Personalised Environmental Monitoring of Building Occupants: Integration of Scalable Technologies

Graham David Coulby

PhD

2022

Personalised Environmental Monitoring of Building Occupants: Integration of Scalable Technologies

Graham David Coulby

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

December 2022

Abstract

Urbanised societies spend most of their time indoor. These are places to conduct habitual activities that impact across the life course and are generating discussions on the built environment and its interplay with health and wellbeing. To understand the effect buildings and their enclosed spaces have on people/occupants, there is a need to monitor Indoor Environmental Quality (IEQ) and occupant responses. State-of-the-art monitoring approaches exist, but they have limited utility outside of bespoke scenarios due to their limited pragmatism and large cost. Other emergent technologies exist but questions remain relating to e.g., validity. Other routine/traditional subjective approaches for evaluating building IEQ often negate to account for the experiences of individual occupants, adding to complications.

This thesis explores current monitoring IEQ trends, uncovering the needs to make the individual the unit of analysis. Research undertaken explores contemporary needs and shifting trends to pragmatic approaches, localised sensors to provide richer data that could enable a better understanding of environmental and occupant changes. Quantitative measurement of the environmental conditions local to individuals are explored to understand whether spatial density in monitoring can 1) reinforce data pertaining to how building occupants experience indoor conditions and 2) provide additional context to current approaches for data capture, which traditionally focus on qualitative approaches.

Through a series of original research this thesis broadly presents the design and development of a multi-modal IEQ monitoring device and a supporting methodological process for monitoring individuals. It identifies that low-cost multi-modal monitoring deployed longitudinally can add significant context to traditional qualitative approaches, with the individual as the unit of analysis. Findings from the thesis present a paradigm shift that could have practical implications for researchers and practitioners, changing the way building performance is assessed and the way its impact on health and wellbeing could be evaluated.

Table of Contents

Abstract 3

Table of Contents 4

List of Tables 13

List of Figures 14

List of Acronyms 17

List of my Publications 20

Acknowledgements 22

Declaration 23

Chapter 1 Introduction 24

1.1 Introduction 25

1.2 Approach and investigatory process 26

1.3 Background: The industrial landscape 26

1.4 Thesis statement 29

1.5 Outline of thesis 29

1.5.1 Chapter 2: Literature review **(PoI1-3)** 29

1.5.2 Chapters 3-4: Exploration into IoT hardware and software **(PoI4)** 29

1.5.3 Chapter 5-7: Development of a holistic multimodal IEQ monitoring device **(PoI5)** 30

1.5.4 Chapters 8 and 9: The design and implementation of a study protocol for longitudinal assessment of individuals **(PoI6)** 30

1.5.5 Chapters 10 and 11: Analysis of data collected from remote, longitudinal assessment of an individual building occupant and the indoor environmental quality they experience. (**PoI7**) 30

1.5.6 Chapters 12: Summary of thesis 31

1.6 Previously Published Work 31

1.7 Contribution to knowledge 31

1.8 Next steps 31

Chapter 2 A Scoping Review of Technological Approaches to Environmental Monitoring. 32

2.1 Introduction 33

2.2 Methods 33

2.2.1 Searching and Selection Strategy 34

2.2.2 Eligibility Criteria and Information Sources 35

2.2.3 Charting Screening and Synthesising Data 35

2.3 Understanding IEQ and how it is measured 42

2.3.1 Indoor air quality 43

2.3.1.1 Carbon dioxide (CO_2) 44

2.3.1.2 Airborne contaminants 45
2.3.2 Thermal comfort 46
2.3.2.1 Predictive Mean Vote (PMV) 46
2.3.2.2 Adaptive comfort model 47
2.3.2.3 Occupant control 47
2.3.3 Visual comfort 48
2.3.4 Acoustic comfort 48
2.4 Understanding state-of-the-art environmental monitoring 49
2.4.1 Data loggers 49
2.4.2 Scalability limits around state-of-the-art solutions 50
2.5 Low-cost alternative technologies 51
2.5.1 Limitations of low-cost sensors 51
2.5.1.1 Accuracy vs. precision 52
2.5.2 Scalability 52
2.5.3 Holistic cloud-based systems 53
2.6 Individualised IEQ approaches for health and wellbeing 54
2.6.1 Holistic IEQ approaches 55
2.6.2 Linking health to wellbeing: Augmenting IEQ approaches 56
2.7 Discussions and conclusions 56
2.7.1 Understanding IEQ 57
2.7.2 Understanding IEQ measurement technology 57
2.7.3 Augmenting current IEQ approaches 58
2.7.4 Limitations 59
2.8 Knowledge Gap 59
2.9 Addressing the PoI 60
2.10 Further research 62
Chapter 3 A narrative review of low-cost sensor technologies for environmental monitoring. 63
3.1 Introduction 64
3.2 Low-cost Sensor Technology 64
3.2.1 Initial prototyping tools: MEMS sensors and bench testing 65
3.2.1.1 Ensuring Fit-for-purpose Monitoring 67
3.3 Communication and control 67
3.3.1 FPGA/ASIC 67
3.3.2 CPU 68
3.3.3 MCU 69
3.3.4 Software 72
3.4 Cloud Connectivity 72

3.4.1 Rate limiting and transactional cost ... 73
3.4.1.1 Microsoft Azure IoT Hub ... 73
3.4.1.2 Amazon Web Services (AWS) ... 74
3.4.1.3 Google Cloud Platform (GCP) ... 74
3.4.2 Communication protocols ... 75
3.5 Serial Processing ... 76
3.6 Discussions ... 76
3.6.1 IoT hardware ... 77
3.6.2 Cloud computing ... 77
3.7 Addressing the PoI ... 78
3.8 Further Research ... 78
Chapter 4 Frameworks for wearable integration ... 79
4.1 Introduction ... 80
4.2 Proprietary WHT Systems ... 80
4.3 Application Programming Interface (API) ... 81
4.4 Software Development Kit (SDK) ... 81
4.5 Interacting with Proprietary Systems ... 82
4.5.1 Security and Ethics ... 83
4.5.2 Cloud Services ... 83
4.5.2.1 Software as a Service (SaaS) ... 84
4.5.2.2 Platform as a Service (PaaS) ... 85
4.5.2.3 Infrastructure as a Service (IaaS) ... 85
4.6 Discussion and Conclusions ... 86
4.7 Addressing the PoI ... 86
4.8 Further Research ... 87
Chapter 5 Case Study One - A pilot study for individualised assessment of IEQ ... 88
5.1 Introduction ... 89
5.2 Study setup ... 90
5.2.1 Technologies and outcomes ... 90
5.2.2 Study setting ... 91
5.3 Taming the wild: Lessons learned ... 91
5.3.1 Integrator Apps: Limited data access ... 91
5.3.2 Proprietary APIs: Restrictions and regulations ... 93
5.3.3 Data Security: Authentication and protection ... 94
5.3.4 Communication and connectivity ... 94
5.3.5 Remote deployment: Access restrictions ... 94
5.4 Discussions and Conclusions ... 95

5.5 Addressing the PoI ... 96
5.6 Further Research ... 96
Chapter 6 Case Study Two – Considerations for data sampling using IoT technologies ... 97
6.1 Introduction ... 98
6.2 Physiological measurement of gait: Current state-of-the-art ... 99
6.3 Exploring IoT approaches to remote assessment ... 100
6.4 Experimental setup and equipment ... 100
6.5 Findings ... 102
6.5.1 Gait: High frequency data ... 102
6.5.2 High frequency analysis via IoT: Gait as a Example ... 103
6.5.3 Inter-device data aggregation ... 103
6.6 Discussions and Conclusions ... 104
6.7 Addressing the PoI ... 105
6.8 Further Research ... 105
Chapter 7 A scalable and multimodal approach to monitor IEQ ... 106
7.1 Introduction ... 107
7.2 Related Work ... 108
7.2.1 Inclusion criteria: Sensor integration ... 108
7.2.2 Low-cost IEQ sensing ... 108
7.2.2.1 Air quality: equivalent Carbon Dioxide (eCO_2) and Volatile Organic Compounds (VOCs) ... 108
7.2.2.2 Air quality: Carbon Dioxide (CO_2) ... 109
7.2.2.3 Air quality: Particulate matter (e.g., $PM_{2.5}$) ... 110
7.2.2.4 Thermal comfort: Temperature and humidity ... 111
7.2.2.5 Light: Ambient light intensity ... 111
7.2.2.6 Sound/noise: Noise levels ... 112
7.2.3 Cloud connectivity ... 112
7.3 Reference devices ... 113
7.3.1 Onset HOBO® MX1102 (CO_2 and eCO_2) ... 113
7.3.2 IQAir Air Visual Pro ($PM_{2.5}$) ... 113
7.3.3 Onset HOBO® MX1104 (Light intensity, temperature, humidity) ... 114
7.3.4 Air Pressure ... 114
7.3.5 Omega HHSL-101 (noise levels) ... 114
7.4 Multimodal device architecture ... 115
7.4.1 Hardware development ... 115
7.4.2 Reading CO_2 data via UART ... 116
7.4.3 Reading PM, temperature, humidity and ambient light intensity data via I^2C ... 117
7.4.4 Reading noise data via I^2S ... 118

7.5 Methods ... 119
7.5.1 Data acquisition and connectivity ... 119
7.5.2 Data processing ... 119
7.5.2.1 Resampling (excluding noise level data) ... 120
7.5.2.2 Missing data ... 120
7.5.2.3 Noise level data ... 121
7.5.2.4 Analytical and statistical procedures ... 121
7.5.2.5 Data visualisation ... 121
7.5.3 Sensor deployment ... 122
7.5.4 Reference standard setup ... 123
7.6 Results ... 123
7.6.1 Equivalent carbon dioxide (eCO_2) ... 125
7.6.2 Carbon dioxide (CO_2) ... 125
7.6.3 Particulate matter ... 126
7.6.4 Temperature, humidity (and air pressure) ... 127
7.6.5 Light ... 128
7.6.6 Noise ... 128
7.7 Discussions and Conclusions ... 129
7.7.1 MEMS sensor selection ... 129
7.7.2 Sensor performance ... 129
7.7.3 Accuracy and precision ... 130
7.7.4 Ventilation sensors for air quality ... 130
7.7.5 Cloud connectivity ... 130
7.7.6 Limitations ... 131
7.7.7 Suitability ... 131
7.8 Contribution to knowledge ... 132
7.9 Addressing the PoI ... 133
7.10 Future Work ... 133
7.11 Data ... 134
Chapter 8 A protocol for longitudinal monitoring of individual building occupants and their environments ... 135
8.1 Introduction ... 136
8.1.1 A protocol for personalised multimodal IEQ ... 137
8.2 Methods ... 138
8.2.1 Study design ... 138
8.2.2 Study setting ... 139
8.2.3 Eligibility Criteria ... 139
8.2.4 Sample Size ... 139

8.2.5 Participant Timeline ... 139
8.2.6 Initial Meeting ... 139
8.2.7 Post study interview and survey ... 140
8.3 Outcomes ... 141
8.3.1 Primary Outcomes ... 141
8.3.2 Predictors of the primary outcome ... 141
8.4 Data collection and management ... 141
8.4.1 Data Collection Methods ... 142
8.4.1.1 Measuring physiological and behavioural responses ... 142
8.4.1.2 Measuring IEQ changes ... 142
8.4.1.3 Measuring IEQ perceptions ... 143
8.4.2 Data management ... 144
8.4.2.1 AppleWatch ... 144
8.5 Statistical methods for analysis ... 145
8.6 Ethics & Dissemination ... 146
8.6.1 Research Ethics Approval ... 146
8.6.2 Informed Consent ... 146
8.6.3 Confidentiality ... 147
8.7 Discussions and Conclusions ... 147
8.7.1 Contribution ... 147
8.7.2 Limitations of Study ... 148
8.8 Address the PoI ... 149
8.9 Further Research ... 150
Chapter 9 Personalising a protocol for individualised monitoring of building occupants ... 151
9.1 Introduction ... 152
9.2 Methods ... 152
9.2.1 Study Setting ... 152
9.2.1.1 Office ... 153
9.2.1.2 Home ... 154
9.2.2 Overview ... 154
9.2.2.1 Initial Meeting: Protocol personalisation ... 154
9.2.2.2 Office Hours ... 155
9.2.3 Wearable: Overnight charging ... 155
9.2.3.1 Survey capture mechanism: Amazon Echo ... 155
9.3 Data collection and management ... 155
9.3.1 Data collection methods ... 155
9.3.1.1 Measuring IEQ perceptions ... 155

9.3.1.2 Measuring IEQ changes ... 157
9.4 Discussions and Conclusions ... 158
9.5 Address the PoI ... 158
9.6 Further Research ... 159
Chapter 10 Macro-level personalised IEQ: Exploring use of quantitative data to contextualise perceptions ... 160
10.1 Introduction ... 161
10.2 Methods ... 161
10.2.1 Synchronicity ... 162
10.2.2 Wearable health data processing ... 163
10.2.2.1 Interpolation of heart rate data ... 164
10.2.3 Statistical methods ... 165
10.3 Findings and discussion ... 165
10.3.1 Temperature ... 166
10.3.2 Humidity ... 167
10.3.3 Light ... 168
10.3.4 Noise ... 170
10.3.5 Air Quality ... 172
10.3.5.1 Carbon Dioxide (CO_2) ... 172
10.3.5.2 Equivalent Carbon Dioxide (eCO_2) ... 174
10.3.5.3 Particulate Matter ($PM_{2.5}$) ... 175
10.3.6 Study Process ... 177
10.4 Answering the PoI ... 178
10.5 Further Research ... 179
Chapter 11 Micro-level personalised IEQ: A feasibility study on the dynamic regression analysis of intra-day data ... 180
11.1 Introduction ... 181
11.2 Methods ... 181
11.3 Application of a micro-level approach ... 182
11.3.1 Step 1: Format Dataset ... 182
11.3.2 Step 2: Imputation ... 182
11.3.2.1 Missing Value Analysis ... 183
11.3.2.2 Filtering and removal of unneeded data ... 183
11.3.2.3 Multiple Imputation ... 184
11.3.3 Step 3: Plot data ... 185
11.3.4 Step 4: Assess stationarity ... 186
11.3.4.1 Assessment of time trends and periodicity ... 188
11.3.5 Step 5: Forecasting ... 188

11.3.6 Step 6: Lagged variables 189
11.3.7 Step 7: Confirm autocorrelation correction 190
11.3.8 Repetition of steps 5 -7 for independent variables 191
11.3.9 Step 8: Dynamic regression 192
11.4 Results 193
11.4.1 Step 9: Interpret the results 193
11.5 Discussion 193
11.5.1 Step 10: Reporting 193
11.5.2 Limitations 194
11.6 Addressing the PoI 194
11.7 Conclusion 195
Chapter 12 Conclusion 196
12.1 Introduction 197
12.2 Addressing the points of inquiry 198
12.2.1.1 PoI1 198
12.2.1.2 PoI2 198
12.2.1.3 PoI3 199
12.2.1.4 PoI4 199
12.2.1.5 PoI5 200
12.2.1.6 PoI6 200
12.2.1.7 PoI7 201
12.3 Addressing the Thesis Statement 201
12.3.1 Research question 201
12.3.2 Answer to research question 201
12.3.3 Hypothesis 202
12.3.4 Addressing the Hypothesis 202
12.4 Discussions and Conclusions 202
12.4.1 Contributions to knowledge 203
12.4.1.1 Significant original contribution to knowledge 204
12.4.2 Impact of SARS-COV-2 204
12.4.3 Practical implications 205
12.4.4 Future Research 206
12.5 Closing Summary 206
References 208
Appendix A Theoretical Perspective and Methodology 235
Appendix B Creative Commons 4.0 License 238
Appendix C PRISMA-ScR Checklist 242

Appendix D Statement of Authorisation for Digital Health Chapter 245
Appendix E Informed Consent Form Example 247
Appendix F Pandas source code for synchronising data 253

List of Tables

Table 1 - Points of Inquiry to drive research 28
Table 2 - List of Search Terms (Filters) 34
Table 3 - Overview of Measurements in Selected Studies (1/3) 36
Table 4 - State-of-the-art sensors (1/2) 39
Table 5 - State-of-the-art sensors (2/2) 40
Table 6 - Low-cost sensors 41
Table 7 - Examples of MEMS sensor use for healthcare 66
Table 8 - Arduino's product range, highlighting architectures and ADC/DAC capabilities .. 70
Table 9 - Example of IoT hub pricing tiers 74
Table 10: Outcomes measured with remote sensors 91
Table 11 – Low-Cost sensors used for development including min/max measurement thresholds and units that are measured 112
Table 12 – Reference devices used, with indicative costs, min/max measurement thresholds and units 114
Table 13 - MH-Z19B UART Commands 117
Table 14 - Results of sensor validation study 124
Table 15 - Format of closeout study question responses 140
Table 16 – Dependant variables of the primary outcomes 141
Table 17 – Covariates to predict the primary outcomes 141
Table 18 – Automated survey questions and responses 144
Table 19 – Sensors in the multimodal environmental monitoring device and the specific outcomes they will monitor 157
Table 20 - Descriptive statistics for data 166
Table 21 – Summary of missing values across multimodal variables 183
Table 22 - Summary of missing values across multimodal variables after filtering 184
Table 23 - Descriptive statistic of multimodal device variables before and after multiple imputation 185
Table 24 – Descriptive statistics for 15-minute data 185
Table 25 – Descriptive statistics of HR for monthly partitions 187
Table 26 – Descriptive statistics of SC for monthly partitions 187
Table 27 – Lagged variables for HR and SC data 190
Table 28 – Autocorrelation adjustments for independent variables 192
Table 29 – dynamic regression coefficients for HR model 192
Table 30 – dynamic regression coefficients for SC model 193

List of Figures

Figure 1 - Example of signposting for previously published work 31
Figure 2 – A holistic and general capture of terminologies and themes used across the literature to discuss IEQ 43
Figure 3 - IEQ in Open Plan Offices 44
Figure 4 - Scale of MEMS sensor breakout board, compared to a 555 Timer chip with 2.54mm pitch. 67
Figure 5 - Diagram of proprietary system interactions 84
Figure 6 – Screenshot of triggers available for Fitbit taken from the IFTTT platform [182] 92
Figure 7 - System Architecture for extracting data from proprietary APIs using Azure Function Apps. 93
Figure 8 - MX1101 light intensity data logger and BH1750 ambient light sensor connected to ESP32 development board. 102
Figure 9 - Free-living tri-axial accelerometer data (AX3). The vertical green and red indicate possible start/stop gait bouts. 103
Figure 10 - Data captured from HOBO MX1101 and BH1750 104
Figure 11 - Low-cost sensors collecting data alongside reference devices. Including a closeup image to show the breadboard configuration. 115
Figure 12 - Schematic diagram for multi-modal IEQ sensor device 116
Figure 13 - Data processing methods 120
Figure 14 - Layout of office showing placement of windows, doors, artificial light sources and radiators. 122
Figure 15 – Snapshot of eCO_2 vs CO_2 events captured by CCS811 (blue) and MX1102 (orange) 125
Figure 16 – Snapshot of MH-Z19B CO_2 sensor (blue) vs reference MX1102 CO_2 sensor (orange) 126
Figure 17 – Bland-Altman plot of PMSA003i means against the Air Visual Pro 127
Figure 18 – Comparison of low-cost BME280 temperature sensor (blue) against MX1104 reference (orange). 128
Figure 19 - Diagram of study setting and passive sensor configuration. 138
Figure 20 - Screenshot of iOS Health Data Parser application 145
Figure 21 – Flowchart representation of the methodological approach - highlighting the two-stage macro and micro-level analyses. 146
Figure 22 - Site plan for study location, showing home, office and surrounding features including the locations of windows, heaters, lights and sensors 153

Figure 23 – Outline of the voiceflow application showing the visual script used to create the Alexa Skill 156

Figure 24 – Aligning data to make a synchronous multivariate dataset. Top (before): IEQ and physiological data captured at various time intervals. Bottom (after): all data synchronised to 1min/60s intervals 162

Figure 25 – Multi-step process for merging data from multiple sources 163

Figure 26 – JSON data format for AppleWatch step data. 163

Figure 27 – JSON data format for AppleWatch heart rate data. 164

Figure 28 – comparison of linear vs non-linear interpolation methods 164

Figure 29 – snapshot of data interpolated using PCHIP interpolant 165

Figure 30 - Average hourly temperatures (per weekday) from each temperature source (Home, Office, Outdoor). Note: Outdoor temperature is presented on a secondary (right) Y axis to highlight the similarity in the shape of the curves between outdoor temperature and office temperature. 167

Figure 31 - Average hourly humidity (per weekday) from each humidity source (Home, Office, External). Note: Outdoor humidity is presented on a secondary (right) Y axis to highlight the similarity in the shape of the curves. 168

Figure 32 - Average Home/Office Light Intensity categorised by Day > Hour. Note: variables are presented across two Y Axes to account for the mean difference across the measured variables. 169

Figure 33 – Randomly selected sun map for the study location to give an idea of where the sun (yellow circle) would be throughout the day in relation to the buildings and rooms. ... 169

Figure 34 - Average Home/Office noise levels between 10:00 – 14:00 categorised by Day > Hour > Minute 171

Figure 35 - Average Home/Office noise levels categorised by Day > Hour for Monday to Friday 171

Figure 36 - Average Home/Office CO2 categorised by Day > Hour 172

Figure 37 - Average Home/Office CO2 categorised by Day > Hour for an average Friday 173

Figure 38 - Average Home/Office CO_2 overlayed with sum of steps and key event markers – vertical green lines signify the point when steps rise or fall. 174

Figure 39 - Average Home/Office CO_2 overlayed with average heart rate 174

Figure 40 - Average Home/Office CO2 categorised by Day > Hour 174

Figure 41 - Average Home/Office CO2 categorised by Day > Hour 175

Figure 42 - Average outdoor air pollution vs indoor particulates categorised by Day > Hour 176

Figure 43 - Average outdoor air pollution vs office particulates categorised by Day > Hour 176

Figure 44 - OpenWeatherMap API endpoint for current weather at given Latitude & Longitude ... 177
Figure 45 - Timeseries plot for HR ... 186
Figure 46 - Timeseries scatter plot for SC ... 186
Figure 47 – HR timeseries correlations identified by Autocorrelation Function (ACF) ... 188
Figure 48 – SC timeseries correlations identified by Autocorrelation Function (ACF) ... 188
Figure 49 – HR lags identified by the Partial Autocorrelation Function (PACF) ... 189
Figure 50 – SC lags identified by the Partial Autocorrelation Function (PACF) ... 189
Figure 51 – Example of linear regression configuration for creating unstandardised residuals from lags of dependent variable ... 190
Figure 52 – ACF results for unstandardised residual of HR data, confirming the correction of autocorrelation. ... 191
Figure 53 – ACF results for unstandardised residual of SC data, confirming the correction of autocorrelation. ... 191
Figure 54 - The research onion diagram by Saunders et al. [280] which depicts the spectrum of epistemological standpoints and philosophical perspectives. ... 236

List of Acronyms

ABC	Automatic Baseline Calibration
ACF	Autocorrelation Function
ADC	Analogue-to-Digital Convertors
AI	Artificial Intelligence
API	Application Program Interface
ASHRAE	The American Society of Heating, Refrigerating and Air-Conditioning Engineers
ASIC	Application-Specific Integrated Circuit
AVP	IQAir Air Visual Pro
AWS	Amazon Web Services
BIM	Building Information Modelling
BLE	Bluetooth Low Energy
BMS	Building Management System
BOM	Bill of Materials
BPM	Beats per Minute
BREEAM	Building Research Establishment Environmental Assessment Method
CO_2	Carbon Dioxide
CPU	Central Processing Unit
DAC	Digital-to-Analogue Convertors
dB	Decibels
dB_{REF}	Decibel Equivalence of 1 Pascal (94dB)
DIY	Do it Yourself
ECG	Electrocardiogram
eCO_2	Equivalent Carbon Dioxide
EEG	Electroencephalogram
EPA	Environmental Protection Agency
ERDF	European Regional Development Fund
FCC	Federal Communications Commission
FPGA	Field Programmable Gate Arrays
GB	Gigabytes
GCP	Google Cloud Platform
GDPR	General Data Protection Regulation
GPIO	General-Purpose Input/Output
HR	Heart Rate
HRV	Heart Rate Variability
HTTP	Hyper Text Transfer Protocol
HVAC	Heating Ventilation and Air Conditioning
IaaS	Infrastructure as a Service
IAQ	Indoor Air Quality
ICC	Intraclass Correlation Coefficients
IEQ	Indoor Environmental Quality
IFTTT	If This Then That
I_{IAQ}	Indoor Air Quality Index
I_{IEQ}	Indoor Environmental Quality Index
IIIP	Intensive Industrial Innovation Programme

IoT	Internet of Things
ISO	International Standards Organization
I_{TH}	Thermal Comfort Index
I_V	Visual Comfort Index
KB	Kilobytes
LCD	Liquid Crystal Display
LDR	Light Dependant Resistor
LEED	Leadership in Energy and Environmental Design
LoA	Limit of Agreement
LUX/lx	Measurement of Light Intensity or Illuminance
MAR	Missing at Random
MB	Megabytes
MCAR	Missing Completely at Random
MCU	Microcontroller Unit
MEMS	Micro, Electro-Mechanical Systems
MPU	Microprocessor Unit
MQTT	Message Queuing Telemetry Transport
MRT	Mean Radiant Temperature
NDIR	Non-dispersal, infra-red
NMAR	Not Missing at Random
PaaS	Platform as a Service
PAQ	Perceived Air Quality
PACF	Partial Autocorrelation Function
PCB	Printed Circuit Board
PCHIP	Piecewise Cubic Hermite Interpolating Polynomial
PD	Parkinson's Disease
PFT	Personal Fitness Tracker
PHP	Hypertext Preprocessor
PM1.0	particulate matter up to 1.0µm in diameter
PM10	particulate matter up to 10µm in diameter
PM2.5	Particulate matter up to 2.5µm in diameter
PMV	Predicted Mean Vote
POE	Post Occupancy Evaluations
PPG	Photoplethysmogram
PRISMA	Preferred reporting items for systematic review and meta-analysis
PRISMA-ScR	Preferred reporting items for systematic review and meta-analysis - scoping reviews
PWM	Pulse Width Modulation
REST	Representational State Transfer
SaaS	Software as a Service
SBS	Sick Building Syndrome
SC	Step Count
SCG	Seismocardiogram
SCL	Serial Clock
SDA	Serial Data
SDK	Software Development Kit
SIM	Subscriber Identity Module

SMB	Server Message Block
SMS	Short Messaging Service
SPL	Sound Pressure Level
SRMS	Root Mean Squared of Samples
SSID	Service Set Identifier
SSL	Secure Sockets Layer
TESS	Thermal Environment Satisfaction Survey
TVOC	Total Volatile Organic Compounds
UART	Universal Asynchronous Receiver-Transmitter
VOC	Volatile Organic Compound
WELL	Well Building Standard
WHT	Wearable Health Technologies
WPA2-Enterprise	Enterprise Wi-Fi Protected Access 2
WSN	Wireless Sensor Network

List of my Publications

Title: A protocol for longitudinal monitoring of individual building occupants and their environments
Authors: Graham Coulby, Adrian K Clear, Oliver Jones, Suzanne McDonald, Alan Godfrey
Publication Type: Journal Article,
Publication Date: 23 September 2022,
Published in: PLOS ONE
DOI: https://doi.org/10.1371/journal.pone.0274015

Title: IoT in the Wild: An expedition of discovery for remote monitoring.
Authors: Graham Coulby, Adrian K Clear, Oliver Jones, Alan Godfrey
Publication Type: Conference Proceedings
Publication Date: 21 September 2021,
Published in: Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers
DOI: https://dl.acm.org/doi/abs/10.1145/3460418.3479364

Title: Frameworks: integration to digital networks and beyond.
Authors: Graham Coulby, Fraser Young
Publication Type: Book Chapter
Publication Date: 9 July 2021,
Published in: Digital Health - Exploring Use and Integration of Wearables
DOI: https://doi.org/10.1016/B978-0-12-818914-6.00003-X

Title: Low-cost, multimodal environmental monitoring based on the Internet of Things.
Authors: Graham Coulby, Adrian K Clear, Oliver Jones, Alan Godfrey
Publication Type: Journal Article
Publication Date: 31 May 2021,
Published in: Building and Environment
DOI: https://doi.org/10.1016/j.buildenv.2021.108014

Title: Lifting Lockdown: Insights and experimentation into IoT opportunities for remote healthcare monitoring.
Authors: Graham Coulby, Adrian K Clear, Oliver Jones, Alan Godfrey
Publication Type: Poster Presentation
Publication Date: 13 November 2020,
Published in: Northeast Postgraduate Conference 2020
DOI: https://doi.org/10.6084/m9.figshare.8299424.v1

Title: Towards remote healthcare monitoring using accessible IoT technology: state-of-the-art, insights and experimental design.
Authors: Graham Coulby, Adrian K Clear, Oliver Jones, Alan Godfrey
Publication Type: Journal Article
Publication Date: 30 October 2020,
Published in: Biomedical Engineering Online
DOI: https://doi.org/10.1186/s12938-020-00825-9

Title: A Scoping Review of Technological Approaches to Environmental Monitoring.
Authors: Graham Coulby, Adrian K Clear, Oliver Jones**,** Alan Godfrey
Publication Type: Journal Article
Publication Date: 4 June 2020,
Published in: International Journal of Environmental Research and Public Health
DOI: https://doi.org/10.3390/ijerph17113995

Title: The Building as a Lab: Towards the development of a toolbox.
Authors: Graham Coulby, Adrian K Clear, Oliver Jones**,** Alan Godfrey
Publication Type: Journal Article
Publication Date: 25 March 2020,
Published in: ARCOM Doctoral Workshop 2020
URL: https://itc.scix.net/pdfs/ADW-2020-paper-06.pdf

Acknowledgements

First and foremost, I would like to thank Ryder Architecture and the European Regional Development fund for providing this funding opportunity to undertake my studies at the Department of Computer and Information Sciences at the University of Northumbria Newcastle.

I would like to extend my gratitude to my team of academic and industry supervisors, Dr Alan Godfrey, Dr Adrian K. Clear and Dr Oliver Jones, for their ongoing support throughout my studies. I would also like to extend further gratitude to my principal supervisor Dr Alan Godfrey for his invaluable supervision, tutelage, and pastoral support throughout the Ph.D. process. Dr Godfrey’s guidance not only directed my Ph.D. work, but it also helped me grow as a person, provided me with exceptional continuing professional development, and it has guided me towards becoming an academic.

I would like to thank here the academic team, Dr Nick Dalton, Prof Wai Lok Woo, Dr Lesley McIntyre, and Prof David Greenwood, who have guided me through the project approvals and annual progressions and provided invaluable feedback throughout each of those stages of the Ph.D.

I am extremely grateful of the help and support offered to me over the last few years from my esteemed colleagues: Dr Graham Kelly, Mr Peter Barker, Mr Yongjun Choi, Ms Shelagh Keogh, Mr Clifford Brown, Ms Jill Bradnum, Professor Stephen Lockley, Ms Andrea Reynolds, Ms Lynne Retallick, Mr Dylan Powell, Mr, Fraser Young, Mr Yunus Çelik, Mr John Robson and Mr Conor Wall.

I would especially like to extend further gratitude to Dr Kay Rogage and Mr Lee Proud. Dr Rogage has been an inspiration to me and a mentor throughout my career both in industry and academia and is someone I highly regard as being a cornerstone to my success. Mr Proud has provided exceptional guidance, support and understanding over the last year, without which I likely would have been unable to complete my studies.

Last, but by no means least, I would also like to express my sincerest gratitude to my parents my wife and my daughter. This process has been tough to manage, but their patience, understanding, guidance and support have been truly appreciated and have been of paramount importance to the completion of my studies.

Thank you all.

Declaration

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

The work was done in collaboration with Ryder Architecture through the Intensive Industrial Innovation Programme (IIIP) funded by the European Regional Development Fund (ERDF), Ryder Architecture and Northumbria University.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought on three occasions and granted by Northumbria University's Research Ethics Committee on the following dates:

Application Reference	Date
13311	14 December 2018
17141	07 June 2019
20481	15 November 2019

I declare that the Word Count of this thesis is **59,799** words.

The declared word count is above that of standard *(35,000 - 45,000 words)* for computer and information science submissions, but it has been approved *(up to 85,000 words)* by the Graduate School at Northumbria University Newcastle, given that this thesis comprises of multidisciplinary work that covers a discursive subject area.

Name: Graham David Coulby

Date: 14 December 2022

Chapter 1 Introduction

1.1 Introduction

Global urbanisation is resulting in a behavioural paradigm whereby people are spending an increasing amount of their time within indoor environments [1]–[3]. Yet, prolonged occupation within inadequate indoor environments (e.g., poor ventilation or elevated noise levels) can have a negative impact on health and wellbeing. For example, respiratory problems, headaches, and skin conditions are some of the many symptoms that can be caused by poor Indoor Environment Quality (IEQ) - a term commonly used to describe the measurement of environmental parameters including air quality as well as visual, thermal, and acoustic comfort [4]. The implications poor IEQ have on health and wellbeing are well documented [3]–[11] but many current monitoring approaches have limited pragmatic (i.e., everyday) utility in buildings. Consequently, although extensive research into IEQ has been conducted over the years, the routine adoption rate for indoor environmental monitoring is low.

This Ph.D. is conducted as part of the Intensive Industrial Innovation Programme (IIIP) that is co-funded by the European Regional Development Fund (ERDF), Northumbria University and Ryder Architecture (www.ryderarchitecture.com, the small-to-medium sized enterprise sponsor). Previously, I worked in the construction sector with Ryder Architecture, where my colleagues and I witnessed a desire across the industry to adopt environmental monitoring from clients in early building design phases. However, as projects progressed into the technical specification phase(s) the desire for environmental monitoring often waned as the large cost and complexity of the equipment was realised. This often resulted in monitoring approaches being removed from the specifications completely, or being scaled down into small, sandboxed environmental setups *(as 'proof of concept' studies)*. This limits continuous assessment of buildings and the monitoring of individual occupants, presenting an unmet industrial need. This industrial need acted as a driver for the PhD as it was recognised that rigorous academic research could advance the science in this area to provide practical implications for the construction sector, to enable real world research in this area.

This chapter starts my thesis by presenting a brief background to the field. It introduces the reader to important IEQ topics, driven by industrial challenges which uncover several points of inquiry to be addressed and answered as part of the research conducted. Subsequently, a central hypothesis is identified to drive investigations to tackle current industrial challenges. This chapter concludes by summarising the complete work undertaken within this thesis, to signpost all content throughout.

1.2 Approach and investigatory process

Challenges from the industrial landscape will be outlined in section 1.3. Accordingly, that gives rise to several points of inquiry, which will be addressed and answered by undertaking academic research. To address and answer the inquiries posed here, I adopted a post-positivist stance (*details and justifications in Appendix A*) to enable me to utilise deductive approaches to collect and analyse objective measurement(s), while also considering subjectivity, interpretation, and human experience to support the findings. Consequently, this Ph.D. aims to thoroughly investigate the current state of the art to address those points of inquiry, supported by a series of case studies driven by real world considerations. For expediency and signposting, the point(s) of inquiry positioned in this thesis will be referenced throughout using the acronyms PoI1, PoI2, etc.

1.3 Background: The industrial landscape

There is growing research evidence [3]–[11] recognising the intricate interactions between IEQ, health and wellbeing which may be attributed to the scale and complexity of parameters being measured. IEQ is not one single measurement factor, nor is there one single building design element that impacts health and/or wellbeing, the latter actively promoted by many green building standards, (WELL building standard (WELL) [12], Leadership in Energy and Environmental Design (LEED) [13] and Building Research Establishment Environmental Assessment Method (BREEAM) [11]). Often, IEQ, health and wellbeing studies exacerbate the challenges by inadequately communicating causes and outcomes [14]. This makes it challenging to explicitly understand and measure what IEQ conditions lead to worsening health and wellbeing issues. Therefore, **PoI1-3** *(Table 1, end of this section – page 28)* were positioned to inform the development of the research to gain a solid understanding of what exactly is IEQ and how it is measured across the literature.

Generally, poor IEQ can impact health, cognitive abilities and productivity, by lowering concentration and alertness, and contributing to self-reported symptoms of sick building syndrome [15]–[17]. The impact these have within a workplace are drivers of IEQ research within commercial buildings. Yet, the indoor environment of all buildings can impact health and wellbeing and research now focuses on IEQ within residential buildings. However, there are many kinds of challenges to why implementing residential IEQ studies in real-world research is challenging such as ethical and cultural, which makes residential building studies more challenging than commercial [18]. Additionally, residential buildings often lack suitable infrastructure, such as IT systems, building-wide environmental controls and Building Management Systems (BMS) that could make IEQ monitoring routine. Furthermore, residential research lacks many drivers found in commercial building studies e.g., employer responsibilities or corporate wellness programmes, which provide direct returns on investment if they are used to maintain workplace productivity [19]. Commercial buildings also typically have complex measurement equipment, when compared with residential buildings, which can make residential measurements more difficult and costly [20]. To monitor multiple buildings simultaneously, more equipment is needed, which may be indicative as to why lower sample sizes are commonly reported as limiting factors in residential building studies [21]–[23]. This exposes a technical challenge surrounding the aggregation of data from different sources and, thus, leads to **PoI4** *(Table 1)*.

Studies in open-plan offices benefit from being able to observe large sample sizes, monitored simultaneously with a small number of devices/sensors. Commercial building studies also leverage state-of-the-art sensing equipment, which has implicit validity [24], acting as a 'gold/reference standard'. Typically, those are built into many Heating Ventilation and Air Conditioning (HVAC) systems or retrofitted to existing systems. However, in buildings without mechanical ventilation, there is a requirement to place in-situ sensors in each room. This creates more measurement points and increases the capital investment required for environmental monitoring. In the UK, a substantial majority of residential properties are heated through hot-water central heating systems that use boilers and radiators [25] reducing the need for mechanical ventilation. Consequently, indoor environment measurement must be conducted using in-situ measurement devices. Furthermore, it is unlikely that participants of residential studies will all live in the same enclosed space, implying individual monitoring would be required.

In research and laboratory studies it is feasible to create sandboxed research environments that use expensive measurement equipment or use small sample sizes and/or smaller spaces to reduce the amount of IEQ monitoring equipment required. Beyond research, there is an industry recognised need for monitoring to be scalable, enabling an analysis of buildings with many occupants. Unfortunately, promoting the use of state-of-the-art sensors on design projects is difficult due to the cost and complexity of the equipment, which limits scalability [7]. This exposes a need for low-cost sensor approaches for capturing environmental data outside of research. Low-cost options range from consumer-grade devices, designed for immediate integration into smart-homes, to 'Do It Yourself' (DIY) devices that make use of low-cost sensing components that integrate with microcontrollers/single-board computers. However, lack of scientific validation of those devices means many studies still require the use of expensive, reference-standard equipment as a baseline/reference measurement [7], [21], [26]–[28]. Accordingly, **PoI5** *(Table 1)* was positioned.

Current IEQ measurement is also limited due to subjective data gathering. It is common for thermal, visual, and acoustic comfort monitoring to involve measurements of an occupant's perception rather than just an empirical assessment of the environment. This leads to many discrepancies across the literature as individuals respond differently to changes in the environment [29]. Additionally, occupants are not always aware of what constitutes an inadequate environment, reporting issues that are directly observable only. This leads some to question the validity of subjective responses [30] but it is important not to discount those data. For example, for a system to be effective in improving wellbeing and comfort, it needs to monitor the environment and measure how occupants respond to environmental changes [31]. Objective measurement of individuals may be able to reinforce and possibly even substitute the subjective nature of occupant comfort measurements [32]. Therefore, addressing **PoI6 and PoI7** *(Table 1)* will help me to understand how to deploy localised monitoring devices and whether those data are able to expose additional context to what is seen in traditional measurement approaches.

Table 1 - Points of Inquiry to drive research

#	**Point of Interest**
PoI1	What constitutes IEQ and how is it currently measured?
PoI2	What is the current state-of-the-art in environmental monitoring?
PoI3	What sensor technologies are used to capture IEQ?
PoI4	What are the optimal approaches to aggregate data from numerous devices and settings, including settings without existing monitoring infrastructures?
PoI5	Can low-cost sensors be used to gather with precision, consistency, and reliability across different environments i.e., home, and commercial settings?
PoI6	How can multiple sensing modalities be pragmatically deployed to gather data for the longitudinal assessment of individual building occupants?
PoI7	If deployed longitudinally, can localised devices with multiple sensing modalities be used to address the subjectivity around environmental perceptions by providing quantitative context to how occupants experience building environments?

1.4 Thesis statement

This thesis will position itself around the PoI throughout, to keep the aims and objectives grounded to current industrial challenges. Accordingly, these inquiries form a body of work throughout the thesis that aims to address an industrial need based upon an academic focused hypothesis to have real world impact. Accordingly:

Industrial requirement: To monitor individual building occupants, beyond sandboxed research environments there is a need to increase the spatial density of measurement equipment, which requires exploration of lower cost monitoring solutions.

Hypothesis: Quantitative measurement of the environmental conditions local to individuals could add spatial density to 1) reinforce data pertaining to how building occupants experience indoor environmental conditions and 2) provide additional context to current approaches for data capture, which traditionally focus on qualitative approaches.

Research Question: Can localised sensors provide richer data that will enable better understanding of causal relationships i.e., how individual building occupants respond to environmental changes?

Impact: Practitioners could monitor buildings with a high spatial density to learn about intra-building variability, as well as the localised environmental conditions experienced by individual building occupants. The approach could also be used to better inform occupant health and wellbeing.

1.5 Outline of thesis

This section will provide an overview of the thesis, outlining the purpose of each chapter and highlighting which chapters address which questions.

*1.5.1 Chapter 2: Literature review **(PoI1-3)***

This thesis starts by conducting a scoping review of the literature on IEQ monitoring, exploring state-of-the-art technologies as well as emergent low-cost approaches across a multidisciplinary body of literature. The review chapter aims to answer **PoI1-3** to add scientific grounding to the hypothesis that state-of-the-art monitoring solutions are too costly or complex to have utility outside of research.

*1.5.2 Chapters 3-4: Exploration into IoT hardware and software **(PoI4)***

PoI4 is quite complex and multifaceted, addressing technical aspects and underlying technologies that enable IEQ and individualised monitoring. Accordingly, **PoI4** over two chapters (Chapter 3 and 4) through a series of technical review work. These chapters explore hardware and software of the Internet of Things (IoT) and look and explore how these technologies can be used to aggregated data from multiple data sources.

*1.5.3 Chapter 5-7: Development of a holistic multimodal IEQ monitoring device **(PoI5)***

Chapters 5-7 embark on a robust technical development of a multimodal monitoring solution that could monitor a wide range of environmental factors from a single device. Chapters 5 & 6 first conduct technical experimentation through case studies to explore the technological prerequisites for developing such a device, before outlining a robust approach for appraising and developing the device itself (Chapter 7). The aim of Chapter 7 was to develop a solution for IEQ monitoring that was low-cost, multimodal and was small enough to deploy on a desk for localised/individualised environmental monitoring. To address **PoI5,** studies were undertaken to validate whether low-cost sensor technologies could be used to capture multivariate data with precision, consistency and reliability.

*1.5.4 Chapters 8 and 9: The design and implementation of a study protocol for longitudinal assessment of individuals **(PoI6)***

With the technological aspects of data collection developed, tested, and validated, Chapter 8 presents the development of a study protocol for the assessment of an individual building occupant and their immediate environment, which utilises the developed multimodal monitoring devices from the previous chapter. Chapter 9 presents an extension to Chapter 8 to showcase how the protocol can be applied/adapted to a specific research project. It is hoped that by presenting the protocol in this way, it makes it more versatile to future researchers. Chapter 9 outlines details specific to the participant and details the study location and the environments measured. Accordingly, Chapters 8 and 9 present an approach for how many sensing modalities can be pragmatically deployed for a more seamless integration to individualised monitoring (**PoI6)**.

*1.5.5 Chapters 10 and 11: Analysis of data collected from remote, longitudinal assessment of an individual building occupant and the indoor environmental quality they experience. (**PoI**7)*

Chapter 10 presents an approach on how to format and sanitise big IEQ data, before conducting a macro-level visual analysis of the study outcomes. The visual macro-level analysis provided a useful mechanism for evaluating the large quantity of data captured throughout the study. Chapter 10 was able to addresses **PoI7** on its own but did not action all the methodological steps outlined in the protocol *(Chapter 8).* Therefore, Chapter 11 extends the analysis of Chapter 10 by focusing on intra-day data using a dynamic regression model with the aim of exposing relationships between the quantitative outcomes. However, rather than providing additional context to **PoI7**, Chapter 11 exposed further gaps in knowledge, warranting further research, outside the scope of this thesis.

1.5.6 Chapters 12: Summary of thesis

Chapter 12 concludes the thesis providing a summary, which briefly includes the findings and conclusions of each chapter. The final chapter also outlines the pragmatic implications of this research and highlights the contributions to knowledge that this body of work provides.

1.6 Previously Published Work

A significant portion of this thesis is adapted from previously published work, which was written and published as part of this PhD research. As such, various End User License Agreements are in place to the protect the copyright holders. Thus, it was deemed necessary to clearly signpost where published material has been used throughout the thesis. For clarity, a box will be placed on the title page of any chapter that utilises content from previously published work *(Figure 1)*.

> This chapter is adapted from previously published work to fit the context of this thesis. **<TITLE>** was adapted from the **<PUBLICATION TYPE>**: **<PUBLICATION TITLE>**, which was published in the **<JOURNAL NAME>** on **<PUBLICATION DATE>.**
> This work was made available online on **<ONLINE DATE>** via:
> <**D.O.I.**>
> **<COPYRIGHT DETAILS>**

Figure 1 - Example of signposting for previously published work

1.7 Contribution to knowledge

This thesis will provide a novel contribution to knowledge by presenting a framework and toolkit on which to construct, deploy and analyse personalised monitoring approaches for IEQ research. A detailed statement of contributions will be presented in Chapter 12, which will outline the key contributions, as well as the unique and substantial contribution to knowledge made within this thesis.

1.8 Next steps

The following chapter will present a comprehensive scoping review of literature, to begin addressing and answering the research questions. Chapter 2 will also outline other gaps in knowledge that will shape the direction of the research undertaken as part of this Ph.D.

Chapter 2 A Scoping Review of Technological Approaches to Environmental Monitoring

This chapter is an adapted from previously published work to fit the context of this thesis. The article: **A Scoping Review of Technological Approaches to Environmental Monitoring**, was published in the **International Journal for Environmental Research and Public Health** and was made available online on **4 June 2020 via:**
https://doi.org/10.3390/ijerph17113995

2.1 Introduction

To first approach the research questions and gain an understanding of the knowledge gaps within relevant literature, a review was conducted to compare and evaluate sensor technologies used for measuring IEQ in buildings. The methodological approach deemed most suitable was a Scoping Review (ScR) as it enables a more rapid integration of evidence from broad research areas, when compared to e.g., systematic literature reviews [33]. Moreover, a ScR provided a useful mechanism for mapping research topics where the extent of literature for this thesis is unclear [34], [35] due to the multi-disciplinary nature of the topic. The rigour of the ScR was strengthened by the adoption of the Preferred Reporting Items for Systematic Reviews *(PRISMA-ScR)* [36]. The PRISMA-ScR methodology provided a systematic reporting framework/mechanism, in the form of an itemised checklist to ensure the final review could be considered robust and thorough.

Box 2.1

The original PRISMA-ScR checklist references elements in the original publication, which have been redacted from this chapter to suit the context and narrative structuring of this thesis. For systematic validity, the checklist has been included in Appendix C, but references page numbers of the original publication. Therefore, readers are directed to the original publication *(details on title page of this chapter)* to view the complete checklist and the original unedited content of the manuscript.

This chapter will outline the methods used to identify literature – detailing the search process and methods for screening. Details and metadata extracted from the literature will be presented in a series of tables that allow for quick synthesis of information. Following this, a structured analysis of studies is presented before findings and conclusions are discussed. The main objectives of this review are to address the points of interest: **PoI1 – 3**:

1. What constitutes IEQ and how it is currently measured?
2. What sensor technologies are used to capture IEQ?
3. What is the current state-of-the-art in environmental monitoring?

2.2 Methods

This section will detail the methods used to conduct this inquiry and are presented in accordance with the PRISMA-ScR framework.

2.2.1 Searching and Selection Strategy

The initial phase of the review involved identifying keywords and filters that would be used to build search terms. The key filters used in this review are shown in Table 2. Combinations of filters were used to build search terms; for example, '(#2 AND #7 AND #14)', which equated to '((IEQ OR "Indoor (Environment OR Environmental) Quality") AND Sensors AND Building'. All queries were joined with an AND clause and OR was only used as indicated in Table 2. Literature was selected initially based on the inclusion of keywords within the titles. The abstracts and findings were then subsequently scanned to identify suitability to the aims and purposes of this review. Building type was also factored in when selecting studies. Focus was on residential and commercial buildings, but educational building studies were also included, given that mix of open-plan and enclosed spaces provides parity to office-based studies. Laboratory studies that analysed the technology were also selected to gain an understanding of the benchmarking process of environmental monitoring devices.

Table 2 - List of Search Terms (Filters)

1	(Well?Being OR Wellbeing)[a]
2	(IEQ OR "Indoor (Environment OR Environmental) Quality")
3	(IAQ OR "Indoor Air Quality")
4	("Sick Building Syndrome" OR SBS)
5	"(Thermal OR Visual OR Acoustic) Comfort"
6	Indoor Pollution
7	(Arduino OR "Raspberry Pi" OR "rPi")
8	Sensors
9	("State?of?the?art" OR Industrial OR "Scientifically Valid*")[a]
10	("Low Cost" OR DIY OR Cheap)
11	(Heating Ventilation Air Conditioning OR HVAC)
12	Wearable
13	(POE OR "Post?Occupancy Evaluation")[a]
14	Building
15	"Building Design"
16	"Green Building"
17	"Built Environment"
18	Office
19	Workplace
20	"Commercial Building"
21	Housing
22	Residential

[a] The 'single-character' search wildcard '?' was used on all databases except Google Scholar, which requires the '~' wildcard instead.

2.2.2 Eligibility Criteria and Information Sources

Given the nature of emergent technology, the scope of this inquiry was limited to publications within the last ten years and to literature which directly used hardware to monitor IEQ and not surveys and occupant feedback only. The primary source of information was peer-reviewed academic journals and conference proceedings and multiple databases were used to search for literature, including ScienceDirect, Scopus, PubMed, IEEE Xplore and Google Scholar.

2.2.3 Charting Screening and Synthesising Data

Data extracted from selected studies were collated and presented. Table 3 provides an overview of building types, environmental factors and the types of technology used to measure them. Where possible, demographic details were also extracted. In total 27 papers were reviewed. Data on measurement devices were subsequently presented in Table 4 and Table 6, which identify specific environmental factors that were measured. Table 4 outlines state-of-the-art sensors, which are devices that are used throughout the literature as a reference standard to either validate low-cost sensors or as standalone measurement devices. Table 6 outlines low-cost devices, electronic components that can be incorporated into DIY monitoring devices. An exception was the Netatmo Weather Station, because it is cheaper than some of the more expensive DIY components such as the GSS COZIR or the Sensorist Wireless Pro T/RH and also because it requires calibration against a reference standard [37]. In total, 33 state-of-the-art and 28 low-cost devices were identified across the 27 papers. The sensors that were used indicated that there is a prevalence of IAQ and thermal comfort across the studies but with many inconsistencies relating to measurement. Throughout the review, the data will be used to communicate the technologies, methodologies and findings from the selected studies and their relationship to state-of-the-art and low-cost environmental sensing. The reader should use Table 3, Table 4 and Table 6 for reference, when reading the data syntheses, as these tables categorise studies according to key data items, technologies and demographics.

Table 3 - Overview of Measurements in Selected Studies (1/3)

	Ref	Year	Building Type	Duration	Sample Size†	Demographics	Research Focus	IAQ‡	VC§	AC¶	TC††	SotA‡‡	LCS§§	DIY¶¶	WS†††	BMS‡‡‡
1	Rogage *et al.* [38]	2019	Residential (Multi-unit)	6 months	-	Residents from 7 flats, multi-unit social home building	IEQ/OC/ STP	-	-	-	✓	✓	-	✓	-	-
2	Clements *et al.* [39]	2019	Commercial (Office)	18 weeks	8	Office workers	OC	✓	✓	✓	✓	✓	-	✓	✓	✓
3	Ghahramani *et al.* [40]	2019	Education (University)	1 day	41	18-24-Year-old students uniformly random mix gender	OP	✓	-	-	-	✓	-	-	✓	-
4	Parkinson *et al.* [20]	2019	Commercial (Office)	3 months	-	-	IEQ	✓	✓	✓	✓	✓	-	✓	-	✓
5	Coleman and Meggars [41]	2018	Education (University)	8 days	-	-	STP	✓	-	-	✓	-	-	✓	-	-
6	Moreno-Rangel *et al.* [26]	2018	Residential (Flat)	4 days	-	-	STP	✓	-	-	-	✓	✓	-	-	-
7	Tiele *et al.* [42]	2018	Laboratory	3 days	-	-	STP	✓	✓	✓	✓	-	-	✓	-	-
8	Tijani *et al.* [43]	2018	Laboratory	1 day	-	-	STP	✓	-	-	✓	-	-	✓	-	-
9	Broderick *et al.* [23]	2017	Residential (Single-Family)	1 day	55	Non-smoking family with one or two adults and children. The average occupancy of 3.7 per household	IEQ	✓	-	-	✓	✓	-	✓	-	-
10	Földváry *et al.* [28]	2017	Residential (Multi-Unit)	1 week, x2	94	One participant from each household	IEQ	✓	-	-	✓	✓	-	-	-	-
11	Li *et al.* [32]	2017	Residential Commercial	6 weeks 3 weeks	3 7	- -	OC	✓	✓	✓	✓	-	✓	-	✓	✓

Table 3 - Overview of Measurements in Selected Studies (2/3)

	Ref	Year	Building Type	Duration	Sample Size†	Demographics	Research Focus	IAQ‡	VC§	AC¶	TC††	SotA‡‡	LCS§§	DIY¶¶	WS†††	BMS‡‡‡
12	MacNaughton *et al.* [16]	2017	Commercial (Office)	5 days	109	Office workers aged 20-70 near equal male:female ratio	IEQ/OP	✓	✓	-	-	✓	✓	-	✓	-
13	Tang *et al.* [10]	2017	Commercial (Office)	3 weeks	-	-	IEQ	✓	-	-	✓	-	-	✓	-	-
14	Tanguy *et al.* [44]	2017	Residential (Single-Family)	-	8	-	STP	✓	-	-	✓	-	-	✓	-	-
15	Tran *et al.* [45]	2017	Laboratory	-	-	-	STP	✓	-	-	✓	-	-	✓	-	-
16	Ali *et al.* [46]	2016	Lab, Office, Outdoor	7 days	-	-	STP	✓	-	-	✓	✓	-	✓	-	✓
17	Coombs *et al.* [21]	2016	Residential (Multi-Unit)	1 year	64	Predominantly African American 7-12-year-old asthmatic children from low-income families	IEQ	✓	-	-	✓	✓	-	-	-	-
18	Allen *et al.* [37]	2016	Commercial (Office)	2 weeks / 6 Days	30 / 24	Knowledge workers (professional grade employees)	IEQ/OP	✓	✓	-	-	✓	✓	-	-	-
19	Marques and Pitarma [6]	2016	Laboratory	-	-	-	STP	✓	✓	-	-	-	-	✓	-	-
20	MiHai and Iordache [47]	2016	Education (University)	5 hours	115	Students and teachers	IEQ	✓	✓	✓	✓	✓	-	-	-	-
21	Mui *et al.* [7]	2016	Commercial (Office)	-	-	-	IEQ	✓	✓	✓	✓	✓	-	✓	-	✓
22	Shan *et al.* [48]	2016	Education (University)	2 days	39	University Students with 6:7 male-female ratio	IEQ/OP	✓	-	-	✓	✓	-	-	-	-
23	Salamone *et al.* [49]	2015	Laboratory	3 days	-	-	STP	✓	✓	✓	✓	-	-	✓	-	-

Table 3 - Overview of Measurements in Selected Studies (3/3)

	Ref	Year	Building Type	Duration	Sample Size†	Demographics	Research Focus	IAQ‡	VC§	AC¶	TC††	SotA‡‡	LCS§§	DIY¶¶	WS†††	BMS‡‡‡
24	Hua *et al.* [27]	2014	Education (University)	4 weeks	46	20 - 50-year-old students and staff members, with the majority being between 20-29 years old	IEQ/OC	✓	✓	✓	✓	✓	-	-	-	-
25	McGill *et al.* [22]	2014	Residential (Multi-Unit)	1 day, x2	13	3 properties with an average of four people per house and at least one smoker in the family - non smoking	IEQ/OC	✓	-	-	-	✓	-	-	-	-
26	De Giuli *et al.* [4]	2012	Education (School)	1 day	-	Primary school children from seven Italian schools	IEQ/OC	✓	✓	✓	✓	✓	-	-	-	-
27	Painter Brown *et al.* [24]	2010	Commercial (Office)	1 month	-	-	STP	✓	-	-	✓	✓	-	-	-	-

† Sample size refers to the number of people measured in each study

‡ Indoor Air Quality

§ Visual Comfort

¶ Acoustic Comfort

†† Thermal Comfort

‡‡ State-of-the-Art

§§ Low-Cost Sensors

¶¶ 'Do It Yourself' Sensors (Standalone electronic sensing components, often run through Arduino/Raspberry Pi)

††† Wearable Sensors

‡‡‡ Building Management System

Research Focus Key

OC	Occupant Comfort
OP	Occupant Performance
IEQ	Indoor Environment Quality
STP	Sensor Technology Performance

Table 4 - State-of-the-art sensors (1/2)

		IAQ[†]					TC[‡]			VC[§]		AC[¶]
Manufacturer	Model	CO_2[††]	CO[‡‡]	H_2CO[§§]	PM[¶¶]	VOC[†††]	Temp	Air Vel.	RH[‡‡‡]	Lux	Light Colour	Sound
SKC	AirChek 2000 [21]	-	-	-	-	-	-	✓	-	-	-	-
Bruel & Kjaer	1213 [4]	-	-	-	-	-	✓	✓	✓	-	-	-
	2250 [47]	-	-	-	-	-	-	-	-	-	-	✓
CO2Meters	CM-0018AA [48]	✓	-	-	-	-	✓	-	✓	-	-	-
Extech	SD800 data logger [27]	✓	-	-	-	-	✓	-	✓	-	-	-
	EA80 data logger [22]	✓	-	-	-	-	✓	-	✓	-	-	-
Fieldpiece	SCM4 [20]	✓	-	-	-	-	-	-	-	-	-	-
GrayWolf	FM-108 [23]	-	-	✓	-	-	✓	-	-	-	-	-
	IQ-410 [26]	✓	✓	-	-	✓	✓	-	✓	-	-	-
	IQ-610 [23]	✓	✓	-	-	✓	✓	-	✓	-	-	-
	PC-3016A [26]	-	-	-	✓	-	✓	-	✓	-	-	-
	TG-502 [23], [26]	-	-	-	-	✓	✓	-	✓	-	-	-
HalTech	HFX205 [20]	-	-	✓	-	-	✓	-	✓	-	-	-
HOBO	U12-012 [46]	-	-	-	-	-	✓	-	✓	✓	-	-
Konica Minolta	CL-500A [39]	-	-	-	-	-	-	-	-	✓	✓	-
Lascar	EL-USB-CO [23]	✓	-	-	-	-	-	-	-	-	-	-
Monnit Corp	Wireless Humidity Sensor [39]	-	-	-	-	-	-	-	✓	-	-	-
	Wireless Temp Sensor [39]	-	-	-	-	-	-	-	-	-	-	-
NTi Audio	XL2 Analyzer [39]	-	-	-	-	-	-	-	-	-	-	✓
Rion	NL-52 [20]	-	-	-	-	-	-	-	-	-	-	✓
Telaire	7000 [46]	✓	-	-	-	-	-	-	-	-	-	-
	7001 [7], [23]	✓	-	-	-	-	-	-	-	-	-	-
TSI	DustTrak II 8532 [20]	-	-	-	✓	-	-	-	-	-	-	-
	Q-Trak 7575 [20], [37]	✓	✓	-	-	✓	✓	-	✓	-	-	-
	Q-Trak 964 [39]	-	-	-	-	-	✓	✓	✓	-	-	-

Table 5 - State-of-the-art sensors (2/2)

		IAQ†					TC‡			VC§		AC¶
Manufacturer	Model	CO_2††	CO‡‡	H_2CO§§	PM¶¶	VOC†††	Temp	Air Velocity	RH‡‡‡	Lux	Light Colour	Sound
Watson	Velocicalc 9545 [48]	-	-	-	-	-	✓	✓	✓	-	-	-
Watson	N-8681 SOLAR [22]	-	-	-	-	-	✓	✓	✓	✓	-	-
Wilks	InfraRan Specific Vapor Analyzer [48]	✓	✓	✓	-	-	-	-	-	-	-	-
Wholër	CO_2 datalogger [22]	✓	-	-	-	-	-	-	-	-	-	-
Wovyn	Lux1000	-	-	-	-	-	-	-	-	✓	-	-
Wovyn	Color Lux1000	-	-	-	-	-	-	-	-	✓	✓	-

Table outlines state-of-the-art sensors used within reviewed studies, outlining the manufacturers, models and measurement factors. These sensors cost range from several hundreds of pounds to several thousand.

† Indoor Air Quality; ‡ Thermal Comfort; § Visual Comfort; ¶ Acoustic Comfort; †† Carbon Dioxide; ‡‡ Carbon Monoxide; §§ Formaldehyde; ¶¶ Particulate Matter (PM1.0/PM2.4/PM10); ††† Volatile Organic Compounds; ‡‡‡ Relative Humidity

Table 6 - Low-cost sensors

			IAQ[†]					TC[‡]		VC[§]	AC[¶]
Manufacturer	Sensor	Cost[††]	CO_2[‡‡]	eCO_2[§§]	CO[¶¶]	PM[†††]	VOC[‡‡‡]	Temp	RH[§§§]	Lux	Sound
Adafruit	DHT22 [41], [49]	£2 - £5	-	-	-	-	-	✓	✓	-	-
	MAX 4466 [42]	£1 - £7	-	-	-	-	-	-	-	-	✓
Amphenol	T6615 [6]	£80	✓	-	-	-	-	-	-	-	-
	T6713 [41], [42]	£70-£75	✓	-	-	-	✓	-	-	-	-
AMS	CCS811 [42]	£6 - £30	-	✓	-	-	✓	-	-	-	-
	iAQ-Core C [42]	£15 - £30	-	✓	-	-	✓	-	-	-	-
	TSL2561 [42], [46]	£4 - £7	-	-	-	-	-	-	-	✓	-
BuildAX	Wireless Building Monitoring System [50]	£90	-	-	-	-	-	✓	✓	✓	-
CO2 Meters.com	K-30 [7], [46], [49]	Price by quotation	✓	-	-	-	-	-	-	-	-
GSS	COZIR [32]	£155	✓	-	-	-	-	-	-	-	-
Hanwei	MQ7 [6], [43]	£2 - £7	-	-	✓	-	-	-	-	-	-
Honeywell	HIH-4030 [43]	£10 - £40	-	-	-	-	-	✓	✓	-	-
	HPMA115S0 [42]	£35 - £45	-	-	-	✓	-	-	-	-	-
Netatmo	Weather Station [16], [37]	£130	✓	-	-	-	-	✓	✓	-	✓
Seeed Technology	MH-Z16 [39]	£65 - £100	✓	-	-	-	-	-	-	-	-
	MH-Z19 [39]	£15	✓	-	-	-	-	-	-	-	-
	AM2302 [7]	£3 - £15	-	-	-	-	-	✓	✓	-	-
	101020030 [7]	£3 - £10	-	-	-	-	-	-	-	✓	-
	101020023 [7]	£4 - £6	-	-	-	-	-	-	-	-	✓
Sensirion	SHT10 [6]	£2 - £7	-	-	-	-	-	✓	✓	-	-
	SHT15 [46]	£4 - £25	-	-	-	-	-	✓	✓	-	-
	SHT31 [42]	£3 - £15	-	-	-	-	-	✓	✓	-	-
Sensorist	Wireless Pro T/RH [32]	£140	-	-	-	-	-	✓	✓	-	-
SGX SensorTech	MiCS-VZ-89TE [42]	£20 - £25	-	✓	-	-	✓	-	-	-	-
Sharp	GP2Y1010AU0F [43]	£10 - £15	-	-	-	✓	-	-	-	-	-
Telaire	T6615 [6]	£80	✓	-	-	-	-	-	-	-	-
	T6713 [41]	£75	✓	-	-	-	-	-	-	-	-
Texas Instruments	LM35 [43]	£1	-	-	-	-	-	✓	-	-	-

Table outlines low-cost sensors used within the reviewed studies, outlining the manufacturers, models, measurement factors and typical costs.
[†] Indoor Air Quality; [‡] Thermal Comfort; [§] Visual Comfort; [¶] Acoustic Comfort; [††] Costs are approximate and taken from Google Shopping Search Engine – prices vary according to manufacturer and retailer; [‡‡] Carbon Dioxide; [§§] Equivalent CO_2 (eCO_2) is the measure used to communicate the global warming potential of combined greenhouse gasses; [¶¶] Carbon Monoxide; [†††] Particulate Matter (PM1.0/PM2.4/PM10) ; [‡‡‡] Volatile Organic Compounds; [§§§] Relative Humidity

2.3 Understanding IEQ and how it is measured

IEQ measurement comes in two forms: (i) measurement of physical environmental changes that can be quantified using objective monitoring equipment and; (ii) subjective data on how occupants perceive indoor environments, via surveys or self-reporting which is referred to as 'comfort factors' [5] *(the reader is directed to that study for a very clear outline of comfort determinant types that are present within offices and residential buildings).* Hanc et al. [14] highlight the importance of clarity surrounding wellbeing and note that environmental studies often fail to make clear distinctions between outcomes and determinants. They note ambiguity between comfort, satisfaction, and wellbeing, found in many environmental studies, exacerbates this issue. This is problematic, as it prevents researchers and practitioners from being able to accurately compare studies in order to ask meaningful questions of IEQ. It also makes it difficult to determine what methods and research design models should be applied when attempting to measure IEQ, health or wellbeing in future research.

