Bocconi

WORKING PAPER SERIES
ISSN 2785-1079

Number 23
MAY 2023

SOIL ARIDIFICATION, PRECIPITATIONS, AND INFANT HEALTH: EVIDENCE FROM AFRICA

Jacopo Lunghi

Maurizio Malpede

Marco Percoco

Università Bocconi
GREEN
Centre for Geography, Resources, Environment, Energy and Networks

Soil Aridification, Precipitations, and Infant Health: Evidence from Africa

Jacopo Lunghi[†], Maurizio Malpede[‡], and Marco Percoco[⋆]

[†] Bocconi University, Milan, Italy

[‡] University of Verona and GREEN

[⋆]Bocconi University and GREEN

This version: May 5, 2023

Abstract

This study shows how soil aridity (proxied with a measure of soil potential evapotranspiration) impacts child wellbeing in Sub-Saharan Africa. Using climate and infant health data from a grid of approximately 4,000 cells in 34 African countries, we find that infants born in arid areas are comparatively more likely to die under the age of 5 and be systematically underweight at birth. In addition, we show how the aridity measure in this study reduces the effect of rainfall on child wellbeing and how aridification drives substantial heterogeneity in the estimated response to increasing precipitation. The findings are combined with model projections of future climate conditions to emphasize the importance of accounting for aridity alongside precipitations when assessing the economic impact of climate.

Keywords: Rainfall, climate change, potential evapotranspiration, child mortality, infant health

JEL Classification: J1, J13, I15, Q54, Q56, O15

1 Introduction

The effects of climate change and global warming have been central issues in the recent scientific and economic literature. Studies have found strong evidence of a close connection between the evolution of weather factors and economic welfare, both at the micro and macro level (Dell et al. 2008, 2012, Burke et al. 2015, Zhang et al. 2017, Peri & Sasahara 2019). Yet, while climate change is now extensively validated by empirical evidence, the nature of this relationship remains unclear. The consequences of this accelerating process for the planet and its inhabitants are still a matter of great discussion, and relatively little consensus has been reached.

Among the numerous implications of climate change, water access for people worldwide remains a key concern. A reliable water source is crucial for developing sustainable human settlements and agricultural and economic systems, and water availability is strongly influenced by precipitation and its intensity. According to current predictions, global rainfall is expected to rise in the next few decades due to climate change. Panel [A] in Figure 1 illustrates the upcoming trend in rainfall using five of the most commonly used earth system models (ESM) for the periods 2040–2079 and 2060–2099: all the models predict a consistent increase in global precipitation levels, roughly between 5.7% and 12.2%.

Most of the economic literature has examined the separate actions of rainfall and temperatures as key determinants of the amount of water available to individuals in developing countries. In this respect, Maccini & Yang (2009) use early-life precipitations in Indonesia to show how fluctuations in agricultural output have a long-lasting impact on women's health, schooling, and socio-economic status. Kudamatsu et al. (2012) used rainfall and temperature to determine how a period of relative drought affects infant mortality in Africa, finding that in-utero drought exposure increases babies' risk of death. In both studies, nutrition and malnutrition (experienced either by a mother or an infant during early life) constitute the primary channel through which climate conditions affect individuals. Conversely, Jayachandran (2006) used rainfall shocks in underdeveloped rural India to explain the impact of productivity shocks to agriculture on wages, migration, and credit constraints.

Based on these findings, rainfall is viewed to have a positive impact on human development in rural settings, primarily through its beneficial effect on the agricultural sector. Nonetheless, precipitations alone do not capture actual soil water availability, which also depends on concurring factors, such as land quality, solar radiation, temperature, air humidity, and wind speed. A direct consequence of these interactions is that as climate change accelerates and temperatures rise globally, the stability of the relationship between precipitations and water availability may change dramatically, with substantial effects on human wellbeing. In this paper, we aim to explore the effect of aridity, as a measure of the capacity of the soil to retain water, on child health in Africa. To this end, we combine data on temperature and rainfall defined on 0.5° latitude x 0.5° longitude grids (ca. 56 km x 56 km at the equator) with an original measure of soil aridification, namely the potential evapotranspiration (henceforth PET) index. PET considers the combination of two sources of soil water loss, namely the soil surface evaporation (i.e., the process whereby liquid water is converted to water vapor and removed from the evaporating surface) and reference crop transpiration (the vaporization of liquid water contained in plant tissues and the vapor removal to the atmosphere; Allen et al. 1998). Following environmental research on the topic, we consider the PET as an indicator of the aridity of an area (Rind et al. 1990, Cherlet et al. 2018) and study the impact of our measures on agriculture and infant wellbeing in Africa, where the increase in precipitations is projected to occur similarly to in the rest of the world (Figure 1, Panel [B]).

Consistent with previous literature, this study focuses on the core indicators of infant health as a proxy for economic wellbeing (Aber et al. 1997, Benshaul-Tolonen 2018). We use geolocalized data from four waves of the Demographic and Health Surveys to assemble information on women and children in 34 countries between 1992 and 2018, with a primary focus on child mortality, self-reported by the women in the survey. Based on a sub-sample of individuals, the study also documents the joint effect of rainfall and PET on body mass index (BMI) at the time of the interviews and in terms of their size at birth in children under five.

The findings suggest that land aridification counteracts the beneficial effects of higher rainfall levels: children in more arid regions are more likely to die before the age of five, tend to be smaller at birth, and have systematically lower BMI values by the time the interviews were conducted. The results are robust for redefining the primary outcome variables with comparable measures, such as one-year mortality and weight-to-height ratios. Moreover, the positive effect of rainfall on child health is shown to be susceptible to aridity and progressive land aridification: children in dry climates and those living in areas that have experienced a stronger increase in PET levels in the last decades benefit disproportionally less from years of abundant rainfall.

Our research contributes to the debate on the relationship between precipitations and human wellbeing, adding to the discussion on the effect of rainfall on health during child-bearing and early life (Kudamatsu et al. 2012, Le & Nguyen 2021, Sivadasan & Xu 2021, Ponnusamy 2022). Moreover, with rainfall being associated with a variety of socioeconomic outcomes (Rose 1999, Barrios et al. 2010, Brückner & Ciccone 2011, Björkman-Nyqvist 2013, Rocha & Soares 2015), the paper offers a new perspective to the literature about heterogeneity in the response to rainfall shocks (Sarsons 2015, Damania 2020, Mary 2022) and, more in general, unexpected weather events (Burgess & Donaldson 2010, Dell et al. 2014). The paper also contributes to the discussion about the present and future impact of climate change have in low-income countries (Rabassa et al. 2014, Bharadwaj et al. 2020, Emediegwu et al. 2022, Randell et al. 2021), providing new evidence on the relationship between climatic factors, soil aridity, and child health.

The remainder of the paper is organized as follows: Section 2 summarizes the debate on the future of precipitations and presents the data. In Section 3, we propose a research design to study the impact of weather conditions on child health. In Section 4, we discuss our main findings, and in Section 5, we complement the evidence by testing for potential urban-rural heterogeneity and assessing robustness to potentially endogenous migration. Finally, Section 6 concludes with a simple back-of-the-envelope calculation to quantify the expected toll of aridification on children's wellbeing.

2 Background and Data Description

2.1 Measures of Soil Aridity

The role of precipitations and humidity within the general pattern of climate change is still disputed. It was initially posited that climate change could lead to a wet-get-wetter and dry-get-drier pattern due to atmospheric moisture convergence and divergence (Held & Soden 2006). Nonetheless, due to a lack of consistent empirical evidence, this hypothesis has seemingly been replaced by a contrasting view in which mean precipitations are expected to increase at high- and mid-latitudes. Still, they will likely not decrease in subtropical regions (Kirtman et al. 2013, Donat et al. 2016). In turn, average rainfall at the global level is expected to increase in the coming years (Cherlet et al. 2018).

The evidence found in the existing literature suggests that higher precipitation levels could help prevent drought and boost agricultural productivity (Gornall et al. 2010). As a result, it is possible that communities relying on agricultural output as a primary source of income and food supply may potentially benefit from foreseen weather variations. In this scenario, climate change could positively affect rural households in developing countries in the upcoming decades. In this paper, we study the direct effect of soil evapotranspiration on infant health measures and rainfall's role in light of the PET. The CRU TS4.04 dataset contains time-series data on month-by-month variations in climate over the period 1901–2019, provided on high-resolution (0.5° longitude x 0.5° latitude) grids.[1] We focus on an

[1] As is usually the case for model-computed weather data, the decision to use the CRU database comes with partial concerns regarding the quality of data. We justify the suitability of this dataset for our purpose in Section A.1 in the Appendix.

area covering almost 40% of the entire African continent. Moreover, this paper focuses on weather conditions between 1951 and 2019 and considers three measures of climate variation: PET (mm/month), precipitations (mm/month), and monthly mean temperature (°C). The PET represents the amount of water lost from a cropped reference surface that is not short of water (a hypothetical grass reference crop with specific characteristics). As such, this measure estimates the evaporative demand of the atmosphere independently of crop type, crop development, and management practices. The PET estimates are calculated using a variant of the Penman-Monteith method, briefly summarized in the Appendix.

We report the evolution of our environmental measures over the sample period in Table 1. The yearly averages of each variable are included in time windows of 15 years each. The predictions regarding increasing rainfall cannot be confirmed using historical data and a limited sample period; however, the study's series points in that direction. Indeed, the level of average precipitations appears to follow a reverting trend, decreasing at first and then rising again in the early 2000s. Conversely, average PET levels and temperatures are on a stable, increasing path. Additionally, the process of aridification and rising temperatures seem to accelerate starting from the 1980s.

Table 1 also reports correlations between the climate variables used in this research. To interpret cross-sectional variability across the grid, the average yearly precipitations and PET levels throughout the sample period are plotted in Figure 2. While the PET shows steady but little volatility over time, the sample retains substantial cross-sectional variation, ranging from areas with almost no evapotranspiration to cells where this measure exceeds the average precipitations.[2]

[2]When the PET exceeds the actual precipitation, it indicates that the soil may eventually dry out unless irrigation is used to offset the loss. However, the effective amount of water dispersed also depends on the type of plants cultivated on the land.

