CLIMATE ANALYSIS IN AFRICAN CITIES (CAIAC)

D2.1 Scientific Methodology Strategy

Study accomplished under the authority of the European Space Agency

Reference: 2025/EI/R/3470

December 2025

vito

Vision on technology for a better world

vito.be

D2.1 Scientific Methodology Strategy

VITO

Boeretang 200
2400 MOL
Belgium
VAT No: BE0244.195.916
vito@vito.be – www.vito.be
IBAN BE34 3751 1173 5490 BBRUBEBB

Nele Veldeman
Project Manager
nele.veldeman@vito.be

Koen De Ridder
Science Lead
koen.deridder@vito.be

Vision on technology for a better world

vito.be

AUTHORS

VITO

Koen De Ridder

Tomas Crols

Benjamin Lanssens

Niels Souverijns

Inge Uljee

Nele Veldeman

University of Ottawa

Mohammad Alobaidi

Hossein Bonakdari

Ousmane Seidou

DISTRIBUTION LIST

Clement Albergel, ESA

SUMMARY

This report describes the scientific methodology developed in the Climate Analysis in African Cities (CAIAC) project to assess present-day and future urban heat stress and flooding in African cities using Earth Observation (EO) data, climate projections, and high-resolution urban modelling. The methodology is designed to support consistent, forward-looking analyses across a representative set of African urban environments, while explicitly accounting for both climate change and urban growth.

The methodological framework targets 54 African cities, selected to represent the diversity of climate zones and city sizes across the continent while reflecting stakeholder priorities. City selection is based on the Global Human Settlement Urban Centres Database and the Köppen climate classification, ensuring a balanced and policy-relevant sample.

At the core of the approach is the integration of urban growth modelling, urban climate modelling, and urban flood modelling. Future urban expansion is simulated using the GeoDynamiX land-use change model, which generates annual land-use and population maps from 2020 to 2100 under selected Shared Socioeconomic Pathways (SSPs). These projections provide scenario-consistent representations of future urban form and population distribution.

Urban heat is simulated using the UrbClim model at 300-m resolution. UrbClim is driven by ERA5 reanalysis for present-day conditions and bias-corrected CMIP6 climate projections for future periods. EO-derived datasets are used to characterise land cover, vegetation, impervious surfaces, building height, terrain, and anthropogenic heat flux. The simulations produce multi-decadal time series of near-surface meteorological variables, from which a suite of urban-climate and heat-stress indicators is derived, including the wet-bulb globe temperature (WBGT) as a key impact-relevant metric.

Urban flooding is addressed through the SAHEL framework, an ensemble-based machine-learning approach that combines EO-derived hydrometeorological variables with physically based hydrological modelling. SAHEL employs convolutional neural networks trained using a Rapid Inundation Labelling methodology to generate high-resolution flood extent maps for historical and future scenarios. Urban growth projections are incorporated to capture changes in exposure associated with expanding built-up areas.

Climate projections are analysed for multiple future time horizons and expressed both in terms of SSP scenarios and Global Warming Levels, enabling a consistent comparison of risks across cities and scenarios. The methodology supports a range of scientific analyses, including the quantification of urban amplification of heat and flooding, the separation of climate and urbanisation effects, the evaluation of green-infrastructure mitigation potential, and the translation of hazards into socio-economic risk indicators.

The report further presents feasibility studies for heat and flooding, demonstrates the operational implementation of the methodology, and provides a critical appraisal addressing uncertainty, validation strategies, risks, and the added value of EO data. Together, this methodology establishes a robust and scalable foundation for comparative urban climate-risk assessments in African cities.

TABLE OF CONTENTS

LIST OF ACRONYMS

Acronym	Definition
AHF	Anthropogenic Heat Flux
CAIAC	Climate Analysis in African Cities (this project)
CCI	Climate Change Initiative (ESA)
CMIP6	Coupled Model Intercomparison Project Phase 6
CNN	Convolutional Neural Network
DEM	Digital Elevation Model
ECV	Essential Climate Variable
ERA5	ECMWF Reanalysis version 5
GCM	Global Climate Model
GDDP	Global Daily Downscaled Projections
GHSL	Global Human Settlement Layer
GHS-BUILT-H	Global Human Settlement Layer – Built-up Height
GHS-BUILT-S	Global Human Settlement Layer – Built-up Surface
GHS-POP	Global Human Settlement Layer – Population
GHS-UCDB	Global Human Settlement – Urban Centres Database
GISA	Global Impervious Surface Area
GloFAS	Global Flood Awareness System
GRDC	Global Runoff Data Centre
GWL	Global Warming Level
HPC	High-Performance Computing
IAM	Integrated Assessment Model
IIASA	International Institute for Applied Systems Analysis
IPCC	Intergovernmental Panel on Climate Change
LST	Land Surface Temperature
ML	Machine Learning
NDVI	Normalised Difference Vegetation Index
OSM	OpenStreetMap
QDM	Quantile Delta Mapping
RIL	Rapid Inundation Labelling
SAR	Synthetic Aperture Radar
SSP	Shared Socioeconomic Pathways
SUHI	Surface Urban Heat Island
SST	Sea Surface Temperature
SWAT+	Soil and Water Assessment Tool Plus
UTCI	Universal Thermal Climate Index
UTM	Universal Transverse Mercator
WBGT	Wet-Bulb Globe Temperature

1 INTRODUCTION

Previously, the Science Requirements Document (deliverable D1.2) emphasised the need for high-resolution, forward-looking modelling of heat and flood risks in rapidly growing African cities. Current evidence shows that most existing studies rely on coarse climate data and historical analyses, providing limited insight into future conditions shaped by both urban expansion and climate change.

To address these gaps, the methodology must enable the integration of detailed urban morphology, vegetation, and other terrain features within urban climate and hydrological models, while combining them with projected climate and urban growth scenarios. The approach should further include quantitative impact metrics – such as thermal comfort indicators and sectoral risk indices – to assess health, labour, and infrastructure implications.

A central enabler of this approach is the systematic use of Earth Observation data in Africa's data-scarce cities, providing consistent inputs, including land cover, vegetation indices, land surface temperature, soil moisture, and hydrological datasets; and supporting model validation where ground observations are limited.

This report defines the scientific methodology, ensuring that it is aligned with the above objectives and the availability of high-quality remote sensing and climate model datasets. The methodology will incorporate downscaling techniques to bridge the resolution gap between global climate models and city-scale climate dynamics.

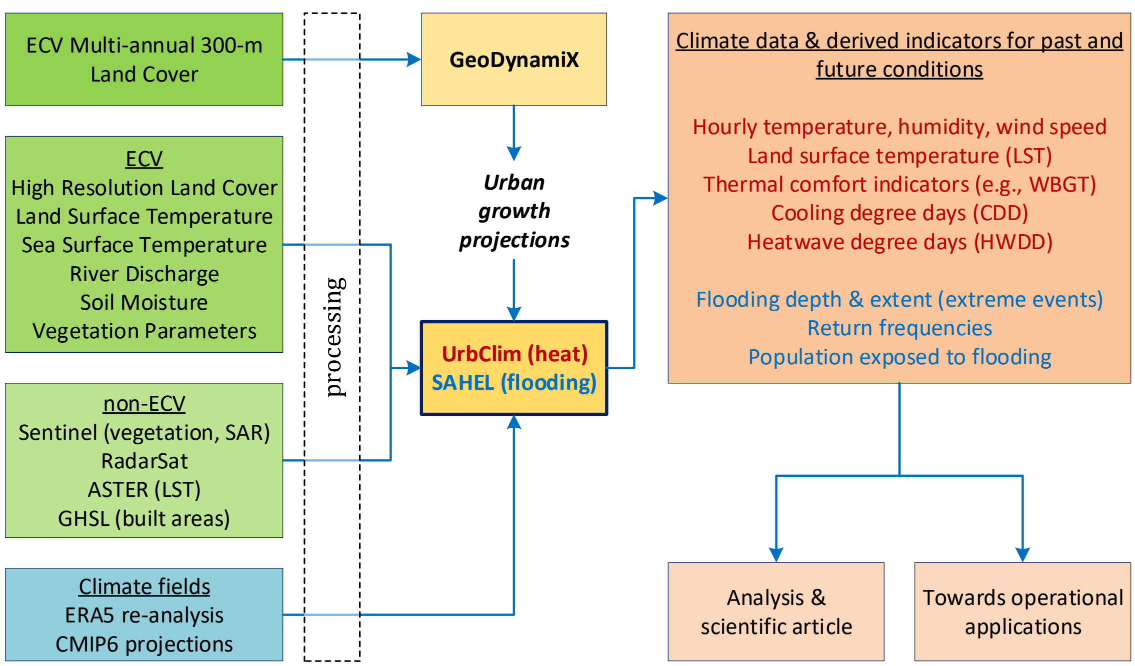


Figure 1. Overview of the methodology's data flow, with the satellite (green) and atmospheric (blue) input data, the models (yellow boxes), and main outcomes in terms of data as well as derived results, analysis and prospects to operational applications (orange).

Figure 1 provides a schematic overview of the modelling workflow that underpins the CAIAC study. It shows how various Earth Observation (EO) datasets and climate fields are combined to generate high-resolution information on urban heat and flooding in African cities. The workflow begins with several EO-based input layers, including the multi-annual 300-m Land Cover product from the ESA Climate

Change Initiative (CCI) and other Essential Climate Variables (ECVs) such as land surface temperature, sea surface temperature, river discharge, soil moisture, and vegetation parameters. Additional non-ECV sources, such as Sentinel (SAR and vegetation), RadarSat, ASTER, and GHSL, provide complementary information on surface and built-up characteristics, while ERA5 reanalysis and CMIP6 projections deliver the large-scale climate boundary conditions.

These inputs feed into two main modelling components: GeoDynamiX, which produces urban growth projections, and the high-resolution models UrbClim (for urban heat) and SAHEL (for flooding). The combined framework is designed to generate a suite of climate and impact indicators for both past and future conditions, including temperature, humidity, wind speed, thermal comfort indices such as WBGT, cooling and heatwave degree days, and flood depth, extent, frequency, and population exposure. The resulting outputs then serve two complementary objectives: advancing scientific analysis through dedicated publications and guiding the development of operational applications to support climate adaptation in African cities.

The remainder of this report explains how the 54 target cities were selected, and briefly describes the modelling strategy, the considered climate scenarios, and the way the scientific analysis will be conducted. It closes with a feasibility (demonstration) study and a critical appraisal, including uncertainty considerations.

2 SELECTION OF TARGET CITIES

The CAIAC project aims to conduct EO–based simulations of urban heat and flooding for 54 cities, i.e., one city in each African country. The simulation results will be used in a series of analytical exercises to generate new insights into the impacts of global climate change on African urban climate futures. Below, a brief description is given of how the 54 cities were selected.

From the outset, the objective was to ensure that the selected cities constitute a representative sample of African urban environments, both in terms of climate zones and city size (expressed by the number of inhabitants). At the same time, the selection needed to reflect stakeholder preferences, as expressed through responses to an online survey conducted in August–September 2025 (see Deliverable D1.2 for details).

These two requirements were reconciled through an iterative selection procedure, supported by a dedicated Python script and based on the following datasets:

- the list of cities prioritized by the stakeholders (the latter were asked to nominate three cities in their own country, ranked by priority) – approximately half of the African countries responded to the survey (see Deliverable D1.2);
- the GHS Urban Centre Database (UCDB) R2024A[1], provided through the Copernicus Global Human Settlement Layer (GHSL), which includes population estimates for each city (Melchiorri, 2022);
- the Köppen climate classification map (Beck et al., 2023)[2].

The selection process started from the list of cities proposed by stakeholders. The availability of three cities per country provided some flexibility to adjust the selection in order to improve representativity across climate zones and city-size categories. One of the three proposed cities was always selected, following the stakeholder ranking wherever possible. However, in a limited number of cases, the ranking was overridden when the selection of a different city substantially improved the overall balance of the sample.

The only exception to this rule concerns Ethiopia. None of the cities proposed by stakeholders across all surveyed countries fell within the Csb (warm-summer Mediterranean) climate zone, which would otherwise have been entirely absent from the selected sample. Ethiopia offered a practical opportunity to address this gap, as suitable Csb cities are present within the country. To ensure representation of this climate class and to improve the overall match with the climate-zone distribution observed in the GHS-UCDB dataset, the city of Naqamte (Nekemte) was therefore selected.

For countries without stated city preferences, cities were selected through a trial-and-error process, giving preference to candidates that brought the overall distribution across climate zones and city-size categories closer to that observed for the full African city sample in the GHS-UCDB database.

The resulting cities and their distribution across climate zones and city size categories are shown in Figure 2and Figure 3 as well as in Annex 1.

It should be noted that the resulting set of selected cities does not constitute an objectively optimal selection. Rather, the adopted procedure aimed to produce a sample that reasonably mirrors the observed African-wide distributions of climate zones and city sizes, while remaining consistent with stakeholder input wherever possible.

---

[1] https://human-settlement.emergency.copernicus.eu/ghs_ucdb_2024.php

[2] https://www.gloh2o.org/koppen/

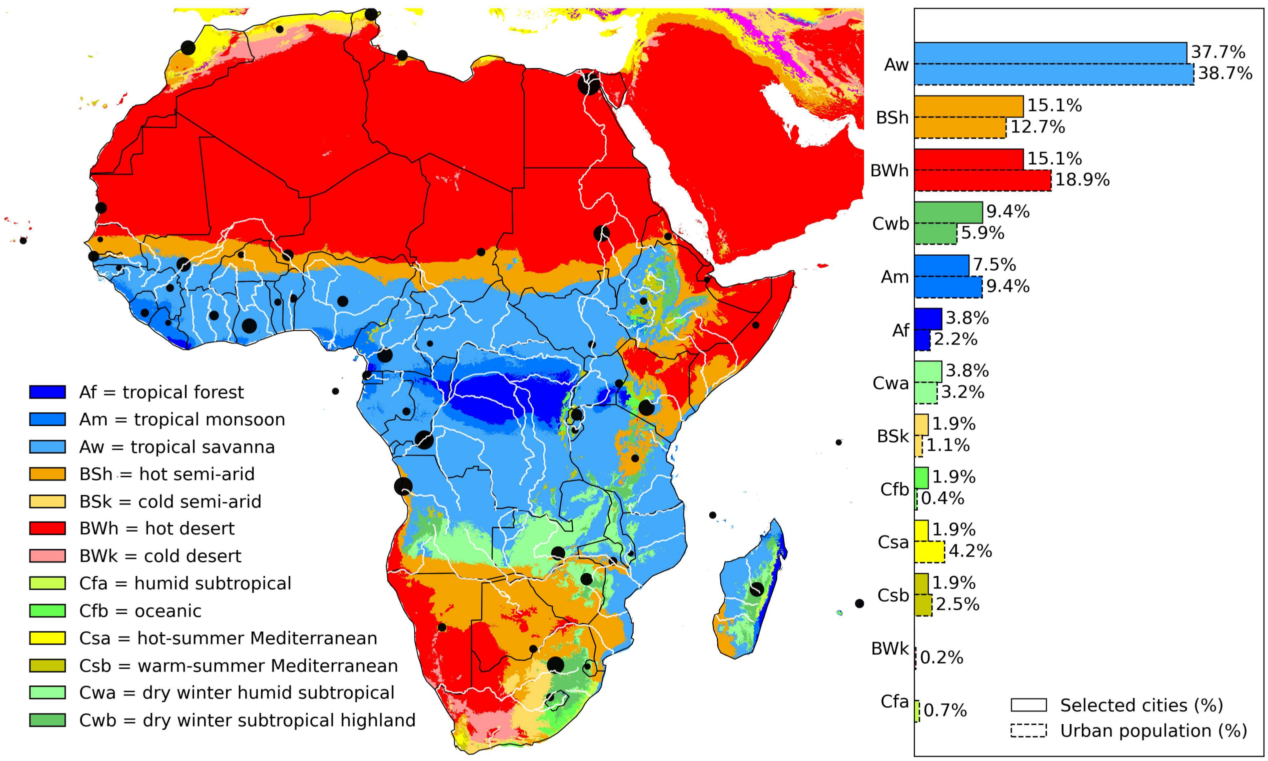


Figure 2. Selected cities, shown as solid circles, their sizes scaling with the size of the city (in terms of number of inhabitants, see Annex 1). The colour scale represents the Köppen climate zones. The bar chart on the right shows the distribution of the selected cities versus the distribution of the actual urban population across African climate zones.

