



Co-funded by the  
Erasmus+ Programme  
of the European Union



# **Prediction of Outdoor Thermal Comfort Changes and Uncovering Mitigation Strategies based on Machine Learning Algorithm**

A decision support tool for climate-sensitive design: A case study of  
Glasgow, UK

New sha Modjrian

A thesis submitted for the Joint programme of  
Master in Urban Climate & Sustainability

September 2022





|  |   |                                |
|--|---|--------------------------------|
| <b>Author</b><br>Modjrian, New sha   | <b>Publication Type</b><br>Thesis               | <b>Completion year</b><br>2022 |
| <b>Number of pages:</b> 126  |   |                                |
| <b>Supervisor I</b><br>Prof. Rohinton Emmanuel   | <b>Supervisor II</b><br>Prof. José Miguel Nieto |                                |
| <b>Title</b><br>Prediction of Outdoor Thermal Comfort Changes and Uncovering Mitigation Strategies based on Machine Learning Algorithm A decision support tool for climate-sensitive design: A case study of Glasgow, UK   |   |                                |
| <b>Degree:</b> Master in Urban Climate & Sustainability  |   |                                |
| <b>Abstract</b><br><p>Cities are becoming increasingly warm as a result of climate change and increasing population (Dimoudi <i>et al.</i>, 2013). It forewarns severe weather and climate catastrophes happening considerably sooner than initially anticipated. There will be more heatwaves, longer warm seasons, and shorter cold seasons. It is predicted that the UK will experience 50% more hot days in the decades ahead (Arnell <i>et al.</i>, 2021). The magnitude of the heat island in Glasgow can approach approximately 4 °C under certain climatic circumstances (Krüger, Drach and Emmanuel, 2018). Such thermal variation in Glasgow demands more action; otherwise, rising UHI intensity may have disastrous impacts on inhabitants' health.</p> <p>In light of managing and controlling the outdoor environment to meet thermal comfort, this research developed a framework for predicting and simulating the thermal comfort proxies depending on the secondary and historical data. This framework will facilitate the decision-making process from a climatic perspective at the initial planning stage by forecasting the heat stress changes in urban settings. This study has considered Generalized Linear Regression (GLR), Exploratory Regression, Spatial Autocorrelation, Geographically Weighted regression (GWR), and Artificial Neural Network (ANN) algorithms to find the best fit model to predict the proxies of thermal comfort.</p> <p>The ANN model achieved better performance predicting Mean Radiant Temperature (MRT) and Land Surface Temperature LST considering the 12 predictors. Among all the factors, SVF was the best factor in predicting MRT, while NDBI was recognised as a significant variable in LST forecasting. The framework was validated by predicting LST for vacant lands in Glasgow. Three scenarios were considered to evaluate the impact of greening strategies (100%-50%-0%).</p> <p>The results from LST prediction depict the modest effects of urban greenery in decreasing heat stress at the surface level under a non-linear trend. The UTCI simulation in ENVI-met was applied in the framework to understand thermal comfort. The impact of shadowing from plants and buildings, for instance, could alter thermal comfort depending on the area's compactness and openness in Glasgow's central district. Heat mitigation measures at the level of lowering the surface temperature do not always meet human thermal comfort. The association between LST and MRT has made it clear that it is impossible to establish a direct connection between them. It is advantageous to provide such a tool that may be used by policymakers who are less proficient in climatology, to forecast changes in heat stress in urban environments swiftly.</p> |   |                                |
| <b>Keywords</b><br>Machine Learning, Universal Thermal Comfort Index, Climate-sensitive Design, Heat Mitigation, Artificial Neural Network, Glasgow  |   |                                |
| <b>Originality statement.</b> I hereby declare that this Master's dissertation is my own original work, does not contain other people's work without this being stated, cited and referenced, has not been submitted elsewhere in fulfilment of the requirements of this or any other award.   | <b>Signature</b>                                |                                |





*Dedicated to all environmental defenders and  
climate activists around the globe who lost  
their lives to save the planet for us.*

# TABLE OF CONTENTS

|  |    |
|--|----|
| Acknowledgements.....  | 5  |
| List of Figures .....  | 7  |
| List of Tables .....   | 9  |
| Abbreviations .....  | 10 |
| Chapter 1: Introduction .....  | 13 |
| 1.1. Problem Statement .....   | 13 |
| 1.2. Aim and Objectives .....  | 14 |
| 1.3. Methodology Outline .....   | 15 |
| 1.4. Structure of Report .....   | 15 |
| Chapter 2: Literature Review .....   | 18 |
| 2.1. Global Warming and Increasing Temperature in Urban Settings.....                | 18 |
| 2.2. Urban Heat Island .....   | 19 |
| 2.3. Land Surface Temperature .....  | 20 |
| 2.4. Thermal Comfort in Outdoor Environment .....                                    | 21 |
| 2.5. Contributing Parameters Contributing to Outdoor thermal Comfort .....           | 22 |
| 2.6. Prediction Model of Outdoor Thermal Comfort.....                                | 23 |
| 2.7. Heat Stress Mitigation Strategies in Policy Making .....                        | 24 |
| 2.8. Research and Practice Gap: Climate Sensitive Decision Support Application ..... | 25 |
| Chapter 3: Study Area .....  | 27 |
| 3.1. City of Glasgow .....   | 27 |
| 3.2. Climate Profile of Glasgow.....   | 27 |
| 3.3. Glasgow Critical Climate Events and Plans .....                                 | 28 |
| 3.4. Vacant Land Condition in Glasgow .....  | 29 |
| Chapter 4: Method and Methodology .....  | 31 |
| 4.1. Research Philosophy, Approach and Framework .....                               | 31 |
| 4.1.1. Research Philosophy .....   | 31 |
| 4.1.2. Research Approach for Theory Development .....                                | 31 |

|  |    |
|--|----|
| 4.1.3. Methodological choice.....  | 31 |
| 4.1.4. Strategies .....  | 31 |
| 4.1.5. Time Perspective.....   | 32 |
| 4.2. Methodology Framework .....   | 32 |
| 4.3. Data Sourcing.....  | 33 |
| 4.4. Data Pre-processing.....  | 34 |
| 4.5. Data Processing.....  | 34 |
| 4.5.1. Independent Variables (Predictors).....                               | 35 |
| 4.5.2. Dependent Variables (Response Variables).....                         | 45 |
| 4.5.3. Standardisation of Dependent Variable .....                           | 49 |
| 4.6. Microclimate Modelling.....   | 49 |
| 4.6.1. ENVI-met.....   | 50 |
| 4.6.2. SOLWEIG .....   | 51 |
| 4.7. Regression-based Prediction Models .....                                | 52 |
| 4.7.1. Pearson Correlation Coefficient.....                                  | 52 |
| 4.7.2. Linear-based Regression Model.....                                    | 53 |
| 4.8. Nonlinear Regression Model: Artificial Neural Network (ANN) .....       | 54 |
| 4.8.1. The Architecture of the ANN network.....                              | 54 |
| 4.8.2. Factor Characterisation and Sensitivity Analysis .....                | 55 |
| 4.9. Evaluation of Models' Performance .....                                 | 55 |
| 4.10. Validation of Final Model: Scenarios for Vacant Lands .....            | 56 |
| Chapter 5: Results and Analysis- Part I .....                                | 59 |
| 5.1. LST: A Proxy for Human Thermal Comfort? .....                           | 59 |
| 5.1.1. Correlation Matrix: LST .....   | 59 |
| 5.1.2. Generalized Linear Regression and Exploratory Regression Results..... | 60 |
| 5.1.3. Summary of Variables Significance .....                               | 61 |
| 5.1.4. Spatial Autocorrelation (Global Moran's I) Test .....                 | 61 |
| 5.1.5. Geographically Weighted Regression (GWR) .....                        | 61 |
| 5.2. Non-linearity by Neural Network.....                                    | 66 |

|            |  |     |
|------------|--|-----|
| 5.2.1.     | Architecture of LST prediction model .....                             | 66  |
| 5.2.2.     | LST NN Model Performance.....  | 67  |
| 5.2.3.     | Sensitivity Analysis .....   | 69  |
| 5.3.       | Validation of NN models: LST Prediction and Human Thermal Comfort..... | 70  |
| 5.3.1.     | Predicted LST and Human Thermal Comfort .....                          | 74  |
| Chapter 6: | Results and Analysis- Part II .....                                    | 81  |
| 6.1.       | MRT and Human Thermal Comfort? .....                                   | 81  |
| 6.1.1.     | Correlation Matrix: MRT .....  | 81  |
| 6.1.2.     | LST and MRT Correlation .....  | 82  |
| 6.1.3.     | Generalized Linear Regression and Exploratory Regression results ..... | 82  |
| 6.1.4.     | Summary of Variables Significance .....                                | 83  |
| 6.1.5.     | Spatial Autocorrelation (Global Moran's I) Test .....                  | 83  |
| 6.2.       | Non-linearity by Neural Network: MRT .....                             | 84  |
| 6.2.1.     | Architecture of MRT prediction model.....                              | 84  |
| 6.2.2.     | MRT NN Model Performance.....  | 84  |
| 6.2.3.     | Sensitivity Analysis .....   | 85  |
| Chapter 7: | Discussion and Reflection .....  | 88  |
| 7.1.       | LST and MRT Prediction depending on Secondary Data .....               | 88  |
| 7.1.1.     | How Regression models work in the case of MRT and LST .....            | 88  |
| 7.1.2.     | Prediction Models Performance in Scenarios.....                        | 90  |
| 7.2.       | Urban Design Factors in LST and MRT Prediction.....                    | 91  |
| 7.2.1.     | Significant Parameters in Prediction.....                              | 91  |
| 7.2.2.     | Significant Dataset in Prediction .....                                | 92  |
| 7.3.       | LST as Proxy of Thermal Comfort Studies? .....                         | 92  |
| Chapter 8: | Conclusions and Recommendations .....                                  | 96  |
| 8.1.       | Summary of Findings.....   | 96  |
| 8.2.       | Limitations.....   | 97  |
| 8.3.       | Recommendation for Further Studies .....                               | 99  |
| Reference  | .....  | 101 |

|                   |     |
|-------------------|-----|
| Appendices.....   | 115 |
| Appendix I.....   | 115 |
| Appendix II.....  | 116 |
| Appendix III..... | 117 |
| Appendix IV.....  | 118 |
| Appendix V.....   | 119 |
| Appendix VI.....  | 120 |
| Appendix VII..... | 121 |

## ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, Professor Rohinton Emmanuel, for his patience, knowledge and support throughout my research. I also want to thank my second supervisor, Professor José Miguel Nieto, for his motivation and support. Your suggestions and critical feedbacks were invaluable in the development of my thesis.

I sincerely thank all MURCS lecturers and staff in Glasgow, Lahti and Huelva who made this journey memorable. Kiitos! Gracias!

In addition, I owe a lot to my friends in MURCS family, for all the special memories we shared. I cherish our friendship and am sincerely grateful for their support in this adventure.

Most importantly, I owe my family a genuine and sincere debt of appreciation. A great distance separated us, but I could always feel their presence by my side. To my wonderful sister, Nadia, for her unlimited support and love throughout my life. To my dearest parents, Shokouh and Alireza, who have been my enduring heroes because of their guidance, love, and happiness.



## LIST OF FIGURES

|  |    |
|--|----|
| Figure 1 A two-layer categorisation of thermal change in a schematic depiction of the urban environment (Oke, 1976).....             | 19 |
| Figure 2 Graphic presentation of different UHI types (Oke et al., 2017) .....  | 20 |
| Figure 3 The illustration of the possible use of machine learning in urban form applications (Koumetio Tekouabou et al., 2022) ..... | 23 |
| Figure 4 The Boundary of Case Study-Glasgow .....  | 27 |
| Figure 5 Thesis Framework .....  | 32 |
| Figure 6 The workflow of pre-processing and processing steps .....   | 35 |
| Figure 7 Spatial distribution of NDBI retrieved from Landsat 8 (2018).....   | 36 |
| Figure 8 Spatial distribution of NDVI retrieved from Landsat 8 (2018).....   | 37 |
| Figure 9 Spatial distribution of DEM.....  | 37 |
| Figure 10 Spatial distribution of DSM .....  | 39 |
| Figure 11 Spatial distribution of CDSM.....  | 39 |
| Figure 12 Spatial distribution of SVF.....   | 40 |
| Figure 13 Spatial distribution of LULC.....  | 40 |
| Figure 14 Spatial distribution of Road Area Ratio .....  | 41 |
| Figure 15 Spatial distribution of Vegetation Surface Fraction.....   | 41 |
| Figure 16 Spatial distribution of Building Height .....  | 42 |
| Figure 17 Spatial distribution of Building Surface Fraction .....  | 43 |
| Figure 18 Left: NO <sub>x</sub> Right: PM <sub>10</sub> .....  | 44 |
| Figure 19 Distribution of Predictors (Left: before normalisation; Right: after normalisation).....                                   | 44 |
| Figure 20 Procedure of LST retrieval.....  | 45 |
| Figure 21 Spatial distribution of Land Surface Temperature retrieved from Landsat 2018 ...   | 46 |
| Figure 22 Flowchart of SOLWEIG model (Lindberg et al., 2022) .....   | 47 |
| Figure 23 Spatial distribution of Mean Daytime MRT(5:00-21:00) .....   | 47 |
| Figure 24 Spatial distribution of MRT at 13:00 (peak).....   | 48 |
| Figure 25 Spatial distribution of MRT at 16:00.....  | 48 |
| Figure 26 Pearson Correlation between Predictors and LST .....   | 59 |
| Figure 27 DEM and LST in GWR .....   | 62 |
| Figure 28 DSM and LST in GWR .....   | 62 |
| Figure 29 CDSM and LST in GWR .....  | 63 |
| Figure 30 SVF and LST in GWR .....   | 63 |
| Figure 31 NDVI and LST in GWR.....   | 64 |
| Figure 32 NDBI and LST in GWR.....   | 64 |



|   |    |
|---|----|
| Figure 33 NOx and LST in GWR .....  | 64 |
| Figure 34 PM10 and LST in GWR .....   | 65 |
| Figure 35 Group of DEM_SVF_NDVI_NOX_PM10 and LST in GWR.....                            | 65 |
| Figure 36 Group of DEM_SVF_NDVI_NDBI_NOX_PM10 and LST in GWR.....                       | 66 |
| Figure 37 Optimisation of Hyperparameters (The Number of Neurons in Hidden Layer) ..... | 67 |
| Figure 38 Architecture of Neural Network .....  | 67 |
| Figure 39 left: Error histogram.....  | 68 |
| Figure 40 Residual plots for the NN based LST model.....                                | 68 |
| Figure 41 Distribution of biased prediction in the LST model .....                      | 69 |
| Figure 42 Feature Significance in LST Prediction Model from sensitivity analysis .....  | 70 |
| Figure 43 a: Best Scenario b: Intermediate Scenario c: Worst Scenario .....             | 71 |
| Figure 44 Predicted LST for two sites in the city centre .....                          | 72 |
| Figure 45 LST Fluctuation of predicted for the proposed Scenarios Site A .....          | 73 |
| Figure 46 LST Fluctuation of predicted for the proposed Scenarios Site B .....          | 74 |
| Figure 47 Site A UTCI index .....   | 75 |
| Figure 48 Site A UTCI absolute difference .....   | 75 |
| Figure 49 Site B UTCI index .....   | 76 |
| Figure 50 Site B UTCI absolute difference .....   | 76 |
| Figure 51 Site A MRT.....   | 77 |
| Figure 52 Site B MRT .....  | 77 |
| Figure 53 Simulated Surface Temperature Top: Site A Bottom: Site B .....                | 78 |
| Figure 54 LST absolute difference Top: Site A Bottom: Site B .....                      | 79 |
| Figure 55 Pearson Correlation between Predictors and MRT .....                          | 81 |
| Figure 56 left: Error histogram.....  | 84 |
| Figure 57 Residual plots for the NN based MRT model.....                                | 85 |
| Figure 58 Distribution of biased prediction in MRT model .....                          | 86 |
| Figure 59 Feature Significance in MRT Prediction Model from sensitivity analysis .....  | 86 |

## LIST OF TABLES

|   |    |
|---|----|
| Table 1 Heat Stress Mitigation Strategies.....  | 24 |
| Table 2 Annual Meteorological Factors in Glasgow (1991-2021) (Climate Data, 2022) ..... | 28 |
| Table 3 Vacant Land in Glasgow .....  | 29 |
| Table 4 Data Sourcing .....   | 33 |
| Table 5 Specification of Landsat 8 images used in this study .....                      | 34 |
| Table 6 Classification by data source .....   | 35 |
| Table 7 Thermal Stress Categories (Błazejczyk et al., 2013).....                        | 49 |
| Table 8 ENVI-met models' settings.....  | 50 |
| Table 9 SOLWEIG settings .....  | 52 |
| Table 10 Data Partitioning .....  | 54 |
| Table 11 Pearson Correlation between Predictors and LST .....                           | 60 |
| Table 12 Summary of GLR Result-LST .....  | 60 |
| Table 13 LST Criteria Passed .....  | 60 |
| Table 14 LST Variables Significance .....   | 61 |
| Table 15 Global Moran's I Summary.....  | 61 |
| Table 16 Combined GWR models .....  | 66 |
| Table 17 Comparison of Scenarios and RS-LS.....   | 70 |
| Table 18 Comparison of LST predictions on selected scenarios.....                       | 73 |
| Table 19 Comparison of LST predictions on selected scenarios.....                       | 74 |
| Table 20 UTCI area changes through georeferenced simulated scenarios .....              | 77 |
| Table 21 Pearson Correlation between Predictors and MRT .....                           | 81 |
| Table 22 Correlation between MRT and LST .....  | 82 |
| Table 23 Summary of GLR Result-MRT .....  | 82 |
| Table 24 MRT Criteria Passed .....  | 83 |
| Table 25 MRT Variables Significance .....   | 83 |
| Table 26 Global Moran's I Summary.....  | 83 |
| Table 27 Comparison of GLR in LST and MRT models .....                                  | 88 |
| Table 28 Comparison of error metrics in MRT and LST prediction models.....              | 89 |

## ABBREVIATIONS

|               |  |
|---------------|--|
| <b>ANN</b>    | Artificial Neural Network  |
| <b>ASHRAE</b> | American Society of Heating, Refrigerating, and Air-Conditioning Engineers |
| <b>BD</b>     | Big Data   |
| <b>BSF</b>    | Building Surface Fraction  |
| <b>CDSM</b>   | Vegetation Canopy of Digital Surface Model                                 |
| <b>CEIP</b>   | Climate Emergency Implementation Plan                                      |
| <b>DEM</b>    | Digital Elevation Model  |
| <b>DSM(b)</b> | Digital Surface Model  |
| <b>EEA</b>    | European Environment Agency  |
| <b>EBK</b>    | Empirical Bayesian Kriging   |
| <b>ECMWF</b>  | European Centre for Medium-Range Weather Forecasts                         |
| <b>GIS</b>    | Geographical Information System  |
| <b>GLR</b>    | Generalized Linear Regression  |
| <b>GWR</b>    | Geographically Weighted Regression   |
| <b>IoT</b>    | Internet of Things   |
| <b>IPCC</b>   | Intergovernmental Panel on Climate Change                                  |
| <b>IUPS</b>   | International Union of Physiological Sciences                              |
| <b>LiDAR</b>  | Light Detection And Ranging  |
| <b>LST</b>    | Land Surface Temperature   |
| <b>LULC</b>   | Land Use and Land Cover  |
| <b>MAE</b>    | Mean Absolute Error  |
| <b>MIR</b>    | Mid Infrared Spectrometers   |
| <b>ML</b>     | Machine Learning   |
| <b>MRT</b>    | Mean Radiant Temperature   |
| <b>MSE</b>    | Mean Squared Error   |
| <b>NIR</b>    | Near Infrared Spectrometers  |
| <b>NOx</b>    | Nitrogen Oxides  |

|              |  |
|--------------|--|
| <b>NDBI</b>  | Normalized Difference Built-up Index                   |
| <b>NDVI</b>  | Normalized Difference Vegetation Index                 |
| <b>OLS</b>   | Ordinary Least Squares                                 |
| <b>OS</b>    | Ordnance Survey  |
| <b>OTC</b>   | Outdoor Thermal Comfort                                |
| <b>PET</b>   | Physiological Equivalent Temperature                   |
| <b>PM10</b>  | Particulate Matter 10 micrometres or less in diameter  |
| <b>PM2.5</b> | Particulate Matter 2.5 micrometers or less in diameter |
| <b>RAR</b>   | Road Area Ratio  |
| <b>REML</b>  | Restricted Maximum Likelihood                          |
| <b>RH</b>    | Relative Humidity                                      |
| <b>RMSE</b>  | Root Mean Squared Error                                |
| <b>RSD</b>   | Remotely Sensed Derived                                |
| <b>SAQD</b>  | Scottish Air Quality Database                          |
| <b>SVF</b>   | Sky View Factor  |
| <b>TIR</b>   | Thermal Infra-Red                                      |
| <b>TS</b>    | Thermal Stress   |
| <b>UHI</b>   | Urban Heat Island                                      |
| <b>UK</b>    | United Kingdom   |
| <b>UKCP</b>  | UK Climate Projections                                 |
| <b>UN</b>    | United Nations   |
| <b>USGS</b>  | United States Geological Survey                        |
| <b>UTCI</b>  | Universal Urban Thermal Comfort                        |
| <b>VIF</b>   | Variance Inflation Factor                              |
| <b>VSF</b>   | Vegetation Surface Fraction                            |



# CHAPTER 1: INTRODUCTION

## 1.1. Problem Statement

According to the United Nations, cities will accommodate almost two-thirds of the world's population by 2050 (United Nations, 2019). Densely populated areas push cities to grow either vertically or horizontally, causing greater heat generation, a wider blockage impact against urban air circulation, larger accumulation of solar radiation, and eventually a reduction in long-wave emission to the atmosphere due to building blockage and usage of artificial materials (Kleerekoper, Van Esch and Salcedo, 2012; Ashtiani, Mirzaei and Haghighat, 2014; Mirzaei, 2015). The lack of greenery, the considerable use of impervious surfaces, the urban spatial structure absorbing solar radiation and the reduction in wind velocity, along with expanding anthropogenic activities, all contribute to the evolution of heat islands (Lee, Kim and Yun, 2016). According to Li *et al.* (2018), global warming and increased urbanization have dramatically aggravated UHI effects. Several challenges linked to UHIs have been identified, including increased air pollution, heavy precipitation, excessive energy consumption, rising energy prices, and increasing greenhouse gas emissions (Synnefa *et al.*, 2008), thermal discomfort, the heat stress, high fatality rate among the physically susceptible population (Alves and Lopes, 2017). Therefore, excessive heat makes the condition for human beings to experience thermal discomfort in outdoor spaces along with increasing environmental disasters, health-related issues.

One of the major challenges in developing a realistic tool to address the effects of climate change and overheating on a city and its citizens is the lack of an appropriate simulation and prediction model of thermal stress as a result of city development (Parsaee *et al.*, 2019). Massive volumes of data have been generated in cities due to the application of modern information technology (Gadgets, IoT, Sensors, and so on) (Koumetio Tekouabou *et al.*, 2022). Machine Learning algorithms can potentially deliver significant improvements in terms of both understanding and forecasting climate consequences. What makes machine learning so powerful tool is its ability to compute massive amounts of data, process data, and exhibit nonlinear behaviour (Koc and Acar, 2021).

Cold spells and heat waves are occurring as a result of ongoing global warming (Seltenrich, 2015). The impacts of increasing global temperature on people's lives, particularly during heat waves, vary across the UK. It is predicted that the UK would experience 50% more hot days in the decades ahead (Arnell *et al.*, 2021). The statistical predictions in UKCP18 report made it clear that in the following years, there would be a greater risk of warmer, wetter winters and

hotter, drier summers with higher frequency and severity of anomalies across the UK and specifically Scotland (Met Office, 2019). According to a recent study by Krüger, Drach and Emmanuel (2018), the magnitude of heat island in Glasgow can approach approximately 4 °C under certain climatic circumstances. Such thermal variation in Glasgow demands more actions; otherwise, rising UHI intensity may have disastrous impacts on inhabitants' health. In this regard, it is expected that the urban population would become substantially vulnerable under the same trend shortly.

In light of managing and controlling the outdoor environment to meet thermal comfort, the concern is to develop a prediction model to alert city planners and designers about the changes in increased temperature in outdoor spaces and their effects on thermal comfort in outdoor areas. Givoni *et al.* (2003) emphasized the importance of the given tool to assess the significance of altering the outdoor environment in a particular direction by specific design details. Additionally, moving toward climate-sensitive planning and designing approaches is hindered by the absence of climate expertise in planner and designer communities (Keith *et al.*, 2021; Kelly Turner *et al.*, 2022). Understanding the heat-related effect of project development requires intensive fieldwork and running microclimate simulations, which are costly and time-consuming, as well as a suitable, strong, and robust system design. The current study seeks to put cutting-edge machine learning algorithms into practice in the case of Glasgow by providing a precise, affordable, and user-friendly tool for policymakers and planners to find noble solutions for managing the extra heat released into the atmosphere.

## 1.2. Aim and Objectives

The aim of the present study is to generate a machine learning-based model in terms of new urban development to facilitate the decision-making process at the initial planning stage from a climatic perspective. This model can forecast thermal comfort and heat stress changes in an urban setting. Urban morphology parameters of urban microclimate, land use and land cover, Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-Up Index (NDBI), Digital Elevation Model (DEM), Digital Surface Model (DSM), Vegetation Canopy of Digital Surface Model (CDSM), along with thermal comfort and heat island indices in case of Glasgow City will be employed to validate the climate-sensitive model. Thus, it is guided by the objectives listed below.

1. Identify the contributing parameters to outdoor thermal comfort (OTC) and heat island derived from historical and secondary data at the spatial and quantitative levels.

2. Investigate the spatial distribution of outdoor thermal comfort indices and proxies in order to explore the linearity and non-linearity relationship with urban design features, air quality factors and remotely sensed derived (RSD) data.
3. Train and evaluate the performance of various prediction models for outdoor thermal comfort indices to identify robust models and significant variables.
4. Simulate scenarios on vacant lands (vacant and derelict) to evaluate the effect of predictors on outdoor thermal comfort.
5. Proposing the holistic workflow and through the use of ArcGIS, MATLAB, and ENVI-met to predict changes in thermal comfort based on secondary data for policymakers.

Finally, the purpose of this research is to provide answers to the following critical questions:

1. How well can LST and MRT be estimated by deploying the historical data on land cover, along with urban physical factors?
2. To what extent can LST, the variable derived from remote sensing and historical data be considered as a proxy of the urban thermal comfort index?
3. To what extent can the changes in urban design factors meet the goal of heat mitigation at the city scale and street canyon?

### 1.3. Methodology Outline

By utilizing a set of tools and software to measure and analyse the urban physical aspects to its thermal comfort levels, as well as to investigate statistical interactions between exploratory and response variables, this project opens up opportunities to estimate outdoor thermal comfort using open source and secondary data at the level of Glasgow city. ArcGIS Pro 2.8, MATLAB 2021b, QGIS 3.24.3, UMEP tool 3.20, SOLWEIG v2022a, SketchUp Pro 2022, ENVI-met 5.0.3, and ENVI-metINX are a handful of programmes that are considered essential for describing this procedure to policymakers.

### 1.4. Structure of Report

The present dissertation is organized into the following major sections:

The first chapter begins with a concise introduction to the research, outlining the study's aim, objectives, and scope, followed by a review of the literature, with a critical evaluation of state of the art to shed light on the main drivers behind the topic. Chapters 2-3 mainly discuss the technical methods, data preparation and retrieval of the major indicators of the study in the City of Glasgow. Chapters 5 and 6 highlight that the primary analysis was resulting from the



critical parameters linked to thermal comfort using linear and nonlinear analysis. Chapter 7 continues the discussion and analysis of the important topics from the investigated data. The closing chapter of this report discusses the key findings, limitations and directions for further study.



## CHAPTER 2: LITERATURE REVIEW

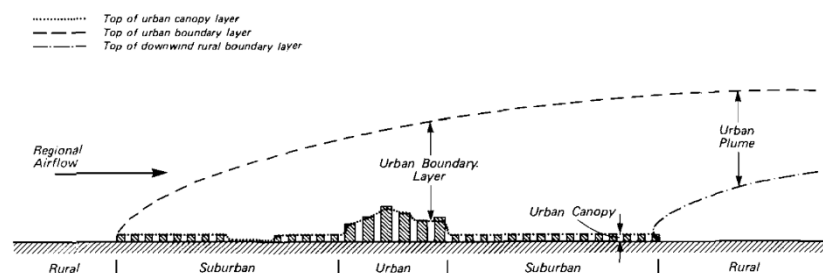
### 2.1. Global Warming and Increasing Temperature in Urban Settings

It is estimated that global temperatures will exceed 1.5 °C within two decades under the current rate of global warming. It shows the importance of reducing GHG emissions immediately as well as limiting global warming rapidly; otherwise, managing the 2 °C would be unachievable (IPCC, 2021). IPCC analysis indicated that anthropogenic GHG emissions account for roughly 1.1 °C since the pre-industrial era. Climate change is expected to worsen in all locations over the next few decades, according to the research anticipation. There will be more heatwaves, longer warm seasons, and shorter cold seasons. On the other hand, cities are expected to accommodate roughly two-thirds of the global population within three decades (United Nations, 2019), exposing more people to the risk of global warming (Alexander and Arblaster, 2008).

The higher temperature in urban areas compared to their surroundings, known as urban heat island (UHI) (Arnfield, 2003; Wilby, 2008), which is more prevalent throughout the night (Wilby, 2003; Memon, Leung and Liu, 2009), can aggravate the excessive heat caused by heatwaves for city dwellers and world's cities (Patz et al., 2005; Watkins, Palmer and Kolokotroni, 2007; Zhao et al., 2018). Extreme weather events, heatwaves and cold spells, which are now recognized risks linked with increased mortality and morbidity, present seasonal threats to the health and well-being of vulnerable people (Seltenrich, 2015) with adverse effects on environmental quality and water supply (Zhao et al., 2018; Venter, Krog and Barton, 2020). Tan et al. (2010) explored that heat-related deaths have been demonstrated to be higher in urban areas than in the countryside. According to Emmanuel and Baker (2012), climate change is the most severe threat that modern humans have ever struggled with. Rising temperatures caused by climate change, heat waves, and the UHI effect necessitate immediate response; otherwise, future generations would suffer from the same repercussions as the current generation.

## 2.2. Urban Heat Island

The term "urban heat islands" (UHI) refers to the extra heat generated in cities as a result of the density of buildings, automobiles, and impervious surfaces. UHI, as a micro-climatic phenomenon, is associated with larger air temperature variations in urban areas than rural regions (Oke, 1976, 1982). Kim (1992) characterized this regional climatic phenomenon with soil albedo and moisture availability playing crucial roles in total boundary forcing and its feedback consequences. Oke (1976) differentiated two forms of atmospheric heat islands: urban boundary layer (UBL) and urban canopy layer (UCL), shown in *Figure 1*. Stewart (2011) added two more categories to Oke's (1976) classification: surface heat island and subsurface heat island. The former exists at the surface, which is identified by the temperature differential between urban roads and buildings, whereas the subsurface lies beneath the city surface (*Figure 2*).



*Figure 1* A two-layer categorisation of thermal change in a schematic depiction of the urban environment (Oke, 1976)

The effect of UHI fluctuates according to changes in city size (Oke, 1973; Dan *et al.*, 2010), urban vegetation (Tan, Lau and Ng, 2016; Leal Filho *et al.*, 2021), surface material (Watkins, Palmer and Kolokotroni, 2007), land use and land cover (LULC) conversion (Jin, Kessomkiat and Pereira, 2011; Yang *et al.*, 2017), urban spatial structure (Jusuf and Hien, 2009), street geometry (Oke *et al.*, 1991), climatic conditions (Kim and Baik, 2002; Eastin *et al.*, 2018), and anthropogenic activities (Lee, Kim and Yun, 2016) and time (Wilby, 2007; Memon, Leung and Liu, 2009).

The morphological and climatic factors that cause UHI can be either (1) Controllable: human-related activities or (2) Uncontrollable: fixed factors with the possibility of prediction (Memon, Leung and Chunho, 2008). Climatic parameters, meteorological variables, and geographical features are considered uncontrollable. In contrast, the other variables are related to urban planning and design, decision-making, building geometry, urban morphology, sky view factor, water bodies, urban vegetation, surface attributes and building

material and transportation (Levermore *et al.*, 2018; Parsaee *et al.*, 2019). In this respect, the UHI is a direct outcome of urbanisation (Levermore *et al.*, 2018).

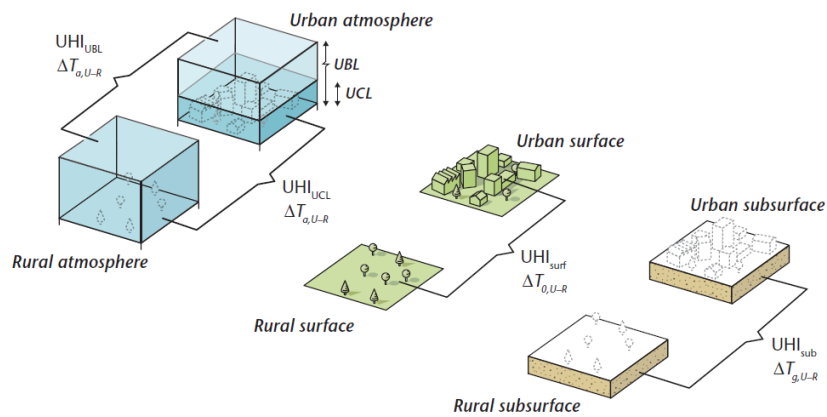


Figure 2 Graphic presentation of different UHI types (Oke *et al.*, 2017)

### 2.3. Land Surface Temperature

Land surface temperature (LST) corresponds to surface energy flow as a thermodynamic state marker. One of the most empirical approaches for continually monitoring the thermal environment in an urban context is remote sensing LST derived from Thermal Infra-Red (TIR) spectral (Li and Meng, 2018; Michel *et al.*, 2021). The prediction and modelling of land surface temperature (LST) result from an interaction between meteorological and geographical parameters from remote sensing datasets, which ultimately shed light on creating resilient and sustainable environments (Şahin *et al.*, 2012; Afrakhteh *et al.*, 2016). As one of the foremost factors for UHI magnitude, the capacity to track LULC modifications within a specified site makes LST a crucial technique in UHI investigations (Bokaie *et al.*, 2016).

One of the significant challenges of LST simulation in UHI mapping is the lack of continuous and heterogeneous data with appropriate resolution (Anderson *et al.*, 2021); however, other parameters such as Mean Radiant Temperature (MRT), a proxy for human thermal experience, and air temperature can be simulated and investigated alongside LST (Kelly Turner *et al.*, 2022). In addition, the land surface temperature calculated from satellite data does not accurately reflect how much heat people feel outside. Other elements like shade, wind velocity, and relative humidity also affect how much heat people are exposed to (Norton *et al.*, 2015). UHI modelling has to be the most visually appealing way of identifying places most exposed to excessive heat risk, but it does not represent all the risk's intricacies. Hence, socio-economic factors and meteorological data should be included in simulation and prediction models (Keith, Meerow and Wagner, 2019).

## 2.4. Thermal Comfort in Outdoor Environment

Several academics who specialise in the domains of climate studies, urban planning, and environmental sustainability have acquired an interest in the topic of outdoor thermal comfort as a result of the major detrimental effects that climate change has had on population health (Jamei *et al.*, 2016; Aram *et al.*, 2020). Thermal comfort is defined by the ASHRAE<sup>1</sup> as a state of mind that displays satisfaction with the thermal environment (ASHRAE, 2010), while The IUPS Thermal Commission<sup>2</sup> described it as 'subjective indifference to the thermal environment.' In other words, integrating the build-up area and climate-sensitive solutions into the environment considers the activities that are expected to take place there.

There are several indices taken into account in research on thermal comfort assessment. The air temperature of the reference condition, which causes the thermophysiological model to respond similarly, serves as the definition of the Universal Thermal Climate Index (UTCI). Frequently mentioned, the temperature derived from the collection of air temperature, wind speed, relative humidity, and radiation contributes to the calculation of the UTCI value (Blazejczyk *et al.*, 2012; Krüger, 2021). In response to wind speed, humidity, and thermal radiation in hot and cold climates, Bröde *et al.* (2012a) described the UTCI as an index broadly relevant to areas of application of human biometeorology. The Physiological Equivalent Temperature (PET), another thermal comfort indicator, is generated from a heat-budget model of the human body that takes the skin's temperature variations in a complicated outdoor environment into account (Höppe, 1999). By assessing the intensities of the short- and long-wave radiation on the human body, the Mean Radiant Temperature (MRT) index was developed to measure the environment's thermal stress (Matzarakis, Rutz and Mayer, 2007). MRT is characterized as "The temperature of an imaginary isothermal "black" enclosure in which a solid body or inhabitant would exchange the same amount of heat through radiation as in the actual non-uniform enclosure" by IUPS Thermal Commission (2001). Compared to other indices, UTCI better captures the temporal changes in temperature conditions at the human scale (Blazejczyk *et al.*, 2012) and is computed by the following equation (Bröde *et al.*, 2012a), where  $T_a$  is air temperature;  $T_r$  is MRT;  $v_a$  is wind velocity;  $p_a$  is the water pressure and offset is a deviation from  $T_a$ .

$$\text{Eq. (1) } UTCI(T_a; T_r; v_a; p_a) = T_a + \text{Offset}(T_a; T_r; v_a; p_a)$$

---

<sup>1</sup> American Society of Heating, Refrigerating, and Air-Conditioning Engineers

<sup>2</sup> Commission for Thermal Physiology of the International Union of Physiological Sciences

Numerous methods have been created in order to make it simpler to assess the impacts of various urban design and planning approaches on microclimate and human thermal comfort, owing to the advantages of numerical simulations and the increasing processing power of computers (Nice, Coutts and Tapper, 2018; Guo *et al.*, 2020). These programmes include ENVI-met, Rayman, and SOLWEIG, which can also calculate MRT and microclimate conditions (Gál and Kántor, 2020). SOLWEIG can simulate the SVF and MRT through horizontal 2-dimensional simulations under shortwave and longwave global radiation data, giving it a substantial edge over rival tools while speeding up simulation times (Guo *et al.*, 2020).

