

Spatiotemporal dynamics of urban health: Physiological data driven strategies for enhancing urban health and wellbeing

by Dimitra Dritsa

Thesis submitted in fulfilment of the requirements for
the degree of

Doctor of Philosophy

under the supervision of
Dr. Nimish Bioria
Dr Ling Chen

University of Technology Sydney
Faculty of Design Architecture and Building

January 2021

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Dimitra Dritsa declare that this thesis, is submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy, in the Faculty of Architecture at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Production Note:

Signature: Signature removed prior to publication.

Date: 18/01/2021

ACKNOWLEDGEMENTS

First and foremost, I am particularly indebted to my supervisor, Associate Professor Nimish Bitoria, for his continued support and constant encouragement, without which this research would not have been possible. Nimish was really influential in shaping the direction that my research taken in the past six years, and I am extremely grateful for the time that he spent sharing his knowledge and guiding me throughout all the aspects of this project. I would also like to thank my co-supervisor, Associate Professor Ling Chen, who was always eager to give me helpful and encouraging feedback.

My eternal gratitude to Anastasia, Anky, Fryni and Despina, who were always there for me despite the sabotaging timezone difference, sending me sourdough recipes and aesthetically questionable postcards with the most heartwarming messages, and giving me the much needed motivation to speed up the submission process so that we can meet again.

I am also indebted to my friends in UTS who shared with me the joys and despairs of the PhD journey and made the school a home away from home. An enormous thanks to Fereshteh and Pejman, Thao, Mukesh and Kerl; also, to Alessia, Biyanka, Binyang, Caroline, Elaine, Elham, Enya, Fiona, Fisher, Gautam, Jeff, Juanwen, Megan, Samuel, Tazy, Xianli, Xiaohang and the other prominent members of the UTS PhD Girls et al. group. I will always cherish our times together.

And finally, my biggest thanks to my family for all the love and support they have shown through this journey, and my partner, Chrysostomos, for making it such a pleasure to return home every day, in each and every home that we have shared all around the world.

THESIS FORMAT STATEMENT

This thesis is formatted following the requirements of a conventional thesis. Two papers submitted for publication are also presented in [Appendix A](#) and [B](#). References to the two papers have been included in the main work whenever appropriate.

TABLE OF CONTENTS

Certificate of original authorship.....	ii
Acknowledgements	iii
Thesis format statement	iv
Table of contents.....	v
List of figures	xiv
Abstract	xxvii
Abbreviations.....	xxix
Publications	xxxii
1 Introduction	1
1.1. Context.....	1
1.2. Mapping physiological responses in the urban environment using continuous monitoring: overview of past studies	6
1.2.1. Methods for study selection.....	6
1.2.2. Description and methodological overview of reviewed studies.....	9
1.2.3. Findings of past studies	16
1.2.3.1. Understanding the link between urban environment and physiological responses	17
1.2.4. Future prospects	21
1.2.5. Issues and challenges	24
1.3. Research aims and objectives.....	31
1.4. Thesis structure	35
2 Research design.....	40
2.1. Introduction	40

2.2.	The methodology for collection and analysis of physiological data in the urban space	40
2.3.	Linking the proposed methodology with the identified issues	43
2.4.	Research design	44
2.4.1.	Overview of research strategy	44
2.4.2.	Experiment design	48
2.5.	Conclusion	55
3	Understanding the connection between urban and environmental features and physiological responses: A review	56
3.1.	Introduction	56
3.2.	Information processing, stress and physical activity	58
3.2.1.	Sensory processing, physiological arousal and the autonomic nervous system	58
3.2.2.	The orienting and defensive response	61
3.2.3.	Stress	62
3.2.4.	Physical activity and exercise	65
3.2.5.	Summary and conclusions	67
3.3.	The effect of different parameters on measures of physiological responses	70
3.3.1.	Stress and physiological responses	73
3.3.2.	Stimulation and physiological responses	74
3.3.3.	Movement and physiological responses	75
3.3.4.	Personal factors, context and interaction between stressors	76
3.3.5.	Summary and conclusions	77
3.4.	Physiological responses during interactions with the urban environment	78
3.4.1.	Parameters that affect the experience of walking in the urban environment	80

