Innovative Data Strategies in Climate and Health:

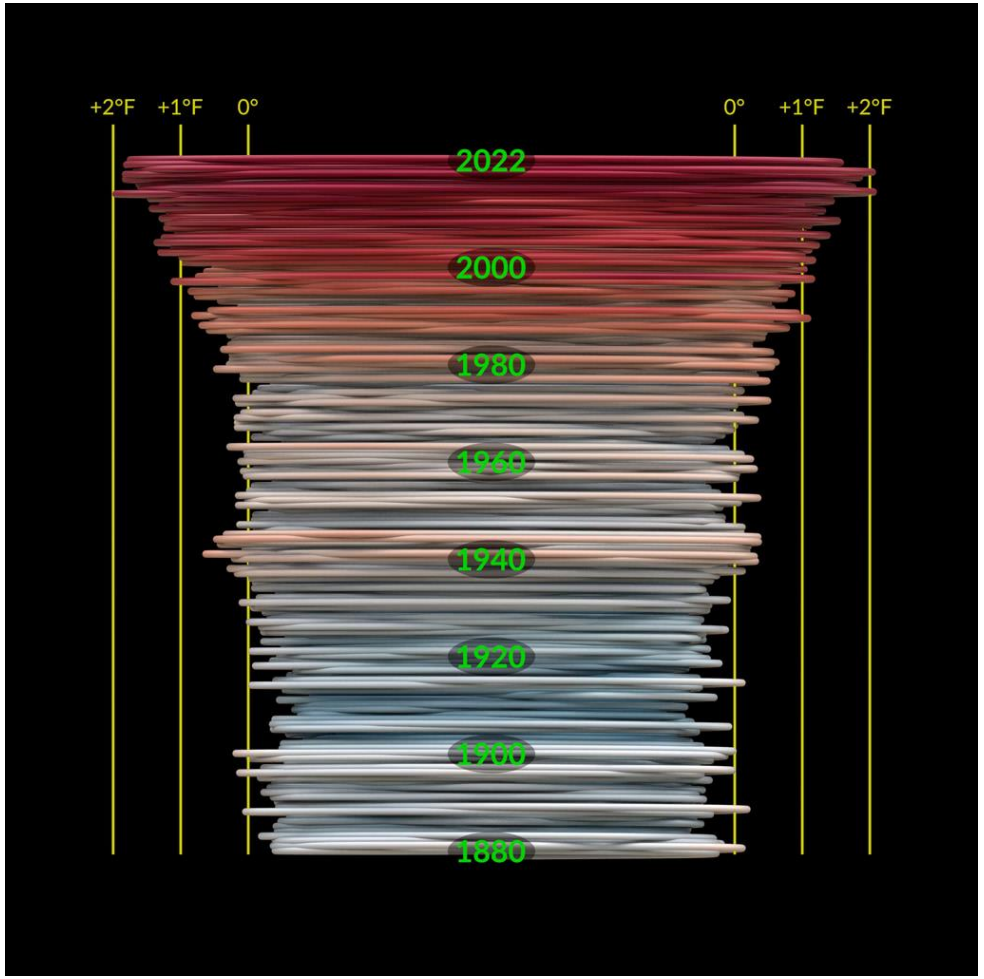
Transforming Research into Action

Presenter: Craig Parker

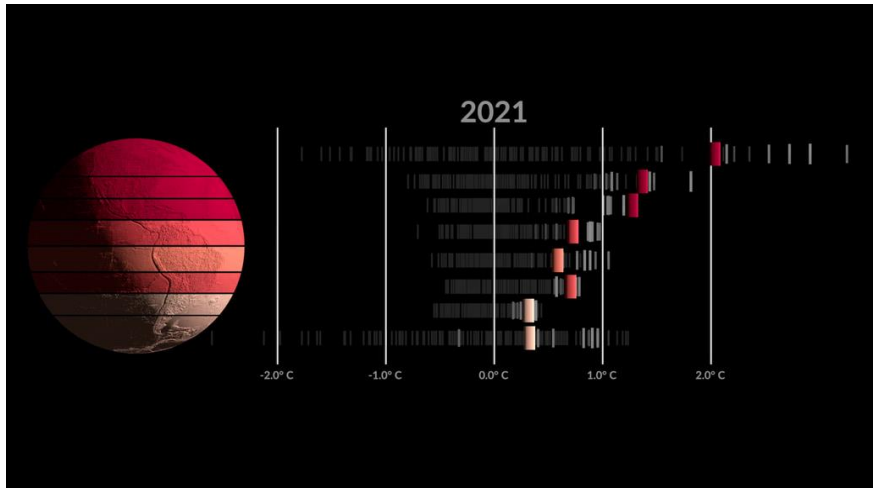
Data Scientist

Wits RHI

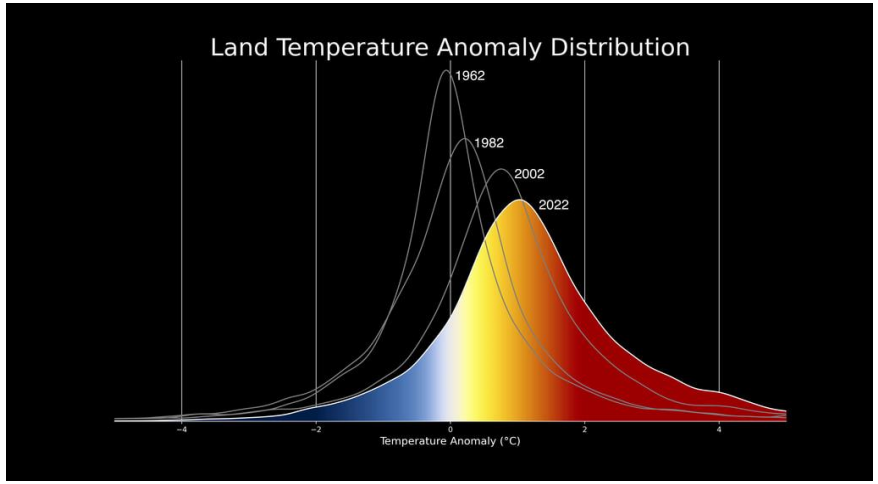
September 5th 2023



The NASA climate spiral 1880-2022



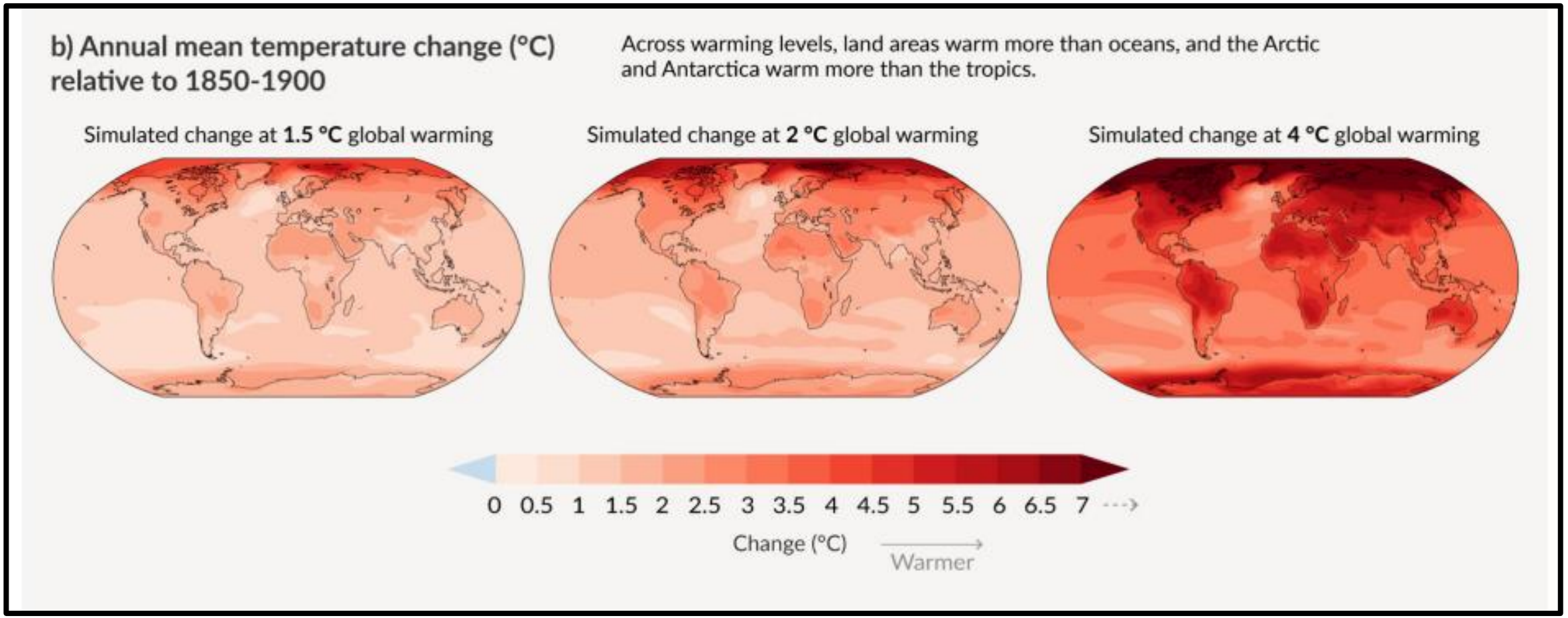
Visualization of temperature anomalies in latitude zones (90N-64N to 64S-90S), calculated against the 1951-1980 baseline

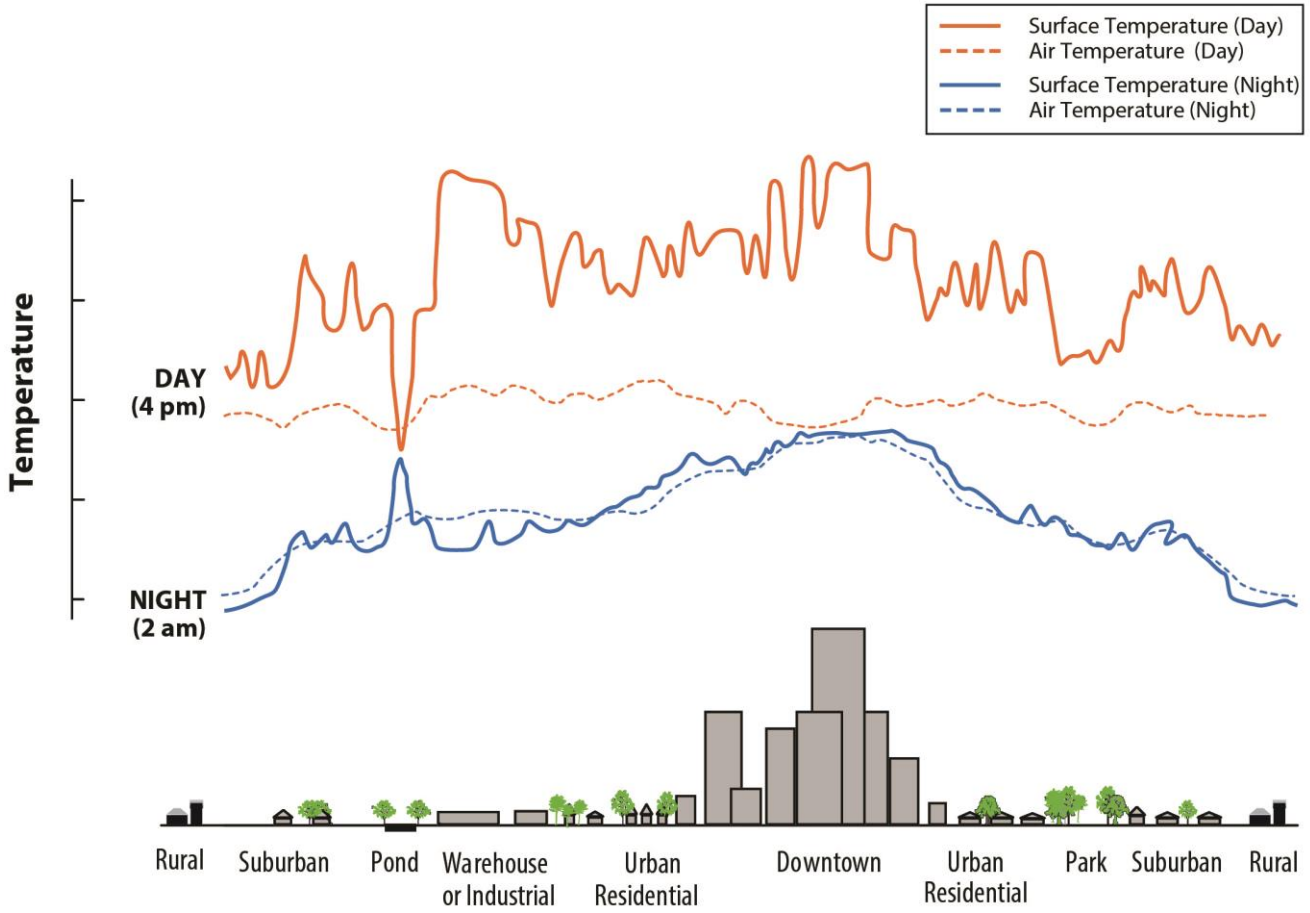


