



UNIVERSITY OF THE WITWATERSRAND

SCHOOL OF PUBLIC HEALTH

**Urban Heat and Health in Johannesburg:
A Multidimensional Analysis of
Vulnerability,
Explanatory Modelling, and Predictive
Outcomes**

PhD Protocol

Version v1.0

Last updated: April 22, 2025

Submitted by:

Craig Parker, MSc, MM-PDM, BS

Wits Planetary Health Research

Supervisors:

Dr. Admire Chikandiwa

Professor Matthew Chersich

Professor Akbar Waljee

Dr. Christopher Jack

April 22, 2025

Contents

Abstract

Acknowledgements

1	Introduction and Background	1
1.1	Climate Change and Heat-Health Impacts in the Johannesburg Context	1
1.2	Socio-Spatial Inequity and Heat Vulnerability	1
1.3	Research Context and Positionality	1
1.4	Conceptual Framework and Research Gaps	2
1.5	Recent Evidence and Research Gaps	3
2	Aims and Objectives	3
2.1	Primary Research Question	3
2.2	Objectives	4
2.2.1	Objective 1: Geographically Weighted Heat Vulnerability Mapping . .	4
2.2.2	Objective 2: Causal Mechanisms of Heat-Health Relationships	4
2.2.3	Objective 3: Predictive Modeling for Early Warning Systems	4
3	Methodology	5
3.1	Study Design	5
3.2	Data Sources and Collection	5
3.2.1	Environmental and Socioeconomic Data	6
3.2.2	Health and Socioeconomic Data	6
3.2.3	Sample Size Justification	7
3.3	Analytical Approaches	7
3.3.1	Geographically Weighted Vulnerability Mapping	7
3.3.2	Causal Machine Learning Approaches	8
3.4	Objective 1: Vulnerability Mapping	8
3.5	Objective 2: Explanatory Modeling	8
3.5.1	Clinical Pathway Validation	8

3.5.2	Temporal Heat Exposure Modeling	9
3.6	Objective 3: Predictive Modeling	9
3.6.1	Model Accuracy and Reliability Measures	9
3.7	Forecast Window Selection	10
4	Ethical Considerations	11
5	Expected Outcomes and Impact	11
5.1	Research Outputs	11
5.2	Anticipated Impact and Infrastructural Justice Framing	11
5.3	Knowledge Translation	13
6	Project Timeline	13
7	Supervision	14
8	References	

Abstract

This research proposal investigates the complex relationship between urban heat and health in Johannesburg, South Africa. As climate change drives increasing temperatures globally, urban populations face heightened health risks, with vulnerable communities disproportionately affected. The study employs a multidimensional approach through three primary objectives: (1) mapping intra-urban heat vulnerability by integrating environmental, socio-economic, and health data; (2) delineating heat-health dynamics through a two-stage explanatory modeling approach that combines hypothesis generation with targeted testing to uncover physiological pathways and temporal effects; and (3) developing a stratified predictive model for heat-related health outcomes. Drawing on clinical trial data, satellite imagery, climate records, and socioeconomic surveys, this research will apply advanced statistical and machine learning techniques to create vulnerability maps, explain complex relationships, and generate predictive models. The findings aim to inform targeted public health interventions, urban planning decisions, and climate adaptation strategies to protect vulnerable populations from increasing heat exposure in Johannesburg and potentially other African urban centers.

Keywords: urban heat, health outcomes, vulnerability mapping, machine learning, climate change, Johannesburg

Acknowledgements

This research was conducted as part of the HEat and HEalth African Transdisciplinary (HE²AT) Center initiative, supported by the NIH Common Fund under Award Number U54 TW 012083.

1 Introduction and Background

1.1 Climate Change and Heat-Health Impacts in the Johannesburg Context

Climate change has intensified heat-related health risks in Johannesburg, where rising temperatures are linked to increasing mortality and morbidity ([Gasparrini et al., 2015](#); [Romanello et al., 2023](#)). With over 5.87 million inhabitants, the city faces substantial warming projections—by 2050, mean temperatures may rise approximately 2°C with hot nights projected to quadruple ([Engelbrecht and et al., 2015](#); [World Bank Cities Support Program, 2024](#)). Current research indicates that above 18.7°C apparent temperature, all-cause mortality rises by 0.9% per 1°C increase, with seniors experiencing 2.1% increases ([Wichmann, 2017](#)). The IPCC warns that beyond +2°C of global warming, heat-attributable health impacts in Africa will sharply rise ([Intergovernmental Panel on Climate Change, 2024](#)).

