	Sig. (2-tailed)	.000	.001	.000	.821
	N	41	41	41	33
CC	Pearson Correlation	.969**	.535**	.868**	-.010
	Sig. (2-tailed)	.000	.000	.000	.955
	N	41	41	41	33

**. Correlation is significant at the 0.01 level (2-tailed).

Table 9.3: Correlations between design variables and EC per GIFA and CC per GIFA

		Building Height	Wall to Floor Ratio	Circulation Ratio
EC per GIFA	Pearson Correlation	.306	.523**	.360*
	Sig. (2-tailed)	.052	.000	.039
	N	41	41	33
CC per GIFA	Pearson Correlation	.389*	.322*	.391*
	Sig. (2-tailed)	.012	.040	.024
	N	41	41	33

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

9.2.5. Investigate the Relationship between the Embodied Carbon and the Capital Cost of Buildings

The research questions RQ5 (Is there a statistically significant association between the EC and the CC of buildings?) was answered by this objective. The correlation coefficient between EC and CC were analysed at the building level and element levels at 0.05 significance level (see, Section 7.6). The analysis at the building level suggested that EC and CC are positively correlated. Pearson's correlation indicated a very strong positive correlation of 0.977 between EC and CC at 99.99% confidence level (or 0.01 significance level). However, it was suspected that this correlation was caused by a third variable GIFA and hence, the correlation coefficient was calculated between EC per GIFA and CC per GIFA, which was found to be 0.645 at 0.01 significance level, which is a moderately strong correlation. The finding suggests that it is possible to achieve lower cost and lower EC at the same time due to the positive association between the EC per GIFA and the CC per GIFA.

However, when investigating EC and CC relationships at element levels different results were found (see, Table 6.50). EC per GIFA and CC per GIFA of Upper Floors and Internal Walls and Partitions were very strongly correlated (>0.80 at 0.01 significance level) while Substructure, Frame, External Walls, and Internal Doors were strongly correlated (between 0.60 and 0.79 at 0.01 significance level). Correlation between EC per GIFA and CC per GIFA was moderate (between 0.40 and 0.59 at 0.05 significance level) in External Windows and Doors and Floor Finishes. Therefore, the findings suggest that is possible to reduce both EC and CC in most of the elements at the same due to its positive correlation, hence, both EC and CC could be optimised simultaneously in office buildings.

9.2.6. Develop Models for Predicting Embodied Carbon and Capital Cost during Early Design Stages

The research question RQ6 (How can an early design stage EC prediction model be developed using design variables of buildings?) was answered by this objective. The objective covers two sub-processes including the development of services and finishes quality Indices and the regression analysis. Design quality indices were developed for internal finishes and services quality to transform the qualitative parameters into quantitative parameters by assigning an ordinal level scale to facilitate regression analysis. A finishes quality index was developed from a conceptual finishes quality index which was verified through a Delphi-based expert forum and a three-tiered finishes quality index was developed for the study for Wall (internal), Floor and Ceiling finishes of office buildings (see, Table 6.11 in Section 6.3.1). Finishes quality of each building was identified as Basic, Moderate or Luxury for Wall, Floor and Ceiling Finishes based on the type of finish used in the building. The final finishes quality index was derived using a weighted average method where the area finished was multiplied by the calculated finishes index and the sum was obtained (see, Table 6.12 and Table 6.13 in Section 6.3.1).

A services quality index was developed by reviewing services quality levels proposed in various price books (refer, Table 5.27). Among which, the services quality levels proposed in the Spon's Mechanical and Electrical Services Price book (Davis Langdon Consultancy, 2014) was considered appropriate due to its adaptability to the study which had three levels of services quality. Further, the services quality levels proposed in the Spon's Mechanical and Electrical Services Price book was improved by adding another level and each quality level was subdivided into two such as 'with lift installations' and 'without lift installations' (see, Table 6.14 in Section 6.3.2). In this way the Finishes Quality Index and the Services Quality Index was developed objectively. Then, the regression analysis was performed to formulate the EC and CC models after deriving the finishes quality and services quality of buildings using the developed indices.

The EC model is presented in Equation 8.1, which explains 48.1% of the variation in EC per GIFA attributable to Wall to Floor ratio and the number of basements. EC per GIFA is increased by 164 $kgCO_2/m^2$ when Wall to Floor ratio is increased by one unit for a given number of basements and adding a basement will increase EC per GIFA by 68 $kgCO_2/m^2$ for a given Wall to Floor ratio. The model was statistically significant and all the variables were significant at 0.05 significance level. Circulation Space Ratio was not significant in the model even though it significantly correlated with EC per GIFA in the bivariate analysis.

On the other hand, CC per GIFA model was presented in Equation 8.2 explains 50.6% of the variation in CC per GIFA, which is a better fit than the EC per GIFA model. The model suggests that the CC per GIFA increases by £18/m^2 for every meter increase in the building height, for a given circulation space ratio. Similarly, the CC per GIFA increases by £8/m^2 for every percentage increase in circulation space ratio for a given building height. However, remaining 49.4% of the change is attributable to other design variables, which were not modelled in the study. In contrast, the bivariate analysis shows that all the design variables (building height, wall to floor ratio, circulation space ratio and no. of basements) significantly correlate with CC per GIFA at a 0.05 significance level. However, during the model fit, not all variables were found to be significant resulting in the elimination of the insignificant variables.

9.2.7. Validate the Decision Support Models with Real-Time Construction Projects

This objective answers the research question RQ7 (How can the developed EC and CC models be validated?). The developed models were validated by assessing their closeness to fit and prediction accuracy with internal data and external data (see, Section 7.1). EC per GIFA model has an R^2 value of 48.1%, which is satisfactory. The CV of the model was found to be 10.65% when predicting for internal data, which is within the desired CV range for early stage estimating. The difference in the estimates to that with the actual EC per GIFA ranges from -25% to 20%. Further, the predictions range from -20% to 11% when predicting for external data (see, Section 7.3.1). Therefore, the model performs fairly well with the data outside of the model and it can be used to predict EC of real time projects.

Similarly, CC per GIFA model fit was found to be 50.6%, which is better than the EC model fit. The CV of the model was 13.2% when predicting for internal data, which is within the desired CV range for early stage estimating though the CV deteriorated when predicting for external data to 24.5%. However, after filtering the buildings up to 6 storeys from the external data CV was improved to 9.85%, which implies the model performs well within the given range with buildings up to 6 storeys (see, Section 7.3.4).

9.3. Contributions to Knowledge

The findings of this research contribute to the body of knowledge of carbon management in buildings both in theoretical and in practical terms which are discussed as follows:

9.3.1. Contributions to Theory

The research findings on design variables and EC relationships add knowledge to the theory of design economics. Theory of design economics is well established in terms of construction cost (see, Seeley, 1996, Ashworth and Perera, 2015, Dell'Isola and Kirk, 1981, Collier, 1984, Morton and Jaggar, 1995, Robinson and Symonds, 2015). However, the other component of the dual currency of construction projects, which is carbon, has not been explored. Hence, the findings of the research provide a different dimension to the design economics theory. The behaviours of both CC and EC with respect to the changes in the design variables in building designs are captured in the selected sample and compared. This knowledge helps to identify the relationship between cost and EC and the design variables affecting both cost and EC (refer to Section 9.2.4). In addition, the strength of the relationships highlights the significant design variables that affect CC and EC.

Further, the relationships between CC and EC at different levels, including building and element level, demonstrate the interaction between the two. In particular, the element level correlations highlight the elements in which both

CC and EC reductions are attainable due to the identified positive correlations (see, Section 9.2.5) which is absent in the literature.