## 2.5. Contributing Parameters Contributing to Outdoor thermal Comfort

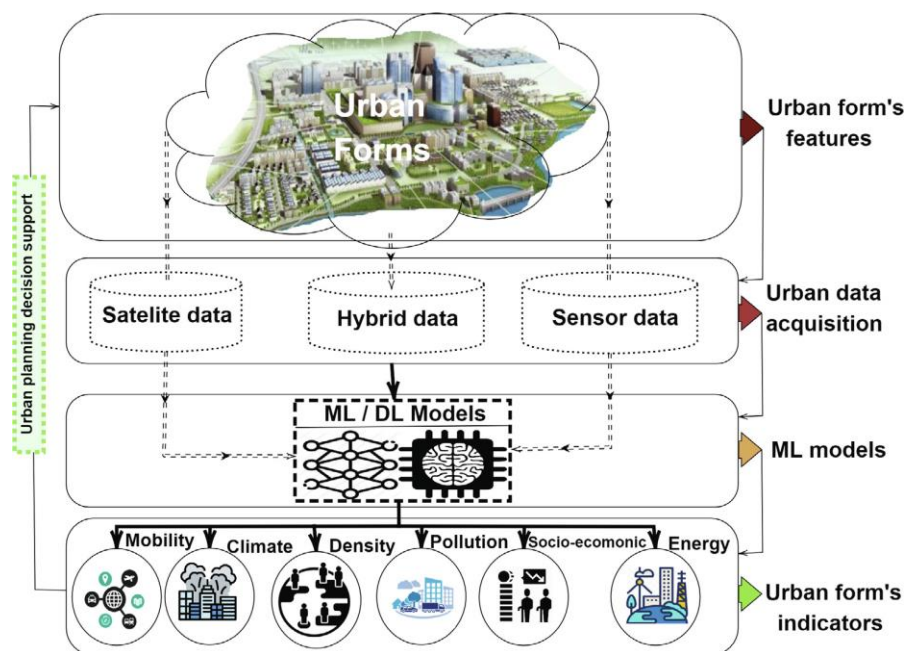
Four environmental aspects, together with two distinct characteristics can be used to sum up the remarkable parameters that influence how comfortable people are outside, such as air temperature, radiant temperature, humidity, air velocity, human activity, and clothing. There are significant changes in thermal conditions and, therefore, in how the human body regulates its internal temperature in dense urban environments due to the high number of commuters and vehicles as well as complex urban geometry (Lau *et al.*, 2021). Such a compact urban shape minimises air velocity and shields solar radiation (Zhang *et al.*, 2022). The contribution of urban geometry is for analysing the openness of urban spaces. Factors like SVF (Souza, Rodrigues and Mendes, 2003; He *et al.*, 2015; Yan *et al.*, 2022), H/W (Krüger, Minella and Rasia, 2011), building density and coverage ratios (Mehrotra, Bardhan and Ramamritham, 2020), floor area ratios (FARs), as well as vegetation density (Alavipanah *et al.*, 2015) and coverage ratios (Duarte *et al.*, 2015) have been studying.

Yilmaz *et al.* (2021) examined the effect of microclimate properties and urban form factors on air quality and thermal comfort in a cold climate in Turkey. It was observed that urban centres with high building density had recorded higher PET and air pollution owing to low wind velocity. The study by Lai and Cheng (2009) showed that increasing intensity of heat islands led to a higher level of air pollutants and severe UHI corresponded to greater quantities of PM<sub>2.5</sub> and PM<sub>10</sub> than weak UHIs did ( $P < 0.05$ ). Research conducted in Athena revealed that levels of PM<sub>10</sub> and NO<sub>2</sub> dramatically rose during a heatwave, which was accompanied by reduced thermal comfort due to the worsening of the air quality and greater concentrations of air pollutants (Papanastasiou, Melas and Kambezidis, 2015). Urban microclimatic features and urban pollution levels can have a substantial impact on human health during a heatwave.

## 2.6. Prediction Model of Outdoor Thermal Comfort

The development and implementation of new methods and policies for monitoring the additional heat trapped in outdoor environments have been made easier by the capacity to predict the influence of heat stress factors on outdoor air temperatures (Mangal, Rajesh and Misra, 2020). While Machine Learning (ML) techniques have the potential to deliver significant advancements in terms of both explanation and prediction of climate repercussions (Koc and Acar, 2021), Big Data (BD) science has speeded up the related technological advancements and real-world application by the addition of a readily available, easily accessible, and quickly growing volume of data (Hassani, Huang and Silva, 2019).

Several methods of prediction in the climate research context have been used, including trend and regression analysis (Kolokotroni and Giridharan, 2008; Liu *et al.*, 2021); multiple regression (Lee, Kim and Yun, 2016; Levermore *et al.*, 2018); neural network (Gobakis *et al.*, 2011; Ashtiani, Mirzaei and Haghghat, 2014; Lee, Kim and Yun, 2016; Equere *et al.*, 2021; Liu *et al.*, 2021); tree-regression (Pena Acosta *et al.*, 2021); time-series analysis (Lee *et al.*, 2020); long short-term memory-recurrent neural network (LSTM-RNN) (Koc and Acar, 2021); support vector regressor (SVR) (Mendez-Astudillo and Mendez-Astudillo, 2021). *Figure 3* illustrates how ML approaches have been at the core of successful modelling of urban indices to help in planning processes that address present and future problems (Koumetio Tekouabou *et al.*, 2022).



*Figure 3* The illustration of the possible use of machine learning in urban form applications (Koumetio Tekouabou *et al.*, 2022)



## 2.7. Heat Stress Mitigation Strategies in Policy Making

Overload of research on UHI mitigation measures has failed to offer valuable information at the level of human thermal comfort due to a focus on temperature differences between urban and suburban regions (Martilli, Krayenhoff and Nazarian, 2020). The main problem is that there is no clear definition of rural area (Martin-Vide, Sarricolea and Moreno-García, 2015). Furthermore, the magnitude of UHI is greater at night, whereas thermal comfort and heat stress are greatest during the day (Martilli, Krayenhoff and Nazarian, 2020).

In theory, lowering the urban heat is in light of reducing the net radiation. On this basis, plants and reflecting materials are extensively employed as passive techniques of heat mitigation in metropolitan areas (Taleghani, 2018). Another efficient method of reducing the excess heat trapped in an urban canyon, which lessens the level of thermal comfort for people, is to create shading through vegetation, buildings or urban furniture (Ketterer and Matzarakis, 2014; Emmanuel, 2021). From an urban design perspective, street orientation along with building density can be useful in considering MRT and thermal comfort as they affect the wind velocity and radiation angles on the surfaces and open areas (Thorsson *et al.*, 2017). The strategies for heat mitigation can be summarised in *Table 1*.

*Table 1 Heat Stress Mitigation Strategies*

| <b>Strategies</b>                | <b>Scale</b>              | <b>Source</b>                                   |
|----------------------------------|---------------------------|---|
| <b>Vegetation</b>                | Park                      | (Taleghani, 2018)                               |
|                                  | Street Tree               |   |
|                                  | Green Roof                |   |
|                                  | Green Wall                |   |
| <b>Material with high albedo</b> | White roof                | (Taleghani, 2018)                               |
|                                  | Reflective Pavement       |   |
| <b>Shading Effect</b>            | Vegetation                | (Ketterer and Matzarakis, 2014)                 |
|                                  | Building (shadow-casting) | (Ketterer and Matzarakis, 2014; Emmanuel, 2021) |
|                                  | Urban furniture           |   |
| <b>Urban Geometry</b>            | Urban Orientation         | (Thorsson <i>et al.</i> , 2017)                 |
|                                  | Building density          |   |

According to Pearsall (2017), green conversion of vacant lots has the potential to reduce imbalances in urban heating in the city's most vulnerable regions of Virginia in the United States; however, only a few studies have attempted to shed light on the role of vacant lands in managing and mitigating extra heat.

## 2.8. Research and Practice Gap: Climate Sensitive Decision Support Application

Cities accommodate more than half of the world's population. Since cities are more prone to heat stress when the urban environment is heavily constructed and densely populated with a scarcity of green spaces (Leal Filho *et al.*, 2021), specific considerations in planning policy should be provided to tackle several environmental issues as consequences of urbanization. Givoni *et al.* (2003) pointed out the value of using the specified tool to evaluate the impact of changing the environment in a certain way through specific design aspects. Numerous studies have been conducted to investigate the mitigation interventions to understand, simulate and minimize the negative impact of heat stress along with scoping the future of urban climate conditions with conservative scenarios; however, most of them have not been included in spatial planning guidelines owing to:

1. Lack of a generalized multi-faceted model (Parsaee *et al.*, 2019) to predict thermal comfort based on historical and secondary data
2. Time-consuming and costly field studies of thermal comfort, specifically at an initial stage
3. Lack of a tool for planners with less knowledge of climatology to apply in the decision-making process (Keith *et al.*, 2021; Kelly Turner *et al.*, 2022)
4. Lack of high-resolution remote sensing data (Anderson *et al.*, 2021) and availability of easy-access data on air temperature (Oertel, Emmanuel and Drach, 2015) to evaluate the model with different contexts to find its reliability and validity

The present research tries to eliminate the first three challenges that have been identified in order to outline the prediction model for three radical scenarios:

- 1- Best-case scenario: the desirable vision based on the application of greenery in mitigation strategies
- 2- Intermediate-case scenario: Combination of both best and worst scenarios by greening 50% of the land
- 3- Worst-case scenario: fast-growing built-up strategies without consideration of urban vegetation.



## CHAPTER 3: STUDY AREA

### 3.1. City of Glasgow

Glasgow, Scotland's most populous city (Britannica, 2022), is one of the financial capitals of the United Kingdom (Glasgow City Council, 2017b). In 2020, Glasgow City recorded a population of 635,640, with growth rates of 8.3% since 1998 (National Records of Scotland, 2021). Glasgow's official boundary (Figure 4) encompasses about 175 km<sup>2</sup> with GPS coordinates of 55° 51' 39.2976" N and 4° 15' 5.1588" W (LatLong, 2022).

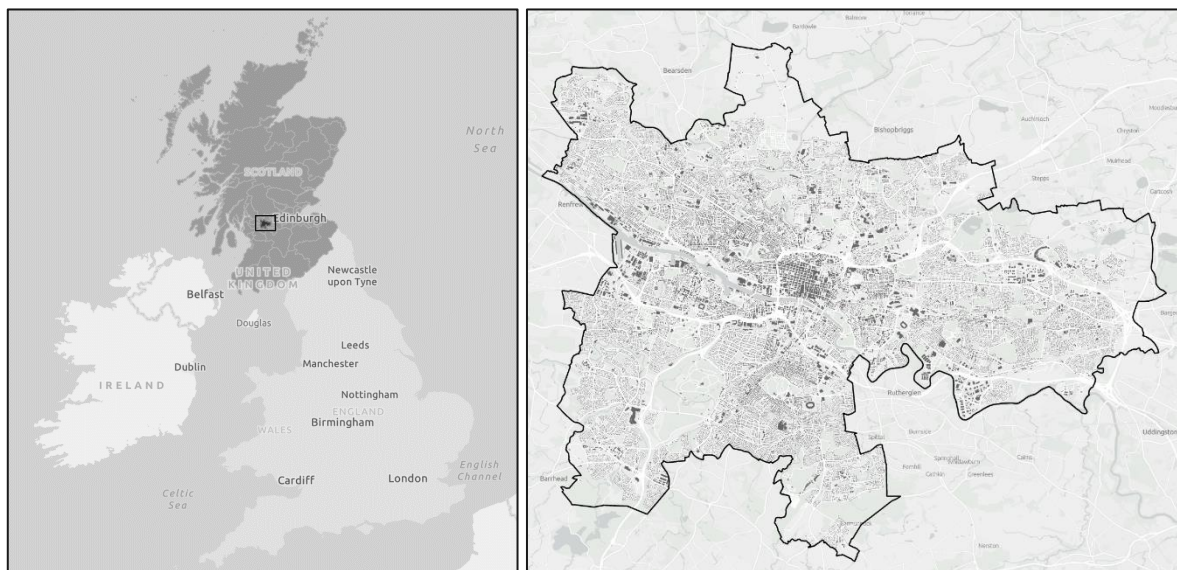


Figure 4 The Boundary of Case Study-Glasgow

In 2017, Glasgow City Development Plan versioned the solutions to beat the climate change severity under green strategies (adaptation and mitigation), including promoting green infrastructure, defining a greenbelt and redevelopment of vacant lands along with energy management (Glasgow City Council, 2017b). Its place-making strategy puts health and well-being at the centre of aspiration. According to the strategic plan for 2017-2022, supporting the sustainable and low-carbon city was a long-term perspective with the prioritization of supporting climate resilience places (Glasgow City Council, 2017a). Upon the City Council's announcement of a climate emergency in 2019, Glasgow has undertaken the ambitious vision of being the UK's first carbon-neutral city by 2030 (Glasgow City Council, 2021a)

### 3.2. Climate Profile of Glasgow

Glasgow's climate is classified as marine temperate. Glasgow typically has moderate winters and cool summers, with high yearly precipitation (Met Office, 2016). The average maximum

temperature occurs in June, July, and August, while January is the year's coldest month summarised in *Table 2* (Climate Data, 2022).

*Table 2 Annual Meteorological Factors in Glasgow (1991-2021) (Climate Data, 2022)*

|                                  | January             | February            | March               | April                | May                  | June                 | July                 | August               | September            | October              | November            | December            |
|----------------------------------|---------------------|---------------------|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|---------------------|
| Avg. Temperature °C (°F)         | 3 °C<br>(37.5) °F   | 3.3 °C<br>(37.9) °F | 4.6 °C<br>(40.3) °F | 6.9 °C<br>(44.5) °F  | 9.9 °C<br>(49.8) °F  | 12.5 °C<br>(54.5) °F | 14.2 °C<br>(57.5) °F | 13.7 °C<br>(56.6) °F | 11.9 °C<br>(53.5) °F | 8.9 °C<br>(47.9) °F  | 5.5 °C<br>(41.9) °F | 3.3 °C<br>(37.9) °F |
| Min. Temperature °C (°F)         | 0.8 °C<br>(33.4) °F | 0.6 °C<br>(33) °F   | 1.5 °C<br>(34.7) °F | 3.2 °C<br>(37.7) °F  | 6 °C<br>(42.8) °F    | 9 °C<br>(48.2) °F    | 10.8 °C<br>(51.4) °F | 10.4 °C<br>(50.8) °F | 8.9 °C<br>(48) °F    | 6.1 °C<br>(43) °F    | 3 °C<br>(37.4) °F   | 0.9 °C<br>(33.5) °F |
| Max. Temperature °C (°F)         | 5.3 °C<br>(41.5) °F | 6 °C<br>(42.7) °F   | 7.7 °C<br>(45.9) °F | 10.5 °C<br>(50.9) °F | 13.4 °C<br>(56.2) °F | 15.6 °C<br>(60.1) °F | 17.3 °C<br>(63.1) °F | 16.7 °C<br>(62.1) °F | 14.8 °C<br>(58.6) °F | 11.5 °C<br>(52.7) °F | 7.8 °C<br>(46) °F   | 5.5 °C<br>(41.9) °F |
| Precipitation / Rainfall mm (in) | 124<br>-4           | 103<br>-4           | 96<br>-3            | 83<br>-3             | 85<br>-3             | 93<br>-3             | 98<br>-3             | 108<br>-4            | 90<br>-3             | 118<br>-4            | 109<br>-4           | 121<br>-4           |
| Humidity (%)                     | 89%                 | 85%                 | 82%                 | 78%                  | 76%                  | 78%                  | 79%                  | 82%                  | 83%                  | 86%                  | 88%                 | 89%                 |
| Rainy days (d)                   | 13                  | 11                  | 11                  | 11                   | 11                   | 11                   | 12                   | 13                   | 11                   | 12                   | 12                  | 11                  |
| avg. Sun hours (hours)           | 2.9                 | 3.8                 | 4.3                 | 5.9                  | 7.2                  | 6.5                  | 6.4                  | 6                    | 4.7                  | 4                    | 3.5                 | 2.8                 |

### 3.3. Glasgow Critical Climate Events and Plans

According to the Met Office, the hottest day of 2018 was June 28<sup>th</sup>, with a temperature of 31.9 degrees Celsius, surpassing the previous record of 31.1 °C in 1950 (The Guardian, 2018). On the same day, BBC Scotland News tweeted proof that the Glasgow Science Centre roof had begun to melt due to the rising temperatures up to 31.9 °C (BBC Scotland News, 2018). In the summer of 2022, Glasgow and UK experienced more than three heatwaves in which the temperature reached almost 31 °C (Sandercock, 2022).

According to the recent Glasgow climate adaptation plan 2022– 2030, Scotland's top ten hottest years have all happened since 1997, given the available records from 1884. The mean temperature in the recent decade was .68 °C, greater than the average from 1960 to 1997 (Glasgow City Council, 2022). The statistical predictions in UKCP18 report made it clear that in the following years, there would be a greater risk of warmer, wetter winters and hotter, drier summers with higher frequency and severity of anomalies across the UK and specifically Scotland (Lowe *et al.*, 2018; Met Office, 2019).

Climate Emergency Implementation Plan (CEIP) study, along with Glasgow's Climate Plan, highlighted the role of the integrated network of green and open spaces to lessen the effect of urban heat island (Glasgow city council, 2020; Glasgow City Council, 2021b). The UHI effect was shown to be responsible for an increase in the average temperature of 4-6 degrees Celsius in 2021, which was recognized as the warmest summer since 1884 (Emmanuel *et al.*, 2021).

### 3.4. Vacant Land Condition in Glasgow

The high number of vacant lands in Glasgow, which covers more than 5 per cent of the whole city, is significant to be considered as the potential for future development and steps toward sustainability. About 667 vacant lots with a surface coverage of almost 9 square kilometres need policy makers' attention for bringing the climate-sensitive strategies into practice as mentioned in Glasgow City Development Plan to be redeveloped. The vacant lands are divided into three main categories on the EDINA Digimap dataset vacant land, derelict land, and vacant with buildings (*Table 3*).

*Table 3 Vacant Land in Glasgow*

| <b>Category of Vacant Land</b>    | <b>Number</b> | <b>Area (m<sup>2</sup>)</b> |
|-----------------------------------|---------------|-----------------------------|
| <b>Vacant Land</b>                | 401           | 3,811,434                   |
| <b>Derelict Land</b>              | 249           | 5,097,740                   |
| <b>Vacant land with buildings</b> | 17            | 120,438                     |
| <b>Glasgow Vacant Lands</b>       | 667           | 9,029,612                   |



## CHAPTER 4: METHOD AND METHODOLOGY

### 4.1. Research Philosophy, Approach and Framework

#### 4.1.1. Research Philosophy

According to the literature review section, minor changes in microclimate properties, urban form characteristics, LULC, and air quality parameters might adversely affect the thermal comfort balance in outdoor spaces. The pragmatic research philosophy was employed to achieve the project's principal aim of developing a holistic model for decision makers since it emphasizes addressing challenges within a real-world situation (Salkind, 2010).

#### 4.1.2. Research Approach for Theory Development

Under the abduction approach, data is utilised to analyse, identify themes, and explain patterns in order to generate a new/modify a current theory, which is then assessed, typically by further data collection (also referred to as retrodution) (Saunders, Lewis and Thornhill, 2018). Thermal comfort proxies retrieved from secondary data will be investigated in this study to understand the patterns underlying the outdoor thermal comfort concept.

#### 4.1.3. Methodological choice

Quantitative research design, which is typically connected with a deductive method (Saunders, Lewis and Thornhill, 2018), is used to generate and utilize various numerical models (linear and nonlinear regression) for the simulation and prediction process.

#### 4.1.4. Strategies

Analysis methods relying on numerical models (linear and nonlinear regression) were identified based on a quantitative research design to uncover the underlying relationship between urban form characteristics, LULC, and air quality parameters and thermal comfort proxies in the case of Glasgow using GIS-based analysis methods and spatial statistical tools.



### 4.1.5. Time Perspective

Since UTCI is tightly correlated to time, this study investigates human thermal comfort variability for different scenarios based on the hottest day during a heatwave. In this sense, the simulation and prediction are based on the significant climate elements to cover 24 hours of a heatwave. However, the data used in this study were collected from the available dataset from the last five years.

## 4.2. Methodology Framework

According to the background literature, urban morphology at both scales of 2D and 3D, urban land use and land cover, along with air quality, all appeared to be related to the land surface temperature and thermal comfort. In light of this knowledge, *Figure 5* depicts the framework that has been taken into consideration.

There are five main steps in the entire procedure, which begin with reviewing of the literature and data gathering. To create the foundational dataset for the prediction models, data pre-processing and processing were put into account. Linear-based regressions and ANN models were employed to shed light relationship between predictors and response variables. Following that, the trained ANN model was used to estimate LST for quick scenarios

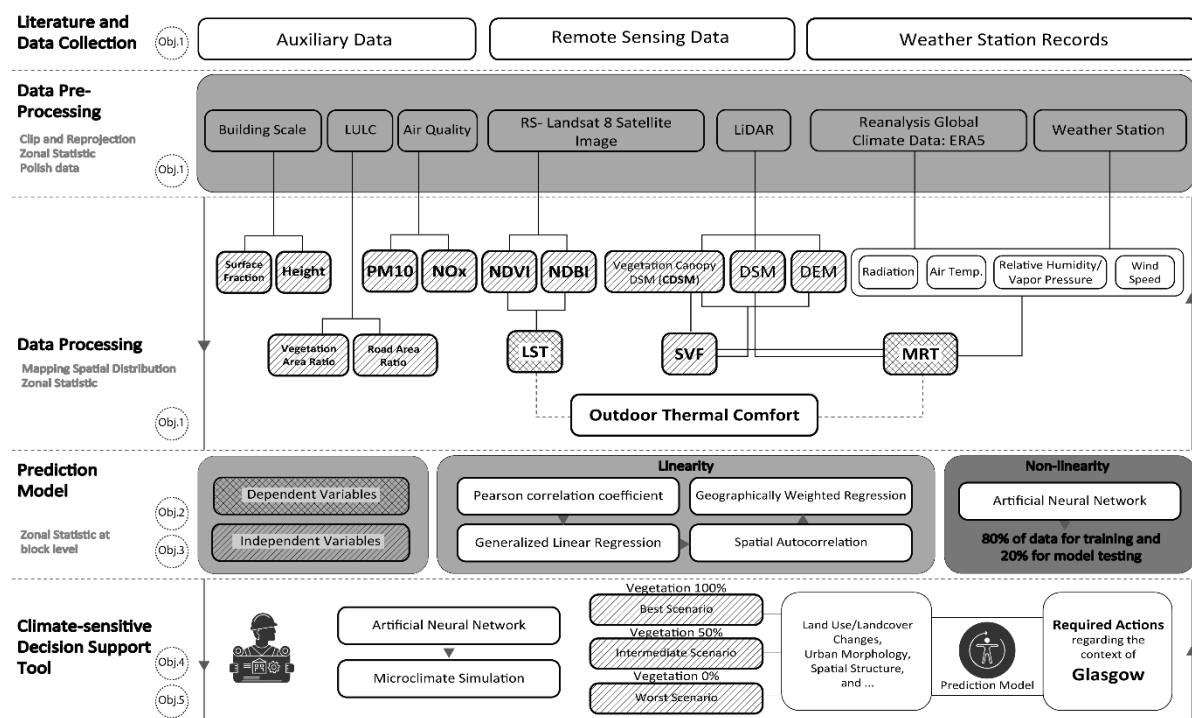


Figure 5 Thesis Framework

throughout the planning process. Eventually, the microclimate simulation was run to evaluate the performance of ANN.

By recognising these elements, the analysis essentially provides the key methodological steps to accomplish its stated goals:

- Identification and description of 3D and 2D urban forms along with vegetation indices, RSD data as independent variables or predictors
- Investigate the spatial distribution of land surface temperature and mean radiant temperature as proxies of human thermal comfort
- Analysing the connections between the predictors and response variable under linear and nonlinear regression techniques both spatially and statistically in order to compare the performance of various prediction models for outdoor thermal comfort
- Simulate scenarios on available lands (vacant and derelict) to evaluate the effect of predictors on outdoor thermal comfort
- Outline a predictive model and framework through the use of ArcGIS and MATLAB to develop a predictive thermal comfort tool based on secondary data

### 4.3. Data Sourcing

Digimap Edina and Ordnance Survey, as the national mapping agencies, offer a plethora of freely available spatial data on a wide range of environmental and geographical topics. These platforms provide data at the scale of cities, blocks, and buildings in various geospatial geodatabase formats. Glasgow’s LiDAR datasets gathered for 2021, which is published by Scotland’s catalogue of spatial data Metadata Portal<sup>3</sup> with free access for the public. Urban Atlas has provided land use and land cover characteristic for 2018 with the update from 2020 under the Copernicus programmes of the European Environment Agency (EEA). All the data needed to deliver the aim of this project is classified in *Table 4*.

*Table 4 Data Sourcing*

| <b>Data</b>                        | <b>Source</b>  | <b>Year</b> | <b>Format</b>  | <b>Application</b> |
|------------------------------------|--|-------------|----------------|--------------------|
| <b>Landsat 8 OLI and TIRS</b>      | United States Geological Survey (USGS) (USGS, 2018)                | 2018        | GeoTIFF        | LST, NDVI, NDBI    |
| <b>Land Use and Land Cover</b>     | VERISK Collection DIGIMAP (Digimap Edina, 2022b)                   | 2022        | SHP            | LULC               |
|                                    | Copernicus Land Monitoring Service - Urban Atlas (EEA) (EEA, 2020) | 2018-2020   | GPKG           |                    |
| <b>LiDAR: DEM, DSMb, CDSM, SVF</b> | Scottish Government (Scottish Government, 2021a)                   | 2021        | LAZ<br>GeoTIFF | 3D model           |

<sup>3</sup> SpatialData.gov.scot

|  |  |      |      |           |
|--|--|------|------|-----------|
| <b>Meteorological data</b>                 | ERA5–ECMWF (Shiny Weather Data)                  | 2018 | EPW  | Tmrt      |
| • <b>Air temperature</b>                   | (Shiny Weather Data, 2018)                       |      |      |           |
| • <b>Relative Humidity</b>                 | Weather Underground                              | 2018 | XLSX | Tmrt      |
| • <b>Wind speed</b>                        | (Weather Underground, 2018)                      |      |      | UTCI      |
| • <b>Incoming short/longwave radiation</b> |  |      |      |           |
| <b>Building Height and Footprint</b>       | OS Collection DIGIMAP (Digimap Edina, 2020)      | 2020 | GDB  | DSMb      |
| <b>Boundary and Network</b>                | OS Collection DIGIMAP (Digimap Edina, 2022a)     | 2022 | SHP  | Boundary  |
| <b>Air Pollution</b>                       | Scottish Air Quality (Scottish Government, 2019) | 2019 | XLSX | NOx, PM10 |

Two satellite images of Glasgow from the summers of 2018 and 2019 were taken into consideration. The dates were picked based on cloud coverage of less than 10%, with records of the highest temperature during the heatwave. *Table 5* lists the features of the Landsat images that were utilised to create the model.

*Table 5 Specification of Landsat 8 images used in this study*

| Date/Time of Acquisition | Sensor Identifier | Spatial Resolution | Path/Row | Number of Bands | Land Cover (%) | Cloud | Radiometric Resolution | UTM Zone/Datum |
|--------------------------|-------------------|--------------------|----------|-----------------|----------------|-------|------------------------|----------------|
| 25-Jun-2018 11:14:55     | OLI_TIRS          | 30 m               | 205/21   | 10              | 0.27           |       | 16-bit                 | 30/WGS84       |
| 28-Jun-2019 11:15:48     | OLI_TIRS          | 30 m               | 205/21   | 10              | 0.49           |       | 16-bit                 | 30/WGS84       |

#### 4.4. Data Pre-processing

The first step is to create a single coordinate system, the OSGB36 National Grid, also known as the National Grid coordinate system, to get reliable findings given the numerous data origins. Then after, the spatial factors were clipped to the area of interest (Glasgow City Boundary) and cleaned the data to deal with noise, missing data, and outliers. To combine the data with a different format, vector and raster-based data, the zonal statistics were used to compute and summarize the individual pixels of raster data (LST, DEM and other variables) into a grid of 30 m as a sense of division of urban areas to be considered as a block.

#### 4.5. Data Processing

The predictive model takes advantage of the correlation between one or more features and the response variable's value. By preserving the characteristics that deliver the best values, feature selection aims to improve the model. Additionally, the final model will be easier to understand and more computationally effective (*Figure 6*).

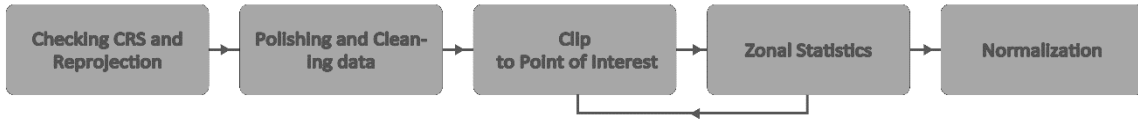


Figure 6 The workflow of pre-processing and processing steps

#### 4.5.1. Independent Variables (Predictors)

In the current study, urban characteristics were used in order to create a prediction model. These components were intended to play a leading role in outdoor thermal comfort, as seen in *Table 6*, was collected from open-source datasets.

Table 6 Classification by data source

| Categories based on source of dataset       | Predictors          |
|---|---------------------|
| Remotely sensed derived (RSD): Landsat 8    | NDVI, NDBI          |
| Urban morphology: LiDAR                     | DEM, DSM, CDSM, SVF |
| LULC  | RAR, VSF            |
| Vector-based at level of building footprint | BSF, Mean Height    |
| Air quality factors                         | PM10, NOx           |

##### 4.5.1.1 Remotely sensed derived (RSD): Landsat 8

The first category belongs to the derived variables from remotely sensed derived (RSD) such as LANDSAT 8 image:

- NDBI

NDBI is calculated using band 5 (NIR) and 6 (MIR), which indicate the reflectance of the near-infrared band and the reflectance of the short-wave infrared (Zheng, Tang and Wang, 2021). The NDBI values varied between -1 and 1. According to studies, positive NDBI values reflect urban built-up areas, whereas negative NDBI values imply non-urban land regions, including water bodies and non-urban land areas (Zha, Gao and Ni, 2003). *Figure 7* displays the spatial distribution of NDBI obtained from Landsat 8.

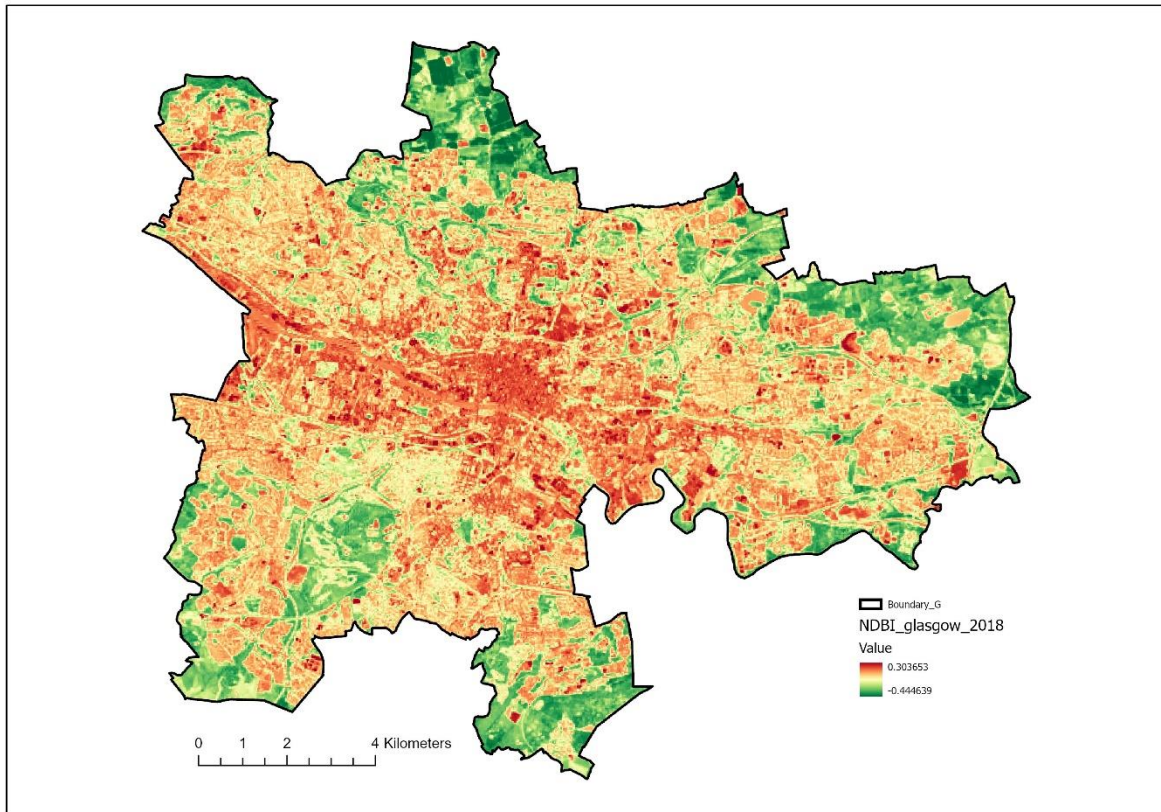
$$\text{Eq. (2) } NDBI = \frac{MIR - NIR}{MIR + NIR}$$

- NDVI

Tracking environmental changes has been becoming a global concern, which makes it possible to better understand the effect of humans' activities on the environment. The Normalized Difference Vegetation Index (NDVI) is considered a useful tool to monitor the ecological changes and effects of climatic disasters. One of its critical points is converting vegetation directly into the local climate and microclimate condition (Pettorelli *et al.*, 2005). In the Landsat-8 satellite image, the NDVI is calculated using band 5 (NIR) and 4 (RED), which

indicate the reflectance of the near-infrared band and the reflectance of the red band. (Portelaa *et al.*, 2020; Taloor, Manhas and Kothiyari, 2021). *Figure 8* displays the spatial distribution of NDVI obtained from Landsat 8.

$$\text{Eq. (3) } NDVI = \frac{NIR - RED}{NIR + RED}$$



*Figure 7 Spatial distribution of NDBI retrieved from Landsat 8 (2018)*

#### 4.5.1.2. Urban 3D data: LiDAR

- DEM (Digital Elevation Model)

From the LiDAR point clouds, a Digital Elevation Model (DEM) was produced (*Figure 9*).

- DSM (Building and ground model)

The DSM generating tool in the UMEP plugin (QGIS) was used to construct the digital surface model of the building, which combines the DEM model with the raster surface of the 3D building elements (*Figure 10*).

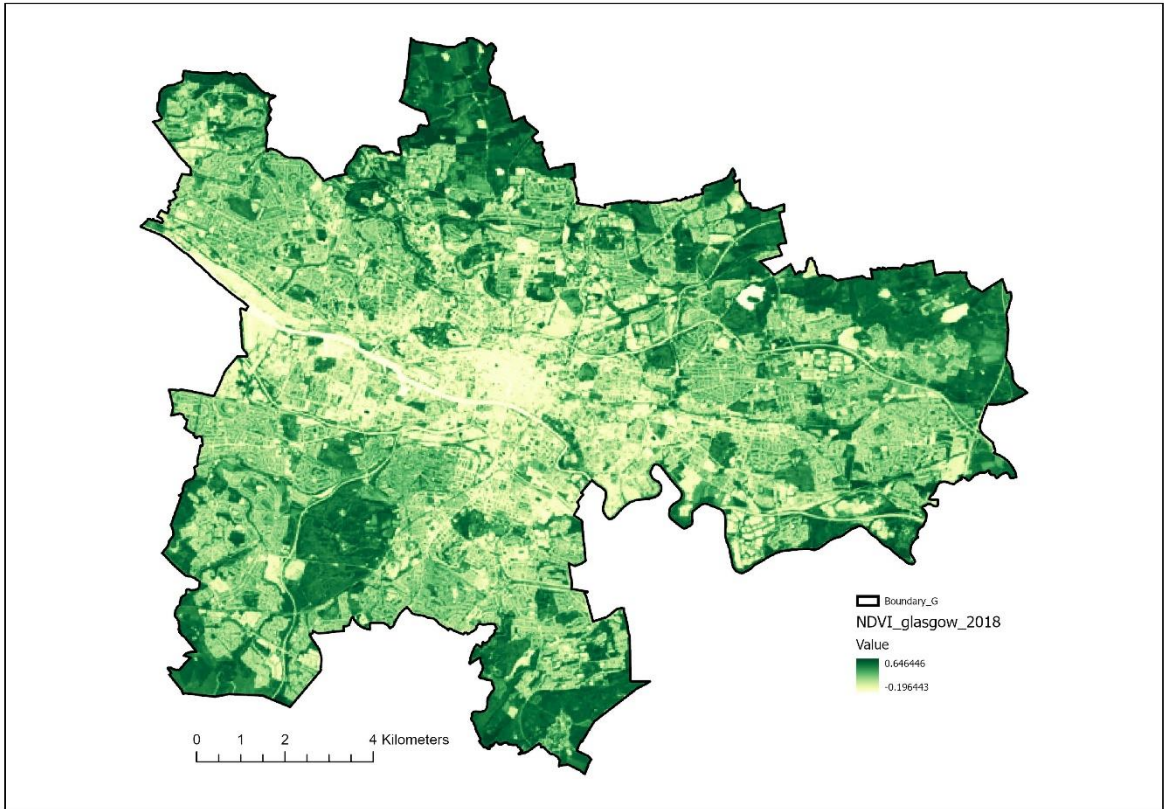


Figure 8 Spatial distribution of NDVI retrieved from Landsat 8 (2018)

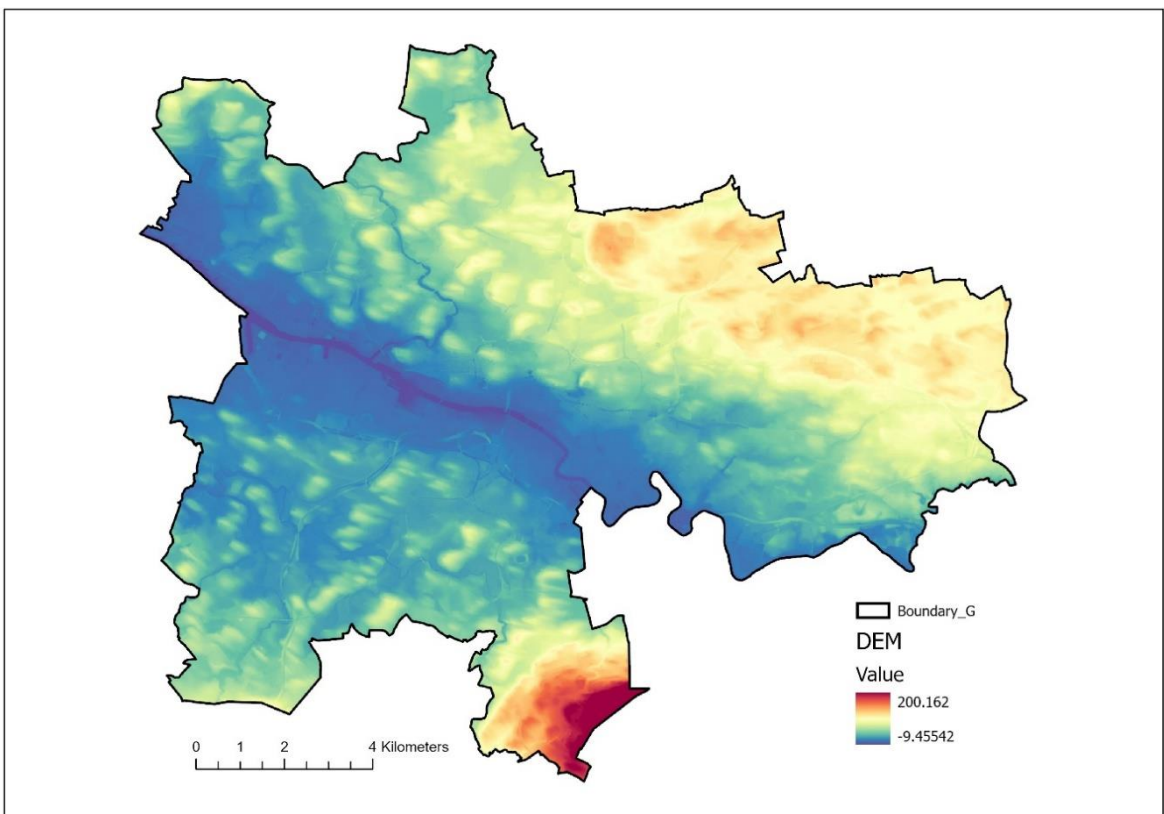


Figure 9 Spatial distribution of DEM

- CDSM

Canopy DSM is the vegetation surface layer of vegetation identified by Lindberg *et al.* (2022) for microclimate simulation. Due to the low-resolution LiDAR dataset (4 ppm), vegetation classification was done through the classification of ground, building, and water bodies, in which an unassigned class was considered as vegetation coverage (*Figure 11*).