2.2 Measures of Infant Health

This study combines data on weather factors and agricultural production with information on child health from the Demographic and Health Surveys (DHS). This program collects nationally representative data on health and population in developing countries, compensating for the lack of high-quality infant health statistics, particularly in Sub-Saharan Africa. In the present study, we construct a dataset using survey waves conducted between 1992 and 2018 for 34 African countries. The survey is stratified into clusters, localized with displacements of up to 2 km for urban and 10 km for rural points. The geographical distribution of clusters available in at least one wave is plotted in Figure 3. Data from the Individual Recode dataset and the Child Recode dataset are utilized. The former contains data on every eligible woman, including individual socio-economic characteristics, birth history, pregnancy, and postnatal care; the latter comprises some core child health indicators for children under five years old and their mothers. Here, we restricted the sample to women who experienced at least one completed pregnancy.

Table 2 summarizes the descriptive statistics on mothers and infant mortality across the sample period.[3] We obtained information for 468,873 women aged 13 to 50, corresponding to more than 2 million births between 1955 and 2019. However, wealth and education characteristics and records on size at birth and BMI were only available for a sub-sample of individuals, leaving us with a final sample of roughly 1.7 million births and over 400,000 children under five.

Since this study considers a large number of countries and a variety of developing settings, the women in the sample are heterogeneous in terms of education and economic wealth scores. The average years of schooling are relatively low: the women in our sample attended school for less than 4 years, with almost 50% declaring having received no formal education. Moreover, roughly 70% to 75% of the sample comprises women living in rural households. As extensively highlighted in previous literature, fertility appears to be high

[3] All the variables relevant to our study are described in Section A.1 in the Appendix.

in the sample, with an average of more than five births per woman. These measures tend to remain homogeneous across different survey waves.

The first outcome variable is child mortality, retrieved from a woman's reported completed pregnancies. The variable is a binary indicator that takes a value of 1 if a woman reports that the child died within 60 months of the date of birth. This methodology of computing infant mortality has already been adopted in the literature employing DHS data (Kudamatsu et al. 2012, Benshaul-Tolonen 2018). As Table 2 indicates, despite a steady decrease, the average probability of a pregnancy resulting in a child's death ranges from 8.9% to 12% across the survey waves. However, infant mortality is highly volatile across the sample, with large standard deviations common to all the rounds. Moreover, there seems to be only a negligible difference between the mortality rates of male and female children in the sample.

In addition to child mortality, we consider other infant health indicators to provide further evidence of the mechanism discussed in this paper. Specifically, we consider BMI, calculated using the new Child Growth Standards (CGS) from the World Health Organization (WHO) at the time of the interviews and a categorical variable indicating whether the child's size at birth was below or above average.[4] These measures allow us to better understand whether aridity could also affect the health of those who survive. These indicators have been collected irregularly throughout the DHS waves; hence, they are only available for a subset of children under five years of age. From Table 2, we notice that the infants in our sample consistently display below-average weight/height ratios (the average child falls in the left tail of the weight/height distribution, around the 40$^{\text{th}}$ percentile). At the same time, regarding BMI, this only appears to occur towards the last waves of survey collection. Unsurprisingly, this is not captured in our measure of size at birth, given that mothers determine the size of their children through comparison with their peers. All three

[4]Later in the analysis, we check the robustness of our findings using a slight variation in our outcome measures. We focus on five-year mortality and the ratio of weight over height, expressed in standard deviations and again calculated using the CGS method.

measures retain substantial volatility across the sample.

3 Methodological Approach

In related literature, soil evapotranspiration has been shown to influence crop productivity and lead to a sensitive loss in agricultural yield, depending on the characteristics of the cultivated land (Mendelsohn & Dinar 2003, Mendelsohn 2009, Malpede & Percoco 2023). In an area such as Sub-Saharan Africa, where most of the population lives in rural areas, agriculture has historically accounted for more than 50% of the gross domestic product (GDP) until the beginning of the 1990s (Diao et al. 2007), and agricultural productivity was vital. Even in 2020, in 34 out of the overall sample of countries available, agriculture, forestry, and fishing accounted for an average of 21.4% of the GDP, with some countries even exceeding 50%.[5] As such, aridity could affect human development by impacting agricultural productivity. This section quantifies the effect of the PET on health indicators that can serve as proxies for economic wellbeing.

We employ a linear fixed-effects panel data model, exploiting annual variations in weather conditions across cells to identify the effects of rainfall and the PET on the infant health indicators. We assembled an unbalanced panel dataset using the DHS to compare the impact of the PET and precipitations. The following equation is estimated:

$$Y_{i,m,q,c,t} = \sum_{r=0}^{1} \beta_{1,r} \mathrm{PRE}^{g}_{c,t^*-r} + \sum_{r=0}^{1} \beta_{2,r} \mathrm{PET}^{g}_{c,t^*-r} + \rho X_{i,m,c} + \sigma_{trend} + \phi_{q(t^*)} + \epsilon_{i,m,q,c,t} \tag{1}$$

where Y_{i,m,q,c,t^*} represents our measure of infant health for child i born from mother m in cluster q in cell c. $t \in \{t^*, \tilde{t}\}$ can represent either a child's year of birth (t^*) or the year of the

[5]Source: World Development Indicators, The World Bank Group: Agriculture, forestry, and fishing, value added (% of GDP). Data for Comoros, Cote d'Ivoire, and Mozambique, are available from 2019. Data for Zimbabwe are available from 2018.

interview ($\tilde{t}$). As the main regressand, we consider a measure of infant mortality (mortality at five years), the BMI in standard deviations recorded at the time of the interview, and the size at birth. We also assess the robustness of our results on a different yet comparable set of outcome variables (mortality at one year and the weight/height ratio).

$\text{PRE}_{c,t}$ and $\text{PET}_{c,t}$ are, respectively, the levels of precipitations and potential evapotranspiration at a child's year of birth (t^*) in cell c. In a variation of our specification, precipitation and PET levels are measured only using growing season months[6]. One-year lagged terms are included. Indeed, precipitation and PET levels may impact infant health by impinging on a mother's physical well-being or ability to take care of her child effectively. Thus, the study controls for weather conditions one year before childbirth, which more clearly isolates a potential effect on the mother that is not directly linked to the child's health.

In our baseline specification, the reference model for the binary outcome variable of infant mortality is a linear probability model (LPM). Using a linear estimator eases the interpretation of the coefficients and allows for immediate comparisons between the different specifications. Moreover, this choice is preferable given the ample set of fixed effects included in the regression.[7]

$X_{i,m,c}$ includes a set of covariates at the birth, woman, and grid level. Firstly, it is controlled for cell temperature, making sure our measure of the PET is not simply a proxy for raw temperatures. Second, since information was obtained on each child's year and month of birth, the year is taken as the time dimension, but the study accounts for the month of birth as a control. This approach allows for seasonality in precipitations and the PET to be smoothened out, while accounting for infant health may be influenced by exposure to weather conditions both in utero and immediately after childbirth. Moreover,

[6] Data on the growing season in each cell are available through the Global Agro-Ecological Zones (GAEZ) v4 database, provided by the Food and Agriculture Organization (FAO). The growing season is calculated using the beginning date of the earliest growing period and the total number of growing period days.

[7] While it is possible to estimate non-linear models through iterative algorithms, such as iteratively reweighted least squares (IRLS), some theoretical challenges exist regarding the existence of a solution to the algorithm or whether relevant parameters are identifiable.

the study controls for a mother's education and self-reported wealth index, the household's main source of water supply, and, again, for temperature in the grid at time t or in the corresponding growing season, for the same reason as above. In addition, $\phi_{c(t)}$ includes two sets of time and cluster fixed effects.[8] Lastly, the term σ_{trend} captures country trends.

Based on previous findings and the existing literature, the PET and precipitations are expected to have opposite effects on infant health. However, it is an empirical question whether controlling for PET in the estimation can also affect the impact of rainfall. To explore this hypothesis more precisely, we focus on the effect of precipitation *vis-à-vis* steadily increasing levels of soil water dispersion. To this end, we estimate the following regression:

$$Y_{i,m,q,c,t} = \sum_{r=0}^{1} \sum_{k=1}^{10} \gamma_{r,k} \mathrm{PRE}^{g}_{c,t^*-r} \times \mathbb{1}\left[\Delta \mathrm{PET}_{dec} = k\right]_c + \rho X_{i,m,c} + \sigma_{trend} + \phi_{q(t^*)} + \epsilon_{i,m,q,c,t} \tag{2}$$

where the term $\mathbb{1}\left[\Delta \mathrm{PET}_{dec} = k\right]$ indicates a binary indicator for the k-th decile in the sample distribution of ΔPET. The latter is a cell-specific measure that aims to capture the long-term variation in evapotranspiration. It is therefore calculated as the growth rate of the 10-year average PET between 1950 and 1959 and 2010 and 2019. The grid distribution of the long-term variation in the PET (Figure 4) shows how systematically arid cells do not necessarily identify areas in which the increase in the PET has been relatively more pronounced in the last decades, with the south-east regions, in particular, experiencing a steep increase in soil aridity. As such, Equation 2 more precisely captures the role of soil aridification, and not simply the overall soil quality, in driving a heterogeneous response of individuals to rainfall.

[8] Adding cluster FEs constitutes a more conservative approach than cell FEs, as a cell in our grid contains at least one cluster.

4 Results

Table 3 reports the estimates of the impacts of precipitations and the PET from Equation 1. As in the previous sections, in Panel [A], the weather variables are computed over the entire year of reference, while in Panel [B], they are calculated based on the corresponding growing season. The PET and precipitation variables are standardized to ease comparison. Column (1) in Panel [A] shows the effect of precipitations on infant mortality within five years from birth. One standard deviation increase in yearly rainfall levels reduces mortality by 0.48 percentage points. The same increase in the year before a child's birth is also expected to reduce the probability of child mortality by 0.33 percentage points, suggesting that part of the effect may be realized during pregnancy. All the coefficients are significant at the 99% level.