3.4.2.	Sensory stimulation, psychological stressors and restoration in the urban space	85
3.5.	Presentation of the theoretical and conceptual framework	94
3.6.	Discussion	104
4	Data mining methods for movement-related, contextual and physiological data: A review	108
4.1.	Introduction	108
4.2.	Data mining methods and techniques	109
4.2.1.	Data preparation	109
4.2.2.	Machine learning	109
4.2.3.	Data mining methods for spatial and temporal data	115
4.3.	Mining movement, physiological and spatial urban data: trends and challenges	118
4.3.1.	Movement data	118
4.3.2.	Physiological data	126
4.3.3.	POI and OSM data	136
4.4.	The proposed scheme for data analysis	146
4.5.	Discussion	148
5	A method for analysis and classification of physiological responses based on movement	150
5.1.	Introduction	150
5.2.	The proposed method for data collection and analysis	151
5.2.1.	Setting up the spatial database: Preparation of POI and OSM data	151
5.2.2.	Movement and physiological data collection	159
5.2.3.	Speed and accelerometer data pre-processing	162
5.2.4.	EDA, HR and skin temperature data pre-processing	168

5.2.5.	Spatial, physiological and movement data fusion for individual analysis	174
5.3.	Demonstration of the method using data from 2 users	179
5.3.1.	User A	180
5.3.2.	User B	189
5.4.	Discussion	193
6	Examining the connection between movement, contextual parameters and physiological responses in the urban environment in Sydney and Zurich	200
6.1.	Introduction	200
6.2.	Dataset characteristics and context analysis	202
6.2.1.	Description of the experiment setup in Sydney and Zürich	202
6.2.2.	Data collection and analysis	204
6.2.3.	Analysis of contextual and activity-related characteristics for the predefined route in Sydney and Zürich	207
6.2.4.	Analysis of contextual characteristics for the free-living activities dataset in Sydney	219
6.3.	Methods	220
6.4.	Results	223
6.4.1.	Sydney: Predefined route	224
6.4.2.	Zürich: Predefined route	225
6.4.3.	Sydney: Free-living activities	226
6.4.4.	Combined data from Sydney and Zürich	226
6.5.	Discussion	227
7	Methods for spatial analysis of physiological responses in the urban environment	234
7.1.	Introduction	234
7.2.	Methods	238
7.2.1.	Dataset characteristics	238

7.2.2.	Network analysis	238
7.2.3.	The proposed method for hotspot identification, cluster separation and user ranking.....	240
7.3.	Results.....	244
7.3.1.	Spatial analysis of physiological responses in Zürich	244
7.3.2.	Spatial analysis of physiological responses in Sydney.....	251
7.4.	Discussion	258
8	Machine learning methods for prediction of physiological responses in the urban environment.....	261
8.1.	Introduction	261
8.2.	Methods.....	265
8.2.1.	Dataset	265
8.2.2.	Data analysis.....	265
8.2.3.	Machine learning methods.....	271
8.2.4.	Feature importance identification.....	273
8.3.	Results.....	273
8.3.1.	Classification.....	273
8.3.2.	Regression	277
8.4.	Discussion	283
9	Linking physiological responses in the urban space to pathfinding: Algorithmic methods for identifying the least stressful route	286
9.1.	Introduction	286
9.2.	Methods.....	289
9.2.1.	Typical approaches to pathfinding in the urban space.....	289
9.2.2.	Incorporating urban and activity-related features as attributes	291
9.2.3.	Possible approaches to the pathfinding problem.....	293