There are many environmental factors that can be measured quantitatively and there are many measurement devices available for this. However, given the complexity of IEQ, it typically can't be defined by a single outcome, though some have tried to encapsulate it in the form of an IEQ index [42], [47], [51], [52]. Tiele et al. [42] use IEQ-Index as a term to measure a range of environmental factors, including: temperature, humidity, Carbon Dioxide, Volatile Organic Compounds (VOCs), Carbon Monoxide, illuminance, sound levels and particulate matter less than 10 micrometres in diameter. Others [51] use the term I_{IEQ} to encapsulate the sample mean from the sum of Indoor Air Quality Index (I_{IAQ}), Thermal Comfort Index (I_{th}), Visual Comfort Index (I_v) and Acoustic Comfort Index (I_a), according to the following equation:

$$I_{IEQ} = \overline{\chi} = \frac{(I_{IAQ} + I_{th} + I_v + I_a)}{\eta}$$

Equation 1: equation for calculating IEQ Index [51].

It is suggested that this number could be used as a 'star rating', but comparing indices used in separate studies [42], [51] would be futile, given the two IEQ indices are actually measuring different environmental factors. Since there is no standardised approach to IEQ indexing, it is likely to further obfuscate the subject area and complicate inter-study comparison. The complexity of IEQ is seen throughout the literature as researchers attempt to provide their own in-depth overviews of what constitutes IEQ, where many authors accept and state that IEQ constitutes four key objective and subjective sub-factors: IAQ, visual comfort, acoustic comfort and thermal comfort [3], [5], [29], [51], [53]. Figure 1 highlights these factors and subfactors of IEQ and demonstrates the relationships between Perceived Environmental Quality factors, which are highlighted in the underlying Venn diagram. The determinants of each sub-factor are also included, but these are what are commonly quoted rather than an exhaustive list.

Figure 2 – A holistic and general capture of terminologies and themes used across the literature to discuss IEQ

2.3.1 Indoor air quality

Except for subjectively measured Perceived Air Quality (PAQ), most air quality measurements can be done objectively. Table 6 shows some of the measurable indoor pollutants that contribute to poor air quality. Common factors of air quality seen across the literature are CO_2 and VOCs which are fine breathable particles that are distributed into the air from building materials, food, viruses and furniture [54] (Figure 3). Many air quality studies [20], [23], [26], [37], [41], [42] measured IAQ from particulate matter and/or VOCs but the most commonly measured factor of air quality across the literature is CO_2 (Table 4 and Table 6). Probably due to the impact it has on workplace productivity as opposed to its impact on health.

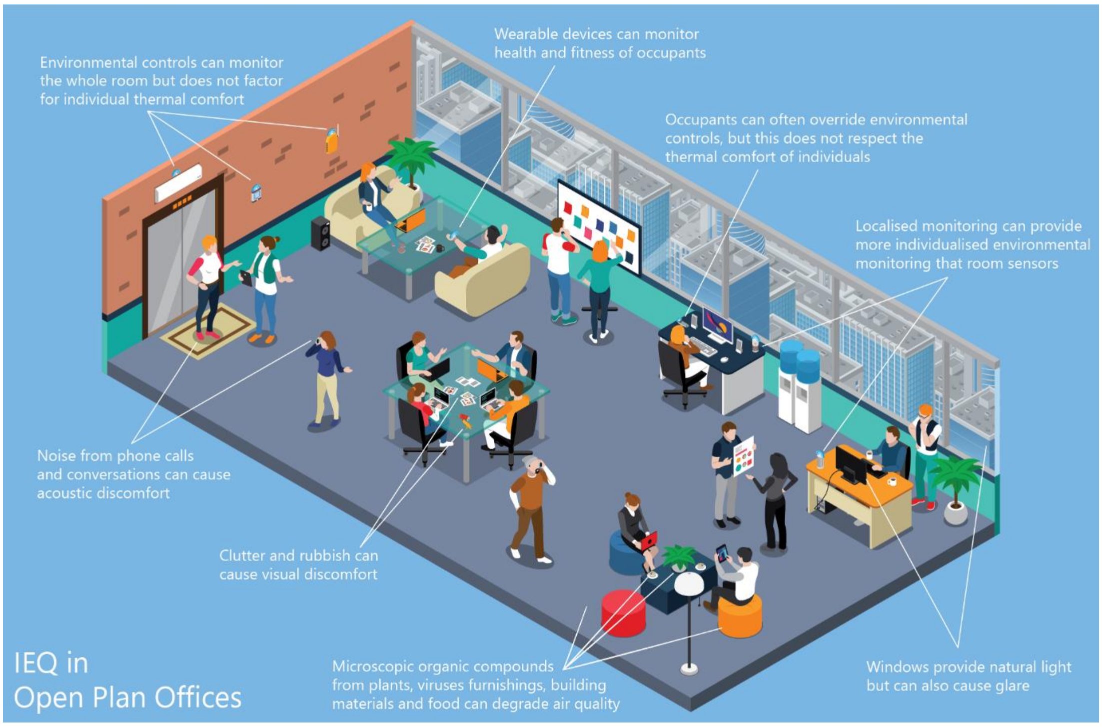

Figure 3 - IEQ in Open Plan Offices

2.3.1.1 Carbon dioxide (CO_2)

Health conditions attributed to CO_2 are not commonly present when exposed to less than 10,000ppm and CO_2 concentrations under 5,000ppm are considered safe for eight-hour exposure [55]. Whilst not immediately threatening to health, exposure to CO_2 levels above 1,000ppm can have an impact on cognitive functioning, productivity and comfort [37].

Allen et al. [37] monitored CO_2 in offices for two-weeks using an off-the-shelf Netatmo Weather Station, calibrated to a reference-standard TSI Q-Trak sensor. Data were supported by surveys and self-reported Sick Building Syndrome (SBS) symptoms, defined by The World Health Organisation as reported health-related symptoms that are caused by poor IEQ [56]. After the initial two-week period, participants were relocated to a building certified as Platinum by the LEED green building standard [13]. The study found participants reported 43% more SBS symptoms when the CO_2 levels rose above 1000ppm. However, authors note that participants were aware of the test conditions including relocation details to a high-standard ‘green building’. Participants reported more SBS symptoms when they were in a ‘non-green building’, even when environmental conditions in the building were optimal. Given how subjective occupant perceptions are, passive sensors can be an important way to reinforce findings through objective measurements.

Shan et al. [48] found links between CO_2 and SBS. Their study monitored air quality and thermal comfort of two rooms using a range of state-of-the-art sensors (Table 4). Thirty-nine participants completed self-reported SBS symptom and thermal comfort questionnaires. Additionally, participants completed a series of tests that would evaluate their cognitive abilities, whilst air quality measurements were conducted. Authors found inverse correlations between cognitive performance and CO_2 concentration levels with CO_2 to be the main cause of SBS symptoms. Those authors suggest that since CO_2 concentrations correlate with SBS symptoms, it is possible that higher CO_2 concentrations attributed to decreased performance because participants were also experiencing discomfort. Their study also identified correlations between CO_2 and other airborne contaminants, making it difficult to establish definitive causal links between their outcomes and CO_2 concentration levels. Some other studies suggest CO_2 is an inadequate measure of Indoor Air Quality (IAQ) [20], [41], [57] and airborne contaminants such as particulate matter and VOCs are a more valuable indicator of IAQ. However, CO_2 concentrations are known to increase when Air Exchange Rates are reduced [20], [38]. This may indicate why increased CO_2 is found to correlate with concentrations of airborne contaminants.

2.3.1.2 Airborne contaminants

Particulates and VOCs are known to accumulate within indoor environments and are regarded as a great environmental risk to health [8]. Building standards such as LEED and BREEAM, provide guidance and accreditation for the management of IAQ. However, only a small amount of accreditation points are awarded for it so there are insufficient incentives to encourage the additional work [58]. Alternatively, energy performance is often more valuable, but studies [21]–[23] show that reduced air-flow and increased air-tightness required to increase energy performance, results in the concentration of contaminants and a reduction of IAQ.

Coombs et al. [21] investigated non-green, multi-residential apartments home to asthmatic children (7-12yrs). The inquiry was conducted as buildings were renovated to comply with green building standards. Airflow and IAQ were monitored in eight homes before and after the renovations using the reference standard SKC AirCheck 2000. Air filters were attached to the latter in order to collect airborne contaminants. As a control, IAQ was simultaneously monitored in a low-income, non-green, multi-residential complex. Authors discovered significant differences in properties before and after renovations and found reduced airflow and increased airtightness, typically required to increase energy performance, resulted in an increased concentration of contaminants and a reduction of IAQ. Similarly, Broderick et al. [23] monitored fifteen, three-bedroom, semi-detached social housing properties. Their study measured airborne contaminants (not airflow) using a range of state-of-the-art sensors (Table 4). IAQ was monitored in the living room and master bedroom of each property before and after an energy performance renovation. Their study revealed an 18-25% increase in CO_2 and VOC concentrations levels and a 40% increase in the concentration of particulate matter up to 2.5µm in diameter (PM2.5) after buildings were renovated. Findings also revealed negative correlations between energy performance and air quality. Furthermore, whilst CO_2 may not provide an adequate determination of IAQ, there are links between PM2.5, VOCs and CO_2 and those links may explain the prevalence of CO_2 in IAQ studies across the literature [21].

2.3.2 Thermal comfort

IAQ is not a standalone factor of environmental quality, being influenced by many other objective, and subjective IEQ factors. Occupants have been found to report poor PAQ when they are thermally uncomfortable [59], explaining why many studies focus on IAQ and thermal comfort. ISO 7730 [60] defines thermal comfort as being associated to a person's thermal balance and is affected by clothing, physical activity, temperature, humidity, movement of air and the average temperature of surfaces in a room (Mean Radiant Temperature, MRT), Table 6. Prevalence of thermal comfort across the literature may be due to its intrinsic influence over PAQ, but it may also attribute to the maturity of building standards that focus on thermal comfort. Those standards specify thermal comfort factors and how to measure it. This standardisation of measurement means that thermal comfort studies can be directly compared.

2.3.2.1 Predictive Mean Vote (PMV)

The ASHRAE Standard 55 widely measures thermal comfort [4], [32], [39] and should be done using occupant satisfaction surveys, point-in-time surveys and electronic sensors measuring the thermal environment [61]. ISO 7730 and ASHRAE 55 standardise a Predictive Mean Vote (PMV) steady-state model [62] used for measuring thermal comfort. It does this by predicting occupants mean thermal perception and predicts the percentage of those who will be dissatisfied by the thermal conditions [63]. However, it was found that this approach did not accurately predict thermal comfort [64], which may be attributed to the fact it wasn't field-tested before being incorporated into the standard [65]. The model has also been criticised for its ineffectiveness in naturally-ventilated buildings [32], [66], it does not account for climatic differences and occupants are given little to no control over their own thermal comfort.

2.3.2.2 Adaptive comfort model

To challenge concerns around the PMV model, an adaptive model was developed [66], which acknowledged that occupants in naturally ventilated buildings have a much broader tolerance threshold for thermal conditions than those in mechanically ventilated buildings. That model is recognised in British Standard EN 16798-1 [67] as well as the deprecated EN 15251 [68]. De Giuli et al. [4] utilised the latter standard, prior to its deprecation, to assess whether children would be able to perceive environmental changes in non-mechanically ventilated schools. This was measured according to the ISO 7730 standard [60] using a Brüel & Kjær climatic analyser. De Giuli et al. [4] found that children were able to perceive poor air quality and noise, perceiving poor thermal comfort in the summer. They also noted that since the environmental conditions of the classroom were set according to the preferences of the teacher, children were found to be unaware of many conditions or behaved as passive users of the environment. This was found to be the case in all schools other than mechanically ventilated schools where students showed they were more aware of the environmental conditions and they had more control over it.

2.3.2.3 Occupant control

It is believed that occupants should be given control over mechanical ventilation systems [32], as environmental control plays a role in personal comfort [69]. Li et al. [32] found that environmental studies were often limited by ventilation systems that lack the capabilities to facilitate such control. However, providing user access could lead to dissatisfaction due to individualised perceptions of comfort (Figure 3) but too much control can also distract workers from their duties [70]. To overcome these issues, Li et al. [32] designed a system which allowed participants to use a smartphone to vote on how comfortable they are and provide details of their clothing and level of activity. This allowed occupants to have control over their environment, without control becoming a distraction. The smartphone application also collected data from a COZIR CO_2 sensor, a Sensorist Wireless Pro Temperature and Humidity sensor. Data were collected and combined with participant votes and were used to alter the HVAC thermostat setpoint. To test, authors conducted studies in single-occupancy rooms and in an open-plan space. By substituting static-setpoint thermostats with their system, they found that reports of thermal discomfort dropped by >50%. Inclusion of physiological data also serves to remove much of the subjectivity from traditional thermal comfort measurements. However, their study failed to take air velocity or MRT into account and instead measured CO_2 as a determinant of IAQ, but since their system was developed to take data from multiple sources, it is likely that it could be adapted to include measurements of air velocity and MRT.

2.3.3 Visual comfort

Visual comfort is highly influential on other subjectively measured environmental factors. Table 6 shows some visual comfort factors and highlights its influence over PAQ and acoustic comfort. Although there are many subjective visual comfort factors, light intensity or illuminance, measured in LUX (lx) is the major objective measurement (Table 3). The threshold for illuminance is dependent on task but for most office tasks, thresholds range from 300lx to 500lx [71]. Yet, ambient illuminance will likely not reflect the light levels at individual workstations. Light reflectance and glare can cause areas of visual discomfort (Figure 3), but the ambient illuminance can be within specified limits.

Many green building standards recognise visual comfort extends beyond base level illuminance and are establishing new parameters of visual comfort. The WELL Building Standard includes a range of factors including glare control, fenestration of daylight, ergonomics of the space design and lighting colour [12]. All but one study measured visual comfort using LUX only (Table 3). The exception [39], measured visual comfort using a combination of LUX and a measure of lighting colour, using the Wovyn Color Lux1000. Their approach measured the impact of blue light on sleep but also the effectiveness of window tinting on a room's ambient colour temperature by placing light sensors at desk level and in elevated positions. This allowed authors to identify the environmental variability of daylight and artificial lighting in buildings. They acknowledged that studies in controlled environments could mitigate this variability, but felt their study provided more natural conditions. By including RGB sensors they identified many key aspects of visual comfort identified by the WELL Building Standard [12].

2.3.4 Acoustic comfort

Noise is a major contributor to discomfort in many naturally ventilated buildings [70]. Noise can come from a number of sources, but any sound that causes distractions to everyday actions, such as relaxation or work, can be considered noise in the context of occupant comfort [72]. A few studies [7], [16], [20], [44] measured noise using a microphone, which provides a measurement of Sound Pressure Level (SPL) in Decibels (dB), Table 3. According to the ASHRAE's guidelines [73], the sound levels in open-plan offices should not exceed 45dB [74]. This is closely mirrored by the WELL standard for sound masking in those spaces [12], which states levels should not exceed 48dB. Whilst those standards provide strict limits on noise levels, they do not translate to acoustic comfort. In offices, noises often come from mechanical or electrical equipment, conversations or phone calls from surrounding occupants [39], [75] (Figure 3).

Notwithstanding the fact that office noises can be a great source of discomfort for building occupants, they are often well within the specified limits. As noted by Tiele et al. [42], this makes SPL measurement ineffective at measuring acoustic comfort, as levels of perceived noise may not match those captured by electronic equipment. Moreover, they indicated that there are more quantitative measurements that should be considered in research, such as sound variations and peaks. Even with objective methods to support measurement, acoustic comfort is predominantly subjective. With 1-in-6 people in the UK suffering from hearing impairment [76], this subjectivity must be assessed on a case by case basis. Unlike thermal and visual comfort, acoustic standards specify the thresholds of the objective measures and do not provide a standardised approach to measuring occupant perceptions of their acoustic environment.

2.4 Understanding state-of-the-art environmental monitoring

BMSs control and manage building assets such as HVAC systems [77]. These systems are typically used by facility managers for scheduling asset maintenance but extend to the collection, storage and transmission of asset data using built-in state-of-the-art sensors. HVAC systems are often retrofitted or preinstalled with sensors that monitor air quality, temperature, humidity and flow. Using a BMS to monitor assets is considered to be a well-established approach that provides useful data [24]. Occasionally, this may not provide a useful monitoring solution when supplying air to multiple spaces. Generally, air distributed in each space can be monitored with a single measurement point, but this does not always provide an accurate portrayal of the environmental conditions experienced by occupants [20]. Handheld monitoring devices may provide a better-individualised approach.

2.4.1 Data loggers

Five studies referenced BMSs and HVAC systems, most of which used state-of-the-art monitoring devices in commercial buildings. However, there are many types of buildings that do not have the supporting assets to warrant using a BMS, such as single-family residential buildings. In those, environmental data can be collected using data loggers, which are portable monitoring devices with built-in storage. Several studies investigated IEQ within multi-occupant spaces using state-of-the-art data loggers [4], [24], [27], [47] (Table 4). Typically, most are designed to be handheld for point-in-time measurements or periodically mounted within buildings for continuous monitoring.

The most common data logger manufacturers were TSI and GrayWolf with prices ranging between a few hundred (Onset Hobo U12-012) and to several thousand pounds sterling (TSI DustTrak 8532). Whilst the accuracy and precision of these can make them extremely valuable tools, there are many drawbacks making them unpragmatic. For example, with the Extech SD800 or the Wholër CDL 210, data are stored within internal memory and later downloaded. Therefore, real-world applications are limited to point-in-time measurement or short-term studies such as Post Occupancy Evaluations (POEs).

Primarily, a POE is the process in which buildings are evaluated, after the point of occupancy, to assess whether the building performs according to the occupants' needs [78]. POE also focuses on post-construction building performance to assess whether it meets design specifications [79]. When buildings are designed according to standards that specify IEQ thresholds there is a need to measure environmental factors to ensure it meets those standards after occupation, typically running for two to eight weeks [80], [81]. Whilst dataloggers of this type may be ideal for conducting such evaluations, they lack the ability to provide real-time feedback making them impractical for continuous environmental monitoring.

2.4.2 Scalability limits around state-of-the-art solutions

Open-plan office studies often mitigate the state-of-the-art cost by measuring multiple participants in a single location, as fewer sensors are needed to measure the space and larger sample sizes can be observed. However, since residential studies measure participants across multiple properties, small sample sizes [32], [39], [44] and short measurement periods [4], [22], [23], [43] are often built into the research design to address budgetary restrictions. McGill et al. [22] measured air quality in buildings that were built according to the German Passivhaus standard, an approach using passive design systems to maintain a balance between environmental quality and energy use [82]. They measured air quality using state-of-the-art data loggers from Extech and Wholër. However, the measurement period was short and the sample size was both limited and split across multiple buildings. This resulted in findings that can only be used to provide insights.

Contrastingly, a commercial office study [39] measured a similar sample size but because participants were within the same environment, sensors could be used to simultaneously measure multiple occupants. Large portions of the building could also be monitored more easily; however, authors note this made it difficult to provide individuals with paralleled IEQ environments. However, the cost of the equipment used in this study means that it is highly unlikely that this methodology could be applied outside of research.

An emergent market of low-cost accessible devices is available, which opens use cases that can drive the future of research, whilst addressing capital investment requirements. It is therefore important to understand and research where low-cost technologies can add value as standalone measurement devices for both researchers and practitioners.

2.5 Low-cost alternative technologies

Low-cost microcontrollers and microcomputers (e.g., Arduino and Raspberry Pi, respectively) are becoming valuable IEQ measurement tools. Many devices use open-source hardware, which advances technological development through a community [83]. This approach means users can become developers, instead of consumers. Furthermore, open-source hardware actively permits the creation of clones, which are cheaper alternatives to the official products or devices that are modified for a specific use [84]. Many of those sensors are now also being incorporated into "breakout" boards, which are low-cost, universal devices designed to interface directly with a serial bus on microcontrollers and microcomputers [85]. This means that technology used to monitor IEQ is becoming accessible, easier to develop and can be significantly cheaper than state-of-the-art counterparts [46]. Prevalence of DIY devices (Table 3), is testament to a paradigm shift that is breaking down IEQ monitoring entry barriers.

2.5.1 Limitations of low-cost sensors

Low-cost sensors also use cheaper components than state-of-the-art equivalents. For example, CO_2 sensors typically use infrared to detect gas concentrations. Sensors such as the MH-Z1x range and CozIR are cheaper alternatives to state-of-the-art CO_2 sensors, such as the Wholër CO2 Data Logger. Whilst all infrared CO2 sensors measure CO2 in the same way, cheaper components are used in low-cost sensors. Other low-cost sensors, such as the CCS811 and the iAQ-Core C, use different technology altogether by detecting gasses that come into contact with a semiconductor that has a Metal Oxide surface [86]. Those sensors typically provide a CO_2 measurement; though not actually a measure of carbon dioxide (Table 4). In fact data from the sensor measures the total concentration of VOCs in the air (TVOCs) [87] which is returned at a different scale factor known as Equivalent Carbon Dioxide (eCO_2). This approach lacks transparency as none of the datasheets [88]–[90] (Table 6) articulate how eCO_2 is calculated, nor do they state whether a standardised calculation method is used. Therefore, it would be difficult to distinguish whether any differences in measurements were caused by the conversions or the sensors. Furthermore, a lack of clarity about what eCO_2 is and how it is calculated has led some to consider it an actual measure of CO_2 [91], [92]. Issues such as these are likely contributing factors to mistrust with low-cost devices.

2.5.1.1 Accuracy vs. precision

Low-cost sensors have been found to be less accurate than state-of-the-art sensors [20] but they have been found to have good precision [86]. This means they may not provide an accurate measurement of environmental factors but will be responsive to changes. For example, if a room contains a CO_2 concentration of 650ppm and the CO_2 rises by 10ppm/minute, an inaccurate but precise sensor may read an initial value of 900ppm, but still, measure 10ppm concentration increases. Ultimately, the suitability of a device depends on the application, as a low-cost device would be unsuitable where precision and accuracy is needed for a building to ensure it meets a government regulation. Conversely, high-accuracy, low-precision devices may be unsuitable for studying CO_2 elevations on the concentration levels of occupants. This is because low accuracy devices can be calibrated against reference standards to mitigate the baseline offset, but removal of imprecision cannot be so easily achieved.

2.5.2 Scalability

Mihai and Iordache [47] highlight how cost can impact IEQ research. In their study, a single (£1500) CALCTM 7525 was used to measure air quality in university classrooms. Measurement of illuminance was conducted using an array of nine LUTRON LM-8102 light meters (£100/each) per room, placed at each student's desk. Whilst the LUTRON sensors are reference-standard equipment, the price difference of the sensors is indicative of the level of granularity in the measurement of the two IEQ factors. The effects of the inconsistencies in measurement granularity can be seen in the visualisations provided in their article which impacted their findings. Each index of IEQ was mapped to the floor plans of the building. IAQ, thermal comfort and acoustic comfort were measured and visualised on a room-by-room basis, where one room may perform better or worse than another. However, visual comfort was measured and visualised at an individual level, meaning certain areas of a single space were found to perform better or worse than others. This is significant as visual comfort was measured at an individual level only. The combined IEQ index was visualised using the buildings floor plans and their visualisation clearly highlights how data were skewed by the visual comfort measurements. The study serves as a good indicator of how scalability can affect study design, whilst clearly highlighting the value of the localised measurement.

Mihai and Iordache [47] measured environmental factors at an individual level but devices could be termed state-of-the-art and considered as relatively expensive when compared to other light sensors (Table 6). Notwithstanding measurement accuracy, low-cost devices may be more suitable for measuring how individuals are affected by environmental changes. By using such devices, it is possible to incorporate more sensors at a very low cost, which will allow measurement resolution to be increased and focused on the individual. To make this increased resolution scalable, there is a need to use a holistic system like a BMS to capture, record and analyse the data from multiple, different sensor sources but such systems are not always available or applicable.

2.5.3 Holistic cloud-based systems

Cloud-based platforms are rapidly increasing in popularity and are often inexpensive, open-source or are delivered as a scalable service but require internet-enabled measurement devices. Reference-standard, portable data-loggers are typically offline devices that store data. Wireless data loggers exist, but often only interface with proprietary web platforms and are more expensive than their offline counterparts [24]. Most of the low-cost sensors in Table 6 are not standalone wireless devices. Instead, they are sensor components that need to be connected using microcontrollers or microcomputers such as Arduino or Raspberry Pi. The underlying infrastructure for these devices is largely driven by Internet of Things (IoT) technologies, which describes a paradigm of interconnected devices that communicate through the internet [6]. Once connected, these devices can read and write data from sensors either to local SD card storage [42], [46] or transmitted wirelessly to cloud platforms to integrated with other devices or services [6], [7], [32], [37], [41], [43], [49], [50]. However, there are several approaches seen across the literature to bridge the gap between the device and the cloud.

Across the literature in Table 6, there are three approaches for wirelessly transmitting data from DIY sensors, the first approach simply involves using Wi-Fi enabled sensors in the first instance [32], [37]. However, the second and third approaches involve developing hardware devices with Wi-Fi capabilities and there are two approaches seen across the reviewed literature for doing this. One method involves using Wireless Sensor Networks (WSNs), which are a network of wireless devices (nodes) that connect to each other to form a network that enables data transmission over large distances with low power consumption [93]. This approach was found to be advantageous, as it facilitated the simultaneous collection of data from different devices through the various nodes [6]. However, additional to the sensor nodes, there is often a requirement for the network to contain access points and gateways, which the nodes must first communicate with [41]. Contrastingly, modern microcontrollers now come included with on-board Wi-Fi chips [7], [41], [43], [50], which allow the devices to directly communicate with wide area networks. Wi-Fi shields can also be used to addon wireless functionality to boards that otherwise would not have it [49]. This approach also has benefits as it removes the need for gateways and hubs, potentially reducing project costs.

Since many of these sensors are integrated into custom-made devices, they do not depend on proprietary systems to visualise or analyse data. This means developers have the freedom to connect to a wide array of cloud-based applications or create custom architectures which is reflected in the literature as no two studies implementing web-based platforms [6], [20], [32], [38], [39], [41], [43], [44], [49] used the same web architecture or visualisation platforms. Need for wireless monitoring has created a competitive market for cloud-based applications and interactive dashboards to display sensors data. Consequently, there is no one standardised approach seen across the literature for storing, recording, and analysing sensor data.

Cloud-based applications allow the creation of complex rules and associations [94], meaning that sensor data can be concurrently associated with a building, a room, and an occupant. That process can be streamlined and improved by incorporating data from 3D models containing Building Information Modelling (BIM) data, as shown in a recent study [38]. Those models contain a wealth of information about buildings including spatial structures and asset information. These data can be integrated into a holistic system that collects data from multiple sources including environmental sensors and subjective occupant feedback. This makes it feasible to monitor individual environments with a wide range of sensors and understand how building design and environmental changes impact occupant health and wellbeing.

2.6 Individualised IEQ approaches for health and wellbeing

Compared to health outcomes, measurement of wellbeing can often be challenging given the subjective nature of what is being measured. and lack of a standardised method for which to collect data. Moreover, the wording of questions, ambiguous responses and inconsistent administration techniques mean that there are many limitations with these measurements [95]. Furthermore, it is common for studies on wellbeing and the indoor environment to lack clarity in the methods used to collect subjective wellbeing data. For example, the clarity of questions is not explicitly detailed and/or there are unclear links to wellbeing outcomes [4], [9], [96], [97]. Yet, whilst ambiguity around research design does not invalidate findings, repeating studies or identifying patterns across the literature is challenging. Therefore, it is difficult to understand whether IEQ measurements have the efficacy of determining wellbeing. Given the tenuity of links between IEQ and wellbeing, there may be value exploring links between IEQ and health, as good health is found to positively impact wellbeing[98].

2.6.1 Holistic IEQ approaches

Researchers have discussed the prevalence of using sensors to monitor the relationship between occupants and their environments [6]–[10], [43] but few make the individual the primary unit of analysis. Measuring individual response to changes is a key factor of occupant comfort and described as an important requirement for environmental monitoring systems [31]. It is proposed that non-invasive, Wearable Health Technologies (WHTs) are an accessible way to monitor a range of psychological and physical health conditions such as depression and hypertension, respectively [99]. These technologies enable researchers to access a vast repository of individualised health biomarkers which could be used to augment passive IEQ measurement by using smartwatches, smartphones, smart-clothes and even smart-tattoos [100]. Three studies [16], [32], [39] used personal fitness trackers to monitor a variety of health data in relation to IEQ. These studies all involved the collection of data from multiple sensors and they each highlighted potential links between occupant physiology and IEQ. Moreover, the methods used in these studies highlighted the need and value of holistic IEQ approaches.

In recent years wearable health and fitness market has become saturated with new devices, so it is not always clear which devices are most appropriate where many are released and discontinued each year. A recent review [101] highlighted that Personal Fitness Trackers (PFTs) such as Fitbit and Garmin feature heavily across the literature, providing a checklist, which outlines eight categories to appraise PFTs, a useful starting point for those considering the use of PFTs in research projects. The wearable market seems to have split consumers into those who want expensive smartwatches that integrate with smartphones and those who want lower-cost PFTs, such as devices by Xaiomi and Huawei. The latter is driving down PFT cost, which means they could be built into scalable monitoring tools.

Many of the applications of PFTs across the literature involve either evaluating the gamification of health or looking at the effects PFTs have on daily routines. Measuring daily steps can have a positive effect on health, as it is indicative of a more active daily routine [102]. This can also serve as an objective measurement of activity levels that can be used to support IEQ studies. Though it is important that users are actively involved in the early stages of health technology research and design to ensure the technology is developed and appraised with a user-centric approach [103]. Moreover, it is important that during this process users are made aware of how their data will be collected, stored and analysed to ensure it is compliant with data protection standards (for example GDPR) and is done so in an ethical manner.

2.6.2 Linking health to wellbeing: Augmenting IEQ approaches

Continuous in-situ measurement via wearables has the potential to provide individualised health measurements, augmenting IEQ approaches by providing quantitative data to support qualitative data, by reducing errors found in subjective measurements of wellbeing [104]. MacNaughton et al. [16] combined wearable health data with data obtained from IEQ sensors and results from health surveys that assessed cognitive function. Whilst their study did not eliminate the need for subjective responses, the inclusion of the additional health data meant that additional insights were formed. The study found patterns between the quality of sleep and cognitive ability and associate the former with environmental conditions such as lighting. Moreover, by objectively monitoring the individual, was authors were able to make definitive links between environmental factors and physiological responses.

Diminishing costs and increased accessibility of environmental sensing technology mean that scalable solutions can be developed that takes a personalised approach to wellbeing measurement. Localised environmental monitoring augmented with WHT data can be fed into holistic monitoring systems, providing meaningful results to both building owners and occupants. Augmenting WHTs with low-cost IEQ approaches will allow research to extend beyond a single environment, as sensors can be placed in multiple environments, such as the home and workplace and WHTs can be worn as occupants move between environments, but to compare different environments the data collected from each environment must be comparable. Unfortunately, studies show that approaches to environmental monitoring differ greatly between residential and commercial buildings, as does the technology used. There is a clear need for research that addresses these knowledge gaps by considering longitudinal IEQ measurements and individualised approaches to monitoring alongside within the home and workplace. This will provide a better holistic picture of how their physiology is affected by those environments. It will also allow researchers to draw conclusions about the impact buildings have on occupant wellbeing.

2.7 Discussions and conclusions

This chapter presented approaches to IEQ measurement in buildings from a range of research domains, whilst exploring technologies and approaches to health and wellbeing trends.

2.7.1 Understanding IEQ

It is evident that IEQ is a complex and multi-faceted area of study and efforts to define it often clutter the definition rather than add clarity. This is exacerbated further as researchers attempt to encapsulate IEQ measurements into a single IEQ-Index [42], [47], [51], [52], which could have a profound impact on future research. Researchers may be inclined to compare like-for-like, but this may lead to inaccuracies unless this encapsulation follows a defined standard. Currently, IEQ indices should be treated with caution until the literature provides a common understanding of what constitutes an IEQ index, and which factors it encapsulates. It is also important, that researchers do not compare IEQ indices directly unless they are sure they are comparable measures.

There is a general acceptance that IEQ consists of four sub-factors: IAQ, visual comfort, acoustic comfort and thermal comfort. This may suggest that Equation 1 [51] is an appropriate index. However, there is further confusion about what constitutes each of those sub-factors. Given the current state of literature, it is not possible to gain a definitive understanding of what factors constitute IEQ, but this review has highlighted many of the factors which have been measured. This list is not exhaustive and the way in which IEQ is measured is inconsistent and conflicting across studies. Table 4 and Table 6 highlight only a snapshot of the environmental factors that make up the sub-factors of IEQ.

There is little commonality across the literature over the quantity or combinations of environmental factors that must be measured to satisfy a measure of air quality or thermal comfort. For example, Tiele et al. [42] measure IAQ with a single measure of CO_2. Whereas Li et al. [32] measure IAQ with multiple measures of VOCs, Carbon Monoxide, CO_2 and particulate matter. Whilst this may suit the needs of the individual inquiries it adds complexity to the subject area. It is pertinent that future research aligns to a common understanding of what constitutes IEQ. Though, more needs to be done to standardise and legislate a homogeneous IEQ measurement. Notwithstanding the confusion about what should be measured, there are a plethora of environmental factors that can be measured, using environmental sensors.

2.7.2 Understanding IEQ measurement technology

This review highlights the range of devices used and outlines how state-of-the-art monitoring devices compare with low-cost sensors. The primary differences between these devices are cost, accuracy, and connectivity. Since the cost of state-of-the-art sensors is prohibiting and difficult to promote on projects [7], it is important to understand the needs of the project before procuring hardware. It is also important to understand how the research is to be applied. If the purpose of the research is to benchmark low-cost sensors [26], then it is feasible to use reference-standard equipment to act as a baseline for measurement. If the study is proposing solutions that could be adopted by practitioners [39], it isn't feasible or pragmatic to propose such expensive equipment.

This review found that the accuracy and precision of devices are regularly questioned, particularly with low-cost approaches. Low-cost devices are found to have lower accuracy, but they have good precision [86]. Yet, whilst reference-standard devices may have higher accuracy in a laboratory test; in practice, equipment costs lead to fewer devices being used. This means it is not possible to gain an accurate indication of what individuals experience [20]. Using low-cost devices, it is possible to counteract the reduction in accuracy by increasing the number of measurements. It is feasible to measure individual environments so that data can be analysed more accurately alongside an individual occupant, but this results in a high degree of data points that need to be stored, visualised, and analysed.

This review discovered that there is a rapid market growth surrounding low-cost hardware and cloud-based applications that these devices can interface with. Such applications are hardware agnostic meaning that any device that can send data could send it to a holistic monitoring system. This may bring the power of high-cost BMSs to small businesses and even individuals. Moreover, since data obtained by these systems are not just limited to building assets, it is possible that occupant surveys and even data from wearable sensors could be included in holistic systems. This makes it possible to understand how occupants respond to environmental changes within their immediate environment.

2.7.3 Augmenting current IEQ approaches

I found that WHTs have demonstratable value to research in this area. Granular measurements can be taken at an individual level that can be used to support traditional environmental monitoring. Localised sensors can measure an individual's immediate environment, but WHTs could also be used to collect an individual's psychological and physiological data, used to better inform or support findings. WHTs also have the potential to monitor for health biomarkers, which could make it possible to understand how individuals are affected by changes in the environment from quantitative data alone. By measuring individualised physiological data, Li et al. [32] were able to demonstrate a personalised approach to monitoring that provides a wealth of data to inform, reinforce or even replace traditional subjective measurements of comfort. However, research is needed to understand, which health biomarkers and wellbeing parameters correlate with each other. This may make it easier to quantitatively associate IEQ with wellbeing, but it will also mean that researchers must pay closer attention to the physiological conditions, as they will have a greater impact on individualised measurements than they would on group studies.

2.7.4 Limitations

The key limitations in this review were establishing the boundaries of the review and the time taken to conduct such a broad range of topics, which spanned multiple research disciplines. Like all reviews unconscious bias could have influenced the search outcomes, which could have possibly been mitigated by a full systematic review. However, given the multi-disciplinary nature of the review, a systematic review would have likely been too focussed on one subject area, or complex to the point it would add little value. While presenting such a broad range of subject areas, in a way that aligned with the objectives of this study, was a time-consuming process, scoping reviews are designed to tackle this kind of broad, multidisciplinary meta-analyses. Therefore, a middle-ground was established to conduct a scoping review that was able to provide the scope of the literature, while applying the PRISMA framework to apply systematic approaches to the process.

2.8 Knowledge Gap

There is a need for a paradigm shift that makes the individual the unit of analysis. The high percentage of time people spend indoors is not only spent in one single environment; yet, most studies only target a single environment, such as schools, offices, or homes. Researchers should take advantage of the scalability of low-cost devices as it will enable them to incorporate more environments into their studies. Measurements could be taken at work, at home and even as they commute, to get a more holistic picture of each individual where notable trends could be observed from longitudinal periods of monitoring. However, there are few IEQ studies that do this. Instead, human participants are typically found to provide contextual data to reinforce environmental data.

The confusion around IEQ and the factors of measurement are present, but there is a wealth of literature that can inform researchers to draw their own arguments and ideas on which their work can be founded. However, the lack of studies that monitor the environment from the point-of-view of an individual is a knowledge gap that must be addressed. Also, given that there are recognised links between health and wellbeing [98], future research should explore these associations to identify any quantitative measurements of health that can be used as an indicator of wellbeing. By addressing these gaps, it may be possible to use low-cost IEQ sensors, WHTs, and cloud-based platforms to create holistic, personalised, and scalable wellbeing monitoring systems.

2.9 Addressing the PoI

Modern adaptive comfort models recognise the value of the individual and to make environmental monitoring feasible outside of research. In this chapter, a research need was identified around the exploration of holistic personalised approaches to environmental monitoring in buildings that analyse the individual as they move between environments.

PoI1:

What constitutes IEQ and how is it currently measured?

There is no standardised definition of IEQ, but it is generally regarded as a multifaceted environmental outcome. To date, it is measured in an ad-hoc manner but constitutes several important environmental domains*: air quality, thermal comfort, visual comfort, and acoustic comfort*, with each of those comprising a series of outcomes including: *temperature, humidity, light, noise, dust, volatile organic compounds, carbon dioxide (CO_2) and equivalent carbon dioxide (eCO_2).*

IEQ factors comprise of both subjective and objective determinants of environmental quality. While subjective (qualitative) data is measured predominantly with pen and paper self-reported surveys, objective determinants can be monitored through data captured with electronic sensors or from pen and paper, self-reported surveys. Although surveys are useful (and highly scalable) their subjective nature means they lack e.g., absolute clarity on personal perceptions and are burdensome due to their often-time-consuming requirements. In contrast, objective (quantitative) data measurements are often omitted, due to cost and complexity of electronic-based monitoring IEQ.

For studies that do use quantitative approaches a range of technologies can measure IEQ depending on the underlying infrastructures available in the building under analysis. If buildings have complex HVAC systems (typically seen in commercial buildings), measurements can be taken from state-of-the-art measurement equipment that is integrated into the air handling units and managed through a BMS. However, if buildings lack these infrastructures environmental monitoring can be retrofit into the building using portable monitoring devices, or data loggers placed in situ. Again, these devices are costly and have limited sensing modalities – meaning multiple devices are often required to comprehensively capture IEQ.

PoI2:

What is the current state-of-the-art in environmental monitoring?

The current state-of-the-art focuses heavily on the use of costly, reference-standard, measurement devices. While these devices were recognised to have implicit accuracy, the latter was often offset due to the lack of spatial density in the measurements, because of device placement. Typically, devices are placed in a central location to monitor a space from a single monitoring device. Consequently, the technology may not respond to localised changes in large spaces meaning that measurements may not reflect the environmental quality experienced by all occupants. Moreover, many of these devices each cost several thousands of pounds (often for a single IEQ outcome), meaning they have limited utility in e.g., non-commercial spaces.

There is a lack of IEQ monitoring from an individual perspective, suggesting there is a need for a paradigm shift that makes the individual the unit of analysis within the context of building monitoring. Accordingly, consideration of low-cost IEQ sensing approaches is warranted. Furthermore, wearables and cloud-based platforms could create holistic, personalised monitoring systems, which could aggregate data from multiple locations. Evidence suggests that such approaches from multiple sources could provide more scalability to achieve better (objective) and more spatially dense IEQ data.

PoI3:

What sensor technologies are used to capture IEQ?

In commercial buildings, state-of-the-art sensors are commonly integrated directly into HVAC systems. These types of sensors can provide accurate monitoring of e.g., air velocity, air quality, temperature and humidity using probes that are fitted into ducts, and air handling units of the HVAC system. These probes contain discrete analogue and digital circuitry that can capture changes in air quality through optical responses or changes in electrical resistance. If buildings lack these infrastructures, monitoring can be retrofit into the building using portable devices, or data loggers placed in situ. These devices can also be costly as they have comparable sensors to those used in HVAC systems. However, instead of probing a HVAC unit these devices have sensors integrated into the device. Unfortunately, many data loggers lack Wi-Fi connectivity, meaning that data must be manually downloaded.

Low-cost sensor technologies are increasing in popularity with multiple studies using both state-of-the-art and low-cost sensors in tandem. Low-cost sensors can measure a wide range of measurements and use similar technologies to those found in data loggers, but the components are often lack the resolution compared to state-of-the-art sensors, are made more cheaply, or use completely different technologies. For example, state-of-the-art CO_2 sensors measure optical responses using infrared beams to detect particle concentrations in a chamber. In contrast, metal oxide sensors are often used in cheaper sensors that may claim to measure CO_2 but measure an electro-chemical resistance that is equated to CO_2.

2.10 Further research

As expected from a ScR, this chapter found a range of approaches are used for conducting measurements of IEQ in buildings, covering a broad range of research domains. The broad ScR nature of this review meant that new research areas were identified that warranted further exploration. This chapter identified a need for scalable monitoring solutions that exploit emergent low-cost technologies for longitudinal periods, but it also highlights the potential of WHTs in this space. This chapter identified state-of-the-art and low-cost sensor technologies that are used to capture IEQ. However, to begin answering **PoI4** *(What are the optimal approaches to aggregate data from numerous devices and settings, including settings without existing monitoring infrastructures?)* further examination of the technologies identified is warranted. Accordingly, the next chapter (Chapter 3) will conduct a deeper exploration of low-cost technologies and how they can be used to monitor the health and wellbeing of individuals. This will be followed by an outline of key concepts, terminologies and frameworks used in WHT research and sensor technology research in general (Chapter 4). Both chapters will investigate how researchers and practitioners could utilise those technologies to communicate the computer science skills/workflows, which are required to interface with them.

Chapter 3 A narrative review of low-cost sensor technologies for environmental monitoring.

This chapter is an adapted from previously published work to fit the context of this thesis. The article: **Towards remote healthcare monitoring using accessible IoT technology: state-of-the-art, insights and experimental design**, was published in the **BioMedical Engineering Online** and was made available online on **30 October 2020** via:
https://doi.org/10.1186/s12938-020-00825-9

3.1 Introduction

Chapter 2 addressed **PoI1 – 3** and identified a need for low-cost solutions for monitoring environmental quality and the effects changes in IEQ have on individuals. This warranted further exploration into both WHTs and emergent low-cost sensor technologies. However, while many studies (identified in Chapter 2) presented the *"Design and development of"* low-cost monitoring approaches, few studies present pragmatic instructions to future researchers on how to develop similar sensors. Studies often reported the development stage at a high level with little more than a Bill of Materials (BOM). This often resulted in studies that provided limited information on the underlying technologies, or any discussion on how these sensors were assembled, programmed, or configured to perform the research.

The aim of this chapter is to begin to address **PoI4**:

4. What are the optimal approaches to aggregate data from numerous devices and settings, including settings without existing monitoring infrastructures?

Here, this chapter will provide a narrative review while surveying the current state-of-the-art of accessible IoT sensor technologies. This chapter specifically examines low-cost technologies and investigates their use within the context IEQ monitoring. An overview of current low-cost devices and technical specifications is presented to outline the possibilities, workflows and limitations presented by these technologies within healthcare applications. It is hoped that this chapter will build upon the findings of Chapter 2 to curate and review a solid body of multidisciplinary literature that will inform the technical developments and studies used in this thesis.

3.2 Low-cost Sensor Technology

Sensors are a prevalent driver of IoT technology, and they serve a multitude of purposes, from measuring people or places to systems or things. Sensors can be used to measure a range of outcomes from air quality or motoric activity, the latter which can help identify symptoms of underlying medical conditions e.g., Parkinson's disease (PD) [105]. Those type of sensors have taken a variety of form factors, from environmental sensors that use printed conductive plastics that can accurately detect the concentration of Carbon Dioxide (CO_2) in the air [106] to smart clothes that integrate tri-axial accelerometers directly into garments [107]. Key to these developments is the increasing technological advancements in microelectromechanical systems (MEMS) [108].

3.2.1 Initial prototyping tools: MEMS sensors and bench testing

MEMS use micro-engineering to integrate circuits and microscopic mechanical components into silicone microchips [109]. In doing so, it is possible to create micro-scale sensors with a range of sensing capabilities. Table 7 highlights the versatility and potential for MEMS technology within healthcare research. Whilst some research focuses on the use of MEMS sensors for specific healthcare applications, researchers are exploiting these technologies to create accessible sensor-fusion eHealth monitoring systems. For example, studies [110]–[114] previously combined a range of low-cost sensors to create monitoring systems that were able to remotely measure a variety of health conditions. Alternatively, Rienzo *et al.* [115] adopted a different sensor-fusion approach to combine three sensors (Electrocardiogram (ECG), Photoplethysmogram (PPG) and Seismocardiogram (SCG)) to simultaneously measure heart rate from 12 sensor nodes (each containing 3 sensors) that could be placed on different anatomical locations. In doing so, they were able to take 36 unique and individualised, high-frequency measurements of heart rate.

One of the most prominent resources available for rapid prototyping electronic circuits are solderless breadboards, which is a device made of interconnected rows and columns designed to temporarily connect circuits. Typically, there are four rows of sockets on a breadboard, which are connected horizontally and are used for supplying power. The remaining sockets are connected vertically and are used for connecting components. The sockets are designed so that components and wires slot in, without needing to solder a permanent connection. Solderless breadboards are a mature approach for prototyping, so component manufacturers typically conform to the width and spacing of sockets when designing hardware. Therefore, by convention, many electronic components are standardised to have a pin spacing (known as pitch) of 2.54mm [116]. This often makes MEMS sensors alone unsuitable for prototyping as they have a much smaller pitch, which vary from sensor to sensor. Sensors (e.g. Table 7) are often integrated onto 'breakout boards', which are small Printed Circuited Boards (PCBs) useful for prototyping and facilitate access to the pins on a microchip [117] by conforming to the 2.54mm convention, Figure 4. Many breakout boards can be used with little to no knowledge about electronic engineering. This is because much of the additional circuitry required to operate a MEMS chip is provided on the breakout board (Figure 4), often exposing only inputs, outputs and voltage control pins. This is the reason why the number of pins on the MEMS component differs from the number of pins on the breakout board.

Table 7 - Examples of MEMS sensor use for healthcare

Author	Year	Healthcare Application	Sensor ID	Sensor Type
Alberto *et al.*[118]	2020	Heart Rate	MAX30003[§]	Electrocardiogram (ECG)
Bakar *et al.* [110]	2020	Body Temperature Heart Rate	MAX30205[§] SEN11574[¶]	Temperature Electrocardiogram (ECG)
Rienzo *et al.* [115]	2020	Heart Rate Pulse	MAX30003[§] MAX30101[§] LSM6DSM[††]	Electrocardiogram (ECG) Photoplethysmogram (PPG) Seismocardiogram (SCG)[†]
Al-Naggar *et al.* [111]	2019	Heart Rate Pulse Body Temperature	MAX30003[§] AFE4490[‡‡] MAX30205[§]	Electrocardiogram (ECG) Pulse Oximeter Temperature
Anisimov *et al.* [119]	2019	Heart Rate	ADS1292R[‡‡] ADAS1000[‡‡] MAX30003[§] AD8232[§§]	Electrocardiogram (ECG)
Portaankorva, A. [112]	2018	Heart Rate Activity Monitoring	MAX30003[§] LASM6DSL[††] LIS3MDL[††]	Electrocardiogram (ECG) Accelerometer / Gyroscope Magnetometer
Yudhana *et al.* [120]	2018	Sign Language Detection	MPU6050[¶¶]	Accelerometer / Gyroscope
Anik *et al.* [121]	2017	Activity Recognition	MPU6050[¶¶]	Accelerometer / Gyroscope
Dawson S. [122]	2017	Medical Implant Security	ADXL362[§§]	Accelerometer / Gyroscope
Fitriani *et al.* [123]	2017	Activity Recognition	MPU6050[¶¶]	Accelerometer / Gyroscope
Kardos *et al.* [124]	2017	Gait Analysis	MPU6050[¶¶]	Accelerometer / Gyroscope
Mohanraj and Keshore [113]	2017	Body Temperature Pulse Heart Rate Emotion Detection	MAX30205[§] SEN11574[¶] AD8232[§§] 101020052[†††]	Temperature Photoplethysmogram (PPG) Electrocardiogram (ECG) Galvanic Skin Response
Mota *et al.* [125]	2017	Gait Analysis	MPU6050[¶¶]	Accelerometer / Gyroscope
Shaji *et al.* [114]	2017	Body Temperature Blood Pressure Pulse Heart Rate Fall Detection	MAX30205[§] HoneyWell 26PC[‡] SEN11574[¶] AD8232[§§] ADXL362[§§]	Temperature Pressure Photoplethysmogram (PPG) Electrocardiogram (ECG) Galvanic Skin Response
Al-Dahan *et al.* [126]	2016	Fall Detection	MPU6050[¶¶]	Accelerometer / Gyroscope
Kim *et al.* [127]	2015	Medical Implant Security	ADXL362[§§]	Accelerometer / Gyroscope
Lei *et al.* [128]	2015	Fall Detection	MPU6050[¶¶]	Accelerometer / Gyroscope
Wang *et al.* [129]	2015	Gait Analysis	MPU6050[¶¶]	Accelerometer / Gyroscope

[†]Seismocardiograph measurements were conducted using a MEMS-based accelerometer / gyroscope.; [‡]HoneyWell have a range of 26PC sensors, but the authors have not declared the specific sensor used in their study.; [§]Maxin Integrated Products; [¶]SparkFun; [††]STMicroelectronics; [‡‡]Texas Instruments; [§§]Analog Devices; [¶¶]TDK InvenSense; [†††]Seeed Studio

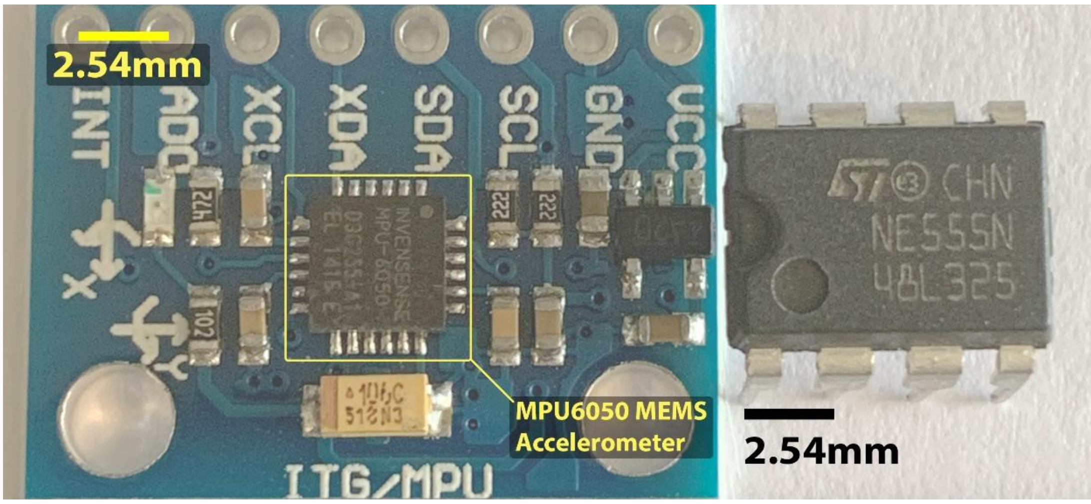

Figure 4 - Scale of MEMS sensor breakout board, compared to a 555 Timer chip with 2.54mm pitch.

3.2.1.1 Ensuring Fit-for-purpose Monitoring

Sensor use within healthcare research is becoming more prevalent, but it has often been reactive rather than proactive as innovation in this field can be quite fractious [130], [131]. With continued uptake of emergent technologies, researchers and practitioners must perform robust and vigorous bench testing (e.g., via tools outlined in section 2.1) to ensure new sensor-based technologies are valid and fit-for-purpose [130]. There is no absolute standard regarding sensor selection, as choosing an appropriate sensor will depend on what the researcher needs to measure and the subsequent digital endpoint(s) that is/are sensitive to the pathology in question [132]. Once fit-for-purpose sensors have been selected, appropriate and equally fit-for-purpose processing units (i.e., what the sensors are integrated into) must also be selected, to send control signals to the sensors as well as read and process sensor data.

3.3 Communication and control

There are a variety of ways to communicate and control sensor technology, which can vary depending on the stage of production, requirements of the hardware or accessibility.

3.3.1 FPGA/ASIC

For applications that require a great deal of power efficiency, whilst executing control algorithms in parallel and at high speeds, an Application-Specific Integrated Circuit (ASIC) may be required [133]. ASICs are microchips that contain an integrated circuit that is designed for a single application and cannot be reprogrammed [134]. This makes them suited to production level devices that do not need to change throughout the device’s lifecycle. Alternatively, Field Programmable Gate Arrays (FPGA) are reprogrammable. FPGAs are similar to ASICs as they contain integrated hardware circuits and once programmed can perform any logical function [135]. However, FPGA architecture differs from an ASIC and is comprised of an array of inputs and outputs (I/Os), logic blocks, interconnects, and connection lanes. These interconnects can be programmed so that the connection lanes bridge a connection between I/Os and a series of logic blocks to form a circuit of components that are suited to a specific application [136].

Since FPGAs and ASICs require the configuration of hardware circuits, they have a steep learning curve and may lack general accessibility to those without circuit design experience. However, it is also possible to interface with sensors using a programmable Central Processing Unit (CPU), which is used for controlling hardware and software [137]. Within IoT applications, CPUs are typically integrated into a Microprocessor Unit (MPU) or into a Microcontroller Unit (MCU) which combines CPU with memory. That enables CPUs to execute processes whilst being able to read and write data during an execution [138]. The key distinction between MPUs and MCUs, is the latter combines the CPU and memory onto a single microchip making it act as a single-chip-computer, capable of executing programmed instructions [139].

3.3.2 CPU

In contrast to FPGAs or ASICs, CPUs process algorithms in series, meaning they are not capable of running concurrent tasks. This can be overcome by using multicore CPUs, which combine multiple CPUs cores into a single processing unit, where each core can concurrently execute commands in series [140]. Another key distinction between CPUs and FPGAs is that whilst both architectures can be programmed, the program used in an FPGA is used to define how the hardware is configured, whereas the CPU executes the code as a series of instructions [136]. Since processing on an FPGA is done using hardware this means they are capable of handling analogue or digital signals, whereas a CPU is only capable of processing digital information. While this may seem like a major limitation for healthcare applications, one of the benefits of MCUs is that they typically contain a bus of General-Purpose Input/Output (GPIO) pins, which allow the device to send or receive both analogue and digital information from peripheral devices such as sensors [141].

Since the underlying CPU is capable of processing digital information only, analogue signals must first be converted to a digital signal or vice-versa. This is done using either analogue-to-digital convertors (ADC) or digital-to-analogue convertors (DAC) depending on the direction of the signal. When considering MCUs for healthcare applications and analogue signal measurement, it is important to consider the performance of the ADC to ensure that the device has sufficient resolution to be fit-for-purpose. This largely comprises of a trade-off between the sample rate, measured in samples per second (sps) and the bit resolution of the ADC, which refers to the number of discrete digital values an analogue signal can be mapped to. The more the bit resolution of the converted signal is lowered, the more the degradation of information is increased. Moreover, as the sample rate is increased, the ADC needs to convert a greater amount of information, which further reduces the bit resolution of the conversion [142]. Therefore, if researchers and practitioners intend to take measurements from analogue sensors, at a high sample rate, it is important that they choose an MCU with an ADC that has a high bit resolution when operating at the desired sample rate. This is to ensure the quality of the digital signal that is converted from the analogue stream is of a high standard for accurate data capture and robust patient assessment

3.3.3 MCU

A key benefit of MCUs is their low-cost and accessibility, largely driven by open-source based Arduinos – a range of inexpensive MCUs that are typically built onto development boards for rapid prototyping [143]. In software development, open-source code is typically distributed with a license that enables other developers to view, modify and share derivative works legally [144]. In much the same way, open-source hardware licenses allow the technology to be modified and distributed legally. This means that manufacturers and developers are free to clone, build, enhance and distribute hardware that builds upon the original infrastructure. Since derivative boards are based on the Arduino architecture, the way in which these microcontrollers are programmed has become standardised. The widespread adoption of these boards has not only incited rapid advancements in the capability of Arduinos, but it has also drastically reduced the costs of associated components.

Since their conception, Arduinos have taken a variety of forms and purposes. These include controllers for smart clothes that use inductive thread to control sensors to compact networked boards that are designed to interface with the IoT. For a full list of options and specifications, readers are directed to Arduinos product range[1], which outline the technical specifications of each board and categorises the boards according to accessibility. Additionally, Nayyar and Puri [145] present a review of Arduino hardware, outlining the technical intricacies of each board. However, the Arduino product range is continuously evolving and many of the boards in that review have subsequently been discontinued, as is the nature of disruptive technology [146]. Whilst the details presented in Arduino's product range provide detailed technical specifications, they lack aggregated information on the ADC/DAC capabilities of each device. To address this gap, Table 8 is provided to further

[1] Arduino product range https://www.arduino.cc/en/main/products

guide researchers and practitioners when choosing boards to suit the needs of their research projects.

Table 8 - Arduino's product range, highlighting architectures and ADC/DAC capabilities

Board	Price †	Processor	Digital / PWM‡	ADC Bit Resolution	ADC CHLs	ADC Sample Rate§	DAC Bit Resolution	DAC CHLs
Entry Level								
UNO R3	$ 23	ATmega328P (8-bit)	14 / 6	10-bit	6	15 ksps	-	0
Nano	$ 21	ATmega328P (8-bit)	22 / 6	10-bit	8	15 ksps	-	0
Leonardo	$ 21	ATmega32U4 (8-bit)	20 / 7	10-bit	12	15 ksps	-	0
Micro	$ 21	ATmega32U4 (8-bit)	20 / 7	10-bit	12	15 ksps	-	0
Nano Every	$ 11	ATMega4809 (8-bit)	22 / 5	10-bit	8	115 ksps	-	0
Enhanced								
MKR Zero	$ 26	SAMD21 (32-bit)	22 / 12	8/10/12-bit	7	350 ksps	10-Bit	1
Zero	$ 43	SAMD21 (32-bit)	20 / 10	12-bit	6	350 ksps	10-Bit	1
Due	$ 41	AT91SAM3X8E (32-bit)	54 / 12	12-bit	16	1000 ksps	12-bit	2
Mega 2560 Rev3	$ 41	ATmega2560 (8-bit)	54 / 15	10-bit	16	15 ksps	-	0
IoT								
Nano 33 IOT	$ 19	SAMD21 (32-bit)	14 / 11	8/10/12-bit	8	350 ksps	10-Bit	1
Nano 33 BLE	$ 21	nRF52840 (32-bit)	14 / 14	12-bit	8	200 ksps	-	0
Nano 33 BLE Sense	$ 32	nRF52840 (32-bit)	14 / 14	12-bit	8	200 ksps	-	0
MKR WAN 1300	$ 41	SAMD21 (32-bit)	8 / 12	8/10/12-bit	7	350 ksps	10-Bit	1
MKR GSM 1400	$ 69	SAMD21 (32-bit)	8 / 13	8/10/12-bit	7	350 ksps	10-Bit	1
MKR WiFi 1010	$ 33	SAMD21 (32-bit)	8 / 13	8/10/12-bit	7	350 ksps	10-Bit	1
MKR NB 1500	$ 77	SAMD21 (32-bit)	8 / 13	8/10/12-bit	7	350 ksps	10-Bit	1
MKR Vidor 4000¶	$ 72	SAMD21 (32-bit)	8 / 13	8/10/12-bit	7	350 ksps	10-Bit	1
MKR 1000	$ 37	SAMD21 (32-bit)	8 / 12	8/10/12-bit	7	350 ksps	10-Bit	1
UNO WiFi Rev2	$ 45	ATMega4809 (8-bit)	14 / 5	10-bit	6	115 ksps	-	0

All information has been sourced from Arduino's product range and the subsequent datasheets provided there.; † prices (as recorded on 10 July 2020) are rounded up to the nearest USD (ex. VAT); ‡ Pulse Width Modulation(PWM) is an emulated analogue signal created with high frequency digital pulses.; § ADC sample rates specified are in kilo-samples per second (ksps) and are achieved at the highest bit resolution of the ADC, lower bit resolutions can achieve sample rates greater than those specified above; ¶ MKR Vidor 4000 has an on-board Intel® Cyclone® 10CL016 FPGA to supplement the SAMD21 MCU.

Boards in Table 8 convert analogue signals to digital with at least a 10-bit resolution. Moreover, the sample rates of modern Arduinos enable them to be applicable for a range of healthcare applications as they exceed requirements for measuring high frequency analogue signals, e.g., Electrocardiographs [147]. As the technology continues to disrupt, modern Arduinos push the boundaries with new processors and higher resolution ADC capabilities. Furthermore, IoT is an increasing driver of technological development and Arduino's own IoT range now come equipped with e.g., a range of wireless capabilities to suit a variety of remote measurement projects via Cloud services or MCU boards with built-in FPGA for additional programmable functionality. However, these come at an increased cost, inhibiting accessibility.

Derivative boards and inexpensive clone boards are an alternative, providing equal functionality much lower cost. For example, an official Arduino Uno R3 costs approx. $23 but a clone built to equal sizes and specifications is as little as $3.00. Although not supported by Arduino, clone boards will function the same as Arduino counterparts and will likely be compatible with Arduino software, as the latter supports third party manufacturers. However, researchers and practitioners using clone boards should be aware that they would be unlikely to receive official support from Arduino for any clones.

The open-source nature of Arduino products means that derivative boards can also be created. Instead of aiming to create clones that offer equal functionality, derivative boards aim to extend the functionality of Arduinos by on-boarding features such as LCD screens, wireless communication and more powerful processors, which can be useful for providing real-time feedback from sensor readings. One example which is gaining popularity [148] is the ESP32[2]. The latter cannot be directly compared to an Arduino development board as it is regarded as a System on Chip (SoC), meaning that it is an entire system on a single microchip. These chips are considered a market leader as they integrate WiFi, Bluetooth Low Energy (BLE), dual-core processing and sensors onto a single chip [149]. Moreover, these chips are now being integrated onto a wide range of development boards that offer similar accessibility as Arduino development boards but with increased functionality and lower costs. One reason why SoCs (and the development boards built upon them) have been so successful within IoT development is that the entire chip can be reconfigured at run time to operate at extremely low power, making them suitable controllers for battery powered IoT devices [150]. Furthermore, the ESP32 chip has 18 multi-resolution ADC channels capable of running 200ksps at 12-bit resolution and two 8-bit DAC channels, which makes the chip comparable to the Arduino Due – one of Arduino's largest form-factor development boards. Of note, while the ESP32 has an 18 channel ADC, two of those channels are occupied by integrated temperature and hall-effect sensors that detect magnetic fields and the temperature of its chip [150]. This means that for applications that do not make use of these sensors, the ESP32 has only 16 usable ADC channels, though this is comparable to the Arduino's Due and Mega 2560 boards.

Unlike FPGAs and ASICs, Arduinos and their derivative microcontrollers were designed to be accessible to beginners yet flexible to accommodate skilled developers [143]. This makes them ideal for those that may not possess the prerequisite knowledge of an e.g. electrical

[2] Espressif Systems https://www.espressif.com/

engineer but wish to gain insights into IoT hardware development or become more knowledgeable about possibilities and limitations of the hardware.

3.3.4 Software

The scale of data across the healthcare sector has been increasing and is expected to continue increasing exponentially as healthcare professionals adopt IoT solutions [151]. As more information is stored into healthcare models, challenges around transmission and storage of those models increases in tandem. IoT adds further complexity to the issues of data scale as devices typically send a telemetry stream, which is continuous data ranging in frequencies from seconds to weeks. Therefore, frequency of data transmission has a direct impact on the level of storage and the type of system that is needed to manage the stream.

Devices like Arduino processors must be programmed with a specific set of commands telling it which pins to read and write to and what to do with the data. Hardware manufacturers (e.g., Arduino, Adafruit, SparkFun) provide searchable databases of open-source code libraries (often accompanied with setup tutorials) that can be accessed from a web browser or their proprietary software[3]. Thus, researchers and practitioners can be more informed about the steps involved and understand the possibilities and challenges the hardware presents through the support of those tutorials and documentation.

3.4 Cloud Connectivity

IoT workflows extend beyond the development of sensor technology by developing software that collects, stores and analyses data streams. Open-source IoT software platforms are also becoming a driving force of accessibility and innovation. These platforms are typically centred on providing a web-based dashboard and a database to collect and display data from IoT devices. Researchers and practitioners should be aware there are more than 600 known IoT platforms [152] and, whilst the sector is largely dominated by large corporations such as Amazon, Google, and Microsoft [153], IoT cloud platforms are continuing to expand and fragment with niche platforms designed for specific use cases [152]. These platforms are typically centred on providing a web-based dashboard and a database to collect and display data. Many of these platforms are complex and feature-rich, with a range of integration protocols that can directly interface with MCUs [94], [154], [155]. However, many of these cloud platforms operate on a quota or a pay-as-you-go model, where users pay for services, storage, or bandwidth they consume [156]. For IoT applications such as smart homes this

[3] Readers are directed to the Arduino Code Library List [200], which is an automatically generated database of libraries. This list contains approx. 3000 libraries, which provide detailed license, author and version information as well as links to download source code.

can be an affordable option as the frequency of events *(when an IoT device uses some of the quota)* can be sporadic or low frequency, e.g., when a light turns on or off. In healthcare research, the frequency of data transmission may often need to be much greater, in the region of hundreds or thousands of samples per second. This currently creates multiple technical obstacles that make cloud-based remote monitoring of patients challenging.