- Sky View Factor

Sky view factor (SVF), a parameter in the analysis of the urban micro-climate, energy balance, local air circulation, and outdoor thermal comfort, quantify the characteristics of surface structures with the ratio of the radiation received by a planar surface (Souza, Rodrigues and Mendes, 2003; He *et al.*, 2015). Scholars have offered a variety of approaches for calculating SVF, including fisheye lens photography, computation of continuous SVF using three types of data: photographs; raster format (DSM and DEM); as well as vector format (LiDAR point clouds: building and tree data) (Heo *et al.*, 2021). In the current study, SVF assessment was performed using continuous SVF methods introduced by Lindberg and Grimmond (2010). The given method uses a shadow-casting algorithm for computation from high-resolution urban DEM, DSM, and CDSM. The SVF values varied between -1 and 1 (*Figure 12*).

#### 4.5.1.3. Land Use and Land Cover

The third category is prepared from the available land use and land cover dataset to demonstrate the landscape metrics along with the other urban morphology factors.

The Urban Atlas prepared the map of LU/LC for Glasgow city employing supervised machine learning methods based on the Copernicus Sentinel-2 time series (2012-2018). Under this classification, there are 22 categories were considered for mapping the LULC (*Figure 13*).

- Road Area Ratio (RAR)/ Impervious Surfaces Fraction

RAR covers the location of impervious surfaces that are responsible for an area's surface heat (*Figure 14*). The overall surface is calculated by counting the fast transit roads, railways, and other roads with their associated lands (Mehrotra, Bardhan and Ramamritham, 2020).

- Vegetation Surface Fraction (VSF)

The percentage of vegetation surface (i.e., bare soil, vegetation, farmlands) to total plan area is referred to as vegetation as part of the pervious surface fraction mentioned by Chen, Zheng and Hu (2020). The LULC classifications for this parameter consist of arable lands, forests, green urban areas, herbaceous vegetation associations, and pastures from the LULC map of Urban Atlas used to retrieve the surface fraction (*Figure 15*).



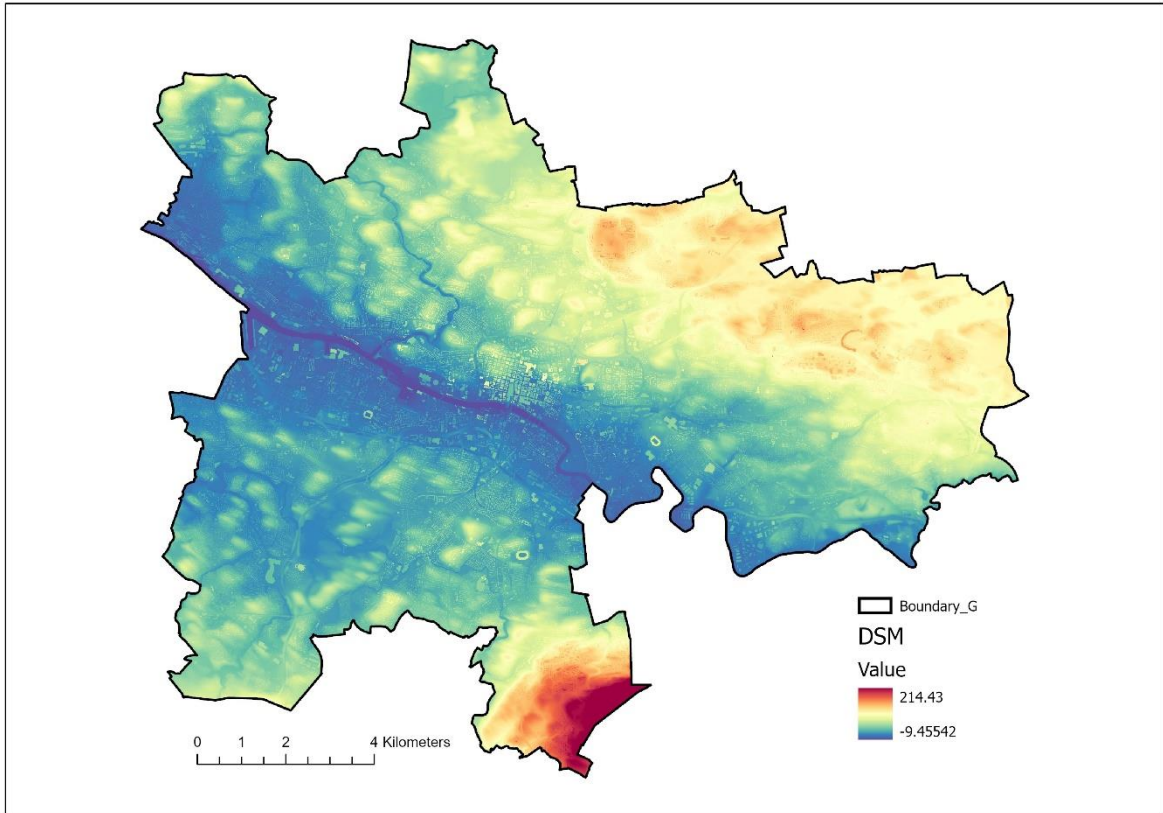


Figure 10 Spatial distribution of DSM

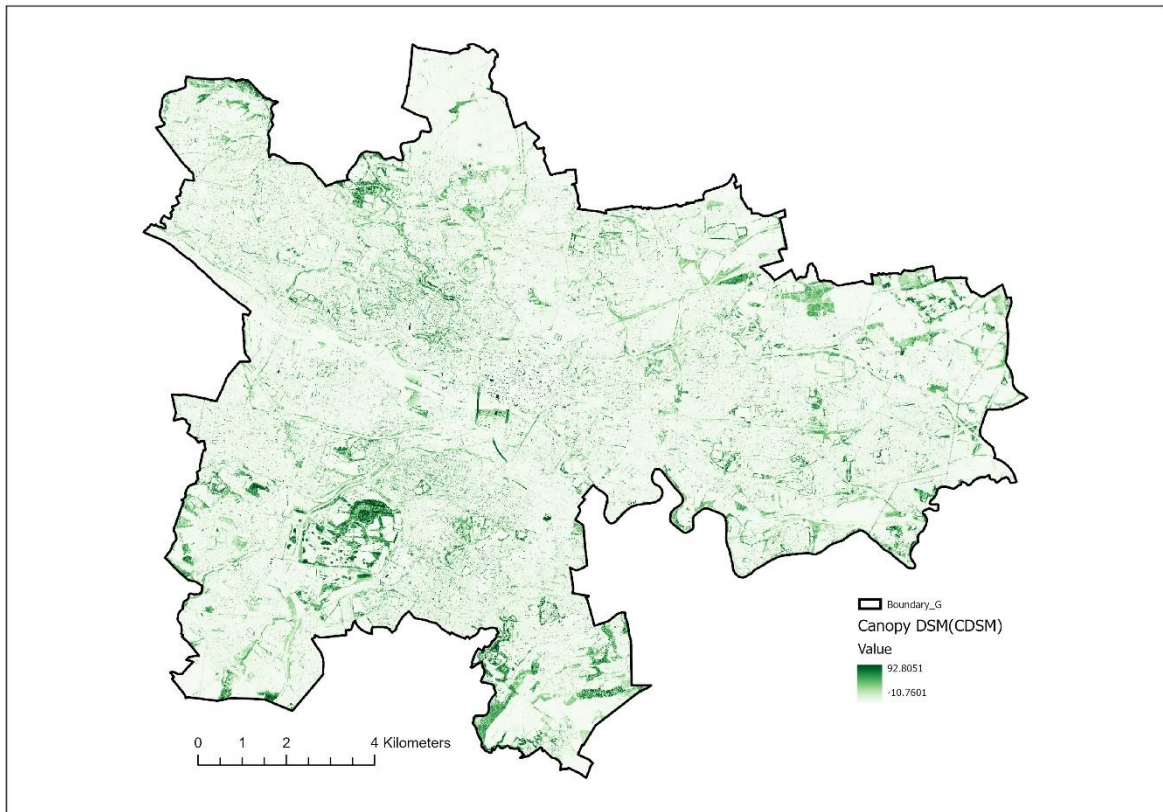


Figure 11 Spatial distribution of CDSM



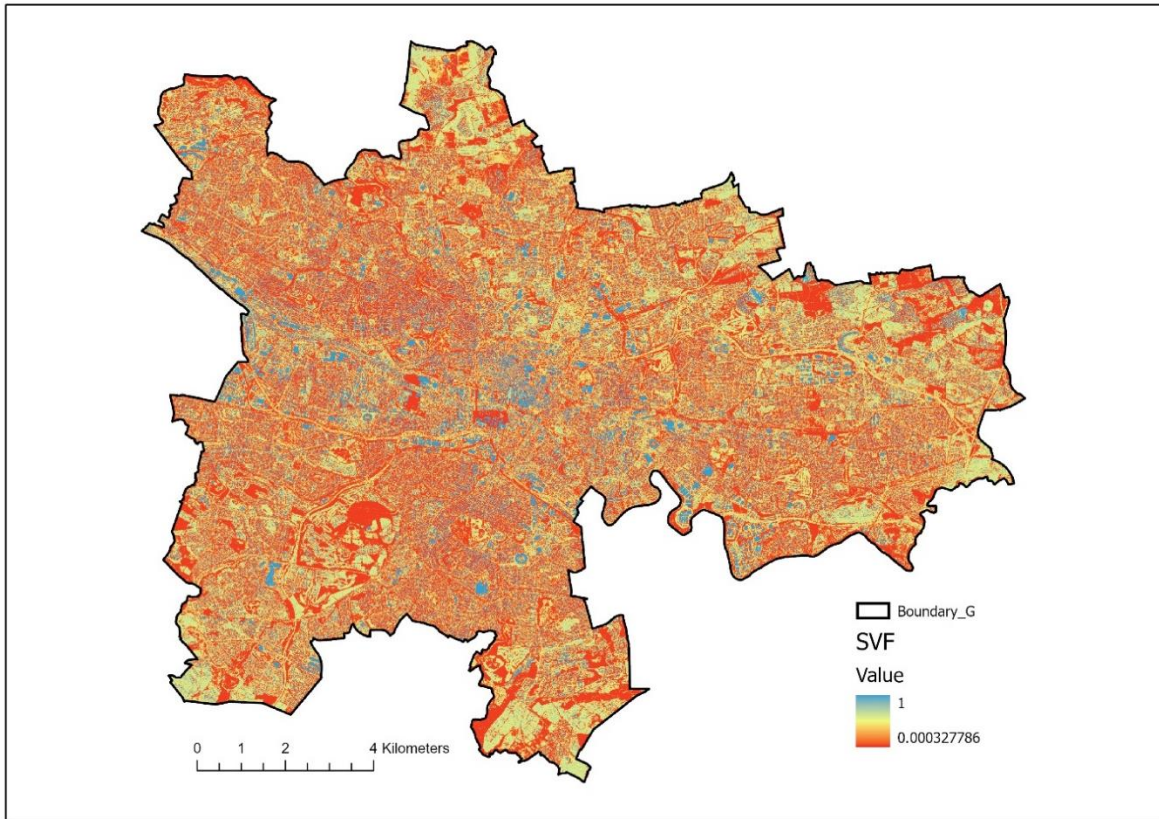


Figure 12 Spatial distribution of SVF

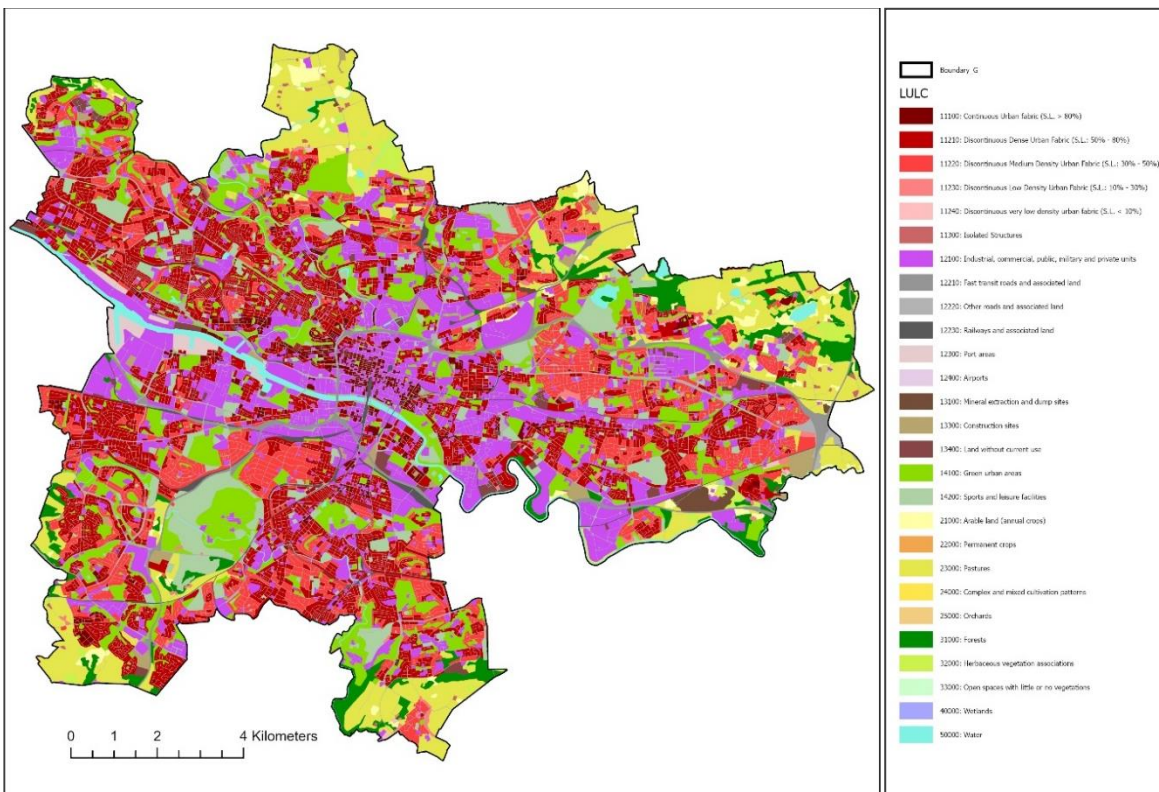


Figure 13 Spatial distribution of LULC



Figure 14 Spatial distribution of Road Area Ratio

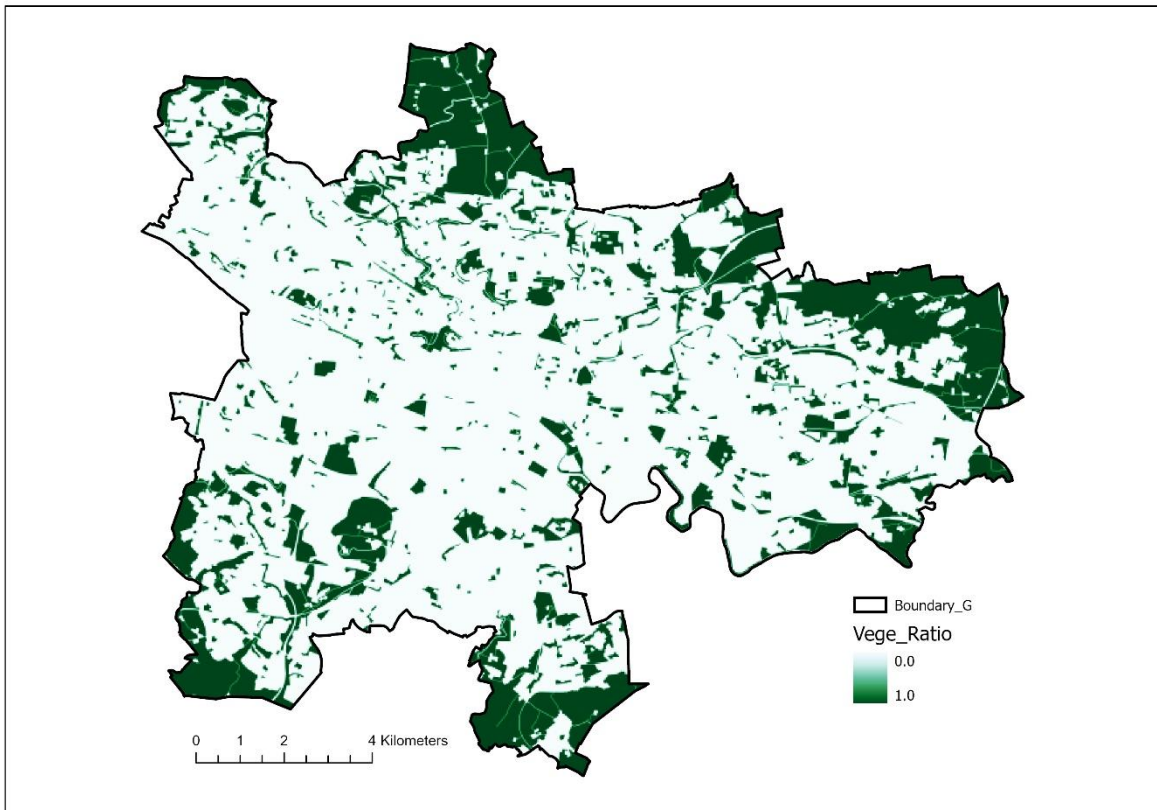


Figure 15 Spatial distribution of Vegetation Surface Fraction

#### 4.5.1.4. Vector-based variables at the level of building footprint

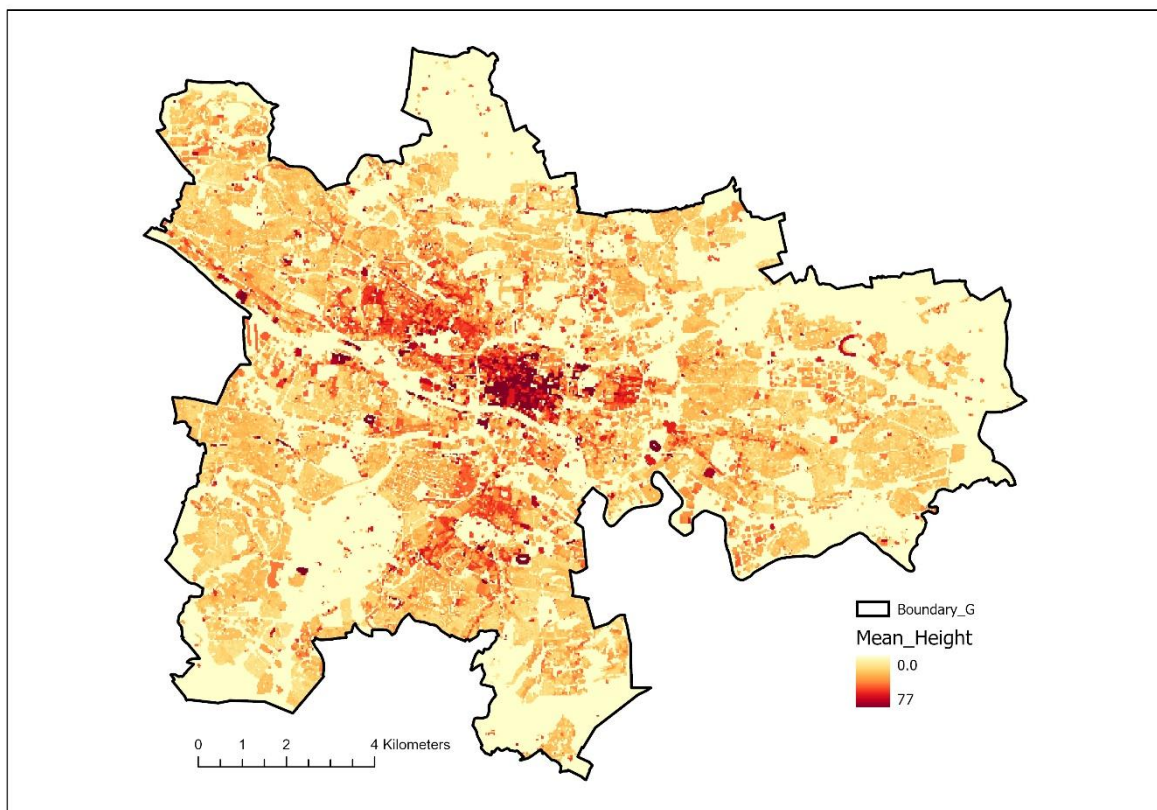
The fourth category illustrates the 2D and 3D aspects of urban settings, which are derived from the vector-based data of Glasgow.

- Building Height

The building height was derived from DIGIMAP-Verisk upon the availability of vector-based data. The highest building stands 77 metres tall (*Figure 16*).

- Built Surface Fraction (BSF)

BSF is a frequently used measure of building morphology for determining the thermal pattern at the level of the urban canopies shown in *Figure 17* (Mehrotra, Bardhan and Ramamritham, 2020).



*Figure 16 Spatial distribution of Building Height*





Figure 17 Spatial distribution of Building Surface Fraction

#### 4.5.1.5. Air quality factors

Last but not least category is created using historical air quality data for the entire of Scotland at a spatial resolution of 1 km. Ricardo Energy & Environment and the Scottish Air Quality Database (SAQD) project deliver plotted pollution levels for Scotland. The most recent maps, which show the spatial patterns of ambient and roadside annual mean concentrations, are PM10, N2O and NOx. The availability of the Glasgow PM2.5 map is currently under assessment as additional PM2.5 observation stations are added to the SAQD. In Scotland, roadway emissions are recorded for urban main road connections, and ambient quantities are mapped at a spatial resolution of one square kilometre (Scottish Government, 2021b).

The annual data of 2019 has been considered in this study due to the highest emission rates compared to the most recent years (impact of COVID and low concentration of pollution particles). The geo-statistical interpolation technique known as the Empirical Bayesian Kriging (EBK) was applied to accurately predict non-stationary data with a resolution of 1 km (Figure 18). The logic behind the semivariogram parameters of power in EBK was to use the restricted

maximum likelihood (REML) as a secure option balancing the performance and accuracy of the interpolation process (ESRI, 2022b).

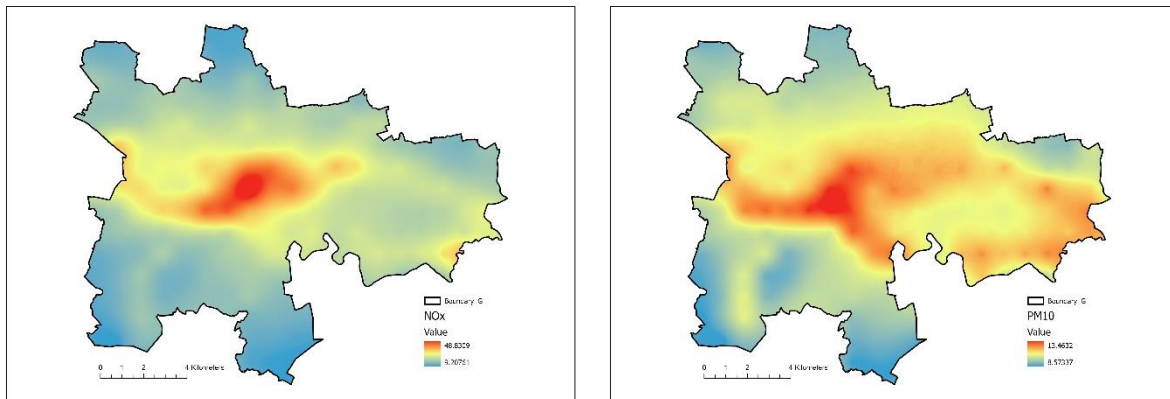


Figure 18 Left: NOx Right: PM10

To prevent any data loss during the data resampling procedure, all the data mentioned above were generated at the highest resolution of 2 meters, excluding Landsat-based variables of NDVI, NDBI and LST. The mean statistics for higher resolution data were used to compute and summarise the individual pixels of raster data into a grid of 30 m.

#### 4.5.1.6. Normalized Distribution of Predictors

Due to the varying numeric ranges of the variables with different units and magnitudes, biases are avoided by normalising all the variables using the scaling technique of normalisation (Eq. 4). Figure 19 shows the distribution of predictors.

$$\text{Eq. (4)} \quad X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

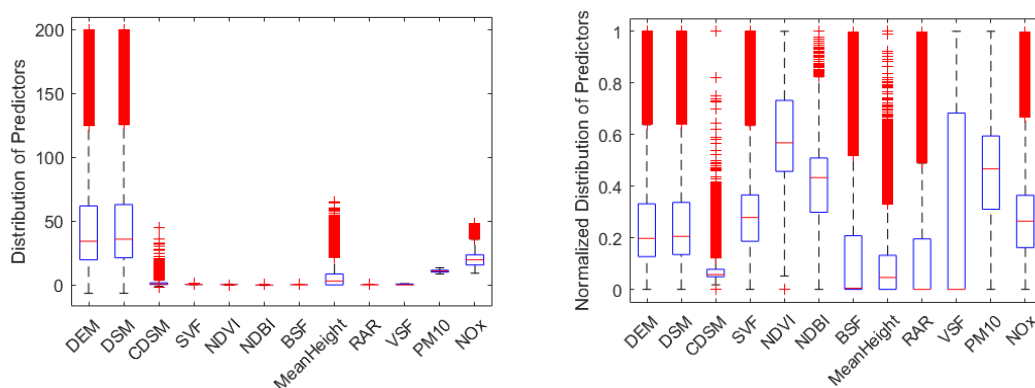


Figure 19 Distribution of Predictors (Left: before normalisation; Right: after normalisation)

## 4.5.2. Dependent Variables (Response Variables)

### 4.5.2.1. Land Surface Temperature Retrieval

The LST was calculated using Landsat-8, band-10 (Thermal infrared) (Portelaa *et al.*, 2020; Taloor, Manhas and Kothyari, 2021). The LST retrieval was run on 25<sup>th</sup> June 2018 during the heatwave 2018. The highest temperature reached 31.9 Celsius on 28<sup>th</sup> June, and due to a 16-day temporal resolution of Landsat 8, the closest day to the peak heatwave day was considered with the cloud coverage of below 10 per cent. LST was calculated for June 2019 (Appendix I), but the highest LST range was seen in 2018.

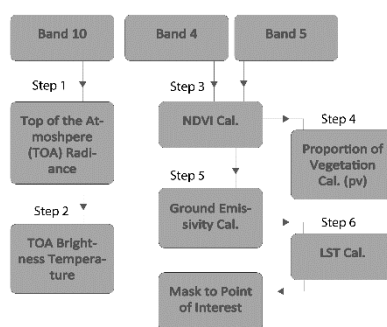


Figure 20 Procedure of LST retrieval

There are 6 steps (shown in Figure 20) to retrieve the LST out of Landsat 8 (Made 4 Geek, 2018).

#### Step 1 conversion to Top of Atmospheric (TOA) spectral radiance

The following equation will be used to compute the sensor radiance for the thermal infrared band (10) of Landsat 8, where  $L_{\lambda}$  = TOA spectral radiance,  $ML$  = Band-specific multiplicative rescaling factor extracted from the metadata = 0.0003342,  $AL$  = Band-specific additive rescaling factor extracted from the metadata = 0.10000, and  $Q_{cal}$  = Quantized and calibrated standard product pixel values = band 10 image

$$\text{Eq. (5)} \quad L_{\lambda} = ML * Q_{cal} + AL$$

**Step 2 Conversion of Radiance to at-Sensor Temperature** (in °C):  $BT$  = for deriving the temperature values the following equation will be considered, where  $L_{\lambda}$  = TOA spectral radiance,  $K1$  =  $K1$  constant band (No.) = 774.8853 and  $K2$  =  $K2$  constant band (No.) = 1321.0789.

$$\text{Eq. (6)} \quad TB = (K2 / \ln ((K1 / L_{\lambda}) + 1)) - 273.15$$

**Step 3 NDVI Extraction:** Considering the band 5 and 4 which are indicating the reflectance of the near infrared band, the reflectance of the red band, the NDVI will be calculated from the following equation (Portelaa *et al.*, 2020; Taloor, Manhas and Kothyari, 2021).

$$\text{Eq. (7) } NDVI = \frac{NIR - RED}{NIR + RED}$$

**Step 4 Calculation of Vegetation Fraction ( $pv$ ):**

$$\text{Eq. (8) } \left[ \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \right]^2$$

**Step 5 Ground emissivity**

$$\text{Eq. (9) } \epsilon_{OLI} = 0.004 \times pv + 0.986481$$

**Step 6 LST Computation (Figure 21)**

Where  $C2 = h \cdot c / s = 14388$ ,  $h = \text{planck's constant} = 6.626 \cdot 10^{-34}$ ,  $s = \text{Boltzmann constant} = 1.38 \cdot 10^{-23}$ , and  $c = \text{velocity of light} = 2.998 \cdot 10^8 \text{ m/s}$ .

$$\text{Eq. (10) } T = \frac{TB}{1 + \left( \frac{10.8 \cdot TB}{14388} \right) \ln(E)}$$

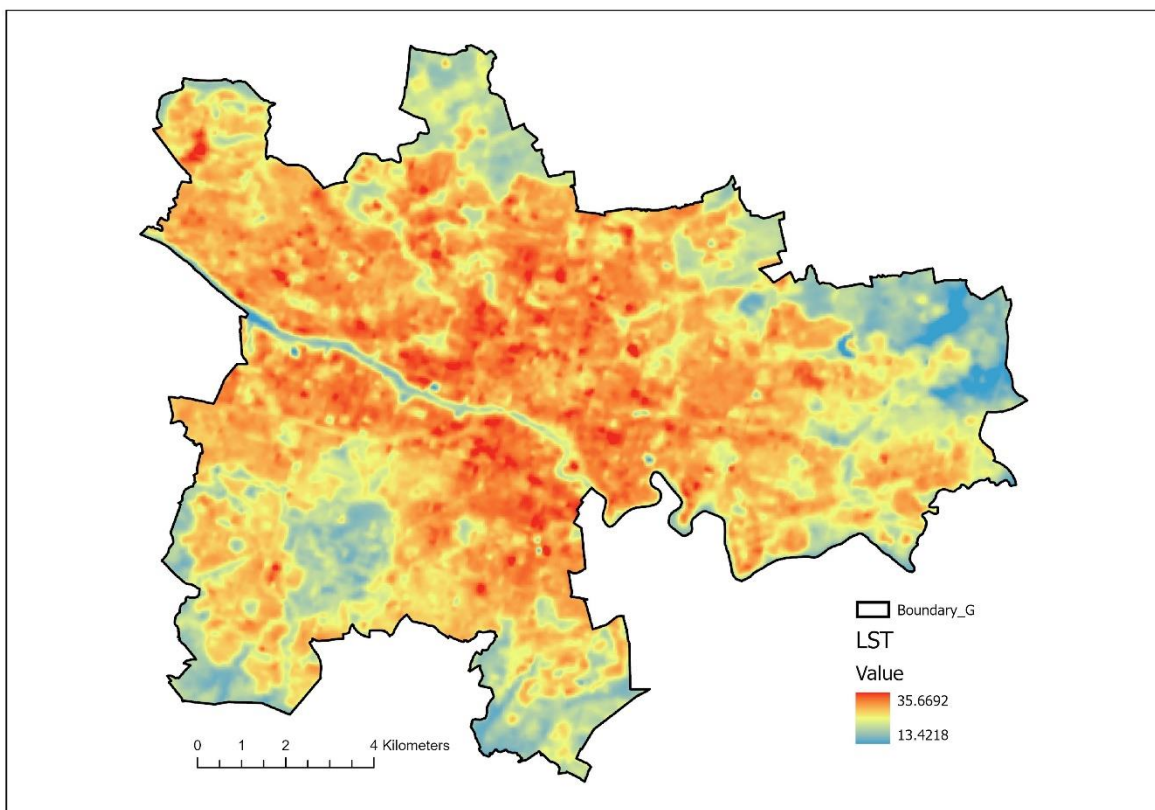


Figure 21 Spatial distribution of Land Surface Temperature retrieved from Landsat 2018

#### 4.5.2.2. Mean Radiant Temperature (MRT/ $T_{mrt}$ )

The SOLar and LongWave Environmental Irradiance Geometry (SOLWEIG) model calculates the  $T_{mrt}$  as one of the most important climatic factors affecting human energy balance and thermal comfort in this study. The computation is based on angular components such as shortwave and longwave radiation fluxes in six directions. The inputs of the model are geographic data, air temperature, relative humidity, urban geometry, and global shortwave radiation with the incorporation of a vegetation scheme (Figure 22) (Lindberg, Holmer and Thorsson, 2008; Lindberg *et al.*, 2022).