9.2.4.	Description of each algorithmic approach.....	295
9.2.5.	Evaluation of the experiments.....	300
9.3.	Simulation analysis for demonstration of the proposed methods	301
9.3.1.	Simulations for Scenario A.....	301
9.3.2.	Simulations for scenario D, and comparison with scenario A	303
9.3.3.	Simulations for scenario B, and comparison with scenario D.....	305
9.3.4.	Quantitative evaluation of the proposed models.....	306
9.4.	Discussion	308
10	Conclusions and future directions	314
10.1.	Introduction	314
10.2.	Revisiting the research aims and objectives.....	315
10.2.1.	The designed methodology as a response to the primary objectives..	315
10.2.2.	Discussion of the findings of the statistical analysis for Sydney and Zürich: Revisiting the conceptual framework.....	318
10.3.	Research contributions	322
10.3.1.	Methodological and practical contributions.....	322
10.3.2.	Contributions to the theoretical field of investigation.....	326
10.4.	Limitations.....	327
10.5.	Research implications for different domains	329
10.5.1.	Implications for urban design and planning.....	329
10.5.2.	Methodological considerations for future studies on continuous physiological data monitoring in the urban environment.....	331
10.6.	Future work.....	331
10.6.1.	Research on the links between urban environment features, activity and physiological responses for different user groups	332
10.6.2.	Incorporation of qualitative data related to user experience.....	332

10.6.3.	Personalisation of the models for prediction of physiological responses	334
10.6.4.	Use of the developed methods for post-occupancy evaluation	334
10.6.5.	Use of the spatial database to identify stressor patterns at a city scale	334
10.6.6.	Enrichment of the spatial database with socioeconomic and other stressors	335
10.6.7.	Adaptability and response to the current circumstances	335
Appendix A	337
Abstract.....	338
1. Introduction.....	338
2. Materials and methods	341
2.1. Spatial data processing: data analysis and data fusion	342
2.2 Measuring complexity from the site characteristics	343
2.3. Statistical analysis.....	345
3. Results	346
3.1. Case study 1	346
3.2. Case study 2	347
4. Discussion	351
Acknowledgements	354
Declaration of interest	354
Appendix B.....	355
Abstract.....	356
Keywords	356
1. Introduction.....	356
2. Materials and methods	359

2.1.	Study design	359
2.2.	Data collection and data fusion.....	361
2.3.	Statistical analysis.....	367
3.	Results	370
3.1.	Analysis of measures during the outdoor route.....	370
4.	Discussion	374
4.1.	Elaboration on the presented methodology and analysis of the findings...	374
4.2.	Implications for urban design and planning.....	376
4.3.	Limitations.....	376
5.	Conclusion	376
	Disclosure statement	377
	Appendix C.....	378
C.1.	Protocol and study design	378
C.2.	Statistical analysis approach	380
C.3.	Results.....	382
C.4.	Graphs of the other physiological measures during the indoor experiments	385
	Appendix D	388
D.1.	Ethics approval letter: ETH19-3752	389
D.2.	Ethics approval letter: ETH20-5253	391
D.3.	Participant information sheet and consent form for the experiments in Sydney 393	
D.4.	Questionnaire	397
D.5.	Guide given to the participants explaining the use of the equipment.....	403
	Appendix E.....	410
E.1.	Accelerometer data analysis: Activity classification.....	410
E.1.1.	Description of the dataset and assignment of the ground truth labels	410

E.1.2. A threshold-based method for activity classification	412
E.1.3. Supervised ML methods for activity classification.....	413
E.1.4. Comparison of the threshold-based approach with the ML model for activity classification	417
E.2. Artefact recognition	418
E.2.1. Dataset	418
E.2.2. Feature extraction	419
E.2.3. Classification approach and ML algorithms.....	420
Appendix F.....	422
Appendix G	425
Appendix H	428
References.....	430

LIST OF FIGURES

Figure 1.1. An overview of the reviewed studies.	8
Figure 1.2. Upper: The increase in studies on physiological mapping in the urban space in the past years. Bottom: The frequency of appearance of different physiological data in studies.	10
Figure 1.3. Frequency of appearance of concepts and theories related to stress and emotions in the studies.	12
Figure 1.4. Frequency of appearance of different environmental data (left) and urban features (right) in the studies.	14
Figure 1.5. The steps of the workflow followed in the reviewed studies.	16
Figure 1.6. The connections between future applications and relevant stakeholders.	24
Figure 1.7. Current challenges and prospects.	31
Figure 2.1. The components of the proposed methodology.	41
Figure 2.2. The research strategy.	47
Figure 2.3. Presentation of the datasets used in this study.	49
Figure 2.4. The predefined route for the outdoor test in Phase A of data collection.	50
Figure 2.5. The wristbands used in the study (left: FitBit Charge 2; right: Empatica E4).	51
Figure 2.6. Demographic characteristics for the data collected in Sydney.	53
Figure 3.1. The aim of the chapter and the connection with the conceptual methodology.	57
Figure 3.2. Schematic diagram situating the topics explored in section 3.2 in relation to the proposed methodology.	58