In addition, the methodology adopted in the study is also considered as a contribution to the theory as this research is relatively new in the field of EC estimating and no similar research are reported. Hence, the methodology proposed in this study can be replicated at different contexts such as in different locations and with different building types.

9.3.2. Contributions to Practice

The most significant applied contribution of this research is the early design stage EC prediction model. Although it is not currently the trend of the UK construction industry, it is expected to become one of the future trends. The EC prediction model is not self-sufficient and there is a need for a CC prediction model to facilitate dual currency evaluation. However, both prediction models developed from the selected sample have limited application in relation to the type and the number of storeys of the buildings. Therefore, EC and CC prediction models for different types of building with different design features will have to be formulated.

In addition, the findings related to the carbon hotspots contribute to practice during the detailed design stage. Designers can be well informed of the building elements that require more attention during the detail design stages, with the knowledge of carbon hotspots, to realise substantial reductions in the EC of buildings. Further, the mapping of cost and carbon hotspots based on their positions (Lead, Special and Remainder Positions) helps to achieve an optimum balance between the CC and the EC of building designs. Finally, the knowledge of elements whose impacts are negligible allows designers to work with these elements more liberally compared to the others.

9.4. Limitations of the Research

The major limitation of the research was the lack of standalone EC databases. Therefore, the study sample for the statistical analysis was obtained from two different sources (the primary data from QS consultancy practices and the secondary data from a special database from another QS consultancy practice) and validated using another dataset from an independent source (WRAP EC Database). However, since the t-Test suggests that there is a significant difference in the EC estimates of Superstructure Structural between Dataset 3 and WRAP dataset. Therefore, there is some form of ambiguity concerning the estimate of Superstructure Structural EC. Hence, the reliability of the estimate of Superstructure Structural could not be verified without any additional information on element specification, which is unknown (see, section 5.7.3).

Inability to seek clarifications about datasets obtained from online databases and BoQs obtained from QS consultancy practices was another challenge faced during the research. For instance, the cost analyses and EC analyses obtained from BCIS, special databases and WRAP EC Database deemed to be assumed as free from errors and manipulations. Further, cost analyses could not be adjusted for certain factors such as mark-up. Furthermore, errors in measurements were noticed in some elements in the BoQs obtained from QS consultancy practices, hence, respective data could not be used for the model validation.

The sample size is another limitation of the study. A larger sample could not be obtained due to the lack of EC databases and only a limited set of data met the data requirement of the study. However, the best available data from different sources were obtained and the study sample was validated.

Some key limitations of the models include: models work best at predicting EC and CC of up to 6 storeys as the sample comprises buildings from 1 to 6 storeys. Models should be adjusted to accommodate the changes in the industry, especially, for the method of manufacturing of construction materials as the assumptions used in the EC databases were adopted in the study by default. Similarly, cost data were rebased to 2016 1Q and a location index of 100. Hence, the cost model predictions need to be adjusted for time and location when predicting the CC of a future project using appropriate indices. In addition, cradle-to-gate system boundary is also a limitation of the models. Even though, it is desirable for the models to cover cradle-to-grave boundary to provide a holistic perspective on designs, it is challenging due to limited EC data. Another limitation of regression models is that they are static models. Regression analysis has to be performed again to derive a new model when new data become available (see, Section 8.2.4 more details).

Finishes and services indices were developed using qualitative data collection and analysis techniques due to the qualitative nature of the variables. Even though there could be possibilities for employing quantitative methods to develop this kind of indices, it is comparatively harder and requires significantly more time and participants. Furthermore, materials imported due to lower cost will have higher EC compared to locally sourced materials. This was not considered in the finishes index development due to the adoption of cradle-to-gate system boundary.

9.5. Recommendations

The key limitation being the lack of a standalone EC database, there is a serious need for publicly available industry governed EC databases to facilitate research in this area. The rising need for EC estimating of construction projects will require a standard practice or models to be in place for a systematic and effective day-to-day running of businesses. The proposed methodology can be adopted by construction businesses if there is an in-house EC database or an industry governed public standalone database such as BCIS. While there is WRAP EC Database, developed and maintained by WRAP and UK-GBC, the database lacks key design data of the projects. Therefore, it is recommended to WRAP and UK-GBC that they improve their existing database with more design data and promote the database so that the industry practices can effectively contribute to the database development which will facilitate research of this kind.

Further, it is recommended that the construction practices (or regulatory bodies such as RICS) develop and manage their in-house EC databases, which can contribute to their in-house research and development. Otherwise, it is recommended to the RICS that they incorporate EC analyses in the BCIS online cost database so that both cost and carbon information can be obtained for a particular project at the same time. Secondly, it is recommended that the BCIS make detailed specification information available (than what is available at present) for the users to allow in-depth studies. In addition, availability of cost analyses excluding mark-up will be an added advantage in standardising the base for the data.

Apart from the need for EC databases, there is a need for industry developed benchmarks for EC-EURs to encourage early stage EC estimating. The use of this kind of benchmarks can be realised during the preparation of early stage EC plans (see, Ashworth and Perera (2015) for the mapping of EC planning process to the NRM1 cost planning process). Especially, quantity surveyors trained in performing early stage cost planning can also produce EC estimates in parallel to cost estimates which lead to dual currency appraisals of construction projects provided that benchmarks for EC-EURs are available similar to CC-EURs.

There is also a need for a standardised finishes and services index for the finishes and services quality of the buildings so that it can be incorporated into the heuristic cost models. This will eliminate the subjectivity of the cost models and the definition of the quality of finishes and services will be universal within the region.

9.6. Further Research Directions

The thesis concludes with three directions for further research.

First, the research outputs (CC and EC models) can be developed into a scalable decision support system, which will constitute the cost and carbon models as the system driver. Such a system should be self-updating as new data are fed into the system, easily manageable and user-friendly. On the other hand, the integration of these models or similar models into a BIM platform can also be studied.

Secondly, similar research can be conducted in different contexts such as different countries and different types of buildings (high rise offices, retail buildings, domestic buildings etc.). The CC and EC data used are from the UK sources, however, EC of materials vary from one country to another. Therefore, there is a scope for similar research in different parts of the world so that the findings can be compared. Further, the function of the building also determines the cost and the EC of the building, hence, research in different types of building will add to the existing knowledge and will create new insights to the problem studied.

Finally, the system boundary of the research can be broadened to cover Cradle-to-Grave so that both EC and OC can be included in the analysis. Similarly, both CC and operation cost should be considered and the total cost of the project can be compared to the total carbon of the project. This final point is very important, as it would provide a holistic picture of the total cost savings and carbon savings of a proposed project.

References

Energy Performance of Buildings Directive 2002 (2002/91/EC).

UK Building regulations Approved Document 2013 (Part L1A).

UK Building regulations Approved Document 2013 (Part L2A).

AACE INTERNATIONAL. 2015. *Cost estimating model for buildings* [Online]. AACE International Available: http://www.aacei.org/resources/BuildingModel.shtml [Accessed 21 October 2015].

AAMODT, A. & PLAZA, E. 1994. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications,* 7**,** 39-59.

ADELI, H. & WU, M. 1998. Regularization neural network for construction cost estimation. *Journal of Construction Engineering and Management,* 124**,** 18-24.

AHRENS, C. D. 2007. *Meteorology today: an introduction to weather, climate, and the environment,* Belmont, CA, Thomson Brooks/Cole.

ALSHAMRANI, O. S. 2016. Construction cost prediction model for conventional and sustainable college buildings in North America. *Journal of Taibah University for Science*.

ANDERSON, J. 2011. Materials, products and carbon - CIRIA Carbon Leaders Briefings RP958 (No 3). UK: PE International.