When adding the PET into the regression [Column (2)], the effect of precipitations drops to 0.26% in the year of birth and 0.13% one year before, with the latter coefficient losing significance. Conversely, one standard deviation increase in the (lagged) PET increases the expected mortality by (0.33%) 0.53%, hinting at a detrimental effect of increased PET on child mortality, although this is never statistically significant. When only focusing on the growing season months (Panel [B]), the coefficients estimated are similar in magnitude, and the same dynamics in the impact of rainfall are detected. The effect of the PET in the year of birth is now significant at the 95% level.

A similar story is depicted in the coefficients estimated for BMI, expressed in 100 standard deviations [Columns (4) and (5)]. When the precipitations alone are included, one standard deviation increase raises the BMI by 0.0746 standard deviations in the year of birth and 0.0331 standard deviations in the previous year. However, when the PET is brought into the regression, the combined effect of the two terms is almost halved. Conversely, PET appears to reduce BMI: one standard deviation increase in PET decreases BMI by 0.0961 standard deviations in the year of birth and by 0.1356 standard deviations in the previous year. The effect in the growing season months is now relatively lower for both regressors, but signs and significance remain considerably stable. One standard deviation increase in rainfall one year prior to and in the year of birth is estimated to boost the size at birth (between a 0.0046 and 0.036 probability of scoring one category higher across Panels [A] and [B]), while an increase in PET has an opposite effect, which in most cases is stronger in magnitude[9].

Second, we turn to the results of Equation 2. We plot the $\gamma_{r,k}$ estimates in Figure 5. Existing trends in the coefficients of our decile interactions can inform about the evolution of rainfall's effect when considering diverging land aridification levels. In Panel [A], we notice how precipitations in the year of birth and one year before tend to reduce child mortality, especially in areas where the PET has been stable in the last decades. In contrast, the effect gets rapidly offset in cells with severe aridification. The downward trend in interaction coefficients for BMI depicts a similar scenario: the positive effect of precipitation on children's body mass tends to dissipate in cells where the PET has been rising more significantly in the last 60 years. The results for size at birth are less transparent: while a similar initial drop is observed in the coefficients, these exhibit some mild spikes toward the end of the PET distribution.

Overall, this evidence supports multiple considerations regarding the relationship between precipitation and the PET. First, the effects of the two climate factors appear to frequently move in opposite directions. Second, whenever the PET is introduced in the analysis, the explanatory power of rainfall systematically decreases, which may hint that the role of rainfall may vary depending on the soil water retention ability. This is confirmed by noticing how the beneficial effect of increased rainfall before a child's birth dissipates significantly when looking at areas that have experienced stronger aridification.

In the remainder of the study, we argue how this could affect considerations regarding the future impact of climate change. Before that, we address some potential endogeneity

[9] In Table A.2, we show how the results observed for infant health are robust when substituting the outcome variables adopted with similar measures of health and mortality. Indeed, repeating the estimation with mortality within one year and the weight-to-height ratio in standard deviations returns remarkably similar coefficients, both in terms of magnitude and significance.

concerns by testing the robustness of the results.

5 Robustness Checks

5.1 Migration

Climate change may push people to move toward more productive regions based on some non-random characteristics. In turn, the results could be shaded by an endogenous adaptation mechanism working through migration flows. To deal with this potential threat to our estimates, we drop from the sample those women reporting with certainty having lived in the same place of residence for less than 15 years at the time of the interviews. Estimates of Equation 1 are then repeated for the sub-sample of long-settled individuals, who are assumed to be fully affected by the impact of aridification. While observing significant discrepancies with the study's baseline estimates would support the hypothesis that individuals tend to react to climate change by relocating endogenously, thus casting doubt on the study's ability to estimate the true role of the PET, the results reported in the Appendix (Table A.2) seem to alleviate this concern. The reduction in sample size translates into only a marginal loss of significance in the estimates, while the order of magnitude and the direction of the effect appear to be roughly unchanged. Such evidence supports the belief that the estimate coefficients can capture the actual effect of yearly rainfall and aridity on child health.

5.2 Heterogeneity across Urban and Rural Clusters

Existing literature on the health and economic impacts of rainfall usually identifies in the agricultural sector a major — if not the only — mechanism directly or indirectly connecting it to individual wellbeing. Under this hypothesis, the adverse effects of soil aridification could be more directly relevant to individuals whose primary income and food resources lie in the primary sector. As a result, the effect estimated with our empirical strategy could be strongly heterogeneous concerning this aspect, with relevant implications for future social and environmental conservation policies. As individuals are expected to move away from rural areas (Cohen 2006, Güneralp et al. 2020), strong differences in the impact of soil aridification in urbanized cities could justify prioritizing measures of urban development over soil quality conservation in the short term. To explore this scenario, we use the information provided by the DHS dataset, which classifies clusters as urban and rural. This classification is country-specific, and the assessment can consider multiple criteria, such as the availability of electricity or piped water and access to healthcare, schools, and transportation. To explore the difference between these two settings, we estimate Equation 1 separately for urban and rural clusters.[10]

The results are reported in Table A.3. Despite the sample being almost equally divided between rural (60%) and urban clusters (40%), the higher number of children per woman in rural areas leads to more than 70% of births occurring in rural households. Looking at the coefficients for precipitation alone in the two panels, the effect of rainfall appears systematically stronger for individuals living in rural clusters. Still, it persists in sizable magnitude in the urban sample, with the only exception of size at birth. In both the urban and rural clusters, the same drop in coefficient magnitude is observed with the introduction of the PET, which systematically counteracts the beneficial effects of higher rainfall, despite having reduced statistical significance. These results reveal how individuals in urbanized areas are still susceptible to similar trends due to aridification and reinforce the importance of tapering future soil quality loss.

6 Conclusion

In this paper, we have studied the combined effect of rainfall and of the capacity of the soil to retain water on child health in Sub-Saharan Africa, finding that the evapotranspiration

[10] For this exercise, we focus only on yearly values of precipitation and the PET. Using growing season definitions of our climate variables leads to comparable results.

of the soil negatively affects infant health, reducing the size of children at birth and their BMIs. This result is particularly important when we consider that, since 1980, PET levels in our sample have increased by an average of 0.928 mm per year. Our estimates suggest how progressive soil aridification in Africa has been, since then, directly responsible for a yearly increase in mortality of around 0.15% and a decrease in BMI and size at birth of 0.028 standard deviations and 0.022 units[11]. These numbers may be even higher when thinking about the impact of increasing PET on the beneficial effect of precipitations. Furthermore, most ESM projections foresee a peak in emissions, which will fuel aridification even further in the upcoming years: different representative pathway scenarios (RCPs) depict an increase in the yearly PET from a minimum of 30 mm/year to a maximum of over 300 mm/year.[12]. Under intermediate scenarios (i.e. RCP 4.5), African households may experience up to 3.8 mm average increase in PET, resulting in an increase in child mortality by 0.6% a year and a BMI loss and infant size loss of 0.119 standard deviations and 0.093 units, respectively. These figures escalate even more under worst-case scenarios (i.e. RCP 8.5, 1.4% increase in mortality, 0.256 and 0.2 losses in BMI size at birth loss). The results presented in this study call for policy actions to tackle soil aridification, which may generate benefits not only in terms of land productivity but also in terms of child health.

[11]These calculations make use of the unstandardized coefficients obtained from estimating Equation 1, which are reported in Table A.4 in the Appendix.

[12]Time series at $0.5^{\circ} \times 0.5^{\circ}$ grid resolutions are available for five ESMs: GFDL-ESM2M, HadGEM2-ES, IPSL-CM54-LR, MIROC-ESM-CHEM, and NorESM1-M. These are all part of the Coupled Model Intercomparison Project Phase 5 (CMIP5). We report the average projections limited to our grid in Table A.5 More information about the nature and specificities of these models can be found in Noce et al. (2020).

References

Aber, J. L., Bennett, N. G., Conley, D. C. & Li, J. (1997), 'The effects of poverty on child health and development', *Annual Review of Public Health* **18**(1), 463–483.

Allen, R. G., Pereira, L. S., Raes, D., Smith, M. et al. (1998), 'Crop evapotranspiration-guidelines for computing crop water requirements – FAO irrigation and drainage paper 56', *FAO, Rome* **300**(9), D05109.

Allen, R., Smith, M., Perrier, A. & Pereira, L. S. (1994), 'An update for the definition of reference evapotranspiration', *ICID bulletin* **43**(2), 1–34.

Barrios, S., Bertinelli, L. & Strobl, E. (2010), 'Trends in rainfall and economic growth in Africa: A neglected cause of the African growth tragedy', *The Review of Economics and Statistics* **92**(2), 350–366.

Benshaul-Tolonen, A. (2018), 'Local industrial shocks and infant mortality', *The Economic Journal* **129**(620), 1561–1592.
URL: *https://doi.org/10.1111/ecoj.12625*

Bharadwaj, P., Fenske, J., Kala, N. & Mirza, R. A. (2020), 'The green revolution and infant mortality in India', *Journal of Health Economics* **71**, 102314.

Björkman-Nyqvist, M. (2013), 'Income shocks and gender gaps in education: Evidence from Uganda', *Journal of Development Economics* **105**, 237–253.
URL: *https://www.sciencedirect.com/science/article/pii/S0304387813001120*

Brückner, M. & Ciccone, A. (2011), 'Rain and the democratic window of opportunity', *Econometrica* **79**(3), 923–947.

Burgess, R. & Donaldson, D. (2010), 'Can openness mitigate the effects of weather shocks? evidence from india's famine era', *American Economic Review* **100**(2), 449–453.

Burke, M., Hsiang, S. M. & Miguel, E. (2015), 'Global non-linear effect of temperature on economic production', *Nature* **527**(7577), 235–239.

Cherlet, M., Hutchinson, C., Reynolds, J., Hill, J., Sommer, S. & Von Maltitz, G. (2018), *World atlas of desertification: Rethinking land degradation and sustainable land management*, Publications Office of the European Union.