1.2 Socio-Spatial Inequity and Heat Vulnerability

Johannesburg's socio-spatial layout—largely a legacy of apartheid-era planning—significantly shapes contemporary heat vulnerability patterns. Historical policies created wealthy suburbs with green spaces alongside dense townships with minimal vegetation, resulting in temperature differentials of approximately 6°C between affluent neighborhoods and informal settlements ([World Bank Cities Support Program, 2024](#)). Housing quality further exacerbates this disparity, with informal dwellings experiencing up to 15°C higher indoor temperatures than formal housing ([Naicker and et al., 2017](#)). This embedded vulnerability continues to cluster heat-health risks in historically marginalized communities ([Strauss, 2019](#)).

1.3 Research Context and Positionality

This research builds upon ongoing work at the HE²AT Center, which has established baseline heat vulnerability assessment frameworks for South African cities ([Jack et al., 2024](#)). As a researcher within this initiative, I bring expertise in data science and machine learning methods to

extend these frameworks through advanced statistical techniques. My position at the intersection of public health and climate science informs a multidisciplinary approach that aims to translate complex data into actionable health interventions. This work specifically addresses gaps in the Heat Center’s existing vulnerability assessments by developing dynamic predictive models that can account for temporal changes in vulnerability patterns across Johannesburg’s diverse socioeconomic landscape.

1.4 Conceptual Framework and Research Gaps

This research employs a comprehensive framework for heat vulnerability encompassing three interconnected dimensions: exposure, sensitivity, and adaptive capacity ([Intergovernmental Panel on Climate Change, 2022](#)). Figure 1 illustrates these relationships and their determinants within Johannesburg’s urban context.

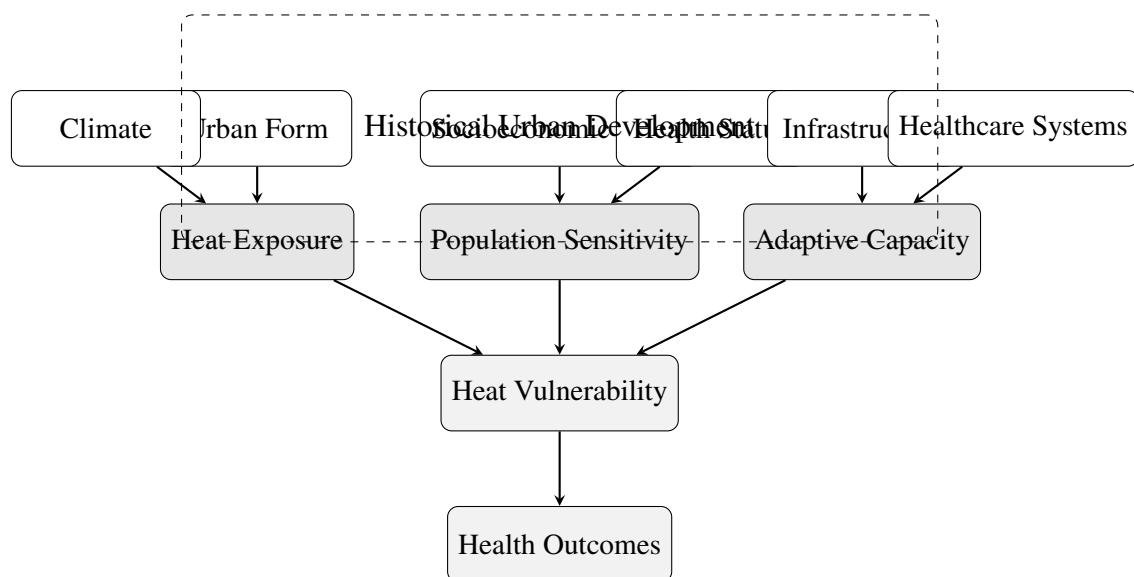


Figure 1: Conceptual framework of heat vulnerability in Johannesburg, highlighting the three primary dimensions (exposure, sensitivity, adaptive capacity) and their determinants. The dashed outline indicates components influenced by historical urban development patterns including apartheid-era planning.

These components interact as follows:

- *Exposure* refers to heat stress degree and duration, with dense urban areas showing significantly higher surface temperatures (up to 5°C) compared to well-vegetated neighbourhoods ([Li and et al., 2017](#); [Santamouris, 2015](#)).

- *Sensitivity* reflects population susceptibility influenced by socio-economic conditions and health status, with chronic conditions significantly increasing heat-related health risks (Watts and et al., 2023; Khosla et al., 2021; Souverijns et al., 2022).
- *Adaptive capacity* depends on access to healthcare, cooling infrastructure, and social support systems (Ansah and et al., 2024), with limited access to healthcare affecting vulnerability to heat, as demonstrated by increased heat-related mortality in areas with restricted medical services (Murage and et al., 2020).