ANDERSON, J., SHIERS, D. & STEELE, K. 2002. The green guide to specification an environmental profiling system for building materials and components. UK: BRE.

ANDERSON, T. & HAMMARBERG, L. 2015. *Consequences of implementing timber in medium high-rise office buildings.* MSc, Chalmers University of Technology.

ASHWORTH, A. 2010. *Cost studies of buildings,* Harlow, England and New York, Pearson.

ASHWORTH, A. & PERERA, S. 2015. *Cost studies of buildings,* Oxon, Routledge.

ASHWORTH, A. & SKITMORE, R. M. 1983. Accuracy in Estimating. *CIOB Occasional Paper No. 27.* London: CIOB.

ASIEDU, Y. & GU, P. 1998. Product life cycle cost analysis: State of the art review. *International Journal of Production Research,* 36**,** 883-908.

ATKINS. 2014. *Carbon tools* [Online]. Available: http://www.atkinsglobal.co.uk/en-GB/corporate-sustainability/an-environment-with-a-future/a-low-carbon-economy/lower-carbon/carbon-tools [Accessed 27 February 2017].

AYAZ, E. & YANG, F. 2009. *Zero carbon isn't really zero: Why EC in materials can't be ignored* [Online]. Available: http://www.di.net/articles/zero_carbon/ [Accessed 27 February 2017].

BBC. 2013. *Why the industrial revolution happened in Britain* [Online]. Available: http://www.bbc.co.uk/history/0/20979973 [Accessed 27 February 2017].

BCIS 2013. *Comprehensive building price book 2013. Major works,* Poole, Dorset, Poole, Dorset : BCIS / RICS.

BOWLBY, R. L. & SCHRIVER, W. R. 1986. Observations on productivity and composition of building construction output in the United States, 1972–82. *Construction Management and Economics,* 4**,** 1-18.

BRANDT, A. R. 2012. Variability and uncertainty in life cycle assessment models for greenhouse gas emissions from Canadian oil sands production. *Environmental Science & Technolog,* 46**,** 1253-1261.

BRE. 2009. *National calculation method* [Online]. Available: http://www.ncm.bre.co.uk/ [Accessed 27 February 2017].

BREEAM 2013. Green guide calculator online tool guidance. UK.

BUILD CARBON NEUTRAL. 2007. *Estimate the embodied CO2 of a whole construction project* [Online]. Available: http://buildcarbonneutral.org/ [Accessed 30th January 2014].

BUILDING RESEARCH ESTABLISHMENT 2014. iSBEM: An Interface for SBEM (Simplified Building Energy Model) (user guide). UK: Communities and Local Government.

CALZADILLA, A., REHDANZ, K., BETTS, R., FALLOON, P., WILTSHIRE, A. & TOL, R. S. J. 2013. Climate change impacts on global agriculture. *Climatic Change,* 120**,** 357-374.

CARBON TRUST 2010. Carbon: Reducing the footprint of the construction process. UK: Carbon Trust.

CATTO, I. 2008. Carbon zero homes UK style. *Renewable Energy Focus,* 9**,** 28-29.

CB RICHARD ELLIS 2009. Who pays for green? The economics of sustainable buildings. UK.

CHAKRAPANI, C. 2011. Statistical reasoning vs. magical thinking, shamanism as statistical knowledge: is a sample size of 30 all you need? *Vue.*

CHEN, T., BURNETT, J. & CHAU, C. 2001a. Analysis of EE use in the residential building of Hong Kong. *Energy,* 26**,** 323-340.

CHEN, T. Y., BURNETT, J. & CHAU, C. K. 2001b. Analysis of EE use in the residential building of Hong Kong. *Energy,* 26**,** 323-340.

CHENG, M.-Y., TSAI, H.-C. & HSIEH, W.-S. 2009. Web-based conceptual cost estimates for construction projects using Evolutionary Fuzzy Neural Inference Model. *Automation in Construction,* 18**,** 164-172.

CLARK, D. H. 2013. *What Colour is your Building?: Measuring and reducing the energy and carbon footprint of buildings,* UK, RIBA Enterprises Limited.

CLAYTON, M. J. 1997. Delphi: a technique to harness expert opinion for critical decision-making tasks in education. *Educational Psychology,* 17**,** 373-386.

COHEN, J. 1988. *Statistical power analysis for the behavioral sciences,* Hillsdale, N.J., Erlbaum Associates.

COLE, R. J. & KERNAN, P. C. 1996. Life-cycle energy use in office buildings. *Building and Enrironment,* 31**,** 307-317.

COLLIER, K. 1984. *Estimating construction costs: a conceptual approach,* Reston, Virginia, Reston Publishing.

COMMITTEE ON CLIMATE CHANGE 2013. Fourth carbon budget review – part 2 the cost-effective path to the 2050 target. UK: Committee on Climate Change.

COMMITTEE ON CLIMATE CHANGE 2014. Meeting carbon budgets – 2014 progress report to parliament. UK: Committee on Climate Change.

CORBETTA, P. 2003. *Social research: theory, methods and techniques,* London, SAGE.

CRAIGHEAD, G. 2009. High-rise building definition, development, and use. *High-Rise Cecurity and Fire Life Safety, 3rd ed.; Butterworth-Heinemann: Boston, MA, USA*.

ČUČEK, L., KLEMEŠ, J. J. & KRAVANJA, Z. 2012. A Review of Footprint analysis tools for monitoring impacts on sustainability. *Journal of Cleaner Production,* 34**,** 9-20.

DAVIS LANGDON CONSULTANCY 2014. *Spon's mechanical and electrical services price book 2014,* London, Taylor & Francis.

DELERS, A. 2015. *Pareto's Principle: Expand your business!*, 50Minutes.com.

DELL'ISOLA, A. J. & KIRK, S. J. 1981. *Life cycle costing for design professisonals,* USA, McGraw Hill

DEPARTMENT FOR COMMUNITIES AND LOCAL GOVERNMENT 2008. Definition of zero carbon homes and non-domestic buildings - consultation. UK: Department for Communities and Local Government.

DEPARTMENT FOR ENVIRONMENT FOOD & RURAL AFFAIRS. 2015. *Greenhouse gas conversion factor repository* [Online]. Crown Available: http://www.ukconversionfactorscarbonsmart.co.uk/ [Accessed 27 February 2017].

DEPARTMENT OF COMMUNITIES AND LOCAL GOVERNMENT 2014. Fire and rescue authorities opeartional guidance, GRA 3.2 fighting fires - in high rise buildings. UK: The Stationery Office.

DEPARTMENT OF ENERGY & CLIMATE CHANGE 2012. The Government's Standard Assessment Procedure for Energy Rating of Dwellings. Watford: BRE.

DIXIT, M. K., FERNÁNDEZ-SOLÍS, J. L., LAVY, S. & CULP, C. H. 2010. Identification of parameters for EE measurement: A literature review. *Energy and Buildings,* 42**,** 1238-1247.

DIXIT, M. K., FERNÁNDEZ-SOLÍS, J. L., LAVY, S. & CULP, C. H. 2012. Need for an EE measurement protocol for buildings: A review paper. *Renewable and Sustainable Energy Reviews,* 16**,** 3730-3743.

DTI 2003. Energy white paper our energy future - creating a low carbon economy. UK: State for Trade and Industry.

DTI 2007. Meeting the energy challenge - A white paper on energy. UK: State for Trade and Industry.

ECO INFO. 2012. *When Will We Run Out Of Fossil Fuels?* [Online]. Available: http://www.eco-info.net/fossil-fuel-depletion.html [Accessed 27 February 2017].

ECOINVENT ASSOCIATION. 2015. *The ecoinvent Database* [Online]. Available: http://www.ecoinvent.org/database/ [Accessed 23 July 2015].