Cohen, B. (2006), 'Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability', *Technology in Society* **28**(1–2), 63–80.

Couttenier, M. & Soubeyran, R. (2014), 'Drought and civil war in sub-saharan africa', *The Economic Journal* **124**(575), 201–244.

Damania, R. (2020), 'The economics of water scarcity and variability', *Oxford Review of Economic Policy* **36**(1), 24–44.

Dell, M., Jones, B. F. & Olken, B. A. (2008), Climate change and economic growth: Evidence from the last half century, Technical report, National Bureau of Economic Research.

Dell, M., Jones, B. F. & Olken, B. A. (2012), 'Temperature shocks and economic growth: Evidence from the last half century', *American Economic Journal: Macroeconomics* **4**(3), 66–95.

Dell, M., Jones, B. F. & Olken, B. A. (2014), 'What do we learn from the weather? The new climate-economy literature', *Journal of Economic literature* **52**(3), 740–798.

Diao, X., Hazell, P. B., Resnick, D. & Thurlow, J. (2007), *The role of agriculture in development: Implications for Sub-Saharan Africa*, Vol. 153, Intl Food Policy Res Inst.

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A. & Maher, N. (2016), 'More extreme precipitation in the world's dry and wet regions', *Nature Climate Change* **6**(5), 508–513.

Ekström, M., Jones, P. D., Fowler, H. J., Lenderink, G., Buishand, T. A. & Conway, D. (2007), 'Regional climate model data used within the SWURVE project: Projected changes in seasonal patterns and estimation of pet', *Hydrology and Earth System Sciences* **11**(3), 1069–1083.
URL: *https://hess.copernicus.org/articles/11/1069/2007/*

Emediegwu, L. E., Wossink, A. & Hall, A. (2022), 'The impacts of climate change on agriculture in Sub-Saharan Africa: a spatial panel data approach', *World Development* **158**, 105967.

Fischer, G., Nachtergaele, F. O., van Velthuizen, H., Chiozza, F., Francheschini, G., Henry, M., Muchoney, D. & Tramberend, S. (2021), 'Global agro-ecological zones (GAEZ v4)–model documentation'.

Gornall, J., Betts, R., Burke, E., Clark, R., Camp, J., Willett, K. & Wiltshire, A. (2010), 'Implications of climate change for agricultural productivity in the early twenty-first century', *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**(1554), 2973–2989.

Güneralp, B., Reba, M., Hales, B. U., Wentz, E. A. & Seto, K. C. (2020), 'Trends in urban land expansion, density, and land transitions from 1970 to 2010: A global synthesis', *Environmental Research Letters* **15**(4), 044015.

Harari, M. & Ferrara, E. L. (2018), 'Conflict, climate, and cells: A disaggregated analysis', *Review of Economics and Statistics* **100**(4), 594–608.

Harris, I., Osborn, T. J., Jones, P. & Lister, D. (2020), 'Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset', *Scientific data* **7**(1), 1–18.

Held, I. M. & Soden, B. J. (2006), 'Robust responses of the hydrological cycle to global warming', *Journal of climate* **19**(21), 5686–5699.

Jayachandran, S. (2006), 'Selling labor low: Wage responses to productivity shocks in developing countries', *Journal of Political Economy* **114**(3), 538–575.

Kirtman, B., Power, S. B., Adedoyin, A. J., Boer, G. J., Bojariu, R., Camilloni, I., Doblas-Reyes, F., Fiore, A. M., Kimoto, M., Meehl, G. et al. (2013), 'Near-term climate change: Projections and predictability'.

Kudamatsu, M., Persson, T. & Strömberg, D. (2012), Weather and infant mortality in Africa, Technical report, CEPR Discussion Papers N. DP9222. Available at: `https://ssrn.com/abstract=2210191`.

Le, K. & Nguyen, M. (2021), 'In-utero exposure to rainfall variability and early childhood health', *World Development* **144**, 105485.

Maccini, S. & Yang, D. (2009), 'Under the weather: Health, schooling, and economic consequences of early-life rainfall', *American Economic Review* **99**(3), 1006–26.

Malpede, M. & Percoco, M. (2023), 'Aridification, precipitation, and crop productivity: evidence from the aridity index', *European Review of Agricultural Economics* p. jbad006.

Mary, S. (2022), 'Dams mitigate the effect of rainfall shocks on hindus-muslims riots', *World Development* **150**, 105731.

Mendelsohn, R. (2009), 'The impact of climate change on agriculture in developing countries', *Journal of Natural Resources Policy Research* **1**(1), 5–19.

Mendelsohn, R. & Dinar, A. (2003), 'Climate, water, and agriculture', *Land economics* **79**(3), 328–341.

Noce, S., Caporaso, L. & Santini, M. (2020), 'A new global dataset of bioclimatic indicators', *Scientific data* **7**(1), 1–12.

Peri, G. & Sasahara, A. (2019), The impact of global warming on rural-urban migrations: Evidence from global big data, Technical report, National Bureau of Economic Research.

Ponnusamy, S. (2022), 'Rainfall shocks, child mortality, and water infrastructure', *Health Economics* **31**(7), 1317–1338.

Rabassa, M., Skoufias, E. & Jacoby, H. (2014), 'Weather and child health in rural Nigeria', *Journal of African Economies* **23**(4), 464–492.

Randell, H., Jiang, C., Liang, X.-Z., Murtugudde, R. & Sapkota, A. (2021), 'Food insecurity and compound environmental shocks in Nepal: Implications for a changing climate', *World development* **145**, 105511.

Rind, D., Goldberg, R., Hansen, J., Rosenzweig, C. & Ruedy, R. (1990), 'Potential evapotranspiration and the likelihood of future drought', *Journal of Geophysical Research: Atmospheres* **95**(D7), 9983–10004.

Rocha, R. & Soares, R. R. (2015), 'Water scarcity and birth outcomes in the Brazilian semiarid', *Journal of Development Economics* **112**, 72–91.

URL: *https://www.sciencedirect.com/science/article/pii/S0304387814001096*

Rose, E. (1999), 'Consumption smoothing and excess female mortality in rural India', *The Review of Economics and Statistics* **81**(1), 41–49.

URL: *http://www.jstor.org/stable/2646784*

Sarsons, H. (2015), 'Rainfall and conflict: A cautionary tale', *Journal of development Economics* **115**, 62–72.

Sivadasan, J. & Xu, W. (2021), 'Missing women in india: Gender-specific effects of early-life rainfall shocks', *World Development* **148**, 105652.

Vicente-Serrano, S. M., Beguería, S. & López-Moreno, J. I. (2010), 'A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index', *Journal of climate* **23**(7), 1696–1718.

Zhang, P., Zhang, J. & Chen, M. (2017), 'Economic impacts of climate change on agriculture: The importance of additional climatic variables other than temperature and precipitation', *Journal of Environmental Economics and Management* **83**, 8–31.

Tables

Table 1: Summary statistics – Climate variables

	1951–1965	1966–1980	1981–1995	1996–2010	2010–2019	Avg growth
Prec	1064.9	1032.9	965.6	1006.1	1016.9	-0.3%
	(545.99)	(527.67)	(523.01)	(520.64)	(516.55)	
	[1.67;3,194.77]	[2.44;3,110.26]	[2.09;3,107.39]	[2.07;3,102.73]	[0.89;2,959.36]	
PET	1402.1	1396.6	1412.0	1431.9	1435.7	0.07%
	(315.89)	(314.28)	(310.26)	(318.03)	(310.12)	
	[809.00;2,664.20]	[811.60;2,680.00]	[823.20;2,704.00]	[822.20;2,794.80]	[836.67;2,713.33]	
Temp	23.231	23.243	23.555	23.905	23.993	0.07%
	(3.58)	(3.61)	(3.58)	(3.58)	(3.56)	
	[8.76;29.68]	[8.90;29.¡93]	[9.35;30.15]	[9.51;30.65]	[9.99;30.68]	

Correlations

	Prec	PET	Prec	PET	Prec	PET	Prec	PET	Prec	PET
PET	-0.58	.	-0.62	.	-0.63	.	-0.63	.	-0.61	.
Tmp	0.21	0.39	0.13	0.41	0.13	0.41	0.13	0.42	0.13	0.40

Notes: summary statistics are shown for a sample of 4,052 grid cells. Precipitations and PET show the total millimetres of rain and water lost by the soil in the year, averaged throughout the indicated period, while temperature is a yearly average (°C). Standard deviations are reported in parentheses; minimum and maximum values are reported in brackets. The average growth column is calculated as $\Delta = 1/T \sum_t^T (X_{t+1} - X_t)/X_t$, where T is the entire sample period (68 years).