Figure 3.3. The relationship between the autonomic nervous system, stressors and sympathetic arousal.	67
Figure 3.4. A categorisation of stressors related to physiological responses based on the factors that initiate the arousal.	69
Figure 3.5. Schematic diagram situating the topics explored in section 3.3 in relation to the proposed methodology.	70
Figure 3.6. Upper: The tonic and the phasic components of EDA. Bottom: EDRs and NS.SCRs elicited after stimulation (acquired from Boucsein 2012).	72
Figure 3.7. A typical EDR and its measures (acquired from Boucsein 2012).	72
Figure 3.8. A summary of typical physiological (EDA and HR) responses to stimulation and movement.	78
Figure 3.9. The topics explored in this section in relation to component 1 of the proposed methodology.	80
Figure 3.10: Urban characteristics and associated changes in speed and activity intensity.	83
Figure 3.11: Spatial parameters which have the potential to act as physical stressors.	84
Figure 3.12. The relationship between urban stimulus characteristics and physiological effects in the urban environment.	93
Figure 3.13. The theoretical framework.	97
Figure 3.14. A presentation of the theoretical framework from the perspective of a user moving in the urban space.	98
Figure 3.15. The conceptual framework linking urban and movement-related features to physiological responses.	100
Figure 3.16. A conceptual scheme for analysis and interpretation of physiological data in the urban domain, based on analysis of movement and context.	103

Figure 3.17. Connection of the designed scheme for the analysis of different data with component 1 of the methodology.	104
Figure 4.1. Common approaches to the analysis of speed data.	121
Figure 4.2. Common steps in the analysis of accelerometer data.	126
Figure 4.3. Heart rate variability (HRV): The variation in successive RR-intervals.	127
Figure 4.4. Common approaches to HR and HRV data analysis.	129
Figure 4.5. Types of overlapping EDR responses.	130
Figure 4.6. Types of artefacts in EDA measurement.	131
Figure 4.7. Common methods implemented in the analysis of EDA data.	132
Figure 4.8. A summary of methods commonly used in POI data analysis.	142
Figure 4.9. A summary of methods commonly used in OSM data analysis.	145
Figure 4.10. A schematic depiction of the proposed data fusion model for the analysis of physiological responses.	146
Figure 5.1. The aim of the chapter and the connection with the conceptual methodology.	151
Figure 5.2. OSM and POI data acquisition, analysis and fusion.	153
Figure 5.3. The scheme used for the analysis of OSM tags.	155
Figure 5.4. Application of the scheme for assessment of traffic intensity in Sydney. ...	156
Figure 5.5. The spatial distribution of POI density at different scales.	157
Figure 5.6. The outcome of combining POI density data with OSM nodes.	158
Figure 5.7. The contents of the spatial database.	159
Figure 5.8. The physiological data collection protocol.	161
Figure 5.9. The analysis of speed and altitude data.	162