1.5 Recent Evidence and Research Gaps

Recent studies have strengthened our understanding of heat-mortality relationships in African urban contexts. Parker et al. (Parker et al., 2023) demonstrated that in Johannesburg, heat vulnerability clusters in historically disadvantaged areas, with environmental exposure explaining 31.5% of variance. The Lancet Countdown (Romanello et al., 2023) reports escalating impacts, with an estimated 5 million people globally dying annually from suboptimal temperatures. Despite this growing evidence, significant research gaps persist in African urban contexts (Khine and Langkulsen, 2023), including: (1) scarcity of region-specific studies; (2) siloed disciplinary approaches failing to capture multifaceted heat-health relationships; and (3) limited recognition of unique urban challenges stemming from historical development patterns and disease profiles. Johannesburg exemplifies these challenges through its urban disparities and accelerated warming projections (Engelbrecht and et al., 2015).

2 Aims and Objectives

2.1 Primary Research Question

This study addresses the following primary research question: *How do urban form, socio-demographic characteristics, and climate factors interact to shape heat vulnerability and health outcomes in Johannesburg, and what interventions could effectively reduce heat-related health impacts?*

Specific sub-questions include: (1) spatial distribution of heat vulnerabilities; (2) mediating

factors in heat-health relationships; (3) effectiveness of current adaptation strategies; and (4) optimal infrastructural and policy interventions.

2.2 Objectives

This research is structured around three interconnected objectives:

2.2.1 Objective 1: Geographically Weighted Heat Vulnerability Mapping

Develop comprehensive spatial vulnerability assessment using geographically weighted approaches that account for spatial non-stationarity. This will identify patterns of heat risk across Johannesburg's diverse urban landscape, examine how historical development patterns shape contemporary vulnerability, and identify priority intervention areas with particular emphasis on infrastructural and socio-spatial inequities.

2.2.2 Objective 2: Causal Mechanisms of Heat-Health Relationships

Identify causal mechanisms underlying observed patterns of heat vulnerability through advanced explanatory modeling. This involves investigating which biological systems are most affected by heat stress, how these impacts vary across demographic groups, and what socioeconomic factors mediate these effects.

2.2.3 Objective 3: Predictive Modeling for Early Warning Systems

Develop predictive models incorporating near-real-time environmental data, meteorological forecasts, and vulnerability factors to predict heat-health risks. This includes identifying critical temperature thresholds and neighborhood-specific warning indicators to support short-term emergency response and longer-term adaptation planning.

These four objectives form an integrated research framework addressing spatial, physiological, temporal, and policy dimensions of heat vulnerability. The research design builds each objective upon the findings of the previous one, creating a cohesive pathway from vulnerability characterization to mechanistic understanding to practical solution development.

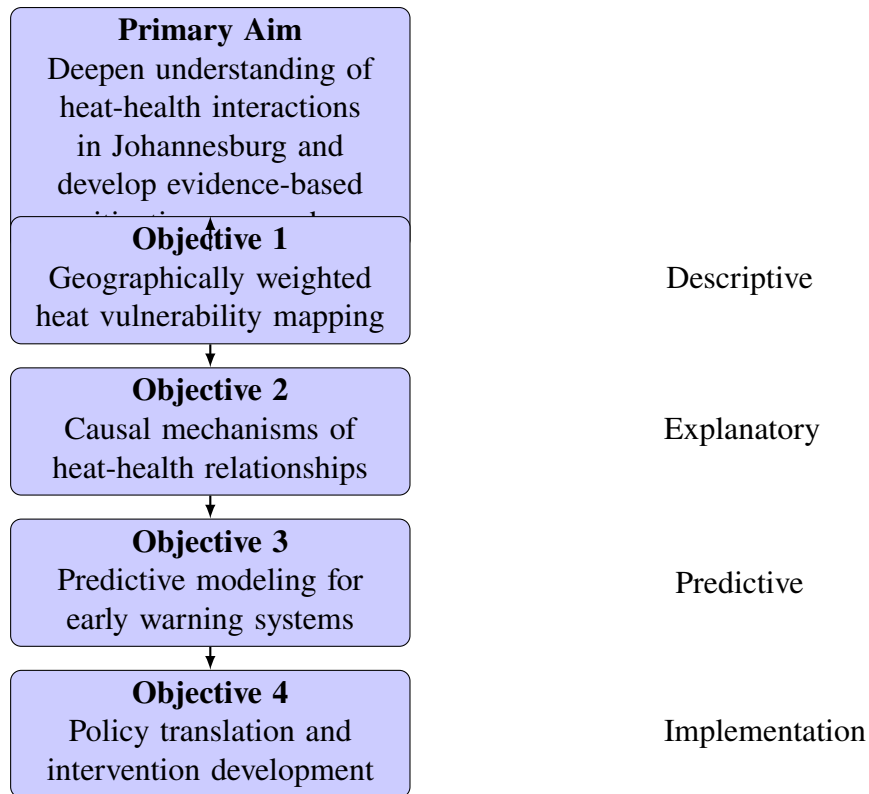


Figure 2: Hierarchical framework of research objectives showing progression from descriptive to implementation approaches.