Table 2: Summary statistics – Mothers and children

	All sample	Wave III	Wave IV	Wave V	Wave VI	Wave VII	Min	Max
*Sample characteristics**								
N. countries	32	3	14	13	28	12		
N. clusters	36842	859	7885	6433	16010	5655		
Mothers	468873	11277	120289	110359	186949	89867		
Births	2358849	46343	511967	501581	979807	319151		
Births u5	697710	12237	132073	149913	296656	106831		
				Mean (sd)				
Mother's characteristics								
Age	35.254	35.630	36.041	35.083	35.061	34.796	15	50
	(8.07)	(8.02)	(8.02)	(8.15)	(8.03)	(8.05)		
Education	3.451	2.469	2.852	3.806	3.558	3.669	0	27
	(4.18)	(3.99)	(4.12)	(4.00)	(4.23)	(4.29)		
Wealth**	2.796	.	2.822	2.804	2.790	2.781	1	5
	(1.40)	.	(1.42)	(1.39)	(1.40)	(1.40)		
Rural household	0.279	0.263	0.266	0.230	0.296	0.330	0	1
	(0.45)	(0.44)	(0.44)	(0.42)	(0.46)	(0.47)		
Births	5.402	5.724	5.673	5.694	5.226	5.004	1	19
	(2.75)	(2.67)	(2.78)	(2.81)	(2.70)	(2.65)		
Age at first birth	18.784	18.225	18.915	18.402	18.883	18.976	3	47
	(3.64)	(3.52)	(3.57)	(3.51)	(3.72)	(3.66)		
Children under 5	1.550	1.694	1.386	1.546	1.623	1.576	0	24
	(1.32)	(1.39)	(1.28)	(1.22)	(1.40)	(1.26)		
N. of living children	4.560	4.494	4.584	4.652	4.543	4.438	0	16
	(2.24)	(2.19)	(2.24)	(2.29)	(2.22)	(2.20)		
Infant mortality								
1 year	0.084	0.120	0.103	0.097	0.073	0.062	0	1
	(0.28)	(0.33)	(0.30)	(0.30)	(0.26)	(0.24)		
5 years	0.120	0.172	0.144	0.140	0.105	0.089	0	1
	(0.33)	(0.38)	(0.35)	(0.35)	(0.31)	(0.28)		
1 year – boys	0.090	0.126	0.110	0.104	0.079	0.068	0	1
	(0.29)	(0.33)	(0.31)	(0.30)	(0.27)	(0.25)		
5 years – boys	0.127	0.177	0.151	0.148	0.111	0.096	0	1
	(0.33)	(0.38)	(0.36)	(0.35)	(0.31)	(0.30)		
Infant nutrition–health								
BMI (SD)***	0.57	.	.	1.82	1.22	-2.72	-500	500
	(142.94)	.	.	(156.37)	(144.29)	(119.50)		
Weight/Height (SD)***	-12.38	.	.	-12.88	-11.83	-13.20	0	99.8
	(140.44)			(152.18)	(140.38)	(124.27)		
Size at birth (cat) ****	2.22	2.45	2.08	2.32	2.23	2.25	0	4
	(0.99)	(0.99)	(0.95)	(0.98)	(0.99)	(1.00)		

Notes: the DHS surveys employed were conducted between 1992 and 2019.
* The number of countries, clusters, mothers and generic births refers to the data available from the Individual Recode DHS survey. The number of births under 5 refers to the data available from the Child Recode DHS survey.
** Wealth is a categorical index ranging from 1 (very poor) to 5 (very rich).
*** The measures are presented with two implied decimal places; the actual value is obtained by dividing the variable by 100.
**** Size at birth is a categorical variable ranging from 0 (very small) to 4 (very large).

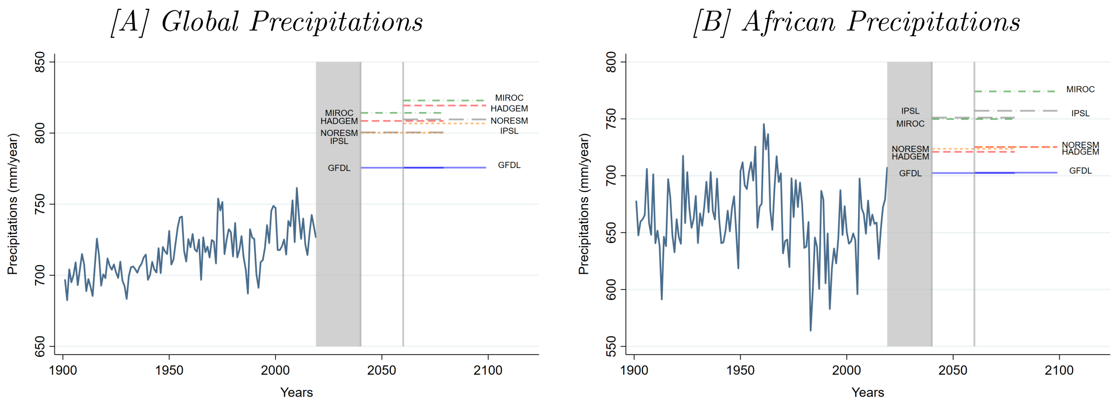
Table 3: Impact of precipitations and the PET on infant health

	Mort (5y)		**BMI (SD)**		**Size at birth**	
	(1)	(2)	(3)	(4)	(5)	(6)
			[A] Yearly values			
Prec_t	-0.0048***	-0.0026***	7.4652***	5.1982***	0.0363***	0.0222***
	(0.0009)	(0.0010)	(1.2166)	(1.2941)	(0.0064)	(0.0067)
$\text{Prec}_{(t-1)}$	-0.0033***	-0.0013	3.3141***	1.2999	0.0308***	0.0172***
	(0.0009)	(0.0010)	(1.1741)	(1.2384)	(0.0059)	(0.0063)
PET_t		0.0053		-9.6136*		0.0139
		(0.0045)		(5.1651)		(0.0266)
$\text{PET}_{(t-1)}$		0.0033		-1.3564		-0.0754***
		(0.0044)		(5.1662)		(0.0266)
Observations	2,059,690	2,059,690	266,469	266,469	518,591	518,591
R^2	0.0601	0.0601	0.1999	0.2000	0.1331	0.1332
			[B] Growing season			
Prec GS $_t$	-0.0027***	-0.0017*	3.4957***	2.4391**	0.0173***	0.0080
	(0.0009)	(0.0009)	(1.0977)	(1.1367)	(0.0059)	(0.0062)
Prec GS $_{(t-1)}$	-0.0034***	-0.0027***	3.2944***	2.2290**	0.0117**	0.0046
	(0.0008)	(0.0009)	(1.0955)	(1.1369)	(0.0058)	(0.0059)
PET GS $_t$		0.0053**		-4.5811*		-0.0416***
		(0.0025)		(2.7278)		(0.0147)
PET GS $_{(t-1)}$		0.0006		-4.5907*		-0.0137
		(0.0024)		(2.7214)		(0.0146)
Observations	1,721,384	1,721,384	227,786	227,786	432,729	432,729
R^2	0.0598	0.0598	0.1833	0.1835	0.1287	0.1289
Controls	Long	Long	Long	Long	Long	Long
Year FE	Y	Y	Y	Y	Y	Y
Cluster FE	Y	Y	Y	Y	Y	Y
Country Trends	Y	Y	Y	Y	Y	Y

Notes: the table presents the estimates of vectors $\beta_\mathbf{1}$ and $\beta_\mathbf{2}$. Long controls include the woman's education in single years, wealth index of the household, kid's month of birth, main source of water, and cell temperature at times t and $t-1$. Columns (1) and (2) are based on a sample of 33 countries and 22,757 clusters. Columns (3) to (6) are based on a sample of 34 countries and 22,909 clusters. In Panel [A], precipitations and the PET are computed over the entire year; in Panel [B], only the growing season months are considered. Robust standard errors are clustered at the DHS cluster level, with significance levels at 10, 5, and 1 percent.

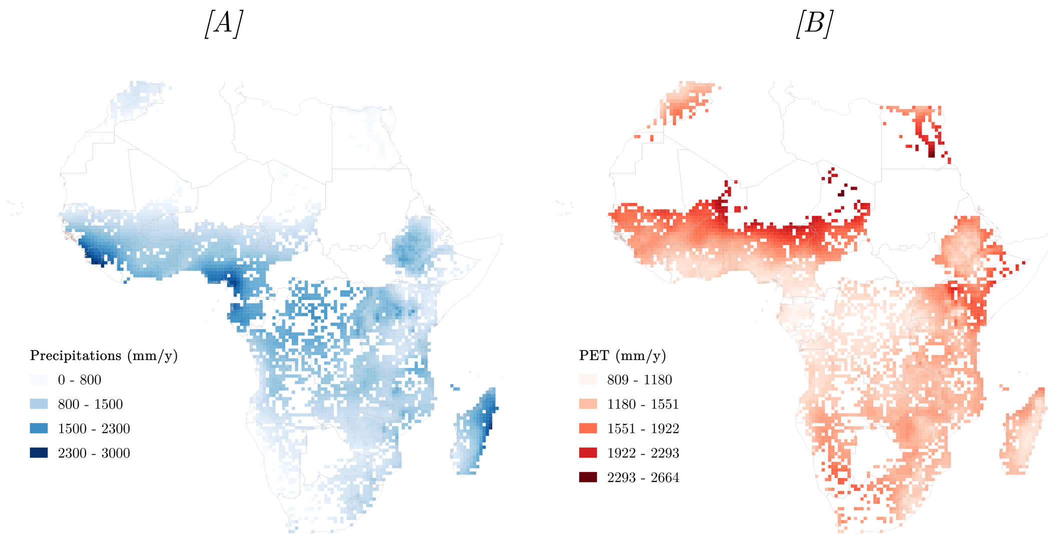
Figures

Figure 1: Historical trend and future projections in yearly precipitation, 1901–2099



