Proposal for Academy Projects
Leverage prior analyses from the Academy - Ajadi Clearly indicate how the project builds on work done during the course and what new elements will be added.
Define a focused research question or hypothesis - Elisabeth - added it below Ensure the research direction is specific, feasible, and grounded in the existing context of the project.
Outline project objectives and expected outcomes - Helyne Describe what the project aims to achieve and ensure the goals are realistic within the ISP's 6-month timeframe.
Define the methodology and tools - Sharon Summarize the analytical approaches and tools used during the Academy, and explain how they will be extended or refined.
Describe the potential impact of the research - Helyne Outline the expected contributions and explain how your findings may be relevant or useful to the scientific or applied community.

Grading:

Category	What it evaluates
Motivation & Research Significance	Does the proposal clearly state why the research question is important? Is the problem well-defined and relevant to the field? Does it demonstrate an understanding of previous work (with references if needed)?
Methodology & Feasibility	Are the data sources, tools, and analysis methods clearly described and appropriate? Is the approach technically sound and achievable within the ISP timeframe? If a partner dataset is used, does the proposal effectively leverage it?
Preliminary Work & Supporting Evidence	**For Academy projects:** Are there relevant preliminary results with key figures that support feasibility? **For Partner projects:** Does the proposal demonstrate understanding of the dataset/project and outline expected results?
Expected Outcomes & Impact	Are the expected outcomes clearly defined? Will the project contribute meaningfully to scientific knowledge, practical applications, or societal impact? Does it align with the ISP's goals?
Proposed Work & Timeline	Does the proposal outline a realistic and structured work plan for the six-month ISP period? Is the scope appropriate given time and computational resources?

Core study questions:

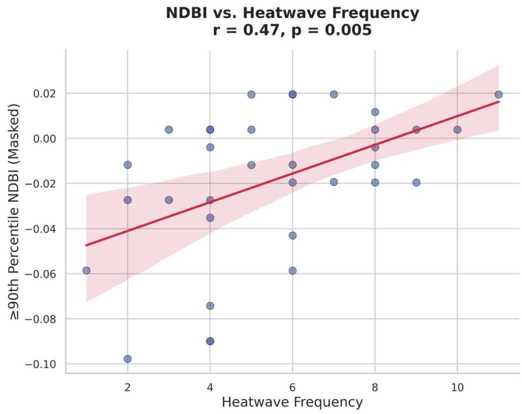
- How transferable are the drivers of urban heat islands across diverse climatic and urban contexts?
- **Are the factors that drive extreme heat universal, or are they specific to a city's geography, climate, and development pattern?**

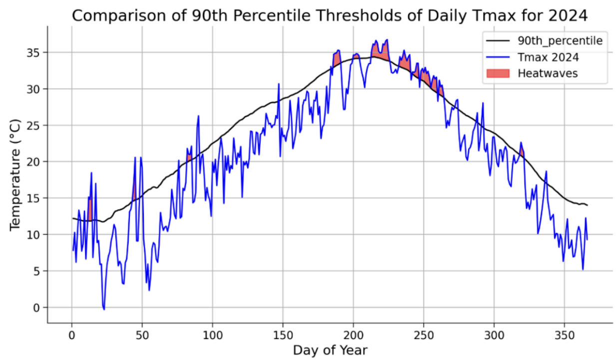
Connection with previous study:

- Pivot in focus from temporal trends to spatial mechanisms
- Logical progression from correlation analysis to prediction and causal explanation
- We saw urban development correlate with heatwave frequency -> now, for a snapshot in time, what are the precise spatial mechanisms driving the worst heat within a city?

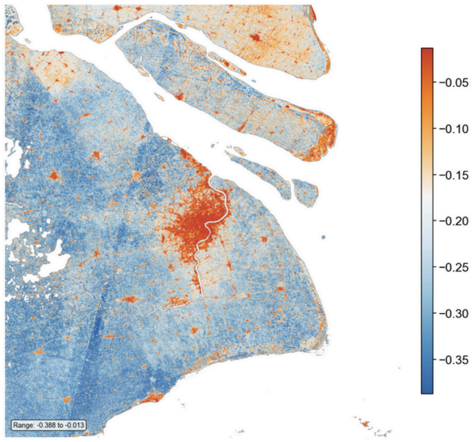
A) Leverage prior analyses from the Academy

Based on the initial proposal submitted during the climatematch programme, the Poisedon group worked on *"Concrete Heat: Contributions of Urban Land Use to Heatwave Intensity in Shanghai, China"*, which will serve as a foundational framework for our current research. In that project, we explored the relationship between impervious surfaces and heatwave intensity in Shanghai, China, by using remote sensing indices (Landsat 5 to 9) and ERA5-Land data. This initial phase focused on correlation analysis and heatwave characterization, providing essential insights into the spatial and temporal dynamics of urban heat (see Fig 1). During our presentation, the audience provided valuable suggestions, and after reviewing our work post-programme, we also identified some limitations. These observations motivated us to further develop and expand the research through the Impact Scholarship programme.