Figure 5.10. Activity analysis after the application of the activity classification model.	167
Figure 5.11. The extraction of different features from the analysis of activity.	168
Figure 5.12. EDA signal processing: artefact removal, extraction of tonic EDA and peak identification for extraction of EDRs.	169
Figure 5.13. The analysis of EDA data.	170
Figure 5.14. The analysis of HR data.	171
Figure 5.15. Presentation of the two methods for the analysis of changes in the HR data.	172
Figure 5.16. An example of the extraction of different features from EDA data.	173
Figure 5.17. The scheme for spatial, physiological and movement data fusion at the final stage of the individual data analysis.	175
Figure 5.18. The calculation of the level of psychological stressors based on the different contextual features. The changes in the traffic and POI density are also added later to this calculation.	177
Figure 5.19. Extraction of the physical stressors based on the analysis of activity.	178
Figure 5.20. The spatial distribution of physiological responses (sum of EDR amplitudes) in routes 2 and 5 for User A.	181
Figure 5.21. A place which was characterised as a space with a low level of psychological stressors in Route 2 of User A.	181
Figure 5.22. Analysis of the physical and psychological stressors for Route 2.	182
Figure 5.23. The upper graph presents a detailed analysis of the contextual and activity- related stressors.	183
Figure 5.24. Photo 2A shows the place where the responses first started appearing in Route 2, for user A.	184

Figure 5.25. Analysis of the physical and psychological stressors for Route 5.	185
Figure 5.26. A detailed analysis of the contextual and movement-related stressors, in parallel to the analysis of the change in the physiological responses.	187
Figure 5.27. Two of the places with high levels of psychological stressors that User A encountered during Route 5. Photo 5A is taken from Google Street View (15 Broadway street).	188
Figure 5.28. Analysis of the physical and psychological stressors for Route 9.	189
Figure 5.29. A detailed analysis of the contextual and movement-related stressors for Route 9, in parallel to the analysis of the change in the physiological responses.	190
Figure 5.30. The spatial distribution of physiological responses (sum of EDR amplitudes) in route 9 for User B.	191
Figure 5.31. William Street-Kings Cross.	192
Figure 5.32. Rushcutters Bay.	192
Figure 6.1. The aim of the chapter and the connection with the conceptual methodology.	200
Figure 6.2a. Feature description for the datasets collected in Sydney.	205
Figure 6.2b. Feature description for the dataset collected in Zürich and the combined dataset.	206
Figure 6.3. Mean temperature data for each participant, for the predefined routes in Zürich and Sydney.	208
Figure 6.4. POI density data in the studied areas in Zürich and Sydney.	209
Figure 6.5. Photos from the route in Zürich. The photos were obtained from Google Street View (Google Maps 2020).	210
Figure 6.6. Photos from the route in Sydney. Photos 8 and 9 were taken by the author; the rest were obtained from Google Street View (Google Maps 2020).	212

Figure 6.7. POI density time-series data for one user in Zürich and another user in Sydney.	213
Figure 6.8. Screenshots from a 3-dimensional model of the studied area in Sydney (3d city models of Sydney 2017).	214
Figure 6.9. Screenshots of the studied area in Zürich (Bing maps 2020).	215
Figure 6.10. Traffic time-series data for one user in Zürich and another user in Sydney.	216
Figure 6.11. Example of the activity intensity data collected during walking on a predefined route in Zürich and Sydney.	217
Figure 6.12. Steady-state walking (time series) data for one user in Zürich and another user in Sydney.	217
Figure 6.13. Raw EDA data from all users in Zürich and Sydney.	218
Figure 6.14. The movement-related and contextual characteristics of the free-living activities dataset.	219
Figure 6.15. The statistical analysis approach followed in this chapter.	221
Figure 6.16. The parameters of the selected linear mixed model for the predefined route in Sydney.	224
Figure 6.17. The parameters of the selected linear mixed model for the predefined route in Zürich.	225
Figure 6.18. The parameters of the selected linear mixed model for the free-living activities in Sydney.	226
Figure 6.19. The parameters of the selected linear mixed model for the combined dataset from Sydney and Zürich.	226
Figure 6.20. The coefficients of the significant features for all models.	227
Figure 6.21. Analysis of the effect size for each parameter.	228