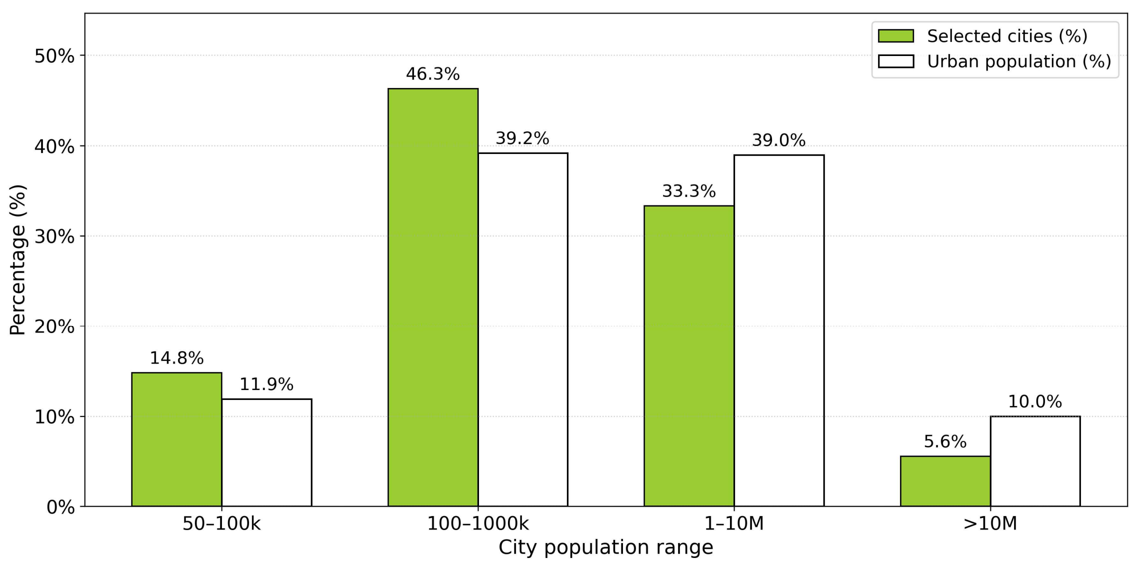


Figure 3. Distribution of the selected cities (green bars) across city size ranges, accounting for small (< 100k inhabitants), medium (100-1000k), large (1-10M) and megacities (> 10M). The actual distribution of the African urban populations is shown with the white bars.

3 MODELLING SETUP

3.1 GeoDynamiX (urban growth)

3.1.1 Model configuration

A key aspect of our approach is the integration of urban growth modelling. Given the rapid expansion of African urban areas, any plausible urban climate projection must account for urban sprawl scenarios. In CAIAC, future urban growth scenarios are simulated with the GeoDynamiX model (White et al., 2012; Crols, 2017). The model combines land-use interactions, physical suitability, zoning and accessibility to compute transition potentials that determine future urban growth (Figure 4).

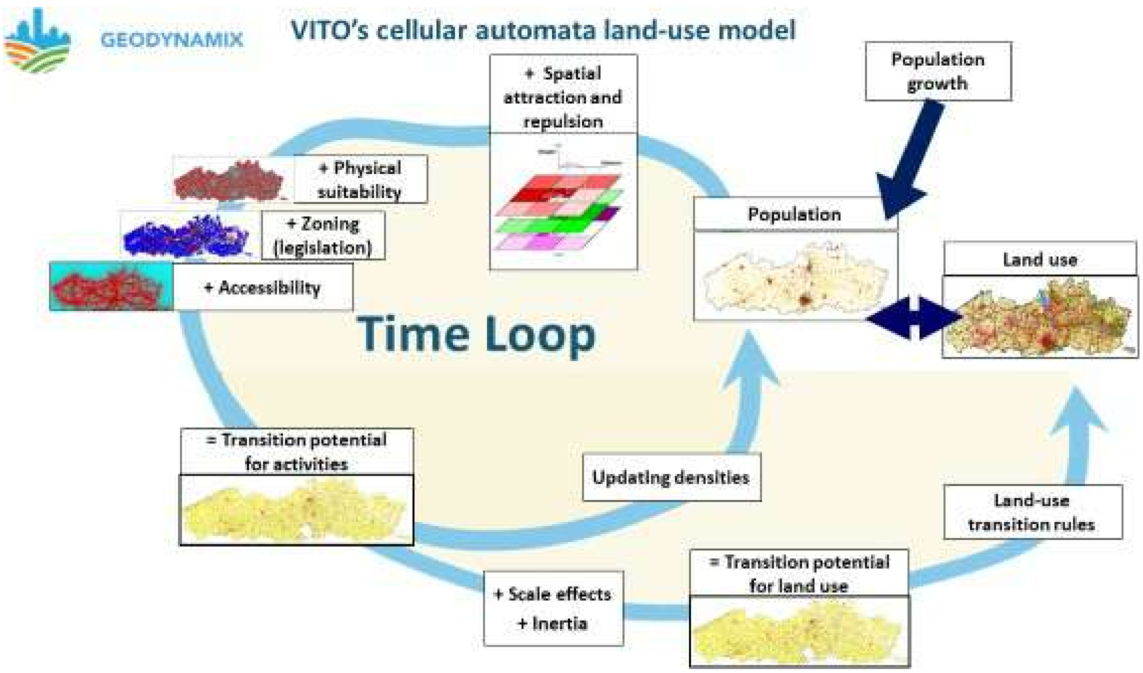


Figure 4. The components of the GeoDynamiX land-use change model.

3.1.2 Integration of EO and other land-use change data

All simulations start from land-use maps that were derived for the years 2005, 2010, 2015 and 2020 by combining ESA CCI Land Cover, Copernicus GHS-BUILT-S, ESA WorldCover and OpenStreetMap (OSM) data. Population maps were copied from Copernicus GHS-POP data. Lower physical suitability is a combination of reclassified steep slopes and high flood risk. Slopes were derived from the Copernicus Global DEM[3], while flood risk was taken from CCI wetlands land cover, OSM wetlands and the Sentinel Global Flood Awareness System (GloFAS). Scenarios with and without flood risk can be taken into account in order to model the maximum and minimum impact of the population exposed to flood risk. Later in the project, we can evaluate if we can replace the global flood risk maps by the ones generated in the project for the selected cities. Zoning maps combine military areas and protected nature from OSM data. Accessibility to the transport network is calculated for each city with network data (roads and railways) taken from OSM.

[3] Due to missing Copernicus DEM data, USGS SRTM 1 Arc-Second data were used for the Seychelles.

3.1.3 Simulation strategy

GeoDynamiX is run in every CAIAC city from 2020 to 2100 in annual time steps for 2 SSP scenarios (SSP2 and SSP3, see section 4 for more details).

The parameters of the spatial rules of urban growth are calibrated for each city by running the model over the 2005-2020 period and validating the results in 2020 compared to the actual present land-use and population maps. For this calibration we use a Covariance Matrix Adaptation Evolution Strategy (CMA-ES) optimisation algorithm.

Future population growth per country per scenario is copied from the 2023 Human Capital Data Explorer (K.C. et al., 2024). Urbanisation rates per country and per scenario are copied from the 2018 IIASA IAM database (Riahi et al., 2017; Jiang and O'Neill, 2017). For population, these scenarios are translated into city scenarios by assuming that the population growth rates and urbanisation rates of the country apply to the cities, but multiplied by a correction factor which is the ratio of (1) the observed city population growth rate during the 2005-2020 calibration period to (2) the population growth rate of the country. For urban growth quantities, we first compute if there was a densification (lower additional urban area per new inhabitant) or an urban sprawl effect (higher additional urban area) during the calibration period. In SSP2, we apply half of this densification or urban sprawl ratio towards the future to create an average scenario. In SSP3 – which is a scenario with a high demographic footprint in the Global South – we define a worst-case scenario: for currently densifying cities the urban area per person is kept constant, while for sprawling cities the past urban sprawl ratio is continued towards the future.

3.1.4 Model output

Urban growth modelling will leverage the following products for each considered SSP scenario:

- Annual land-use maps (2020-2100)
- Annual population maps (2020-2100)

3.2 UrbClim (heat)

3.2.1 Model configuration

The UrbClim model (De Ridder et al., 2015) is applied in CAIAC to simulate urban climate conditions at 300-m spatial resolution, providing multi-decadal datasets of near-surface temperature, humidity, wind and radiation for African cities. The model solves simplified boundary-layer equations coupled to a surface energy- and water-balance scheme.

Surface characteristics are represented through four main categories (built-up, vegetation, bare soil and water), each following its own parameterisation in UrbClim, and each being subdivided into several sub-classes. The physical parameters (albedo, emissivity, roughness length, root depth, vegetation resistance, etc.) are prescribed per land-cover class following UrbClim's standard lookup tables, as detailed in D1.3 *Data Inventory Document*. The merged land-use map driving this parameterisation combines ESA CCI Land Cover, Copernicus WorldCover, GHSL Built-up and OSM data. For urban cells, aerodynamic parameters are adjusted according to built-up density and representative building height.

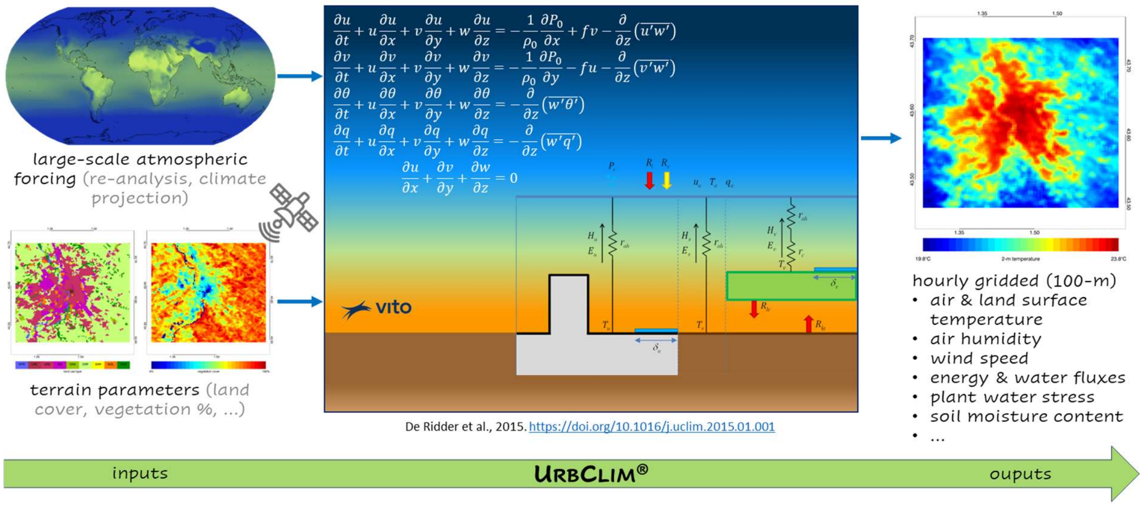


Figure 5. Schematic overview of the data flow in an UrbClim simulation.

3.2.2 Integration of EO and climate data

The processing of input data is based on the UrbClim pre-processing chain described previously in D1.3:

- Earth-Observation data are harmonised to a common 300-m grid and used to specify terrain, surface and vegetation parameters.
- Forcing data come from ERA5 reanalysis (hourly, 0.25° × 0.25°) for the period 2001–2020, providing temperature, humidity, wind components, pressure, radiation, soil-moisture and sea-surface temperature fields.
- Future climate projections are generated using bias-corrected CMIP6 GCM outputs. Bias correction applies quantile-delta mapping (QDM) to ERA5-based reference time series, ensuring that projected changes retain the statistical structure of each GCM while correcting historical biases.

The resulting ERA5 and CMIP6 fields are interpolated to the UrbClim domain in space and time and used as lateral and top boundary forcing. Soil and vegetation states are initialised from EO-derived maps and ERA5 soil-layer fields. All inputs and outputs are managed in NetCDF format and documented through metadata standards ensuring traceability within the CAIAC data infrastructure.

3.2.3 Simulation strategy

For each selected city, UrbClim is run for a multi-year baseline (2001–2020) and three future horizons (2021–2040, 2041–2060, 2081–2100), mainly considering IPCC scenarios SSP2-4.5 and SSP3-7.0, since these also give access – in combination with suitably selected time horizons – to the Global Warming Levels 1.5-2.0-3.0°C. More details on the climate scenarios are relegated to Section 4.

A common 300-m land-use representation is used for the baseline simulations; scenario-dependent land-use updates, derived from GeoDynamiX urban-growth projections (Section 3.1), will be used for the urban climate projections.

The simulations deliver continuous hourly meteorological fields, which are post-processed into climatological indicators (means, percentiles, extremes, aggregated quantities, etc.). This approach allows robust statistics while keeping storage requirements manageable.

3.2.4 Outputs and indicators

From the UrbClim raw hourly outputs, a number of key indicators is derived, including mean, maximum and minimum air temperature (2 m); day- and night-time Urban Heat Island (UHI) intensity; frequency and duration of heat-wave events; cooling- and heating-degree-day indices. A complete overview of the indicators, including units and mathematical expressions, is provided in Annex 2. If required, these indicators can be spatially aggregated to urban sub-zones (e.g., formal vs. informal, or over administrative units) for integration into socio-economic and health impact analyses defined elsewhere in the project (see D1.2 Science Requirements Document, Table 1).

3.3 SAHEL (flooding)

3.3.1 Model configuration

The Satellite-based African Hydrological Ensemble Learner (SAHEL) developed by uOttawa is an ensemble-based machine learning model designed to generate flood maps for African cities using past and future hydroclimatic time series (primarily precipitation and streamflow). It is being adapted to integrate multiscale hydrometeorological data from Essential Climate Variables (ECVs), satellite products, and ground-based observations. SAHEL employs statistical and deep learning models, prioritizing convolutional neural networks (CNNs) due to their effectiveness in processing spatial data. Once trained, SAHEL projects future flood risks by integrating climate model outputs and urban expansion projections.

Within the SAHEL configuration, CNN sub-models are implemented as encoder–decoder mapping networks that learn a transformation from multichannel gridded inputs to pixel-wise flood-map outputs, using established architectures such as U-Net and ResNet-based variants (Figure 6). Inputs are arranged as stacked spatial-temporal layers (e.g., precipitation, soil moisture, land cover, DEM-derived indices, and SWAT-derived hydrologic variables). The encoder extracts hierarchical spatial features that capture the terrain controls and the urban forms, while the decoder reconstructs the inundation likelihood (flood/no flood) field. Multiscale feature fusion (e.g., skip connections) preserves fine urban-scale detail while retaining basin-scale hydrologic context, and multiple CNN variants trained with different initializations and feature subsets form an ensemble whose aggregated outputs provide flood predictions and associated uncertainty.

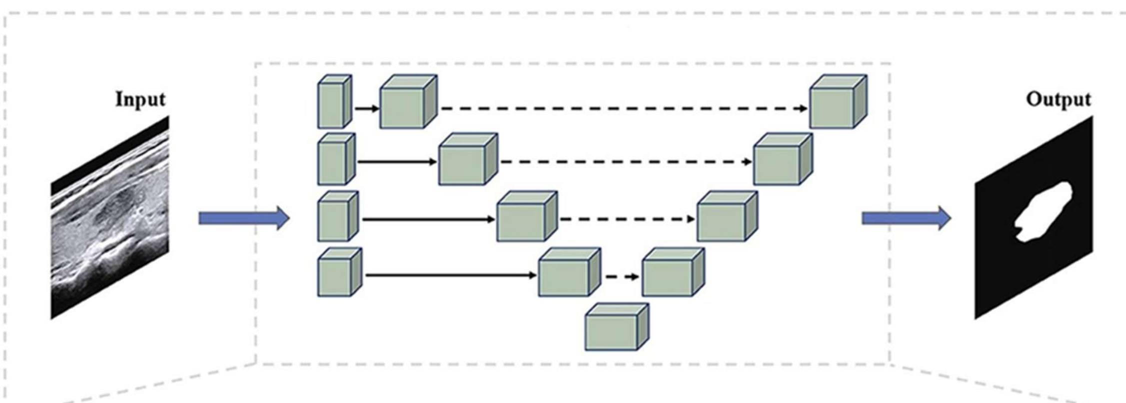


Figure 6. General Multi-Scale U-Net CNN Architecture.

A physical rainfall–runoff model is used to support the SAHEL architecture by simulating the basin-scale hydrologic response. The model solves water-balance and flow-routing processes to generate continuous time series of surface runoff and river discharge from hydroclimatic forcing. These simulated variables represent catchment-scale hydrologic response processes that are not directly observable from Earth Observation products and contribute to maintaining physical consistency within the learning framework.

The hydrologic model outputs are used both to inform inundation dynamics and to support the construction of training datasets for the machine-learning components. In particular, simulated runoff and discharge are used to support a Rapid Inundation Labelling (RIL) approach, which combines dynamic hydrometeorological conditions with flood-risk characteristics to generate instantaneous inundation maps.

