Documentation for the Africa Climate Mobility Initiative (ACMI): Internal Migration Projections

January 2025

CUNY Institute for Demographic Research (CIDR), City University of New York
Center for Integrated Earth System Information (CIESIN), Columbia University
Global Centre for Climate Mobility (GCCM)

Abstract

The Africa Climate Mobility Initiative (ACMI) Internal Migration Projections project internal migration flows at 5-year intervals from 2020 to 2050 for a combination of 2 sets of Shared Socioeconomic Pathways (SSPs) scenarios and 3 sets of Representative Concentration Pathways (RCPs) scenarios. The unit of analysis is 7.5 arc-minutes, covering the entire African continent (approximately 15 km at the equator). These grid cells are drawn using a vector representation, and not a tessellation of grid cells (rasters). While the calibration and projection took place at a grid cell resolution of 2.5 arc-minutes, these data have been aggregated to a 7.5 arc-minute resolution. These data underpin the African Shifts Report, produced by the ACMI and enabled by the Global Centre for Climate Mobility (GCCM). The ACMI was launched in September 2021 as a joint initiative of the GCCM, the African Union Commission (AUC), the United Nations System (UN), and the World Bank.

Data set citation: CUNY Institute for Demographic Research (CIDR), City University of New York, Center for Integrated Earth System Information (CIESIN), Columbia University, and Global Centre for Climate Mobility (GCCM). 2025. Africa Climate Mobility Initiative (ACMI): Internal Migration Projections. Palisades, New York: NASA Socioeconomic Data and Applications Center. https://doi.org/10.7927/5tv3-ff20. Accessed DAY MONTH YEAR.

Suggested citation for this document: CUNY Institute for Demographic Research (CIDR), City University of New York, Center for Integrated Earth System Information (CIESIN), Columbia University, and Global Centre for Climate Mobility (GCCM). 2025. Documentation for the Africa Climate Mobility Initiative (ACMI): Internal Migration Projections. Palisades, New York: NASA Socioeconomic Data and Applications Center. https://doi.org/10.7927/0z6j-px84. Accessed DAY MONTH YEAR.

We appreciate feedback regarding this data set, such as suggestions, discovery of errors, difficulties in using the data, and format preferences.
Please contact:
NASA Socioeconomic Data and Applications Center (SEDAC)
Center for Integrated Earth System Information (CIESIN)
Columbia University

Phone: 1 (845) 365-8920
Email: ciesin.info@ciesin.columbia.edu

Contents

I. Introduction

This documentation describes the data and methodology behind the projections of climate change-induced internal migration projections for 2020 to 2050 at five-year intervals. The data are released as grid cell-based vector data with a resolution of 7.5 arc-minutes. The Africa Climate Mobility Model was developed from 2020 to 2021 with contributions from the ACMI partners, including the Global Centre for Climate Mobility (GCCM), the African Union Commission (AUC), the United Nations Development Fund (UNDP), the UN Framework Convention on Climate Change (UNFCCC), the International Organization for Migration (IOM), the World Bank, and the Technical Advisory Group (TAG) of experts of ACMI. This modeling work builds on pioneering methods developed for the World Bank's Groundswell series of reports (Rigaud et al., 2018, Clement et al., 2021, Rigaud et al., 2021a, and Rigaud et al., 2021b), with some important improvements.

II. Data and Methodology

In this section, the methodology behind the calibration and projections of climate-induced internal migration to 2050 at five-year intervals is described. In large part, the text in this section is based on Appendix A of the following:

Amakrane, K., S. Rosengaertner, N. P. Simpson, A. de Sherbinin, J. Linekar, C. Horwood, B. Jones, F. Cottier, S. Adamo, B. Mills, G. Yetman, T. Chai-Onn, J. Squires, J. Schewe, B. Frouws, and R. Forin. 2023. African Shifts: The Africa Climate Mobility Report, Addressing Climate-Forced Migration & Displacement. New York: Global Centre for Climate Mobility. https://africa.climatemobility.org/report.

Input data

The Africa Climate Mobility Model projections are informed by combinations of development and emissions scenarios. The scenarios are discussed, and then the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) model data used as modeling inputs, followed by other inputs are described.

II.1.1 The Development Scenarios

The development scenarios informing the climate mobility model are based on the Shared Socioeconomic Pathways (SSPs), a framework for describing socioeconomic and demographic developments in Africa and globally. Two contrasting SSPs were chosen. The first is a "sustainability" scenario (SSP1) that is characterized by low population growth, high urbanization, medium Gross Domestic Product (GDP), and high education across Africa. Under SSP1, rapid economic growth in low-income countries reduces the number of people below the poverty line. The world is characterized by an open, globalized economy, with relatively rapid technological change directed toward environmentally friendly processes, including clean energy technologies and yield-enhancing technologies for land. This is an optimistic development pathway for the African continent, and results in a total continental population in 2050 of 1.75 billion people (up from 1 billion in 2010) that is relatively concentrated in urban areas.

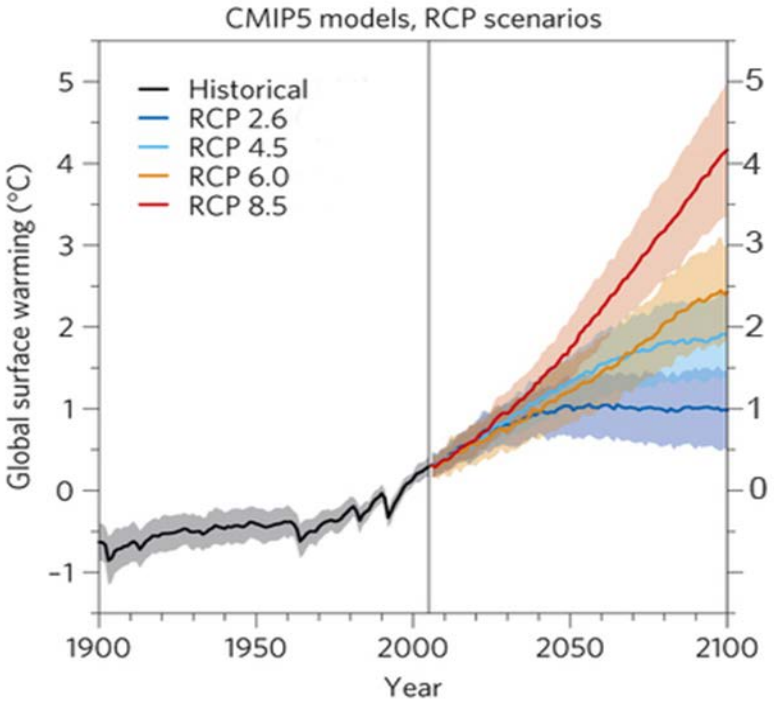
The second scenario is the "regional rivalry" scenario (SSP3), which is characterized by high population growth and low urbanization, as well as low GDP and education across much of sub-Saharan Africa. This is a world failing to achieve global development goals, and with little progress in reducing resource intensity, fossil fuel dependency, or addressing local environmental concerns such as air pollution. Inequality remains high both across and within countries, and economies are relatively isolated, leaving large, poor populations in developing regions highly vulnerable to climate change with limited adaptive capacity. The world has de-globalized, and international trade, including energy and agricultural trade, is severely restricted. By contrast with the low income countries of Africa, middle income countries (South Africa and northern Africa) are characterized by low population growth rates, high urbanization, moderate GDP, and low education levels. For this scenario, Africa's population grows to 2.3 billion people by 2050 (500 million more than under SSP1) and remains largely rural.

II.1.2 The Warming Scenarios

Turning to emissions or warming scenarios, the magnitude of future global warming is framed by the Representative Concentration Pathways (RCPs) (van Vuuren et al., 2014). RCPs are trajectories of greenhouse gas concentrations resulting from human activity corresponding to a specific level of radiative forcing in 2100. For the ACMI modeling work, two RCPs were chosen, a lower greenhouse gas concentration of RCP2.6 and the higher greenhouse gas concentration of RCP6.0. These imply futures where radiative forcing of 2.6 and 6.0 watts/meters^2, respectively, are achieved by the end of the century. From a baseline in the year 1990, the additional warming implied by these RCPs ranges from a low of 0.5°C (RCP2.6) to a high of 2.0°C (RCP6.0) by 2050, with far more warming anticipated (about 2.5°C on average) by 2100 under RCP6.0 (Figure 1).

RCPs do not rely on a fixed set of scenario-specific assumptions on economic development, technological change, or population growth. Many different socioeconomic futures or pathways may lead to the same level of radiative forcing. This framework allows researchers to consider alternative policy decisions with combinations of social, economic, and technological change. A future with high population growth but rapid development of clean energy technology may achieve the same level of radiative forcing as a world characterized by low population growth but continued reliance on fossil fuels. This framework allows researchers to specify certain levels of temperature change and then explore alternative policy options to achieve greenhouse gas concentration levels consistent with the goal.

Figure 1: Projected global average surface temperature change by RCP



Source: CMIP5 models, https://cimss.ssec.wisc.edu/climatechange/modeling

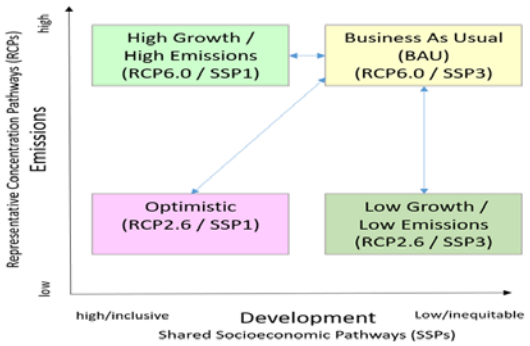
Some may question the choice of scenarios. RCP2.6 is considered unrealistic by many in a world that is poised to hurtle past the Paris Agreement target of limiting additional warming to 1.5°C, and where nationally determined contributions are largely insufficient to reach that target (Rogelj et al., 2016). This RCP serves mainly as a contrast to the higher emission pathway, and serves also to demonstrate that even at this lower level of emissions, the consequences for human mobility may be equal to or higher than RCP6.0. As for RCP6.0, while RCP8.5 was chosen as a high-end scenario for past work (e.g. Rigaud et al., 2018), partly because it was often portrayed as a "Business As Usual" (BAU) scenario, in reality it reflects the very high-end of the BAU emissions pathway, and is considered by some to be implausible (Hausfather and Peters, 2020). Furthermore, it could not be used for this work because (a) not all ISIMIP crop models have been run under RCP8.5, and (b) it is only compatible with SSP5 (a world characterized by rapid conventional development that leads to an energy system dominated by fossil fuels), according to the current literature (Tebaldi et al., 2021).

II.1.3 The four climate mobility modeling scenarios

The combination of SSPs and RCPs create four plausible future internal climate mobility scenarios (Figure 2):

- **BAU** (SSP3 and RCP6.0), in which low-income countries are characterized by moderate population growth, low rates of urbanization, low GDP growth, and low education levels. Urban growth is poorly planned, and high emissions drive greater climate impacts. This scenario poses high barriers to adaptation because of the slow pace of development and isolation of regional economies.
- **Low Growth/Low Emissions** (SSP3 and RCP2.6), which reduces climate impacts, but holds the development scenario consistent with the BAU scenario.
- **High Growth/High Emissions** (SSP1 and RCP6.0), which holds emissions where they are in the BAU scenario but provides a development scenario that is more optimistic, and the potential for adaptation is higher than under SSP3. Population growth is lower than in SSP3 for low-income countries, and urbanization is more rapid, resulting in more concentrated populations.
- **Optimistic** (SSP1 and RCP2.6), which reduces climate impacts and provides a development scenario that is more optimistic.

Figure 2: Africa Climate Mobility Model scenarios



II.1.4 Climate Impacts Data

Climate models based on the two warming scenarios (RCP2.6 and RCP6.0) drive the indicators of water, agricultural, and ecosystem sector change as well as flood risk provided by the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP), which are incorporated in projections of future population distributions. ISIMIP is a climate-impact modeling initiative aimed at contributing to a quantitative and cross-sectoral synthesis of the differential impacts of climate change, including uncertainties. ISIMIP has compiled a database of state-of-the-art computer model simulations of biophysical climate impacts. It offers a framework for consistently projecting the impacts of climate change across affected sectors and spatial scales.