Copernicus Era 5 Land, 2m daily maximum data



Normalized Difference Built-Up Index (NDBI) (1991)

Fig. 1 Results from our initial study. Previously, we established a temporal relationship between urbanization and heatwave trends in Shanghai, finding that increased Normalized Difference Built-Up Index (NDBI) correlated with increased heatwave frequency.

In preparation, we reviewed several related studies with similar research aims and objectives. Notably, a recent study in Da Nang, Vietnam (Hoang et al., 2025), employed a two-step machine learning approach to map and analyze urban heat hotspots during heatwave events and identify the most influential urban and environmental factors. Interestingly, they also found NDBI to be a key driver of urban heat. Building on this, our research replicates this methodology in Ouagadougou, Burkina Faso, to examine urban heat dynamics in a Sahelian context and compare key drivers of hotspot formation between the two cities. *Building on this, we decided to focus on Ouagadougou, Burkina Faso, due to the recent reports of heatwaves in the country as well as its history of extreme heatwave events.* Unlike our previous proposal, this new study will incorporate additional variables such as built-up density, green space density, elevation, and proximity to water bodies, coastlines, and roads. These are some of the variables that were not comprehensively considered in our earlier work. We also plan to include a reference rural area to standardize heatwave intensity calculations, thereby improving comparability and accuracy. Furthermore, we will apply machine learning approaches, specifically XGBoost and Random Forest, to compare the performance of these models in predicting future heatwave intensity in the study area. This will provide an avenue for us to implement some of the computational tools we were introduced to during the summer school.

Ultimately, this approach will enable us to develop a robust framework that integrates machine learning with Earth Observation variables to improve the understanding of spatial variations in heatwave intensity across the study area. The significance of this research lies in its potential to serve as an early warning tool for urban planning authorities, helping them identify vulnerable hotspots. In addition, the study will provide evidence-based adaptation and mitigation strategies aimed at reducing the impacts and footprint of heatwaves in Burkina Faso.

B) Define a focused research question or hypothesis

This comparative study replicates the methodology of the Da Nang study in Vietnam and applies it to Ouagadougou, Burkina Faso, to investigate urban heat hotspots in a Sahelian urban context. Specifically, it aims to evaluate the effectiveness of machine learning models in identifying hotspots during heatwave events, determine the most influential environmental and urban factors driving hotspot formation, and directly compare these results with those from Da Nang to evaluate which patterns are consistent and which differ due to Ouagadougou's specific climate and urban structure.

Hypothesis: We expect to find similar drivers of hot spots within the city of Ouagadougou as Da Nang - in particular, we expect hottest spots within Ouagadougou are driven by factors like **built-up density** and lack of green space.

C) Outline project objectives and expected outcomes

This project aims to build directly on our Climatematch Academy work, which investigated the correlation between impervious surfaces and heatwave intensity in Shanghai over time. Previously, we established a temporal correlation between urbanization and heatwave trends in Shanghai, and we will now build on this work by advancing from correlation to prediction and causal explanation. We move from asking "if" urban factors correlate with heat to "how, where, and why" they drive extreme heat risk at a local scale. We aim to determine the core drivers of intra-urban heat patterns in Ouagadougou, Burkina Faso, and seek to test whether these drivers are consistent across cities with vastly different urban environments. As a stretch goal, we would then perform a simple policy scenario test (e.g., adding green space) to simulate the effects of different urban planning interventions.

Our specific objectives for the 6-month program are:

1. Replicate and adapt Hoang et al. (2025) methodology for Ouagadougou
2. Identify and rank key drivers of urban heat vulnerability
3. Compare and interpret key drivers between cities
4. Generate actionable outputs (e.g. for urban planning)

Our specific expected outcomes from this project are:

- A processed geospatial dataset and a validated, interpretable machine learning model for Ouagadougou.
- A comparison report between Ouagadougou and Da Nang, highlighting universal vs. localized drivers of urban heat, and a qualitative comparison with our previous results from Shanghai.
- A public-facing heat vulnerability map for Ouagadougou.
- A final micropublication and seminar presentation.

Our scope of work focuses on a spatial snapshot analysis for one primary city (Ouagadougou), using the proven analysis methods from Hoang et al. (2025) instead of developing a novel technique from scratch. Our team has already demonstrated proficiency with satellite images and developed a preprocessing pipeline for Landsat data using Google Earth Engine (GEE) during the Climatematch Academy, providing a strong foundation for this follow-up study. Using these already established methods will keep our work feasible within the 6-month timeline.