3 Methodology

3.1 Study Design

This study employs a mixed-methods design integrating quantitative geospatial analysis with qualitative community-engaged research across four phases:

1. **Vulnerability Mapping** - Developing geographically weighted heat vulnerability indices
2. **Causal Analysis** - Identifying key causal pathways and mechanisms
3. **Intervention Assessment** - Evaluating effectiveness of existing interventions
4. **Policy Translation** - Developing evidence-based policy recommendations

3.2 Data Sources and Collection

Health data spanning 2000–2022 will be drawn from clinical trials and cohort studies conducted in Johannesburg that meet defined inclusion criteria: cohorts or trials with 200 adult participants, prospectively collected data, comprehensive clinical and laboratory variables, and appropriate

ethics approvals. The dataset will include clinical indicators and laboratory tests covering renal, metabolic and inflammatory markers and demographic/socioeconomic factors. Given potential confounding effects, 2020–2022 data will undergo stratified analysis to account for pandemic-related influences on healthcare utilization and outcomes. A comprehensive overview of all data sources is provided in Appendix Table 1.

3.2.1 Environmental and Socioeconomic Data

Environmental parameters will be derived from multiple validated sources, including Landsat 8 ([USGS and NASA, 2013](#)), Sentinel-2 ([European Space Agency, 2015](#)), ERA5 reanalysis ([Hersbach et al., 2017](#)), and the Copernicus Climate Data Store ([Copernicus, 2020](#)). These data streams will provide high-resolution measurements of Land Surface Temperature, vegetation indices, air temperature, and derived heat metrics. Socioeconomic determinants will be incorporated through ward-level data from the Gauteng City-Region Observatory Quality of Life Surveys ([Gauteng Observatory, 2021](#)), capturing housing quality, infrastructure adequacy, and healthcare access. The data processing and integration workflow is detailed in Appendix Table 2, with specific security measures outlined in Appendix Table 3.

3.2.2 Health and Socioeconomic Data

We will integrate: daily mortality records and hospital admission data (with focus on cardiovascular, respiratory, and heat-specific diagnoses), demographic variables from Statistics South Africa, social vulnerability indicators (housing quality, utility access), pre-existing health condition prevalence, and COVID-19 data to assess potential interactive effects with heat vulnerability. All datasets will be harmonized to comparable spatial units for integrated analysis.

3.2.3 Sample Size Justification

The study's statistical power is ensured through comprehensive data coverage and adequate sample sizes. For vulnerability mapping, data from all 135 wards provide complete spatial coverage. For the explanatory and predictive modelling components, the combined dataset of approximately 5,000–7,000 records substantially exceeds the minimum requirement of 10-20

observations per predictive feature (with 20-25 features anticipated). This sample size provides >80% power to detect clinically meaningful effects, calculated as:

$$n = \frac{2(Z_{\alpha/2} + Z_{\beta})^2 \sigma^2}{\Delta^2} \quad (1)$$

where $Z_{\alpha/2}$ and Z_{β} are standard normal deviates for type I and II errors, σ^2 is the expected variance, and Δ is the minimal detectable effect size.

3.3 Analytical Approaches

3.3.1 Geographically Weighted Vulnerability Mapping

A key innovation in this study is the application of Geographically Weighted Principal Component Analysis (GWPCA), which accounts for spatially heterogeneous relationships between vulnerability indicators (Quispe et al., 2024; Praharaj et al., 2024). This is crucial in Johannesburg given its extreme socio-spatial fragmentation. The GWPCA approach includes: (1) indicator selection across exposure, sensitivity, and adaptive capacity domains; (2) spatial bandwidth optimization through cross-validation; (3) local component extraction allowing vulnerability factors to vary spatially; (4) construction of location-specific vulnerability indices; (5) spatial mapping; and (6) comparative analysis between global and local statistical approaches. Implementation will utilize the GWmodel package in R.

3.3.2 Causal Machine Learning Approaches

To move beyond correlation and identify causal mechanisms linking urban form, socioeconomic conditions, and heat-health outcomes, we will employ: (1) causal structure learning using the PC algorithm to discover potential causal relationships; (2) causal effect estimation using double machine learning and causal forests to estimate heterogeneous treatment effects of specific urban features; and (3) causal mediation analysis to quantify direct and indirect pathways through which urban characteristics influence health outcomes. Implementation will use causal-learn and econml Python packages, with sensitivity analyses to assess robustness.