The change in the distribution of land temperature anomalies over the years 1962 to 2022

‘The Heat will kill you first’

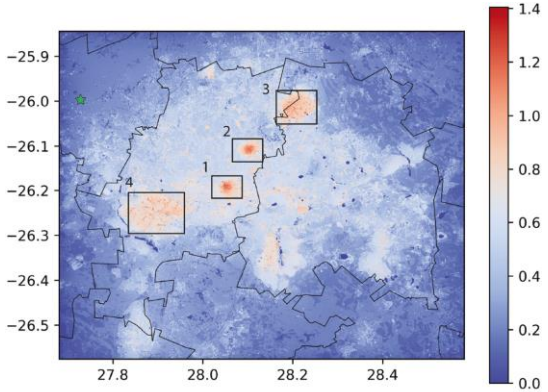
Non-uniform heating of the planet



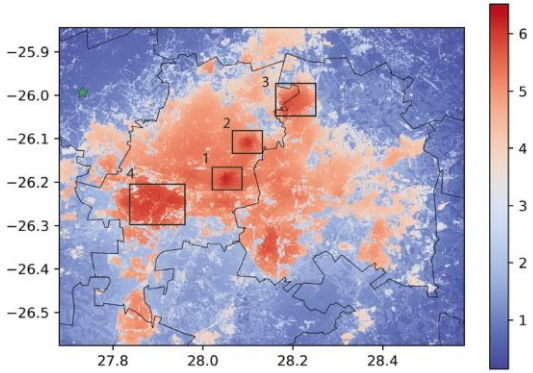


Parks, open land, and bodies of water can create cooler areas within a city. Temperatures are typically lower at suburban-rural borders than in downtown areas. Credit: EPA

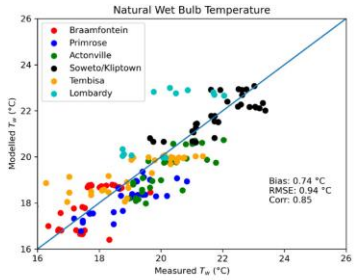
Strong evidence of an Urban Heat Island in Johannesburg