The ACMI modeling uses outputs of the ISIMIP2a modeling work, which covers the historical period from 1970-2010, and the ISIMIP2b modeling effort, which has projections for 2010-2100 (Frieler et al., 2017). Under the 2b modeling effort, the future sectoral impact models are driven by a range of General Circulation Models (GCMs). The 2b modeling effort has the advantage over the prior ISIMIP Fast Track in that the models are bias-corrected, meaning they better capture historical means and variability in temperature and precipitation. This project uses two GCMs that provide a good spread for the temperature and precipitation parameters of interest: the Hadley Centre Global Environment Model version 2-Earth System (HadGEM2-ES) climate model developed in the United Kingdom, and the Princeton University Geophysical Fluid Dynamics Laboratory Earth Systems Models (GFDL-ESM2M) produced in the United States. For the rationale behind our model selection, see Section II.1.6.

The ISIMIP2b collection of sectoral models includes a range of systems and sectors, such as health, coastal infrastructure, forests, and other ecosystems. The focus of this study is on crop production, water availability, ecosystem impacts, and riverine floods. The global crop, water, and ecosystem simulations—at a relatively coarse spatial scale (0.5 degrees or roughly 55 km at the equator)—are an advance over purely climate model-based indicators of rainfall and temperature, because they represent actual resources of relevance to development. The flood impact model is at 500 meter resolution, and is based on projected flood depth. In this work, flood impact is used as a mask to reduce the potential of affected grid cells, and therefore the likelihood of future migration into areas that are projected to suffer increasing flood risks.

These climate impacts are selected because the literature shows that water scarcity, declining crop yields, declines in pasturage, and flood impacts are among the major potential climate impacts facing low-income countries, and that these impacts will also be very important drivers of migration and displacement. Finally, sea level rise is included as a spatial mask that does not permit people to live in areas likely to experience inundation. Each of these input layers is described in greater detail below.

The models are better at assessing long-term trends rather than individual extreme events such as drought, extreme rainfall, and cyclones. As devastating as they may be for rural

livelihoods, short-duration, fast-onset events are not directly included. That said, the proposed five-year time step does capture the combined effects of repeated extreme events better than the original ten-year time step used in Groundswell (Rigaud et al., 2018, Clement et al., 2021), where extremes in either direction are more likely to counterbalance each other over the course of a decade. To further assess the impact of extremes, projected flood impacts are included (described below).

Data on water availability and crop production are integrated into the Africa Climate Mobility Model using the following approach. The water sector model outputs represent river discharge, measured in cubic meters per second in daily/monthly time increments. The crop sector model outputs represent crop yield in tons per hectare on an annual time step at a 0.5° x 0.5° grid cell resolution. Crops include maize, wheat, rice, and soybeans for the GIS-based Environmental Policy Integrated Climate (GEPIC) model, and those crops plus cassava/tropical roots, groundnut, millet/tropical cereals, field pea/pulses, rapeseed, sugarcane, sugar beet/temperate roots, sunflower for the Lund-Potsdam-Jena managed Land (LPJmL) model. For regions with multiple cropping cycles, yield reflects only the major crop production period. In conformity with the Groundswell work, the data are converted to five-year average water availability and crop production (in tons) per grid cell.

Climate change impact is measured by calculating at each 0.5 x 0.5 degree grid cell an index of 5-yearly deviations from a baseline period, for the following variables: annual mean discharge (water), annual crop yield (agriculture), annual mean total Net Primary Productivity (NPP, biomes/ecosystems). Note that crop yield is the sum over all considered crops, weighted by estimated growing areas around the year 2000 (Portmann, Siebert, and Döll, 2010). Let t_0 be the baseline period (1970-2010), t a 5-year time period (1971-1975, 1976-1980, etc.), and $x(t)$ the average of one of the above variables over t. Then the reported index $D_x(t)$ is calculated as:

$$D_x(t) = \frac{x(t) - x(t_0)}{x(t_0)}$$

(Equation 1)

That is, D is a dimensionless number ranging from -1 to +∞, where 0 means no change compared to the baseline. A value of -0.5 means a reduction by 50% compared to the baseline, while a value of +1 means a doubling (increase by 100%) compared to the baseline.

The ISIMIP crop and water model outputs are based on different combinations of climate, crop, and water models. Applying the combinations—two global climate models driven by two different emissions scenarios, which in turn drives two sets of sectoral impact models (described below)—provides a range of plausible population projections. It also gives a sense of the level of agreement across scenarios. Because the population modeling process is time consuming and computationally intensive, a reduced set of ISIMIP inputs was needed. The modeling employs the HadGEM2-ES and GFDL-

ESM2M global climate models, which drive combinations of the two water and crop models: the LPJmL and GEPIC crop models and the WaterGAP2 and MATSIRO water models (Table 1). Note that because the crop models only cover parts of Africa where cropping is prevalent, the climate data is gap-filled with two models of Net Primary Productivity (NPP)—Organising Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) and LPJmL—that are intended to represent changes in pasture-land productivity.

The crop and water models are selected based on several criteria, including model performance over the historical period, diversity of model structure, diversity of signals of future change, and availability of both observationally driven historical (ISIMIP2a) and global climate model-driven historical and future (ISIMIP2b) simulations. Table 1 presents the combinations of crop and water models that will be used. Section II.1.6 provides detailed information on model selection.

Table 1: Matrix of global climate models and crop and water model combinations

			ISIMIP Crop Models			
	Global Climate Models (CMIP5)		**HadGEM2-ES**		**GFDL-ESM2M**	
		ISIMIP Models	GEPIC*	LPJmL**	GEPIC*	LPJmL**
ISIMIP Water Models	**HadGEM2-ES**	WaterGAP2	Model 1			
		MATSIRO		Model 2		
	GFDL-ESM2M	WaterGAP2			Model 3	
		MATSIRO				Model 4
			* GEPIC crop model coverage is gap-filled with the ORCHIDEE NPP model			
			** LPJmL crop model coverage is gap-filled with the LPJmL NPP model			

The ISIMIP flood risk projections use an ensemble of the H08, LPJmL, Max Planck Institute for Meteorology's Hydrology Model (MPI-HM), ORCHIDEE, PCRaster GLOBal Water Balance (PCR-GLOBWB) and WaterGAP2 models under RCP6.0. According to Zhao (personal communication), the differences between flood risk for RCP2.6 and RCP6.0 are not that great, so projections only under RCP6.0 were chosen. A maximum reduction of 20% was chosen, meaning all things being equal, a flooded grid cell would be 20% less likely to attract new population than a neighboring unflooded grid cell.

Lastly, the model incorporates projected sea level rise impacts. Parts of the African continent’s coastline, particularly the Niger and Nile deltas, are particularly vulnerable to sea level rise impacts (Hinkel et al., 2012). The analysis also considers Sea Level Rise (SLR) projections from the IPCC Fifth Assessment Report, augmented by an increment for storm surges. According to Dasgupta et al., (2007, 6), “Even a small increase in sea level can significantly magnify the impact of storm surges, which occur regularly and with devastating consequences in some coastal areas.” A comprehensive assessment of the likely levels of storm surge for all the coastal areas will be beyond the scope of this project. In any case, according to IPCC’s Fifth Assessment Report, Working Group II, Chapter 5, the habitability of coastal areas not immediately within the Low Elevation

Coastal Zone (LECZ) will be negatively impacted through increased coastal flooding, erosion, and saltwater intrusion into estuaries and deltas, as well as increases to the water table. For simplicity in this work, it is assumed that SLR + storm surge under RCP2.6 will amount to a 1 meter inundation in LECZ by 2050, and that under RCP6.0, it will amount to a 2 meter inundation. These levels of inundation are progressively diminished working backwards from 2050 to 2010. The effect of SLR + storm surge is to remove land in each 2.5 arc-minute grid cell from circulation, which results in a reduction in population potential in those grid cells in the 1 or 2 meter LECZ. The source data set is NASA's Shuttle Radar Topography Mission (SRTM).

II.1.5 Other Data Inputs

A full set of data inputs are in Table 2, and a few more important data sets used in the modeling are described here. The Africa Climate Mobility Model innovates in some other important respects. A major advance was the incorporation of modeled population grids for the calibration and the baseline for projections. Past work under Groundswell and Groundswell Africa used an un-modeled population surface, Gridded Population of the World, Version 4 (GPWv4): Population Count Revision 11 (CIESIN, 2018), which takes as its basis census inputs that are provided by countries for widely varying geographies. It uses a uniform distribution or proportional allocation that does not make use of any other geographic data in order to spatially disaggregate the census population. In the case of Africa, some countries only have provincial level census inputs (admin1), whereas others have much higher resolution inputs (admin4 or admin5). Coarser resolution inputs mean that GPWv4 tends to overestimate rural populations because populations are allocated (spread out) over large census units. To rectify this situation, a modeled population surface was used, Global Human Settlement Layer: Population (GHS-POP), available in time series for 1990, 2000, and 2015 (JRC and CIESIN, 2021). GHS-POP consists of census data from GPWv4, Revision 11 (GPWv4.11), spatially-allocated within census units based on the percent built-up areas from GHS-BUILT, a layer constructed from Landsat and later Sentinel satellite imagery. The native resolution of GHS-POP is 30 arc-seconds (or 1 km), but in order to reduce the potential for artifacts to affect the modeling work, the data were aggregated to several different resolutions, and 2.5 arc-minutes was chosen because it presented the best balance between higher resolution and lower amounts of errors.[1]

Owing to the prevalence of conflict on the African continent, and the potential for conflict to affect mobility through forced displacement, data on conflict incidence are also included in the model. Data from the Uppsala Conflict database were used, which has a longer and more consistent time series (1989-present) than the Armed Conflict Locations and Events Database (ACLED), which is 1998 and onward. The historical

[1] GHS-BUILT tends to have higher errors of commission, meaning it finds settlements in areas where there are actually no settlements or sparsely populated settlements owing to the spectral signature of certain kinds of land covers (rocky outcrops, lake beds, etc.), than errors of omission. Potential modeling uncertainties introduced with the use of GHS-POP are addressed in Box A2 on "Sources of uncertainty in modeling climate migration".

conflict data in the calibration process (two periods 1990-2000 and 2000-2010) was included, to assess the sensitivity of past changes in population distribution to conflict events. In the absence of data on projected future conflict locations, the 2000-2010 conflict surface in the gravity model through 2050 was included. While it is an unlikely assumption that future conflict will remain stationary—indeed, recent outbreaks in the Sahel are showing how volatile some regions are—there was no way to project future conflict spatially without heroic assumptions.

The model applies a mask to avoid rapid population growth in regions where population densities currently are below 1 person per square kilometer (e.g. in parts of the Saharan desert).

Additional possible data layers to include in the model were investigated. One was a surface of population displacement owing to disasters based on the Geocoded Disasters (GDIS) Dataset (Rosvold and Buhaug, 2021). The theory was that past disaster displacement may have repelled populations from areas frequently affected by meteorological disasters (flood and drought). In reality, however, this calibration work revealed that the displacement surface always produced large and positive coefficients, which would have resulted in a "pull" towards disaster displacement regions. This may be because flood disasters are a major type of disasters in the database and waterbodies (particularly river valleys) tend to attract populations over time, or it could be that the coarse resolution of many reported displacements resulted in geolocation errors. Ultimately, disaster displacement was not included in the model.

The remainder of the data sets included in the model listed in Table 2 are for the purposes of removing certain areas ("masking") from future settlement. Masks effectively set the population potential of a grid cell to zero (0), meaning that no migration will occur in those areas.