3.4 Objective 1: Vulnerability Mapping

Geospatial data will undergo rigorous preprocessing, including normalization, completeness assessments, and spatial harmonization. The methodology employs Geographically Weighted Principal Component Analysis (GWPCA) ([Harris et al., 2011](#); [Quispe et al., 2024](#)) to develop a comprehensive Heat Vulnerability Index that accounts for spatial non-stationarity in vulnerability relationships.

3.5 Objective 2: Explanatory Modeling

The explanatory modelling framework employs a two-stage clinical-computational approach examining both physiological pathways and temporal effects of heat exposure. This framework harmonizes data from multiple cohort studies to investigate underlying mechanisms across different physiological systems.

3.5.1 Clinical Pathway Validation

Working with clinical experts, we will develop comprehensive physiological pathway diagrams representing hypothesized heat-health mechanisms at multiple biological levels. These diagrams will be rigorously validated through expert consensus panels and systematic literature review, establishing physiologically plausible causal pathways before computational analysis.

3.5.2 Temporal Heat Exposure Modeling

To capture the dynamic nature of heat exposure, we will develop spatiotemporal models that characterize: (1) diurnal patterns of day-night temperature variations; (2) seasonal variations between summer and winter; (3) extreme heat events and their spatiotemporal distributions; and (4) long-term trends under climate change scenarios. These analyses will employ time series modeling, extreme value theory, and climate downscaling approaches.

3.6 Objective 3: Predictive Modeling

The predictive framework integrates environmental, socioeconomic, and health data to forecast heat-related health outcomes within a 24–72 hour horizon. The methodology employs advanced feature engineering combining domain knowledge with statistical selection techniques including mutual information analysis and recursive feature elimination with cross-validation to identify optimal predictive variables while controlling for collinearity.

Rather than relying on a single algorithm, the approach employs a carefully calibrated ensemble that combines gradient boosting methods, random forests, and other algorithms through a meta-learning framework. For the gradient boosting component, the objective function being optimized is:

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

where l is the loss function comparing the prediction \hat{y}_i with the target y_i , and Ω represents the regularization term controlling model complexity.

The validation strategy employs forward-chaining time-series cross-validation to realistically simulate forecasting scenarios, with performance evaluated using metrics including area under the receiver operating characteristic curve (AUC-ROC):

3.6.1 Model Accuracy and Reliability Measures

To ensure predictive model accuracy and reliability, particularly across Johannesburg’s diverse urban contexts, we will implement a comprehensive validation framework. This includes: (1) spatially stratified cross-validation to test model performance across different urban morphologies; (2) out-of-distribution detection methods to identify and address anomalous predictions; (3) uncertainty quantification through conformal prediction intervals; and (4) regular recalibration procedures to maintain model accuracy over time. Our model evaluation will employ metrics beyond traditional accuracy measures (AUC-ROC, F1) to include reliability diagrams, expected calibration error, and variable-specific sensitivity analyses. This approach ensures that predictions remain valid across the city’s diverse built environments and socioeconomic con-

texts, with explicit reliability assessments for informal settlements and areas undergoing rapid urban transformation.

3.7 Forecast Window Selection

The 24–72 hour forecast window was selected based on three key considerations. First, physiological evidence indicates that significant heat-health impacts manifest within 1–3 days of exposure (Gasparri et al., 2015; Kinney et al., 2020), making this the critical intervention window for preventative healthcare measures. Second, meteorological forecast accuracy significantly degrades beyond 72 hours, particularly for urban microclimates, limiting the actionable reliability of longer-term predictions. Third, stakeholder consultations with healthcare providers indicated that a 1–3 day window optimally balances advance warning with operational response capabilities in Johannesburg’s healthcare system. Sensitivity analysis will test alternative windows (12–24 hours and 72–120 hours) to assess the stability of predictive relationships across different temporal scales.

$$\text{AUC-ROC} = \int_0^1 \text{TPR}(\text{FPR}^{-1}(t)) dt \quad (3)$$

where TPR is the true positive rate and FPR is the false positive rate at different classification thresholds.

The modelling output translates into actionable intelligence through stratified risk profiles, geospatial risk mapping, and population-specific threshold recommendations. This translation from statistical prediction to practical application ensures research findings can directly inform heat-health interventions and resource allocation for vulnerable populations in Johannesburg.

4 Ethical Considerations

This study received ethical approval from the Wits Human Research Ethics Committee (reference 220606) and complies with U.S. Department of Health and Human Services regulations (45 CFR 46). Privacy protection includes data minimization, secure servers with restricted access, and geographical jittering/aggregation in compliance with South Africa’s Protection of Personal

Information Act (POPIA, 2013).