D) Define the methodology and tools

The study area

We selected the capital city of Ouagadougou, Burkina Faso, a major urban center in West Africa's Sahel region, as our comparative study area. Crucially, Ouagadougou shares key characteristics with Da Nang (e.g. similar latitude, population size, and a significant March-May 2024 heatwave), while representing a starkly different environment. Unlike coastal Da Nang's humid climate, Ouagadougou's semi-arid inland conditions create a climatic contrast, allowing us to investigate which urban heat drivers are universal or context-specific. This comparison can reveal how local climate modulates urban heat island effects.

Research material

To characterize heatwave dynamics and intra-urban variations in temperature, we will collect and preprocess a combination of remote sensing and vector-based spatial datasets (summarised in Table 1). We will primarily use satellite imagery from Landsat 8 and Sentinel-2, digital elevation data from Copernicus, and road network data extracted from OpenStreetMap.

Table 1 Research materials and data sources

Dataset	Retrieved period	Bands	Resolution	Data source
Landsat 8 Level 2, Collection 2, Tier 1	03/01/2024–05/30/2024	SR_4, SR_5, and ST_B10	30 m	Google Earth Engine's data catalog
NASA SRTM Digital Elevation 30 m		Elevation	30 m	Google Earth Engine's data catalog
Sentinel-2	01/01/2024–12/31/2024	B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12	10 m (B2, B3, B4, and B8) 20 m (B5, B6, B7, B8A, B11, and B12)	Google Earth Engine's data catalog
Road vector data	–	–	–	OpenStreetMap extracted via https://extract.bbbike.org/

Table 1 - Research material and data sources

Surface reflectance and surface temperature products (30 m) from Landsat-8 and Sentinel-2 provide fine-scale spatial information of both land surface temperature (LST) and land use land cover (LULC) indices, including urban, vegetation, water, and soil indices (e.g. NDBI, NDVI, MNDWI, and Bare Soil Index). Topographical features are derived from the Copernicus global digital elevation model (DEM) at a resolution of 30m. OpenStreetMap (OSM) vector data is used for urban infrastructure features, such as proximity to roads and water bodies. Global Human Settlement Layer (GHSL) datasets are used for population metrics.

The analysis methods

We will apply a machine learning approach to examine the precise spatial drivers of intra-urban heat patterns during extreme events and then test whether these drivers are consistent across two different urban environments. The study will replicate the XGBoost and SHAP analysis framework from Hoang et al. (2025) for Da Nang, Vietnam in 2024 (as shown in Fig. 2), and adapt it for the novel context of Ouagadougou, Burkina Faso, a rapidly urbanizing Sahelian city that also experienced severe heatwave events in 2024. We will test the generalizability of urban heat drivers by comparing the results from Ouagadougou with the published results from Da Nang, Vietnam.

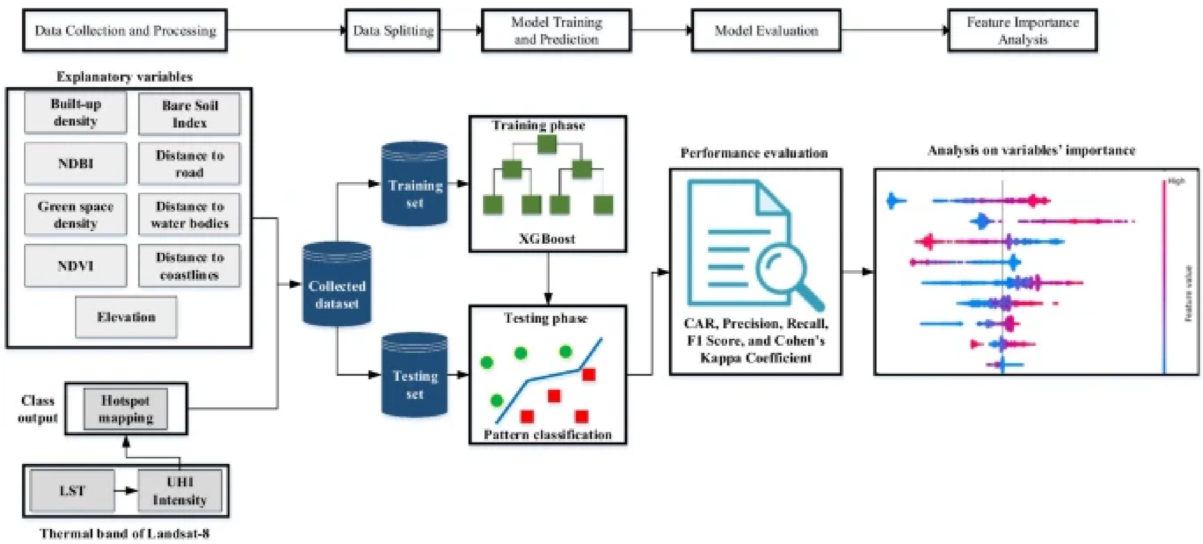


Fig. 2 - Schematic of methodological approach for hotspot mapping adapted from Hoang et al. (2025)