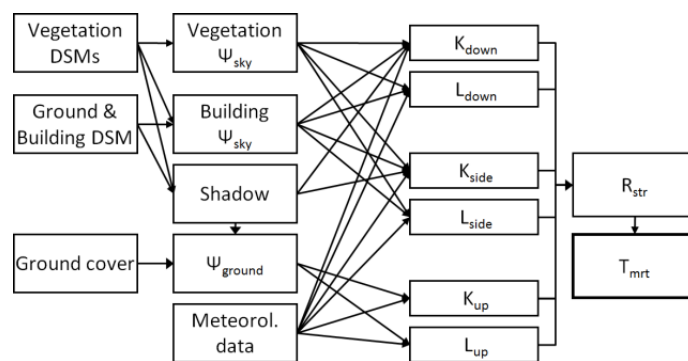


Figure 22 Flowchart of SOLWEIG model (Lindberg *et al.*, 2022)

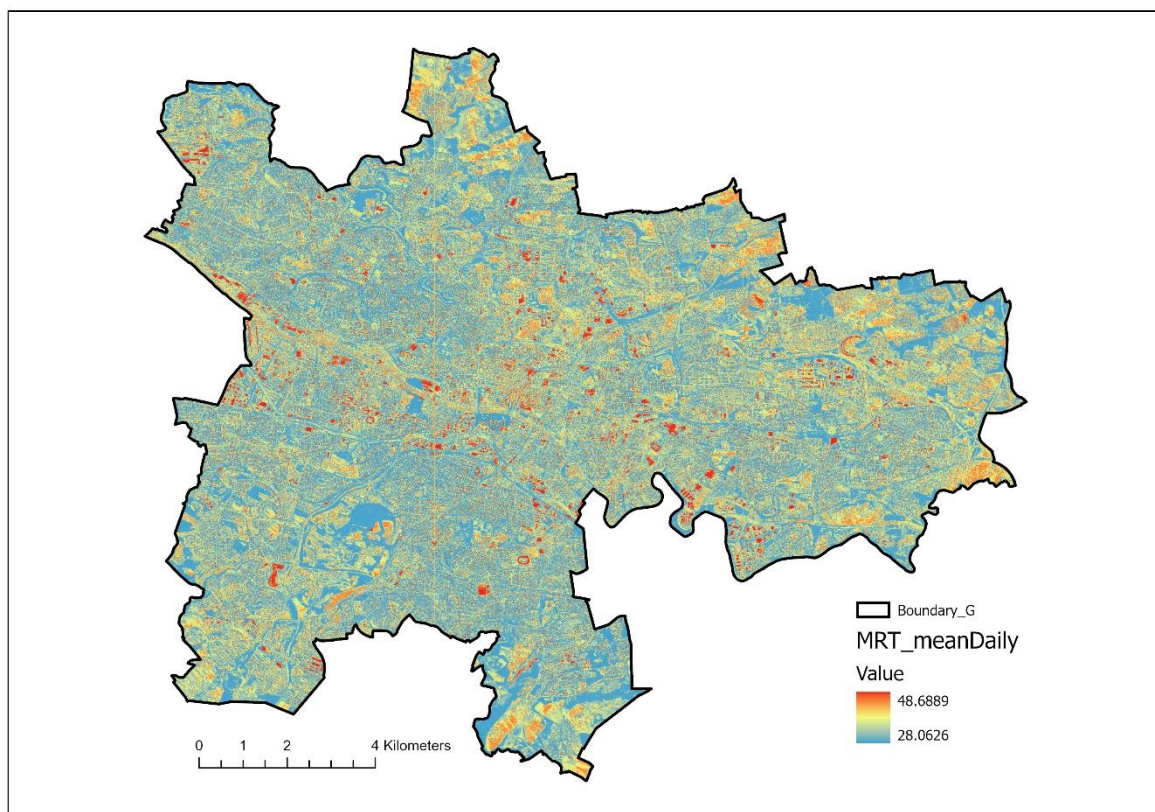


Figure 23 Spatial distribution of Mean Daytime MRT(5:00-21:00)



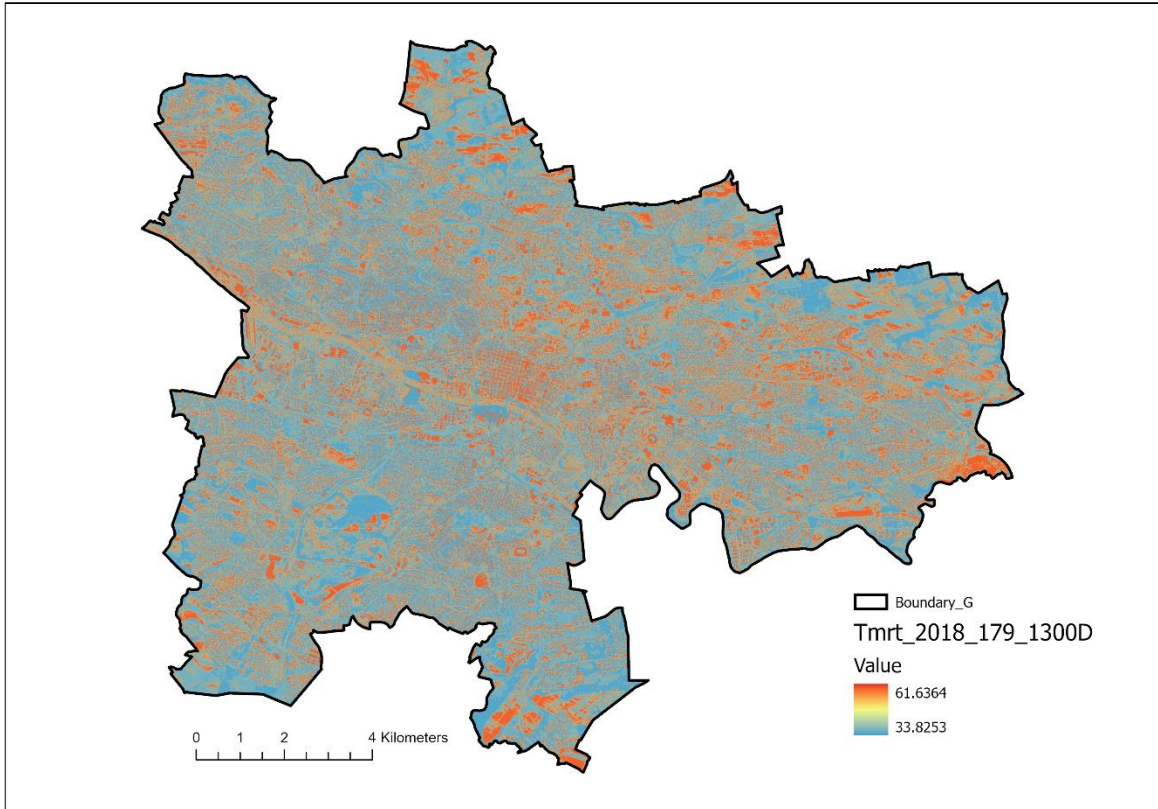


Figure 24 Spatial distribution of MRT at 13:00 (peak)

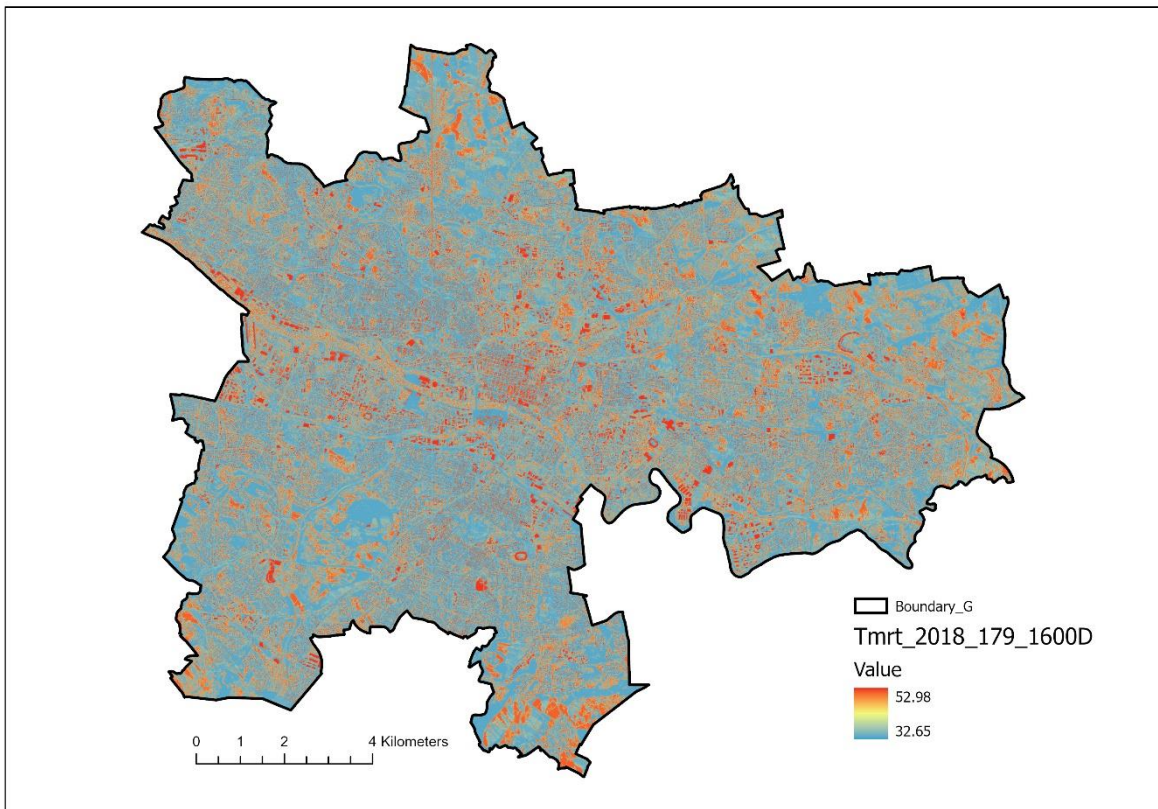


Figure 25 Spatial distribution of MRT at 16:00

As MRT is dependent on solar radiation (Lai, Maing and Ng, 2017; Thorsson *et al.*, 2017), it fluctuates during the day, and the mean daily map (from 5:00 AM to 21:00) shown in *Figure 23*, 13:00 (*Figure 24*), and 16:00 (*Figure 25*) for the 28<sup>th</sup> June have been considered. The 16:00 MRT has been considered in the following processing analysis to avoid the dramatic difference between RS-LST collection time and the clustering effect of the hot spot and cold spot effect (13:00 was the peak temperature).

#### 4.5.3. Standardisation of Dependent Variable

The *z\_score* function in MATLAB transforms the independent variables, which returns new values for each observation across all variables to have a standard deviation of 1 and are centred on having a mean of 0. To standardise the measurement scales of the network input, the following formula is used where *x* is the observed dependent variable,  $\mu$  is the mean, and  $\sigma$  is the standard deviation (Olden and Jackson, 2002).

$$\text{Eq. (11) } z = (x - \mu) / \sigma$$

### 4.6. Microclimate Modelling

SOLWEIG and ENVI-met microclimate models were utilised to analyse the impacts of changes in urban design factors on human thermal comfort. Both models simulate the microclimate in a specific period by integrating urban design factors with thermodynamics elements, including wind direction and velocity, air temperature, relative humidity, and solar radiation. In this study, UTCI and MRT were calculated via BioMet (ENVI-met tool) and SOLWEIG (UMEP processing tool), respectively. In this study, the UTCI assessment scale (*Table 7*) in terms of thermal stress (TS) by Błazejczyk *et al.* (2013) was followed.

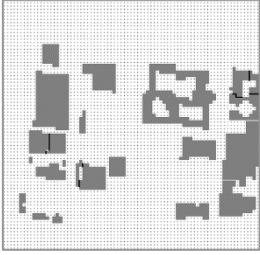
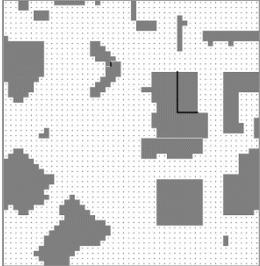
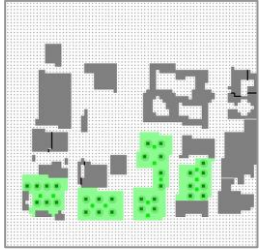
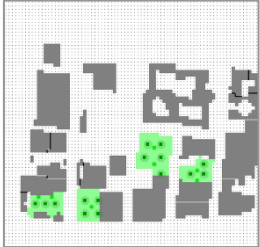
*Table 7 Thermal Stress Categories (Błazejczyk et al., 2013)*

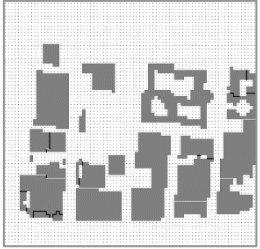
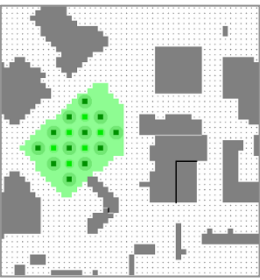
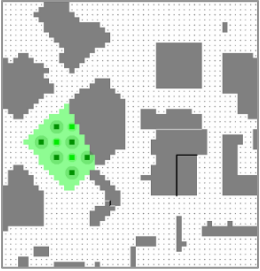
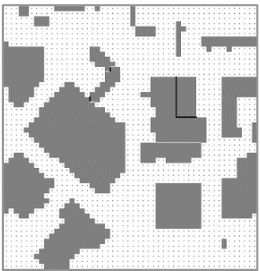
| <b>Heat Stress</b>             | <b>UTCI</b> |
|--------------------------------|-------------|
| <b>Extreme heat stress</b>     | Above +46   |
| <b>Very strong heat stress</b> | +38 to +46  |
| <b>Strong heat stress</b>      | +32 to +38  |
| <b>Moderate heat stress</b>    | +26 to +32  |
| <b>No thermal stress</b>       | +9 to +26   |
| <b>Slight cold stress</b>      | 0 to +9     |
| <b>Moderate cold stress</b>    | -13 to 0    |
| <b>Strong cold stress</b>      | -27 to -13  |
| <b>Very strong cold stress</b> | -40 to -27  |
| <b>Extreme cold stress</b>     | Below -40   |

#### 4.6.1. ENVI-met

Table 8 presents the details on ENVI-met models and the inputs for different scenarios.

Table 8 ENVI-met models' settings

| Parameters          |   | Model layout   |   |
|---------------------|---|--|---|
| <b>Main model</b>   |   |  |   |
| Grid Size           | 5m x 5m x 3m  | Domain_Site 1: 77x75x23<br> |   |
| Date                | 28 <sup>th</sup> June 2018                                  |  |   |
| Simulation period   | 30 h  |  |   |
| Model starting time | 27 <sup>th</sup> June 2018<br>18:00                         |  |   |
| Model level         | Simple forcing (Appendix II)                                |  |   |
| Model rotation      | 0   |  |   |
| Wind speed          | 9 m/s   |  |   |
| Wind direction      | 141.75  |  |   |
| Roughness length    | 0.01  |  |   |
| Min-Max Temp        | 13-31 C   |  |   |
| Min-Max RH          | 29%-94%   |  |   |
| Soil type           | Asphalt road<br>Sandy Loam<br>Loamy soil                    |  | Domain_Site_2: 50x52x14<br> |
| Material            | Building: default<br>Pavement: asphalt                      |  |   |
| Vegetation          | Trees<br>Hanging Birch(middle):15 m<br>Acer Campestre: 12 m |  |   |
|                     | Grass 25 cm   |  |   |
|                     |   |  |   |
| <b>Site 1</b>       |   |  |   |
| Nesting grid        | 3   |  |   |
| Best Case           | 30 Birch tree<br>12 Acer tree                               |                           |   |
| Intermediate Case   | 15 Birch tree<br>7 Acer tree                                |                           |   |

|               |                   |                               |   |
|---------------|-------------------|-------------------------------|---|
|               | Worst Case        | No Birch tree<br>No Acer tree |    |
| <b>Site 2</b> | Nesting grid      | 15                            |   |
|               | Best Case         | 12 Birch tree<br>6 Acer tree  |    |
|               | Intermediate Case | 6 Birch tree<br>3 Acer tree   |   |
|               | Worst Case        | No Birch tree<br>No Acer tree |  |

For intermediate scenarios, 50% of the land was randomly assigned to a vegetated region by dividing along an axis perpendicular to the nearby main street. In order to avoid the effects of attributes that were not included in the chosen predictors, tree orientation was also done randomly for all green cases.

#### 4.6.2. SOLWEIG

*Table 9* shows the details of the SOLWEIG model and the inputs for retrieving the mean radiant temperature at the scale of Glasgow city. The meteorological data, which was

considered for simulating MRT from ERA5 reanalysed climate data, was summarised in Appendix III.

*Table 9 SOLWEIG settings*

|                                 |                                   |                                     |
|---------------------------------|-----------------------------------|-------------------------------------|
| <b>Main model parameters</b>    | Grid Size                         | 9 grids of 3501*3000 m <sup>2</sup> |
|                                 | Date                              | 28 <sup>th</sup> June 2018          |
|                                 | UTC offset                        | 0                                   |
|                                 | Simulation period                 | 24 h                                |
|                                 | Model starting time               | 28 <sup>th</sup> June 2018 00:00    |
| <b>Environmental parameters</b> | Albedo (walls)                    | 0.20                                |
|                                 | Albedo (ground)                   | 0.15                                |
|                                 | Emissivity (wall)                 | 0.90                                |
|                                 | Emissivity (ground)               | 0.95                                |
| <b>Tmrt parameters</b>          | Absorption of shortwave radiation | 0.70                                |
|                                 | Absorption of longwave radiation  | 0.95                                |
|                                 | Body posture                      | standing                            |

One base case model was run with SOLWEIG calculations on ENVI-met to validate the MRT calculation since the meteorological data for ENVI-met and SOLWEIG models were derived from two different sources. The comparison of the two base cases is available in Appendix IV.

## 4.7. Regression-based Prediction Models

To reach the goal of the project to develop the prediction model, the following non-linear and linear statistical models were considered to find the relation between the predictors and the response variables, MRT and LST, which have been taken into account as proxies of outdoor thermal comfort.

### 4.7.1. Pearson Correlation Coefficient

The scale-less covariance, which stands for Pearson's correlation coefficient, is a metric used to determine the linear dependence of the chosen variables. The correlation coefficient, which assesses the strength and direction of the association between two continuous variables, ranges in value from -1 to +1. The positive number demonstrates that as one variable increases, the other rises. Regardless of the correlation's power, it cannot establish causality (Brown and Hambley, 2002; Frost, 2019).

$$\text{Eq. (12) } r = \frac{\sum (X_i - X_{mean})(Y_i - Y_{mean})}{\sqrt{\sum (X_i - X_{mean})^2 \sum (Y_i - Y_{mean})^2}}$$

## 4.7.2. Linear-based Regression Model

### 4.7.2.1. *Generalized Linear Regression and Exploratory Linear Regression*

The GeoAnalytics toolset of ArcGIS Pro- generalized linear regression (GLR) and Exploratory Regression were used to make predictions of dependent variables depending on their connection to a collection of explanatory factors (ESRI, 2022a). Based on the nature of input data, the Ordinary Least Squares (OLS) approach for GLR delivers appropriate regression analysis findings.

### 4.7.2.2. *Spatial Autocorrelation*

Spatial autocorrelation develops as geography is fundamental (Griffith and Chun, 2018). Spatial autocorrelation statistics, which quantify the similarity of neighbouring observations, are frequently used to measure spatial linkage or spatial dependency. The positive spatial autocorrelation happens when the significant value of an attribute seems to be clustered together in some portions of the study and the low value of other attributes are clustered together in other parts. On the contrary, the proximity of low and high values in clusters displays a negative autocorrelation. Furthermore, if no relationship is discovered between the high and low attributes, zero spatial autocorrelation will result. The logic behind spatial autocorrelation statistics can be summarised in the concept of proximity and spatial weighting (Fotheringham, Brunsdon and Charlton, 2002).

### 4.7.2.3. *Geographically Weighted Regression (GWR)*

The GWR, as a spatial regression method, emphasises non-stationary data. One of the main facts in GWR studies is the determination of zones during the regression analysis. It is crucial to finely divide the area where the relationships between the explanatory and dependent variables will be analysed. Most spatial processes are continuous and do not have distinct and frequently arbitrary borders. These zones may be those where data is collected, independently chosen locations within the area used as a foundation for additional mapping, or even the combination of both methods (Fotheringham, Brunsdon and Charlton, 2002). The spatial distribution of the local coefficient of determination that follows GWR analysis tends to elicit extensive discussions of the patterns identified (Páez, Farber and Wheeler, 2011). The lower AICc value provides a better fit to the observed data, while the higher value for R-Squared is a measure of goodness of fit (ESRI, 2018).

## 4.8. Nonlinear Regression Model: Artificial Neural Network (ANN)

A collection of interconnected units or nodes resembling a biological brain's neurons. Each processor is in charge of processing data, and all parts cooperate to create a network and carry out simultaneous interactions (Yegnanarayana, 2005). Each network will calculate the required output as a result of the interaction between the input values (Lee, Kim and Yun, 2016). To avoid overfitting, the model's complexity should be determined, the validation dataset should be considered, and the model should be regularized. In this regard, the data is divided into three main categories: training, validation, and test data. The training dataset helps build the model, whereas the validation data guides the model selection, parameter optimisation and any iterative pre-processing steps. Test data finally evaluates how the model performs based on unseen data for the first time. Splitting a subset of data for final testing as well as using validation to prevent a selected model that over-fits or under-fits. Due to the large dataset, the validation method chosen for the fitting model is hold-out by considering 15 per cent of the training data, which is considered 80 per cent of the whole data. The remaining data belongs to the test data test (*Table 10*). In order words, the training, validation, and test dataset consist of 70%, 15% and 15% of all data, respectively. Although cross-validation would perform better by dividing each set of learning for training and validation, the processing time would be time-consuming (Josse and Husson, 2012). For the given reason, it was not considered in this study. The method used for dividing datasets was random.

*Table 10 Data Partitioning*

| Training (80%) |            |      | Test (20%) |
|----------------|------------|------|------------|
| Train          | Validation | Test |            |
| 70%            | 15%        | 15%  |            |

### 4.8.1. The Architecture of the ANN network

The input layer (corresponding variables or predictors), output layer (independent or response variable), and a hidden layer or layers make up the ANN's architecture. The LST value, as a proxy for outdoor thermal comfort in the current study, along with the MRT, is the output layers in two separate networks. The ANN models were optimised to yield the fewest errors between anticipated and actual values. This was accomplished by adjusting the number of neurons within the hidden layer according to Eq. (13) (Chan and Chau, 2019). Where  $N_i$  is the number of input variables (12),  $N_o$  is the number of output(1), and;  $N_h$  shows the number

of neurons in the hidden layer. This way, the number of neurons in hidden layers should be between 6 and 25.

$$\text{Eq. (13)} \quad 2 \times \sqrt{N_i + N_o} \leq N_h \leq 2 \times N_i + 1$$

In machine learning the hyperparameter is critical in optimising the model's accuracy. Hyperparameters are employed to manage the learning process under the external configuration. Levenberg-Marquardt algorithm (trainlm in MATLAB) as the most effective method for training moderate-sized feedforward networks has been considered in this study (Liu, 2010). Appendix V summarises the training parameters and hyperparameters applied in ANN.

#### 4.8.2. Factor Characterisation and Sensitivity Analysis

ANN, as a black box predictive model, does not reveal the information about how the variables are connected to one another to arrive at a final prediction (Olden and Jackson, 2002). There are a few statistical techniques (sensitivity analysis, Neural Interpretation Diagram, Garson's algorithm) to uncover the significance of independent variables. In the current study, the model will be considered for analysing the effect of each variable (focusing mostly on the impact of greenery and built-up) by excluding one predictor in each round of training (sensitivity analysis) and comparing their performance through different error metrics.

### 4.9. Evaluation of Models' Performance

The regression-based prediction model quantifies the relationship between predictors and the dependent variable, and the following metrics (Naser and Alavi, 2021) are taken into account to determine how well a model fits in order to assess how effectively the model is adapting to new data. The objective is to identify the "best" model—one that can produce predictions that are most closely related to the true number values from the dataset—with the lowest RMSE value (Botchkarev, 2018).

**Error (Residual-R):** indicate the difference between the predicted and observed values.

$$\text{Eq. (14)} \quad R = (y_p - y_o)$$

**MSE (Mean Squared Error):** It determines the square root of the errors' mean.

$$\text{Eq. (15)} \quad MSE = \Sigma(y_p - y_o)^2 / n$$



**RMSE (Root Mean Squared Error):** It displays the errors' standard deviation, which illustrates how dispersed the residuals are.

$$\text{Eq. (16) } RMSE = \text{Sqrt} (\Sigma(y_p - y_o)^2 / n)$$

**R<sup>2</sup>:** R-squared is understood as the proportion of the response variable's variance that the model can faithfully predict. (Where  $y_o$  is observe;  $y_p$  is predicted and  $y_m$  is the mean value of observed data)

$$\text{Eq. (17) } R^2 = 1 - (\Sigma(y_p - y_o)^2 / \Sigma(y_p - y_m)^2)$$

**MAE (Mean Absolute Error):** It calculates the mean error in a range of forecasts without taking into account the direction (positive/negative) of the mistakes.

$$\text{Eq. (18) } MAE = \Sigma |y_p - y_o| / n$$

#### 4.10. Validation of Final Model: Scenarios for Vacant Lands

According to the current situation, there are around 5% of vacant lands in the city have the potential to be developed in the future. Through the use of a trained ANN model and ENVI-microclimate met's simulation, the current study aims to completely eliminate the possibility of climate-related risk to future development in such lands.

1. Best case scenario: the desirable vision based on increased greenery (100%)
2. Intermediate scenario: Combination of both greenery and built-up (50-50 %)
3. Worst case scenario: All the lands converted for the building construction

All the cells with vacant land surface fractions are selected to develop the scenarios. All the predictors, excluding pollution factors, DEM, CDSM, and DSM will take the new values for each scenario randomly by following similar trained cells.

For example, in the best scenario, each cell will take the new values for MeanHeight(0), BSF(0), VSF(1), and RAR(0), while in the case of NDVI(), NDBI(), and SVF(), the new values will be randomly chosen based on the same configuration of cells with MeanHeight (0), BSF(0), VSF(1), and RAR(0) values from training dataset.

In the intermediate scenario, BSF(.5) and VSF(.5) will be constant for the vacant lands, while MeanHeight(), RAR(), NDVI(), NDBI(), and SVF() will take values randomly considering the trained dataset.

In the worst scenario, the cell will take the new values for BSF(1), VSF(0), SVF(1) will be considered consistent and only MeanHeight(), RAR(), NDVI(), and NDBI() will be chosen randomly.



## CHAPTER 5: RESULTS AND ANALYSIS- PART I

Chapter 5 is divided into three sections. The relationship between the LST and its chosen predictors was first investigated. Then, in order to identify the specifically suited model, the regression prediction models were conducted. Regression-based approaches were compared, the most significant contributing factors were determined, and the LST best-fit prediction model was validated.

### 5.1. LST: A Proxy for Human Thermal Comfort?

#### 5.1.1. Correlation Matrix: LST

The Pearson correlation technique was used to calculate the correlation values in *Figure 26*. The coefficient correlation employed in multiple regression analysis reveals that the association between NDVI and NDBI, and LST are the strongest, while Road Ratio and SVF are the least correlated ones. Additionally, among all the predictors, NDVI, VSF, CDSM, DEM, and DSM showed a negative correlation (*Table 11*).

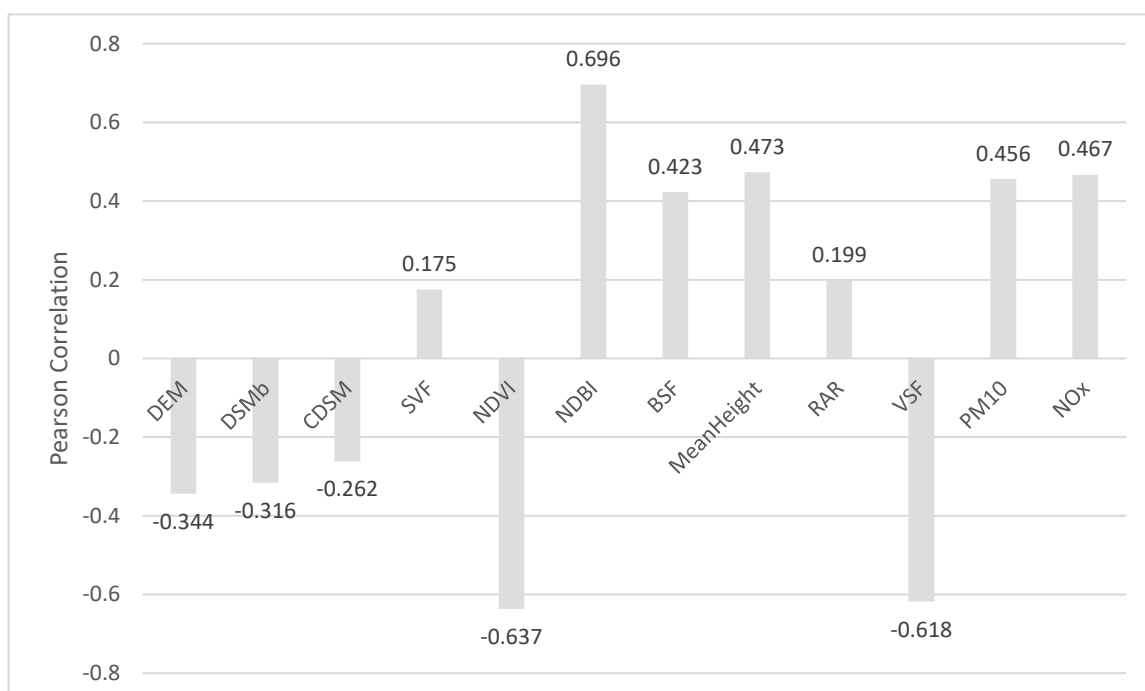


Figure 26 Pearson Correlation between Predictors and LST

Table 11 Pearson Correlation between Predictors and LST

| Pearson CC | DEM    | DSMb   | CDSM   | SVF   | NDVI   | NDBI  | BSF   | MeanHeight | RAR   | VSF    | PM10  | NOx   |
|------------|--------|--------|--------|-------|--------|-------|-------|------------|-------|--------|-------|-------|
| LST        | -0.344 | -0.316 | -0.262 | 0.175 | -0.637 | 0.696 | 0.423 | 0.473      | 0.199 | -0.618 | 0.456 | 0.467 |

### 5.1.2. Generalized Linear Regression and Exploratory Regression Results

The results from Generalized Linear Regression (GLR) model in *Table 12* express that all variables are statistically associated with LST in linear regression. The variance inflation (VIF) of each predictor ranges from 1.20 to 528.99, demonstrating the impact of multicollinearity and redundancy among independent variables. In this case, DEM and DSM are the independent variables that are highly collinear with the other variables in the model.

Table 12 Summary of GLR Result-LST

| Variable    | Coefficient [a] | StdError | t-Statistic | Probability [b] | Robust_SE | Robust_t   | Robust_Pr [b] | VIF [c]    |
|-------------|-----------------|----------|-------------|-----------------|-----------|------------|---------------|------------|
| Intercept   | 20.826022       | 0.076598 | 271.888518  | 0.000000*       | 0.080360  | 259.157987 | 0.000000*     | -----      |
| DEM         | 0.123169        | 0.002612 | 47.158723   | 0.000000*       | 0.003560  | 34.595821  | 0.000000*     | 528.990332 |
| DSMB        | -0.128765       | 0.002610 | -49.331317  | 0.000000*       | 0.003558  | -36.186083 | 0.000000*     | 515.631618 |
| CDSM        | -0.156234       | 0.002574 | -60.692667  | 0.000000*       | 0.002864  | -54.551464 | 0.000000*     | 2.119121   |
| SVF         | -1.592273       | 0.040079 | -39.728119  | 0.000000*       | 0.049559  | -32.128850 | 0.000000*     | 3.026691   |
| NDVI        | 4.375922        | 0.071235 | 61.429642   | 0.000000*       | 0.115758  | 37.802427  | 0.000000*     | 8.638775   |
| NDBI        | 15.395456       | 0.096010 | 160.352694  | 0.000000*       | 0.148403  | 103.740540 | 0.000000*     | 8.353480   |
| BSF         | 3.085725        | 0.039477 | 78.164350   | 0.000000*       | 0.049959  | 61.765531  | 0.000000*     | 4.628908   |
| MEAN_HEIGHT | 0.052510        | 0.000971 | 54.092938   | 0.000000*       | 0.001108  | 47.384361  | 0.000000*     | 3.001814   |
| ROAD_RATIO  | 0.365542        | 0.018946 | 19.294089   | 0.000000*       | 0.021105  | 17.320190  | 0.000000*     | 1.200761   |
| VEGE_RATIO  | -1.185666       | 0.012305 | -96.353949  | 0.000000*       | 0.013641  | -86.918504 | 0.000000*     | 2.243475   |
| PM10        | 0.670956        | 0.008947 | 74.994966   | 0.000000*       | 0.008868  | 75.662656  | 0.000000*     | 6.305819   |
| NOX         | -0.039077       | 0.001371 | -28.500809  | 0.000000*       | 0.001375  | -28.412640 | 0.000000*     | 6.964888   |

The Exploratory Regression Global Summary explains that 80.37% of the criteria calculated based on 12 predictors got the passing mark in the case of LST to be considered as a robust model for prediction with a minimum Adjusted R-squared of above .5 (*Table 13*). The Global Moran’s I test was the next step to quantify how values are similar to their neighbours.

Table 13 LST Criteria Passed

| Search Criterion                    | Percentage of Search Criteria Passed |          |        |          |
|-------------------------------------|--------------------------------------|----------|--------|----------|
|                                     | Cutoff                               | Trials # | Passed | % Passed |
| Min Adjusted R-Squared              | > 0.50                               | 4095     | 3291   | 80.37    |
| Max Coefficient p-value             | < 0.05                               | 4095     | 3858   | 94.21    |
| Max VIF Value                       | < 7.50                               | 4095     | 2304   | 56.26    |
| Min Jarque-Bera p-value             | > 0.10                               | 4095     | 0      | 0.00     |
| Min Spatial Autocorrelation p-value | > 0.10                               | 37       | 0      | 0.00     |

### 5.1.3. Summary of Variables Significance

The significance of exploratory variables in LST is summarised in *Table 14*. The strongest predictors are NDBI, MeanHeight, CDSM, BSF, and PM10 for their consistent significance and stable relationship with LST. However, all the recorded scores are above 93 per cent, and all the predictors are significant to be considered in the models. Among all significant exploratory variables in the LST set, the NDVI and NOx are the inconsistent ones for their both negative and positive relationship.

*Table 14 LST Variables Significance*

| Summary of Variable Significance |               |            |            |
|----------------------------------|---------------|------------|------------|
| Variable                         | % Significant | % Negative | % Positive |
| CDSM                             | 100.00        | 100.00     | 0.00       |
| NDVI                             | 100.00        | 50.00      | 50.00      |
| NDBI                             | 100.00        | 0.00       | 100.00     |
| BSF                              | 100.00        | 0.00       | 100.00     |
| MEAN_HEIGHT                      | 100.00        | 0.00       | 100.00     |
| VEGE_RATIO                       | 100.00        | 100.00     | 0.00       |
| PM10                             | 100.00        | 0.00       | 100.00     |
| DSMB                             | 99.85         | 87.40      | 12.60      |
| DEM                              | 99.66         | 62.79      | 37.21      |
| NOX                              | 99.61         | 48.68      | 51.32      |
| SVF                              | 95.36         | 68.60      | 31.40      |
| ROAD_RATIO                       | 93.41         | 15.72      | 84.28      |

### 5.1.4. Spatial Autocorrelation (Global Moran's I) Test

ArcGIS Pro's Global Moran's I is used to test the spatial autocorrelation theory. The outcome is displayed in *Table 15*. Given the z-score of 486.186243 and Moran's Index of .77, there is a probability of spatial autocorrelation and spatial clustering, in which geographically weighted regression would shed light on this.

*Table 15 Global Moran's I Summary*

|            | Moran's Index | Expected Index | Variance | z-score    | p-value  |
|------------|---------------|----------------|----------|------------|----------|
| <b>LST</b> | 0.775091      | -5E-06         | 0.000003 | 486.186243 | 0.000000 |

### 5.1.5. Geographically Weighted Regression (GWR)

Given the multicollinearity and redundancy between variables, utilising the GWR with all variables was not feasible for the regression-based prediction. Here are the individual GWR models and a few combined models to better understand the spatial non-stationarity in their relationship. The GWR technique was considered under the optimisation of Distance Band to produce a robust model.

### 5.2.5.1. GWR and DEM

The local  $R^2$  from considering the DEM variable ranged between 0 and .97, and the results explain 75-100% variation between DEM and LST for only 19.08 per cent of the whole data (Figure 27).

| Local R2         |       |
|------------------|-------|
| <b>0-0.25</b>    | 9.20  |
| <b>0.25-0.50</b> | 26.41 |
| <b>0.50-0.75</b> | 45.31 |
| <b>0.75-1.00</b> | 19.08 |

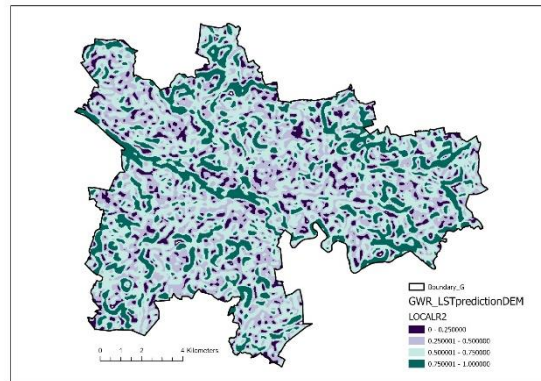


Figure 27 DEM and LST in GWR

### 5.2.5.2. GWR and DSM

The local  $R^2$  of DSM varied between 0 and .97, and the results depict that the GWR model could explain 75-100% of the variation between DSM and LST for only 17.83 per cent of the whole data (Figure 28).

| Local R2         | %     |
|------------------|-------|
| <b>0-0.25</b>    | 9.72  |
| <b>0.25-0.50</b> | 26.96 |
| <b>0.50-0.75</b> | 45.49 |
| <b>0.75-1.00</b> | 17.83 |

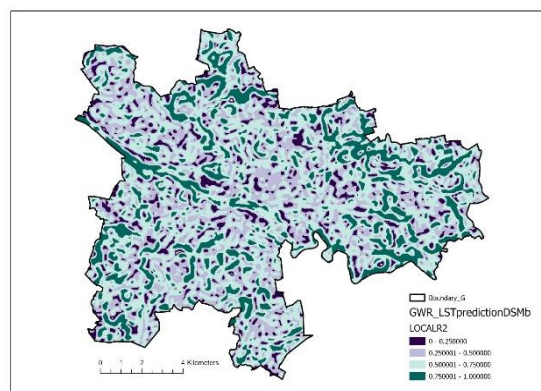


Figure 28 DSM and LST in GWR

### 5.2.5.2. GWR and CDSM

The local  $R^2$  altered between 0 and .94, and the GWR result shows that the regression model has the ability to explain 75-100% of the variation between CDSM and LST for only 17.68 per cent of the whole data (Figure 29).

| Local R2         |       |
|------------------|-------|
| <b>0-0.25</b>    | 7.83  |
| <b>0.25-0.50</b> | 26.18 |
| <b>0.50-0.75</b> | 48.31 |
| <b>0.75-1.00</b> | 17.68 |

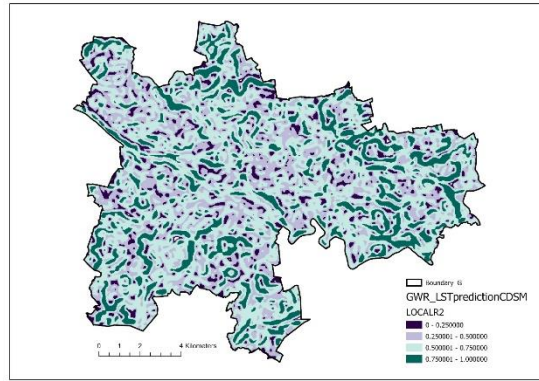


Figure 29 CDSM and LST in GWR

#### 5.2.5.4. GWR and SVF

The local  $R^2$  results came out between 0 and .94, which demonstrates that the GWR model can predict 75-100% of the variation between SVF and LST for only 16.57 per cent of the whole data (Figure 30).

| Local R2         |       |
|------------------|-------|
| <b>0-0.25</b>    | 8.33  |
| <b>0.25-0.50</b> | 26.23 |
| <b>0.50-0.75</b> | 48.87 |
| <b>0.75-1.00</b> | 16.57 |

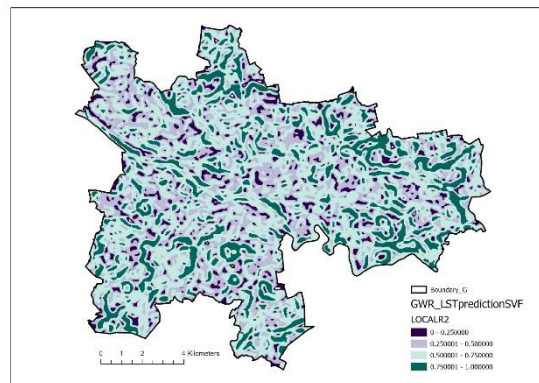


Figure 30 SVF and LST in GWR

#### 5.2.5.5. GWR and NDVI

The local  $R^2$  varied between 0 and .94, and the results show that the GWR model can explain 75-100% of the variation between NDVI and LST for only 21.36 per cent of the whole city (Figure 31).



| Local R2         |       |
|------------------|-------|
| <b>0-0.25</b>    | 5.8   |
| <b>0.25-0.50</b> | 21.88 |
| <b>0.50-0.75</b> | 50.96 |
| <b>0.75-1.00</b> | 21.36 |

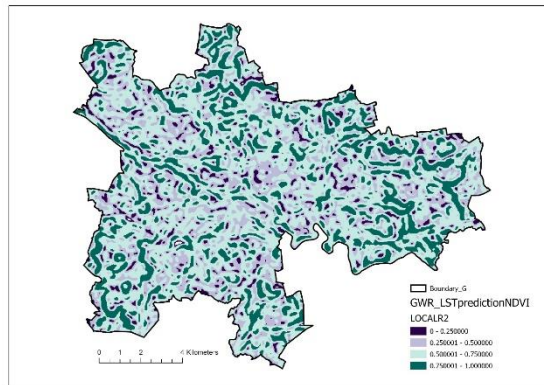


Figure 31 NDVI and LST in GWR

### 5.2.5.6. GWR and NDBI

The local  $R^2$  ranges between 0 and .93, and the result demonstrates that the GWR model is able to explain 75-100% of the variation between NDBI and LST for only 22.01 per cent of all data (Figure 32).

| Local R2         |       |
|------------------|-------|
| <b>0-0.25</b>    | 5.56  |
| <b>0.25-0.50</b> | 21.15 |
| <b>0.50-0.75</b> | 51.28 |
| <b>0.75-1.00</b> | 22.01 |

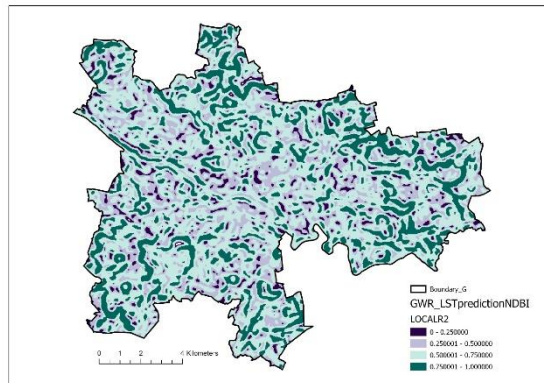


Figure 32 NDBI and LST in GWR

### 5.2.5.7. GWR and NOx

The local  $R^2$  fluctuates between 0 and .97, and the results depict that the GWR model could explain 75-100% of the variation between NOx and LST for only 16.09 per cent of the whole Glasgow (Figure 33).

| Local R2         |       |
|------------------|-------|
| <b>0-0.25</b>    | 11.88 |
| <b>0.25-0.50</b> | 29.65 |
| <b>0.50-0.75</b> | 42.38 |
| <b>0.75-1.00</b> | 16.09 |

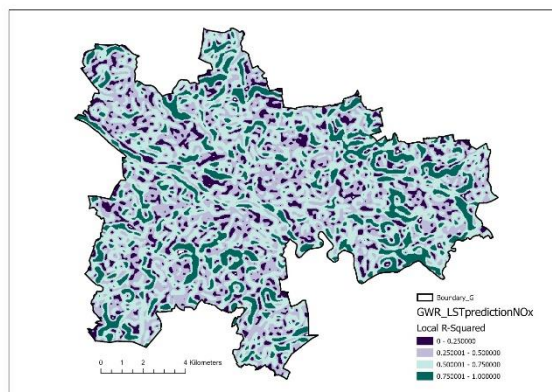


Figure 33 NOx and LST in GWR

### 5.2.5.8. GWR and PM10

The local  $R^2$  spans between 0 and 0.97, and the results express that the GWR model has the ability to explain 75-100% of the variation between PM10 and LST for only 15.80 per cent of the whole data (Figure 34).

| Local R2         |       |
|------------------|-------|
| <b>0-0.25</b>    | 9.72  |
| <b>0.25-0.50</b> | 29.75 |
| <b>0.50-0.75</b> | 42.53 |
| <b>0.75-1.00</b> | 15.80 |

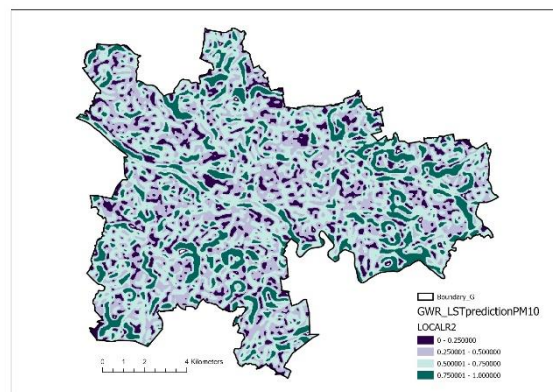


Figure 34 PM10 and LST in GWR

The combined GWR models express the impact of spatial factors in a group of 5(6)-predictors on LST changes. For two chosen models with 5–6 predictors, the GWR results are compared together in Table 16.