For secondary data usage, contractual guarantees from data providers confirm appropriate consent practices. Ethical risks—such as re-identification, secondary consent issues, and community stigmatization—are mitigated through data aggregation, secure storage, broad consent waivers, and responsible community engagement.

Detailed POPIA compliance framework and data protection procedures are provided in Appendix B.

5 Expected Outcomes and Impact

5.1 Research Outputs

This doctoral research will generate several significant outputs, including a spatially explicit heat vulnerability index for Johannesburg, a comprehensive explanatory model of urban heat-health relationships, and a validated predictive framework for heat-health outcomes under varied climate scenarios. The work will also produce evidence-based policy recommendations for city planning and public health systems, culminating in at least three peer-reviewed publications addressing vulnerability mapping, explanatory modelling, and predictive applications.

5.2 Anticipated Impact and Infrastructural Justice Framing

This research will generate both scholarly and applied outcomes, advancing methodology while informing practical adaptation solutions. The primary outcomes include:

1. **Novel Methodological Approaches:** Integrating geographically weighted principal component analysis with causal machine learning for contexts with high socio-spatial inequality.
2. **Evidence-Based Adaptation Framework:** A prioritization system for heat interventions based on vulnerability reduction potential and implementation feasibility.
3. **Policy Translation:** Targeted recommendations for municipal authorities and health systems addressing both emergency measures and infrastructural solutions.

This research advances an infrastructural justice perspective on heat vulnerability ([Romanello et al., 2023](#)), examining how historical development patterns shape contemporary risk

distributions. This framing positions heat vulnerability within socio-political processes that determine whose communities receive protection from climate hazards ([Intergovernmental Panel on Climate Change, 2024](#)).

Research outputs will be made available through open access repositories and disseminated via academic publications, policy briefs, and community materials in local languages. This framing recognizes that vulnerability to heat is not merely a function of temperature but a manifestation of systemic infrastructure decisions that have unequally distributed adaptive capacity across the urban landscape.

The project's findings will inform a justice-oriented approach to infrastructure development that prioritizes addressing historical imbalances in resource allocation, building on methodological advances in identifying causal pathways in environmental justice contexts ([Velasquez et al., 2022](#)). Specifically, results will identify areas where structural interventions (green infrastructure expansion, cooling centers, healthcare facility development) are most urgently needed to reduce disproportionate vulnerability as emphasized in the latest IPCC synthesis report ([Intergovernmental Panel on Climate Change, 2024](#)). This focus on infrastructural justice extends beyond technical solutions to incorporate community governance structures and local knowledge systems in designing context-appropriate adaptations. By framing climate resilience through an equity lens, this research contributes to transformative urban planning that addresses current vulnerability while preventing the perpetuation of historical injustices in future climate adaptation efforts—an approach increasingly recognized as essential for sustainable urban development in the context of accelerating climate change ([Praharaj et al., 2024](#)).

5.3 Knowledge Translation

Research findings will be strategically disseminated through peer-reviewed publications in journals spanning public health, climate science, and urban planning disciplines. Additional dissemination channels include presentations at national and international conferences, targeted policy briefs for local and national authorities, community engagement workshops in vulnerable areas, and interactive data visualization tools to facilitate stakeholder decision-making. Detailed impact assessment metrics and comprehensive dissemination strategies are provided in Appendix

D.

6 Project Timeline

This 36-month PhD research will progress through four key phases, with critical milestones at months 3 (M1: protocol finalization), 9 (M2: data acquisition completion), 18 (M3: vulnerability mapping), and 30 (M4: model validation), culminating in thesis submission at month 36. Figure 4 presents a detailed timeline of research activities including faculty assessment milestones. The research timeline includes four major publications: a protocol paper titled “Leveraging