3.4.1 Rate limiting and transactional cost

When transmitting high frequency sensor data to the cloud, a large volume of data can be accumulated in a short space of time. This will require large amounts of cloud storage and may require a great deal of bandwidth. Before adopting a cloud solution, researchers and practitioners must be aware of how a user is charged for data, with regards to both storage and bandwidth. Given the number of available cloud solutions, a complete breakdown of costs involved with each service is beyond the scope of this chapter. Instead, indicative costs associated with different subscription models from the key providers, Microsoft Azure, Amazon' Web Services (AWS) and Google's Cloud Platform (GCP) are presented [153].

To demonstrate the speed in which message quotas would be consumed using cloud platforms, I extracted several ten-second samples of raw tri-axial data from a low-cost commercial MEMS based wearable accelerometer (AX3, Axivity, Newcastle, UK) in CSV format with timestamp information included. The sample rate was set at 100sps (100Hz) and so each sample contained 1000 rows (100Hz × 10s) of values. The average file size of the CSV data were approximately 33 kilobytes (KB). This file size was then used to compare the pricing for the three major cloud IoT platforms.

3.4.1.1 Microsoft Azure IoT Hub

Microsoft Azure's IoT Hub has a range of pricing options and quotas (Table 9). Users of the service are billed monthly and charged according to the amount of messages/day. For device-to-cloud messaging, the maximum of a single message equals 256KB [157], meaning no single device can send more than that at any one time. However, that message size is far greater than the meter size for each tier, which is capped at a maximum of 4KB for paid tiers and 0.5KB for the free tier, Table 9. Therefore, while a single 256KB message can be sent from an IoT device to the cloud, this message is segmented into 0.5KB/4KB segments and charged accordingly. Thus, a 256KB message will expend 64 messages from the daily quota on paid tiers and 512 messages from the daily quota on the free tier. For high-frequency data this quota can be quickly consumed. Using the example set out in Section 3.4.1, a 67.1KB message would consume 9 messages from paid tier subscriptions and 66 messages from a free tier subscription. At that rate, to monitor tri-axial data, values at around 100sps (100Hz) for 24-hours, approximately 71,280 messages would be consumed on a paid tier subscription. This would mean either the S1 or the B1 tier would be applicable. However, the daily message quota on free tier subscription would be completely consumed in around 20 minutes.

Table 9 - Example of IoT hub pricing tiers

	Tier	Monthly cost	Messages / day	Meter size
Azure	Free Tier	$0	8000	0.5 KB
	Basic Tier 1 (B1)	$10	400,000	4 KB
	Basic Tier 2 (B2)	$50	6,000,000	4 KB
	Basic Tier 3 (B3)	$500	300,000,000	4 KB
	Standard Tier 1 (S1)	$25	400,000	4 KB
	Standard Tier 2 (S2)	$250	6,000,000	4 KB
	Standard Tier 3 (S3)	$2500	300,000,000	4 KB
	Monthly Messages	**Price†**	**Meter Size**	**Connection cost‡**
AWS	<1 billion	$1	5 KB	$0.08
	1 billion – 5 billion	$0.80	5 KB	$0.08
	More than 5 billion	$0.70	5 KB	$0.08
	Data usage	**Price / MB**	**Minimum charge**	
GCP	Up to 250 MB	$0.00	1024 bytes	
	250 MB to 250 GB	$0.0045	1024 bytes	
	250 GB to 5 TB	$0.0020	1024 bytes	
	5 TB and above	$0.00045	1024 bytes	

Data relating to tiers, pricing and message quotas was obtained from the pricing pages of Microsoft Azure [158], Amazon Web Services [159] and Google Cloud Platform [160] on 17 July 2020. †Per million messages. ‡Per million minutes.

3.4.1.2 Amazon Web Services (AWS)

Similar to Azure, AWS IoT Core service involves chunking large messages and charging according with a maximum message size of 128KB and a 5KB meter size. Yet, unlike Azure, AWS tiers decrease in price as more messages are transmitted. If 10 seconds of tri-axial accelerometer recordings creates 33KB of data, AWS would bill for 57,204 messages in 24-hours. This equates to 1,710,720 messages over a 30-day period, where each million messages will be billed at $0.80 – equalling $1.37 per month. Additionally, AWS also charge $0.08 per million minutes of connection, but for a single device the price change is negligible as a device connected continuously for 30 days would cost $0.003456.

3.4.1.3 Google Cloud Platform (GCP)

GCP adopts a different quota system to Azure and AWS, instead charging according to the total amount of data transmitted rather than the total number of messages (Table 9). Additionally, instead of charging in data segments according to a meter size, GCP adopt a minimum charge approach when billing for transactions. Consequently, GCP encourage users to send fewer large messages rather than many small messages (unlike Azure and AWS). If 33KB of tri-axial accelerometer data were sent from a device to GCP, instead of it being segmented and metered, prices would be calculated per megabyte (MB). In this instance, continuous data for 30 days would equate to 8.55 Gigabytes (GB) of data, costing

\$0.0045/MB. Therefore, total cost *(including first 250MB free)* for 30 days would be \$37.37. It is important to note that this cost only considers data being sent from the sensor, as GCP also have costs associated to the communication protocols used to send data. It is important for researchers and practitioners to understand which protocols are available on a chosen platform as they can significantly impact the cost of data transmission.

3.4.2 Communication protocols

Many cloud-platforms accept a range of communication protocols, with two of the most popular protocols used within IoT platforms are Hyper Text Transfer Protocol (HTTP) and Message Queuing Telemetry Transport (MQTT). HTTP is a mature protocol for requesting and received data over the internet [161]. Within IoT, devices can send data over HTTP by attaching the data (known as payload) to the HTTP request being sent to a server. When the server receives the request, it returns a response to indicate the success or failure of the request/response lifecycle [162]. However, each request requires authentication and once the request/response lifecycle is completed the connection to the server is then closed [163]. This uses a lot of bandwidth and creates overheads for IoT devices that need to send high-frequency data to the Cloud. Contrastingly, instead of using a request/response lifecycle, MQTT protocol uses a publish/subscribe approach, where data is published to a server (message broker) and made available for subscription [162]. For example, an IoT device can publish a sensor reading to the broker and an IoT application (subscribed to the broker) can receive that data. A key benefit of MQTT over HTTP, for IoT applications, is that a persistent connection can be made to a broker, which allows devices to send multiple data payloads with a single authentication [161].

The fundamental differences between HTTP and MQTT have a substantial impact on cost within GCP. This is because GCP charges for each connection. For MQTT, monthly costs depend on how long the connection from a device is kept active. For example, if each device refreshes the connection every fifteen minutes, 96 daily requests will be made to broker. Yet, whilst each request will be extremely small, GCP's minimum charge means that every request is charged at 1024 bytes (1KB), which equates to approx. 3MB/month. Alternatively, HTTP makes a request and response every time data is sent. If 33KB of data were transmitted every 10 seconds, 8640 messages would be sent daily. Since GCP would apply the minimum charge of 1KB to each response, the HTTP responses alone would use the entirety of the 250MB free quota. For this reason, in contrast to AWS and Azure, it would be important to send considerably larger amounts of data and to transmit less often when using GCP.

3.5 Serial Processing

Whichever cloud platform is adopted for an IoT solution, technological inadequacies of processing units can be a limitation when attempting to collect, store and transmit high-frequency data. As discussed previously, MCUs process data in series, meaning they execute each command one after another. Therefore, single core processors are unable to initiate the next command until the previous one is complete. On a single-core MCU, data transmission must therefore interrupt the data collection and the MCU will be unable to read sensor data until the data has been transmitted. This could also involve waiting for a response if transmitting over HTTP. Since it would be problematic to transmit every reading from a sensor running at a frequency of 100Hz (100sps), the MCU must read data from the sensor, perform analogue-to-digital conversion (if required), and store that data in memory. This whole process must also be executed within 10 milliseconds (ms) to maintain a sample rate of 100sps. When enough data has been collected in memory, the MCU must then send the data to the cloud. However, this instruction must also be executed within one of the 10ms windows allocated to data collection, otherwise the sample rate will drop. This problem could be mitigated by using multi-core MCUs such as the ESP32, or devices that combine FPGAs with MCUs such as the MKR Vidor 4000. These devices would allow an uninterrupted data stream to be collected and stored, while simultaneously transmitting the data to the Cloud.

3.6 Discussions

This chapter presented a narrative review and survey of current state-of-the-art for accessible and low-cost IoT sensor technology. In doing so, I presented pragmatic insights of current technologies and the technical specifications that could present opportunities or limitations to researchers and practitioners. One of the key benefits to these technologies is their low-cost, meaning it is feasible to create scalable sensor-fusion devices that incorporate a range of sensors for monitoring patients. Moreover, such sensor fusion devices could enable researchers and practitioners to augment wearable sensing devices with environmental sensors to provide more context to topical patient assessment outcomes e.g., gait , which would help them move their research beyond the laboratory and into free-living conditions.

Given the prevalence of IoT technologies, many research domains and industry practitioners will become more reliant on multi-disciplinary teams to break ground with these disruptive technologies. This thesis will aim to present technological developments in later chapters with a much greater detail than is typically seen across the literature to address the gap in knowledge surrounding the design and development of low-cost hardware. Given the multi-disciplinary nature of this thesis, it is important to explore the core concepts of IoT technologies to better inform the direction of thesis and to communicate these concepts with the reader, who may not be familiar with the computer science or electrical engineering fundamentals that underpin IoT developments.

3.6.1 IoT hardware

Advancements in MEMS technologies allow a range of sensing capabilities that can aid researchers and practitioners. However, many of these devices deal with high frequency analogue signals, which present a new set of challenges, which must be considered when specifying both the sensors and the processing units that will collect data from the sensor. For many healthcare applications, such as ECGs and electroencephalogram (EEGs), high-frequency sampling is a requirement [164]. For this reason, it may be necessary to exploit the technological capabilities of ASICs or FPGAs, which can capture multiple high-frequency analogue signals simultaneously. However, opensource microcontrollers such as Arduino have driven the industry to develop boards that are demonstrably capable within this field. Whilst microcontrollers were traditionally limited by being unable to execute tasks concurrently, multi-core microcontrollers are now becoming more prevalent. Moreover, whilst MCUs cannot process analogue signals directly, due to the limitations of the internal CPU, this chapter has demonstrated how advancements in ADC/DAC technologies are enabling MCUs to perform continually higher-resolution conversions of analogue signals at high frequencies. Yet, for these devices to be considered IoT devices there is a need to connect these devices to the internet. Networked MCU development boards are becoming more prevalent, boasting a range of wireless connection options that enable these devices to not only collect and process sensor data, but also transmit these data to the cloud IoT platforms.

3.6.2 Cloud computing

In many cases it may be suitable for researchers to use off-the-shelf, cloud-based applications within their researcher. Here, three of the major cloud providers were compared to provide more clarity around the cost of cloud hosting and provide a mechanism to compare and appraise cloud providers.

From the three major platforms AWS was found to be the cheapest platform overall, especially when using many devices. Contrastingly, GCP was found to be significantly more expensive. Nevertheless, the unique pricing model adopted by Google, means that the platform is better suited for transmitting large amounts of data infrequently, as opposed to Azure and AWS, which favour regular small amounts. Breaking down the transactional costs of cloud platforms in this way was beneficial as it provided a much clearly picture of the indicative costs that would be incurred using each platform. Pricing data across the platforms were found to be obfuscated and difficult to compare. This obfuscation was not just prevalent when comparing different service providers, but it was also difficult to compare the tiered services and scaling costs from single providers.

3.7 Addressing the PoI

Given the pervasiveness of IoT technologies, monitoring of occupant health and wellbeing will become more reliant on use of these disruptive technologies. For this reason, it was important that the core concepts of popular commercial IoT technologies were explored to better inform the technological capabilities and the challenges in this research area. Here, pragmatic insights of current technologies and the technical specifications were presented to highlight opportunities or limitations available to the development of this thesis.

One of the key benefits to emergent IoT technologies is their low-cost, meaning it is feasible to create scalable monitoring solutions that incorporate a range of sensors and can be deployed at an individual level due to their cost. However, the cost of hardware only forms one part of the equation in IoT infrastructures as a symbiosis of hardware and software is required. To create a low-cost IoT system in general, it is important to evaluate the on-going costs of data collection and explore scalable low-cost software solutions as well.

This chapter began to explore **PoI4** by examining the underlying technologies that are used to collect and process data that drive IoT hardware. It first outlined emergent IoT technologies that can be used to collect, store, aggregate, transmit and analyse data from low-cost sensors. It then explored the computational hardware infrastructures of IoT technologies, such as MCUs, CPUs, and FPGAs, which drive the processing and aggregation of data in sensing systems. These are key to understanding **PoI4** as the digital processing infrastructures can limit the type of data that can be captured or can reduce the resolution and quality of the data output. This chapter also explored how IoT devices can integrate with cloud technologies and explored the transactional costs of transmitting data to the cloud.

3.8 Further Research

While this chapter provided useful findings surrounding the use of low-cost hardware for data aggregation. Further research is required into the challenges surrounding data aggregation from proprietary devices. The following chapter (4) will expand upon **PoI4** by reviewing the software infrastructures that are required to interface with WHTs and proprietary monitoring devices.

Chapter 4 Frameworks for wearable integration

This section is adapted from previously published work to fit the context of this thesis. The chapter: **Frameworks for wearable integration - Digital Networks and Beyond**, was published by **Elsevier** on 9 July 2021 for the book: **Digital Health: Exploring Use and Integration of Wearables**. Permission to reuse up to 8x 500-word excerpts of the published work was obtained from Elsevier on 28 September 2022 – License Number: 5397780628489. The declaration of authorisation is included in *(Appendix D)*

*The DOI chapter is **https://doi.org/10.1016/B978-0-12-818914-6.00003-X** and it is available through Academic Press.*

This original chapter published in Digital Health: Exploring Use and Integrations of Wearables was co-authored by an academic peer. However, none of the text they submitted for the original publication has been included in this thesis.

4.1 Introduction

Chapter 3 begun to address **PoI4** by exploring low-cost technologies suitable for IEQ monitoring and presented foundational work for a series of case studies exploring the use of low-cost technologies and WHTs. However, many WHTs are often tightly coupled with proprietary applications that interface with e.g., smartphones, creating a symbiosis between the proprietary software and hardware. For daily monitoring, this can present challenges due to the limited integrations offered by proprietary applications. As Chapter 2 highlighted that there is a need for low-cost approaches, these limitations could prevent or complicate research that use bespoke, low-cost platforms for data collection.

This chapter will further my investigation to **PoI4** by exploring the frameworks, and cloud computing technologies that sensor technologies can interact with when developing bespoke monitoring solutions or when aggregating data from proprietary hardware. The primary purpose of this chapter is to present key concepts and terminologies, to inform the rationale in the selection processes adopted throughout the remainder of this thesis and to understand the technical, ethical and security implications of interacting with proprietary hardware. This chapter will present and explore the processes involved with communicating and collecting data from WHTs; as well as exploring the limitations that could be faced when attempting to do so. In presenting these principles, this chapter will outline key considerations that must be considered within this thesis when conducting research using WHTs, such as, privacy, security and ethical compliance.

4.2 Proprietary WHT Systems

Digital Health Snippet 1: 297 Words

Many WHT brands provide proprietary mobile applications and web-based dashboards on which to interface with their devices. These proprietary applications often provide a tailored experience that offer a seamless and unique symbiosis between the hardware and the software platform. For most consumers, this experience will provide the user with a plethora of data and information to suit their needs. However, research needs differ greatly from consumer needs, and this provides a unique set of challenges, which need to be carefully considered. Many of the data provided by proprietary platforms are used to deliver real-time information to wearers. However, when conducting research there is a need to collect these data longitudinally for analysis. Whilst it is possible to export data from devices such as Fitbit/Garmin/Apple fitness trackers, the user often has little control over this process. For example, Fitbit provide real-time access to data both through a mobile application and a web-based platform. The platform provides a mechanism for users to export data for a given time-period into various formats, but the user has little control over the level of granularity within the export process. The data exported is not the raw *(sample level)* data that were collected. Instead, it is aggregated and summarised in a way that is controlled by Fitbit.

These devices do collect and store granular, real-time data for various health parameters such as step count, heart rate and sleep data. However, to access these data there is a need to develop bespoke applications that either interface with an Application Program Interface (API) for the proprietary system or build upon Software Development Kits (SDKs) provided by the manufacturer. Here these terminologies will be explained and discussed in the context of this thesis.

4.3 Application Programming Interface (API)

Digital Health Snippet 2: 213 Words

In its simplest form, an API is a way to interface with functions in a program. When developing software, functions can have varying degrees of accessibility *(in the sense of access, not mobility).* Functions can be given *private*, *protected,* or *internal* access modifiers, meaning they can only be called by the program itself. Alternatively, they can be made *public,* meaning they can be accessed from anywhere and interfaced with from outside the program. This is like having a document on a laptop vs publishing it on the web. This process of interfacing with external functions is widely used and there are many different types of APIs that are available to developers. However, over recent years, one form of API called a Representational State Transfer (REST) API has gained a lot of popularity. Without delving too deep into the intricacies of software development, A REST API is built upon the REST framework [165] and provides a set of conventionalised HTTP endpoints, which are interfaces to a program that are accessed using a URL hyperlink. However, instead of providing direct access to functions within an application, REST APIs provide simplified access to predefined operations within a program, that can be accessed without knowledge of the programming language used to create the operations [166].

4.4 Software Development Kit (SDK)

Digital Health Snippet 3: 101 Words

An SDK is another way to interface with software applications. Like an API, an SDK can provide access to a series of functions within a target application. However, SDKs can contain several APIs within them and are targeted specifically towards other software developers. Therefore, an SDK typically contains pre-packaged code, APIs and documentation that are provided as an accelerator for software developers wishing to extend the functionality provided in the target application [167]. As such, to interact with an SDK there are several prerequisites, which often makes them less accessible than a standalone API.

4.5 Interacting with Proprietary Systems

Digital Health Snippet 4: 425 Words

APIs and SDKs can considerably improve interoperability of WHT devices. However, these are not universal and are developed by the WHT manufacturers to interact solely with their proprietary software. As such, there is little standardisation about how data is accessed, retrieved, or presented to the user. For example, despite the conventions laid out in the REST framework [165], there are many ways to interact with a REST API. Some APIs require authentication, others do not; the level of authentication required differs from one API to the next and the formats in which data are returned to the user are unique to each application.

Additional to considering proprietary APIs and SDKs, there are also third-party APIs such as Google Fit and Apple HealthKit. These APIs are primarily designed for Android and iOS based WHT devices; however, Apple and Google have extended their services to include support for third party WHTs. Since many WHTs are designed to interact with smartphones, it is possible to sync health data to these APIs; though, this adds a further layer of complexity. For example, Fitbit simply do not support this process and require users to interface using their proprietary API for data retrieval, as they can better control the security and ethics. Contrastingly, data retrieval from Xiaomi devices is only achievable through this process; though, it first depends on the WHT syncing with its proprietary application on the smartphone first, which then syncs with the Apple or Google API. This process can be discontinuous and range from near-real-time to several days [168] and on iOS devices this can be even more complicated because while data can be posted to the Apple HealthKit API, there are no APIs to retrieve data. To retrieve data from the HealthKit API, developers must implement the HealthKit SDK and develop a native iOS application. This application can then sync with HealthKit and retrieve any data that has already been synced by the WHT. Developers can mitigate this process by installing Google Fit for iOS and use that app to sync HealthKit data, which can be retrieved through Google Fit's retrieval API. However, this is also not supported by Fitbit and further highlights the complexities of data retrieval from WHTs and proprietary platforms.

It is important to note, that whilst the challenges of WHT interoperability can be difficult or frustrating, these levels of complexity are not arbitrary. Instead, they are purposefully implemented to protect the health data captured by WHTs and mitigate data security and ethical implications.

4.5.1 Security and Ethics

Digital Health Snippet 5: 380 Words

Data security and ethics is an important consideration. However, the General Data Protection Regulation (GDPR) put a microscope over all sectors that work or operate in the European Union. The changes around how information is collected and stored has forced all sectors to audit and improve their data collection practices.

One of the key changes in GDPR is the definition of personal data, which has a profound impact on the way in which data is collected from WHT devices. Many fitness trackers collect an array of information about the wearer, including heart rate, weight and exercise data. These data, and their combination, can be used to identify a 'natural' living person *(a living individual and not a business or professional entity)*. This is particularly relevant for intraday data. For example, on its own, heart rate (in beats per minute (BPM)) does not constitute as personal data. However, by measuring and analysing real-time heart-rate data, it is possible to calculate Heart Rate Variability (HRV), which has unique properties and can be used as a biometric fingerprint to identify a person [169]. These data are often collected by WHT devices and stored alongside names, email addresses and even global positioning data. Therefore, WHT systems become warehouses of personal data, which must be protected to maintain GDPR compliancy.

The levels of control imposed by WHT proprietary systems can be challenging as applications must always validate that a WHT wearer has authorised access to the platform and allowed access to the required data. This is akin to how validation works in online signup forms; when a user does not enter the required information, they are forced to resubmit the form with those fields completed. Advancements in cloud computing and IoT technologies, has led to a paradigm of accessible IoT platforms that are beginning to also focus on WHTs and may provide solutions that mitigate the need to develop bespoke applications and integrations with WHTs.

4.5.2 Cloud Services

WHTs can be integrated with web based IoT platforms and cloud services in several ways, depending on the technical capabilities of the research team [170]. Fundamentally there are three levels of service offered by cloud service providers, which include Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). These services offer a tiered approach to cloud computing but require significantly more development for each service layered (Figure 5).

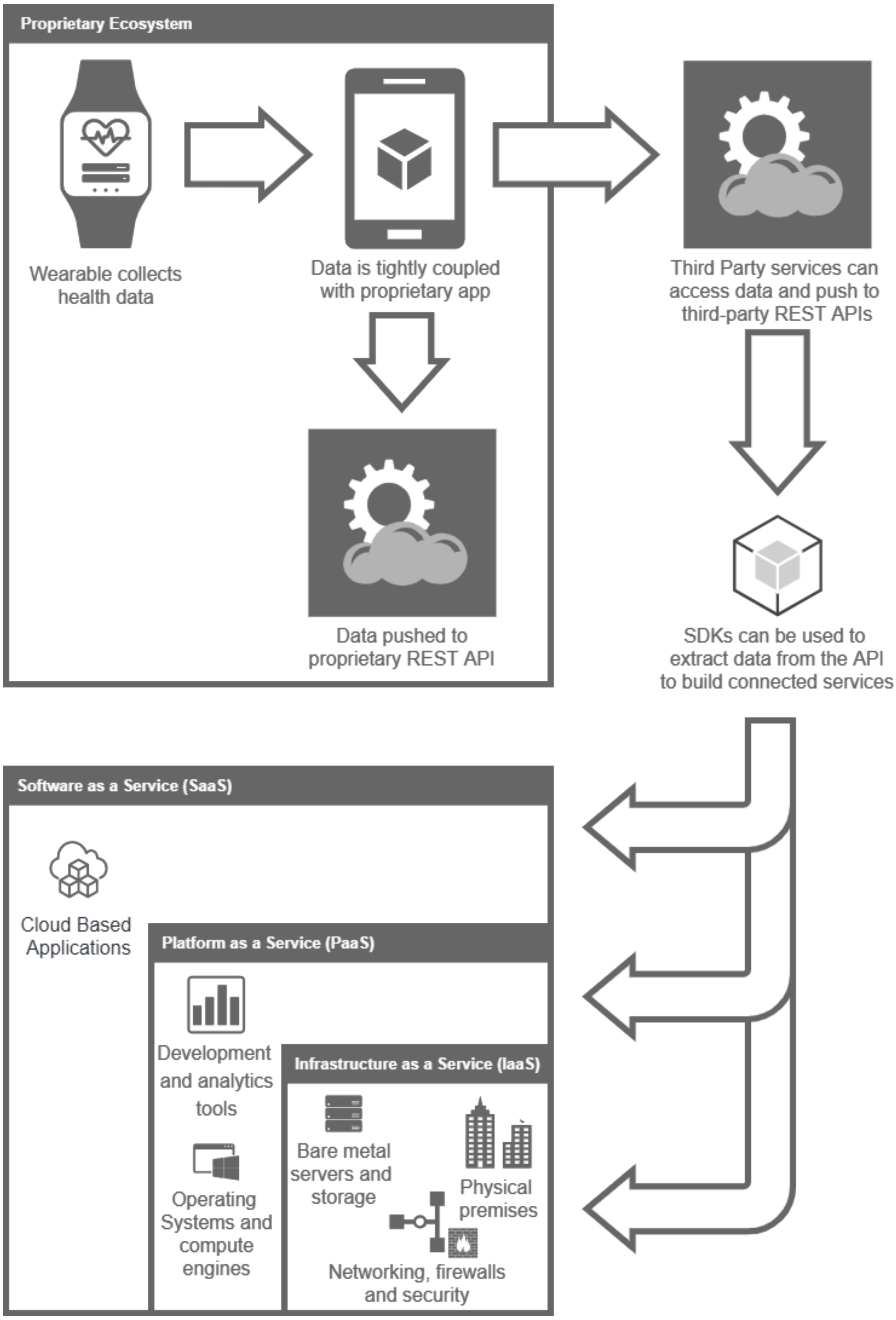

Figure 5 - Diagram of proprietary system interactions

4.5.2.1 Software as a Service (SaaS)

Digital Health Snippet 6: 229 Words

Software as a Service (SaaS) provides the highest level of abstraction in cloud computing. Many of the solutions in this domain are often software systems that build upon a cloud computing PaaS, to deliver a Software as a Service (SaaS) solution. The primary difference between a PaaS and a SaaS is that a PaaS offers a platform on which software can be developed, but a SaaS offers a software solution, which may or may not build upon a PaaS [171]. Since SaaS IoT systems offer a software solution, these often mitigate the need for software development, providing an out-of-the-box approach for interacting with wearables. IoT SaaS systems may be appraised based on prevalence within the literature. However, that approach is not applicable to IoT platforms. A recent review on environmental monitoring in buildings [172] found that whilst there was commonality around most aspects of the studies reviewed; including methodologies, IoT hardware, measurement factors and even building types, there was absolutely no commonality surrounding the software platforms used to capture data, due to the vast array of options and hardware agnosticism. However, within cloud computing there are various types of service layers that often need to be considered when identifying/developing IoT solutions.

4.5.2.2 Platform as a Service (PaaS)

Digital Health Snippet 7: 150 Words

Platform as a Service (PaaS) removes a layer of abstraction from the SaaS service and provides platforms that are designed and delivered as a service on which to build a SaaS or hosted application. Consequently, these apps are capable of catering to thousands of use cases, but only a small percentage of these use cases can be solved using the platforms out-of-the-box. Instead, software is built on top of these platforms using and paying for only the features required for a system. Cloud PaaS providers *(such as Amazon, Google & Microsoft)* are largely paralleled and offer an almost identical feature set and experience and they all offer both an IoT solution and the ability to communicate with external proprietary APIs [170]. As such, this method of interacting with wearables still involves communicating with APIs and building upon SDKs, which ultimately requires software development capabilities, the only difference here being that the developed solution will have the benefits of cloud computing.

4.5.2.3 Infrastructure as a Service (IaaS)

Infrastructure as a Service (IaaS) provides the lowest level of abstraction in cloud computing but provides the most bespoke functionality. This service layer provides physical services such as datacentres, physical networking, security systems and firewalls - requiring all other services to be developed. The primary advantage of IaaS is that developers can build either PaaS or SaaS offerings built upon bespoke IaaS infrastructures. However, due to the development complexity, it is unlikely that this approach would be suitable for smaller, individual research projects, but could suit larger research bids, which involve setting up frameworks and toolkits on which to develop smaller research projects.

4.6 Discussion and Conclusions

This chapter has presented an overview of concepts and terminology required when interacting with WHTs. This work was part of a wider contribution to healthcare research, but a selection of snippets was presented here that were deemed of value to the intended multi-disciplinary readership of this thesis. The aim here was to explore and present the underlying infrastructures that surround/govern the use of WHTs to inform the development of this thesis. Here, it was important to understand the technical, ethical and security implications of interacting with proprietary APIs and SDKS and how these can limit choices when considering IoT technologies and WHTs. For example, nuanced approaches for augmenting data from IoT devices and smartphones may require bespoke software that builds upon APIs, SDKs or a PaaS. This will mean that software development capability will become a prerequisite for this project. In this example, this would be unavoidable; however, for project teams that lack the technical capability, it may be possible to amend project goals so that they are less nuanced and able to utilise off-the-shelf SaaS solutions.

It was also valuable, for this stage of the thesis, to explore and understand the potential implications of collecting health related outcomes. As health data becomes more prevalent and the data becomes richer, advancements in algorithm developments can turn ordinary WHT outcomes into a means of personal identification. Therefore, when collecting continuous WHT data it is important to treat the data as if it could be used to identify a ‘natural’ person and ensure data management strategies are in place in case future developments in data science advance in such a way as to establish biomarkers in stored data.

4.7 Addressing the PoI

Chapter 3 explored the current state-of-the-art of accessible IoT sensor technologies, from which a series of case studies arose that utilise both IEQ technologies and WHTs. This chapter has been positioned as an extension of the previous chapter to present approaches for utilising WHTs and the challenges this involves, before exploring the case studies in detail. While the current WHT market is largely comprised of proprietary, off-the-shelf technologies, there are still many challenges surrounding the development, implementation, and management of these technologies in the wild. This chapter provided a deeper exploration of the frameworks and infrastructures of cloud solutions, exploring the types of cloud solutions available for data collection as well as the tools and frameworks available to support this.

PoI4:

What are the optimal approaches to aggregate data from numerous devices and settings, including settings without existing monitoring infrastructures?
Microcontrollers were identified as optimal approaches to aggregate data from a multitude of sensors in near real-time at a relatively low cost. They enable the intra-device aggregation of data from multiple sensors and can transmit data to cloud services for continued integration with other devices. It is recognised that other solutions exist, such as off-the-shelf monitoring solutions and FPGAs, but their increased cost makes them less scalable and thus, less fit-for-purpose within the context of the overarching industrial requirements of the Ph.D.

4.8 Further Research

Having now addressed **PoI4**, the following chapters (chapters 5-7) will explore **PoI5** from multiple angles to comprehensively understand whether low-cost sensors and IoT technologies can be used to gather data with precision, consistency, and reliability across different building environments. Chapters 5 & 6 will first present two case studies that arose from chapters 3 & 4, which will explore the use of IEQ monitoring solutions and WHTs using both PaaS and SaaS solutions for data aggregation. In doing so, they will present challenges and lessons-learned that have informed the direction of this thesis. Chapter 7 will then conduct a comprehensive study to demonstrate the development, testing and validation of a bespoke multimodal monitoring device developed using low-cost sensors and IoT technologies.

Chapter 5 Case Study One - A pilot study for individualised assessment of IEQ

This chapter is adapted from previously published work to fit the context of this thesis. Case Study 1 was adapted from the conference proceeding: **IoT in the Wild: An expedition of discovery for remote monitoring**, which was presented at the **UBICOMP: Wild by Design virtual workshop** in **September 2021.**
This work was made available online on **21 September 2021** via:
https://doi.org/10.1145/3460418.3479364
The copyright for this work was retained so that it could be used within this thesis.

5.1 Introduction

Chapters 3 & 4 identified a range of technologies that could be used to monitor buildings in response to Chapter 2. These included low-cost microcontrollers, which can interface with a plethora of sensors, but require development, and proprietary off-the-shelf monitoring devices, which can be rapidly deployed. However, chapter 4 identified that it can often be challenging to interact with off-the-shelf devices as they are often tightly coupled with proprietary applications that interface with e.g., smartphones, creating a symbiosis between the proprietary software and hardware. For daily monitoring of individuals, this can present challenges due to the limited integrations offered by proprietary applications.

As uncovered in Chapter 2, emergent, disruptive, and accessible sensor technologies are reducing costs associated with remote sensor deployment [173], which is increasing feasibility of remote localised assessment of individuals. These technologies have a range of applications for remote healthcare monitoring with WHTs and passive environmental monitoring. For example, Fitbit can provide affordable and direct mechanisms for free-living monitoring that have potential to provide new digital biomarkers in research [174]. Whereas environmental sensors can be used to monitor IEQ, with potential to provide environmental context to individualised monitoring. Poor IEQ can impact general health [175], so augmenting those data with information from WHT can provide a wider health context [176]. However, as learned from Chapters 3 & 4, multi-modal data capture from many different devices can create unique challenges even in laboratory conditions, but these can be exacerbated during remote deployment and more so when deployed longitudinally. This depends on devices having remote access capabilities, but also ensuring environments have suitable communication/connectivity infrastructures. Moreover, many systems/devices are tied to closed eco-systems that can be challenging to utilise within research [170].

This chapter will begin to address **PoI5** to understand whether low-cost sensors and IoT technologies can be used to gather with precision, consistency, and reliability across different environments. Though this investigation, will be conducted over a series of experimental work across multiple chapters *(Chapters 5 – 7),* this will first start by exploring the remote deployment of off-the-shelf monitoring equipment to evaluate the limitations surrounding how to simultaneously interact with multiple proprietary systems. This chapter presents the first of two case studies emerging from Chapters 3 & 4. These studies are deemed important to examine the use of accessible technologies for remote occupant monitoring. The following chapter *(Chapter 6)* will explore the use of low-cost sensors and SaaS solutions; but here, the focus is on the use of off-the-shelf technologies, which were deployed in a pilot study that aimed to gather data from remotely deployed WHT and IEQ devices via a bespoke PaaS solution.

This study was prematurely halted due to the SARS-COV-2 pandemic, but the primary aim of this study was to examine how to interact with proprietary systems, so the setup and mobilisation phases provided noteworthy evidence for the development of this thesis. Moreover, the challenges faced during the setup and mobilisation phases are worth communicating as they presented valuable lessons learned that not only informed the development of this thesis but could also provide greater awareness in future studies.

5.2 Study setup

The purpose of the pilot study was to capture data from a variety of remote devices to explore possible links between WHT outcomes and IEQ. The primary purpose was to examine collection protocols while aggregating data from different off-the-shelf devices within a single case (female, 44 years) home and office setting. A single case method was chosen as it can inform many types of research but are useful in exploratory research and pilot studies for monitoring individuals (especially during longitudinally situation) to gain a wider context on health outcomes and patterns of behaviour, when compared to group-based studies [177]. Ethical consent was granted by Northumbria Research Ethics committee (REF: 17141) and the participant gave written informed consent before the study commenced.

5.2.1 Technologies and outcomes

The health and environmental outcomes, and the devices used to capture them, are listed in Table 10. WHT outcomes were measured using a Fitbit Charge 3 and IEQ outcomes were measured using a Netatmo[4] Healthy Home Coach and a Foobot[5] Air Quality Monitor. Each device was configured to use the participant's smartphone (Apple iPhone 11) where data were synced periodically to the proprietary mobile apps. Data were then accessible from a cloud account and could be accessed via the web platform, integration apps *(e.g., Alexa and If This Then That, IFTTT)* and proprietary APIs.

[4] https://www.netatmo.com/

[5] https://foobot.io/

Table 10: Outcomes measured with remote sensors

	Netatmo Healthy Home Coach	Foobot Air Quality Monitor	Fitbit
Temperature *(°C)*	✓	✓	-
Humidity *(%)*	✓	✓	-
CO2 *(ppm)*	✓	-	-
eCO2 *(ppm)*	-	✓	-
PM2.5 *(μg/m³)*	-	✓	-
VOCs *(ppb)*	-	✓	-
Outdoor Pollution *(AQI†)*	-	✓	-
Noise *(dBA)*	✓	-	-
Steps	-	-	✓
Calories *(kcal)*	-	-	✓
Distance Travelled *(km)*	-	-	✓
Heart Rate *(bpm)*	-	-	✓

† *Air quality index taken from geolocation data gathered by connected mobile device*

5.2.2 Study setting

IEQ sensors were placed in the participant's home and on their desk within their office in a multi-occupant room on a university campus. The pilot study was initially designed to run through spring 2020 for 8 weeks. It was mobilised on 18 March 2020 but was subsequently halted on 23 March due to COVID-19 UK lockdown restrictions, forcing the participant to work at home only. Pilot projects are often only reported, when successful, as a means to support future work, but the outcomes *(whether positive or negative)* can be of benefit to future researchers [178].

5.3 Taming the wild: Lessons learned

Irrespective of the project being halted, a series of challenges had to be addressed within the setup and mobilisation of technology [178]. This section outlines what steps were taken to mitigate and/or overcome the impact of these challenges.

5.3.1 Integrator Apps: Limited data access

Consumer grade IoT devices are typically marketed to highlight a multitude of options to connect data. Cross-platform mobile support, web-based access from a browser, and tight integration with smart home platforms (e.g., Amazon Alexa) make these devices appealing. However, those data can often lack detail suitable for robust health When direct, built-in integrations are not available it is possible to use code-free automation/integration platforms, with one of the most popular being IFTTT[6][179], [180]. IFTTT gets its name ***(If This Then That)*** from the fact that it is a trigger-action programming platform, whereby **I**f a trigger (**T**his) is called **T**hen perform an action (**T**hat). The advantage of IFTTT is that it is a code-free automation platform meaning it has no prerequisites for Computer and Information Sciences.

Here, IFTTT was explored to test the efficacy of platform for the collection of research data. IFTTT was used to extract data from a FitBit fitness tracker (Fitbit Blaze) and log the data to Google Sheets. However, there were only two automated categories of triggers available for FitBit – daily summaries and threshold triggers (Figure 6). Daily summaries would capture an aggregated summary of the previous day's data from the FitBit API, and threshold triggers would trigger IFTTT if set goals for *e.g., Steps, Calories, Floors,* are reached. Thus, it is not possible to capture continuous data from these services or make requests for data during a specific period. Furthermore, consumer activity monitors utilising accelerometers *(the same sensors used by clinicians to detect sensitive spatio-temporal gait characteristics* [181]*)* will rarely output raw *(sample level)* data. Instead, data are processed and output as e.g., step count. Therefore, the IFTTT platform was not fit-for-purpose within the context of this thesis.

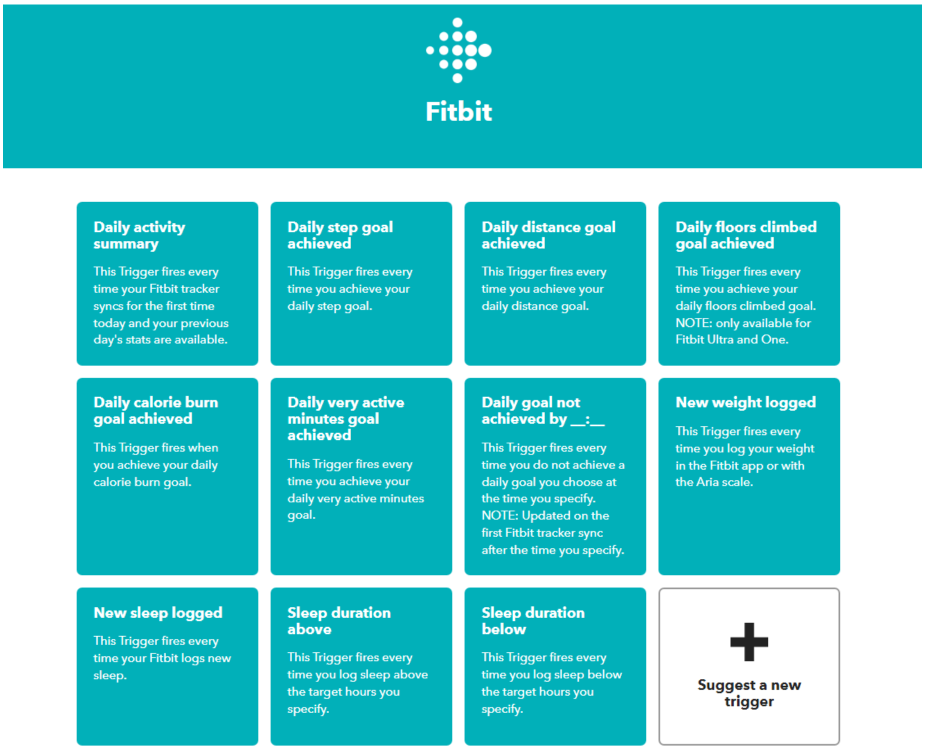

Figure 6 – Screenshot of triggers available for Fitbit taken from the IFTTT platform [182]

[6] https://ifttt.com/

5.3.2 *Proprietary APIs: Restrictions and regulations*

Data captured from proprietary hardware are often stored in the manufacturers closed eco-system. Although these eco-systems may provide access to backlogs, resulting data are often collated to e.g., hourly aggregations [170]. Data access varies by device, manufacturer, data type and is also dependant on the level of authorisation. Data are often synced to a mobile device, web-based platform or both but there is often a requirement for software development to create apps that can interface with the proprietary API and provide authenticated requests for data.

IEQ data from the Netatmo and Foobot devices were accessible via an API key setup within the associated account for each device. Data were collected using Azure Function Apps (developed in .NET Core version 3.1) which made periodic authenticated calls to the respective APIs (Figure 7). An Azure function app was also used to obtain data from the participant's Fitbit, but an intermediary API (developed in PHP using the Simple API framework) was used to provide an authentication layer that the Function App could use to fetch data.

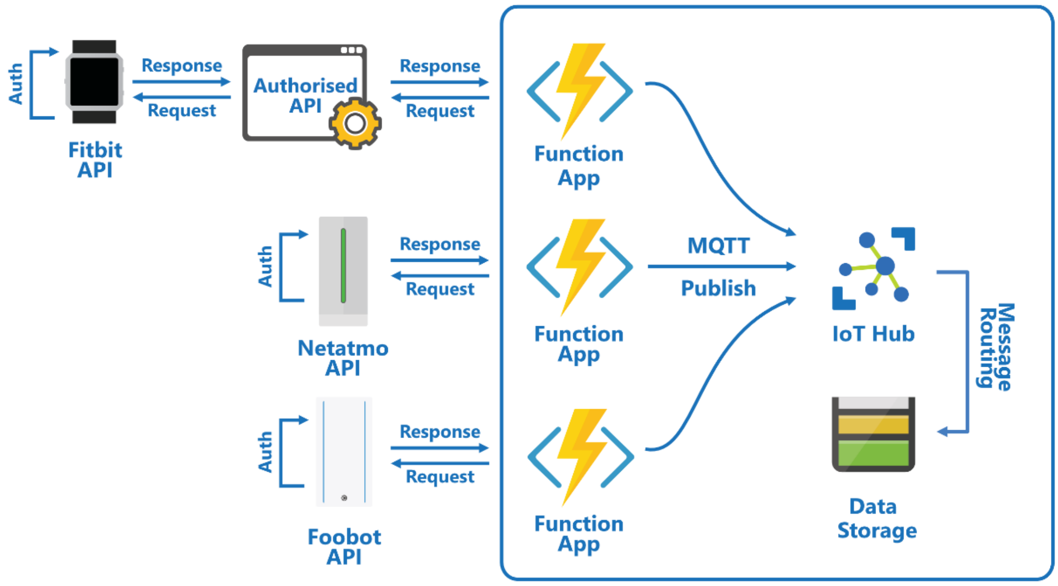

Figure 7 - System Architecture for extracting data from proprietary APIs using Azure Function Apps.

Azure Function apps communicated with the proprietary APIs. Initially these were developed in Python, but it became apparent when hosting the app that Azure can only host Python-based Function Apps on Linux. However, Linux hosted apps require an additional service plan to be set up, incurring additional costs. Therefore, apps were subsequently developed in .NET Core and hosted on a Windows server.

5.3.3 Data Security: Authentication and protection

Under GDPR, Fitbit provides access to historical data via the Web API only. To access intra-day data there was a requirement to make an additional ethical application to Fitbit outlining the purpose of the data collection, as well as indicate how the data would be collected, used, stored and managed. Within this process it became apparent that using Azure Function Apps to directly communicate with the Fitbit API was not appropriate. Fitbit authentication protocol requires an application to be developed with a specific call-back that is protected by a Secure Sockets Layer (SSL). That enables Fitbit users to securely connect their accounts and authorise the application to access specific health outcomes.

Call-backs are not available in Azure Function Apps and require additional web services to provide an interface. Therefore, Function Apps were unsuitable for connecting directly to Fitbit. Consequently, an intermediary API was developed in Hypertext Preprocessor (PHP) using the Simple API framework. The API had a single endpoint that could be called by the Azure Function App to provide an authenticated layer required to access Fitbit data.

5.3.4 Communication and connectivity

IEQ monitoring devices were initially brought to a university campus for testing, with the intention of connecting the devices to the university's Wi-Fi network. However, IoT devices were configured to accept Wi-Fi credentials in the form of a Network Name Service Set IDentifier (SSID) and Password pair, and the university used Enterprise Wi-Fi Protected Access 2 (WPA2-Enterprise). This meant that it was not possible to connect the IEQ monitoring devices to the internet via the university network.

The use of 4G Subscriber Identity Module (SIM) card routers overcame these challenges, but they also provided additional protection regarding ethics and governance. Since the intention was to deploy IoT devices both on a university campus and in the participant's home, the sand-boxed environment provided a suitable mitigation against the security risks involved with using IoT devices, which are becoming increasingly targeted by malware [183].

5.3.5 Remote deployment: Access restrictions

Prior to the UK lockdown there were established procedures within the university to prohibit unnecessary travel or meetings. Consequently, sensors were given to the participant, having been pre-configured prior to deployment. This meant the setup and mobilisation of the study had to be done remotely. However, shortly after deployment, the SIM cards became unresponsive and had to be reconfigured remotely, by assisting the participant over video conferencing. This was problematic due to the type of SIM card that was used.

Researchers conducting short-term/pilot projects may be inclined to choose a pay-as-you-go (PAYG) SIM card so that they are not constrained with contracts that extend beyond the study period. Unlike contracted SIM cards, which typically have accounts associated to them, PAYG SIM cards are not always designed for use in 4G routers. For example, a TESCO Mobile[7] PAYG SIM card was initially selected, but this had to be switched to another SIM as there was no online access. All configuration of the SIM was done via a Short Messaging Service (SMS), which made remote deployment unfeasible. Despite not having a contract, GiffGaff[8] provided online access to the SIM account, which allowed data consumption and renewals to be monitored and managed.

5.4 Discussions and Conclusions

Although SARS-COV-2 had a major impact on the length of this pilot, some challenges were identified relating to the setup and mobilisation of off-the-shelf, free-living/remote monitoring equipment. Here, solutions to overcome those challenges were presented.

When establishing in-the-wild research projects, it can be useful to understand the pragmatic technical issues once equipment is deployed. Failure to do so can delay research, increase costs, or require in-the-field modifications that could increase patient and researcher burden. Here, many challenges occurred during the setup phase. This meant contingencies *(e.g., use of 4G routers)* and alternative solutions *(e.g., use of Azure Function Apps and intermediary APIs)* could be developed before mobilisation. The challenges identified *(e.g., additional ethics from FitBit and procurement of 4G routers)* delayed the project many weeks and had these challenges been foreseen, it may have been possible to conduct the pilot a number of weeks prior to lockdown.

Transparency surrounding project limitations is vital. Nuanced technical challenges may present themselves throughout projects, yet these same challenges could have a greater impact on other researchers with different skill sets. Here, notable challenges in a similar technology deployment scenario are presented. It should be noted that lack of control and potential need for remote data access can present challenges specific to individual projects. Proprietary systems or black-boxed eco-systems can exacerbate challenges, and often present a need for bespoke solutions requiring software development. Also, due to commercial interest, proprietary device manufacturers often hide the underlying technologies to prevent competition, which can limit transparency in research.

[7] https://www.tescomobile.com/
[8] https://www.giffgaff.com/

Consumer grade monitoring equipment is becoming more accessible and is permeating into multiple disciplines of research including healthcare. Many devices are targeted towards the IoT market and integration with cloud-based services are commonplace. Integration services such as Amazon Alexa, Google Home, or IFTTT can provide access to proprietary platforms, but integration is limited due to consumer focus and are often of little use to researchers. There is a need for software development when creating interactions with proprietary APIs. However, this creates an opportunity to create bespoke implementations that can collect data from multiple data sources for holistic remote monitoring.

5.5 Addressing the PoI

This chapter presented a case study that explored the augmentation of IEQ monitoring devices and WHTs in both the home and workplace. While the pandemic significantly impacted the study, it was still possible to address the original aims of the study. The technical developments, methodological processes and the methods of study mobilisation presented a series of challenges and limitations that were deemed valuable to the outputs of this thesis.

This chapter explored how remote IoT monitoring devices and WHTs could be deployed in a multi-modal context to capture a range of IEQ and health-related outcomes. In doing so, it highlighted that there are nuanced challenges with dealing with proprietary hardware and there are often symbioses between hardware and software, which can be limiting in research. Integrations apps such as IFTTT are often promoted to solve such limitations, but these typically provide a further layer of abstraction over the raw data. The way in which data is accessed, can be limited or black-boxed, meaning it is not always possible to get raw data or get data that compares with other manufacturers.

This chapter presented challenges and limitations surrounding the deployment of remote IoT-enabled, monitoring equipment, which is useful for answering **PoI5**. However, a second case study was undertaken to provide further experimental context.

5.6 Further Research

The next chapter will continue the exploration of **PoI5** by presenting the second case study, which will investigate specific challenges highlighted in Chapter 3 surrounding high-frequency data transmission in an IoT context, which can be present when collecting data from WHTs and can have sever cost implications on projects.

Chapter 6 Case Study Two – Considerations for data sampling using IoT technologies

This chapter is adapted from previously published work to fit the context of this thesis.

Case study 2 was adapted from the article: **Towards remote healthcare monitoring using accessible IoT technology: state-of-the-art, insights and experimental design**, was published in the **BioMedical Engineering Online** on **30 October 2020** and is distributed under a **Creative Commons 4.0** license *(Appendix B)*.

This work was made available online on **30 October 2020** via: **https://doi.org/10.1186/s12938-020-00825-9**

6.1 Introduction

Chapters 3 & 4 outlined a body of research in response to **PoI4**, but highlighted a series of challenges, which may influence how **PoI5** is addressed. Predominantly, challenges surrounding the use of proprietary systems and the transactional costs of data. The previous chapter specifically presented challenges surrounding the use of off-the-shelf measurement devices, highlighting nuanced limitations that are specific to individual devices or manufacturers. This chapter will continue the investigation surrounding **PoI5** by presenting a second case study that focuses on high resolution sensors to explore more common challenges faced when collecting data from WHTs in an IoT context. This chapter, by means of experimental work, explores approaches for remote environmental and physiological monitoring - specifically focussing on a contemporary and important health related outcome, gait/walking analysis. The latter was chosen because the underlying technologies used to assess gait are also used in smart watches and PFTs for monitoring general physical activity levels. Gait was also chosen due to the frequency required for robust data capture, which is typically conducted with millisecond intervals, typically 100hz (or 100 samples per second). As discussed in Chapter 3, this can be challenging in an IoT context due to the underlying technological limitations *(e.g., serial microprocessors),* transactional costs of data transmission, and bandwidth during cloud transmission. Gait is also a valuable outcome in healthcare research, and it is commonly referred to the sixth vital sign [184] due to its ability to provide pragmatic insights to neurological conditions such as PD [185].

The current state-of-the-art in longitudinal remote gait assessment predominantly aligns to placing an inertial-based wearable (typically tri-axial accelerometer) on the lower back for extended periods (up to 7-days when also considering ambulatory behaviours). Upon completion of recording, the wearable is collected in person or returned to the researcher by post. This is extremely inefficient, costly and may often result in damage (or loss) of wearables (and data). Thus, there is a need to investigate how future habitual gait assessment could be best facilitated and maintained using IoT technologies. Addressing this need is outside the scope of this thesis, but the challenges are aligned with the overall investigation. By capturing gait data in an IoT context, it is possible to explore in greater detail the challenges presented in Chapter 3, but with a real-world example of data.

While Chapter 3 presented a conceptual model of data to communicate indicative costs of Cloud computing, this chapter will explore the specific challenges surrounding sending high volumes of high-frequency data. This chapter will also present a comparative analysis of low-frequency light intensity sensors *(1 sample in 5 minutes)* chosen at random from available IEQ outcomes, while exploring the aggregation of low-cost IoT technologies, with offline data loggers that interact with proprietary systems. These experiments extend the work conducted in Chapter 5 to explore further limitations that have been identified in Chapters 3 & 4.

This chapter also aims to extend the exploration of cloud technologies from the previous chapter, which used an Azure-based PaaS solution to capture data. This chapter will instead explore the use of a SaaS application called ThingSpeak™. ThingSpeak™ was chosen because it is developed by the creators of MATLAB®, which is a tool traditionally used to perform algorithmic analysis of gait related outcomes [186]. Consequently, ThingSpeak™ can run MATLAB® code in cloud on data collected from IoT devices. Doing this also provides, by the means of a case study, the use of an off-the-shelf SaaS application, to explore the benefits and limitations this entails.

As chapter 3 identified, the cloud computing market is extremely saturated, so it is not possible to compare all PaaS/SaaS solutions. Instead, this chapter will showcase an adopted SaaS platform, to complement the previous case study with the purpose of demonstrating the differences in the methodological processes involved in using PaaS vs SaaS solutions.

6.2 Physiological measurement of gait: Current state-of-the-art

Given the fabrication of modern inertial-based wearables due to MEMS technology, they can generally be worn on any anatomical location but placement on the lower back conforms to harmonisation of two principal algorithms for gait quantification [187], [188] to generate 14 spatial and temporal characteristics [189] of clinical utility [190]. In brief, use of the continuous wavelet transform helps identify timings of the initial (heel strike) and final contact (toe off) for each step from the vertical acceleration of MEMS based wearables, such as the AX3 (Axivity, York, UK). The AX3 has been widely used for validated gait analysis studies in various clinical cohorts [191]–[194]. Those contact times coupled with the inverted pendulum model [195], which estimates change in height of the wearable due to attachment near the wearers centre of mass, provide pragmatic gait characteristics. Furthermore, identifying periods of gait (bouts of walking) from longitudinal assessment is feasible from a heuristic approach of (i) wearable location (accelerometer orientation) and (ii) recognising periods of interest from combined tri-axial inertial signals to define when the wearer is upright (mean accelerometer output) and moving (threshold to standard deviation). Once those periods of interest are located, they are analysed for initial and final contacts to deduce that the wearer is walking [196].

Previously, it was shown that accessible IoT-based technology (smartphone, inertial wearable and Raspberry Pi) could be used beyond the clinic to gather robust gait data under observation when compared to routine procedures of analysing, via manual data download and processing through MATLAB® based gait algorithms [186]. Although the latter platform is being used less by data scientists, it remains popular due to its extensive toolboxes and formally arranged documentation and so may be perceived as the standard reference for processing sensor data. Nevertheless, more popular approaches involving use of Python or Octave have been shown to be comparable to MATLAB® for gait characteristic analysis [186], [197].

6.3 Exploring IoT approaches to remote assessment

ThingSpeak™ is an open-source Cloud platform built upon MATLAB® meaning it can run its code in the Cloud to perform real-time analysis and visualisations on incoming data streams from IoT devices.

Like many Cloud platforms, ThingSpeak™ imposes rate limits and quotas and these could be a major limitation for longitudinal assessment and multi-patient monitoring. When transmitting data to ThingSpeak™, data can be sent as individual messages where one message could comprise a reading from up to eight sensors. Alternatively, those data can be batched and sent collectively (i.e., in bulk) but regardless of transmission method the rate cannot be greater than one every 15 seconds. Nonetheless, ThingSpeak™ limits the number of readings that can be transmitted in a bulk update message, with free users being limited to 960 rows and paid subscriptions being limited to 14,400 [198]. Given a sample rate of 100Hz, each 15 second period would consume 1500 messages, equating to 8,640,000 messages/day. ThingSpeak™ charges in units where each unit includes a quota of data channels and messages. For academic subscriptions, costing $250/unit, a single unit has a message quota of 33 million messages. Therefore, a single unit would last just under four days if data were continuously transmitted. For longitudinal and/or multi-patient monitoring, these costs could grow exponentially. However, if the platform were used to analyse snapshots of data, biomedical engineers could fine tune their algorithms throughout a study and monitor the progress without waiting until the end of the sampling period. For environmental monitoring, high frequency transmission is not always necessary, so these limits are not a factor.

6.4 Experimental setup and equipment

To test the feasibility of ThingSpeak™ as a SaaS data aggregation platform, two experiments were conducted. The first experiment was conducted to compare MATLAB® and ThingSpeak™, within the context of gait analysis to examine the suitability of the platform for high-frequency data capture and to evaluate whether the ThingSpeak™ platform has feature parity with MATLAB®. The second experiment evaluates ThingSpeak™ using environmental data captured at lower-frequency to aggregate data from an IoT sensor and a reference-standard data logger, to evaluate the suitability of the platform as a general data aggregation SaaS platform.

For the purposes of the first experiment, AX3 data from a single user in their habitual setting is presented. The participant wore a single AX3 (100Hz, ±8g) on the lower back for 1-hour during which time they were free to perform their normal activities. Ethical consent was granted by the Northumbria University Research Ethics Committee (REF: 16335/335) and the participant gave informed written consent before participating in this study.

Since the AX3 lacks wireless connectivity, the device was plugged into a desktop computer and the data were extracted and exported to CSV format. Those data were then analysed in MATLAB® using a usual approach and validated algorithm [196]. Subsequently, a MATLAB® analysis application was created on ThingSpeak™ that contained the same code as on the desktop. The CSV file was then imported into ThingSpeak™ and analysis of these data was performed in the Cloud.

Whilst the ability to run MATLAB® code in the Cloud is one of the primary benefits of ThingSpeak™, the platform also provides supported integration and code libraries for Arduino based devices. Therefore, to test the potential of using the platform as a data aggregation platform *(for the second experiment)*, a low-cost MEMS-based light intensity sensor (BH1750) was used to collect and transmit data to ThingSpeak™ every 5 minutes using a Heltec ESP32 Wi-Fi 32 development board (Figure 8). This enabled the evaluation of ThingSpeak™ for real-time capture of low-frequency data. The frequency of data transmission was set to match a reference device, the HOBO MX1101 light intensity data logger, which was simultaneously logging data on local storage to validate data from the BH1750. Data were captured from both devices consecutively for five days. Data from the HOBO were pushed to ThingSpeak™ after the collection period and were used to cross examine the low-cost MEMs-bases sensor with the reference standard device.

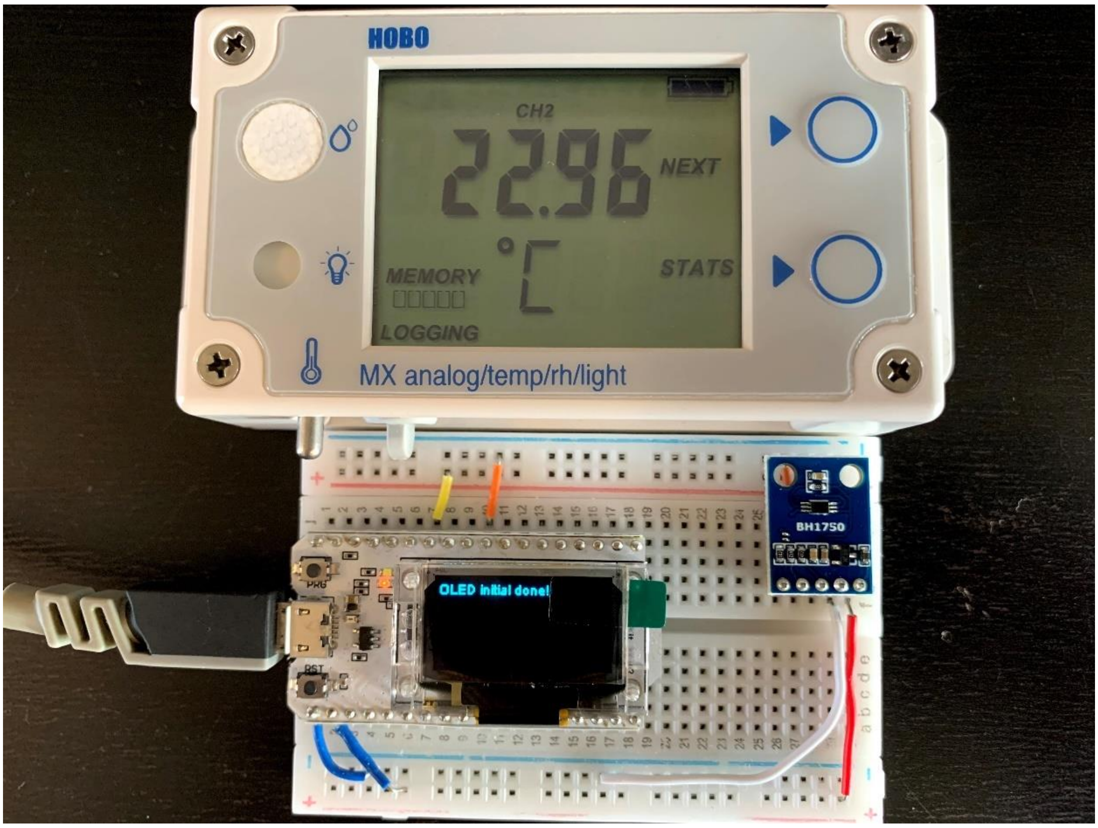

Figure 8 - MX1101 light intensity data logger and BH1750 ambient light sensor connected to ESP32 development board.

6.5 Findings

Here the findings of both the high-frequency and low-frequency capture experiments are presented.

6.5.1 Gait: High frequency data

Individualised gait data was successfully gathered and download via the usual desktop approach. The algorithm successfully segmented and identified gait events (Figure 9, each bout was examined for initial and final contact times) and generated spatial and temporal outcomes, presented previously [181], [196], [199]. In contrast, I found that while ThingSpeak™ could collect, store, visualise and analyse data from low-frequency environmental sensors, its ability to be used for existing gait assessment approaches within an IoT context highlighted some major limitations. Although the rate limits imposed allow up to 14,400 readings to be sent every 15 seconds, the platform is capable of processing high-frequency data akin to similar approaches via a desktop. However, during a bulk update ThingSpeak™ checks no duplicate rows exist by comparing the timestamp of each reading. While this validation process accepts milliseconds and microseconds resolution timestamps, ThingSpeak™ rounds these to the nearest second, making it unsuitable for high frequency data. Given the 1Hz frequency limitation, to test how the Cloud-based MATLAB® Analysis compared with desktop approach, I circumvented the timestamp checks by changing timestamps to epochs in (seconds). This allowed high frequency gait data upload for analysis.

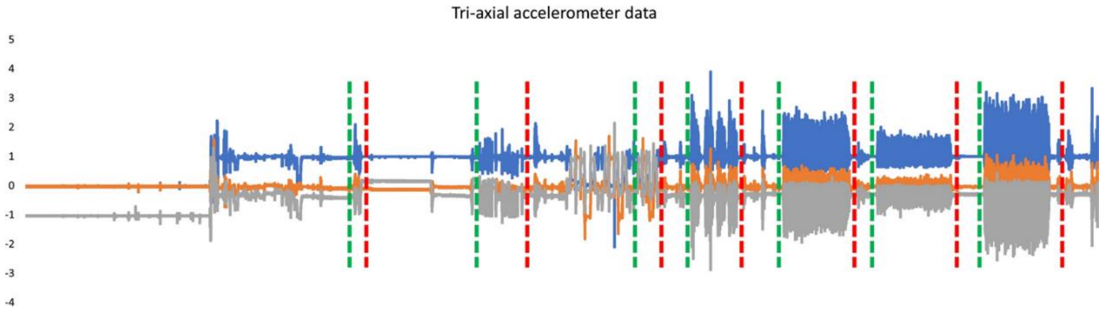

Figure 9 - Free-living tri-axial accelerometer data (AX3). The vertical green and red indicate possible start/stop gait bouts.

6.5.2 High frequency analysis via IoT: Gait as a Example

Reading data via a ThingSpeak™ channel instead of from a CSV file stored on a desktop uncovered further limitations. Firstly, ThingSpeak™ limits readable data to 8000 rows, which meant that analysis had to be batched into 80 second sample windows. Once complete, a further error was encountered as the code utilised (e.g., filtering) functions from MATLAB® toolboxes that were not present in ThingSpeak™. Despite the removal of filtering processes, further errors were encountered, which highlighted fundamental differences between the two computation engines. While attempts were made to evaluate the IoT approach to gait assessment, in its current state, ThingSpeak™ is currently unsuitable for collecting high-frequency biomedical research data.

6.5.3 Inter-device data aggregation

The official Arduino support from ThingSpeak™ made connecting the BH1750 to the cloud a seamless process. Data were transmitted directly from the Heltec development board which was connected to the internet via Wi-Fi. Each time data was sent to ThingSpeak™, live graphs were updated allowing data from the IoT device to be quickly visualised. During data collection it was also feasible to download data ad-hoc (as a CSV file) as well as analyse the data directly in the Cloud. Data transmission frequency meant there was no need to consider any rate limits imposed by ThingSpeak™ and the ESP32 was more than capable of transmitting the data at such a low frequency. Data from the HOBO was exported from the device using the proprietary mobile applications *(HOBOmobile)* and imported into ThingSpeak™ after the data collection period to perform inter-device data comparison, as the platform supports both real-time capture and historic data imports.

Regarding the data validation, the BH1750 was found to be highly correlative to the HOBO MX1101 sensor, with a Pearson correlation of 0.799. Moreover, Figure 10 shows that whilst the accuracy of the BH1750 is slightly lower than the MX1101, the BH1750 is comparably responsive to changes in light intensity. The results of this experiment highlight the potential low-cost MEMS light sensors have in measuring ambient light intensity. They also highlight the potential of Cloud platforms such as ThingSpeak™ for remote monitoring of an individual's environment. For real-time capture of low-frequency data capture, ThingSpeak™ doubled as both a solution for long-term data management as well as data aggregation, even allowing up to eight sensors to be captured for free.