### 5.2.5.8. GWR and DEM\_SVF\_NDVI\_NOX\_PM10

The local  $R^2$  fluctuates between 0 and 0.97, and the results illustrate that the GWR model can explain 75-100% of the variation between DEM and LST for only 45.32 per cent of the whole data (Figure 35).

| Local R2         | %     |
|------------------|-------|
| <b>0-0.25</b>    | 1.60  |
| <b>0.25-0.50</b> | 10.49 |
| <b>0.50-0.75</b> | 42.59 |
| <b>0.75-1.00</b> | 45.32 |

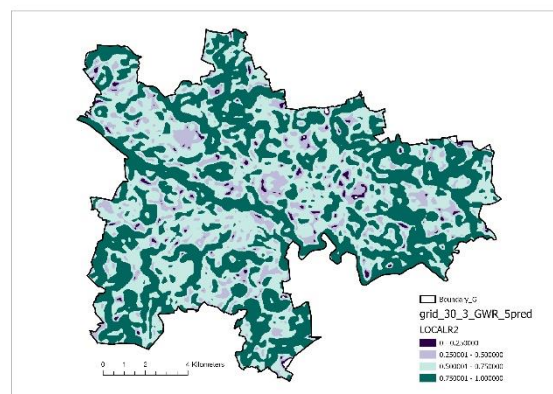


Figure 35 Group of DEM\_SVF\_NDVI\_NOX\_PM10 and LST in GWR

### 5.2.5.9. GWR and DEM\_SVF\_NDVI\_NDBI\_NOX\_PM10

The local  $R^2$  ranges between 0 and 0.98, and the results show that the GWR model is able to explain 75-100% of the variation between DEM and LST for only 60.84 per cent of the whole data (Figure 36).



Figure 36 Group of DEM\_SVF\_NDVI\_NDBI\_NOX\_PM10 and LST in GWR

The comparison of combined GWR models shows the significance of NDBI in prediction, which revealed better performance as a non-stationary predictor for LST. The improvement was reflected in increasing the local coefficient of determination from 45.32 to 60.84, explaining the 75-100% variation.

Table 16 Combined GWR models

| Model Type                      | Predictors   | R-Squared | Adjusted R-Squared | Akaike's Information Criterion (AICc) |
|---------------------------------|--------------|-----------|--------------------|---------------------------------------|
| GWR: DEM_SVF_NDVI_NOX_PM10      | 5 Predictors | 0.9404    | 0.9379             | 389854.9805                           |
| GWR: DEM_SVF_NDVI_NDBI_NOX_PM10 | 6 Predictors | 0.9620    | 0.9591             | 310199.1428                           |

## 5.2. Non-linearity by Neural Network

### 5.2.1. Architecture of LST prediction model

The ANN model was optimised to yield the fewest errors between anticipated and real values. This was accomplished by adjusting the number of neurons within the hidden layer between 6 and 25. The optimum number of 22 was selected due to the low error value for all three portions of the training dataset (70, 15, and 15 for training, validation, and testing) aside from the 20 per cent of the whole data assigned for testing. Figure 37 illustrates the optimum

number of neurons (22) for the prediction, which causes the least MSE without causing overfitting and underfitting.

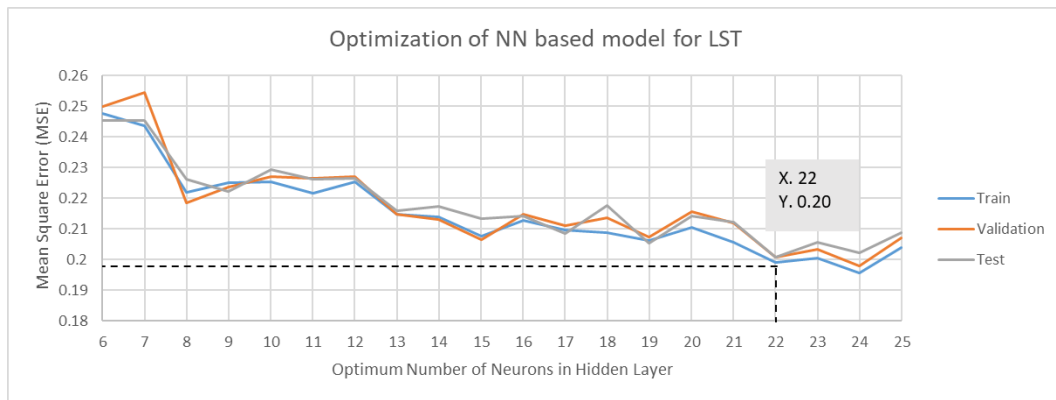


Figure 37 Optimisation of Hyperparameters (The Number of Neurons in Hidden Layer)

To avoid the complication of NN architecture for generalizing the prediction process, the Neural Network model is made of three main layers: input layer, hidden layer and an output layer as seen in *Figure 38*.

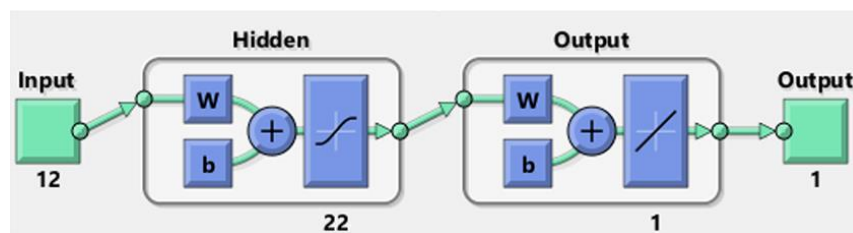


Figure 38 Architecture of Neural Network

### 5.2.2. LST NN Model Performance

The error histogram for the trained NN model illustrates the normality distribution in the majority of errors falling within the range of +1 and -1, indicating that LST models have reasonable predictability. In addition, the zero-error line crossed the error bar with the highest instances (*Figure 39*).

The residual plots for three partitions of the trained model in *Figure 40* show the best-fit correlation lines, which are almost close to a 45-degree angle. The overall residual plot demonstrates the remarkable correlation between observed (true) values and predicted values. Additionally, the coefficient of determination computed for the model provides

details on the model's ability to predict 80 per cent of the variance (the R-value of .88-.89) in LST with the selected predictors.

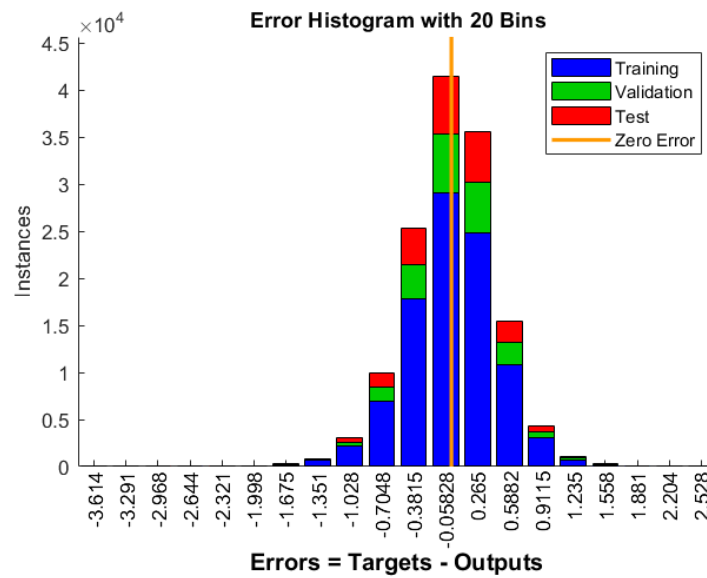


Figure 39 left: Error histogram

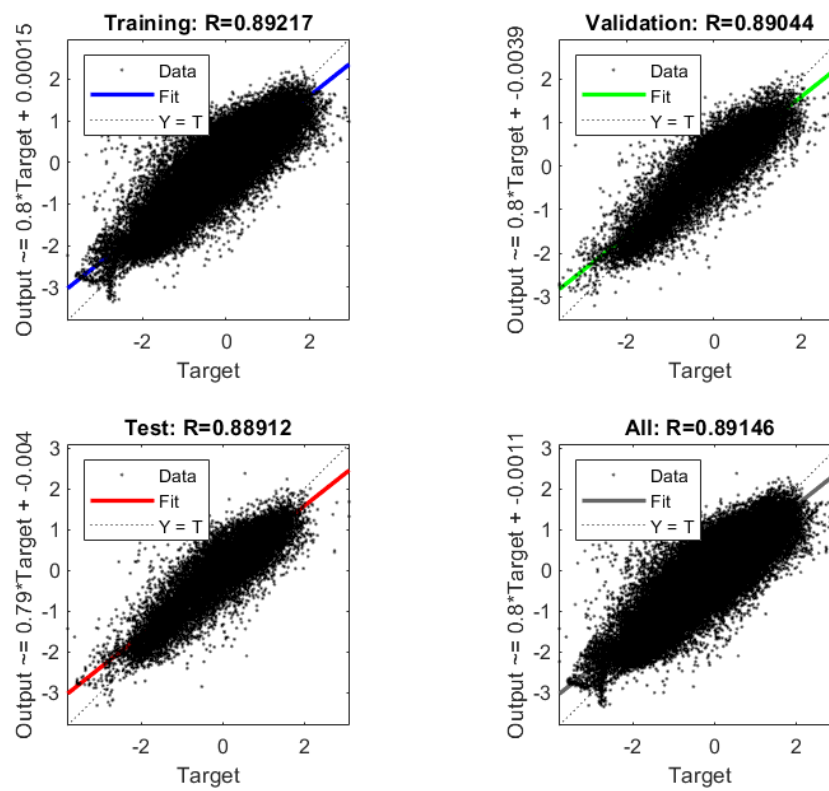
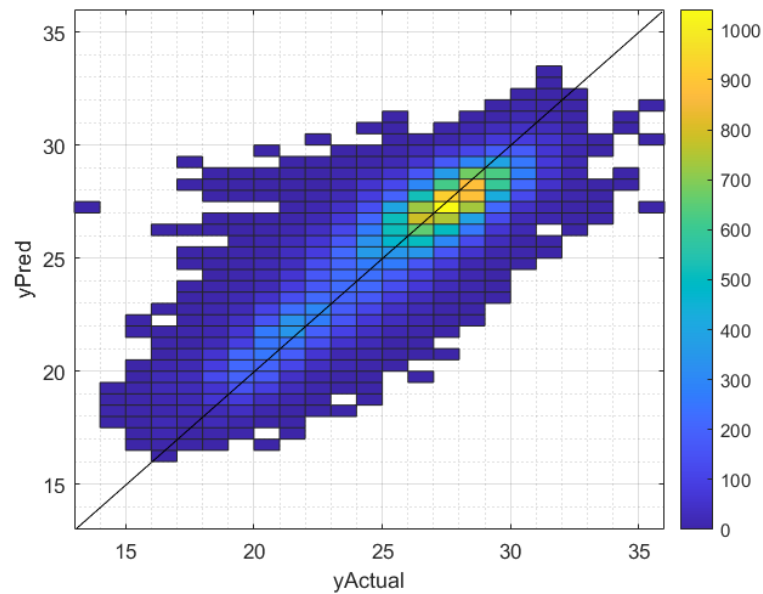


Figure 40 Residual plots for the NN based LST model

The 2D histogram shown in *Figure 41* clarifies that the highest proportion of observed values are located on the 45-degree fit-line. The bias could be found for the outliers, which have the lowest number of incidents for the training process. To avoid the influence of outliers, two models trained through oversampling (the outliers) and undersampling (the observation with the highest instances) techniques and results were acceptable in the training process but weak performance in predicting the test dataset.



*Figure 41 Distribution of biased prediction in the LST model*

### 5.2.3. Sensitivity Analysis

The result of sensitivity analysis illustrates the following key points based on the importance of the predictor through the elimination of each predictor from the training process and the elimination of predictors from one data source (Appendix VI). NDBI has the highest significance among the predictors due to the greatest difference in coefficient of determination. The essential elements in forecasting the LST based on secondary data are NDVI, DEM, DSM, SVF, BSF, and pollution variables. Landsat 8 is the most influential dataset for predicting LST using non-linear regression of Neural Network, followed by the LiDAR dataset, pollution, LULC, and vector-based data. *Figure 42* exhibits the importance of all predictors in predicting over 90% of the LST. Furthermore, sensitivity analysis suggested that the most significant dataset for LST prediction was the RSD dataset.

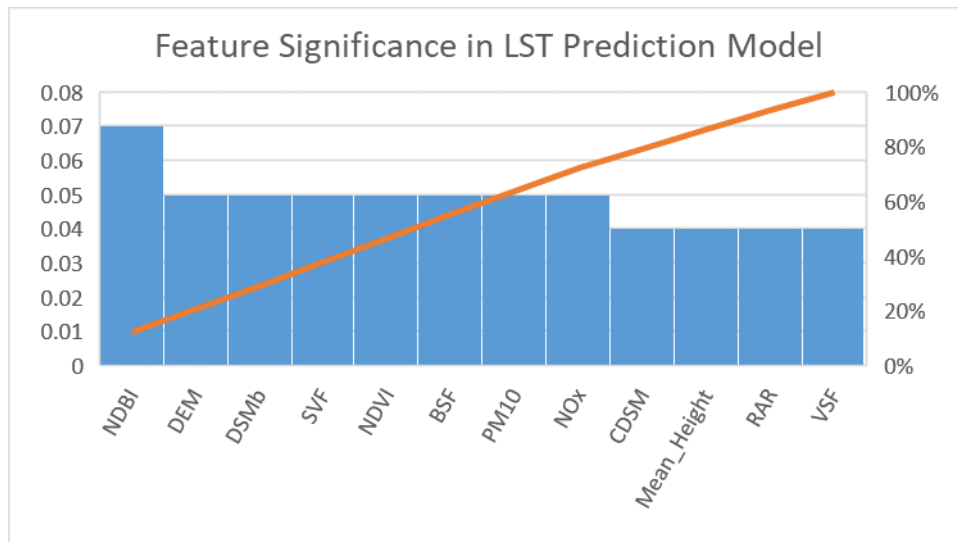


Figure 42 Feature Significance in LST Prediction Model from sensitivity analysis

### 5.3. Validation of NN models: LST Prediction and Human Thermal Comfort

The scenarios listed below are taken into account for validating the trained ANN prediction model (Figure 43).

**Best Case:** All the vacant lands are becoming fully vegetated areas in which all the contributing predictors (vegetation, built-up area, urban form) will take new data due to the land modification.

**Worst Case:** All the lands converted into built-up areas.

**Intermediate Case:** Due to the changes in almost all predictors, vegetation, built-up area, and urban form factors are getting new values (50-50 vegetation and built-up).

Table 17 demonstrates the impact of outliers in prediction as in all cases; the model had weak performance in the prediction of minimum and maximum LST.

Table 17 Comparison of Scenarios and RS-LS

|                              | Min LST | Max LST |
|------------------------------|---------|---------|
| <b>LST 2018</b>              | 13.42   | 35.67   |
| <b>Best Scenario</b>         | 18.33   | 31.13   |
| <b>Intermediate Scenario</b> | 18.10   | 31.65   |
| <b>Worst Scenario</b>        | 18.51   | 31.19   |



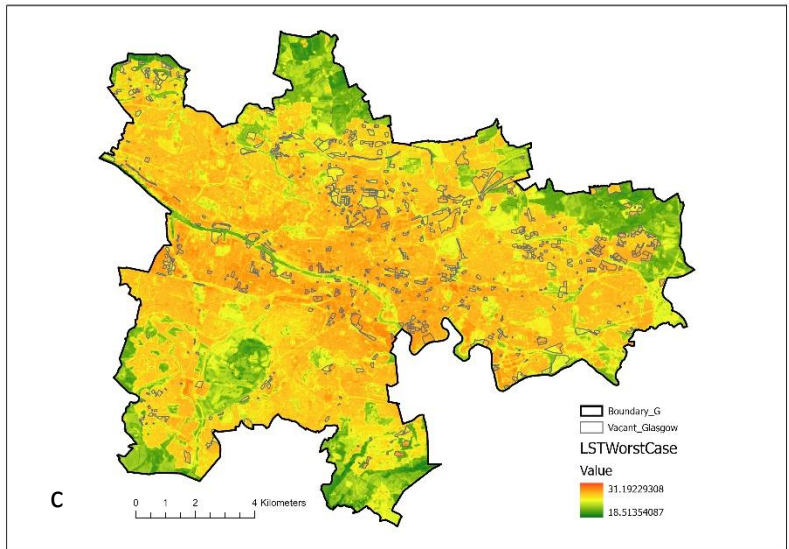
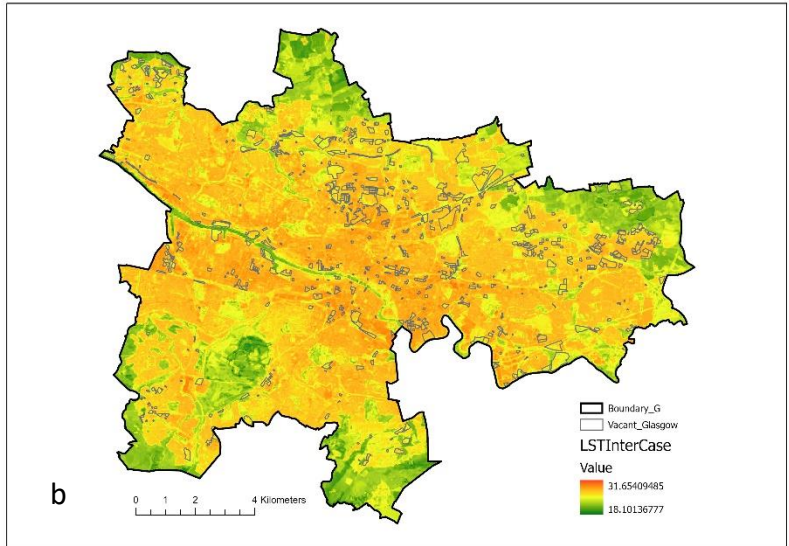
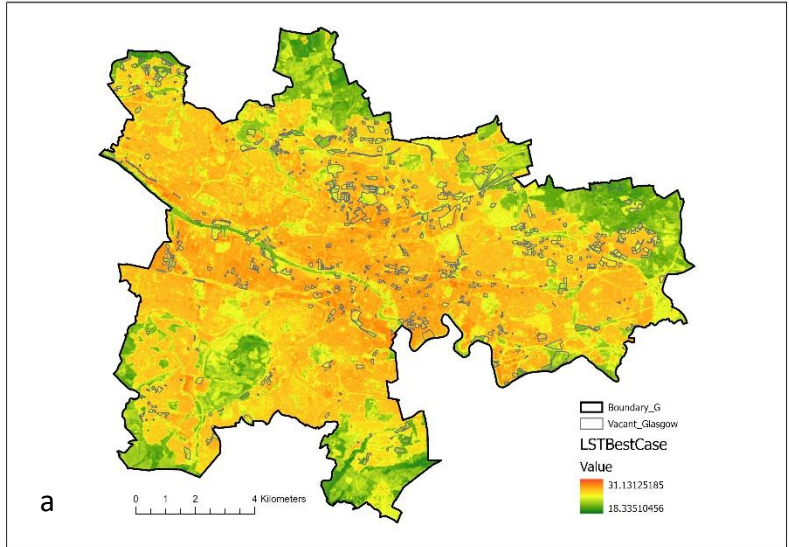
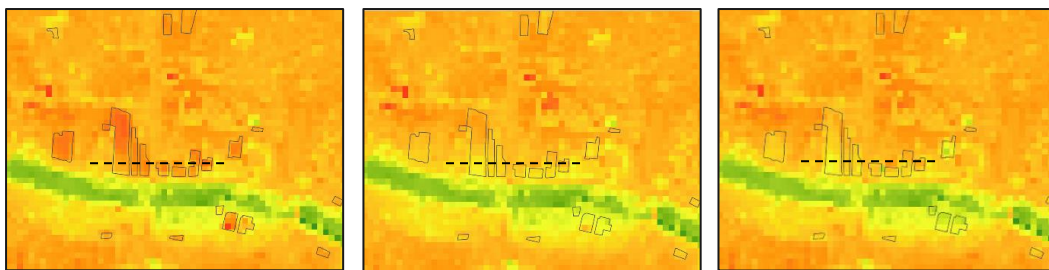


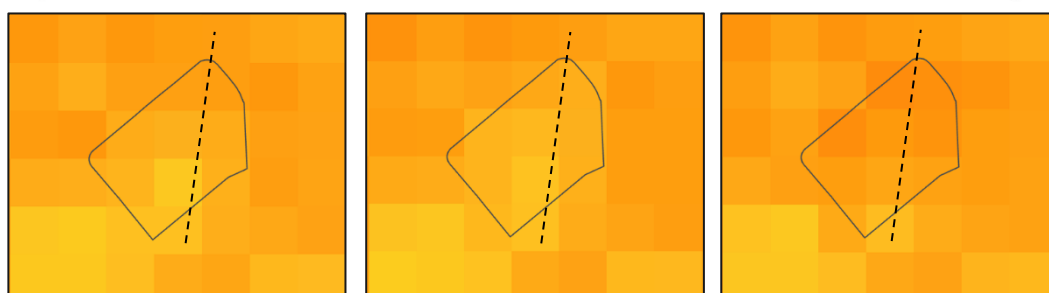
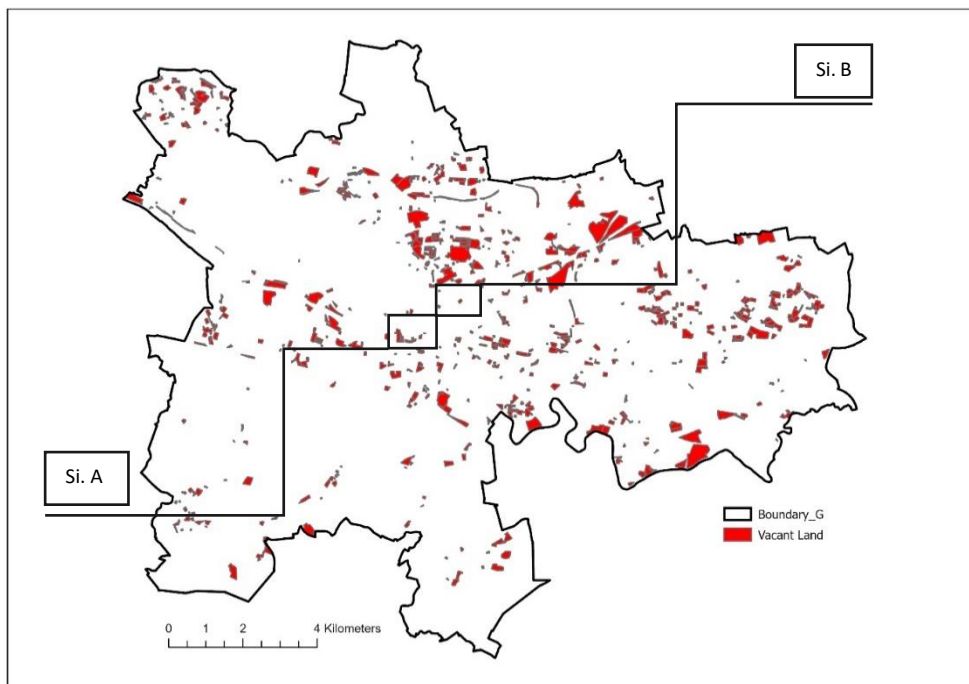
Figure 43 a: Best Scenario b: Intermediate Scenario c: Worst Scenario



To make the difference clear in the scenarios, *Figure 44* provides the higher resolution of three scenarios for a few vacant lands in the city centre area. The ANN model results provide valuable information about changes in LST based on different scenarios. It is seen that the vegetation proposed for the vacant lands in the denser part of the city was able to decrease the temperature at the level of the surface layer. At the same time, the worst scenario increased the temperature of the vacant land by increasing the built-up area. In addition, combining the given scenarios, the intermediate provides predictable results (Site A). The results from the prediction for Site B were more challenging.



Site A: Left to Right: Worst Case, Intermediate Case, and Best Case



Site B: Left to Right: Worst Case, Intermediate Case, and Best Case

*Figure 44 Predicted LST for two sites in the city centre*

Figure 45 demonstrates the LST fluctuation in detail to provide how the trained ANN model works for each grid pixel in Site A. LST fluctuation plot is a great tool to find how the model works specifically in case of comparison with the LST retrieved from the remote-sensing data. The worst-case behaviour shows the highest predicted value for LST, while the Best and Inter scenarios are equivocal due to high variation in comparison with the RS-LST. The fluctuation ranged from -1.76 to 1.38 for the best-case scenario, -1.33 to .88 for the intermediate situation, and .51 to 2.62 for the worst case compared to RS-LST.

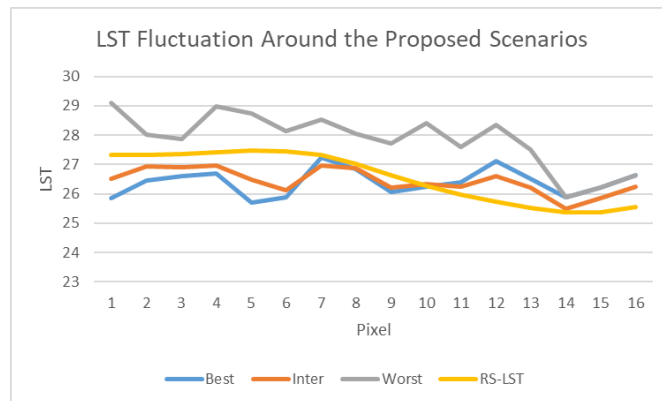


Figure 45 LST Fluctuation of predicted for the proposed Scenarios Site A

Table 18 represents the average predicted values for Site A. The best scenario resulted in the reduction of RS-LST of .177 °C and was recorded as the lowest temperature, followed by intermediate and worst cases. The intermediate scenario also reduced .133 °C, while the worst scenario revealed an increase in LST by 1.289 °C.

Table 18 Comparison of LST predictions on selected scenarios

| Best     | Inter    | Worst    | RS-LST |
|----------|----------|----------|--------|
| 26.39183 | 26.43581 | 27.85728 | 26.568 |

For Site B, the LST projected under each scenario showed a reduction in RS-LST. The best-case scenario was predicted to have the highest values for LST in contrast to the worst-case and intermediate scenarios. The notable point about Site B is that the outcome was dissimilar to what was seen at Site A. The fluctuation ranged from -1.22 to 1.44 for the best-case scenario, -1.57 to 1.53 for the intermediate situation, and -1.39 to 1.44 for the worst case compared to RS-LST (Figure 46). All in all, the reduction of LST was observed as a result of all scenarios.

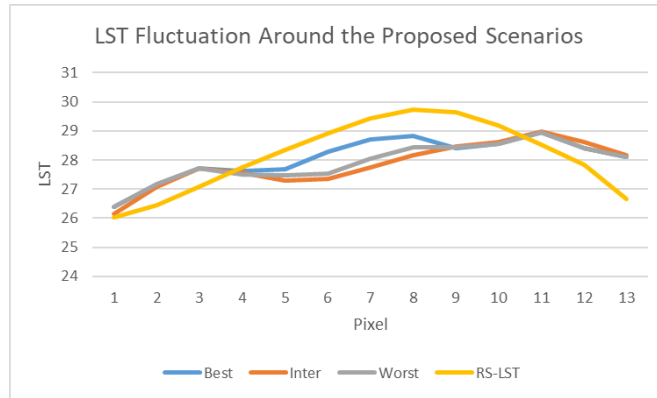


Figure 46 LST Fluctuation of predicted for the proposed Scenarios Site B

Table 19 illustrates the average predicted values for Site B, where the best scenario reduced RS-LST by .054 °C while intermediate and worst cases resulted in a reduction of .281 °C and .217 °C, respectively. Additionally, the LST was lower for the scenario with 50% vegetation than for the scenario with 100% built-up land.

Table 19 Comparison of LST predictions on selected scenarios

| Best     | Inter    | Worst    | RS-LST   |
|----------|----------|----------|----------|
| 28.06529 | 27.83872 | 27.90251 | 28.11971 |

### 5.3.1. Predicted LST and Human Thermal Comfort

The results from LST prediction depict the modest impact of urban greenery in decreasing the heat stress at the surface level under a non-linear trend. In order to clarify the impacts of LST changes, the UTCI and MRT simulation were considered for the two selected sites to compare the impacts of heat mitigation strategies on human thermal comfort.

#### 5.3.1.1. ENVI-met microclimate simulation

- UTCI Simulation

The outcome of the UTCI simulation for both locations shows that the UTCI index can be classified into four main categories: No-TS<sup>4</sup>, Moderate-TS<sup>5</sup>, Strong-TS<sup>6</sup>, and Very Strong-TS<sup>7</sup>. For the first location, the highest proportion (over 43%) belonged to the Moderate-TS, while the lowest was recorded as the Very Strong-TS (less than 0.5%), which was seen in intermediate and worst-case scenarios as well. In the best scenario, the Very Strong-TS was

<sup>4</sup> No Thermal Stress

<sup>5</sup> Moderate Thermal Stress

<sup>6</sup> Strong Thermal Stress

<sup>7</sup> Very Strong Thermal Stress

not observed at all. The increasing built-up areas had an increasing impact on both No-TS areas and Strong-TS due to the shading effect of nearby high-rise buildings and increasing surfaces with higher albedo respectively (Figure 47).

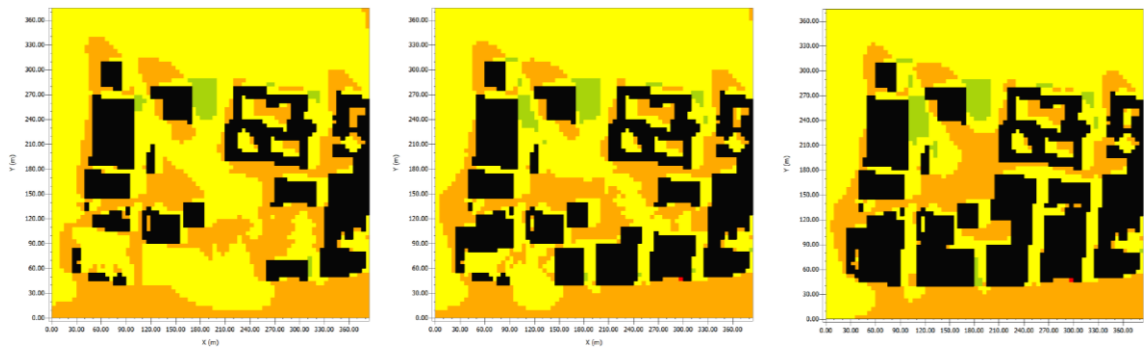


Figure 47 Site A UTCI index  
 Left: Best Scenario Mid: Intermediate Scenario Right: Worst Scenario

| UTCI                       |                   |
|----------------------------|-------------------|
| No thermal stress          | below 14.00 °C    |
| No thermal stress          | 14.00 to 20.00 °C |
| No thermal stress          | 20.00 to 26.00 °C |
| Moderate thermal stress    | 26.00 to 32.00 °C |
| Strong thermal stress      | 32.00 to 38.00 °C |
| Very strong thermal stress | above 38.00 °C    |

The absolute difference of the UTCI index (with base case) in Figure 48 expresses the favourable impact of vegetation in reducing the UTCI effect in an open space by more than 3K. The new buildings' shading also influences the UTCI at the street canyon. Where the shading effect of buildings was not strong, the UTCI increased up to 6K.

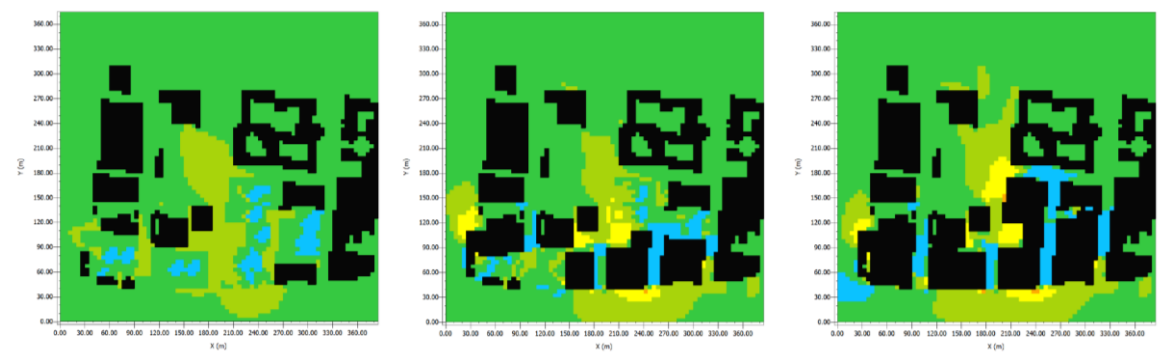


Figure 48 Site A UTCI absolute difference  
 Left: Best Scenario Mid: Intermediate Scenario Right: Worst Scenario

| absolute difference UTCI |                 |
|--------------------------|-----------------|
|                          | below -3.00 K   |
|                          | -3.00 to 0.00 K |
|                          | 0.00 to 3.00 K  |
|                          | 3.00 to 6.00 K  |
|                          | 6.00 to 9.00 K  |
|                          | above 9.00 K    |

In the second site, more areas were simulated under the Very Strong-TS category, and a higher proportion of sites were simulated as Very Strong-TS for all scenarios in comparison to Site A. The Very Strong-TS class has been observed in every scenario. As observed in all scenarios with green spaces, the suggested vegetated area in the vacant land was able to maintain the open space under Moderate-TS conditions (Figure 49).

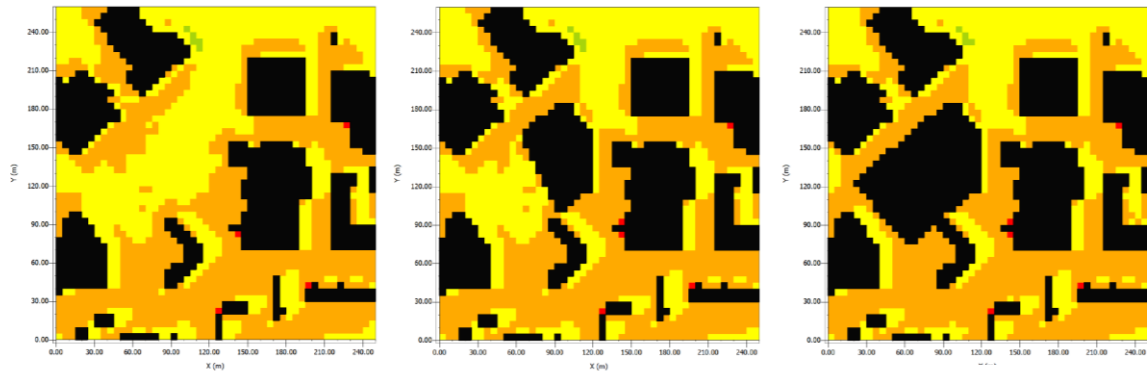
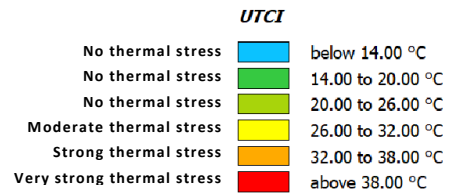


Figure 49 Site B UTCI index

Left: Best Scenario Mid: Intermediate Scenario Right: Worst Scenario



The absolute difference of the UTCI index in Figure 50 demonstrates the beneficial impact of vegetation in reducing the UTCI at the street canyon by more than 3.39K; however, the worst scenario caused an increasing UTCI up to 4.61K at street level.

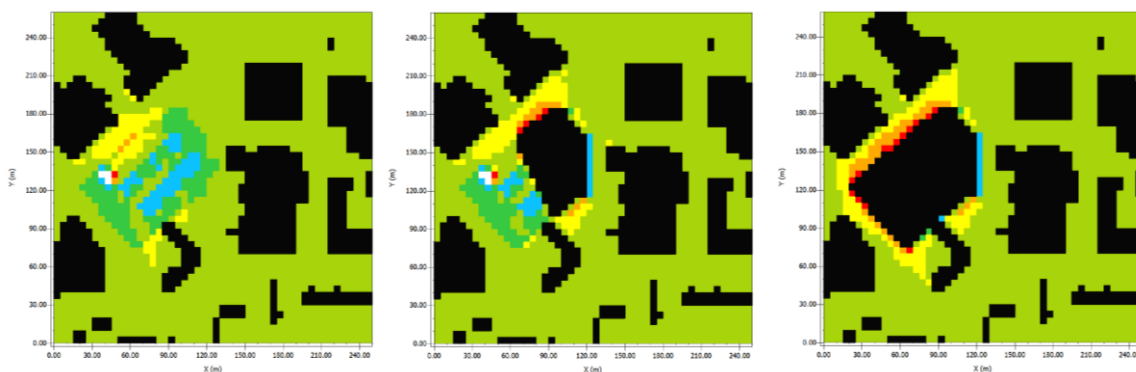


Figure 50 Site B UTCI absolute difference

Left: Best Scenario Mid: Intermediate Scenario Right: Worst Scenario

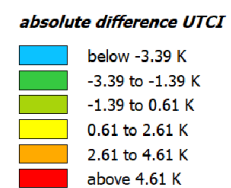


Table 20 presents insight into how thermal stress alters through different scenarios areas rose while the areas subject to strong and very strong classification reduced (UTCI index of between +26 and +32). Due to the increase in built-up areas that ENVI-met does not include in its simulation of thermal stress, the overall area of the thermal stress zone decreased.