Heatwave days are identified using the ETCCDI framework, specifically the TX90p and Warm Spell Duration Index (WSDI), calculated from ERA5-Land daily maximum temperatures against a 1991–2020 baseline climatology. Events are defined as periods with at least six consecutive days exceeding the 90th percentile threshold. For each identified heatwave event (e.g., March–May 2024 and March 2025), corresponding cloud-free Landsat scenes are composited to generate representative LST fields.

Heatwave intensity is quantified as the LST anomaly relative to a rural reference area located 15–30 km outside Ouagadougou, restricted to non-urban land cover types. Two target variables are constructed: (i) a continuous anomaly field (°C), and (ii) a binary hotspot classification, where pixels exceeding the 85th percentile of anomaly within the city are labeled as hotspots. This dual target approach enables both regression and classification analyses.

Predictors are selected to represent the physical and socio-environmental drivers of heatwave amplification. At the pixel level, these include: NDVI (vegetation density), NDBI (built-up density), MNDWI (water presence), elevation and slope, local green space density (share of pixels with NDVI > 0.3 within a 300 m window), built-up density (GHSL share within 300 m), and Euclidean distances to major roads and surface water bodies. Population density is considered as an auxiliary variable to examine exposure patterns.

To predict and interpret spatial variations in heatwave intensity, we will employ machine learning models. Model training uses stratified sampling of hotspot and non-hotspot pixels, with spatial block cross-validation (2 km blocks) to reduce spatial autocorrelation bias. Performance metrics include accuracy, F1-score, and Cohen's Kappa for classification, and R^2 and mean absolute error (MAE) for regression. Model interpretability is assessed using SHapley Additive exPlanations (SHAP), enabling both global rankings of variable importance and local attributions at the neighborhood scale.

Our multi-city comparison will analyze the relative importance and effect direction of urban heat drivers (e.g., built-up density, green space) during similar 2024 heatwaves. We will directly compare the ranked drivers of heat hotspots from our Ouagadougou analysis against the published results from Da Nang, Vietnam. By contrasting the SHAP-derived variable importance rankings and effect directions between the two cities, we will identify which drivers are universal and which are unique to each city's local climate and urban form. Furthermore, by qualitatively comparing these results with our prior analysis of Shanghai - which established a strong temporal link between urbanization (NDBI) and heatwaves - we can assess if the primacy of built-up areas is a consistent driver across all three cities.

Stretch goal: To test mitigation policies, we will digitally alter Ouagadougou's input maps - for example, artificially increasing the 'green space density' value by 10% in the city center - and then feed this modified map into our already-trained model to generate a new heat risk forecast. Comparing this new forecast to the original map will show us exactly how much and where the heat risk would be modified by an intervention.

E) Describe the potential impact of the research

Our study will have both scientific and applied policy impacts. From a scientific perspective, this study will contribute a new data point from the under-studied Sahelian urban context, testing the generalizability of machine learning frameworks for urban heat island analysis. Our comparative approach will help move beyond single-city case studies towards a more nuanced, global understanding of how urban heat drivers vary by geography, climate, and urban structure. Determining which drivers are consistent (e.g., built-up density) and which are context-specific (e.g., distance to coastline) is important for building robust, global theories of urban heat island formation. Furthermore, the use of SHAP analysis ensures our model is interpretable, moving from a "black box" prediction to a clear, causal explanation of complex urban systems. This provides a replicable methodology for other researchers to apply in under-studied cities, particularly in the Global South.

From a policy perspective, our research has the potential to assist urban planners and policymakers in vulnerable regions to mitigate the effects of extreme heat events in the city. One primary output will be a high-resolution heat vulnerability map of Ouagadougou. This can be a practical decision-support tool for city planners, disaster risk management agencies, and NGOs. It will identify priority neighborhoods for interventions such as urban greening projects, cool roof programs, and the protection of existing water bodies. Our findings will provide evidence-based, locally-specific guidance for developing Heat Action Plans in Ouagadougou and possibly contribute to better early warning systems for heat waves. Understanding whether heat risk is driven more by a lack of greenery, building density, or other factors allows policymakers to allocate limited resources efficiently and effectively to save lives and reduce energy demand. The scenario analysis stretch goal also contributes to proactive, rather than reactive, planning. Urban developers can test "what-if" scenarios to evaluate the cooling potential of new parks or revised zoning laws before they are implemented.