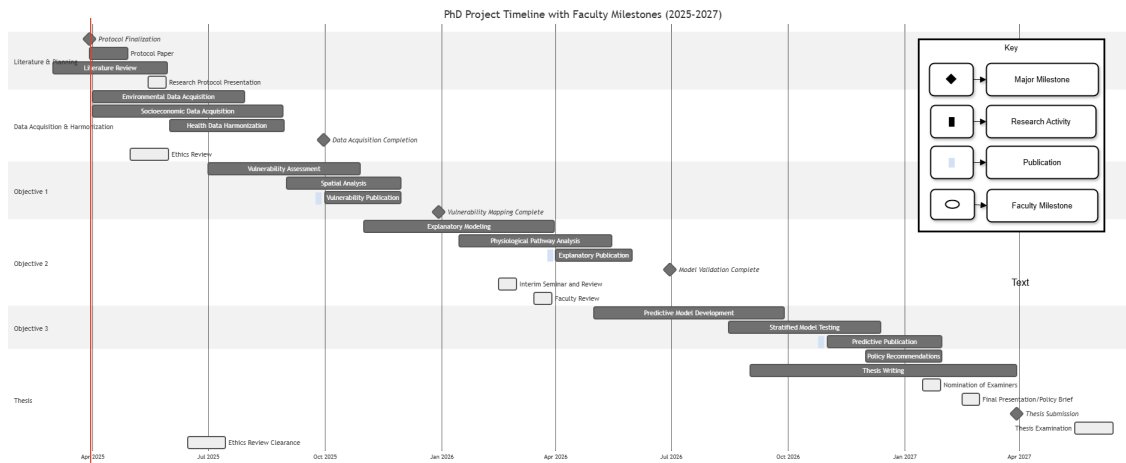


Figure 3: Project Timeline with Key Research Milestones and Publications

data science and machine learning for urban climate adaptation in two major African cities: a HE2AT Center study protocol” [Jack et al. \(2024\)](#) (Month 4), followed by three objective-specific publications: “Quantifying intra-urban socio-economic and environmental vulnerability to extreme heat events in Johannesburg, South Africa” (Month 9), “Uncovering Heat-Health Pathways in Johannesburg: An Explanatory Machine Learning Approach Using Harmonized Urban Cohort Data” (Month 19), and “Developing Stratified Heat-Health Early Warning Systems for Johannesburg: Integration of Vulnerability Mapping and Predictive Modeling” (Month 33). Primary risks include data accessibility challenges, computational constraints, and model performance issues. Detailed timelines, activities, and risk mitigation strategies are provided in Appendix C.

7 Supervision

This research will be conducted under the guidance of a multidisciplinary supervisory team with expertise spanning clinical epidemiology, public health, statistical modeling, and climate science. The team includes Dr. Admire Chikandiwa (Wits University), Professor Matthew Chersich (Trinity College/Wits University), Professor Akbar Waljee (University of Michigan), and Dr. Christopher Jack (University of Cape Town).

Monthly progress meetings and bi-weekly consultations will ensure consistent guidance throughout the research process. A detailed supervision framework, including supervisor profiles and advisory committee information, is provided in Appendix E.

8 References

References

- Ansah, E. and et al. (2024). Health systems response to climate change adaptation: a scoping review of global evidence. *BMC Public Health*, 24(1):2015.
- Copernicus (2020). Climate Data Store. <https://cds.climate.copernicus.eu/>. Accessed: 2025-03-30.
- Engelbrecht, F. and et al. (2015). Projections of rapidly rising heat extremes in south africa. *Journal of Climate*.
- European Space Agency (2015). Sentinel-2 Mission. <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2>. Accessed: 2025-03-30.
- Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklöv, J., Forsberg, B., et al. (2015). Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet*, 386(9991):369–375.
- Gauteng Observatory (2021). Quality of Life Survey. <https://gcro.ac.za/research/project/detail/quality-of-life-survey>. Ward-level data. Accessed: 2025-03-30.
- Harris, P., Brunsdon, C., and Charlton, M. (2011). Geographically weighted principal components analysis. *International Journal of Geographical Information Science*, 25:1717–1736.
- Hersbach, H. et al. (2017). ERA5 Reanalysis Data. <https://cds.climate.copernicus.eu/>. Copernicus Climate Change Service. Accessed: 2025-03-30.
- Intergovernmental Panel on Climate Change (2022). Climate change 2022: Impacts, adaptation and vulnerability. contribution of working group ii to the sixth assessment report of the intergovernmental panel on climate change.
- Intergovernmental Panel on Climate Change (2024). Climate change 2024: Synthesis report. Contribution to the Seventh Assessment Report.

- Jack, C., Parker, C., Kouakou, Y. E., Joubert, B., McAllister, K. A., Ilias, M., Maimela, G., Chersich, M., Makhanya, S., Luchters, S., Makanga, P. T., Vos, E., Ebi, K. L., Koné, B., Waljee, A. K., and Cissé, G. (2024). Leveraging data science and machine learning for urban climate adaptation in two major african cities: a he2at center study protocol. *BMJ Open*, 14(6).
- Khine, M. M. and Langkulsen, U. (2023). The implications of climate change on health among vulnerable populations in south africa: A systematic review. *International Journal of Environmental Research and Public Health*, 20(4):3468.
- Khosla, R., Jani, A., and Perera, R. (2021). Health risks of extreme heat. *BMJ*, 375:n2438.
- Kinney, P. L., Schwartz, J., Pascal, M., Petkova, E., Le Tertre, A., Medina, S., and Vautard, R. (2020). Climate change, heat, and mortality in the tropical urban area of san juan, puerto rico. *International Journal of Environmental Research and Public Health*, 17(11):4037.
- Li, X. and et al. (2017). The surface urban heat island response to urban expansion: A panel analysis for the conterminous united states. *Science of The Total Environment*, 605-606:426–435.
- Murage, P. and et al. (2020). What factors increase the risk of heat-related mortality? a case-crossover study in london. *Environment International*, 134:105292.
- Naicker, N. and et al. (2017). Indoor temperatures in low-cost housing in johannesburg, south africa. *International Journal of Environmental Research and Public Health*, 14(11):1410.
- Parker, C., Chersich, M., Waljee, A., and Jack, C. (2023). Quantifying intra-urban socio-economic and environmental vulnerability to extreme heat events in johannesburg, south africa. *Environment International*, 174:107788.
- Praharaj, S., Choi, M., and Han, H. (2024). A novel urban heat vulnerability analysis: Integrating machine learning and remote sensing for enhanced insights. *Remote Sensing*, 16(16):3032.
- Quispe, E., Smith, J., and Li, W. (2024). Assessing heat vulnerability in philadelphia using