Table 2: Data inputs used in the Africa Climate Mobility Model

Product	Source Data	Resolution	Time Series	Time Step	Indicator / Purpose
Population Grids	GHS-POP	30 arc-seconds converted to 2.5 arc-minutes	1990, 2000, 2015		Calibration; population count
Urban Mask	GHS-SMOD	30 arc-seconds converted to 2.5 arc-minutes			Calibration; dummy variable
Water Availability	ISIMIP2b	0.5 degrees	1990-2050	5-year	Calibration; deviations from baseline
Crop Production	ISIMIP2b	0.5 degrees	1990-2050	5-year	Calibration; deviations from baseline
Net Primary Productivity	ISIMIP2b	0.5 degrees	1990-2050	5-year	Calibration; deviations from baseline

Conflict Deaths	Uppsala	Buffered points converted to 2.5 arc-minutes	2000-2015	annual	Calibration
Flood Hazard	ISIMIP2b	1 kilometer	1990-2040	5-year	Mask out flooding
Sea-Level Rise	SRTM	1 kilometer	2020-2050	5-year	Mask out coastal SLR
Water Bodies	Esri	vector	2021	n/a	Mask out perennial water bodies to future settlement
Protected Areas (PAs)	WDPA	vector	2021	n/a	Mask out PAs to future settlement, SSP3 includes IUCN categories 1-1a-2-3, SSP1 adds category 4
Slope					Mask - 25% maximum slope for settlement
Elevation					Mask - highest existing settlement

II.1.6 Rationale for Model Selection

Global climate models

Four Global Climate Models (GCMs) were considered for use in the ACMI modeling work: GFDL-ESM2M, HadGEM2- ES, IPSL-CM5A-LR: Institut Pierre Simon Laplace Coupled Model 5A Low Resolution (IPSL-CM5A-LR), and Model for Interdisciplinary Research on Climate version 5 (MIROC5). These models were selected for ISIMIP2b from the larger Coupled Model Intercomparison Project Phase 5 (CMIP5) model ensemble, and the data were bias-corrected to remove systematic deviations of the models' historical mean climate state from observations, while preserving simulated long-term trends (Frieler et al., 2017). The four models cover a range of climate sensitivities; IPSL-CM5A-LR and HadGEM2-ES being on the warm side (Equilibrium Climate Sensitivity (ECS) above 4°C), while GFDL-ESM2M and MIROC5 are on the cool side (ECS around 2.5°C; Schlund et al., 2020). For this work, two models were selected with a view to representing a range of possible climate responses in Africa; keeping in mind that such a small number of models cannot be representative of the spread found in larger model ensembles for many important climate variables.

Changes in regional temperature relative to global mean temperature are similar across all models. The projected drying trend in the two subtropical regions of the continent, the Mediterranean and southern Africa, is also relatively robust across CMIP5 models, and well-established in atmospheric science (IPCC AR5 WG1 chapter 12). The precipitation response in tropical West Africa and East Africa is more uncertain, yet very important for the societal impacts. For instance, some models project strongly increasing summer (wet season) rainfall across the central and eastern Sahel region, while others project only small changes there (Schewe and Levermann, 2017). Of the four ISIMIP2b models, MIROC5 and HadGEM2-ES both project a substantial increase in Sahel rainfall under

global warming, while IPSL-CM5A-LR projects little change, and GFDL-ESM2M even projects a decrease.

In East Africa, IPSL-CM5A-LR projects by far the strongest increase in precipitation, but has been shown to be an outlier compared to the rest of the CMIP5 ensemble, and likely overestimates a feedback between sea surface temperatures and cloud cover over the Indian Ocean (Rowell, 2019). MIROC 5 and GFDL-ESM2M both project increasing precipitation over East Africa, while HadGEM2-ES projects relatively stable precipitation (Frieler et al., 2017, Supplement). Thus, the HadGEM2-ES and GFDL-ESM2M models were chosen because they cover both high and low ECS (i.e. overall "intensity" of global warming), contrasting precipitation responses in both West Africa and East Africa.

ISIMIP Crop Models

A comprehensive analysis of global crop model responses to projected future global warming shows that the GIS-based Environmental Policy Integrated Climate (GEPIC) model is typically one of the more pessimistic models (i.e. predicting stronger declines in crop yield, on average), while LPJmL and the Python-based Environmental Policy Integrated Climate (PEPIC) model tend to fall near the center of the ensemble (Müller et al., 2021). A detailed benchmarking study, comparing historical model simulations with reported national crop yield data, indicates that LPJmL, GEPIC, and PEPIC have relatively low mean bias for maize yields in most African countries, while the Community Land Model (CLM-Crop) often has larger positive biases (Müller et al., 2017, Fig. S23). For wheat, GEPIC and PEPIC exhibit larger negative biases in some countries such as Zambia, Namibia, and Botswana, while LPJmL and CLM-Crop have positive biases in individual countries such as Egypt and Burundi (Müller et al., 2017, Fig. S24). In terms of year-to-year variability in yields, correlation between models and data varies and is rather low in many African countries for all of the models, however, this is at least partly due to high uncertainty in the reported yield data (Müller et al., 2017, Fig. 9). In the present work, annual yield data are not used directly, but aggregated over five-year periods, placing less importance on the timing of individual annual yields.

A global study of future risk of crop failure including LPJmL, GEPIC, and PEPIC, shows that the projected increases in population exposure to crop failure under global warming do not differ greatly between the models when forced with the HadGEM2-ES or MIROC5 climate models (Lange et al., 2020). In the simulations forced with the other two climate models, GEPIC tends to project larger increases in exposure to crop failure than PEPIC and LPJmL.

LPJmL is the only model that simulates a number of additional crop types, some of which are widely grown in Africa: cassava (representing tropical roots), groundnut, millet (tropical cereals), field pea (pulses), rapeseed, sugarcane, sugar beet (temperate roots), and sunflower. GEPIC and the other models only provide maize, wheat, rice, and soybean yields. For this reason, LPJmL and GEPIC will be used.

ISIMIP Water Models

A recent evaluation of several global hydrological models, including five from the available ISIMIP2b ensemble (H08, LPJmL, MATSIRO, PCR-GLOBWB, and WaterGAP2), found that WaterGAP2 is best at simulating mean annual runoff in almost all hydrobelts (hydro-geographical regions on Earth), and second-best in the northern subtropical hydrobelt (which in Africa includes the Niger river basin), showing relatively small deviations from observed streamflow at observational stations around the globe (Zaherpour et al., 2018). This is partly expected because of the extensive calibration applied to this model, and does not necessarily imply that projected future trends in the other models are less plausible. The MATSIRO model comes second place in this evaluation of mean annual runoff.

A separate study of historical changes in global water scarcity using those same five models indicates that the portion of people estimated to be affected by water scarcity is largest with MATSIRO and H08, and lowest with PCR-GLOBWB, while the estimates from LPJmL and WaterGAP2 fall in between (Veldkamp et al., 2017). Under future global warming, WaterGAP2 and JULES-W1 tend to project smaller changes in drought exposure, while PCR-GLOBWB, LPJmL, and H08 tend to project larger relative increases in drought exposure (Lange et al., 2020). However, these results depend to some extent on the underlying climate model. For instance, Community Land Model version 4.5 (CLM4.5) projects rather large increases with GFDL-ESM2M, but relatively small increases with HadGEM2-ES. MATSIRO was not analyzed in this study. For this reason, MATSIRO and WaterGAP2 will be used.

An additional note is warranted about the inclusion of human impacts on the water cycle such as damming and irrigation. During calibration of the water models, the past response of population distribution to changes in water availability is calculated based on simulations that did not include changes in Human Impacts (HI) on the water cycle, other than those related to greenhouse gas emissions (so called no societal or “nosoc” simulations). Such HI include the construction of dams and reservoirs, and the withdrawal of water for purposes such as irrigation. Changes in such impacts were excluded because they could confound the effect of climate change on water availability, which is the effect being pursued to identify. In other words, this work is looking to quantify the effect of climate change on human migration, not the effect of other anthropogenic interferences with the water cycle.

Accordingly, for the future projections, ISIMIP2b simulations are used that also exclude any potential changes in HI in the future, by keeping all HI fixed at 2005 levels (“2005soc”). For the historical period in ISIMIP2b (which is still driven with climate-model output and must not be confused with the observations-based ISIMIP2a simulations), only simulations accounting for changing HI (“histsoc”) are available. This is not a significant problem because the results of the population modeling are presented relative to the baseline year 2015, and thus any changes related to HI prior to 2015 do not have any imprint on the modeling results. (Note: the baseline against which deviations in water availability are being measured is defined as the average of 1971-2010, and thus

includes some HI-related variations, but this is not relevant for the population modeling results for the aforementioned reason).

ISIMIP ecosystem models
As discussed above, the ecosystem model, more properly understood as a model of Net Primary Productivity (NPP), is used to gap-fill areas that do not have agricultural model outputs owing to the absence of cropping. A global evaluation of carbon fluxes in the ISIMIP2a biome modes shows that the magnitude of historical Net Biome Productivity (NBP) simulated by JULES-W1, VEGAS, ORCHIDEE, and LPJmL falls well within the observationally constrained range, while the Dynamic Land Ecosystem Model (DLEM), Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS), Vegetation Integrative SImulator for Trace gases (VISIT), and CARbon Assimilation In the Biosphere (CARAIB) simulations fall partly or largely outside that range, and overestimate NBP (Chang et al., 2017). On the other hand, the observed trend in global NBP is most closely reproduced by CARAIB, ORCHIDEE, and LPJmL, with VEGAS and JULES-W1 most strongly underestimating the trend, and the other models falling in between. ORCHIDEE runs on a 1 x 1 degree grid, whereas LPJmL and most other models run on a 0.5 x 0.5 degree grid. Based on this evaluation, ORCHIDEE and LPJmL model outputs were used.

II.2 Methods

II.2.1 Introduction to the Model

Climate impacts on crop production, water availability, ecosystem productivity, and flood depth and extent have important impacts on the population potential of locations in the gravity model, as described in this section. The modeling work is based on a modified version of the National Center for Atmospheric Research (NCAR) - CUNY Institute for Demographic Research (CIDR) INCLUDE gravity model (Jones and O'Neill, 2016).

The original INCLUDE model is a gravity-based approach that downscales national population projections to subnational raster grids (Jones and O'Neill, 2013, 2016) as a function of geographic, socioeconomic, and demographic characteristics of the landscape and existing population distribution. Gravity-type approaches are commonly used in geographic models of spatial allocation and accessibility. They take advantage of spatial regularities in the relationship between population agglomeration and spatial patterns of population change. These relationships can then be characterized as a function of the variables known to correlate with spatial patterns of population change.[2]

[2] Data for the original SSP-only population projections, using a different baseline population and set of modeling approaches, the Global Population Projection Grids Based on Shared Socioeconomic Pathways (SSPs), 2010-2100 data set, is available for download via the NASA Socioeconomic Data and Applications Center (SEDAC) at https://doi.org/10.7927/H4RF5S0P. Projections produced using the INCLUDE model for the Groundswell report series, the Groundswell Spatial Population and Migration Projections at One-Eighth Degree According to SSPs and RCPs, 2010-2050 data set, are available at https://doi.org/10.7927/c5kq-fb78.

The INCLUDE model uses a modified form of population potential, a distance-weighted measure of the population taken at any point in space that represents the relative accessibility of that point (for example, higher values indicate a point more easily accessible by a larger number of people). Population potential can be interpreted as a measure of the influence that the population at one point in space exerts on another point. Summed over all points within an area, population potential represents an index of the relative influence that the population at a point within a region exerts on each point within that region, and can be considered an indicator of the potential for interaction between the population at a given point in space and all other populations (Rich, 1980). This potential will be higher at points closer to large populations; potential is thus also an indicator of the relative proximity of the existing population to each point within an area (Warntz and Wolff, 1971). Such metrics are often used as a proxy for attractiveness, under the assumption that agglomeration is indicative of the various socioeconomic, geographic, political, and physical characteristics that make a place attractive.

In the Africa Climate Mobility Model, the calculation of population potential is modified primarily by adding variables that describe local conditions, including climate impacts, and weighting the attractiveness of each location (grid cell) as a function of the historical relationship between these variables and observed population change. Figure 4 is a flowchart of the modeling steps; boxes in red show the addition of climate impacts (or results incorporating climate impacts), demographic characteristics, and conflict related fatalities. Population potential is, conceptually, a relative measure of agglomeration, indicating the degree to which amenities and services are available. In the original model, this value shifts over time as a function of population, assumptions regarding spatial development patterns (for example, sprawl vs. concentration), and certain geographic characteristics of the landscape. The choice of SSPs influences each of these factors. For example, in the model the agglomeration effect is enhanced or muted as a function of the characteristics discussed above that aid in differentiating between places, as well as the SSPs. For example, SSP1 results in higher concentration of population than SSP3.