3.3.2 Model Integration of EO and climate data

Earth Observation (EO) variables and climate forcing data feed directly into the physical rainfall runoff model. Historical hydroclimatic records establish a precise hydrological baseline that captures the unique response characteristics of the city watershed system. Future climate scenarios are then imposed upon this established baseline. This core methodology guarantees that the physical model strictly maintains the inherent hydrological proportions and relationships of the system, while simultaneously permitting a rigorous and internally consistent examination of alterations driven by climate.

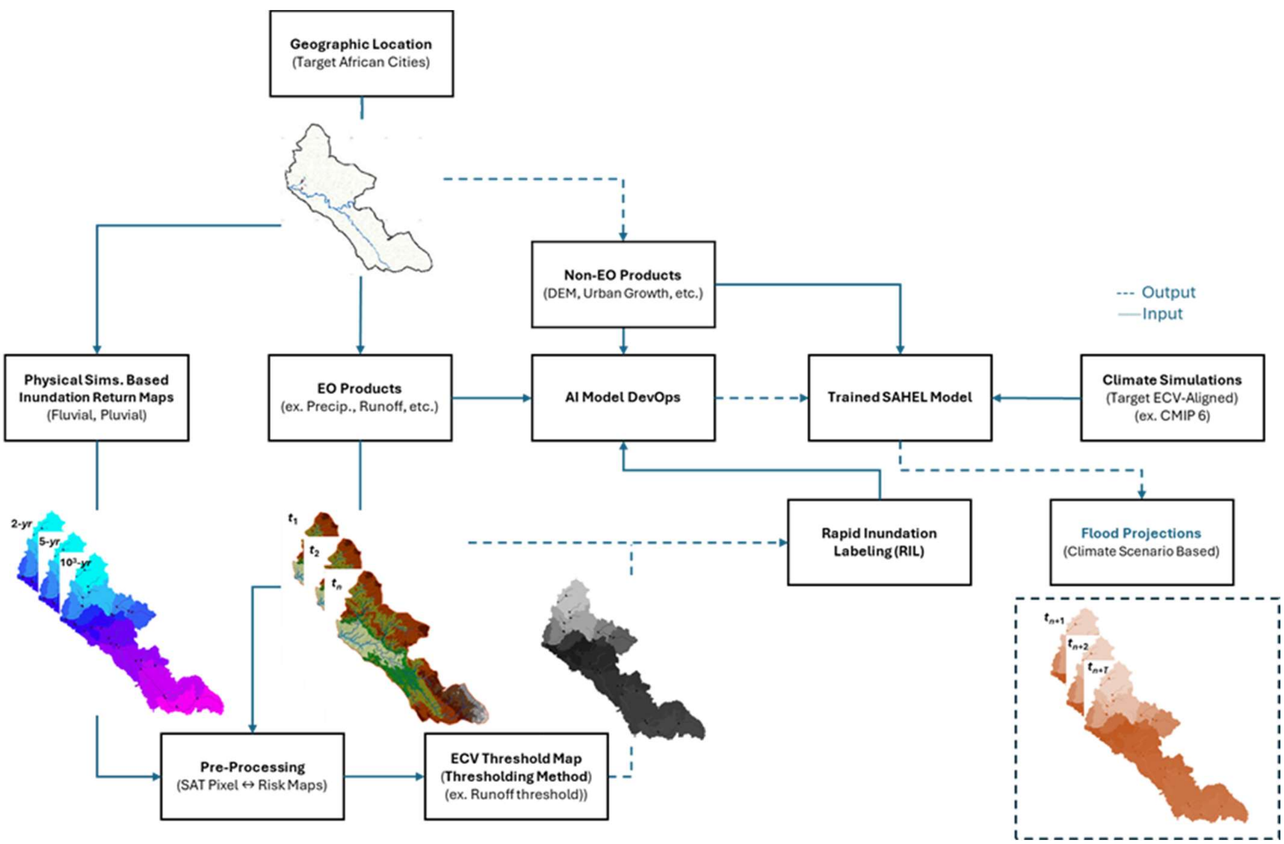


Figure 3. The components of the SAHEL development framework, with emphasis on the RIL approach.

Within this framework, Essential Climate Variables (ECVs) provide the indispensable dynamic backbone. They deliver continuously evolving hydrometeorological information required to characterize the state of the system. The consistent availability of ECV based inputs therefore becomes critical, as it represents the enabling precondition for the Rapid Inundation Labelling (RIL) methodology. Generating accurate and instantaneous representations of inundation fundamentally depends on coherent and synchronous descriptions of precipitation intensity, land surface condition, and atmospheric drivers. In the absence of this stream of EO derived information, the construction of physically meaningful and temporally responsive inundation labels would not be feasible, as the underlying data foundation would be incomplete.

Thus, the synthesis of EO derived variables, climate forcing data, and physics based hydrological simulation creates a unified and coherent representation of the evolving hydroclimatic state of the city watershed system. This integrated representation is not a passive output; it constitutes the essential foundation for all subsequent analytical steps. It provides a physically consistent and dynamically responsive basis that enables and informs downstream learning processes and scenario based exploration. All subsequent modelling phases derive their integrity from this foundational representation, ensuring that projections and analyses remain anchored in a realistic and climate-responsive physical context. This approach explicitly accounts for projected urban growth and enables the joint assessment of hydroclimatic hazards and evolving urban exposure under plausible future scenarios.

3.3.3 Simulation strategy

SAHEL, as depicted in Figure 3, simulates future inundation patterns using a scenario-based workflow that couples climate forcing, urban growth projections, and physically consistent hydrologic response. A suite of climate model projections (GDDP data) is first selected and harmonized with the EO/ECV input space after bias adjustment. In parallel, Land Use/Land Cover layers are obtained for each city using the GeoDynamix urban growth model to represent alternative urbanization pathways. For each city and scenario, a physical rainfall–runoff model (SWAT+) is forced with the corresponding hydroclimatic inputs to produce continuous runoff and river discharge time series, providing catchment-scale response variables that are not directly observable from EO data. These variables, together with a DEM, are then ingested by the SAHEL CNN ensemble to generate pixel-wise inundation maps. Finally, results will be presented as a binary future flood extent maps.

4 CLIMATE SCENARIOS

Climate projections will be based on the CMIP6 Shared Socioeconomic Pathways (SSPs). Due to the unavailability of CORDEX CMIP6 downscaled datasets for Africa during the project timeframe[4], we will not rely on CORDEX Africa simulations. Instead, we will directly employ CMIP6 Global Climate Model (GCM) outputs to drive the urban climate and flood models. For the flood modelling, use is also made of the NASA NEX-GDDP-CMIP6[5] dataset, offering calibrated daily precipitation and temperature at a 0.25° resolution.

Apart from running the models for the present-day situation (2001-2020) based on re-analysis fields, we will consider the following future time horizons for the climate projections: 2021-2040 (near term); 2041-2060 (mid-term); 2081-2100 (long term). Next to this, the project targets Global Warming Levels (GWLs) of 1.5-2.0-3.0°C as a basis for analysis.

Regarding the IPCC scenarios, we propose to focus primarily on SSP2-4.5, which best represents the most likely future trajectory. According to UNEP (2023), "Fully implementing and continuing mitigation efforts of unconditional NDCs made under the Paris Agreement for 2030 would put the world on course for limiting temperature rise to 2.9°C this century," consistent with SSP2-4.5. Similarly, IPCC (2023, §A.4.3) estimates a 2.8°C rise by 2100 under current policies – again close to SSP2-4.5 – while noting that incomplete implementation could push warming toward 3.2°C, i.e., somewhere between SSP2-4.5 and SSP3-7.0. As shown in Table 1, SSP2-4.5 also aligns well with the 1.5°C and 2°C warming levels, whereas the 3°C level cannot be reached within this scenario. To represent this higher 3°C warming, SSP3-7.0 will be used together with the 2066–2085 period (see Table 1).

Table 1. Correspondence between global warming levels versus SSP scenario and time horizons. The upper three rows show projected global mean temperature changes (°C) relative to 1850–1900 for different time horizons, while the lower three rows indicate the first 20-year period during which each specified global warming level is reached. Values in brackets denote the very likely range (5–95 %), and "n.c." indicates that the corresponding warming level is not reached between 2021 and 2100. Cells highlighted in yellow mark the SSP/time combinations considered in CAIAC. Cells with red text identify the SSP/time combinations used to define the 1.5 °C, 2.0 °C, and 3.0 °C Global Warming Levels. Adapted from Table 1 in IPCC (2021).

	SSP1-1.9	SSP1-2.6	SSP2-4.5	SSP3-7.0	SSP5-8.5
Near term, 2021–2040	1.5 [1.2 to 1.7]	1.5 [1.2 to 1.8]	**1.5 [1.2 to 1.8]**	1.5 [1.2 to 1.8]	1.6 [1.3 to 1.9]
Mid-term, 2041–2060	1.6 [1.2 to 2.0]	1.7 [1.3 to 2.2]	**2.0 [1.6 to 2.5]**	2.1 [1.7 to 2.6]	2.4 [1.9 to 3.0]
Long term, 2081–2100	1.4 [1.0 to 1.8]	1.8 [1.3 to 2.4]	**2.7 [2.1 to 3.5]**	3.6 [2.8 to 4.6]	4.4 [3.3 to 5.7]
1.5°C	2025–2044 [2013–2032 to n.c.]	2023–2042 [2012–2031 to n.c.]	2021–2040 [2012–2031 to 2037–2056]	2021–2040 [2013–2032 to 2033–2052]	2018–2037 [2011–2030 to 2029–2048]
2°C	n.c. [n.c. to n.c.]	n.c. [2031–2050 to n.c.]	2043–2062 [2028–2047 to 2075–2094]	2037–2056 [2026–2045 to 2053–2072]	2032–2051 [2023–2042 to 2044–2063]
3°C	n.c. [n.c. to n.c.]	n.c. [n.c. to n.c.]	n.c. [2061–2080 to n.c.]	**2066–2085 [2050–2069 to n.c.]**	2055–2074 [2042–2061 to 2074–2093]

[4] WCRP/CORDEX expects at best initial, incomplete results by the end of 2025

[5] https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-gddp-cmip6

5 SCIENTIFIC ANALYSIS

In the CAIAC project, we will take advantage of the large and “representative-by-design” sample of African cities (Section 2) to conduct analysis exercises and derive statistical information from the sample. The core scientific questions identified in Deliverable D1.2 will be addressed as briefly described below.

Quantifying amplification of heat and flooding by the urban fabric. To assess how the urban fabric amplifies extreme heat and flooding, CAIAC will generate 300-m urban climate and flood maps for the selected representative set of African cities spanning different climate zones and city sizes. Urban characteristics such as built-up density, surface materials, and vegetation cover will be explicitly represented in the modelling framework using EO-derived land cover, vegetation indices, and settlement datasets. By applying consistent modelling approaches across cities, CAIAC will systematically analyse how hazard amplification varies with city size and climatic context, including differences between humid, arid, and temperate African regions. Validation will be supported through EO products such as LST and, where available, in situ observations.

Disentangling the roles of urban growth and climate change. To quantify the relative contributions of urban growth and climate change to future heat and flood risk, CAIAC will employ scenario-based modelling. Urban climate and flood simulations will be driven by combinations of (i) present-day urban form with future climate projections, and (ii) future urban growth scenarios combined with both present and future climate conditions. Urban growth projections will be derived from EO time series and settlement datasets, allowing spatially explicit representation of expanding built-up areas. This factorial design enables the separation of risk changes attributable to climate forcing from those driven by urbanisation, thereby identifying cities and contexts where adaptation priorities should focus on land-use planning versus climate resilience measures.

Evaluating the mitigation potential of urban green infrastructure. CAIAC will explicitly test the effectiveness of urban green infrastructure, such as trees and parks, and green corridors, in mitigating heat and flood risks. Rather than relying solely on air temperature, the project will apply impact-relevant thermal comfort indicators (e.g., WBGT and related indices) to quantify reductions in heat stress under different greening scenarios. For flooding, the influence of increased infiltration and delayed runoff will be assessed through modified surface and drainage representations. EO-derived vegetation metrics (e.g., NDVI) will support the spatial characterisation of green infrastructure.

Translating hazards into socio-economic risk. To move from physical hazards to impacts, CAIAC will integrate vulnerability curves and demographic data with urban climate and flood model outputs. Hazard intensity metrics (e.g., thermal stress levels or flood depths) will be linked to sector-specific impact functions addressing health outcomes, labour productivity losses, and infrastructure damage. Spatially explicit population data, settlement typologies, and indicators of informality will be used to differentiate risks across population groups and neighbourhood types. This integrated framework enables estimates of economic and social risk, supports comparison between cities, and provides an evidence base for prioritising adaptation options and investments.

6 FEASIBILITY STUDY

In this section we present feasibility exercises for heat and flooding, focusing on present-day conditions. These serve as a blueprint and demonstration of subsequent work on all 54 target cities later in the project. Both the heat and flooding demonstration studies are elaborated for the city of Niamey (Niger), considering the prior expertise of the contractors.

6.1 Urban heat

6.1.1 Domain selection

The UrbClim simulations for Niamey start with the definition of the model grid. We specify a city-centre coordinate, a pixel resolution, a projected coordinate system in metres, and the number of grid cells in each direction. The city-centre point is chosen as close as possible to the centre of an ESA CCI Land Cover pixel, so that reprojection artefacts are minimal in the urban core. All CAIAC cities are simulated at 300-m resolution (600-m if the city is very large). We use the local UTM projection, which provides metric coordinates and preserves distances and areas sufficiently well for urban-scale modelling.

The domain size is selected so that (i) the present-day city is fully contained, (ii) there is a sufficient rural buffer around the city to resolve the urban heat island signal, and (iii) the projected urban expansion remains within the domain. UrbClim can deal with domains covering up to 301 × 301 grid cells, which sets an upper bound on the domain size.

For the demonstration case of Niamey, the resulting grid parameters are:

- City-centre coordinate: 13.5264° N, 2.1181° E
- Grid resolution: 300 m
- Grid size: 301 × 301 cells
- Projection: UTM zone 31N (EPSG:32631)
- Domain extent: approximately 90 km × 90 km

6.1.2 Meteorological forcing data

Meteorological forcing for Niamey is derived from ERA5 reanalysis fields. In a first step, we extract a compact ERA5 subset defined on a 3° × 3° latitude-longitude window centred on the city. The extracted variables are as defined in Table 1 of the D1.3 data inventory document.

In a second step, we convert these ERA5 subsets into UrbClim-ready input. All 2D and 3D atmospheric fields are horizontally interpolated to the Niamey domain centre. Full-level pressure and geopotential are computed from the ECMWF hybrid coefficients, and potential temperature is derived from temperature and pressure. Vertical profiles are linearly interpolated to the UrbClim output levels, and the time axis is resampled to the model step.

Soil and surface initial states are constructed on the Niamey grid from ERA5 soil layers and a multi-depth soil-texture map (using Clapp–Hornberger relations with light smoothing). For coastal cities, sea-surface temperature (SST) is obtained from an area-mean over a 0.5° × 0.5° box around the city. These steps result in one UrbClim forcing file per year, containing centre-point meteorological profiles, surface fluxes / precipitation time series, and soil / vegetation initialization.

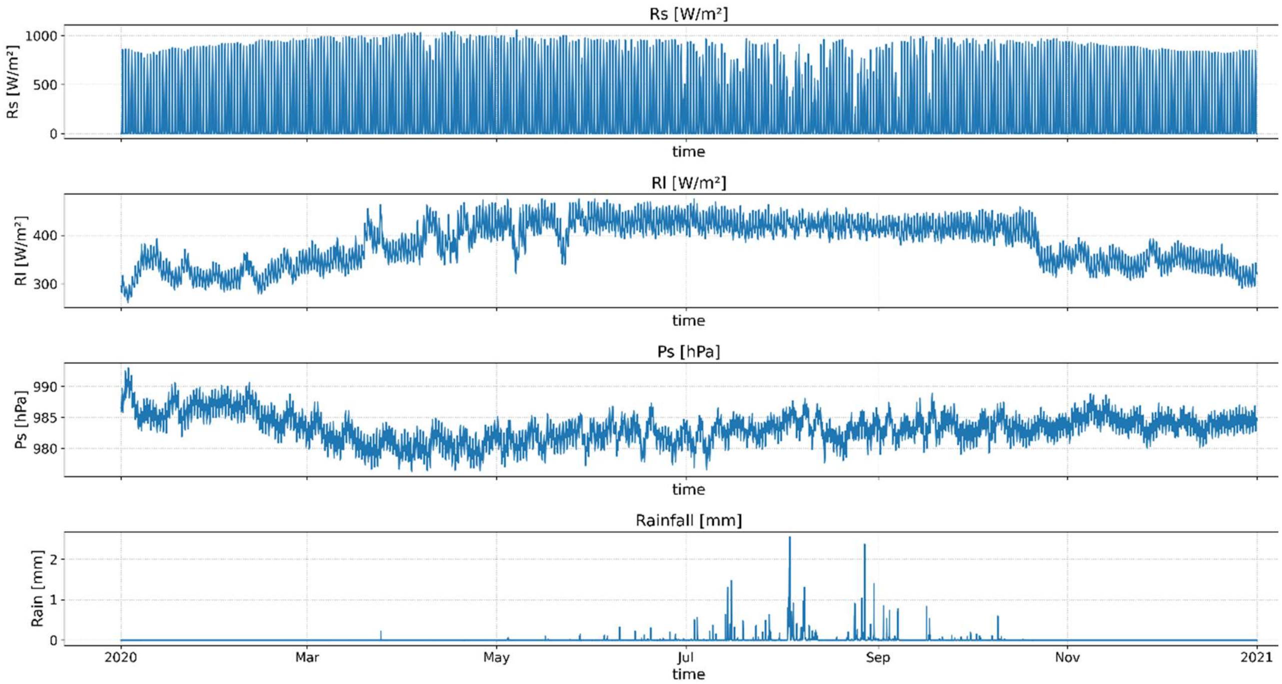


Figure 7. Example of a time series of ERA5-derived surface forcing variables for Niamey in 2020: downwelling surface shortwave radiation (Rs), downwelling surface longwave radiation (Rl), surface atmospheric pressure (Ps), and rainfall (Rain).