Table 20 UTCI area changes through georeferenced simulated scenarios

| Site                | Min-Max UTCI °C | No TS % | Moderate TS % | Strong TS % | Very strong TS % | Cumulative TS % | Built-up |
|---------------------|-----------------|---------|---------------|-------------|------------------|-----------------|----------|
| <b>1</b>            |                 |         |               |             |                  |                 |          |
| <b>Best</b>         | 22.37-37.96     | 1.35    | 57.78         | 15.56       | 0                | 73.34           | 25.31    |
| <b>Intermediate</b> | 22.15-38.11     | 1.94    | 50.94         | 16.07       | 0.17             | 67.13           | 30.93    |
| <b>Worst</b>        | 22.95-38.11     | 2.45    | 43.01         | 17.77       | 0.17             | 60.95           | 36.60    |
| <b>2</b>            |                 |         |               |             |                  |                 |          |
| <b>Best</b>         | 25.73-38.41     | 0.85    | 33.25         | 36.61       | 0.68             | 60.54           | 28.61    |
| <b>Intermediate</b> | 25.75-38.43     | 0.85    | 29.03         | 36.87       | 0.85             | 66.75           | 32.40    |
| <b>Worst</b>        | 25.79-38.49     | 0.68    | 23.87         | 36.93       | 0.85             | 61.65           | 37.67    |

- MRT Simulation

The simulated MRT showed the same pattern as UTCI in the positive effect of shading in Site B and positive impacts of vegetation in Site A (Figure 51-52).

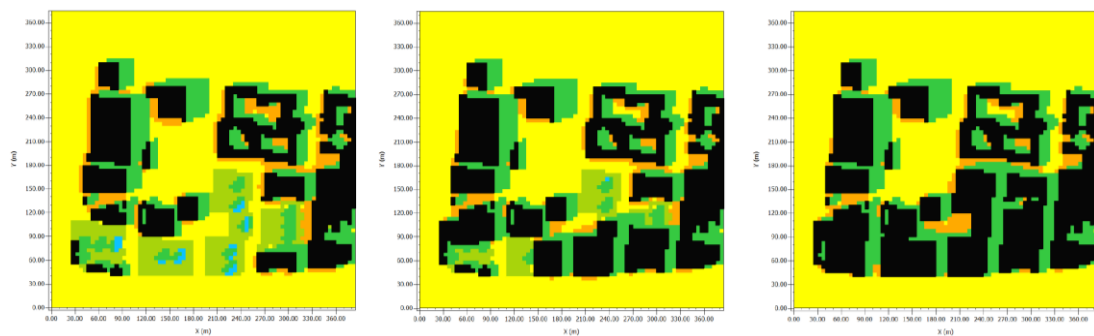


Figure 51 Site A MRT

Left: Best Scenario Mid: Intermediate Scenario Right: Worst Scenario

Mean Radiant Temp.

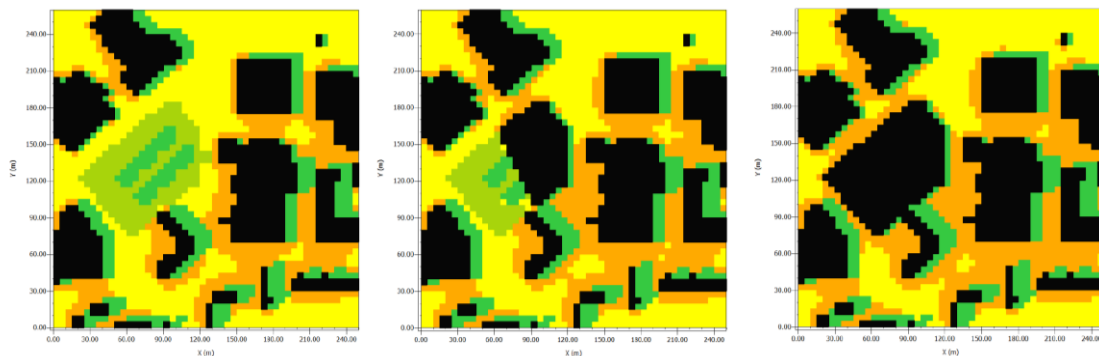
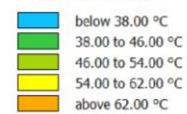


Figure 52 Site B MRT

Left: Best Scenario Mid: Intermediate Scenario Right: Worst Scenario

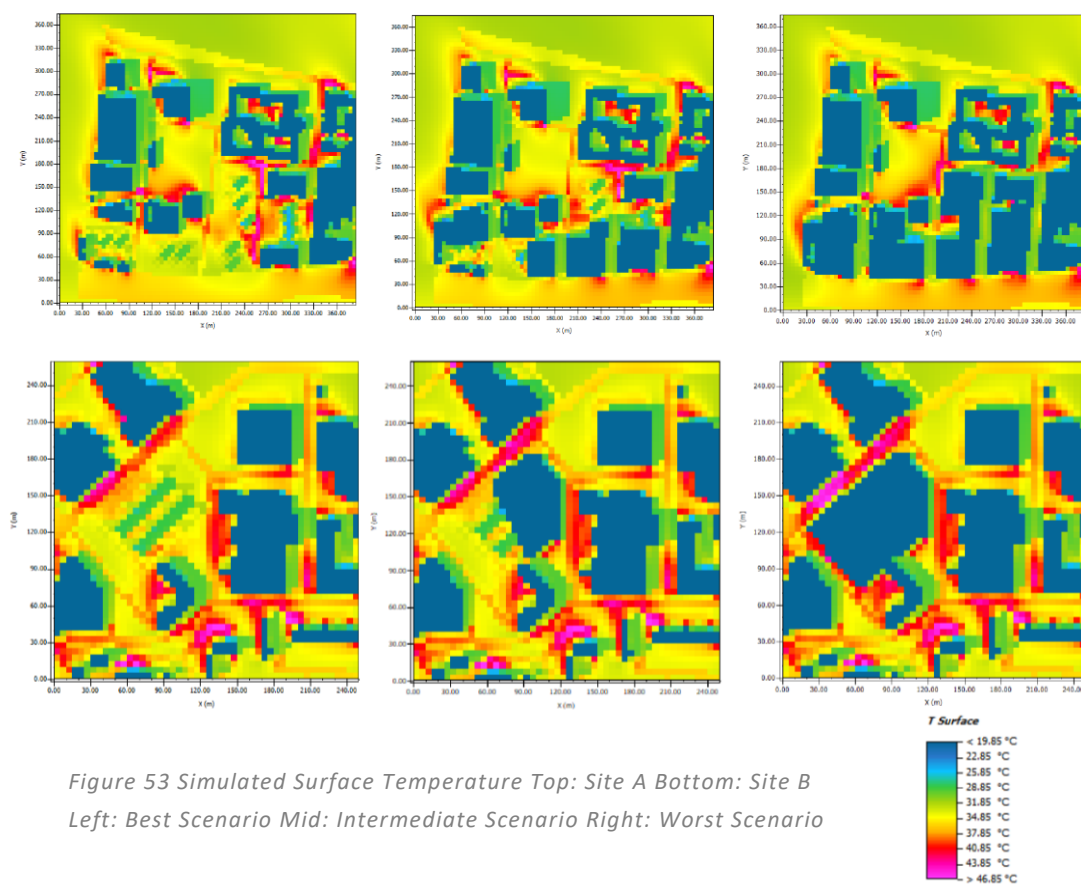
Mean Radiant Temp.



The comparison of the three scenarios shows that when more lands are developed through construction, the overall UTCI index will rise. In other words, the surrounding area will experience more severe heat stress following physical development unless their shading effect appears strong.

- LST Simulation

The simulated LST from ENVI-met for two sites is shown in *Figure 53*. The overall simulated LST was recorded higher than the predicted LST and RS-LST. The vegetation affected the LST to keep the surface cooler than the surrounding. In Site A, the shading effect helped the surface to keep cool, while for Site B, the surface for the worst scenario simulated over 46.65 °C.



Changes in simulated LST compared to the base case for each scenario are shown in *Figure 54*. It details about the positive effect of increasing built-up surface on simulated LST in dense and compact contexts. However, the worst scenario in Site B had recorded the increasing surface temperature of up to 11K in some parts of the street canyon. Vegetation showed positive impact in less compact urban areas with higher spatial openness.

For the best scenario in Site A, the LST increased 3-5K in vegetated lands. For the second location, avoiding the built-up surfaces, the worst scenario caused the increasing rate of LST, while in the intermediate scenario, the increasing temperature at the level of the surface was

not as severe as the worst scenario (the opposite of predicted LST). The removal of temporary structures from the land in the Best and Intermediate scenarios (Site B) resulted in the large absolute difference from Base Case, increased LST above 13K in several parts.

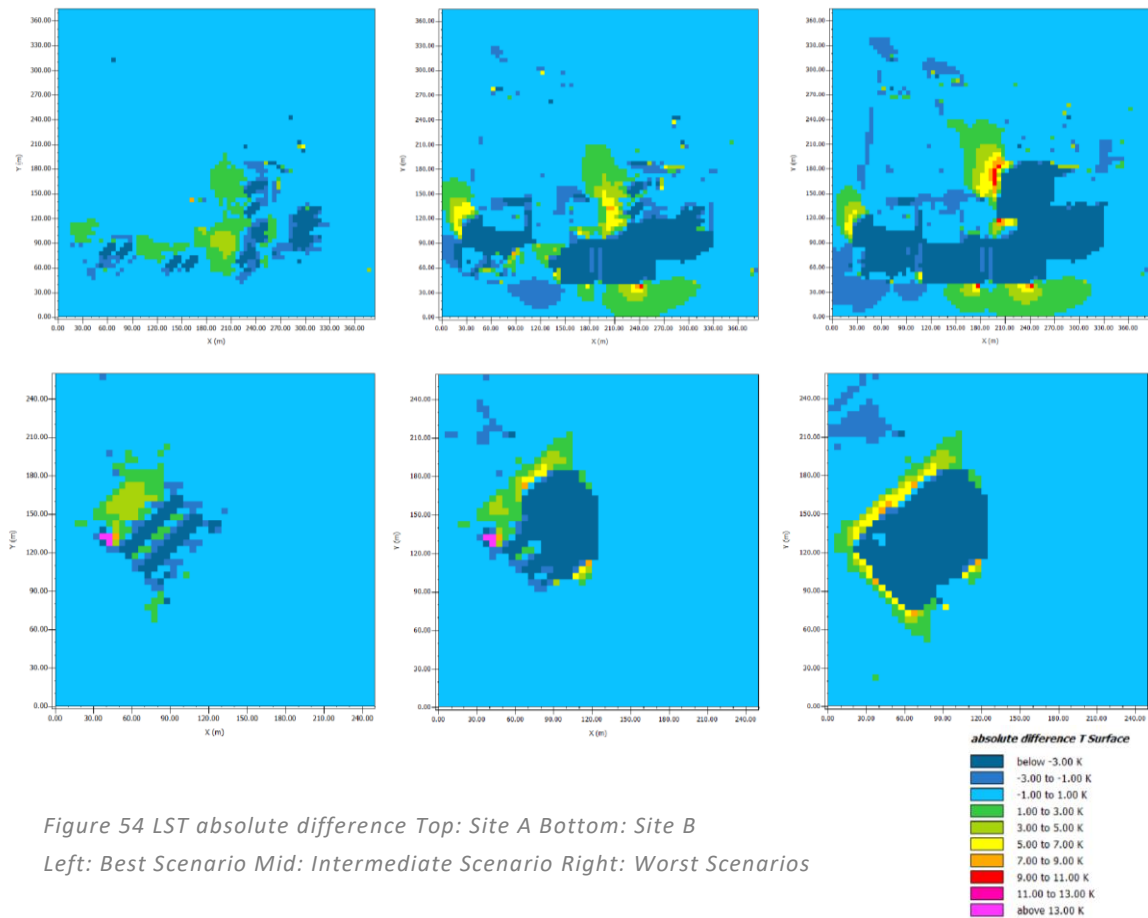


Figure 54 LST absolute difference Top: Site A Bottom: Site B  
 Left: Best Scenario Mid: Intermediate Scenario Right: Worst Scenarios

The predicted LST based on the Neural Network model showed the stark variation between the worst and the best scenarios observed in the simulation of LST in microclimate modelling for site A. Besides, the fluctuation between intermediate and worst scenarios was similar to the comparison of best and intermediate scenarios. For Site B, under vegetated areas, the LST decreased, while for the other areas, LST increased, which was stronger in the worst scenario.





## CHAPTER 6: RESULTS AND ANALYSIS- PART II

The second section of the result chapter describes how the MRT may be predicted by LST-based predictors.

### 6.1. MRT and Human Thermal Comfort?

#### 6.1.1. Correlation Matrix: MRT

The Pearson correlation values are displayed in *Figure 55*. The coefficient correlation employed in multiple regression analysis reveals the association between SVF and MRT is the strongest, while DEM is the least correlated factor. Additionally, among all the predictors, NDVI, VSF, CDSM, DEM, and DSM showed a negative correlation, as was observed in the LST correlation. The comparison between the MRT correlations with predictors at two separate times of the day depicts the clear difference for VSF, RAR, DSM and SVF predictors (*Table 21*).

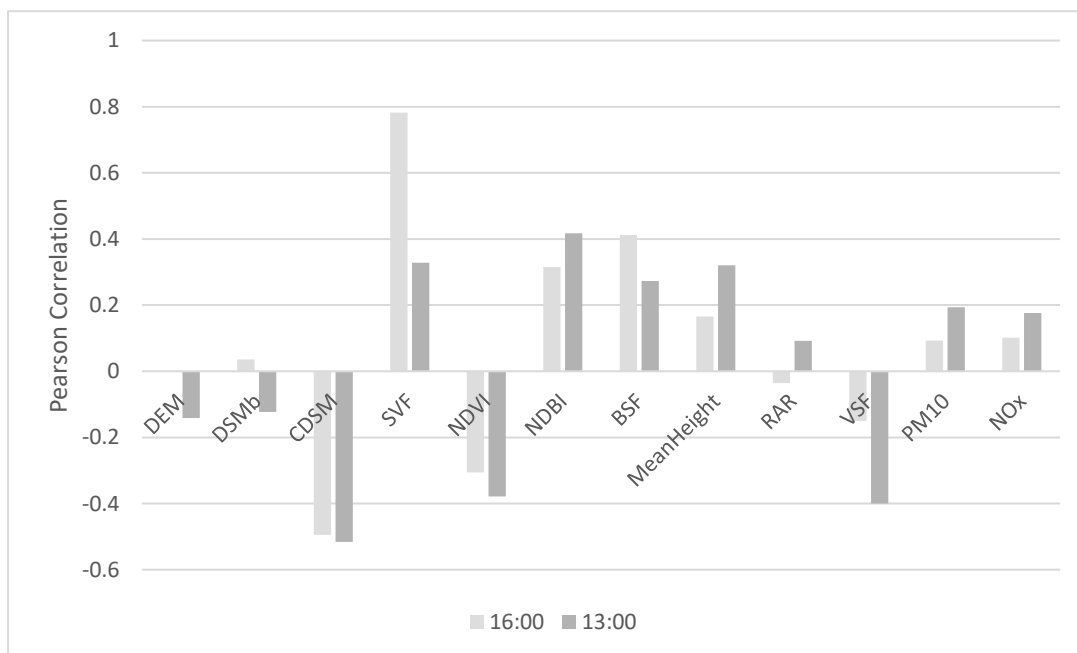


Figure 55 Pearson Correlation between Predictors and MRT

Table 21 Pearson Correlation between Predictors and MRT

| Pearson CC       | DEM    | DSMb   | CDSM   | SVF   | NDVI   | NDBI  | BSF   | MeanHeight | RAR    | VSF    | PM10  | NOx   |
|------------------|--------|--------|--------|-------|--------|-------|-------|------------|--------|--------|-------|-------|
| <b>MRT 16:00</b> | -0.001 | 0.036  | -0.495 | 0.782 | -0.306 | 0.316 | 0.412 | 0.166      | -0.036 | -0.150 | 0.093 | 0.102 |
| <b>MRT 13:00</b> | -0.141 | -0.123 | -0.516 | 0.328 | -0.378 | 0.417 | 0.273 | 0.320      | 0.092  | -0.400 | 0.194 | 0.176 |

### 6.1.2. LST and MRT Correlation

The linear regression between simulated MRT and derived LST shows a weak relationship between the two response variables, where only 8 percent of LST can be estimated by MRT (13:00) summarised in *Table 22*.

*Table 22 Correlation between MRT and LST*

| Input Features:             | grid_30_7      | Dependent Variable:                         | MAX_MRT        |
|-----------------------------|----------------|---|----------------|
| Number of Observations:     | 198310         | Akaike's Information Criterion (AICc) [d]:  | 1184447.249499 |
| Multiple R-Squared [d]:     | 0.086778       | Adjusted R-Squared [d]:                     | 0.086774       |
| Joint F-Statistic [e]:      | 18844.056958   | Prob(>F), (1,198308) degrees of freedom:    | 0.000000*      |
| Joint Wald Statistic [e]:   | 7212.896972    | Prob(>chi-squared), (1) degrees of freedom: | 0.000000*      |
| Koenker (BP) Statistic [f]: | 15224.301978   | Prob(>chi-squared), (1) degrees of freedom: | 0.000000*      |
| Jarque-Bera Statistic [g]:  | 1335077.761393 | Prob(>chi-squared), (2) degrees of freedom: | 0.000000*      |

### 6.1.3. Generalized Linear Regression and Exploratory Regression results

The results from Generalized Linear Regression (GLR) model in *Table 23* express that all variables are statistically associated with MRT in linear regression. The variance inflation (VIF) of each predictor ranges from 1.20 to 528.99, demonstrating the impact of multicollinearity and redundancy among independent variables.

*Table 23 Summary of GLR Result-MRT*

| Variable   | Coefficient [a] | StdError | t-Statistic | Probability [b] | Robust_SE | Robust_t   | Robust_Pr [b] | VIF [c]    |
|------------|-----------------|----------|-------------|-----------------|-----------|------------|---------------|------------|
| Intercept  | 54.172613       | 0.197838 | 273.823474  | 0.000000*       | 0.203082  | 266.752179 | 0.000000*     | -----      |
| DEM        | 0.253701        | 0.006746 | 37.608729   | 0.000000*       | 0.004872  | 52.070876  | 0.000000*     | 528.990332 |
| DSMB       | -0.259885       | 0.006742 | -38.549047  | 0.000000*       | 0.004891  | -53.130531 | 0.000000*     | 515.631618 |
| CDSM       | -1.132185       | 0.006649 | -170.288116 | 0.000000*       | 0.014693  | -77.053733 | 0.000000*     | 2.119121   |
| SVF        | -0.480820       | 0.103517 | -4.644825   | 0.000005*       | 0.164391  | -2.924853  | 0.003455*     | 3.026691   |
| NDVI       | 5.616758        | 0.183986 | 30.528151   | 0.000000*       | 0.165394  | 33.959964  | 0.000000*     | 8.638775   |
| NDBI       | 7.559429        | 0.247976 | 30.484498   | 0.000000*       | 0.249076  | 30.349912  | 0.000000*     | 8.353480   |
| BSF        | 3.136801        | 0.101963 | 30.764144   | 0.000000*       | 0.069523  | 45.118961  | 0.000000*     | 4.628908   |
| MEAN_HEIGH | 0.172041        | 0.002507 | 68.617612   | 0.000000*       | 0.002293  | 75.036465  | 0.000000*     | 3.001814   |
| ROAD_RATIO | 0.404105        | 0.048934 | 8.258247    | 0.000000*       | 0.037288  | 10.837438  | 0.000000*     | 1.200761   |
| VEGE_RATIO | -1.852686       | 0.031782 | -58.292848  | 0.000000*       | 0.038296  | -48.378291 | 0.000000*     | 2.243475   |
| PM10       | 0.376931        | 0.023108 | 16.311961   | 0.000000*       | 0.022479  | 16.768191  | 0.000000*     | 6.305819   |
| NOX        | -0.011272       | 0.003541 | -3.183068   | 0.001473*       | 0.002956  | -3.812642  | 0.000149*     | 6.964888   |

The Exploratory Regression Global Summary in the case of MRT explains that none of the models calculated based on 12 predictors got the passing mark to be considered as a robust model for prediction (*Table 24*). This could be due to the outliers, spatial autocorrelation, and non-linearity relationship between variables and response variables owing to the Jarque-Bera (JB) test result. The Global Moran's I test is the next step to diagnosing the cause of model failure as a result of spatial autocorrelation.

Table 24 MRT Criteria Passed

| Percentage of Search Criteria Passed |        |        |          |          |
|--------------------------------------|--------|--------|----------|----------|
| Search Criterion                     | Cutoff | Trials | # Passed | % Passed |
| Min Adjusted R-Squared               | > 0.50 | 4095   | 0        | 0.00     |
| Max Coefficient p-value              | < 0.05 | 4095   | 3414     | 83.37    |
| Max VIF Value                        | < 7.50 | 4095   | 2304     | 56.26    |
| Min Jarque-Bera p-value              | > 0.10 | 4095   | 0        | 0.00     |
| Min Spatial Autocorrelation p-value  | > 0.10 | 36     | 0        | 0.00     |

#### 6.1.4. Summary of Variables Significance

The significance of exploratory variables in MRT models is summarised in *Table 25*. CDSM, NDBI, MeanHeight, VSF and PM10 are among the highest significance variables. However, all the recorded scores are above 88 per cent, and all the predictors are significant to be considered in the models. Inconsistent relationships have also been seen in the MRT model for NDVI, SVF, DEM, and NOx.

Table 25 MRT Variables Significance

| Summary of Variable Significance |               |            |            |
|----------------------------------|---------------|------------|------------|
| Variable                         | % Significant | % Negative | % Positive |
| CDSM                             | 100.00        | 100.00     | 0.00       |
| NDBI                             | 100.00        | 0.00       | 100.00     |
| MEAN_HEIGHT                      | 100.00        | 0.00       | 100.00     |
| VEGE_RATIO                       | 100.00        | 100.00     | 0.00       |
| PM10                             | 100.00        | 0.00       | 100.00     |
| NDVI                             | 98.10         | 43.85      | 56.15      |
| SVF                              | 97.22         | 46.53      | 53.47      |
| DSMB                             | 96.58         | 85.25      | 14.75      |
| DEM                              | 95.75         | 53.37      | 46.63      |
| BSF                              | 95.51         | 21.53      | 78.47      |
| NOX                              | 93.02         | 52.05      | 47.95      |
| ROAD_RATIO                       | 88.57         | 16.41      | 83.59      |

#### 6.1.5. Spatial Autocorrelation (Global Moran's I) Test

ArcGIS Pro's Global Moran's I is used to test the spatial autocorrelation theory. The outcome is displayed in *Table 26*. Given the z-score of 367.38, there is a less than 1% likelihood that this clustered pattern could result from random chance. Global Moran's I spatial autocorrelation test for MRT showed a lower value for both z-score and Moran's Index, in which the MRT prediction is not statistically and spatially significant compared to LST.

Table 26 Global Moran's I Summary

|            | Moran's Index | Expected Index | Variance | z-score    | p-value  |
|------------|---------------|----------------|----------|------------|----------|
| <b>MRT</b> | 0.585688      | 5E-06          | 0.000003 | 367.379920 | 0.000000 |

## 6.2. Non-linearity by Neural Network: MRT

### 6.2.1. Architecture of MRT prediction model

The same model configuration was considered in case of MRT (mentioned in Chapter 5).

### 6.2.2. MRT NN Model Performance

The error histogram (*Figure 56*) for the trained NN-based MRT model demonstrates the normal distribution, where the majority of errors fall within the range of +1 and -1. The broad range of errors in the model indicates that MRT has a poor predictive capacity, and the zero-error line crossed the error bar with the greatest occurrences with errors ranging from 0 to 0.1298 (on a scale of normalized data).

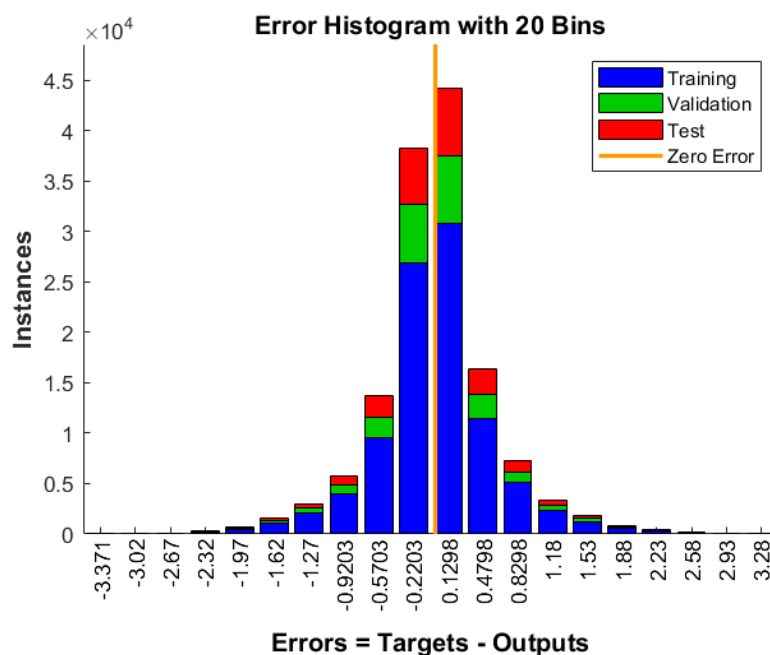


Figure 56 left: Error histogram

On the other hand, the residual plots for three partitions of the trained model in *Figure 57* indicate the best-fit correlation line, which is almost close to a 35-degree angle. In other terms, the overall residual plot demonstrates the poor correlation between observed (true) values and predicted values due to the high error rate and residual.

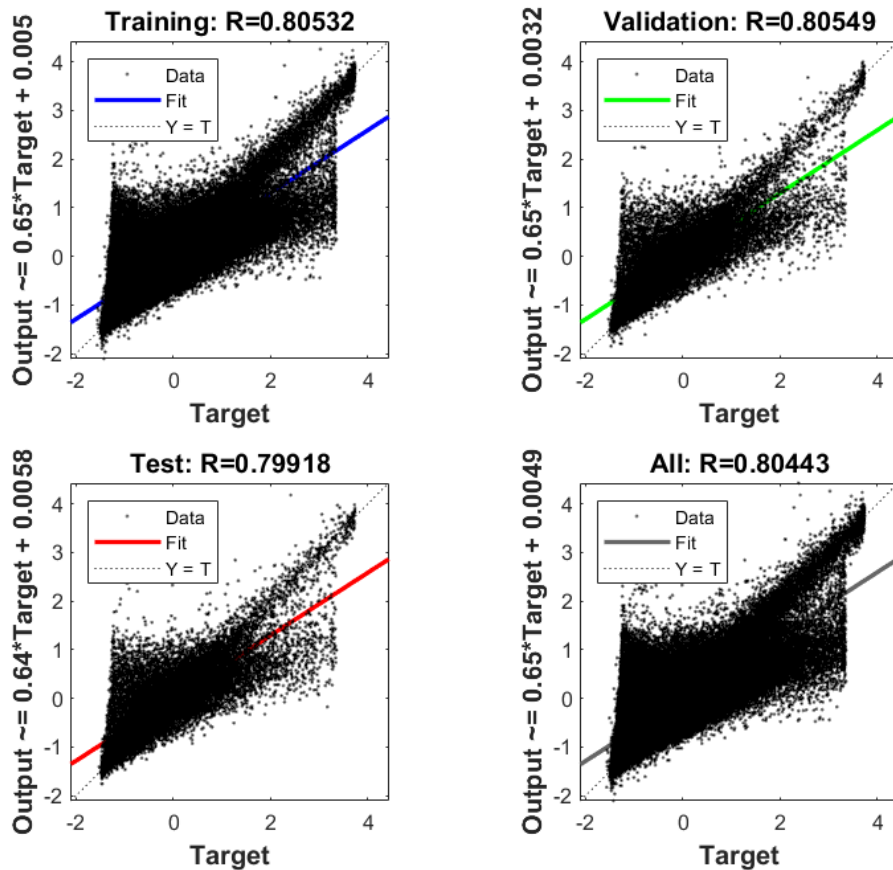


Figure 57 Residual plots for the NN based MRT model

The 2D histogram shown in *Figure 58* clarifies that the highest proportion of observed values is located on the 45-degree fit line; however, the bias could be found for the outliers where the model underestimates and overestimates the MRT for the highest (above 50) and lowest (below 40) range temperature observed.

### 6.2.3. Sensitivity Analysis

The result of sensitivity analysis shows the following key points based on the importance of the predictor through the elimination of each predictor from the training process and the elimination of predictors from one data source (Appendix VII). SVF has the highest significance among the predictors due to the greatest difference in coefficient of determination. The essential elements in forecasting the LST based on secondary data are NDBI, NDVI and RAR. The other noteworthy point is that excluding CDSM and DSM from the forecast improved MRT prediction. *Figure 59* displays the importance of SVF, NDVI, NDBI, and RAR in predicting

over 70% of the MRT. Additionally, sensitivity analysis revealed that the LiDAR dataset, LULC, and air pollution variables were the most notable datasets for MRT prediction.

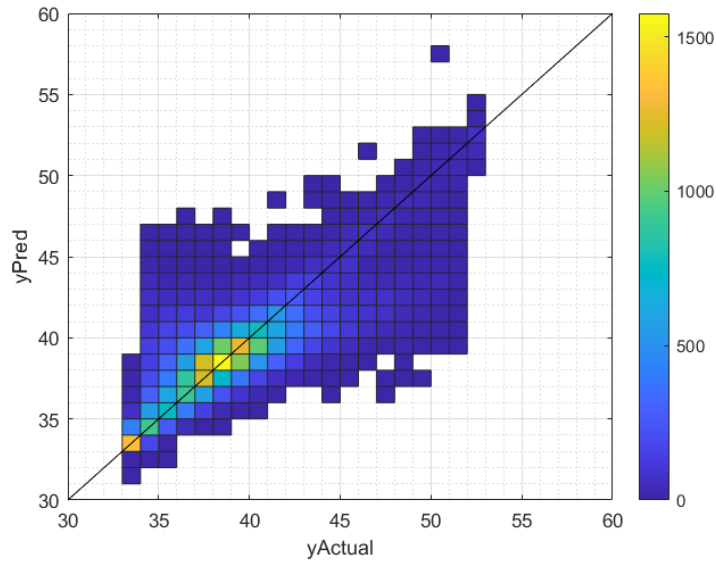


Figure 58 Distribution of biased prediction in MRT model

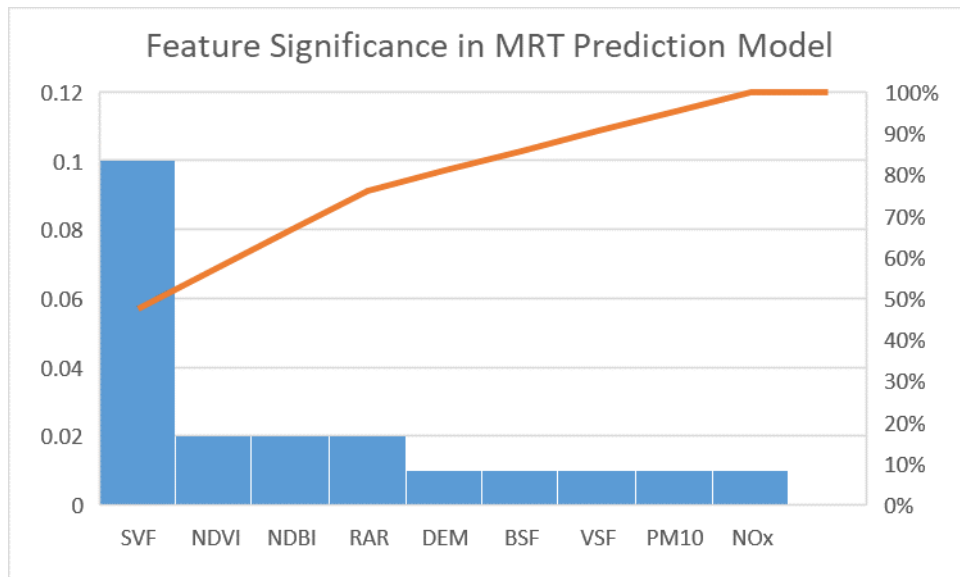


Figure 59 Feature Significance in MRT Prediction Model from sensitivity analysis





## CHAPTER 7: DISCUSSION AND REFLECTION

This chapter covers further interpretation and comments on the simulation and prediction model findings in order to investigate the significance of discoveries to answer the project's main questions concerning what was previously known. In response to the critical questions of the current study, this chapter was set up to demonstrate how secondary data on thermal comfort proxies and urban design elements can be used to make predictions.

### 7.1. LST and MRT Prediction depending on Secondary Data

The rising popularity and demand for remote sensing data, along with geo-processing approaches in decision-making, are mostly due to a shortage of reliable data at the fine scale and an expensive data collection procedure (Do Nascimento *et al.*, 2022). The effective use of such readily available datasets might create opportunities for addressing environmental concerns by policymakers.

#### 7.1.1. How Regression models work in the case of MRT and LST

The GLR results demonstrate the value of  $R^2$  (.64) which is not significant to be considered a reliable and robust linear model in the prediction of LST, while in the case of MRT, the low value of  $R^2$  (.37) shows other factors might be needed to be considered in finding a good fit model with the significant contribution (*Table 27*). The MRT's poor performance could be a result of its strong dependency on radiation and meteorological factors in computation (Guo *et al.*, 2020). In addition, non-linear regression could be a better fit for the chosen predictors. Overall, the GLR model had less explanatory power than the ANN model since it had a lower capability of predicting data with higher levels of variance; nevertheless, the predicted LST range for the scenarios was significantly closer to the RS-LST. In all conditions and cases, relying on R-squared in model performance was owing to its scale less value spanning from 0 to 1. The high score was only given if the models could predict a larger fraction of the variation in the response variable (Chicco, Warrens and Jurman, 2021). Other error metrics, including MAE, MSE, and RMSE, were attributed to finding the better model under the same settings (e.g., for all LST NN models).

*Table 27 Comparison of GLR in LST and MRT models*

| Model No. | Model Type | Predictors | R-Squared | Adjusted R-Squared | Akaike's Information Criterion (AICc) | Probability |
|-----------|------------|------------|-----------|--------------------|---------------------------------------|-------------|
|-----------|------------|------------|-----------|--------------------|---------------------------------------|-------------|

|               |                   |               |          |          |                |          |
|---------------|-------------------|---------------|----------|----------|----------------|----------|
| <b>1. LST</b> | GLR(Gaussian/OLS) | 12 Predictors | 0.642756 | 0.642734 | 732447.003358s | 0.000000 |
| <b>2. MRT</b> | GLR(Gaussian/OLS) | 12 Predictors | 0.376488 | 0.376450 | 1108792.066108 | 0.000000 |

Additionally, the GWR model used in the LST scenario performed better when several predictors were taken into account. The model with 12 predictors failed to develop due to the redundancy and correlation between the input variables. To avoid redundancy, working on a larger grid size might be helpful as the importance of different resolutions in GWR studies was examined by Zhao, Ren and Tan (2018) and Luo and Peng (2016). The grid of 90 m was used in this study to see the difference, although the outcomes were biased. GWR was not applicable for MRT because of the lower value of Moran's I. It is worth mentioning that MRT mostly relies on radiation, and geographically close predictor values may not anticipate its variation.

Overall, GWR Regression performed as a powerful prediction model based on each independent and the response variable (LST). Due to the high volume of grids (193810 cells), ignoring the redundancy and multicollinearity between the independent variable made it impossible to develop one generalised model based on all the predictors. A few models were run by bringing 5-6 predictors, resulting in over 90% of total data can be predicted with satisfactory accuracy. Luo and Peng (2016) also evaluated the output of combined GWR with reasonable accuracy.

Both LST and MRT were investigated using the same optimised ANN model. The models above indicated that LST is more likely to be predicted than MRT using the identified 12 independent parameters. Approximately 80% of the variance in LST could be predicted using an ANN model with a Mean Absolute Error of 1.1750 °C. The MRT model estimated 65% of untrained data with a Mean Absolute Error of 1.5144 °C (*Table 28*).

*Table 28 Comparison of error metrics in MRT and LST prediction models*

| Model No.             | Model Type | Predictors    | Error Metrics |        |        |           |
|-----------------------|------------|---------------|---------------|--------|--------|-----------|
|                       |            |               | RMSE          | MSE    | MAE    | R-squared |
| <b>1. LST</b>         | ANN        | 12 Predictors | 1.5157        | 2.2974 | 1.1750 | 0.8004    |
| <b>2. MRT (16:00)</b> | ANN        | 12 Predictors | 2.2227        | 4.9405 | 1.5144 | 0.6597    |

The weak performance of the MRT model in both linear and non-linear regressions could be the lack of significant predictors to provide a robust prediction. It is expected that including fine-resolution meteorological data will result in a superior prediction model, as Shah, Pandit, and Gaur (2022) revealed the substantial capability of microclimatic factors in forecasting UTCI and PET using an ANN algorithm. MRT is being calculated through Eq. (1), which reflects

the impacts of radiation, temperature, humidity, and wind on the level of human thermal comfort (Bröde *et al.*, 2012). Another explanation for the poor performance might be the high volatility of MRT during the day. The spatial distribution of MRT shows that during peak temperature hours, the city undergoes cold and hot patches, causing the prediction model to fail to function on dispersed (non-normalized) data. Training various models for each hour is recommended to find the most robust prediction model fully based on the selected predictors.

The previous ANN prediction models in UHI and thermal comfort field could be compared with the outcome of this study. Integrating the terrain factor (TF) significantly improved LST forecasts made using an ANN model for substantial terrain variation areas, with average RMSE decreasing from 1.26 to .90 °C and R<sup>2</sup> increasing from .74 to .81 (Equere *et al.*, 2021). Chan and Chau (2019) found that the inclusion of psychological factors in the prediction of outdoor thermal comfort along with microclimate parameters in the ANN model can positively improve the performance (R-value) by .33 and .3 in summer and winter, respectively. The outcome of a study by Shah, Pandit and Gaur (2022) showed a high coefficient of correlation of .87 to .99 and a low RMSE of .11 to 1.72, demonstrating the ANN model's exceptional predictive accuracy for UTCI prediction.