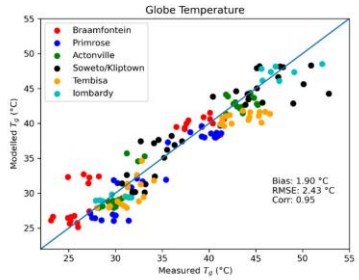
(a) Daytime UHI



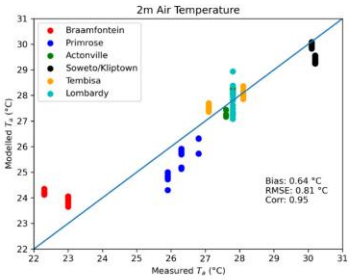
(b) Nighttime UHI



(a) T_w



(b) T_g



(c) T_a

Souverijns, N., De Ridder, K., Veldeman, N., Lefebre, F., Kusambiza-Kiingi, F., Memela, W., & Jones, N. K. (2022). Urban heat in Johannesburg and Ekurhuleni, South Africa: A meter-scale assessment and vulnerability analysis. *Urban Climate*, *46*, 101331. https://doi.org/10.1016/j.uclim.2022.101331

University of the Witwatersrand
WITS RHI

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

Heat impacts on health: example of South Africa

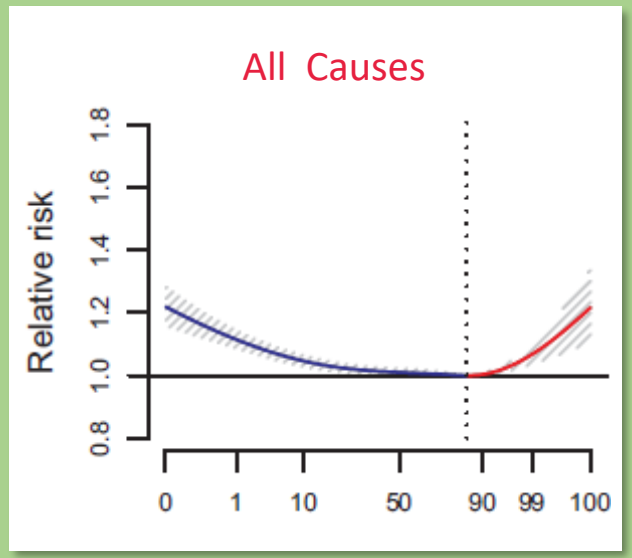
Contents lists available at ScienceDirect

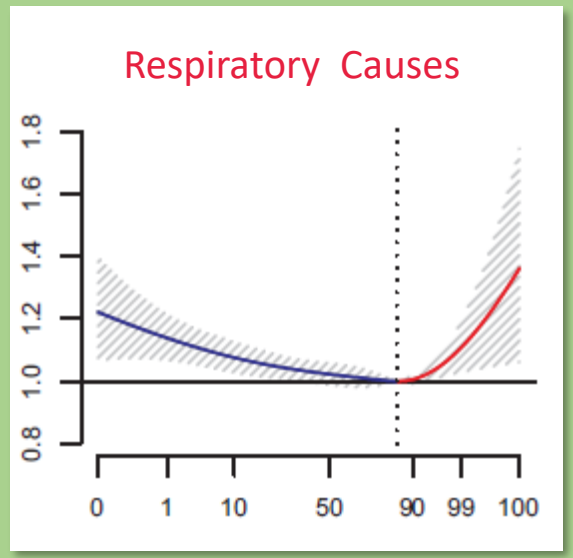
Environmental Research

journal homepage: www.elsevier.com/locate/envres

The association between ambient temperature and mortality in South Africa: A time-series analysis

Noah Scovronick[a,*], Francesco Sera[b], Fiorella Acquaotta[c], Diego Garzena[c], Simona Fratianni[c], Caradee Y. Wright[d], Antonio Gasparrini[b]





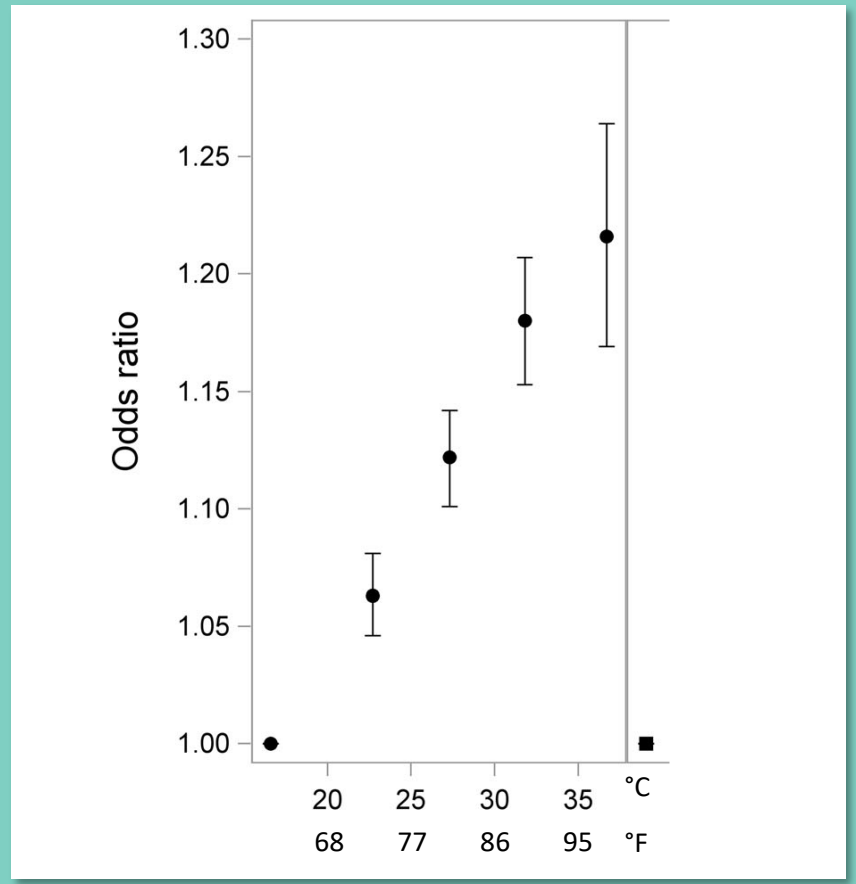
Gates et al. Environmental Health (2019) 18:109
https://doi.org/10.1186/s12940-019-0549-4