6.1.3 Terrain input data

6.1.3.1 Land cover

For the land-use field, we combine ESA 10-m WorldCover with the 2020 land-use map produced by the GeoDynamiX workflow (Section 3.1). Starting from the UrbClim grid defined for Niamey, we first cut out the corresponding WorldCover tile, reproject it to the city UTM projection and aggregate this to the model resolution (300 m) using a majority filter. The WorldCover classes are then reclassified into the internal UrbClim land-use categories.

Because UrbClim has no internal “airport” land cover class, we correct the “airport” class in the GeoDynamiX land-use map. Airport pixels (class 12) are replaced by the dominant reclassified WorldCover class in the same grid cell, so that large airport areas are treated consistently with the surrounding land cover rather than as a separate artificial class.

In a final step, we refine the treatment of water. Using a high-resolution global coastline dataset, we derive a land–sea mask at the UrbClim resolution and separate inland water bodies from the open ocean. Pixels classified as “water” that fall on the ocean are labelled as a dedicated “sea” class, while lakes and rivers remain in the “inland water” class. This prevents coastal ocean from being mixed with inland water in the land-use input. The “wetland” class is also not present in UrbClim, hence we remap it to the “grassland” or “woodland class depending on the dominant WorldCover class (“Herbaceous grassland” or “Mangrove”). The resulting GeoTIFF file provides the final UrbClim land-use map.

Table 2. Correspondence between GDX land cover and UrbClim land cover

GDX class	UrbClim class	Refinement step
1 Bare soil	6 Bare soil	-
2 Grassland	7 Grassland	-
3 Cropland	8 Cropland	-
4 Shrubland	9 Shrubland	-
5 Woodland	10 Woodland	-
6 Broadleaf	11 Broadleaf trees	-
7 Needleleaf	12 Needleleaf trees	-
8 Wetland	7 Grassland 10 Woodland	• Herbaceous wetland (class 90) dominant → *7 Grassland* • Mangrove (class 95) dominant → *10 Woodland*
9 Urban	1 Urban	-
10 Sub-urban	2 Sub-urban	-
11 Water	14 Inland water 15 Sea water	refine with coastline mask: • over land → *14 Inland water bodies* • over sea → *15 Sea water*
12 Airport	1 Urban 6 Bare soil	refine with WorldCover: airport grid cells are replaced by the dominant WorldCover-based class (mainly urban & bare soil)

6.1.3.2 Vegetation cover fraction

Vegetation cover for UrbClim is derived from a multi-year NDVI climatology based on Sentinel-2 Level-3 cloud-filtered mosaics from the Copernicus Data Space Ecosystem (product type S2MSI_L3_MCQ) for the period 2021–2024. For each year and calendar quarter (Q1–Q4), we collect all Sentinel-2 tiles intersecting a 1° × 1° box around the city, mosaic them, and compute NDVI at 10-m resolution from the Sentinel-2 B08 and B04 reflectance bands. This yields four quarterly NDVI mosaics per year. We then construct a quarterly NDVI climatology by averaging each quarter over the years 2021–2024.

From this quarterly climatology we derive a monthly NDVI climatology by weighted linear interpolation between adjacent quarters. For each calendar month, NDVI is obtained as a time-weighted combination of the two neighbouring quarterly means, with larger weight given to the quarter whose centre date is closest to the middle of the month. This produces 12 monthly NDVI fields at 10-m resolution in the Sentinel-2 native projection.

These monthly NDVI mosaics are then projected to the local UTM projection of the UrbClim domain and resampled to the UrbClim grid (city-specific EPSG, model resolution 300 m or 600 m). At this stage, NDVI values are clipped to the physical range [−1, 1], missing values (NaNs) are set to zero, and a land mask is applied so that only land pixels are retained. Isolated saturated pixels with NDVI = 1 are replaced by a local neighbourhood statistic computed in a 5 × 5 window, to limit artefacts.

Fractional vegetation cover is derived from NDVI following the methodology of Gutman and Ignatov (1998). We apply a linear re-scaling between a land-only minimum NDVI (NDVI_min) and an upper NDVI value representative of dense vegetation (NDVI_max = 0.85)

$$f_{veg} = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}},$$

with f_{veg} constrained to the interval [0,1]. Negative NDVI values are set to zero before the transformation. The final product consists of a set of 12 monthly vegetation cover fraction maps on the UrbClim grid.

6.1.3.3 Terrain elevation (DEM)

We use the Copernicus DEM GLO-30 global digital elevation model (30-m resolution) to specify terrain height. We clip the DEM to the UrbClim domain over Niamey, reproject that to the local UTM projection (EPSG:32631), and resample it to the UrbClim grid using bilinear interpolation. We then set any negative elevation values (rare artefacts below sea level on land) to zero. The resulting field provides the terrain height (DEM) for each UrbClim grid cell.

6.1.3.4 Soil sealing (impervious surface fraction)

We derive soil sealing from the GISA-10m global impervious surface area dataset (Huang et al., 2022). We clip the GISA raster to the Niamey domain, reproject it to EPSG:32631, and resample that to the UrbClim resolution. We replace values above 1 with a local mean estimated from a 10 × 10 neighborhood and mask the field with the land mask so only land pixels remain.

To handle local gaps in the imperviousness product, we fill missing values in urban classes. For each UrbClim land-use class, we compute the mean soil-sealing fraction over pixels with valid values, and we assign this class mean to urban pixels where the sealing fraction is missing or ≤ 0. Outside urban classe, we set soil sealing to zero. In this way, the final map represents the fraction of impervious surface in each grid cell while remaining consistent with the land-use classes.

6.1.3.5 Roughness length in built-up areas

We obtain building information from two GHSL products for 2018:

- GHS-BUILT-S: built-up surface area at 10 m, which we use as a proxy for building fraction.
- GHS-BUILT-H: average building height at 100 m.

We clip both products to the Niamey domain, reproject them to EPSG:32631, and resample them to the UrbClim grid. We normalise built-up surface to a 0–1 building fraction, and we enforce a minimum building height of 10 m to avoid unrealistically small roughness elements.

Using these fields, we recompute the aerodynamic roughness length for momentum (z_{0m}) in built-up areas. We use:

- the monthly vegetation fraction of the peak-greenness month
- an “average urban tree cover” coefficient $b = 0.2$, and
- a simple combination of building fraction, building height and peak-season vegetation to estimate an effective roughness for buildings and trees:

$$z_{0m,bld} = f_{bld} \frac{H_{bld}}{10} + 2b f_{veg,peak}.$$

We clip this combined roughness to the range 0.3–2 m and overwrite the table-based z_{0m} for urban land-use classes. For all other land-use classes, we keep the roughness length value from the land-use parameter table.

6.1.3.6 Anthropogenic heat flux (AHF)

We derive anthropogenic heat flux from a global AHF dataset from Wang et al. (2022). We clip the AHF raster to the Niamey domain, reproject it to EPSG:32631, and resample it to the UrbClim grid. We do not apply any further scaling or temporal disaggregation, so the resulting field is a static annual-mean AHF map that we use as a constant anthropogenic heat source in the UrbClim simulations.

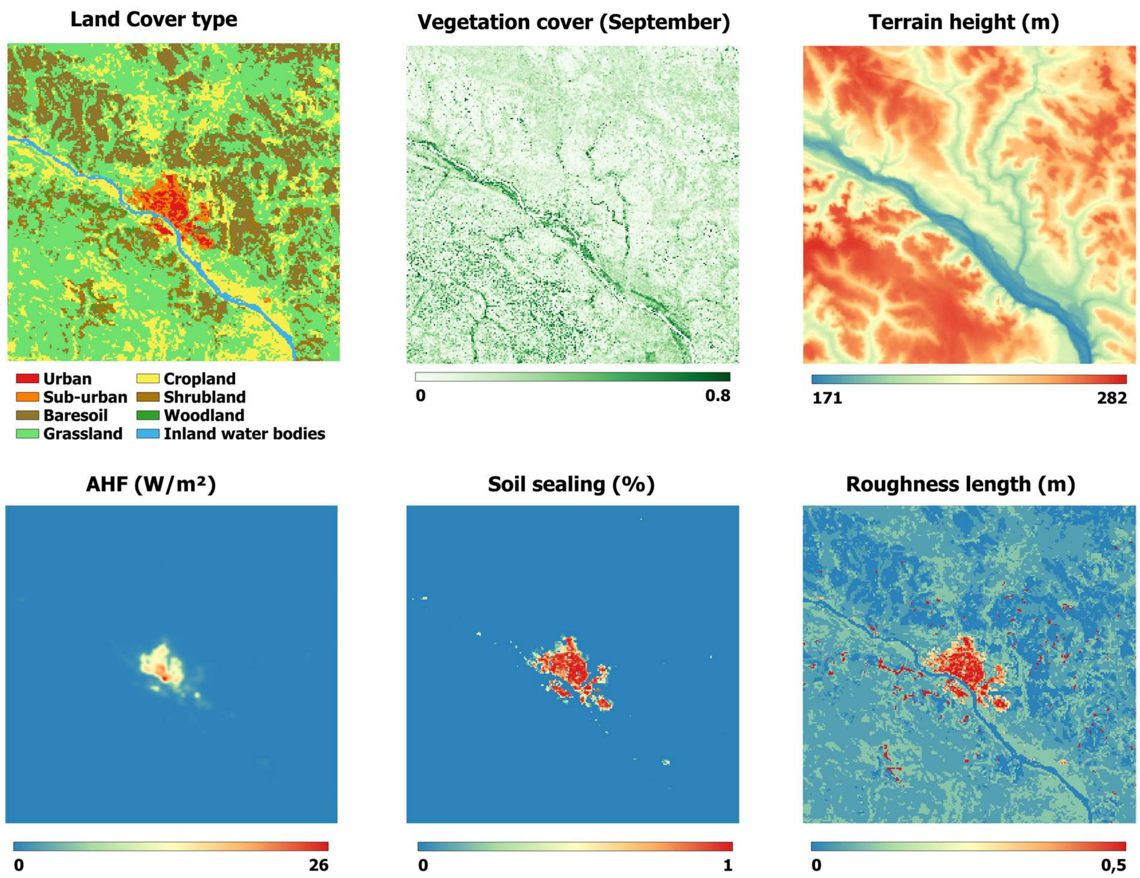


Figure 8. Map containing terrain parameters for Niamey at 300-m resolution, required as input for the UrbClim model. While the domain may appear over-dimensioned for a city as Niamey, this rather large size anticipates the urban growth scenarios.

6.1.3.7 Soil texture and hydraulic parameters

We take soil texture from a global USDA-style texture map at 250-m resolution. This map classifies each pixel into 12 texture classes (clay, silty clay, sandy clay, clay loam, loam, sandy loam, sand, …). We clip the texture map to the Niamey domain, reproject it, and resample it to the UrbClim grid using nearest-neighbour interpolation.

For each texture class, we assign soil hydraulic parameters following standard Clapp–Hornberger relationships and ECMWF land-surface documentation: volumetric porosity, saturation matric potential, saturated hydraulic conductivity, pore-size distribution index (see De Ridder et al., 2015 for details).

We use default “medium-texture soil” values where the texture map is undefined, and we overwrite these defaults with class-specific values where texture information is available. To remove small isolated artefacts introduced by re-sampling, we apply a weak Gaussian smoothing (σ = 3 grid cells) to each parameter field.

6.1.4 UrbClim simulation for Niamey

Once the meteorological forcing and terrain datasets prepared as described above, we use them as input to UrbClim, following the standard CAIAC model configuration outlined in section 3.2. For Niamey we run UrbClim on the 300-m grid defined in section 6.1.1, using ERA5-based boundary conditions for the period 2001–2020, and the land-use, vegetation and surface parameters derived in section 6.1.3. We do not apply any city-specific tuning of the physical parameterizations.

The resulting simulation provides a 20-year time series of hourly near-surface meteorological fields (air temperature, specific humidity, wind speed, land surface temperature) over the full Niamey domain. These fields form the basis for subsequent post-processing steps, where we derive wet bulb globe temperature (WBGT), heat-stress indicators and additional urban climate metrics that we downscale to 10-m resolution.

6.1.5 Wet bulb globe temperature (WBGT)

The WBGT is a composite heat-stress index that estimates the thermal strain on humans by combining the effects of air temperature, humidity, wind, and radiant heat into a single metric. We compute this quantity in a post-processing step based on the hourly UrbClim output, following Liljegren et al. (2008) and employing standard micrometeorological relations (e.g., Monteith and Unsworth, 2013).

The WBGT depends on the following input variables:

- simulated 2-m air temperature, 2-m specific humidity, 2-m wind speed, surface (skin) temperature, surface pressure, and downward short and longwave radiation at the surface;
- a pre-computed sky-view factor (SVF)[6], the land cover map and the monthly vegetation cover fraction.

Note that among the above-mentioned meteorological quantities, exposure to downwelling shortwave radiation is a key driver of high WBGT values (high heat stress). In the shade of a tree, downwelling radiation is strongly attenuated, thus decreasing the WBGT considerably.

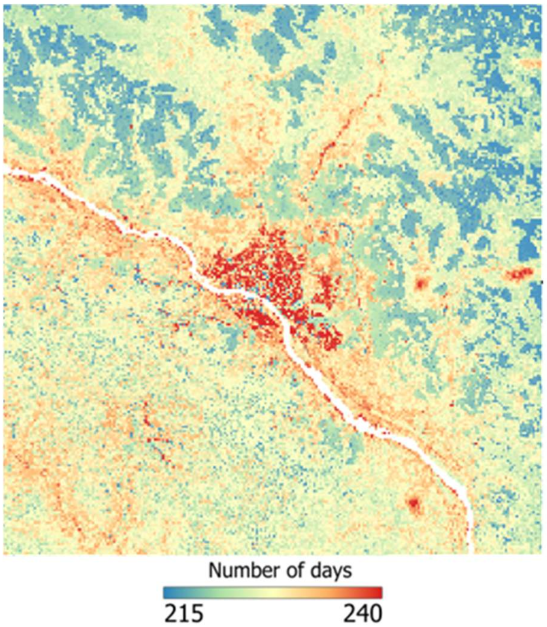
In our procedure, within each UrbClim grid cell, the WBGT is calculated separately for sun-exposed and tree-shaded locations. The grid-average WBGT value is calculated as a weighted average of both, using as weights the fractional vegetation (tree) cover. The latter is estimated using satellite-derived fractional vegetation cover, based on the NDVI (see above). Hence, the WBGT is crucially dependent on EO data through its dependence on vegetation cover. More information regarding the calculation of the WBGT and its dependence on satellite-based vegetation cover is provided in Section 7.1.1.

6.1.6 Indicator maps and downscaling

After computing UrbClim hourly fields and WBGT, we derive time-aggregated (*not* spatially aggregated) indicators, an overview of which is given in Annex 2. For each grid cell and year, we

[6] portion of the sky that is visible from a given location

aggregate the hourly time series to the relevant statistics. We then compute and average these yearly indicators over the 2001-2020 baseline period to obtain climatological indicator maps at the UrbClim resolution (300 m). Figure 9 shows examples of indicator maps generated with UrbClim for Niamey for the period 2001-2020.