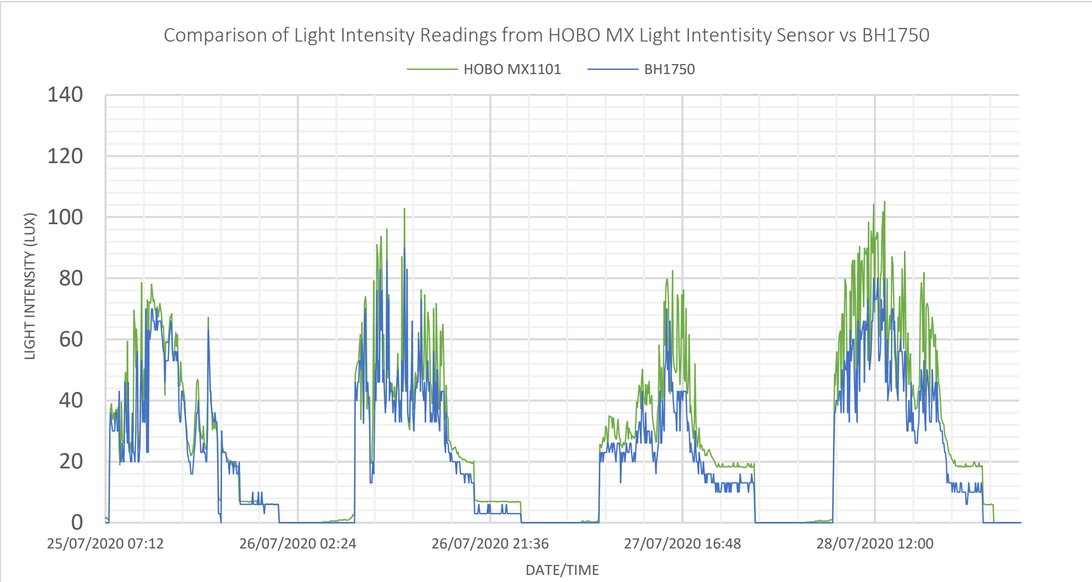

Figure 10 - Data captured from HOBO MX1101 and BH1750

6.6 Discussions and Conclusions

This case study focused specifically on ThingSpeak™ as it could be reasoned as an accessible and logical next step to explore Cloud IoT platforms for this thesis. This study found that low-cost infrequent data collection is feasible using ThingSpeak™, which makes the platform suitable for environmental data collection. Yet, there are serious limitations that make the platform unsuitable for physiological monitoring, namely rate limiting that curtails data logging to 1Hz which is unsuitable for intensive physiological monitoring e.g. spatio-temporal gait analysis.

This study focused the experiment on gait analysis and the augmentation of environmental data, due to emergence of the former as a pragmatic patient monitoring outcome. Our evaluations found that current limitations with the ThingSpeak™ platform make the platform unsuitable to process high-frequency data.

However, it could be a useful and inexpensive way to augment environmental data with healthcare data to provide more context during habitual assessment. Our experiment highlighted that low-cost MEMS technology can provide valid data which can be suitably collected, analysed, and visualised via ThingSpeak™.

Consumer grade monitoring equipment is becoming more accessible and is permeating into multiple disciplines of research including healthcare. Many devices are targeted towards the IoT market and integration with cloud-based services are commonplace. Integration services such as Amazon Alexa, Google Home, or IFTTT can provide access to proprietary platforms, but integration is limited due to consumer focus.

6.7 Addressing the PoI

This chapter continued the exploration into **PoI5** and conducted a measurement into accelerometers with a medical context to capture data at a high frequency and explore the findings in Chapter 3 surrounding the transactional costs/limitations of high-frequency data processing. Given the prevalence of MATLAB® in this field of research, ThingSpeak™ was chosen as a free cloud platform, with an open data schema, meaning multiple sensors could transmit to the platform without requiring data processing.

However, while ThingSpeak™ can run MATLAB® code, there was limited feature parity between the two platforms, so it cannot be used in place of MATLAB® at this time. Another consideration is that the transactional rate limits in ThingSpeak™ imposed limitations that made the platform unsuitable for high resolution data capture (>1Hz). This could potentially be solved with edge computing, though this would be outside the scope of this thesis. Regardless, the platform was identified as fit-for-purpose for the capture of IEQ data, which is typically captured at a much lower capture frequency.

6.8 Further Research

At this point, this thesis has presented a series of reviews and case-studies that position the research questions, aims and objectives of this thesis in context. In doing so, it provides suitable background research and experimentation required to begin to fully address **PoI5**. To build upon Chapters 5 & 6, Chapter 7 will continue to explore **PoI5** with a comprehensive evaluation study that documents the development and validation of a low-cost, multimodal, environmental monitoring device that uses Wi-Fi enabled MCU hardware, and transmits data to ThingSpeak™. In doing so, it will present the development, testing and validation of a bespoke device to evaluate whether low-cost sensors can be used to gather with precision, consistency, and reliability across different environments i.e., home, and commercial settings.

Chapter 7 A scalable and multimodal approach to monitor IEQ

This chapter is adapted from previously published work to fit the context of this thesis. The article: **Low-cost, multimodal environmental monitoring based on the Internet of Things**, was published in the **Building and Environment** Volume 203 in **October 2021.**
It was made available online on 31 May 2021 via: https://doi.org/10.1016/j.buildenv.2021.108014.

7.1 Introduction

Chapter 2 identified that the dominant approach to monitoring IEQ is with research grade devices, which offer excellent quality data but have many limitations. They are often unimodal, measure only one aspect of IEQ and can cost thousands of US dollars [176]. Given that IEQ is determined by many different factors, this requires multiple devices to gain a holistic picture. The review also found that as a result of the cost and complexity of the measurement equipment, indoor spaces are often measured from a single location, reducing the spatial density of IEQ measurements [10], [48], limiting what can be inferred about the health and wellbeing of a spatially distributed set of occupants. Research grade devices can be moved between locations, but these limit long-term measurements in any one location, meaning data from different locations will not be temporally comparable.

Occasionally it may be beneficial to sacrifice precision and accuracy of measurement for gains in spatial density and longitudinal monitoring. Chapter 3 explored low-cost sensor technologies that could positively impact on building sciences (and the wider healthcare, architecture, engineering and construction fields), by enabling the development of multimodal monitoring devices that can measure a range of IEQ factors from a single low-cost device [173]. Low-cost sensors would be a likely requirement in creating devices that could be scalable to support localised and therefore, increased spatially dense IEQ monitoring. Such devices could capture remote longitudinal IEQ data in many locations and could even be positioned near individuals to provide greater insights into the long-term effects of indoor environments on individual occupants [39].

While low-cost technologies are enabling the development of pragmatic and affordable solutions for remote monitoring, there is a general lack of acceptance towards low-cost sensor technologies [176]. This issue is further complicated by a lack of transparency from manufacturers of consumer-grade monitoring equipment, that utilise low-cost sensors. While manufacturers typically publish accuracy information and technical specifications, they often negate to publish which sensors are used in their devices, which results in the need to explore internal photographs from Federal Communications Commission (FCC) reports or dismantle devices to identify sensors [26]. This means there are requirements to continually benchmark, verify and validate these technologies to reference (gold) standard devices. This can be extremely costly and impracticable to implement [7], [21], [26]–[28].

The main objective of this chapter is to bring together pilot work in chapter 5 and 6 to complete **PoI5**:

- *Can low-cost sensors be used to gather with precision, consistency, and reliability across different environments i.e., home, and commercial settings?*

To investigate, the chapter will propose and develop a multimodal, IoT-enabled, monitoring device that is low-cost and scalable, to support localised environmental monitoring. A secondary aim is to compare the accuracy and precision of the low-cost sensors with reference standard devices, used in IEQ research, to determine suitability of low-cost sensors for applications in this space. That will be achieved by presenting a suitable and robust analytical/statistical analysis methodology/process which could be considered within the field. The multimodal approach will be constructed to measure a wide range of important IEQ factors from a single low-cost device.

7.2 Related Work

Air quality, thermal comfort, sound/noise and light are the four most commonly measured factors of IEQ [3], [5]. Low-cost sensors can be used to measure a range of data relating to those factors, but differences exist in the technologies and methods used to derive them [176]. This section will provide a rationale for the selection of low-cost sensors in the proposed multimodal device by comparing to reference standards.

7.2.1 Inclusion criteria: Sensor integration

Cost and accessibility of IoT technology has resulted in a vast array of instruments. Many have a great deal of support (manufacturers and the IoT community of users [173]) where, e.g., hardware is supported by official code-libraries to expedite integration between sensors and microcontrollers [200]. Many libraries provide mechanisms to read and process sensor data, with minimal coding. In addition to low-cost, availability of libraries and functionality within was another inclusion factor within this study when selecting low-cost sensing instruments.

7.2.2 Low-cost IEQ sensing

7.2.2.1 Air quality: equivalent Carbon Dioxide (eCO_2) and Volatile Organic Compounds (VOCs)

Non-dispersal, infra-red (NDIR) sensors are most commonly used as the reference/gold standard for measuring Carbon Dioxide (CO_2) [201]. However, sensors that measure eCO_2 *(known as equivalent or estimated CO_2)* are becoming more popular due to their low-cost (*approx US$5-10)* [173]. Yet eCO2 sensors are often used in place of CO_2 sensors and incorrectly reported as CO_2 [92], [93], [202].

eCO_2 sensors use a heated Metal Oxide (MOx) semiconductor on which oxygen reacts with gasses to change the resistance, proportional to gas concentrations [203]. However, MOx sensors are highly sensitive to environmental conditions and a wide range of gases and pollutants, which can have a major influence on measurement accuracy [204]. As the name suggests, eCO_2 is not an actual measure of carbon dioxide but an estimation. The value is derived from a measurement of the Total VOCs (TVOCs), which describe the total concentration of organic, carbon-based compounds that evaporate into the air at room temperature [205]. There is a lack of clarity around what TVOC sensors actually measure and also little information about how the readings are calculated or whether readings/outcomes are standardised [88]–[90]. Algorithms that estimate TVOC values are typically implemented within the on-board microcontroller of the sensors and are often black boxed [176]. General assumptions are also made during the calculation of both TVOC and eCO_2. For example, TVOC sensors assume the primary source of VOC is from humans and validity can be questioned when measuring in environments where this is not case [88].

Naepelt [206] provided transparent algorithms for eCO_2/TVOC calculations, but they depend on knowing the proportion of human generated VOCs. For their calculations, they assert that human generated VOCs increase gradually, and VOCs released from aerosols or cooking increase rapidly. However, TVOCs are most commonly released by finishes and furnishings, especially those made using artificial materials such as solvents and adhesives [207]. As room temperature affects the concentration of such VOCs [205], it cannot be assumed that all artificial VOCs increase TVOC concentrations rapidly. Moreover, there can potentially be hundreds of VOCs within indoor air [208], so making assumptions based on one or two of the most volatile compounds is likely to have inaccuracies.

Given the limitations of TVOC sensors, they will not be used as a measure of air quality in the proposed multimodal device. However, since eCO_2 is often misrepresented as a measurement of CO_2 [92], [93], [202], a MOx eCO_2 sensor was selected for comparing eCO_2 readings with readings from reference CO_2 sensors. This was to identify whether eCO_2 readings have any correlation with CO_2 concentrations and to understand whether it is suitable to use them in place of NDIR CO_2 sensors. Thus, an AMS CCS811 *(Table 11)* was selected for our multimodal device as it is regarded as a reliable solution for measuring gas concentrations [209]–[211]. Consequently, a CCS811 breakout board was chosen as it provides an I^2C interface and is supported by an official Adafruit Arduino code library [212].

7.2.2.2 Air quality: Carbon Dioxide (CO_2)

High concentrations of CO_2 can have a range of impacts on productivity and cognitive performance [37], It is therefore often used in workplace IEQ studies. However, extremely high concentrations of CO_2 are required before it becomes detrimental to health [37]. Thus, CO_2 is often regarded as a poor indicator of Indoor Air Quality (IAQ) [20], [41], [57]. Nevertheless, it has become more commonly used in IAQ/IEQ monitoring as a proxy outcome for ventilation [38], [57]. Focus on CO_2 sensors for ventilation monitoring has increased rapidly since the outbreak of the SARS-COV-2 pandemic [213], [214], to monitor the circulation of fresh air within buildings and to help stop the spread of the virus. This greatly increases their importance in IEQ monitoring and provides a need for monitoring solutions that can be deployed at scale.

NDIR sensors set the standard for CO_2 measurements [201] and technological advancements are driving a growing market of low-cost sensors that can also measure CO_2 using the same NDIR technology [215]. Due to the complexity of NDIR components, even the lowest costing NDIR sensors (approx. US$20) is greater than the cost of most MOx sensors. However, the price range of low-cost NDIR CO_2 sensors is much broader than MOx sensors. For example, a previous study [176] identified several low-cost, NDIR CO_2 sensors that ranged from US$20 (MH-Z19, Winsen Electronics) to US$200 (CozIR, Gas Sensing Solutions), with most costing approx. US$100. Here, the MH-Z19 (Table 11) was the most viable option for use in the proposed multimodal device. Despite its low cost, the MH-Z19 is regarded as a reliable, stable and accurate sensor for CO_2 measurement [216]–[218]. The MH-Z19 is supported by Dempsey's Arduino library [219], which provides many functions for interfacing.

7.2.2.3 Air quality: Particulate matter (e.g., $PM_{2.5}$)

$PM_{2.5}$ is used to describe airborne particulate matter. The '2.5' refers to particles that are up to 2.5µm *(microns)* in diameter. The term PM_{10} is also used to describe particles >2.5 microns, but <=10 microns. $PM_{1.0}$ is used to describe the smallest range of particles (up to 1.0 microns), but is not standardised by global environmental protection agencies (EPAs) [220]. The current specification of the US EPA is that within no 24-hour period should $PM_{2.5}$ exceed 35µg/m^3 [221].

Measurement of particulates and the pollutants within them requires expensive equipment that collects particles in a filter and measures the weight increase using a Tapered Element Oscillating Microbalance (TEOM) [222]. Lower-cost, optical, sensing technologies can also be used to measure PM, but these cannot detect particles <0.25µm and can miss some sources of pollution such as a cooking that does not involve frying or heating oil [222].

A previous review [26] identified the scientific potential of the FooBot Air Quality monitor, which uses a Sharp GP2Y1010AU0F (GP2YX) optical dust sensor [26]. Therefore, the GP2YX was selected for use in the low-cost multimodal device, but initial testing with the GP2YX produced erratic and highly inaccurate data. Alternatively, PlanTower sensors are often used in many commercial devices [222] and have been found to report data that correlates with reference equipment when measuring $PM_{2.5}$ [223]–[225]. Therefore, a (PlanTower) PMSA003 *(Table 11)* sensor was selected here to measure $PM_{2.5}$. The Adafruit

PMSA003i variant of the sensor was chosen as it provides an I^2C interface and an official Arduino code library to interact with the sensor [226].

7.2.2.4 Thermal comfort: Temperature and humidity

Temperature and humidity have an influential role in IEQ monitoring as optical sensors that detect airborne particles and molecules (such as those used to measure CO_2 and PM) can be highly impacted by thermal changes [227]. Therefore, many commercial devices include a temperature and humidity sensor to support and/or calibrate the primary sensors [176].

Hygrometers are commonly used to measure humidity and temperature simultaneously. Hygrometers are typically small *(approx. 1cm²)*, low-cost *(approx. $2-5)* devices that output signals which can be read by analogue inputs on microcontrollers and processed with an Analogue-to-Digital Convertor (ADC). MEMS technologies are also used to measure temperature and humidity, providing a range of benefits over traditional hygrometers. For example, they are significantly smaller than analogue hygrometers *(approx. 2mm²)* and have integrated amplification and ADC circuitry.

Given the role temperature and humidity play on other sensor technologies, there are a plethora of low-cost sensors. However, the Bosch BME280 *(Table 11)* sensor was chosen as it is a low-cost, multimodal, MEMS sensor that is used within healthcare applications [228]. Moreover, the sensor also captures barometric air pressure, which can also impact readings from optical sensors [227]. A BME280 breakout board was chosen as it provides an I^2C interface and is supported by an official Adafruit Arduino code library [229].

7.2.2.5 Light: Ambient light intensity

Light intensity can be a source of discomfort for occupants, causing distractions, eye pain and skin conditions [230]. There are two common approaches for measuring light intensity *(in lux):* (i) Light Dependant Resistor (LDR), that reduces the resistance across a circuit as light intensity increases [231] and (ii) photodiodes, which converts light intensity into an electrical current [232]. LDRs have a response delay between light exposure and resistance decrease, which can be a limitation in high frequency measurement [232]. Photodiodes can use filters to target specific frequencies bands in a light spectrum and can obtain more precise measurements across a broader range of light intensities compared to LDRs [233]. They are often incorporated into integrated circuits that contain amplification circuitry and an ADC [234]. This can provide more control of the output measurements. Previous work examined the *ROHM BH1750* photodiode sensor *(Table 11)* [173], where it was found to be highly correlative to research standard sensors. Thus, it was selected here for the multimodal device. The BH1750 is also available as an I^2C breakout board and was setup according to the installation instructions and code libraries provided with the sensor [235].

7.2.2.6 Sound/noise: Noise levels

Microphones work by converting sound pressure into a linear electrical signal, meaning the latter directly correlates with the sound signal [236]. To measure loudness, decibels (dB) logarithmically scale to mirror human hearing sensitivity [237], [238]. Therefore, to measure dBs with a microphone, the output voltage from microphones is converted to the logarithmic dB scale. The complexity of the logarithmic calculation is highly dependent on the sensitivity of the microphone. Yet how the sensitivity of a microphone is determined is influenced by the selection of an analogue or a digital microphone.

The voltage of analogue microphones typically needs to be routed through both a pre-amp and an audio codec that converts the analogue signal to digital using an ADC [239]. This three-stage approach can be affected by other circuitry and communication signals such as Wi-Fi and Bluetooth [240]. Therefore, it is often appropriate to use digital microphones, built using MEMS, as they have an in-built ADC that converts the signal directly from the microphone and is therefore not as susceptible to circuit noise [240].

Since the microphone within the proposed multimodal device will be housed in close proximity to the microcontroller *(WiFi and Bluetooth enabled)*, a digital MEMS microphone was selected to minimise interference [240]. The InvenSense INMP441 *(Table 11)* microphone was chosen here, which is supported by Kostoski's ESP32 library for I^2S digital microphones [241]. This library is specifically designed for digital MEMS microphones and calculates A-weighted decibel readings from MEMS microphones and provides the necessary filters and equalisation to do so, which are based on precalculated analyses conducted in MATLAB®.

Table 11 – Low-Cost sensors used for development including min/max measurement thresholds and units that are measured

Measure	**Instrument**†	**Protocol**	**Cost**‡	Min	Max
eCO2 (ppm)	**CCS-811** [88]	I^2C	$5	400	8192
CO_2 (ppm)	**MH-Z19B**§ [242]	UART/ PWM/ DAC	$20	0	5000
$PM_{2.5}$ (μg/m^3)	**PMSA003i** [243]	I^2C	$16	0	500
Temp (°C) Humidity (%) Pressure (hPa)	**BME280** [244]	I^2C	$2	-40 0 300	85 100 1100
Light (lux)	**BH1750** [245]	I^2C	$1	1	65535
Noise (dB SPL)	**INMP441** [246]	I^2S	$2	33	120

† Detailed technical specifications for each sensor (including voltage requirements, accuracies and working temperature/humidity) can be found within the referenced datasheets; ‡ All prices (rounded to the nearest USD) are taken from AliExpress.com (10 November 2020).; § The MH-Z19B supports two measurement ranges (0-2000ppm and 0-5000ppm) over the UART protocol only.

7.2.3 Cloud connectivity

IoT cloud computing is largely dominated by Amazon, Google and Microsoft [153], but there are hundreds more readily *(low-cost/free)* accessible platforms, many tailored for unique use cases [152]. Many IoT platforms provide free but limited functionality for testing and prototyping. Equally there are often hidden costs with these services that need to be considered before choosing platforms for production [173].

Previously, ThingSpeak® [173] was identified as a fit-for-purpose IoT/Cloud platform when conducting IEQ monitoring. It is developed by the creators of MATLAB® and supports real-time transmission and visualisations of data from IoT devices and if required, Cloud-based analysis using MATLAB® code [173]. For the purposes of this study, the quota provided in the free package was enough to evaluate the feasibility of real-time transmission from the prototype, without any costs. ThingSpeak® also provides an official code library to transmit data from various microcontrollers to the Cloud. This meant that a simple interface could be created to encapsulate the data transmission function of the code library, meaning that ThingSpeak® could be easily switched out to different IoT platforms, if required.

7.3 Reference devices

7.3.1 Onset HOBO® MX1102 (CO_2 and eCO_2)

The HOBO® MX1102 *(Table 12)* datalogger was selected to measure CO_2 due to high accuracy at room temperature [247], [248]. Although not an IoT device, it has a large storage capacity and is able to gather data continuously for several months. Since eCO_2 sensors will also be included in the multimodal device, to evaluate the validity of eCO_2 readings, the HOBO® MX-1102 sensor was also used as a reference for eCO_2 readings.

7.3.2 IQAir Air Visual Pro ($PM_{2.5}$)

Due to cost and complexity of reference standard PM monitoring equipment, where single units can cost several thousands of dollars [176], it was not possible to obtain a true reference within the budget of this study. However, a previous study [222] evaluated several lower-cost monitors *(defined as devices <US$300),* concluding that those lower-cost monitors may be used to efficiently detect $PM_{2.5}$ events. In that study, the validity of six low-cost, optical $PM_{2.5}$ were examined against a TEOM, which measured the actual mass of dust particles. Of the sensors examined, two were available at the time of the study *(Kaiterra: Laser Egg 2 and IQAir: Air Visual Pro).* The ratio of both device measurements against the TEOM was ≤1:2 but for most test events the IQAir Air Visual Pro (AVP) reported closer to the true mass concentrations captured by the TEOM. Additionally, IQAir calibrate each device against a Grimm 11-A [222] sensor, which is commonly used as a research standard measure of $PM_{2.5}$ [249]–[251]. Based on these findings, the AVP *(Table 12)* was selected as a $PM_{2.5}$ reference for this study.

7.3.3 Onset HOBO® MX1104 (Light intensity, temperature, humidity)

An Onset HOBO® datalogger was also used as a reference standard to measure ambient light intensity. The MX1104 *(Table 12)* features a similar interface to the MX1102 but measures light intensity alongside temperature and humidity. Like the MX1102, the MX1104 has large internal storage and is also highly accurate at room temperature. Temperature and humidity were validated against the HOBO® MX1104 as both HOBO® devices measure these factors, but (*according to the manufacturer's specifications)* the MX1104 has a slightly higher accuracy for temperature and humidity than the MX1102.

7.3.4 Air Pressure

To validate results from the BME280's air pressure sensor, data were compared to outdoor air pressure extracted from the weather.com API. The following endpoint was used to acquire data from the study location (Newcastle Upon Tyne, UK) during the sample period:

https://api.weather.com/v1/location/EGNT:9:GB/observations/historical.json?apiKey=<API_KEY>&units=m&startDate=20201101&endDate=20201130.

7.3.5 Omega HHSL-101 (noise levels)

The Omega HHSL-101 *(Table 12)* sound level meter was selected as it has a similar dynamic range *(100dB SPL)* to the INMP441 *(87dB)*. Many sound level meters are designed for real-time measurements, but to validate the INMP411 it was important that data could be logged, extracted and analysed. The Omega HHSL-101 logs with a decimal resolution *(0.1dB)* to internal storage, up to 32,000 samples. At 10 second (s) intervals, this is not a large amount of storage, but it enables data capture to run a comparative analysis.

Table 12 – Reference devices used, with indicative costs, min/max measurement thresholds and units

Measure	**Instrument**†	**Cost**‡	Min	Max	Units
CO_2 (ppm) Temp (°C) Humidity (%)	**Onset HOBO® MX-1102** [252]	\$595	0 0 1	5000 50 70	ppm °C %
Light (lux) Temp (°C) Humidity (%)	**Onset HOBO® MX-1104** [253]	\$185	0 -20 0	167,731 70 100	lx °C %
$PM_{2.5}$ (μg/m^3) CO_2 (ppm) Temp (°C) Humidity (%)	**IQAir Air Visual Pro** [254]	\$269	0 400 -10 0	1,798 10,000 40 95	μg/m^3 ppm °C %
Noise (dB SPL)	**Omega HHSL-101** [255]	\$149	30	130	dB SPL

† Detailed technical specifications for each sensor (including voltage requirements, accuracies and working temperature/humidity) can be found within the referenced datasheets; ‡ All prices (rounded to the nearest USD) are taken from the manufacturer's websites via Google Shopping - with the region set to United States (10 November 2020).

7.4 Multimodal device architecture

7.4.1 Hardware development

For testing, three multimodal devices were constructed to test low-cost devices/sensors against reference standards, and to test inter-sensor reliability. The low-cost sensors were connected to a Heltec Wi-Fi Kit 32 ESP32 microcontroller on a solderless breadboard (Figure 11). Additional schematic diagrams and detailed breadboard configurations have also been included in the supplementary material.

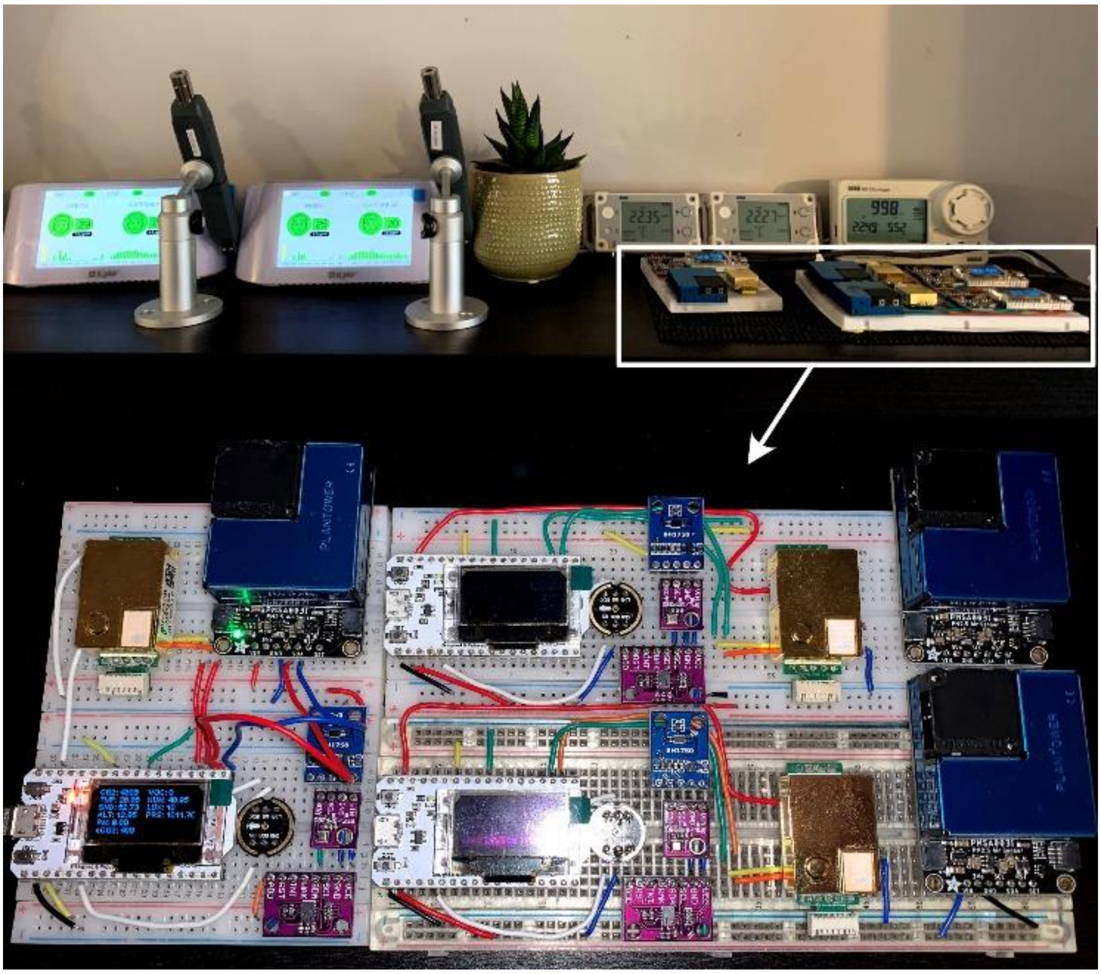

Figure 11 - Low-cost sensors collecting data alongside reference devices. Including a closeup image to show the breadboard configuration.

The Heltec was selected as it included Wi-Fi for Cloud *(Figure 12)* communication and a built-in OLED display for real-time feedback. The specifications for the Heltec also made it a suitable choice for the intended application as the included communication interfaces *(3x UART; 2x I^2C; 1 x I^2S)* can support a simultaneous connection to all sensors. Despite its size, the Heltec also has 3V and 5V output, when powered with USB, meaning it can power the MH-Z19B without additional voltage boosting circuity.

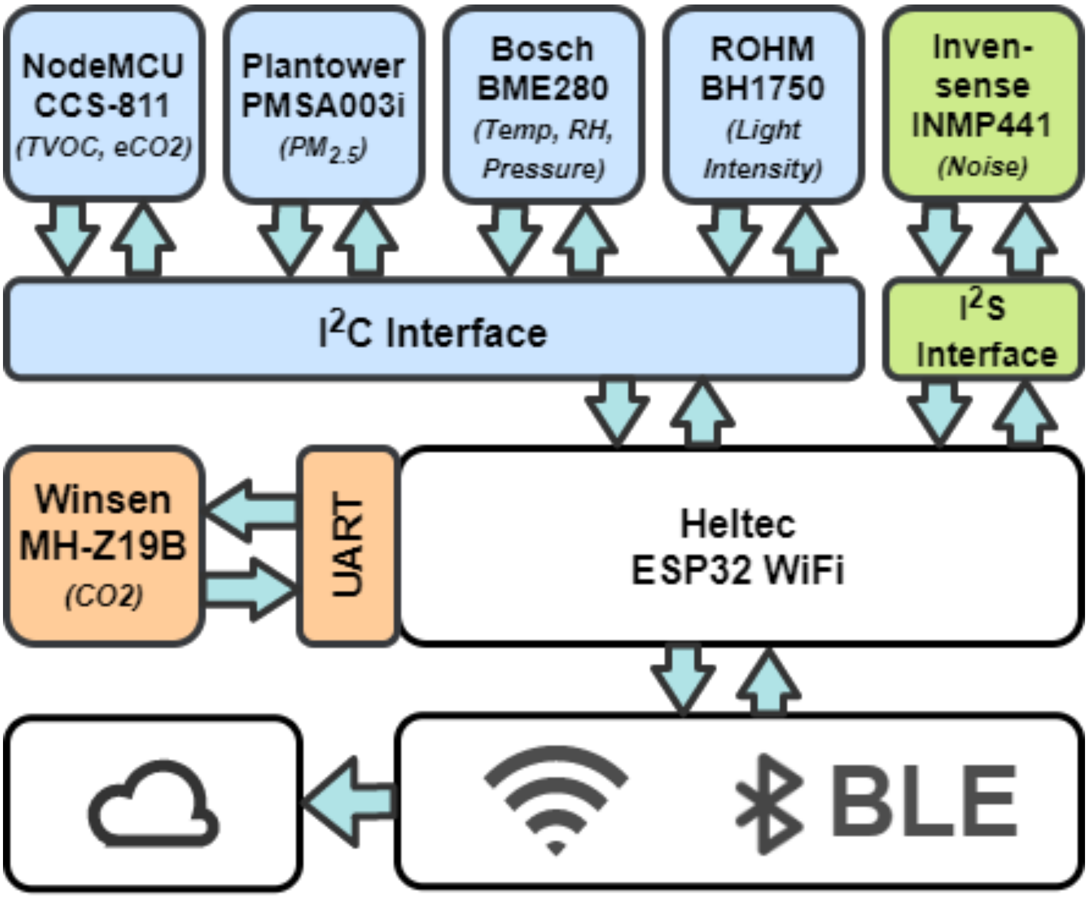

Figure 12 - Schematic diagram for multi-modal IEQ sensor device.

A mix of the Arduino IDE and Visual Studio Code (with the Arduino extension) were used to program the ESP32 with C++ code. The ESP32 can be programmed with both Arduino code and Micro Python code, but the Arduino workflow was chosen because of its maturity and wider support for sensors.

7.4.2 Reading CO_2 data via UART

The MH-Z19B outputs data via Pulse Wave Modulation (PWM) Digital Analogue Conversion (DAC) or Universal Asynchronous Receiver-Transmitter (UART). UART was selected as it enabled the read/write of byte commands to request and receive data, which facilitates a range of additional functionality, Table 13. These commands also allow the device to be configured to utilise the full measurement range of the sensor *(0-5000ppm)*.

To read CO_2 data from the MH-Z19B a byte command is first sent to the sensor's microcontroller. After processing, the MH-Z19B sends data back to the requester as a return byte command. These commands will be sent as Byte3 when sending a command to the sensor and returned as Byte2 when receiving data from the sensor, Table 13.

Table 13 - MH-Z19B UART Commands

Action	Byte0	Byte1	Byte2	Byte3	Byte 4 – 7	Byte8
Send Command	0xFF	0x01	CMD†	0x00	0x00	0x79
Receive Command	Start	CMD†	High Level	Low Level	-	Check

† UART Byte commands available on the MHZ19B - **0x78:** Recovery Reset; **0x79:** ABC Mode ON/OFF; **0x84:** Raw CO2; **0x85:** Temp float; **0x86:** Temp integer; **0x87:** Zero Calibration; **0x88:** Span Calibration; **0x99:** Range; **0x9B:** Get Range; **0x9C:** Get Background CO2; **0xA0:** Get Firmware Version; **0xA2:** Get Last Response; **0xA3:** Get Temp Calibration.

The MH-Z19B uses Automatic Baseline Calibration (ABC), enabled by factory default, that calibrates 400ppm to the lowest measured PPM in the last 24-Hour cycle. The sensor also supports zero-point calibration, whereby the sensor can be manually set to 400ppm. Manual calibration was done to all MH-Z19B sensors on the first connection, due to high initial output values. To do this, sensors were connected to a microcontroller and placed outdoors. After exposing the sensors to 400ppm for 20 minutes (mins), the zero-point calibration was set by connecting the **Hd** pin to **GND** on each sensor for approx. 7s.

7.4.3 Reading PM, temperature, humidity and ambient light intensity data via I^2C

Four sensors communicated with the microcontroller via an I^2C bus, Table 11. This was chosen because it is a serial communication protocol that uses a two-wire interface: (i) the Serial Data *(SDA)* wire sends data across the bus, and the (ii) Serial Clock *(SCL)* wire synchronises communication between the master (microcontroller) and slaves (sensors) [256]. As the protocol requires two wires only to form the serial bus, it is optimal in microcontroller applications. Thus, it has become standardised for ARM microcontrollers [257]. Moreover, the ability to read from multiple sensors from a single two-wire bus makes this protocol extremely useful for multimodal devices [258].

Each I^2C slave communicates on its own unique I^2C address. Some sensor manufacturers develop sensors with multiple I^2C buses to allow more than one of the same slaves to communicate with an I^2C master. Addresses for CCS811 and BH1750 can be changed with a software modification, typically by specifying the address, when declaring a new instance of sensor within a code library. However, the BME280 requires a hardware modification to switch addresses. On the front of the breakout board there are three solder points (i.e. jumpers). If no jumpers are joined, *(or the two left-most jumpers are joined)* the device defaults to address 0x76. However, if the two right-most jumpers are soldered the address is changed to 0x77. I^2C multiplexors can also be used which have separate busses for communicating with sensors. Each bus works the same way as the I^2C bus on an MCU, in that multiple devices can be connected if there are no address conflicts. However, each bus of the multiplexor has its own access address and each bus is separated from one another [259]. This means that address conflicts can be resolved by connected devices with conflicts to separate busses. For this study, there were no I^2C address conflicts in the multimodal device. Consequently, multiplexors were not required and the default addresses for each sensor were used.

7.4.4 Reading noise data via I^2S

To calculate the Sound Pressure Level (SPL) in dB with a microphone, a logarithmic calculation is required:

$$dB = 20 \times log_{10}\left(\frac{S_{RMS}}{ref}\right) \tag{2}$$

Where S_{RMS} is the root mean squared of the samples captured by the microphone over a given sample period *(*e.g., *1000ms)* and *ref* is the peak amplitude of the microphone. To calculate S_{RMS}, Kostoski's library first applies equalisation and filtering to the samples and calculates the sum of squared, weighted samples:

$$S_w = \sum_{0}^{N} y^2 \tag{3}$$

Where N is the number of samples captured in the sample period and y is the samples after weighting and equalisation have been applied. The sum of samples is then used to calculate the S_{RMS}:

$$S_{RMS} = \sqrt{\frac{S_w}{N}} \tag{4}$$

To calculate the peak amplitude *(ref*, Eq. 1*)* the sensitivity of the microphone must first be calculated. For digital microphones, the sensitivity of the microphone should be pre-specified in the microphone's datasheet and is calculated as:

$$Sens_{dBFS} = dB_{REF} - dB_{MAX} \tag{5}$$

Where dB_{REF} is 94dB *(equivalent to 1 Pascal),* and dB_{MAX} is the maximum acoustic input for the given microphone. For the INMP441, the specified maximum acoustic input is 120dB (Table 11) so the resulting sensitivity is -26dBFS.

Once sensitivity is calculated, it is then used to calculate the peak amplitude of the microphone that is used as the *ref, Eq* (2). The INMP441 datasheet [246] specifies that the peak amplitude for this microphone is calculated as:

$$ref = \left(2^{(bitrate-1)} - 1\right) \times \left(10^{(dBFS/20)}\right) \tag{6}$$

Since the INMP441 transmits data via a 24-bit I^2S interface [246], the peak amplitude of the microphone therefore maps to 420,426 discreet digital values.

7.5 Methods

7.5.1 Data acquisition and connectivity

All data from low-cost sensors were read and processed by a HELTEC ESP32 Wi-Fi microcontroller (a dual-core microcontroller with Wi-Fi, Bluetooth, Bluetooth Low Energy (BLE) and an integrated Liquid Crystal Display (LCD) display). A reading for each sensor was collected every 15s and data were written to ThingSpeak®, which allows up to eight sensor readings *(per channel)* to be written to the Cloud simultaneously. The data/channel quota included with a free subscription was suitable to test the prototype and conduct validation of sensors in each device.

7.5.2 Data processing

The intervals between measurements, for each device (low-cost and reference), were determined from when the devices were initially configured to initiate logging. Therefore, it was not possible to synchronise the sample rate across devices. Consequently, there was a need to resample data extracted from the measurement instruments to ensure they were comparable. A series of steps were undertaken for data processing and analysis (Figure 13).

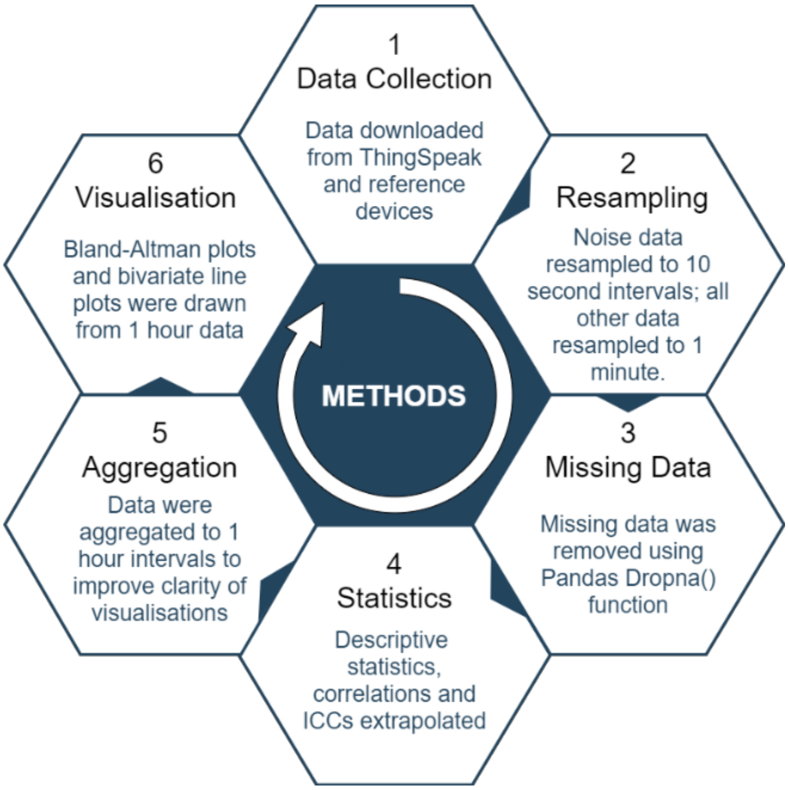

Figure 13 - Data processing methods

7.5.2.1 Resampling (excluding noise level data)

A sample rate of 1min was chosen for resampling to reduce the amount of interpolated data in the final dataset. Since the sample rate of the reference standard sensors was 5mins this meant up-sampling all reference standard sensors by interpolating the values for the missing intervals. This was done using the pre-processing package from the Sci-Kit Learn library for Python, using a linear interpolation method. Conversely, down-sampling was required to resample sensor data from low-cost sensors, which captured sensor readings with a 15s sample rate. Sci-Kit Learn was also used to resample those data, but the mean function was used on the data series to calculate the mean for each 1min period.

7.5.2.2 Missing data

Some data was missing due to connection interruptions. So, after resampling data, the rows which contained missing data were dropped using the Pandas' Dropna function so that a bivariate analysis could be performed on any two columns without conflicts. A dataset was created that contained a total of 15,828 samples. Those data were combined into a single Pandas DataFrame for processing. (Since there were many gaps in the HHSL-101 data, due to the three-day reconfiguration cycle, noise data from the HHSL-101 and INMP441 were not included here). Since rows with missing data in were dropped figures throughout this chapter can exhibit straight lines between measurements, akin to linear regression lines. However, this is simply the result of data between two points being removed.

7.5.2.3 Noise level data

INMP441 data were combined with HHSL-101 data separately using the same methods used for other sensors. Data were resampled with a 10s sample rate and the resulting dataset contained 47,602 samples that were also combined into a single DataFrame.

7.5.2.4 Analytical and statistical procedures

To analyse both data sets, Pandas was used in conjunction with Matplotlib, Seaborn, Sci-Kit Learn, Pingouin and the Statsmodel API libraries for Python. Pandas was used to provide descriptive statistics and to process datasets (resampling, missing data removal, Pearson correlation statistics and data aggregation). The Seaborn library was used in conjunction with Matplotlib to create plots and visualisations and, finally, Sci-Kit Learn, Pingouin and the Stats Models API performed agreement analysis (Intraclass Correlation Coefficients (ICCs), and Bland-Altman) on bivariate pairs.

Correlation and agreement (absolute and consistency) of each low-cost sensor was validated against reference devices. ICC estimates and their 95% confidence intervals were calculated using the Pingouin v0.3.9 [260] library for Python. ICCs were used to assess the reliability of sensor data taken from low-cost sensors against reference devices [261]. Predefined acceptance ratings for ICC were: excellent (>0.900), good (0.750–0.899), moderate (0.500–0.749) and poor (<0.500) [261]–[263].

The Two-Way Random-Effects Model, against a single rater ($ICC_{2,1}$), was used to determine the reliability of randomly chosen low-cost sensors [262]. In this model, the low-cost sensors were evaluated to test the absolute agreement between measurements from each low-cost sensor against the reference device/sensor. This model can be used to generalise findings and evaluate the potential reliability of other sensors (from the same manufacturer) [262]. For example, here I selected three low-cost BH1750 light sensors, and can use $ICC_{2,1}$ to determine the potential reliability of other BH1750 sensors.

The Two-Way Mixed-Effects Model, against a single rater ($ICC_{3,1}$), was used to assess the reliability of this specific sample of sensors [262]. In this model, the consistency of each low-cost sensor is evaluated against a reference. Since this model focuses specifically on the sampled low-cost sensors, the results cannot be generalised to other similar low-cost sensors, even if they share the same characteristics [262]. For example, the consistency of the three BH1750 sensors can be evaluated, but the results cannot be used to determine the consistency of other BH1750 sensors.

7.5.2.5 Data visualisation

To improve clarity of bivariate visualisations, datasets were resampled to a 1-hour frequency before plots were generated, Figure 13. Bivariate line plots and regression plots were generated to visually inspect data pairs. Finally, the Statsmodel API was used to calculate the mean difference statistics for each data pair. These mean differences were then plotted to Bland-Altman plots, using the Statsmodel Graphic Utilities, to quantify agreement between each data pair [264].

7.5.3 Sensor deployment

All devices (low-cost multimodal and reference) were placed next to each other on a *(165cm high)* shelf within an office, above a computer desk *(Figure 14).* The office *(located in Newcastle Upon Tyne, UK)* was a south-facing shared office, occupied by two people. Occupants had control over the windows, blinds and heating and the core operating hours were typically between 8am-6pm, Monday to Friday. The position of the shelf meant that no direct light from computer monitors was able to enter the light sensors, so light captured was a mixture of natural daylight and ambient artificial lighting from LED bulbs. Except for the reference sound level meter (Omega HHSL-101), all devices logged sensor readings continuously between 1st – 30th of November 2020. However, the storage capacity of the HHSL-101 meant that data were downloaded, and the sensors reconfigured every three days.

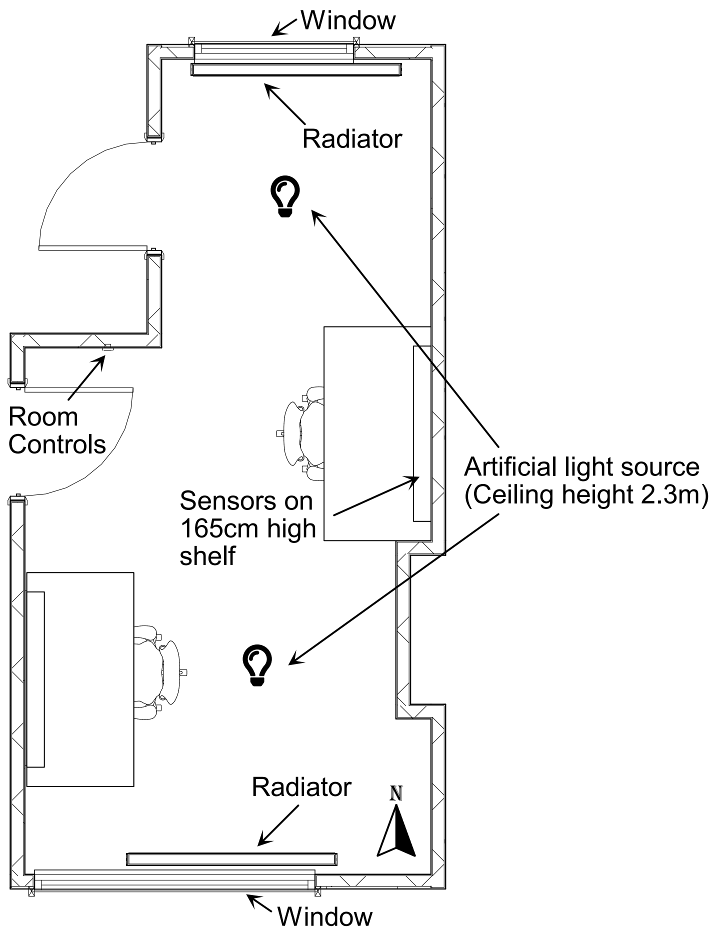

Figure 14 - Layout of office showing placement of windows, doors, artificial light sources and radiators.

7.5.4 Reference standard setup

All reference standard equipment stored data locally and data were downloaded at the end of the study via the proprietary interfaces. Both HOBO® devices recorded from each of their internal sensors every 5min and stored data internally. Data were then downloaded using the HOBOMobile app.

The AirVisual Pro provides a Server Message Block (SMB) interface. This allows a user to connect to the device, via an IP address, to obtain the stored data files, which are stored in .CSV format. The device has the capability of storing data in the Cloud that can be accessed using the IQAir dashboard. However, to obtain data from the Cloud, a paid subscription is required.

The HHSL-101 also requires use of an application (Omega's Sound DataLogger) to initialise, record and download data. However, this application is only available for Microsoft Windows (XP or greater). The sample rate for the HHSL-101 was set to 10s. From options available, that was the closest to the sample rate of the low-cost multimodal device. At the chosen sample rate, the HHSL-101 was able to capture data for three days only. Therefore, multiple sample periods from this device were conducted throughout November 2020.

7.6 Results

Due to the number of sensors being evaluated, it is not possible to provide a complete set of data and visualisations. Consequently, a summary of data is presented (Table 14) while data and visualisations used to inform the analysis are included in the online supplementary material.

Table 14 - Results of sensor validation study

Src	Dev†	Descriptive Statistics					P'son Corr	Agreement (Bland-Altman)			ICC§	
		Mean	Std	50%	75%	Max		Mean Dif	Std	LoA‡	$ICC_{2,1}$	$ICC_{3,1}$
eCO_2	ESP_A_eCO2	652.01	353.28	552.88	727.5	7899	0.38	-50.87	244.54	479.29	0.79	0.89
(ppm)	ESP_B_eCO2	932.83	549.25	769.33	1052.38	7992	0.38	268.57	474.5	930.01	*[0.49 - 0.9]*	*[0.88 0.89]*
CCS811	ESP_C_eCO2	964.95	520.06	779	1077.38	7632.75	0.36	311.39	456.52	894.79		
	Ref: MX1102	700.43	178.35	678.7	866.8	999	-	-	-	-		
CO_2	ESP_A_CO2	574.03	132.13	564	693	1260	0.97	-138.86	66	129.36	0.72	0.95
(ppm)	ESP_B_CO2	748.38	144.28	743	873.75	1485.67	0.95	39.22	64.64	126.7	*[0.1 - 0.9]*	*[0.95 - 0.96]*
MH-Z19	ESP_C_CO2	632.84	185.49	617.67	793.25	1611	0.96	-67.52	48.77	95.58		
	Ref: MX1102	700.43	178.35	678.7	866.8	999	-	-	-	-		
$PM_{2.5}$	ESP_A_PM25	4.92	17.19	0.00	2.33	414.67	0.23	3.37	20.15	39.50	0.99	0.99
(µg/m3)	ESP_B_PM25	5.49	17.89	0.33	3.00	405.00	0.24	3.99	20.56	40.30	*[0.97 - 0.99]*	*[0.99 - 0.99]*
	ESP_C_PM25	5.61	17.79	0.75	3.33	351.75	0.24	4.24	21.58	42.30		
PMSA003i	Ref: IQAIR AVP	3.75	10.81	1.50	3.00	340.40	-	-	-	-		
Temp	ESP_A_TEMP	24.86	0.73	24.80	25.51	26.63	0.97	3.64	0.16	0.31	0.75	0.98
(°C)	ESP_B_TEMP	25.50	0.87	25.47	26.29	27.44	0.94	4.30	0.29	0.57	*[0.08 - 0.91]*	*[0.98 - 0.99]*
	ESP_C_TEMP	24.59	0.79	24.57	25.29	26.39	0.96	3.39	0.21	0.40		
BME280	Ref: MX1104	21.22	0.70	21.20	21.83	22.88	-	-	-	-		
RH	ESP_A_RH	37.13	3.75	37.35	39.83	50.94	0.99	-10.84	1.14	2.24	0.89	1.00
(%)	ESP_B_RH	36.32	3.58	36.51	39.01	49.74	0.98	-11.67	1.42	1.42	*[0.19 - 0.97]*	*[1.0 - 1.0]*
BME280	ESP_B_RH	39.22	3.85	39.38	42.05	54.13	0.99	-8.78	1.10	2.16		
	Ref: MX1104	47.93	4.79	48.34	51.44	64.96	-	-	-	-		
Air Prs	ESP_A_PRS	992.14	11.81	992.69	1003.36	1011.06	1.00	-4.13	0.45	0.88	1.00	1.00
(hPa)	ESP_A_PRS	991.97	11.84	992.48	1003.20	1010.97	1.00	-4.30	0.43	0.85	*[0.94 - 1.0]*	*[1.0 - 1.0]*
BME280	ESP_A_PRS	992.95	11.81	993.51	1004.12	1011.90	1.00	-3.32	0.45	0.89		
	Ref: Weather API	996.25	12.03	996.37	1007.27	1015.19	-	-	-	-		
Light	ESP_A_LUX	11.07	23.79	0.00	15.00	294.00	0.95	-4.29	6.04	11.84	0.98	0.99
(lux)	ESP_C_LUX	13.39	23.97	3.00	18.00	300.50	0.95	-1.92	5.59	10.96	*[0.91 - 0.99]*	*[0.99 - 0.99]*
	ESP_C_LUX	8.15	21.86	0.00	9.00	275.25	0.93	-7.71	8.08	15.83		
BH1750	Ref: MX1104	14.45	25.88	4.31	22.04	269.84	-	-	-	-		
Noise	ESP_A_SOUND	45.17	8.86	42.49	51.57	92.01	0.51	-1.48	6.49	6.49	0.73	0.73
(dB SPL)	ESP_B_SOUND	46.63	8.21	44.00	52.17	84.49	0.49	0.04	6.65	13.04	*[0.68 - 0.77]*	*[0.69 - 0.78]*
	ESP_C_SOUND	45.07	9.33	42.76	52.13	85.66	0.51	-1.05	7.81	15.31		
INMP441	Ref: HHSL_101	45.42	5.82	43.20	47.30	102.90	-	-	-	-		

† **Device:** Ref = reference device, ESP_{device}_{Measure} labels refer to the labels used in the dataset to identify multimodal devices; ‡ **Limit of Agreement:** from Bland-Altman analysis; § **ICCs:** reported with 95% Confidence Intervals that are displayed as *[Lower – Upper]* bounds. The bounds define a range where there is a 1 in 20 chance the true mean should exist. Thus, wider ranges or a low upper-bound indicates a lower reliability.

7.6.1 Equivalent carbon dioxide (eCO_2)

The MOX eCO_2 sensors had a poor correlation with the reference (≤0.38) and divergence between measurements can be seen across all percentiles (Table 14). When approaching the upper limits of measurements, the measured values are approx. 7000ppm greater than the reference CO_2. There is also little commonality across the mean values and the standard deviations are significantly large. Despite the significant difference in reported values, the MOx sensor data did mostly rise and fall at the same time as the CO_2 reference sensor (Figure 15). However, there were noticeable spikes in the data, which often moved in the opposite direction to CO2 sensor data. This could likely be caused by differences in the measurement technology, but it is likely the result of unmeasured influences, given that eCO2 sensors are highly sensitive to a wide range of outcomes [204].

$ICC_{2,1}$ were good (0.79) with broad confidence intervals (0.41 difference between lower and upper bounds). $ICC_{3,1}$ were also good (0.89) but with much narrower confidence intervals (0.01 difference). However, eCO2 sensors had significantly large mean differences against reference CO2 sensors and had and had a larger LoA *(between 350-900ppm)* than the NDIR CO_2 sensors.

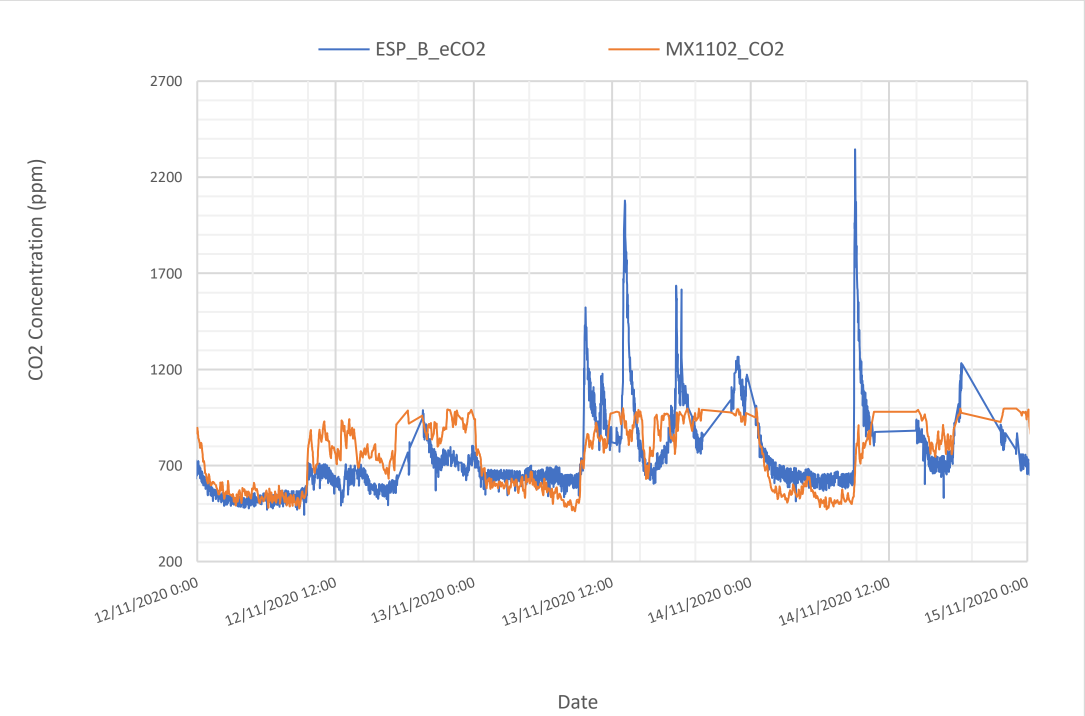

Figure 15 – Snapshot of eCO_2 vs CO_2 events captured by CCS811 (blue) and MX1102 (orange)

7.6.2 Carbon dioxide (CO_2)

CO_2 measurements across all low-cost sensors were found to strongly correlate with reference data (≥0.95). Low-cost CO_2 sensors appear to generally agree across the lower percentiles but diverge above the 75th percentile. NDIR CO_2 sensor data were closer to the reference data than eCO_2 sensors and had a similar precision to the reference (Figure 16).

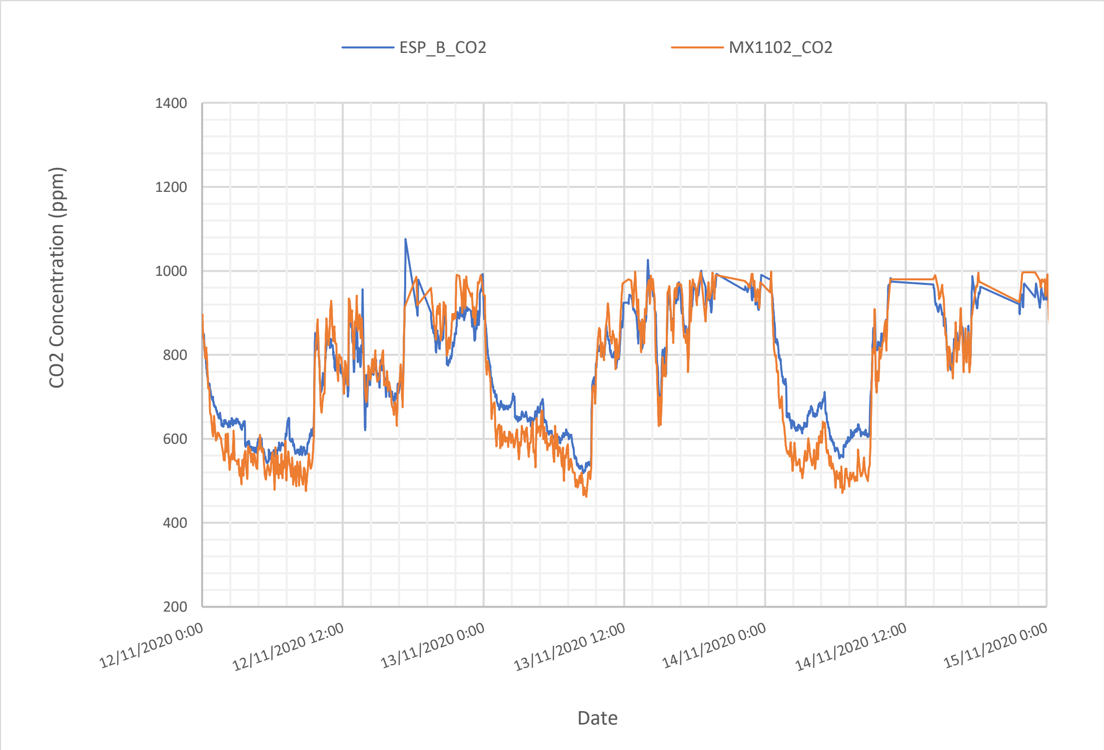

Figure 16 – Snapshot of MH-Z19B CO_2 sensor (blue) vs reference MX1102 CO_2 sensor (orange)

$ICC_{2,1}$ were moderate (0.72), but the confidence intervals were broad (0.8 difference between lower and upper bounds) indicating that other sensors of the same manufacturer and model may have a higher variability than the sensors sampled here. Contrastingly, $ICC_{3,1}$ were excellent (0.95), with a high and narrow confidence interval (0.1 difference between lower and upper bounds). However, the Limit of Agreements (LoAs) were large across all low-cost devices. That notwithstanding, all CO_2 sensors *(including the reference)* have a high standard deviation *(between 132 - 186),* which may have impacted the LoA.

7.6.3 Particulate matter

Low-cost $PM_{2.5}$ were found to have poor correlations (≤0.38) to the reference standard. However, there was little divergence across the percentiles. Despite the low Pearson correlations, the low-cost sensors had excellent ICCs for both $ICC_{2,1}$ and $ICC_{3,1}$ (both 0.99). There was also very narrow range between the lower and upper bounds of the confidence intervals (>0.02). This indicates that that there is good inter-sensor reliability.

There was a significantly low mean difference *(between 2 – 4),* but relatively high LoAs were seen across the low-cost sensors *(≥35µg/m^3 more than the reference mean)*. However, the Bland-Altman analysis (Figure 17) highlights that there is a significantly strong agreement between the low-cost and reference sensors during events where dust concentrations are lower. As the dust concentrations increase, the agreement between sensors is reduced, which is likely causing the low correlations. While low-cost $PM_{2.5}$ monitoring devices are considered effective indicators of $PM_{2.5}$ events [222], it is possible that the optical technology used in low-cost $PM_{2.5}$ is not suitable for accurately determining higher concentrations of particulate matter. However, for longitudinal monitoring of IEQ, optical sensors could provide a fit-for-purpose indicator of $PM_{2.5}$ events despite not being able to accurately report values when concentrations exceed around 10-15µg/m^3. The reduction of agreement at higher concentrations is likely a major contributing factor to the low correlation between the low-cost sensor and the reference device.

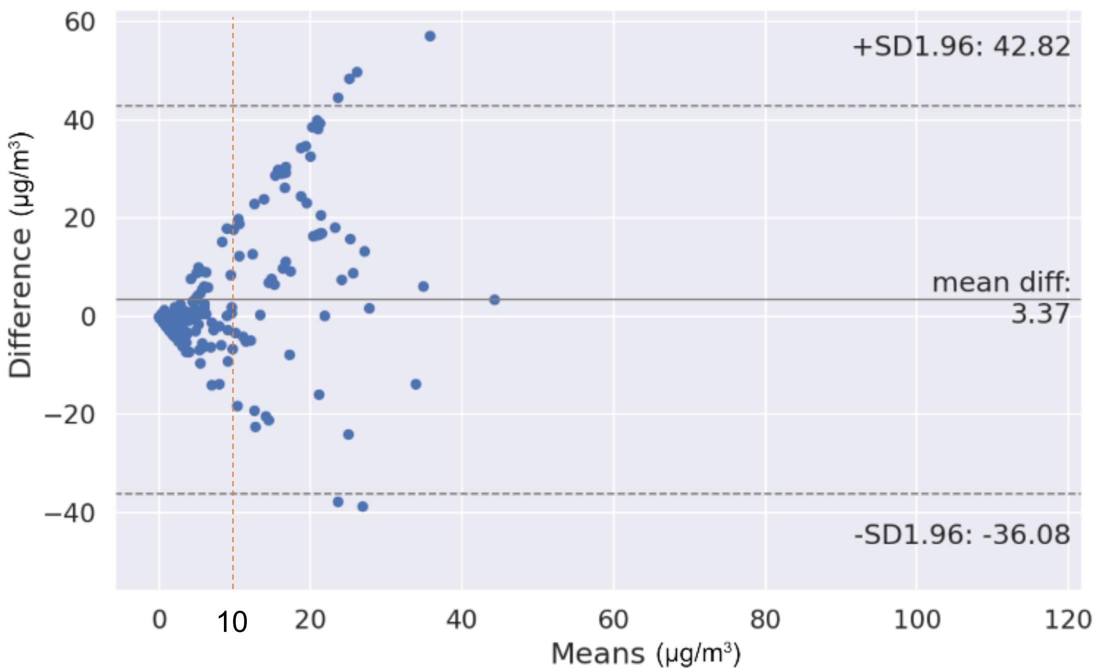

Figure 17 – Bland-Altman plot of PMSA003i means against the Air Visual Pro

7.6.4 Temperature, humidity (and air pressure)

The multimodal sensor for measuring temperature, relative humidity and air pressure performed well across all measurement factors and were found to have significantly high correlations with the selected reference (≥0.94). While the mean values measured by low-cost sensors did not exactly match the reference means, the values did not diverge through the percentiles and measured with consistent precision *(around 3-4°C mean difference, Figure 18), which is likely caused by heat from the internal circuitry.*

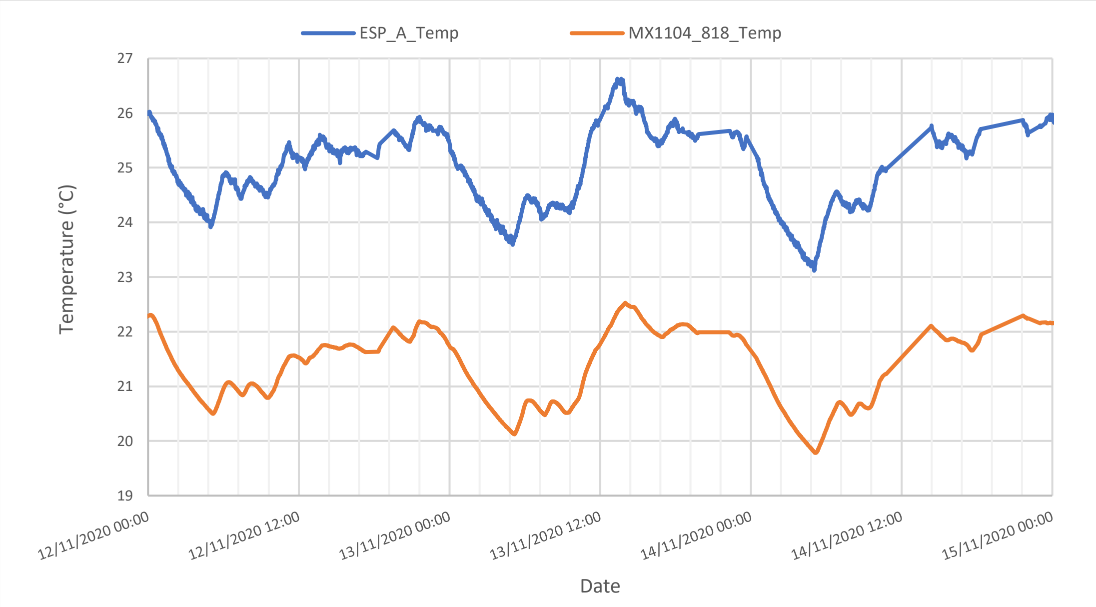

Figure 18 – Comparison of low-cost BME280 temperature sensor (blue) against MX1104 reference (orange).