EKUNDAYO, D., PERERA, S., UDEAJA, C. & ZHOU, L. 2012. Carbon review and qualitative comparison of selected carbon counting tools. *RICS COBRA 2012 The Construction, Building and Real Estate Research Conference of the Royal Institution of Chartered Surveyors.* Las Vegas, Nevada USA: RICS.

ENVIRONMENT AGENCY. 2012. *Carbon calculator for construction projects* [Online]. Available: https://www.gov.uk/government/publications/carbon-calculator-for-construction-projects [Accessed 20th October 2014].

EUROPEAN COMMISSION 2011. Communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions a roadmap for moving to a competitive low carbon economy in 2050. Brussels: European Commission.

EVANS, J. D. 1996. *Straightforward statistics for the behavioral sciences,* Pacific Grove, CA, Brooks/Cole Publishing.

FIELD, A. P. 2013. *Discovering statistics using IBM SPSS : and sex and drugs and rock 'n' roll,* London, SAGE.

FRANKLIN & ANDREWS 2011. *Hutchins UK Building Blackbook: The Cost and Carbon Guide : Hutchins' 2011: Small and Major Works,* Croydon, Franklin & Andrews.

FRANKLIN AND ANDREWS 2012. *Griffiths complete building price book 2012 : a complete guide for estimating, checking and forecasting building and civil engineering costs,* Croydon, Croydon : Franklin + Andrews.

FRENCH, S., PAPAMICHAIL, N. & MAULE, J. 2009. *Decision behaviour, analysis and support,* Cambridge, UK & New York, Cambridge University Press.

GAMBHIR, A. & VALLEJO, L. 2011. The UK's progress towards its carbon budgets. London: Grantham Institute for Climate Change.

GRAY, D. E. 2014. *Doing research in the real world,* London, SAGE Publications.

GUBA, E. G. & LINCOLN, Y. S. 1994. Competing paradigms in qualitative research. *Handbook of qualitative research,* 2.

GUPTA, U. G. 1994. How case-based reasoning solves new problems. *Interfaces,* 24**,** 110-119.

HACKER, J. N., DE SAULLES, T. P., MINSON, A. J. & HOLMES, M. J. 2008. Embodied and OC dioxide emissions from housing: A case study on the effects of thermal mass and climate change. *Energy and Buildings,* 40**,** 375-384.

HAIR, J. F. 1998. *Multivariate data analysis,* Upper Saddle River, N.J, London, Prentice Hall.

HALCROW YOLLES 2010a. EC Sustainable Offices: a supplementary report. UK: South West Regional Development Agency.

HALCROW YOLLES 2010b. Sustainable offices – EC. South West Regional Development Agency.

HAMMOND, G. & JONES, C. 2008a. Inventory of carbon energy (ICE). UK: University of Bath.

HAMMOND, G. & JONES, C. 2011. A BSRIA guide EC The Inventory of Carbon and Energy (ICE). UK: BSRIA.

HAMMOND, G. P. & JONES, C. I. 2008b. EE and carbon in construction materials. *Proceedings of the Institution of Civil Engineers - Energy,* 161**,** 87-98.

HEGAZY, T. & AYED, A. 1998. Neural network model for parametric cost estimation of highway projects. *Journal of Construction Engineering and Management,* 124**,** 210-218.

HERNANDEZ, P. & KENNY, P. 2010. From net energy to zero energy buildings: Defining life cycle zero energy buildings (LC-ZEB). *Energy and Buildings,* 42**,** 815-821.

HITCHIN, R. 2013. CIBSE Research Report 9: EC and Building Services. UK: CIBSE.

HM TREASURY 2013. Infrastructure carbon review. London: HM Treasury.

HOX, J. J. & BOEIJE, H. R. 2005. Data collection, primary vs. secondary. *Encyclopedia of social measurement,* 1**,** 593-599.

HUTCHINS 2011. The UK Building Blackbook - The Cost and Carbon Guide. *In:* ANDREWS, F. (ed.). UK.

IBN-MOHAMMED, T., GREENOUGH, R., TAYLOR, S., OZAWA-MEIDA, L. & ACQUAYE, A. 2013. Operational vs. embodied emissions in buildings—A review of current trends. *Energy and Buildings,* 66**,** 232-245.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE 2001. Climate change, 2001: The Scientific basis: contribution of Working Group I to the third assessment report of the Intergovernmental Panel on Climate Change. *In:* HOUGHTON, J. T. (ed.). Cambridge: Cambridge University Press.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE 2007. Climate Change 2007: the physical science basis: contribution of working group I to the Fourth Assessment Report of the Intergovernmental Panel Climate Change. *In:* SOLOMON, S. (ed.). Cambridge: Cambridge University Press.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE 2012. *Managing the risks of extreme events and disasters to advance climate change adaption: special report of the Intergovernmental Panel on Climate Change,* New York, Cambridge, Cambridge University Press.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. *In:* STOCKER, T. F., QIN, D., PLATTNER, G.-K., TIGNOR, M. M. B., ALLEN, S. K., BOSCHUNG, J., NAUELS, A., XIA, Y., BEX, V. & MIDGLEY, P. M. (eds.). United Kingdom and New York, NY, USA,.

INTERNATIONAL ENERGY AGENCY 2006. World energy outlook 2006. France: Organisation for Economic Co-operation and Development and International Energy Agency.

JOHNSON, V. B. 2008. *Laxton's building price book 2008 : major & small works,* London, Laxton's.

KARANCI, H. 2010. *A comparative study of regression analysis, neural networks and case – based reasoning for early range cost estimation of mass housing projects.* Master of Science, Middle East Technical University.

KARSHENAS, S. 1984. Predesign cost estimating method for multistory buildings. *Journal of Construction Engineering and Management,* 110**,** 79-86.

KELLY, S., CRAWFORD-BROWN, D. & POLLITT, M. G. 2012. Building performance evaluation and certification in the UK: Is SAP fit for purpose? *Renewable and Sustainable Energy Reviews,* 16**,** 6861-6878.

KHOUKHI, M. & AL-MAQBALI, A. 2011. Stack Pressure and Airflow Movement in High and Medium Rise buildings. *Energy Procedia,* 6**,** 422-431.

KIM, G.-H., AN, S.-H. & KANG, K.-I. 2004a. Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Building and environment,* 39**,** 1235-1242.

KIM, G.-H., YOON, J.-E., AN, S.-H., CHO, H.-H. & KANG, K.-I. 2004b. Neural network model incorporating a genetic algorithm in estimating construction costs. *Building and Environment,* 39**,** 1333-1340.

KOCH, R. 2011. *The 80/20 Principle*, Nicholas Brealey Publishing.

KOUSKOULAS, V. & KOEHN, E. 2005. Predesign cost estimating function for buildings. *In:* SKITMORE, M. & MARSTON, V. (eds.) *Cost Modelling.* London: E & FN Spon.

LANGSTON, Y. L. & LANGSTON, C. A. 2008. Reliability of building EE modelling: an analysis of 30 Melbourne case studies. *Construction Management and Economics,* 26**,** 147-160.

LÉLÉ, S. M. 1991. Sustainable development: A critical review. *World Development,* 19**,** 607-621.

LEVERS, M.-J. D. 2013. Philosophical Paradigms, Grounded Theory, and Perspectives on Emergence. *SAGE Open,* 3.

LIN, C.-T. & LEE, C. G. 1991. Neural-network-based fuzzy logic control and decision system. *Computers, IEEE Transactions on,* 40**,** 1320-1336.

LIOR, N. 2008. Energy resources and use: The present situation and possible paths to the future. *Energy,* 33**,** 842-857.