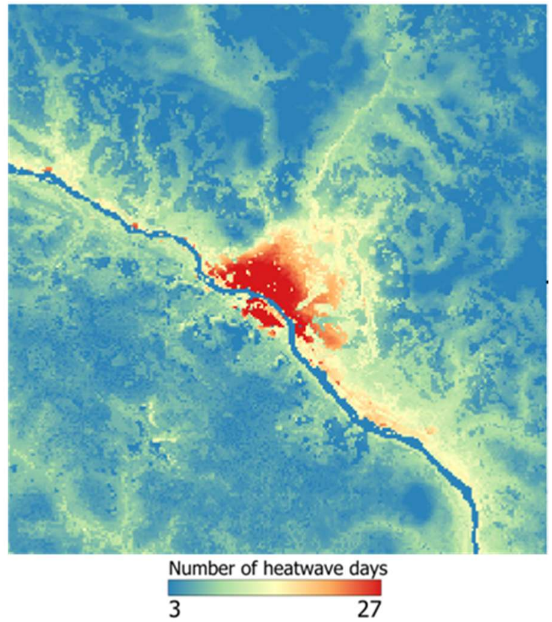


Figure 9. Example UrbClim indicator maps for Niamey, showing the number of days per year with WBGT above 28 °C (left) [7] *and the annual number of heatwave days (right).*

To better match local exposure data and urban morphology, we statistically downscale selected indicators from the UrbClim grid (300 m) to a finer-resolution grid (10 m), using a Random Forest regression model and Sentinel-2 NDVI data. For the regression, we prepare this set of high-resolution features:

- Land-use class from WorldCover;
- Mean vegetation fraction over the year from terrain input;
- Soil sealing from terrain input;
- Digital Elevation Model (DEM) from terrain input;
- Local slope magnitude derived from the DEM.

6.1.7 Computation time

We run all steps on an HPC cluster with a Slurm job management system. Each script is multi-thread optimised and usually loops over multiple simulation years. The Niamey domain configuration uses the maximum UrbClim domain size (301 × 301 grid cells), so its run times are high. Cities with smaller domains generally run faster. Actual wall clock times vary depending on queue load and on the specific node allocated (newer CPUs are faster).

[7] Caution is required when interpreting WBGT values, as these cannot be related to the regular air temperature ('dry bulb') temperature. The threshold WBGT value of 28°C – which may appear low – corresponds to conditions under which many occupational health guidelines already recommend limiting physical activity.

Table 3. Computing time required for a full-run UrbClim simulation.

Step	CPU time	Notes
Process ERA5	50 min	Extract ERA5 subset for 2001–2020 around Niamey.
Meteo preprocessor	18 min	Convert ERA5 subset to UrbClim-ready forcing files.
NDVI / vegetation fraction	90 min	Download S2 mosaics (2021–2024), build NDVI climatology and monthly vegetation fraction.
Terrain processor	3 min	Merge GDX 2020 land use with WorldCover, fix airports and inland water/sea discrimination. Assemble DEM, soil sealing, buildings, AHF and soil parameters into UrbClim surface file.
UrbClim simulation	129 h	20-year baseline simulation (2001-2020), 300 m resolution
post-processing	30 min	Aggregate hourly fields to climatological indicators on the UrbClim grid.
WBGT calculation	40 min	Prepare the necessary fields from UrbClim output, meteo file and terrain for WBGT. Compute hourly WBGT fields.
Indicator downscaling	20 min	Random-Forest downscaling of indicators to the 10 m grid.

6.2 Urban flooding

The University of Ottawa developed a Rapid Inundation Labelling (RIL) approach that uses dynamic essential climate variables (ECVs) together with static flood-risk layers to generate instantaneous inundation maps. To our knowledge, this is the first framework that derives per-pixel flood thresholds directly from hydrological forcing variables rather than relying exclusively on SAR-derived flood extents, which can be limited in urban environments. As a proof of concept, the RIL framework was applied to a case study in Niamey.

Within the RIL framework, river discharge is adopted as the driving variable for fluvial flood threshold estimation, whereas surface runoff is used for pluvial flood threshold estimation for each pixel. Surface runoff is simulated using the physically based hydrological model SWAT+, whereas river discharge is derived from observed streamflow records provided by the Global Runoff Data Centre (GRDC). For each spatial pixel, inundation depths for selected flood return periods obtained from the Aqueduct database are paired with the corresponding predictor values. A functional relationship (e.g., exponential function) is then fitted between each predictor and the inundation response. Model skill (e.g., coefficient of determination, R^2) is used as a measure to evaluate the accuracy of the function fit. From the selected predictor–response curve, a pixel-specific inundation threshold is subsequently extracted to distinguish between inundated and non-inundated states.

Urban areas are delineated using a land-use/land-cover dataset, and thresholds are estimated independently for each urban pixel. This ensures that the RIL method accounts for the distinct, often rapid, hydrological response observed in built environments. The performance of the RIL model will be assessed by comparing simulated inundation maps with observed historical flood events, and by benchmarking against an alternative model configuration to evaluate robustness and relative skill.

To ensure the physical realism of the derived thresholds, the fitted predictor–response curves are inspected for characteristic behaviours. A well-behaved curve should exhibit low inundation probability at low forcing values, a steep transition near the threshold, and saturation under high forcing conditions. Urban pixels are expected to display sharper transitions due to enhanced runoff generation. Conversely, curves that are flat, noisy, or inconsistent indicate weak hydrological control and may signal pixels where threshold estimation is unreliable.

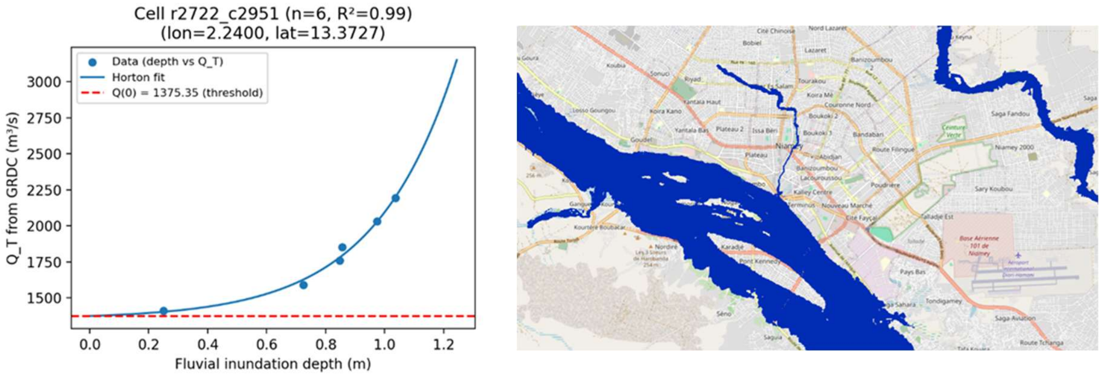


Figure 10. Left: Fluvial Discharge–Inundation Depth Fit (representative Niamey pixel). Right: RIL METHOD map for the University area in Niamey.

As an illustrative example, the fitted curve is presented for a representative pixel in Niamey located near the river and primarily influenced by fluvial flooding, where the y-axis corresponds to river discharge and the x-axis represents inundation depth. The fit and the corresponding plot show a clear and consistent alignment between river discharge and inundation depth, indicating a strong correlation. Based on the derived threshold, the daily inundation depth in the time series is compared against this value, and during fluvial flood events this pixel is expected to become inundated. Inundation depth is currently derived from the Aqueduct dataset; however, a key limitation is its relatively low spatial resolution. To address this, our team concluded that high-resolution inundation depths should be generated by applying a HEC-RAS modelling workflow. Nevertheless, a snapshot of the current RIL output for Niamey is presented in Figure 10, in which inundated and dry cells can be clearly distinguished.

7 CRITICAL APPRAISAL

This section describes an assessment of the methodology, considering (1) uncertainty, (2) validation, (3) risks, and (4) added value of Earth Observation.

7.1 Uncertainty

The CAIAC project considers two different sorts of time frames:

- historical (recent past or 'present-day', typically the period 2001-2020);
- future projection (e.g., 2041-2060).

The uncertainty associated with each (past versus future) is fundamentally different: past heat and flooding simulations make direct use of satellite and re-analysis data. This means that the errors on the final outcome (e.g., heat stress indicator map) are directly associated with the physical errors that come with the satellite imagery and re-analysis data.

In the case of the future projections, uncertainty largely stems from the uncertainty surrounding the scenarios (urban growth, climate scenario). This scenario-related uncertainty will be addressed later in the project, making use of the uncertainty contained in the CMIP6 ensemble forcing. In the remainder of this section, we consider the uncertainty associated with present-day conditions.

7.1.1 UrbClim

Considering UrbClim's chain of processing and modelling steps (Section 6.1), uncertainty of the historical simulations could in principle be addressed by incorporating each stage of the modelling chain in a Monte Carlo analysis, from EO data to urban growth to urban climate simulation and, ultimately, climate indicators. While such a full end-to-end Monte Carlo simulation is theoretically ideal, computational constraints[8] motivate a more pragmatic approach.

In the remainder of this section, we describe a simplified yet relevant Monte Carlo–based uncertainty analysis. The analysis focuses on uncertainty propagated to the wet-bulb globe temperature (WBGT) arising from two sources: (1) uncertainty in the satellite-derived vegetation cover fraction (f_{veg}) and (2) uncertainty in the atmospheric input variables (Figure 11). We focus on WBGT because it constitutes the core heat-stress indicator in CAIAC. In addition, WBGT is particularly sensitive to vegetation cover, owing to the latter's modulation of the radiative environment through shading.

The WBGT is calculated as

$$WBGT = 0.7\,T_w + 0.2\,T_g + 0.1\,T_g,$$

with T_w the natural wet bulb temperature, T_g the black globe temperature, and T_a the air temperature.

Note that we apply a Monte Carlo analysis, rather than a purely analytical / theoretical error propagation, because the WBGT is a highly non-linear function of the atmospheric variables through T_w and T_g. Both require an iterative calculation, which excludes the WBGT from a simple theoretical analysis of error propagation.

As mentioned before, the WBGT is calculated separately for sun-exposed and tree-shade conditions, and brought to grid level using f_{veg} as a weight:

[8] This would require running the UrbClim model potentially hundreds to thousands of times, to generate robust Monte Carlo statistics.

$$WBGT = f_{veg} WBGT_{tree} + (1 - f_{veg}) WBGT_{sun}.$$

This expression confirms the strongly direct impact of vegetation cover (and its uncertainty) on the WBGT. The difference between the sun-exposed and tree-shade values of the WBGT arise from the fact that, below a tree canopy, incoming solar radiation is strongly attenuated, thus greatly reducing the radiation load. Moreover, exposure to longwave radiation is reduced under tree canopies, because the shaded ground surface radiates much less thermal infrared radiation than heated-up soils under the influence of solar radiation loading. The radiation reaching below the canopy is calculated as in De Ridder (1997) and Widlowski et al. (2011).

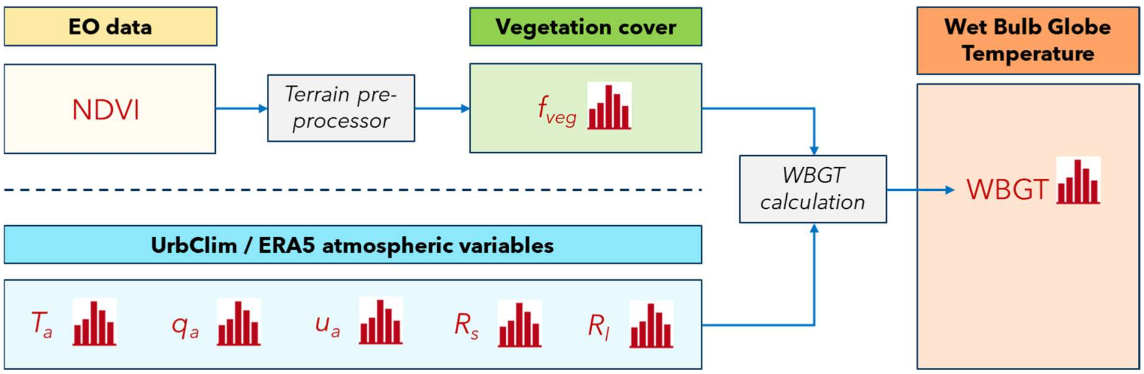


Figure 11. Schematic layout of the Monte Carlo uncertainty analysis. The satellite-(derived) data are in the upper left quadrant, the atmospheric variables in the lower left. f_{veg} represents the vegetation cover fraction, and the atmospheric variables in the blue box correspond to, from left to right, air temperature, specific humidity, wind speed, and downwelling shortwave and longwave radiation. Both the satellite and atmospheric variables contribute to the WBGT (right). Note that we do not assign a probability distribution to the NDVI, but instead assume that this is absorbed in f_{veg}.

In the remainder of this section, the uncertainty on the WBGT is estimated from the uncertainty and probability distribution functions of the following input variables:

- Uncertainty on the vegetation cover fraction (f_{veg}) is assigned a value of 0.17. This is loosely based on values presented in the literature, notably Filipponi et al. (2018), citing an uncertainty of 0.15; Liu et al. (2019), citing uncertainty values in the range 0.12-0.22; Fuster et al. (2020), who cite an uncertainty of 0.21. The probability distribution function assigned to f_{veg} is a beta-distribution, since vegetation cover fraction is constrained to the range [0,1].
- Air temperature is assigned an uncertainty of 1.5°C, based on results in De Ridder et al. (2015), together with a gaussian distribution.
- For specific humidity we assume an uncertainty of 10%. This is derived from a validation exercise in Souverijns et al. (2022) for Johannesburg. We assign a gaussian probability distribution.
- Wind speed – notoriously difficult to simulate in urban environments – is assigned an uncertainty of 50% (De Ridder et al., 2015). We also impose a lower limit on wind of 0.5 m s^{-1} to avoid some of the equations in the WBGT to blow up. This can be further justified by invoking that even at a very low mean wind speed, turbulence still causes ventilation to occur (especially during daytime conditions under a high solar load as will be the case here). Wind is represented through a Weibull distribution, ackowledging that the wind magnitude cannot become negative.

- For the radiation components (which stem from ERA5), little uncertainty information appears to be available, at least not for daytime hourly clear-sky conditions. We assign a 10% uncertainty to both shortwave and longwave radiation components, and a gaussian distribution function.

As base climatological conditions, we consider (1) a hot and dry situation and (2) a warm humid situation, as shown in Table 4, considering these as representative for a variety of African climate zones. As a base value for the fractional vegetation cover, we use f_{veg} = 0.35. The Monte Carlo propagation analysis was applied in Python, yielding a probability distribution for the WBGT as shown in Figure 12, which represents the hot-dry case. The warm-humid case (not shown here) yields a fairly similar result.

Table 4. Base values of the meteorological variables used in the Monte Carlo analysis. Note that the relative humidity values presented here are converted to specific humidity before use in the WBGT calculations.

	hot-dry	**warm-humid**
Air temperature (T_a)	40°C	30°C
Relative humidity (RH)	20%	80%
Land surface temperature (T_s)	55°C	40°C
Downward SW radiation (R_s)	900 W m^{-2}	
Downward LW radiation (R_l)	400 W m^{-2}	

The observed spread in WBGT values (σ = 1.55 °C, Figure 12) should be compared with the granularity of the WBGT heat-stress categories, which are separated by similar WBGT intervals, i.e., WBGT categories are typically 1.5-3 °C apart, implying that the uncertainty derived from the Monte Carlo analysis could result in a classification error of (at most) one category away from the true category.

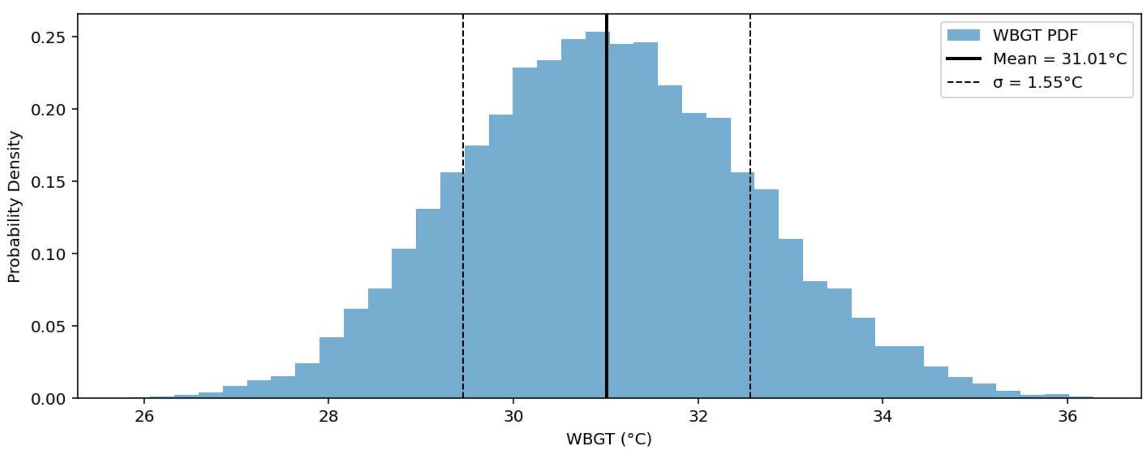


Figure 12. Probability distribution of the WBGT resulting from uncertainty on the vegetation cover and a set of meteorological quantities affecting the WBGT. This result pertains to the hot-dry case (Table 4).