Environmental Health

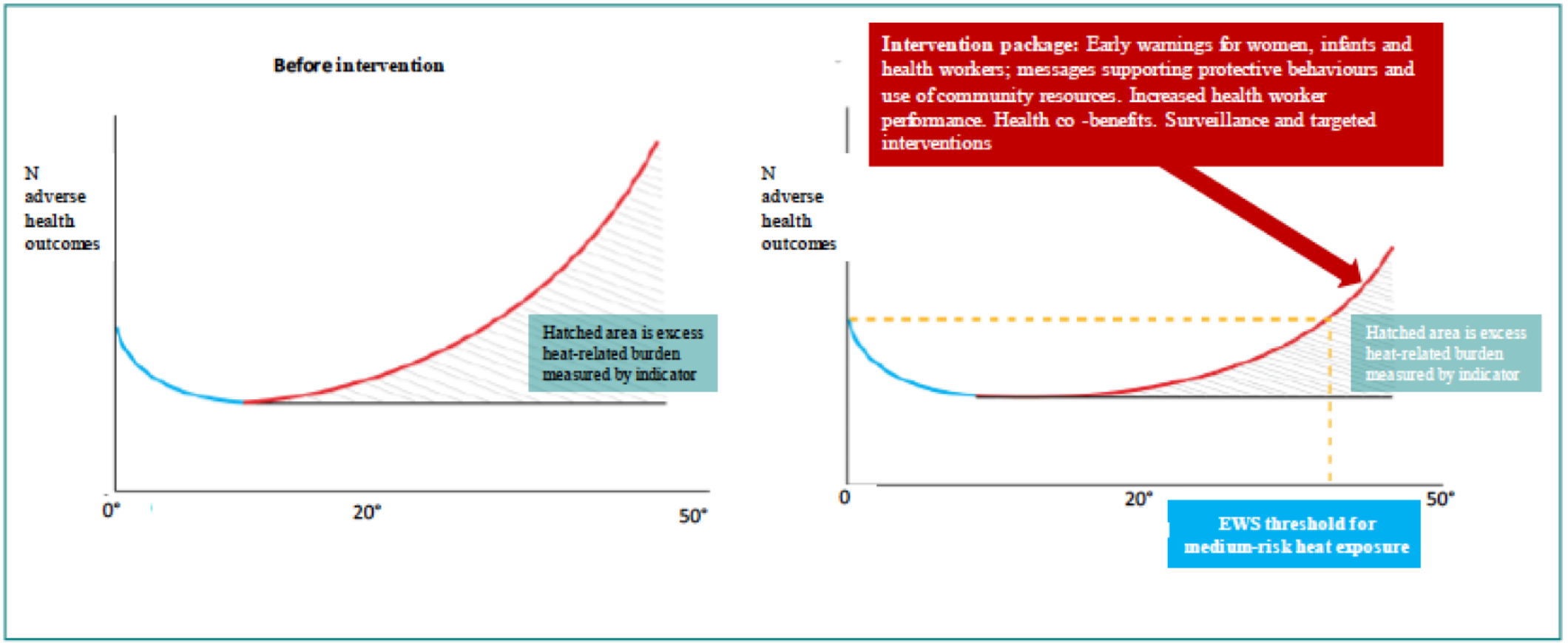
RESEARCH Open Access

Short-term association between ambient temperature and homicide in South Africa: a case-crossover study

Abigail Gates[1*], Mitchel Klein[1], Fiorella Acquaotta[2], Rebecca M. Garland[3,4,5] and Noah Scovronick[1]



‘Flattening the Heat Curve’



Innovative Data Utilization

Heat Impacts: Quantifying effects on pregnant women, neonates; utilizing existing African data, Swedish birth cohort; historical, global insights

- **Large Language Model**: Develop codebook; streamline data harmonisation.
- **Heat Vulnerability Mapping**: Multi-level framework; Principal component analysis.
- **Data Integration**: Satellites, weather, pollution, health outcome data.
- **Geolocation Analysis**: Measure exposure to heat.
- **Machine Learning**: Early Warning System; risk-stratified warnings.
- **Risk Communication**: Smartphone applications.
- **Data Analytics**: Optimize climate & health projects; prototype Early Warning Systems.

- **Adaptation-Mitigation Approach**: Natural solutions; greening, solar-powered cooling.
- **Modelling**: Thermal, Carbon Emissions, Cost-Effectiveness.
- **Heat Stress Warnings**: Smartphone App; time-series, machine learning.
- **Monitoring**: Indoor Climate, Wellbeing.
- **Predictive Modelling**: Utilizing Existing Health Records

- **Comprehensive Approach**: Heat adaptation in Southern Africa.
- **Innovation**: Co-Production; behavioral changes, environment modifications, policy shifts.
- **Multi-Level Intervention**: Individual, household, community, facility, policy levels.
- **Action-Research**: Continuous Improvement.
- **Capacity Building**: Networking.
- **Impact Assessment**: Measurable; Feasibility.
- **Timely Response**: Urban South Africa, rural Zimbabwe; heat-related health risks.
- .

Innovative Data Utilization

CHAMNHA
CLIMATE, HEAT AND MATERNAL AND NEONATAL HEALTH IN AFRICA

Heat Impacts: Quantifying effects on pregnant women, neonates; utilizing existing African data, Swedish birth cohort; historical, global insights

HE²AT center

- **Large Language Model**: Develop codebook; streamline data harmonisation.
- **Heat Vulnerability Mapping**: Multi-level framework; Principal component analysis.
- **Data Integration**: Satellites, weather, pollution, health outcome data.
- **Geolocation Analysis**: Measure exposure to heat.
- **Machine Learning**: Early Warning System; risk-stratified warnings.
- **Risk Communication**: Smartphone applications.
- **Data Analytics**: Optimize climate & health projects; prototype Early Warning Systems.

HIGH Horizons
European Union

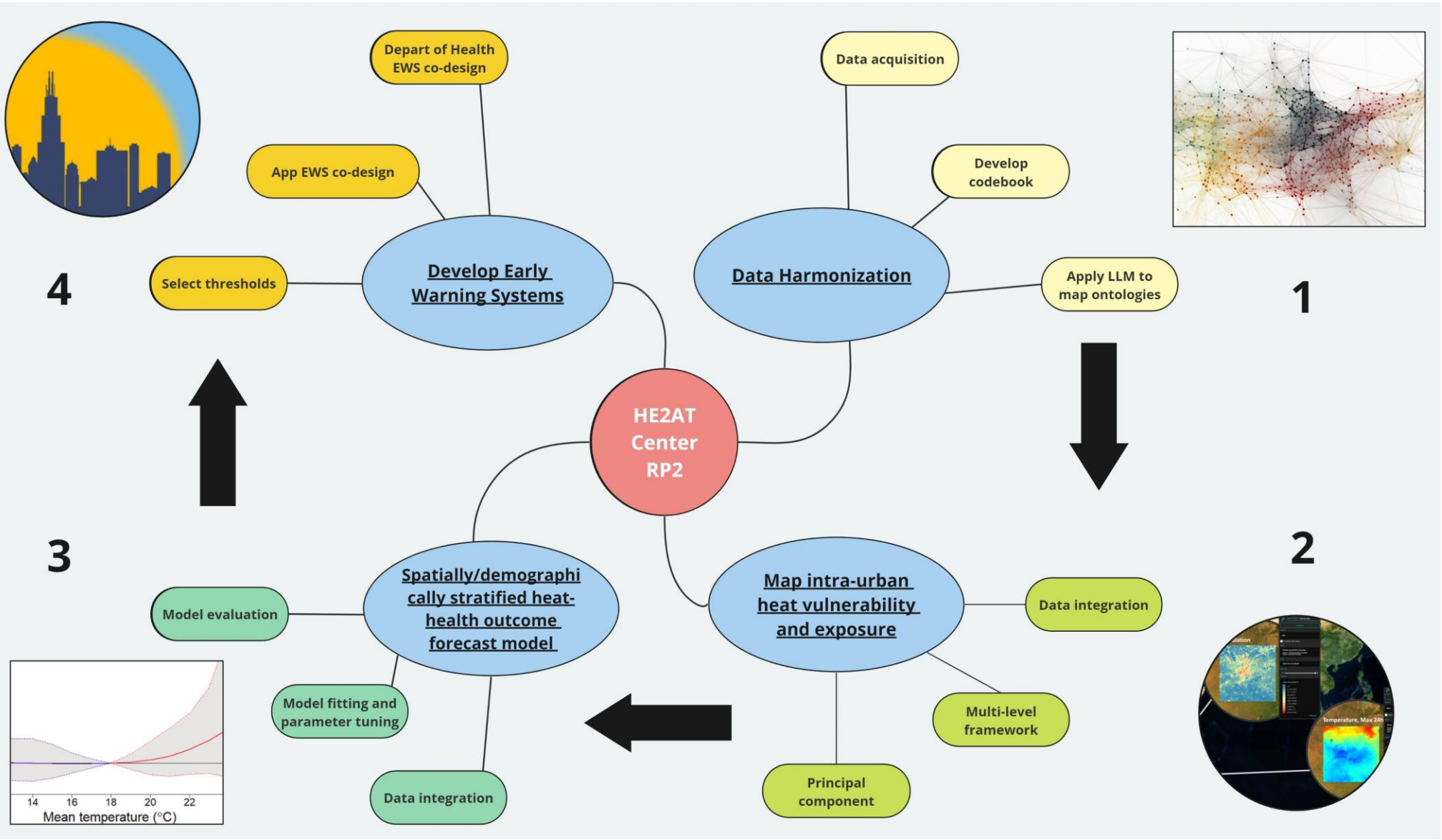
- **Adaptation-Mitigation Approach**: Natural solutions; greening, solar-powered cooling.
- **Modelling**: Thermal, Carbon Emissions, Cost-Effectiveness.
- **Heat Stress Warnings**: Smartphone App; time-series, machine learning.
- **Monitoring**: Indoor Climate, Wellbeing.
- **Predictive Modelling**: Utilizing Existing Health Records