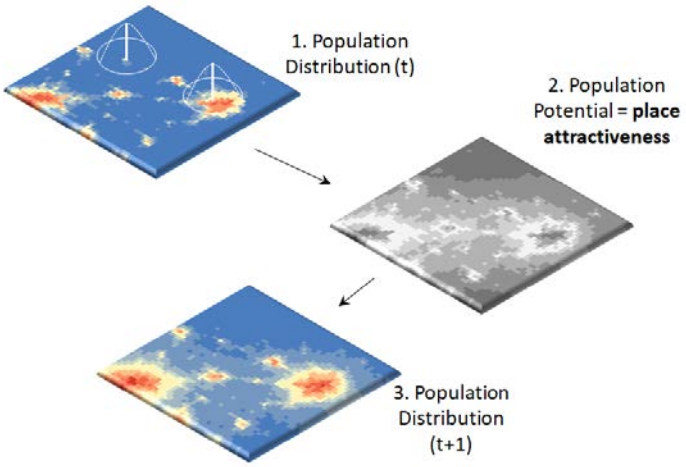
Beginning with the 2010 gridded urban/rural population distribution for each country,[3] the proposed modeling incorporates the influence of climate impacts on relative attractiveness in the following manner:

- Calculate an urban population potential surface (a distribution of values reflecting the relative attractiveness of each grid cell).

[3] Urban and rural population change need to be calculated separately because the factors that influence growth of urban and rural areas are distinct. Data on the evolution of population distributions show that historically urban and rural populations exhibit very different patterns of spatial population change (Jones and O'Neill, 2013). The former tend towards agglomeration over smaller geographic areas that can take several different forms (e.g. dispersion/concentration), while the latter occurs over larger geographic areas, varies across a wider range of patterns (including uniform and proportional) than urban populations, and is subject to periods of substantial population decline. Furthermore, in fitting the model to historical data, substantial variation in many of the parameters driving spatial population change is apparent. These two factors, taken together, suggest that modeling urban and rural populations as separate but interacting components of the total population is advantageous in comparison to considering the entire population as a single entity.

- Calculate a rural population potential surface.
- Allocate projected urban population change to grid cells proportionally based on their urban potentials.
- Allocate changes in the projected rural population to grid cells proportionally based on their rural potential.
- Because the allocation procedure can lead to some redefinition of population from rural to urban (e.g. rural population allocated to a grid cell with an entirely urban population is redefined as urban), this step entails redefining population as urban or rural as a function of density and contiguity of fully urban/rural grid cells to match projected national-level totals.

Figure 3: Gravity modeling approach



These steps are then repeated for each 5-year time interval. Figure 3 illustrates steps 3 and 4 for a hypothetical population distribution. Note that population potential surfaces, urban and rural, are continuous across all grid cells; each grid cell may thus contain urban and rural populations.

Based on the modified INCLUDE model, population potential (vi) is calculated as a parametrized negative exponential function:

$$v_i = l_i \left[\left(\sum_{j=1}^{m} P_j e^{-\beta d_{ij}} \right) + a_i P_i \right]$$

Spatial Mask · Distance Parameter · Distance · Population Parameter - Local Characteristics · Population

(Equation 2)

It is weighted by a spatial mask[4] (l) that prevents population from being allocated to areas that are protected from development or unsuitable for human habitation, including areas that will likely be affected by floods and sea level rise between 2010 and 2050. P_j is the population of all grid cells *j* within distance *m* of grid cell *i*, and P_i is the population of grid cell *i*; *d* is the distance between two grid cells. The distance parameter (β) is estimated from observed patterns of historical population change (for the urban and rural populations, separately). The β parameter is indicative of the shape of the distance-density gradient describing the broad pattern of the population distribution (for example, sprawl versus concentration), typically a function of the cost of travel (with lower costs leading to residential patterns more indicative of sprawl). The population parameter α_i is a weight on the population of grid cell *i* that reflects the relative attractiveness of each grid cell *i* as a function of the socioeconomic, demographic, political, and climate-related characteristics that make a place more or less attractive. Both β and *a* are calibrated from historical data, however the former is a universal parameter while the latter is grid cell-specific.

The SSPs include no climate impacts on aggregate total population, urbanization, or the subnational spatial distribution of the population. The INCLUDE approach is modified by incorporating additional spatial data including the ISIMIP sectoral impacts and the distribution of conflict-related and disaster-related fatalities, all of which are likely to affect population outcomes. The index α_i is a weight on population potential that is calibrated to represent the influence of these factors on the agglomeration effect that drives changes in the spatial distribution of the population. All of the data are incorporated into the model as 15 kilometer gridded spatial layers. The ISIMIP data represent 5-year deviations from long-term baseline conditions, and conflict-related and disaster-related fatalities are interpolated from point data. The value α_i is calculated as a function of these indicators. Numerically, it represents an adjustment to the relative attractiveness of (or aversion to) specific locations (grid cells), reflecting current water availability, crop yields, and ecosystem services relative to "normal" conditions, and the likelihood of dangerous conflict and disaster. As previously noted, the model is calibrated over two time periods (1990-2000 and 2000-2015) of observed population change relative to observed climatic and demographic conditions as well as safety (e.g. conflict-related fatalities).

II.2.2 Model Calibration

The value α_i (from Equation 2) is calculated as a function of the climate impact indicators, and represents an adjustment to the relative attractiveness of (or aversion to) specific locations (grid-cells) reflecting projected water availability, crop yield, and net primary production relative to "normal" conditions, in addition to flood risk, sex ratio, median age, and risk of conflict. In order to carry out the procedure, an estimate of the β

[4] Spatial masks are used in geospatial processing to exclude areas from consideration. The effect is that the algorithm is not applied in these areas. Examples in this instance would include protected areas or places where the terrain is too rugged to inhabit.

parameter for the urban and rural populations is necessary, and (Equation 2) must be fully calibrated. Two separate procedures are employed and carried out both for the urban and rural population distributions separately. As mentioned in Section II.2.1, urban and rural populations interact in the model, but changes in both are projected separately at the grid-cell level in the same manner. Here the procedure is described once, and, unless otherwise noted, the process is redundant for urban/rural components.

The β parameter is designed to capture broad-scale patterns of change found in the distance-density gradient, which is represented by the shape/slope of the distance decay function (parabolas) depicted in Equation 2. The negative exponential function described by Equation 2 is very similar to Clark's (1951) negative exponential function which has been shown to accurately capture observed density gradients throughout the world (Bertaud and Malpezzi, 2003). To estimate β, the model in Equation 2 is fitted to the 1990-2000 urban and rural population change from GHS-POP and to the 2000-2010 urban and rural population change data from GHS-POP, and then the value β is computed that minimizes the sum of absolute deviations:

$$S(\beta) = \sum_{i=1}^{n} \left| P_{i,t}^{mod} - P_{i,t}^{obs} \right|$$

(Equation 3)

where $P_{i,t}^{mod}$ *and* $P_{i,t}^{obs}$ are the modeled and observed populations in cell *i*, and *S* is the sum of absolute error across all grid cells. The model is fitted for two decadal time steps (1990-2000 and 2000-2010) and takes the average of the β estimates.

In this modified version of the population potential model, the index is a grid cell-specific metric that weights the relative attractiveness of a location (population potential) as a function of environmental and/or socioeconomic conditions. The modeling approach requires to empirically estimate the association between the relative attractiveness and the different sectoral impacts, flood risk, and conflict. The relative attractiveness of each grid cell is then assumed to be associated with population change. When β is estimated from historical data (e.g. observed change between 2000 and 2010), a predicted population surface is produced that reflects the optimized value of β, such that absolute error is minimized. Figure 5 includes a cross-section (one-dimension) of grid cells illustrating observed and predicted population for 10 grid cells. Each grid cell contains an error term that reflects the error in the population change projected for each grid cell over a 10-year time step. It is hypothesized that this error can at least partially be explained by a set of omitted variables, including environmental/sectoral impacts.

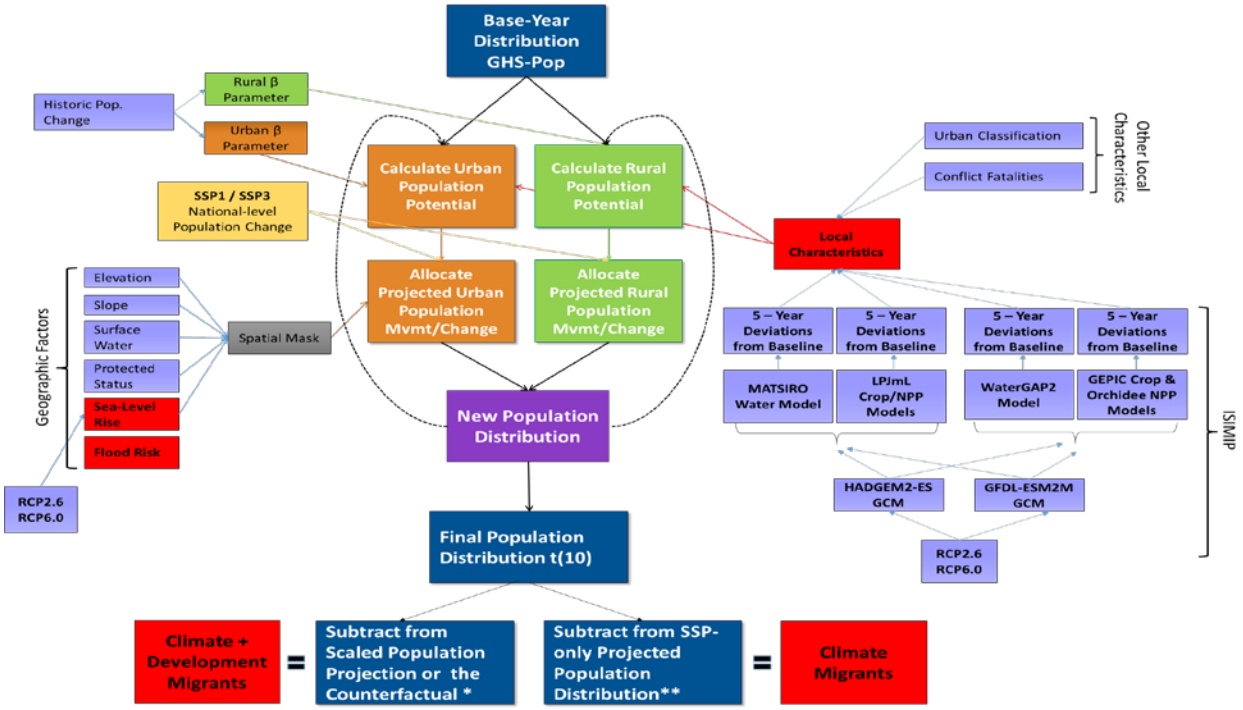
To incorporate these effects, the value of α_i is calculated such as to eliminate ε_i (from Figure 5) for each individual grid cell (which is labeled observed):

$$\Delta P_{i,t}^{obs} = obs\_a_i * \Delta P_{i,t}^{mod}$$

(Equation 4)

where $P_{i,t}^{obs}$ and $P_{i,t}^{mod}$ are the observed and modeled population change for each grid cell *i* and α_i is the factor necessary to equate the two.

Figure 4: Flowchart of Modeling Steps

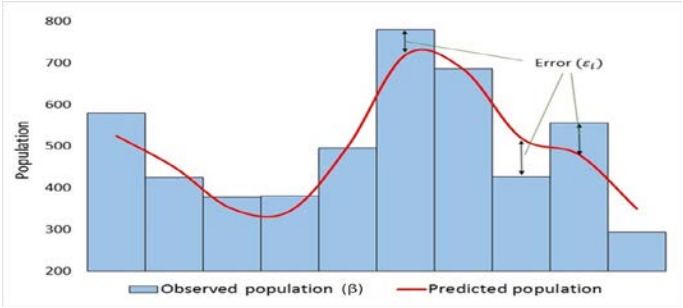


Note: Boxes in red represent the addition of climate impacts into the modeling framework or results that reflect climate impacts.

*The counterfactual population projection simply scales the population distribution in 2010 to country-level population totals appropriate to each SSP.

**The no climate impacts population projection represents the population projection without climate impacts (i.e. based only on the development trajectories embodied in the SSPs, and the conflict, age, and sex characteristics of the baseline population).