Like CO_2, the $ICC_{2,1}$ for temperature were good (0.75), but with broad confidence intervals (0.83 difference between lower and upper bounds). However, $ICC_{3,1}$ values indicate excellent reliability among the sample (0.9), with 0.01 difference between the confidence bounds. The BME280 also had the lowest observed LoAs (≤2.3) across all the evaluated sensors.

7.6.5 Light

Light sensors also performed well against the reference. Low-cost sensors performed consistently across the percentiles and had significant correlations (≥0.93). Mean values and standard deviations agreed with the reference device. Although mean values for light-intensity were lower than expected, it is noted that office users had control over the blinds so were able to mitigate glare (and as such control light intensity) from south-facing windows.

Light sensors were excellent for both $ICC_{2,1}$ (0.98) and $ICC_{3,1}$ (0.99), with high, narrow confidence intervals for both, (≥0.90 with ≤0.08 difference across the lower and upper bounds). This suggests that other sensors of the same manufacturer and model should have similarly low variability across devices.

7.6.6 Noise

The correlation of noise sensors was approx. 0.50, mildly significant. However, the descriptive statistics agreed with the reference device. LoAs were also found to be relatively low (≤15.5dB) and the ICCs for both $ICC_{2,1}$ and $ICC_{3,1}$ showed moderate reliability (0.73), with only 0.09 difference between the lower and upper bounds for both ICCs. This indicates good inter-sensor reliability. The sensors also performed consistently across all percentiles showing the lowest recorded mean difference against the reference ($\leq$1.48).

7.7 Discussions and Conclusions

The study within this chapter proposed and developed a multimodal IEQ monitoring device to address identified research gaps surrounding scalable and individualised IEQ monitoring. A background of related work was presented in previous chapters that outlined underlying technologies that could be used to develop a low-cost IEQ monitoring device and the cumulation of which was this chapter. Specifically, it built upon that work to identify a range of specific sensors and technologies that were used to construct and configure a low-cost, IoT enabled device. An analytical methodology was proposed and presented to compare variations across low-cost sensors and to reference standards.

7.7.1 MEMS sensor selection

Use of I^2C MEMS sensors meant that fewer General-Purpose Inputs/Outputs (GPIOs) were required on the microcontroller. As a result, smaller microcontrollers with less GPIOs could be used. Moreover, selected MEMS sensors had integrated ADCs negating the requirement for a microcontroller with a high-resolution ADC, which can be more costly [173]. Also, using a digital I^2S MEMS microphone meant that the proximity with WiFi and Bluetooth components *(required by a small form-factor, multimodal, IoT device)* would not interfere with the microphone's performance.

Use of MEMS sensors did identify certain drawbacks. For example, many MEMS sensors *(e.g., CCS811, BH1750, BME280)* have built-in controllers that enable the ADC conversions and signal processing. However, this can often result in a sensor that performs 'hidden', black boxed algorithms on data, which can be problematic in research settings where transparency is key. Understanding, *e.g.,* how the CCS811 calculates TVOC and eCO_2 could be beneficial to researchers and help to better assess calculation validity or to develop algorithms against the unreported raw data.

7.7.2 Sensor performance

$ICC_{2,1}$ indicated good-to-excellent reliability for all sensors. However, the range between lower and upper bounds of the 95% confidence intervals was as much greater in some cases *(CO_2: 0.9; Temperature: 0.83).* This indicates that the sensors sampled here cannot be used to determine the reliability of other sensors with the same characteristics.

Most sensors however *(MH-Z19, PMSA003i, BME280, BH1750)* had excellent $ICC_{3,1}$ ($\geq$0.95) with a 95% confidence range of 0.01 between the upper and lower bounds. This indicates that this sample of sensors had excellent inter-sensor reliability. The sound and eCO_2 sensors did not perform as well, but $ICC_{3,1}$ was still good (0.89) with a 95% confidence range of 0.01 between the upper and lower bounds.

7.7.3 Accuracy and precision

This study confirmed findings [176] that low-cost sensors can often have high precision, but with a reduced accuracy. A clear misalignment between the two datasets (i.e., low accuracy) was found, but the low-cost sensor responded with the same precision as the reference. This was not the case for all sensors, but CO2, Temperature, Relative Humidity, Air Pressure and Light all had high precision.

Due to the high precision, it is possible to calibrate the devices against the mean difference between the low-cost sensor and the reference. However, since light intensity dropped to 0lx at night, adjusting against the mean difference alone would result in negative values, so measurement ranges need to be considered when calibrating sensors.

7.7.4 Ventilation sensors for air quality

Here, MOx eCO_2 sensor data was erratic, when compared against actual CO_2 measurements. Reported values from the MOx sensors did mostly rise and fall at the same time as the CO_2 reference sensor, but the reported values for eCO_2 were often far greater than the reference. Since MOx sensors are highly sensitive to a range of environmental factors [204], It is possible that these sensors are responding to the accumulation of unventilated air *(as opposed to actual CO_2).* In this way, they may be useful as proxies for ventilation measurement in the same way CO_2 sensors are used [38], [57].

More research is required to assess whether eCO_2/TVOC sensors are fit-for-purpose in ventilation monitoring. However, they could provide an affordable and scalable solution *(compared to NDIR CO_2 sensors)* to address the growing need for ventilation monitoring brought on by the SARS-COV-2 pandemic. Therefore, there are strong practical implications for identifying eCO_2/TVOC sensors as proxies for ventilation.

If eCO_2 sensors are found to be suitable for indoor ventilation monitoring, it would be preferable to report data from these sensors with more appropriate terminology, as the eCO_2 terminology implies the measurement is related to carbon dioxide, which cannot be confirmed by this study.

7.7.5 Cloud connectivity

The use of ThingSpeak® in this study was beneficial as it provided an IoT platform to capture low-resolution data in real-time, without needing to connect to devices or download data from internal storages. This was advantageous as it meant a sample could be downloaded during the data collection phase and used to develop code needed for analysis, reducing the workload at the end of the project. ThingSpeak® also allowed for real-time monitoring of the data, providing graphs and visualisations. While the platform was previously deemed currently unsuitable for real-time medical monitoring [173], this study confirmed that it is suitable for prototyping/small scale IEQ monitoring projects. The platform is also scalable, adopting a pay-as-you-grow model. However, the quota strategies used by larger Cloud computing platforms (e.g., Amazon Web Services) may be better suited for larger projects as resources can be shared and distributed amongst multiple devices more easily [173]. It is also worth noting that the IoT cloud platform market is rapidly growing and there are more than 600 dedicated platforms available, each designed with nuanced use-cases [152], [173]. It is recognised here that while ThingSpeak® was suitable for this study, the quotas and limitations it imposes may create a requirement for researchers to conduct an analysis of the IoT platform market to assess available platforms.

7.7.6 Limitations

Only three low-cost multimodal devices were examined in this study. With more resources, more devices could have been examined, which would have provided a greater sample size for ICC analyses. While three sensors are enough to conduct a reliability study, a greater number of sensors would reduce the potential lack of variability between sensors, which may impact ICC estimations [262].

All devices were connected to a premium residential network package (Virgin Media, Reading, UK). However, on occasions during the study period the internet was heavily interrupted, which caused the devices to disconnect and resulted in lost data. Regardless, there was ample data to conduct an informed analysis and comparison of all sensors.

The measurement ranges captured by sensors in this study were measured under normal operating conditions so there was no necessity to test the upper/lower limits of the sensors (i.e., at extreme conditions). Therefore, it was not possible to evaluate the LoA for sensors in those ranges. Nevertheless, Bland-Altman analyses did highlight *(for sensors that approached upper office comfort limits)* that LoAs did diverge as values increased (Figure 17). Researchers wishing to use these sensors under more extreme conditions should be mindful of this and further evaluate the sensors under the desired measurement conditions.

7.7.7 Suitability

Sensors used in this study would be suitable for continuous monitoring to provide building occupants with an indication of environmental quality and changes. Despite inaccuracies in certain sensors, the high precision witnessed means that in most cases the sensors can be calibrated easily against the mean differences recorded in this study. However, the $ICC_{2,1}$ showed that the variability of sensors seen in this study may not be representative of other sensors with the same make/model. Therefore, with the findings from this study alone, it would not be possible to provide a general calibration offset that would be applicable for other sensors.

In certain cases, it may be preferable to report data from sensors as a red/amber/green system instead of using the numerical output. For example, eCO2 was found to be unsuitable as a measurement of carbon dioxide but showed potential as a proxy measure for ventilation. While further research would be needed to confirm this, reporting data as Parts Per Million, the unit of measurement for CO_2, may not be applicable for this application.

Given the potential accuracy biases found in this study, it would not be possible to ensure the scientific validity of the sensors for use in applications such as occupational health assessments or standard compliance. However, the intended use-case for the devices proposed here is to provide building occupants with a general indication of environmental changes. For this application, this study found evidence that the specific sensors sampled in this study are fit-for-purpose. Consequently, the multimodal device developed here could provide a viable solution for localised, continuous monitoring that could be pragmatically deployed at scale.

7.8 Contribution to knowledge

The primary purpose of this chapter was to develop, and evaluate the development of, a multimodal device that was sufficiently low in cost that it could be deployed at scale for monitoring IEQ at an individual level. The reviews and experimentation across previous chapters have provided a broad yet focussed lens on the technologies, challenges and knowledge gaps that have informed the development of this chapter and the multimodal sensor that was developed here. To address knowledge gaps surrounding high-level detailing of sensor technologies within building studies (identified in Chapter 2), this chapter presented the development of the multimodal device with a much greater level of detail than is typically seen across the literature. This demonstrated the complexity and technological requirements of building such a device, highlighting the need for MDRTs, which include computer science/electrical engineering disciplines, within building science research. The presentation of sensor technology at a low level, may lead to the false expectations.

This chapter also took a novel approach to appraising the developed device, by using biomedical analysis methods (Bland Altman agreement analysis and ICCs) to validate the individual sensors and the combined multimodal device. In doing so, this chapter presented demonstrable value from these approaches. Sensor technologies are typically evaluated to assess the precise scientific accuracy, measured under strict laboratory conditions, which may attribute to the general lack of acceptance towards low-cost sensor technologies as their low cost can often lead to inaccuracies, reading drifts, and cross parametric influences (such as temperature/humidity effects on readings). It is possible that some of the adopted sensors could have failed traditional methods of sensor validation. However, by taking the approach to measure accuracy, precision, absolute agreement, consistency, and reliability separately, it was possible to identify use-case dependant value in the evaluated sensors. For example, the sensors validated in this study were deemed fit-for-purpose within the context of this thesis. This is because in real-world buildings it is more important to have general indicators of environmental conditions, measured at scale, than to have no measurements at all, due to the cost and complexity of gold-standard monitoring equipment. In scenarios where measuring accuracy to the n^{th} decimal place is imperative, some of these sensors would be unsuitable. However, the high precision and inter-sensor reliability witnessed in these devices means that they could be suitable within the context of this thesis, especially considering buildings are traditionally only monitored by evaluating the subjective responses of individuals.

7.9 Addressing the PoI

PoI5:

Can low-cost sensors be used to gather with precision, consistency, and reliability across different environments i.e., home, and commercial settings?

Depending on the purpose of the measurements, yes, low-cost sensors can be used to gather with precision, consistency, and reliability across different environments i.e., home, and commercial settings. However, low-cost sensors have recognised limitations, which may make them unsuitable for certain measurement scenarios. If highly accurate measurements *(to 'nth' decimal place)* are required, it was demonstrated that the low-cost sensor technologies would likely be unsuitable. However, for the purposes of providing a general indicator of environmental conditions *(for the purposes of continuous monitoring solutions),* low-cost sensors were shown to have pragmatic use. Moreover, low-cost sensors can demonstrate good inter-sensor reliability, meaning that they could be deployed across different environments and the measurements of two sensors could be compared with one another.

7.10 Future Work

The affordability and multimodality of the device proposed here identifies a scalable solution for occupant monitoring that can provide a guidance around IEQ in buildings. Consequently, this Ph.D. will exploit the affordability of these sensors to deploy multimodal devices for longer periods of time and address the needs for localised monitoring of building occupants, by deploying sensors at an individual level.

The following two chapters *(8 & 9)* will explore and answer **PoI6** to understand how multiple sensing modalities be pragmatically deployed for a more seamless integration to personalised monitoring. To do this, they will present the design of a general study protocol that will outline some of the methods and methodologies that will be adopted in the major study of this Ph.D. For clarity, Chapter 8 will present a protocol that will serve as a framework for future studies in this area. However, the protocol is generalised for widespread dissemination and considers the need for flexibility pertaining to bespoke research requirements such as participant considerations (discussed in more detail later). Subsequently, Chapter 9 showcases how the protocol can be modified and as such is described how it was deployed for this Ph.D. By presenting the methods in this format, future researchers can apply the methods set out in Chapter 8 to their own research questions and deploy them following Chapter 9 as an example.

7.11 Data

Box 7.1

Additional data and visualisations used in this study are available via online supplementary material linked to the original publication *(details on title page of this chapter).* Due to copyright, it is not possible to include those supplementary materials within this thesis.

Chapter 8 A protocol for longitudinal monitoring of individual building occupants and their environments.

This chapter is adapted from previously published work to fit the context of this thesis. The article: **A protocol for longitudinal monitoring of individual building occupants and their environments**, was published in **PLOS ONE** on **23 September 2022.**
This work was made available online on **23 September 2022** via:
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0274015

8.1 Introduction

The previous chapters have identified a need for low-cost (scalable) monitoring that have an individualised approach. The use of low-cost sensors and wearable health technologies has been proposed and tested through a series of case studies.

Chapter 2 identified a need for longitudinal and objective monitoring of environmental outcomes to capture how environmental outcomes change over time as longitudinal data can provide a holistic picture that exposes timeseries trends. However, the review also found many studies utilised complex and expensive, research-grade measurement devices and as a result capture periods were often short snapshots as opposed to longitudinal assessments. That notwithstanding, comprehensive IEQ monitoring requires the measurement of a wide range of outcomes, comprising of subjective comfort factors *(e.g., noise, privacy, or control over the environment)* and objective environmental conditions *(e.g., the thermal environment or indoor air quality)* [5]. This often results in short study periods where devices are placed at a single location *(in multi-occupant spaces)*, resulting in coarse-grained spatial data. However, IEQ can vary in buildings, floors, and rooms in a non-uniform manner. Therefore, the distance between an individual building occupant and a monitoring device can lead to inconsistencies between the conditions measured and the conditions experienced by the individual. Furthermore, approaches used to capture occupant experiences often generalise the views of multiple occupants *(based on common demographics, activities, locations etc.)* [265]. However, while occupants could be of the same demographic or sharing the same space, they could be experiencing different IEQ conditions, or different experiences of the same IEQ conditions.

Study methodologies can also impact the symmetry between comfort factors and environmental conditions. For example, Andargie *et al*. [1] note that studies were found to either measure IEQ at a different time to when surveys were conducted or that objective and subjective outcomes were measured in isolation with assumptions being made about the unmeasured outcomes. Thus, there is a need for research that conducts measurement of environmental conditions and comfort factors simultaneously, while measuring IEQ at an individual level. However, research grade measurement devices lack the scalability required to conduct such research and advocating such technologies on projects *(at scale)* can be challenging [7]. As identified in chapter 2, low-cost monitoring is seen to be a requirement for future research, where the technology could be used to develop scalable solutions that are able to monitor buildings at a more granular level, providing individualised monitoring.

Subsequent chapters have explored the feasibility of low-cost monitoring solutions in research by increasing the accessibility of sensor technologies and reducing development costs [173]. Accessible low-cost sensors enable the development of multimodal monitoring solutions that can be tailored to the specific needs of a project [266]. Consequently, IoT technologies can be regarded as a development toolkit for researchers to create more engaging and reactive research [173]. Since these devices can be made at low-cost with multimodality, they could be deployed at scale and left in-situ for continuous synchronous monitoring of individuals and their immediate environments.

Here, there is now a need to define how those IoT and sensor-based technologies could be deployed to understand causal relationships between the environment and occupant. Accordingly, this thesis proposes use of a longitudinal single participant observational methodology. Specifically, this type of study focuses on an individual as the unit of analysis, which can be useful for identifying causal relationships between measured variables [177] and can identify time-differential phenomena that relate to those causal links [267]. Due to the longitudinal and observational nature of those studies, it is possible to observe naturally occurring phenomena over time [267].

8.1.1 A protocol for personalised multimodal IEQ

This chapter will begin to explore and broadly answer **PoI6** to rationalise how multiple sensing modalities can be pragmatically deployed to gather data for the longitudinal assessment of individuals building occupants. This chapter presents a generalised protocol that will serve as a framework for future researchers. This protocol will present methodologies, technologies and workflows that will be used to collect perceptions of an individual's environmental conditions within the spaces they occupy as well as objective measurements of an individual's immediate IEQ. Since the purpose of this protocol is to serve as a framework for future researchers, who may conduct their research several months/years after this publication, specific low-cost IEQ sensors will not be explicitly detailed in this protocol. This is due to the rapid deprecation of low-cost sensor technologies that was discussed in Chapter 3. Instead, this protocol will specify the environmental outcomes that should be monitored, with the intention that future researchers can identify, appraise and validate suitable sensors/devices that can address these outcomes at the point when they conduct their studies.

Since this protocol will present a methodological approach for collecting and comparing an individual's perception of environmental conditions against objective measurements of IEQ, the unit of analysis will not be the building, but rather individual responses to changes within a building. Thus, multiple environmental settings could be monitored to understand how occupants respond to environmental changes and change of environment. It is hoped that the methodologies outlined in this protocol could help identify individual levels of comfort by collecting data from multiple sources to gain new insights on the impact of environmental changes on individuals. There are many factors that can be used to determine comfort, but for the purposes of this protocol it will be defined as the participant's general satisfaction with the extrinsic environment and intrinsic personal physiological conditions.

8.2 Methods

8.2.1 Study design

In this study, a mixed methods approach will be adopted to explore the interactions and relationships between an individual's physiological and behavioural responses to environmental changes. Additionally, a comparative analysis of (traditional) survey data and sensor data will be adopted to measure the relationship between perceptions of environmental conditions and objective environmental measurements (from passive environmental sensors). This approach will enable the exploration of **PoI7** to see whether localised sensors with multiple sensing modalities *(deployed longitudinally)* can be used to address the subjectivity around environmental perceptions by providing quantitative context to how occupants experience building environments. The protocol described here will quantify regular physical activity levels (e.g., walking) and heart rate alongside environmental factors with the aim to identify causal relationships between sedentary and active behaviour during environmental changes (Figure 19). Given the relationships between objective measurements and perceived IEQ, it is possible that the individualised, mixed methods approach could also be used to identify relationships between perceptions and measurements. This has the potential to provide additional context to how occupants experience indoor environmental conditions.

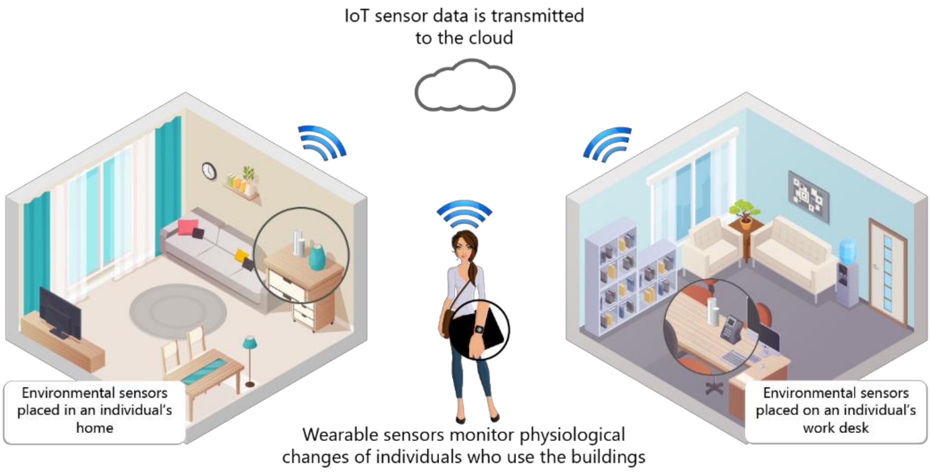

Figure 19 - Diagram of study setting and passive sensor configuration.

8.2.2 Study setting

This protocol is designed to simultaneously monitor different environments with passive environmental sensors and focus on the individual to pose a unique opportunity for IEQ research. This could involve monitoring multiple spaces within a single building that an individual would occupy *(e.g., occupant's office, meeting rooms or breakout areas)*, or it could involve monitoring multiple buildings that an individual would occupy *(e.g., home and office).* This protocol could be applied to either scenario. To mitigate the risks SARS-COV-2 could have on the study, this protocol will be applied to monitoring multiple environments within a home setting. Consequently, a participant is required that works from home with a home office that has distinct separation from their living area.

8.2.3 Eligibility Criteria

There are no inclusion/exclusion criteria pertaining to participant health for this study, but the participant will be of working age and will predominantly work from their office. The participant should be willing to wear a fitness tracker, which will monitor physical activity levels and heart rate during the working day *(09:00-17:00)* throughout the study period. The participant should also be willing to have sensors installed in their home and office, which will passively monitor IEQ and transmit data to the cloud *(Figure 19)*.

8.2.4 Sample Size

This is an observational study that will focus on drawing conclusions about an individual over time with the aim of exposing causal relationships between measurement outcomes. Therefore, only one participant is required for this study, but that participant will be monitored longitudinally to ensure statistical power to detect relationships between variables comes from repeated measurement of those variables over time.

8.2.5 Participant Timeline

This protocol outlines the design of a 16-week longitudinal observational study that will use the individual as the unit of analysis. The exact timeline for the participant will also include three additional days that will be used for an initial meeting with the participant as well as one day on either side of the study period for mobilising and demobilising the study.

8.2.6 Initial Meeting

As alluded to at the end of Chapter 7, there is a need for protocols to be flexible. Specifically, due to the longitudinal nature of the study, requiring long-term, active, participant involvement, it is important to negotiate and personalise the protocol framework to suit individual participant(s). This is typically done in the form of a pre-study interview, conducted to tailor data capture to meet the expectations of the participant *(due to their in-depth, longitudinal involvement in the study)* [177], [268] or to personalise healthcare interventions to individual patients [269].

The participant will receive full details of the study and provide written informed consent prior to the study commencing. The participant will be given the opportunity to discuss and negotiate the pragmatics of the study procedures, within the boundaries of an approved ethics application. The purpose of this was to offer the participant a sense of ownership of the study procedures so they could tailor the procedures to reduce the burden their day-to-day activities and ensure to ensure participant adherence.

The participant will receive an Apple Watch and two identical sets of environmental sensors. One set of sensors will be placed near the participant's office workstation and the other within the living room of their home. Since the study will monitor the participant's home, sensors will be connected on a dedicated 4G mobile network, to ensure no security vulnerabilities are exposed to their personal network.

8.2.7 Post study interview and survey

At the end of the study period an interview will be conducted where the participant will be asked a series of questions from a modified version of the ASHRAE Standard 55 Thermal Environment Satisfaction Survey (TESS) [61]. Modifications were made to that survey so that questions could be asked about the outcomes being measured in this study but framed so they followed the same format temperature is investigated in the ASHRAE Standard (Table 15). In line with the TESS, the participant was also asked when each case was most prevalent *(morning, mid-day, afternoon, evening, night).*

Table 15 - Format of closeout study question responses

Temperature	Humidity	Light	Sound
always too hot	always too humid	always too light	always too noisy
often too hot	often too humid	often too light	often too noisy
occasionally too hot	occasionally too humid	occasionally too light	occasionally too noisy
occasionally too cold	occasionally too dry	occasionally too light	occasionally too quiet
often too cold	often too dry	often too light	often too quiet
always too cold	always too dry	always too light	always too quiet

After the survey, an open-ended interview will be conducted to gather additional retrospective, subjective and self-reported feedback on the environmental conditions during the study period.

8.3 Outcomes

8.3.1 Primary Outcomes

The primary outcome for this study will be the physiological and behavioural responses to environmental changes (Table 16).

Table 16 – Dependant variables of the primary outcomes

Outcome	Source	Sample Period
Movement (Steps)	Apple Watch	After event
Heart Rate	Apple Watch	After event

‡ AppleWatch does not record at a fixed sample rate but instead records aggregated data retrospectively after event bouts.

8.3.2 Predictors of the primary outcome

This protocol will also explore a series of secondary outcomes (Table 17). These outcomes will provide objective measurements for each of the environmental outcomes, which will be obtained from validated and passive, multimodal, environmental sensors [266]. The secondary outcomes will also explore physiological and behavioural changes, which will be obtained from the Apple Watch.

Table 17 – Covariates to predict the primary outcomes

Outcome	Source	Sample Period
Temperature	Passive IEQ Sensor	1 minute
Humidity	Passive IEQ Sensor	1 minute
Air Pressure	Passive IEQ Sensor	1 minute
Light	Passive IEQ Sensor	1 minute
Noise	Passive IEQ Sensor	1 minute
Carbon Dioxide (CO_2)	Passive IEQ Sensor	1 minute
Particulate Matter up to 2.5μm in diameter ($PM_{2.5}$)	Passive IEQ Sensor	1 minute
Local Weather	OpenWeatherMap API †	1 hour
Outdoor Air Pollution	OpenWeatherMap API †	1 hour

† Historical data will be captured retrospectively from the OpenWeatherMap API, the API supports real-time connections and is therefore scalable, but this is not required for this study;

8.4 Data collection and management

Quantitative environmental data will be obtained from passive sensors in-situ and physiological data will be captured from a wearable wrist-worn fitness tracker. To adhere to recommendations on study length, data will be collected for a total of 16 weeks to ensure it is sufficiently longitudinal for individualised timeseries analysis [177]. Data on perceptions of environmental conditions will be captured using short, real-time surveys. Sensor data from this study will be used to investigate the data captured from surveys during analysis.

8.4.1 Data Collection Methods

8.4.1.1 Measuring physiological and behavioural responses

An Apple Watch Series 3 will be used to capture quantitative physiological outcomes, which will be stored on the participant's phone during the study via the iOS Health app. These data will also be combined with passive environmental sensor data to monitor heart rate and activity levels in relation to environmental changes. The wearable will be used to examine if physiological changes correlate with environmental changes and to assess whether environmental changes impact activity levels. The wearable will also monitor physiological responses to comparable environmental quality captured from different environments.

8.4.1.2 Measuring IEQ changes

IEQ sensors will need to be placed near the study participant, ensuring the participant’s perceptions of IEQ are reflective of the data being captured. For this study, it is important to ensure any identified solutions can be pragmatically deployed outside of research. Therefore, low-cost IoT sensors (identified as fit-for-purpose in a previous study [266], Chapter 7) will be used to measure IEQ in both environments.

IoT sensors should be considered a toolkit to develop bespoke multimodal monitoring equipment that is tailored to researcher requirements. As a result, the specific sensors used to measure the IEQ outcomes (Table 17) will likely be project specific. For example, monitoring measured/perceived noise levels would not be required if the participant were deaf. The same would apply to light levels for blind participants. Therefore, appraising and identifying sensors *(or discussing how they should be configured)* is deemed outside the scope of this protocol. However, Chapter 7 outlined this process and presented a bespoke multimodal device that was developed using low-cost sensors and IoT technologies. This multimodal device will be used to capture all the outcomes listed in Table 17 to provide a quantitative measurement of indoor environmental quality within the home and home office. Though, researchers should note that the IoT market is rapidly developing [173], so they should conduct a review of the market prior to selecting sensors for their study. Thus, the specific sensors have not been detailed here, as the study outcomes are deemed more important than the specific sensors used to capture the data. Please see Chapter 7 for details of the sensors deployed in this study.

8.4.1.3 Measuring IEQ perceptions

Perceived environmental quality and comfort assessment will be captured using surveys informed by traditional pen and paper-based approaches. However, surveys can often be burdensome, generally asking occupants to rate perceptions against scales, such as *e.g., temperature from "Cold" to "Hot"* [270]. Ratings are typically based on subjective responses that will likely differ between occupants, meaning data could only be used for generalising the perceptions of a population. Qualitative scale-based responses may also lead to situations where occupants are unable to accurately determine the difference between *e.g., "Slightly Warm"* and *"Warm",* meaning test-retest reliability could be affected. By complimenting survey data with sensor data, it could be possible to achieve greater meaning from comfort assessments, whereby sensor data could be used to reinforce data pertaining to perceived IEQ [32]. If the environmental conditions are known to the researcher *(from passive sensors),* the mechanism for capturing perceptions may be altered to reduce the number of subjective responses required by participants. This could reduce survey burden and allow for more prolonged/longitudinal feedback to be captured.

Here, an alternative solution is proposed that exploits the reinforcement of multimodal IEQ outcomes. Instead of asking occupants to rate the environmental factors using a scale, occupants will instead be asked if they are *e.g., "too cold"*, *"too hot"* or *"comfortable"*. In isolation, these data are highly subjective and less meaningful than traditional scales. Yet, the combination of these data with quantitative data from environmental sensors mean that this approach could help identify an individual's perception of *e.g.,* hot and cold, without burdening the occupant with multi-page surveys.

Surveys will investigate the participant's perceptions of IEQ. To ensure data from surveys can be combined with passive sensor data, questions will be related to the specific measurements being captured by IEQ sensors. Since this study is intended to last many months, the survey is designed to be short and quick to complete, so regular feedback can be captured without study fatigue. Since high frequency survey capture can be burdensome to the participant, the exact time and frequency of data capture should be negotiated during the initial meeting with the participant. Table 18 presents a list of questions and responses that will be used to determine the perceived environmental quality outcomes.

Table 18 – Automated survey questions and responses

Outcome (Perceived)	Question	Response 1	Response 2	Response 3
Temperature	How is the temperature?	Too Cold	Comfortable	Too hot
Humidity	How is the humidity?	Too Dry	Comfortable	Too humid
Light	How is the light?	Too Dark	Comfortable	Too light
Noise	How is the noise?	Too Quiet	Comfortable	Too noisy
Air Quality	How is the air circulation?	Too Draughty	Comfortable	Too stuffy
Air Quality	Is it Dusty?	Yes	No	-
Air Quality	Are the any odours?	Yes	No	-

8.4.2 Data management

8.4.2.1 AppleWatch

At the end of the study, the participant will export their data using the Apple Health app and will be provided access to a web app [11] that has been developed for this study. The app was developed as the iOS health app that exports all data and includes personal data that would be unrequired for this study. Therefore, the app will provide the participant with a transparent system for converting data into the format required here (CSV). The app has been designed so that all processing is done on the participant's computer to ensure data are not uploaded to the internet. The app allows participants to choose which data to export, allowing them to remain in control over what they are sharing *(Figure 20)*. Oftentimes, these data can still be used to identify the device owner as Apple device names are typically named after the account holder e.g., Graham's iPhone/AppleWatch. Consequently, device names in the exported files will be simplified to iPhone/AppleWatch before the data is stored in Microsoft OneDrive.

It should be noted that while this chapter presents the use of an Apple Watch and presents a tool app [11], which specifically interacts with AppleHealth data, this protocol is not specifying that an Apple device must be used. The outcomes being measured by the wearable in this study are related to walking, heart rate and caloric activity data, most personal fitness trackers are able to measure these outcomes. While future researchers could explore Apple devices within their studies using the processes and tools outlined here, it is recommended that they consider exploring the work by Henriksen *et al.* [101], who provide a case study analysis that identifies a checklist of eight categories for selecting and appraising WHTs for monitoring health outcomes.

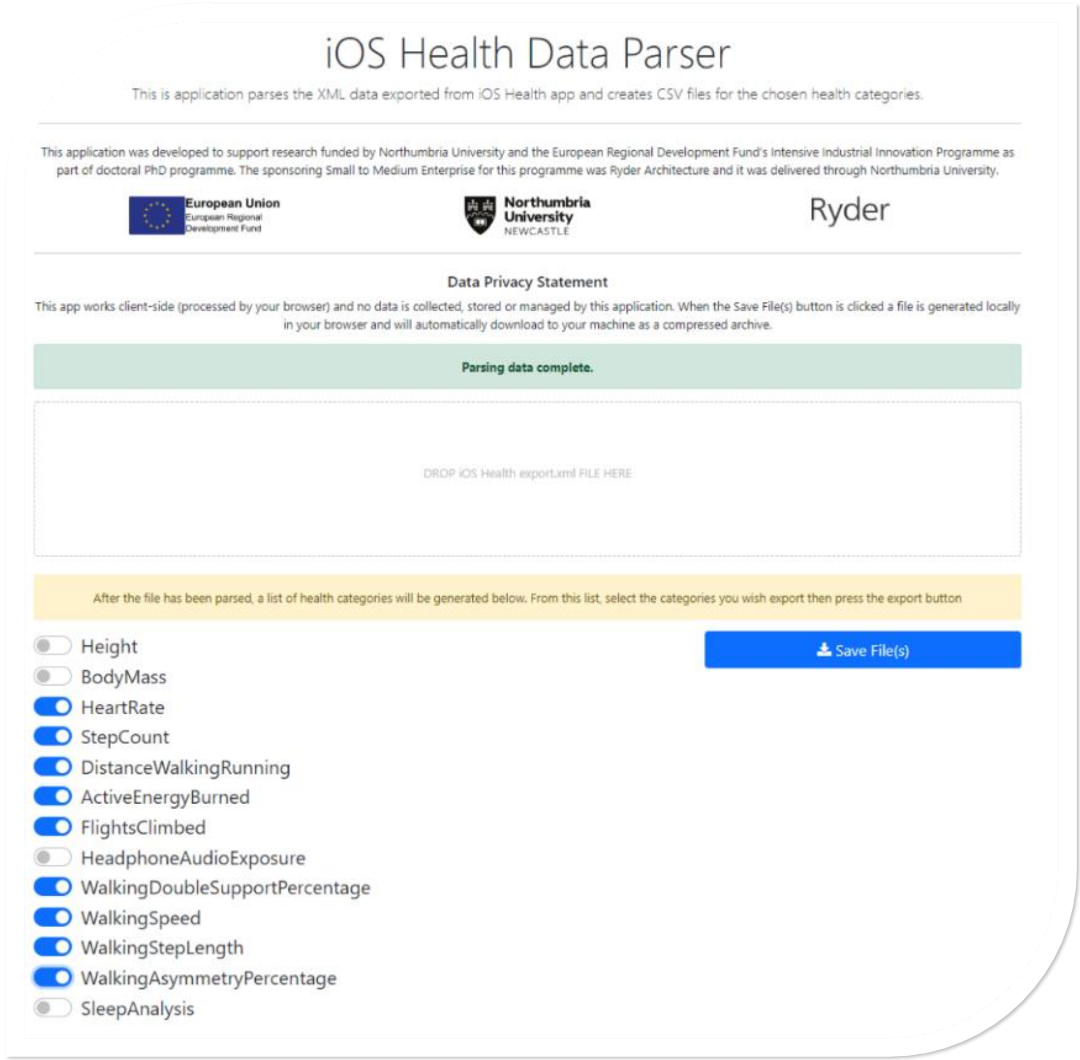

Figure 20 - Screenshot of iOS Health Data Parser application

8.5 Statistical methods for analysis

Due to the volume of data that will be captured during longitudinal measurements, a two-stage approach will be adopted to explore the data at different timescales. An initial macro-level assessment will be conducted to evaluate the data at high-level exploring daily/weekly averages of sensor data. This will involve comparing qualitative data captured from the post study interview and TESS survey against daily/weekly averages of sensor data.

The second stage of the analysis process will involve a micro-level assessment that explores the minute-by-minute data in more detail. Combining these approaches will form a 10-step process on which to address the overarching research question outlined in the thesis statement (Figure 21), with the aim of addressing whether localised sensors can provide richer data that will enable better understanding causal relationships i.e., how individual building occupants respond to environmental changes.

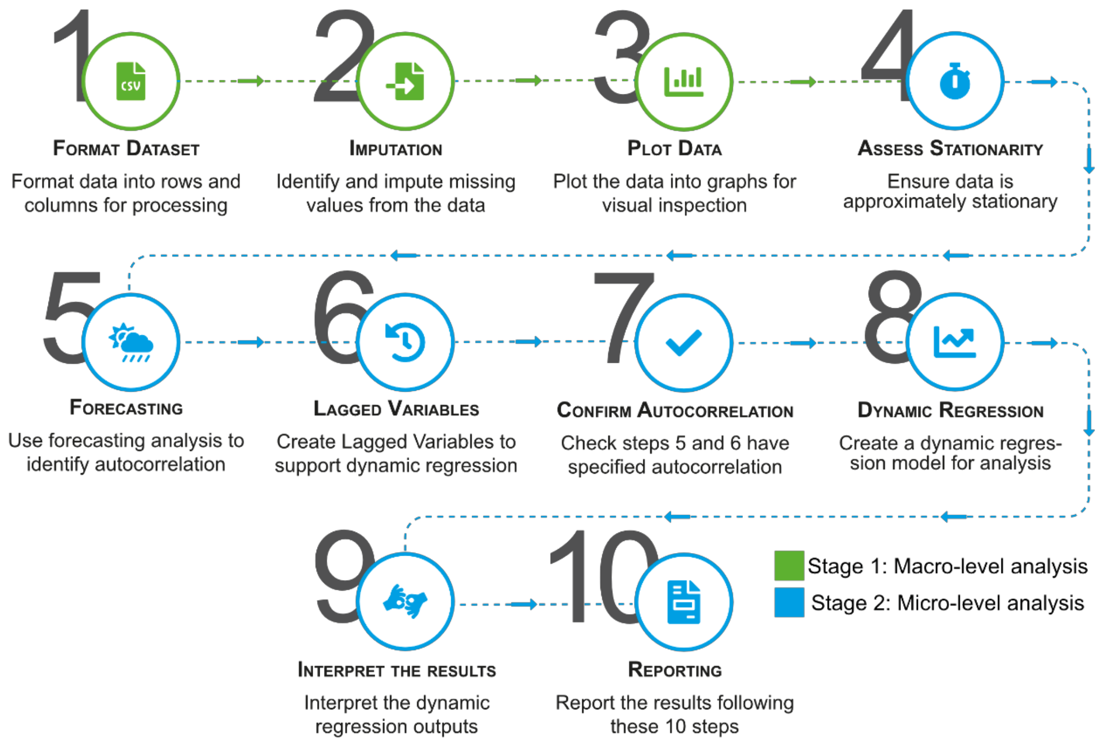

Figure 21 – Flowchart representation of the methodological approach - highlighting the two-stage macro and micro-level analyses.

Dynamic regression analysis will be conducted following a 10-step process (Figure 21), to identify causal relationships between the dependant variables (*Table 16*) and the independent covariates (*Table 17*). This process will involve preparing the dataset ready for dynamic regression modelling, which will include formatting the dataset, assessing how stationary the data is and evaluating the extent of autocorrelation in across the data, which are important first steps in dynamic modelling [272]. Figure 21 provides an indicative high-level overview of the steps that will be involved for the purposes of this study when conducting the statistical analysis, but readers are directed to a comprehensive tutorial for conducting such analyses in IBM SPSS Statistics for a more in-depth outline of the statistical methods involved [267].

8.6 Ethics & Dissemination

8.6.1 Research Ethics Approval

The protocol outlined in this chapter was granted ethical approval from the Northumbria University Research Ethics Committee *(Submission Reference: 20481)*.

8.6.2 Informed Consent

Informed consent will be obtained from the participant by the principal investigator during the initial meeting. The participant will consent to the researchers to obtain data from wearable sensors, environmental sensors and surveys. The participant will consent to the terms of the study as negotiated during the initial meeting.

8.6.3 Confidentiality

Confidentiality will be maintained throughout this study. While it will not be possible to blind the participant from the research or vice versa, all data collected by the researchers will be anonymous or anonymised prior to analysis. Demographic data on the participant will be included in the write-up and analysis of data, but this be presented in a way that provides no means to identify a natural living person.

8.7 Discussions and Conclusions

The protocol outlined here has the potential to explore a unique avenue of IEQ research. Monitoring IEQ in an individual context presents opportunities to not only gather spatially dense IEQ data within buildings, but the data captured will be representative of what is experienced by individuals near the monitoring equipment. By using emergent low-cost technologies, researchers can develop scalable and more objective/insightful and personalised monitoring solutions that could be used to address a range of challenges for building scientists [266]. This protocol has the potential to provide unique contributions to building science research by addressing current gaps in literature around the effect environmental changes have on individual building occupants.

8.7.1 Contribution

The proposed approach of using personal assessment by longitudinally measuring an individual, could be used to examine relationships between environmental conditions, activity behaviour and physiological changes *(e.g., change in heart rate)*. This may help identify individual thresholds of comfort in future work. For example, if these thresholds are known with a degree of confidence, it could be possible to remove the need for survey feedback entirely *(or heavily reduce the data capture frequency).* Furthermore, longitudinal capture of these thresholds could also be used to inform personal comfort models trained with machine-learning, comfort thresholds with real-time data capture [265].

If environmental measurements detect changes in IEQ outside of an individual's comfort thresholds, it is possible that those data could drive building management systems that manage environmental conditions at a local level [273]. These data could potentially enable the development of automated systems that can provide personalised actionable feedback to occupants based on environmental conditions. Systems exist that provide actionable advice to building occupants based on building information and IoT-based environmental monitoring, but there is a need for a wider context to understand the more subjective factors of IEQ [38]. With additional context, such systems could provide steps to control the IEQ, if conditions rise or fall outside of an occupant's comfort threshold and if they have control over the environment. If occupants do not have control over the environmental conditions *(e.g., in a shared office environment)*, the system could provide *e.g.,* alternative work locations, based on measurements obtained from other sensors.

It should be noted that when applying this protocol to a single case study, the perceptions of the participant would likely be unsuitable for evaluating building performance as they would be too subjective in isolation. To conduct such a study, multiple, synchronous, individualised studies would need to be conducted on a wider building population. It is still suitable to use this individualised approach for group studies, but each individual would be measured as a single case [267]. The results could still be used to get indications of average comfort levels, but the individual focus in the data would provide added value compared to traditional group studies. Where traditional group studies would generalise the views/opinions of the population, multiple individualised studies have the potential to expose variations among the study sample. However, while multiple individualised studies could be used to assess generalised findings within a population, each single-case study should be treated as its own study. In doing so, causal relationships could be assessed between intra-case variables, but cross-examination should not be conducted on inter-case variables. Instead, the results from each case could be used to form a new dataset, that could be evaluated as a population study. This is particularly important with regard to the use of wrist-worn fitness trackers, as the heart rate sensors can report variations when used to measure different demographics [274]. These variations would not impact individual single-case studies but would likely impact wider population comparisons if measured within an inter-case, multivariate analysis. This protocol has the potential to provide a holistic methodology for understanding personalised thresholds of comfort while gaining quantitative insights into the effect buildings and environments have on occupants. This has the potential to provide a deeper understanding of individual building occupants and move away from the generalised measurements of occupant populations seen within current building standards [60], [67], [270]. This approach could also help to acknowledge variations across individuals, which could add value to building performance assessments.

8.7.2 Limitations of Study

The small sample size in single-case studies could be perceived as a limitation and is often faced with barriers and resistance in practice [275]. However, single-case studies are specifically designed with a view to using longitudinal timeframes in the examination of an individual, which can provide greater insights on changes in health and behaviour over time, when compared with studies of larger sample sizes [177]. Alternatively, multiple individualised studies can run simultaneously with results used to evaluate how generalised the findings of individual cases are [269]. Upon validating the approach and evaluating the findings of this protocol, future research could scale the methods up to multiple individualised studies.

The use of in-situ sensors outlined in this protocol means that this protocol is ideally suited to office workers, who would generally work from a fixed location. Researchers wishing to apply this protocol to workers with a more mobile profession would need to develop multimodal measurement devices that are portable and potentially wearable. This was deemed outside the scope of this protocol as it presents many complications including, *(but not limited to)* how sensors are calibrated to deal with participants navigating between indoor and outdoor environments. The participant selection process will account for this limitation, but it should be considered in future research. Where participants have more active jobs, which involve moving between spaces or changing working locations, more in-situ sensors could be used to account for these transitions.

8.8 Address the PoI

From longitudinal monitoring of individuals, time-differential outcomes can be observed that in future work could help to determine individualised thresholds of comfort. Additionally, robust methods for data collection and analysis of individuals could expose causal relationships between environmental changes and physiological responses. These comfort levels could be used to train machine learning models and aid the development of automated feedback mechanisms that provide individualised actionable advice to building occupants. The approach proposed here has the potential to counter generalisation in occupant comfort studies by exposing variations in research groups while providing a deep understanding of the effect environmental changes have on building occupants at an individual level. This could have practical implications for building owners as it could provide them with a better understanding of how buildings could be adapted to suit individual variations in comfort. Overcoming generalisations in comfort assessment could extend this protocol to work as an interventional model for evaluating the impact energy performant buildings have on individuals. For example, longitudinally monitoring individuals during future energy performance renovation projects. In this way, this protocol could have the potential to identify causal pathways between energy performant buildings and the health and wellbeing of building occupants.

To answer **PoI6**, this chapter has presented generalised methods for deploying the multimodal sensor developed as part of chapter 7 and has demonstrated a methodological process for deploying a range of data capturing devices that can be used to collect data on the individual, their immediate environments (home/office) and the outdoor environments that influence the buildings they occupy. However, this chapter has been positioned to present the protocol as a generalised framework for future researchers, which must be expanded upon before conducting the major study.

8.9 Further Research

Chapter 9 will build upon this protocol by presenting an outline of; the study location, specifications of the technologies used to capture the outcomes, and details of how this protocol has been adapted to suit the individual needs of the participant.

Chapter 9 Personalising a protocol for individualised monitoring of building occupants

9.1 Introduction

An initial meeting was conducted with the study participant to understand their expectations and to discuss the methods of data collection that would be adopted. The placement of sensors, survey-data capture mechanisms and participant involvement were discussed, leading to several protocol methodological adjustments. Specifically, the aim of this chapter is to present how the study framework outlined in the previous chapter was personalised to suit the participant that was recruited as part of this thesis and to define criteria pertaining to:

1. study location
2. timeline for data collection
3. adopted sensor technologies
4. mechanisms adopted for survey-based data collection.

This chapter will continue to explore **PoI6** by expanding and applying the generalised protocol *(Chapter 8)* with details of how the methods were personalised for the studies in the subsequent chapters. This chapter will enable future researchers to understand the nuanced flexibility of individualised studies. It will also serve as a basis for the field to implement computing science-based approaches for pragmatic deployment of studies that make an individual the unit of analysis. Accordingly, this chapter will present a series of study specific hypotheses, which were identified after deployment but before analysis. Hypotheses will provide a point of investigation when beginning to explore the data in Chapter 10.

9.2 Methods

9.2.1 Study Setting

The study location for this research was a residential property in Tyne and Wear, UK. The property is located near a major transport infrastructure and contains two buildings within the same property boundaries (Figure 22). A participant was chosen for this study that had a dedicated home office, as the COVID restrictions were still in place at the time of the study, so the mitigations specified in the protocol were adhered to.

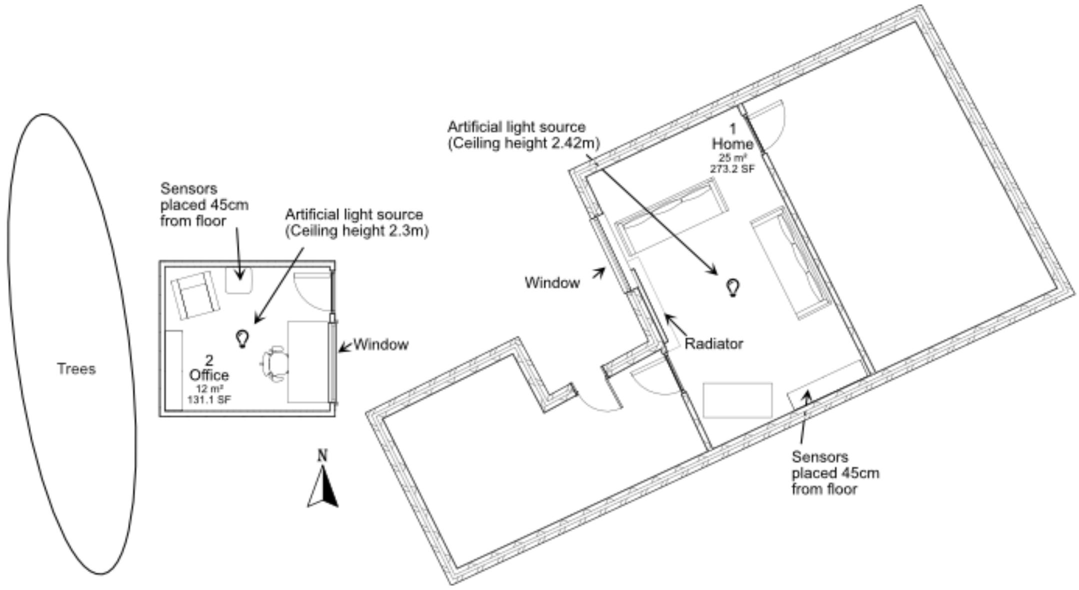

Figure 22 - Site plan for study location, showing home, office and surrounding features including the locations of windows, heaters, lights and sensors

9.2.1.1 Office

The office was single occupancy, in a converted summer house, located at the end of the rear garden and situated within a tree line. The building was raised from the garden by a height of one meter on a decked platform and was constructed of 23mm walls, which comprised of 15mm solid timber panels and 8mm of internal insulation. The roof was constructed of 19mm tongue and groove boards on timber beams, with a green mineral roofing felt on the outer surface and no internal insulation. The floor was also constructed with 19mm tongue and groove boards on timber beams and was carpeted with 9mm underlay. Since the roof lacked insulation and there was no cavity wall insulation, it raised the question as to how effective the internal layer of wall insulation would be for indoor thermal comfort (Hypothesis 9.1). In the office there was an electric heater that was manually controlled to provide additional heat; however, there was no air conditioning so natural ventilation was used to cool the office.

Hypothesis 9.1

The construction of the home office lacked cavity wall insulation and had no insulation on the roof. Thus, it is hypothesised that outdoor weather conditions will have a large effect on indoor thermal comfort.

The office had east-facing doubled-glazed windows and French patio door i.e., the east side of the building is almost entirely glazed *(Hypothesis 9.2).* The property had no blinds, so the participant was unable to block out light. There was also a manually controlled artificial light source in the centre of the room.

Hypothesis 9.2

Since the office has no blinds, it is hypothesised that, due to geographical positioning of the office, morning sunlight will saturate the office.

9.2.1.2 Home

The home was a 1930s, 2-bed, semi-detached bungalow.. The total floor area for the property was 82m^2 and the room chosen for the study, within the home, was the living room – a 25m^2 room situated in the centre of the property. The property was a brick-built construction with cavity walls, but no information could be obtained about the presence or thickness of cavity wall insulation. The roof was a pitched, tiled roof with 100mm of insulation in the loft space *(Hypothesis 9.3).* The room was heated with central heating, which was controlled by a thermostat located in that room and was manually controlled by the participant. The room also had a gas fire in an existing fireplace, which was flued through an existing chimney. There was no air conditioning in the property, so natural ventilation was also used to cool the home.

Hypothesis 9.3

Due to the brick construction of home, the presence of cavity walls and the insulated loft space, it is hypothesised that the home should provide a greater thermal stability, when compared to the office.

The home was entirely double-glazed and the windows for the living room were west-south-west facing. The room had several small lamps, but the primary light source for the room was an artificial light source in the centre of the room.

9.2.2 *Overview*

An initial meeting took place one week before the study to discuss/adjust the requirements with the participant. The monitoring technologies were deployed one day prior to the study period and were removed from the study location one day after data collection stopped. After the sensors were removed from the study location a close out interview was conducted with the participant, which lasted approximately one hour.

9.2.2.1 Initial Meeting: Protocol personalisation

An initial meeting with the participant was conducted a week prior to data collection. During this meeting the Informed Consent Form (Appendix E) was explained to the participant, and they were presented with an opportunity to tailor the protocol with the researcher to ensure the study had minimal disruption to their home and work life. In this meeting, the participant had several requests, which resulted in modifications to the data collection methods as described here.

9.2.2.2 Office Hours

The participant requested that data be collected on weekdays only. Consequently, the decision was made to focus primarily on the office and collect data during office hours. The participant also requested that surveys be conducted once per day at 10:30am. The participant had the option to defer and complete the survey at a more convenient time in the same day. While qualitative data were captured during the study period were solely from the office, passive quantitative sensor data were continuously collected from the home and office to compare both locations.

9.2.3 Wearable: Overnight charging

The participant explained that they didn't find the AppleWatch comfortable to wear when sleeping, so they used this time to charge the device, which was required daily. This meant that it would not be possible to capture sleep data in this study. This highlights limitations around the prevalence of smart watches in this domain. Many PFTs have extended battery lives (>1 week) and short charging cycles (<1 hour) meaning that they can be used to monitor health outcomes day and night. However, smart watches integrate many additional features, which have a significant impact on battery life and require daily charging. Since sleep was not defined as an outcome of this study, this did not impact the collection of data required for this investigation.

9.2.3.1 Survey capture mechanism: Amazon Echo

The participant felt that manual data entry (paper and/or digital) surveys would be disruptive to their work and habitual routines. Therefore, the participant requested surveys be audible and voice based. Accordingly, surveys were conducted using Alexa via an Amazon Echo.

9.3 Data collection and management

9.3.1 Data collection methods

9.3.1.1 Measuring IEQ perceptions

To capture survey data for IEQ perception measurement, an *Alexa Skill* was created using *VoiceFlow*[9], which is an online, web-based visual scripting tool for creating voice enabled applications for smart assistants e.g., *Amazon*, *Google*. The skill enabled voice-controlled events for (i) completing the survey or (ii) halting the survey if the participant was busy. The

[9] https://www.voiceflow.com/

development of the skill involved creating logic blocks that can be chained together to create a program (Figure 23). Loops and conditions can be used to create conditional flow and they can trigger Amazon Alexa's Text-to-Speech engine to voice commands to the user. Microphone capture nodes can be used to listen to voice samples for keywords known as intents, which can be captured and stored as variables into Google Sheets. Of note, it is important to fully specify which intents the capture node should listen for to ensure there are no intent conflicts. For example, when testing this application, the question *"How is the humidity?"* could be answered freely without intent specification as the responses "too dry", "too humid" or "comfortable" did not conflict with Alexa. However, for sound and light answers "too light" or "too loud" would not be captured, given one can control light and sound on the Alexa device with voice intents it was assumed that these were reserved keywords. By specifying that a capture node should expect these responses, Alexa allows the skill to use those keywords in that instance.

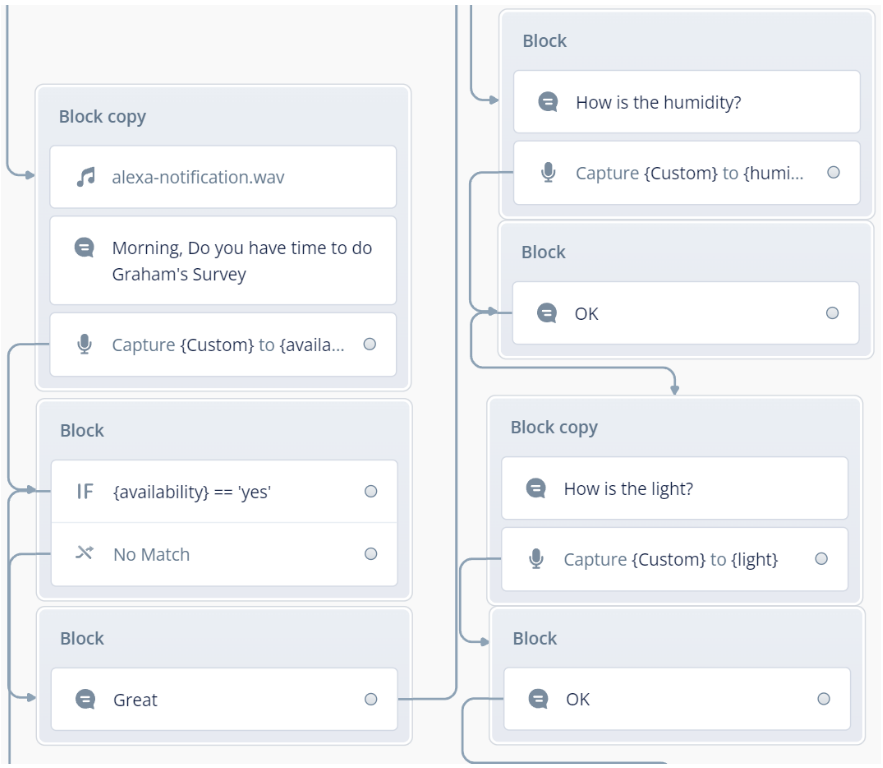

Figure 23 – Outline of the voiceflow application showing the visual script used to create the Alexa Skill

The skill was distributed using an *Alexa Routine* so that the survey prompt was delivered to the participant at 10:30am Monday-Friday (in accordance with the participant's request). The process of deploying the application to Amazon was inhibited by Amazon's beta testing policy, which prevented long-term deployment of the skill to third-party users. Such limitations were not present to the developer account *(the amazon account owner)*, so a dedicated Amazon account was setup whereby the participant was the developer giving them full, uninhibited access to the skill. Amazon provides the capability for skills to be hosted privately and distributed in the same way as public skills. However, this feature is only available to business organisations. Therefore, small research groups may need to consider publishing the skill publicly if larger sample-groups were recruited outside of business use.

9.3.1.2 Measuring IEQ changes

The sensors used in this study are the same sensors, specified, tested and validated in Chapter 7. Given the accessibility and low-cost nature of sensors is causing rapid deprecation of hardware [276] *(Chapter 3)*, the decision was made to exclude specific sensor make/models recommendations from the protocol given that the intended use-case for the protocol is to act as a framework for future research. Alternatively, the outcomes that should be measured (by the sensors) were specified to enable researchers select the most appropriate sensor(s) at the time of research. Table 19 presents a breakdown of the specific sensors used in this thesis for the multimodal monitoring device and presents these data alongside the specific outcomes defined in the protocol.

Table 19 – Sensors in the multimodal environmental monitoring device and the specific outcomes they will monitor

Outcome	Sensor	Outcome	Unit
	CCS811	eCO_2	ppm
Air Quality	MH-Z19	CO_2	ppm
	PMSA003i	$PM_{2.5}$	µg/m3
Temperature	BME280	Temp	°C
Humidity	BME280	RH	%
Light	BH1750	Light	lux
Noise	INMP441	Noise	dB SPL

All data are captured as per the protocol, but it should be noted that noise data are captured at the same frequency as other measurements taken by the device.

1.2.3 *Technical issues during deployment*

Visits to the study location beyond routine data collection were required. The first arose when a pet knocked the office based IEQ device off the table, damaging the microcontroller. This resulted in data loss for several days due to availability of the participant to be at home to enable access to the property. However, the IEQ devices were designed so components could be substituted in the field, enabling a speedy repair. Additionally, there were issues with SIM card connectivity throughout this study. Although contingencies were put in place *(based on the lessons learned, Chapter 5)* to restart the sensors in the field *(using smart power strips to remotely restart devices when they became unresponsive),* the smart functionality of the power strips was halted in the event of a data outage. However, the participant was able to restart the 4G router by turning the smart switch off and on manually, which removed the need to make multiple visits to the study location. Due to the data outages, the study ran from 15 March 2021 until 05 August 2021 (>20weeks, 143 days *and 4 weeks longer than specified in Chapter 8)* to ensure sufficient data were captured.

9.4 Discussions and Conclusions

This chapter presents details of the study site, including the approximate location, a description of the buildings' construction, the orientation of the properties/windows; survey of the site boundaries, including details such as surrounding trees and a trainline. Details of the internal spaces were also presented, showing details of sensor place heights location of light sources (both artificial and natural), location of heating systems and ventilation sources. This process was valuable and should be considered when conducting similar research as it raised a series of hypotheses which will be used as a starting point when exploring the data in the next chapter.

This chapter also discussed how the protocol was personalised to suit participant expectations/requirements. It must be noted that while the protocol is intended to serve as a framework, the ability to personalise it should not be seen as a license to change the methods in such a way that the aims of the study cannot be achieved.

The steps taken in this chapter to personalise the protocol do not change the outcomes of the study or impact the aims or objectives. The use of an *Amazon Alexa Skill* changed the mechanism for capture to adhere to participant requirements, but the resulting data produced from the skill is the deemed appropriate and an adequate replacement compared to paper-based survey capture. This process also demonstrated the ability of using computer science-based approaches in the field of building sciences to ensure greater participant adherence and streamlined data capture.

As per the protocol, qualitative data was collected once per day, but these data were only captured during office hours, as requested by the participant. This meant that no qualitative data were captured at weekends and primarily focused on office assessment. However, both environments were passively monitored continuously and could still be compared. Qualitative data was captured at the end of the study to assess the participant's perceptions of the two environments.

9.5 Address the PoI

This chapter, when combined with the previous chapter, presents a methodological approach to deploy a range of sensing modalities for the remote assessment of an individual. It is felt that these chapters serve as a framework from which to conduct such research.

PoI6:
How can multiple sensing modalities be pragmatically deployed to gather data for the longitudinal assessment of individuals building occupants?
Wi-Fi-enabled, multi-modal monitoring devices could provide a mechanism to pragmatically deploy multiple sensing modalities local to individual occupants across a range of building environments. While data from intra-device sensors can be aggregated internally, data can be transmitted to cloud platforms for aggregation with other devices. The use of mobile 4G routers and smart power solutions provides increased security and control over the hardware when it is deployed longitudinally in the field, to reduce the need for participant interventions and to provide the researcher with a mechanism to monitoring, control and troubleshoot sensors remotely. Including the participant in the study design can enable a tailored methodological approach, whereby individuals can contribute to the study design to provide a sense of ownership and control to reduce burden and increase participant adherence.

9.6 Further Research

The first stage of the two-stage analysis approach *(Chapter 8)* involves formatting, sanitising, and visualising the dataset, to conduct a macro-level investigation of the longitudinal data. To do this, Chapter 10 will present an investigation of daily/weekly averages, with the aim to evaluate whether trends appeared across the course of the whole study. This provided an opportunity to investigate the utility of the data to provide meaningful insights into the health and wellbeing of occupants at an individual level. Since the TESS was used to gather retrospective, subjective perceptions about the environmental quality across the whole of the study period, the preliminary investigation also provided the opportunity to explore the average trends in quantitative sensor data against the qualitative retrospective evaluations made by the participant. This was deemed a valuable step, given that the literature review *(Chapter 2)* identified that many building performance assessments are conducted in this retrospective, self-assessment manner.

Chapters 10 and 11 will explore **PoI7** to understand whether wearable sensors can be used to address the subjectivity around environmental perceptions by providing quantitative context to how occupants are affected by environmental changes within buildings. Chapter 10 will first present the results of the macro level analysis of daily/weekly averages which will serve as the foundational steps of the 10-step analysis process. This will then be followed by Chapter 11, which will continue the comprehensive 10-step analysis of data to form stage 2 of the analysis process by conducting a micro level analysis using dynamic regression modelling to understand relationships between variables.

Chapter 10 Macro-level personalised IEQ: Exploring use of quantitative data to contextualise perceptions

10.1 Introduction

Due to the quantity of personalised IEQ data gathered, a macro-level investigation was performed first before conducting a micro-level analysis (Chapter 11). To align with current practice, this investigation was driven by qualitative findings obtained from data gathered in a closeout interview that used a customised version of the ASHRAE Standard 55 Thermal Environment Satisfaction Survey (TESS) [61], gathering subjective perceptions of environmental quality from the participant. Specifically, TESS prefixes thermal comfort questions with the following statement:

> *"Please respond to the following questions based on overall or average experience in the past [six] month"*

Accordingly, the investigation conducted here was to obtain a broad understanding of the IEQ sensor data, such as uncovering any broad trends to showcase utility of macro-level data in providing meaningful insights into the health and wellbeing of individual occupants [61]. This was deemed appropriate as the outcomes of the self-reported survey were retrospective i.e., comparison between survey outcomes and continuous high-resolution sensor data would not be a like-for-like. Therefore, the aim of this chapter was to answer **PoI7**, by presenting a data aggregation workflow and to examine whether localised IEQ devices with multiple sensing modalities can be used to address the subjectivity around environmental perceptions by providing quantitative context to how occupants experience building environments?

10.2 Methods

For continuity, methods described here form the basis the 10-step process that will be explored in depth in Chapter 11. For the purposes of this thesis, the TESS question (above, *introduction*) was rephrased from "*in the past [six] months"* to *"during the course of the study"*, so the subjective responses gathered from the participant are based on generalised opinions over the course of the study and not based on any real-time, or point-in-time measurement, which averaged sensor data would be more aligned to. This process could be of value since it is common for studies to measure both objective and subjective outcomes of IEQ with a lack of synchronicity between measurements [5]. It is hypothesised that this investigation will present high level, but insightful findings aligned to the list of primary outcomes in Table 1 of Chapter 8. Averaged, objective IEQ sensor outcomes will be compared against (i) the qualitative data captured from the end of study interview and (ii) the quantitative data from the weather API and wearable device. Additionally, an outline of the analysis and statistical methods will be presented.

10.2.1 Synchronicity

The multimodal IEQ sensors captured/sampled at 40 second (s) intervals (0.025Hz), but were not aligned to one another, nor aligned to the measurements taken from either the wearable or the weather API (Figure 24). Synchronisation was needed for parity during the statistical analysis of a multivariate dataset, ensuring each measurement could be treat as a variable of a single measurement case. It also enabled like-for-like comparisons (for initial investigation and visualisation), which is an important early step in the 10-step analysis process.

To achieve synchronisation, data were resampled to 1-minute intervals – in line with the approach used during the testing and validation of the multimodal sensors (Chapter 7). Data were downloaded from *ThingSpeak* and renamed to appropriate variable names e.g., HOME_CO2, HOME_Temp for easy identification. Data from the two multimodal sensors *(home and office)* were then merged into a single CSV file using the timestamp *(generated by ThingSpeak, when data were sent from sensors)* as the common field to join the two datasets. Appendix F outlines the functions that were used in this process. In short, each of the files were parsed from CSV format into a *Pandas DataFrame*, which were then resampled according to the frequency of the initial data capture.