7.1.2 SAHEL

The uncertainty of the SAHEL framework will be assessed through a multi-stage evaluation strategy that explicitly considers the main sources of uncertainty in the modelling chain, including (i) scenario forcing uncertainty (differences across climate projections and urbanization pathways), (ii) physical-model uncertainty (uncertainty in SWAT+ simulations of runoff and discharge used to inform inundation dynamics and training data generation), (iii) model uncertainty (convolutional neural network (CNN) ensemble structure, parameterization, and training stochasticity), and (iv) data and label uncertainty (uncertainty in EO/ECV inputs and RIL-generated inundation targets). SAHEL outputs will be compared against documented historical flood events and satellite-derived inundation observations to evaluate how well the model reproduces past flood patterns, and to examine whether prediction errors and uncertainty are larger under specific hydroclimatic conditions or urban settings. This evaluation will consider overall prediction performance as well as the consistency of inundation outputs and the spread of the ensemble predictions.

Model uncertainty will be assessed by comparing SAHEL's baseline CNN ensemble with alternative deep-learning architectures trained on the same inputs and targets (e.g., U-Net/ResU-Net backbones versus a temporal CNN variant such as CNN–ConvLSTM, or an attention-based segmentation backbone). Differences in predicted inundation maps across these configurations will be analyzed across climate and urbanization scenarios to quantify sensitivity to architectural choices and modelling assumptions, and to identify where structural uncertainty dominates the projections.

In addition, SAHEL will assess scenario, input, and label uncertainty through general sensitivity analyses. The model will be tested under plausible variations in climate forcing and urban growth layers, as well as variations in the input data and training labels, to examine how stable inundation labeling outputs remain when data quality, availability, or scenario assumptions change. This includes evaluating the influence of uncertainty in EO/ECV drivers, the influence of SWAT+-derived hydrologic variables, and the uncertainty associated with RIL-generated inundation maps used for training.

7.2 Validation strategy

UrbClim **urban heat** simulation results for historical periods (2001-2020) will be validated against ground-based observations of key climate variables, including air temperature, humidity, and downwelling shortwave radiation, where available. ACMAD is currently providing these ground-based measurements, which will be used to assess model accuracy by comparing modelled and observed values. This evaluation will generate uncertainty estimates for past and present conditions, providing insight into potential biases and model performance under historical scenarios.

Given the scarcity of urban climate stations and the limited availability of archived climate data in Africa, validation will rely heavily on satellite-derived datasets. In particular, Land Surface Temperature (LST) data will be used to evaluate the Surface Urban Heat Island (SUHI) effect simulated by UrbClim.

The validation of the SAHEL **flooding** framework will be carried out by comparing model outputs with documented historical flood events and satellite-derived inundation observations, in order to assess how well the model reproduces past flood patterns. This evaluation will consider overall prediction performance, the consistency of the simulated inundation outputs, and the spread of the ensemble predictions.

Given the limited availability of in situ flood observations, the validation will rely on EO-derived inundation data. In addition, the robustness of the SAHEL framework will be assessed by comparing inundation maps produced by the baseline convolutional neural network (CNN) ensemble with those obtained using alternative deep-learning architectures trained on the same inputs and targets, in order to evaluate sensitivity to modelling choices across different climatic scenarios and urbanisation projections.

7.3 Risks and mitigation approach

Risk	***Mitigation***
Sparse in-situ climate monitoring infrastructure limits the ability to validate climate model outputs, potentially affecting confidence in results.	*Complement ground-based data with land surface temperature from Earth observation imagery.*
Apart from native CMIP6 climate projections, we will use the derived / calibrated NASA NEX-GDDP-CMIP6 dataset – i.e., a different data set.	*We will take care to ensure consistency by enforcing common SSP / time horizon combinations.*
Inclusion of urban growth up to 2100 demands caution, considering the large uncertainties in the urban growth modelling associated with projecting demographic, economic, and land-use patterns over longer timescales.	*This is unavoidable, yet we are aware of the limitations and will communicate transparently about it.*
Global Warming Levels (GWLs) are not tied to a specific point in time but are rather climate states that may be reached at different times under different SSPs, depending on emission trajectories. In contrast, urban growth projections are time-bound, typically given for specific years (e.g., 2050) along defined SSP narratives.	*While strict alignment between a GWL analysis and urban growth projections is not always feasible, we have established a credible correspondence between GWL and scenario / time horizon (Section 4).*
The urban growth modelling (GeoDynamiX) avoids creating new built areas in flood-prone zones. However, such areas tend to attract low-income and vulnerable populations, and are likely to be occupied regardless, unless strong legislation is in place. While we could consider two growth trajectories (one in which the occupation of flood-prone areas is discouraged, and another in which it is not), this is not so easy to implement in GeoDynamiX.	*We are aware of this limitation and will communicate transparently about it.*

7.4 Added value of Earth Observation

The best and most straightforward way to understand and appreciate the importance of EO data for **urban climate (heat)** modelling is to consider model input and output fields such as those presented in Figure 8 and Figure 9. While the magnitude of the heat indicators stems from the model physics, it is clear that the spatial layout of the urban heat indicators (Figure 9) strongly mirrors the EO-based terrain input fields (Figure 8). Additionally, EO data provide another important added value. As shown above (Section 7.1.1), EO data – and satellite-based vegetation cover fraction in particular – play a key role in determining heat-stress intensity, as expressed by the wet-bulb globe temperature (WBGT), or any other heat stress indicator (e.g., UTCI).

More fundamentally, EO data form the essential basis for any meaningful urban climate simulation. In models such as UrbClim, the spatial variability in the simulated fields arises entirely from the underlying EO-derived representation of land cover, urban form, and surface properties. The atmospheric forcing (e.g. ERA5) is too coarse to resolve urban areas and does not contain any explicit urban signal. Without EO-based input, the model would therefore lack any description of urban structure. As a consequence, the simulations would collapse to spatially homogeneous, “flat” fields, with only weak contrasts between broad natural surface types. In such a configuration, characteristic urban climate features, such as the urban heat island, would not emerge at all. In this sense, **EO data are not only indispensable, but a prerequisite for the model to function as an urban climate model.**

This effect is illustrated in Figure 13 for a night with strong UHI intensity over Niamey, Niger, by comparing UrbClim output with ERA5 for a transect over the city. While the ERA5 results display nearly no temperature variation over the domain, UrbClim shows UHI intensities of around 6°C. It is also clear that UrbClim closely follows the EO-based land cover patterns: for instance, the intrusion of non-urban land cover around 5 km east of the centre (left panel in Figure 13) corresponds to the dip at x = 5 km in the right panel of Figure 13.

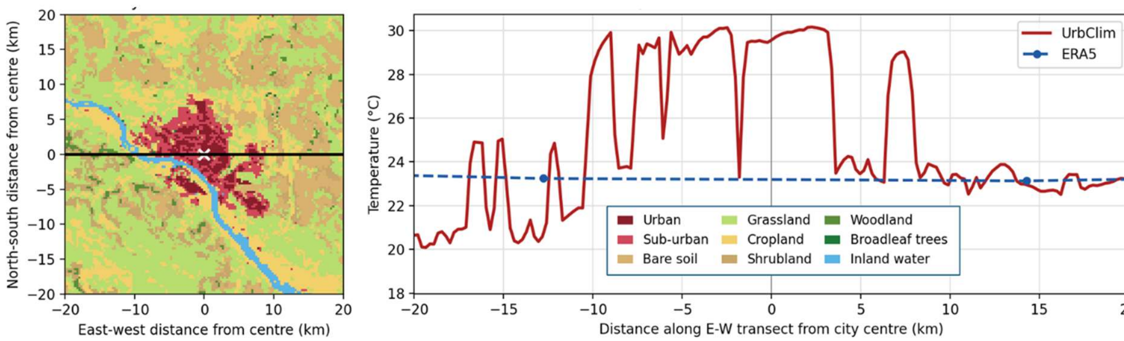


Figure 13. Nighttime temperature transect over Niamey for a selected moment (2 April 2001, 05UTC), for UrbClim (red line) and ERA5 (dashed blue line).

In a similar way, the importance of EO data for understanding and modelling **urban flooding** can be illustrated by considering key hydrometeorological input fields and the resulting flood indicators. While the magnitude and timing of flooding ultimately depend on hydrological model physics, it is evident that the spatial patterns of urban inundation strongly reflect EO-derived forcing fields such as precipitation, soil moisture, and land-surface characteristics. EO data provide a further crucial added value for urban flooding applications. As shown previously, satellite-based observations of rainfall and antecedent surface wetness play a key role in determining flood occurrence and intensity, particularly in data-scarce urban environments. When combined with physically based hydrological modelling, EO-derived variables enable a physically consistent representation of inundation processes and provide a robust basis for flood-risk assessment under present and future climate conditions.

8 CONCLUSIONS

This report has defined a scientifically robust methodology for analysing urban heat stress and flooding in African cities through an integrated use of Earth Observation data, climate projections, urban growth modelling, and impact models. The methodology demonstrates that it is technically feasible to generate spatially consistent and impact-relevant indicators for African cities despite data scarcity and strong heterogeneity in climate, urban form, and socio-economic conditions.

A central achievement of the methodology is the explicit coupling of future urban growth with future climate forcing, enabling the respective contributions of urbanisation and climate change to be disentangled in a consistent manner. This approach allows urban risk trajectories to be analysed in a way that goes beyond purely climate-driven assessments and provides a stronger basis for understanding where and why risks are expected to increase.

The feasibility studies confirm that the full modelling workflow – from EO data ingestion to indicator generation – is operational and computationally tractable. The critical appraisal further clarifies the main sources of uncertainty, the validation strategy, and the limitations inherent to long-term projections, ensuring that results can be interpreted transparently and responsibly.

With the methodology now specified and demonstrated, the project is ready to enter the next phase, in which the complete modelling chain will be applied to all 54 selected African cities. This phase will generate a unique dataset of urban heat and flood indicators under present and future scenarios. The resulting multi-city analysis will yield new insights into how urban climate risks vary across climate zones and city sizes, how urban growth amplifies or mitigates these risks, and how the effectiveness of adaptation measures such as green infrastructure depends on local context.

In this way, the methodology described in this report provides a solid foundation for the forthcoming analysis phase of CAIAC, supporting both scientific discovery and the development of actionable knowledge for climate adaptation in African cities.

REFERENCES

Beck, H.E., McVicar, T.R., Vergopolan, N., Alexis, B., Lutsko, N.J., Dufour, A., Zeng, Z., Jian, X., van Dijk, A.I.J.M., Miralles, D.G, 2023. High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections. Sci Data 10, 724. https://doi.org/10.1038/s41597-023-02549-6

Crols, T., 2017. Integrating network distances into an activity based cellular automata land-use model. Semi-automated calibration and application to Flanders, Belgium. PhD thesis, Vrije Universiteit Brussel and VITO, Brussels, Belgium, and Mol, Belgium. https://geodynamix.eu/sites/geodynamix/files/PhD_thesis_Tomas_Crols_Fin.pdf

De Ridder, K., 1997. Radiative transfer in the IAGL land surface model. Journal of Applied Meteorology, 36, 12-21.

De Ridder, K., D. Lauwaet, B. Maiheu, 2015. UrbClim – a fast urban boundary layer climate model. Urban Climate, 12, 21-48. https://doi.org/10.1016/j.uclim.2015.01.001

Filipponi, F., Valentini, E., Nguyen Xuan, A., Guerra, C. A., Wolf, F., Andrzejak, M., & Taramelli, A., 2018. Global MODIS Fraction of Green Vegetation Cover for Monitoring Abrupt and Gradual Vegetation Changes. Remote Sensing, 10, 653. https://doi.org/10.3390/rs10040653

Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., Weiss, M., Baret, F., & Smets, B., 2020. Quality Assessment of PROBA-V LAI, fAPAR and fCOVER Collection 300 m Products of Copernicus Global Land Service. Remote Sensing, 12, 1017. https://doi.org/10.3390/rs12061017

García-Díez, M., D. Lauwaet, H. Hooyberghs, J. Ballester, K. De Ridder, and X. Rodó, 2016. Advantages of using a fast urban boundary layer model as compared to a full mesoscale model to simulate the urban heat island of Barcelona. Geoscientific Model Development, 9, 4439–4450. https://gmd.copernicus.org/articles/9/4439/2016

Gutman, G., Ignatov, A., 1998. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. International Journal of Remote Sensing 19, 1533–1543. https://doi.org/10.1080/014311698215333

Huang, X., Yang, J., Wang, W., Liu, Z., 2022. Mapping 10 m global impervious surface area (GISA-10m) using multi-source geospatial data. Earth Syst. Sci. Data 14, 3649–3672. https://doi.org/10.5194/essd-14-3649-2022

IPCC, 2021. Technical Summary. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 33–144. https://doi.org/10.1017/9781009157896.002

IPCC, 2023. Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the 6th Assessment Report of the IPCC [Core Writing Team, H. Lee and J. Romero (eds.)]. Intergovernmental Panel on Climate Change, Geneva, Switzerland, 184 pp., doi: 10.59327/IPCC/AR6-9789291691647. https://www.ipcc.ch/report/ar6/syr/

Jiang, L., and O'Neill, B. C. (2017). Global urbanization projections for the Shared Socioeconomic Pathways. *Global Environmental Change*, *42*, 193-199. https://doi.org/10.1016/j.gloenvcha.2015.03.008

K.C., S., Dhakad, M., Potancokova, M., Adhikari, S., Yildiz, D., Mamolo, M., Sobotka, T., Zeman, K., Abel, G., Lutz, W., and Goujon, A. (2024). Updating the Shared Socioeconomic Pathways (SSPs) Global Population and Human Capital Projections. IIASA Working Paper. Laxenburg, Austria:

WP-24-003. https://pure.iiasa.ac.at/19487.
Available at: http://www.wittgensteincentre.org/dataexplorer

Liljegren, J.C., Carhart, R.A., Lawday, P., Tschopp, S., Sharp, R., 2008. Modeling the Wet Bulb Globe Temperature Using Standard Meteorological Measurements. Journal of Occupational and Environmental Hygiene 5, 645–655. https://doi.org/10.1080/15459620802310770

Liu, D., Jia, K., Wei, X., Xia, M., Zhang, X., Yao, Y., Zhang, X., & Wang, B., 2019. Spatiotemporal Comparison and Validation of Three Global-Scale Fractional Vegetation Cover Products. Remote Sensing, 11, 2524. https://doi.org/10.3390/rs11212524

Melchiorri, M., 2022. The global human settlement layer sets a new standard for global urban data reporting with the urban centre database. Frontiers in Environmental Science, 10:1003862. https://doi.org/10.3389/fenvs.2022.1003862

Monteith J., Unsworth M., Principles of Environmental Physics, 2013. 4th ed. Amsterdam: Elsevier. https://doi.org/10.1016/C2010-0-66393-0

Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., ... and Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global environmental change*, *42*, 153-168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Souverijns, N., K. De Ridder, N. Veldeman, F. Lefebre, F. Kusambiza-Kiingi, W. Memela, N. K.W. Jones, 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.

UNEP, 2023. Emissions Gap Report 2023: Broken Record – Temperatures hit new highs, yet world fails to cut emissions (again). United Nations Environment Programme, Nairobi, 108 pp. https://doi.org/10.59117/20.500.11822/43922

Wang, S., Hu, D., Yu, C., Wang, Y., Chen, S., 2022. Global mapping of surface 500-m anthropogenic heat flux supported by multi-source data. Urban Climate 43, 101175. https://doi.org/10.1016/j.uclim.2022.101175

White, R., Uljee, I., and Engelen, G. (2012). Integrated modelling of population, employment and land-use change with a multiple activity-based variable grid cellular automaton. Int. J. Geogr. Inf. Sci., 26, 1251–1280. https://doi.org/10.1080/13658816.2011.635146

Widlowski, J.-L., B. Pinty, M. Clerici, Y. Dai, M. De Kauwe, K. De Ridder, A. Kallel, H. Kobayashi, T. Lavergne, W. Ni-Meister, A. Olchev, T. Quaife, S. Wang, W. Yang, Y. Yang, and H. Yuan, 2011. RAMI4PILPS: An intercomparison of formulations for the partitioning of solar radiation in land surface models, J. Geophys. Res., 116, G02019, doi:10.1029/2010JG001511.