Figure 5: Cross-section of grid cells illustrating observed and projected population distributions. Note: The error term is used to calibrate the index ai



The second step is to estimate the relationship between observed index α_i and the different potential drivers of spatial population metrics by fitting a spatial lag model:

$$obs\_a_{i,t} = \rho W A_{i,t+} \beta_1 C_{i,t} + \beta_2 H_{i,t} + \beta_3 N_{i,t} + \beta_4 U_{i,t} + \beta_5 K_{i,t} + \varepsilon_{i,t}$$

(Equation 5)

where *C*, *H* and *N* are the five-year deviations from the historical baseline on crop yield, water availability, and net primary production, *U* is a dummy variable reflecting the status of each grid cell as urban (1) or rural (0), and *K* is the conflict-related fatalities metric. Together, these five variables and their respective coefficients constitute the set of explanatory variables that go into producing the index α_i. Note that for any grid cell in which *C* (crop yield) is a non-zero value, the value of *N* (net primary production) is automatically set to zero, so that only one of the two variables is contributing to the index a_i. Finally, ρ is the spatial autocorrelation coefficient and *W* is a spatial weight matrix. From this procedure, a set of grid cell specific α values is estimated for both urban and rural population change.

For future projections (for urban and rural populations), projected values are used of $C_{i,t}$, $H_{i,t}$ and $N_{i,t}$ and current values of $U_{i,t}$ and $K_{i,t}$ are used along with their respective coefficient estimates from Equation 5 to estimate spatially and temporally explicit values of a_i. Finally, to produce a spatially explicit population projection, estimates of β are adjusted to reflect the SSPs (e.g. the SSP4 storyline implies a more concentrated pattern of development than SSP5, see Jones and O'Neill, 2016) to produce estimates of the agglomeration effect, to which the spatio-temporally variant estimates of for the RCPs described above are applied, and finally exogenous projections of national urban and rural population change are incorporated in the model applied as specified in Equation 2.

It is important to note that, as a result of testing, grid cells meeting certain criteria are excluded from the calibration procedure. First, grid cells that are 100% restricted from future population growth by the spatial mask (*l*, Equation 2) are excluded, as the value of *l* in these grid cells (0), renders the observed value of a_i inconsequential. Second, the rural and urban distributions of observed α_i were found to include significant outliers that skewed coefficient estimates in Equation 2. In most cases, these values were found to correspond with very lightly-populated grid cells where a small over/under-prediction of the population in absolute terms (e.g. 100 persons) is actually quite large relative to the total population within the grid cell (e.g. large percent error). The value of α_i (the weight on potential), necessary to eliminate these errors, is often proportional to the size of the error in percentage terms, and thus can be quite large even though a very small portion of the total population is affected. Including these large values in Equation 5 would have a substantial impact on coefficient estimates. To combat this problem, the most extreme 2.5% of observations are eliminated on either end of the distribution. Third, because the model is calibrated to urban and rural change separately, grid cells in which rural population was reclassified as 100% urban over the decade (2000-2010) were excluded, as the effect would be misleading (in the rural distribution of change, it would appear an entire grid cell was depopulated, while in the urban change distribution the same grid cell

would appear to grow rapidly). It would be incorrect to attribute these changes to sectoral impacts when, in fact, they are the result of a definitional change. In most cases, these exclusions eliminate 5-10% of grid cells.

The coefficients in Table 3 are the results of applying the calibration procedure to countries for which the appropriate, high-resolution census data were available over at least two consecutive time-periods. Positive values indicate that the driver has a positive influence on local attractiveness (e.g. improved water availability or crop yields leads to increasing attractiveness), and a negative value indicates a negative influence on attractiveness (e.g. a larger number of conflict related deaths leads to a less attractive location). In general, the larger the value (positive or negative), the larger the influence of the driver. However, the coefficients are taken in conjunction with future estimates of each climate-related driver, or in the case of conflict, on present day conditions, and thus a larger coefficient does not necessarily indicate a larger weight on *P(i)*. The value of α_i for each grid cell at each time *t* is calculated as:

$$a_{i,t} = 1 + \beta_1 C_{i,t} + \beta_2 H_{i,t} + \beta_3 N_{i,t} + \beta_4 U_{i,t=1} + \beta_5 K_{i,t=1}$$

(Equation 6)

where the values *β(n)* are the coefficients on each driver, $C_{i,t}$, $H_{i,t}$ and $N_{i,t}$ are the projected deviations in crop yields, water stress, and net primary production for each grid cell *i* at each time *t*, and $U\_\{i, t = 1\}$ and K_{$i, t = 1$} are the present day values of urban classification and conflict fatalities. The assumed value of α_i in the absence of any local attractive or repulsive characteristics is 1, thus any positive values will increase local attractiveness, and negative values will decrease local attractiveness.

Table 3: Coefficients derivers from historical calibration

	Urban													
	Cote d'Ivoire	Egypt	Ethiopia	Gabon	Ghana	Kenya	Malawi	Morocco	Senegal	S. Africa	Zambia	Zimbabwe	Mean	Std. Dev
Water	0.653	2.648	2.514	0.581	0.346	0.694	1.638	0.852	0.570	0.055	0.281	0.422	0.938	0.823
Conflict	-0.005	n/a	-0.002	n/a	0.000	-0.003	-0.061	-0.045	-0.002	n/a	-0.006	-0.012	-0.015	0.021
	Rural													
	Cote d'Ivoire	Egypt	Ethiopia	Gabon	Ghana	Kenya	Malawi	Morocco	Senegal	S. Africa	Zambia	Zimbabwe	Mean	Std. Dev
Crop/NPP	2.082	1.345	1.342	1.727	1.433	0.480	1.069	1.820	0.915	0.206	1.642	2.124	1.349	0.572
Water	0.973	2.353	1.876	1.429	0.404	1.552	0.419	2.225	2.948	2.178	1.070	1.833	1.605	0.751
Conflict	-0.005	n/a	-0.031	n/a	-0.002	-0.020	-0.035	-0.358	-0.011	-0.069	-0.169	-0.096	-0.080	0.105

Note: Coefficient estimates derived from fitting the spatial autoregressive model to historical population distribution change data for the periods 1990-2000 and 2000-2010 for each of the potential drivers of spatial population change. Coefficients for Water and Crop/NPP can be interpreted similarly, but the coefficients are not normalized for Conflict.

II.2.3 Characterizing the Model

This modeling provides credible, spatially explicit estimates of changes in the population distribution (and indirectly migration) as a function of climate, demographic, and development trends. It is important to understand what the model does and does not do.

Gravity models, in their simplest form, can be used to reconstruct and quantify the past evolutions of population distributions based on observed agglomeration effects over large geographic regions, under varying conditions, and at alternative spatial scales. They can also be refined and expanded to incorporate additional details, such as environmental parameters that affect the relative attractiveness of locations, typically improving the capacity of the model to accurately replicate past trends and thus, theoretically, project into the future.

Gravity models do not directly model internal migration. Instead, internal migration is assumed to be the primary driver of deviations between population distributions in model runs that include climate impacts (in this model, crop production, ecosystem productivity, water availability, and flood risk) and the development-only (also referred to as the SSP or "no climate" models that include only the demographic and conflict metrics). Both types of models include the agglomeration effect. Migration is a "fast" demographic variable compared with fertility and mortality; it is responsible for much of the decadal-scale redistributions of population. Without significant variation in fertility/mortality rates between climate-migrant populations and non-migrant populations, it is fair to assume that differential population change between the climate impact scenarios and the development-only scenarios occur as a function of migration. Another way of saying this is that the model assumes that fertility and mortality rates are relatively consistent across populations in a locale. Note that the model does not provide any information about the directionality of migration. In other words, it cannot be inferred that migrants are moving from a given area of out-migration (e.g. a "hotspot" of climate out-migration) to a given area of in-migration. Rather, the model reflects broader changes in the spatial distribution of population as a result of climate impacts, with the distribution changing incrementally with each time step.

For each climate migration scenario, the model produces a range of estimates that reflect variation in the underlying inputs to the model, which in turn reflects scientific uncertainty over likely future climate projections and impacts and development trajectories. In any scenario, outcomes are a function of the global climate models and the sectoral impact models that drive climate impacts on population change. For each of the four scenarios, there are four models, consisting of different global climate model/ISIMIP combinations. The ensemble mean (or average) of the four models is reported as the primary result for each scenario. Uncertainty is reflected in the range of outcomes (across the four models) for each grid cell and at different levels of aggregation. While some may prefer to have just one figure, in a complex issue like climate-related migration, a scenario-based approach of plausible outcomes is preferable. It would be desirable to have even more scenarios, to better assess the uncertainty (or conversely confidence) in the results.

The model is analyzed at spatial and temporal scales that capture migration well. With grid cells of about 15 square kilometers at the equator, population shift can be considered a form of short-distance migration. The temporal scale of 5-year increments from 2015 to 2050 is adequate to capture the longer-term shifts in population caused by changes in

water availability, crop conditions, ecosystem productivity, and flood risk. The five-year temporal resolution of the model corresponds to the temporal resolution most national censuses consider when attempting to capture and quantify migration trends.[5] Shorter-term and/or seasonal migration are not captured by the model.

The focus is on the 30 years between 2020 and 2050. This period represents a meaningful planning horizon, especially when considering social dimensions of migration. Chapter 4 of Rigaud et al. (2018) considers water and agriculture sector impacts beyond 2050 by examining ISIMIP outputs for 2050-2100. They suggest that, if anything, the climate signal will become far stronger toward the end of the 21st century.

The model cannot forecast all future adaptation efforts or conflict, cultural, political, institutional, or technological changes. Discontinuities are likely to arise as a result of political events and upheavals that can heavily influence migration behavior. Armed conflict itself may have non-linear links to climate variability and change, but models are generally not yet sophisticated enough to forecast the changing nature of armed conflict or state failure with any precision. The scenario framework is not designed to predict shocks to any socioeconomic or political system, such as large-scale war or market collapse. The models can also not anticipate new technologies that may dramatically affect adaptation efforts to the degree that climate impacts become negligible. The SSPs, as well as output from the global climate model and ISIMIP, reflect plausible futures that span a wide range of global trajectories, with the caveat that extremely unpredictable or unprecedented events are explicitly excluded. The SSPs assume certain levels of adaptation and a continuation of the business as usual, and the projected scale of migration is not cast in stone. The scenario-based results should be seen as a plausible range of outcomes rather than precise forecasts–to spur policy and action to counter distress driven climate migration.

III. Data Set Description(s)

The projections for internal migration flows are available at 5-year intervals from 2020 to 2050 for a combination of 2 sets of Shared Socioeconomic Pathways (SSPs) scenarios and 3 sets of Representative Concentration Pathways (RCPs) scenarios. The unit of analysis is 7.5 arc-minute grid cells or ‘pixel’, covering the entire African continent (approximately 15 km at the equator). These grid cells are drawn using a vector representation, and not a tessellation of grid cells (rasters). While the calibration and projection took place at a grid cell resolution of 2.5 arc-minutes, these data have been aggregated to a 7.5 arc-minute resolution.

[5] Migration data are sporadic in national censuses, but when present, they are typically based on a “five year question”, which prompts respondents to indicate where they lived five years ago.

Data set web page:
SEDAC URL: https://sedac.ciesin.columbia.edu/data/set/climmig-acmi-internal-migration-projections
Permanent URL: https://doi.org/10.7927/5tv3-ff20

The data are available in two different versions: country specific and Africa-wide. The first version of the data set provides grid cell-level data for all countries individually, while the second version combines the data from all the countries in a single file. Both versions are available in Esri file Geodatabase (GDB) and OGC GeoPackage (GPKG) formats.

Data set format:
The data are available in GDB and GPKG formats. The downloadable is a compressed zip file containing: 1) GDB or GPKG file, and 2) PDF Documentation.

- Country files
  - <iso3c countrycode>.gdb [Esri file Geodatabase]
    - *Example for Angola:* AGO.gdb
    - 51 countries will have .gdb files
  - <iso3c countrycode>.gpkg [OGC GeoPackage
    - *Example for Angola:* AGO.gpkg)
    - 51 countries will have .gpkg files
- Africa continent-wide files
  - continent.gdb [Esri file Geodatabase]
  - continent.gpkg [OGC GeoPackage]

Data set downloads (Country-level): 51 countries will have "-gdb and -gpkg" files
- climmig-acmi-internal-migration-projections-2020-2050-iso3ccountrycode-gdb
- climmig-acmi-internal-migration-projections-2020-2050-iso3ccountrycode-gpkg

Data set downloads (Continent-level):
- climmig-acmi-internal-migration-projections-2020-2050-africa-gdb
- climmig-acmi-internal-migration-projections-2020-2050-africa-gpkg

Users of the Africa-wide data set should be mindful that grid cells are systematically assigned to a specific country. In the situation of grid cell border overlap between two or more countries, *two or more grid cells with the exact same geometries are generated, one for each country*. The data in the attribute table assigned to each grid cell (population, area, etc.) only refers to the part of a grid cell that extends over the territory of the country to which it is assigned to.