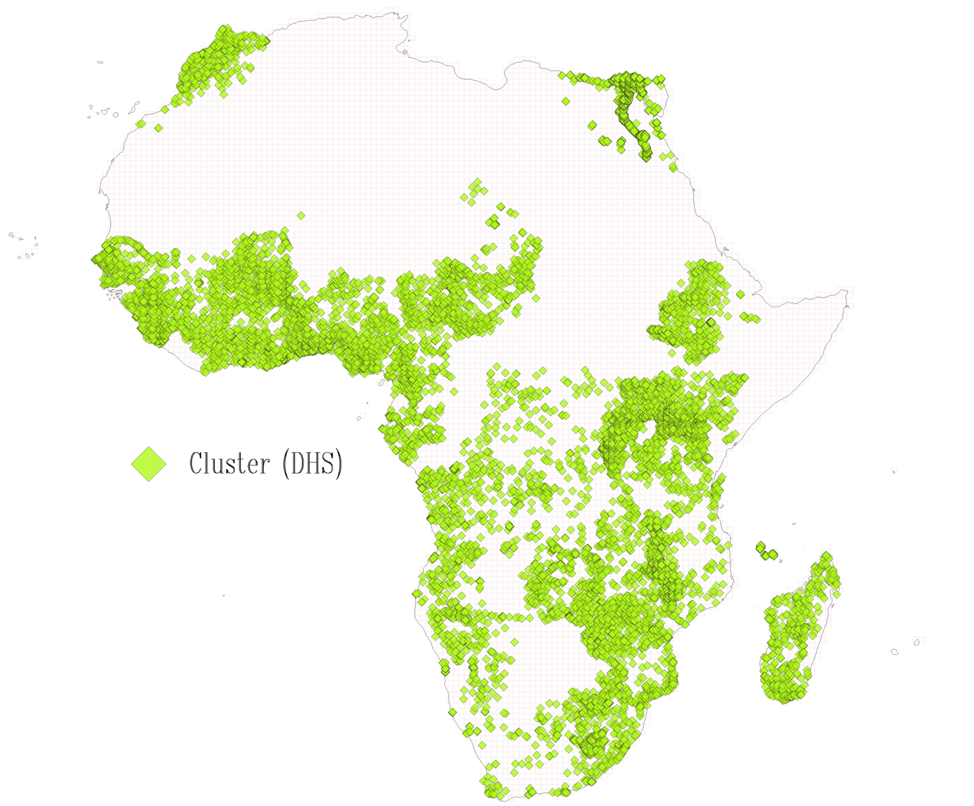
Notes: the figure depicts the trend in yearly precipitations (mm/year) starting form 1901. The series between 1901 and 2019 is computed from the CRU TS4.04 dataset. Projections for two time horizons (2040–2079 and 2060–2099) are accessed from five commonly employed earth system models (ESMs): GFDL-ESM2M, HadGEM2-ES, IPSL-CM54-LR, MIROC-ESM-CHEM, and NorESM1-M. Bias-corrected projections are plotted under the representative concentration pathway (RCP) 4.5, a greenhouse gas concentration trajectory, which possibly constitutes the most probable baseline scenario by taking into account the exhaustible character of non-renewable fuels. Panel [A] plots average global precipitations levels; Panel [B] focuses on the African continent. Source: CMCC-BioClimInd dataset.

Figure 2: Geographical variation in precipitations and the PET (1951-2018)



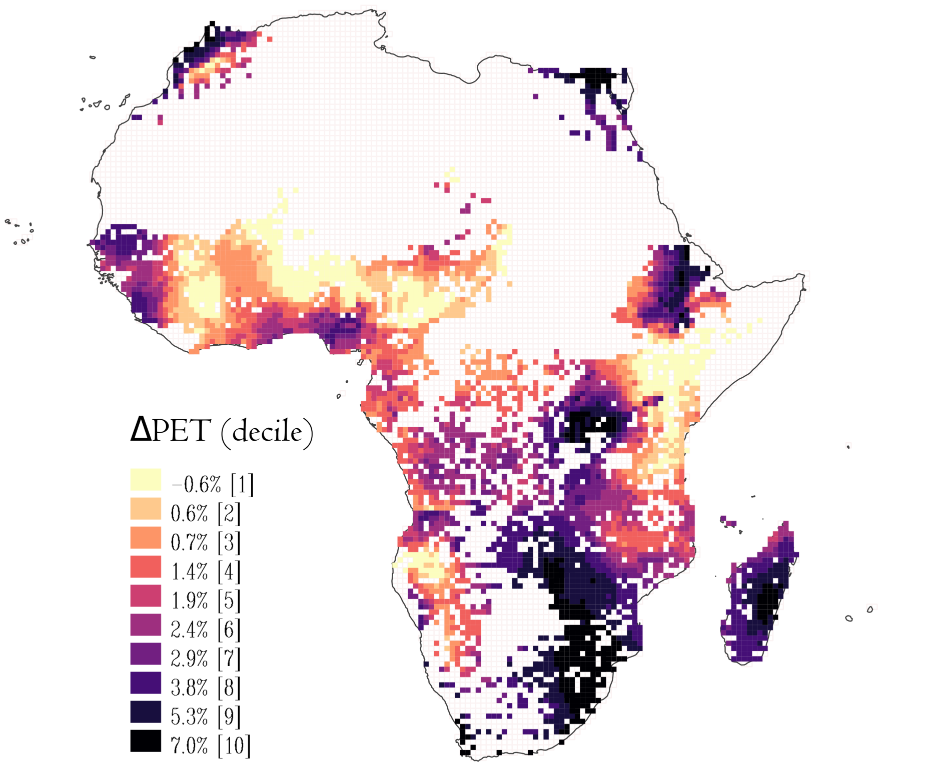
Notes: Panel [A] depicts precipitations (mm/year) over the sample grid. Panel [B] shows the PET (mm/year) for the same cells. In Panel [A], lighter cells identify areas of scarce precipitations. In Panel [B], darker cells identify arid regions.

Figure 3: Sampled DHS clusters



Notes: the image reports all the clusters appearing at least once in the sample. The sample includes up to 22,756 unique clusters, with an average (standard deviation) number of mothers per cluster-survey year of 53.3 (34.1). Geolocation displacement can be up to 2 km for urban and 10 km for rural clusters.

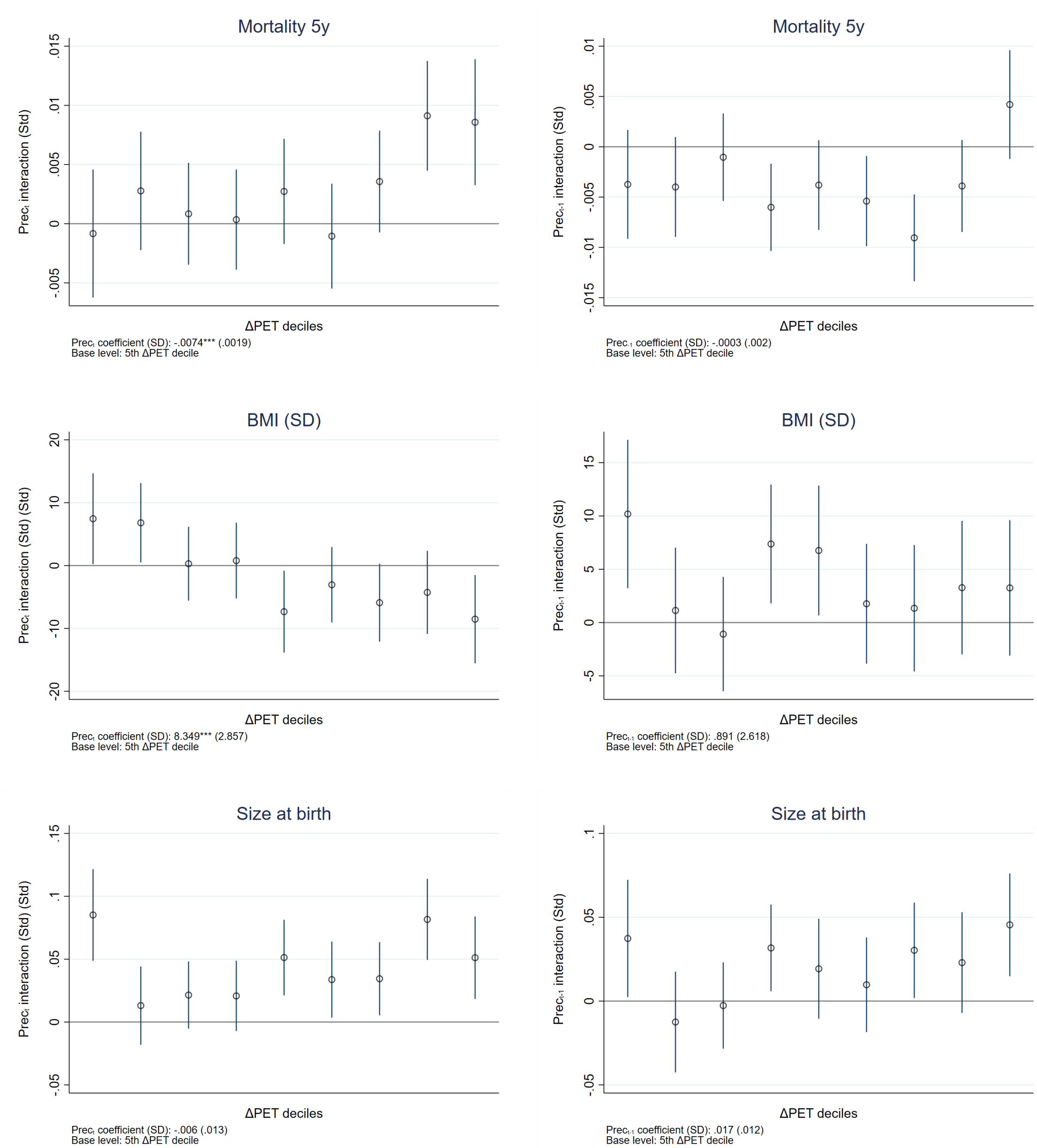
Figure 4: PET variation deciles–Cell distribution



Notes: the image plots the sample grid classified by deciles in the distribution of PET long-term variation. Long-term variation is calculated as the growth rate of the 10-year average PET in the decades between 1950 and 1959 and 2010 and 2019. The legend reports the average percentage variation in the PET per decile (in parentheses).

Figure 5: Effect of rainfall by deciles of land aridification

[A] $\sum_k \gamma_{0,k}$ coefficients (PRE_t) [B] $\sum_k \gamma_{1,k}$ coefficients (PRE_{t-1})



Notes: the image reports $\gamma_{r,k}$ (with $r \in \{0, 1\}$ and $k \in \mathbb{N} \cap [1, 10]$) from Equation 2. Panel [A] reports interaction terms with precipitation at time t. In Panel [B], the same decile indicators are interacted with lagged precipitation. The sample distribution of ΔPET is found in Figure 4. 95% confidence intervals clustered at DHS cluster level are reported.

A.1 Data Appendix

A.1.1 Computation of the Potential Evapotranspiration

We access a measure for the potential evapotranspiration from the CRU TS.04 dataset. This is calculated using a modeling scheme based on climate simulations developed by the Hadley Centre (HadRM3H). A full description of the relevant regional climate models can be found in Ekström et al. (2007). Here, we report a short summary explaining the computation of the PET.