HAPI

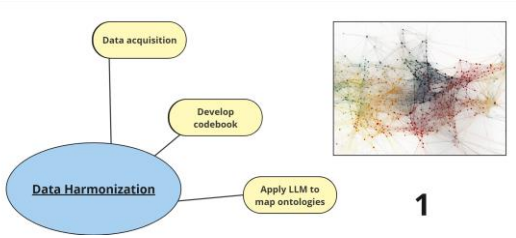
- **Comprehensive Approach**: Heat adaptation in Southern Africa.
- **Innovation**: Co-Production; behavioral changes, environment modifications, policy shifts.
- **Multi-Level Intervention**: Individual, household, community, facility, policy levels.
- **Action-Research**: Continuous Improvement.
- **Capacity Building**: Networking.
- **Impact Assessment**: Measurable; Feasibility.
- **Timely Response**: Urban South Africa, rural Zimbabwe; heat-related health risks.
- .

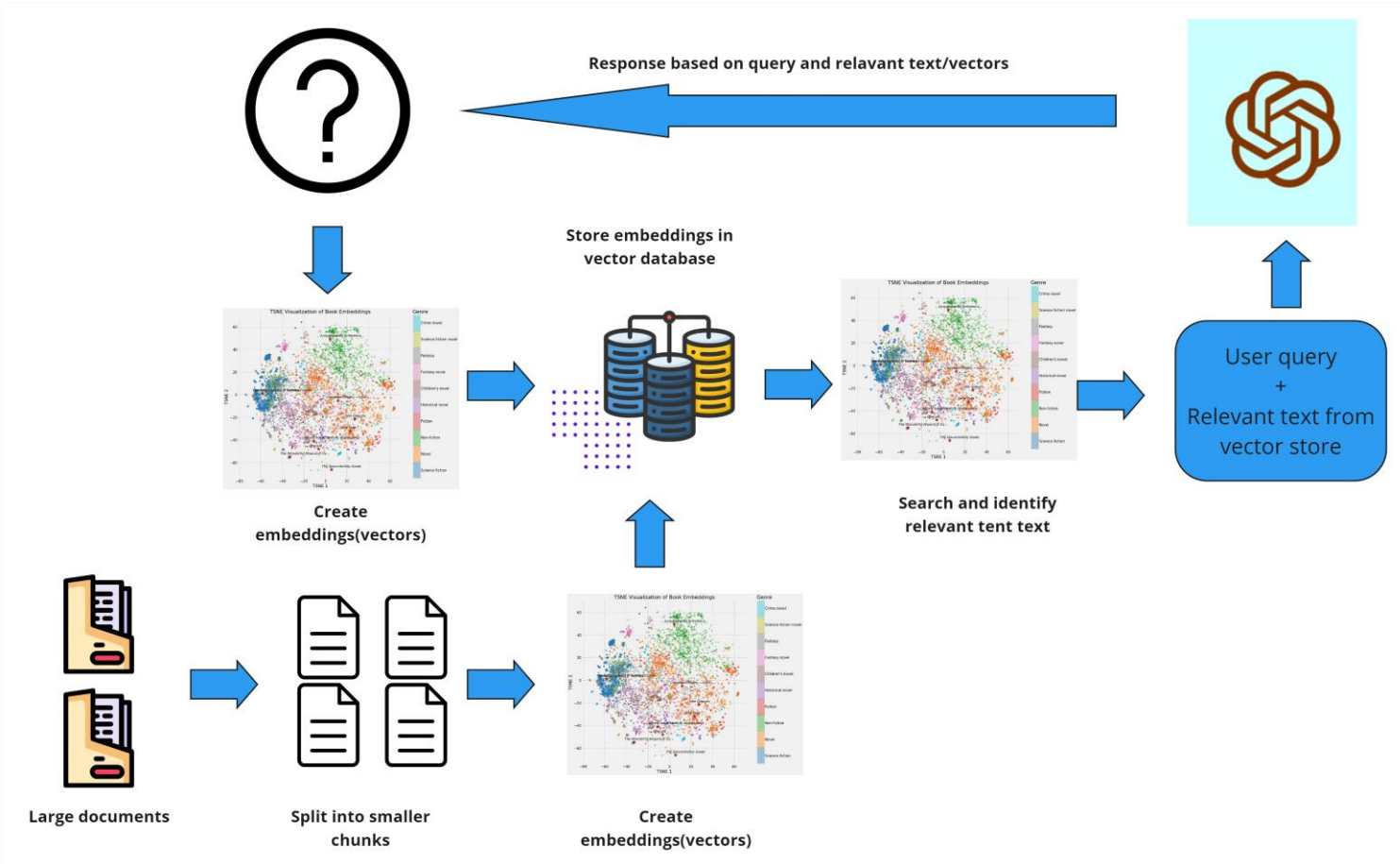
University of the Witwatersrand
WITS RHI

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG


Depart of Health EWS co-design
App EWS co-design
Select thresholds
Develop Early Warning Systems
4
Data acquisition
Develop codebook
Data Harmonization
Apply LLM to map ontologies
1
HE2AT Center RP2
3
Model evaluation
Spatially/demographically stratified heat-health outcome forecast model
Model fitting and parameter tuning
Data integration
14 16 18 20 22
Mean temperature (°C)
Map intra-urban heat vulnerability and exposure
Data integration
Multi-level framework
Principal component
2
Temperature, Max 24h