III.1 Field Naming Conventions

In creating fields in the attribute tables, the following are the naming conventions:

- **Individual scenario population projections**
  - <ISIMIP2bModel [*variable characters length*]><RCP [*2 digits*]><SSP [*1 digit*]><Year [*2 digits*]>
  - Thus, field *GeWGH26120* refers to the projected population counts in the grid cell for the GEPIC WaterGAP2 HADGEM2 ISIMIP2b scenario assuming Global Carbon Dioxide (CO2) emissions are consistent with an RCP 2.6 scenario and SSP1 population projections for the year 2020.
- **Population projections ignoring climate change**
  - <NC><RCP [*2 digits*]><SSP [*1 digit*]><Year [*2 digits*]>
  - Thus, field *NC26120* refers to the projected population count in the grid cell under an SSP1 population scenario for the year 2020, assuming no climate change impact (climatic variables are maintained at the historical averages for the period 1970-2010).
- **Ensemble means projections**
  - <EM><RCP [*2 digits*]><SSP [*1 digit*]><Year [*2 digits*]>
  - Thus, field *EM26120* refers to the projected number of populations in the grid cell averaging across all four ISIMIP2b scenarios considered here for the year 2020.
- **Climate migration projections**
  - Climate-induced internal migration projections are derived by subtracting the population projections from the ensemble means scenarios from the scenarios holding climatic conditions constant to their historical average.
  - They are labelled as follows: <EM><RCP [*2 digits*]><SSP [*1 digit*]> <Year [*2 digits*]>_<NC><RCP [*2 digits*]><SSP [*1 digit*]><Year *[2 digits*]>_diff.
  - In other words, field *EM26120_NC120_diff* provides ensemble means projections of internal climate migration under SSP1 and RCP 2.6 for the year 2020.

Note: Users interested in computing ISIMIP2 specific climate migration scenarios can simply subtract the projected populations in each grid cell under the historical climate change scenario from their ISIMIP2 projection population scenario of interest. In other words, subtracting the population projection *NC120* from the *GeWGH26120* population projections will provide projections of climate migrants under the GEPIC WaterGAP2 HADGEM2 ISIMIP2 scenario assuming Global CO2 emissions are consistent with an RCP 2.6 scenario and SSP1 population projections for the year 2020.

III.2 Data Dictionary

Field	Description	Unit	RCP	SSP	YEAR
Reference data					
PIXELID	Unique Pixel identifier		--	--	--
BL10	Baseline Population in 2010	Population Count	--	--	2010

Field	Description	Unit	RCP	SSP	YEAR
ISMIP2b model projections for combination of SSP and RCP population scenarios					
GeWGH26120	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	1	2020
GeWGG26120	GEPIC WaterGAP2 GFDL	Population Count	2.6	1	2020
LpMsH26120	LPJmL MATSIRO HADGEM2	Population Count	2.6	1	2020
LpMsG26120	LPJmL MATSIRO GFDL	Population Count	2.6	1	2020
GeWGH26125	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	1	2025
GeWGG26125	GEPIC WaterGAP2 GFDL	Population Count	2.6	1	2025
LpMsH26125	LPJmL MATSIRO HADGEM2	Population Count	2.6	1	2025
LpMsG26125	LPJmL MATSIRO GFDL	Population Count	2.6	1	2025
GeWGH26130	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	1	2030
GeWGG26130	GEPIC WaterGAP2 GFDL	Population Count	2.6	1	2030
LpMsH26130	LPJmL MATSIRO HADGEM2	Population Count	2.6	1	2030
LpMsG26130	LPJmL MATSIRO GFDL	Population Count	2.6	1	2030
GeWGH26135	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	1	2035
GeWGG26135	GEPIC WaterGAP2 GFDL	Population Count	2.6	1	2035
LpMsH26135	LPJmL MATSIRO HADGEM2	Population Count	2.6	1	2035
LpMsG26135	LPJmL MATSIRO GFDL	Population Count	2.6	1	2035
GeWGH26140	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	1	2040
GeWGG26140	GEPIC WaterGAP2 GFDL	Population Count	2.6	1	2040
LpMsH26140	LPJmL MATSIRO HADGEM2	Population Count	2.6	1	2040
LpMsG26140	LPJmL MATSIRO GFDL	Population Count	2.6	1	2040
GeWGH26145	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	1	2045
GeWGG26145	GEPIC WaterGAP2 GFDL	Population Count	2.6	1	2045
LpMsH26145	LPJmL MATSIRO HADGEM2	Population Count	2.6	1	2045
LpMsG26145	LPJmL MATSIRO GFDL	Population Count	2.6	1	2045
GeWGH26150	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	1	2050
GeWGG26150	GEPIC WaterGAP2 GFDL	Population Count	2.6	1	2050
LpMsH26150	LPJmL MATSIRO HADGEM2	Population Count	2.6	1	2050
LpMsG26150	LPJmL MATSIRO GFDL	Population Count	2.6	1	2050
GeWGH60120	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	1	2020
GeWGG60120	GEPIC WaterGAP2 GFDL	Population Count	6.0	1	2020
LpMsH60120	LPJmL MATSIRO HADGEM2	Population Count	6.0	1	2020
LpMsG60120	LPJmL MATSIRO GFDL	Population Count	6.0	1	2020
GeWGH60125	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	1	2025
GeWGG60125	GEPIC WaterGAP2 GFDL	Population Count	6.0	1	2025
LpMsH60125	LPJmL MATSIRO HADGEM2	Population Count	6.0	1	2025
LpMsG60125	LPJmL MATSIRO GFDL	Population Count	6.0	1	2025
GeWGH60130	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	1	2030
GeWGG60130	GEPIC WaterGAP2 GFDL	Population Count	6.0	1	2030
LpMsH60130	LPJmL MATSIRO HADGEM2	Population Count	6.0	1	2030

Field	Description	Unit	RCP	SSP	YEAR
LpMsG60130	LPJmL MATSIRO GFDL	Population Count	6.0	1	2030
GeWGH60135	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	1	2035
GeWGG60135	GEPIC WaterGAP2 GFDL	Population Count	6.0	1	2035
LpMsH60135	LPJmL MATSIRO HADGEM2	Population Count	6.0	1	2035
LpMsG60135	LPJmL MATSIRO GFDL	Population Count	6.0	1	2035
GeWGH60140	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	1	2040
GeWGG60140	GEPIC WaterGAP2 GFDL	Population Count	6.0	1	2040
LpMsH60140	LPJmL MATSIRO HADGEM2	Population Count	6.0	1	2040
LpMsG60140	LPJmL MATSIRO GFDL	Population Count	6.0	1	2040
GeWGH60145	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	1	2045
GeWGG60145	GEPIC WaterGAP2 GFDL	Population Count	6.0	1	2045
LpMsH60145	LPJmL MATSIRO HADGEM2	Population Count	6.0	1	2045
LpMsG60145	LPJmL MATSIRO GFDL	Population Count	6.0	1	2045
GeWGH60150	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	1	2050
GeWGG60150	GEPIC WaterGAP2 GFDL	Population Count	6.0	1	2050
LpMsH60150	LPJmL MATSIRO HADGEM2	Population Count	6.0	1	2050
LpMsG60150	LPJmL MATSIRO GFDL	Population Count	6.0	1	2050
GeWGH26320	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	3	2020
GeWGG26320	GEPIC WaterGAP2 GFDL	Population Count	2.6	3	2020
LpMsH26320	LPJmL MATSIRO HADGEM2	Population Count	2.6	3	2020
LpMsG26320	LPJmL MATSIRO GFDL	Population Count	2.6	3	2020
GeWGH26325	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	3	2025
GeWGG26325	GEPIC WaterGAP2 GFDL	Population Count	2.6	3	2025
LpMsH26325	LPJmL MATSIRO HADGEM2	Population Count	2.6	3	2025
LpMsG26325	LPJmL MATSIRO GFDL	Population Count	2.6	3	2025
GeWGH26330	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	3	2030
GeWGG26330	GEPIC WaterGAP2 GFDL	Population Count	2.6	3	2030
LpMsH26330	LPJmL MATSIRO HADGEM2	Population Count	2.6	3	2030
LpMsG26330	LPJmL MATSIRO GFDL	Population Count	2.6	3	2030
GeWGH26335	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	3	2035
GeWGG26335	GEPIC WaterGAP2 GFDL	Population Count	2.6	3	2035
LpMsH26335	LPJmL MATSIRO HADGEM2	Population Count	2.6	3	2035
LpMsG26335	LPJmL MATSIRO GFDL	Population Count	2.6	3	2035
GeWGH26340	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	3	2040
GeWGG26340	GEPIC WaterGAP2 GFDL	Population Count	2.6	3	2040
LpMsH26340	LPJmL MATSIRO HADGEM2	Population Count	2.6	3	2040
LpMsG26340	LPJmL MATSIRO GFDL	Population Count	2.6	3	2040
GeWGH26345	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	3	2045
GeWGG26345	GEPIC WaterGAP2 GFDL	Population Count	2.6	3	2045
LpMsH26345	LPJmL MATSIRO HADGEM2	Population Count	2.6	3	2045

Field	Description	Unit	RCP	SSP	YEAR
LpMsG26345	LPJmL MATSIRO GFDL	Population Count	2.6	3	2045
GeWGH26350	GEPIC WaterGAP2 HADGEM2	Population Count	2.6	3	2050
GeWGG26350	GEPIC WaterGAP2 GFDL	Population Count	2.6	3	2050
LpMsH26350	LPJmL MATSIRO HADGEM2	Population Count	2.6	3	2050
LpMsG26350	LPJmL MATSIRO GFDL	Population Count	2.6	3	2050
GeWGH60320	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	3	2020
GeWGG60320	GEPIC WaterGAP2 GFDL	Population Count	6.0	3	2020
LpMsH60320	LPJmL MATSIRO HADGEM2	Population Count	6.0	3	2020
LpMsG60320	LPJmL MATSIRO GFDL	Population Count	6.0	3	2020
GeWGH60325	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	3	2025
GeWGG60325	GEPIC WaterGAP2 GFDL	Population Count	6.0	3	2025
LpMsH60325	LPJmL MATSIRO HADGEM2	Population Count	6.0	3	2025
LpMsG60325	LPJmL MATSIRO GFDL	Population Count	6.0	3	2025
GeWGH60330	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	3	2030
GeWGG60330	GEPIC WaterGAP2 GFDL	Population Count	6.0	3	2030
LpMsH60330	LPJmL MATSIRO HADGEM2	Population Count	6.0	3	2030
LpMsG60330	LPJmL MATSIRO GFDL	Population Count	6.0	3	2030
GeWGH60335	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	3	2035
GeWGG60335	GEPIC WaterGAP2 GFDL	Population Count	6.0	3	2035
LpMsH60335	LPJmL MATSIRO HADGEM2	Population Count	6.0	3	2035
LpMsG60335	LPJmL MATSIRO GFDL	Population Count	6.0	3	2035
GeWGH60340	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	3	2040
GeWGG60340	GEPIC WaterGAP2 GFDL	Population Count	6.0	3	2040
LpMsH60340	LPJmL MATSIRO HADGEM2	Population Count	6.0	3	2040
LpMsG60340	LPJmL MATSIRO GFDL	Population Count	6.0	3	2040
GeWGH60345	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	3	2045
GeWGG60345	GEPIC WaterGAP2 GFDL	Population Count	6.0	3	2045
LpMsH60345	LPJmL MATSIRO HADGEM2	Population Count	6.0	3	2045
LpMsG60345	LPJmL MATSIRO GFDL	Population Count	6.0	3	2045
GeWGH60350	GEPIC WaterGAP2 HADGEM2	Population Count	6.0	3	2050
GeWGG60350	GEPIC WaterGAP2 GFDL	Population Count	6.0	3	2050
LpMsH60350	LPJmL MATSIRO HADGEM2	Population Count	6.0	3	2050
LpMsG60350	LPJmL MATSIRO GFDL	Population Count	6.0	3	2050
Projections for SSP only with climate data kept to its historical average					
NC120	SSP-only	Population Count	--	1	2020
NC125	SSP-only	Population Count	--	1	2025
NC130	SSP-only	Population Count	--	1	2030
NC135	SSP-only	Population Count	--	1	2035
NC140	SSP-only	Population Count	--	1	2040
NC145	SSP-only	Population Count	--	1	2045