### 7.1.2. Prediction Models Performance in Scenarios

The Neural Network prediction model on LST showed a modest influence at the city level, which opens the fields for further investigations. Three scenarios were developed for the ANN model to predict the changes when vegetation and built-up are considered in different configurations. Two vacant lands were selected to discover the changes in detail at street level. One showed the positive impacts of applying 100% vegetation, while the other site experienced an insignificant decrease (less than .06 °C) of LST as a result of the same portion of vegetation. In the 100% built-up cases predicted, LST rose in one context while dropping in the other. There are a few reasons behind the performance of the trained ANN model, including the data availability and data quality which play a pivotal role in machine learning prediction, which has been raised by (Gobakis *et al.*, 2011) in the development of the ANN model for UHI prediction. Gathering spatiotemporal data from multiple sources was considered a challenge throughout the procedure. Working with coarse resolution and resampling, for example, resulted in some data being missing or carrying bias throughout the workflow. In addition, even though both locations are near the city centre, they are in quite distinct urban settings. One is in a low-density neighbourhood, while the other is in a high-density region.

## 7.2. Urban Design Factors in LST and MRT Prediction

### 7.2.1. Significant Parameters in Prediction

#### 7.2.1.1. LST

Pearson correlation coefficient proved the contribution of each variable for each response variable of this study. Under Pearson correlation of LST, NDBI, NDVI and VSF were the most crucial factors with a correlation of .696, -.637, and -.618, respectively, which was aligned with previous studies (Guha *et al.*, 2018; Malik, Shukla and Mishra, 2019; Goldblatt *et al.*, 2021).

In linear regression, the strongest predictors are NDBI, MeanHeight, CDSM, BSF, and PM10 for their consistent importance and steady connection with LST. NDBI was identified as a significant predictor in the GWR model as well.

It is believed that ANN is a black box that does not provide the contribution of each factor in the regression model, the same as linear regression models. The sensitivity analysis provided the most contributing factors in the model by the elimination of each factor by the difference of error metrics. In the case of LST, NDBI got the highest error difference in the elimination process. In the second place, DEM, DSM, PM10, NO<sub>x</sub>, NDVI, BSF and SVF got the same contributions. The overall difference between the changes in error metrics was not that significant to say there is a dramatic difference between what has been received from each round of sensitivity analysis.

#### 7.2.1.2. MRT

MRT Pearson correlation (16:00) revealed SVF as the most significant correlated component, as shown in a study by Gál and Kántor (2020). For the 13:00 (peak MRT), the CDSM, NDBI and NDVI were the principal factors indicating the impacts of vegetation and built-up. The difference is based on the variation of MRT during the day, which is mainly caused by the changes in radiation on heat stress (Thorsson *et al.*, 2017). The minimal impact of RAR was not stable due to both positive and negative correlations with MRT at 13:00 and 16:00. Instability correlation was seen in the case of DSM as well. The instability of DSM and RAR needs further study to make a clearer outcome.

The most powerful predictors of MRT (16:00) in linear regression are CDSM, NDBI, MeanHeight, VSF, and PM10.

In the ANN back box of the MRT model, SVF was by far the best factor in prediction. Research by Xie *et al.* (2022) also portrayed the importance of SVF in ANN model prediction of MRT. Other notable factors are NDBI, NDVI, and RAR. It is important to mention that CDSM and DSM elimination showed decreasing errors for 16:00, which might be accounted for the impacts of shading after the peak time. In other words, the factors which have been found most significant in a specific hour in regression can have negative impacts on the other hours due to their impact on shading and exposure to the radiation.

### 7.2.2. Significant Dataset in Prediction

In the non-linear prediction model of LST, the RSD dataset showed the best contribution compared to the other source of data which has been outperformed in the study of Goldblatt *et al.* (2021). On the other hand, the MRT ANN model showed the LiDAR dataset, LULC along with air pollution parameters were the most remarkable datasets for the prediction of MRT (at 16:00 PM). According to Yilmaz, Irmak and Qaid (2022), air pollution has an adverse impact on thermal comfort and may therefore be more important to consider in MRT prediction than LST.

## 7.3. LST as Proxy of Thermal Comfort Studies?

Land Surface Temperature, being one of the primary elements contributing to ambient temperature, plays a vital role in human lives, which is most impacted by unpredictability in air temperature variations (Irmak, Yilmaz and Dursun, 2017). The two recent studies by Kelly Turner *et al.* (2022) and Goldblatt *et al.* (2021) have tried to illustrate the potential use of LST as a proxy of thermal comfort at the local and hyper-local levels in the hot semi-arid area of Tuscan (Arizona) and arid region of Jeddah (Saudi Arabia). To generalise the significance of LST in thermal comfort, they argued that heat mitigation methods might have diverse effects in different contexts, which calls for more studies to illustrate the clearer result.

The findings from LST-ANN reveal that increasing urban vegetation has a minor influence on reducing heat stress at the surface level. Vacant land conversion to fully green space can decrease the LST by .177 °C, while the built-up area could rise by 1.289 °C. For the same location, the UTCI modifications were seen to drop 3K by greening strategy (moving from Strong-TS to Moderate-TS and No-TS). The shading effect of surrounding high-rise buildings reduced thermal stress as the built-up area increased. Compared to other studies' findings, the impact of vegetation on achieving climate comfort conditions was shown to be modest, which was discovered by Stepani and Emmanuel (2022). Furthermore, shading through plants

or buildings is a great way to manage heat stress at the street level, which improves human thermal comfort by blocking solar exposure (Ketterer and Matzarakis, 2014). Under the simulated condition, the effect of buildings' shading was stronger than the vegetation in site A, which is located in denser urban settings surrounded by tall buildings. The results show that the LST mitigation strategy does not necessarily lead to thermal stress mitigation (Martilli, Krayenhoff and Nazarian, 2020), where the correlation between LST and MRT proved this. Enormous cooling spots are produced by the shadows of high-rise structures, which may have stronger impacts than vegetation and greenery. Furthermore, it is impractical in some regions of the city to correspond to thermal comfort using the same heat mitigation strategies (Stepani and Emmanuel, 2022).

However, the weaker impact of greening vacant land in reducing LST compared to intermediate and worst scenarios appeared to have less thermal stress by microclimate modelling. The contrast has been seen in the role of vegetation and built-up in two selected vacant lands. In site B, the vegetation decreased LST by .057, while the combination of both built-up and greenery ended with decreasing the LST by .281. It is worth mentioning that the worst scenario operated to mitigate LST by .217 (better than the best case). The UTCI index for this location reveals that vegetation had a beneficial effect on lowering UTCI by more than 3.39K at street canyon. In contrast, the worst-case scenario caused UTCI to rise to 4.61K at the street level. Due to the higher rate of openness in site B, surrounded by the tallest building of 18 meters, the buildings' shading appeared to have less impact than the vegetation under thermal stress. As a case in point, a tall building in a compact urban setting might have a better shading effect at a street level compared to a tree at the peak of solar radiation. It resulted in research conducted in Jakarta (Stepani and Emmanuel, 2022), where the LST favourable strategy in heat mitigation is a combination of building and vegetation together, and the UTCI responds positively to the greenery intervention.

The negligible vegetation influence result for Site B scenarios from the ANN model might be because of the coarse resolution of RS-LST (30 m), which was found that underestimates the high and overestimates the low LST (Vanos *et al.*, 2016; Kelly Turner *et al.*, 2022) and comparison of LST changes from both ANN and microclimate simulation for scenarios justified this fact. Additional parameters that apply to street orientation, optimisation of street canyon layout, and urban form configuration would be useful in urban heat mitigation (Emmanuel, 2021) and incorporating them in training ANN models are needed to clarify this distinction in results. It was found that the east-west canyon receives more solar radiation than the north-south canyon, according to calculations of irradiance on the various canyons and has the largest impact on the microclimate (Ketterer and Matzarakis, 2014; Taleghani *et al.*, 2021). In addition, street orientation, along with building density, can affect the wind velocity and

radiation angles on the surfaces and open areas (Thorsson *et al.*, 2017), which leads to minimizing the frequency of high MRT and the spatial expansion of hot spots. Furthermore, the efficient tree orientation at the street level was attributed to managing shading and cooling effect (Chatzidimitriou and Yannas, 2017).

A comparison of predicted and simulated LST showed a difference in the effect of vegetation and built-up role in surface temperature. In light of the fact that the simulated LST in ENVI-met is not carried into consideration at the same time as RS-LST, the general trend of LST changes from ENVI-met simulation contrasted with the results of the ANN model. Additionally, the building footprint, which typically has the lowest LST, is not taken into account by ENVI-met when calculating LST. It is worth noting that in the ANN model, each cell (grid of 30 m) was predicted without the influence of its neighbouring cells, while the ENVI-met computation of surface temperature is continuous and small changes in one part might affect the whole domain.





## CHAPTER 8: CONCLUSIONS AND RECOMMENDATIONS

Apart from reducing energy usage and reliance on fossil fuels, city development methods can help to reduce heat stress caused by climate change. The contributing parameters to outdoor thermal comfort (OTC) and urban heat derived from historical and secondary data could be divided into five main categories: Remotely sensed derived (RSD): Landsat 8 dataset; Urban morphology: derived from LiDAR, LULC; Vector-based at the level of building footprint; and Air quality factors. Among all the factors, SVF was by far the best factor in the prediction of MRT followed by vegetation-based predictors, including CDSM and VSF; however, it was found that the significant parameters depend on the time of the day. For the peak hour of MRT (13:00), CDSM and NDBI were recognized as significant variables. It reveals the importance of shading effect by greenery and build-up area in MRT mitigation. On the other hand, NDBI and NDVI were identified as the most significant predictors of LST.

It was planned to train and evaluate the performance of various prediction models for outdoor thermal comfort indices. Generalized Linear Regression, Geographically Weighted Regression, Pearson Correlation Coefficient, as well as Artificial Neural Network were run. Given its unbiased results and reduced residuals, the nonlinear model offers a better fit. However, the outliers were predicted with higher error. GWR was found as a substitute for GLR in LST prediction, although the redundancy and multicollinearity of predictors should be resolved for better performance.

The microclimate modelling in ENVI-met for two selected sites in the city centre shed light on comparing the mitigation strategies for LST and thermal comfort. The concept of generating radical situations was to comprehend the quick changes and effects of Glasgow's rising vegetation, as well as how these modifications affect heat stress and people's comfort levels. The ML prediction of LST was validated by ENVI-met, which could somewhat enhance overall thermal comfort. It was intended to interpolate UTCI changes to the whole city based on the results from two sites, but MRT simulation made it clear that even minor changes and improvements in LST mitigation can cause different trends in MRT and UTCI improvement, and a complicated model is needed to generalize the changes.

### 8.1. Summary of Findings

The present study's primary findings are as follows:

- The strongest predictor for LST is the NDBI variable in both linear and nonlinear regressions. Other significant predictors in linear relationship are MeanHeight, CDSM,

BSF, and PM10 for their consistent significance and stable relationship with LST. On the other hand, NDVI, DEM, DSM, SVF, BSF and air pollution variables are performing better in a nonlinear manner.

- The most powerful predictors of MRT (non-peak time of 16:00) in a linear regression are CDSM, NDBI, MeanHeight, VSF, and PM10, whereas the most contributing variables in the nonlinear model are SVF, NDBI, NDVI, and RAR.
- ANN as a powerful model with a wider range of applications, needs high-quality data with precise accuracy to be able to create a valid and generalized prediction. Aside from how accurate is the training process, there is no guarantee that the estimation of new data would be exactly the same under the same realistic settings. However, the outcomes will be expressed as Inhibitory actions.
- Furthermore, the amount of heat that people can actually feel outside is not correctly reflected by the land surface temperature that is measured using satellite data. This is because other factors, including shade, wind speed, and relative humidity, can have an impact on how much heat people are exposed to. Because of the complexity of thermal stress, particularly at the level of human comfort, heat mitigation solutions must be researched in many contexts and locations for generalizing the mitigation approach.
- The comparison of the three scenarios shows that when more lands are developed through construction, the overall UTCI index will rise. In other terms, the surrounding area will experience more severe heat stress following physical development unless their shading effect appears to be strong.
- The impact of shadowing from plants and buildings, for instance, could alter thermal comfort depending on the area's compactness and openness in Glasgow's central district. Heat mitigation measures at the level of lowering the surface temperature do not always meet human thermal comfort. The association between LST and MRT has made it clear that it is impossible to establish a direct connection between them.

## 8.2 Limitations

In light of the study's intricate approach, the following is a list of the study's limitations and shortcomings in data and the process.

1- In order to remove biases and develop an accurate and effective machine learning-based prediction model, data quality is crucial.

- The lack of continuous and heterogeneous data with an appropriate resolution for LST owing to a specific spatiotemporal resolution of Landsat 8 (Zhao, Ren and Tan, 2018; Anderson *et al.*, 2021). A single date, the closest time to the heatwave was selected, which was taken before noon which might have not recorded the heatwave impact on surface temperature level.
- Aside from the Landsat 8 dataset, the low resolution and unclassified point clouds of the LiDAR dataset were challenging and time-consuming.
- The coarse resolution of air quality data (1 km) for the city of Glasgow, along with unavailability of PM<sub>2.5</sub> data to cover the whole area of interest, made it impossible to consider all fine particulate matters in the air that is a concern for human's health.
- Lack of a homogeneous urban block zone for Glasgow ended in considering the grid size of 30 and 90 m, which causes data redundancy and multicollinearity.
- The meteorological data considered for both ENVI-met and SOLWEIG models were chosen from two dissimilar sources (one from the fifth generation ECMWF reanalysis for the global climate with a resolution of 30 km; the other from Paisley meteorological station).
- The unavailability of some key factors like the diverse types of vegetation along with inaccessibility of urban geometry parameters(i.e. H/W ratio) made the analysis incomplete in some parts.

## 2- Complexity of Applications and Models

- SOLWEIG model is a developing software which is being upgraded through the online platform by connecting its end-users to developers. The application bugs made the process of this project longer through modelling the thermal comfort simulation.
- SOLWEIG is suited for thermal comfort and heat management research that focuses on the hottest time of the day. However, because the model has a likelihood to exaggerate  $T_{mrt}$  in the shaded and underestimate it in the sun, heat mitigation studies may underestimate the influence of shading (Gál and Kántor, 2020). Due to the clustering impact of the result for peak time—hotspots versus cold spots, which are not normally distributed and suited for running regression models—the hottest hour was not taken into account in the predictive models of the study.
- ENVI-met, as a complex and precise microclimate, took a long time to simulate the thermal comfort index in which the time limit could not allow running UTCI simulation for more sites in various parts of Glasgow.

## 3- Time-limit

- Hourly MRT prediction models could not be run due to the high number of occurrences (198310) and the lengthy training times for ML models in MATLAB.

### 8.3. Recommendation for Further Studies

Due to the nature of this study on prediction modelling and microclimate simulation, further research is recommended through the following points.

#### **Training new prediction models under following conditions:**

- Incorporating the outcomes of the ENVI-met microclimatic simulation into the ANN prediction model to assess the effectiveness of new predictors
- As the results showed the controversial impacts of the built-up area led to heat mitigation for LST and UTCI index, one recommendation should be the examination of heat mitigation at the scale of facade and surfaces for a built-up area to be included in the process of prediction. The façade material was chosen from the default category, which is one of the key parts to be considered.
- It is recommended to consider street orientation as a predictor which has shown notable impacts on MRT and thermal comfort by affecting the wind velocity and radiation angles on the surfaces and open spaces.
- Diverse kinds of vegetation can reveal different trends in reducing heat stress, and these trends should be investigated as an independent variable when training the ANN model.
- The air temperature should be considered in the training model due to its strong relationship with thermal sensation, specifically helpful in windy temperate metropolitan (Oertel, Emmanuel and Drach, 2015), which is recommended to be collected in fine resolution. For instance, the method of crowdsourcing by citizens in the study of Zumwald *et al.* (2021) with reliable accuracy could be valuable in the prediction process.
- The appropriate comparison between simulated LST by ENVI-met and the prediction of the ANN model was not possible due to the different computation processes for a target land. The ANN model was powerful in predicting the inside of the vacant land LST and the less concentration on its surroundings because of the small grid of 30 m. It is recommended to train a model at a large scale which consists of street orientation and heterogeneous urban blocks.

- ANN has been chosen in this study due to its strong performance in previous studies. The suggestion is to follow the same workflow under different ML regression models to find the impacts of other algorithms in thermal comfort prediction.

### **Model Validation**

- The same model configuration should be applied in similar cities under the same data source with a temperate climate to compare the performance of the prediction model
- Field study could provide a ground-based validation dataset in the simulation and prediction process.

### **Data collection**

- Centralized data collection system should be designed for outdoor thermal comfort considering air quality, urban geometry, and remotely sensed derived datasets,. Further research should be done on creating a data sourcing agency to supervise this data hub for future usage by decision-makers and scholars.
- Although it is time-consuming and complicated to have people contribute to share their sense and experience of thermal stress, a wide range of studies on outdoor thermal comfort have attempted to involve human contact through survey and fieldwork studies. It is suggested to include people's perceptions of heat stress in the process to establish a more comprehensive framework.

## REFERENCE

Afrakhteh, R. *et al.* (2016) 'Evaluating the strategy of integrated urban-rural planning system and analyzing its effects on land surface temperature in a rapidly developing region', *Habitat International*, 56, pp. 147–156. doi: 10.1016/j.habitatint.2016.05.009.

Alavipanah, S. *et al.* (2015) 'The role of vegetation in mitigating urban land surface temperatures: A case study of Munich, Germany during the warm season', *Sustainability (Switzerland)*, 7, pp. 4689–4706. doi: 10.3390/su7044689.

Alexander, L. V. and Arblaster, J. M. (2008) 'Assessing trends in observed and modelled climate extremes over Australia in relation to future projections', *International Journal of Climatology*, 29, pp. 417–435. doi: <https://doi.org/10.1002/joc.1730>.

Alves, E. D. L. and Lopes, A. (2017) 'The urban heat island effect and the role of vegetation to address the negative impacts of local climate changes in a small Brazilian City', *Atmosphere*, 8(2). doi: 10.3390/atmos8020018.

American society of heating refrigerating and air-conditioning engineers (ASHRAE) (2010) *ANSI/ASHRAE Standard 55-2010, Thermal Environmental Conditions for Human Occupancy*, American society of heating refrigerating and air-conditioning engineers. Atlanta, GA.

Anderson, V. *et al.* (2021) 'Technological opportunities for sensing of the health effects of weather and climate change: a state-of-the-art-review', *International Journal of Biometeorology*, 65, pp. 779–803. doi: 10.1007/s00484-020-02063-z.

Aram, F. *et al.* (2020) 'Urban heat resilience at the time of global warming: evaluating the impact of the urban parks on outdoor thermal comfort', *Environmental Sciences Europe*, 32(1). doi: 10.1186/s12302-020-00393-8.

Arnell, N. W. *et al.* (2021) 'Changing climate risk in the UK: A multi-sectoral analysis using policy-relevant indicators', *Climate Risk Management*, 31(December 2020), pp. 1–25. doi: 10.1016/j.crm.2020.100265.

Arnfield, A. J. (2003) 'Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island', *International Journal of Climatology*, 23(1), pp. 1–26. doi: 10.1002/joc.859.

Ashtiani, A., Mirzaei, P. A. and Haghghat, F. (2014) 'Indoor thermal condition in urban heat island: Comparison of the artificial neural network and regression methods prediction', *Energy and Buildings*, 76, pp. 597–604. doi: 10.1016/j.enbuild.2014.03.018.

BBC Scotland News (2018) *How Hot?*, *Twitter*. Available at: [https://twitter.com/BBCScotlandNews/status/1012331862956609536?ref\\_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwtterm%5E1012331862956609536%7Ctwgr%5E%7Ctwtcon%5Es1\\_&ref\\_url=https%3A%2F%2Fwww.theguardian.com%2Fuk-news%2F2018%2Fjun%2F28%2Fuk-heatwave-glasgow-bak](https://twitter.com/BBCScotlandNews/status/1012331862956609536?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwtterm%5E1012331862956609536%7Ctwgr%5E%7Ctwtcon%5Es1_&ref_url=https%3A%2F%2Fwww.theguardian.com%2Fuk-news%2F2018%2Fjun%2F28%2Fuk-heatwave-glasgow-bak).

Blazejczyk, K. *et al.* (2012) 'Comparison of UTCI to selected thermal indices', *International Journal of Biometeorology*, 56(3), pp. 515–535. doi: 10.1007/s00484-011-0453-2.

Blazejczyk, K. *et al.* (2013) 'An introduction to the Universal thermal climate index (UTCI)', *Geographia Polonica*, 86(1), pp. 5–10. doi: 10.7163/GPol.2013.1.

Bokaie, M. *et al.* (2016) 'Assessment of Urban Heat Island based on the relationship between land surface temperature and Land Use / Land Cover in Tehran', *Sustainable Cities and Society*, 23, pp. 94–104. doi: 10.1016/j.scs.2016.03.009.

Botchkarev, A. (2018) 'Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology', *Interdisciplinary Journal of Information, Knowledge, and Management*, 14, pp. 45–79. doi: 10.48550/arXiv.1809.03006.

Bröde, P., Fiala, D., *et al.* (2012) 'Deriving the operational procedure for the Universal Thermal Climate Index ( UTCI )', *International Journal of Biometeorology*, (56), pp. 481–494. doi: 10.1007/s00484-011-0454-1.

Bröde, P., Krüger, E. L., *et al.* (2012) 'Predicting urban outdoor thermal comfort by the Universal Thermal Climate Index UTCI — a case study in Southern Brazil', *International Journal of Biometeorology*, 56, pp. 471–480. doi: 10.1007/s00484-011-0452-3.

Brown, P. M. B. L. C. and Hambley, D. F. (2002) *Statistics for Environmental Engineers*. Second, *Environmental & Engineering Geoscience*. Second. LEWIS PUBLISHERS. doi: 10.2113/8.3.244.

Chan, S. Y. and Chau, C. K. (2019) 'Development of artificial neural network models for predicting thermal comfort evaluation in urban parks in summer and winter', *Building and Environment*, 164(April). doi: 10.1016/j.buildenv.2019.106364.

Chatzidimitriou, A. and Yannas, S. (2017) 'Street canyon design and improvement potential for urban open spaces; the influence of canyon aspect ratio and orientation on microclimate and outdoor comfort', *Sustainable Cities and Society*, 33(May), pp. 85–101. doi: 10.1016/j.scs.2017.05.019.

Chen, Y., Zheng, B. and Hu, Y. (2020) 'Numerical Simulation of Local Climate Zone Cooling Achieved through Modification of Trees , Albedo and Green Roofs — A Case Study of Changsha , China', *Sustainability*, 12(2752), pp. 1–23.

Chicco, D., Warrens, M. J. and Jurman, G. (2021) 'The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation', *PeerJ Computer Science*, 7, pp. 1–24. doi: 10.7717/PEERJ-CS.623.

Climate Data (2022) *CLIMATE GLASGOW*, *Climate Data Organization*. Available at: <https://en.climate-data.org/europe/united-kingdom/scotland/glasgow-20/#climate-table>.

Dan, S. *et al.* (2010) 'Evolution of urban heat island effect in middle and small cities in Sichuan Basin', *2010 2nd IITA International Conference on Geoscience and Remote Sensing, IITA-GRS 2010*, 1, pp. 158–160. doi: 10.1109/IITA-GRS.2010.5603103.

Digimap Edina (2020) *Land and Height data* [dataset], *ORDNANCE SURVEY*. Available at: <https://digimap.edina.ac.uk/roam/download/os>.

Digimap Edina (2022a) *Boundary and Location Data* [dataset], *ORDNANCE SURVEY*. Available at: <https://digimap.edina.ac.uk/roam/download/os>.

Digimap Edina (2022b) *VERISK Collection: Land Use and Land Cover* [dataset], *VERISK Collection DIGIMAP*. Available at: <https://digimap.edina.ac.uk/verisk>.

Dimoudi, A. *et al.* (2013) 'Investigation of urban microclimate parameters in an urban center', *Energy and Buildings*, 64, pp. 1–9. doi: 10.1016/j.enbuild.2013.04.014.

Duarte, D. H. S. *et al.* (2015) 'The impact of vegetation on urban microclimate to counterbalance built density in a subtropical changing climate', *Urban Climate*, 14, pp. 224–239. doi: 10.1016/j.uclim.2015.09.006.

Eastin, M. D. *et al.* (2018) 'Temporal Variability of the Charlotte ( Sub ) Urban Heat Island', *Journal of Applied Meteorology and Climatology*, 57(1), pp. 81–102. doi: 10.1175/JAMC-D-17-0099.1.

EEA (2020) *Urban Atlas 2018* [dataset], *EUROPEAN ENVIRONMENT AGENCY*. Available at: <https://land.copernicus.eu/local/urban-atlas/urban-atlas-2018>.

Emmanuel, R. *et al.* (2021) *Climate Proofing Glasgow Adaptation Strategies for Urban Overheating*. Edited by R. Emmanuel. Glasgow: Glasgow Caledonian University.

Emmanuel, R. (2021) 'Urban microclimate in temperate climates: a summary for practitioners', *Buildings and Cities*, 2(1), pp. 1–9. doi: 10.5334/bc.109.

Emmanuel, R. and Baker, K. (2012) *Carbon Management in the Built Environment*. Routledge.

Equere, V. *et al.* (2021) 'Integration of topological aspect of city terrains to predict the spatial distribution of urban heat island using GIS and ANN', *Sustainable Cities and Society*, 69. doi: 10.1016/j.scs.2021.102825.

ESRI (2018) *Interpreting GWR results, ArcGIS Desktop*. Available at: <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/interpreting-gwr-results.htm>.

ESRI (2022a) *Generalized Linear Regression (GeoAnalytics), ArcGIS Pro geoprocessing tool reference*. Available at: [https://pro.arcgis.com/en/pro-app/latest/tool-reference/big-data-analytics/generalized-linear-regression.htm#:~:text=Summary-,Summary,and count \(Poisson\) models](https://pro.arcgis.com/en/pro-app/latest/tool-reference/big-data-analytics/generalized-linear-regression.htm#:~:text=Summary-,Summary,and count (Poisson) models).

ESRI (2022b) *What is empirical Bayesian kriging?, ArcGIS Pro help*. Available at: <https://pro.arcgis.com/en/pro-app/2.8/help/analysis/geostatistical-analyst/what-is-empirical-bayesian-kriging-.htm>.

Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2002) *Geographically weighted regression: the analysis of spatially varying relationships*, John Wiley & Sons. John Wiley &



Sons. Available at:  
[https://www.researchgate.net/publication/269107473\\_What\\_is\\_governance/link/548173090cf22525dcb61443/download%0Ahttp://www.econ.upf.edu/~reynal/Civilwars\\_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625](https://www.researchgate.net/publication/269107473_What_is_governance/link/548173090cf22525dcb61443/download%0Ahttp://www.econ.upf.edu/~reynal/Civilwars_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625).

Frost, J. (2019) *Regression Analysis: An Intuitive Guide for Using and Interpreting Linear Models*. First. Jim Frost. doi: 10.1007/978-3-030-36323-9\_26.

Gál, C. V. and Kántor, N. (2020) 'Modeling mean radiant temperature in outdoor spaces, A comparative numerical simulation and validation study', *Urban Climate*, 32(April 2019), p. 100571. doi: 10.1016/j.uclim.2019.100571.

Givoni, B. *et al.* (2003) 'Outdoor comfort research issues', *Energy and Buildings*, 35(1), pp. 77–86. doi: 10.1016/S0378-7788(02)00082-8.

Glasgow city council (2020) *Glasgow's Climate Plan Our Response to the Climate and Ecological Emergency*, Glasgow city council. Glasgow.

Glasgow City Council (2017a) *Glasgow City Council Strategic Plan 2017-2022*. Glasgow.

Glasgow City Council (2017b) *Glasgow City Development Plan*. Glasgow. Available at: <https://www.glasgow.gov.uk/index.aspx?articleid=16186>.

Glasgow City Council (2021a) *City Centre Strategic Development Framework*. Glasgow. Available at: <https://www.glasgow.gov.uk/CHttpHandler.ashx?id=53335&p=0>.

Glasgow City Council (2021b) *Climate Emergency Implementation Plan*, Glasgow City Council. Glasgow. Available at: <https://www.glasgowconsult.co.uk/Project/861>.

Glasgow City Council (2022) *Glasgow climate adaptation plan 2022 – 2030*. Glasgow.

Gobakis, K. *et al.* (2011) 'Development of a model for urban heat island prediction using neural network techniques', *Sustainable Cities and Society*, 1(2), pp. 104–115. doi: 10.1016/j.scs.2011.05.001.

Goldblatt, R. *et al.* (2021) 'Remotely sensed derived land surface temperature (Lst) as a proxy for air temperature and thermal comfort at a small geographical scale', *Land*, 10(4). doi: 10.3390/land10040410.

Griffith, D. A. and Chun, Y. (2018) '3.01 GIS and Spatial Statistics/Econometrics: An Overview', in Huang, B. (ed.) *Comprehensive geographic information systems-VOLUME 3: GIS APPLICATIONS FOR SOCIO-ECONOMICS AND HUMANITY*. Elsevier, pp. 1–27.

Guha, S. *et al.* (2018) 'Analytical study of land surface temperature with NDVI and NDBI using Landsat 8 OLI and TIRS data in Florence and Naples city, Italy', *European Journal of Remote Sensing*, 51(1), pp. 667–678. doi: 10.1080/22797254.2018.1474494.

Guo, H. *et al.* (2020) 'On the understanding of the mean radiant temperature within both the indoor and outdoor environment, a critical review', *Renewable and Sustainable Energy*

*Reviews*, 117(April 2019), p. 109207. doi: 10.1016/j.rser.2019.06.014.

Hassani, H., Huang, X. and Silva, E. S. (2019) 'Big Data and Climate Change', *Big Data and COGNITIVE cOMPUTING*, (February). doi: 10.3390/bdcc3010012.

He, X. *et al.* (2015) 'Influence of sky view factor on outdoor thermal environment and physiological equivalent temperature', *International Journal of Biometeorology*, 59(3), pp. 285–297. doi: 10.1007/s00484-014-0841-5.

Heo, H. K. *et al.* (2021) 'Sky view factor calculation in complex urban geometry with terrestrial LiDAR', *Physical Geography*, 42(4), pp. 374–394. doi: 10.1080/02723646.2020.1778156.

Höppe, P. (1999) 'The physiological equivalent temperature - a universal index for the biometeorological assessment of the thermal environment.', *International journal of biometeorology*, 43(2), pp. 71–75. Available at: <http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L129347950>.

Intergovernmental Panel on Climate Change (IPCC) (2021) *Report 6 Climate Change 2021: The Physical Science Basis*. Available at: <https://www.ipcc.ch/report/ar6/wg1/>.

Irmak, M. A., Yilmaz, S. and Dursun, D. (2017) 'Effect of different pavements on human thermal comfort conditions', *Atmosfera*, 30(4), pp. 355–366. doi: 10.20937/ATM.2017.30.04.06.

IUPSThermal Commission (2001) *Glossary of terms for thermal physiology Third Edition*.

Jamei, E. *et al.* (2016) 'Review on the impact of urban geometry and pedestrian level greening on outdoor thermal comfort', *Renewable and Sustainable Energy Reviews*, 54, pp. 1002–1017. doi: 10.1016/j.rser.2015.10.104.

Jin, M. S., Kessomkiat, W. and Pereira, G. (2011) 'Satellite-Observed Urbanization Characters in Shanghai, China: Aerosols, Urban Heat Island Effect, and Land–Atmosphere Interactions', *Remote Sensing*, 3, pp. 83–99. doi: 10.3390/rs3010083.

Josse, J. and Husson, F. (2012) 'Selecting the number of components in principal component analysis using cross-validation approximations', *Computational Statistics and Data Analysis*, 56(6), pp. 1869–1879. doi: 10.1016/j.csda.2011.11.012.

Jusuf, S. K. and Hien, W. N. (2009) 'Development of empirical models for an estate level air temperature prediction in Singapore', in *In Proceedings of the Second International Conference on Countermeasures to Urban Heat Islands*.

Keith, L. *et al.* (2021) 'Deploy heat officers, policies and metrics', *Nature*, 598(7879), pp. 29–31. doi: 10.1038/d41586-021-02677-2.

Keith, L., Meerow, S. and Wagner, T. (2019) 'Planning for Extreme Heat : A Review', *Journal of Extreme Events*, 6(03n04), pp. 1–27. doi: 10.1142/S2345737620500037.

Kelly Turner, V. *et al.* (2022) 'More than surface temperature: mitigating thermal exposure in

hyper-local land system', *Journal of Land Use Science*, 00(00), pp. 1–21. doi: 10.1080/1747423X.2021.2015003.

Ketterer, C. and Matzarakis, A. (2014) 'Human-biometeorological assessment of heat stress reduction by replanning measures in Stuttgart , Germany', *Landscape and Urban Planning*, 122, pp. 78–88. doi: 10.1016/j.landurbplan.2013.11.003.

Kim, H. H. (1992) 'International Journal of Remote Sensing Urban heat island', *International Journal of Remote Sensing*, 13(12), pp. 2319–2336. doi: 10.1080/01431169208904271.

Kim, Y.-H. and Baik, J.-J. (2002) 'Maximum Urban Heat Island Intensity in Seoul', *Journal of Applied Meteorology and Climatology*, 41(6), pp. 651–659.

Kleerekoper, L., Van Esch, M. and Salcedo, T. B. (2012) 'How to make a city climate-proof, addressing the urban heat island effect', *Resources, Conservation and Recycling*, 64, pp. 30–38. doi: 10.1016/j.resconrec.2011.06.004.

Koc, M. and Acar, A. (2021) 'Investigation of urban climates and built environment relations by using machine learning', *Urban Climate*, 37(23). doi: 10.1016/j.uclim.2021.100820.

Kolokotroni, M. and Giridharan, R. (2008) 'Urban heat island intensity in London : An investigation of the impact of physical characteristics on changes in outdoor air temperature during summer', *Solar Energy*, 82(11), pp. 986–998. doi: 10.1016/j.solener.2008.05.004.

Koumetio Tekouabou, S. C. *et al.* (2022) 'Reviewing the application of machine learning methods to model urban form indicators in planning decision support systems : Potential , issues and challenges', *Journal of King Saud University - Computer and Information Sciences*, 34, pp. 5943–5967. doi: 10.1016/j.jksuci.2021.08.007.

Krüger, E., Drach, P. and Emmanuel, R. (2018) 'Atmospheric Impacts on Daytime Urban Heat Island', *Air, Soil and Water Research*, 11, pp. 1–3. doi: 10.1177/1178622118810201.

Krüger, E. L. (2021) 'Chapter 1 Introduction', in Krüger, E. L. (ed.) *Applications of the Universal Thermal Climate Index UTCI in Biometeorology Latest Developments and Case Studies*. Curitiba, Brazil: Springer. doi: <https://doi.org/10.1007/978-3-030-76716-7>.

Krüger, E. L., Minella, F. O. and Rasia, F. (2011) 'Impact of urban geometry on outdoor thermal comfort and air quality from field measurements in Curitiba, Brazil', *Building and Environment*, 46(3), pp. 621–634. doi: 10.1016/j.buildenv.2010.09.006.

Lai, A., Maing, M. and Ng, E. (2017) 'Observational studies of mean radiant temperature across different outdoor spaces under shaded conditions in densely built environment', *Building and Environment*, 114, pp. 397–409. doi: 10.1016/j.buildenv.2016.12.034.

Lai, L. W. and Cheng, W. L. (2009) 'Air quality influenced by urban heat island coupled with synoptic weather patterns', *Science of the Total Environment*, 407(8), pp. 2724–2733. doi: 10.1016/j.scitotenv.2008.12.002.

LatLong (2022) *Glasgow, the UK, LatLong*. Available at:

<https://www.latlong.net/place/glasgow-the-uk-28006.html>.

Lau, K. K.-L. *et al.* (2021) *Outdoor thermal comfort in urban environment. Assessments and applications in urban planning and design*.

Leal Filho, W. *et al.* (2021) 'Addressing the urban heat islands effect: A cross-country assessment of the role of green infrastructure', *Sustainability (Switzerland)*, 13(2), pp. 1–20. doi: 10.3390/su13020753.

Lee, K. *et al.* (2020) 'Trend Analysis of Urban Heat Island Intensity According to Urban Area Change in Asian Mega Cities', *Sustainability (Switzerland)*, 12(1), p. 11.

Lee, Y. Y., Kim, J. T. and Yun, G. Y. (2016) 'The neural network predictive model for heat island intensity in Seoul', *Energy and Buildings*, 110, pp. 353–361. doi: 10.1016/j.enbuild.2015.11.013.

Levermore, G. *et al.* (2018) 'The increasing trend of the urban heat island intensity', *Urban Climate*, 24, pp. 360–368. doi: 10.1016/j.uclim.2017.02.004.

Li, G. *et al.* (2018) 'Urban heat island effect of a typical valley city in China : Responds to the global warming and rapid urbanization', *Sustainable Cities and Society*, 38(October 2017), pp. 736–745. doi: 10.1016/j.scs.2018.01.033.

Li, T. and Meng, Q. (2018) 'A mixture emissivity analysis method for urban land surface temperature retrieval from Landsat 8 data', *Landscape and Urban Planning*, 179(July 2017), pp. 63–71. doi: 10.1016/j.landurbplan.2018.07.010.

Lindberg, F. *et al.* (2022) *UMEP Tutorial*. Available at: <https://umep-docs.readthedocs.io/projects/tutorial/en/latest/>.

Lindberg, F. and Grimmond, C. S. B. (2010) 'Continuous sky view factor maps from high resolution urban digital elevation models', *Climate Research*, 42(3), pp. 177–183. doi: 10.3354/cr00882.

Lindberg, F., Holmer, B. and Thorsson, S. (2008) 'SOLWEIG 1.0 - Modelling spatial variations of 3D radiant fluxes and mean radiant temperature in complex urban settings', *International Journal of Biometeorology*, 52(7), pp. 697–713. doi: 10.1007/s00484-008-0162-7.

Liu, H. (2010) 'On the Levenberg-Marquardt Training Method for Feed-Forward Neural Networks', (Icnc), pp. 456–460.

Liu, S. *et al.* (2021) 'Simulating and mitigating extreme urban heat island effects in a factory area based on machine learning', *Building and Environment*, 202(June), p. 108051. doi: 10.1016/j.buildenv.2021.108051.

Lowe, J. A. *et al.* (2018) *UKCP18 Science Overview Report, Met Office Hadley Centre*. Exeter. Available at: <https://www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/science-reports/UKCP18-Overview-report.pdf>.

Luo, X. and Peng, Y. (2016) 'Scale effects of the relationships between urban heat islands and

impact factors based on a geographically-weighted regression model', *Remote Sensing*, 8(760), pp. 1–19. doi: 10.3390/rs8090760.