LOCKIE, S. 2012. *EC – new guidelines from the RICS* [Online]. Available: http://www.fgould.com/uk-europe/articles/embodied-carbon-guidelines/ [Accessed 27 February 2017].

LUO, Z., YANG, L. & LIU, J. 2015. EC Emissions of Office Building: A Case Study of China's 78 Office Buildings [online]. *Building and Environment*.

LÜTZKENDORF, T., FOLIENTE, G., BALOUKTSI, M. & WIBERG, A. H. 2014. Net-zero buildings: incorporating embodied impacts. *Building Research & Information,* 43**,** 62-81.

MARSZAL, A. J., HEISELBERG, P., BOURRELLE, J. S., MUSALL, E., VOSS, K., SARTORI, I. & NAPOLITANO, A. 2011. Zero Energy Building – A review of definitions and calculation methodologies. *Energy and Buildings,* 43**,** 971-979.

MCGARRITY, R. J. 1988. *Parametric estimating: an equation for estimating buildings.* MSc, Georgia Institute of Technology.

MCLEOD, R. S., HOPFE, C. J. & REZGUI, Y. 2012. An investigation into recent proposals for a revised definition of zero carbon homes in the UK. *Energy Policy,* 46**,** 25-35.

MEBRATU, D. 1998. Sustainability and sustainable development: Historical and conceptual review. *Environmental Impact Assessment Review,* 18**,** 493-520.

MICHALEWICZ, Z. 2013. *Genetic algorithms+ data structures= evolution programs*, Springer Science & Business Media.

MILES, J. & SHEVLIN, M. 2001. *Applying regression & correlation: a guide for students and researchers,* London, SAGE.

MONAHAN, J. & POWELL, J. C. 2011. An EC and energy analysis of modern methods of construction in housing: A case study using a lifecycle assessment framework. *Energy and Buildings,* 43**,** 179-188.

MONCASTER, A. M. & SYMONS, K. E. 2013. A method and tool for 'cradle to grave' EC and energy impacts of UK buildings in compliance with the new TC350 standards. *Energy and Buildings,* 66**,** 514-523.

MONTGOMERY, D. C., VINING, G. G. & PECK, E. A. 2012. *Introduction to linear regression analysis,* Oxford, Wiley-Blackwell.

MORTON, R. & JAGGAR, D. 1995. *Design and the economics of building,* London, Spon.

MUNNS, A. K. & AL-HAIMUS, K. M. 2000. Estimating using cost significant global cost models. *Construction Management and Economics,* 18**,** 575-585.

NATIONAL RESEARCH COUNCIL 2011. *Climate Stabilization Targets: Emissions, Concentrations, and Impacts over Decades to Millennia,* United States of America, The National Academies Press.

OSMANI, M. & O'REILLY, A. 2009. Feasibility of zero carbon homes in England by 2016: A house builder's perspective. *Building and Environment,* 44**,** 1917-1924.

OXFORD DICTIONARY 2010. Oxford Dictionary of English. UK: Oxford University Press.

PALIWAL, M. & KUMAR, U. A. 2009. Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications,* 36**,** 2-17.

PAN, W. & GARMSTON, H. 2012. Building regulations in energy efficiency: Compliance in England and Wales. *Energy Policy,* 45**,** 594-605.

PATTON, M. Q. 2015. *Qualitative research and evaluation method,* USA, SAGE Publicayions.

PE INTERNATIONAL. 2014. *GaBi software* [Online]. Available: http://www.gabi-software.com/uk-ireland/index/ [Accessed 27 February 2017].

PEURIFOY, R. L. & OBERLENDER, G. D. 2002. Estimating construction costs. 5th ed. Boston, Mass, London: McGraw-Hill.

PHAOBUNJONG, K. 2002. *Parametric cost estimating model for conceptual cost estimating of building construction projects.* PhD, University of Texas at Austin.

PHLORUM. 2011. *Embodied CO2 Estimator* [Online]. Available: http://eco2.phlorum.com/calculator/save [Accessed 20th October 2014].

PICKEN, D. & ILOZOR, B. 2015. The relationship between building height and construction costs. *In:* ROBINSON, H., SYMONDS, B., GILBERTSON, B. & ILOZOR, B. (eds.) *Design economics for the built environment: impact of sustainability on project evaluation.* UK: Wiley-Blackwell.

POLIT-O'HARA, D., HUNGLER, B. P. & BECK, C. T. 2001. *Essentials of nursing research: methods, appraisal, and utilization,* Philadelphia, Lippincott.

POPPER, K. R. 1975. *Systematic Zoology,* 24**,** 233-243.

PRÉ CONSULTANTS. 2014. *Simapro LCA Software* [Online]. Available: http://www.pre-sustainability.com/lca-software-simapro?gclid=CLCqi76CprwCFdQPtAodTkoACw [Accessed 30th January 2014].

RAMESH, T., PRAKASH, R. & SHUKLA, K. K. 2010. Life cycle energy analysis of buildings: An overview. *Energy and Buildings,* 42**,** 1592-1600.

RAWLINSON, S. & WEIGHT, D. 2007. Sustainability: EC. *Building Magazine.*

RIBA. 2013. *RIBA plan of work 2013* [Online]. Available: http://www.ribaplanofwork.com/ [Accessed 20th October 2014].

RICS 2008a. *BCIS wessex estimating price book 2008*, BCIS Wessex.

RICS 2008b. Carbon management of real estate: a guide to best practice. 1 ed. UK: RICS.

RICS 2009. RICS oceania - expert witness registration scheme. UK: RICS.

RICS 2011. Non-domestic real estate climate change model: quantifying the sensitivity of carbon emissions and operating costs to regional climatic impacts. UK: RICS.

RICS 2012a. Elemental standard form of cost analysis principles, instructions, elements and definitions 4th (NRM) edition. London, UK: Building Cost Information Service.

RICS 2012b. Methodology to calculate EC of materials. 1 ed. UK: RICS.

RICS 2014. Methodology to calculate EC. 1 ed. UK: RICS.

RICS 2015. Assessment of professional competence quantity surveying and construction. London: RICS.

RICS 2016. Building Cost Information Services (BCIS). RICS.

ROBINSON, H. & SYMONDS, B. 2015. Theories and principles of design economics *In:* ROBINSON, H., SYMONDS, B., GILBERTSON, B. & ILOZOR, B. (eds.) *Design economics for the built environment: impact of sustainability on project evaluation.* UK: Wiley-Blackwell.

ROBSON, C. 2011. *Real world research: a resource for users of social research methods in applied settings,* Chichester, Wiley.

ROSCOE, J. T. 1975. *Fundamental research statistics for the behavioural sciences,* New York, Holt Rinehart & Winston.

SANSOM, M. & POPE, R. J. 2012. A comparative EC assessment of commercial buildings. *The Structural Engineer,* October**,** 38-49.

SAUNDERS, M., LEWIS, P. & THORNHILL, A. 2009. *Research methods for business students,* Essex, Pearson Education Limited.

SAWALHI, N. I. E. 2012. Modeling the parametric construction project cost estimate using fuzzy logic. *International Journal of Emerging Technology and Advanced Engineering,* 2.

SEELEY, I. H. 1996. *Building economics: appraisal and control of building design cost and efficiency,* Basingstoke, Macmillan.

SEO, K.-K., PARK, J.-H., JANG, D.-S. & WALLACE, D. 2002. Approximate estimation of the product life cycle cost using artificial neural networks in conceptual design. *The international journal of advanced manufacturing technology,* 19**,** 461-471.

SHAFIEE, S. & TOPAL, E. 2009. When will fossil fuel reserves be diminished? *Energy Policy,* 37**,** 181–189.