Applying LLMs to developing a codebook

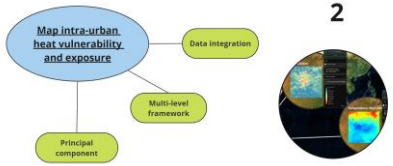


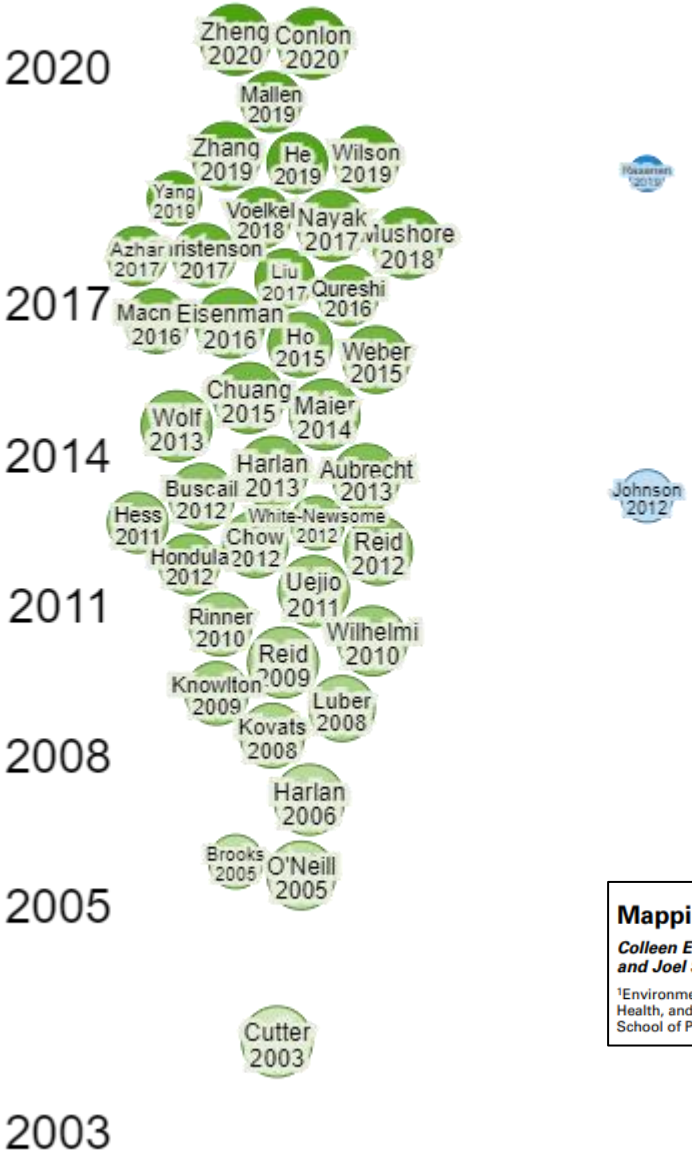


University of the Witwatersrand
WITS RHI

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

Map intra-urban heat vulnerability and exposure





International Journal of *Environmental Research and Public Health*

MDPI

Article

Mapping Heat-Related Risks in Northern Jiangxi Province of China Based on Two Spatial Assessment Frameworks Approaches

Minxuan Zheng [1,2], Jiahua Zhang [1,2,*], Lamei Shi [1,2], Da Zhang [1,2], Til Prasad Pangali Sharma [1,2] and Foyez Ahmed Prodhan [1,2,3]

[1] Key Laboratory of Digital Earth Sciences, Aerospace Information Research Institute (AIR), Chinese Academy of Sciences (CAS), Beijing 100094, China; zhengminxuan19@mails.ucas.ac.cn (M.Z.); lameis@126.com (L.S.); einsteinzhang@foxmail.com (D.Z.); tilsharma@radi.ac.cn (T.P.P.S.); foyez@bsmrau.edu.bd (F.A.P.)
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] Department of Agricultural Extension and Rural Development, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur-1706, Bangladesh
* Correspondence: zhangjh@radi.ac.cn

Received: 24 July 2020; Accepted: 3 September 2020; Published: 10 September 2020

check for updates

Research

A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP4030.

Mapping Human Vulnerability to Extreme Heat: A Critical Assessment of Heat Vulnerability Indices Created Using Principal Components Analysis

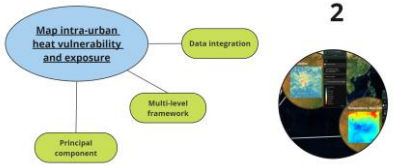
Kathryn C. Conlon,[1,2] Evan Mallen,[3,4] Carina J. Gronlund,[1,5] Veronica J. Berrocal,[6] Larissa Larsen,[3] and Marie S. O'Neill[1]

[1]University of Michigan School of Public Health, Ann Arbor, Michigan, USA
[2]School of Medicine, University of California Davis, Davis, California, USA
[3]University of Michigan Taubman College of Architecture and Urban Planning, Ann Arbor, Michigan, USA
[4]Georgia Institute of Technology School of City and Regional Planning, Atlanta, Georgia, USA
[5]University of Michigan Institute for Social Research, Ann Arbor, Michigan, USA
[6]School of Information and Computer Science, University of California Irvine, Irvine, California, USA

Mapping Community Determinants of Heat Vulnerability

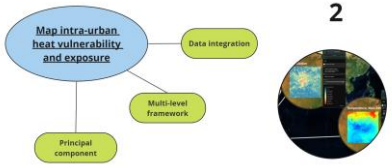
Colleen E. Reid,[1,] Marie S. O'Neill,[2] Carina J. Gronlund,[2] Shannon J. Brines,[3] Daniel G. Brown,[3] Ana V. Diez-Roux,[2] and Joel Schwartz[4]*

[1]Environmental Health Sciences Division, School of Public Health, University of California at Berkeley, California, USA; [2]School of Public Health, and [3]School of Natural Resources and the Environment, University of Michigan, Ann Arbor, Michigan, USA; [4]Harvard University School of Public Health, Boston, Massachusetts, USA

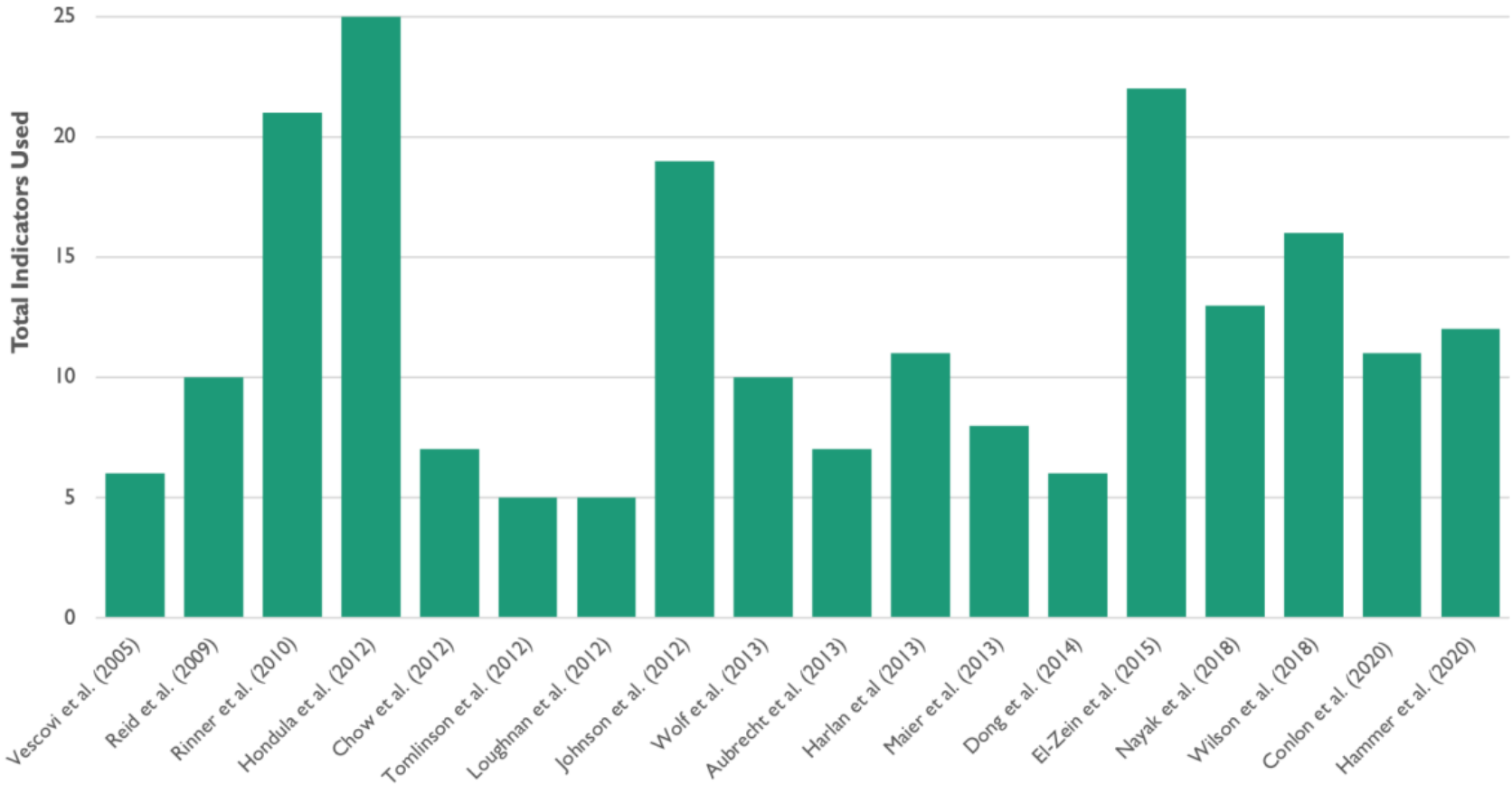


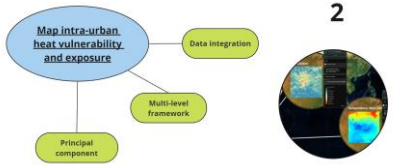
Methods used to create heat vulnerability maps in the literature