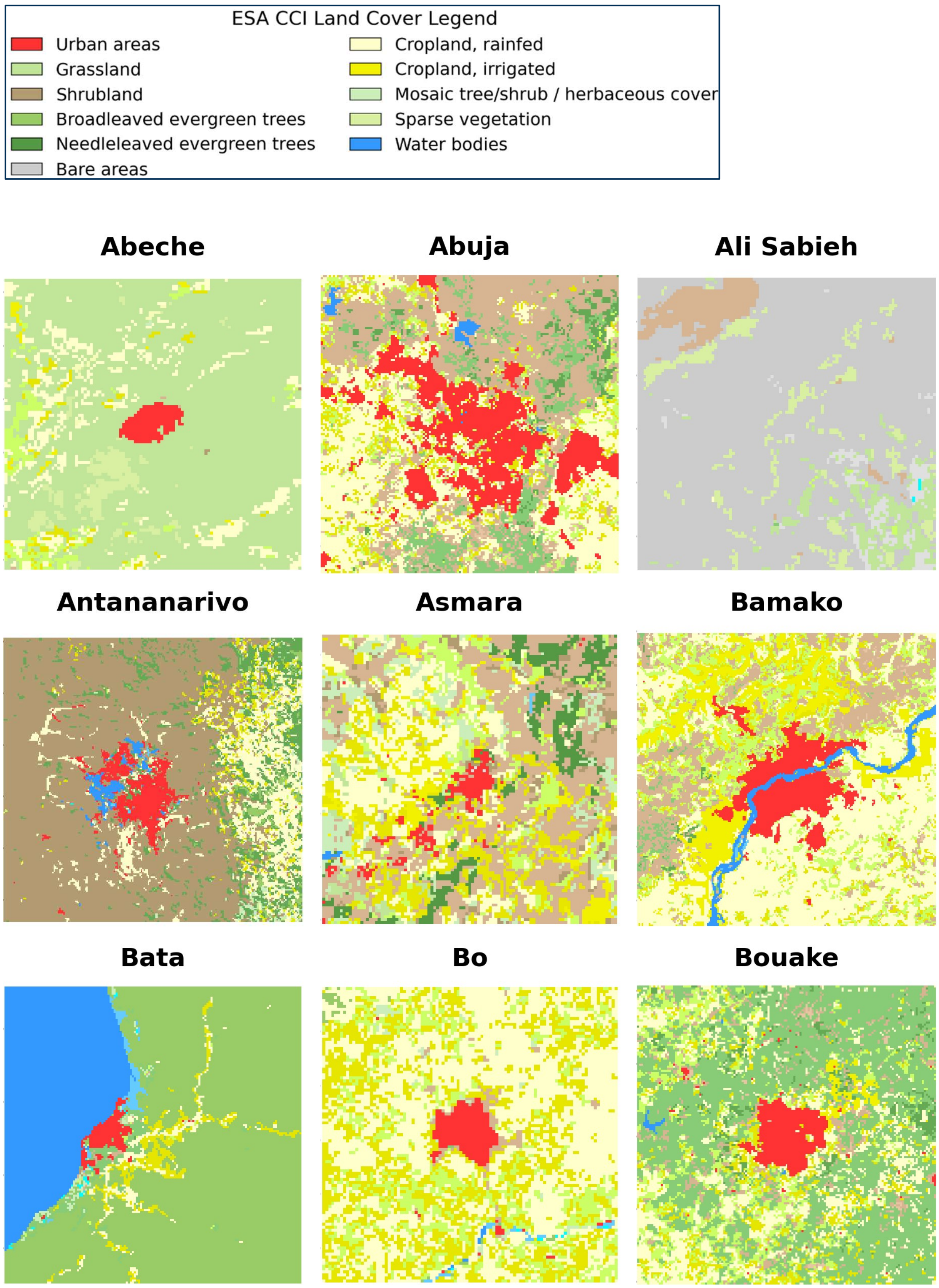
ANNEX 1: SELECTED CITIES

The table below shows the CAIAC project's selected cities, including their country code (ISO3), latitude-longitude coordinates, number of inhabitants, and Köppen climate zone. All cities in this list will have heat simulations performed. The cities annotated with a clear blue background are selected for pluvial flooding (flash floods), the dark blue cities for pluvial *and* fluvial (river) flooding.

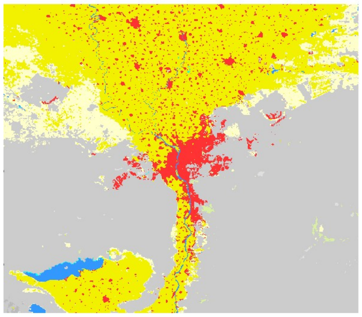
#	city	country	iso3	lat	lon	# inhs.	Köppen
1	Tiaret	Algeria	DZA	35.363	1.325	246774	BSk
2	Luanda	Angola	AGO	-8.881	13.295	11672134	BSh
3	Parakou	Benin	BEN	9.350	2.622	392541	Aw
4	Gaborone	Botswana	BWA	-24.613	25.876	298993	BSh
5	Ouahigouya	Burkina Faso	BFA	13.564	-2.416	129506	BSh
6	Gitega	Burundi	BDI	-3.429	29.923	176123	Aw
7	Praia	Cabo Verde	CPV	14.918	-23.509	159050	BWh
8	Yaounde	Cameroon	CMR	3.840	11.517	5479204	Aw
9	Carnot	Central African Republic	CAF	4.943	15.876	90317	Aw
10	Abeche	Chad	TCD	13.825	20.830	395145	BWh
11	Moroni	Comoros	COM	-11.707	43.256	175832	Am
12	Brazzaville	Congo	COG	-4.246	15.271	3190743	Aw
13	Bouake	Cote d'Ivoire	CIV	7.701	-5.033	704949	Aw
14	Kinshasa	DR Congo	COD	-4.383	15.327	12945683	Aw
15	Ali Sabieh	Djibouti	DJI	11.153	42.706	86781	BWh
16	Cairo	Egypt	EGY	30.069	31.255	25230325	BWh
17	Bata	Equatorial Guinea	GNQ	1.848	9.786	789800	Af
18	Asmara	Eritrea	ERI	15.332	38.929	251980	Cfb
19	Mbabane	Eswatini	SWZ	-26.317	31.133	94827	Cwb
20	Naqamte	Ethiopia	ETH	9.090	36.557	217687	Csb
21	Franceville	Gabon	GAB	-1.625	13.592	300424	Aw
22	Serrekunda	Gambia	GMB	13.408	-16.687	1313465	Aw
23	Kumasi	Ghana	GHA	6.691	-1.614	4923115	Aw
24	Kankan	Guinea	GIN	10.365	-9.295	287892	Aw
25	Gabu	Guinea-Bissau	GNB	12.285	-14.231	86491	Aw
26	Nairobi	Kenya	KEN	-1.265	36.852	6646257	Aw
27	Maseru	Lesotho	LSO	-29.327	27.516	296516	Cwb
28	Gbarnga	Liberia	LBR	6.999	-9.469	77865	Am
29	Tripoli	Libya	LBY	32.845	13.200	1248354	BSh
30	Antananarivo	Madagascar	MDG	-18.867	47.536	4255254	Cwb
31	Zomba	Malawi	MWI	-15.401	35.322	77471	Cwa
32	Bamako	Mali	MLI	12.622	-7.980	3960563	Aw
33	Nouakchott	Mauritania	MRT	18.087	-15.966	1625384	BWh
34	Port Louis	Mauritius	MUS	-20.235	57.476	557431	Am
35	Casablanca	Morocco	MAR	33.581	-7.543	4577677	Csa
36	Tete	Mozambique	MOZ	-16.158	33.593	470794	BSh
37	Windhoek	Namibia	NAM	-22.528	17.047	417842	BWh
38	Niamey	Niger	NER	13.527	2.117	1600656	BSh
39	Abuja	Nigeria	NGA	9.069	7.434	1322229	Aw
40	Kigali	Rwanda	RWA	-1.951	30.095	1621760	Aw
41	Sao Tome	Sao Tome and Principe	STP	0.333	6.728	154878	Aw
42	Darou Mousty	Senegal	SEN	15.040	-16.049	53487	BSh
43	Victoria	Seychelles	SYC	-4.620	55.455	78045	Af
44	Bo	Sierra Leone	SLE	7.955	-11.736	376731	Am
45	Galkayo	Somalia	SOM	6.770	47.427	187162	BWh
46	Johannesburg	South Africa	ZAF	-26.182	28.039	8592843	Cwb
47	Juba	South Sudan	SSD	4.854	31.568	296087	Aw
48	Khartoum	Sudan	SDN	15.606	32.508	7114789	BWh
49	Sokode	Togo	TGO	8.987	1.143	184962	Aw
50	Tunis	Tunisia	TUN	36.812	10.185	2790169	Ocean
51	Mbale City	Uganda	UGA	1.083	34.192	380225	Aw
52	Dodoma	Tanzania	TZA	-6.164	35.744	267333	BSh
53	Lusaka	Zambia	ZMB	-15.398	28.290	3658064	Cwa
54	Harare	Zimbabwe	ZWE	-17.882	31.042	2076037	Cwb

	pluvial
	pluvial + fluvial

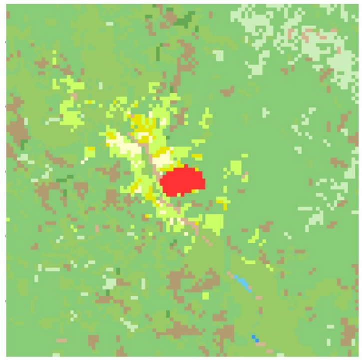
In the remainder of this Annex we provide an overview of the CCI land cover throughout the domains foofr all 54 selected cities. (Note that Kinshasa and Brazzaville are in a single common domain, which explains why there are only 53 images.)