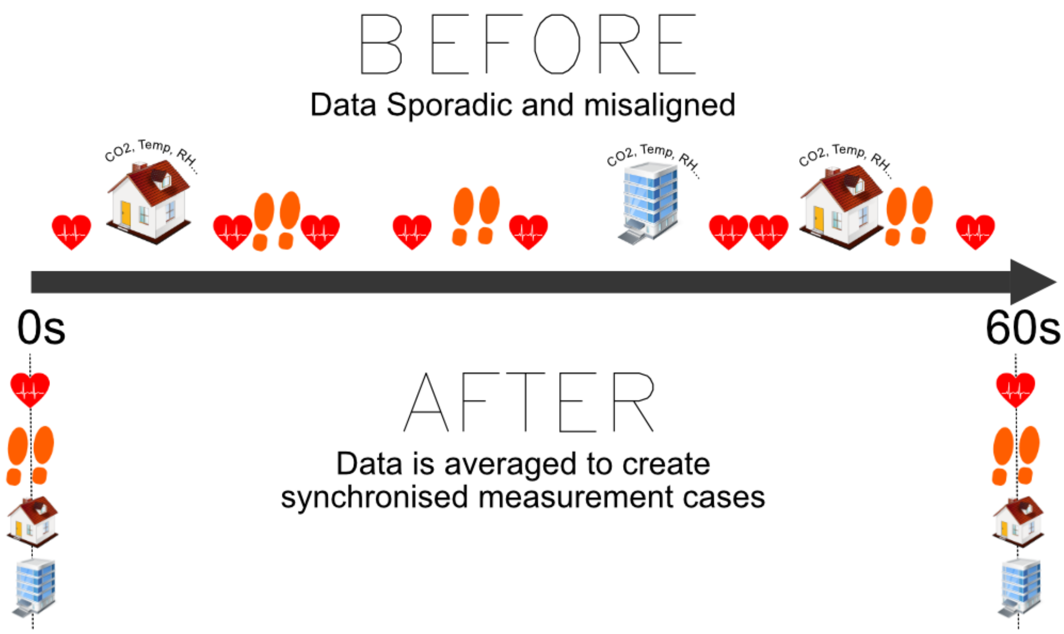

Figure 24 – Aligning data to make a synchronous multivariate dataset. Top (before): IEQ and physiological data captured at various time intervals. Bottom (after): all data synchronised to 1min/60s intervals

A multi-step process was used to create a single dataset suitable for conducting multivariate analysis. Since the multimodal sensors sampled every 40s, mean resampling was used to create a dataset for both sensors with the resulting dataset containing less cases than the original sets, but with each newly created cases equating to the mean value for each minute (Figure 25).

Figure 25 – Multi-step process for merging data from multiple sources

Hourly weather/pollution data were extracted from the OpenWeatherMap.org API, using a free-tier account. The free tier provides a 60 calls/s quota, which allows for up to 60 geolocations to be monitored per hour, since data are updated once per hour only. To merge the data (with the multimodal sensor data), it was padded to replicate the hourly reading for each minute within the hour. These data were then merged with the data from the multimodal devices to create a single dataset.

10.2.2 Wearable health data processing

Data were extracted from the wearable using a bespoke application/app (Chapter 8*, iOS Health Data Parser).* The app created individual CSV files for each outcome, which were processed before analysis. To ensure compatibility with data from the weather API and multimodal sensors, timestamps were changed to a standardised format. The 'yyyy-mm-dd hh:mm:ss' *(e.g., 2021-11-01 12:05:15)* date format was chosen to parse the dates, as SPSS (v27.0.1.0, IBM, "New York") is limited in the date formats that can be used.

Additional processing was done to the health data to ensure it could be analysed in the same context as IEQ data. For example, the step count is recorded by the *AppleWatch* for a given period of sustained activity/walking (Figure 26). To enable minute by minute comparison, data were divided into one-minute bins e.g., 50 steps for 12:05 and 12:06, then 25 steps for 12:07. Where no steps were recorded for a given minute, those data were padded with a 0.

```
{
   "start_time": "2021-11-01 12:05:01",
   "end_time": "2021-11-01 12:07:31",
   "total_steps": 175
}
```

Figure 26 – JSON data format for AppleWatch step data.

Heart rate was also captured at random times in the same way as step events. However, the **startDate** and **endDate** were the same for all heart rate measurements as the heart rate is a point-in-time measurement *(Figure 26).* Since AppleWatch heart rate measurements are frequent, but not continuous, heart rate data is recorded with variable measurement increments. For example, on occasion, multiple heart rate measurements were taken in one minute, followed by gaps in the measurements between 2 – 20 minutes.

```json
{
   "startDate": "2021-11-01 19:12:01",
   "endDate": "2021-11-01 19:12:01",
   "value": 84
}
```

Figure 27 – JSON data format for AppleWatch heart rate data.

Where it was possible to pad measurement intervals in the steps data with zeroes, as step events would only be recorded after a bout of steps, the same could not be said about heart rate. This is because 0bpm would not be a valid measurement of heart rate. To remedy this, interpolation of heart rate was required to remove data missing due to measurement intervals.

10.2.2.1 Interpolation of heart rate data

To impute missing data for heart rate, several interpolation methods were explored, but these methods either resulted in linear algorithms, aptly, producing linear transitions through the data or non-linear algorithms (e.g., spline, polynomial) producing datasets outside of acceptable ranges (Figure 28).

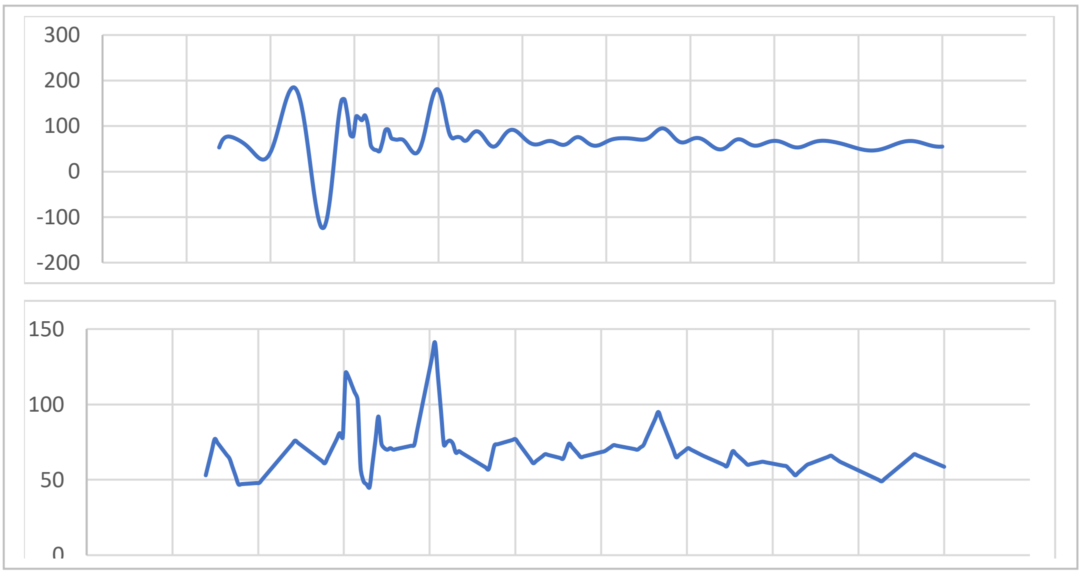

Figure 28 – comparison of linear vs non-linear interpolation methods

While linear interpolation methods produced a valid range in the data, it resulted in a dataset that is smoothed to a degree where it would not be possible to detect sudden changes in heart rate. However, when applying non-linear interpolation methods, the resulting interpolants contained values, which were invalid due to not only being significantly outside the extrema of the original data, but outside the extrema of acceptable heart rate data *(e.g. -120bpm)*. Barker and McDougall [277] acknowledge these phenomena when interpolating timeseries data with variable time measurement increments, noting that non-linear splines can produce nonviable data that exceeds realistic bounds. They suggest the use of Piecewise Cubic Hermite Interpolating Polynomial (PCHIP), which produces polynomial interpolants that are bound by the extrema of the original data. The PCHIP method was applied to the heart rate data, producing data that remained within the extrema of the original data, while retaining a non-linear curve across the interpolants (Figure 29).

Figure 29 – snapshot of data interpolated using PCHIP interpolant

10.2.3 Statistical methods

A combination of SPSS and Excel (v2112, Microsoft) were used to generate pivot tables, graphs and statistics. This enabled broad descriptive statistics (minimum, maximum, mean values, standard deviation) and visualisations. Bivariate analyses of data pairs were generated from SPSS, so that key data pairs could be evaluated. The primary statistical approach for this chapter is visual analysis, conducted by graphing out variables for comparison over averaged time frames.

10.3 Findings and discussion

Table 20 presents descriptive statistics for all variables and total volume of cases (N) i.e., the total number of minutes for the given study period (142 days: 16/03/2021 – 04/08/2021) was 205016, so this number is the maximum Valid N (listwise) that could be in the final dataset. Table 20 confirmed that the sensors were reporting data within the expected range of the sensors and were performing within expected norms (according to their evaluation in Chapter 7). The process of collecting data from the weather API and Apple Health Data

Parser meant these data were already padded to included missing values; therefore, these variables had a maximum Valid N (listwise). When collecting data from the multimodal sensors (home and office), these data were not interpolated to account for large gaps in data so these data have approximately 20% data loss (Table 20).

Table 20 - Descriptive statistics for data

	N	Minimum	Maximum	Mean	Std. Deviation
HOME_TEMP ***(°C)***	168849	16.85	32.36	25.10	1.81
HOME_RH ***(%)***	168785	25.30	59.00	41.44	4.89
HOME_LIGHT ***(lx)***	168851	0.00	650.00	10.72	31.29
HOME_SOUND ***(dBA)***	168846	34.30	87.42	39.09	3.05
HOME_CO2 ***(ppm)***	168851	390.00	1,317.00	572.27	141.20
HOME_eCO2 ***(ppm)***	168851	400.00	7,992.00	1,187.73	1,111.21
HOME_PM25 ***(µg/m3)***	168851	0.00	212.00	1.10	4.84
OFFICE_TEMP ***(°C)***	154396	3.64	42.17	22.02	6.57
OFFICE_RH ***(%)***	154362	13.43	57.95	36.00	5.38
OFFICE_LIGHT ***(lx)***	154396	0.00	26,460.00	177.40	839.41
OFFICE_SOUND ***(dBA)***	122532	24.86	109.62	45.23	15.17
OFFICE_CO2 ***(ppm)***	154396	258.00	2,644.50	618.18	191.85
OFFICE_eCO2 ***(ppm)***	154396	400.00	7,992.00	595.56	459.93
OFFICE_PM25 ***(µg/m3)***	154396	0.00	141.00	0.89	2.38
Outdoor Humidity ***(%)***	205016	21.00	100.00	77.00	15.45
Outdoor Temp ***(°C)***	205016	-2.89	27.01	10.70	5.42
Air Quality Index	205016	1.00	4.00	1.42	0.57
Step Count	205016	0.00	218.00	5.78	18.75
Heart Rate	205016	39.00	160.00	65.435	11.21

10.3.1 Temperature

TESS-based qualitative findings for temperature indicate that there is a link between the office temperatures and outdoor temperatures:

> *"When it's hot outside it gets too hot in office and I have to open door. When it's cold outside, I sometimes have to put the heater on before going into the office. If it's too cold, I would take a hot water bottle in.*

Hypothesis 9.1 identified a lack of insulation in the office building and hypothesised that thermal conditions of the building would be impacted by outdoor conditions. By inspecting the data from the IEQ devices in the home and office against outdoor weather data *(Figure*

30), the temperature from the office IEQ device demonstrates a large degree of variations (range: 15 to 29°C) during day and night cycles and these readings strongly correlated with outdoor temperatures. When compared to the data from the home (range variation: 24 to 26°C), highlighting that the office was significantly less insulated than the home. This also confirms **Hypothesis 9.3:** that the home should provide greater thermal stability due to the increase insulation.

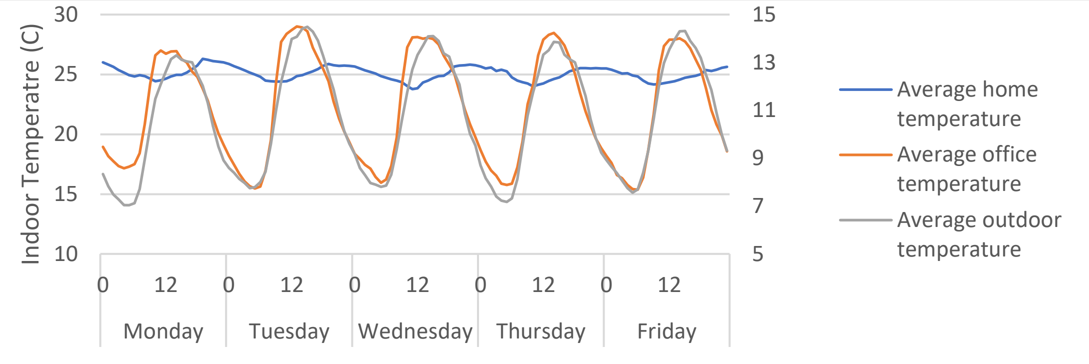

Figure 30 - Average hourly temperatures (per weekday) from each temperature source (Home, Office, Outdoor). Note: Outdoor temperature is presented on a secondary (right) Y axis to highlight the similarity in the shape of the curves between outdoor temperature and office temperature.

While the office does appear to be significantly less insulated than the home, the average indoor temperature in the office *(left Y axis)* was shown to be around double that of outdoor *(right Y axis)*. This means that the office building did provide some insulation as the office did not lose all of its heat during the night, when the building was unoccupied.

As the participant highlighted, the source of heating between the two properties differs. While the home used central heating, controlled by a thermostat, the office heat was controlled by an electric heater. This heater was only used when the office was occupied and only when it was cold in the office. Element heaters cool faster than radiator systems, which could indicate why the temperature drops off so suddenly in the office. However, since the home is controlled by a central thermostat, this will keep the property at a given set-point temperature. That notwithstanding, the temperature drops in the home are much more gradual than that of the office, further confirming **Hypothesis 9.3.**

10.3.2 Humidity

Qualitative data for humidity identified that the participant found that:

> *"[the office is] always more humid than the home".*

However, the sensor data *(Figure 31)* shows the opposite to be true. One possible explanation for this is that the human body is better able to manage core temperature within stable thermal environments and transient conditions can significantly affect thermal sensations [278]. This means people are better able to provide qualitative assessment of environmental conditions under stable thermal conditions. Given that qualitative assessment through self-reporting is one of the primary mechanisms for the assessment of building performance [5], [176], this raises questions over the efficacy of these methods. Therefore, the lack of thermal stability, in the office environment, may be an influencing factor as to why the participant was unable to provide an accurate self-reported assessment of the humidity in the office. These findings present the case for augmenting quantitative data into building performance and comfort assessment, to support and/or validate the findings obtained from subjective assessment.

That notwithstanding, when comparing the data from the sensors against outdoor data, there are further notable correlations between outdoor conditions and the conditions recorded in the office. This comparison provides further evidence of insufficient insulation in the office. Humidity in the home maintained relative stability, whereas the changes in the office reflected those of outdoor measurements. These findings match expectations, based on the evidence seen so far, alluding to assumption of poor insulation in the office building. However, these findings do not align with the qualitative data.

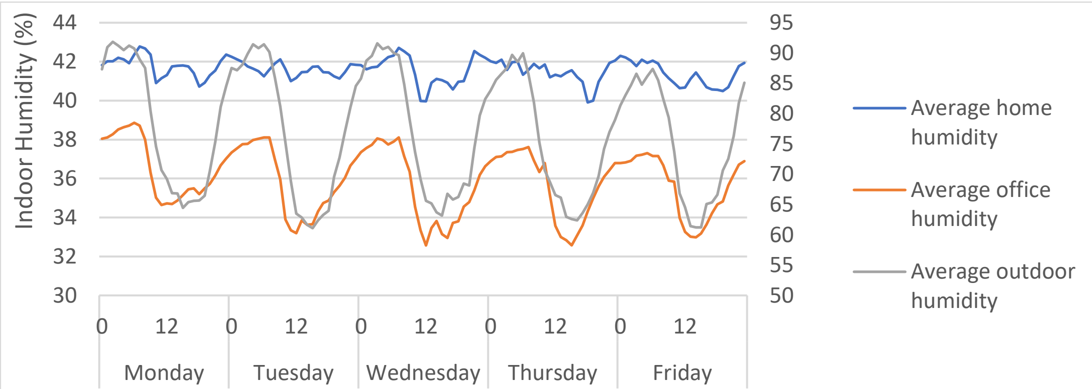

Figure 31 - Average hourly humidity (per weekday) from each humidity source (Home, Office, External). Note: Outdoor humidity is presented on a secondary (right) Y axis to highlight the similarity in the shape of the curves.

10.3.3 Light

The participant expressed a clear dissatisfaction with the lighting in the office, remarking that it was *"often too light in the morning"* and *"[the office is] always lighter than the home"*. The participant also noted:

> *"Sun glares through the [office] window, which makes it difficult to see the screen and there aren't any shades to block the sun out. Though, I can move seating position to get out of direct sunlight."*

Figure 32 shows that the measurement ranges for each light sensor were vastly different and light in the office was consistently, significantly greater than that of the home. Figure 32 also shows that the office consistently measured high levels of light intensity during the morning period, which is in line with the findings from the qualitative data and with the expectations of **Hypothesis 9.2**.

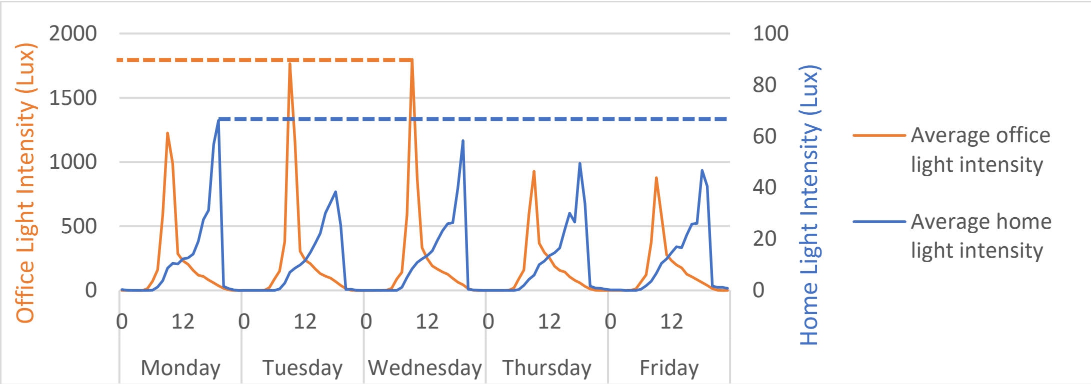

Figure 32 - Average Home/Office Light Intensity categorised by Day > Hour. Note: variables are presented across two Y Axes to account for the mean difference across the measured variables.

Hypothesis 9.2 also posed that the geographical positioning was the primary cause of the reported phenomena. The window for the home environment faced due West-South-West, whereas the windows in the office faced due East. A sun map was created using Autodesk Revit (v.2022.2.1, *Figure 33)* to track the sun movements during a randomly selected day during the study period. The sun map shows that after midday the sun would move behind the office and begin to cast light on the home environment. However, the shape of the home, and the treeline meant that the sun did not have a direct line-of-sight until around 18:00, by which time it was setting.

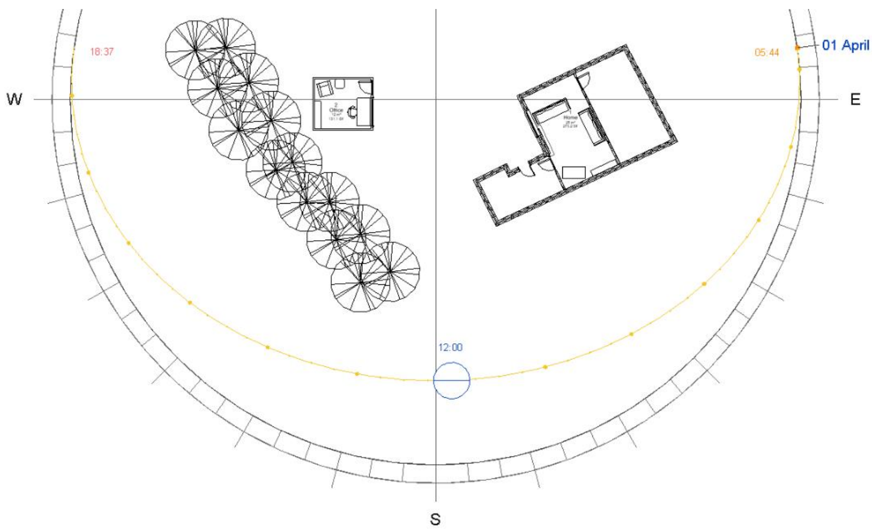

Figure 33 – Randomly selected sun map for the study location to give an idea of where the sun (yellow circle) would be throughout the day in relation to the buildings and rooms.

The light sensors performed with a high degree of accuracy during the controlled conditions of the validation study (Chapter 7). These sensors also show a demonstratable consistency with self-reporting data obtained from the qualitative assessment in this study and the modelled sun maps. The qualitative assessment did identify that the participant was dissatisfied with the light in the space, but they did not indicate the extent of the problem. Quantitative data in this instance, was extremely useful for identifying times where light intensity would be disruptive to work and the specific levels of light intensity the participant was exposed to. These data were also able to determine window orientations based on the cross comparison of light sensors with sun models. This presents a strong case for the longitudinal capture of localised, quantitative data in building performance and comfort assessments.

10.3.4 Noise

The participant explained that there was a trainline that passed close to the office building, stating:

> *"There is a train line that runs near the office. This does not bother me personally but can be distracting during remote meetings."*

Due to the proximity of the railway, and the regularity of trains passing, averaged data for sound was explored to see if regular high SPLs were noticeable that may indicate regularity in passing transportation. Unfortunately, the frequency of data capture meant that this inquiry could not be observed at this level of interrogation (Figure 34). Since data were originally captured at 40 second intervals, the exact time would not be synchronous between days. If a train were to pass the office at e.g., 10:33 every day, passing for a total of 20 seconds, this would not be captured every day for the 10:33 measurement window. This may spread the data in such a way as to not cause an increased average for a specific time. Therefore, for this interrogation, the measurement frequency reduced the statistical power of the data. To circumvent this issue, a measurement frequency of >1Hz would need to be adopted to ensure sufficient data were captured to calculate the sound pressure level for each second. However, this would increase the ethical risk surrounding audio capture, increase the cost of data capture/storage and may even become a source of discomfort for occupants.

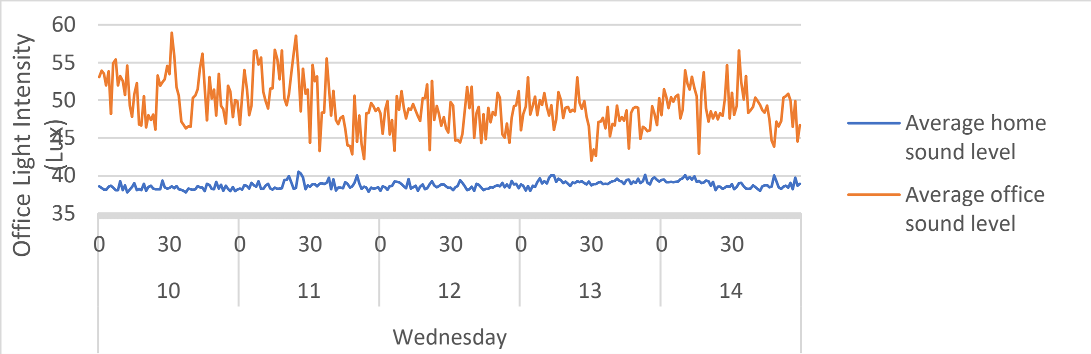

Figure 34 - Average Home/Office noise levels between 10:00 – 14:00 categorised by Day > Hour > Minute

Figure 35 shows the average noise levels were regularly between 5dB and 15dB louder in the office than in the home. Daily average trends also show that the highest average values in the office were always during office hours, indicating occupancy related noise levels. Comparatively, the highest average values in the home were recorded in the evening outside of office hours. This indicates that sound sensors could also serve as a proxy measure for occupancy. However, the use of audio recording equipment for detecting human occupancy could raise issues of ethics, privacy and trust, whether voice is recorded or not.

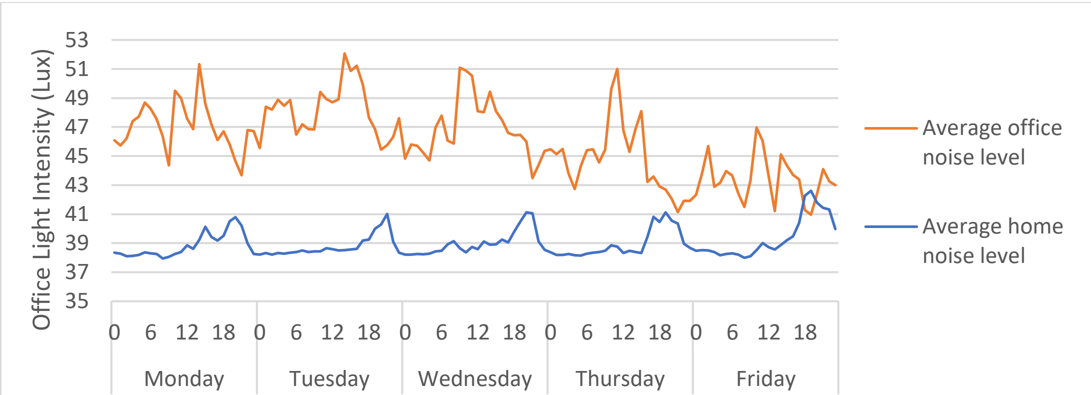

Figure 35 - Average Home/Office noise levels categorised by Day > Hour for Monday to Friday

In this study, no audio was recorded, and sound pressure levels were measured at the same sample rate as other measurements captured by the multimodal device (40s). So, while it may be possible to detect occupancy, it would not be possible to understand spoken words or determine who was present in a space. If a greater sampling rate is required, it is important to consider the ethical implications this may have especially if the resulting data could be deciphered in such a way where spoken words could be observed.

As identified in previous chapters *(Chapter 5 and 7)*, high sample rates could be problematic in a multimodal device, such as the one proposed, as it would mean the processor of the Microcontroller Unit (MCU) would potentially be continuously blocked by sound pressure level calculations. Multicore MCUs (discussed in Chapter 7) could enable synchronous, continuous SPL calculations, but the processing bottleneck would be pushed to the networking functionality (WiFi/Bluetooth/BLE/Zigbee etc.) instead, when data is transmitted to the cloud. Moreover, to record this level of data, a large amount of storage would be consumed in a short space of time and bandwidth/messaging quotas would also be rapidly consumed. The 40 second sample may not have been able to address specific lines of enquiry at this stage of the investigation, but this data capture frequency meant that both the unit costs and the running costs of the multimodal were both low, while still being able to provide findings useful to the aims and objectives of this Ph.D.

10.3.5 Air Quality

To obtain quantitative measurements of air quality, carbon dioxide was used as a proxy measure for ventilation and subsequently used to determine occupancy, and air circulation. $PM_{2.5}$ was used to measure the presence of dust, pollution and odours.

10.3.5.1 Carbon Dioxide (CO_2)

Qualitative data surrounding air circulation indicated that:

> "*[the office is]* is often stuffier than the home *[and]* stuffier in cold weather because I can't open the door and the office is small".

Given that the home is slightly more than double the square meterage of the office, the participant's comment on the circulation is not unexpected. Figure 36 shows that even during the evenings when the home environment is occupied, the maximum average CO_2 never exceeds that of the office despite that the participant shares that space with their partner, yet they are the sole occupant of their office.

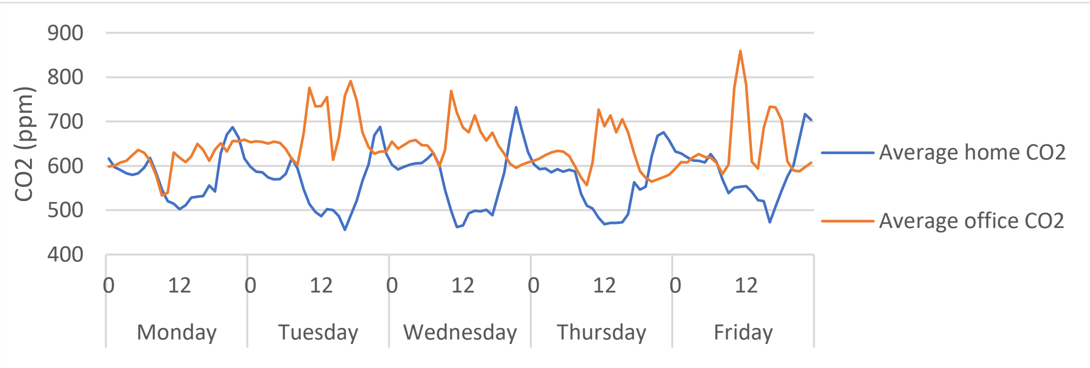

Figure 36 - Average Home/Office CO2 categorised by Day > Hour

Given that the participant felt that the room is stuffier when the door to the office was closed, it is worth exploring the links between occupancy and CO_2 levels as CO_2 sensors are useful indicators of occupancy and can be used as a surrogate measurement to determine the level of ventilation within indoor environments [38], [57], [266]. As already seen in Figure 36, patterns can be observed that indicate transitions from the home to the office during office hours. During these times CO_2 can be seen to increase in the office, while decreasing in the home.

By inspecting the averaged data of a single day (Figure 37), it is possible to see these transitions more clearly. Inverse relationships between office CO_2 and home CO_2 can be observed, which are also aligned with office hours and the qualitative accounts of working hours provided by the participant. Interestingly, Figure 37 also shows that on average the CO_2 shows a significant drop on Friday afternoons. This could indicate that the participant regularly left the office during these times.

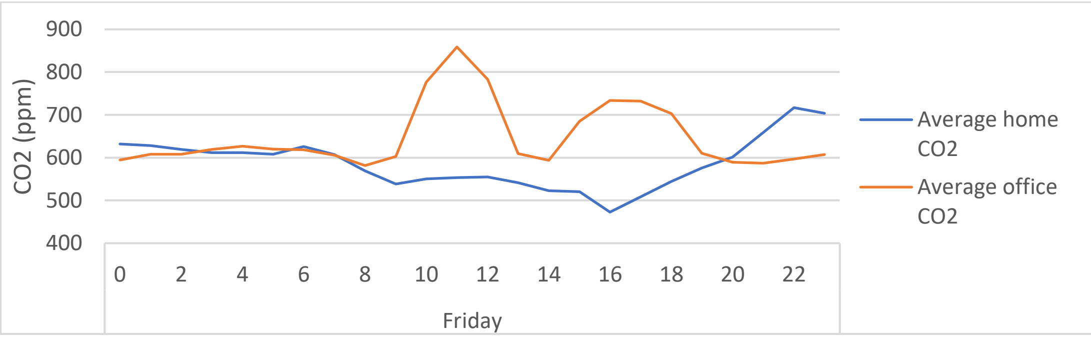

Figure 37 - Average Home/Office CO2 categorised by Day > Hour for an average Friday

It was hypothesised, that if the CO_2 levels observed here are indicative of occupancy, that there could be a relationship between CO_2 measurements and participant activity. To investigate this, step data was included and overlayed over the CO_2 averages (Figure 38). In doing so, it was possible to see that there were a series of inverse relationships between CO_2 averages and summed step count. As the participant transitioned from walking to resting (step count rising and falling) CO_2 levels in the office would begin to rise. Conversely, As the CO_2 levels in the office began to fall, an increase in step count can also be seen. By plotting these events, it is possible to see events when the participant may have entered and exited the office. This demonstrates that the combination of step data and CO2 can be used to detect occupancy, but it is anticipated that this phenomenon could only clearly be observed in a single-occupancy space.

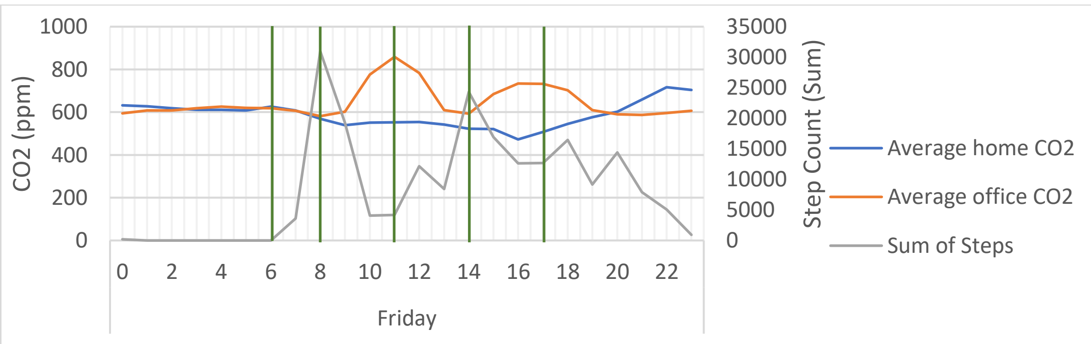

Figure 38 - Average Home/Office CO_2 overlayed with sum of steps and key event markers – vertical green lines signify the point when steps rise or fall.

This link between CO_2 and activity can be affirmed by looking at average heart rate for the same period (Figure 39). CO_2 has been shown to rise when steps decrease and fall when steps increase, and the same is true for heart rate. This would indicate that the CO_2 is linked to sedentary behaviour of the participant.

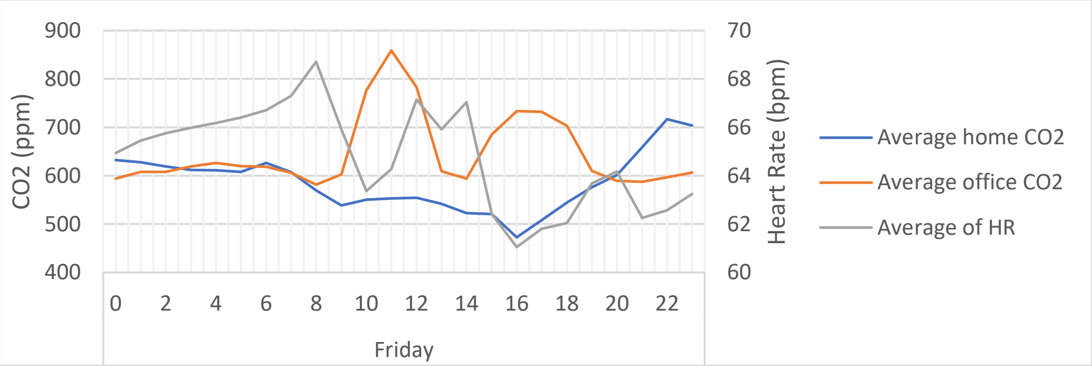

Figure 39 - Average Home/Office CO_2 overlayed with average heart rate

These findings provide demonstrable evidence of the capabilities of CO_2 sensors as proxy measurements for ventilation and occupancy, as identified in Chapter 7. Therefore, it is worth also exploring the eCO2 data within this context as Chapter 7 highlighted that they eCO_2 sensors also have the potential to be used as lower-cost proxy measures for ventilation despite recording erratic measurements in comparison to CO_2.

10.3.5.2Equivalent Carbon Dioxide (eCO_2)

By including the eCO_2 measurements *(from the office),* similar trends can be seen as those observed in Chapter 7 (Figure 40). The eCO_2 sensors provide more erratic measurements when compared to CO_2, yet the measurements also provide a similar indicator of office-based occupancy.

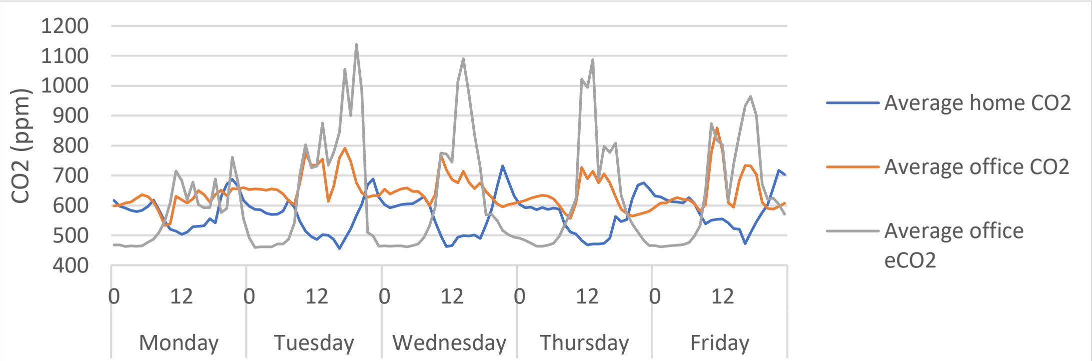

Figure 40 - Average Home/Office CO2 categorised by Day > Hour

When the eCO2 data from the home is also included (Figure 41), the movement in eCO_2 data becomes much more exaggerated. If this were used as a proxy measure for ventilation it would be safe to assume that the participant left the home at around 6:00am but did not enter the office until 9:00. There is also a degree of crossover at around 15:00. Consequently, the same assumptions cannot be drawn from Figure 41 as they can from Figure 37. Given that eCO_2 sensors are highly sensitive to a wide range of environmental conditions, pollutants and gases [204], the erratic behaviour of the sensor could indicate that the increased quantity of fixtures, fittings and furnishings *(present in a home environment)* are saturating the indoor eCO_2 sensors with an increased concentration of airborne pollutants. Thus, these sensors may be less useful for providing a proxy measure of ventilation in environments with a high concentration of pollutants.

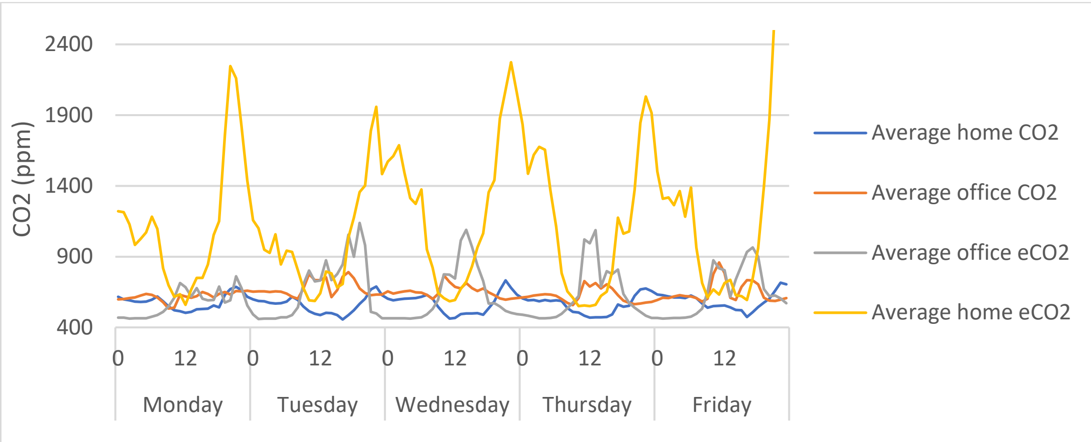

Figure 41 - Average Home/Office CO2 categorised by Day > Hour

10.3.5.3 Particulate Matter ($PM_{2.5}$)

Qualitative findings for dust, pollution and odours indicate that:

> *"On average, the office often has odours in warm weather and occasionally has odours in cold weather, [but] never has more odours than the home, due to the proximity of the living room to the kitchen."*

> *"[The office] is always dustier than the home as trains cause a lot of dust when they pass. However, most of the trains are electric so likely not to emit pollution."*

The recorded highs within the home were consistent with mealtimes (Figure 42), which may indicate that that the sensor was affected by cooking from the kitchen which neighboured the living room. As was identified in Chapter 7, $PM_{2.5}$ sensors can be highly sensitive to certain types of cooking, especially when frying or cooking with oils or fats [222], so this finding is in line with expectations. The highest averages were recorded were typically around 18:00, which is consistent with when an evening meal would be cooked - based on the average time the participant left the office. Therefore, it is important to consider the proximity of PM2.5 sensors from pollution sources. While the data highlights links between cooking and $PM_{2.5}$,

Figure 42 does not indicate baseline values that would suggest the office is dustier than the home, nor did the $PM_{2.5}$ averages indicate regularity in the data that would suggest correlations between timetabled train services. However, if the dust in the office is predominantly caused by airborne debris from passing trains, it is likely that the diameter of the micro particles are larger than the $PM_{2.5}$ sensor can measure.

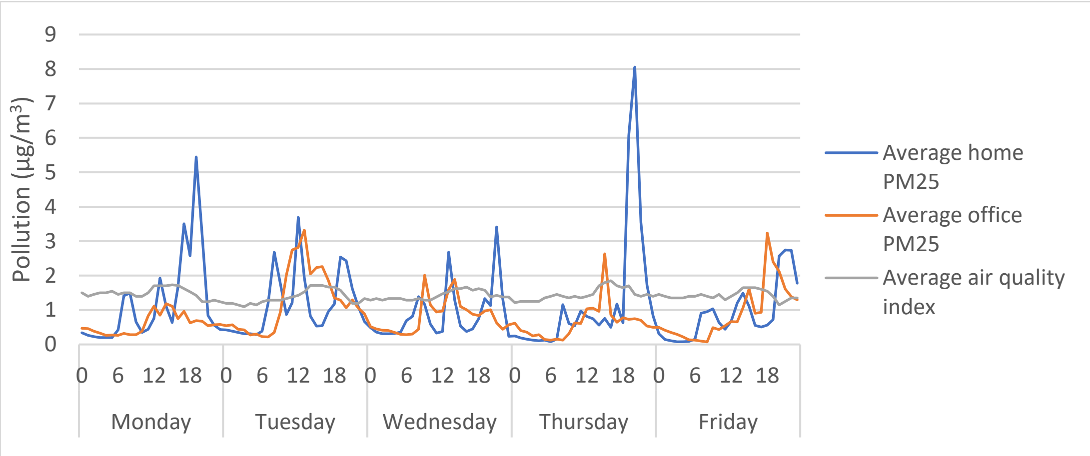

Figure 42 - Average outdoor air pollution vs indoor particulates categorised by Day > Hour

The visual inspection of $PM_{2.5}$ data showed that the both the office $PM_{2.5}$ and outdoor pollution increased around midday. Due to the scaling in Figure 42, it is not clear to see the extent of this, but these trends can be seen more clearly when the home is removed from the graph *(Figure 43)*. Since temperature and humidity were affected by poor insulation, it was expected that the indoor pollution could be influenced by outdoor conditions as well. However, the influence outdoor pollution had over indoor measurements was less significant than either temperature or humidity. However, it is possible that this was due to the distance between the study location and the weather station.

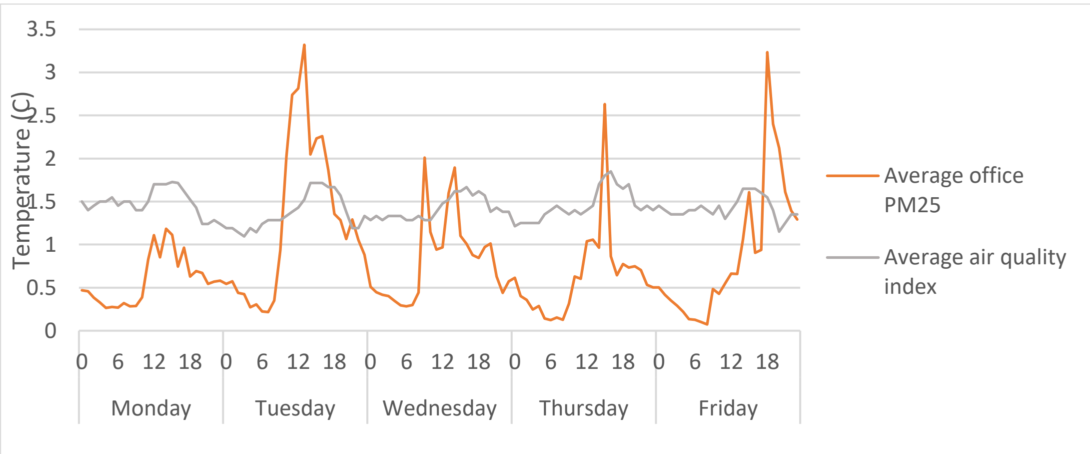

Figure 43 - Average outdoor air pollution vs office particulates categorised by Day > Hour

The *openweathermap* API provides data on both pollution and weather via supplied latitude and longitude coordinates, which are mapped to the nearest weather station. Using the current weather API (Figure 44) the nearest weather station to the study location was Gosforth, Newcastle Upon Tyne, UK. This station is less than 5 miles away from the study *(measured linearly from a to b),* but it is located near one of the central traffic routes leaving to the north of Newcastle City Centre. Thus, it is not possible to determine the accuracy of outdoor pollution for the given study location for data obtained from a weather API. Weather data *(e.g., temperature, humidity)* obtained from this API is unlikely to deviate greatly over this distance, but the same should not be assumed about pollution data.

```
api.openweathermap.org/data/2.5/weather?lat={lat}&lon={lon}&appid={API key}
```

Figure 44 - OpenWeatherMap API endpoint for current weather at given Latitude & Longitude

10.3.6 Study Process

During the closeout interview the participant was also given the opportunity to provide feedback on the level of involvement, this presented the opportunity to gain feedback on the process of collecting data using Alexa, which was a request made by the participant. The participant found the use of Alexa based survey capture a preferable experience overall stating:

> *"I almost created a little relationship with Alexa, which I presume was a sort of indirect relationship with you as the researcher, where it was like I'm going to do that thing today I'm going to do it for Graham for part of his survey"*

This was apparent while the participant was talking about Alexa, as they regularly personified the device making statements such as:

> *"Occasionally, she would interrupt me in the middle of a meeting and ask is I was ready to do the survey."*

The participant also went on to say that the experience of conducting the survey vocally allowed them to engage with the survey without having to switch tasks, reflecting on the personability of the device and the relationship with the researcher throughout:

> *"[when asked if able to do survey during a meeting] I would say no to her, but I think the majority of the time I always made the effort to pick up on our survey and do it later. I don't know if I would have done that if it had been an online or written survey, (...) partly because there was less effort involved on my part to fill the survey in, I could do it whilst I was doing other things, but also that relationship of*

> *talking to a voice almost created an incentive to not let that person down and complete that task right."*

This provides an interesting angle on the problem of reducing the burden towards participants. The three-point assessment conducted daily was originally chosen to complement the TESS approach conducted at the end to evaluate whether shorter surveys, which quantified experience could reduce the burden experienced during longitudinal assessment. The participant highlighted that:

> *"If I had to do an online questionnaire or written questionnaire every day, I would be pretty cheesed off by now, but now I am not as it has been quite an exciting process and I was discussing with my family yesterday saying, 'Graham is coming round tomorrow, he's been collecting my data' and I was quite enthusiastic about it"*

This highlights that the method of data collection had a more notable reduction of burden to the participant than the reduction of responses within the survey. In fact, the participant stated:

> *"If there was an option to add additional information, if you really wanted to, in a free-text way, that would have been useful because I could have said 'it's raining really heavily outside' or 'I'm not feeling well today' (...) that will give you additional information (...) but also it does get a little bit annoying when you have got somebody repeating the same words to you every day."*

This highlighted that the use of the three-point assessment, made the participant feel like they could not provide justifications as to why they were responding in a particular way. They expressed desire throughout for a natural language system, where their responses did not need to be scripted, allowing for qualitative data capture. This provides an interesting line of inquiry as upon inspecting the daily survey data, the three-point responses reduced the statistical power due to missing data and a lack of variation across the variables. The survey data also contained many inconsistencies such as saying they were too hot when it was 17°C in the office, but saying they were too cold when it was 23°C. Additional qualitative data may have provided needed context to make these data useful, but in their current state, they provide little utility and will not be explored further in this thesis.

10.4 Answering the PoI

This study sought to answer **PoI7** through a two-stage analysis approach comprising of a macro and micro level of investigation. However, this chapter was able to provide a substantial contribution to knowledge and sufficient findings to answer this question through the macro-level analysis alone.

PoI7

If deployed longitudinally, can localised sensors with multiple sensing modalities be used to address the subjectivity around environmental perceptions by providing quantitative context to how occupants experience building environments?

Depending on the qualitative data in question, yes, localised sensors *(if the sensors are deployed longitudinally, with multiple sensing modalities)*, can be used to address the subjectivity around environmental perceptions by providing quantitative context to how occupants experience building environments. However, that does not apply to all subjective data.

Electronic sensor data were able to add confirmation to the individual's retrospective experience *(gathered from post study survey)*. Conversely, the participant was able to describe environmental conditions and compare buildings in line with data gathered from IEQ sensors. However, the subjective experiences of the participant did not align to point-in-time measurements. A lack of context surrounding participant specific e.g., general health (i.e., medical history), hormones, nutrition etc, meant that the data lacked statistical power in this context. The regular capture of all the potential determinants of comfort would be incredibly burdensome to people and would interfere with day-to-day activities. Thus, the subjective, open-ended responses offered in the TESS format provide a valuable opportunity to gain insight on individual experience, without overloading the individual with an unrealistic and burdensome line of questioning.

10.5 Further Research

This chapter conducted a macro level investigation of the multimodal IEQ data to evaluate quantitative data captured from sensors within the context qualitative data captured from surveys. In doing so it was able to wholly answer **PoI7**. The next chapter will continue with the 10-step methodology outlined in Chapters 8 & 9 to develop a dynamic regression model for conducting analysis of intra-day outcomes, a micro approach. It is proposed that dynamic regression modelling can expose further time-differential (causal) phenomena and relationships between variables above and beyond what has been exposed here. This could provide further insight into **PoI7** to help better understand the value of longitudinally deployed, mixed-method approaches, which make the individual the unit of analysis.

Chapter 11 Micro-level personalised IEQ: A feasibility study on the dynamic regression analysis of intra-day data

11.1 Introduction

This chapter will explore intra-day data at a micro level, focussing on data obtained from the multimodal sensor, during office hours. The previous chapter (10) demonstrated that quantitative context can be applied to qualitative data on IEQ by longitudinally deploying localised measurement solutions, with multiple sensing modalities. This highlighted value in individualised measurement approaches that focused on a mixed methods approach. This chapter will further address **PoI7** by conducting a suggested dynamic regression analysis on the intraday data with the aim to identify if causal relationships exist between primary and secondary outcomes. The aim of this preliminary, feasibility study in this chapter is to explore whether dynamic regression modelling can increase the statistical power of the quantitative data to identify similar trends through passive sensing or whether additional context can be acquired through the analysis of multimodal intra-day data.

The methods adopted for this chapter were selected due to their focus on individuals and how they can be applied in single participant observational studies, which involve repeated and longitudinal measurements of individuals to identify casual relationships and patterns of behaviour [269]. Although these methods have been exploited for use in many single-case study designs within epidemiological/psychological contexts, no research has applied these methods to the longitudinal monitoring of the environments local to individual building occupants. Thus, this chapter serves two purposes: (i) a detailed undertaking of the novel data handling process and (ii) as an exploratory/feasibility study of the methods proposed for individualised IEQ monitoring. It is suggested that the regression models adopted here could have utility in this domain due to the individualised and longitudinal focus of the data capture.

11.2 Methods

The dynamic regression approach outlined in Chapter 8 requires a single primary outcome. However, since this study is aimed to understand physiological *and* behavioural response to environmental changes, there will be two outcomes and therefore two separate regression models. One model will be constructed with Heart Rate (HR) as the dependent variable and another model will be constructed with Step Count (SC) as the dependent variable. In this way, it should be possible to identify if there are causal relationships between each variable in relation to environmental changes. As the dynamic regression is a novel application, a detailed descriptive process (section 11.3) is provided to outline the purpose of this chapter.

11.3 Application of a micro-level approach

11.3.1 Step 1: Format Dataset

The formatting methods applied in the previous chapter served to provide a dataset for both the previous macro-level analysis and the micro-level analysis that is used in this chapter. The methodological differences in this chapter stem from the need for greater granularity in the timeseries analysis. While it was desirable for this chapter to explore minute-by-minute data without any averaging, methods for synchronising data such as PCHIP *(Chapter 10)* result in data points that are influenced by previous measurements. One of the steps of the dynamic regression modelling is to ensure there is no evidence of autocorrelation in the data and adjust the data to account for it if present. Preliminary examination of the minute-by-minute data resulted in heavily autocorrelated data that could not be feasibly adjusted. For this chapter data were resampled to 15-minute intervals to ensure each interval contained measurements captured by the sensors and to ensure autocorrelation of the dataset could be adjusted.

11.3.2 Step 2: Imputation

To impute missing data, it was important to first understand the causes of the missing data. Moritz *et al.* [279] highlight that there are three degrees of randomness, which can cause distinct distributions in missing data. The term "Missing completely at random" (MCAR) is used when the probability of a missing value is unrelated to other variables *e.g., when a sensor fails to transmit data due to a random technical/connectivity issue.* "Missing at random" (MAR) is used to describe scenarios where the probability of missing data is still independent of the missing measurement, but where an outside influence can be attributed to the event *e.g., a wearable device is taken off for charging.* The last category, "Not missing at random" (NMAR), is used to describe events where there is no degree of randomness resulting in data loss*, e.g., if the value '-1' is always recorded in the event of a measurement outside the parameters of a sensor.*

A portion of the missing values in this study can be categorised as MCAR, due to random connectivity issues resulting in sporadic short periods missing data. For the most part, were relating to the multimodal device losing connection with the router for whatever reason. The devices themselves had built in failovers and would repeatedly, and infinitely, try to re-establish a Wi-Fi link, if the connection was broken.

Issues with the SIM card were also present during this study, which resulted in dropouts where the 4G router required bouncing. The SIM cards were on a month-to-month rolling plan, and the dropouts typically happened during the transition of months. Since these events were caused by an outside influence, but still independent of the measurements, these data were MAR.

During the study *(23/04/2021)* the participant's pet was in their home office and knocked the multi-modal device onto the floor. This resulted in the microcontroller becoming unresponsive and remaining offline until a replacement microcontroller could be installed (which was done on 26/04/2021), this resulted in the data being MAR.

11.3.2.1 Missing Value Analysis

To impute the data for the multimodal monitoring devices, it was important to evaluate to the extent of the data that was missing. When there are <10% missing values, it is applicable to use simple imputation, which replaces the missing values with the mean/median of nearby data [267]. Due to the events causing data to be MAR and MCAR, there were higher percentages of missing data (24-41%) across the variables captured *(Table 21)*. Multiple imputation can be used to impute data when there are >10% missing values, but imputation, *be it simple or multiple*, is an area of research that is generally underexplored [267]. Whichever method of imputation is used, it is important to clearly define the quantity of missing data *(Table 21)* and the cause of the missing data, which can cause distinct distributions in data [279].

Table 21 – Summary of missing values across multimodal variables

	Missing				
	N	**Percent**	**Valid N**	**Mean**	**Std. Deviation**
OFFICE_SOUND	82481	40.3%	122026	45.222	15.201
OFFICE_RH	50651	24.8%	153856	36.006	5.389
OFFICE_PRESSURE	50617	24.8%	153890	1009.846	10.199
OFFICE_PM25	50617	24.8%	153890	0.885	2.381
OFFICE_LIGHT	50617	24.8%	153890	177.901	840.742
OFFICE_TEMP	50617	24.8%	153890	22.031	6.574
OFFICE_eCO2	50617	24.8%	153890	594.983	459.783
OFFICE_CO2	50617	24.8%	153890	617.932	191.474

a. Minimum percentage of missing values for variable to be included: 1.0%

11.3.2.2 Filtering and removal of unneeded data

Before imputing missing values, several cases were removed from the dataset. Since the wearable was not worn at night and the primary focus of the study was the participant in the office environment, data outside of the hours of 06:00 – 23:59 were removed. This time window was chosen as the previous chapter highlighted that, on average, the participant was active during these hours. Equally, since the participant requested the survey capture was only done during office hours, Monday – Friday, all data from weekends were removed from the dataset.

Three days of data were also removed from the dataset due to the recorded event of the damaged equipment. However, since this occurred on a Friday and was rectified by Monday so only two days of valid data were removed *(23/04/2021 and 26/04/21)*. Despite, the removal of filtered data, the percentage of missing data in the dataset still remained >10% for all variables, meaning multiple imputation would be required.

Table 22 - Summary of missing values across multimodal variables after filtering

	Missing		Valid N	Mean	Std. Deviation
	N	Percent			
OFFICE_SOUND	41901	38.80%	66099	46.23562	15.45730
OFFICE_RH	25241	23.40%	82759	35.31714	5.64565
OFFICE_PRESSURE	25230	23.40%	82770	1009.74175	9.84693
OFFICE_PM25	25230	23.40%	82770	0.99133	2.93633
OFFICE_LIGHT	25230	23.40%	82770	229.79925	902.00627
OFFICE_TEMP	25230	23.40%	82770	23.83326	5.91431
OFFICE_eCO2	25230	23.40%	82770	682.30544	607.07773
OFFICE_CO2	25230	23.40%	82770	648.90813	234.39746

11.3.2.3 Multiple Imputation

A 5-pass multiple imputation was performed using linear interpolation to determine the missing values. Initial experiments with the parameters returned promising results, but the imputation process produced outliers that were outside the bounds of any recorded measurements throughout the study. To rectify this, constraint parameters were specified, to ensure the multiple imputation algorithm could not interpolate values outside the original bounds. For example, the eCO2 sensors could not record below 400ppm therefore imputed values below this concentration would be invalid; thus, a minimum constraint of 400 was set. Across all the variables, the minimum imputed value was constrained according to the minimum recorded value *(rounded down to the nearest integer)* and the maximum imputed value was constrained to the maximum recorded value *(rounded up to the nearest integer)*.

Table 23 provides descriptive statistics of the data capture before and after the multiple imputation, highlighting that the minimum and maximum values are within the specified constraints and that the mean/standard deviations also closely match the original dataset. To address the missing data resulting from the EWMA calculations the first 10 values were removed from the dataset resulting a Valid N (listwise) of 108000.

Table 23 - Descriptive statistic of multimodal device variables before and after multiple imputation

Source	Data	N	Mean	Std. Deviation	Minimum	Maximum
OFFICE_CO2	Original Data	82770	648.91	234.40	277.00	2644.50
	Imputed Data	108000	660.57	228.60	258.00	2644.50
OFFICE_eCO2	Original Data	82770	682.31	607.08	400.00	7992.00
	Imputed Data	108000	754.34	582.37	400.00	7992.00
OFFICE_TEMP	Original Data	82770	23.83	5.91	4.29	39.33
	Imputed Data	108000	24.17	5.95	3.83	42.99
OFFICE_RH	Original Data	82759	35.32	5.65	17.61	57.95
	Imputed Data	108000	35.28	5.64	14.14	57.95
OFFICE_LIGHT	Original Data	82770	229.80	902.01	0.00	21150.00
	Imputed Data	108000	359.45	868.58	0.00	21150.00
OFFICE_SOUND	Original Data	66099	46.24	15.46	31.52	109.62
	Imputed Data	108000	48.85	15.36	24.01	109.73
OFFICE_PM25	Original Data	82770	0.99	2.94	0.00	141.00
	Imputed Data	108000	1.34	2.82	0.00	141.00
OFFICE_PRESSURE	Original Data	82770	1009.74	9.85	979.85	1030.82
	Imputed Data	108000	1010.73	9.57	979.85	1031.00

With the missing data imputed the data were then resampled down to 15-minute intervals using the same processes used in Chapter 10 (Appendix F). This is shown in Table 24.

Table 24 – Descriptive statistics for 15-minute data

	N	Minimum	Maximum	Mean	Std. Deviation
OFFICE_CO2	7200	306.933	2632.733	660.568	203.172
OFFICE_eCO2	7200	430.533	7338.367	754.337	526.425
OFFICE_TEMP	7200	4.892	39.196	24.166	5.424
OFFICE_RH	7200	18.516	54.954	35.276	5.187
OFFICE_LIGHT	7200	0.000	16778.667	359.446	742.323
OFFICE_SOUND	7200	33.724	99.351	48.848	12.074
OFFICE_PM25	7200	0.000	51.367	1.338	2.420
OFFICE_PRESSURE	7200	979.960	1030.796	1010.734	9.483
HR	7200	42.294	130.462	64.969	9.462
STEP_COUNT	7200	0.000	113.733	8.012	18.667
Valid N (listwise)	7200				

11.3.3 Step 3: Plot data

The next step involves creating time series plots to visually inspect for variance across the variables. Since the previous chapter presented an extensive visual inspection of the data, this section will only present plots for the dependent variables HR *(Figure 45)* and SC *(Figure 46)*.

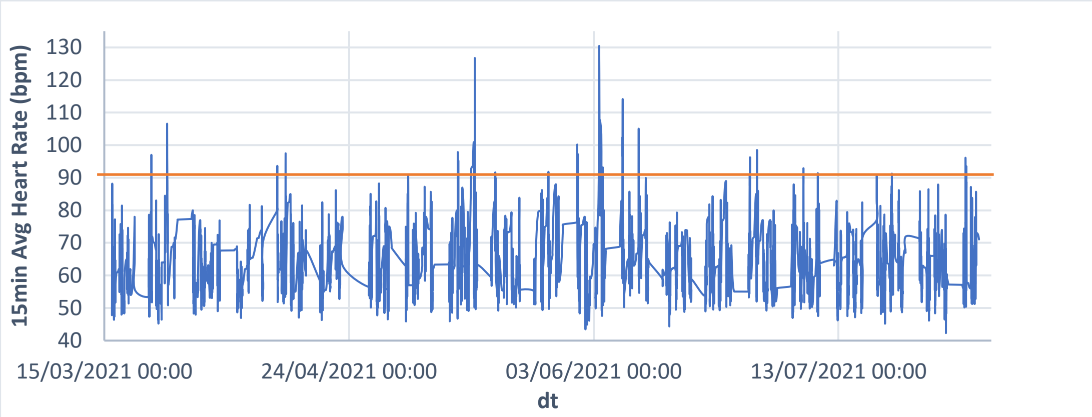

Figure 45 - Timeseries plot for HR

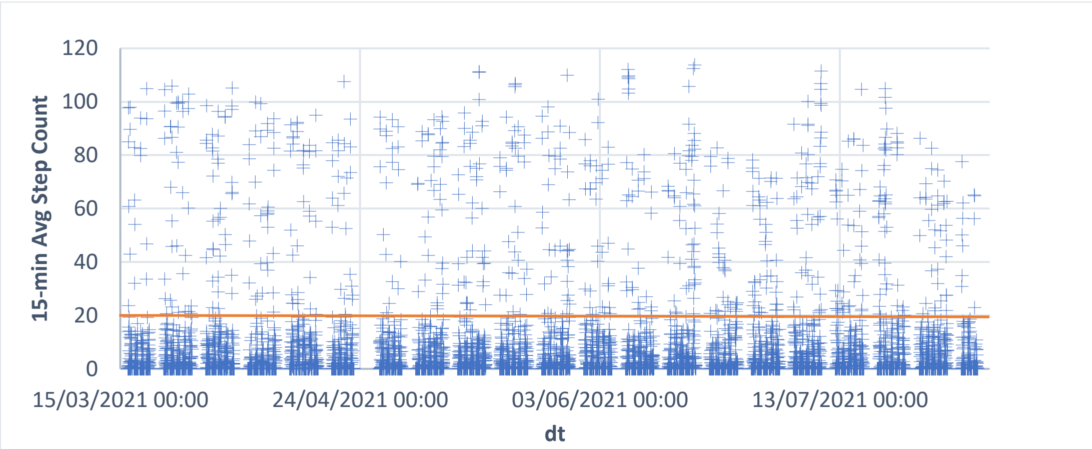

Figure 46 - Timeseries scatter plot for SC

Figure 45 & Figure 46 show that both HR and SC data have a high degree of variation across the dataset, while remaining relatively stationary across the overall timeseries. This is important as to perform dynamic regression modelling, it is important that the data under investigation has relative stationarity, meaning the mean, variance and auto correlation are constant across the timeseries.

11.3.4 Step 4: Assess stationarity

To Assess the stationarity of the outcome variables (HR/SC), the data were partitioned by month to create six separate partitions *(3-8)* and descriptive statistics for each partition were created to evaluate stationarity across the means (Table 25 & Table 26).

Table 25 – Descriptive statistics of HR for monthly partitions

Partition	*N*	Minimum	Maximum	Mean	Std. Dev
March	864	45.2173	106.5846	63.23566	8.618778
April	1440	46.29205	97.50761	64.62398	8.001923
May	1512	45.88045	126.7614	67.10166	10.85441
June	1584	43.45137	130.4623	66.0657	10.46009
July	1584	42.2944	92.92619	63.18645	8.053185
August	216	50.19406	96.13333	64.30387	9.077843

HR can be seen to have approximate stationarity with only 3.92bpm difference across the monthly means and 2.85bpm difference across the standard deviations. While recorded minimums stayed relatively stationary *(7.90bpm difference)* across the partitions, there was more recorded variance across the maximum values *(37.54bpm),* between May and July, which can also be seen in Figure 45.

Table 26 – Descriptive statistics of SC for monthly partitions

Partition	*N*	Minimum	Maximum	Mean	Std. Dev
March	864	0	105.8667	8.468981	20.04528
April	1440	0	107.4667	7.543426	18.5108
May	1512	0	111.2	8.209656	19.37474
June	1584	0	113.7333	7.650337	17.80291
July	1584	0	111.4667	8.773737	18.87382
August	216	0	77.53333	5.003086	12.25665

SC also shows approximate stationarity with only 3.77 step difference across the monthly means but has a 7.25 step difference across the standard deviations. Though this is likely the result of the smaller sample size in March and August. Recorded minimums for steps were all identical (0), which is to be expected based on how missing data were handled for SC. However, the maximum step count varied by 36.20 steps, which can be seen in Figure 46. However, this is also likely due to the smaller sample size of August.

11.3.4.1 Assessment of time trends and periodicity

Since there is a slight variation in HR/SC across the monthly partitions and since the previous chapter identified daily/hourly activity trends, the Month, Day and Hour will be included in the final regression model.

11.3.5 Step 5: Forecasting

Before conducting linear regression, it is important to identify and control autocorrelation in the data. The Autocorrelation Function (ACF) identified correlations across both the HR *(Figure 47)* and SC *(Figure 48)* lags, which is identified by lags that exceed the 95% confidence interval.