Method	Description
Heat Vulnerability Index (HVI)	Utilizes factors like poverty, nonwhite population, social isolation, elderly population, and poor health to determine vulnerability at county level.
Social Vulnerability Index (SoVI)	Constructed from county-level socioeconomic and demographic data to identify social vulnerability to environmental hazards.
Raster-Based Modeling	Integrates heat exposure and vulnerability data to reduce spatial scale issues, enabling fine to coarse spatial mapping.
GIS-Based Spatial Information System	Utilizes remote sensing data and socio-economic information through PCA to create a heat vulnerability index for specific urban areas.
Vehicle-Traverse Collection Method	Records temperature data during heat waves to determine heat exposure at the census block group level. Assesses socio-demographic factors.
Spatial Generalized Linear and Mixed Models	Addresses spatial autocorrelation and determines the importance of exposure, built environment, and socioeconomic vulnerability for heat mortality or distress.

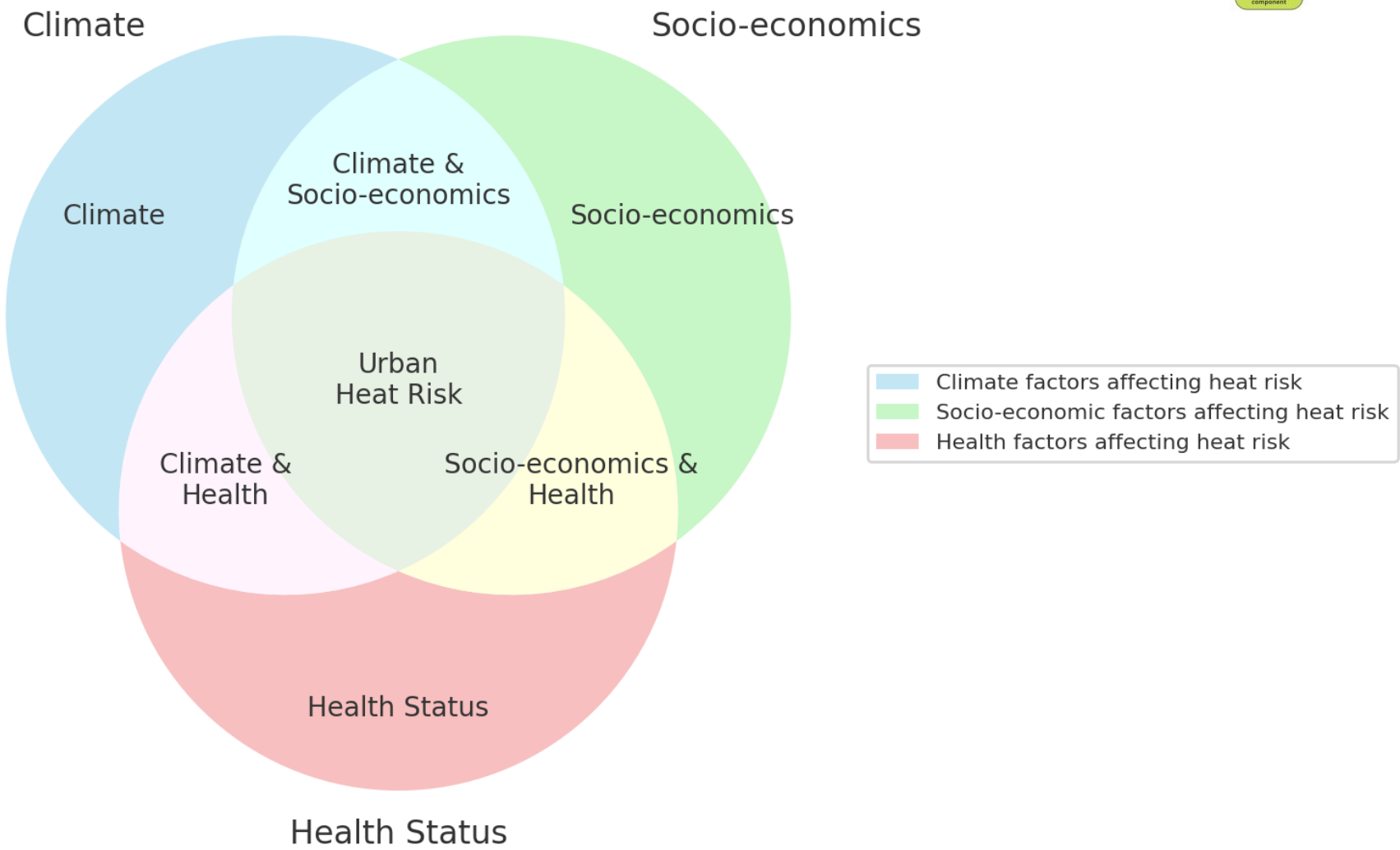


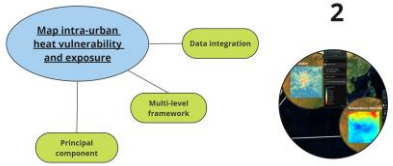
Heat Vulnerability Indicators





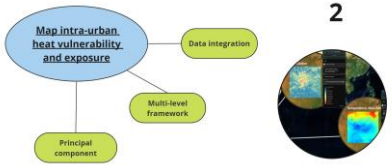
Relationships Between Climate, Socio-economics, and Health Status in Urban Heat Risk





Exposure	Sensitivity	Adaptive Capacity	= Vulnerability
Hot/heatwave days	Older adults	Air conditioning	
Consecutive hot days	Infants, young age	Living alone	
Min/Mean/Max temp	Sex	Inome/wealth	
Impervious surfaces	Diabetes	Rental/homeowner	
Vegetation	Cardiopulmonary	Unhoused	
Urban density	Renal Respiratory	Education	
Land cover	Obesity	Ethnicity	
Land use		Language	
Homes without AC		Cognitive impairment	
Population density		Mobility/transportation	