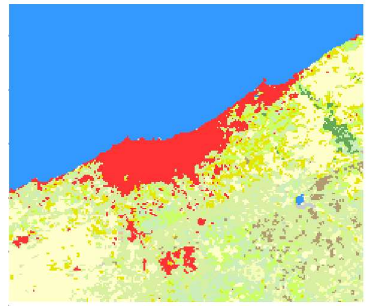
Cairo



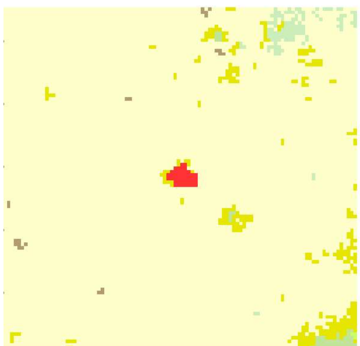
Carnot



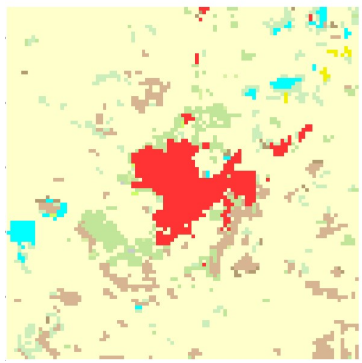
Casablanca



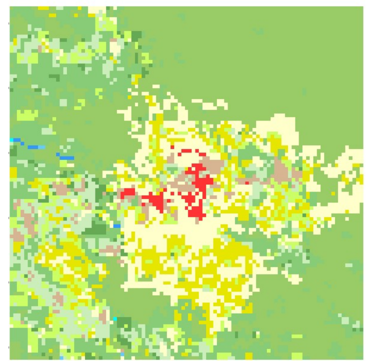
Darou Mousty



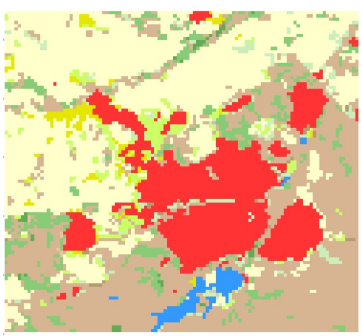
Dodoma



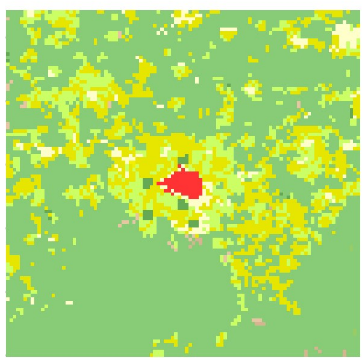
Franceville



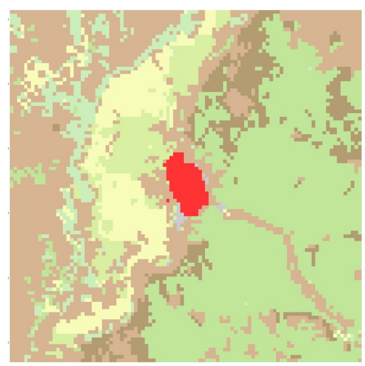
Gaborone



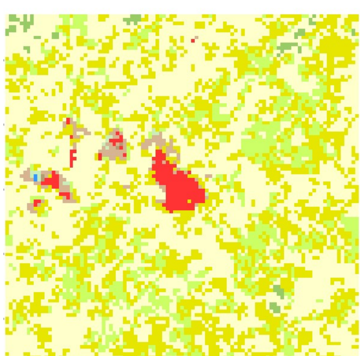
Gabu



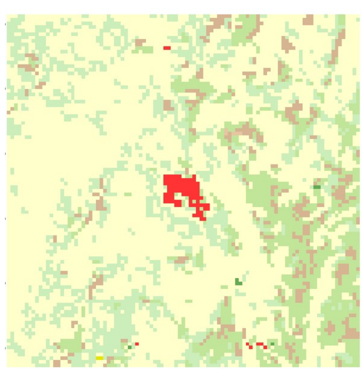
Galkayo



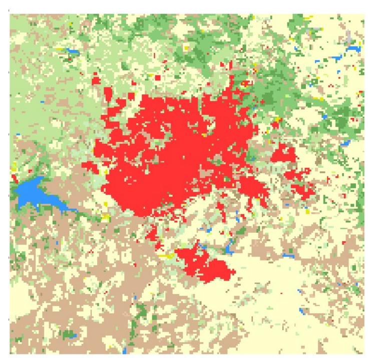
Gbarnga



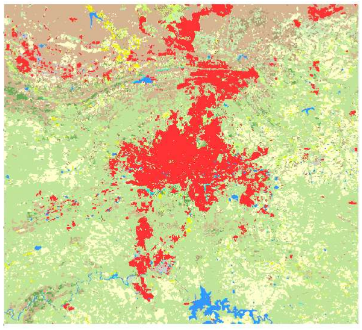
Gitega



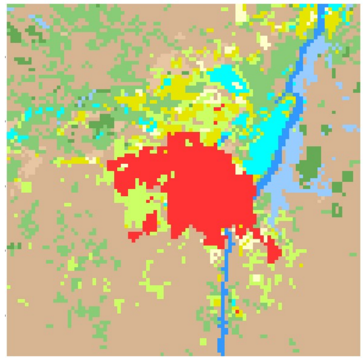
Harare



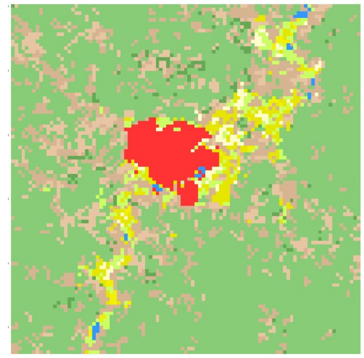
Johannesburg



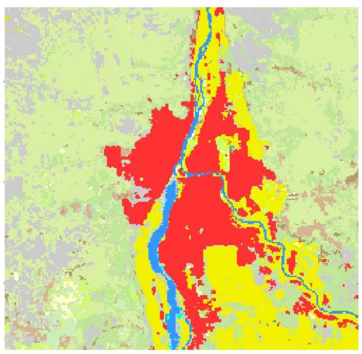
Juba



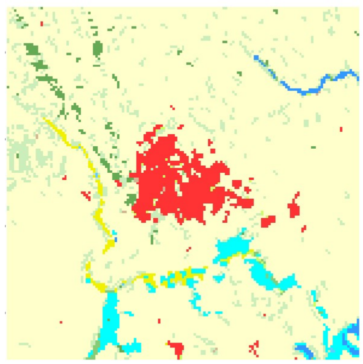
Kankan



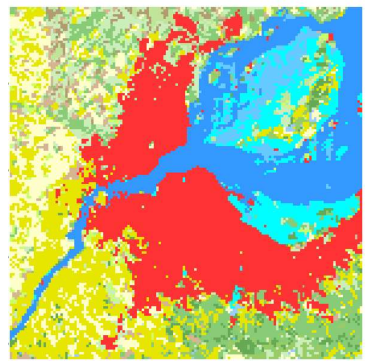
Khartoum



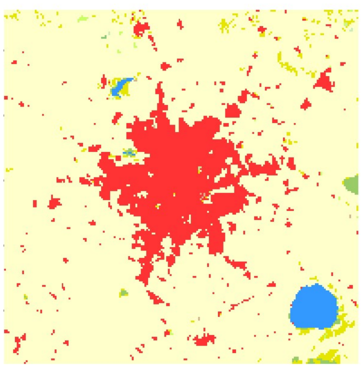
Kigali



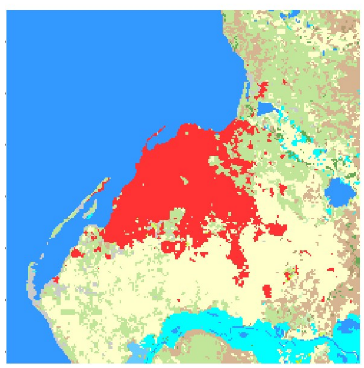
Kinshasa-Brazzaville



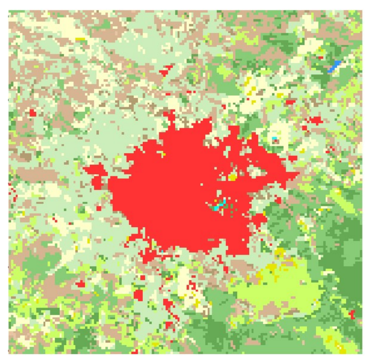
Kumasi



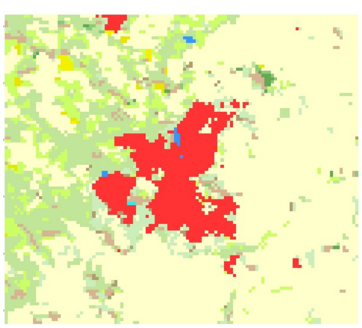
Luanda



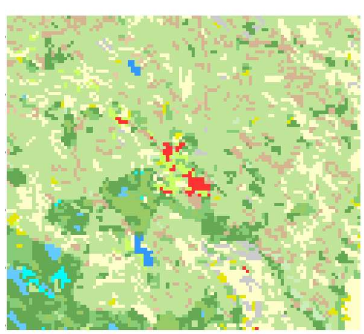
Lusaka



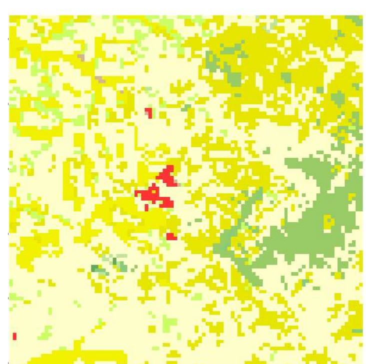
Maseru

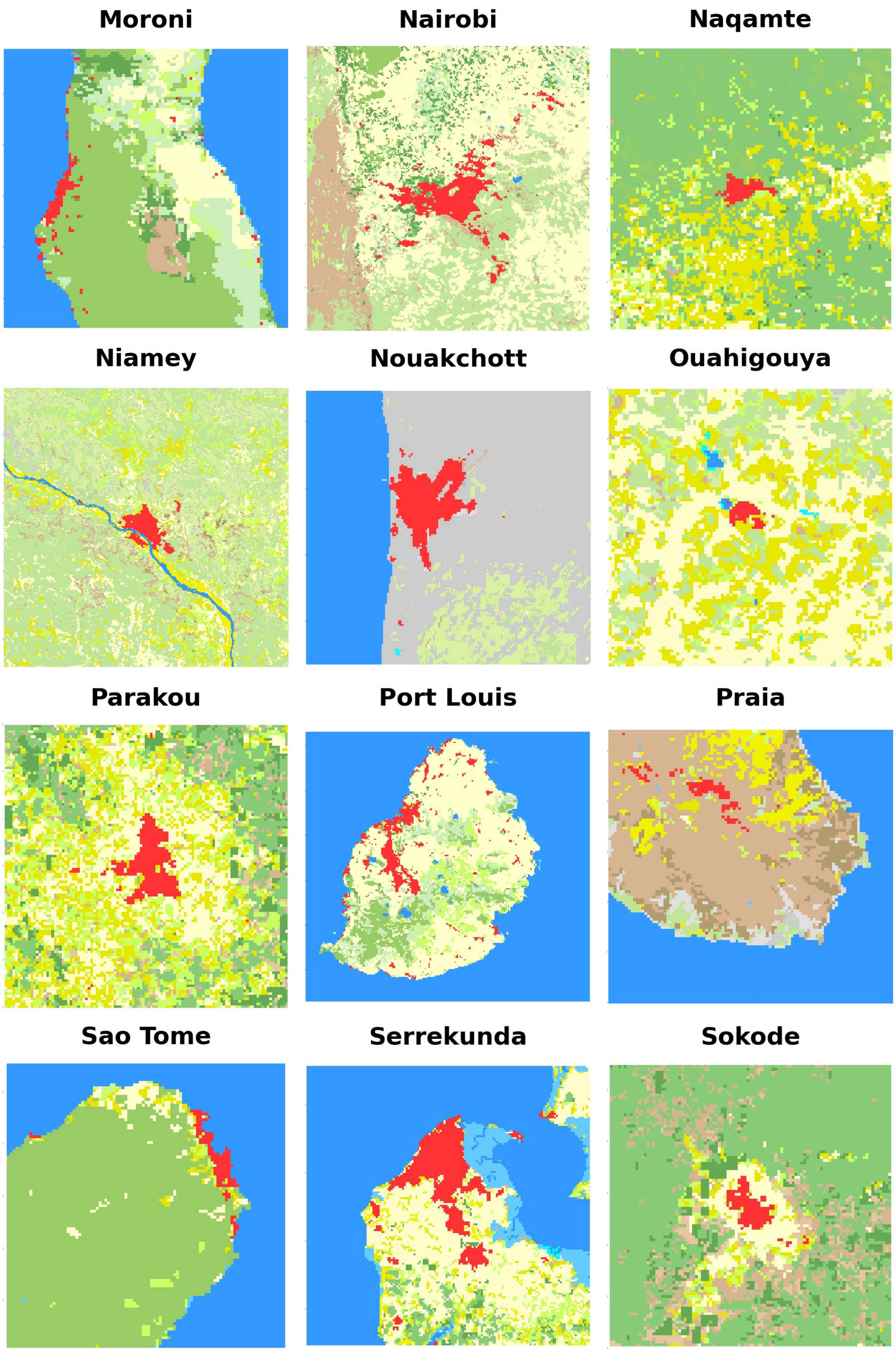


Mbabane

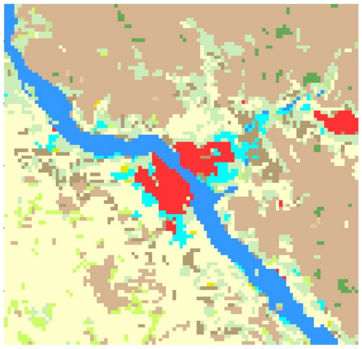


Mbale City

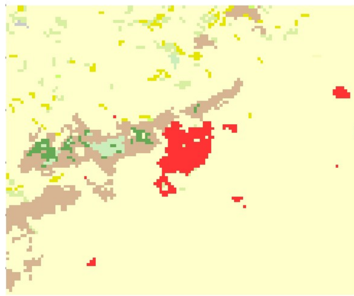



Moroni
Nairobi
Naqamte
Niamey
Nouakchott
Ouahigouya
Parakou
Port Louis
Praia
Sao Tome
Serrekunda
Sokode

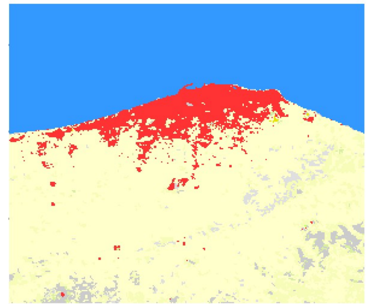
Tete



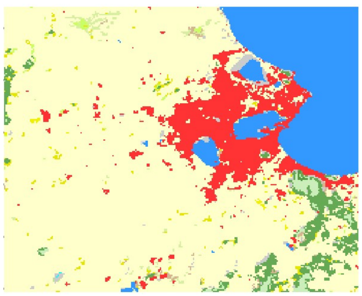
Tiaret



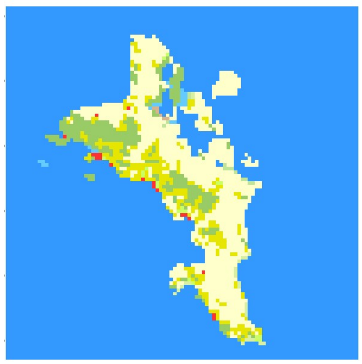
Tripoli



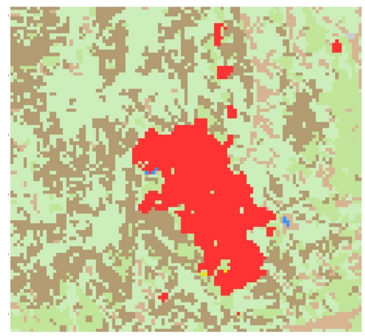
Tunis



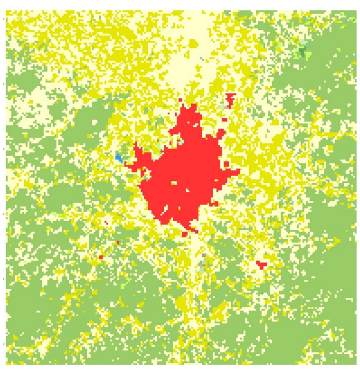
Victoria



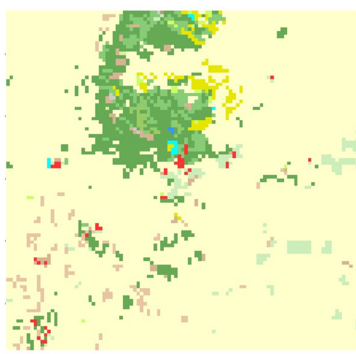
Windhoek



Yaounde



Zomba



ANNEX 2: INDICATORS GENERATED BY URBCLIM

Below we present a table containing all UrbClim indicators, including units, descriptions, and formulas.

Notation used in the table:

- $T(t)$: hourly 2-m air temperature in Kelvin.
- $W(t)$: hourly WBGT (wet-bulb globe temperature) in °C.
- Days are indexed by d, months by m.
- N_d: number of days in year. N_t: number of timesteps.
- $q90_{day}$: 90th percentile of all $T_{\max}(d)$ over the base period.
- $q90_{night}$: 90th percentile of all $T_{\min}(d)$ over the base period.

Indicator	Units	Description	Formula
$T2M_{mean}$	K	Mean annual 2-m air temperature	$\frac{1}{N_t}\sum_t T(t)$
$T2M_{daily\_mean\_max}$	K	Mean of daily maximum 2-m temperature	$\frac{1}{N_d}\sum_d T_{\max}(d)$
$T2M_{daily\_mean\_min}$	K	Mean of daily minimum 2-m temperature	$\frac{1}{N_d}\sum_d T_{\min}(d)$
UHI_{day}	K	Mean daytime UHI: daily max temperature, corrected for topography, relative to rural grassland	Apply lapse-rate correction: $T'(x) = 2M_{daily\_mean\_max} + h(x) \cdot 0.0065$. $UHI_{day(x)}$ = $T'(x) - T_{\min}$ with $T_{\min} = \min_{x \in \text{grassland}} T'(x)$.
UHI_{night}	K	Mean nighttime UHI: daily min temperature, corrected for topography, relative to rural grassland	Same as UHI_{day} but using $T2M\_daily\_mean\_min$
$T2M_{dayover25}$	# of days	Number of days with daily max T ≥ 25 °C	$\sum_d (T_{max}(d) > 273.15 + 25)$
$T2M_{dayover30}$	# of days	Number of days with daily max T ≥ 30 °C	$\sum_d (T_{max}(d) > 273.15 + 30)$
$T2M_{dayover35}$	# of days	Number of days with daily max T ≥ 35 °C	$\sum_d (T_{max}(d) > 273.15 + 35)$
$T2M_{nightover20}$	# of nights	Number of nights with daily max T ≥ 20 °C	$\sum_d (T_{min}(d) > 273.15 + 20)$
$T2M_{nightover25}$	# of nights	Number of nights with daily max T ≥ 25 °C	$\sum_d (T_{min}(d) > 273.15 + 25)$

$T2M_{nightover28}$	# of nights	Number of nights with daily max T ≥ 28 °C	$\sum_{d}(T_{min}(d) > 273.15 + 28)$
$Heatwave_{days}$	# of days	Total number of heatwave days in the year, where a heatwave is ≥ 3 consecutive hot days	Hot day: $H(d) = \left(T_{\max}(d) > q90_{\text{day}} \wedge T_{\min}(d) > q90_{\text{night}}\right)$. We Identify consecutive sequences of hot days of length $L \geq 3$. $Heatwave_{days}$ = sum of lengths of all such sequences
$MTWM$	K	Mean Temperature of the Warmest Month	$max_m(m)\ \overline{T_{max}}$ with $\overline{T_{max}} = mean_{d\in m}\mathrm{T}_{\max}(d)$
$MTCM$	K	Mean Temperature of the Coldest Month	$min_m(m)\ \overline{T_{min}}$ with $\overline{T_{min}} = mean_{d\in m}\mathrm{T}_{\min}(d)$
$Cooling\ degree\ hours$	degree-hours above 25 °C	Cooling degree hours relative to 25 °C, used as a proxy for cooling energy demand	$\sum_t \Delta T(t)$ with $\Delta T(t) = \max\left(0, T(t) - (273.15 + 25)\right)$
$WBGT_{dayover25}$	# of days	Number of days with daily max WBGT ≥ 25 °C	$\sum_{d}(W_{max}(d) > 25)$
$WBGT_{dayover28}$	# of days	Number of days with daily max WBGT ≥ 28 °C	$\sum_{d}(W_{max}(d) > 28)$
$WBGT_{dayover295}$	# of days	Number of days with daily max WBGT ≥ 29.5 °C	$\sum_{d}(W_{max}(d) > 29.5)$
$WBGT_{dayover31}$	# of days	Number of days with daily max WBGT ≥ 31 °C	$\sum_{d}(W_{max}(d) > 31)$
$WBGT_{nightover25}$	# of nights	Number of nights with daily min WBGT ≥ 25 °C	$\sum_{d}(W_{min}(d) > 25)$
$WBGT_{nightover28}$	# of nights	Number of nights with daily min WBGT ≥ 28 °C	$\sum_{d}(W_{min}(d) > 28)$
$WBGT_{hourover25}$	# of hours	Number of hours with WBGT ≥ 25 °C	$\sum_{d}(W_{max}(h) > 25)$
$WBGT_{hourover28}$	# of hours	Number of hours with WBGT ≥ 28 °C	$\sum_{d}(W_{max}(h) > 28)$
$WBGT_{hourover295}$	# of hours	Number of hours with WBGT ≥ 29.5 °C	$\sum_{d}(W_{max}(h) > 29.5)$
$WBGT_{hourover31}$	# of hours	Number of hours with WBGT ≥ 31 °C	$\sum_{d}(W_{max}(h) > 31)$
LWH_{light}	# of hours lost / year	Lost working hours for light manual work, derived from WBGT as a	$\sum_t f(t)$ with $f(t) = 0.5W(t) - 15.5$

		loss per hour summed over the year	• If $W(t) < 31,\ f(t) = 0.3$ • If $W(t) > 33,\ f(t) = 1.4$
LWH_{mod}	# of hours lost / year	Lost working hours for moderate manual work, derived from WBGT as a loss per hour summed over the year	$\sum_t f(t)$ with $f(t) = 0.22W(t) - 6.2$ • If $W(t) < 28.3,\ f(t) = 0.3$ • If $W(t) > 32.8,\ f(t) = 1.4$
LWH_{int}	# of hours lost / year	Lost working hours for intensive manual work, derived from WBGT as a loss per hour summed over the year	$\sum_t f(t)$ with $f(t) = 0.17W(t) - 4.4$ • If $W(t) < 26.6,\ f(t) = 0.3$ • If $W(t) > 32.6,\ f(t) = 1.4$
$WBGT_{dayly\_mean\_max}$	°C	Mean of daily maximum WBGT over summer	$\text{mean}_d W_{max}^{summer}(d)$

vision on technology
for a better world

vito