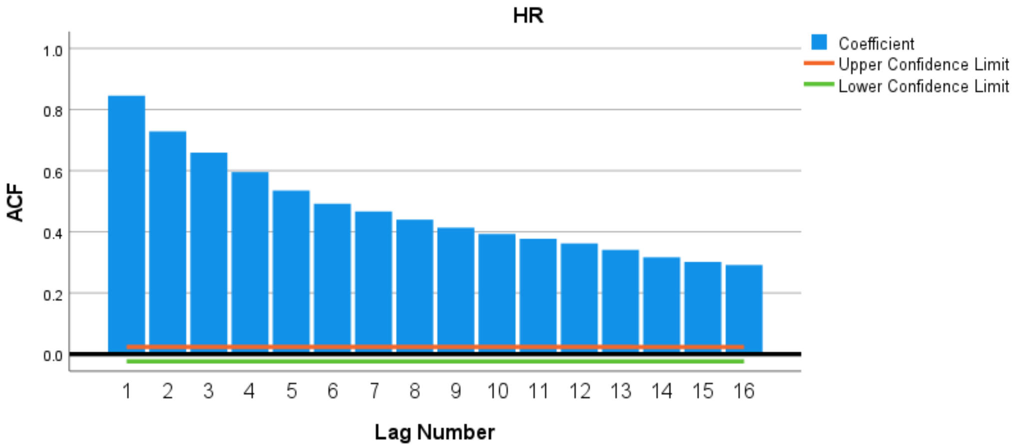

Figure 47 – HR timeseries correlations identified by Autocorrelation Function (ACF)

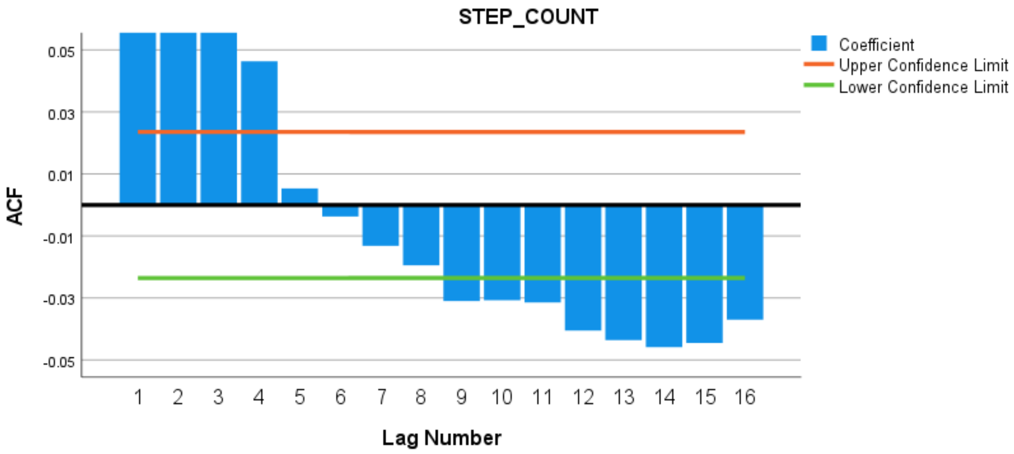

Figure 48 – SC timeseries correlations identified by Autocorrelation Function (ACF)

To correct autocorrelation in the HR data, a Partial Autocorrelation Function (PACF) was used to identify significant lags that should be included in the final model. Figure 49 shows that HR lags 1,2,3,6,7,10 and 11 are outside the bounds of the 95% confidence interval meaning that lagged variables for these lags are required in the final model. However, to remove the autocorrelation fully, lag 16 also needs to be included as this too had an influence on the correlation.

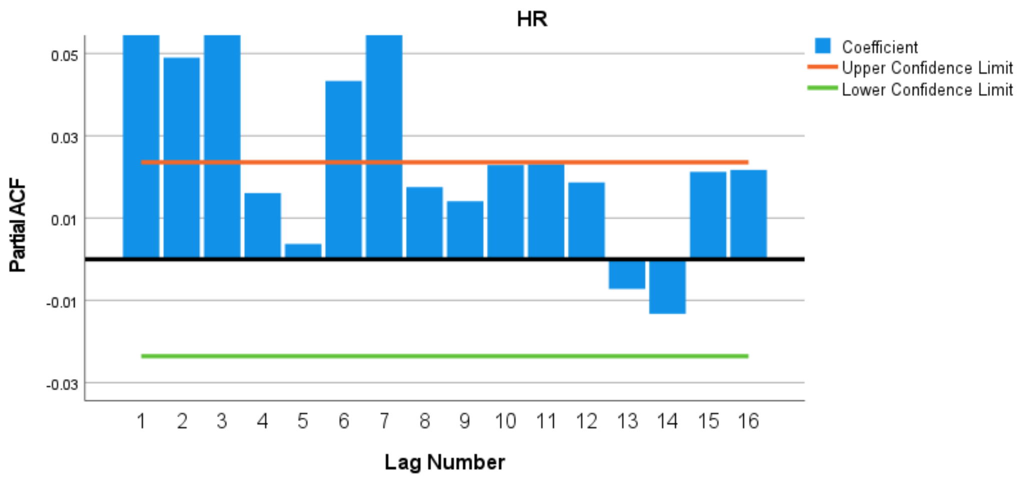

Figure 49 – HR lags identified by the Partial Autocorrelation Function (PACF)

The PACF for SC data highlighted that the lags 1,2,3 and 12 were outside the bounds of the 95% confidence interval *(Table 27)*. Lag 9 was also found to have an influence on the autocorrelation.

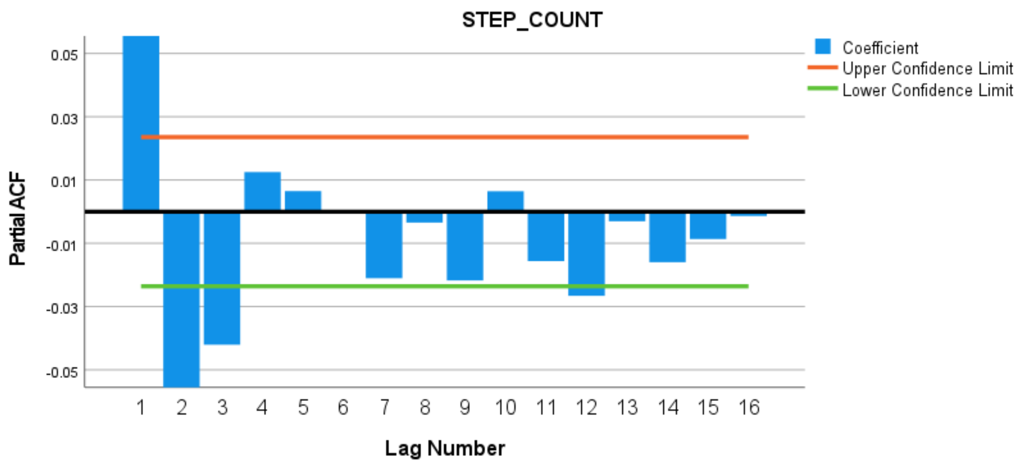

Figure 50 – SC lags identified by the Partial Autocorrelation Function (PACF)

11.3.6 Step 6: Lagged variables

Lagged variables were created for HR lags 1,2,3,6,7,10,11 and 16, to account for the lags that were influential to the autocorrelation of HR data *(Table 27)*. Lagged variables were also created for SC lags 1,2,3,9,12 to account for the lags that were influential to the autocorrelation of SC data.

Table 27 – Lagged variables for HR and SC data

	N	Minimum	Maximum	Mean	Std. Dev
LAGS(HR,1)	7199	42.294	130.462	64.968	9.462
LAGS(HR,2)	7198	42.294	130.462	64.967	9.463
LAGS(HR,3)	7197	42.294	130.462	64.966	9.463
LAGS(HR,6)	7194	42.294	130.462	64.964	9.464
LAGS(HR,7)	7193	42.294	130.462	64.963	9.465
LAGS(HR,10)	7190	42.294	130.462	64.960	9.465
LAGS(HR,11)	7189	42.294	130.462	64.959	9.466
LAGS(HR,16)	7184	42.294	130.462	64.954	9.467
LAGS(SC,1)	7199	0.000	113.733	8.013	18.668
LAGS(SC,2)	7198	0.000	113.733	8.015	18.669
LAGS(SC,3)	7197	0.000	113.733	8.016	18.670
LAGS(SC,9)	7191	0.000	113.733	8.022	18.677
LAGS(SC,12)	7188	0.000	113.733	8.025	18.680
Valid N (listwise)	7184				

11.3.7 Step 7: Confirm autocorrelation correction

With the lagged variables created they can be used to confirm the correction of autocorrelation. To do this, linear regression is used to create unstandardised residuals whereby the dependent variable in the linear regression is one of the primary outcome variables (HR/SC) and the dependent variable is their corresponding lags (Figure 51).

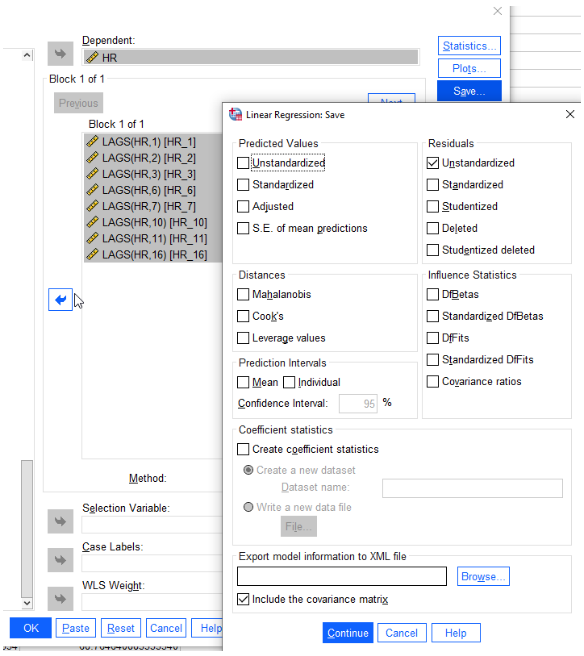

Figure 51 – Example of linear regression configuration for creating unstandardised residuals from lags of dependent variable

The unstandardised residual variable can then be used to determine if autocorrelation has been corrected with the inclusion of the lagged variables. If autocorrelation is accounted for then ACF/PACF functions should return 16 lags for each variable that are all within the 95%CI, which can be seen in both Figure 52 and Figure 53.

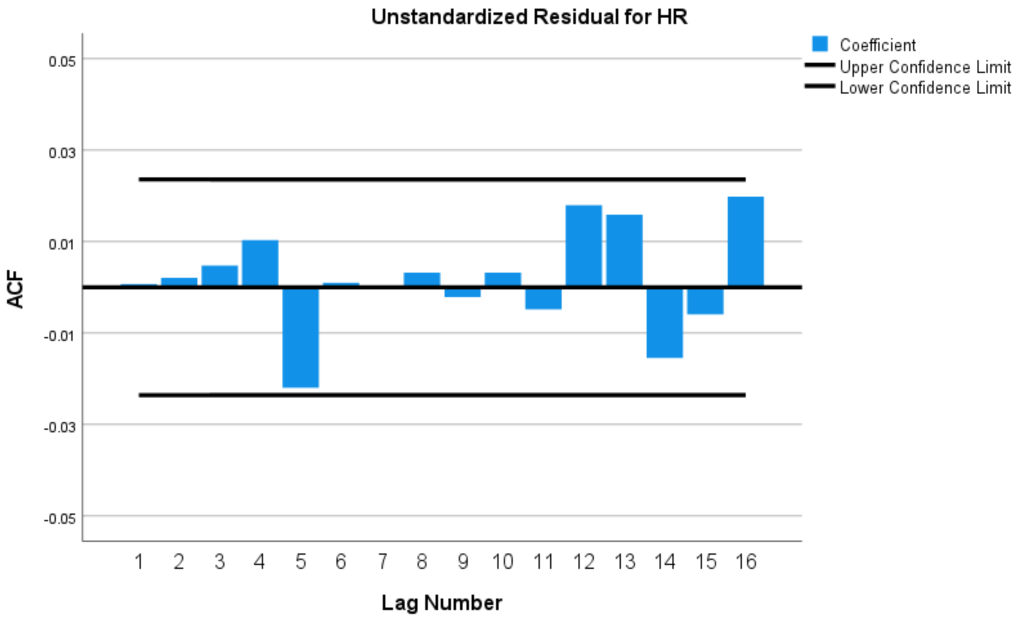

Figure 52 – ACF results for unstandardised residual of HR data, confirming the correction of autocorrelation.

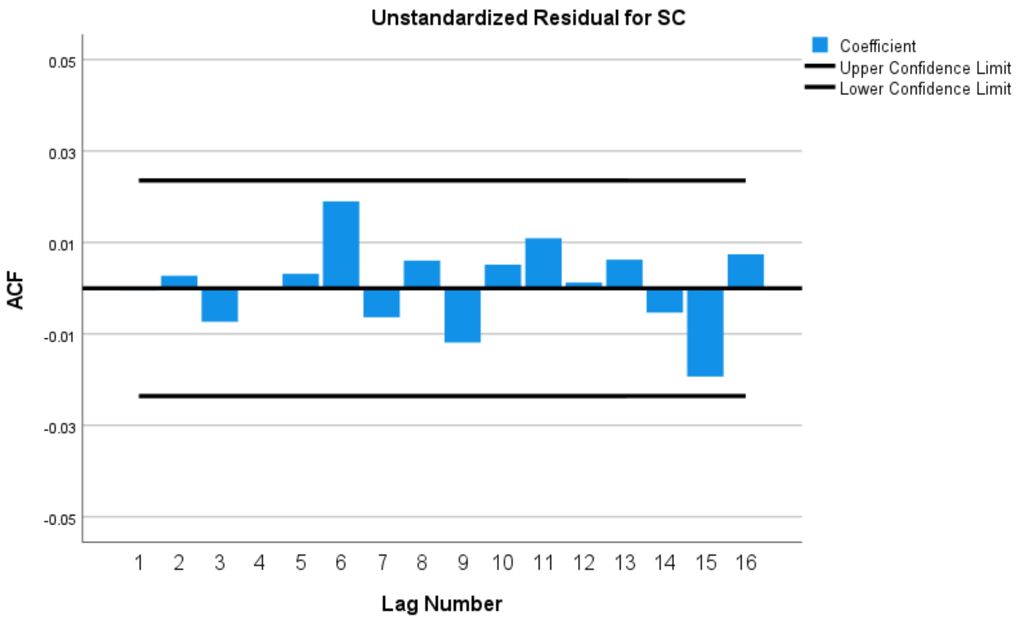

Figure 53 – ACF results for unstandardised residual of SC data, confirming the correction of autocorrelation.

11.3.8 Repetition of steps 5 -7 for independent variables

Before building the dynamic regression model it is important to also adjust any autocorrelation found in the independent variables. Thus, steps 5-7 were repeated for each of the independent variables under analysis. Table 28 outlines which lags will be included in the final model for each variable.

Table 28 – Autocorrelation adjustments for independent variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
CO2	✓	✓	-	-	✓	-	✓	✓	✓	-	-	-	-	-	✓	-	-	-	-	-
eCO2	✓	✓	-	-	✓	✓	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-
TEMP	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	-	-	-	✓	-	✓	-	✓	✓
LIGHT	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LIGHT	✓	✓	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SOUND	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	-	✓	-	-	-	-
PM25	✓	✓	✓	✓	✓	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-
PRESSURE	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	✓	-	-	-

11.3.9 Step 8: Dynamic regression

Two dynamic regression models were created. One for each of the dependent variables (HR/SC), which included Month, Day, Hour to account for time trends and periodicity. Lagged variables were also included for each dependent and independent variable in each model to account for autocorrelation identified in steps 5-7. Many outputs were created from the dynamic regression modelling, but the primary indicator of relationships between the dependent and independent variables is the coefficients output *(Table 29 &* Table 30*)*.

Table 29 – dynamic regression coefficients for HR model

	Unstandardized Coefficients		*Standardized Coefficients*			*95.0% Confidence Interval for B*	
	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound
Constant (HR)	7.740	7.429		1.042	0.298	-6.823	22.303
CO2	0.000	0.001	-0.007	-0.380	0.704	-0.002	0.001
eCO2	0.000	0.000	-0.009	-0.657	0.511	-0.001	0.000
TEMP	-0.110	0.050	-0.063	-2.197	0.028	-0.209	-0.012
RH	0.051	0.059	0.028	0.850	0.395	-0.066	0.167
LIGHT	0.000	0.000	0.016	1.704	0.088	0.000	0.000
SOUND	-0.003	0.014	-0.004	-0.221	0.825	-0.031	0.025
PM25	-0.023	0.049	-0.006	-0.468	0.640	-0.118	0.072
PRESSURE	0.017	0.067	0.017	0.252	0.801	-0.115	0.148

Table 30 – dynamic regression coefficients for SC model

	Unstandardized Coefficients		*Standardized Coefficients*			*95.0% Confidence Interval for B*	
	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound
Constant (SC)	4.886	19.412		0.252	0.801	-33.167	42.939
CO2	-0.014	0.002	-0.155	-6.235	0.000	-0.019	-0.010
eCO2	0.000	0.001	-0.008	-0.411	0.681	-0.002	0.001
TEMP	0.516	0.132	0.150	3.909	0.000	0.257	0.775
RH	0.383	0.156	0.107	2.451	0.014	0.077	0.689
LIGHT	0.000	0.000	-0.012	-0.918	0.358	-0.001	0.000
SOUND	-0.156	0.038	-0.101	-4.143	0.000	-0.230	-0.082
PM25	0.028	0.127	0.004	0.220	0.826	-0.221	0.277
PRESSURE	0.018	0.176	0.009	0.102	0.919	-0.327	0.363

11.4 Results

11.4.1 Step 9: Interpret the results

To determine if independent variables are predictors of the constant, the span of the 95%CI must not cross over 0; if this is satisfied, the regression coefficient can then be used to determine a positive or negative relationship between the variables and can determine the significance of the relationship [267].

In the HR model, Office Temperature was the only found to have 95%CI bounds that did not cross over 0 (95%CI: [-0.209, -0.012]). Yet, while this would indicate that temperature is a predictor of HR change, the low regression coefficient (-0.110) is not indicative of a significant relationship. Moreover, temperature was heavily autocorrelated and required lags beyond the default 16 to be created. In the SC model, more variables were found to be potential predictors of SC change *(CO2: [-0.019, -0.010], Temperature: [0.257, 0.775], Relative Humidity: [0.077, 0.689] and Sound [-0.230, -0.082])*. However, like HR, there was limited significance in the regression coefficients.

11.5 Discussion

11.5.1 Step 10: Reporting

The aim of this chapter was to detail the application of a micro approach while extending the inquiry of **PoI7**. Specifically, to understand whether a micro approach of dynamic regression modelling could increase the statistical power of the quantitative data to identify similar trends through passive sensing or whether additional context acquired through the analysis of multimodal intra-day data. As this is a novel application of that approach in this field, it was necessary to provide a detailed account of the step-by-step process for clarity. The results of the preliminary study presented in this chapter suggest that the methods adopted, may have limited utility in the approach taken. Specifically, when applied in the context of multi-modal and personalised IEQ monitoring there seemed to be no additional context to the findings obtained through a macro approach (Chapter 10). Though, it should be noted that this may not be a limitation of the methods themselves, which have been used across a range of single participant observational studies in other fields [177].

11.5.2 Limitations

The data capture frequency used in this study was likely faster than environmental changes in the study location, therefore, repeated similar measurements are recorded across the timeseries. This would likely cause autocorrelation across the data, which can impede the success of the dynamic regression modelling process. This was likely exacerbated by events leading to data being missing at random, which resulted in the need for varying degrees of data imputation.

Many of the independent variables were severely autocorrelated, requiring most lags to be included in the final model. Moreover, for temperature and pressure additional lags (beyond the default 16 lags) were introduced to try and control the autocorrelation. McDonald et al. [267], report that more than two lags are rarely required, but note that recurring autocorrelation patterns could exist that require more lags. The significant number of lags used to ‘correct’ the autocorrelation cannot be ignored here, though again there is a significant difference in the frequency of measurements between this work *(1 measurement approximately every 40 seconds, resampled to 1 measurement in 15 minutes)* and the single-case studies highlighted [177] *(one daily measurement)*.

11.6 Addressing the PoI

It was suggested that micro-based (dynamic regression) modelling could expose further time-differential (causal) phenomena and relationships between variables in addition to a macro approach (Chapter 10). While the feasibility (preliminary) statistical methods used to analyse data in this chapter were not able to provide any additional context from what was already obtained in chapter 10, the methods used to gather data for this chapter ultimately led to the exploration conducted in chapter 10.

Clear value was noted in the last chapter regarding the collection of individualised and longitudinal data, demonstrated by the ability to wholly answer **PoI7** from those methods alone. The mixed methods approach provided clear value against the dynamic modelling methods applied in this chapter. However, there is a clear need to explore n-of-1 methods in more detail within this domain, as they are a form of single-case study, but they focus on longitudinal, repeated measurements of individuals, which enable the exposure of time-differential phenomena and causal links between measured outcomes [269] and could be useful in this domain.

11.7 Conclusion

Since this was an initial exploration of these methods *(which are being applied outside of their intended context)*, the findings of this chapter by no means suggest that these methods have no utility in this research domain. On the contrary, it is felt that the statistical analysis methods *(often adopted in n-of-1 studies)* could provide a lot of value to building science research, due to their ability to analyse causal relationships between measured variables as well as behavioural changes in individuals. This opens a whole new line of enquiry, which could pave the way of future research in this domain.

Chapter 10 highlighted that clear links between the participant and the building can be obtained from the outcomes under interrogation. Therefore, further research would need to be conducted to understand how these outcomes can be best applied in an n-of-1 context, though this work clearly extends beyond the scope of this thesis. To do this, the feasibility of repeated intra-day measurements from multimodal devices would need to be explored to assess the impact these data have on the dynamic regression modelling approaches traditionally used in n-of-1 studies.

Having now provided a body of work that has explored each of the points of inquiry in detail, the following chapter will summarise the thesis and explore the practical implications and contributions to knowledge that this work has provided and its industrial applications.

Chapter 12 Conclusion

12.1 Introduction

People spend an increasing amount of their time within indoor environments, so there is a need to better understand IEQ through pragmatic and attainable methods. Achieving robust and scalable IEQ measurement could better inform the health and wellbeing of building occupants. Through, my experience of working in industry (with the industrial sponsor of this Ph.D.), I often recognised a desire to adopt environmental monitoring approaches through digital technologies. Typically, that was met with reluctance due to the e.g., cost and complexity of the available options. Accordingly, this research was positioned around a common industrial requirement, routine objective monitoring of environments with scalable (i.e., affordable) IEQ monitoring approaches.

Several points of interest (PoI) were posed which formulated a research path on which to conduct this Ph.D. Those PoI were built upon the current industrial challenges, where their academic credibility was evidenced from the literature, which explored the state-of-the-art and emergent solutions for monitoring IEQ. Accordingly, a central hypothesis was positioned for my thesis, that quantitative measurement of the environmental conditions local to individuals could add spatial density to 1) reinforce data pertaining to how building occupants experience indoor environmental conditions and 2) provide additional context to current approaches for data capture, which traditionally focus on qualitative data capture.

Evidence to support that hypothesis could provide direct research impact by enabling practitioners to monitor buildings with a high spatial density and to learn not only about intra-building variability, but the localised environmental conditions experienced by individual building occupants. It is plausible to suggest that the work arising from a confirmation of my hypothesis could provide a technical pathway/framework for monitoring in buildings, which could provide a better understanding of the effect environmental changes have on the health and wellbeing of building occupants and/or associated research. Moreover, this would address an industrial need to go beyond sandboxed research environments, by exposing pragmatic approaches for quantitative data collection in buildings, which is traditionally too costly or complex to implement.

Throughout this thesis the chapters have been constructed such that they address and answer specific PoI (presented in Chapter 1). Each subsequent chapter addressed a series of PoI that were evident and rooted in industrial challenges. In some cases, multiple chapters were dedicated to a single PoI due to the rigor needed for a comprehensive approach. Here, a synopsis of the thesis is presented according to each PoI, to draw it to a close. This chapter will outline the contributions to knowledge and discuss the practical implications this thesis has on future researchers and industry practitioners.

12.2 Addressing the points of inquiry

Here, I provide a reminder to all the PoI with corresponding answers. That will lead me to answer the fundamental research question and hypothesis posed within this thesis.

12.2.1.1 PoI1

What constitutes IEQ and how is it currently measured?

There is no standardised definition of IEQ, but it is generally regarded as a multifaceted environmental outcome. To date, it is measured in an ad-hoc manner but constitutes several important environmental domains*: air quality, thermal comfort, visual comfort, and acoustic comfort*, with each of those comprising a series of outcomes including: *temperature, humidity, light, noise, dust, volatile organic compounds, carbon dioxide (CO_2) and equivalent carbon dioxide (eCO_2).*

IEQ factors comprise of both subjective and objective determinants of environmental quality. While subjective (qualitative) data is measured predominantly with pen and paper self-reported surveys, objective determinants can be monitored through data captured with electronic sensors or from pen and paper, self-reported surveys. Although surveys are useful (and highly scalable) their subjective nature means they lack e.g., absolute clarity on personal perceptions and are burdensome due to their often-time-consuming requirements. In contrast, objective (quantitative) data measurements are often omitted, due to cost and complexity of electronic-based monitoring IEQ.

For studies that do use quantitative approaches a range of technologies can measure IEQ depending on the underlying infrastructures available in the building under analysis. If buildings have complex HVAC systems (typically seen in commercial buildings), measurements can be taken from state-of-the-art measurement equipment that is integrated into the air handling units and managed through a BMS. However, if buildings lack these infrastructures environmental monitoring can be retrofit into the building using portable monitoring devices, or data loggers placed in situ. Again, these devices are costly and have limited sensing modalities – meaning multiple devices are often required to comprehensively capture IEQ.

12.2.1.2PoI2

What is the current state-of-the-art in environmental monitoring?

The current state-of-the-art focuses heavily on the use of costly, reference-standard, measurement devices. While these devices were recognised to have implicit accuracy, the latter was often offset due to the lack of spatial density in the measurements, because of device placement. Typically, devices are placed in a central location to monitor a space from a single monitoring device. Consequently, the technology may not respond to localised changes in large spaces meaning that measurements may not reflect the environmental quality experienced by all occupants. Moreover, many of these devices each cost several thousands of pounds (often for a single IEQ outcome), meaning they have limited utility in e.g., non-commercial spaces.

There is a lack of IEQ monitoring from an individual perspective, suggesting there is a need for a paradigm shift that makes the individual the unit of analysis within the context of building monitoring. Accordingly, consideration of low-cost IEQ sensing approaches is warranted. Furthermore, wearables and cloud-based platforms could create holistic, personalised monitoring systems, which could aggregate data from multiple locations. Evidence suggests that such approaches from multiple sources could provide more scalability to achieve better (objective) and more spatially dense IEQ data.

12.2.1.3 PoI3

> *What sensor technologies are used to capture IEQ?*

In commercial buildings, state-of-the-art sensors are commonly integrated directly into HVAC systems. These types of sensors can provide accurate monitoring of e.g., air velocity, air quality, temperature and humidity using probes that are fitted into ducts, and air handling units of the HVAC system. These probes contain discrete analogue and digital circuitry that can capture changes in air quality through optical responses or changes in electrical resistance. If buildings lack these infrastructures, monitoring can be retrofit into the building using portable devices, or data loggers placed in situ. These devices can also be costly as they have comparable sensors to those used in HVAC systems. However, instead of probing a HVAC unit these devices have sensors integrated into the device. Unfortunately, many data loggers lack Wi-Fi connectivity, meaning that data must be manually downloaded.

Low-cost sensor technologies are increasing in popularity with multiple studies using both state-of-the-art and low-cost sensors in tandem. Low-cost sensors can measure a wide range of measurements and use similar technologies to those found in data loggers, but the components are often lack the resolution compared to state-of-the-art sensors, are made more cheaply, or use completely different technologies. For example, state-of-the-art CO_2 sensors measure optical responses using infrared beams to detect particle concentrations in a chamber. In contrast, metal oxide sensors are often used in cheaper sensors that may claim to measure CO_2, but measure an electro-chemical resistance that is equated to CO_2.

12.2.1.4PoI4

> *What are the optimal approaches to aggregate data from numerous devices and settings, including settings without existing monitoring infrastructures?*

Microcontrollers were identified as optimal approaches to aggregate data from a multitude of sensors in near real-time at a relatively low cost. They enable the intra-device aggregation of data from multiple sensors and can transmit data to cloud services for continued integration with other devices. It is recognised that other solutions exist, such as off-the-shelf monitoring solutions and FPGAs, but their increased cost makes them less scalable and thus, less fit-for-purpose within the context of the overarching industrial requirements of the Ph.D.

12.2.1.5PoI5

> *Can low-cost sensors be used to gather with precision, consistency, and reliability across different environments i.e., home, and commercial settings?*

Depending on the purpose of the measurements, yes, low-cost sensors can be used to gather with precision, consistency, and reliability across different environments i.e., home, and commercial settings. However, low-cost sensors have recognised limitations, which may make them unsuitable for certain measurement scenarios. If highly accurate measurements *(to 'nth' decimal place)* are required, it was demonstrated that the low-cost sensor technologies would likely be unsuitable. However, for the purposes of providing a general indicator of environmental conditions *(for the purposes of continuous monitoring solutions),* low-cost sensors were shown to have pragmatic use. Moreover, low-cost sensors can demonstrate good inter-sensor reliability, meaning that they could be deployed across different environments and the measurements of two sensors could be compared with one another.

12.2.1.6PoI6

> *How can multiple sensing modalities be pragmatically deployed to gather data for the longitudinal assessment of individuals building occupants?*

Wi-Fi-enabled, multi-modal monitoring devices could provide a mechanism to pragmatically deploy multiple sensing modalities local to individual occupants across a range of building environments. While data from intra-device sensors can be aggregated internally, data can be transmitted to cloud platforms for aggregation with other devices. The use of mobile 4G routers and smart power solutions provides increased security and control over the hardware when it is deployed longitudinally in the field, to reduce the need for participant interventions and to provide the researcher with a mechanism to monitoring, control and troubleshoot sensors remotely. Including the participant in the study design can enable a tailored methodological approach, whereby individuals can contribute to the study design to provide a sense of ownership and control to reduce burden and increase participant adherence.

12.2.1.7PoI7

If deployed longitudinally, can localised sensors with multiple sensing modalities be used to address the subjectivity around environmental perceptions by providing quantitative context to how occupants experience building environments?

Depending on the qualitative data in question, yes, localised sensors *(if the sensors are deployed longitudinally*, *with multiple sensing modalities)*, can be used to address the subjectivity around environmental perceptions by providing quantitative context to how occupants experience building environments. However, that does not apply to all subjective data.

Electronic sensor data were able to add confirmation to the individual's retrospective experience *(gathered from post study survey)*. Conversely, the participant was able to describe environmental conditions and compare buildings in line with data gathered from IEQ sensors. However, the subjective experiences of the participant did not align to point-in-time measurements. A lack of context surrounding participant specific e.g., general health (i.e., medical history), hormones, nutrition etc, meant that the data lacked statistical power in this context. The regular capture of all the potential determinants of comfort would be incredibly burdensome to people and would interfere with day-to-day activities. Thus, the subjective, open-ended responses offered in the TESS format provide a valuable opportunity to gain insight on individual experience, without overloading the individual with an unrealistic and burdensome line of questioning.

12.3 Addressing the Thesis Statement

By presenting a comprehensive body of work, this thesis has gathered enough evidence, through primary and secondary research, to address the overarching hypothesis and research question outlined in the Thesis Statement.

12.3.1 Research question

Can localised sensors provide richer data that will enable better understanding of causal relationships i.e., how individual building occupants respond to environmental changes?

12.3.2 Answer to research question

Yes, the use of localised sensors, if deployed longitudinally, can provide richer data that can enable a better understanding of causal relationships between environmental changes and occupant responses. There were clear demonstratable links between physiological responses *(Heart Rate, Steps)* and environmental conditions (Temperature, Humidity, CO_2). The sensor data were also able to provide affirmation of the participant's subjective, comparative feedback of both building environments. These links were achieved through adoption of a mixed-method approach and longitudinal monitoring of an individual. Realisation of this question could not have been achieved from traditional approaches to building monitoring. This has the potential to change how building performance is evaluated, as well as providing evidence towards the confirming the overarching hypothesis of this thesis.

12.3.3 Hypothesis

Quantitative measurement of the environmental conditions local to individuals could add spatial density to 1) reinforce data pertaining to how building occupants experience indoor environmental conditions and 2) provide additional context to current approaches for data capture, which traditionally focus on qualitative approaches.

12.3.4 Addressing the Hypothesis

The quantitative and multi-modal monitoring devices developed as part of this thesis provide a scalable solution for monitoring a wide range of IEQ outcomes that has demonstrable benefit within this domain. Their low cost makes it feasible to deploy them in a pragmatic manner, and at an individual level. In doing so, this hypothesis has been evidenced throughout this thesis, demonstrating that objective and quantitative data can be used to reinforce understanding pertaining to how building occupants experience indoor environmental conditions, and can also provide rich context to traditional, qualitative approaches for building performance assessment.

12.4 Discussions and Conclusions

This thesis presents a novel approach for IEQ measurement including the development, validation, and deployment of a multimodal device with IoT technologies to provide scalable and individualised approaches. The thesis proposes new methodologies for a pragmatic but contemporary approach in the field of building sciences, exposing causal relationship which could prove useful in future health and wellbeing research. This is especially important now as recent times have shown why there is a need for better monitoring enclosed spaces at individual level (e.g., SARS-CoV-2 pandemic of 2019). Of note, the thesis identifies that while there are use cases where the scientific validity of low-cost sensors could be questioned, this does not mean they should be disregarded as unfit for all use cases.

Evidence gathered in this thesis highlights that localised monitoring of IEQ can provide sufficient context to qualitative data if the individual is made the unit of analysis. If done in this way, traditional studies *(that require large sample sizes to normalise the subjectivity in the data)*, could be adapted to use passive sensors to validate the responses, thus preserving the nuances of individual experience. This provided a direct response to the initial hypothesis and confirmed that low-cost sensing solutions can provide scalable and robust approaches to monitor individual building occupants and their IEQ.

12.4.1 Contributions to knowledge

This thesis presents the following contributions to knowledge:

- The identification of an industrial and research requirement for paradigm shifts in the construction sector that place individual building occupants as the unit of analysis within the context of environmental monitoring of buildings.
- Demonstrable processes for aggregating data from a multitude of a sensor sources ranging from reference standard equipment, proprietary consumer devices and bespoke devices.
- A detailed framework for the development of a low-cost multimodal monitoring device capable of capturing a wide range of environmental factors that contribute to IEQ.
- A novel approach to sensor validation within building sciences was explored that utilised statistical analysis from other fields e.g., epidemiology. This will help extend upon traditional approaches for validating sensor technologies. Bland Altman agreement analysis was used to determine the consistency and absolute agreements of measurements against a reference standard. Intra-class correlations were used to determine the reliability of multiple randomly selected low-cost sensors. By evaluating the accuracy, precision, absolute agreement, consistency, and reliability of low-cost sensors, it was possible to identify their use-case specific applications and identify whether they were fit-for-purpose. This approach to sensor validation provides evidence to challenge the stigma surrounding the use of low-cost sensors within research.
- The demonstrated need to separate “validity” and “fit-for-purpose”, when exploring emergent hardware and software.
- The development of a two-stage methodological protocol for analysing micro and macro level data captured longitudinal from a single building occupant to understand causal covariate relationships.
- The demonstration and evidencing that the combination of qualitative and quantitative data, collected in a longitudinal individualised context – from a single participant, can provide deep insights into both personal experience and building performance.
- The identified value remote control contingencies on longitudinal remote assessment projects, to interact, monitoring and power-cycle remote monitoring equipment with minimal/no participant involvement.

- The identification of methodological limitations surrounding the use of dynamic regression modelling for analysing longitudinal sensor data captured at much greater frequencies than data typically captured in n-of-1 observational studies.
- The exposure of knowledge gaps warranting further research into the design, specification and applicability of n-of-1 study designs within building science research.

12.4.1.1Significant original contribution to knowledge

This thesis has answered several important PoIs, driven by industry but supported by the literature. Subsequently, my thesis identifies a significant original contribution to knowledge.

> The longitudinal deployment of low-cost, multimodal, monitoring devices can add a high degree of context to subjective environmental assessment surveys *(designed to capture information on IEQ and building performance)*, if the sensors are localised to individuals and the study makes the individual the unit of analysis. IEQ sensor data not only provides context to open-ended, subjective feedback, but can provide quantitative affirmation of responses, which could be used to reduce the number of participants required to evaluate building performance or reduce the generalisation of individual responses.

12.4.2 Impact of SARS-COV-2

Since this work was conducted throughout the SARS-COV-2 pandemic, there were several ways the pandemic impacted this Ph.D. The first was during the case study #1 *(Chapter 5)*, which was mobilised only 6 days before the UK went into an extensive national lockdown. The study was initially designed to monitor a participant in both their home and their workplace, which became unfeasible task after lockdown and closure of the latter. Regardless, informative findings were gathered during the setup, mobilisation, and aggregation of data from off-the-shelf devices which proved useful for later work. Secondly, supply chain issues made obtaining sensors difficult for the major study (Chapter 10 and 11). Additionally, I originally envisioned the final study to be a comparative analysis of multiple employees from a single commercial office, while also monitoring their homes. The original format of the study mirrored that of the final study I deployed (a single-case observational study), but repeated in parallel across multiple participants in a single workplace but multiple homes. To overcome, I amended the study design to focus on a single participant with an annexed home office.

Despite some major setbacks, the UK lockdown(s) challenged my technical abilities for the better. For example, I placed additional contingencies in place on the final study to ensure I had full remote control over the devices if other lockdown restrictions were put in place. This enabled me to remotely conduct in-the-field modifications, without having to involve the participant. The pandemic afforded me first-hand experience of real-world challenges, allowing me to become agile and adaptive to outside influences. I feel those experiences not only helped the progression of this thesis, but also prepared me for real-world research.

12.4.3 Practical implications

The outcomes of this Ph.D. could have practical implications for future researchers and industry practitioners alike. This thesis has outlined evidence from literature to support the claims of the sponsors, that state-of-the-art measurement equipment have limited utility outside of research and are not feasible in real-world building projects, due to the cost and complexity of the equipment. Consequently, this thesis has explored low-cost and emergent technologies that could have practical implications in industry. While the development of a mass-produced, multimodal monitoring device was deemed outside the scope of this thesis, the thesis has identified a methodological approach for designing, developing, and validating a multimodal monitoring device that uses sensors and microcontrollers that are low cost and accessible. This thesis also demonstrates nuanced approaches to deploy that technology, ensuring a more seamless integration to individual monitoring with low burden. This has the potential to influence the development of bespoke monitoring equipment that designed around the outcomes of this thesis, but that are designed around the specific requirements of buildings.

The contribution to knowledge has significant practical implications for industry practitioners. The findings from later studies in this thesis have identified the need for a paradigm shift in building assessments, whereby the individual becomes the unit of analysis. The traditional approaches of survey-based capture are still recommended, but the addition of localised IEQ monitoring and wearable devices has been shown to provide a wealth of context surveys that can reduce (and in some cases remove) the subjectivity typical in qualitative data capture. Practitioners should explore ways to integrate longitudinal, localised monitoring devices into buildings to provide a spatially dense, context to building performance assessments. The additional context can also be used to compare different building types, so it is feasible to hypothesise that if this became a standard practice in building developments, future buildings could become rich data stores that can be analysed to challenge current building standards and drive innovation across the construction industry. Machine learning and artificial intelligence (AI) could then be adopted to utilise big data arising from buildings as they evolve over time.

12.4.4 Future Research

The practical implications for industrial practitioners (outlined above), also have the potential to drive a great deal of research impact. Researchers should seek out to support industry practitioners with the research and development of bespoke hardware, as well as seek to support the deployment of spatially dense, low-cost monitoring devices. Consequently, researchers can maximise the impact of their research and ensure research excellence. It is important that researchers challenge the stigma around low-cost devices and identify solutions that are fit-for-purpose, while ensuring that their findings and outputs have pragmatic value within industry led projects.

This thesis demonstrated that making the individual the unit of analysis is an important step in creating a paradigm shift in building assessments. This thesis focused on the specific technologies involved to feasibly conduct such research. However, there needs to be comprehensive research into what types of single-case study methodologies exist and how they could be applied building sciences to support the identified technological approaches. It is anticipated that this thesis could lay the foundations for a variety of future Ph.D. theses. These could undertake a deeper exploration of the measurement hardware from an electrical engineering perspective, or explore methodological approaches from architecture, engineering, and health perspectives.

While the specific single-case methods applied in this research provided more questions than answers, there are a wealth of successful single-case studies that utilise more robust n-of-1 methods with a wide range of methodological processes that could be explored in this domain. There is the potential here to explore an n-of-1 approach tailored specifically for building science research add a lot of value to IEQ research and building performance assessments, as the longitudinal approach to data collection and the individualised focus of the research feels ideally suited to the principles of n-of-1. However, it is anticipated that this would require a comprehensive literature review of n-of-1 studies to sufficiently understand the rules and methodological processes involved for conducting such a study, and how they could be applied to building science research. Given their application within epidemiological domains, robust n-of-1 methods could even open up the potential to explore more detailed health metrics in this context, by the capturing detailed medical history within initial interviews, while capturing additional data from an array of WHTs. This opens new lines of questioning that could be the foundation of one or more Ph.D. theses.

12.5 Closing Summary

This thesis has embarked on a journey to better understand how building occupants experience environmental conditions within buildings and whether emergent technologies can provide additional context to existing approaches to assess environmental quality in buildings. In addressing the research questions and hypotheses positioned for this body of work, this thesis has uncovered several contributions to knowledge and arrived at a unique and substantial contribution to knowledge that has practical implications for future researchers and practitioners alike. By focusing on individuals, it is hoped that future buildings could be designed around the health and wellbeing of individual occupants and that human experience becomes a key component of this paradigm shift. This work has demonstrated that emergent low-cost technologies can provide the technological infrastructures to underpin the future of research in this field. It is hoped that through a combination of technology, methodology and real-world research, the indoor environments where people spend the majority of their time can be healthier environments that are tailored and adaptable to the comfort and wellbeing of individuals.

References

[1] P. Kumar *et al.*, ‘Real-time sensors for indoor air monitoring and challenges ahead in deploying them to urban buildings’, *Science of The Total Environment*, vol. 560–561, no. April, pp. 150–159, Aug. 2016, doi: 10.1016/j.scitotenv.2016.04.032.

[2] C. Chen and B. Zhao, ‘Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor’, *Atmos Environ*, vol. 45, no. 2, pp. 275–288, Jan. 2011, doi: 10.1016/j.atmosenv.2010.09.048.

[3] Y. Al horr, M. Arif, M. Katafygiotou, A. Mazroei, A. Kaushik, and E. Elsarrag, ‘Impact of indoor environmental quality on occupant well-being and comfort: A review of the literature’, *International Journal of Sustainable Built Environment*, vol. 5, no. 1, pp. 1–11, Jun. 2016, doi: 10.1016/j.ijsbe.2016.03.006.

[4] V. De Giuli, O. Da Pos, and M. De Carli, ‘Indoor environmental quality and pupil perception in Italian primary schools’, *Build Environ*, vol. 56, pp. 335–345, Oct. 2012, doi: 10.1016/j.buildenv.2012.03.024.

[5] Andargie MS; Touchie M; O’Brien W, ‘A review of factors affecting occupant comfort in multi-unit residential buildings’, *Build Environ*, vol. 160, no. April, p. 106182, 2019, doi: 10.1016/J.BUILDENV.2019.106182.

[6] G. Marques and R. Pitarma, ‘An indoor monitoring system for ambient assisted living based on internet of things architecture’, *Int J Environ Res Public Health*, vol. 13, no. 11, p. 1152, Nov. 2016, doi: 10.3390/ijerph13111152.

[7] K. W. Mui, L. T. Wong, H. C. Yu, and T. W. Tsang, ‘Development of a user-friendly indoor environmental quality (IEQ) calculator in air-conditioned offices’, *IAQVEC 2016 - 9th International Conference on Indoor Air Quality Ventilation & Energy Conservation In Buildings*, no. June 2017, 2016.

[8] R. Pitarma, G. Marques, and B. R. Ferreira, ‘Monitoring Indoor Air Quality for Enhanced Occupational Health’, *J Med Syst*, vol. 41, no. 2, p. 23, Feb. 2017, doi: 10.1007/s10916-016-0667-2.

[9] X. Shan, A. N. Melina, and E. H. Yang, ‘Impact of indoor environmental quality on students’ wellbeing and performance in educational building through life cycle costing perspective’, *J Clean Prod*, vol. 204, pp. 298–309, Dec. 2018, doi: 10.1016/j.jclepro.2018.09.002.

[10] Y. Tang, C. Shrubsole, and H. Altamirano, ‘Indoor Air Quality and Thermal Comfort: is all well with the Well Standard?’’, 2017.

[11] T. Taylor and H. Pineo, 'Health and wellbeing in BREEAM', 2015.

[12] Well Building Institute, 'WELL Building Standard ® v1 Copyright', 2019. https://www.wellcertified.com/en/system/files/WELL v1 with 2018 Q4 Addenda.pdf (accessed Mar. 08, 2019).

[13] U.S Green Building Council, 'LEED v4 for Building Design and Construction', 2019.

[14] M. Hanc, C. McAndrew, and M. Ucci, 'Conceptual approaches to wellbeing in buildings: a scoping review', *Building Research and Information*, vol. 47, no. 6, pp. 767–783, Sep. 2019, doi: 10.1080/09613218.2018.1513695.

[15] P. Wargocki, D. P. Wyon, Y. K. Baik, G. Clausen, P. O. Fanger, and D. Wargocki, Pawel; Wyon, 'Perceived Air Quality , Sick Building Syndrome (SBS) Symptoms and Productivity in an Office with Two Different Pollution Loads', *Indoor Air*, vol. 9, no. May, pp. 165–179, Sep. 1999, doi: 10.1111/j.1600-0668.1999.t01-1-00003.x.

[16] P. MacNaughton *et al.*, 'The impact of working in a green certified building on cognitive function and health', *Build Environ*, vol. 114, pp. 178–186, Mar. 2017, doi: 10.1016/j.buildenv.2016.11.041.

[17] M. Zhu, W. Liu, and P. Wargocki, 'Changes in EEG signals during the cognitive activity at varying air temperature and relative humidity', *J Expo Sci Environ Epidemiol*, p. 1, Jun. 2019, doi: 10.1038/s41370-019-0154-1.

[18] H. Thomson, M. Petticrew, and D. Morrison, 'Health effects of housing improvement: systematic review of intervention studies.', *BMJ*, vol. 323, no. 7306, pp. 187–90, 2001.

[19] L. Giddens, E. Gonzalez, and D. Leidner, *I Track, Therefore I Am: Exploring the Impact of Wearable Fitness Devices on Employee Identity and Well-being*, no. 1. 2016.

[20] T. Parkinson, A. Parkinson, and R. de Dear, 'Continuous IEQ monitoring system: Performance specifications and thermal comfort classification', *Build Environ*, vol. 149, no. February, pp. 241–252, 2019, doi: 10.1016/j.buildenv.2018.12.016.

[21] K. C. Coombs *et al.*, 'Indoor air quality in green-renovated vs. non-green low-income homes of children living in a temperate region of US (Ohio)', *Science of the Total Environment*, vol. 554–555, pp. 178–185, 2016, doi: 10.1016/j.scitotenv.2016.02.136.

[22] G. McGill, M. Qin, and L. Oyedele, 'A case study investigation of indoor air quality in UK Passivhaus dwellings', *Energy Procedia*, vol. 62, pp. 190–199, 2014, doi: 10.1016/j.egypro.2014.12.380.

[23] A. Broderick, M. Byrne, S. Armstrong, J. Sheahan, and A. M. Coggins, 'A pre and post evaluation of indoor air quality, ventilation, and thermal comfort in retrofitted co-operative social housing', *Build Environ*, vol. 122, pp. 126–133, 2017, doi: 10.1016/j.buildenv.2017.05.020.

[24] B. Painter Brown, N., and Cook, M., 'Evaluation of sensors for post-occupancy building monitoring', *Improving energy efficiency in commercial buildings (IEECB) 2010*, no. 1, 2010.

[25] R. Hoggett, J. Ward, and C. Mitchell, 'Heat in Homes: customer choice on fuel and technologies Study for Scotia Gas Networks Energy Policy Group Heat in Homes: customer choice on fuel and technologies', 2011.

[26] A. Moreno-Rangel, T. Sharpe, F. Musau, and G. McGill, 'Field evaluation of a low-cost indoor air quality monitor to quantify exposure to pollutants in residential environments', *Journal of Sensors and Sensor Systems*, vol. 7, no. 1, pp. 373–388, 2018, doi: 10.5194/jsss-7-373-2018.

[27] Y. Hua, Ö. Göçer, and K. Göçer, 'Spatial mapping of occupant satisfaction and indoor environment quality in a LEED platinum campus building', *Build Environ*, vol. 79, pp. 124–137, 2014, doi: 10.1016/j.buildenv.2014.04.029.

[28] V. Földváry, G. Bekö, S. Langer, K. Arrhenius, and D. Petráš, 'Effect of energy renovation on indoor air quality in multifamily residential buildings in Slovakia', *Build Environ*, vol. 122, pp. 363–372, 2017, doi: 10.1016/j.buildenv.2017.06.009.

[29] M. Frontczak and P. Wargocki, 'Literature survey on how different factors influence human comfort in indoor environments', *Build Environ*, vol. 46, no. 4, pp. 922–937, Apr. 2011, doi: 10.1016/j.buildenv.2010.10.021.

[30] C. Shorter *et al.*, 'Indoor visible mold and mold odor are associated with new-onset childhood wheeze in a dose-dependent manner', *Indoor Air*, vol. 28, no. 1, pp. 6–15, Jan. 2018, doi: 10.1111/ina.12413.

[31] H. Lawson, *Wearables and Wellbeing in Buildings- the Story So Far*. Berkshire, UK: BSRIA Limited, 2018.

[32] D. Li, C. C. Menassa, and V. R. Kamat, 'Personalized human comfort in indoor building environments under diverse conditioning modes', *Build Environ*, vol. 126, pp. 304–317, Dec. 2017, doi: 10.1016/J.BUILDENV.2017.10.004.

[33] T. Osama *et al.*, 'Teaching the relationship between health and climate change: a systematic scoping review protocol', *BMJ Open*, vol. 8, p. 20330, 2018, doi: 10.1136/bmjopen-2017-020330.

[34] A. Weir, S. Kitto, J. Smith, J. Presseau, I. Colman, and S. Hatcher, 'Barriers and enablers to conducting cluster randomized control trials in hospitals: A theory-informed scoping review', *Eval Program Plann*, vol. 80, no. October 2019, p. 101815, 2020, doi: 10.1016/j.evalprogplan.2020.101815.

[35] D. Gough, J. Thomas, and S. Oliver, 'Clarifying differences between review designs\rand methods art:10.1186/2046-4053-1-28', *Syst Rev*, pp. 1–9, 2012.

[36] A. C. Tricco *et al.*, 'PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation', *Annals of Internal Medicine*, vol. 169, no. 7. American College of Physicians, pp. 467–473, Oct. 02, 2018. doi: 10.7326/M18-0850.

[37] J. G. Allen, P. MacNaughton, U. Satish, S. Santanam, J. Vallarino, and J. D. Spengler, 'Associations of cognitive function scores with carbon dioxide, ventilation, and volatile organic compound exposures in office workers: A controlled exposure study of green and conventional office environments', *Environ Health Perspect*, 2016, doi: 10.1289/ehp.1510037.

[38] K. Rogage, A. Clear, Z. Alwan, T. Lawrence, and G. Kelly, 'Assessing Building Performance in Residential Buildings using BIM and Sensor Data', *International Journal of Building Pathology and Adaptation*, 2019, doi: 10.1108/IJBPA-01-2019-0012.

[39] N. Clements *et al.*, 'The Spatial and Temporal Variability of the Indoor Environmental Quality during Three Simulated Office Studies at a Living Lab', *Buildings*, vol. 9, no. 3, p. 62, Mar. 2019, doi: 10.3390/buildings9030062.

[40] A. Ghahramani *et al.*, 'Personal CO 2 bubble: Context-dependent variations and wearable sensors usability', *Journal of Building Engineering*, vol. 22, no. September 2018, pp. 295–304, 2019, doi: 10.1016/j.jobe.2018.11.015.

[41] J. R. Coleman and F. Meggers, 'Sensing of Indoor Air Quality—Characterization of Spatial and Temporal Pollutant Evolution Through Distributed Sensing', *Front Built Environ*, vol. 4, p. 28, Aug. 2018, doi: 10.3389/fbuil.2018.00028.

[42] A. Tiele, S. Esfahani, and J. Covington, 'Design and Development of a Low-Cost, Portable Monitoring Device for Indoor Environment Quality', *J Sens*, vol. 2018, pp. 1–14, Mar. 2018, doi: 10.1155/2018/5353816.

[43] I. B. Tijani, A. D. Almannaee, A. A. Alharthi, and A. M. Alremeithi, 'Wireless sensor node for indoor air quality monitoring system', in *2018 Advances in Science and Engineering Technology International Conferences, ASET 2018*, Feb. 2018, pp. 1–6. doi: 10.1109/ICASET.2018.8376839.

[44] P. Tanguy, C. Lohr, and J. Kerdreux, 'A transparent home sensors/actuators layer for health and well-being services', in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, 2017, vol. 181 LNICST, pp. 29–35. doi: 10.1007/978-3-319-49655-9_5.

[45] T. V. Tran, N. T. Dang, and W.-Y. Chung, 'Battery-free smart-sensor system for real-time indoor air quality monitoring', *Sens Actuators B Chem*, vol. 248, pp. 930–939, Sep. 2017, doi: 10.1016/j.snb.2017.01.198.

[46] A. S. Ali, Z. Zanzinger, D. Debose, and B. Stephens, 'Open Source Building Science Sensors (OSBSS): A low-cost Arduino-based platform for long-term indoor environmental data collection', *Build Environ*, vol. 100, pp. 114–126, May 2016, doi: 10.1016/j.buildenv.2016.02.010.

[47] T. Mihai and V. Iordache, 'Determining the Indoor Environment Quality for an Educational Building', 2016. doi: 10.1016/j.egypro.2015.12.246.

[48] X. Shan, J. Zhou, V. W. C. Chang, and E. H. Yang, 'Comparing mixing and displacement ventilation in tutorial rooms: Students' thermal comfort, sick building syndromes, and short-term performance', *Build Environ*, vol. 102, pp. 128–137, Jun. 2016, doi: 10.1016/j.buildenv.2016.03.025.

[49] F. Salamone *et al.*, 'Design and Development of nEMoS, an All-in-One, Low-Cost, Web-Connected and 3D-Printed Device for Environmental Analysis', *Sensors*, vol. 15, no. 6, pp. 13012–13027, Jun. 2015, doi: 10.3390/s150613012.

[50] J. Clarke and J. Hand, 'An overview of the EnTrak/ BuildAX eService delivery platform', 2015.

[51] T. Catalina and V. Iordache, 'IEQ assessment on schools in the design stage', *Build Environ*, 2012, doi: 10.1016/j.buildenv.2011.09.014.

[52] A. Marchetti *et al.*, ‘Indoor environmental quality index for conservation environments: The importance of including particulate matter’, *Build Environ*, vol. 126, pp. 132–146, Dec. 2017, doi: 10.1016/J.BUILDENV.2017.09.022.

[53] I. Mujan, A. S. Anđelković, V. Munćan, M. Kljajić, and D. Ružić, ‘Influence of indoor environmental quality on human health and productivity - A review’, *J Clean Prod*, vol. 217, pp. 646–657, Apr. 2019, doi: 10.1016/j.jclepro.2019.01.307.

[54] British Standards Institution, ‘BS ISO 16814:2008 Building environment design - Indoor air quality - Methods of expressing the quality of indoor air for human occupancy’, 2008.

[55] National Research Council (U.S.), ‘Carbon Dioxide’, in *Emergency and Continuous Exposure Guidance Levels for Selected Submarine Contaminants*, Washington, DC: The National Academies Press, 2007, pp. 46–66. doi: 10.17226/12741.

[56] P. Skov *et al.*, ‘Influence of personal characteristics, job-related factors and psychosocial factors on the sick building syndrome’, *Scand J Work Environ Health*, vol. 15, no. 4, pp. 286–295, Aug. 1989, doi: 10.5271/sjweh.1851.

[57] J. M. Daisey, W. J. Angell, and M. G. Apte, ‘Indoor air quality, ventilation and health symptoms in schools: an analysis of existing information.’, *Indoor Air*, vol. 13, no. 1, pp. 53–64, 2003.

[58] A. Steinemann, P. Wargocki, and B. Rismanchi, ‘Ten questions concerning green buildings and indoor air quality’, *Build Environ*, vol. 112, pp. 351–358, Feb. 2017, doi: 10.1016/j.buildenv.2016.11.010.

[59] L. Fang, D. P. Wyon, G. Clausen, and P. O. Fanger, ‘Impact of indoor air temperature and humidity in an office on perceived air quality, SBS symptoms and performance’, *Indoor Air*, vol. 14, no. s7, pp. 74–81, Aug. 2004, doi: 10.1111/j.1600-0668.2004.00276.x.

[60] International Organization for Standardization, *ISO 7730:2005(E) : Moderate Thermal Environments-Determination of the PMV and PPD Indices and Specification of the Conditions for Thermal Comfort*. ISO, 2005.

[61] ASHRAE, ‘ANSI/ASHRAE Addenda o, p, and to ANSI/ASHRAE Standard 55-2010’, p. 7, 2013.

[62] P. O. Fanger, *Thermal comfort: Analysis and applications in environmental engineering*. Copenhagen: Danish Technical Press., 1977. doi: 10.1016/s0003-6870(72)80074-7.

[63] R. F. Rupp, N. G. Vásquez, and R. Lamberts, 'A review of human thermal comfort in the built environment', *Energy Build*, vol. 105, pp. 178–205, Oct. 2015, doi: 10.1016/j.enbuild.2015.07.047.

[64] F. Jazizadeh, A. Ghahramani, B. Becerik-Gerber, T. Kichkaylo, and M. Orosz, 'User-led decentralized thermal comfort driven HVAC operations for improved efficiency in office buildings', *Energy Build*, vol. 70, pp. 398–410, 2014, doi: 10.1016/j.enbuild.2013.11.066.

[65] R. Yao, B. Li, and J. Liu, 'A theoretical adaptive model of thermal comfort - Adaptive Predicted Mean Vote (aPMV)', *Build Environ*, vol. 44, no. 10, pp. 2089–2096, Oct. 2009, doi: 10.1016/j.buildenv.2009.02.014.

[66] R. J. de Dear, G. Brager, and D. Cooper, 'Developing an Adaptive Model of Thermal Comfort and Preference -RP 884 final report', *Database*, no. March, 1997.

[67] British Standards Institution, 'BS EN 16798-1:2019 BSI - Energy performance of buildings: Presentation of measured energy use of buildings', 2019.

[68] British Standards Institution, 'BS EN 15251:2007 - Indoor Environmental Input Parameters for Design and Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics', 2007.

[69] A. Boerstra, T. Beuker, M. Loomans, and J. Hensen, 'Impact of available and perceived control on comfort and health in European offices', *Archit Sci Rev*, vol. 56, no. 1, pp. 30–41, 2013, doi: 10.1080/00038628.2012.744298.

[70] A. Leaman and B. Bordass, 'Are users more tolerant of "green" buildings?', *Building Research and Information*, vol. 35, no. 6, pp. 662–673, 2007, doi: 10.1080/09613210701529518.

[71] British Standards Institution, 'BS EN 12464-1:2011 - Indoor work places', 2011.

[72] A. C. K. Lai, K. W. Mui, L. T. Wong, and L. Y. Law, 'An evaluation model for indoor environmental quality (IEQ) acceptance in residential buildings', *Energy Build*, vol. 41, no. 9, pp. 930–936, Sep. 2009, doi: 10.1016/j.enbuild.2009.03.016.

[73] R. Cohen, 'Noise and Vibration Control', in *The Journal of the Acoustical Society of America*, 21st ed., vol. 51, no. 1A, Atlanta, GA, 2005, pp. 116–116. doi: 10.1121/1.1981324.

[74] J. H. Choi and K. Lee, 'Investigation of the feasibility of POE methodology for a modern commercial office building', *Build Environ*, vol. 143, pp. 591–604, Oct. 2018, doi: 10.1016/j.buildenv.2018.07.049.

[75] Y. S. Lee and S.-K. Kim, 'Indoor Environmental Quality in LEED-Certified Buildings in the U.S.', *Journal of Asian Architecture and Building Engineering*, vol. 7, no. 2, pp. 293–300, 2008, doi: 10.3130/jaabe.7.293.

[76] NHS England and Department of Health, 'Action Plan on Hearing Loss', London, 2015.

[77] H. H. Y. Lee and D. Scott, 'Overview of maintenance strategy, acceptable maintenance standard and resources from a building maintenance operation perspective', *Journal of Building Appraisal*, vol. 4, no. 4, pp. 269–278, Mar. 2009, doi: 10.1057/jba.2008.46.

[78] A. Zimmerman and M. Martin, 'Building Research and Information Post-occupancy evaluation: benefits and barriers', 2010, doi: 10.1080/09613210010016857.

[79] J. Vischer, 'Post-Occupancy Evaluation: a multifaceted tool for building improvement', in *Learning from out buildings: A state-of-the-pracfice summary of post-occupancy evaluation*, Washington, DC: NATIONAL ACADEMY PRESS, 2002, pp. 23–34.

[80] Y. Davara, I. A. Meir, and M. Schwartz, 'Architectural design and IEQ in an office complex: On research, politics and their dynamics Architectural Design and IEQ In An Office Complex On Research, Politics and Their Dynamics', doi: 10.13140/2.1.4431.9682.

[81] L. Frenkel, S. Fundaminsky, I. Meir, and L. Morhayim, 'Post-Occupancy Evaluation of a Scientists Village Complex in the Desert-Towards a Comprehensive Methodology', 2006.

[82] R. Schiano-phan, R. Schiano-phan, B. Ford, M. Gillott, and L. T. Rodrigues, 'The Passivhaus standard in the UK : Is it desirable ? Is it achievable ?', in *PLEA 2008 – 25th Conference on Passive and Low Energy Architecture*, 2008, no. January.

[83] M. Moritz, T. Redlich, P. P. Grames, and J. P. Wulfsberg, 'Value creation in open-source hardware communities: Case study of Open Source Ecology', *PICMET 2016 - Portland International Conference on Management of Engineering and Technology: Technology Management For Social Innovation, Proceedings*, pp. 2368–2375, 2017, doi: 10.1109/PICMET.2016.7806517.

[84] A. D'Ausilio, 'Arduino: A low-cost multipurpose lab equipment', *Behav Res Methods*, vol. 44, no. 2, pp. 305–313, 2012, doi: 10.3758/s13428-011-0163-z.

[85] Š. Kubínová and J. Šlégr, 'ChemDuino: Adapting Arduino for Low-Cost Chemical Measurements in Lecture and Laboratory', *J Chem Educ*, vol. 92, no. 10, pp. 1751–1753, 2015, doi: 10.1021/ed5008102.

[86] N. Barsan, D. Koziej, and U. Weimar, 'Metal oxide-based gas sensor research: How to?', *Sens Actuators B Chem*, vol. 121, no. 1, pp. 18–35, 2007, doi: 10.1016/j.snb.2006.09.047.

[87] L. Molhave *et al.*, 'Total Volatile Organic Compounds (TVOC) in Indoor Air Quality Investigations*', *Indoor Air*, vol. 7, no. 4, pp. 225–240, Dec. 1997, doi: 10.1111/j.1600-0668.1997.00002.x.

[88] AMS, 'CCS811 Datasheet', 2016. http://ams.com/eng/Products/Environmental-Sensors/Air-Quality-Sensors/CCS811 (accessed Aug. 28, 2019).

[89] AMS, 'iAQ-Core Datasheet', 2016. https://www.mouser.co.uk/datasheet/2/588/iAQ-core_DS000334_1-00-1512544.pdf (accessed Aug. 28, 2019).

[90] Sensortech, 'MiCS-VZ-89TE Data sheet'. https://sgx.cdistore.com/datasheets/e2v/MiCS-VZ-89TE_V1.0.pdf (accessed Aug. 28, 2019).

[91] J. Lozano, 'Towards the Miniaturization of Electronic Nose as Personal Measurement Systems', *Proc West Mark Ed Assoc Conf*, vol. 14, no. 1, p. 30, 2019, doi: 10.3390/proceedings2019014030.

[92] S. Chen and Q. Chen, 'Self-Regulation and Parameters Monitoring System for Culturing Chamber', *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, pp. 1–6, 2019, doi: 10.1109/ccaa.2018.8777716.

[93] S. S. Chiang, C. H. Huang, and K. C. Chang, 'A minimum hop routing protocol for home security systems using wireless sensor networks', *IEEE Transactions on Consumer Electronics*, vol. 53, no. 4, pp. 1483–1489, 2007, doi: 10.1109/TCE.2007.4429241.

[94] L. T. De Paolis, V. De Luca, and R. Paiano, 'Sensor data collection and analytics with thingsboard and spark streaming', *EESMS 2018 - Environmental, Energy, and Structural Monitoring Systems, Proceedings*, pp. 1–6, Jun. 2018, doi: 10.1109/EESMS.2018.8405822.

[95] C. Mukuria, D. Rowen, T. Peasgood, and J. Brazier, 'An empirical comparison of well-being measures used in UK', 2014.

[96] Y. Al Horr, M. Arif, A. Kaushik, A. Mazroei, E. Elsarrag, and S. Mishra, 'Occupant productivity and indoor environment quality: A case of GSAS', *International Journal of Sustainable Built Environment*, vol. 6, no. 2, pp. 476–490, Dec. 2017, doi: 10.1016/j.ijsbe.2017.11.001.

[97] J. Lee, H. Je, and J. Byun, 'Well-Being index of super tall residential buildings in Korea', *Build Environ*, vol. 46, no. 5, pp. 1184–1194, May 2011, doi: 10.1016/J.BUILDENV.2010.12.010.