Figure 6.22. Calculation of effect size based on the std values for the sum of EDR amplitude data.	229
Figure 7.1. Flowchart outlining the aim of the chapter and the connection with the conceptual methodology.	234
Figure 7.2. The results of betweenness centrality analysis for the studied area in Sydney.	240
Figure 7.3. The proposed workflow for spatial analysis of physiological responses.	241
Figure 7.4. The results of the hotspot identification based on the Local Moran's I values for the Zürich dataset.	245
Figure 7.5. Experimentation with different resampling rates and cut-off values for the change in the sum of EDR amplitudes data.	246
Figure 7.6. The results of parameter testing for the cluster separation phase with the DBSCAN algorithm applied to the Zürich dataset.	248
Figure 7.7. Application of cluster ranking methods on the route in Zürich.	249
Figure 7.8. The locations of the most important clusters in Zürich.	250
Figure 7.9. Photos of the locations corresponding to the critical clusters in Zürich (Google Maps 2020).	250
Figure 7.10. The results of the hotspot identification with Local Moran's I values for the free-living activities dataset in Sydney.	251
Figure 7.11. Result of the DBSCAN algorithm applied on the sum of EDR amplitudes, with user-based ranking applied.	252
Figure 7.12. The proposed workflow for spatial analysis applied to the change in the sum of EDR amplitudes data collected in Sydney.	253
Figure 7.13. The locations of the most important clusters in the studied area in Sydney.	254

Figure 7.14. Photos depicting the contextual circumstances in the significant clusters. In Sydney. The photos are taken from Google Street View (Google Maps 2020), apart from photos A and 3 that were taken by the author.	255
Figure 7.15. Extraction of contextual parameters for each cluster.	257
Figure 8.1. The aim of the chapter and the connection with the conceptual methodology.	261
Figure 8.2. Description of the workflow adopted for the prediction of the target variable.	268
Figure 8.3. Presentation of the two approaches that were followed in the regression task.	269
Figure 8.4. An alternative problem framing that can be considered in the future.	270
Figure 8.5. CNN architecture.	272
Figure 8.6. LSTM architecture.	272
Figure 8.7. Presentation of different metrics for all models at three window sizes.	274
Figure 8.8. Comparison of accuracy levels for all models.	274
Figure 8.9. Performance metrics at different time windows for all models.	275
Figure 8.10. Accuracy score of the RF classifier for different time windows.	275
Figure 8.11. Results of feature ranking according to their importance in different time windows.	276
Figure 8.12. The sum of importance scores for each feature, based on the scores presented in Figure 8.11.	276
Figure 8.13. Detailed importance score for each feature.	277
Figure 8.14. Presentation of the MSE scores for each model.	278

Figure 8.15. Presentation of the MSE scores for each model, at different window sizes.	278
Figure 8.16. The actual and predicted values from the different models that were tested.	279
Figure 8.17. The MSE scores for each user group, using the XGBoost model, at different window sizes.	279
Figure 8.18. Comparison of the actual and predicted values for the selected XGBoost model.	281
Figure 8.19. Feature importance analysis for the regression task, for each time window.	281
Figure 8.20. Visualisation of feature importance ranking for the regression task.	282
Figure 8.21. Visualisation of feature importance ranking for the regression task: The sum of ranks for each feature, based on Figure 8.19.	282
Figure 9.1. Flowchart outlining the aim of the chapter and the connection with the conceptual methodology.	287
Figure 9.2. The procedure typically followed for solving pathfinding problems in the urban space.	290
Figure 9.3. The selection of relevant attributes for the pathfinding algorithm, and their grouping in two categories.	292
Figure 9.4. The different scenarios for pathfinding for the identification of the least stressful route.	294
Figure 9.5. An example of Scenario A applied to a network constructed for demonstration purposes.	296
Figure 9.6. An example of nodes coloured according to the generated isochrones.....	299