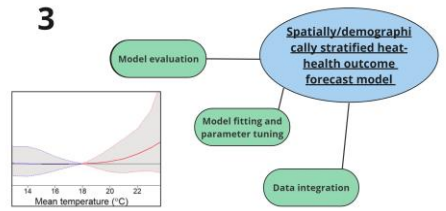
Our approach so far

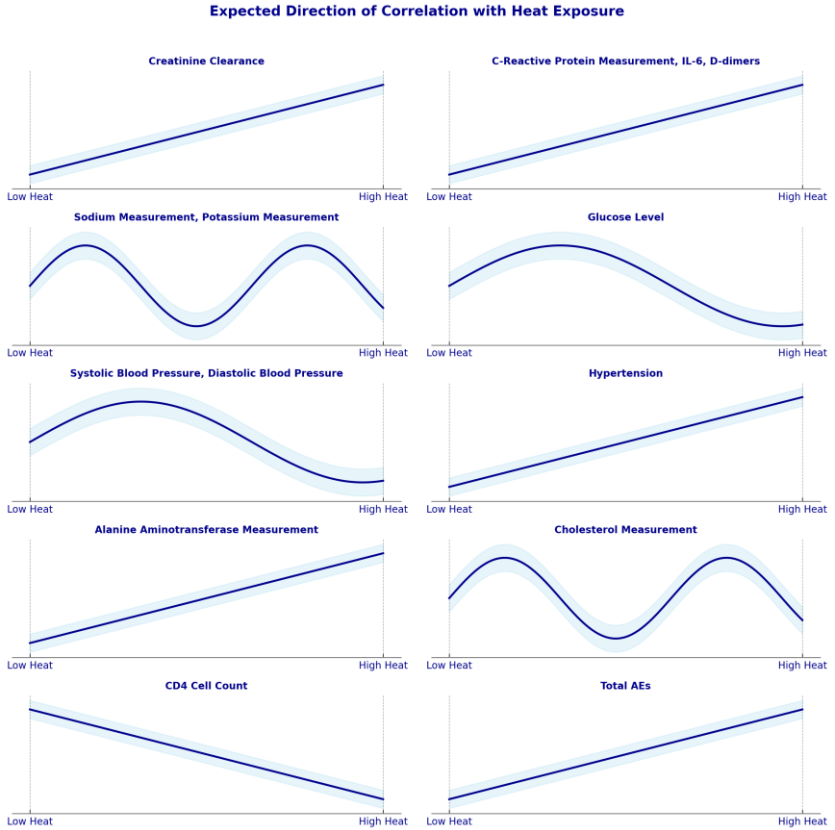


- Insert GEO-den images here (create process map)

Heat Health Hypotheses

Variable	Hypothesis
Creatinine Clearance	Heat exposure may adversely affect renal function, with an increase in creatinine indicating potential kidney dysfunction.
C-Reactive Protein Measurement, IL-6, D-dimers	Increased levels may be observed following heat exposure, indicating an inflammatory response to heat stress.
Sodium Measurement, Potassium Measurement	Abnormal electrolyte levels may reflect dehydration associated with heat exposure.
Glucose Level	Heat exposure may negatively impact blood sugar control in individuals with diabetes.
Systolic Blood Pressure, Diastolic Blood Pressure	Decreases in blood pressure may be associated with increasing temperatures.
Hypertension	Individuals with hypertension may experience exacerbated symptoms during high heat.
Alanine Aminotransferase Measurement	Elevated levels may indicate liver stress related to heat exposure.
Cholesterol Measurement	Changes may be observed in response to heat exposure, reflecting potential impacts on lipid metabolism.
CD4 Cell Count	CD4 cell count levels may reduce at higher temperatures.
Total AEs	Heat exposure may increase the risk of adverse events.

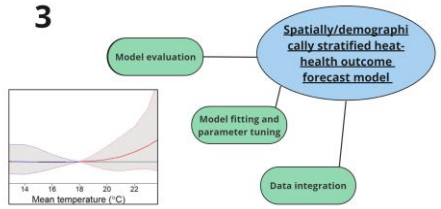




University of the Witwatersrand
WITS RHI

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

Statistical and Machine Learning Methods to explore and model he and health associations



Data preparation

Data Integration

- Clinical trials
- Weather
- Climate

Dataset Cleaning

- Missing values
- Data Transformation

Split the data for test and validation

- Train(80%)
- Test(20%)

Modeling

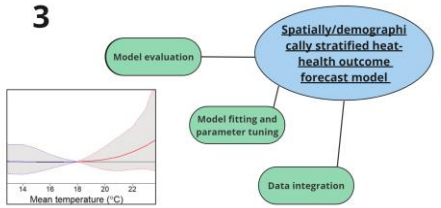
Model fitting and parameter tuning

- **Tree-based methods**
  - Decision tree
  - Random forest
  - Extreme Gradient Boosting
- **Neural networks**
  - Recurrent Neural Networks (RNNs)
  - Long Short-Term Memory Networks (LSTMs)
- **Statistical models**
  - Generalized Additive Models (GAM)
  - Distributed Lag Non-linear Models (DLNM

Results

Model evaluation

- **Explainability**
  - Direct interprebility
  - Feature importance
- **Performance**
  - R squared RMSE MAE
  - F -score
- **Partial validation**
  - Heatwave validation



Interpretability

Generalised linear models

Decision trees

K Nearest Neighbours

Random Forest

Support Vector Machines

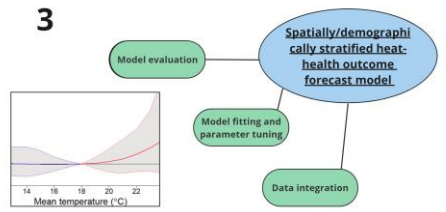
XGboost

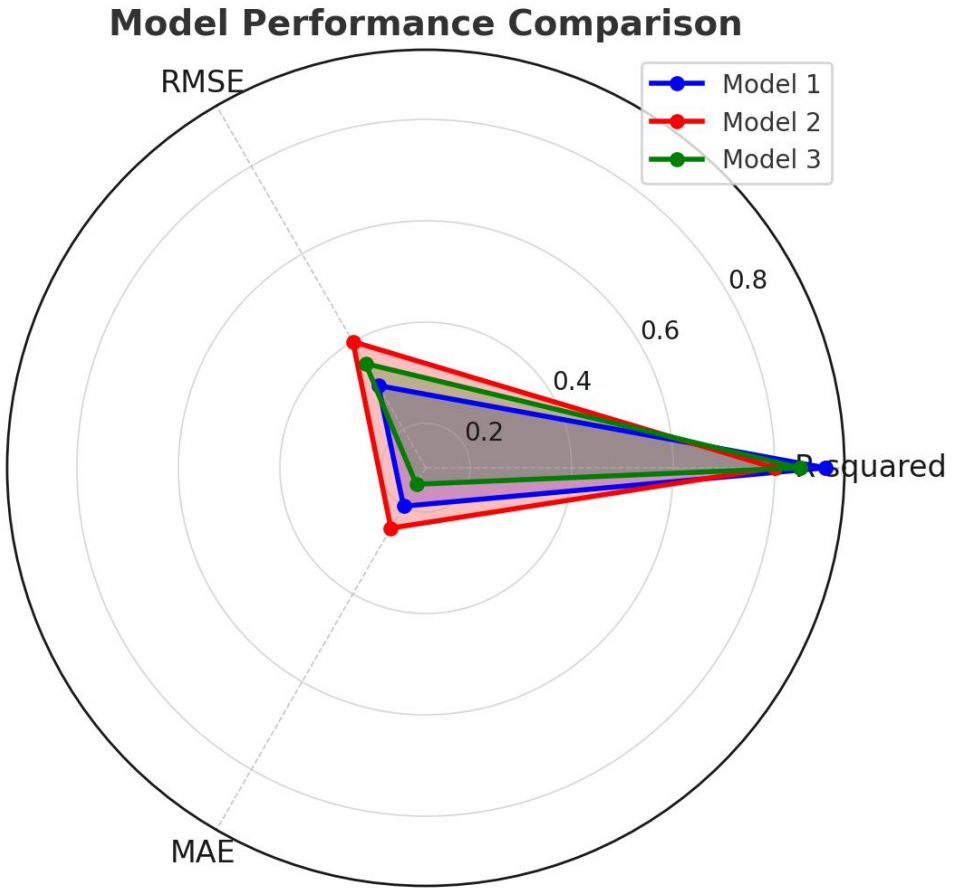
Neural Networks

Predictive Performance

Ciatto, Giovanni & Omicini, Andrea & Schumacher, Michael & Calvaresi, Davide. (2020). Agent-Based Explanations in AI: Towards an Abstract Framework.

Model performance comparison





University of the Witwatersrand
WITS RHI

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

Early warning system

Acknowledgments