SHAFIQ, N., NURRUDIN, M. F., GARDEZI, S. S. S. & KAMARUZZAMAN, A. B. 2015. Carbon footprint assessment of a typical low rise office building in Malaysia using building information modelling (BIM). *International*

Journal of Sustainable Building Technology and Urban Development, 1-16.

SINGH, B. R. & SINGH, O. 2012. Global trends of fossil fuel reserves and climate change in the 21st century, fossil fuel and the environment. *In:* KHAN, D. S. (ed.) *InTech.* Croatia and China.

SONMEZ, R. 2004. Conceptual cost estimation of building projects with regression analysis and neural networks. *Canadian Journal of Civil Engineering,* 31**,** 677-683.

SPAIN, B. J. D. 2010. *Spon's first stage estimating handbook,* Hoboken, Hoboken : Taylor and Francis.

STATISTICS SOLUTIONS. 2016. *Assumptions of Linear Regression* [Online]. Available: http://www.statisticssolutions.com/assumptions-of-linear-regression/ [Accessed 27 February 2017].

STERN, N. H. 2007. *The economics of climate change: the Stern review,* Cambridge, Cambridge University Press.

STURGIS ASSOCIATES 2010. Carbon profile: ropermark place. UK: Sturgis carbon profiling.

STURGIS, S. & ROBERTS, G. 2010. Redefining zero: Carbon Profiling as a solution to whole life carbon emission measurement in buildings. UK: RICS.

SUNDARAKANI, B., DE SOUZA, R., GOH, M., WAGNER, S. M. & MANIKANDAN, S. 2010. Modeling carbon footprints across the supply chain. *International Journal of Production Economics,* 128**,** 43-50.

SUTRISNA, M. Research methodology in doctoral research: understanding the meaning of conducting qualitative research. Proceedings of the Association of Researchers in Construction Management (ARCOM) Doctoral Workshop held in Liverpool John Moores University. Conducted by ARCOM Liverpool, UK: ARCOM, 2009.

TARGET ZERO 2012. Guidance on the design and construction of sustainable, low carbon office buildings. UK: Target zero.

TAS, E. & YAMAN, H. 2005. A building cost estimation model based on cost significant work packages. *Engineering, Construction and Architectural Management,* 12**,** 251-263.

TATA STEEL 2014. Steel construction EC. UK.

THE GREEN CONSTRUCTION BOARD 2013. Low carbon routemap for the uk built environment. London: The Green Construction Board, WRAP, The Climate Centre and ARUP.

THE QUALITY OF LIFE POLICY GROUP 2007. Blueprint for a green economy,. London.

TORCELLINI, P., PLESS, S. & LEACH, M. 2014. A pathway for net-zero energy buildings: creating a case for zero cost increase. *Building Research & Information,* 43**,** 25-33.

TROCHIM, W. M. K. 2006. *Deduction & Induction* [Online]. Available: http://www.socialresearchmethods.net/kb/dedind.php [Accessed 27 February 2017].

TURBAN, E., SHARDA, R. & DELEN, D. 2011. *Decision support and business intelligence systems,* New Jersey, Pentice Hall.

UK-GBC 2008. The definition of zero carbon. UK: UK-GBC.

UK-GBC 2014a. Building zero carbon – the case for action examining the case for action on zero carbon non domestic buildings. UK: UK-GBC.

UK-GBC. 2014b. *Key statistics: construction industry and carbon emissions* [Online]. Available: http://www.ukgbc.org/resources/additional/key-statistics-construction-industry-and-carbon-emissions [Accessed 27 February 2017].

UNITED NATIONS 1992. United Nations Framework Convention on Climate Change (UNFCCC).

VAN DE VEN, A. H. & DELBECQ, A. L. 1974. The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes. *Academy of Management Journal,* 17**,** 605-621.

VANVOORHIS, C. R. W. & MORGAN, B. L. 2007. Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology,* 3**,** 43-50.

VICTORIA, M. F., PERERA, S., ZHOU, L. & DAVIES, A. 2015. Estimating EC: a dual currency approach. *Sustainable Buildings and Structures.* CRC Press.

WAN OMAR, W.-M.-S., DOH, J.-H. & PANUWATWANICH, K. 2014. Variations in EE and carbon emission intensities of construction materials. *Environmental Impact Assessment Review,* 49**,** 31-48.

WATSON, I. 1999. Case-based reasoning is a methodology not a technology. *Knowledge-Based Systems,* 12**,** 303-308.

WATSON, I. & MARIR, F. 1994. Case-based reasoning: A review. *The knowledge engineering review,* 9**,** 327-354.

WILLIAMS, P. L. & WEBB, C. 1994. The Delphi technique: a methodological discussion. *Journal of Advanced Nursing,* 19**,** 180-186 7p.

WILMOT, C. G. & CHENG, G. 2003. Estimating future highway construction costs. *Journal of Construction Engineering and Management,* 129**,** 272-279.

WOODCOCK, J., EDWARDS, P., TONNE, C., ARMSTRONG, B. G., ASHIRU, O., BANISTER, D., BEEVERS, S., CHALABI, Z., CHOWDHURY, Z., COHEN, A., FRANCO, O. H., HAINES, A., HICKMAN, R., LINDSAY, G., MITTAL, I., MOHAN, D., TIWARI, G., WOODWARD, A. & ROBERTS, I. 2009. Public health benefits of strategies to reduce greenhouse-gas emissions: urban land transport. *Health and Climate Change,* 2**,** 1930-1943.

WRAP & UK-GBC. 2014. *EC database* [Online]. WRAP. Available: http://www.wrap.org.uk/content/embodied-carbon-database [Accessed 27 February 2017].

WU, H., YUAN, Z., ZHANG, L. & BI, J. 2012. Life cycle energy consumption and CO2 emission of an office building in China. *The International Journal of Life Cycle Assessment,* 17**,** 105-118.

XU, L. D. 1994. Case based reasoning. *Potentials, IEEE,* 13**,** 10-13.

YIN, R. K. 2014. *Case study research design and methods,* USA, Sage Publications.

YOHANIS, Y. G. & NORTON, B. 2002. Life-cycle operational and EE for a generic single-storey office building in the UK. *Energy,* 27**,** 77-92.

YU, W. D. 2006. PIREM: a new model for conceptual cost estimation. *Construction management and economics,* 24**,** 259-270.

ZADEH, L. A. 1975. Fuzzy logic and approximate reasoning. *Synthese,* 30**,** 407-428.

ZADEH, L. A. 1994. Fuzzy logic, neural networks, and soft computing. *Communications of the ACM,* 37**,** 77-84.

ZERO CARBON HUB 2013. Zero carbon strategies for tomorrow's new homes. UK: Zero Carbon Hub.

ZERO CARBON HUB 2014a. Zero carbon homes and nearly zero energy buildings uk building regulations and eu directives. UK: Zero Carbon Hub,.

ZERO CARBON HUB. 2014b. *Zero carbon policy* [Online]. Available: http://www.zerocarbonhub.org/zero-carbon-policy/zero-carbon-policy [Accessed 27 February 2017].

ZERO CARBON HUB & SWEETT 2014. Cost analysis: meeting the zero carbon standard. UK: Zero Carbon Hub.