The estimates for the PET are provided using a variant of the Penman-Monteith method, as proposed by the FAO. This indicator is addressed as *potential* since it employs a grass reference crop. [13] The PET is computed according to the following equation:

$$PET = \frac{0.408\Delta\left(R_n - G\right) + \gamma + \frac{900}{T+273.16}U_2\left(e_a - e_d\right)}{\Delta + \gamma\left(1 + 0.34U_2\right)} \tag{3}$$

where R_n represents the net radiation at the crop surface (MJ m^{-2} per day), G is soil heat flux (MJ m^{-2} per day), T is the mean temperature, U is wind speed ($\frac{m}{s}$), $(e_a - e_d)$ and Δ are the vapour pressure deficit and the relative slope of the vapour pressure curve ($\frac{kPa}{\circ C}$) respectively, and γ is a psychrometric constant. While wind speed and temperature are direct outputs from the HadRM3H, the other constants in the formula are calculated using the model data.

As Equation 3 suggests, while the temperature is indeed relevant in the computation of the PET, which justify an average positive correlation of around 40% between PET and temperature, it is only part of the story. As such, by controlling for the yearly average temperature in our main specifications, we are able to isolate the effect of soil water availability without capturing potential noise coming from heat volatility.

A.1.2 Datasets and Variables Description

Precipitations (PRE): total, mm/year

Potential evapotranspiration (PET): total, mm/year. See Section A.1.2 for the computational details

[13] For the original contribution on this computation, see Allen et al. (1994).

Temperature: °C, average monthly value at 2 m altitude.

Growing Season: growing season months are calculated using the start date of the earliest growing period (day of year) for the time period 1981–2010 and the total number of growing period days. Both measures are accessed through the GAEZ v4 dataset, which employs the climate data source HadGEM2-ES. More information can be found in Fischer et al. (2021).

Infant mortality (1–5 years): binary indicator computed using the mother-reported time of death after birth (in months).

BMI (SD): Body mass index, defined as the weight in kilograms divided by the square of the height in meters (W/H^2), and then expressed in standard deviations.

Weight/height: Weight for height standard deviations from the reference median based on the DHS reference standard.

Education: Highest year of education to give the years of education completed.

Wealth index: a composite measure of a household's cumulative living standard, calculated using easy-to-collect data on a household's ownership of selected assets, such as televisions and bicycles and materials used for housing construction. Generated by principal components analysis, the index separates individual households into five wealth quintiles.

Source of drinking water: Main source of drinking water for members of the household (major categories).

A.1.3 CRU TS.04 Choice and Comfront with Alternative Datasets

When it comes to environmental studies, researchers in need of high-frequency data on weather and climate conditions have more than one alternative. Researchers have compared and highlighted the peculiarities of these different sources. However, no rule of thumb exists to guide them through the adoption of one particular dataset.

In their paper on weather shocks, malaria, and child mortality, Kudamatsu et al. (2012) access observations on monthly rainfall through the 45-Year European Centre for Medium-Range Weather

Forecasts (ECMWF) Re-Analysis (ERA-40) data archive. The authors justify their choice by claiming its superiority over the more well-known CRU dataset. Their main argument relies on the fact that rainfall gauge data in Africa lack the necessary quality and show bias in arid and semi-arid areas, where departures from standard seasonal fluctuations are more pronounced. A similar argument is provided by Harari & Ferrara (2018) to justify the adoption of the ERA-40. Other authors have instead deemed gauge data suitable for the purpose of their studies, and have, thus, turned at the CRU dataset, usually in its previous versions (Vicente-Serrano et al. 2010, Couttenier & Soubeyran 2014).

While the concerns about gauge data are surely legitimate, significant drawbacks are also implied by the choice of reanalysis data. First, the ERA-40 dataset is provided at more than twice the resolution of the CRU TS.04 one, at 1.25×1.25 degress (roughly 139 km $\times$ 139 km), which is a significant loss in terms of spatial variation. As our sample is an unbalanced panel, variation across grid cells is of great importance, and, as such, this could impinge on the detection of an effect of precipitations and the PET on agricultural productivity and infant health. Aanother alternative available to researchers is the ERA-5 dataset, in which near-surface meteorological variables have been re-gridded to a half-degree resolution. Yet, in addition to using monthly-scale bias corrections still based on CRU data, this dataset is only available from 1980.

Second, reanalysis relies on a variety of sources, including weather stations, ships, aircraft, and satellites. To provide the corresponding weather measures, recorded data are analyzed through an atmospheric circulation model (IFS CY23r4). Compared to gauge data, this augments the risk of measurement error.

It is indeed true that CRU stations are partially dispersed in Sub-Saharan Africa, and, since they cannot provide full direct coverage, the resulting data rely on interpolation. However, the data on station location and resulting cover contained in Harris et al. (2020), in which stations are included if they contribute at least 75% of observations in the decade, show that the problem of loss of variability due to interpolation may be more of a concern in the areas of scarce coverage (in this case, historical data would have a greater role in filling in for missing observations). In

Figure A.1, we plot the spatial distribution of the clusters available from the DHS together with the spatial coverage of the CRU stations in the decades 1970-79 and 2000-09. Coverage in the CRU dataset is defined as an area that experiences the direct measurement of at least 75% of all potential observations in the decade. We notice that most of the clusters in the sample are actually within the declared coverage, which provides reassurance, at least regarding the probability of errors generated by the stations. As we use the area identified by the DHS cluster to run the analysis on crop productivity, a similar reasoning applies to crop yield observations.

This evidence ultimately helps reduce the concern that our data may not properly capture the variability in precipitations and PET. This, in turn, is essential to our research strategy and contributes to justifying the choice of the CRU TS.04 dataset.

A. Tables

Table A.1: Impact of precipitations and the PET on infant health – alternative measures

	Mort (1y)		**W/H ratio (SD)**	
	(1)	(2)	(3)	(4)
	[A] Yearly values			
Prec_t	-0.0025***	-0.0020**	7.9419***	5.4943***
	(0.0008)	(0.0008)	(1.1803)	(1.2505)
$\text{Prec}_{(t-1)}$	-0.0010	-0.0005	3.1379***	0.8670
	(0.0008)	(0.0008)	(1.1443)	(1.2045)
PET_t		0.0022		-6.9669
		(0.0039)		(5.0209)
$\text{PET}_{(t-1)}$		-0.0001		-5.1389
		(0.0038)		(4.9873)
Observations	2,059,690	2,059,690	267,083	267,083
R^2	0.0391	0.0391	0.1964	0.1966
	[B] Growing season			
Prec GS $_t$	-0.0014*	-0.0012*	3.0697***	1.9820*
	(0.0007)	(0.0007)	(1.0663)	(1.1028)
Prec GS $_{(t-1)}$	-0.0017**	-0.0017**	2.9896***	1.7785
	(0.0007)	(0.0007)	(1.0694)	(1.1072)
PET GS $_t$		0.0014		-4.1160
		(0.0021)		(2.7023)
PET GS $_{(t-1)}$		-0.0005		-5.8206**
		(0.0021)		(2.6913)
Observations	1,721,384	1,721,384	228,387	228,387
R^2	0.0386	0.0386	0.1811	0.1812
Controls	Long	Long	Long	Long
Year FE	Y	Y	Y	Y
Cluster FE	Y	Y	Y	Y
Country Trends	Y	Y	Y	Y

Notes: the table presents the estimates of β_1 and β_2 for comparable outcome variables (infant mortality and weight/height ratio). Long controls include the woman's education in single years, wealth index of the household, child's month of birth, main source of water, and cell raw temperature in t and $t-1$. In Panel [A], precipitations and PET are computed over the entire year; in Panel [B], only the growing season months are considered. Robust standard errors are clustered at the DHS cluster level, with significance levels at 10, 5, and 1 percent.

Table A.2: Impact of precipitations and the PET on infant health - long-settled sample

	Mort (5y)		BMI (SD)		Size at birth	
	(1)	(2)	(3)	(4)	(5)	(6)
[A] Yearly values						
Prec_t	-0.0047***	-0.0023**	7.3400***	5.3897***	0.0336***	0.0169**
	(0.0010)	(0.0011)	(1.4110)	(1.5072)	(0.0077)	(0.0083)
$\text{Prec}_{(t-1)}$	-0.0037***	-0.0014	4.0424***	2.2763	0.0321***	0.0166**
	(0.0010)	(0.0011)	(1.3594)	(1.4348)	(0.0071)	(0.0076)
PET_t		0.0050		-7.4546		0.0280
		(0.0050)		(6.0538)		(0.0315)
$\text{PET}_{(t-1)}$		0.0048		-2.4066		-0.0998***
		(0.0050)		(6.0461)		(0.0315)
Observations	1,673,718	1,673,718	199,674	199,674	378,859	378,859
R^2	0.0647	0.0647	0.2291	0.2292	0.1527	0.1529
[B] Growing season						
Prec GS $_t$	-0.0003	0.0007	3.8901***	2.8326**	0.0159**	0.0061
	(0.0009)	(0.0009)	(1.3120)	(1.3546)	(0.0074)	(0.0077)
Prec GS $_{(t-1)}$	-0.0013	-0.0005	3.6749***	2.7837**	0.0059	-0.0020
	(0.0009)	(0.0009)	(1.2832)	(1.3257)	(0.0070)	(0.0072)
PET GS $_t$		0.0036		-5.3823*		-0.0393**
		(0.0027)		(3.1412)		(0.0172)
PET GS $_{(t-1)}$		0.0021		-3.1989		-0.0229
		(0.0027)		(3.1077)		(0.0172)
Observations	1,384,393	1,384,393	166,138	166,138	311,944	311,944
R^2	0.0386	0.0387	0.2100	0.2101	0.1465	0.1467
Controls	Long	Long	Long	Long	Long	Long
Year FE	Y	Y	Y	Y	Y	Y
Cluster FE	Y	Y	Y	Y	Y	Y
Country Trends	Y	Y	Y	Y	Y	Y

Notes: the table presents the estimates of β_1 and β_2 for the sample of women who have been living in the same place of residence for at least 15 years. Long controls include the woman's education in single years, wealth index of the household, child's month of birth, main source of water, and cell temperature in t and $t-1$. Columns (1) to (2) are based on a sample of 34 countries and 22,900 clusters. Columns (3) to (6) are based on a sample of 33 countries and 22,757 clusters. In Panel [A], precipitations and PET are computer over the entire year; in Panel [B], only the growing season months are considered. Robust standard errors are clustered at the DHS cluster level, with significance levels at 10, 5, and 1 percent.