Figure 9.7. Comparison between the baseline model and three applications of Scenario A for a random pair of nodes in Sydney.	302
Figure 9.8. The image shows all the routes from Figure 9.7 together. A route generated by incorporating all four criteria from equation 1 is also presented for comparison....	303
Figure 9.9. Comparison between Scenario A and Scenario D (the isochrone-based approach) applied to the selected pair of starting and ending nodes.	304
Figure 9.10. Two places which are included in the route with- and without the isochrone-based (time correction) penalty (Google Maps 2020).	305
Figure 9.11. Comparison of the results of the isochrone-based scenario D, with and without the integration of stress hotspots (scenario B).	306
Figure 9.12. Comparison of the results of the different scenarios.	307
Figure 9.13. The comparison of exposure to stressors for 50 simulated routes.	308
Figure A1. Upper: The outcome of the calculation of POI density for the Sydney CBD. Lower: POI density in the area selected for Case Study 2.	342
Figure A2. Examples of the locations which were assessed in terms of complexity.	344
Figure A3. An example of the attributes extracted from the images in relation to complexity. The resulting degree of complexity is also displayed for each location, as well as the corresponding POI density.	345
Figure A4. The results of the spatial regression models.	346
Figure A5. Left: Graphs presenting the calculated degree of complexity and POI density for each point. Right: A kernel density plot and a regression plot showing the two-dimensional relationship between the two variables.	347
Figure A6. Maps showing the spatial distribution of complexity, POI density and the separate predictors of complexity.	348

Figure A7. Maps showing the variation in pedestrian activity and traffic. The points with highest variation are depicted with red colour, while low variations are depicted with light yellow.	349
Figure A8. Variations in the number of pedestrians in two selected spots.	350
Figure B1. The outdoor route.	360
Figure B2. Demonstration of the application of the artefact recognition algorithm.	361
Figure B3. An example of extraction of different features from EDA data.	363
Figure B4. Map displaying the POI density in the studied area, following the analysis of the contextual data.	365
Figure B5. An example of the extraction and analysis of movement-related and contextual features for one route.	366
Figure B6. Photos from the route.	367
Figure B7. A description of the features used in the analysis.	369
Figure B8. Graphs describing the EDA measures and the contextual and movement variables at each segment, for all participants.	370
Figure B9. The results of the hotspot analysis for the sum of EDR amplitudes and the changes in this measure, in parallel to contextual and movement-related parameters.	372
Figure B10. The results of the linear mixed model analysis, with the sum of EDR amplitudes as the dependent variable.	373
Figure B11. Graphs showing the positive and negative affect score of all participants for each segment.	374
Figure C1. Description of the indoor activities.	379
Figure C2. Boxplots showing the tonic EDA and sum of EDR amplitudes for the indoor activities, before the outdoor activity ('indoor A') and after it ('indoor B').	382

Figure C3. The results of the linear mixed model with the best performance.	383
Figure C4. The results of the PANAS questionnaire for each participant for the indoor test, before ('_PRE') and after ('_POST') the outdoor route.	384
Figure C5. The distribution of other physiological measures related to EDA (EDR frequency measured in peaks/minute; mean EDR amplitude measured in μ S) during each activity stage in the indoor experiments.	385
Figure C6. The distribution of other physiological signals (heart rate (HR), skin temperature (ST)) during each activity stage in the indoor experiments.	386
Figure C7. The median and standard deviation values for each measure during each activity.	387
Figure E1. The data used for the construction of the activity classification model.	411
Figure E2. An example of labelled accelerometer data.	412
Figure E3. The number of data belonging to each level of activity intensity.	412
Figure E4. Application of the threshold-based algorithm for activity recognition.	413
Figure E5. Accuracy scores for each model for the activity classification task.	414
Figure E6. Performance metrics for each activity classification model.	415
Figure E7. Comparison of DNN models with different layers for the activity classification task.	416
Figure E8. An example of activity classification using the selected DNN model.	417
Figure E9. Comparison between the threshold-based approach and the ML model for activity classification, using the test data.	417
Figure E10. The data used in the EDA artefact recognition model.	418
Figure E11. The distribution of the data used for training the artefact recognition algorithm.	419

Figure E12. The method used for visualising the data in segments and labelling them as artefacts or clean data.	419
Figure E13. Accuracy scores for each model for the artefact recognition task.	420
Figure E14. Performance metrics for each artefact recognition model.	420
Figure E15. Example of the application of the constructed DNN model for artefact recognition.	421
Figure H1. Frequency table for the categorical variables in the data collected in Zürich and the combined dataset.	428
Figure H2. Frequency table for the categorical variables in the data collected in Sydney.	429

ABSTRACT

The starting point for this research was the emergence of physiological data as a source of information that can help us understand how our interactions with the urban environment affect the human body. There is significant potential in extending existing methods for physiological data analysis in the urban domain in a way that maximises the benefits at the individual and the city scale. Physiological data could be used to identify the least stressful route, but there is currently a lack of research on their incorporation in pathfinding studies. The area of prediction of physiological responses during outdoor walking has also been understudied.