Appendix 1: Details of the Pilot Case from BCIS

Customer Service Centre, Brewery Lane - #21402

Rebased to 2Q 2014 (248; forecast)

Summary

Customer Service Centre, Brewery Lane

Location: Bridgend, Mid Glamorgan

Date: 27-Nov-2002

Building cost: £3,037,440 rebased

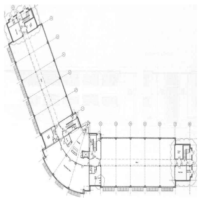
Cost/m²: £986 rebased

Floor area: 3,080m^2

Main construction: Steel framed

Storeys: 2

Level of analysis: Elemental



Ground Floor Plan
Image 1 of 2

© RICS 2014

Detail

Customer Service Centre, Brewery Lane

Building function: 320. - Offices

Type of work: New build

District: Ogwr (Bridgend)

Grid reference: SS9080

Receipt date: 27-Nov-2002

Base date: 17-Nov-2002

Date of acceptance: 10-Dec-2002

Date of possession: 13-Jan-2003

Project details: 2 storey office block together with external works including block paving, landscaping, services, drainage and site lighting.

Site conditions: Level car park site with good ground conditions. Excavation above water table. Unrestricted working space and access.

Market conditions: Competitive.

Project tender price index: 156 on 1985 BCIS Index Base

Client: Welsh Development Agency

Tender documentation: Bill of Quantities

Selection of contractor: Selected competition

Number of tenders issued: 6

Number of tenders received: 6

Contract: JCT Private 1998 contractors designed portion

Contract period (months): Stipulated: 9; Offered: 9; Agreed: 9

Cost fluctuations: Firm

Contract breakdown

Measured work:	£1,976,676	rebased
Provisional sums:	£281,284	rebased
Prime cost sums:	£868,241	rebased
Preliminaries:	£454,879	rebased
Contingencies:	£104,911	rebased
Contract sum:	£3,685,991	rebased

Tender list (lowest first)

£3,685,991	(-)
£3,690,019	(0.1%)
£3,860,290	(4.7%)
£3,895,681	(5.7%)
£4,224,069	(14.6%)
£4,463,975	(21.1%)

Accommodation and design features

V shaped 2 storey customer service centre with open plan offices. Mass concrete fill, RC pad foundations and ground slab; PCC upper floor and stairs. Steel frame, felt covered flat and slate covered pitched roof. Rendered block walls; aluminium curtain walling and windows; Brise Soleil. Block partitions. Flush doors. Plaster to walls; carpet, tiles and access flooring; mineral fibre suspended ceilings. Fittings. Sanitaryware. Gas HW central heating, comfort cooling, ventilation, electric light and power. Lift. Lightning protection, fire/intruder alarms, CCTV, BMS.

Areas

Areas

© RICS 2014

Basement:	$0m^2$
Ground floor:	$1,448m^2$
Upper floors:	$1,632m^2$
Gross floor area:	$3,080\ m^2$

Usable area:	$2,073m^2$
Circulation area:	$601m^2$
Ancillary area:	$341m^2$
Internal divisions:	$65m^2$
Gross floor area:	$3,080\ m^2$

External envelope / floor heights

Area of external walls:	$2,930m^2$
Wall to floor ratio:	95.13%
Average storey heights (ground):	4.20m
Average storey heights (upper):	3.60m

Floor area percentages

2 storey (100.00%)

Credits

Submitted by: Hills

Client: Welsh Development Agency

Architect: Wigley Fox

Quantity Surveyor: Hills

Structural Engineer: Bingham Hall O'Hanlan

Services Engineer: White Young Green

Planning Supervisor: SPR Hooper

General Contractor: Stradform Ltd

© RICS 2014

Elements rebased

Element	Total cost	Cost per m²	Element unit qty	Element unit rate	Percent age
1 Substructure	£117,313	£38	1448 m2	£81	3%
2A Frame	£152,228	£48	3080 m2	£49	4%
2B Upper Floors	£97,980	£31	1632 m2	£60	3%
2C Roof	£248,125	£80	1948 m2	£127	7%
2D Stairs	£22,565	£7	4 No	£5,641	1%
2E External Walls	£329,040	£106	2290 m2	£144	9%
2F External Windows and Doors	£320,719	£103	640 m2	£501	9%
2G Internal Walls and Partitions	£74,065	£23	2228 m2	£33	2%
2H Internal Doors	£81,305	£26	80 No	£1,016	2%
2 Superstructure	£1,326,026	£429			36%
3A Wall Finishes	£59,983	£18	3057 m2	£20	2%
3B Floor Finishes	£201,080	£65	2775 m2	£72	5%
3C Ceiling Finishes	£57,104	£18	2715 m2	£21	2%
3 Finishes	£318,167	£103			9%
4 Fittings and Furnishings	£159,739	£51			4%
5A Sanitary Appliances (Costs include other elements)	£59,405	£18	54 No	£1,100	2%
5B Services Equipment	£0	£0			
5C Disposal Installations (Costs included in 5A)					
5D Water Installations	£19,753	£5			1%
5E Heat Source	£0	£0			
5F Space Heating and Air Conditioning	£219,584	£70			6%
5G Ventilating Systems	£131,867	£42			4%
5H Electrical Installations	£201,337	£65			5%
5I Fuel Installations	£439	£0			
5J Lift and Conveyor Installations	£28,226	£9			1%
5K Fire and Lightning Protection	£33,708	£10			1%
5L Communications and Security Installations	£8,872	£3			
5M Special Installations	£13,053	£4			
5N Builder's Work in Connection	£14,127	£4			
5O Management of the Commissioning of Services	£0	£0			
5 Services	£730,370	£236			20%
Building Sub-total	£2,651,616	£860			72%
6A Site Works	£277,380	£89			8%
6B Drainage	£67,644	£21			2%
6C External Services	£129,562	£42			4%
6D Minor Building Works	£0	£0			
6E Demolition and Work Outside the Site	£0	£0			
6 External Works	£474,586	£154			13%
7 Preliminaries	£454,879	£147			12%
8 Contingencies	£104,911	£34			3%
Total (less Design Fees)	£3,685,991	£1,196			100%
9 Design Fees	£0	£0			
Total Contract sum	£3,685,991	£1,196			100%

Specification

© RICS 2014

Element	Specification
1 Substructure	Mass concrete fill. RC GEN3 pad foundations to BS 5328 and grade 40 bed.
2A Frame	Steel frame.
2B Upper Floors	Contractor designed PCC upper floor, 3-9m spans.
2C Roof	Steel pitched roof with profiled sheet deck, Tactray 90, and 600x300mm fibre cement slates on battens and counterbattens; Celotex insulation. Steel flat roof with single layer polymer warm deck covering.
2D Stairs	Contractor designed PCC stairs.
2E External Walls	100mm blockwork with 15mm proprietary render; cast stone features to openings.
2F External Windows and Doors	Contractor designed double glazed aluminium curtain walling and windows. Brise Soleil to south and east elevations. Stainless steel frames and surrounds to entrance door.
2G Internal Walls and Partitions	Non-loadbearing 4N/mm2 blockwork.
2H Internal Doors	Flush cherry veneered solid core doors in hardwood frames and linings.
2 Superstructure	
3A Wall Finishes	13mm lightweight plaster.
3B Floor Finishes	600x600mm, 269mm cavity access floor. 300x300x8mm ceramic tiles on screed; carpet tiles.
3C Ceiling Finishes	Armstrong Orcal Tegular Microlook Prelude 24 grid, 600x600x16mm micro-perforated metal tiles.
3 Finishes	
4 Fittings and Furnishings	Shelving.
5A Sanitary Appliances	Sanitaryware.
5B Services Equipment	
5C Disposal Installations	Soil and waste pipes.
5D Water Installations	Hot and cold water services.
5E Heat Source	
5F Space Heating and Air Conditioning	Gas HW central heating. Comfort cooling and heating.
5G Ventilating Systems	Fresh air supply and extract.
5H Electrical Installations	Electric light and power.
5I Fuel Installations	Gas supply.
5J Lift and Conveyor Installations	Lift.
5K Fire and Lightning Protection	Lightning protection.
5L Communications and Security Installations	CCTV, security and fire alarms.
5M Special Installations	BMS.
5N Builder's Work in Connection	General builder's work in connection with services.
5O Management of the Commissioning of Services	
5 Services	
Building Sub-total	
6A Site Works	200x100x53mm clay block paving on sand; 219x109x80mm concrete sett paving on sand. Planting shrubs and trees.
6B Drainage	150 and 225mm clay pipes, concrete beds and surrounds. Polypropylene inspection chambers; PCC manholes.