Made 4 Geek (2018) *Calculating Land Surface Temperature Landsat8 by ArcGIS*. YouTube. Available at: <https://www.youtube.com/watch?v=G09MOTHFikM&list=PLcKS-BFfI9ED6IL6F991rvAvqww5btZwR&index=4>.

Malik, M. S., Shukla, J. P. and Mishra, S. (2019) 'Relationship of LST , NDBI and NDVI using Landsat-8 data in Kandaihimmat Watershed, Hoshangabad, India', *Indian Journal of Geo Marine Sciences*, 48(01), pp. 25–31.

Mangal, P., Rajesh, A. and Misra, R. (2020) 'Big Data in Climate Change Research : Opportunities and Challenges', *2020 International Conference on Intelligent Engineering and Management (ICIEM) Big*, pp. 321–326.

Martilli, A., Krayenhoff, E. S. and Nazarian, N. (2020) 'Is the Urban Heat Island intensity relevant for heat mitigation studies?', *Urban Climate*, 31(January 2019), pp. 1–4. doi: 10.1016/j.uclim.2019.100541.

Martin-Vide, J., Sarricolea, P. and Moreno-García, M. C. (2015) 'On the definition of urban heat island intensity: The "rural" reference', *Frontiers in Earth Science*, 3(June), pp. 2–4. doi: 10.3389/feart.2015.00024.

Matzarakis, A., Rutz, F. and Mayer, H. (2007) 'Modelling radiation fluxes in simple and complex environments - Application of the RayMan model', *International Journal of Biometeorology*, 51(4), pp. 323–334. doi: 10.1007/s00484-006-0061-8.

Mehrotra, S., Bardhan, R. and Ramamritham, K. (2020) 'Urban form as policy variable for climate- sensitive area planning under heterogeneity : a geographically weighted regression approach', *Area Development and Policy ISSN:*, 5(2), pp. 167–188. doi: 10.1080/23792949.2019.1609368.

Memon, R. A., Leung, D. Y. C. and Chunho, L. I. U. (2008) 'A review on the generation , determination and mitigation of Urban Heat Island', *Journal of Environmental Sciences*, 20, pp. 120–128.

Memon, R. A., Leung, D. Y. C. and Liu, C. (2009) 'An investigation of urban heat island intensity ( UHII ) as an indicator of urban heating', *Atmospheric Research*, 94(3), pp. 491–500. doi: 10.1016/j.atmosres.2009.07.006.

Mendez-Astudillo, J. and Mendez-Astudillo, M. (2021) 'A Machine Learning Approach to Monitoring the UHI From GNSS Data', *IEEE Transactions on Geoscience and Remote Sensing*, PP, pp. 1–11. doi: 10.1109/tgrs.2021.3091949.

Met Office (2016) 'Western Scotland: climate', *Met Office*. Available at: [https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/weather/learn-about/uk-past-events/regional-climates/western-scotland\\_-climate---met-office.pdf%0Ahttp://www.metoffice.gov.uk/climate/uk/regional-climates/ws](https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/weather/learn-about/uk-past-events/regional-climates/western-scotland_-climate---met-office.pdf%0Ahttp://www.metoffice.gov.uk/climate/uk/regional-climates/ws).

Met Office (2019) *UK Climate Projections: Headline Findings*. Exeter. Available at: <https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp-headline-findings-v2.pdf>.

Michel, A. *et al.* (2021) 'A new material-oriented tes for land surface temperature and suhi retrieval in urban areas: Case study over madrid in the framework of the future TRISHNA mission', *Remote Sensing*, 13(24), pp. 1–35. doi: 10.3390/rs13245139.

Mirzaei, P. A. (2015) 'Recent challenges in modeling of urban heat island', *Sustainable Cities and Society*, 19, pp. 200–206. doi: 10.1016/j.scs.2015.04.001.

Do Nascimento, A. C. L. *et al.* (2022) 'Comparison between Air Temperature and Land Surface Temperature for the City of São Paulo, Brazil', *Atmosphere*, 13(3), pp. 1–21. doi: 10.3390/atmos13030491.

Naser, M. Z. and Alavi, A. H. (2021) 'Error Metrics and Performance Fitness Indicators for Artificial Intelligence and Machine Learning in Engineering and Sciences', *Architecture, Structures and Construction*. doi: 10.1007/s44150-021-00015-8.

National Records of Scotland (2021) *Glasgow City Council Area Profile, National Records of Scotland*. Available at: <https://www.nrscotland.gov.uk/files/statistics/council-area-data-sheets/glasgow-city-council-profile.html>.

Nice, K. A., Coutts, A. M. and Tapper, N. J. (2018) 'Development of the VTUF-3D v1.0 urban micro-climate model to support assessment of urban vegetation influences on human thermal comfort', *Urban Climate*, 24(November 2017), pp. 1052–1076. doi: 10.1016/j.uclim.2017.12.008.

Norton, B. A. *et al.* (2015) 'Planning for cooler cities: A framework to prioritise green infrastructure to mitigate high temperatures in urban landscapes', *Landscape and Urban Planning*, 134, pp. 127–138. doi: 10.1016/j.landurbplan.2014.10.018.

Oertel, A., Emmanuel, R. and Drach, P. (2015) 'Assessment of predicted versus measured thermal comfort and optimal comfort ranges in the outdoor environment in the temperate climate of Glasgow , UK', *Building Services Engineering Research and Technology*, 36(4), pp. 482–499. doi: 10.1177/0143624414564444.

Oke, T. R. (1973) 'City Size and The Urban Heat Island', *Atmospheric Environment Pergamon Press*, 7, pp. 769–779.

Oke, T. R. (1976) 'The distinction between canopy and boundary - layer urban heat islands', *Atmosphere*, 14(4), pp. 268–277. doi: 10.1080/00046973.1976.9648422.

Oke, T. R. (1982) 'The energetic basis of the urban heat island', *Quarterly Journal of the Royal Meteorological Society*, 108(455), pp. 1–24. doi: 10.1002/qj.49710845502.

Oke, T. R. *et al.* (1991) 'Simulation of Surface Urban Heat Islands Under "Ideal" Conditions at Night Part 2: Diagnosis of Causation', *Boundary-Layer Meteorology*, 56, pp. 339–358. doi: 10.1007/BF00119211.

Oke, T. R. *et al.* (2017) 'Urban Heat Island', in *Urban Climates*. Cambridge: Cambridge University Press, pp. 197–237.

Olden, J. D. and Jackson, D. A. (2002) 'Illuminating the "black box": A randomization approach for understanding variable contributions in artificial neural networks', *Ecological Modelling*, 154(1–2), pp. 135–150. doi: 10.1016/S0304-3800(02)00064-9.

Páez, A., Farber, S. and Wheeler, D. (2011) 'A Simulation-Based Study of Geographically Weighted Regression as a Method for Investigating Spatially Varying Relationships', *Environment and Planning A*, 43, pp. 2992–3010. doi: 10.1068/a44111.

Papanastasiou, D. K., Melas, D. and Kambezidis, H. D. (2015) 'Air quality and thermal comfort levels under extreme hot weather', *Atmospheric Research*, 152, pp. 4–13. doi: 10.1016/j.atmosres.2014.06.002.

Parsaee, M. *et al.* (2019) 'Urban heat island, urban climate maps and urban development policies and action plans', *Environmental Technology and Innovation*, 14. doi: 10.1016/j.eti.2019.100341.

Patz, J. A. *et al.* (2005) 'Impact of regional climate change on human health', *Nature*, 438(7066), pp. 310–317. doi: 10.1038/nature04188.

Pearsall, H. (2017) 'Staying cool in the compact city: Vacant land and urban heating in Philadelphia, Pennsylvania', *Applied Geography*, 79, pp. 84–92. doi: 10.1016/j.apgeog.2016.12.010.

Pena Acosta, M. *et al.* (2021) 'How to bring UHI to the urban planning table? A data-driven modeling approach', *Sustainable Cities and Society*, 71(April). doi: 10.1016/j.scs.2021.102948.

Pettorelli, N. *et al.* (2005) 'Using the satellite-derived NDVI to assess ecological responses to environmental change', *TRENDS in Ecology and Evolution*, 20(503–510). doi: 10.1016/j.tree.2005.05.011.

Portelaa, C. I. *et al.* (2020) 'Impact of urban and industrial features on land surface temperature : Evidences from satellite thermal indices', *Sustainable Cities and Society*, 56, pp. 1–13. doi: 10.1016/j.scs.2020.102100.

Şahin, M. *et al.* (2012) 'Modelling and Remote Sensing of Land Surface Temperature in Turkey', *Journal of the Indian Society of Remote Sensing*, 40, pp. 399–409. doi: 10.1007/s12524-011-0158-3.

Salkind, N. J. (2010) *Encyclopedia of Research Design, Dictionary of Statistics & Methodology*. doi: 10.4135/9781412983907.n439.

Sandercock, H. (2022) *When is the next UK heatwave? Met Office weather forecast - will there be another hot spell in August 2022*, *National World*. Available at: <https://www.nationalworld.com/news/weather/when-is-the-next-uk-heatwave-met-office-weather-forecast-will-there-be-another-heatwave-august-2022-3774591> (Accessed: 3 September 2022).

Saunders, M., Lewis, P. and Thornhill, A. (2018) *Research Methods for Business Students*. Eight, Pearson Education Ltd. Eight. United Kingdom. Available at: [https://www.amazon.com/Research-Methods-for-Business-Students/dp/1292208783/ref=sr\\_1\\_2?dchild=1&qid=1614706531&refinements=p\\_27%3AAdrian+Thornhill+%2F+Philip+Lewis+%2F+Mark+N.+K.+Saunders&s=books&sr=1-2&text=Adrian+Thornhill+%2F+Philip+Lewis+%2F+Mark+N.+K.](https://www.amazon.com/Research-Methods-for-Business-Students/dp/1292208783/ref=sr_1_2?dchild=1&qid=1614706531&refinements=p_27%3AAdrian+Thornhill+%2F+Philip+Lewis+%2F+Mark+N.+K.+Saunders&s=books&sr=1-2&text=Adrian+Thornhill+%2F+Philip+Lewis+%2F+Mark+N.+K.)

Scottish Government (2019) *Data for Local Authority Review and Assessment purposes [dataset], Air Quality in Scotland*. Available at: <https://www.scottishairquality.scot/data/mapping/data>.

Scottish Government (2021a) *LiDAR for Scotland Phase 5 - LAS [dataset], Scottish Remote Sensing Portal*. Available at: <https://spatialdata.gov.scot/geonetwork/srv/eng/catalog.search#/metadata/63c46ac1-b10f-4453-968e-b1e95be0e47d>.

Scottish Government (2021b) *Maps of annual concentrations*. Available at: <https://www.scottishairquality.scot/data/mapping>.

Seltenrich, N. (2015) 'Between extremes: Health effects of heat and cold', *Environmental Health Perspectives*, 123(11), pp. 276–280. doi: 10.1289/ehp.123-A275.

Shah, R., Pandit, R. K. and Gaur, M. K. (2022) 'Urban physics and outdoor thermal comfort for sustainable street canyons using ANN models for composite climate', *Alexandria Engineering Journal*, 61(12), pp. 10871–10896. doi: 10.1016/j.aej.2022.04.024.

Shiny Weather Data (2018) *ERA5–ECMWF [dataset], Shiny Weather Data*. Available at: <https://www.shinyweatherdata.com/>.

Souza, L. C. L., Rodrigues, D. S. and Mendes, J. F. G. (2003) 'Sky-view factors estimation using a 3D-gis extension', *Eighth International IBPSA Conference*, (2001), pp. 1227–1234.

Stepani, H. M. N. and Emmanuel, R. (2022) 'How Much Green Is Really “Cool”? Target Setting for Thermal Comfort Enhancement in a Warm , Humid City ( Jakarta , Indonesia )', *Atmosphere*, 13(184), pp. 1–16. doi: 10.3390/atmos13020184.

Stewart, I. D. (2011) *Redefining the Urban Heat Island*. The University of British Columbia (Vancouver).

Synnefa, A. *et al.* (2008) 'On the use of cool materials as a heat island mitigation strategy', *Journal of Applied Meteorology and Climatology*, 47(11), pp. 2846–2856. doi: 10.1175/2008JAMC1830.1.

Taleghani, M. (2018) 'Outdoor thermal comfort by different heat mitigation strategies- A review', *Renewable and Sustainable Energy Reviews*, 81(March 2017), pp. 2011–2018. doi: 10.1016/j.rser.2017.06.010.

Taleghani, M. *et al.* (2021) 'Urban cooling : Which façade orientation has the most impact on a microclimate?', *Sustainable Cities and Society*, 64, pp. 1–10. doi:



10.1016/j.scs.2020.102547.

Taloor, A. K., Manhas, D. S. and Kothiyari, G. C. (2021) 'Retrieval of land surface temperature, normalized difference moisture index, normalized difference water index of the Ravi basin using Landsat data', *Applied Computing and Geosciences*, 9(August 2020), pp. 1–11. doi: 10.1016/j.acags.2020.100051.

Tan, J. *et al.* (2010) 'The urban heat island and its impact on heat waves and human health in Shanghai', 1850, pp. 75–84. doi: 10.1007/s00484-009-0256-x.

Tan, Z., Lau, K. K. and Ng, E. (2016) 'Urban tree design approaches for mitigating daytime urban heat island effects in a high-density urban environment', *Energy & Buildings*, 114, pp. 265–274. doi: 10.1016/j.enbuild.2015.06.031.

The Editors of Encyclopaedia Britannica (Britannica) (2022) *Glasgow, Encyclopaedia Britannica*. Available at: <https://www.britannica.com/place/Glasgow-Scotland>.

The Guardian (2018) *Taps aff: Glasgow bakes on city's hottest day on record*, *The Guardian*. Available at: <https://www.theguardian.com/uk-news/2018/jun/28/uk-heatwave-glasgow-bakes-on-hottest-day-on-record>.

Thorsson, S. *et al.* (2017) 'Present and projected future mean radiant temperature for three European cities', *International Journal of Biometeorology*, 61(9), pp. 1531–1543. doi: 10.1007/s00484-017-1332-2.

United Nations Department of Economic and Social Affairs Population Division (2019) *World Urbanization Prospects: The 2018 Revision, United Nations*. New York. Available at: <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf>.

USGS (2018) *Landsat 8 OLI and TIRS [dataset]*, *EarthExplorer*. Available at: <https://earthexplorer.usgs.gov/>.

Vanos, J. K. *et al.* (2016) 'Landscape and Urban Planning Hot playgrounds and children ' s health : A multiscale analysis of surface temperatures in Arizona , USA', *Landscape and Urban Planning*, 146, pp. 29–42. doi: <http://dx.doi.org/10.1016/j.landurbplan.2015.10.007>.

Venter, Z. S., Krog, N. H. and Barton, D. N. (2020) 'Linking green infrastructure to urban heat and human health risk mitigation in Oslo, Norway', *Science of the Total Environment*, 709, p. 136193. doi: 10.1016/j.scitotenv.2019.136193.

Watkins, R., Palmer, J. and Kolokotroni, M. (2007) 'Increased temperature and intensification of the urban heat island: Implications for human comfort and urban design', *Built Environment*, 33(1), pp. 85–96. doi: 10.2148/benv.33.1.85.

Weather Underground (2018) *Paisley, Scotland, United Kingdom Weather History [dataset]*, *Weather Underground*. Available at: <https://www.wunderground.com/history/daily/EGPF/date/2018-6-28>.

Wilby, R. L. (2003) 'Past and projected trends in London's Urban heat island', *Weather*, 58(7),

pp. 251–260. doi: 10.1256/wea.183.02.

Wilby, R. L. (2007) 'A Review of Climate Change Impacts on the Built Environment', *CLIMATE CHANGE AND CITIES*, 33(1), pp. 31–45. Available at: <https://www.jstor.org/stable/23289471>.

Wilby, R. L. (2008) 'Constructing climate change scenarios of urban heat island intensity and air quality', *Environment and Planning B: Planning and Design*, 35(5), pp. 902–919. doi: 10.1068/b33066t.

Xie, Y. *et al.* (2022) 'Prediction of mean radiant temperature distribution around a building in hot summer days using optimized multilayer neural network model', *Sustainable Cities and Society*, 84, pp. 1–15. doi: <https://doi.org/10.1016/j.scs.2022.103995>.

Yan, H. *et al.* (2022) 'Influence of view factors on intra-urban air temperature and thermal comfort variability in a temperate city', *Science of the Total Environment*, 841(March), p. 156720. doi: 10.1016/j.scitotenv.2022.156720.

Yang, C. *et al.* (2017) 'Mapping the Influence of Land Use / Land Cover Changes on the Urban Heat Island Effect — A Case Study of Changchun , China', *Sustainability*, 9(312), pp. 1–17. doi: 10.3390/su9020312.

Yegnanarayana, B. (2005) *Artificial Neural Networks*, PHI Learning Pvt. Ltd. Delhi, India: PHI Learning Pvt. Ltd. Available at: <http://cdn.iiit.ac.in/cdn/speech.iiit.ac.in/svlpubs/book/Yegna1999.pdf>.

Yilmaz, S. *et al.* (2021) 'Analysis of outdoor thermal comfort and air pollution under the influence of urban morphology in cold-climate cities: Erzurum/Turkey', *Environmental Science and Pollution Research*, 28(45), pp. 64068–64083. doi: 10.1007/s11356-021-14082-3.

Yilmaz, S., Irmak, M. A. and Qaid, A. (2022) 'Assessing the effects of different urban landscapes and built environment patterns on thermal comfort and air pollution in Erzurum city, Turkey', *Building and Environment*, 219(May), p. 109210. doi: 10.1016/j.buildenv.2022.109210.

Zha, Y., Gao, J. and Ni, S. (2003) 'Use of normalized difference built-up index in automatically mapping urban areas from TM imagery', *International Journal of Remote Sensing ISSN*, 24(3), pp. 583–594. doi: 10.1080/01431160304987.

Zhang, J. *et al.* (2022) 'Assessment of macroclimate and microclimate effects on outdoor thermal comfort via artificial neural network models', *Urban Climate*, 42(March). doi: 10.1016/j.uclim.2022.101134.

Zhao, H., Ren, Z. and Tan, J. (2018) 'The spatial patterns of land surface temperature and its impact factors: Spatial non-stationarity and scale effects based on a Geographically-Weighted regression model', *Sustainability (Switzerland)*, 10(7). doi: 10.3390/su10072242.

Zhao, L. *et al.* (2018) 'Interactions between urban heat islands and heat waves', *Environmental Research Letters*, 13(3). doi: 10.1088/1748-9326/aa9f73.

Zheng, Y., Tang, L. and Wang, H. (2021) 'An improved approach for monitoring urban built-up

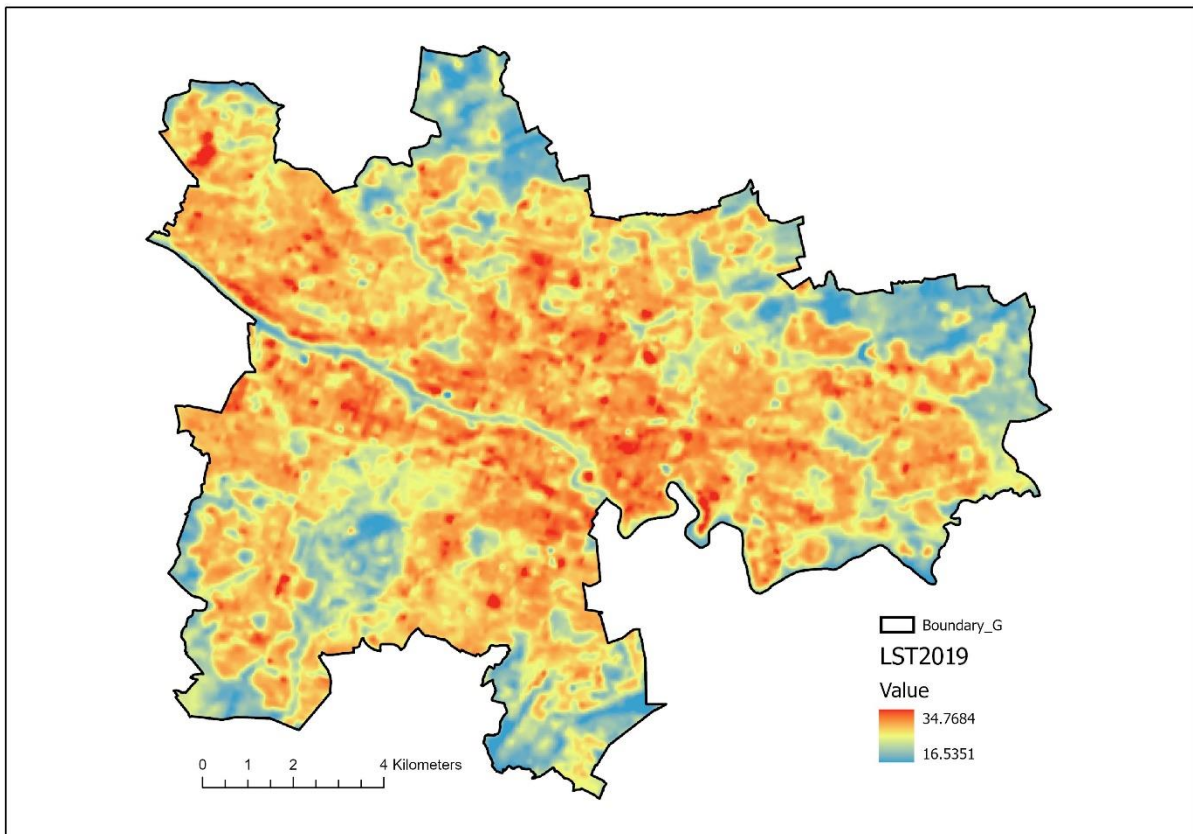
areas by combining NPP-VIIRS nighttime light, NDVI, NDWI, and NDBI', *Journal of Cleaner Production*, 328, pp. 1–13. doi: 10.1016/j.jclepro.2021.129488.

Zumwald, M. *et al.* (2021) 'Mapping urban temperature using crowd-sensing data and machine learning', *Urban Climate*, 35, pp. 1–14. doi: 10.1016/j.uclim.2020.100739.

# APPENDICES

## Appendix I

LST retrieved from Landsat 8- 28/06/2019



## Appendix II

### Meteorological data for SOLWEIG model

| iy   | id  | it | imin | Wind | RH | Td   | press | rain | Kdn | ldown | Wd  |
|------|-----|----|------|------|----|------|-------|------|-----|-------|-----|
| 2018 | 179 | 0  | 0    | 3.3  | 67 | 17.5 | 102.1 | 0    | 0   | 340   | 201 |
| 2018 | 179 | 1  | 0    | 3.2  | 72 | 16.4 | 102.1 | 0    | 0   | 336   | 209 |
| 2018 | 179 | 2  | 0    | 3.1  | 76 | 15.8 | 102.1 | 0    | 0   | 333   | 212 |
| 2018 | 179 | 3  | 0    | 3.1  | 79 | 15.4 | 102   | 0    | 0   | 329   | 212 |
| 2018 | 179 | 4  | 0    | 3.1  | 78 | 16.2 | 102   | 0    | 30  | 325   | 213 |
| 2018 | 179 | 5  | 0    | 3.1  | 72 | 18.3 | 102   | 0    | 108 | 323   | 216 |
| 2018 | 179 | 6  | 0    | 3    | 62 | 21   | 101.9 | 0    | 218 | 325   | 217 |
| 2018 | 179 | 7  | 0    | 3.4  | 54 | 23.2 | 101.9 | 0    | 343 | 330   | 220 |
| 2018 | 179 | 8  | 0    | 3.8  | 45 | 25.4 | 101.8 | 0    | 469 | 339   | 221 |
| 2018 | 179 | 9  | 0    | 4.3  | 39 | 27.2 | 101.7 | 0    | 586 | 345   | 225 |
| 2018 | 179 | 10 | 0    | 5    | 35 | 28.5 | 101.7 | 0    | 684 | 350   | 234 |
| 2018 | 179 | 11 | 0    | 5.4  | 31 | 29.4 | 101.6 | 0    | 752 | 355   | 245 |
| 2018 | 179 | 12 | 0    | 5.7  | 30 | 30   | 101.5 | 0    | 787 | 358   | 244 |
| 2018 | 179 | 13 | 0    | 5.9  | 29 | 30   | 101.4 | 0    | 767 | 363   | 241 |
| 2018 | 179 | 14 | 0    | 6.2  | 29 | 29.7 | 101.3 | 0    | 708 | 366   | 241 |
| 2018 | 179 | 15 | 0    | 6.5  | 28 | 29.3 | 101.2 | 0    | 631 | 365   | 246 |
| 2018 | 179 | 16 | 0    | 6.5  | 29 | 28.5 | 101.1 | 0    | 491 | 367   | 250 |
| 2018 | 179 | 17 | 0    | 5.8  | 30 | 27.6 | 101   | 0    | 392 | 369   | 252 |
| 2018 | 179 | 18 | 0    | 5    | 33 | 26.5 | 100.9 | 0    | 275 | 364   | 253 |
| 2018 | 179 | 19 | 0    | 4.2  | 34 | 25.4 | 100.9 | 0    | 118 | 367   | 259 |
| 2018 | 179 | 20 | 0    | 5.2  | 40 | 23.3 | 101   | 0    | 25  | 400   | 311 |
| 2018 | 179 | 21 | 0    | 6.5  | 68 | 18.3 | 101.2 | 0    | 5   | 387   | 353 |
| 2018 | 179 | 22 | 0    | 4.4  | 77 | 15.9 | 101.2 | 0    | 0   | 373   | 354 |
| 2018 | 179 | 23 | 0    | 3.8  | 72 | 14.8 | 101.3 | 0    | 0   | 344   | 339 |

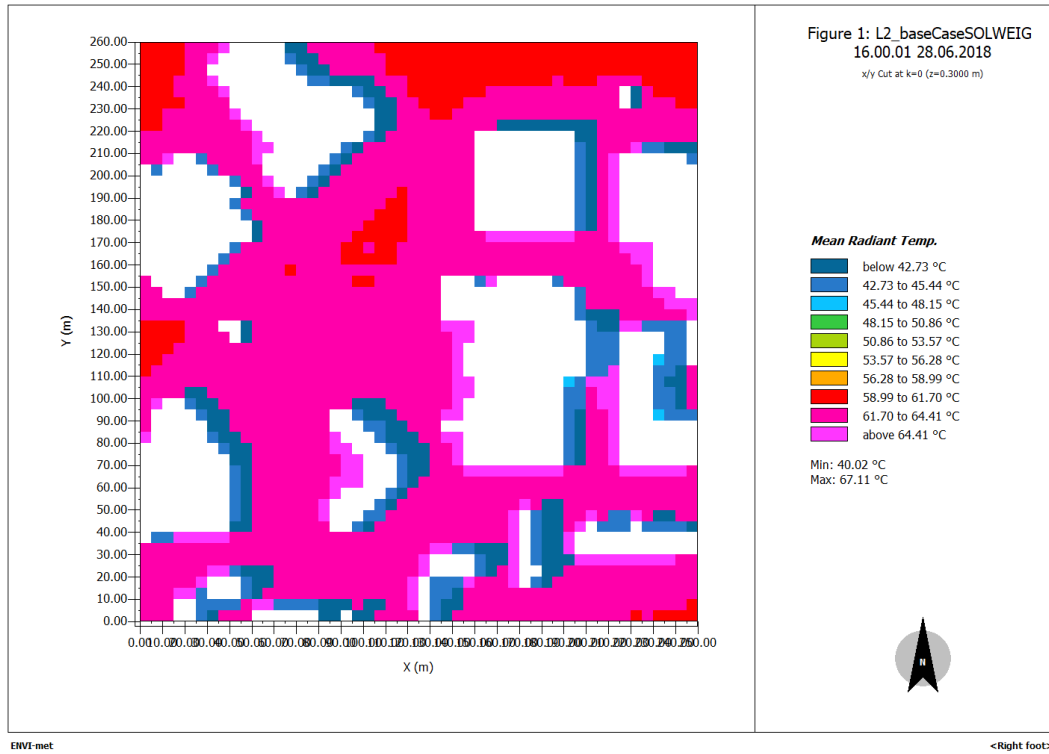
## Appendix III

### Meteorological data for ENVI-met

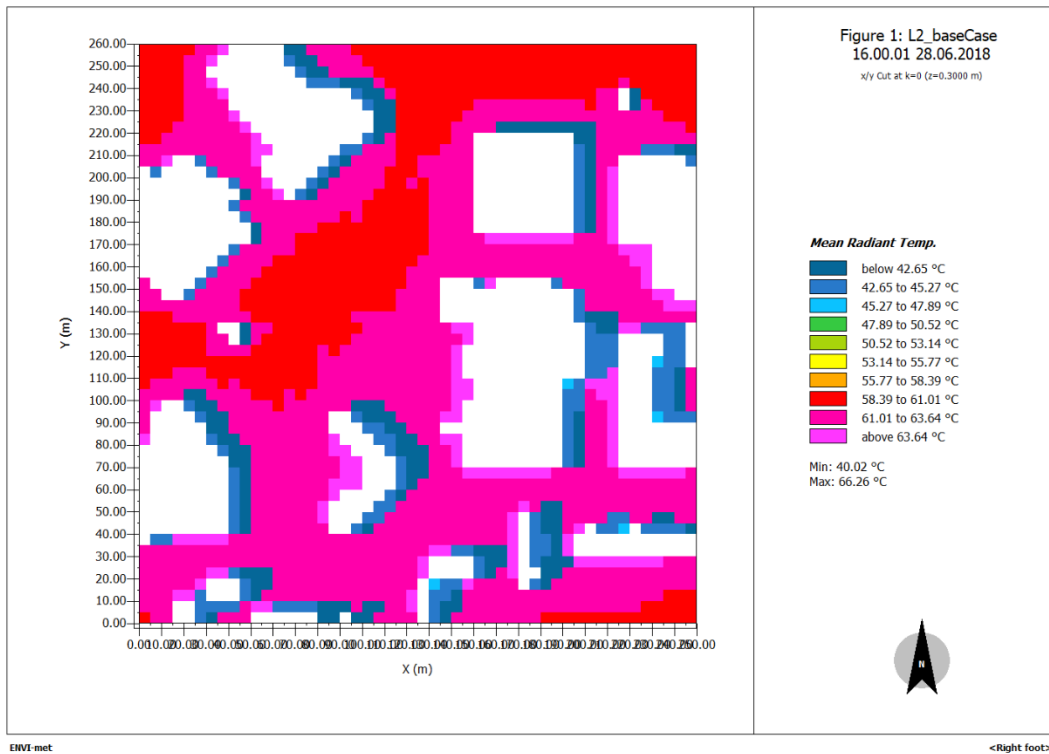
| Time     | Temperature | Humidity | Wind | Wind Speed |
|----------|-------------|----------|------|------------|
| 12:00 AM | 20 °C       | 35 %     | ENE  | 20 km/h    |
| 1:00 AM  | 17 °C       | 72 %     | E    | 7 km/h     |
| 2:00 AM  | 16 °C       | 77 %     | E    | 13 km/h    |
| 3:00 AM  | 15 °C       | 82 %     | ENE  | 9 km/h     |
| 4:00 AM  | 13 °C       | 94 %     | ENE  | 7 km/h     |
| 5:00 AM  | 13 °C       | 94 %     | NE   | 6 km/h     |
| 6:00 AM  | 14 °C       | 94 %     | NE   | 4 km/h     |
| 7:00 AM  | 16 °C       | 82 %     | VAR  | 2 km/h     |
| 8:00 AM  | 18 °C       | 77 %     | VAR  | 2 km/h     |
| 9:00 AM  | 20 °C       | 68 %     | VAR  | 4 km/h     |
| 10:00 AM | 23 °C       | 61 %     | VAR  | 4 km/h     |
| 11:00 AM | 25 °C       | 54 %     | VAR  | 4 km/h     |
| 12:00 PM | 27 °C       | 48 %     | VAR  | 6 km/h     |
| 1:00 PM  | 28 °C       | 37 %     | W    | 13 km/h    |
| 2:00 PM  | 30 °C       | 31 %     | WNW  | 11 km/h    |
| 3:00 PM  | 31 °C       | 29 %     | VAR  | 6 km/h     |
| 4:00 PM  | 31 °C       | 29 %     | VAR  | 6 km/h     |
| 5:00 PM  | 31 °C       | 29 %     | W    | 9 km/h     |
| 6:00 PM  | 31 °C       | 33 %     | WNW  | 7 km/h     |
| 7:00 PM  | 31 °C       | 35 %     | WNW  | 9 km/h     |
| 8:00 PM  | 29 °C       | 40 %     | NW   | 11 km/h    |
| 9:00 PM  | 25 °C       | 54 %     | NE   | 13 km/h    |
| 10:00 PM | 23 °C       | 47 %     | NE   | 22 km/h    |
| 11:00 PM | 21 °C       | 46 %     | ENE  | 19 km/h    |

# Appendix IV

## Base case for Site B with SOLWEIG calculation



## Base case for Site B without SOLWEIG calculation



## Appendix V

Optimisation and hyperparameter for ANN training

| <b>Training Parameters Levenberg-Marquardt Algorithm and Hyperparameter</b> |          |
|---|----------|
| <b>goal</b>   | 0        |
| <b>show</b>   | 25       |
| <b>epochs</b>   | 1000     |
| <b>max_fail</b>   | 6        |
| <b>min_grad</b>   | 1.0e-07  |
| <b>time</b>   | Infinite |
| <b>mu</b>   | 1.0e-03  |
| <b>mu_dec</b>   | 0.1      |
| <b>mu_inc</b>   | 10       |
| <b>mu_max</b>   | 1.0e+10  |



## Appendix VI

### Sensitivity Analysis for LST

| Exclusion                             |  | RMSE    | MSE     | MAE     | R-Squared | R-Squared Difference |
|---------------------------------------|--|---------|---------|---------|-----------|----------------------|
| <b>None</b>                           | -None                                      | 0.43617 | 0.19024 | 0.32862 | 0.80      | -                    |
| <b>Feature elimination: predictor</b> | -DEM                                       | 0.49714 | 0.24714 | 0.37921 | 0.75      | -0.05                |
|                                       | -DSMb                                      | 0.49926 | 0.24926 | 0.38062 | 0.75      | -0.05                |
|                                       | -CDSM                                      | 0.49505 | 0.24508 | 0.37873 | 0.76      | -0.04                |
|                                       | -SVF                                       | 0.50059 | 0.25059 | 0.38156 | 0.75      | -0.05                |
|                                       | -NDVI                                      | 0.50233 | 0.25233 | 0.38457 | 0.75      | -0.05                |
|                                       | -NDBI                                      | 0.5193  | 0.26967 | 0.39321 | 0.73      | -0.07                |
|                                       | -BSF                                       | 0.50111 | 0.25111 | 0.3836  | 0.75      | -0.05                |
|                                       | -Mean_Height                               | 0.4908  | 0.24089 | 0.3749  | 0.76      | -0.04                |
|                                       | -RAR                                       | 0.49291 | 0.24296 | 0.3766  | 0.76      | -0.04                |
|                                       | -VSF                                       | 0.49284 | 0.24289 | 0.3756  | 0.76      | -0.04                |
|                                       | -PM10                                      | 0.50079 | 0.25079 | 0.38392 | 0.75      | -0.05                |
|                                       | -NOx                                       | 0.50289 | 0.25289 | 0.38524 | 0.75      | -0.05                |
|                                       | <b>Feature elimination: source of data</b> | -DEM    | 0.51693 | 0.26721 | 0.39624   | 0.73                 |
| -DSM                                  |  |         |         |         |           |                      |
| -CDSM                                 |  |         |         |         |           |                      |
| -SVF                                  |  |         |         |         |           |                      |
| -NDVI                                 |  | 0.54164 | 0.29338 | 0.41455 | 0.70      | -0.10                |
| -NDBI                                 |  |         |         |         |           |                      |
| -RoadRatio                            |  | 0.52261 | 0.27313 | 0.39953 | 0.73      | -0.07                |
| -VegeRatio                            |  |         |         |         |           |                      |
| -BSF                                  |  | 0.49623 | 0.24625 | 0.38039 | 0.75      | -0.05                |
| -MeanHeight                           |  |         |         |         |           |                      |
| -NOx                                  | 0.52147                                    | 0.27193 | 0.40117 | 0.73    | -0.07     |                      |
| -PM10                                 |  |         |         |         |           |                      |

## Appendix VII

### Sensitivity Analysis for LST

| Exclusion                             |  | RMSE    | MSE     | MAE      | R-Squared | R-Squared Difference |
|---------------------------------------|--|---------|---------|----------|-----------|----------------------|
| <b>None</b>                           | -None                                      | 0.57814 | 0.33425 | 0.39548  | 0.66      | -                    |
| <b>Feature elimination: predictor</b> | -DEM                                       | 0.59313 | 0.3518  | 0.4039   | 0.65      | -0.01                |
|                                       | -DSMb                                      | 0.57373 | 0.32916 | 0.38909  | 0.67      | +0.01                |
|                                       | -CDSM                                      | 0.57418 | 0.32969 | 0.39216  | 0.67      | +0.01                |
|                                       | -SVF                                       | 0.6629  | 0.43944 | 0.48127  | 0.56      | -0.10                |
|                                       | -NDVI                                      | 0.59695 | 0.35635 | 0.41131  | 0.64      | -0.02                |
|                                       | -NDBI                                      | 0.59585 | 0.35504 | 0.40903  | 0.64      | -0.02                |
|                                       | -BSF                                       | 0.58889 | 0.34679 | 0.40399  | 0.65      | -0.01                |
|                                       | -Mean_Height                               | 0.58047 | 0.33694 | 0.3999   | 0.66      | -0.00                |
|                                       | -RAR                                       | 0.60232 | 0.36278 | 0.40966  | 0.64      | -0.02                |
|                                       | -VSF                                       | 0.5915  | 0.34987 | 0.40512  | 0.65      | -0.01                |
|                                       | -PM10                                      | 0.5949  | 0.3539  | 0.40526  | 0.65      | -0.01                |
|                                       | -NOx                                       | 0.59288 | 0.3515  | 0.40364  | 0.65      | -0.01                |
|                                       | <b>Feature elimination: source of data</b> | -DEM    | 0.776   | 0.602018 | 0.57702   | 0.39                 |
| -DSM                                  |  |         |         |          |           |                      |
| -CDSM                                 |  |         |         |          |           |                      |
| -SVF                                  |  |         |         |          |           |                      |
| -NDVI                                 |  | 0.58725 | 0.34486 | 0.40105  | 0.65      | -0.01                |
| -NDBI                                 |  |         |         |          |           |                      |
| -RoadRatio                            |  | 0.61292 | 0.37567 | 0.41697  | 0.62      | -0.04                |
| -VegeRatio                            |  |         |         |          |           |                      |
| -BSF                                  |  | 0.59409 | 0.35295 | 0.41064  | 0.65      | -0.01                |
| -MeanHeight                           |  |         |         |          |           |                      |
| -NOx                                  | 0.6099                                     | 0.37198 | 0.41556 | 0.63     | -0.03     |                      |
| -PM10                                 |  |         |         |          |           |                      |