This study aims to address these issues by designing a methodology for collection and analysis of physiological data in the urban space. The methodology incorporates three components: (1) the collection and analysis of physiological data at an individual level, (2) the hotspot analysis of physiological responses at a city scale, and (3) the utilisation of the collected data in models for prediction of physiological responses, and pathfinding methods for the identification of the least stressful route. The methods and algorithms for each component of the methodology are calibrated using data collected in Sydney from experiments organised by the author, and publicly available data from a previous study conducted in Zürich.

The study acts as a pilot project that will pave the way towards large-scale experiments in this area. Its main contribution is that it supports the construction of tools for individuals who want to understand how different routes might affect their physiological responses, and have a calm experience while walking in the urban environment. It can also help researchers identify which parts of the city are associated with an increased intensity of physiological responses, possibly indicating increased stress levels. The construction of a theoretical and conceptual framework supporting the construction of the methodology also enriches current research on the links between urban environment, activity and physiological responses. Other methodological and practical contributions include the development of methods for analysing how movement may influence physiological responses as a physical stressor, and their incorporation in the designed methodology; also, the development of

methods for identifying physical and psychological stressors from contextual data, based on freely available OpenStreetMap and Point of Interest data, as an alternative to image-based analysis which was used in previous studies.

ABBREVIATIONS

ADAM	Adaptive Moment Estimation
AHR	Additional Heart Rate
ANS	Autonomic Nervous System
API	Application Programming Interface
CNN	Convolutional Neural Network
DBSCAN	Density-Based Spatial Clustering of Application with Noise
DNN	Deep Neural Network
DT	Decision Tree
ECG	Electrocardiogram
EDA	Electrodermal Activity
EDR	Electrodermal Response
EEG	Electroencephalography
FFT	Fast Fourier Transform
GIS	Geographic Information System
GPS	Global Positioning System
GSR	Galvanic Skin Response
HMM	Hidden Markov Model
HR	Heart Rate
HRV	Heart Rate Variability
KDE	Kernel Density Estimation
k-NN	k-Nearest Neighbors

LISA	Local Indicators of Spatial Association
LSTM	Long-Short Term Memory (Network)
MAE	Mean Absolute Error
ML	Machine Learning
MSE	Mean Squared Error
NS.SCR	Non-Specific Skin Conductance Response
OSM	OpenStreetMap
PANAS	Positive And Negative Affect Schedule
POI	Point Of Interest
STD	Standard Deviation
SVM	Support Vector Machine
RNN	Recurrent Neural Networks
RF	Random Forests
RMSE	Root Mean Square Error
PPG	Photoplethysmography
ST	Skin Temperature

PUBLICATIONS

Materials presented in chapter 1 have generated the following research output:

Dritsa, D. & Bioria, N. 2021, 'Mapping the urban environment using real-time physiological monitoring', *Archnet-IJAR*, Vol. ahead-of-print No. ahead-of-print, doi: <https://doi.org/10.1108/ARCH-02-2021-0041>.

Materials presented in Appendix A have generated the following research output:

Dritsa, D. & Bioria, N. 2021, 'Analysing the relationship between POI density and stimulus complexity in the urban environment', *Journal of Urban Design*, Vol. ahead-of-print No. ahead-of-print, doi: [10.1080/13574809.2021.1903306](https://doi.org/10.1080/13574809.2021.1903306).

Materials presented in chapter 1 are based on findings from the following research output:

Dritsa, D. & Bioria, N. 2018, 'Towards a Multi-Scalar Framework for Smart Healthcare', *Smart and Sustainable Built Environment*, vol.7, no.1, pp.33-52, doi: <https://doi.org/10.1108/SASBE-10-2017-0057>.