© RICS 2014

Element	Specification
6C External Services	External services and site lighting.
6D Minor Building Works	
6E Demolition and Work Outside the Site	
6 External Works	
7 Preliminaries	14.55% of remainder of Contract Sum (excluding Contingencies).
8 Contingencies	3.36% of remainder of Contract Sum (excluding Preliminaries).
Total (less Design Fees)	
9 Design Fees	
Total Contract sum	

© RICS 2014

Appendix 2: Detailed Calculations of the Pilot Study

Item Nr	Description	Unit	Quantity	Unit Emb. Carbon ($kgCO_2$)	Total Emb. Carbon ($kgCO_2$)
1	**Substructure**				
	No details at all				
2	**Superstructure**				
2A	**Frame**				
	Columns Beams Fixings — No details about the sections and the span				
2B	**Upper Floors**	**m^2**	**1632**	**86.387**	**140,983.58**
	Contractor designed PCC floors				
E6010003A/B	Standard JJ1/RJ1	m^2	1632	74.046	120,843.07
	OR				
E6010003C/D	Jetplus JP1/RP1	m^2	1632	98.728	161,124.10
2C	**Roof**				
	Pitched and flat roof quantity combined together				
2D	**Stairs**				
Rate build-up	Contractor designed PCC stairs	**No**	**4**	**1,680.95**	**6,723.80**
E101325A	Reinforced concrete C30	m^3	2	350.14	654.09
E200105A	Formwork general finish to steps	m	25	1.77	44.25
E200112E	Formwork general finish to soffit of stairs	m^2	9	5.73	51.81
E301105G	Bar reinforcements, 25 mm bars fixed with tying wire	t	0.43	1,722.16	739.95
M505502A	Carpet tiles, 500 x 500 mm Heuga interloop	m^2	13	10.448	133.03

Item Nr	Description	Unit	Quantity	Unit Emb. Carbon ($kgCO_2$)	Total Emb. Carbon ($kgCO_2$)
M202101A & M601001G	Carlite plaster 10 mm thick two coats to concrete background and one mist coat and two full coats of emulsion paint to soffits	m^2	9	2.15	19.91
P207153B	Balustardes, 25 x 25 mm housed at 100 mm centres; 50 x 75 mm moulded handrail, 25 x 75 mm string capping	m	6	6.341	37.92
2E	**External Walls**	m^2	**2290**	**26.011**	**59,565.19**
F100107D	100mm Block work, 7N/mm^2, in cement motar (1:3)	m^2	2290	20.573	47,112.17
M201501A	Cement, lime, sand (1:1:6) screed finish, 15mm thick, over 300m wide	m^2	2290	4.370	10,007.30
M601001B	Emulsion paint, one mist coat and two full coats, plastered backgroung, 3.5- 5.0 m	m^2	2290	1.068	2,445.72
2F	**Windows and External Doors**	**m^2**	**840**	**26.971**	**22,655.60**
L105150	Steel framed glased screens, 6mm clear toughned glass; fixing with screws	m^2	640	32.249	20,639.36
L105160C	Brise-soleil, 7.5 x 2.0 m	Nr	12	112.400	1,348.80
L105160C	Brise-soleil, 10.0 x 2.0 m	Nr	1	667.440	667.44
2G	**Internal Walls & Partitions**	**m^2**	**2228**	**16.541**	**36,853.35**
F100106A	100mm Block work, 3.5N/mm^2, in cement motar (1:3)	m^2	2228	16.541	36,853.35
2H	**Internal Doors**	**Nr**	**80**	**15.887**	**1,270.96**
L202313	Wood veneered interior flush doors, clear laquer finish, 40mm thick	Nr	80	15.887	1,270.96

Item Nr	Description	Unit	Quantity	Unit Emb. Carbon ($kgCO_2$)	Total Emb. Carbon ($kgCO_2$)
3	**Finishes**				
3A	**Wall Finishes**	**m^2**	**3057**	**6.018**	**18,396.26**
M201202A	Cement, sand (1:3) screed finish, 13mm thick, over 300m wide	m^2	3057	4.950	15,131.39
M601001B	Emulsion paint, one mist coat and two full coats, plastered backgroung, 3.5- 5.0 m	m^2	3057	1.068	3,264.88
3B	**Floor Finishes**	m^2	**2248.5**	**60.684**	**54,750.82**
Iaufacturers' rat	600 x 600 cavity access floor	m^2	2073	25.027	51,880.97
)1206A & M4054	300 x 300x 8 mm ceramic tiles on screed	m^2	70	25.209	1,769.67
M505502A	Carpet tiles, 500 x 500 mm Heuga interloop	m^2	105	10.448	1,100.17
3C	**Ceiling Finishes**				
K401312C	Armstrong Orcal Tegular Microlook Prelude 24 grid, 600x600x16mm micro-perforated metal tiles.	m^2	**2715**	**25.027**	**67,948.31**
4	**Fittings & Furnishings**				
	Lack of details				
5	**Services**				
	Lack of details				

Appendix 3: Table for Durbin-Watson Test

Table for Durbin Watson Test

Sample Size	Probability in Lower Tail (Significance Level= α)	k = Number of Regressors (Excluding the Intercept) 1		2		3		4		5	
		d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	.01	.81	1.07	.70	1.25	.59	1.46	.49	1.70	.39	1.96
	.025	.95	1.23	.83	1.40	.71	1.61	.59	1.84	.48	2.09
	.05	1.08	1.36	.95	1.54	.82	1.75	.69	1.97	.56	2.21
20	.01	.95	1.15	.86	1.27	.77	1.41	.63	1.57	.60	1.74
	.025	1.08	1.28	.99	1.41	.89	1.55	.79	1.70	.70	1.87
	.05	1.20	1.41	1.10	1.54	1.00	1.68	.90	1.83	.79	1.99
25	.01	1.05	1.21	.98	1.30	.90	1.41	.83	1.52	.75	1.65
	.025	1.13	1.34	1.10	1.43	1.02	1.54	.94	1.65	.86	1.77
	.05	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	.95	1.89
30	.01	1.13	1.26	1.07	1.34	1.01	1.42	.94	1.51	.88	1.61
	.025	1.25	1.38	1.18	1.46	1.12	1.54	1.05	1.63	.98	1.73
	.05	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
40	.01	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
	.025	1.35	1.45	1.30	1.51	1.25	1.57	1.20	1.63	1.15	1.69
	.05	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
50	.01	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
	.025	1.42	1.50	1.38	1.54	1.34	1.59	1.30	1.64	1.26	1.69
	.05	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
60	.01	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
	.025	1.47	1.54	1.44	1.57	1.40	1.61	1.37	1.65	1.33	1.69
	.05	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
80	.01	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
	.025	1.54	1.59	1.52	1.62	1.49	1.65	1.47	1.67	1.44	1.70
	.05	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
100	.01	1.52	1.56	1.50	1.58	1.48	1.60	1.45	1.63	1.44	1.65
	.025	1.59	1.63	1.57	1.65	1.55	1.67	1.53	1.70	1.51	1.72
	.05	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78