Table A.3: Impact of precipitations and the PET on infant health – Urban vs rural areas

	Mort (5y)		**BMI (SD)**		**Size at birth**	
	(1)	(2)	(3)	(4)	(5)	(6)
			[A] Rural Sample			
Prec_t	-0.0043***	-0.0027**	9.1989***	6.9904***	0.0488***	0.0319***
	(0.0012)	(0.0012)	(1.6058)	(1.6906)	(0.0088)	(0.0092)
$\text{Prec}_{(t-1)}$	-0.0021*	-0.0006	2.9441*	0.9855	0.0408***	0.0249***
	(0.0012)	(0.0012)	(1.5243)	(1.5885)	(0.0080)	(0.0083)
PET_t		0.0016		-8.9526		0.0163
		(0.0055)		(6.3722)		(0.0326)
$\text{PET}_{(t-1)}$		0.0052		-3.4781		-0.0932***
		(0.0055)		(6.3143)		(0.0322)
Observations	1,477,984	1,477,984	186,831	186,831	367,939	367,939
R^2	0.0634	0.0634	0.2093	0.2094	0.1408	0.1409
			[B] Urban Sample			
Prec_t	-0.0030**	-0.0014	6.9626***	5.0379**	0.0052	-0.0029
	(0.0015)	(0.0016)	(2.3029)	(2.4781)	(0.0107)	(0.0116)
$\text{Prec}_{(t-1)}$	-0.0035**	-0.0019	6.5940***	4.8267**	0.0074	-0.0005
	(0.0015)	(0.0016)	(2.3223)	(2.4551)	(0.0105)	(0.0114)
PET_t		0.0071		-9.9607		0.0082
		(0.0078)		(9.8640)		(0.0493)
$\text{PET}_{(t-1)}$		0.0012		-2.5634		-0.0508
		(0.0079)		(9.8445)		(0.0495)
Observations	581,679	581,679	79,505	79,505	150,620	150,620
R^2	0.0602	0.0602	0.2276	0.2277	0.1563	0.1563
Controls	Long	Long	Long	Long	Long	Long
Country FE	N	N	N	N	N	N
Year FE	Y	Y	Y	Y	Y	Y
Cluster FE	Y	Y	Y	Y	Y	Y
Country Trends	Y	Y	Y	Y	Y	Y

Notes: the table presents the estimates of vectors $\beta_{\mathbf{1}}$ and $\beta_{\mathbf{2}}$ separately for rural (Panel [A]) and urban (Panel [B]) clusters. Climate variables are measured using the entire years t and $t-1$. Long controls include the woman's education in single years, wealth index of the household, child's month of birth, main source of water, and cell temperature at times t and $t-1$. Columns (1) and (2) are based on a sample of 16,319 rural and 11,292 urban clusters. Columns (3) to (6) are based on a sample of 16,560 rural and 11,368 urban clusters. Robust standard errors are clustered at the DHS cluster level, with significance levels at 10, 5, and 1 percent.

Table A.4: Precipitations, the PET and child health – Unstandardized coefficients

	Mort (5y)		BMI (SD)		Size at birth	
	(1)	(2)	(3)	(4)	(5)	(6)
			[A] Yearly values			
Prec_t	-0.0008***	-0.0004***	1.2536***	0.8729***	0.0061***	0.0037***
	(0.0002)	(0.0002)	(0.2043)	(0.2173)	(0.0011)	(0.0011)
$\text{Prec}_{(t-1)}$	-0.0006***	-0.0002	0.5567***	0.2184	0.0052***	0.0029***
	(0.0002)	(0.0002)	(0.1972)	(0.2080)	(0.0010)	(0.0011)
PET_t		0.0017		-3.0913*		0.0045
		(0.0014)		(1.6609)		(0.0086)
$\text{PET}_{(t-1)}$		0.0010		-0.4352		-0.0242***
		(0.0014)		(1.6575)		(0.0085)
Observations	2,059,690	2,059,690	266,469	266,469	518,591	518,591
R^2	0.0601	0.0601	0.1999	0.2000	0.1331	0.1332
			[B] Growing season			
Prec GS_t	-0.0001***	-0.0000*	0.0700***	0.0489**	0.0003***	0.0002
	(0.0000)	(0.0000)	(0.0220)	(0.0228)	(0.0001)	(0.0001)
Prec GS_{t-1}	-0.0001***	-0.0001***	0.0658***	0.0445**	0.0002**	0.0001
	(0.0000)	(0.0000)	(0.0219)	(0.0227)	(0.0001)	(0.0001)
PET GS_t		0.0083**		-7.2105*		-0.0655***
		(0.0038)		(4.2934)		(0.0231)
PET GS_{t-1}		0.0010		-7.2090*		-0.0216
		(0.0038)		(4.2735)		(0.0230)
Observations	1,721,384	1,721,384	227,786	227,786	432,729	432,729
R^2	0.0598	0.0598	0.1833	0.1835	0.1287	0.1289
Controls	Long	Long	Long	Long	Long	Long
Year FE	Y	Y	Y	Y	Y	Y
Cluster FE	Y	Y	Y	Y	Y	Y
Country Trends	Y	Y	Y	Y	Y	Y

Notes: the table presents unstandardized estimates of β_1 and β_2. Rainfall and PET are rescaled to represent a 100 mm/month variation. Long controls include the woman's education in single years, wealth index of the household, child's month of birth, main source of water, and cell temperature in t and $t-1$. In Panel [A], precipitations and PET are computer over the entire year; in Panel [B], only growing season months are considered. Robust standard errors are clustered at the DHS cluster level, with significance levels at 10, 5, and 1 percent.

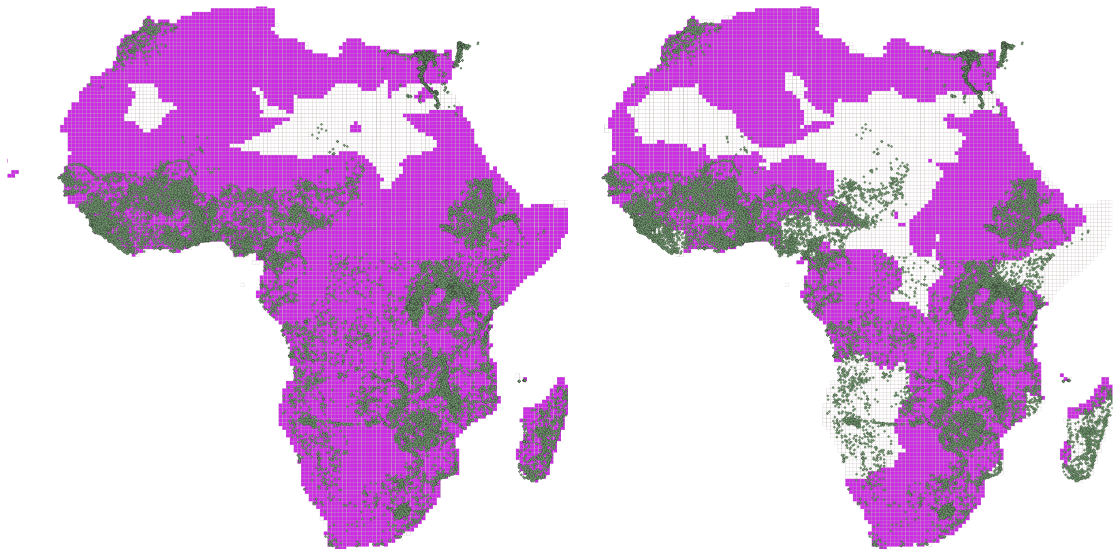
Table A.5: Precipitations and the PET projections – Sample averages

	RCP 4.5				**RCP 8.5**			
	Rainfall		PET		Rainfall		PET	
	2040–2079	2060–2099	2040–2079	2060–2099	2040–2079	2060–2099	2040–2079	2060–2099
GFDL	1094.3	1098.4	1437.5	1458.7	1099.7	1105.2	1520.6	1600.6
HadGEM2	1117.4	1123.7	1524.2	1566.7	1116.7	1114.1	1627.9	1743.6
IPSL	1176.0	1189.0	1500.6	1533.6	1222.3	1255.6	1608.7	1730.0
MIROC	1150.9	1180.3	1458.6	1488.6	1183.7	1224.0	1547.1	1547.1
NorESM1	1107.0	1113.4	1419.4	1441.1	1111.9	1128.0	1496.6	1578.9
Sample	1021.1	1021.1	1412.9	1412.9	1021.1	1021.1	1412.9	1412.9

Notes: the table presents the projections for yearly precipitations and the PET (mm/year) averaged over our sample grid from five ESMs (GFDL-ESM2M, HadGEM2-ES, IPSL-CM54-LR, MIROC-ESM-CHEM, NorESM1-M) for the time intervals 2040–2079 and 2060–2099. Data are displayed for two RCP scenarios: 4.5 (decreasing "intermediate" emission levels by 2100) and 8.5 (non-decreasing "worst-case" emission levels by 2100). Source: CMCC-BioClimInd dataset.

A. Figures

Figure A.1: DHS clusters and CRU TS.04 sensor coverage



[1970–1979] [2000–2009]

Notes: green dots represent the spatial distribution of DHS clusters in our sample. The purple-shaded area identifies the grid cell station coverage of the CRU dataset in the decades 1970–2079 and 2000–2009.
Source: Harris et al. (2020).

GREEN
Centre for Geography, Resources, Environment, Energy and Networks
via Röntgen, 1
20136 Milano - Italia

www.green.unibocconi.eu