Field	Description	Unit	RCP	SSP	YEAR
NC150	SSP-only	Population Count	--	1	2050
NC320	SSP-only	Population Count	--	3	2020
NC325	SSP-only	Population Count	--	3	2025
NC330	SSP-only	Population Count	--	3	2030
NC335	SSP-only	Population Count	--	3	2035
NC340	SSP-only	Population Count	--	3	2040
NC345	SSP-only	Population Count	--	3	2045
NC350	SSP-only	Population Count	--	3	2050
Ensemble means of different combination of SSP and RCP scenarios					
EM26120	Ensemble Mean Population Total	Population Count	2.6	1	2020
EM26125	Ensemble Mean Population Total	Population Count	2.6	1	2025
EM26130	Ensemble Mean Population Total	Population Count	2.6	1	2030
EM26135	Ensemble Mean Population Total	Population Count	2.6	1	2035
EM26140	Ensemble Mean Population Total	Population Count	2.6	1	2040
EM26145	Ensemble Mean Population Total	Population Count	2.6	1	2045
EM26150	Ensemble Mean Population Total	Population Count	2.6	1	2050
EM26320	Ensemble Mean Population Total	Population Count	2.6	3	2020
EM26325	Ensemble Mean Population Total	Population Count	2.6	3	2025
EM26330	Ensemble Mean Population Total	Population Count	2.6	3	2030
EM26335	Ensemble Mean Population Total	Population Count	2.6	3	2035
EM26340	Ensemble Mean Population Total	Population Count	2.6	3	2040
EM26345	Ensemble Mean Population Total	Population Count	2.6	3	2045
EM26350	Ensemble Mean Population Total	Population Count	2.6	3	2050
EM60120	Ensemble Mean Population Total	Population Count	6.0	1	2020
EM60125	Ensemble Mean Population Total	Population Count	6.0	1	2025
EM60130	Ensemble Mean Population Total	Population Count	6.0	1	2030
EM60135	Ensemble Mean Population Total	Population Count	6.0	1	2035
EM60140	Ensemble Mean Population Total	Population Count	6.0	1	2040
EM60145	Ensemble Mean Population Total	Population Count	6.0	1	2045
EM60150	Ensemble Mean Population Total	Population Count	6.0	1	2050
EM60320	Ensemble Mean Population Total	Population Count	6.0	3	2020
EM60325	Ensemble Mean Population Total	Population Count	6.0	3	2025
EM60330	Ensemble Mean Population Total	Population Count	6.0	3	2030
EM60335	Ensemble Mean Population Total	Population Count	6.0	3	2035
EM60340	Ensemble Mean Population Total	Population Count	6.0	3	2040
EM60345	Ensemble Mean Population Total	Population Count	6.0	3	2045
EM60350	Ensemble Mean Population Total	Population Count	6.0	3	2050
Climate migration projections (based on ensemble means population projections)					
EM26120_NC120_diff	Climate Migrants Total	Climate Migrants	2.6	1	2020
EM26125_NC125_diff	Climate Migrants Total	Climate Migrants	2.6	1	2025

Field	Description	Unit	RCP	SSP	YEAR
EM26130_NC130_diff	Climate Migrants Total	Climate Migrants	2.6	1	2030
EM26135_NC135_diff	Climate Migrants Total	Climate Migrants	2.6	1	2035
EM26140_NC140_diff	Climate Migrants Total	Climate Migrants	2.6	1	2040
EM26145_NC145_diff	Climate Migrants Total	Climate Migrants	2.6	1	2045
EM26150_NC150_diff	Climate Migrants Total	Climate Migrants	2.6	1	2050
EM26320_NC320_diff	Climate Migrants Total	Climate Migrants	2.6	3	2020
EM26325_NC325_diff	Climate Migrants Total	Climate Migrants	2.6	3	2025
EM26330_NC330_diff	Climate Migrants Total	Climate Migrants	2.6	3	2030
EM26335_NC335_diff	Climate Migrants Total	Climate Migrants	2.6	3	2035
EM26340_NC340_diff	Climate Migrants Total	Climate Migrants	2.6	3	2040
EM26345_NC345_diff	Climate Migrants Total	Climate Migrants	2.6	3	2045
EM26350_NC350_diff	Climate Migrants Total	Climate Migrants	2.6	3	2050
EM60120_NC120_diff	Climate Migrants Total	Climate Migrants	6.0	1	2020
EM60125_NC125_diff	Climate Migrants Total	Climate Migrants	6.0	1	2025
EM60130_NC130_diff	Climate Migrants Total	Climate Migrants	6.0	1	2030
EM60135_NC135_diff	Climate Migrants Total	Climate Migrants	6.0	1	2035
EM60140_NC140_diff	Climate Migrants Total	Climate Migrants	6.0	1	2040
EM60145_NC145_diff	Climate Migrants Total	Climate Migrants	6.0	1	2045
EM60150_NC150_diff	Climate Migrants Total	Climate Migrants	6.0	1	2050
EM60320_NC320_diff	Climate Migrants Total	Climate Migrants	6.0	3	2020
EM60325_NC325_diff	Climate Migrants Total	Climate Migrants	6.0	3	2025
EM60330_NC330_diff	Climate Migrants Total	Climate Migrants	6.0	3	2030
EM60335_NC335_diff	Climate Migrants Total	Climate Migrants	6.0	3	2035
EM60340_NC340_diff	Climate Migrants Total	Climate Migrants	6.0	3	2040
EM60345_NC345_diff	Climate Migrants Total	Climate Migrants	6.0	3	2045
EM60350_NC350_diff	Climate Migrants Total	Climate Migrants	6.0	3	2050
Climate migration Hotspots					
Hotspot20	Hotspot of Climate Mobility	Number of Scenarios at Top 5th Percentile of Distribution*	--	--	2020
Hotspot25	Hotspot of Climate Mobility	Number of Scenarios at Top 5th Percentile of Distribution*	--	--	2025
Hotspot30	Hotspot of Climate Mobility	Number of Scenarios at Top 5th Percentile of Distribution*	--	--	2030
Hotspot35	Hotspot of Climate Mobility	Number of Scenarios at Top 5th Percentile of Distribution*	--	--	2035
Hotspot40	Hotspot of Climate Mobility	Number of Scenarios at Top 5th Percentile of Distribution*	--	--	2040
Hotspot45	Hotspot of Climate Mobility	Number of Scenarios at Top 5th Percentile of Distribution*	--	--	2045

Field	Description	Unit	RCP	SSP	YEAR
Hotspot50	Hotspot of Climate Mobility	Number of Scenarios at Top 5th Percentile of Distribution*	--	--	2050
Urban-rural pixel classification					
Urban20	Pixel Classed as Urban	Binary Variable (0–1)	--	--	2020
Urban25	Pixel Classed as Urban	Binary Variable (0–1)	--	--	2025
Urban30	Pixel Classed as Urban	Binary Variable (0–1)	--	--	2030
Urban35	Pixel Classed as Urban	Binary Variable (0–1)	--	--	2035
Urban40	Pixel Classed as Urban	Binary Variable (0–1)	--	--	2040
Urban45	Pixel Classed as Urban	Binary Variable (0–1)	--	--	2045
Urban50	Pixel Classed as Urban	Binary Variable (0–1)	--	--	2050
Ancillary data					
ANTHROME	Anthropogenic Biome (Livelihood Zone)	Recode of Zones	--	--	--
Borderland	If pixel proximity to a border, which country	Recode of Borderland	--	--	--
Farm_Sys	Type of crop cultivation / pastoralism practiced within pixel	Recode of farming type	--	--	--
Urban center	Name of urban agglomeration, if applicable	Recode of urban areas	--	--	--
Coast50km	Is pixel within 50km of a coast?	Binary Variable (0–1)	--	--	--
Coast100km	Is pixel within 100km of a coast?	Binary Variable (0–1)	--	--	--
Delta	Is pixel part of a Delta region?	Recode of delta region	--	--	--
Country	Iso3c codes of country	Recode of country	--	--	--

IV. How to Use the Data

The data files can be used to access and visualize projections of population change and migration patterns at different timescales (from 2020 to 2050) and under different combinations of RCP and SSP scenarios at a 7.5 arc-minute resolution across the African continent. The data are available in two file versions: individual country files (51) and continent files. The continent files are simply an aggregation of individual files for all 51 countries in the data set. For each version (country and continent files), the data are available in Esri file Geodatabase (GDB) and OGC GeoPackage (GPKG) format.

As a warning, the continent files should be approached cautiously as border grid cells for two countries are not merged, but rather overlayed with each other with one assigned to each country.

V. Potential Use Cases

For a discussion of the results and illustration on how the projections can be used to inform either country, regional or continental population and migration projections, see the following report:

Amakrane, K., S. Rosengaertner, N. P. Simpson, A. de Sherbinin, J. Linekar, C. Horwood, B. Jones, F. Cottier, S. Adamo, B. Mills, G. Yetman, T. Chai-Onn, J. Squires, J. Schewe, B. Frouws, and R. Forin. 2023. African Shifts: The Africa Climate Mobility Report, Addressing Climate-Forced Migration & Displacement. New York: Global Centre for Climate Mobility. https://africa.climatemobility.org/report.

VI. Limitations

The climate migration modeling results incorporate six main sources of uncertainty that can affect the estimated number of climate migrants or the differences between the four scenarios and the development-only scenario.

1. ISIMIP impacts vary across models. The differences result in different effects in the gravity model; models with the highest negative impacts repel more people from affected areas than those projecting less extreme outcomes. Similarly, in isolated cases (a small number of grid cells), different ISIMIP models can disagree on the positive/negative nature of changes, leading one model to attract population and the other to repel.

2. Variations between the two global climate models—HadGEM2-ES and GFDL-ESM2M—can amplify the ISIMIP differences. The global climate models were selected in part because their future precipitation trends differ substantially in magnitude, and partly even in sign (see Section II.1.6). This variance in precipitation has an impact on the water, crop, and NPP models.

3. The modeling has a temporal component that can influence population distribution trajectories. Stronger sectoral impacts early in the 30-year projection period will have greater influence than the same impacts later in that period, because those early impacts affect the gravitational pull of locations, creating "temporal" momentum over which later climate impacts may have less influence. Similarly, the timing of population change (growth or decline) projected by the SSPs relative to the development of sectoral impacts can influence outcomes. For example, for most countries in the study, projected population growth is greatest during the first decade; if conditions are also predicted to deteriorate severely during that period, the impact on migration will be greater than if the deterioration took place during a more demographically stable period.

4. If the "no climate impacts" model finds that a place is relatively attractive and the sectoral climate impacts are positive or neutral (relative to other areas that see negative impacts), it will have the effect of reinforcing the attractiveness of that area. Conversely, in remote areas experiencing population decline and negative climate impacts, "push" factors will be reinforced. This phenomenon creates spatial momentum.

5. Model parameterization affects the results. The model is calibrated using actual population changes in association with actual climate impacts (represented by ISIMIP model outputs) for two periods, 1990-2000 and 2000-2010. This calibration was done using the two separate sets of model combinations: the MATSIRO and WaterGAP2 water models, the LPJmL-Crop and GEPIC crop models, and the LPJmL-NPP and ORCHIDEE ecosystem models. Different parameters correspond to the different models. If the parameter estimates are close together across the different crop or water models, there will be less variation in the population distribution projected by each model; the uncertainty around the ensemble mean (measured using the coefficient of variation) will therefore be lower. Conversely, if parameter estimates are not close together, there will be greater uncertainty around the ensemble mean.