[98] M. Haller and M. Hadler, 'How Social Relations and Structures can Produce Happiness and Unhappiness: An International Comparative Analysis', *Soc Indic Res*, vol. 75, no. 2, pp. 169–216, Jan. 2006, doi: 10.1007/s11205-004-6297-y.

[99] E. J. Topol, 'Transforming medicine via digital innovation.', *Sci Transl Med*, vol. 2, no. 16, p. 16cm4, Jan. 2010, doi: 10.1126/scitranslmed.3000484.

[100] E. Elenko, L. Underwood, and D. Zohar, 'Defining digital medicine', *Nat Biotechnol*, vol. 33, no. 5, pp. 456–461, May 2015, doi: 10.1038/nbt.3222.

[101] A. Henriksen *et al.*, 'Using fitness trackers and smartwatches to measure physical activity in research: Analysis of consumer wrist-worn wearables', *J Med Internet Res*, vol. 20, no. 3, p. e110, Mar. 2018, doi: 10.2196/jmir.9157.

[102] L. Giddens, D. Leidner, and E. Gonzalez, 'The Role of Fitbits in Corporate Wellness Programs : Does Step Count', pp. 3627–3635, 2017.

[103] H. J. Warraich, R. M. Califf, and H. M. Krumholz, 'The digital transformation of medicine can revitalize the patient-clinician relationship', *NPJ Digit Med*, vol. 1, no. 1, p. 49, Dec. 2018, doi: 10.1038/s41746-018-0060-2.

[104] M. Rabbi, S. Ali, T. Choudhury, and E. Berke, 'Passive and In-Situ assessment of mental and physical well-being using mobile sensors', in *Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11*, 2011, p. 385. doi: 10.1145/2030112.2030164.

[105] Y. A. Qadri, A. Nauman, Y. Bin Zikria, A. V. Vasilakos, and S. W. Kim, 'The Future of Healthcare Internet of Things: A Survey of Emerging Technologies', *IEEE Communications Surveys and Tutorials*, vol. 22, no. 2, pp. 1121–1167, 2020, doi: 10.1109/COMST.2020.2973314.

[106] W. Y. Chuang, C. C. Wu, S. S. Lu, and C. T. Lin, 'A printable conductive polymer CO2 sensor with high selectivity to humidity', *TRANSDUCERS 2017 - 19th International Conference on Solid-State Sensors, Actuators and Microsystems*, pp. 1501–1503, 2017, doi: 10.1109/TRANSDUCERS.2017.7994344.

[107] K. Niazmand *et al.*, 'Freezing of Gait detection in Parkinson's disease using accelerometer based smart clothes', *2011 IEEE Biomedical Circuits and Systems Conference, BioCAS 2011*, pp. 201–204, 2011, doi: 10.1109/BioCAS.2011.6107762.

[108] S. Kaminaga, 'Trillion sensors and MEMS', *Sensors and Materials*, vol. 30, no. 4, pp. 723–731, 2018, doi: 10.18494/SAM.2018.1814.

[109] N. Mahmoudian, G. N. Jazar, M. R. Aagaah, and M. Mahinfalah, 'Dynamics of a Micro Electro Mechanical System (MEMS)', *Proceedings - 2004 International Conference on MEMS, NANO and Smart Systems, ICMENS 2004*, pp. 688–693, 2004, doi: 10.1109/icmens.2004.1509039.

[110] A. Abu Bakar, S. A. Shaikh Rahim, A. Rashidy Razali, E. Noorsal, R. Radzali, and A. Farhanah Abd Rahim, 'Wearable Heart Rate and Body Temperature Monitoring Device for Healthcare', *J Phys Conf Ser*, vol. 1535, p. 12002, 2020, doi: 10.1088/1742-6596/1535/1/012002.

[111] N. Q. Al-Naggar, H. Mohammed Al-Hammadi, A. Mohammed Al-Fusail, and Z. Ali AL-Shaebi, 'Design of a Remote Real-Time Monitoring System for Multiple Physiological Parameters Based on Smartphone', 2019, doi: 10.1155/2019/5674673.

[112] A. PORTAANKORVA, 'EVALUATION OF BLUETOOTH LOW ENERGY TECHNOLOGY FOR TEAM SPORTS REAL-TIME PHYSIOLOGICAL MONITORING Examiner: Assistant Prof. Antti Vehkaoja Examiner and topic approved by the Fac-ulty Council of Computing and Electrical Engineering 2nd of', 2018.

[113] T. Mohanraj and Raj. S. N. ; Keshore, 'Patient Monitoring System Using LabVIEW', vol. 24, no. 4, pp. 67–71, 2017.

[114] J. E. Shaji, B. Varghese, and R. Varghese, 'A Health Care Monitoring System with Wireless Body Area Network using IoT', *International Journal of Recent Trends in Engineering and Research*, vol. 3, no. 11, pp. 112–117, 2017, doi: 10.23883/ijrter.2017.3499.edo9r.

[115] M. Di Rienzo, G. Rizzo, Z. M. Işilay, and P. Lombardi, 'SeisMote: A Multi-Sensor Wireless Platform for Cardiovascular Monitoring in Laboratory, Daily Life, and Telemedicine', *Sensors*, vol. 20, no. 3, p. 680, Jan. 2020, doi: 10.3390/s20030680.

[116] D. J. Im, B. S. Yoo, M. M. Ahn, D. Moon, and I. S. Kang, 'Digital electrophoresis of charged droplets', *Anal Chem*, vol. 85, no. 8, pp. 4038–4044, 2013, doi: 10.1021/ac303778j.

[117] T. M. Rideout, J. R. LaCourse, and P. L. McWilliam, 'Force sensing syringe as a teaching tool for injection', *2012 38th Annual Northeast Bioengineering Conference, NEBEC 2012*, pp. 364–365, 2012, doi: 10.1109/NEBC.2012.6207115.

[118] J. Alberto *et al.*, 'Fully Untethered Battery-free Biomonitoring Electronic Tattoo with Wireless Energy Harvesting', *Sci Rep*, vol. 10, no. 1, pp. 1–11, Dec. 2020, doi: 10.1038/s41598-020-62097-6.

[119] A. Anisimov, B. Alekseev, and D. Egorov, 'Development of Portable Cardiograph Using Novel Front-End Solutions', in *Conference of Open Innovation Association, FRUCT*, Nov. 2019, pp. 32–37. doi: 10.23919/FRUCT48121.2019.8981497.

[120] A. Yudhana, J. Rahmawan, and C. U. P. Negara, 'Flex sensors and MPU6050 sensors responses on smart glove for sign language translation', *IOP Conf Ser Mater Sci Eng*, vol. 403, no. 1, 2018, doi: 10.1088/1757-899X/403/1/012032.

[121] M. A. I. Anik, M. Hassan, H. Mahmud, and M. K. Hasan, 'Activity recognition of a badminton game through accelerometer and gyroscope', in *19th International Conference on Computer and Information Technology, ICCIT 2016*, Feb. 2017, pp. 213–217. doi: 10.1109/ICCITECHN.2016.7860197.

[122] S. P. Dawson, 'IMPLEMENTING VIBRATION-BASED COMMUNICATION CHANNELS FOR SECURE IMPLANTABLE MEDICAL DEVICES A Dissertation', 2017.

[123] D. A. Fitriani, W. Andhyka, and D. Risqiwati, 'Design of Monitoring System Step Walking With MPU6050 Sensor Based Android', *JOINCS (Journal of Informatics, Network, and Computer Science)*, vol. 1, no. 1, p. 1, 2017, doi: 10.21070/joincs.v1i1.799.

[124] S. Kardos, P. Balog, and S. Slosarcik, 'Gait dynamics sensing using IMU sensor array system', *Advances in Electrical and Electronic Engineering*, vol. 15, no. 1, pp. 71–76, 2017, doi: 10.15598/aeee.v15i1.2019.

[125] F. A. O. Mota, V. H. M. Biajo, H. O. Mota, and F. H. Vasconcelos, 'A wireless sensor network for the biomechanical analysis of the gait', Jul. 2017. doi: 10.1109/I2MTC.2017.7969859.

[126] Z. T. Al-Dahan, N. K. Bachache, and L. N. Bachache, 'Design and implementation of fall detection system using MPU6050 Arduino', in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016, vol. 9677, pp. 180–187. doi: 10.1007/978-3-319-39601-9_16.

[127] Y. Kim, W. S. Lee, V. Raghunathan, N. K. Jha, and A. Raghunathan, 'Vibration-based secure side channel for medical devices', *Proc Des Autom Conf*, vol. 2015-July, 2015, doi: 10.1145/2744769.2744928.

[128] L. Lei, Z. Dongxue, and L. Xiaojun, 'Design of fall detection system for elderly people based on MPU6050 Sensor', *Zhongguo Yi Liao Qi Xie Za Zhi*, vol. 39, no. 5, pp. 327–330, 2015.

[129] B. Wang, K. S. Rajput, W. K. Tam, A. K. H. Tung, and Z. Yang, 'FreeWalker: A smart insole for longitudinal gait analysis', in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Nov. 2015, vol. 2015-November, pp. 3723–3726. doi: 10.1109/EMBC.2015.7319202.

[130] J. C. Goldsack *et al.*, 'Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs)', *NPJ Digit Med*, vol. 3, no. 1, 2020, doi: 10.1038/s41746-020-0260-4.

[131] A. Godfrey *et al.*, 'BioMeT and algorithm challenges: A proposed digital standardized evaluation framework', *IEEE J Transl Eng Health Med*, vol. 8, no. May, 2020, doi: 10.1109/JTEHM.2020.2996761.

[132] A. Coravos, S. Khozin, and K. D. Mandl, 'Developing and adopting safe and effective digital biomarkers to improve patient outcomes', *NPJ Digit Med*, vol. 2, no. 1, pp. 1–5, 2019, doi: 10.1038/s41746-019-0090-4.

[133] J. Y. Yeom, I. Defendi, H. Takahashi, K. Zeitelhack, M. Nakazawa, and H. Murayama, 'A 12-channel CMOS preamplifier-shaper-discriminator ASIC for APD and gas counters', in *IEEE Transactions on Nuclear Science*, Aug. 2006, vol. 53, no. 4, pp. 2204–2208. doi: 10.1109/TNS.2006.878573.

[134] O. Laakkonen, A. Penttinen, J. Luukko, and O. Pyrhonen, ‘Single chip implementation of the frequency converter controller using FPGA circuit and soft processor’, in *2005 European Conference on Power Electronics and Applications*, 2005, vol. 9, pp. 8 pp.-P.8. doi: 10.1109/EPE.2005.219396.

[135] M. H. Rais, M. H. Al Mijalli, and M. Nisar, ‘FPGA realization of braun’s multipliers’, *Lecture Notes in Engineering and Computer Science*, vol. 2195, no. December, pp. 275–277, 2012.

[136] E. L. Graas, E. A. Brown, and R. H. Lee, ‘An FPGA-based approach to high-speed simulation of conductance-based neuron models’, *Neuroinformatics*, vol. 2, no. 4, pp. 417–435, 2004, doi: 10.1385/NI:2:4:417.

[137] R. Aoki, S. Oikawa, R. Tsuchiyama, and T. Nakamura, ‘Hybrid OpenCL: Connecting different OpenCL implementations over network’, *Proceedings - 10th IEEE International Conference on Computer and Information Technology, CIT-2010, 7th IEEE International Conference on Embedded Software and Systems, ICESS-2010, ScalCom-2010*, no. Cit, pp. 2729–2735, 2010, doi: 10.1109/CIT.2010.457.

[138] J. T. Wunderlich, ‘Session 3547 Focusing on the Blurry Distinction between Microprocessors and Microcontrollers’, 1999.

[139] R. Berlia and P. Santosh, ‘Mouse brace: A convenient computer mouse using accelerometer, flex sensors and microcontroller’, *Proceedings of 2014 International Conference on Contemporary Computing and Informatics, IC3I 2014*, pp. 558–561, 2014, doi: 10.1109/IC3I.2014.7019684.

[140] S. M. Lee, E. N. R. Ko, and S. E. Lee, ‘A Hardware Scheduler for Multicore Block Cipher Processor’, *Proceedings - 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, PDP 2016*, pp. 750–754, 2016, doi: 10.1109/PDP.2016.59.

[141] C. Duran *et al.*, ‘A 32-bit RISC-V AXI4-lite bus-based microcontroller with 10-bit SAR ADC’, in *LASCAS 2016 - 7th IEEE Latin American Symposium on Circuits and Systems, R9 IEEE CASS Flagship Conference*, Apr. 2016, pp. 315–318. doi: 10.1109/LASCAS.2016.7451073.

[142] R. H. Walden, ‘Analog-to-digital converter technology comparison’, *Technical Digest - GaAs IC Symposium (Gallium Arsenide Integrated Circuit)*, pp. 217–219, 1994, doi: 10.1109/gaas.1994.636970.

[143] R. Singh, A. Gehlot, L. R. Gupta, B. Singh, and M. Swain, *Internet of Things with Raspberry Pi and Arduino*. First edition. | New York, N.Y. : CRC Press/Taylor & Francis Group, 2019.: CRC Press, 2019. doi: 10.1201/9780429284564.

[144] J. R. Ackermann, 'Toward Open Source Hardware.', *University of Dayton Law Review*, vol. 34, no. 2, pp. 183–222, 2009.

[145] A. Nayyar and V. Puri, 'A review of Arduino board's, Lilypad's & Arduino shields', in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2016, pp. 1485–1492.

[146] K. Ebersold, 'THE IMPACT OF DISRUPTIVE TECHNOLOGY: THE INTERNET OF THINGS', 2015.

[147] J. J. Segura-Juáirez, D. Cuesta-Frau, and L. Samblas-Pena, 'A microcontroller-based portable electrocardiograph system', *Proceedings of the IEEE International Conference on Electronics, Circuits, and Systems*, vol. 2, no. 9, pp. 922–925, 2003, doi: 10.1109/ICECS.2003.1301938.

[148] M. Babiuch, P. Foltynek, and P. Smutny, 'Using the ESP32 microcontroller for data processing', *Proceedings of the 2019 20th International Carpathian Control Conference, ICCC 2019*, pp. 9–14, 2019, doi: 10.1109/CarpathianCC.2019.8765944.

[149] O. Barybin, E. Zaitseva, and V. Brazhnyi, 'Testing the security ESP32 internet of things devices', *2019 IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, PIC S and T 2019 - Proceedings*, pp. 143–146, 2019, doi: 10.1109/PICST47496.2019.9061269.

[150] A. Maier, A. Sharp, and Y. Vagapov, 'Comparative analysis and practical implementation of the ESP32 microcontroller module for the internet of things', *2017 Internet Technologies and Applications, ITA 2017 - Proceedings of the 7th International Conference*, pp. 143–148, 2017, doi: 10.1109/ITECHA.2017.8101926.

[151] M. Bilal *et al.*, 'Big Data in the construction industry: A review of present status, opportunities, and future trends', *Advanced Engineering Informatics*, vol. 30, no. 3. Elsevier Ltd, pp. 500–521, Aug. 01, 2016. doi: 10.1016/j.aei.2016.07.001.

[152] K. L. Lueth, 'IoT Platform Companies Landscape 2020', *IOT Analytics*, 2019. https://iot-analytics.com/iot-platform-companies-landscape-2020/ (accessed Jul. 14, 2020).

[153] P. Pierleoni, R. Concetti, A. Belli, and L. Palma, ‘Amazon, Google and Microsoft Solutions for IoT: Architectures and a Performance Comparison’, *IEEE Access*, vol. 8, pp. 5455–5470, 2020, doi: 10.1109/ACCESS.2019.2961511.

[154] L. O. Aghenta and M. T. Iqbal, ‘Low-cost, open source IoT-based SCADA system design using thinger.IO and ESP32 thing’, *Electronics (Switzerland)*, vol. 8, no. 8, pp. 1–24, 2019, doi: 10.3390/electronics8080822.

[155] S. Pasha, ‘Thingspeak Based Sensing and Monitoring System for IoT with Matlab Analysis’, *Int. J. New Technol. Res.*, vol. 2, no. 6, pp. 19–23, 2016, doi: ISSN : 2454-4116.

[156] D. Cunha, P. Neves, and P. Sousa, ‘PaaS manager: A platform-as-a-service aggregation framework’, *Computer Science and Information Systems*, vol. 11, no. 4, pp. 1209–1228, 2014, doi: 10.2298/CSIS130828028C.

[157] ‘Understand Azure IoT Hub quotas and throttling | Microsoft Docs’. https://docs.microsoft.com/en-us/azure/iot-hub/iot-hub-devguide-quotas-throttling#other-limits (accessed Jul. 17, 2020).

[158] ‘Pricing – IoT Hub | Microsoft Azure’. https://azure.microsoft.com/en-gb/pricing/details/iot-hub/ (accessed Jul. 17, 2020).

[159] ‘AWS IoT Core Pricing - Amazon Web Services’. https://aws.amazon.com/iot-core/pricing/ (accessed Jul. 17, 2020).

[160] ‘Pricing | Cloud IoT Core Documentation | Google Cloud’. https://cloud.google.com/iot/pricing (accessed Jul. 17, 2020).

[161] T. Yokotani and Y. Sasaki, ‘Comparison with HTTP and MQTT on required network resources for IoT’, *ICCEREC 2016 - International Conference on Control, Electronics, Renewable Energy, and Communications 2016, Conference Proceedings*, pp. 1–6, 2017, doi: 10.1109/ICCEREC.2016.7814989.

[162] N. Naik, ‘Choice of effective messaging protocols for IoT systems: MQTT, CoAP, AMQP and HTTP’, *2017 IEEE International Symposium on Systems Engineering, ISSE 2017 - Proceedings*, 2017, doi: 10.1109/SysEng.2017.8088251.

[163] B. Wukkadada, K. Wankhede, R. Nambiar, and A. Nair, ‘Comparison with HTTP and MQTT in Internet of Things (IoT)’, in *Proceedings of the International Conference on Inventive Research in Computing Applications, ICIRCA 2018*, Dec. 2018, pp. 249–253. doi: 10.1109/ICIRCA.2018.8597401.

[164] E. Urrestarazu, J. D. Jirsch, P. LeVan, J. Hall, and J. Gotman, ‘High-frequency intracerebral EEG activity (100-500 Hz) following interictal spikes’, *Epilepsia*, vol. 47, no. 9, pp. 1465–1476, 2006, doi: 10.1111/j.1528-1167.2006.00618.x.

[165] R. T. Fielding, *Architectural Styles and the Design of Network-based Software Architectures (Doctoral Dissertation)*. Irvine, Calafornia, 2000.

[166] M. Van Mill, A. Belloum, M. Baranowski, and A. S. Z. Belloum, ‘A Cookery extension to sim-plify cloud service integrations (Thesis)’, 2017.

[167] E. Lusk, N. Desai, R. Bradshaw, A. Lusk, and R. Butler, ‘An Interoperability Approach to System Software, Tools, and Libraries for Clusters’, *Int J High Perform Comput Appl*, vol. 20, no. 3, pp. 401–407, Aug. 2006, doi: 10.1177/1094342006067473.

[168] F. de Arriba-Pérez, M. Caeiro-Rodríguez, and J. Santos-Gago, ‘Collection and Processing of Data from Wrist Wearable Devices in Heterogeneous and Multiple-User Scenarios’, *Sensors*, vol. 16, no. 9, p. 1538, Sep. 2016, doi: 10.3390/s16091538.

[169] S. Pirbhulal *et al.*, ‘An Efficient Biometric-Based Algorithm Using Heart Rate Variability for Securing Body Sensor Networks’, *Sensors*, vol. 15, no. 7, pp. 15067–15089, Jun. 2015, doi: 10.3390/s150715067.

[170] G. Coulby and F. Young, ‘Frameworks: integration to digital networks and beyond’, in *Digital Health*, A. Godfrey and S. Stuart, Eds. Newcastle Upon Tyne: Elsevier, 2021, pp. 163–176. doi: 10.1016/B978-0-12-818914-6.00003-X.

[171] D. Rani and R. K. Ranjan, ‘A Comparative Study of SaaS, PaaS and IaaS in Cloud Computing’, 2014.

[172] G. Coulby, A. K. Clear, O. Jones, and A. Godfrey, ‘A Scoping Review of Technological Approaches to Environmental Monitoring’, *International Journal of Environmental Research and Public Health 2020*, 2020.

[173] G. Coulby, A. Clear, O. Jones, and A. Godfrey, ‘Towards remote healthcare monitoring using accessible IoT technology: State-of-the-art, insights and experimental design’, *Biomed Eng Online*, vol. 19, no. 1, pp. 1–24, 2020.

[174] A. Godfrey, ‘Wearables for independent living in older adults: Gait and falls’, *Maturitas*, vol. 100, pp. 16–26, 2017, doi: 10.1016/j.maturitas.2017.03.317.

[175] S. Kubba, ‘Indoor Environmental Quality (IEQ)’, in *LEED v4 Practices, Certification, and Accreditation Handbook*, vol. 169, New York : Routledge, 2018.: Elsevier, 2016, pp. 303–378. doi: 10.1016/B978-0-12-803830-7.00007-4.

[176] G. Coulby, A. Clear, O. Jones, and A. Godfrey, ‘A Scoping Review of Technological Approaches to Environmental Monitoring’, *International Journal of Environmental Research and Public Health 2020, Vol. 17, Page 3995*, vol. 17, no. 11, p. 3995, Jun. 2020, doi: 10.3390/IJERPH17113995.

[177] S. McDonald, R. Vieira, A. Godfrey, N. O’Brien, M. White, and F. F. Sniehotta, ‘Changes in physical activity during the retirement transition: A series of novel n-of-1 natural experiments’, *International Journal of Behavioral Nutrition and Physical Activity*, vol. 14, no. 1, pp. 1–12, 2017, doi: 10.1186/s12966-017-0623-7.

[178] E. R. Van Teijlingen, A. M. Rennie, V. Hundley, and W. Graham, ‘The importance of conducting and reporting pilot studies: The example of the Scottish Births Survey’, *J Adv Nurs*, vol. 34, no. 3, pp. 289–295, 2001, doi: 10.1046/j.1365-2648.2001.01757.x.

[179] X. Mi, Y. Zhang, F. Qian, and X. Wang, ‘An empirical characterization of IFTTT: Ecosystem, usage, and performance’, *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, vol. Part F131937, pp. 398–404, 2017, doi: 10.1145/3131365.3131369.

[180] R. Xu, Q. Zeng, L. Zhu, H. Chi, X. Du, and M. Guizani, ‘Privacy Leakage in Smart Homes and Its Mitigation: IFTTT as a Case Study’, *IEEE Access*, vol. 7, pp. 63457–63471, 2019, doi: 10.1109/ACCESS.2019.2911202.

[181] S. Del Din, A. Godfrey, B. Galna, S. Lord, and L. Rochester, ‘Free-living gait characteristics in ageing and Parkinson’s disease: impact of environment and ambulatory bout length’, *J Neuroeng Rehabil*, vol. 13, no. 1, p. 46, Dec. 2016, doi: 10.1186/s12984-016-0154-5.

[182] If This Then That, ‘Create your own - IFTTT’. https://ifttt.com/create (accessed May 19, 2022).

[183] C. Kelly, N. Pitropakis, S. McKeown, and C. Lambrinoudakis, ‘Testing And Hardening IoT Devices Against the Mirai Botnet’, in *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, Jun. 2020, pp. 1–8. doi: 10.1109/CyberSecurity49315.2020.9138887.

[184] A. Middleton, S. L. Fritz, and M. Lusardi, 'Walking speed: The functional vital sign', *J Aging Phys Act*, vol. 23, no. 2, pp. 314–322, 2015, doi: 10.1123/japa.2013-0236.

[185] A. J. Espay *et al.*, 'A roadmap for implementation of patient-centered digital outcome measures in Parkinson's disease obtained using mobile health technologies', *Movement Disorders*, vol. 34, no. 5, pp. 657–663, 2019, doi: 10.1002/mds.27671.

[186] C. Ladha *et al.*, 'Toward a low-cost gait analysis system for clinical and free-living assessment', in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2016, vol. 2016-Octob, pp. 1874–1877. doi: 10.1109/EMBC.2016.7591086.

[187] J. McCamley, M. Donati, E. Grimpampi, and C. Mazzà, 'An enhanced estimate of initial contact and final contact instants of time using lower trunk inertial sensor data', *Gait Posture*, vol. 36, no. 2, pp. 316–318, 2012, doi: 10.1016/j.gaitpost.2012.02.019.

[188] W. Zijlstra and A. L. Hof, 'Assessment of spatio-temporal gait parameters from trunk accelerations during human walking', *Gait Posture*, vol. 18, no. 2, pp. 1–10, 2003, doi: 10.1016/S0966-6362(02)00190-X.

[189] A. Godfrey, S. Del Din, G. Barry, J. C. Mathers, and L. Rochester, 'Instrumenting gait with an accelerometer: A system and algorithm examination', *Med Eng Phys*, vol. 37, no. 4, pp. 400–407, 2015, doi: 10.1016/j.medengphy.2015.02.003.

[190] S. Lord, B. Galna, J. Verghese, S. Coleman, D. Burn, and L. Rochester, 'Independent domains of gait in older adults and associated motor and nonmotor attributes: Validation of a factor analysis approach', *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, vol. 68, no. 7, pp. 820–827, 2013, doi: 10.1093/gerona/gls255.

[191] R. Z. Ur Rehman *et al.*, 'Comparison of walking protocols and gait assessment systems for machine learning-based classification of parkinson's disease', *Sensors (Switzerland)*, vol. 19, no. 24, 2019, doi: 10.3390/s19245363.

[192] S. A. Moore, A. Hickey, S. Lord, S. Del Din, A. Godfrey, and L. Rochester, 'Comprehensive measurement of stroke gait characteristics with a single accelerometer in the laboratory and community: A feasibility, validity and reliability study', *J Neuroeng Rehabil*, vol. 14, no. 1, pp. 1–10, Dec. 2017, doi: 10.1186/s12984-017-0341-z.

[193] A. Hickey *et al.*, ‘Validity of a wearable accelerometer to quantify gait in spinocerebellar ataxia type 6’, *Physiol Meas*, vol. 37, no. 11, pp. N105–N117, 2016, doi: 10.1088/0967-3334/37/11/N105.

[194] S. Del Din, A. Godfrey, and L. Rochester, ‘Validation of an Accelerometer to Quantify a Comprehensive Battery of Gait Characteristics in Healthy Older Adults and Parkinson’s Disease: Toward Clinical and at Home Use’, *IEEE J Biomed Health Inform*, vol. 20, no. 3, pp. 838–847, 2016, doi: 10.1109/JBHI.2015.2419317.

[195] W. Zijlstra and A. L. Hof, ‘Displacement of the pelvis during human walking: Experimental data and model predictions’, *Gait Posture*, vol. 6, no. 3, pp. 249–262, 1997, doi: 10.1016/S0966-6362(97)00021-0.

[196] A. Hickey, S. Del Din, L. Rochester, and A. Godfrey, ‘Detecting free-living steps and walking bouts: Validating an algorithm for macro gait analysis’, *Physiol Meas*, vol. 38, no. 1, pp. N1–N15, 2017, doi: 10.1088/1361-6579/38/1/N1.

[197] S. Del Din *et al.*, ‘Accelerometer-based gait assessment: Pragmatic deployment on an international scale’, in *IEEE Workshop on Statistical Signal Processing Proceedings*, Aug. 2016, vol. 2016-Augus. doi: 10.1109/SSP.2016.7551794.

[198] MathWorks®, ‘Write many entries to a channel in JSON format with a single HTTP POST - MATLAB - MathWorks United Kingdom’. https://uk.mathworks.com/help/thingspeak/bulkwritejsondata.html;jsessionid=0a4f4fc08e4613831ed872a8a61a (accessed Jul. 27, 2020).

[199] S. Del Din *et al.*, ‘Analysis of Free-Living Gait in Older Adults With and Without Parkinson’s Disease and With and Without a History of Falls: Identifying Generic and Disease-Specific Characteristics’, *The Journals of Gerontology: Series A*, vol. 74, no. 4, pp. 500–506, Mar. 2019, doi: 10.1093/gerona/glx254.

[200] Arduino, ‘Arduino Library List - Arduino Libraries’, 2020. https://www.arduinolibraries.info/ (accessed Apr. 22, 2020).

[201] J. Kwon, G. Ahn, G. Kim, J. C. Kim, and H. Kim, ‘A Study on NDIR-based CO2 Sensor to apply Remote Air Quality Monitoring System’, *2009 ICCAS-SICE*, pp. 1683–1687, 2009.

[202] J. Saini, ‘Internet of Things Based Environment Monitoring and PM 10 Prediction for Smart Home’, 2020.

[203] P. T. Moseley, ‘Solid state gas sensors’, 1997.

[204] R. Piedrahita *et al.*, 'Article in Atmospheric Measurement Techniques · February', *Atmos. Meas. Tech*, vol. 7, pp. 3325–3336, 2014, doi: 10.5194/amtd-7-2425-2014.

[205] J. W. Bennett and A. A. Inamdar, 'Are some fungal volatile organic compounds (VOCs) mycotoxins?', *Toxins*, vol. 7, no. 9. MDPI AG, pp. 3785–3804, Sep. 22, 2015. doi: 10.3390/toxins7093785.

[206] F. Naepelt, R. Schreiber, M. Groening, and C. Meyer, 'An Indoor Air Quality System for IoT applications', *Sensoren und Messsysteme - 19. ITG/GMA-Fachtagung*, pp. 581–584, 2018.

[207] I. Siti Hamimah, D. Baba, and L. Abd Mutalib, 'Indoor Air Quality Issues for Non-Industrial Work Place', 2010.

[208] G. Guyot, M. H. Sherman, and I. S. Walker, 'Smart ventilation energy and indoor air quality performance in residential buildings: A review', *Energy Build*, vol. 165, pp. 416–430, Apr. 2018, doi: 10.1016/J.ENBUILD.2017.12.051.

[209] A. Rusu and P. Dobra, 'Using adaptive transmit power in wireless indoor air quality monitoring', *2019 23rd International Conference on System Theory, Control and Computing, ICSTCC 2019 - Proceedings*, pp. 543–548, 2019, doi: 10.1109/ICSTCC.2019.8885849.

[210] Y. Yang, 'Design and Application of Intelligent Agriculture Service System with LoRa-based on Wireless Sensor Network', *Proceedings - 2020 International Conference on Computer Engineering and Application, ICCEA 2020*, pp. 712–716, 2020, doi: 10.1109/ICCEA50009.2020.00155.

[211] S. Kulkarni, N. G. Kalayil, J. James, S. Parsewar, and R. Shriram, 'Detection of Parkinson's Disease through Smell Signatures', *Proceedings of the 2020 IEEE International Conference on Communication and Signal Processing, ICCSP 2020*, vol. 411052, pp. 808–812, 2020, doi: 10.1109/ICCSP48568.2020.9182283.

[212] Adafruit, 'adafruit/Adafruit_CCS811: Arduino driver for CCS811 digital gas sensor', 2021. https://github.com/adafruit/Adafruit_CCS811 (accessed Feb. 01, 2021).

[213] M. Freiburg, 'CO2 measurements in instrumental and vocal closed room settings as a risk reducing measure for a Coronavirus infection b', 2020.

[214] Umwelt-Campus, 'COVID-19 prevention: CO2 measurement and demand'. https://www.umwelt-campus.de/forschung/projekte/iot-werkstatt/ideen-zur-corona-krise (accessed Jan. 22, 2021).

[215] F. H. L. Koppens, T. Mueller, P. Avouris, A. C. Ferrari, M. S. Vitiello, and M. Polini, 'Photodetectors based on graphene, other two-dimensional materials and hybrid systems', *Nat Nanotechnol*, vol. 9, no. 10, pp. 780–793, 2014, doi: 10.1038/nnano.2014.215.

[216] A. Martín-Garín, J. A. Millán-García, A. Baïri, J. Millán-Medel, and J. M. Sala-Lizarraga, 'Environmental monitoring system based on an Open Source Platform and the Internet of Things for a building energy retrofit', *Autom Constr*, vol. 87, pp. 201–214, Mar. 2018, doi: 10.1016/j.autcon.2017.12.017.

[217] G. Marques, C. R. Ferreira, and R. Pitarma, 'Indoor Air Quality Assessment Using a CO 2 Monitoring System Based on Internet of Things', *J Med Syst*, vol. 43, no. 3, 2019, doi: 10.1007/s10916-019-1184-x.

[218] H. Nijzink, 'A Homemade Air Quality Sensor: Is it viable?', Bachelor's thesis, University of Twente, 2018.

[219] J. Dempsey, 'WifWaf/MH-Z19: For Arduino Boards (&ESP32). Additional Examples/Commands., Hardware/Software Serial', *Github*, 2021. https://github.com/WifWaf/MH-Z19 (accessed Feb. 01, 2021).

[220] S. C. Lee *et al.*, 'PM1.0 and PM2.5 characteristics in the roadside environment of Hong Kong', *Aerosol Science and Technology*, vol. 40, no. 3, pp. 157–165, 2006, doi: 10.1080/02786820500494544.

[221] Environmental Protection Agency, 'National Ambient Air Quality Standards for Particulate Matter; Final Rule', *Fed Regist*, vol. 78, no. 10, pp. 3086–3287, 2013.

[222] Z. Wang, W. W. Delp, and B. C. Singer, 'Performance of low-cost indoor air quality monitors for PM2.5 and PM10 from residential sources', *Build Environ*, vol. 171, p. 106654, Mar. 2020, doi: 10.1016/j.buildenv.2020.106654.

[223] B. Laquai, 'Particle Distribution Dependent Inaccuracy of the Plantower PMS5003 low- cost PM-sensor Particle Distribution Dependent Inaccuracy of the Plantower PMS5003 low- cost PM-sensor', 2017.

[224] M. Levy Zamora, F. Xiong, D. Gentner, B. Kerkez, J. Kohrman-Glaser, and K. Koehler, 'Field and Laboratory Evaluations of the Low-Cost Plantower Particulate Matter Sensor', *Environ Sci Technol*, vol. 53, no. 2, pp. 838–849, 2019, doi: 10.1021/acs.est.8b05174.

[225] H. Mei *et al.*, 'Field evaluation of low-cost particulate matter sensors in Beijing', *Sensors (Switzerland)*, vol. 20, no. 16, pp. 1–17, 2020, doi: 10.3390/s20164381.

[226] Adafruit, 'adafruit/Adafruit_PM25AQI: Arduino library for PMS* Air Quality Sensors', *Github*, 2021. https://github.com/adafruit/Adafruit_PM25AQI (accessed Feb. 01, 2021).

[227] Y. Liu, X. Ni, Y. Wu, and W. Zhang, 'Study on effect of temperature and humidity on the CO2 concentration measurement', *IOP Conf Ser Earth Environ Sci*, vol. 81, no. 1, 2017, doi: 10.1088/1755-1315/81/1/012083.

[228] P. Dani, P. Adi, and A. Kitagawa, 'Performance Evaluation of E32 Long Range Radio Frequency 915 MHz based on Internet of Things and Micro Sensors Data', 2019.

[229] Adafruit, 'adafruit/Adafruit_BME280_Library: Arduino Library for BME280 sensors', *Github*, 2020. https://github.com/adafruit/Adafruit_BME280_Library (accessed Feb. 01, 2021).

[230] M. J. Jafari and A. Asghar, 'Association of Sick Building Syndrome with Indoor Air Parameters', vol. 14, no. 1, pp. 55–62, 2015.

[231] F. P. Jooste, A. Kumar, and G. P. Hancke, 'Energy efficient irrigation scheduling system based on the ISO/IEC/IEEE 21451 standards', *Proceedings of the IEEE International Conference on Industrial Technology*, pp. 1441–1446, 2017, doi: 10.1109/ICIT.2017.7915577.

[232] L. H. M. de Castro, B. L. Lago, and F. Mondaini, 'Damped Harmonic Oscillator with Arduino', *Journal of Applied Mathematics and Physics*, vol. 03, no. 06, pp. 631–636, 2015, doi: 10.4236/jamp.2015.36075.

[233] L. Sánchez, I. Elicegui, J. Cuesta, L. Muñoz, and J. Lanza, 'Integration of Utilities Infrastructures in a Future Internet Enabled Smart City Framework', *Sensors*, vol. 13, no. 11, pp. 14438–14465, Oct. 2013, doi: 10.3390/s131114438.

[234] PeterEmbedded, 'BH1750FVI - Arduino Libraries', 2018. https://www.arduinolibraries.info/libraries/bh1750-fvi (accessed Apr. 22, 2020).

[235] ELEXHUB, 'ELEXHUB Tutorial: BH1750 Digital Light Intensity Sensor', 2016.

[236] R. Osborne, M. Ward, and K. Dawkins, 'A micro-machined acoustic sensor array for fuel level indication', *Sens Actuators A Phys*, vol. 115, no. 2-3 SPEC. ISS., pp. 385–391, 2004, doi: 10.1016/j.sna.2004.04.044.

[237] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, and S. Yaacob, 'A stereo image processing system for visually impaired', *International Journal of Computer,*

Electrical, Automation, Control and Information Engineering, vol. 2, no. 8, pp. 2794–2803, 2008.

[238] B. Bank, 'Logarithmic frequency scale parallel filter design with complex and magnitude-only specifications', *IEEE Signal Process Lett*, vol. 18, no. 2, pp. 138–141, 2011, doi: 10.1109/LSP.2010.2093892.

[239] J. Lewis and B. Moss, 'MEMS Microphones, the Future for Hearing Aids', *Analog Dialogue*, vol. 47, no. 11, pp. 3–5, 2013.

[240] J. Lewis, 'MS-2472: Analog and Digital MEMS Microphone Design Considerations', pp. 1–4, 2013.

[241] I. Kostoski, 'ikostoski/esp32-i2s-slm: Sound Level Meter with ESP32 and I2S MEMS microphone', *Github*, Nov. 22, 2019. https://github.com/ikostoski/esp32-i2s-slm (accessed Jul. 31, 2020).

[242] Z. Winsen and E. Technology, 'Intelligent Infrared CO2 Module (Model: MH-Z19B)', pp. 1–9, 2016.

[243] Z. Yong and Z. Haoxin, 'Digital universal particle concentration sensor PMS6003 series data manual Writer', *Product data manual of PLANTOWER*, vol. c, p. 15, 2016.

[244] Bosch Sensortec, 'BME280 Combined humidity and pressure sensor', 2020.

[245] ROHM, 'Digital 16bit Serial Output Type Ambient Light Sensor IC BH1750FVI', 2011.

[246] InvenSense, 'INMP441 Omnidirectional Microphone with Bottom Port and I 2 S Digital Output', vol. 1, no. 408, pp. 1–20, 2014.

[247] T. Collins, 'EQuALS – Environmental quality in active learning spaces'.

[248] S. Liu *et al.*, 'Predicted percentage dissatisfied with vertical temperature gradient', *Energy Build*, vol. 220, 2020, doi: 10.1016/j.enbuild.2020.110085.

[249] D. Licina, Y. Tian, and W. W. Nazaroff, *Experimental investigation of inhalation intake fraction from localized indoor particle sources*. 2017.

[250] M. A. Molina, 'Design and development of a methodology to monitor PM10 dust particles produced by industrial activities using UAV's Sustainable Mineral Institute'.

[251] S. H. Kim *et al.*, 'The Impact of Composition in Non-steel and Low-Steel Type Friction Materials on Airborne Brake Wear Particulate Emission', *Tribol Lett*, vol. 68, no. 4, p. 118, Dec. 2020, doi: 10.1007/s11249-020-01361-2.

[252] Onset Computer Corporation, 'HOBO® MX CO 2 Logger (MX1102) Manual', 2017.

[253] Onset Computer Corporation, 'HOBO MX1104 Data Logger Specifications', 2021.

[254] IQAir, 'Technical Specifications: AirVisual Pro', 2021. https://www2.iqair.com/sites/default/files/documents/AVP_TS_NA.pdf (accessed Apr. 14, 2021).

[255] Omega, 'Digital Sound Level Data Logger HHSL-101', 2021. https://br.omega.com/omegaFiles/das/pdf/HHSL-101.pdf (accessed Apr. 14, 2021).

[256] F. Leens, 'An introduction to I2C and SPI protocols', *IEEE Instrumentation and Measurement Magazine*, vol. 12, no. 1, pp. 8–13, 2009. doi: 10.1109/MIM.2009.4762946.

[257] M. Wu and H. Zhang, 'Interface Technology', vol. 119, no. Icacie, pp. 26–30, 2017.

[258] E. T. Louokdom *et al.*, 'Small-scale modular multilevel converter for multi-terminal DC networks applications: System control validation', *Energies (Basel)*, vol. 11, no. 7, 2018, doi: 10.3390/en11071690.

[259] V. Jain and R. M. Patrikar, 'A Low-Cost Portable Dynamic Droplet Sensing System for Digital Microfluidics Applications', *IEEE Trans Instrum Meas*, vol. 69, no. 6, pp. 3623–3630, 2020, doi: 10.1109/TIM.2019.2932526.

[260] R. Vallat, 'Pingouin: statistics in Python', *J Open Source Softw*, vol. 3, no. 31, p. 1026, Nov. 2018, doi: 10.21105/joss.01026.

[261] P. E. Shrout and J. L. Fleiss, 'Intraclass correlations: uses in assessing rater reliability.1. Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979, 86:420–8.', *Psychol Bull*, vol. 86, no. 2, pp. 420–8, 1979.

[262] T. K. Koo and M. Y. Li, 'A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research', *J Chiropr Med*, vol. 15, no. 2, pp. 155–163, Jun. 2016, doi: 10.1016/j.jcm.2016.02.012.

[263] A. Hartmann, S. Luzi, K. Murer, R. A. de Bie, and E. D. de Bruin, ‘Concurrent validity of a trunk tri-axial accelerometer system for gait analysis in older adults’, *Gait Posture*, vol. 29, no. 3, pp. 444–448, 2009, doi: 10.1016/j.gaitpost.2008.11.003.

[264] D. Giavarina, ‘Understanding Bland Altman analysis’, *Biochem Med (Zagreb)*, vol. 25, no. 2, pp. 141–151, 2015, doi: 10.11613/BM.2015.015.

[265] J. Kim, S. Schiavon, and G. Brager, ‘Personal comfort models – A new paradigm in thermal comfort for occupant-centric environmental control’, *Build Environ*, 2018, doi: 10.1016/j.buildenv.2018.01.023.

[266] G. Coulby, Adrian. K. Clear, O. Jones, and A. Godfrey, ‘Low-cost, multimodal environmental monitoring based on the Internet of Things’, *Build Environ*, vol. 203, p. 108014, Oct. 2021, doi: 10.1016/j.buildenv.2021.108014.

[267] S. McDonald, R. Vieira, and D. W. Johnston, ‘Analysing N-of-1 observational data in health psychology and behavioural medicine: a 10-step SPSS tutorial for beginners’, *Health Psychol Behav Med*, vol. 8, no. 1, pp. 32–54, Jan. 2020, doi: 10.1080/21642850.2019.1711096.

[268] N. Duan, R. L. Kravitz, and C. H. Schmid, ‘Single-patient (n-of-1) trials: A pragmatic clinical decision methodology for patient-centered comparative effectiveness research’, *J Clin Epidemiol*, vol. 66, no. 8 SUPPL.8, pp. S21–S28, 2013, doi: 10.1016/j.jclinepi.2013.04.006.

[269] S. McDonald *et al.*, ‘The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview’, 2017, doi: 10.1080/17437199.2017.1316672.

[270] ASHRAE, ‘ANSI/ASHRAE Standard 55-2020: Thermal Environmental Conditions for Human Occupancy’, 2021.

[271] G. Coulby, ‘Apple Health Data Parser v1.0 (gcoulby/apple-health-data-parser)’, Jul. 2021, doi: 10.5281/ZENODO.5112807.

[272] R. Vieira, S. McDonald, V. Araújo-Soares, F. F. Sniehotta, and R. Henderson, ‘Dynamic modelling of n-of-1 data: powerful and flexible data analytics applied to individualised studies’, *Health Psychol Rev*, vol. 11, no. 3, pp. 222–234, 2017, doi: 10.1080/17437199.2017.1343680.

[273] Z. Sultan, C. Spanos, E. Goh, G. C. Ping, and M. Sapar, ‘Living lab integrates IEQ technologies in singapore’, *ASHRAE J*, vol. 62, no. 2, pp. 68–71, 2020.

[274] M. Etiwy *et al.*, 'Accuracy of wearable heart rate monitors in cardiac rehabilitation', *Cardiovasc Diagn Ther*, vol. 9, no. 3, pp. 262–271, 2019, doi: 10.21037/cdt.2019.04.08.

[275] J. Nikles *et al.*, 'Establishment of an International Collaborative Network for N-of-1 Trials and Single-Case Designs', *Contemp Clin Trials Commun*, vol. 23, no. July, p. 100826, Sep. 2021, doi: 10.1016/j.conctc.2021.100826.

[276] G. Coulby, A. Clear, O. Jones, and A. Godfrey, 'Towards remote healthcare monitoring using accessible IoT technology: State-of-the-art, insights and experimental design', *Biomedical Engineering Online2*, 2020.

[277] P. M. Barker and T. J. McDougall, 'Two interpolation methods using multiply-rotated piecewise cubic hermite interpolating polynomials', *J Atmos Ocean Technol*, vol. 37, no. 4, pp. 605–619, 2020, doi: 10.1175/JTECH-D-19-0211.1.

[278] K. B. Velt and H. A. M. Daanen, 'Thermal sensation and thermal comfort in changing environments', *Journal of Building Engineering*, vol. 10, no. February, pp. 42–46, 2017, doi: 10.1016/j.jobe.2017.02.004.

[279] S. Moritz, A. Sardá, T. Bartz-Beielstein, M. Zaefferer, and J. Stork, 'Comparison of different Methods for Univariate Time Series Imputation in R', 2015.

[280] M. Saunders, P. Lewis, and A. Thornhill, *Research methods for business students*, Fifth., no. 5. Pearson Education, 2009. Accessed: Jan. 07, 2019. [Online]. Available: http://www.americanbankhttps//www.academia.edu/23374295/Research_Methods_for_Business_Students_5th_Editionid/cir991%5Cnhttp://www.scielo

[281] C. Robson, *Real world research: A Resource for Social Scientists and Practitioner-Researchers*, 2nd ed. Oxford, U.K.: Oxford: Blackwell Publishers Ltd., 2002. [Online]. Available: http://search.ebscohost.com/login.aspx?direct=true&db=rfh&AN=ATLA0001462268&site=ehost-live

[282] A. B. Ryan, 'Post-Positivist Approaches to Research', *Researching and Writing your thesis: a guide for postgraduate students.*, pp. 12–26, Jun. 2006, Accessed: Jan. 07, 2019. [Online]. Available: http://eprints.maynoothuniversity.ie/874/

Appendix A

Theoretical Perspective and Methodology

Previous literature around measuring health and wellbeing with sensor technology highlights a need for more objective measurements of wellbeing that require more deductive approaches to research. By aligning this approach with the research onion created by Saunders *et al.* [280], shown in Figure 54, it is indicative that a positivist philosophy would be most applicable for conducting this research. However, there is a steady decline in positivism as an applicable philosophical standpoint for social research, as positivism holds a distinction between the researcher and the researched subject [281]. This stems from the scientific origins of positivism, whereby the only knowledge available to scientific researchers is the objective knowledge that forms the basis on which to test all hypotheses. This commonly causes positivist philosophies to be too binary or dualistic for social research, as research participants are often treated as objects that can be observed. Since human knowledge cannot simply be observed, nor does it adhere to a universal law, an approach is needed that allows the human experience to support objective empirical evidence.

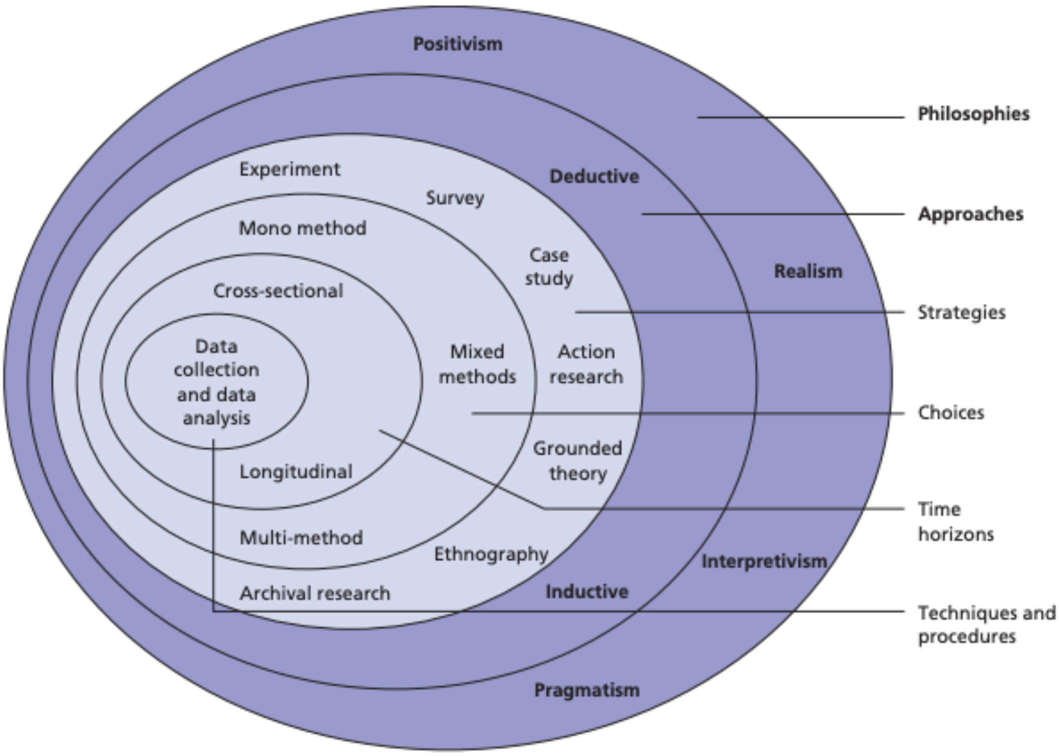

Figure 54 - The research onion diagram by Saunders et al. [280] *which depicts the spectrum of epistemological standpoints and philosophical perspectives.*

Ryan [282] puts forward the argument that a post-positivist philosophy allows researchers to adopt a learning role instead of a testing role when conducting research that requires learning from human participants. Post-positivism allows researchers to use their own knowledge and experience to interpret ideas and understand the complexity of the human mind. With all that said, post-positivist researchers still commit to objectivity but instead, they understand that human participants and they themselves are subject to biases that influence their research [281].

While I always considered my own philosophical standpoint to be more closely aligned with positivism, the dualistic nature of the philosophy was not completely aligned to my own. I still find it a difficult task to label my epistemological standpoint, and personally feel that I inhibit traits from across the spectrum. Though, I would argue that post-positivism suitably captures that alignment towards scientific truth, while accepting that there is value in the human experience. Thus, the studies that will be undertaken throughout this PhD will involve collecting data from electronic devices. However, the analysis of the data will involve interpreting and learning from human participants about how these environmental or biometric changes affect their wellbeing. For this reason, adopting a post-positivist stance will allow the research to utilise deductive approaches to collect and analyse objective measurements of wellbeing and use subjectivity, interpretation, and human experience to support the findings.

Appendix B

Creative Commons 4.0 License

Creative Commons Legal Code

Attribution 4.0 International

Official translations of this license are available in other languages.
Creative Commons Corporation ("Creative Commons") is not a law firm and does not provide legal services or legal advice. Distribution of Creative Commons public licenses does not create a lawyer-client or other relationship. Creative Commons makes its licenses and related information available on an "as-is" basis. Creative Commons gives no warranties regarding its licenses, any material licensed under their terms and conditions, or any related information. Creative Commons disclaims all liability for damages resulting from their use to the fullest extent possible.

Using Creative Commons Public Licenses

Creative Commons public licenses provide a standard set of terms and conditions that creators and other rights holders may use to share original works of authorship and other material subject to copyright and certain other rights specified in the public license below. The following considerations are for informational purposes only, are not exhaustive, and do not form part of our licenses.

Considerations for licensors: Our public licenses are intended for use by those authorized to give the public permission to use material in ways otherwise restricted by copyright and certain other rights. Our licenses are irrevocable. Licensors should read and understand the terms and conditions of the license they choose before applying it. Licensors should also secure all rights necessary before applying our licenses so that the public can reuse the material as expected. Licensors should clearly mark any material not subject to the license. This includes other CC-licensed material, or material used under an exception or limitation to copyright.

Considerations for the public: By using one of our public licenses, a licensor grants the public permission to use the licensed material under specified terms and conditions. If the licensor's permission is not necessary for any reason–for example, because of any applicable exception or limitation to copyright–then that use is not regulated by the license. Our licenses grant only permissions under copyright and certain other rights that a licensor has authority to grant. Use of the licensed material may still be restricted for other reasons, including because others have copyright or other rights in the material. A licensor may make special requests, such as asking that all changes be marked or described. Although not required by our licenses, you are encouraged to respect those requests where reasonable.

Creative Commons Attribution 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

Section 1 – Definitions.

a. **Adapted Material** means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.
b. **Adapter's License** means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.
c. **Copyright and Similar Rights** means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in Section 2(b)(1)-(2) are not Copyright and Similar Rights.
d. **Effective Technological Measures** means those measures that, in the absence of proper authority, may not be circumvented under laws fulfilling obligations under Article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.
e. **Exceptions and Limitations** means fair use, fair dealing, and/or any other exception or limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.
f. **Licensed Material** means the artistic or literary work, database, or other material to which the Licensor applied this Public License.
g. **Licensed Rights** means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.
h. **Licensor** means the individual(s) or entity(ies) granting rights under this Public License.
i. **Share** means to provide material to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.
j. **Sui Generis Database Rights** means rights other than copyright resulting from Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, as amended and/or succeeded, as well as other essentially equivalent rights anywhere in the world.
k. **You** means the individual or entity exercising the Licensed Rights under this Public License. **Your** has a corresponding meaning.

Section 2 – Scope.

a. **License grant**.
   1. Subject to the terms and conditions of this Public License, the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive, irrevocable license to exercise the Licensed Rights in the Licensed Material to:
      A. reproduce and Share the Licensed Material, in whole or in part; and
      B. produce, reproduce, and Share Adapted Material.

Appendix C

PRISMA-ScR Checklist

Table S1: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	2
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	3
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	3
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	4
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	4
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	3
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	4
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	4 – 7
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	n/a
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	4

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	4
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	4 - 7
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	n/a
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	5 – 7
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	8 - 17
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	17 – 20
Limitations	20	Discuss the limitations of the scoping review process.	19
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	19 - 20
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	20

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.
* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.
† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).
‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting.
§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. Ann Intern Med. 2018;169:467–473. doi: 10.7326/M18-0850.

Appendix D

Statement of Authorisation for Digital Health Chapter

Order Completed

Thank you for your order.

This Agreement between Mr. Graham Coulby ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

License Number	5397780628489
License date	Sep 28, 2022

Printable Details

Licensed Content

Licensed Content Publisher	Elsevier
Licensed Content Publication	Elsevier Books
Licensed Content Title	Digital Health
Licensed Content Author	Graham Coulby,Fraser Young
Licensed Content Date	Jan 1, 2021
Licensed Content Pages	14

Order Details

Type of Use	reuse in a thesis/dissertation
Portion	excerpt
Number of excerpts	8
Format	both print and electronic
Are you the author of this Elsevier chapter?	Yes
How many pages did you author in this Elsevier book?	13
Will you be translating?	No

About Your Work

Title	Personalised Environmental Monitoring Building of Occupants: Integration of Scalable Technologies
Institution name	Northumbria University
Expected presentation date	Jan 2023

Additional Data

Order reference number	GCOU_DHCR2022
Portions	short excerpts of text content only (no images, tables or figures) from following pages: 166,167,168,169,172,173

Requestor Location

Requestor Location	Mr. Graham Coulby Northumbria University Newcastle Upon Tyne, Tyne and Wear NE18ST United Kingdom Attn: Mr. Graham Coulby

Tax Details

Publisher Tax ID	GB 494 6272 12

$ Price

Total	0.00 GBP

Appendix E

Informed Consent Form Example

Participant Informed Consent Form

Part One: Information for Participants

Research Project Title

Personalised Health and Wellbeing: A Longitudinal Approach to Monitoring Building Occupants using an n-of-1 Methodology.

Researcher Details

Graham Coulby *(PhD Student)*
g.d.coulby@northumbria.ac.uk
Department of Computer and Information Sciences, Faculty of Engineering and Environment, Northumbria University

Introduction

Part one of this consent form will serve to provide you with information and invite you to be part of this research. You do not have to decide today whether you will participate in the research. Before you decide, you can talk to anyone you feel comfortable with about the information provided here. Should this consent form contain words that you do not understand, further clarification of any points can be provided upon request. If you have questions at any stage of the research, you may ask them directly of the researcher.

Purpose of Research

This project will investigate health and wellbeing of building occupants using wearable sensor data, responses from surveys and data collected from sensors that measure Indoor Environmental Quality (IEQ) factors such as air quality, thermal comfort, noise and light levels. Whilst there has been extensive research into the effects of IEQ on the health and wellbeing of building occupants, few studies have taken a personalised approach to measure by holistically measuring and individual. This project will adopt an n-of-1 methodology to gain a longitudinal assessment of the individual. The n-of-1 methodology involves repetition around the measurement of an individual over a longer period of time compared to traditional observational studies (Mcdonald and Johnston, 2019). n-of-1 methods can inform many types of research design, but they can also be particularly useful in exploratory research and early-phase trials (Lillie *et al.*, 2011). The versatility of n-of-1 methods is acknowledged across disciplines and they enable the measurement of high resolution data, which can be used to identify individualised patterns of behaviour (Mcdonald *et al.*, 2017). For the purposes of this study, n-of-1 will be used to identify individualised perceptions of IEQ and to understand personalised thresholds of comfort by linking subjective survey data responses with objective readings from sensors.

It is hoped that insights may be drawn from the data obtained from wearables into the study. The key aims of this project are to evaluate

1. the potential of n-of-1 methods in identifying individualised perceptions of IEQ and
2. to understand personalised thresholds of comfort by linking subjective survey data responses with objective readings from sensors
3. whether personalised monitoring with n-of-1 methods can be used to better evaluate building performance, by using people to measure buildings instead of using buildings to measure people.
4. whether wearable sensors can be used to address the subjectivity around health and wellbeing in environmental studies by providing quantitative context to how occupants are affected by environmental changes within buildings.

Type of Research Intervention

The research will require participants to wear a fitness tracker/smart watch (FT) that links to their smartphone and tracks heart rate, sleep data and activity data. Participants will also be required to place passive sensors in their home office and living room at their home. These sensors will passively collect data about the environment. No personal data will be collected from any device. Surveys will also be used to collect additional data on your perceptions of the environmental quality. These will be distributed via a provided Amazon Echo Dot and delivered as a voic- controlled application.

Participant Requirements

As a participant of the study, you will be required to:

- Wear an FT that tracks heart rate and activity data throughout the day.
- Provide exported data from Apple Health at the end of the study. Apple Health contains more data than is required/ethically approved for this study. Therefore, an extraction method will be provided to you that will allow you to extract, approve and submit on the required data.
- Complete survey/interview at the beginning and end of the project that should take around 15-20 minutes to complete.
- Complete surveys throughout the measurement period. These surveys will be short surveys consisting of three-point responses only and should not take more than one minute to complete. These will be delivered by an Alexa routine that will be run at 10:30am and will prompt you first with an auditory notification and then it will ask you if you have time to conduct the survey. A 'no' response at this stage will end the survey for that day. You are encouraged to manually run the survey on such occasions by saying "Alexa, open Graham's PhD survey".

Equipment

The Participant will receive the following equipment on loan from the researcher for use throughout the study:

1. 2 x IQAir Air Visual Pro PM2.5 Monitor
2. 2 x HOBO MX1102 CO2 Logger
3. 2 x HOBO MX 1104 Light Intensity Monitor
4. 2 x Prototype multimodal IEQ monitors
5. 2 x TP Link 4G Routers with SIM cards
6. 2 x TP Link Smart Power Strips
7. 1 x Amazon Echo Dot

One of each device will be placed in both the home office and living spaces. All devices are preconfigured to connect to the TP Link router, which will be loaded with pay-monthly sim cards. This will ensure no IoT devices will communicate using your home WiFi connection. Therefore, if security issues are present in any of the networked devices, any security breaches will be contained within a sandboxed environment.

Survey Questions and Responses

1. How is the temperature? (Cold | Comfortable | Hot)
2. How is the Humidity? (Dry | Comfortable | Humid)
3. How is the light? (Dark | Comfortable | Bright/light)
4. How is the noise? (Quiet | Comfortable | Noisy)
5. How is the air circulation? (Draughty | Comfortable | Stuffy)
6. Is it Dusty? (Yes | No)
7. Are there any odours? (Yes | No)

Voluntary Participation and Right to Withdraw

1. **Your decision to participate in this pilot study is voluntary**. You may choose not to participate or withdraw from the study for any reason, at any time, without penalty. Prior to rescinding your participation, you should tell the research team if you decide to leave the study so that the time and date when participation stopped can be logged. If a substantial amount of data has been collected at that point you may be asked to conduct a close out survey, though, you are not obligated to do this.

2. The research team can stop the study or your participation in the study at any time without your consent if you fail to follow directions for participating in the study, if the study is cancelled, or for administrative reasons.

Procedures

The study will be broken down into four phases; initialisation, monitoring, closeout and analysis. In the initialisation phase and the closeout phase, participants will be required to conduct an interview.

1. The initialisation phase and the closeout phase will be used to instruct the participant on how to set up the equipment within their homes.
2. Throughout the monitoring phase, data will be passively captured from the FT and IEQ sensors.
3. Participants will be required to fill short, auditory, daily surveys throughout the monitoring period, which should take approximately one minute to complete. The survey will ask questions about participants perceptions on environment quality.
4. During the closeout period, participants will be required to submit data extracted from AppleHealth app.

Duration

The duration of the study will be approximately four months.

Confidentiality

1. You will be given an anonymised participant ID at the start of the study. All data collected on the study will be linked to that ID to maintain anonymity.
2. The data controller can provide an example of the data collected upon request.

Legal Basis

1. During the study participants retain the right of access as outlined in Article 12 of the EU general data protection regulation 2016/679, which states "The controller shall take appropriate measures to provide any information referred to in Articles 13 and 14 and any communication under Articles 15 to 22 and 34 relating to processing to the data subject in a concise, transparent, intelligible and easily accessible form, using clear and plain language."
2. Participants who choose to withdraw, or are involuntarily withdrawn, from the study early have the right to exercise Article 17 of the EU general data protection regulation 2016/679, which states "The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the [sub clauses of Article 17] applies".
3. In accordance with Article 5(1)e "Personal data will be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed".
4. The data being processed for this research is in the public interest and as such is covered by Article 6(1)e of the EU general data protection regulation 2016/679, which states that "Processing shall be lawful [if the] processing is necessary for the performance of a task carried out in the public interest".
5. Article 9(2)a of the EU general data protection regulation 2016/679, states that the prohibitions regarding the processing of health data shall not apply if "the data subject has given explicit consent to the processing of those personal data for one or more specified purposes;" and Article 9(2)j of the EU general data protection regulation 2016/679, which states that the prohibitions regarding the processing of health data shall not apply if the "processing is necessary for... scientific or historical research purposes... which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject."

Part Two: Contact Information

If you have any questions, you can ask them now or later. If you wish to ask questions later, you may contact the data controller. This proposal has been reviewed and approved by Northumbria University's Research Ethics Committee REC, which is a committee whose task it is to make sure that research participants are protected from harm. These data are collected for legal reasons and will not be stored with or linked in any way to data captured during the study.

PhD Researcher

Graham Coulby

PhD Supervisors

Dr Alan Godfrey *(Principal Supervisor) – alan.godfrey@northumbria.ac.uk*

Dr Oliver Jones *(Industry Supervisor) – ojones@ryderarchitecture.com*

Affiliated University

Northumbria University
Sutherland Building
Newcastle upon Tyne
NE1 8ST

PhD Sponsor

Ryder Architecture
Cooper's Studio
14-18 Westgate Road
Newcastle Upon Tyne
NE1 3NN

Participant

Participant Name

Study Location

Participant's Home Address

Address of Participant's Work

Data Controller

Graham Coulby
g.d.coulby@northumbria.ac.uk

Part Three: Certificate of Consent

Research Project Title

Personalised Health and Wellbeing: A Longitudinal Approach to Monitoring Building Occupants using an n-of-1 Methodology.

Researcher Details

Graham Coulby *(PhD Student)*
g.d.coulby@northumbria.ac.uk
Department of Computer and Information Sciences, Faculty of Engineering and Environment, Northumbria University

Consent

Please tick where applicable:

I have carefully read and understood the Participant Information Sheet. ☐

I have had an opportunity to ask questions and discuss this study and I have received satisfactory answers. ☐

I understand I am free to withdraw from the study at any time, without having to give a reason for withdrawing, and without prejudice. ☐

I agree to take part in this study. ☐

I also consent to the retention of this data under the condition that any subsequent use also be restricted to research projects that have gained ethical approval from Northumbria University. ☐

Signature of Participant: ______________________________

NAME IN BLOCK LETTERS: ______________________________

Date: ______________________________

Signature of Researcher: ______________________________

NAME IN BLOCK LETTERS: ______________________________

Date: ______________________________

Appendix F

Pandas source code for synchronising data

```python
def parse_dt(x):
  return datetime.strptime(x, '%d/%m/%Y %H:%M:%S')

def get_series_data(file):
  series = read_csv(file, header=0, parse_dates=[0], in-
dex_col=0, squeeze=True, na_values=[''], date_parser=parse_dt)
  return pd.DataFrame(series.apply(pd.to_numeric, errors='coerce'))

def resample(series, span, sample_direction = 'up'):
  resampled = series.resample(span)

  if (sample_direction == 'up'):
    return resampled.interpolate(method='linear', limit=5)
  elif (sample_direction == 'down'):
    return resampled.mean()
  elif (sample_direction == 'pad'):
    return resampled.interpolate(method='pad')
  else:
    return None

def read_and_resample(file, span, sample_direction = 'up'):
  series = get_series_data(file)
  series = series[~series.index.duplicated()]
  return resample(series, span, sample_direction)

def resample_and_merge_files(files):
  dfs = []
  for file in files:
    dfs.append(read_and_resample(file["file"], file["span"], file["direction"]))
  return dfs

files = [
  {
    "file": "[PATH]",
    "span": '1T',
    "direction": 'pad'
  },
]

dfs = resample_and_merge_files(files)
```