geographically weighted principal component analysis (gwpca): A geospatial big data-driven approach. *Computers, Environment and Urban Systems*, 111:102012.

Romanello, M., Di Napoli, C., Green, C., et al. (2023). The 2023 report of the lancet countdown on health and climate change: the imperative for a health-centred response in a world facing irreversible harms. *The Lancet*, 402(10417):2022–2060.

Santamouris, M. (2015). Analyzing the heat island magnitude and characteristics in one hundred asian and australian cities and regions. *Science of The Total Environment*, 512-513:582–598.

Souverijns, N., De Ridder, K., and Veldeman, N. (2022). Urban heat in johannesburg and ekurhuleni, south africa: A meter-scale assessment and vulnerability analysis. *Urban Climate*, 46:101331.

Strauss, M. (2019). A historical exposition of spatial injustice and segregated urban settlement in south africa. *Fundamina*, 25(2):135–168.

USGS and NASA (2013). Landsat 8 Data. <https://www.usgs.gov/landsat-missions/landsat-8>. Accessed: 2025-03-30.

Velasquez, D., Hayes, A., Sherburne, M., and Hill, J. (2022). Machine learning-aided causal inference framework for environmental justice. *Environmental Science & Technology*, 56(4):2126–2138.

Watts, N. and et al. (2023). The 2023 report of the lancet countdown on health and climate change: health at the mercy of fossil fuels. *The Lancet*, 402(10397):1610–1654.

Wichmann, J. (2017). Heat effects of ambient apparent temperature on all-cause mortality in cape town, durban and johannesburg, south africa: 2006-2010. *Science of the Total Environment*, 587:266–272.

World Bank Cities Support Program (2024). Heat mapping by citizen scientists points the way to a cooler future. Business Day (11 March 2024).

Appendices

NOTE:

All appendices and supplementary tables are provided in a separate document titled:

"Appendices: Data Management and Methodological Tables for PhD Proposal"

This separation has been implemented to comply with page count requirements.

The appendix document includes:

- **Appendix F: Supplementary Figures**
 - F.1 Project Framework: Heat Vulnerability Assessment
 - F.2 Maximum Temperature Distribution
 - F.3 Seasonal Heat Pattern Analysis
 - F.4 Global Temperature Trends
 - F.5 Urban Heat Stress Patterns
 - F.6 Heat Vulnerability Index Components
 - F.7 Heat-Health Mechanisms Pathways
 - F.8 Data Sources and Integration Framework
- **Data Management and Methodological Tables**
 - Table 1: Key Data Sources for Heat-Health Research
 - Table 2: Data Processing and Integration Workflow
 - Table 3: Data Access Levels and Security Measures
 - Table 4: Geographic De-identification Techniques
 - Table 5: Data Request and Approval Process
 - Table 6: Methodological Approaches for Research Objectives
 - Table 7: Time-Lag Analysis of Heat-Health Effects
- **Appendix A: Roles and Responsibilities in Data Management**
- **Appendix B: POPIA Compliance Framework**
- **Appendix C: Detailed Project Timeline and Risk Assessment**
 - C.1 Comprehensive Timeline of Research Activities

- C.2 Detailed Milestone Descriptions
- C.3 Risk Assessment and Contingency Planning
- **Appendix D: Comprehensive Outcomes Impact Assessment and Dissemination Strategy**
 - D.1 Detailed Research Output Framework
 - D.2 Extended Impact Assessment Framework
 - D.3 Comprehensive Knowledge Dissemination Strategy
 - D.4 Key Research Milestones and Quality Assurance
- **Appendix E: Supervision Framework and Team**
 - E.1 Supervisory Team
 - E.2 Supervisor Profiles
 - E.3 Supervisory Arrangements
 - E.4 Advisory Committee