6. The use of GHS-POP, which is a modeled population surface where population is allocated based on remote sensing imagery, may have introduced issues in the model calibration, whereby the GHS-POP population surface recorded false positives ("built-up" areas that were in-fact rock outcrops or dried lake beds) or false negatives (places where small settlements were missed). These problems affected a relatively very small fraction of country territory, and countries for calibration were chosen that had fewer of such issues. False positives with large values, typically a function of the GHS-POP algorithm placing the majority of an administrative units population in a small number of grid cells where there is no large settlement, have the potential to skew the calibration results. To avoid this problem, grid cells were spot-checked with large populations outside of known urban centers using Esri base map imagery (satellite imagery of the landscape). If grid cells with large populations did not correspond to large settlements in the base map imagery, they were eliminated from the calibration procedure.

For these reasons, users are advised to exercise particular caution when examining data from specific ISIMP2b projection models. Likewise, users should avoid interpreting specific pixel values, and consider general trends across a region of interest.

VII. Acknowledgments

This data set was created as part of the Africa Climate Mobility Initiative (ACMI) (Amakrane et al., 2023, see Scientific publication in section X). The ACMI is a joint collaboration between the African Union Commission, the United Nations Development

Programme (UNDP), the United Nations Framework Convention on Climate Change (UNFCCC), the International Organization for Migration (IOM), and the World Bank. This work benefited from the generous financial support of the Federal Foreign Office of the Federal Republic of Germany, the Ford Foundation, Mayors Migration Council, Open Society Foundations, Porticus foundation, and the Robert Bosch Stiftung foundation.

Researchers at the CUNY Institute for Demographic Research (CIDR), City University of New York, the Center for Integrated Earth System Information (CIESIN), Climate School, Columbia University, and the Lamont-Doherty Earth Observatory (LDEO), Climate School, Columbia University contributed to the calibration and the projections of climate mobility under the guidance of the Global Centre for Climate Mobility (GCCM).

Funding for dissemination of this data set was provided under the U.S. National Aeronautics and Space Administration (NASA) contract 80GSFC18C0111 for the continued operation of the Socioeconomic Data and Applications Center (SEDAC), which is operated by the Center for Integrated Earth System Information (CIESIN) of Columbia University.

VIII. Disclaimer

CIESIN follows procedures designed to ensure that data disseminated by CIESIN are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact SEDAC User Services at ciesin.info@ciesin.columbia.edu. Neither CIESIN nor NASA verifies or guarantees the accuracy, reliability, or completeness of any data provided. CIESIN provides this data without warranty of any kind whatsoever, either expressed or implied. CIESIN shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided by CIESIN.

IX. Use Constraints

X. Recommended Citation(s)

Data set(s):

CUNY Institute for Demographic Research (CIDR), City University of New York, Center for Integrated Earth System Information (CIESIN), Columbia University, and Global Centre for Climate Mobility (GCCM). 2025. Africa Climate Mobility Initiative (ACMI): Internal Migration Projections. Palisades, New York: NASA Socioeconomic Data and Applications Center. https://doi.org/10.7927/5tv3-ff20. Accessed DAY MONTH YEAR.

Scientific publication:

Amakrane, K., S. Rosengaertner, N. P. Simpson, A. de Sherbinin, J. Linekar, C. Horwood, B. Jones, F. Cottier, S. Adamo, B. Mills, G. Yetman, T. Chai-Onn, J. Squires, J. Schewe, B. Frouws, and R. Forin. 2023. African Shifts: The Africa Climate Mobility Report, Addressing Climate-Forced Migration & Displacement. New York: Global Centre for Climate Mobility. https://africa.climatemobility.org/report.

XI. Source Code

No source code is provided.

XII. References

Amakrane, K., S. Rosengaertner, N. P. Simpson, A. de Sherbinin, J. Linekar, C. Horwood, B. Jones, F. Cottier, S. Adamo, B. Mills, G. Yetman, T. Chai-Onn, J. Squires, J. Schewe, B. Frouws, and R. Forin. 2023. African Shifts: The Africa Climate Mobility Report, Addressing Climate-Forced Migration & Displacement. New York: Global Centre for Climate Mobility. https://africa.climatemobility.org/report.

Brown, S., R. J. Nicholls, J. A. Lowe, and J. Hinkel. 2016. Spatial variations of sea-level rise and impacts: An application of DIVA. *Climatic Change*, 134(3), 403-416. https://doi.org/10.1007/s10584-013-0925-y.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2018. Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4JW8BX5.

Chang, J., et al. 2017. Benchmarking carbon fluxes of the ISIMIP2a biome models. *Environmental Research Letters*, 12(4), 45002. https://doi.org/10.1088/1748-9326/aa63fa.

Clark, W. A. V., and S. D. Davies. 2002. Changing jobs and changing houses: Mobility outcomes of employment transitions. *Journal of Regional Science* 39 (4): 653–673. https://doi.org/10.1111/0022-4146.00154.

Clement, V., K. K. Rigaud, A. de Sherbinin, B. Jones, S. Adamo, J. Schewe, S. Nian, and E. Shabahat. 2021. Groundswell Part 2: Acting on Internal Climate Migration. Washington, DC: The World Bank. https://openknowledge.worldbank.org/handle/10986/36248.

Dasgupta, S., B. Laplante, C. Meisner, D. Wheeler, and J. Yan. 2007. The Impact of Sea Level Rise on Developing Countries: A Comparative Analysis. Policy Research Working Paper 4136, Washington, DC: The World Bank. http://documents.worldbank.org/curated/en/156401468136816684/The-impact-of-sea-level-rise-on-developing-countries-a-comparative-analysis.

Frieler, K., et al. 2017. Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geoscientific Model Development*, 10(12), 4321–4345. https://doi.org/10.5194/gmd-10-4321-2017.

Gustafson, P. 2001. Retirement migration and transnational lifestyles. *Ageing & Society* 21 (4): 371–394. https://doi.org/10.1017/S0144686X01008327.

Hausfather, Z. and G. P. Peters. 2020. Emissions–the 'business as usual' story is misleading. *Nature*. Comment, 29 January 2020. https://doi.org/10.1038/d41586-020-00177-3.

Hinkel, J., S. Brown, L. Exner, R. J. Nicholls, A. T. Vafeidis, and A. S. Kebede. 2012. Sea-level rise impacts on Africa and the effects of mitigation and adaptation: an application of DIVA. *Regional Environmental Change*, 12(1), 207-224. https://doi.org/10.1007/s10113-011-0249-2.

Joint Research Centre (JRC), European Commission, and Center for International Earth Science Information Network (CIESIN), Columbia University. 2021. Global Human Settlement Layer: Population and Built-Up Estimates, and Degree of Urbanization Settlement Model Grid. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/h4154f0w.

Kim, T.-K., M. W. Horner, and R. W. Marans. 2005. Life cycle and environmental factors in selecting residential and job locations. *Housing Studies* 20 (3): 457–473. https://doi.org/10.1080/02673030500062335.

Jones, B., and B. C. O'Neill. 2016. Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environmental Research Letters*, 11(8), p.084003. https://doi.org/10.1088/1748-9326/11/8/084003.

Jones, B., and B. C. O'Neill. 2013. Historically Grounded Spatial Population Projections for the Continental United States. *Environmental Research Letters*, 8 (4): 44021. https://doi.org/10.1088/1748-9326/8/4/044021.

Lange, S., et al. 2020. Projecting Exposure to Extreme Climate Impact Events Across Six Event Categories and Three Spatial Scales. *Earth's Future*, 8(12). https://doi.org/10.1029/2020ef001616.

Müller, C., et al. 2017. Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications. *Geoscientific Model Development*, 10(4), 1403–1422. https://doi.org/10.5194/gmd-10-1403-2017.

Müller, C., et al. 2021. Exploring uncertainties in global crop yield projections in a large ensemble of crop models and CMIP5 and CMIP6 climate scenarios. *Environmental Research Letters*. https://doi.org/10.1088/1748-9326/abd8fc.

Niedomysl, T., and U. Fransson. 2014. On distance and the spatial dimension in the definition of internal migration. *Annals of the Association of American Geographers*, 104(2), pp.357-372. https://doi.org/10.1080/00045608.2013.875809.

Portmann, F. T., S. Siebert, and P. Döll. 2010. MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, 24(1), https://doi.org/10.1029/2008GB003435.

Rich, D. C. 1980. Potential Models in Human Geography (Concepts and Techniques in Modern Geography). 26. ISBN 10: 0860940446, ISBN 13: 9780860940449.

Rigaud, K. K., A. de Sherbinin, B. Jones, S. Adamo, D. Maleki, N. E. Abu-Ata, A. T. Casals Fernandez, A. Arora, T. Chai-Onn, and B. Mills. 2021. Groundswell Africa: Internal Climate Migration in West African Countries. Washington, DC: The World Bank. https://openknowledge.worldbank.org/handle/10986/36404.

Rigaud, K. K., A. de Sherbinin, B. Jones, S. Adamo, D. Maleki, N. E. Abu-Ata, A. T. Casals Fernandez, A. Arora, T. Chai-Onn, and B. Mills. 2021. Groundswell Africa: Internal Climate Migration in the Lake Victoria Basin. Washington, DC: The World Bank. https://openknowledge.worldbank.org/handle/10986/36403.

Rigaud, K. K., A. de Sherbinin, B. Jones, J. Bergmann, V. Clement, K. Ober, J. Schewe, S. Adamo, B. McCusker, S. Heuser, and A. Midgley. 2018. Groundswell: Preparing for Internal Climate Migration. Washington DC: The World Bank. https://openknowledge.worldbank.org/handle/10986/29461.

Rogelj, J., M. den Elzen, N. Höhne, T. Fransen, H. Fekete, H. Winkler, R. Schaeffer, F. Sha, K. Riahi, and M. Meinshausen. 2016. Paris Agreement climate proposals need a boost to keep warming well below 2 C. *Nature*, 534(7609), pp.631-639. https://doi.org/10.1038/nature18307.

Rosvold, E., and H. Buhaug. 2021. Geocoded Disasters (GDIS) Dataset. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/zz3b-8y61.

Rowell, D. P. 2019. An Observational Constraint on CMIP5 Projections of the East African Long Rains and Southern Indian Ocean Warming. *Geophysical Research Letters*, 46(11), 6050–6058. https://doi.org/10.1029/2019GL082847.

Schewe, J., and A. Levermann. 2017. Non-linear intensification of Sahel rainfall as a possible dynamic response to future warming. *Earth System Dynamics*, 8(3), 495–505. https://doi.org/10.5194/esd-8-495-2017.

Schlund, M., A. Lauer, P. Gentine, S. C. Sherwood, and V. Eyring. 2020. Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6? *Earth System Dynamics*, 11(4), 1233–1258. https://doi.org/10.5194/esd-11-1233-2020.

Tebaldi, C., et al. 2021. Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6. *Earth System Dynamics*, 12(1), pp.253-293. https://doi.org/10.5194/esd-12-253-2021.

van Vuuren, D. P., E. Kriegler, B. C. O'Neill, K. L. Ebi, K. Riahi, T. R. Carter, J. Edmonds, S. Hallegatte, T. Kram, R. Mathur, and H. Winkler. 2014. A new scenario framework for Climate Change Research: scenario matrix architecture. *Climatic Change* 122 (3): 373–386. https://doi.org/10.1007/s10584-013-0906-1.

Veldkamp, T. I. E., Y. Wada, J. C. J. H. Aerts, P. Döll, S. N. Gosling, J. Liu, Y. Masaki, T. Oki, S. Ostberg, Y. Pokhrel, Y. Satoh, H. Kim and P. J. Ward. 2017. Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century. *Nature Communications*. https://doi.org/10.1038/ncomms15697.

Warntz, W., and P. Wolff (eds.). 1971. Breakthroughs in geography. New York: New American Library. https://search.worldcat.org/title/Breakthroughs-in-geography/oclc/1148006817.

Zaherpour, J., et al. 2018. Worldwide evaluation of mean and extreme runoff from six global-scale hydrological models that account for human impacts. *Environmental Research Letters*, 13(6), 065015. https://doi.org/10.1088/1748-9326/aac547.

Zhao, F. 2021. Personal Communication on 13 October 2021.

XIII. Documentation Copyright and License

Appendix 1. Data Revision History

No revisions have been made to this data set.

Appendix 2. Contributing Authors & Documentation Revision History

Revision Date	ORCID	Contributors	Revisions
January 31, 2025	0000-0002-8875-4864	Fabien Cottier, Alex de Sherbinin	This document is the 1st instance of documentation.